

Using Ensemble Data Assimilation for Predictability and Dynamics

Ryan D. Torn

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2007

Program Authorized to Offer Degree: Atmospheric Sciences

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Ryan D. Torn

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

Gregory J. Hakim

Reading Committee:

Gregory J. Hakim

Clifford F. Mass

Robert A. Houze

Date: _____

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature_____

Date_____

University of Washington

Abstract

Using Ensemble Data Assimilation for
Predictability and Dynamics

Ryan D. Torn

Chair of the Supervisory Committee:
Professor Gregory J. Hakim
Atmospheric Sciences

Atmospheric predictability depends in part on the sources and evolution of errors in numerical weather prediction models. As a consequence, it is important to initialize a model with the best estimate of the atmosphere and understand how errors in this initial condition will affect the forecast. The ensemble Kalman filter (EnKF) is an attractive method of initializing a forecast model because this technique uses state-dependent error statistics to spread observation information to model grid points. In addition, output from an EnKF system can be used to quantify how changes to the initial conditions and observation assimilation affect scalar functions of forecast variables.

A pseudo-operational EnKF system is implemented for a limited-area domain that includes the eastern Pacific Ocean to test the benefit of ensemble analyses and forecasts in a region characterized by sparse in-situ data and complex topography. Comparisons against rawinsondes indicate that ensemble forecasts from this system have comparable skill to other major global NWP forecasts, even though it does not consider satellite radiance data. Forecasts of average pressure over western Washington state from this EnKF system show a region of maximum sensitivity to the west of this region. The accuracy of ensemble predictions of observation impact is verified by

comparing forecasts where observations are assimilated with the control case where no observations are used. These experiments indicate that the impact of thousands of observations can be estimated by a subset of $O(100)$ most-significant observations.

These ensemble techniques are applied to understand the initial condition sensitivity and observation impact during forecasts of western Pacific extratropical transition (ET) events, which are often characterized by large short-term forecast errors. ET forecasts are most sensitive to the position of the tropical cyclone (TC) and to upstream mid-latitude troughs that interact with the transitioning storm and other downstream features. Observation impact calculations indicate that assimilating $O(50)$ observations near the TC and upstream troughs can have nearly the same impact as all 12 000 available observations. Furthermore, the amount of downstream ridging that occurs during these events depends on the lower-tropospheric moisture flux east of the TC.

TABLE OF CONTENTS

	Page
List of Figures	iii
List of Tables	vii
Chapter 1: Introduction	1
1.1 Ensemble Forecasting	2
1.2 Data Assimilation	5
1.3 Sensitivity Analysis	10
1.4 Extratropical Transition	14
Chapter 2: Derivation of the Ensemble Kalman filter and Ensemble Sensitivity Analysis	20
2.1 Ensemble Kalman Filter	20
2.2 Ensemble Sensitivity	27
Chapter 3: Real-Time Data Assimilation System	33
3.1 Experiment Setup	33
3.2 Ensemble Verification	40
3.3 Forecast Errors	45
3.4 Comparison with Control Forecasts	53
Chapter 4: Ensemble-based Sensitivity Analysis	62
4.1 Climatological results	62
4.2 Most sensitive cases	66
4.3 Observation Denial Experiments for Single Observations	70
4.4 Observation Denial Experiments for multiple Observations	76
4.5 Impact of Observations on Forecast Verification	81

Chapter 5:	Application of Ensemble Sensitivity to Extratropical Transition Forecasts	86
5.1	Experiment Setup	86
5.2	Overview of Cases	91
5.3	Ensemble Forecasts of Transition	95
5.4	Forecast Sensitivities	107
5.5	Perturbed Initial Condition Experiments	119
5.6	Observation Impact	129
5.7	Using Ensemble Statistics to Understand the Dynamics of Downstream Ridge Development	138
Chapter 6:	Summary and Conclusions	160
Bibliography	173

LIST OF FIGURES

Figure Number	Page
3.1 Observation distribution for the UW EnKF system.	35
3.2 Verification rank histograms for UW EnKF 6, 12 and 24-hour temperature forecasts	42
3.3 Verification rank histograms for UW EnKF 24-hour forecasts of zonal wind, meridional wind, geopotential height and dew-point temperature.	44
3.4 Verification rank histograms for UW EnKF forecasts of surface fields.	46
3.5 RMS error and bias in temperature forecasts from the UW EnKF system and other operational centers.	49
3.6 RMS error and bias in 24-hour forecasts of zonal wind, meridional wind, geopotential height and dew-point temperature.	51
3.7 RMS error in the ensemble-mean and RMS spread in UW EnKF forecasts of temperature, wind, geopotential height and dew-point temperature.	52
3.8 RMS error and bias in forecasts of surface fields a function of forecast hour.	54
3.9 RMS error and bias in the ensemble-mean and “no assimilation” forecast of temperature, wind, geopotential height and dew-point temperature	56
3.10 RMS error and bias in the 24-hour ensemble-mean and WRF-GFS forecasts of temperature, wind, geopotential height and dew-point temperature	58
3.11 Mean vertical correlation between water vapor mixing ratio and zonal wind, meridional wind, temperature, and water vapor mixing ratio at different mandatory pressure levels.	61
4.1 Percentage of cycles that the average SLP over western Washington is sensitive to the analysis of SLP.	64
4.2 Percentage of cycles that the average SLP over western Washington is sensitive to the analysis of 850 hPa temperature and 500 hPa height.	65

4.3	Percentage of cycles that the average precipitation over western Washington is sensitive to various analysis fields.	67
4.4	Sensitivity of 24-hour forecast of western Washington SLP to the analysis of various fields.	69
4.5	Sensitivity of 24-hour forecast of western Washington precipitation to the analysis of various fields.	71
4.6	Forecast sensitivities, and analysis and forecast differences for one case of withholding buoy 46036.	74
4.7	Comparison of ensemble prediction of how withholding buoy 46036's SLP observation changes the 24-hour forecast of western WA SLP against the actual change obtained from forecasts.	75
4.8	Ensemble predictions of the change in average western WA SLP due to the assimilation of statistically significant surface observations versus the actual change obtained from perturbed WRF forecasts.	77
4.9	RMS difference between the ensemble prediction and actual change in the expected value and spread for various statistical significance levels.	79
4.10	Ensemble predictions of the change in the expected value and spread of western WA and NAC SLP forecasts due to assimilating all observations versus the actual change obtained from perturbed WRF forecasts.	80
4.11	Histograms of the impact of individual observations on the average SLP within the western WA metric.	82
4.12	Ensemble predictions of the impact of all observations on the RMS error in western WA and NAC SLP forecasts versus the actual change obtained from the WRF model.	84
4.13	Histograms of the impact of individual observations on the RMS error in SLP forecasts within the western WA box.	85
5.1	Tropical cyclone best track position and intensity data for cases described here.	93
5.2	RMS error in six-hour WRF EnKF and GFS forecasts of temperature and meridional wind.	96
5.3	WRF EnKF tropical cyclone track forecasts for each ET case explored here.	99
5.4	WRF EnKF tropical cyclone intensity forecasts for each ET case explored here.	100
5.5	WRF EnKF forecast of sea-level pressure and 500 hPa height initialized on 12 UTC 19 October 2004 (Tokage).	102

5.6	WRF EnKF forecast of sea-level pressure and 500 hPa height initialized on 00 UTC 6 September 2005 (Nabi).	104
5.7	WRF EnKF forecast of sea-level pressure and 500 hPa height initialized on 00 UTC 17 October 2005 (Kirogi).	106
5.8	WRF EnKF forecast of sea-level pressure and 500 hPa height initialized on 12 UTC 30 November 2003 (Lupit).	108
5.9	Sensitivity of Tokage's minimum SLP forecast to the analysis of SLP at each grid point.	111
5.10	Sensitivity of Tokage's minimum SLP forecast to the analysis of 500 hPa height at each grid point.	112
5.11	Comparison of 48-hour SLP and 850 hPa temperature for the ensemble member with the most (least) intense cyclone.	113
5.12	Sensitivity of the 48-hour forecast of cyclone minimum SLP to the analysis of 500 hPa height for the Nabi, Kirogi and Lupit forecast. . .	115
5.13	Sensitivity of the RMS error in SLP forecasts surrounding the cyclone to the analysis of 500 hPa height for the Tokage forecast.	118
5.14	Sensitivity of the RMS error in SLP forecasts to the analysis of 500 hPa height for the Nabi, Kirogi, and Lupit forecasts.	120
5.15	Difference between the perturbed and control forecast of SLP and 500 hPa height at hour 0 and hour 48 at 12 UTC 19 October. The perturbed forecast is scaled to increase the 48-hour minimum SLP forecast by 18 hPa.	123
5.16	Change in the 48-hour cyclone minimum SLP forecast due to applying diagnostic corrections to the initial conditions.	125
5.17	Difference between the perturbed and control forecast of SLP and 500 hPa height at hour 0 and hour 48 at 12 UTC 19 October. The perturbed forecast is scaled to decrease the RMS error in 48-hour SLP forecast by 11 hPa.	128
5.18	Change in the RMS error in 48-hour SLP forecasts due to applying diagnostic corrections to the initial conditions.	130
5.19	Change in the expected value of the 48-hour cyclone minimum SLP forecast due to assimilating individual observations.	133
5.20	Change in the spread of the 48-hour cyclone minimum SLP forecast due to assimilating individual observations.	135
5.21	Evolution of the SLP, precipitation, 850 hPa temperature, frontogenesis and upper-tropospheric PV during the transition of Tokage.	140

5.22	Change in the six-hour 250 hPa PV forecast during Tokage's transition associated with adjusting the amount of precipitation.	142
5.23	Change in the six-hour 250 hPa PV forecast during Tokage's transition associated with the 850 hPa frontogenesis or 850 hPa moisture flux. . .	144
5.24	Change in the 12-hour 250 hPa PV forecast during the Tokage transition obtained when the initial conditions are adjusted to increase the analysis of 850 hPa moisture flux.	147
5.25	Evolution of the SLP, precipitation, 850 hPa temperature, frontogenesis and upper-tropospheric PV during the transition of Nabi.	149
5.26	Change in the six-hour 250 hPa PV forecast during Nabi's transition associated with adjusting the amount of precipitation.	150
5.27	Change in the six-hour 250 hPa PV forecast during Nabi's transition associated with the 850 hPa frontogenesis or 850 hPa moisture flux. . .	152
5.28	Change in the 12-hour 250 hPa PV forecast during the Nabi transition obtained when the initial conditions are adjusted to increase the analysis of 850 hPa moisture flux.	153
5.29	Evolution of the SLP, precipitation, 850 hPa temperature, frontogenesis and upper-tropospheric PV during the transition of Lupit.	155
5.30	Change in the six-hour 200 hPa PV forecast during Lupit's transition associated with adjusting the amount of precipitation.	156
5.31	Change in the six-hour 200 hPa PV forecast during Lupit's transition associated with the 850 hPa frontogenesis or 850 hPa moisture flux. . .	158
5.32	Change in the 12-hour 250 hPa PV forecast during the Lupit transition obtained when the initial conditions are adjusted to increase the analysis of 850 hPa moisture flux.	159

LIST OF TABLES

Table Number		Page
3.1	Number and type of observations assimilated by the UW EnKF system.	34
3.2	List of global models used in the UW EnKF comparison.	47
3.3	RMS error and bias for UW EnKF, “no assimilation” and WRF-GFS surface field forecasts.	57
5.1	Number and type of observations assimilated during the ET forecasting experiments.	88
5.2	List of extratropical transition forecasts studied here.	90
5.3	RMS error in the ensemble-mean analyses and six-hour forecast of tropical cyclone position and intensity.	97
5.4	Change in the 48-hour cyclone minimum SLP forecast due to assimilating individual statistically significant observations.	131
5.5	Multi-variate linear regression coefficients where the 850 hPa frontogenesis and moisture flux are the predictors and the average 250 hPa PV is the predictand.	145

ACKNOWLEDGMENTS

I would like to extend my gratitude to the many people who have helped me reach this stage. In particular, I want to thank my advisor Greg Hakim, for teaching me how to be a successful scientist, sharing his ideas and knowledge of the atmosphere, and allowing me the latitude to explore various ideas. The remaining members of my committee, Dennis Hartmann, Bob Houze, Cliff Mass, and Chris Snyder provided valuable insight on the results obtained during this work and the encouragement to branch out into new directions. Over the last four years, my amazing wife Heather has been the most patient and loving person, especially during the preparation of this dissertation. Steven Cavallo, Lucas Harris, Rahul Mahajan, and Helga Schaffrin Huntley always provided constructive comments during our weekly group meetings. Daily conversations on predictability and ensemble data assimilation with Alex Reinecke have been thought provoking and helpful. This work would not have been possible without countless hours of computer support from Harry Edmon and David Warren. The rest of the support staff in the Atmospheric Sciences Department has been helpful and professional throughout my time here. Finally, I wish to thank my family and friends, especially my parents, who have always been supportive and encouraging of a career in science.

Chapter 1

INTRODUCTION

Although there have been significant improvements to modern numerical weather prediction (NWP) models over the past 20 years, large short-term forecast errors still exist for many high impact weather events in the mid-latitudes. These events include North Pacific cyclones (e.g., McMurdie and Mass 2004), east coast cyclones (e.g., Langland et al. 2002), and the extratropical transition of tropical cyclones (e.g., Jones et al. 2003). Numerical model output with large errors can lead to inadequate preparation for these events and loss of life and property.

Understanding the sources and evolution of errors in NWP models is critical to improving forecasts of various atmospheric phenomenon. Errors can originate from two primary sources: the model initial conditions (i.e., the analysis), or errors in the formulation of the model (i.e., model error); this dissertation focuses on exploring how the former affects the ability of a NWP model to forecast the evolution of the atmosphere. Lorenz (1963) and subsequent papers showed that for chaotic systems, such as the atmosphere, small differences in the model's initial conditions can lead to the exponential divergence of forecast solutions over time. As a consequence, several tools have been independently developed to overcome and understand the challenges presented by chaotic systems. Ensemble forecasting acknowledges the forecast errors that result from initial condition deficiencies by providing a probabilistic estimate of the forecast. Data assimilation, the process by which a short-term forecast is combined with observations to produce a best estimate of the atmospheric state, is used to generate initial conditions for a model. Finally, sensitivity analysis is

a method of quantifying how changes to the initial conditions affect the subsequent forecast and suggests where additional observations could improve a forecast. In order to gain maximum utility, each of the aforementioned techniques should consider how the others will impact the forecast. For example, the initial conditions for ensemble forecasting should account for the density of the assimilated observations. This thesis describes a consistent framework for understanding atmospheric predictability and dynamics.

The remainder of this chapter proceeds as follows. An overview of previous applications of ensemble forecasting is given in section 1.1. A short description of various data assimilation techniques is discussed in section 1.2. In section 1.3, a summary of previous research on quantifying how initial condition changes affect a forecast is given. Finally, section 1.4 describes the extratropical transition of tropical cyclones, which is often characterized by large forecast errors resulting from initial condition deficiencies.

1.1 Ensemble Forecasting

For chaotic system such as the atmosphere, a single deterministic forecast can be misleading because arbitrary small errors in the initial conditions can lead to exponentially divergent solutions at some later time. As a consequence, numerical forecasting models must take into account the stochastic nature of the atmosphere by solving for a probability density function (PDF) of the forecast, rather than for a single deterministic solution. In addition, probabilistic forecasting also gives an *a priori* estimate of the error in the forecast; if many ensemble members are predicting the same event to occur, it lends confidence to the forecast.

The most accurate method of obtaining the forecast PDF is through stochastic-dynamic prediction whereby an analysis PDF is integrated forward using the probability continuity equation (e.g., Epstein 1969). Even for a simple low-order model, the integration of the probability continuity equations is computationally intractable,

thus a Monte Carlo approximation is made whereby a small number of ensemble members are used to approximate the full forecast PDF. In the limit as the number of ensemble members approaches infinity, the PDF estimated from the ensemble should approach the stochastic-dynamic prediction.

Several different techniques have emerged for generating an ensemble of initial conditions from a deterministic global analysis. Hoffman and Kalnay (1983) proposed generating an initial condition ensemble using lagged forecasts verifying at the same time. This idea was extended by Toth and Kalnay (1993) who constructed an ensemble of initial conditions by subtracting a control deterministic forecast from a short-term forecast ensemble verifying at the analysis time, scaling the perturbations to reflect the initial condition uncertainty, and adding these perturbations to the deterministic analysis. These perturbations, known as “bred vectors”, represent the fastest growing modes during the previous forecast cycle, thus Toth and Kalnay (1993) have argued that analysis errors should project strongly onto the bred vectors. Breeding has been used to produce ensemble forecasts by the National Centers for Environmental Prediction (NCEP) since 1993 (Toth et al. 1997), though recent improvements to this system have included using ensemble transform methods to scale the initial-time bred vectors (Wei et al. 2006).

Other operational forecasting centers, such as the European Center for Medium Range Weather Forecasting (ECMWF), have developed ensemble forecasting systems based on linear combinations of initial-time singular vectors (Molteni and Palmer 1993, Molteni et al. 1996). Singular vectors are a set of orthogonal structures that have different growth rates for a given forecast interval and forecast metric, where leading singular vectors represent the directions of fastest error growth in the model. For ensemble forecasting, it is desirable to have initial perturbations that grow with the current atmospheric flow; these growing perturbations can be determined by integrating the leading forecast singular vectors backward in time using a linearized version of the model and adding a linear combination of the initial time singular

vectors to the deterministic analysis. The choice of forecast metric, which is a function of the forecast state variables, can result in very different singular vector structures. Palmer et al. (1998) have argued that initial time singular vectors based on the total energy norm grossly represent the analysis-error variance, thus making it the norm of choice for ensemble forecasting. The initial condition ensemble constructed by Barkmeijer et al. (1998), based on the analysis-error covariance norm, lead to initial time singular vectors whose amplitude varies with the density of the observation network.

Several authors have also investigated methods of producing mesoscale ensemble forecasts. Stensrud et al. (1999) applied the breeding technique to the Meso-eta model to create a short-range ensemble prediction system. Although this system produced a significant amount of ensemble spread, their results showed little correlation between the error in the ensemble-mean and the ensemble spread. Eckel and Mass (2005) found that ensemble forecasts with the fifth-generation Pennsylvania State University-National Center for Atmospheric Research (PSU-NCAR) Mesoscale Model (MM5) initialized from eight different operational analyses were dominated by model error, which must be accounted for to maximize the utility of the ensemble. In fact, multi-model ensemble forecasts based on the same initial conditions, but different physics parameterizations, performed better than the initial condition ensemble because it considered the impact of model errors. Wandishin et al. (2001) showed that a multi-model ensemble system that includes members from NCEP's short-range ensemble forecasting system contained a lack of mesoscale error growth in fields such as precipitation and convective indices. In contrast, Stensrud and Yussouf (2003) found that a multi-model ensemble system based on different physics packages is able to predict the skill in surface air and dew-point temperature forecasts.

All of the aforementioned ensemble techniques are meant to obtain accurate ensemble forecasts; however, they do not contain the actual analysis uncertainty at the appropriate scales. Having a state-dependent estimate of the analysis variance can

then describe where the analysis has potentially large errors. The following section contains an alternative method of generating ensemble forecasts based on analysis-error statistics estimated from data assimilation.

1.2 Data Assimilation

It is impossible to know the true state of the atmosphere at any one instant in time, thus we generate a best estimate of the state by combining all information on the atmospheric state via data assimilation. This information usually comes from observations of the atmosphere, such as from surface stations, rawinsonde profiles, or satellite radiances, and from the physical laws that govern the evolution of the flow, which are usually in the form of a numerical model. This analysis is produced by combining these two pieces of information together in a manner that takes the error in each into account. Observation errors result from instrument error and how well an observation represents the surrounding area (i.e., representativeness error), which is often difficult to define. For example, a surface wind observation over the ocean has smaller representativeness error than a surface wind observation near a forest because the trees will shelter the observation platform. Forecast-error statistics are typically not as well known as observation error statistics, but are the key component to any data assimilation system because they determine the relative weighting of the background forecast, and how to distribute the information from observations onto a model grid. Systems that use a flow-dependent estimate of the error statistics are more likely to use the observation information in an optimal way. Two different classes of data assimilation schemes based on statistical estimation theory are used in the atmospheric sciences: direct solve methods which determine the analysis that minimizes the error variance, and the variational methods which find the most likely analysis.

Whereas early techniques of incorporating observations into models involved empirical procedures such as nudging, optimum interpolation (OI) was one of the first

operational data assimilation methods based on statistical estimation theory (e.g., Lorenc 1981, Mitchell 1990). OI computes an analysis via a weighted linear combination of a short-term forecast and the available observations at a single time (i.e., sequential). The weights for the background forecast and the observation data are determined from a set of fixed error statistics, thus the observation information is spread in a sub-optimal manner. Forecast-error statistics are based on horizontal and vertical correlations between forecast errors, determined from the difference between a short-term forecast and rawinsonde observations (e.g., Hollingsworth and Lonnberg 1986). Furthermore, OI requires a number of additional approximations which include determining a “radius of influence” for observations. Restricting the effect of observations in this way can lead to horizontal shocks for grid points influenced by different observations.

The various approximations and associated errors with OI prompted many operational NWP centers to adopt variational data assimilation schemes (e.g., Parrish and Derber 1992, Rabier et al. 1998). Variational methods find the best analysis of the state by minimizing a cost function that includes the forecast and observation error statistics. Moreover, these techniques generate an analysis by considering all data simultaneously, thus it avoids many of the practical approximations needed to implement OI, such as limiting the influence of observations. Forecast-error statistics are computed by taking the time-average difference between short-range model forecasts verifying at the same time, which compared to OI, will give a better estimate of the forecast-error covariance; however, these statistics are still not flow-dependent. In three dimensional variational assimilation (3D-Var), the cost function includes information about the forecast and observations at a single time. An extension of this technique, four dimensional variational assimilation (4D-Var) uses the model as a strong constraint to find an initial condition that best fits observations over a given time interval (e.g., Courtier and Talagrand 1987, Rabier et al. 2000). ECMWF forecasts initialized from an analysis created by 4D-Var had greater skill than forecasts

initialized from the operational 3D-Var analyses, especially for baroclinic cyclones and in the tropics (Klinker et al. 2000). Lorenc (2005) hypothesizes that 4D-Var is superior to 3D-Var because observations can be considered over a window, and are assimilated with time-evolved covariances. 4D-Var requires the adjoint of a tangent-linear model to find the trajectory that best fits the observations. Adjoint models suffer from a number of limitations including the difficulty of coding the adjoint of a tangent linear model, especially for irreversible on-off processes contained in boundary layer or microphysical parameterizations. Moreover, adjoint models are also constrained by having to be updated every time the model formulation is changed and the inherent linearity assumptions, and thus may not work as well in situations that are governed by highly non-linear dynamics (e.g., Ancell and Mass 2006).

The Kalman filter (Kalman and Bucy 1961) is a sequential method of producing an analysis state that is the most likely estimate and has the minimum error variance. Unlike OI and variational methods, the forecast error statistics are flow-dependent because they are advanced using the model itself, rather than being estimated from a constant matrix. Similar to OI, the analysis is computed via a linear combination of the short-term forecast and observation data, weighted by the observation and forecast-error statistics. For NWP applications, advancing the forecast-error covariance with a model is computationally prohibitive, thus Evensen (1994) proposed the ensemble Kalman filter (EnKF), which is a Monte Carlo approximation to the Kalman filter. Rather than integrating the forecast-error covariance matrix, this method approximates the forecast-error statistics using an ensemble of short-term forecasts. Moreover, the EnKF produces an ensemble of analyses by performing parallel data assimilation cycles on the short-term forecast ensemble. The resulting analysis ensemble is consistent with the analysis-error statistics and thus provides a natural combination of ensemble forecasting and data assimilation.

Several authors have demonstrated the benefit of using flow-dependent error statistics to assimilate observations in regions of sparse observations. Snyder and Hamill

(2000) compared analyses and forecasts for a hybrid 3D-Var EnKF system that uses a combination of fixed and flow-dependent error statistics. Their results show that the analysis system performs the best when flow-dependent error statistics are primarily used, especially in regions of reduced observation density. Whitaker et al. (2004) explored the feasibility of performing a reanalysis prior to the rawinsonde era using only surface pressure observations. Their experiments assimilate a sub-sampled version of the 2001 surface pressure observation network using both an EnKF and 3D-Var scheme. Analysis errors for the EnKF system are 50% lower than the analysis errors for the 3D-Var system, especially in regions of fewer observations.

Given the attractive properties of the EnKF, several groups are actively working on applying this technique to operational global forecast models. Whitaker et al. (2007) implemented an EnKF data assimilation system with the NCEP Global Forecasting System (GFS) and compared the resulting EnKF forecasts to those initialized with 3D-Var. These two data assimilation systems only assimilate conventional observations and are verified against operational analyses, which includes satellite data. Over a six-week period, forecasts from the EnKF data assimilation system outperformed the 3D-Var initialized forecast, especially in the data sparse regions such as the Southern Hemisphere. In contrast, Houtekamer et al. (2005) found that for the Canadian Meteorological Center's global model, EnKF-initialized forecasts have errors that are comparable to forecasts initialized using 3D-Var. The similarity of the errors in 3D-Var and EnKF forecasts most likely results from the large number of observations assimilated and their parameterization of model error, which includes the 3D-Var forecast-error statistics.

The EnKF has been shown to be particularly useful on the convective scale, where relationships between variables and error statistics are continuously evolving and not well known. Snyder and Zhang (2003) investigated the potential of using an EnKF to assimilate radar observations on the convective scale in a perfect model context. Their experiments demonstrate that simulated Doppler radar observations assimilated with

flow-dependent error statistics constrain the error in both observed and unobserved variables. In a followup study, Zhang et al. (2004) showed that assimilating data from a hypothetical mesonet with an EnKF can also simulate the storm properly. Caya et al. (2005) compared the analyses generated by using 4D-Var and an EnKF to assimilate radar observations in a perfect model scenario. For a limited period of time, the 4D-Var analyses are better than the EnKF; however, at longer times, the EnKF is superior because it uses continuously evolved background-error statistics. Dowell et al. (2004) explored the feasibility of retrieving temperature and wind fields for an isolated convective storm by using an EnKF to combine actual radar reflectivity and wind data with an anelastic model. Although the location of the main updraft and meso-cyclone are determined rather accurately, the low-level temperature field contains large errors due to the lack of observations near the surface and model biases.

Whereas most of the previous implementations of the EnKF have focused on planetary to synoptic and convective scales, relatively few studies have explored the EnKF at the synoptic scale to mesoscale, which are of practical importance for both forecasting and analysis purposes. One reason for the lack of studies at these scales is the need for a properly posed ensemble of lateral boundary conditions. Convective-scale case studies avoid the use of an ensemble of lateral boundary conditions because they are performed over short periods of time and global models do not require boundary conditions. Torn et al. (2006) proposed and tested several methods of generating an ensemble of lateral boundary conditions even when a global ensemble does not exist or is of the wrong size. Dirren et al. (2007) performed EnKF data assimilation experiments on a limited-area domain over the Pacific Northwest region of the United States, using the Weather Research and Forecasting (WRF) model in a perfect model context. Their results show that observations systematically reduce the error in both observed and unobserved variables, especially along the North American coast. One of the goals of this dissertation is to demonstrate the benefit of assimilating observations with flow-dependent error statistics in a region characterized by sparse *in situ* data

and complex topography by evaluating two years of output from a pseudo-operational EnKF system on the domain used by Dirren et al. (2007).

1.3 Sensitivity Analysis

One method of improving forecasts is to identify where small changes to the initial conditions can have a significant impact on the subsequent forecast and to assimilate supplementary observations in that area. Several objective procedures, such as singular vectors and adjoint sensitivity, have been used to quantify how small changes to the initial conditions will affect a forecast metric. As discussed previously, singular vectors are a set of orthogonal perturbation structures that, for a given metric and time interval, represent the fastest growing modes of the model. Adjoint-based sensitivity methods compute the gradient of a forecast metric with respect to the initial state using the adjoint of the linearized forecast model. Both of these techniques are based on linear perturbations about a non-linear forecast trajectory, and thus are only useful when non-linearities are small.

Several authors have applied adjoint and singular vector techniques to determine the initial condition sensitivity of mid-latitude cyclogenesis. Errico and Vukicevic (1992) used the adjoint of the MM4 mesoscale model to determine the sensitivity of the 36-hour forecasts of several mid-latitude cyclones. The largest sensitivities were associated with sub-synoptic scale structures that tilted upstream with height and are maximized in the middle troposphere. Zou et al. (1998) showed that five-day forecasts of an eastern United States cyclone were most sensitive to the lower-tropospheric temperature over the Rocky Mountains and an upper-level PV anomaly in the Gulf of Alaska, though the distribution was quite complicated. Langland et al. (2002) found that for an explosively deepening cyclone along the east coast of the United States, the vertically integrated sensitivities for a 72-hour Navy Operational Global Analysis and Prediction System (NOGAPS) forecast are largest in a broad region well upstream in a zonally-propagating wave packet. Rabier et al. (1996), Palmer et al.

(1998), and Gelaro et al. (1998) applied singular vectors to explore initial condition sensitivity for cyclones. Their results indicate that the fastest growing modes are also associated with sub-synoptic scale, upshear tilted structures that are maximized in the lower troposphere. Over the course of the forecast, these structures subsequently grow to synoptic-scale size.

Regions of high sensitivity indicate where small initial condition errors can rapidly grow during the forecast; therefore, they can be used to apply *a posteriori* diagnostic corrections to the initial conditions that will decrease the forecast error. Over a one month period, Rabier et al. (1996) obtained a 10% reduction in the 72-hour forecast error when corrections consistent with the leading singular vectors are applied to the initial conditions. Their results also indicate that longer-term forecasts (> 72 hours) also have greater skill than forecasts where corrections are not made to the initial conditions. Zou et al. (1998) and Klinker et al. (1998) introduced perturbations to the analysis based on minimizing the error in 12-hour forecasts. In both studies, the perturbed initial conditions correct key analysis errors, and thus the resulting forecast significantly outperforms the control forecast. In the case studied by Langland et al. (2002), diagnostic corrections to the initial conditions yielded a 75% improvement in the 72-hour forecast of cyclone track and intensity.

Given that diagnostic corrections to the initial conditions based on objective sensitivity analysis can improve forecasts, several field campaigns have been undertaken to determine how additional “targeted” observations in regions of high initial condition sensitivity could reduce the error in forecasts. The 1997 Fronts and Atlantic Storm-Track Experiment (FASTEX) was conducted to improve short-range forecasts of cyclones making landfall in England and Ireland by deploying dropsondes in regions of large forecast sensitivity (Joly et al. 1999). Bergot (2001) evaluated the impact of targeted dropsonde data in 20 FASTEX cases using a 4D-Var data assimilation system. On average, there was a 10% reduction in 24-hour forecast errors when dropsondes are included, although there were cases where observations had no impact on

the forecast because the sensitive region was inadequately sampled. The North Pacific Experiment (NORPEX), conducted during January-February 1998, used total energy singular vectors as guidance for where to deploy dropsondes to improve cyclone forecasts in the Northeast Pacific. Langland et al. (1999) obtained a 10% improvement in 2 day forecasts when dropsonde data is included in the NOGAPS analysis; however, Cardinali and Buizza (2003) showed this data yielded minimal improvement to ECMWF forecasts. The 2003 North Atlantic THORPEX Regional Campaign (NA-TReC) was used to test many of the objective targeting methods described above in the context of the current observational network. Although the targeted dropsonde data during NA-TReC has about three times more impact per observation than regular rawinsonde data (Langland 2005a), the total impact of dropsondes was smaller than the FASTEX experiment, probably because of substantial increases in the amount of regular observations (Langland 2005b). Furthermore, Aberson (2003) and Wu et al. (2005) describe improvement in tropical cyclone guidance when targeted dropsondes are deployed in the synoptic environment surrounding the tropical cyclone. The interested reader is directed to Langland (2005b) for a more extensive review of observation targeting to date.

Since 1999, the United States National Oceanographic and Atmospheric Administration (NOAA) has conducted the Winter Storm Reconnaissance Program (WSRP), whose goal is to “reduce uncertainty in 24-96 hour forecasts for synoptic-scale weather events associated with potentially large societal impact over the continental United States and Alaska” (Toth et al. 2002). Guidance for where to deploy aircraft-based targeted observations is obtained from the ensemble transform Kalman filter (ETKF, Bishop et al. 2001). The ETKF determines a transformation matrix that maps a short-term forecast ensemble into an analysis ensemble based on the observations that would be assimilated. Rather than describing how the forecast metric value responds to initial condition errors, this technique shows how the forecast-error covariance matrix responds to assimilating observations. Regions where large variance

reduction could be achieved by assimilating additional data suggest where additional targeted observations should be deployed. Majumdar et al. (2001) found significant overlap between the ETKF and total energy singular vector guidance during NORPEX; however, Leutbecher et al. (2004) and Majumdar et al. (2006) showed notable differences between these two techniques for NA-TReC and tropical cyclone targeting, respectively. Szunyogh et al. (2000, 2002) and Toth et al. (2002) report that in 70% of the WSRP cases, the forecast errors are reduced, though the average improvement in forecasts is about 10%, or a 12 hour gain in forecast skill.

An ensemble approach to sensitivity analysis has been proposed by Hakim and Torn (2007) and further refined by Ancell and Hakim (2007). Ensemble sensitivity analysis uses analysis and forecast ensemble data from an EnKF system, and thus combines ensemble forecasting, data assimilation and sensitivity analysis. Initial condition sensitivity is computed via linear regression of the ensemble estimate of a forecast metric and each element of the analysis state vector, thus unlike adjoint and singular vector methods, the computation is trivial. Their derivation showed that ensemble sensitivity is related to the adjoint sensitivity by the analysis-error covariance matrix. Furthermore, ensemble sensitivity can be used to compute the impact of observations, and the optimal observation locations for an EnKF system. Unlike singular vectors and adjoint sensitivity, this technique may provide an optimal strategy for observation targeting because it incorporates information on the analysis error, observation error, dynamical error growth and data assimilation (Berliner et al. 1999).

Ancell and Hakim (2007) compared ensemble and adjoint-based sensitivity for a wintertime flow pattern near the west coast of North America. Adjoint-based sensitivity is characterized by mesoscale lower-tropospheric structures that tilt strongly upshear with height. In contrast, ensemble sensitivities emphasize synoptic-scale features that have modest tilt and correspond to the significant weather features at the analysis time. Whereas Ancell and Hakim (2007) explored ensemble sensitivities

for one case, this dissertation will apply this technique to output from a pseudo-operational EnKF system over a long period of time to illustrate how ensemble sensitivities can be used to compute climatological sensitivities and estimate the impact of observations on a forecast.

1.4 Extratropical Transition

Although the extratropical transition (ET) of a warm core tropical cyclone (TC) into a cold-core baroclinic system is often characterized by large forecast errors due to initial condition deficiencies, there have been few studies that have attempted to quantify the sensitivity of these forecasts to the initial conditions. The process of ET occurs in every ocean basin with tropical cyclones, although they are most frequent in the western North Pacific Ocean (≈ 6 per year), and have a great deal of interannual variability (Hart and Evans 2001, Klein et al. 2000, Sinclair 2002). During the first stage of ET, increased vertical wind shear, lower ocean heat content and larger baroclinicity cause the TC to weaken and lose its tropical characteristics (Klein et al. 2000). The reintensification of the TC remnants into a baroclinic system depends on the relative positioning of the TC and dominant mid-latitude features. When an upper-level trough is directly upstream of the TC, the TC remnants can explosively deepen into a strong cyclone via a Petterssen type B process (Petterssen and Smebye 1971). The resulting baroclinic storms are a major hazard for marine interests because they move faster than TC, and are characterized by large regions of storm force winds ($> 25 \text{ m s}^{-1}$) and waves heights greater than 10 m (Jones et al. 2003). In contrast, when the upper trough is downstream of the TC, the cyclone does not reintensify and is sheared apart by the mid-latitude flow (Harr et al. 2000).

Previous research on these storms indicate that errors in ET forecasts are dependent on how the tropical cyclone phases with mid-latitude features. Jones et al. (2003) computed anomaly correlation scores for NOGAPS forecasts of 500 hPa geopotential heights over the North Pacific during August of 1996, when three western Pacific TC

underwent transition. The skill of NOGAPS forecasts initialized while the tropical cyclones were undergoing transition was significantly lower than the skill of NOGAPS forecasts during the remainder of the month. During all three transitions, 72-hour forecasts were no more skillful than climatology, thus forecast errors associated with ET can impact forecasts across an ocean basin on relatively short timescales. By comparing the forecast and verifying analysis, Jones et al. (2003) found that the loss of predictability is related to misplacing the significant circulation features in the basin.

Errors in ET forecasts may also result from how these hybrid systems are represented in the model analysis. Using the cyclone phase space of Hart (2003), Evans et al. (2006) evaluated errors in GFS and NOGAPS forecasts of cyclone structure initialized when a storm was a tropical cyclone, hybrid transitioning cyclone, and a baroclinic system. Their results indicate that the largest errors in cyclone structure occur when the forecast is initialized during the transition of the tropical cyclone, and are lower when the storm is a tropical cyclone or mid-latitude system. They attributed the increased error to how the TC is represented in the model analyses. Operational models require techniques such as vortex bogusing (e.g., Kurihara 1995) to initialize a TC. These schemes assume an axi-symmetric vortex; however, transitioning TC are quite asymmetric, thus the current TC initialization methods are not valid for ET.

Klein and Harr (2002) investigated how displacing the TC vortex within the initial conditions changes the evolution of three different western Pacific ET events. Vortex displacement is achieved by computing the potential vorticity (PV) of the TC, removing the wind and temperature anomaly associated with the storm PV, and placing the anomaly in the desired location. The forecast from the initial conditions with the displaced vortex are then compared to the control forecast where the TC is in the correct location. While two of the events showed minimal sensitivity to shifting the TC position, displacing Typhoon Ginger (1997) 250 km to the southwest in the

initial conditions lead to a forecast cyclone that was 1000 km to the west and 30 hPa deeper than the control after 54 hours. These large differences were attributed to a more favorable interaction between the TC and an upstream upper-level trough.

Several papers have further demonstrated the importance of the upper-level trough on the evolution of ET. McTaggart-Cowan et al. (2001) and McTaggart-Cowan et al. (2003) performed several simulations of Hurricane Earl's transition where the tropical cyclone and upper-level trough are removed from the initial conditions. The value of each of these features is determined by comparing the modified initial condition simulations to the control forecast where no changes are made. Their results indicate that the upstream trough is the most important factor; when the tropical cyclone is removed, a baroclinic cyclone still develops, but when the trough is removed, the reintensification of the TC does not occur until several days later when another upper-level trough phases with the TC remnants. Ritchie and Elsberry (2003) further demonstrated the relative role of the trough and TC during ET in an idealized framework. Their results show that whereas the TC contributes only to the rate of reintensification, the upper-level trough is a necessary condition for baroclinic cyclogenesis during ET. Furthermore, the amplitude of the trough does not limit the reintensification; forecasts initialized with a weak trough had a similar baroclinic cyclone to forecasts with a moderate or strong trough.

Two different groups have shown that forecasts of baroclinic cyclones that result from ET are most sensitive to the initial conditions near the upper-level troughs that phase with the TC remnants. Browning et al. (2000) computed initial condition singular vectors for a poorly forecast mid-latitude cyclone that originated as Hurricane Lili (1996). They determined that the fastest growing singular vectors were maximized near two tropopause-based disturbances that subsequently phase with the TC remnants. Shifting the initial position of the upper-level trough to the south, in a manner consistent with the initial-time singular vectors, leads to improved track and intensity forecasts for this cyclone. Rabier et al. (1996) explored forecast sensitivities

for several high impact weather events in Europe, which included a major cyclone that resulted from the remnants of Hurricane Floyd (1993). Much like the Lili case, the largest sensitivities are associated with the upstream flow and precursor low-level circulation in the Atlantic. Moreover, forecasts of this cyclone were improved by adjusting the initial conditions based in the magnitude of the forecast error and the initial time singular vectors.

The high sensitivity of ET forecasts to the initial conditions and large errors in deterministic forecasts suggest that ensemble forecasting may be particularly beneficial. Ma et al. (2003) evaluated Canadian Meteorological Center (CMC) global model ensemble forecasts of Hurricane Earl's transition for three different initialization times. At each time, some of the ensemble members erroneously weaken the storm before it became a baroclinic system, while others showed the TC remnants becoming a strong extratropical system. Ensemble members that accurately resolve the initial TC circulation and upper-level features were more likely to have a good forecast of ET. However, the initial ensemble perturbations produced by operational ensemble prediction systems are typically largest in the mid-latitudes and not near the TC environment, thus ensemble forecasts of these events may not contain the appropriate uncertainty (Anwender et al. 2006).

Although there are several studies that describe the initial condition sensitivity for ET, they do not focus on the period when transition begins, which has been described as having large errors. Furthermore, these previous experiments do not consider the impact of observations or the actual analysis error associated with the key features during ET. Here, ensemble sensitivity analysis will be applied to four recent ET events that are characterized by various levels of predictability to determine the initial condition sensitivity and observation impact for forecasts initialized at the onset of transition.

Rapidly intensifying ET events are often accompanied by the amplification of the upper-level flow downstream of the cyclone in a manner similar to mid-latitude

cyclogenesis events (e.g., Jones et al. 2003, Chang and Yu 1999). Previous studies (Bosart and Lackmann 1995) have demonstrated that this downstream ridge can aid in the baroclinic development of the TC remnants by sharpening the horizontal vorticity gradient. This process leads to increased vertical motion directly above the surface cyclone and larger height falls within the cyclone itself. In addition, the rapid amplification of the downstream ridge leads to an eastward-propagating Rossby wave packet that travels faster than individual troughs and ridges in the mid-latitude flow (e.g., Chang and Orlanski 1993, Hakim 2003) and the development of downstream baroclinic cyclones (e.g., Orlanski and Sheldon, 1995).

While it is well known that wave packets can be generated by ET in the western Pacific, the physical mechanism responsible for the downstream ridging is not well known (Jones et al. 2003). Two different mechanisms have been proposed in the literature that use “potential vorticity (PV) thinking” to explain this process: the adiabatic interaction between the PV of a tropical cyclone and the mid-latitude jet, and the impact of diabatically modified PV on the mid-latitude circulation. Through action at a distance (Bishop and Thorpe 1994), the circulation associated with a tropical cyclone is expected to excite Rossby waves on the upper-level PV gradients associated with the jet, which has been shown to occur in a shallow-water model by Ferreira and Schubert (1999). This second mechanism involves the modification of the upper-level potential vorticity by convection and latent heat release in the tropical cyclone. In regions of large latent heat release, such as the TC core, the PV is redistributed based on the orientation of the vorticity vector; PV is generated (destroyed) by a gradient of heating in the direction of the absolute vorticity vector (Hoskins et al. 1985). Whereas the area around the TC is characterized by positive relative vorticity, the mid-latitude flow is dominated by large vertical shear in the zonal wind, and thus a poleward-directed vorticity vector. For latent heating near the TC, one would then expect PV reduction in the upper-troposphere downstream of the transitioning cyclone. Recent studies have shown that enhanced downstream

ridging due to latent heat release is often underestimated in many NWP models (Henderson et al. 1999, Atallah and Bosart 2003).

Accurate forecasts of the downstream state during ET require a firm understanding of the features that contribute to the amplification of the downstream flow. Dynamical relationships between the transitioning TC and the mid-latitude flow are explored here by applying a series of techniques proposed by Hakim and Torn (2007), which they call “Ensemble Synoptic Analysis”, to analyze and understand atmospheric phenomenon. These methods use the information of the probabilistic analysis data that comes from an EnKF system to determine dynamical relationships between various features. Hakim and Torn (2007) demonstrated that for an extratropical cyclone, this approach confirmed linkages between surface cyclones and upper-level disturbances, but also suggests less obvious relationships. Here the short-term forecast ensemble generated from the ET predictability experiments will be used to explore how the TC, specifically the latent heat release associated with precipitation, interacts with the mid-latitude flow to create the downstream ridge.

The outline of this dissertation is as follows. Chapter 2 derives the EnKF and ensemble sensitivity from basic principles. The performance of a pseudo-operational EnKF system over the northeast Pacific is described in Chapter 3, while the output is used to determine initial condition sensitivities for this region in Chapter 4. The EnKF and ensemble sensitivities are applied to understand the predictability of four recent western Pacific ET events in Chapter 5. A concluding summary is given in Chapter 6.

Chapter 2

DERIVATION OF THE ENSEMBLE KALMAN FILTER AND ENSEMBLE SENSITIVITY ANALYSIS

This chapter derives the ensemble Kalman filter and ensemble sensitivity analysis equations from basic principles and describes the inherent assumptions made in each using the notation of Ide et al. (1997). The level of detail presented here is not exhaustive, rather it provides a framework for which the techniques and results in the following chapters can be described and understood. For a more rigorous derivation, the interested reader is directed to Kalnay (2002), Hamill (2005), and Ansell and Hakim (2007).

2.1 *Ensemble Kalman Filter*

Since the true state of the atmosphere can never be exactly known, the best estimate of the state is determined through Bayesian statistical estimation theory. In a Bayesian framework, the probability density function (PDF) of the current state of the atmosphere given all current and past observations can be expressed as the product of the observation likelihood PDF and a “prior” PDF, which is the model state given all past observations up to the previous analysis time, under the assumption that the observation and prior are uncorrelated. In a fully probabilistic framework, the prior, or background forecast PDF, is obtained by integrating the analysis PDF forward in time from the previous analysis time using the probability continuity equation. Integration of the probability continuity equation can be very computationally intensive for any system with multiple state variables, thus the PDFs are approximated by multivariate Gaussian distributions, which are described by a mean value and covariance

matrix. After substituting Gaussian distributions into the Bayesian analysis PDF equation, a cost function $J(\mathbf{x})$ can be defined that measures the distance between an arbitrary $N \times 1$ state vector, \mathbf{x} and the observations and background forecast. This cost function has the form

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x})), \quad (2.1)$$

where \mathbf{P}^b is the background-error covariance matrix ($\mathbf{x}^b \mathbf{x}^{bT}$), which is of size $N \times N$ and gives the relationship between state variables. \mathbf{x}^b is the background forecast state vector, \mathbf{y} is a $L \times 1$ vector of observation values, \mathcal{H} is an operator that maps from model state space to observation space. For observations of model state variables, \mathcal{H} can represent bi-linear interpolation; however, in the case of satellite radiance data, this operator would include complicated radiative transfer code. The quantity $\mathbf{y} - \mathcal{H}(\mathbf{x}^b)$ is the observation innovation and is the new information added by observations. \mathbf{R} is the $L \times L$ observation-error covariance matrix, which gives the relationship between observations, and for simplicity is often assumed to be diagonal, which means that individual observation errors are uncorrelated. The above equation indicates that the error in \mathbf{x} is given by the linear combination of the difference between \mathbf{x} and background forecast, weighted by the forecast errors, and the difference between the observation and the model estimate of the observation, weighted by the observation errors.

Although both direct solve and variational data assimilation techniques both derive from (2.1), they differ in how they determine the analysis. Using a minimization algorithm such as conjugate gradient or quasi-Newton methods, variational schemes find an analysis that minimizes the scalar cost function J by computing the gradient in the cost function with respect to the model state \mathbf{x} . Note that in 4D-Var, there is an extra term in (2.1) that represents fitting the model trajectory to observations over a period of time, but is omitted here. These variational techniques employ a

quasi-fixed \mathbf{P}^b for (2.1), thus the observation information may be spread to the model grid points in a sub-optimal manner, though this is less likely to be a problem for 4D-Var since it finds a model trajectory that fits the observations over a period of time (e.g., Lorenc 2005).

Sequential, direct solve schemes, such as the Kalman filter, compute a maximum likelihood analysis by evaluating the derivative of (2.1) with respect to \mathbf{x} and setting the resulting equation equal to zero

$$\frac{\partial J}{\partial \mathbf{x}} = -2 \frac{\partial}{\partial \mathbf{x}} [\mathcal{H}(\mathbf{x})] \mathbf{R}^{-1} [\mathbf{y} - \mathcal{H}(\mathbf{x})] + 2 \mathbf{P}^{b-1} [\mathbf{x} - \mathbf{x}^b] = 0, \quad (2.2)$$

where,

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = 2 \mathbf{A} \mathbf{x},$$

if \mathbf{A} is symmetric, as is the case for covariance matrices. In general, \mathcal{H} is a non-linear function; however, it is possible to expand this function about \mathbf{x}^b using a Taylor series. A truncated Taylor expansion of $\mathcal{H}(\mathbf{x}^b)$ yields the relationship

$$\left. \frac{\partial}{\partial \mathbf{x}} \mathcal{H}(\mathbf{x}) \right|_{\mathbf{x}^b} [\mathbf{x} - \mathbf{x}^b] = \mathbf{H} [\mathbf{x} - \mathbf{x}^b] = \mathcal{H}(\mathbf{x}) - \mathcal{H}(\mathbf{x}^b), \quad (2.3)$$

where \mathbf{H} is a matrix of size $L \times N$ that represents a linearized map of \mathcal{H} about \mathbf{x}^b and thus transforms a model state vector into observation space.

The Kalman update equation, which determines the optimal weight that should be given to a short-term forecast and observations, is found by substituting (2.3) into (2.2) and performing some algebraic manipulations, which are excluded here for brevity (Kalman and Bucy 1961). The analysis state is given by

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathcal{H}(\mathbf{x}^b)), \quad (2.4)$$

with

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1}. \quad (2.5)$$

This equation indicates that the analysis is a linear combination of the background forecast and observation information, where \mathbf{K} is the weight given to the new observation information and also determines how to spread it to the model grid points. When the observation (forecast) has comparatively larger errors, \mathbf{K} will be larger (smaller), and thus the new information from observations will be weighted more (less). For a single observation, \mathbf{R} becomes a scalar, the variance in the model estimate of the observation ($\mathbf{H} \mathbf{P}^b \mathbf{H}^T$) reduces to a scalar, and $\mathbf{P}^b \mathbf{H}^T = \mathbf{x}^b (\mathbf{H} \mathbf{x}^b)^T$ is the covariance between the model estimate of the observation and each model state variable. In these scenario, \mathbf{K} is the linear regression coefficient between the observation and state variables where the new information from observations is the independent variable and the analysis is the dependent variable.

In addition to determining the best estimate of the state, data assimilation also leads to a reduction in the average error. Let $E[.]$ denote the expected value of a quantity, and ϵ indicate the error in a state vector \mathbf{x} ($\epsilon = \mathbf{x} - \mathbf{x}^t$). Moreover, assume that the errors in all fields are unbiased, such that $E[\epsilon] = 0$, while $E[\epsilon \epsilon^T]$ defines an error covariance matrix. Although truth is not explicitly known, the analysis error statistics are found by subtracting the true state from (2.4) and adding and subtracting $\mathcal{H}(\mathbf{x}^t)$ to the right hand side

$$\epsilon^a = \epsilon^b + \mathbf{K}[\mathbf{y} - \mathcal{H}(\mathbf{x}^b) + \mathcal{H}(\mathbf{x}^t) - \mathcal{H}(\mathbf{x}^t)]. \quad (2.6)$$

Substituting \mathbf{x}^t for \mathbf{x} into (2.3) gives

$$\mathcal{H}(\mathbf{x}^t) - \mathcal{H}(\mathbf{x}^b) = \mathbf{H}(\mathbf{x}^t - \mathbf{x}^b) = -\mathbf{H}\epsilon^b.$$

Combining the above expression with (2.6) gives an expression for the linearized analysis error

$$\epsilon^a = \epsilon^b + \mathbf{K}[\epsilon^o - \mathbf{H}\epsilon^b] = (\mathbf{I} - \mathbf{KH})\epsilon^b + \mathbf{K}\epsilon^o, \quad (2.7)$$

where \mathbf{I} is the identity matrix (ones along the main diagonal, zeros everywhere else), and $\epsilon^o = \mathbf{y} - \mathcal{H}(\mathbf{x}^t)$. The analysis error covariance is then found by computing $E[\epsilon^a \epsilon^{aT}]$, under the assumption that the background and observation errors are uncorrelated ($E[\epsilon^b \epsilon^{oT}] = 0$). After performing some algebraic manipulations, the analysis-error covariance matrix is thus given by

$$\mathbf{P}^a = \mathbf{P}^b - \mathbf{KHP}^b. \quad (2.8)$$

For simplicity, consider a single observation; under this scenario $\mathbf{HP}^b\mathbf{H}^T$ and \mathbf{R} reduce to positive scalars σ^b and σ^o , which represent the variance in the model-estimate of the observation and observation error variance respectively. The analysis-error covariance due to assimilating a single observation is then

$$\mathbf{P}^a = \mathbf{P}^b - \frac{\sigma^b}{\sigma^b + \sigma^o} \mathbf{P}^b. \quad (2.9)$$

The coefficient on the second term on the right-hand side is bounded between zero and one, thus assimilating an observation reduces the magnitude of each element in the background-error covariance.

A background forecast for the next time where observations are available, $t + \delta t$, is found by integrating the analysis at time t (\mathbf{x}_t^a) forward using a model. The process of advancing the model forward can be written in the compact form

$$\mathbf{x}_{t+\delta t}^b = \mathcal{M}(\mathbf{x}_t^a) + \eta, \quad (2.10)$$

where \mathcal{M} represents the non-linear forecast model, and η denotes the model error term, which describes deficiencies in the model's formulation of atmospheric processes and is assumed to be Gaussian and uncorrelated in time. The background forecast error covariance matrix at time $t + \delta t$ is determined by linearizing \mathcal{M} about the

non-linear trajectory from t to $t + \delta t$ (\mathbf{M}) and computing $E[\epsilon_{t+\delta t}^b \epsilon_{t+\delta t}^{b\top}]$. Under the assumption that the model error and analysis state are uncorrelated, the background-error covariance at time $t + \delta t$ is

$$\mathbf{P}_{t+\delta t}^b = \mathbf{M} \mathbf{P}_t^a \mathbf{M}^\top + \mathbf{Q}, \quad (2.11)$$

where \mathbf{Q} is the model error covariance matrix, given by $E[\eta\eta^\top]$. Forecasts can have a significant model error component and thus should be taken into account. Several authors have proposed ways to account for the model error (e.g., Dee and DaSilva, 1998); however, this quantity is often difficult to define. This formulation allows the error covariances to be propagated along with the model itself, and thus will give a flow-dependent estimate of forecast errors, and optimally spread observation information during data assimilation. The combination of (2.4), (2.8), (2.10), and (2.11) provide an iterative procedure for generating an analysis of the atmospheric state at regular intervals.

For NWP applications, \mathbf{P}^b is often $O(10^6) \times O(10^6)$, thus it is computationally prohibitive to evaluate (2.11). Furthermore, in situations where non-linearity is strong, the Kalman filter may not provide an accurate analysis because it propagates the error covariance matrix with a linearized version of the model. Evensen (1994) proposed a Monte Carlo alternative to the Kalman filter, known as the *ensemble* Kalman filter (EnKF), that is meant to partially reduce the computational cost of the Kalman filter, but retain some of the favorable properties. In the EnKF, the error covariances are estimated from a finite number of ensemble forecasts generated from parallel data assimilation and short-range forecast cycles. Assume that an ensemble of forecasts exists that randomly sample the model background errors. This ensemble can be denoted by a matrix \mathbf{X} , whose columns contain the M ensemble state vectors. The flow-dependent background error covariance matrix can then be estimated from the ensemble via

$$\hat{\mathbf{P}}^b = \frac{1}{M-1} \mathbf{X}'^b \mathbf{X}^b{}^T, \quad (2.12)$$

where \mathbf{X}^b is the background forecast ensemble state matrix with the ensemble mean removed. Note that in the limit as the number of ensemble members goes to infinity, $\hat{\mathbf{P}}^b$ should approach \mathbf{P}^b . Since the error covariances are estimated from a finite number of ensemble members (typically $O(100)$), sampling errors can create spurious correlations between grid points, which need to be accounted for. Each of the ensemble members in the EnKF algorithm are updated independently with observations using (2.4) and $\hat{\mathbf{P}}^b$. The resulting analysis ensemble is then integrated forward to the next time when observations are available using the non-linear forecast model, whereby the entire process is repeated.

Several authors have proposed modifications to the algorithm described by Evensen (1994) that are designed to improve the performance of the EnKF. Burgers et al. (1998) showed that when each ensemble member is updated with the same observation value, the ensemble can lose variance with time, thus they suggest updating each ensemble member with a perturbed version of the observations. Adding random noise to the observations can produce errors in the analysis statistics because the perturbations do not fully sample the observation-error distribution (e.g., Whitaker and Hamill 2002), thus several techniques have been used that generate an analysis ensemble that is consistent with the analysis-error covariance matrix obtained from the Kalman filter, but does not require perturbed observations. These methods, such as the ensemble adjustment filter (Anderson 2001), ensemble transform Kalman filter (Bishop et al. 2001), and the serial ensemble square-root filter (Whitaker and Hamill 2002) update the ensemble mean via (2.4), but have different algorithms for determining the ensemble deviations. These three methods are all square-root filters (Bierman 1977) and will produce non-unique analysis ensembles that are consistent with the analysis-error covariance (Tippett et al. 2003). This dissertation uses the ensemble square-root filter of Whitaker and Hamill (2002), thus the algorithm is described in

more detail below.

A few modifications to (2.4) are required to assimilate with unperturbed observations. Whereas the ensemble mean is updated with (2.4), the deviations from the ensemble mean are updated with a “reduced” Kalman gain matrix $\widetilde{\mathbf{K}}$ via

$$\mathbf{x}'^a = \mathbf{x}'^b - \widetilde{\mathbf{K}}\mathbf{H}\mathbf{x}'^b. \quad (2.13)$$

If \mathbf{K} is used in (2.13), the analysis-error covariance, computed from $E[\mathbf{x}'^a\mathbf{x}'^{a\text{T}}]$, will not match the analysis-error covariance given by (2.8); therefore, a reduced Kalman gain is required that will give the correct analysis statistics. When uncorrelated observations are serially assimilated, Potter (1964) noted that

$$\widetilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}}}\right)^{-1} \mathbf{K}. \quad (2.14)$$

The quantity multiplying \mathbf{K} in (2.14) is thus a scalar between 0 and 1. The above equation states that in order to assimilate observations without perturbed observations, a reduced magnitude version of the traditional Kalman gain is used to update the deviations from the ensemble mean. Furthermore, the computation scales linearly with the number of observations.

2.2 Ensemble Sensitivity

The following section relates ensemble sensitivity analysis to the ensemble Kalman filter and to adjoint-based sensitivity analysis. In addition, equations for estimating how observations can change a forecast metric using ensemble sensitivities will also be described. Consider a small change to the initial conditions at time t_o , $\delta\mathbf{x}_o$; the change in the model state vector at time t can be determined by making the tangent linear approximation

$$\delta\mathbf{x}_t = \mathbf{R}_{t,t_o}\delta\mathbf{x}_o \quad (2.15)$$

where \mathbf{R}_{t,t_o} is the resolvent matrix, which maps the perturbation column vector at the initial time, \mathbf{x}_o to the perturbation column vector at a later time \mathbf{x}_t . Next, consider a forecast metric J , which is a function of the model state vector \mathbf{x} . For small changes about the control solution trajectory, the forecast metric can be estimated using a Taylor expansion about the control solution

$$\delta J \doteq J(\mathbf{x}_t + \delta \mathbf{x}_t) - J(\mathbf{x}) = \left[\frac{\partial J}{\partial \mathbf{x}_t} \right]^T \delta \mathbf{x}_t. \quad (2.16)$$

Using (2.15), and the algebraic properties of the transpose, this expression becomes

$$\delta J \doteq \left[\mathbf{R}_{t,t_o}^T \frac{\partial J}{\partial \mathbf{x}_t} \right]^T \delta \mathbf{x}_o. \quad (2.17)$$

The expression \mathbf{R}_{t,t_o}^T represents the adjoint of the tangent linear model, which will act to map the sensitivity gradient at time t backwards to yield a sensitivity gradient at the initial time

$$\mathbf{R}_{t,t_o}^T \frac{\partial J}{\partial \mathbf{x}_t} = \frac{\partial J}{\partial \mathbf{x}_o}, \quad (2.18)$$

Substitution of (2.18) into (2.17) yields an equation for how changes to the initial conditions will affect the forecast metric at a later time

$$\delta J = \left[\frac{\partial J}{\partial \mathbf{x}_o} \right]^T \delta \mathbf{x}_o. \quad (2.19)$$

Here $\frac{\partial J}{\partial \mathbf{x}_o}$ is the sensitivity of the forecast metric to the initial conditions; elements of the state vector for which $\frac{\partial J}{\partial \mathbf{x}_o}$ is large indicate where initial condition errors can grow rapidly during the forecast. Adjoint-based sensitivity techniques compute this quantity using the adjoint of a tangent linear model. Adjoint-based methods are often difficult to code, especially for moist processes, thus an alternative method of determining initial condition sensitivity is desired.

The focus of this dissertation is using an ensemble of independent samples of the initial and final state to statistically estimate how initial condition changes affect the

chosen metric. The relationship between ensemble and adjoint sensitivity is obtained by right-multiplying (2.19) by $\delta \mathbf{x}_o^T$ and taking the expected value (Ancell and Hakim 2007)

$$E \left[\delta J \delta \mathbf{x}_o^T \right] = \left[\frac{\partial J}{\partial \mathbf{x}_o} \right]^T \delta \mathbf{x}_o \delta \mathbf{x}_o^T. \quad (2.20)$$

Since $\left[\frac{\partial J}{\partial \mathbf{x}_o} \right]^T$ is a deterministic quantity about a control trajectory, this equation may be re-written as

$$\text{cov}(\delta J, \delta \mathbf{x}_o) = \left[\frac{\partial J}{\partial \mathbf{x}_o} \right]^T \mathbf{P}^a, \quad (2.21)$$

where cov is the covariance between the two arguments. Note that for both δJ and $\delta \mathbf{x}_o$, it is assumed that these vectors have zero mean. The expectation performed above is approximated using an ensemble of size M , and thus is subject to sampling errors. In addition, the above equation gives the relationship between the adjoint and ensemble sensitivity; the covariance alone does not give the linear relationship between the initial state and forecast metric. Multivariate linear regression is achieved by right multiplying (2.21) by \mathbf{P}^{a-1} yielding

$$\left[\frac{\partial J_e}{\partial \mathbf{x}_o} \right]^T = \text{cov}(\delta J, \delta \mathbf{x}_o) \mathbf{P}^{a-1}, \quad (2.22)$$

where $\frac{\partial J_e}{\partial \mathbf{x}_o}$ is the ensemble sensitivity vector, which to within sampling error is identical to the adjoint sensitivity vector $\frac{\partial J}{\partial \mathbf{x}_o}$. Implementation of (2.22) has practical problems, which are related to the rank deficiency of \mathbf{P}^a and the ability of computers to store this matrix in core memory.

Instead of trying to invert \mathbf{P}^a , ensemble sensitivity is defined as the univariate regression of the forecast metric on all degrees of freedom in the initial condition

$$\frac{\partial J_e}{\partial x} = \frac{\text{cov}(\mathbf{J}, \mathbf{X}_j)}{\text{var}(\mathbf{X}_j)} = \text{cov}(\mathbf{J}, \mathbf{X}_j) \mathbf{D}_j^{-1} \quad (2.23)$$

where \mathbf{X}_j is one row of the ensemble state matrix and represents the ensemble estimate of the state variable x , \mathbf{J} is the $M \times 1$ vector of forecast metric values estimated from the M ensemble members, and \mathbf{D}_j is the j th diagonal element of the matrix \mathbf{D} , which is comprised of the diagonal entries of \mathbf{P}^a and zeros everywhere else. Here the independent variable is an analysis state variable and the dependent variable is the forecast metric. The difference between these two types of regressions is expected to be small if the covariance matrix is dominated by the diagonal (variance) elements. For a given metric, this formulation can be rapidly evaluated and is thus computationally more efficient than adjoint sensitivity. Moreover, it is possible to relate (2.23) to adjoint sensitivity by combining it with (2.21)

$$\frac{\partial J_e}{\partial \mathbf{x}} = \mathbf{D}^{-1} \mathbf{P}^a \frac{\partial J_a}{\partial \mathbf{x}}, \quad (2.24)$$

where J_a is the adjoint sensitivity. This shows that these two sensitivity methods are thus related by the analysis-error covariance matrix.

Since ensemble sensitivity analysis combines data assimilation and sensitivity analysis, it is possible to estimate how observation assimilation changes a forecast metric value. When observations are assimilated, the change in the model state vector is given by (2.4), thus the change in the forecast metric expected value can be determined by substituting this equation into (2.19) for \mathbf{x}_o , (2.22) for $\frac{\partial J}{\partial \mathbf{x}_o}$, and expanding

$$\delta J = \mathbf{J}(\mathbf{H}\mathbf{x}^b)^T [\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}]^{-1} [\mathbf{y} - \mathcal{H}(\mathbf{x}^b)]. \quad (2.25)$$

This equation represents linear regression, where the independent variable is the innovation, $\mathbf{y} - \mathcal{H}(\mathbf{x}^b)$, the dependent variable is the forecast metric, and the “slope” is given by the covariance between the forecast metric and the model estimate of the observation, $\mathbf{J}(\mathbf{H}\mathbf{x}^b)^T$, divided by the covariance of the independent variables (innovation covariance). For a single observation, the innovation, innovation covariance and slope are all scalars, and the calculation can be evaluated rapidly. When the

metric is a function of the forecast state only, δJ represents a change in the metric; however, when the metric is a function of the forecast and verification, δJ represents the impact of the observation. Furthermore, comparison of (2.25) and (2.4) shows that the update in the forecast metric expected value has similar form to the update in a model state variable.

In addition to predicting how observation assimilation will change the expected value of a metric, it is also possible to estimate how observation assimilation will reduce the metric variance. Much like how the analysis and forecast-error covariances are calculated, the variance in the forecast metric before or after data assimilation is found by computing $E[\delta \mathbf{J} \delta \mathbf{J}^T]$. Substituting (2.19) for $\delta \mathbf{J}$ yields

$$\sigma_J = \left[\frac{\partial J}{\partial \mathbf{x}} \right] \mathbf{P} \left[\frac{\partial J}{\partial \mathbf{x}} \right]^T. \quad (2.26)$$

where \mathbf{P} is either the background or analysis-error covariance matrix. An expression for the change in the forecast metric variance associated with observation assimilation is found by subtracting the forecast metric variance before assimilation from the forecast metric variance after assimilation

$$\delta \sigma = \sigma_{J^a} - \sigma_{J^b} = \left[\frac{\partial J}{\partial \mathbf{x}} \right] (\mathbf{P}^a - \mathbf{P}^b) \left[\frac{\partial J}{\partial \mathbf{x}} \right]^T. \quad (2.27)$$

The difference between the analysis and background error covariance matrix is given by (2.8), thus substituting it into (2.27) and expanding \mathbf{K} gives the expression for the reduction in metric variance associated with observation assimilation

$$\delta \sigma = -\mathbf{J}(\mathbf{H}\mathbf{x}^b)^T (\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{x}^b \mathbf{J}^T. \quad (2.28)$$

For a single observation, this expression can be evaluated as a product of two scalars: the inverse of the innovation variance, $(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})^{-1}$, and the forecast-metric-observation-estimate covariance, $\mathbf{J}(\mathbf{H}\mathbf{x}^b)^T$. Furthermore, observe that (2.28) is negative definite since the righthand-side is proportional to the square of the forecast-

metric–observation–estimate covariance. Equations (2.23), (2.25), and (2.28) are used extensively in Chapters 4 and 5.

Chapter 3

REAL-TIME DATA ASSIMILATION SYSTEM

3.1 *Experiment Setup*

The following chapter describes the performance of a pseudo real-time ensemble Kalman filter (EnKF) system from 00 UTC 1 January 2005 to 00 UTC 1 January 2007. The University of Washington EnKF (UW EnKF) system generates a 90 member analysis ensemble every 6 hours (00, 06, 12, and 18 UTC) on the domain used by Dirren et al. (2007), which includes the eastern Pacific Ocean and western North America and is characterized by sparse *in situ* data and complex topography. This system uses version 2.0.3.1 of the Advanced Research version (ARW) of Weather Research and Forecasting (WRF), a non-hydrostatic primitive equation, mesoscale model on a numerical grid with 45 km horizontal grid-spacing and 33 vertical levels (Skamarock et al. 2005). The model uses the WRF 3-class microphysics scheme (Hong et al. 2004), Kain-Fritsch cumulus parameterization (Kain and Fritsch 1990), Mellor-Jamada-Janjic boundary-layer scheme (Janjic 2002) and the Noah land-surface model (Ek et al. 2003).

This system assimilates observations from Automated Surface Observing System (ASOS) stations, ships, fixed and drifting buoys, rawinsondes, Aircraft Communications Addressing and Reporting System (ACARS) and cloud motion vectors (Velden et al. 2005) serially using a square-root version of the EnKF (Whitaker and Hamill 2002). These observation types are used here because the model estimate of the observation can be obtained via a bilinear interpolation of model grid points. Satellite radiances, which compose a large fraction of the observations assimilated by many operational forecasting systems, have complex observation operators that are a func-

Table 3.1: Observation types and average number of observations assimilated during each forecast cycle by the UW EnKF system.

		Analysis Time			
Observation	type	00 UTC	06 UTC	12 UTC	18 UTC
Surface	Alt., u, v	430	420	420	440
Rawinsonde	u, v, T, RH	1000	0	1000	0
ACARS	u, v, T	1650	1390	740	1860
Cloud Winds	u, v	2030	1740	1670	1510
Total		5110	3550	3830	3810

tion of the temperature and moisture profile in a column and are thus not used. Table 3.1 summarizes the number and type of observations assimilated during each analysis time and Figure 3.1 shows their horizontal distribution. The largest number of observations are assimilated at 00 UTC when there are frequent takeoffs and landings near many of the airports along the west coast of North America and cloud motion vectors can be estimated from visible satellite images. Observation errors are assumed to be uncorrelated and are obtained from European Center for Medium Range Weather Forecasting (ECMWF) statistics. Assimilating erroneous observations can lead to significant analysis errors, thus observations are quality controlled prior to being assimilated by comparing the observation innovation ($\mathbf{y} - \mathcal{H}(\mathbf{x}_b)$) with the standard deviation of the model estimate of the observation; if the innovation is greater than four times the standard deviation of the model estimate, the observation is not assimilated.

Observations are pre-processed prior to assimilation to reduce the number of observations and remove observations that may contain large representativeness errors. The course horizontal resolution of this domain does not resolve many of the horizontally narrow valleys in this region, thus observations from surface stations are

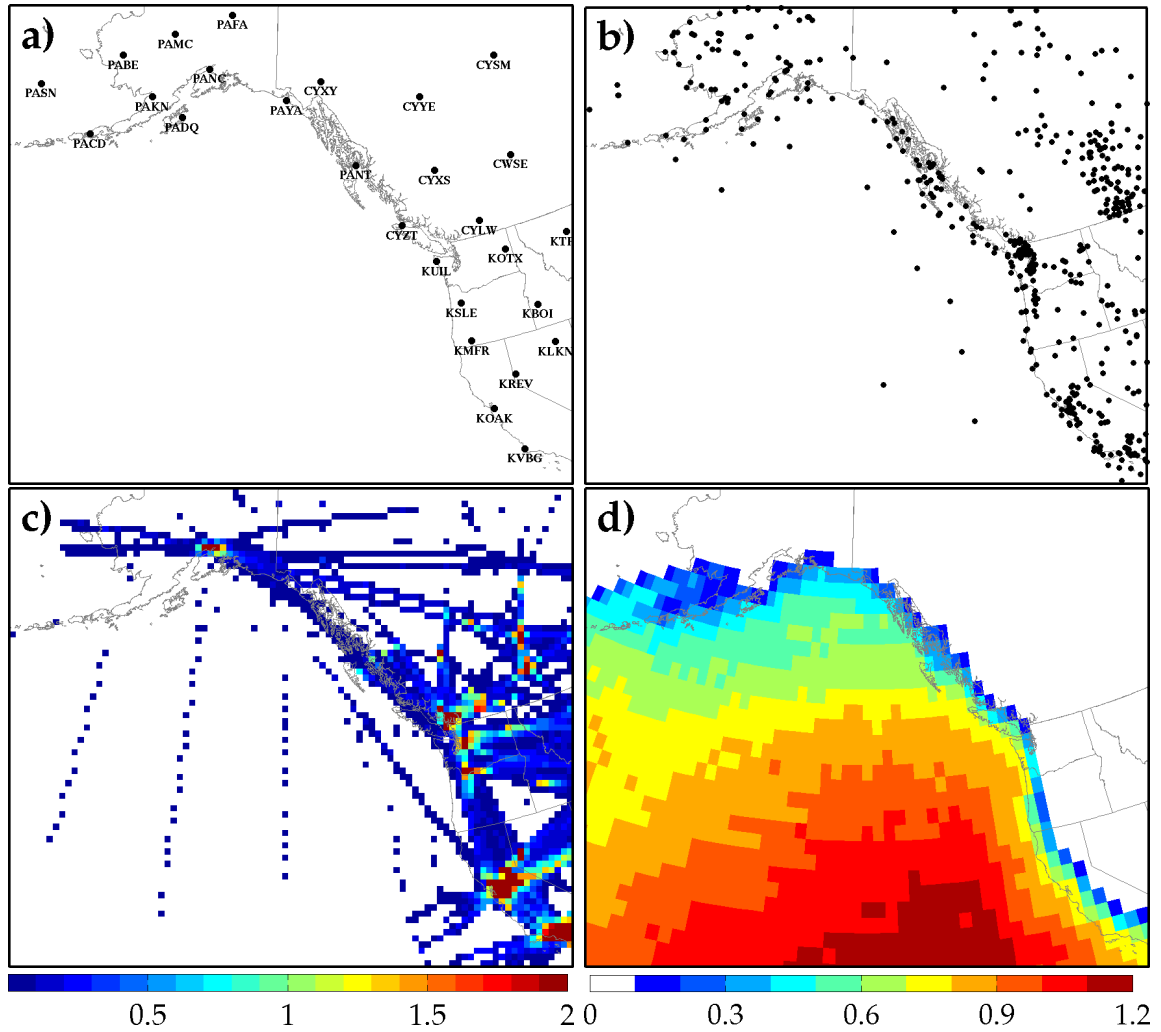


Figure 3.1: Distribution of (a) rawinsonde, and (b) fixed surface stations assimilated by the UW EnKF system. The observations shown in these two panels are also used to verify the forecasts. Panels c and d show the average density of ACARS and cloud wind vectors, respectively, at a particular location for each forecast cycle (units observations/cycle).

only assimilated if the difference between the station altitude and the model estimate of the station’s elevation is less than 300 m. For domains with coarse horizontal spacing, temperature and wind observations from land-based stations often contain large representativeness errors and are thus not assimilated. ACARS observations¹ are thinned by averaging all observations within one horizontal grid point and 25 hPa of each other. This procedure is performed to reduce the redundant information in the multiple takeoffs and landing reports in the area surrounding major airports (Fig. 3.1c). Cloud motion vectors estimated from geostationary visible, infrared and water vapor imagery are exclusively over the ocean; their density is a function of the scan angle and the percentage of times that clouds exist. Errors in individual cloud motion vectors are often quite large; therefore, cloud wind observations are “super-obed” by averaging all observations within 1° of latitude and longitude and 25 hPa of each other.

For any finite-sized ensemble, there is sampling error in the estimation of the background-error covariances, which are used to spread the information from observations to state-variables. These sampling errors can produce spurious noisy corrections to the state at long distances from the observation and artificially reduce the variance of the ensemble. One method of diminishing the influence of these long-distance correlations is to apply a distance dependent weighting function to the ensemble covariances that will restrict the influence of observations to nearby grid points (e.g., Houtekamer and Mitchell 1998, Hamill et al. 2001). In the UW EnKF system, the influence of observations is localized using the Gaspari and Cohn (1999) fifth-order piecewise rational function, given by their equation (4.10), which in this implementation, reduces to zero 2000 km from the observation location. Through trial-and-error testing, Dirren et al. (2007) found this radius to be the optimal trade-off between the desire to have a short enough radius to avoid far-field sampling errors and a long

¹Only ACARS reports within one hour of the analysis time are considered.

enough radius to avoid shocks in the model. Although it has been used in other EnKF applications (e.g., Whitaker et al. 2007), vertical localization is not employed because it can produce severe shocks to the model, possibly because of WRF’s vertical coordinate system (Dirren et al. 2007).

Covariance localization alone cannot completely overcome the effect of undersampling due to using a small ensemble, thus the deviations from the ensemble-mean are artificially boosted during each assimilation time. Moreover, this “covariance inflation” is also meant to be a crude parameterization of model error, which is not otherwise accounted for. If covariance inflation is not applied, the ensemble can lose variance with time and become over-confident in its estimate of the state. The ensemble is inflated using the covariance averaging technique of Snyder and Zhang (2003) where each ensemble member’s deviation from the ensemble-mean analysis is replaced by a weighted average of the deviations before and after data assimilation using

$$\mathbf{x}'_a = a\mathbf{x}'_a + (1 - a)\mathbf{x}'_b, \quad (3.1)$$

where \mathbf{x}'_a is an analysis member’s deviation from the ensemble mean, \mathbf{x}'_b is a background forecast member’s deviation from the ensemble mean, and a is the weighting factor for the analysis perturbations (0.2). This weighting factor is empirically determined by cycling a comparable WRF EnKF system on the same domain during October 2004². For the optimal value of a , the error in the ensemble mean should match the spread of the ensemble; the following section describes the procedure used to determine this. The covariance averaging technique is particularly advantageous in areas of inhomogeneous observations because it will only increase the ensemble spread in areas where observations are assimilated; if there are no observations in a particular region, the variance in the analysis ensemble will match the variance in

²This period is characterized by a variety of weather systems moving through the domain, and thus provides a good *a priori* test of how the UW EnKF system would perform during various weather regimes.

the background ensemble. Previous experiments with EnKF systems (e.g., Anderson and Anderson 1999, Whitaker and Hamill 2002) have inflated the ensemble by multiplying each ensemble member’s deviation from the mean by a small constant prior to observation assimilation. When a constant inflation factor (1.2) is applied to the October 2004 period, a majority of the ensemble members crashed after two weeks of observation cycling. The southwest corner of this domain is characterized by weak flow due to its proximity to the mean sub-tropical high, and because there are few observations, the ensemble spread in this region will continuously grow to unreasonable values solely due to covariance inflation.

Previous research with limited area ensembles (e.g., Nutter 2004, Torn et al. 2006) has shown that applying the same deterministic lateral boundary conditions to each ensemble member will cause the ensemble to lose variance over time, thus an ensemble of lateral boundary conditions is required. As of 1 January 2005, a 90 member global ensemble did not exist, thus an ensemble of lateral boundary conditions for this system is generated using the fixed covariance perturbation (FCP) technique of Torn et al. (2006). This method adds a randomly generated perturbation, consistent with a fixed error covariance model, to an ensemble-mean lateral boundary condition. Boundary condition perturbations are obtained by running the WRF VAR system (Barker et al. 2004) in “randomcv” mode, which generates balanced perturbations from the National Centers for Environmental Prediction (NCEP) covariance model.

Since these perturbations are randomly selected, they are not likely to have temporal continuity, which could cause shocks to the model. As a consequence, Torn et al. (2006) model the boundary perturbations as an autoregressive process in time (their eqn. 4). The autocorrelation coefficient for a six-hour forecast (0.4) is taken from Torn et al. (2006), who determined this value from a pseudo-global ensemble. To ensure that the variance of the boundary perturbations is representative of the error in the ensemble mean, the boundary perturbations are multiplied by 1.6 so that the standard deviation of the WRF VAR perturbations match the RMS error in 6-

hour forecasts from the NCEP Global Forecasting System (GFS). The ensemble-mean 6-hour forecast on the lateral boundaries is obtained from the 12-hour NCEP GFS forecast from the previous forecast cycle; using the forecast from the previous cycle allows the UW EnKF system to proceed without having to wait for the completion of the GFS forecast.

At 00 UTC and 12 UTC, 24-hour ensemble forecasts are obtained by advancing all 90 ensemble members forward in time. The forecasts are limited to 24 hours because for lead times greater than 24 hours, features that would have an effect the Pacific Northwest region of the United States moving faster than 35 m s^{-1} would be outside of the domain in the analysis³. An ensemble of lateral boundary conditions for these forecasts are obtained using the FCP technique where the scaling factor linearly increases from 1.6 for a six hour forecast to 2.4 for a 24-hour forecast. These scaling factors are also determined by matching the RMS error in 24-hour GFS forecasts to the standard deviation of the WRF VAR perturbations. Ensemble-mean lateral boundary conditions are obtained from the GFS forecast from the previous forecast cycle (i.e., at 12 UTC, the UW EnKF system uses the GFS forecast from 06 UTC).

The ensemble is initialized on 12 UTC 22 December 2004 by adding scaled, fixed covariance perturbations from the WRF-VAR system to the 36-hour GFS forecast initialized on 00 UTC 21 December. The perturbations are scaled by 1.9 prior to being added to this forecast. Initializing an EnKF system with a large amount of error and spread helps prevent the ensemble from under-weighting observations during the first few forecast cycles (Dirren et al. 2007). After assimilating observations for three days, the statistics of the ensemble come into equilibrium and there is little memory of the initial ensemble.

³This value is the average speed of Northern Hemisphere Rossby wave packets (Hakim 2003).

3.2 *Ensemble Verification*

Output from the UW EnKF system is verified against mandatory level wind, air temperature and dew-point temperature data from the 30 rawinsonde stations and the wind, temperature, humidity and pressure data from the 300 ASOS stations and fixed buoys in this domain. Figures 3.1a and b show the distribution of rawinsonde and surface stations, respectively, used for this verification; most of the stations are in the northern and eastern part of the domain, which is characterized by complex terrain. As a consequence, many of the rawinsonde observations are launched from an elevation well above mean sea-level, thus the number of verification points at 925 and 1000 hPa is smaller than in the middle troposphere. All verification statistics in this chapter are calculated for the two year period starting 00 UTC 1 January 2005.

Several different metrics are available for determining how well the ensemble spread matches the error in the ensemble mean. Two of the most popular methods involve computing the ratio of the RMS error in the ensemble-mean to either the RMS spread of the ensemble or the average RMS error of each ensemble member (Murphy 1988); however, when model forecasts are verified against imperfect observations or the model contains systematic biases, these ratios can give incorrect estimates of how well the ensemble spread matches the error. As a consequence, verification rank histograms (Anderson 1996, Hamill and Collucci 1997, Harrison 1995, Talagrand et al. 1997) are instead used to determine the skill of UW EnKF ensemble forecasts. This technique is particularly attractive because it accounts for observational errors and will expose persistent biases in a model forecast. Rank histograms are constructed by sorting the ensemble member's estimate of the verification value from low to high value, determining the rank of the verification value within the sorted ensemble, and repeating this for many values and times. If the ensemble is unbiased and drawn from the same distribution as the truth, the population in each of the ranks should be uniform. Excessive population in the outer (middle) ranks of the histogram indicate that truth

falls outside (inside) the ensemble too often, and thus the ensemble has too little (much) variance. Consistent biases in the ensemble forecasts show up as a sloped rank histogram; in these diagrams, larger population on the left (right) side of these figures indicates that the ensemble forecast values are consistently higher (lower) than truth. Hamill (2001) found that when observation error is not considered when computing the rank, the rank histograms can be misleading. Observation error is thus accounted for by ranking the observation relative to the sorted ensemble with random observational noise added to each ensemble member's estimate of the observation.

Figure 3.2 shows rank histograms for UW EnKF forecast of temperature as a function of pressure level and forecast hour. These figures are generated by independently computing the rank histogram for each pressure level and forecast hour. Regions of cold (warm) colors indicate that the population of that rank is less than 80% (greater than 120%) the value of a uniform distribution. In general, there are only small differences, related to sampling, between the rank histograms among the various forecast hours, thus only the 24-hour forecasts are described. Below 700 hPa, the rank histogram shows lower (greater) population in the left (right) ranks, thus the temperature is consistently too cold at those levels. Above 500 hPa the opposite pattern is observed; the histogram has lower (greater) population in the right (left) ranks, which indicates the model's forecast of temperature at these levels is consistently too warm. Although these diagrams are dominated by the model biases, they do not show excessive population in either the inner or outer ranks.

Rank histograms of 24-hour ensemble forecasts of winds, geopotential height and dew-point temperature also show that the UW EnKF ensemble forecasts are characterized by model biases (Fig. 3.3); diagrams of other forecast hours are qualitatively similar and are thus not shown. For zonal wind (Fig 3.3a), the rank histogram has nearly equal population among all ranks below 500 hPa; however, above this level, there is lower (higher) population in the left (right) ranks, indicating the model is consistently underestimating the speed of the upper-tropospheric westerlies. In con-

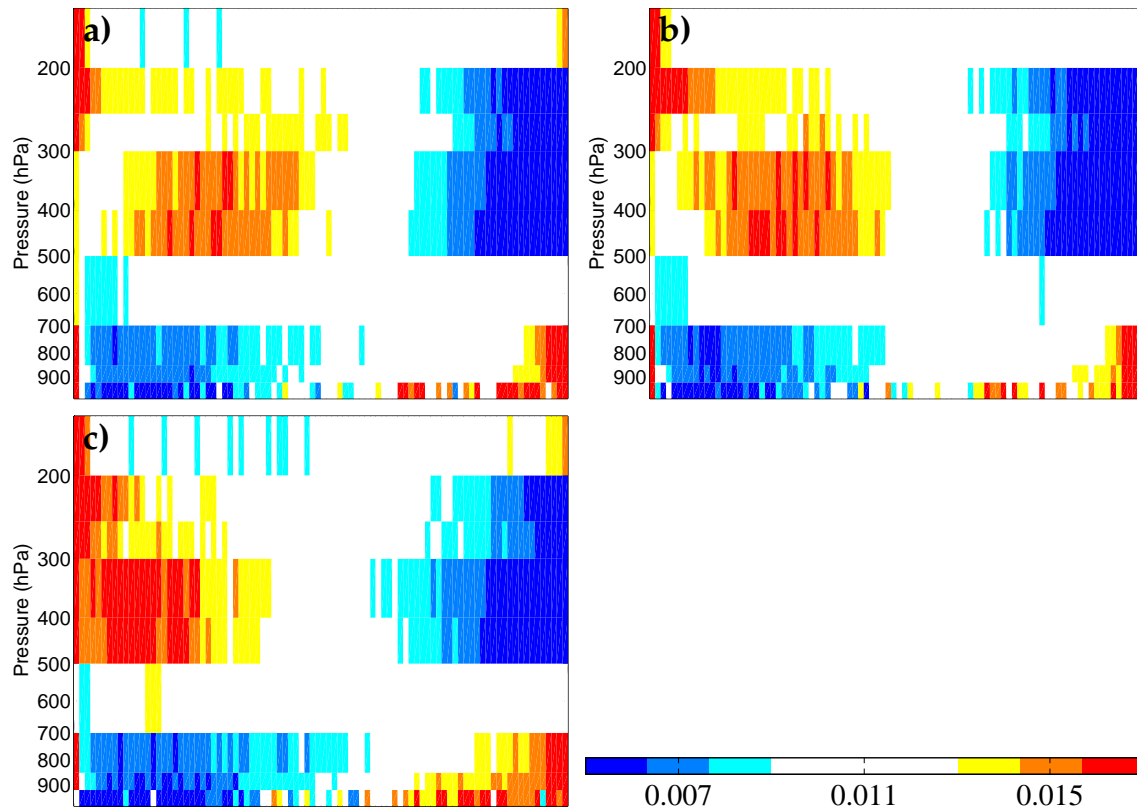


Figure 3.2: Rank histograms for UW EnKF (a) 6 hour, (b) 12 hour, and (c) 24-hour temperature forecasts at various pressure levels from 1 January 2005 to 1 January 2007 verified against the rawinsonde observations shown in Fig. 3.1a (shading). This figure is created by combining the rank histogram for temperature forecasts at each mandatory pressure level.

trast to the zonal winds, the meridional winds show approximately equal population among all ranks below 250 hPa, thus, at these levels, the forecasts generally contain the appropriate amount of variance (Fig. 3.3b); above 250 hPa, the ensemble's meridional winds are generally too high. Geopotential height rank histograms, as compared to wind and temperature figures, show a more complicated distribution (Fig. 3.3c). Between 250 and 850 hPa, the left (right) ranks have lower (higher) population than a uniform distribution, thus the ensemble is biased toward low heights; however, the lower population at middle ranks indicate that the ensemble also does not have enough variance. In contrast to the other fields, the excessive population in the middle ranks indicates that dew-point temperature forecasts below 500 hPa have too much variance (Fig. 3.3d). Above that level, the equal distribution among all bins shows that the ensemble has the appropriate amount of spread.

Similar to height and dew-point temperature, rank histograms for surface variables show that UW EnKF forecasts contain strong biases, and for some variables, excessive ensemble variance (Fig. 3.4). Six-hour altimeter forecasts show excessive population in the middle and outer ranks, which indicates that, at some locations, the ensemble contains too much variance and other areas contains too little variance (Fig 3.4a). At longer forecast hours, the distribution becomes skewed toward high ranks, thus, the model is more likely to have lower altimeter values than observations with increasing lead time. This low altimeter bias is consistent with the previously described low geopotential height bias and suggests that WRF model forecasts could be losing mass with time. Rank histograms for both components of the wind show that these fields contain too much variance at all forecast hours (Fig. 3.4b). For 2 meter temperature, the observations generally fall outside the ensemble too often and the model tends to be colder than the observations (Fig. 3.4c). This bias is not homogeneous among all times of the day; rank histograms of temperature forecasts valid at 00 UTC, 06 UTC, 12 UTC, and 18 UTC show that the model temperatures are too cold (warm) during the day (night), thus the surface parameterization has a weaker diurnal cycle than

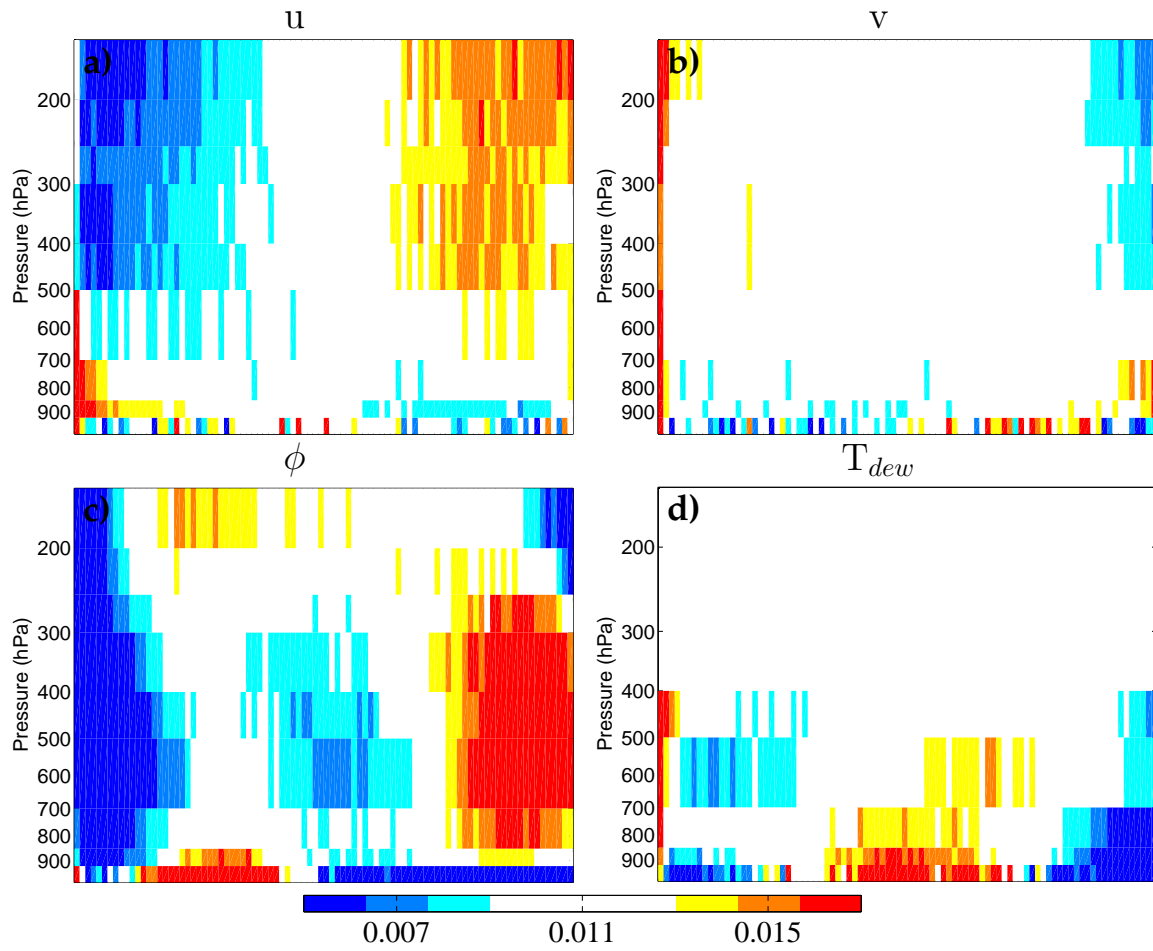


Figure 3.3: As in Fig. 3.2c, but for the UW EnKF 24-hour forecast of (a) zonal wind, (b) meridional wind, (c) geopotential height, and (d) dew-point temperature.

what is actually observed (not shown). Two dominate signals are visible in the rank histograms for 2 meter dew-point temperature forecasts; the ensemble has too much variance and is biased toward lower values (Fig. 3.4d).

The excessive population in the middle ranks of rank histograms for surface field forecasts actually results from the observation error rather than the actual spread in the ensemble; whereas observations of surface wind, air temperature, and dew-point temperature have assumed errors of 2.5 m s^{-1} , 2.5 K and 2.5 K , the RMS ensemble standard deviation for six hour forecasts is 1.3 m s^{-1} , 1.1 K and 1.5 K respectively. Surface parameterizations are a function of the temperature and wind at lower model levels, for which each ensemble member has a different value, and other parameters, such as the solar radiation and land use, that are fixed for each ensemble member. The fixed parameters will apply an equal forcing to each ensemble member and thus can act to constrain the variance in these fields.

3.3 Forecast Errors

In this section, the RMS errors in ensemble-mean forecasts of various model fields are computed with respect to rawinsonde and surface observations within the domain and compared to the RMS error in operational center forecasts in this domain over this two-year period. Errors are computed for 6, 12 and 24-hour forecasts from the NCEP GFS model, Canadian Meteorological Center (CMC) General Environmental Model, United Kingdom Meteorological Office (UKMO) unified model, and the Navy Operational Global Atmospheric Prediction System (NOGAPS); Table 3.2 gives a brief description of the resolution, data assimilation system and output used to compute the errors for each model. With the exception of NOGAPS and CMC, the effective horizontal resolution of each operational model is similar to the UW EnKF system; however the output files are often degraded from the native model resolution (1°).

Figure 3.5 shows that the RMS error in UW EnKF forecasts of temperature are comparable to the error in other operational forecasting systems. For all models, the

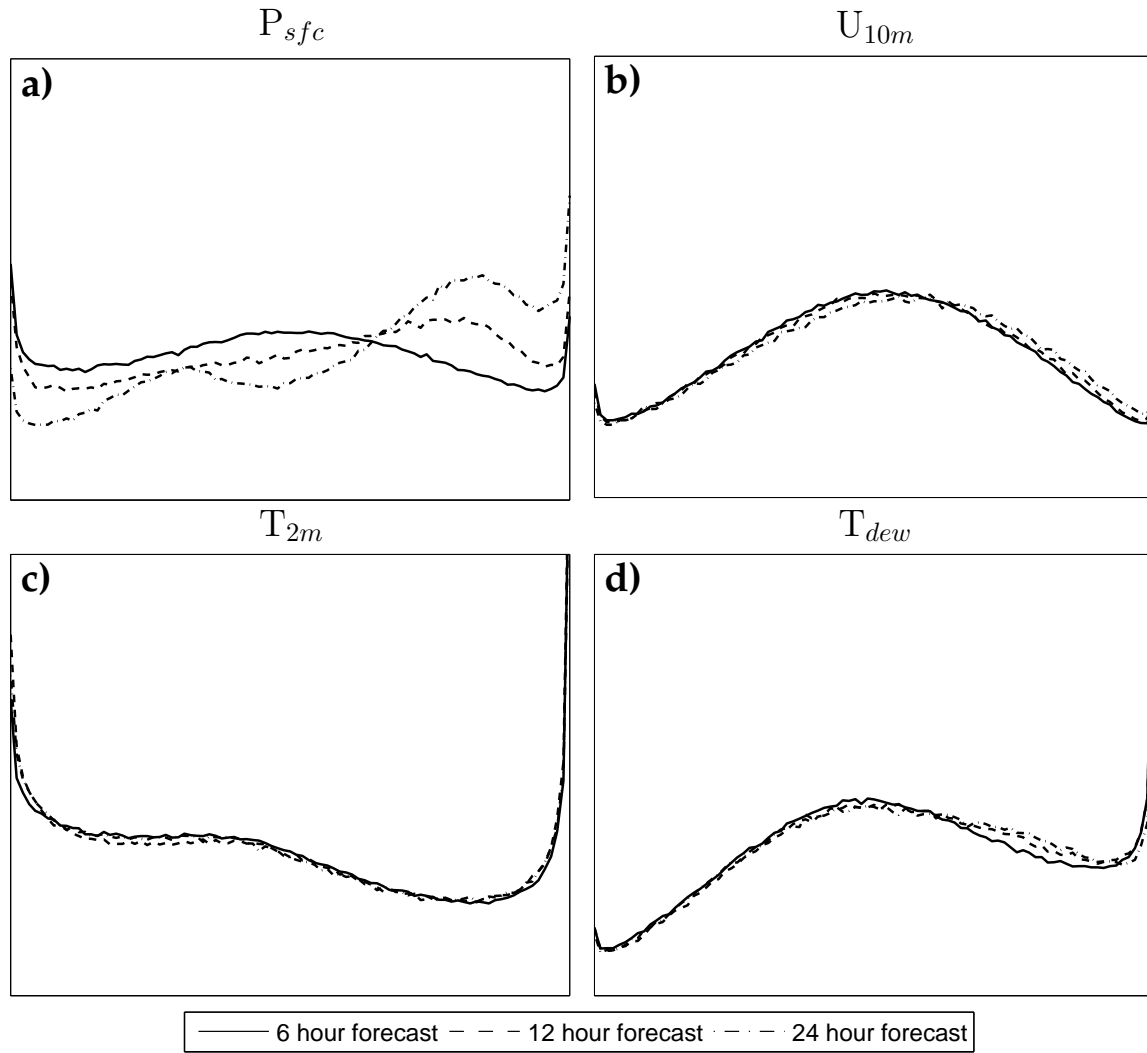


Figure 3.4: Rank histograms for UW EnKF 6, 12, and 24 hour forecast of (a) altimeter, (b) 10 meter zonal wind, (c) 2 meter temperature, and (d) 2 meter dew point temperature from 1 January 2005 to 1 January 2007 verified against the surface observations shown in Fig. 3.1b.

Table 3.2: Resolution, data assimilation technique, and output file resolution of the models used in this study.

Model	Resolution	Data Assimilation	Output Resolution
GFS	T382 (≈ 35 km)	3D-VAR	1°
CMC	≈ 100 km	4D-VAR	1°
NOGAPS	T239 (≈ 55 km)	3D-VAR	1°
UKMO	$0.5625^\circ \times 0.375^\circ$ (≈ 40 km)	4D-VAR	$5/6^\circ$ lon., $5/9^\circ$ lat.

largest errors are below 700 hPa and above 300 hPa, where the planetary boundary layer scheme and mis-specification of the tropopause, respectively, can lead to significant errors. In general, the GFS and CMC forecasts yield the smallest errors among the NWP systems described here. UW EnKF temperature forecasts have RMS errors that are approximately 0.3 K larger than GFS forecasts at all lead times; however, below 300 hPa, error in 12 and 24-hour temperature forecasts are nearly identical to UKMO and NOGAPS. This result is quite remarkable because the UW EnKF system only assimilates a small fraction of the observations used by these other operational centers since satellite radiance data are not considered. The potential reasons for why the errors are comparable include assimilating observations with flow-dependent error statistics, model dynamics and lateral boundary conditions; the role of each factor will be explored in the next section.

Throughout much of the troposphere, UW EnKF temperature forecasts have biases of up to 0.5 K; the model is too cold at lower levels and too warm above 700 hPa. To determine the skill of 6-hour UW EnKF 850 hPa temperature forecasts at different locations in the domain, the RMS error and bias are computed for each station. These calculations show that the magnitude of error, and especially the bias, is proportional to the elevation of the station. Stations near sea-level have bias values less than 0.1 K; however, 6-hour UW EnKF temperature forecasts for rawinsonde stations

above 1000 m elevation, which include Fort Nelson (CYYE) and Edmonton (CWSE), are on average 1.5 K too cold. Hollingsworth and Lonnberg (1986) showed that the time-average correlation between the observation innovations ($\mathbf{y} - \mathcal{H}(\bar{\mathbf{x}}_b)$) at two different locations can indicate systematic errors in the model, since the observations themselves are not correlated. Whereas the correlation between 850 hPa temperature innovations between most stations is zero, the correlation for the Fort Nelson and Edmonton stations is 0.75, thus the model boundary layer parameterization may contain problems for high elevation stations downstream of mountain ranges. At 300 hPa, each station has a negative temperature bias of 0.4 K; however the innovations are uncorrelated at this level. Although observation assimilation acts to remove the model bias, the use of covariance averaging to inflate the ensemble members can re-introduce the bias to the model. The extent to which this is true will be explored in the next section where UW EnKF forecasts are compared to WRF forecasts initialized from GFS analyses.

Similar to temperature, UW EnKF forecasts of other tropospheric variables have RMS errors comparable to the other operational centers described here (Fig. 3.6); for brevity, only the 24-hour forecast errors are presented. Errors in UW EnKF forecasts of zonal and meridional winds (Fig. 3.6a, b), which increase from 4 m s⁻¹ at 850 hPa to 6 m s⁻¹ at 300 hPa, are 1 m s⁻¹ greater than GFS forecasts, but identical to those obtained for UKMO forecasts. Whereas the bias in meridional wind forecasts is less than 0.7 m s⁻¹, the zonal wind bias is greater than 1 m s⁻¹ above 500 hPa. Computing the bias for each station shows that UW EnKF zonal wind forecasts are too slow at all locations; however, the observation innovations are uncorrelated. For geopotential height, the error in UW EnKF forecasts are larger than any of the operational centers below 300 hPa; above that level, the errors are comparable to the UKMO and lower than NOGAPS forecasts (Fig. 3.6c). From 925 hPa to 200 hPa, the WRF geopotential height forecasts are up to 10 m too low and, in a qualitative sense, seem to be hydrostatically related to the temperature bias in Fig 3.5d. In

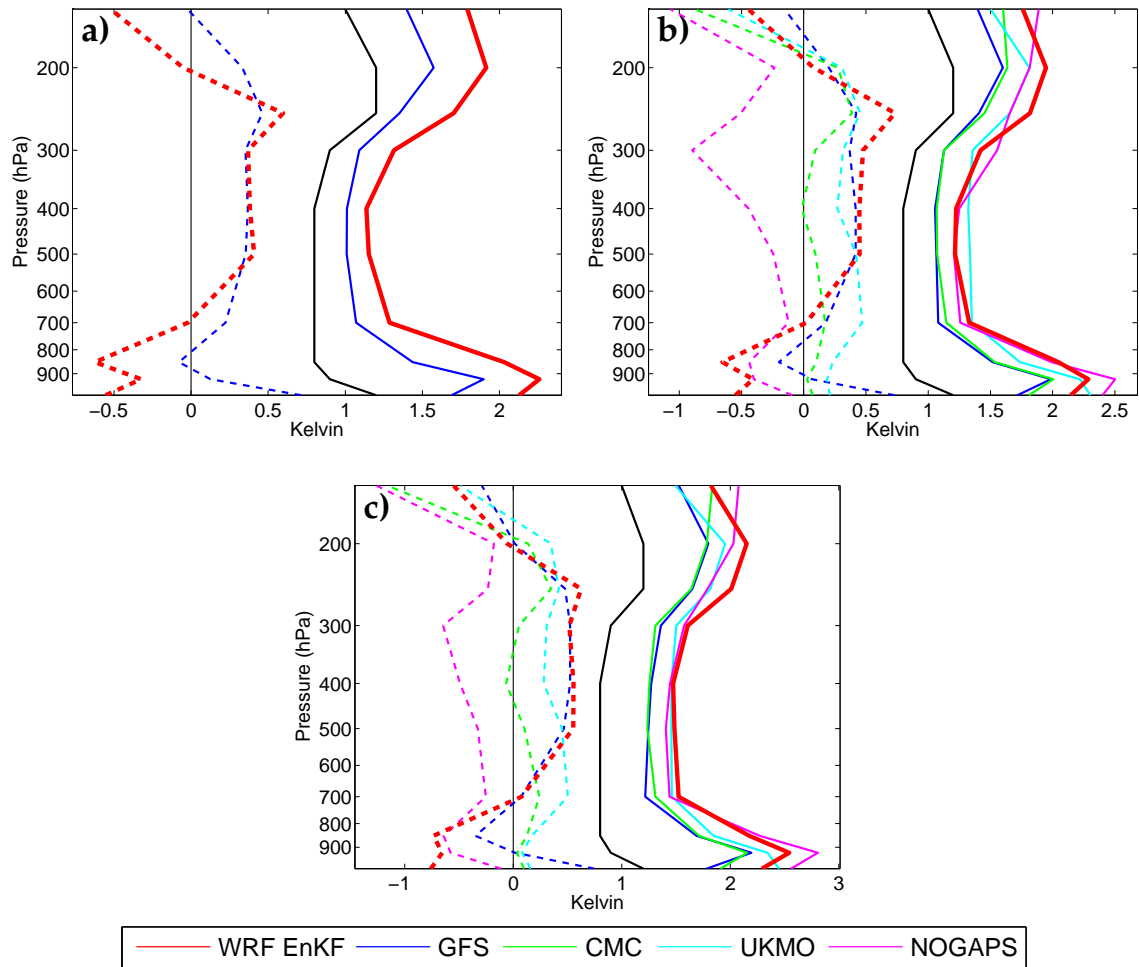


Figure 3.5: RMS error (solid) and bias (forecast-observation, dashed) in (a) 6 (b) 12 and (c) 24-hour temperature forecasts over the UW EnKF domain from 1 January 2005 to 1 January 2007 verified against rawinsonde observations. The black line denotes the ECMWF rawinsonde observation error variance used during data assimilation.

addition, the 10 m low bias is consistent among all stations and the innovations are highly correlated, thus the geopotential height bias is a systematic problem in WRF. UW EnKF dew-point temperature forecast errors increase from 3 K at 1000 hPa to 6 K at 500 hPa (Fig. 3.6d); quite surprisingly, the errors are less than or equal to all other models, especially in the mid-troposphere. Moreover, 24-hour forecasts of dew-point temperature from the UW EnKF system show the largest improvement over GFS forecasts for stations along the North American coast. Potential reasons for why this occurs will be explored in the next section.

Previous implementations of the EnKF have been characterized by a decrease in ensemble spread with increasing time, and suggests that the analysis error does not project onto the unstable modes of the model (Houtekamer et al. 2005, their Fig. 8). To determine the rate of error growth in UW EnKF forecasts, the RMS error in the ensemble-mean forecast and the RMS spread in various fields and vertical levels is computed as a function of forecast hour. Fig. 3.7 shows that for temperature, wind, and height forecasts, the ensemble spread, and to a lesser extent, the ensemble-mean error, increase by 50% over the course of 24 hours. Although the ensemble spread is generally less than the ensemble-mean error, these two quantities are growing at the same rate. In contrast to the other fields, the increase in the error and spread in dew-point temperature are more modest; while there is a 15% increase for 500 hPa forecasts, the 850 hPa and 300 hPa forecasts show an increase of 5-10% (Fig. 3.7d). Although the growth rate for dew-point temperature is smaller than other fields, it is consistent with GFS forecasts.

The reason for the difference in error growth between the UW EnKF system and the Houtekamer et al. (2005) setup may result from how each implementation parameterizes model error and how that projects onto the growing modes of the model. Houtekamer et al. (2005) account for model error by adding a random perturbation drawn from the CMC 3D-VAR background error covariance matrix to each analysis ensemble member. Observations significantly reduce the variance in the CMC system,

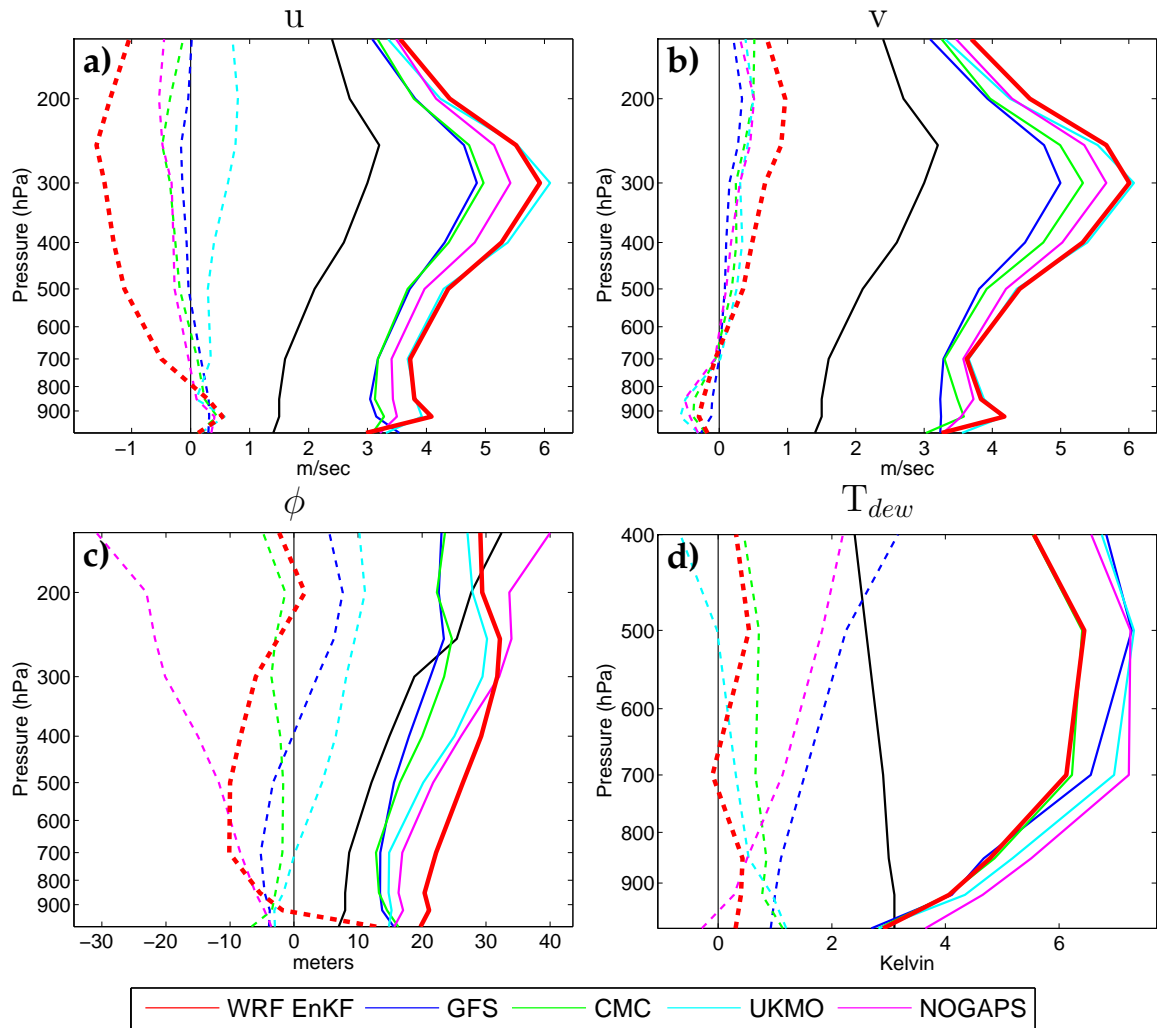


Figure 3.6: As in Fig. 3.5c, but for the 24-hour forecast of (a) zonal wind, (b) meridional wind, (c) geopotential height, and (d) dew-point temperature.

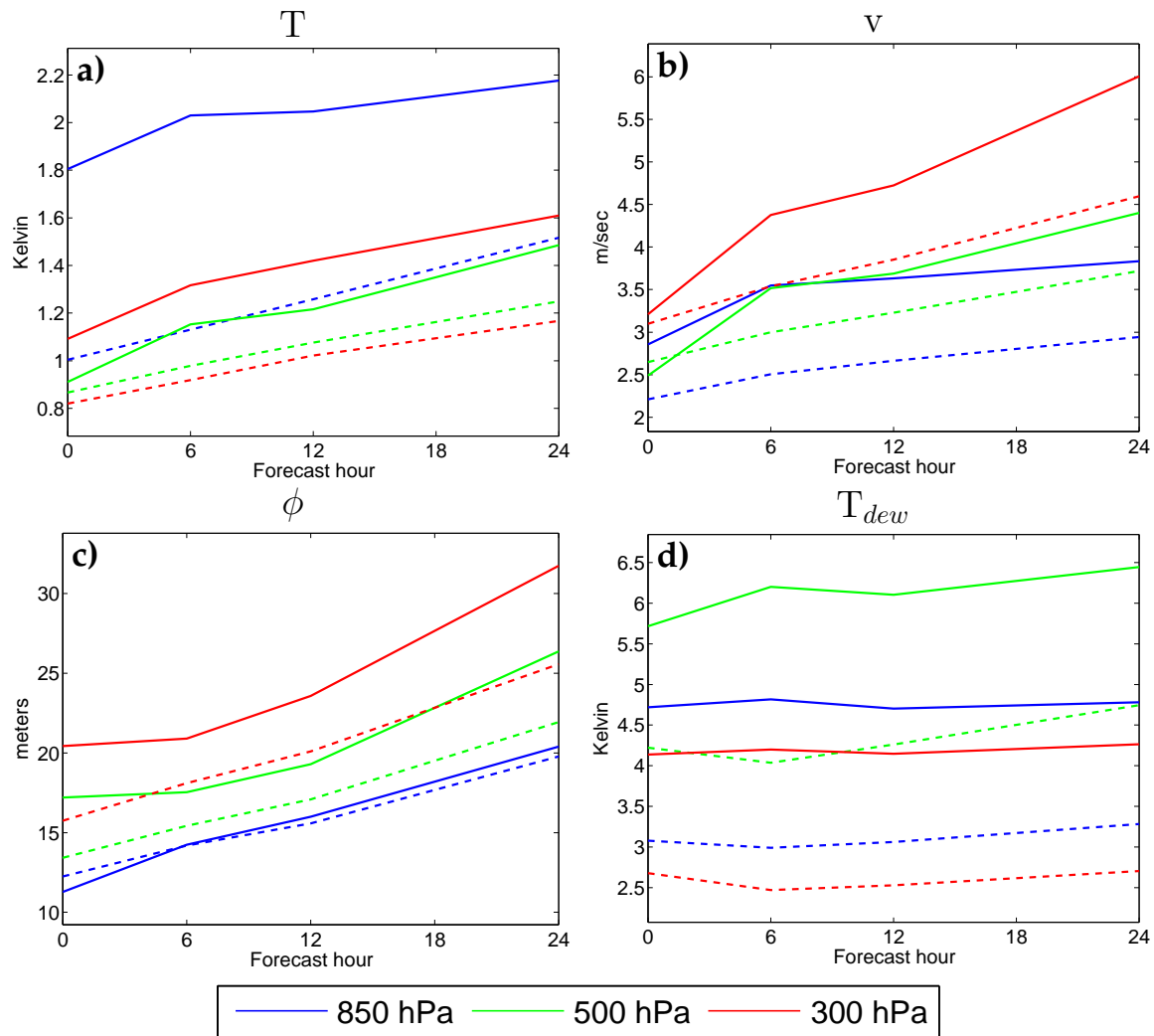


Figure 3.7: RMS error in the ensemble mean (solid line) and RMS ensemble spread (dashed line) in UW EnKF forecasts of (a) temperature, (b) meridional wind, (c) geopotential height, and (d) dew-point temperature as a function of forecast hour from 1 January 2005 to 1 January 2007 verified against rawinsonde observations.

thus a large fraction of the analysis spread is due the model error parameterization. They conclude that, during the first 24 hours, the ensemble perturbations are projecting onto the decaying modes of the model; however, for longer lead times, the amplitude of the growing modes start to dominate the amplitude of the decaying modes and the ensemble variance grows. Recall that the UW EnKF system parameterizes model error via covariance averaging, where the largest weight is given to the prior forecast's deviation from the mean. As a consequence, the analysis perturbations partially project onto the growing modes of the model during the previous six hours and thus the ensemble variance will increase during the first few time steps.

With the exception of altimeter setting, the bias and error in the various surface fields do not show much difference among all forecast hours (Fig. 3.8). The error in UW EnKF altimeter forecasts increases from 2 to 2.7 hPa from the 6 to 24-hour forecasts, which is 0.6 hPa greater than GFS forecasts, but is still up to 0.5 hPa smaller than NOGAPS and UKMO forecasts (Fig. 3.8a). In contrast to the other models, the bias in UW EnKF altimeter forecasts increase from -0.2 hPa for 6-hour forecasts to -0.7 hPa for 24-hour forecasts, and is consistent with the model losing mass in time. For 10 m wind, 2 m air temperature, and 2 m dew-point temperature, the RMS error in UW EnKF forecasts is approximately 2.8 m s^{-1} , 3.7 K and 3.9 K, respectively, at all forecast hours and is similar to or smaller than what is obtained from CMC, UKMO and NOGAPS forecasts. Errors in parameterized surface fields can result from errors in the input dynamical fields and errors in the model formulation. Since the errors and spread in the input dynamical fields are increasing with time (Fig. 3.7), the relative constancy of surface field errors is suggestive of a problem in the model formulation.

3.4 Comparison with Control Forecasts

Several potential reasons exist for why the errors in the UW EnKF forecasts are comparable to other operational centers. One possibility is that the information

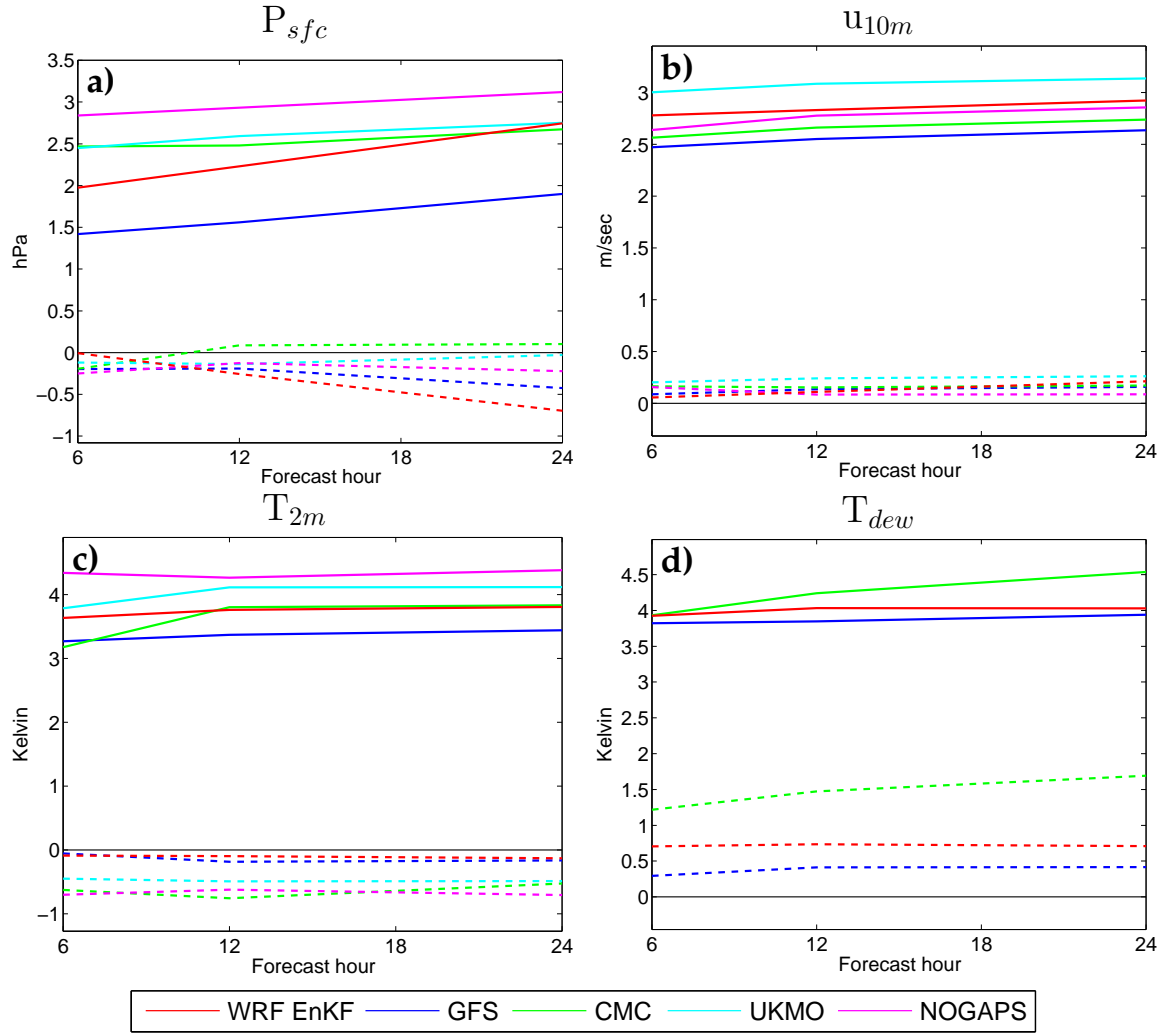


Figure 3.8: RMS error (solid line) and bias (forecast-observation, dashed) in various model forecasts of (a) altimeter, (b) 10 meter zonal wind, (c) 2 meter temperature, and (d) 2 meter dew-point temperature within the UW EnKF domain as a function of forecast hour. Forecasts are verified against the fixed surface observations shown in Fig. 3.1b.

obtained from the GFS lateral boundary conditions is constraining error growth in the domain. This hypothesis is tested by running an additional WRF forecast on this domain over the two year period that uses the same lateral boundary conditions as the ensemble-mean six-hour forecast, but is not updated with observations. Differences between this “no assimilation” and the ensemble-mean six-hour forecast are thus solely attributable to observation assimilation by the UW EnKF system.

Figure 3.9 shows that observation assimilation leads to systematically lower errors than the case where no data assimilation is performed; RMS errors in the ensemble-mean forecast are up to 50% smaller than the “no assimilation” forecast, thus the skill of the UW EnKF system forecasts is not solely due to the information coming from the GFS lateral boundary conditions. The largest (smallest) differences between these two forecasts is in the middle and upper (lower) troposphere where UW EnKF forecasts tend to have errors comparable to other operational centers. Furthermore, the bias in both UW EnKF and “no assimilation” forecasts are similar, except for height for which the bias magnitude is 50% of value yielded by the “no assimilation” forecast. This result shows that the mass bias is due to the model itself and that observation assimilation tries to overcome this problem.

The results presented above showing limited error reduction, and thus the smallest impact of observations near the ground extends to surface fields, whose dynamics are governed by parameterizations of boundary layer processes (Table 3.3). Whereas the error in 6-hour UW EnKF altimeter forecasts are 43% smaller than the “no assimilation” forecast, more modest differences of 18%, 15% and 15% are obtained for 10 meter winds and 2 meter temperature and 2 meter dew-point temperature, respectively. The difference between these forecasts further suggest that the errors in UW EnKF surface field forecasts are mostly due to the formulation of the surface layer parameterization scheme, rather than errors in the input low-level temperature, wind and moisture fields.

Finally, the benefit of performing data assimilation directly with the WRF model

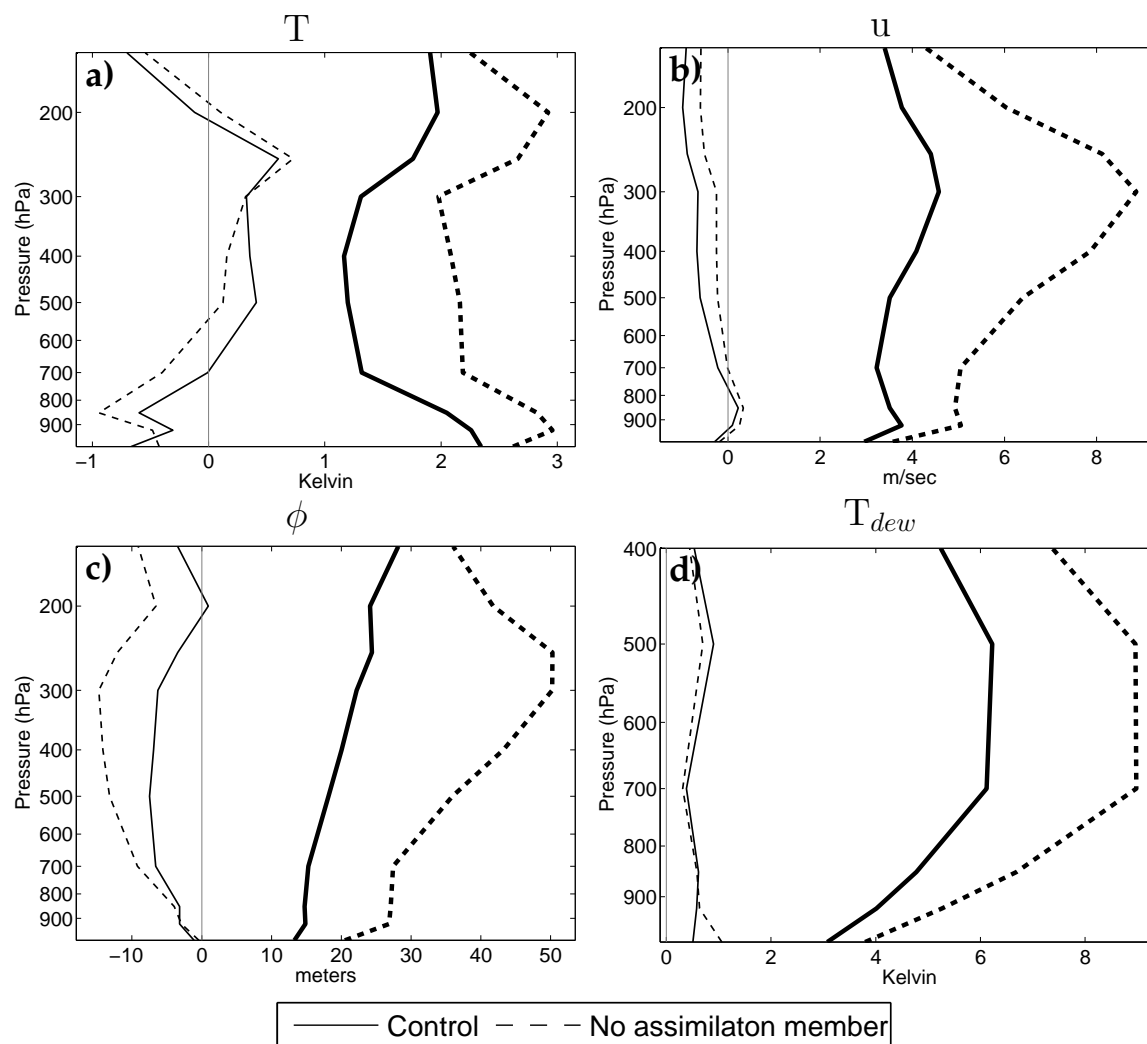


Figure 3.9: RMS error (thick lines) and bias (forecast - observation, thin lines) in the “no assimilation” and the 6-hour ensemble-mean forecast of (a) temperature, (b) zonal wind, (c) geopotential height, and (d) dew-point temperature verified against rawinsonde observations as a function of pressure from 1 January 2005 to 1 January 2007.

Table 3.3: RMS error (bias) in ensemble-mean, “no assimilation”, and the WRF-GFS forecasts of surface fields verified against fixed surface observations.

	6-hour forecast		24-hour forecast	
	Ens. Mean	No assim.	Ens. Mean	WRF-GFS
Altimeter	2.0 (0.0) hPa	3.5 (0.0) hPa	2.7 (-0.7) hPa	2.6 (-0.7) hPa
10 m u-wind	2.8 (0.1) m s ⁻¹	3.4 (0.1) m s ⁻¹	2.9 (0.2) m s ⁻¹	3.1 (0.2) m s ⁻¹
2 m temp.	3.6 (-0.1) K	4.1 (-0.0) K	3.8 (-0.1) K	3.8 (-0.4) K
2 m dew point	3.9 (0.7) K	4.6 (0.7) K	4.0 (0.7) K	4.2 (1.3) K

grid, rather than interpolating a global model’s analysis, is evaluated by comparing UW EnKF forecasts to a deterministic WRF forecast on the same domain. At 00 UTC and 12 UTC, a 24-hour WRF forecast on this domain is generated using the GFS analysis as initial conditions and the same lateral boundary conditions as the ensemble-mean forecast (denoted WRF-GFS). All model settings are kept the same, thus differences between the ensemble-mean and WRF-GFS forecasts are solely due to initial conditions of the two forecasts.

Although the initial conditions are generated by two different data assimilation techniques and different observation sets, Fig. 3.10 shows that the RMS error in 24-hour forecasts from the UW EnKF system are generally within 10% of the error in WRF-GFS forecasts, except for upper-tropospheric winds and geopotential height. Furthermore, comparison of Fig. 3.6 and Fig. 3.10 indicates that the difference between 24-hour WRF-GFS forecast errors and GFS forecast errors is less than 5%, except for geopotential height. These results seem to indicate that the bias toward low geopotential heights exists no matter what initial condition is applied, and thus do not result from covariance averaging or EnKF data assimilation.

One potential reason for why the RMS error in UW EnKF system dew-point temperature forecasts is lower than GFS forecasts relates to WRF’s treatment of micro-

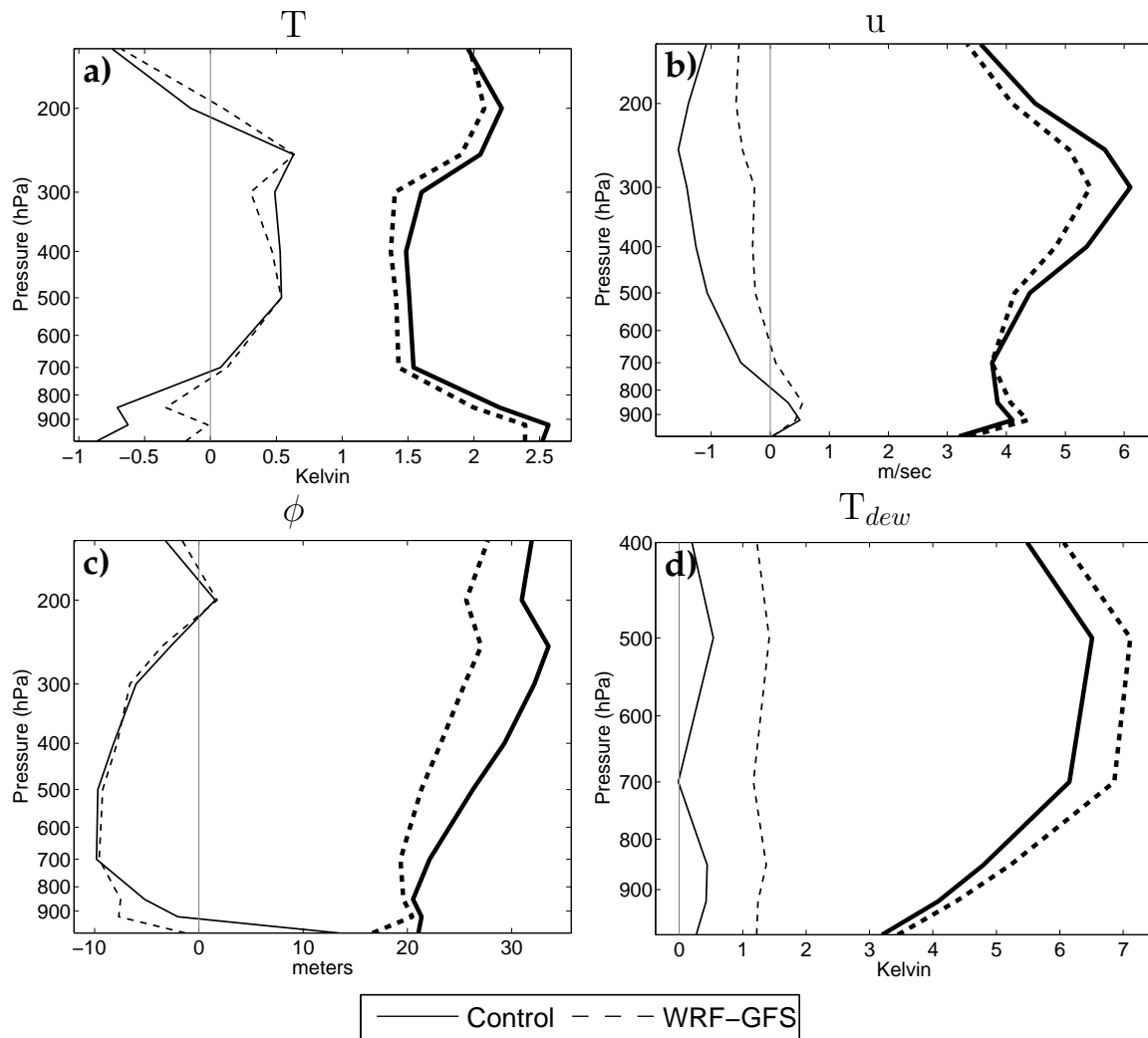


Figure 3.10: RMS error (thick lines) and bias (forecast - observation, thin lines) in 24-hour WRF-GFS and ensemble-mean forecasts of (a) temperature, (b) zonal wind, (c) geopotential height, and (d) dew-point temperature verified against rawinsonde observations as a function of pressure from 1 January 2005 to 1 January 2007.

physical processes; however, Fig. 3.10d shows that even when GFS initial conditions are used with WRF, forecasts from the UW EnKF system still have lower dew-point temperature errors, thus the initial conditions are responsible for the higher forecast skill. In the UW EnKF system, flow-dependent error statistics computed from the ensemble of short-term forecasts are used to spread the information from all observations to the water vapor field. In contrast, the fixed error statistics used in the GFS analysis system only allow observations that directly measure moisture, such as from rawinsondes or satellite estimates of column total precipitable water, to adjust the water vapor field (e.g., Parrish and Derber 1992).

To better understand what observation types may be updating the water vapor mixing ratio, time-average correlations between temperature, wind, and the water-vapor mixing ratio field for a column at the center of the domain are computed every six-hours during the two-year period (Fig. 3.11). This figure is generated by computing the correlation between the six-hour forecast of water-vapor mixing ratio at one pressure level (e.g., 500 hPa) at a single time to the six-hour forecast of temperature, wind and water-vapor mixing ratio at another pressure level (e.g., 300 hPa) at the same time using the ensemble data, and repeating for all possible combinations of variables and mandatory pressure levels. This calculation is then repeated for all analysis times during the two year period and averaged. Although there is little correlation between the water-vapor mixing ratio and the zonal component of the wind (Fig. 3.11a), time-average correlations in excess of 0.12 exist between the meridional wind and water-vapor field in the upper troposphere (Fig. 3.11b). Moreover, these levels correspond with places where the error in dew-point temperature forecasts from the UW EnKF system are 15% lower than GFS forecasts (Fig 3.6d). Recall that the largest improvement over GFS forecasts are for stations along the North American coast and that most of the upstream observations are cloud motion vectors (Fig. 3.1d), thus even though cloud winds have large errors (Bormann 2003), they could be very beneficial to water vapor forecasts. The average correlation between

mid-tropospheric temperature field and mid-tropospheric water-vapor mixing ratio is 0.16; however, non-zero correlations also extend above and below that level, thus, on average, upper-tropospheric temperature observations can impact the water-vapor mixing ratio throughout the column (Fig. 3.11c). For comparison, Fig. 3.11d shows the correlation between water-vapor mixing ratio at various pressure levels; by definition, the correlation between the mixing ratio at the same level is one. In general, the correlation between water vapor mixing ratio at two different levels is greater than 0.16 for pressure levels that are within 200 hPa of each other; however, in comparison to the lower-troposphere, the vertical extent of upper-tropospheric water vapor observations is more limited.

Comparison of the RMS error in UW EnKF and WRF-GFS surface field forecasts indicates that the forecast skill does not depend on the initial conditions (Table 3.3). WRF-GFS forecasts of altimeter (surface wind, dew-point) have slightly smaller (larger) errors than the corresponding UW EnKF forecasts. In contrast, the bias in UW EnKF temperature and dew-point temperature forecasts is smaller than WRF-GFS, thus the UW EnKF initial conditions seem to be more compatible with the surface parameterization schemes.

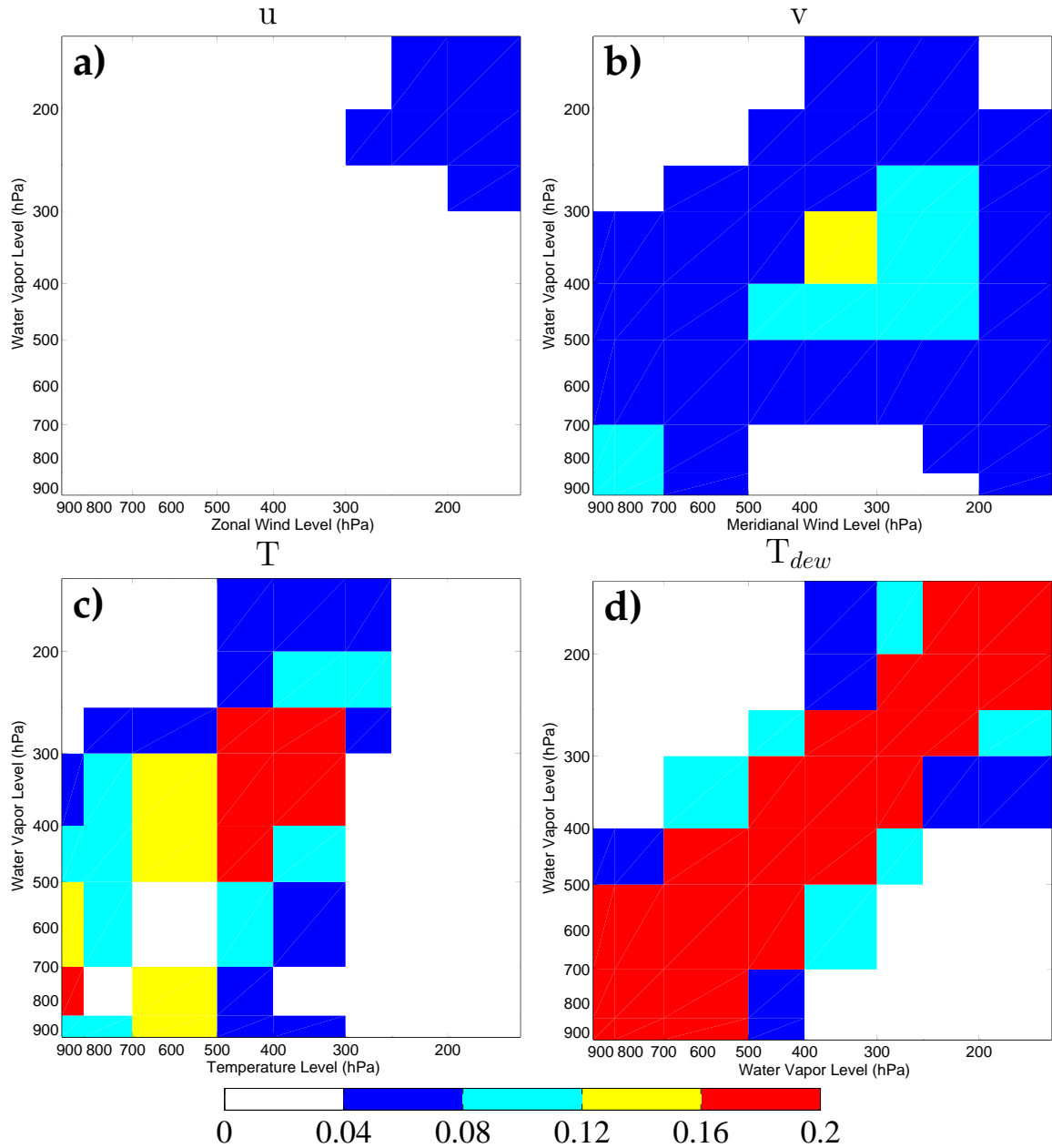


Figure 3.11: Mean correlation between the water vapor mixing ratio and (a) zonal wind, (b) meridional wind, (c), temperature, and (d) water vapor mixing ratio at each mandatory pressure level (925 hPa-150 hPa) averaged from 1 January 2005 to 1 January 2007.

Chapter 4

ENSEMBLE-BASED SENSITIVITY ANALYSIS

Ensemble-based initial condition sensitivity for the West Coast of North America is evaluated using data drawn from the University of Washington ensemble Kalman filter (UW EnKF) system during 1 January 2005 to 30 June 2005. This period is chosen because it contains several different weather regimes and samples both the highly variable winter season and the more tranquil summer season. Throughout this chapter, initial condition sensitivities will be computed for forecasts of the average sea-level pressure (SLP) and precipitation within a box that includes the western half of Washington (WA) state (“western Washington”). Furthermore, the primary focus of this chapter will be on 24-hour forecasts since that is the longest lead-time forecast the UW EnKF system generates. In general, the sensitivity technique is not limited to the metrics and forecast hours described here.

4.1 *Climatological results*

Data drawn from the UW EnKF system is used to determine the climatological sensitivity of pressure and precipitation averaged in a box over western Washington. Climatological sensitivity is defined here as the percentage of analysis cycles for which the ensemble sensitivity of the forecast metric with respect to an analysis grid point is different from zero at a certain level of confidence. Specifically, a state variable will produce a statistically significant change in the forecast metric if

$$\left| \frac{\partial J}{\partial x} \right| > \delta_s, \quad (4.1)$$

where δ_s is the confidence interval on the linear regression coefficient (e.g., Wilks

2005, section 6.2.5). For example, taking δ_s to be the 95% confidence interval (the value used for the climatological results given below), if (4.1) is satisfied, one can reject the null hypothesis that changes to x have no impact on the forecast metric with 95% confidence since a slope of zero would not be included in the confidence bounds. Climatological sensitivities are computed with this form to prevent a few cycles with large magnitude sensitivities from biasing the distribution. Regions with a high percentage of sensitive forecast cycles may be regarded as potential locations for siting new observations.

Figure 4.1 shows results for the sensitivity of the average SLP in the box over western WA to the SLP analysis as a function of forecast hour. For short-term forecasts (6, 12 hours), the region with the largest percentage of sensitive cycles is near the forecast metric box; however, at longer lead times, this region moves further to the west. Whereas the 6-hour forecast is most often sensitive to the analysis of SLP at (46°N and 125°W) 90% of the time, the 24-hour forecast has a maximum near (45°N, 132°W) during 44% of forecasts. This pattern qualitatively reflects the progression of weather systems from west to east during a forecast at a mean translation speed of 9 m s^{-1} , which is determined by computing the distance between the center of the metric box and the point of most frequent sensitivity. This value is roughly consistent with the phase speed of individual eddies in the Northern Hemisphere mid-latitude flow (Hakim 2003). In addition to the region of highest sensitivity moving to the west with lead time, it also becomes more diffuse due to the variety of locations weather systems can originate from.

Regions of persistent sensitivity in Figure 4.1 predict where additional SLP observations would have the greatest impact on the SLP forecast metric on average. Coincidentally, the location of maximum sensitivity for the 24-hour forecast is close to buoy 46005 (white dot). In fact, this buoy failed on 26 December 2004 and therefore observations from this location were not available during the time period of this experiment. This suggests that the absence of observations from buoy 46005 may

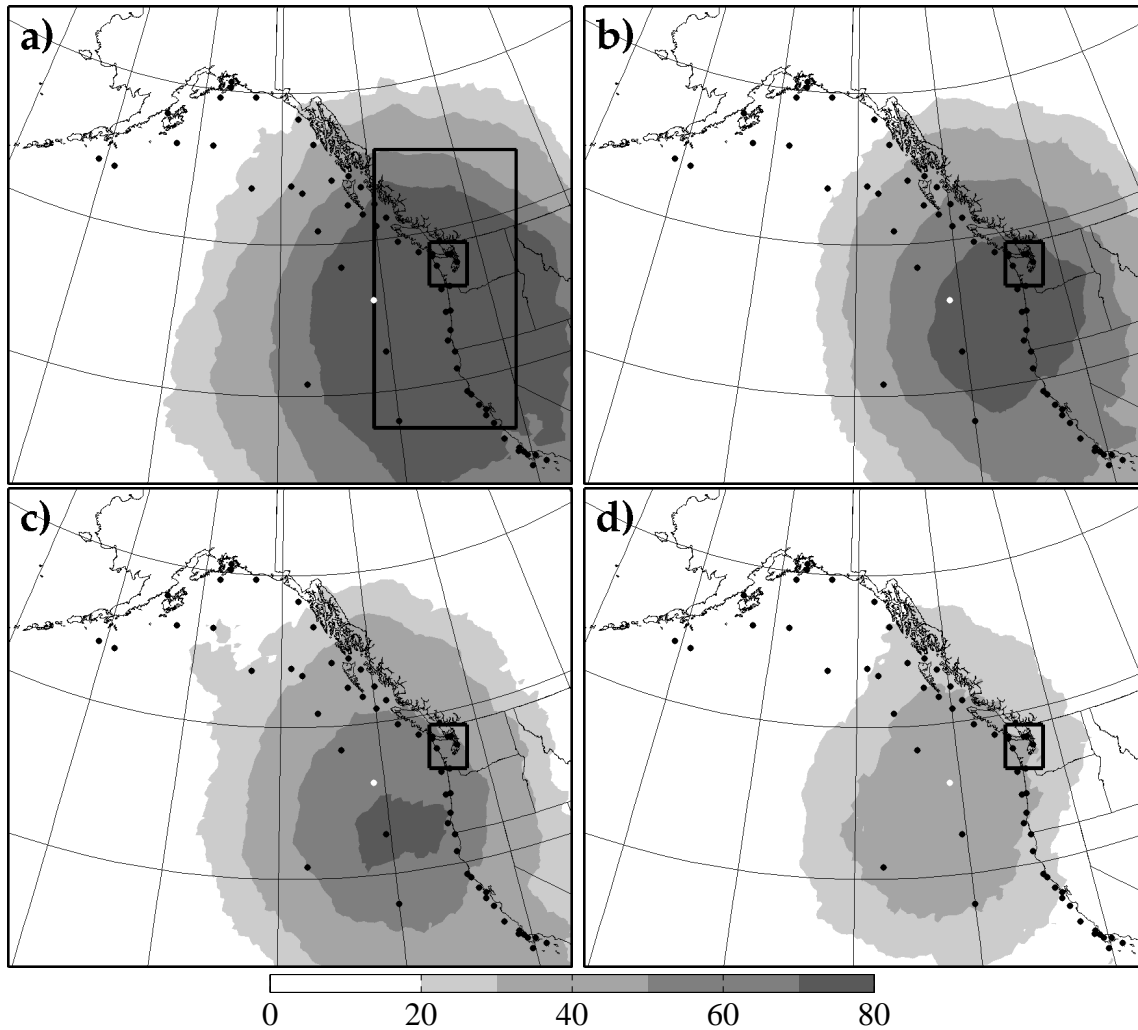


Figure 4.1: Percentage of forecast cycles with grid point sensitivity values that are statistically significant at the 95% confidence level for western Washington (a) 6, (b) 12, (c) 18, and (d) 24-hour SLP forecast to the SLP analysis. Forecasts are initialized at 00 UTC and 12 UTC from 1 January to 30 June. The forecast SLP is averaged over the region indicated by the smaller box in (a). Dots indicate the position of fixed buoys, and the larger box denotes the North American Coast (NAC) metric region used in Figs. 4.10, 4.12, and 4.13.

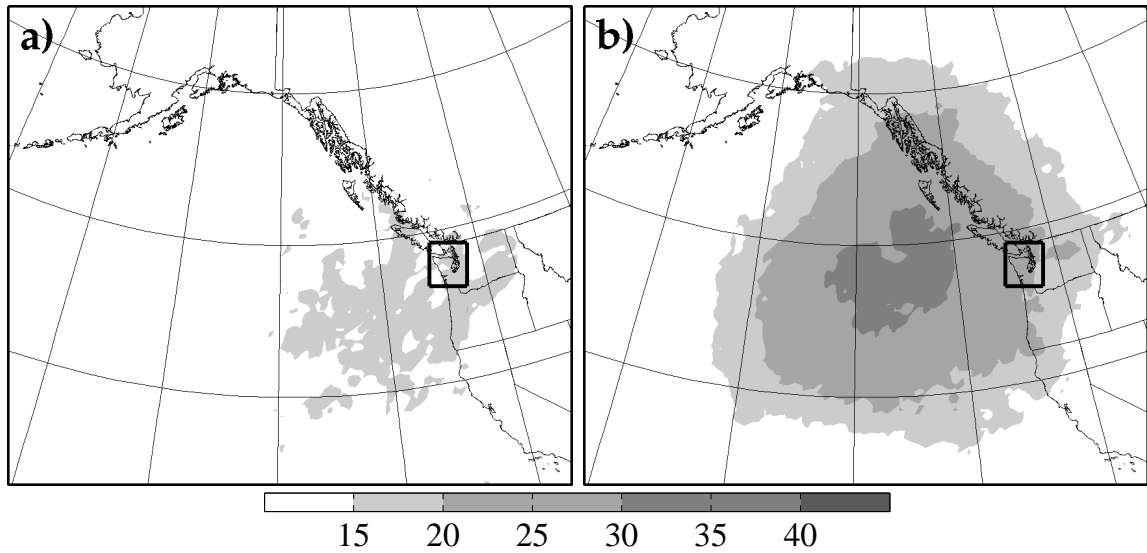


Figure 4.2: Percentage of forecast cycles with grid point sensitivity values that are statistically significant at the 95% confidence level for western Washington 24-hour SLP forecast to the (a) 850 hPa temperature, and (b) 500 hPa height analysis. Forecasts are initialized at 00 UTC and 12 UTC from 1 January to 30 June.

have adversely affected forecasts over western WA during these six months. In light of this possibility, this problem will be revisited later in this chapter where the change in the 24-hour SLP forecast due to the missing buoy is quantified by withholding a nearby buoy from the analysis and comparing the predicted and actual changes in the forecast metric.

The forecast SLP metric is also frequently sensitive to analyses of 850 hPa temperature and 500 hPa geopotential height. Figure 4.2a shows that there are two main sensitive regions for temperature; one to the south-west of Washington state near 43°N , 130°W , and another to the east of the metric box. For 500 hPa geopotential height (Fig. 4.2b), the forecast metric is sensitive more than 20% of the time to the region bounded by 40°N – 60°N and 120°W – 160°W . This region is located a few hundred kilometers upstream of the region of maximum SLP sensitivity, which reflects the moderate upstream tilt typical of baroclinic waves in the westerlies.

The second metric considered is the precipitation averaged in the box over western WA. Since sensitivity can only be determined when the forecast metric variance is nonzero, the percentage of sensitive cycles is computed with respect to the total number of cycles where the ensemble-mean, horizontally-averaged precipitation in the box exceeds 1 mm for forecast hour 18-24; 58% of all forecasts exceed this threshold. Sensitivity of this metric to SLP shows a maximum of 40% over the Pacific Ocean (Fig. 4.3a). Whereas the northern half is relatively well observed by the near-shore buoy network, the southern half is observed by fewer buoys.

For 850 hPa temperature analyses, the precipitation metric is sensitive more than 20% of the time to the southwest of western WA, with maximum sensitivity near 132°W (Fig. 4.3b). Precipitation forecast sensitivity to 500 hPa geopotential height is similar to other fields, with maximum values along 132°W (Fig. 4.3c). The main sensitive region for 500 hPa height is slightly upstream of the region of maximum SLP sensitivity, thus the sensitivity fields for this metric are also titled westward with height.

4.2 *Most sensitive cases*

Whereas the previous section showed how often a forecast metric is sensitive to changes in a state variable, composite averages are used here to determine the spatial pattern of sensitivity that occurs for forecasts having the largest sensitivity. These distributions show locations where small initial condition errors can, on average, lead to large changes in the metric. The most sensitive western WA SLP and precipitation forecasts are determined by computing a domain-average forecast sensitivity (*DAS*)

$$DAS = \frac{1}{N_h} \sum_{i=1}^{N_h} \left| \frac{\partial J}{\partial x_{SLP,i}} \right|, \quad (4.2)$$

where $x_{SLP,i}$ is the SLP at a grid point i , and N_h is the number of horizontal grid points. This norm is used to determine the most sensitive cycles because SLP is a column integrated quantity, and since the forecast metric is SLP, is expected to

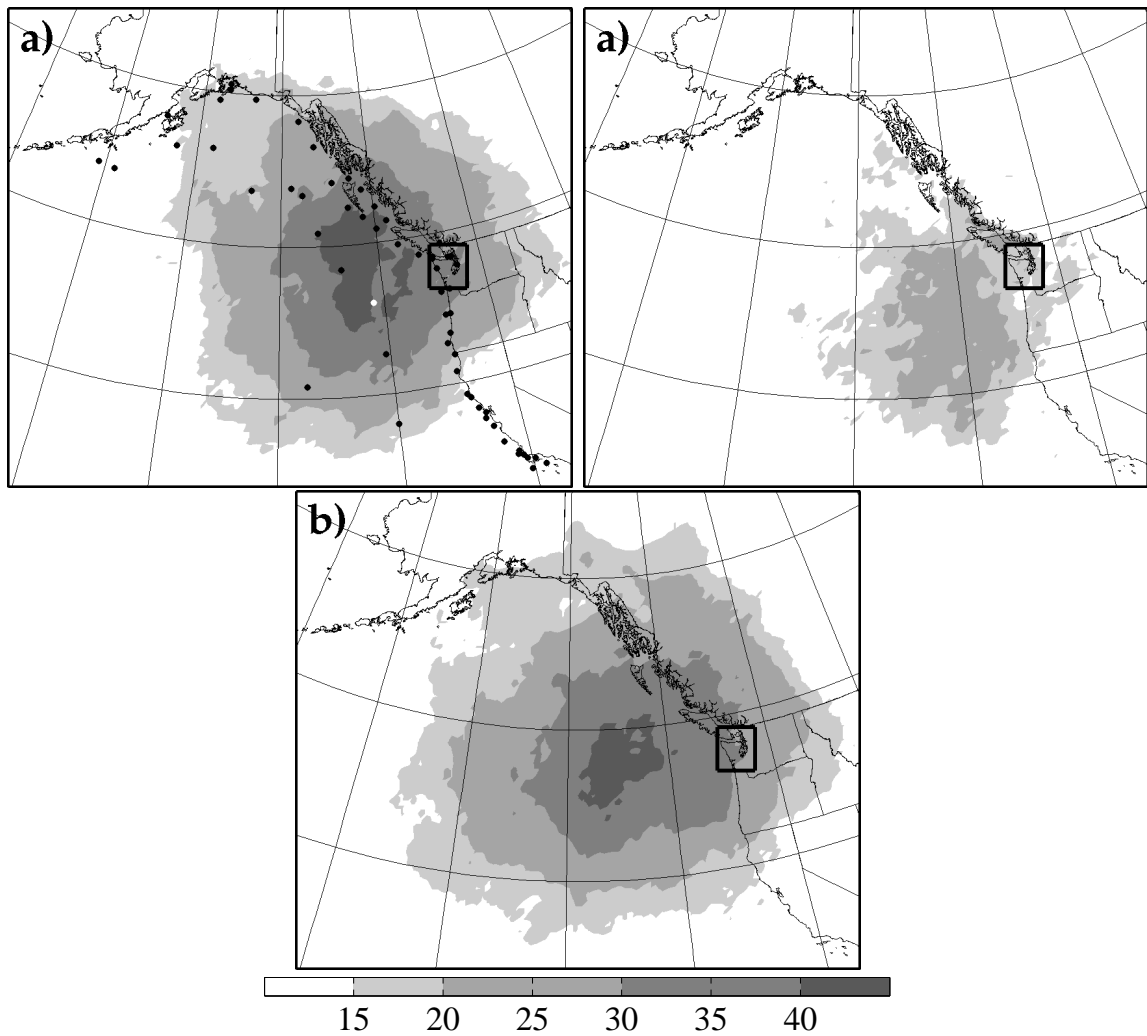


Figure 4.3: As in Fig. 4.1d and 4.2, but for the 24-hour forecast of precipitation averaged over the Western Washington region. Here the percentage of forecast cycles is computed with respect to the number of cycles where the precipitation in the box exceeds 1 mm for forecast hour 18-24.

have the largest sensitivity values. Composite patterns of forecast sensitivity for the 30 cycles¹ with the largest *DAS* values are calculated by:

$$\overline{\Delta J} = \frac{1}{N_t} \sum_{t=1}^{N_t} \left(\frac{\partial J}{\partial x_t} \sigma_{x_t} \right), \quad (4.3)$$

where σ_{x_t} is the standard deviation of \mathbf{x} at time t , and N_t is the number of cycles used in the composite. In the above equation, the sensitivity of a forecast metric to an analysis grid point is multiplied by the standard deviation at that grid point and averaged for all times considered. Multiplying $\frac{\partial J}{\partial x}$ by σ_{x_t} gives the change in J brought about by a one standard deviation change in x_t , and thus a quantitative comparison of how various analysis fields change J . Regions of high composite sensitivity indicate where additional observations could lead to the largest metric changes during the most sensitive forecasts.

Figure 4.4a shows the composite sensitivity of the 24-hour forecast of average SLP in the box over western WA to SLP analyses. Increasing (decreasing) x at a single grid point by one standard deviation within regions of largest sensitivity values implies a 0.9 hPa increase (decrease) in the forecast metric. The region of largest sensitivity is in an area characterized by few buoys at (47°N 135°W), and is nearly co-located with the region having the largest percentage of sensitive cycles (Fig. 4.1d).

For 850 hPa temperature, the average sensitivity is less coherent than for SLP, although sensitivity appears both east and west of the forecast-metric box (Fig 4.4b). Furthermore, this region of high sensitivity is to the north of the region of consistent sensitivity in Fig. 4.2a, thus while the minimum SLP is more often sensitive to the 850 hPa temperature to southwest of the metric box, the largest magnitude sensitivities are to the northwest. Increasing (decreasing) the temperature in the regions with the largest values by one standard deviation only leads to a 0.5 hPa decrease (increase) in the SLP in the box 24-hours later. For 500 hPa height, the SLP forecast

¹This number of cycles is chosen so that a few cycles with large sensitivities do not bias the horizontal distribution and magnitude of the sensitivity pattern.

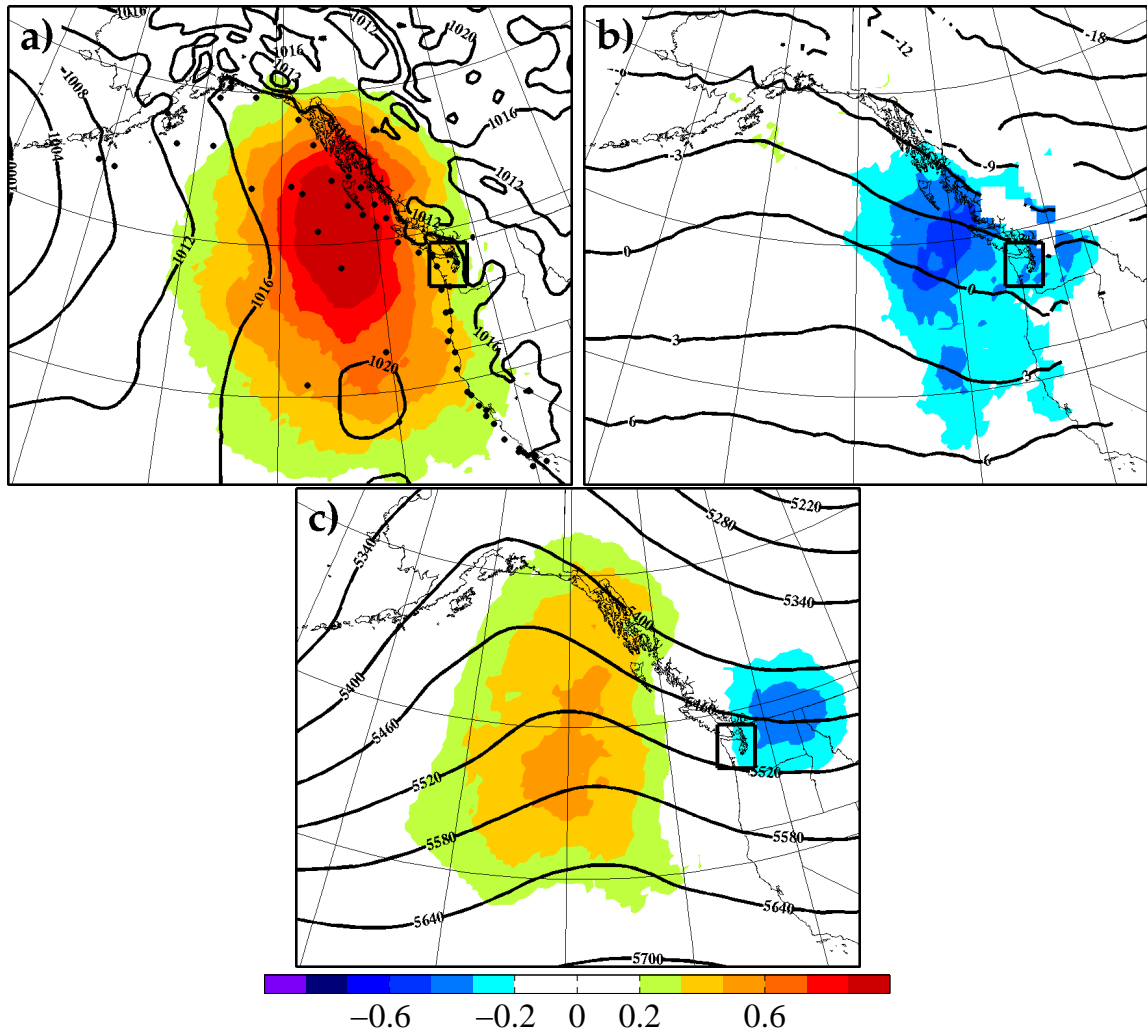


Figure 4.4: Composite sensitivity patterns for western Washington 24-hour SLP forecasts (shading, hPa) to analyses of (a) SLP, (b) 850 hPa temperature and (c) 500 hPa height. Each field represents the sensitivity multiplied by the analysis standard deviation at each analysis grid point for the 30 most sensitive western Washington SLP forecasts between 1 January and 30 June 2005. Contours denote the composite-average ensemble-mean analysis for these 30 cases (hPa).

is sensitive to a meridionally-elongated region near 140°W ; a one standard deviation change in x within the regions of largest sensitivity is associated with a 0.6 hPa change in the forecast metric (Fig. 4.4c). This region is located a few hundred kilometers upstream of the region of maximum SLP sensitivity, indicating that the moderate tilt of the sensitivity field is a common property among these results.

The composite average sensitivities for the 30 most sensitive precipitation forecasts are also computed in a manner similar to the 30 most sensitive SLP forecasts described above. Although one should expect precipitation to have a non-Gaussian distribution and be bounded below by zero, Gaussian statistics are assumed in these calculations. Results for the precipitation metric show a more pronounced composite-average low pressure system in the Gulf of Alaska and maximum sensitivity to SLP near $(44^{\circ}\text{N}, 133^{\circ}\text{W})$, just south of the results for the SLP metric (Fig. 4.5a). Sensitivity to 850 hPa temperature (Fig. 4.5b) falls within a relatively small region near a thermal ridge to the southeast of the composite cyclone. Sensitivity to 500 hPa height (Fig. 4.5c) exhibits largest sensitivity a few hundred kilometers west of the region of maximum sensitivity to SLP and downstream of a composite trough in the height field. For periods when the average precipitation in the box is greater than 2 mm, a one standard deviation change in the region of largest sensitivity is predicted to change the precipitation metric by 0.4 mm, and for 850 hPa temperature by 0.3 mm; thus it can be concluded that, as for SLP forecasts, precipitation forecasts are less sensitive to 850 hPa temperature than SLP or 500 hPa height.

4.3 Observation Denial Experiments for Single Observations

Recall from section 4.1 that buoy 46005 is located in a region of frequent sensitivity, but was not functional during the period considered. The change in western WA 24-hour SLP forecasts due to a missing buoy within the persistently sensitive region is assessed by withholding a nearby reliable buoy (buoy 46036, 42.3°N , 133.8°W) from the analysis. In addition to providing an estimate of the importance of offshore

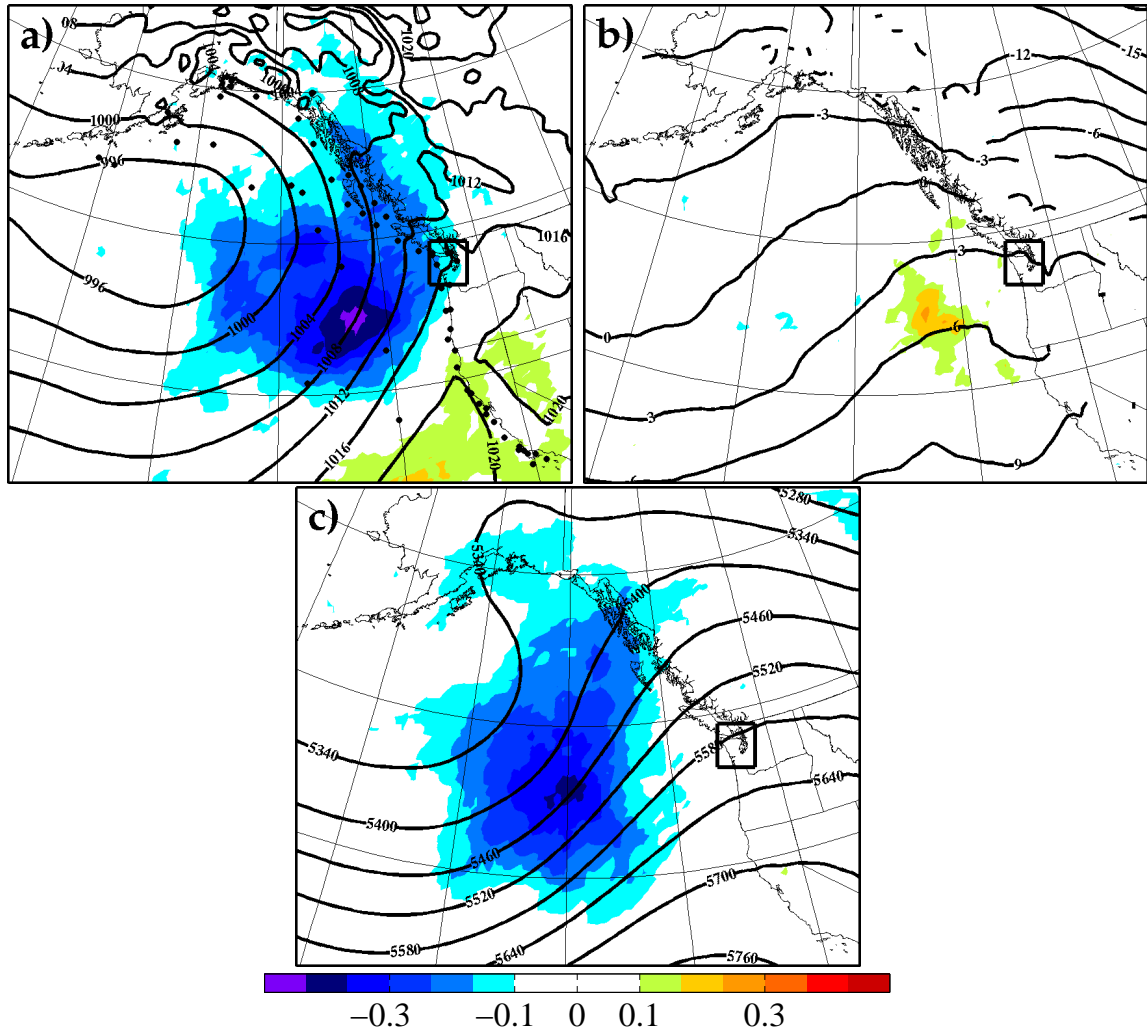


Figure 4.5: As in Fig. 4.4, but for the 24-hour forecast of precipitation (mm) for the western Washington metric box.

buoy observations, these data-denial experiments are used to quantify the accuracy of ensemble-based estimates of the change in a forecast metric due to an observation.

The change in 24-hour western WA SLP forecasts due to buoy 46036’s SLP observation is assessed for the 30 forecast cycles for which this forecast metric is most sensitive to this buoy’s observation using the following method. A “control” analysis is generated by assimilating all observations available to the UW EnKF system using the procedure described in the previous chapter, with the exception of buoy 46036’s SLP observation, which is assimilated without applying covariance localization. When covariance localization is applied, the magnitude of the ensemble estimated change is consistently larger than the actual change obtained from non-linear forecasts. Assimilating an observation with a localization radius may prevent all of the observation information from being spread to the model grid points in a manner consistent with the error statistics. Furthermore, the forecast ensemble does not have knowledge of localization, thus the observation change may not project onto the analysis grid points that have an impact on the forecast metric at some later time. Hamill and Snyder (2002) found that the predicted reduction analysis-error variance reduction well matched the actual reduction in analysis-error variance even when localization is used, but this most likely occurred because this metric did not involve running the forecast model.

A “no-buoy” analysis is generated by the identical procedure as the “control”, but without buoy 46036’s SLP observation; therefore, the differences between these two analyses is due solely to the assimilation of buoy 46036. The change in the expected value and spread of the 24-hour forecast of western WA SLP due to buoy 46036’s SLP observation is estimated by (2.25) and (2.28), respectively. These ensemble predictions, computed from the ensemble without the buoy, are verified here against ensemble forecasts generated from the “control” and “no-buoy” analyses, thus differences between the two forecasts are also due to buoy 46036. This section proceeds by describing the change in the forecast metric due to the buoy’s observation during one

case characterized by an eastern Pacific cyclogenesis event before summarizing all 30 cases.

Fig. 4.6a shows the UW EnKF ensemble-mean SLP analysis and forecast sensitivities for 12 UTC 5 February 2005. A frontal wave is situated on the eastern edge of a deeper cyclone near the international dateline; during the next 24 hours, this wave undergoes cyclogenesis as it moves east toward the North American coast. Forecast sensitivities are maximized along the eastern edge of the frontal wave near buoy 46036 (dot). Increasing (decreasing) the SLP in this region of the analysis by 1 hPa, which amounts to shifting the frontal wave to the northwest (southeast), leads to a 1.5 hPa increase (decrease) in the forecast metric.

The difference between the control and no-buoy analysis and their resulting 24-hour forecast differences are shown in Figs. 4.6b and c, respectively. For the control analysis, the SLP is 0.4 hPa lower to the south of the wave and 0.2 hPa higher to the north of the wave; thus the buoy's observation shifts the wave to the south. The largest 24-hour forecast differences are associated with the resulting cyclone along the Washington Coast; the forecast initialized from the control analysis has SLP values that are up to 0.8 hPa lower. The ensemble-based prediction of a 0.60 hPa (0.15 hPa) decrease in the expected value (spread) of the metric compares closely with the 0.63 hPa (0.18 hPa) reduction obtained from the non-linear forecasts.

Repeating the above process for the remaining 29 forecast cycles indicates that ensemble-based predictions provide accurate estimates of the changes in both the expected value and spread of the forecast metric. Figure 4.7 shows that the ensemble-based prediction of the change in the expected value and spread is in good agreement with the actual change obtained from the non-linear model ($R^2=0.985$); in 90% of cases considered, the error in the expected value (spread) is less than 0.1 hPa (0.05 hPa). Moreover, these results indicate that a buoy within the most sensitive region could produce up to a 0.8 hPa change in the expected value and a 0.5 hPa reduction in the spread of the 24-hour western WA area-averaged SLP forecasts.

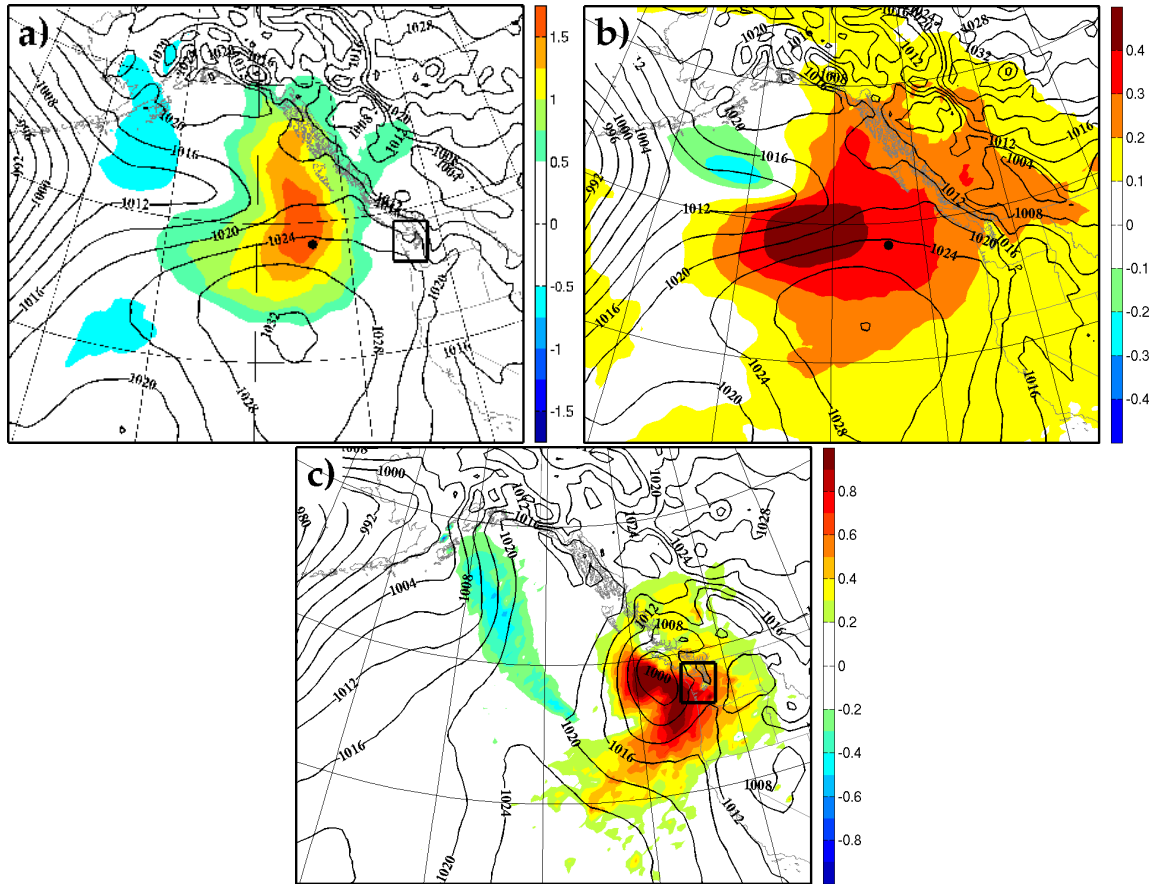


Figure 4.6: (a) Sensitivity of the western Washington 24-hour SLP forecast to the SLP analysis (shading, units hPa hPa^{-1}) and the UW EnKF ensemble mean analysis of SLP (contours, hPa) for the forecast initialized 12 UTC 5 February 2005. (b) Difference between the no-buoy ensemble mean analysis SLP field and the control ensemble mean analysis SLP field at 12 UTC 5 February 2005 (shading, hPa). The no buoy ensemble mean analysis of SLP is given by the solid lines (hPa). (c) as in (b), but for the 24-hour forecast of SLP valid 12 UTC 6 February 2005.

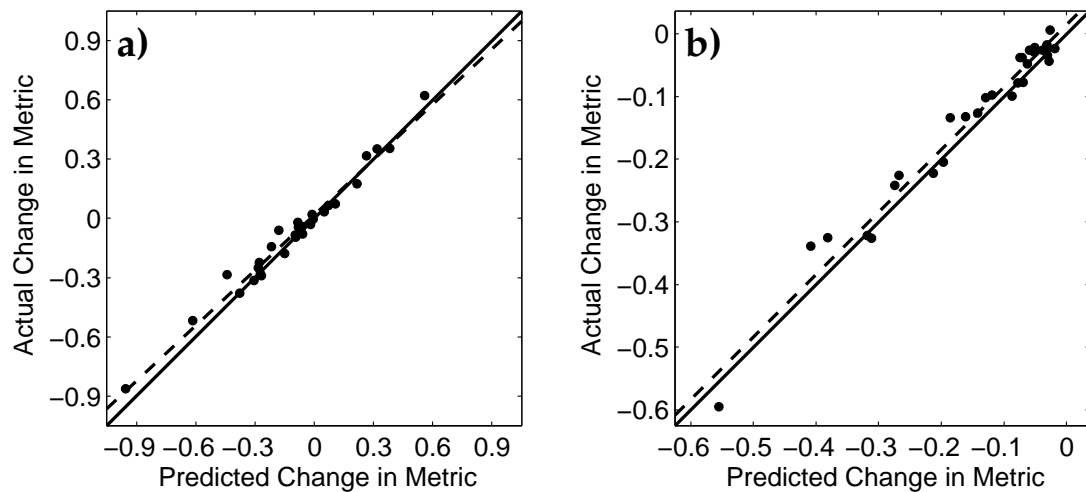


Figure 4.7: Change in the (a) expected value (hPa), and (b) spread (hPa) of the 24-hour western WA SLP forecast due to assimilating buoy 46036's SLP observation. Ensemble predictions (abscissa) and compared with results for differences between perturbed WRF forecasts (ordinate) for the 30 most-sensitive forecast cycles during January-July 2005. The dashed line is the linear least squares fit to the data. The solid line indicates perfect agreement between the ensemble-based prediction and the WRF model solutions.

4.4 *Observation Denial Experiments for multiple Observations*

The results of the previous section are extended from a single observation to larger sets. These experiments are performed to quantify the accuracy of ensemble-based estimates of how an observation changes a forecast metric and to evaluate the value of various observation platforms in the UW EnKF system. The objective here is similar to that of Langland and Baker (2004), who use an adjoint-based technique to estimate the impact of observations on the error in global forecasts.

This section considers 12-hour forecast cycles, and compares forecast metrics for the test case where observations are assimilated at forecast hour 6, with the control case where no observations are not assimilated at hour 6. Specifically, the impact of observations assimilated at 06 UTC and 18 UTC on forecasts valid 6 hours later is assessed during March 2005. This setup is chosen to utilize the existing 00 UTC and 12 UTC forecasts from the UW EnKF system during a period of large variability in weather systems that affected the west coast of North America. Observations are assimilated one at a time using the procedure described in section 3.1. Furthermore, rather than solve (2.28), the ensemble forecast metric values are updated in the same manner as the analysis state variables. The change in ensemble spread may then be evaluated from the updated ensemble metric values, which reflect the impact of all prior observations. The procedure used to update the forecast metric is similar to what Evensen (2003) proposed for EnKF parameter estimation.

Before assessing how all observations change the western WA SLP, the change due to assimilating select surface observations is estimated. These experiments are meant to be intermediary between the single observation experiments of the previous section and the experiments with all observations that will be described later in this section. For each 06 UTC and 18 UTC analysis time, the sensitivity of the forecast metric expected value to the model estimate of each surface observation is computed using (2.23) and tested for statistical significance at the 99% confidence level using

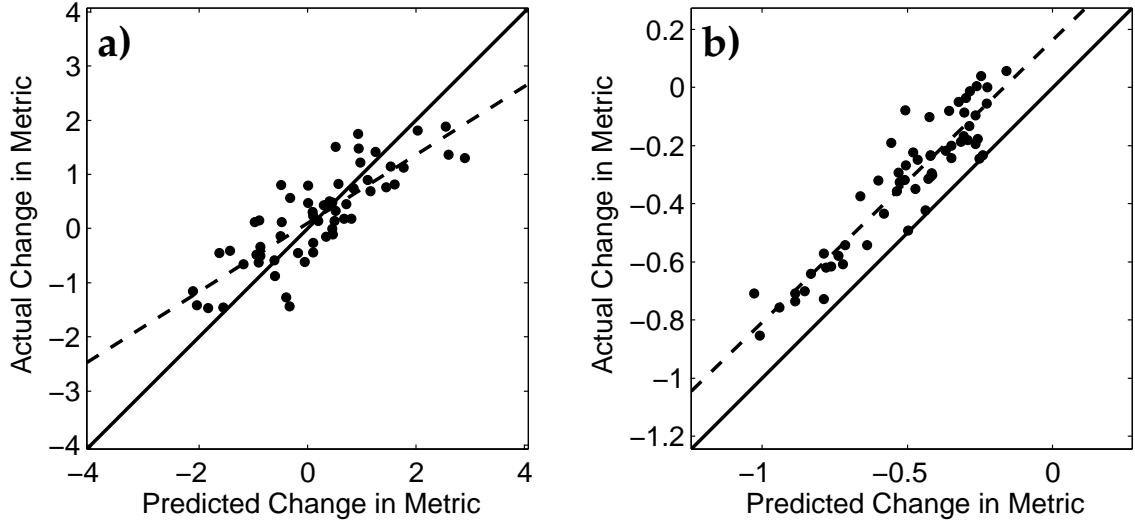


Figure 4.8: Change in the (a) expected value, and (b) spread of 6-hour forecasts of western Washington SLP due to assimilating all statistically significant (at the 99% confidence level). Ensemble predictions (abscissa) and compared with results for differences between perturbed WRF forecasts (ordinate) during March 2005. Dashed lines give the linear least-squares fit, while the solid line denotes perfect agreement between the predicted and actual change. Units are hPa.

(4.1). If this sensitivity value is significant, the observation is assimilated and the change in the forecast metric is evaluated (≈ 20 observations are assimilated during each cycle). Ensemble-based predictions of δJ and $\delta\sigma$ are verified by advancing the resulting analysis ensemble forward six hours using the WRF model.

Results show good agreement between the ensemble predictions and WRF verification of δJ and $\delta\sigma$ (Fig. 4.8). The correlation between the predicted and actual δJ and $\delta\sigma$ is 0.82 and 0.87, respectively. Whereas the bias in δJ is small, the ensemble-based estimate of $\delta\sigma$ is consistently larger than the actual value by 0.19 hPa. Differences between the predicted and actual impact are due to sampling error and nonlinearity; sampling error will be further discussed later.

The change in the western WA SLP forecast metric due to all observations available to the UW EnKF system is now assessed. For each analysis time, all obser-

vations are serially assimilated using the UW EnKF system procedure (≈ 3700 per cycle); however, the estimated change in the forecast metric is computed only if the sensitivity to the model estimate of the observation is significant at the 99% confidence interval (≈ 100 per cycle). This confidence interval is determined by repeating the above procedure with various confidence intervals. Figure 4.9 shows the RMS difference between the predicted and actual values of δJ and $\delta\sigma$ when the statistical significance value is varied. In these experiments, the state vector is updated with the same set of observations, but the number of observations used to estimate the change in the forecast metric varies depending on the statistical confidence level. Whereas the RMS difference between the predicted and actual change in the expected value decreases with increasing significance level, the difference between the predicted and actual spread is minimized at the 99% confidence level. When lower confidence intervals are used, spurious covariances will cause the ensemble sensitivities to adjust the estimated forecast metric for an observation that actually does not change the metric. Confidence intervals in excess of 99% can exclude observations that actually have an impact on the metric. In addition, covariance localization is used in these experiments, where the influence of observations reduces to zero 5000 km from the observation. This broad localization function prevents spurious correlations from adjusting state variables with observations at long distances. Moreover, for localization radii greater than 5000 km, there is no change in the agreement between the predicted and actual change in the metric. For simplicity, covariance inflation is not considered.

Figures 4.10a and b indicate that the predicted impact of observations on western WA SLP is in good agreement with the actual difference. On average, observations change the expected value of this metric by 0.86 hPa and reduce the spread by 0.59 hPa. Whereas the correlation between the predicted and actual δJ and $\delta\sigma$ is 0.49 and 0.93 respectively, there is more scatter about the line of perfect agreement when compared to Fig. 4.8. In contrast to the select surface observation results, the ensemble-based prediction of $\delta\sigma$ here exhibits little bias.

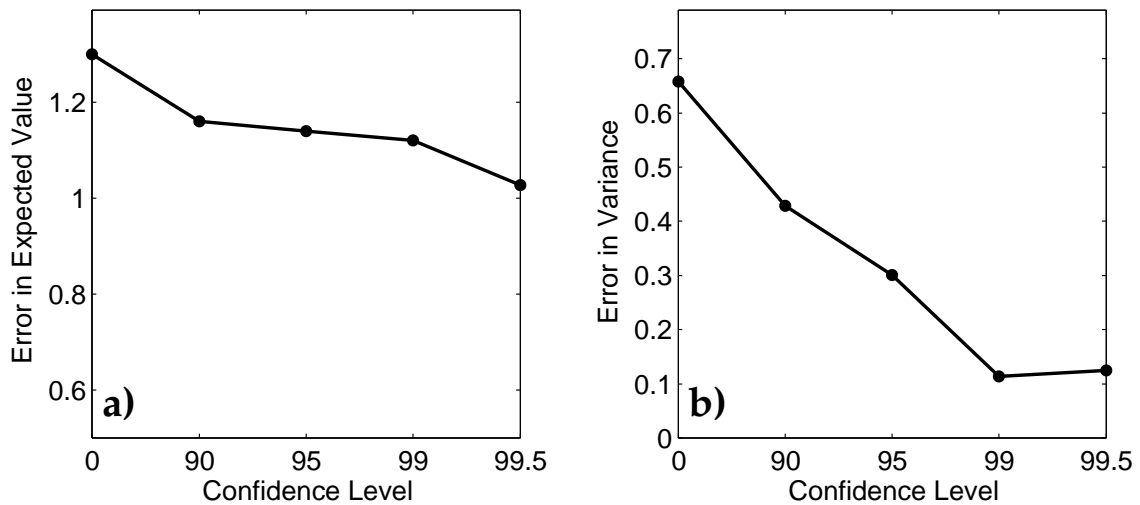


Figure 4.9: RMS difference between the ensemble prediction of the change in the (a) expected value, and (b) spread (units hPa) of western WA SLP due to observation assimilation and the actual difference between WRF forecasts as a function of the statistical significance level used to screen the observations.

A second test is performed to address how the size of the forecast metric box affects the results obtained above. The change in the forecast of average SLP in a box over the western North American coast (NAC; region given by the larger box in Fig. 4.1a) due to observations is determined by repeating the procedure used for the western WA SLP metric. Ensemble-based predictions of the change in NAC show comparable skill to the results obtained for western WA SLP (Fig. 4.10c, d). The correlation between the predicted and actual change in δJ and $\delta\sigma$ is 0.42 and 0.71 respectively. Observations produce slightly smaller changes in the expected value (0.75 hPa) and spread (0.37 hPa) of the NAC SLP metric as compared to the western WA SLP metric because there is less variability in SLP when averaged over a larger area.

The ensemble-based estimates of the change in the average SLP within the western WA box are partitioned by observation type to determine which observations produce the largest changes in this metric. Figure 4.11 shows probability density functions

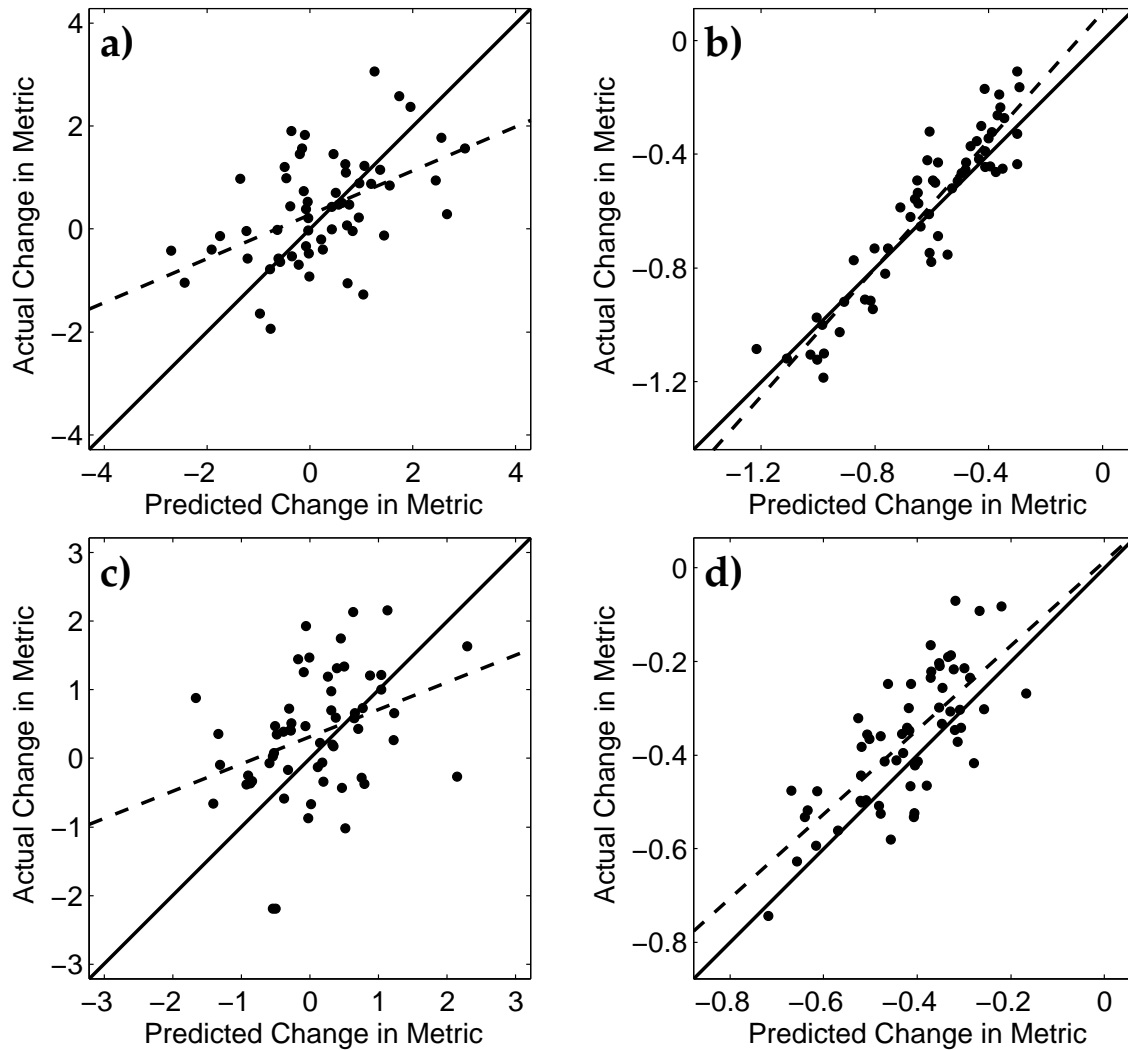


Figure 4.10: Change in the (a) expected value, and (b) spread of 12-hour forecasts of western Washington SLP due to assimilating all available observations. Ensemble predictions (abscissa) and compared with results for differences between perturbed WRF forecasts (ordinate) during March 2005. Dashed lines give the linear least squares fit. Panels (c) and (d) are similar to (a) and (b), but apply to the average SLP within the larger NAC box (see Fig. 4.1a). Units are hPa.

(PDFs) of the change in the expected value and spread of western WA SLP forecasts due to individual statistically significant observations (99% confidence). The results for NAC SLP forecasts are qualitatively similar to the western WA results except that the magnitudes are smaller. For all observation types, the PDFs are sharply peaked near zero, which indicates that most observations produce small changes to the average SLP within this region. The long tails in the surface observation PDFs indicate that, in comparison to ACARS and cloud winds, these observations are more likely to be associated with a large change. Furthermore, the surface observation PDF is skewed toward positive changes, thus surface observations are more likely to increase the average SLP within the box. This tendency toward positive increments is consistent with the observations trying to correct the negative bias in UW EnKF altimeter forecasts described in the previous chapter.

4.5 Impact of Observations on Forecast Verification

Forecast metrics that are only a function of the forecast state vector may be well predicted by this technique; however, this does not guarantee that the observations are actually improving the forecasts. As a consequence, the experiments in the previous section are repeated to assess how observations impact the RMS error in SLP within the western WA and NAC regions. In this case, J is the RMS error in the box and, unlike the previous calculations, this metric can only be evaluated *a posteriori* when an analysis is available for verification. An ensemble of RMS error values within each box is determined based on each ensemble member's forecast verified against the appropriate ensemble-mean analysis; negative values of δJ indicate that observation assimilation decreases the RMS error.

Fig. 4.12a and b indicate that ensemble-based impact predictions for western WA SLP error have skill comparable to the western WA average SLP forecast metric; the correlation between the predicted and actual δJ and $\delta\sigma$ is 0.48 and 0.79 respectively. On average, assimilating observations reduces the RMS error in WA SLP forecasts by

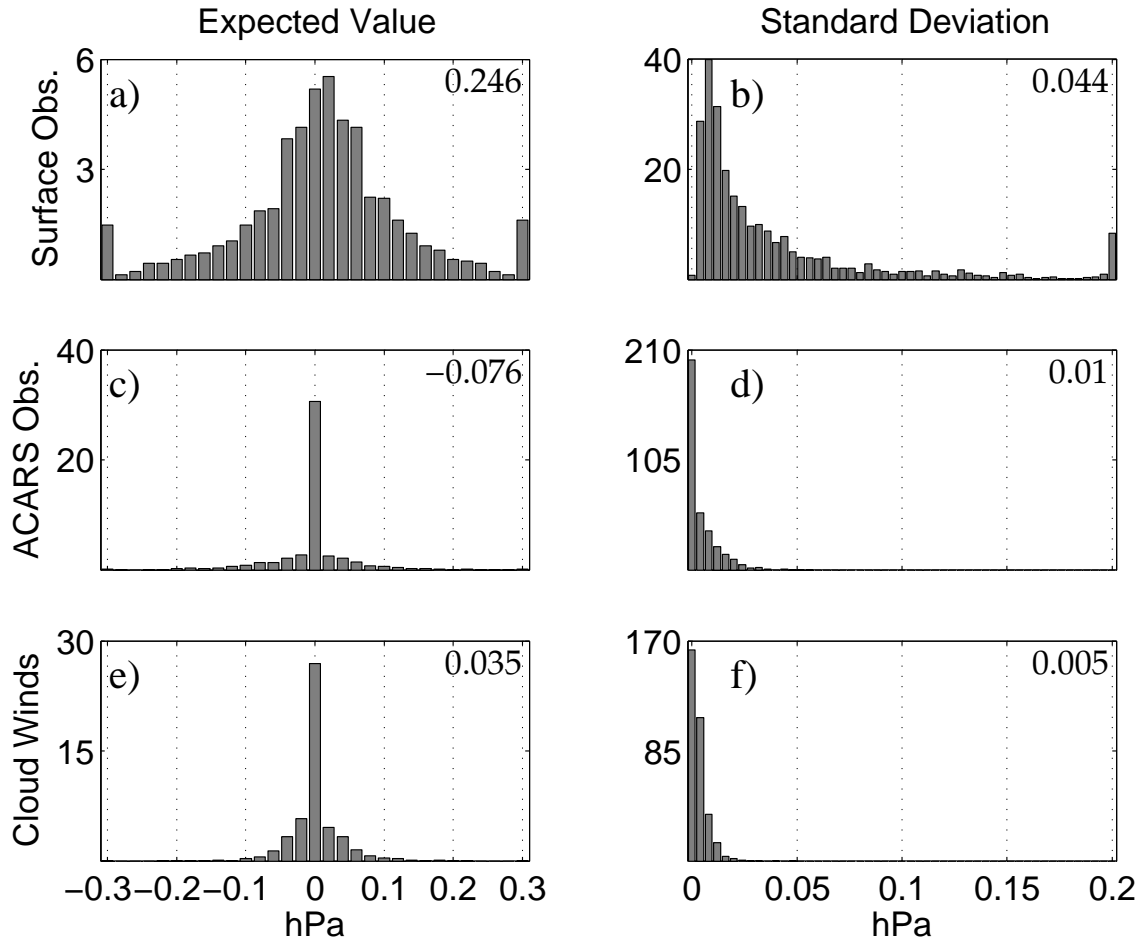


Figure 4.11: Probability density functions (PDFs) of the impact of individual statistically significant (99% confidence) surface (top row), ACARS (middle row) and cloud wind (bottom row) observations assimilated at 06 UTC and 18 UTC on the expected value (left column) and spread (right column) on the average SLP within the western WA region valid 6 hours later during March 2005. The value at the top of each panel indicates the average impact of each observation type during a data assimilation cycle, which has units of hPa per cycle.

0.67 hPa. Comparable results are obtained for the RMS error in SLP over the NAC region. In this case, the correlation between the predicted and actual δJ (0.42) and $\delta\sigma$ (0.71) are similar to the results obtained from the average NAC SLP metric. For both regional metrics, there are multiple cycles where the ensemble-based estimate of δJ is off by at least 1 hPa. Each of these forecasts is characterized by a cyclone undergoing rapid cyclogenesis or cyclolysis near the edge of the respective box (not shown).

Observation histograms for the RMS error in western WA SLP forecasts are shown in Fig. 4.13; histograms for the RMS error in NAC SLP forecasts are similar. In general, some observations improve SLP forecasts in this domain, while others will make the forecast worse. Similar to Fig. 4.11, surface observation PDFs have larger tails and thus are more likely to impact the RMS error in SLP. In Figs. 4.13c and 4.13e, the PDFs are symmetric about zero, which implies that ACARS and cloud wind observations are equally likely to have a positive or negative impact on the RMS error in SLP. In contrast, the surface observation PDF (Fig. 4.13a) is skewed toward negative values, indicating that surface observations are more likely to reduce, rather than increase, the RMS error in SLP. This distribution is qualitatively similar to histograms of the impact of how adding additional observations impact forecast errors for a quasi-geostrophic channel model (Morss and Emanuel 2002, their Fig. 8). The surface observations having the highest impact are given by the buoys located approximately 500 km offshore, which is consistent with earlier results indicating this is a region of high sensitivity and that SLP observations can produce the largest changes in the average SLP. It should be noted that the results obtained in this chapter will vary depending on the chosen metric, observations assimilated, NWP model and season. While ACARS and cloud wind observations may not have a significant impact in the SLP metrics, these observation types may have larger impact on metrics that measure upper-tropospheric forecast fields.

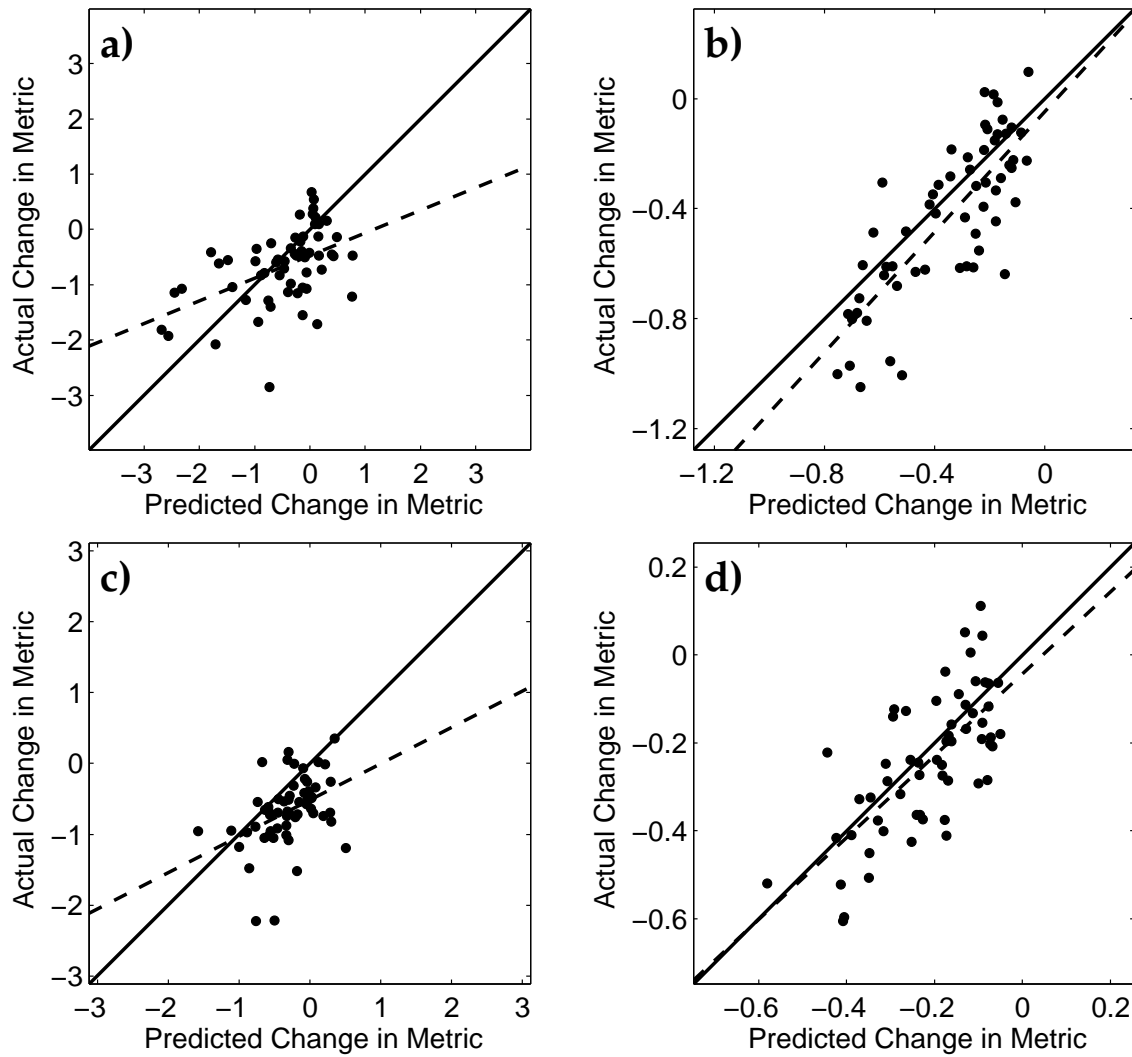


Figure 4.12: As in Fig. 4.10, but for the RMS error in SLP forecasts within the western Washington region (panels a and b) and the NAC region (panels c and d) valid 6 hours later. Units are hPa.

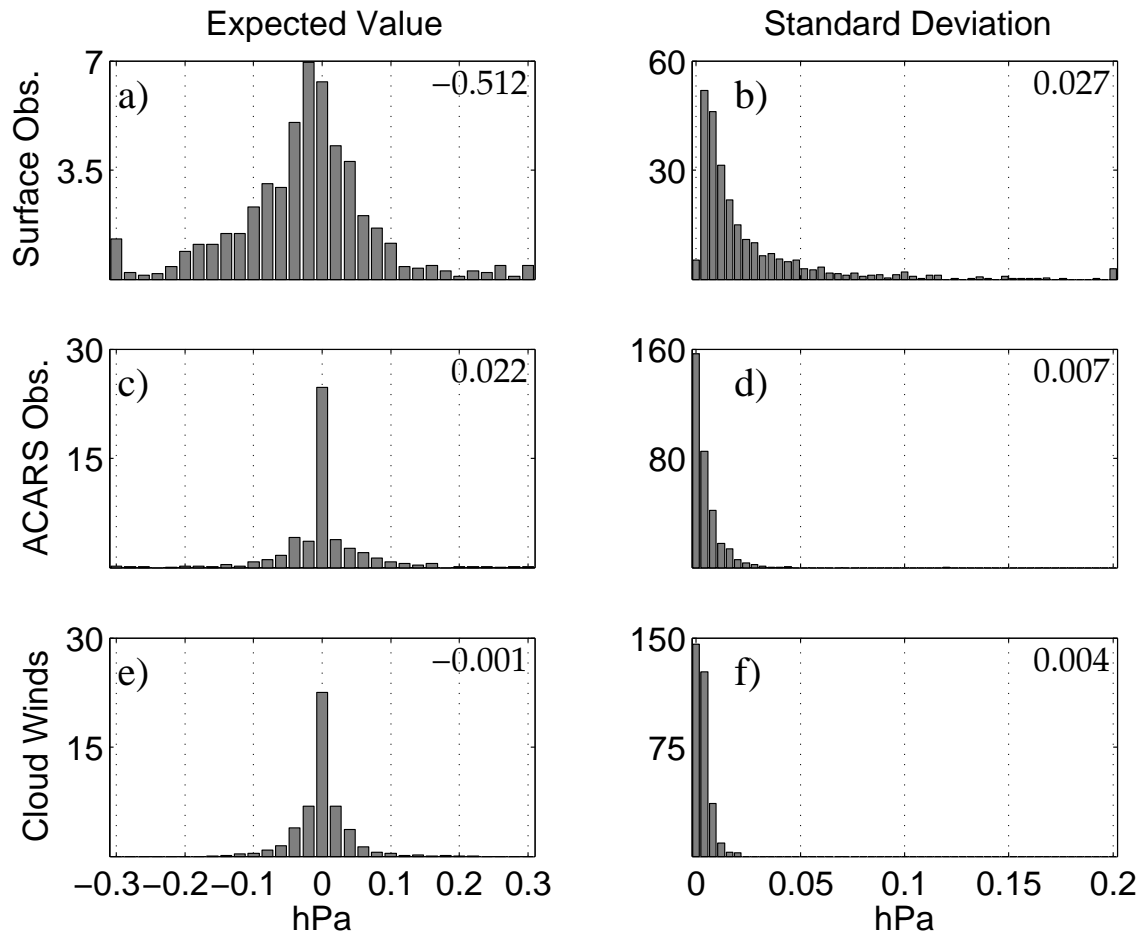


Figure 4.13: As in Fig. 4.11, but for the RMS error in SLP forecasts within the western WA box.

Chapter 5

APPLICATION OF ENSEMBLE SENSITIVITY TO EXTRATROPICAL TRANSITION FORECASTS

5.1 *Experiment Setup*

Here output from a WRF EnKF system is analyzed using the ensemble-based sensitivity techniques described in the previous chapter to understand the predictability and dynamics of four recent western Pacific extratropical transition events. This WRF EnKF system is similar to the UW EnKF setup described in Chapter 3 since the forecasts from the UW EnKF system have comparable skill to guidance generated by other operational centers. Unlike the UW EnKF system which continuously cycled with observations and a short-term forecast over a long period of time, this WRF EnKF setup generates an analysis ensemble each six-hours for a seven-day period prior to, and during the transition of Typhoons Lupit (2003), Tokage (2004), Nabi (2005), and Kirogi (2005). The following section describes each of these storms in detail and why they were chosen for this study.

Each of the 90 ensemble members is advanced using the ARW WRF (Skamarock et al. 2005) on a numerical grid with 45 km horizontal grid spacing with 30 vertical levels over a domain that includes eastern Asia and the western Pacific Ocean. This particular domain and grid spacing represents a trade-off between ensuring that the tropical cyclone and important mid-latitude features are well removed from the lateral boundaries during the forecast time of interest, and memory constraints of the machines used. Furthermore, the horizontal grid spacing used here is similar to the resolution of global models operational during this period (cf. Table 3.2). Although this grid spacing does not allow the model to resolve the mesoscale features of the

tropical cyclone such as the eyewall or rainbands, which could potentially lead to additional model error, the goal of these experiments is to understand the synoptic-scale predictability of ET. For consistency, this implementation of WRF uses the same parameterizations used in the UW EnKF system: WRF 3-class microphysics scheme (Hong et al. 2004), Kain-Fritsch convection (Kain and Fritsch 1990) and, Noah land surface parameterization scheme (Ek et al. 2003). The lateral boundary conditions are perturbed using the fixed covariance perturbation technique of Torn et al. (2006) with a scaling factor and autocorrelation coefficient of 1.6 and 0.4, respectively. Ensemble-mean forecasts on the lateral boundaries are obtained from the 6-hour NCEP GFS forecast valid at the appropriate time.

Similar to the UW EnKF system, observations from surface stations, rawinsondes, ACARS and cloud track winds are assimilated serially using a square-root version of the EnKF (Whitaker and Hamill 2002) for a 90 member ensemble. When the tropical cyclone (TC) is classified as such, the Japan Meteorological Agency (JMA) cyclone best track position (latitude and longitude) data is assimilated in a manner similar to what Chen and Snyder (2007) used for a 2D barotropic vortex. When assimilating the best-track data, the model estimate of the best track position is computed by finding each ensemble member’s latitude of lowest SLP near the TC, assimilating that observation, and then repeating for the longitude of lowest pressure. A significant fraction of the western Pacific Ocean is devoid of *in situ* observations that could constrain the TC location, thus assimilating the TC best-track position ensures that the TC is in the correct location in the WRF EnKF analyses. Table 5.1 displays the average number of observations assimilated at each analysis time by this WRF EnKF system; a large fraction of observations on this domain are in the form of rawinsondes and cloud motion vectors. The influence of observations is localized using the Gaspari and Cohn fifth-order piecewise rational function, which for these simulations, reduces to zero 2500 km from the observation location. Moreover, the tendency for small ensembles to underestimate covariance magnitude is treated by

Table 5.1: Observation types and average number of observations assimilated during each forecast cycle by the WRF EnKF system used for these ET predictability experiments. There are an average of 170 rawinsonde launches at 00 UTC and 12 UTC.

		Analysis Time			
Observation	type	00 UTC	06 UTC	12 UTC	18 UTC
Surface	Alt., u, v	330	310	270	250
Rawinsonde	u, v, T, RH	6250	200	5950	200
ACARS	u, v, T	400	550	500	550
Cloud Winds	u, v	5100	4800	5400	5200
Total		12 080	5860	12 120	6200

inflating the deviations from the ensemble mean by replacing the posterior ensemble perturbations from the mean with a linear combination of the prior and posterior perturbations where the prior (posterior) is weighted by 0.75 (0.25) (Snyder and Zhang 2003). The weighting factor is empirically determined by cycling over the seven-day period during Tokage’s transition with various scaling factors and verifying the six-hour forecasts against rawinsonde data, in a manner similar to what is described for the UW EnKF system in section 3.2.

For each of the extratropical transition events studied here, 48-hour ensemble forecasts are initialized at the onset of each TC’s transition; Table 5.2 shows the initialization time of each 48-hour forecast that will be explored here. Forecasts close to the onset of transition are explored here because Evans et al. (2006) found that forecasts initialized during this time are often characterized by large errors in the structure and position of the storm. The onset of transition is objectively determined by constructing a cyclone phase-space diagram (Hart 2003) from the WRF EnKF ensemble-mean analysis data. Extratropical transition is defined to begin when the average 900-600 hPa thickness to the left of the cyclone track is 10 m less than the

average thickness to the right of the cyclone track, which is akin to the cyclone going from the lower-right to upper-right portion of the phase-space diagram. Evans and Hart (2003) found that no major hurricane (winds greater than 60 m s^{-1}) had a thermal asymmetry that exceeded this value. Lateral boundary conditions for the 48-hour ensemble forecasts are generated using the Torn et. al (2006) fixed covariance perturbation technique, where the scaling factor increases linearly with time to a maximum value of 2.8 at hour 48. The scaling factor for the lateral boundary conditions at the 48-hour lead time is determined by computing the ratio of the RMS error in 48-hour GFS forecasts to the spread of the fixed perturbations generated by the WRF-VAR system (Barker et al. 2004).

For each ET case described here, the 90 member ensemble is initialized at the time shown in Table 5.2 (four days prior to the onset of each TC's transition) by adding fixed covariance perturbations from the WRF-VAR system to a 36-hour GFS forecast valid at the initialization time. Prior to being added to the ensemble-mean forecast, the ensemble perturbations are multiplied by 2.0, so that the spread of the ensemble is greater than the error in the 36-hour forecast. Recall from Chapter 3 that Dirren et al. (2007) found that initializing an ensemble data assimilation system with large error and spread can prevent the ensemble from under-weighting observations during the first few assimilation times while flow-dependent covariances develop. RMS errors in six-hour forecasts of wind, temperature, and geopotential height come into equilibrium 48 hours after the ensemble is initialized, thus the initial ensemble does not have a large impact on the results here (not shown).

This chapter proceeds by giving a short summary of the ET events explored in section 5.2, while section 5.3 describes WRF EnKF forecasts of each storm at the onset of transition. Section 5.4 displays the initial-condition sensitivity for these forecasts, followed by section 5.5 where these structures are used to produce diagnostic corrections to the initial conditions that will improve 48-hour cyclone minimum SLP and SLP forecasts near the TC. In section 5.6, the observations that lead to the largest

Table 5.2: List of extratropical transition forecasts studied in this chapter. The text contains information on how to interpret each category.

Tropical Cyclone	Tokage	Nabi
Experiment Start	12 UTC 15 October 2004	00 UTC 2 September 2005
Experiment End	12 UTC 22 October 2004	00 UTC 9 September 2005
Forecast of Interest	12 UTC 19 October 2004	00 UTC 6 September 2005
Type of Transition	weakening	reintensifying
Predictability	low	high
Downstream Impact	high	low

Tropical Cyclone	Kirogi	Lupit
Experiment Start	00 UTC 12 October 2005	12 UTC 26 November 2003
Experiment End	00 UTC 19 October 2005	12 UTC 03 December 2003
Forecast of Interest	00 UTC 17 October 2005	12 UTC 30 November 2003
Type of Transition	weakening	reintensifying
Predictability	low	medium
Downstream Impact	low	high

changes in the 48-hour minimum SLP forecast will be determined. Finally, section 5.7 uses the statistics from the short-term forecast ensemble in these predictability experiments to understand the dynamical mechanisms responsible for downstream ridging during ET.

5.2 Overview of Cases

The four western Pacific extratropical transition events explored here are selected based on whether the TC completes its transition into a baroclinic system, as determined from GFS analysis data, the skill of the GFS forecast at the onset of transition and the predicted impact on the downstream state; Table 5.2 summarizes the characteristics of each ET forecasts. A reintensifying system is defined as a tropical cyclone whose minimum SLP decreases by 10 hPa and transitions into baroclinic system within 72 hours of the onset of transition, while a weakening storm is a transition where the minimum SLP continuously rises (Klein et al. 2000). The predictability of GFS forecasts initialized at the onset of transition is evaluated based on three objective standards. A storm is deemed to have low predictability if the 48-hour GFS forecast at the onset of transition meets two of the following are met: the error in minimum SLP is greater than 10 hPa¹, the error in cyclone position is larger than 250 km², or the RMS error in 500 hPa height in the western Pacific basin (25°N–60°N and 120°E–180°E) exceeds 22 m³. TC track and intensity forecasts are verified against JMA best-track data, while 500 hPa heights are verified against GFS analyses. The ET is said to have a large downstream impact in the GFS forecast initialized at the onset of transition if a Hovmoller diagram of the average 250 hPa meridional winds

¹This value was used by McMurdie and Mass (2004) to define a large error in 48-hour SLP forecasts in the Northeast Pacific.

²Froude et al. (2007) found this to be the average error in 48-hour global model forecasts of mid-latitude cyclone position.

³The average RMS error in GFS 48-hour forecasts of Northern Hemisphere 500 hPa geopotential height.

between 30°N and 60°N shows a wave packet originating from where the TC enters the mid-latitudes. Prior to evaluating the analyses and forecasts of each transition event, a short summary of each storm's lifecycle is provided.

5.2.1 *Tokage*

Typhoon Tokage was first classified a tropical system on 00 UTC 12 October 2004 60 km west of Pohnpei and thereafter moved to the west along the southern side of the subtropical high (Fig. 5.1a). As Tokage began to move northwest on 16 October, the storm passed within 40 km of Guam, began a period of intensification, and reached its lowest SLP (940 hPa) on 00 UTC 17 October. In response to a mid-latitude trough to its north, Tokage turned to the northeast, accelerated and began to slowly weaken due to the increased vertical wind shear (Atangan and Preble 2004). GFS forecasts initialized at the onset of transition (12 UTC 19 October) had Tokage transitioning into a sub-970 hPa baroclinic cyclone within 48 hours; however, the storm actually continued to weaken as it moved across Japan on 20 October and back out to sea on 21 October. Moreover, the rapid intensification of the baroclinic storm in the GFS forecast lead to a downstream wave packet, thus the RMS error in 48-hour GFS 500 hPa height forecast in the western Pacific box described in the previous paragraph is 36 m.

5.2.2 *Nabi*

Similar to Tokage, Typhoon Nabi became a tropical cyclone near the Mariana Islands on 29 August 2005 (Fig 5.1b). During the next six days, this TC moved to the northwest, then north toward the southern Japanese coast as it intensified into a 935 hPa TC. By the onset of transition (00 UTC 6 September), the TC cloud field took on a more asymmetric appearance just prior to making landfall on the Japanese island of Kyushu (Atangan and Preble 2005). Upon moving into the sea of Japan on 7 September, the storm accelerated to the north and transitioned into a 970 hPa

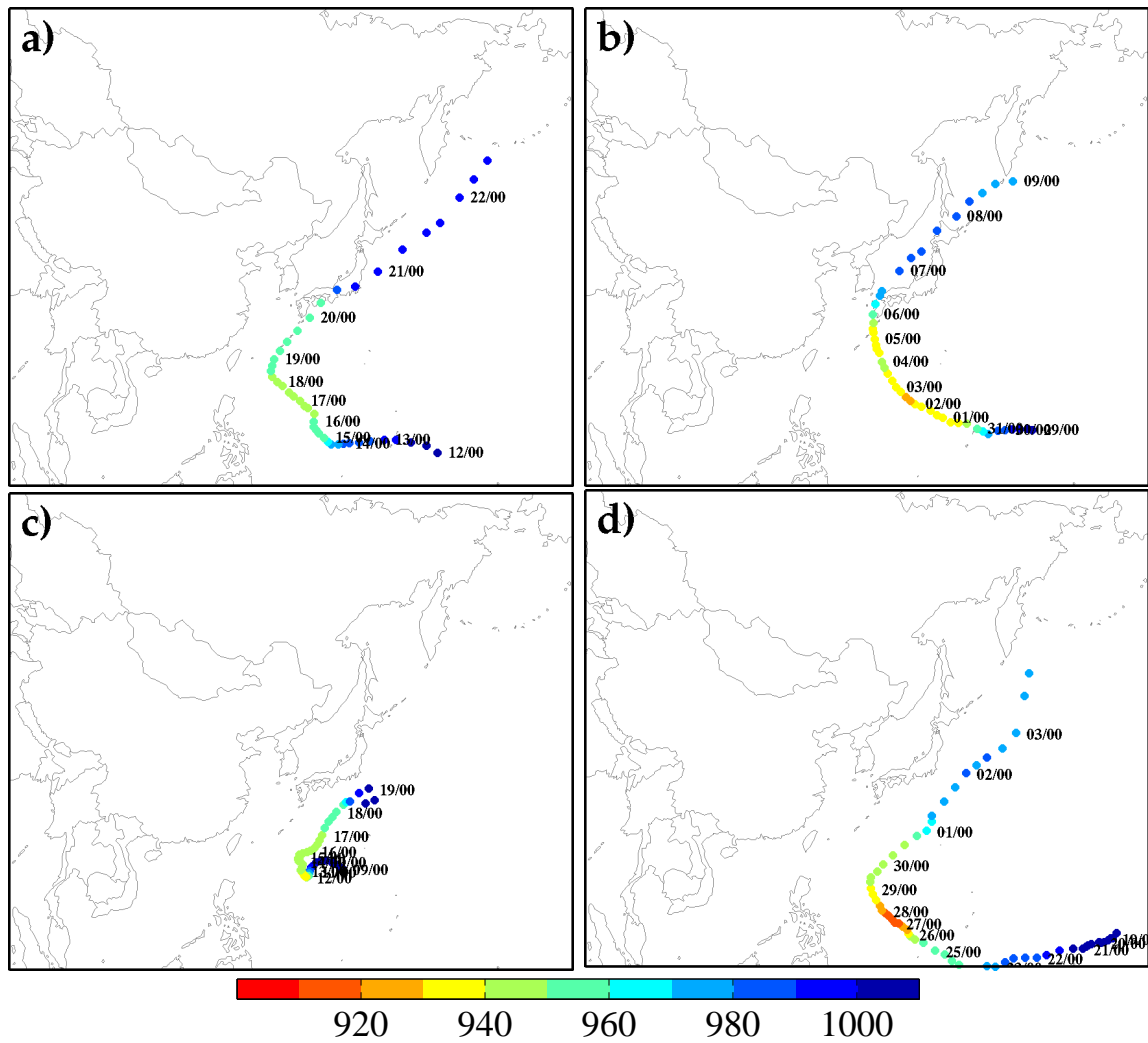


Figure 5.1: Japan Meteorological Agency tropical cyclone best track position and intensity estimates for Typhoons (a) Tokage, (b) Nabi, (c) Kirogi, and (d) Lupit.

baroclinic cyclone on 9 September south of the Kamchatka peninsula. Although Nabi underwent rapid intensification, the GFS forecast initialized on 00 UTC 6 September accurately captures the storm's evolution, so the error in 48-hour GFS 500 hPa height forecast in the western Pacific box is 16 m.

5.2.3 *Kirogi*

Unlike the previous two cases, Typhoon Kirogi formed in the middle of a subtropical ridge on 00 UTC 9 October 2005 and remained nearly stationary over the next five days, whereby it slowly intensified (Fig. 5.1c). This storm reached a maximum intensity of 930 hPa on 12 UTC 12 October. During this period, most global forecasting systems had difficulty determining the track of the cyclone; GFS forecasts of Kirogi alternated between moving the storm west or east due to the weak steering winds. By 00 UTC 15 October, Kirogi began to slowly move to the northeast in response to an approaching synoptic-scale trough. The tropical cyclone began to rapidly weaken on 00 UTC 17 October as the storm was influenced significant vertical wind shear in a mid-latitude jet streak on the eastern side of a synoptic-scale trough (Atangan and Preble 2005). By 48 hours later, Kirogi was not identifiable in visible or infrared satellite imagery. Since Kirogi never reintensified as a baroclinic system, it had minimal impact on the mid-latitude flow and the RMS error in this 48-hour GFS western Pacific 500 hPa height forecast is 22 m.

5.2.4 *Lupit*

Typhoon Lupit was the most intense TC and resulted in the strongest baroclinic cyclone among the four cases described here (Fig. 5.1d). During the first five days of Lupit's lifetime, the storm slowly strengthened into a 970 hPa TC as it moved to the southwest toward the Caroline Islands in response to the subtropical ridge to the north. Beginning 24 November, Lupit underwent a slow strengthening to a 915 hPa TC as it turned to the northwest. After the storm reached its maximum intensity in

the Philippine sea, Lupit made a northward turn in response to a weakness in the subtropical ridge. As the environment became more unfavorable to the TC on 12 UTC 30 November, the storm began to undergo extratropical transition and weakened into a 985 hPa cyclone by 2 December; 36 hours later, Lupit's remnants had transitioned into a 970 hPa baroclinic cyclone to the south of the Kamchatka Peninsula (Furze and Preble 2003). As the TC underwent transition, the mid-latitude flow in the western Pacific amplified, which lead to a wave packet that was associated with high impact weather throughout the Northern Hemisphere, including a sub-960 hPa cyclone in the eastern Pacific on 3 December and a early-season snowstorm on the east coast of the United States on 6 December. In contrast to the Tokage forecast, the RMS error in the 48-hour GFS 500 hPa height forecast in the western Pacific box described above is 25 m, thus Lupit's forecast has a higher degree of predictability.

5.3 Ensemble Forecasts of Transition

Before describing the WRF EnKF forecasts initialized at the onset of transition, a verification of the ensemble-mean background forecasts against rawinsonde and JMA best track data during all transition events is described. Figure 5.2 shows vertical profiles of the RMS error in six-hour WRF EnKF and GFS forecasts of temperature and wind over this domain. Similar to the verification performed for the UW EnKF system (cf., Fig. 3.5a), the RMS error in WRF EnKF background temperature forecasts on this domain are 0.3 K greater than the corresponding time GFS background forecasts, except in the upper troposphere, where the RMS error and bias in UW EnKF forecasts is 0.6 K greater. The magnitude of temperature errors on this domain are 30% larger than what is observed in the eastern Pacific (cf., Fig. 3.5a). For meridional winds, the RMS error in UW EnKF forecasts increases from 3.5 m s^{-1} in the lower troposphere to 5 m s^{-1} at 150 hPa (Fig. 5.2b). The difference between the error in WRF EnKF and GFS forecasts is less than 1 m s^{-1} throughout the column, thus one can conclude that six-hour forecasts on this domain

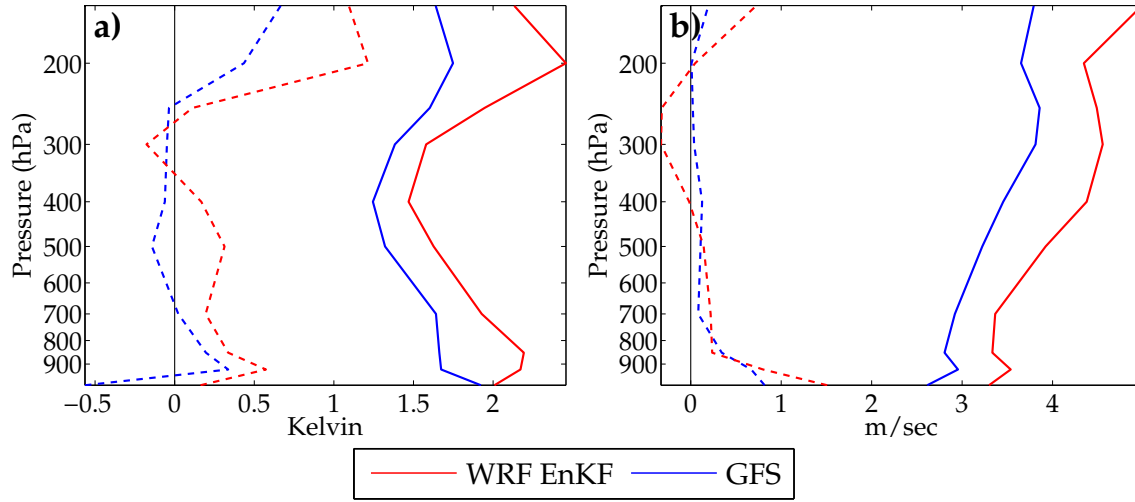


Figure 5.2: RMS error (solid) and bias (dashed) in six-hour ensemble-mean WRF EnKF and GFS forecasts of (a) temperature, and (b) meridional wind in the WRF EnKF domain validated against rawinsonde observations for the duration of the ET cases explored here.

are only slightly worse than the GFS, and are similar to the UW EnKF system.

Validation of ensemble-mean analyses and forecasts of cyclone position and intensity against best track data indicate that the WRF EnKF captures the evolution of each cyclone without applying special techniques that use a “bogused” vortex (e.g., Kurihara 1995) to represent the TC in the analysis (Table 5.3). When the storm is classified as a tropical cyclone, the verification is performed against JMA best track data; however, after the cyclone is declared extratropical, the verification position and intensity are obtained from 1° resolution GFS analysis data. For all cases explored here, the analysis position errors are slightly larger than the distance between grid points, but are consistently lower than background forecast position errors, thus observations systematically reduce the error in TC position. Moreover, the smallest (largest) position errors occur before (after) the cyclone has undergone transition. Larger errors are obtained subsequent to the onset of transition for two different reasons. After the tropical cyclone is classified as extratropical, best track

Table 5.3: RMS Error in the ensemble-mean analysis and six-hour forecast of tropical cyclone track and intensity averaged over the seven day period where cycling is performed for each ET event.

	TC Track		TC Intensity	
	analysis	background	analysis	background
Tokage	90 km	110 km	18 hPa	18 hPa
Nabi	53 km	77 km	14 hPa	14 hPa
Kirogi	60 km	60 km	25 hPa	24 hPa
Lupit	62 km	81 km	25 hPa	21 hPa

position observations are no longer assimilated, thus there are only a limited number of observations that can correct errors in the storm's location. Furthermore, the post-transition verification position is estimated from 1° resolution GFS analysis, thus the storm's location is given within ± 50 km, as compared to ± 10 km for JMA best track position.

The coarse horizontal resolution of this model prevents the WRF EnKF analyses and forecasts from providing an accurate estimate of the TC intensity. For all cases, the RMS error in analysis TC intensity is between 15-25 hPa and are inversely proportional to the minimum central pressure of the tropical cyclone; when the storm is stronger (weaker) the intensity errors tend to be larger (smaller). After the onset of transition, the cyclone's dynamics are governed by processes the model can resolve, thus the intensity errors are smaller. With the exception of Lupit, the analyses and six-hour forecasts have similar intensity errors, thus observations do not have a systematic impact on storm intensity. Nevertheless, it should be noted that while the WRF EnKF analyses and forecasts may not yield the exact intensity of each storm, they do replicate the trend.

The remainder of this chapter focuses on the 48-hour forecasts initialized at the

onset of each storm’s transition and their initial condition sensitivity. Fig. 5.3 shows 48-hour probabilistic track forecasts of each storm’s transition computed from the 90-member WRF EnKF forecasts. These probabilities are determined by calculating the percentage of ensemble members where the storm’s center passes within 90 km of each model grid point. Probabilistic intensity forecasts out to 48-hours from the WRF EnKF ensemble forecasts are shown in Fig. 5.4. For comparison, Figs. 5.3 and 5.4 also show the operational GFS forecast and the best-track data for the corresponding times, which will be used for verification. In each of the forecasts shown here, the WRF EnKF forecasts have problems simulating the intensity of the storm. This could result from the coarse resolution of this model or from initial condition errors. It should be noted that these WRF EnKF intensity forecasts are similar to the GFS forecast, thus the large intensity errors do not seem to be specific to the WRF EnKF system.

The evolution of the SLP and 500 hPa height fields and ensemble spread for 48-hour WRF EnKF forecasts of Tokage’s (Fig. 5.5), Nabi’s (Fig. 5.6), Kirogi’s (Fig. 5.7) and Lupit’s (Fig. 5.8) transition are also shown below. These figures display the ensemble-member closest to the ensemble-mean, rather than the ensemble-mean field, which often does not contain a great amount of detail in individual features. The ensemble member closest to the mean is determined by computing the RMS difference between each ensemble member’s 48-hour forecast of SLP and the ensemble-mean 48-hour forecast. For a given storm, the member with the smallest difference is shown in all panels.

5.3.1 *Tokage*

WRF EnKF forecasts of Tokage’s transition initialized on 12 UTC 19 October contain a great deal of variability in the evolution of the storm and downstream state. At 12 UTC 19 October (forecast hour 0), Tokage is located 1000 km to the south of Japan and is associated with a large amount of variance due to the large SLP gradi-

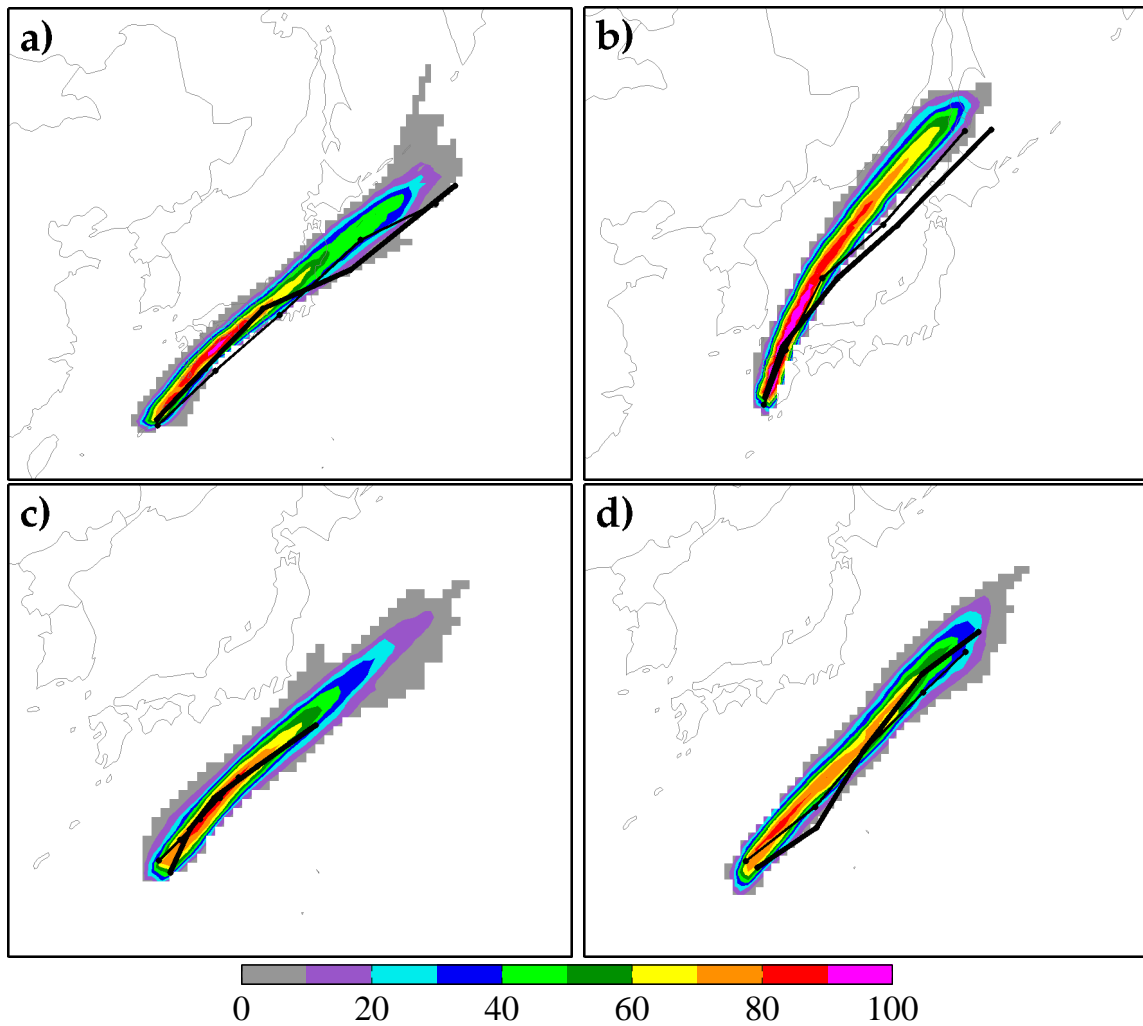


Figure 5.3: Percentage of WRF EnKF forecast ensemble members where the tropical cyclone center passes within 90 km of each model grid point during the 48 hour period initialized on (a) 12 UTC 19 October 2004 (Tokage), (b) 00 UTC 6 September 2005 (Nabi), (c) 00 UTC 17 October 2005 (Kirogi), and (d) 12 UTC 30 November 2003 (Lupit). The thick black line is the JMA best track position and the thin solid line applies to the GFS forecast.

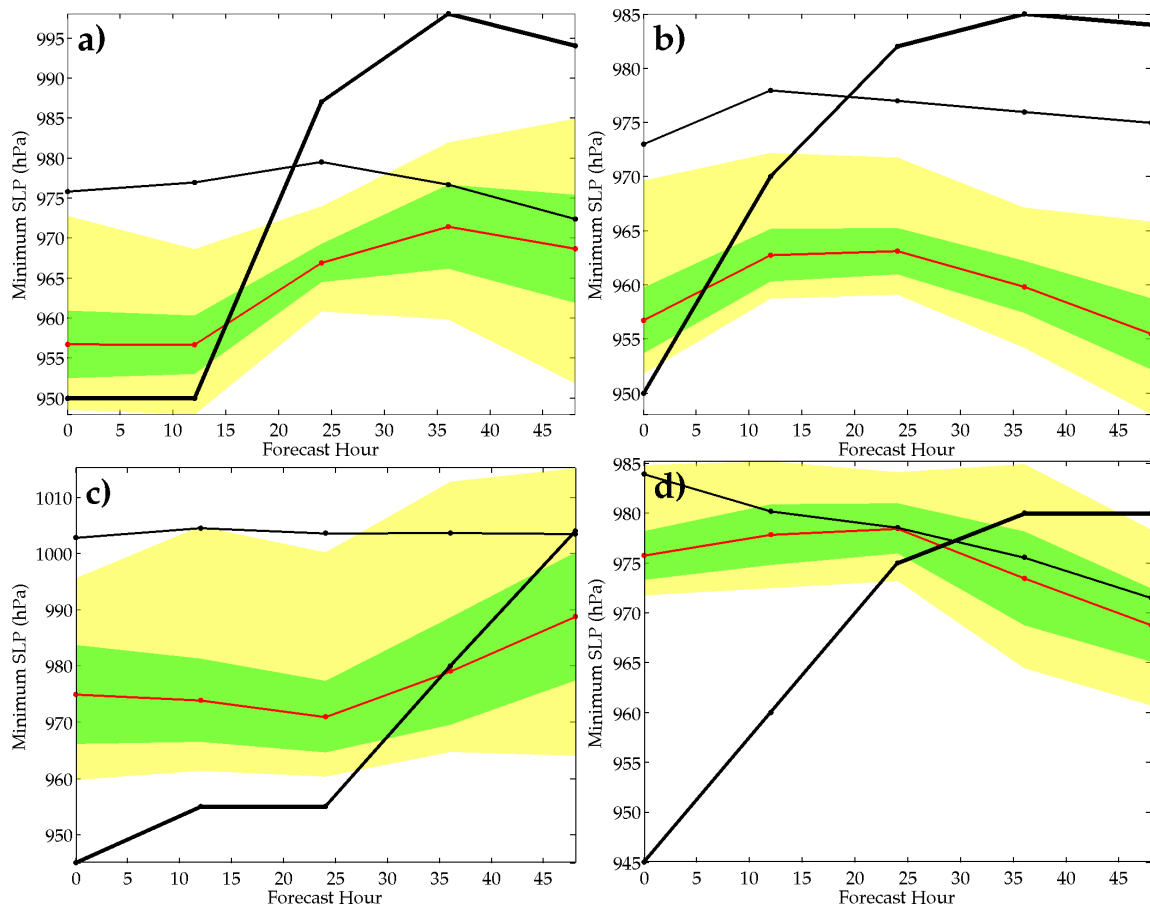


Figure 5.4: WRF ensemble-mean forecast of tropical cyclone minimum central pressure (red line) initialized on (a) 12 UTC 19 October 2004 (Tokage), (b) 00 UTC 6 September 2005 (Nabi), (c) 00 UTC 17 October 2005 (Kirogi), and (d) 12 UTC 30 November 2003 (Lupit). The green shading denotes values within one standard deviation of the ensemble-mean, while the yellow shading denotes values that are within the span of the ensemble. The thick black line is the JMA best track position and the thin solid line applies to the GFS forecast.

ents. During the next 48 hours, the WRF forecast shows the tropical cyclone moving to the northeast over Japan, weakening, and upon re-emerging over the Pacific ocean east of Japan, reintensifying into a 975 hPa baroclinic cyclone (Fig. 5.5). The baroclinic reintensification occurs when an upper-level trough, initially over Mongolia and moving at 20 m s^{-1} , becomes situated immediately upstream of Tokage's remnants in the 48-hour forecast, which results in a favorable interaction between the two features similar to a Petterssen type B process (Petterssen and Smeybe 1971). During the same period, a second 1010 hPa cyclone east of the China-Siberia border slowly intensifies as a trough initially located over Siberia interacts with it. This cyclone moves toward the Sea of Okhotsk during the next 24 hours, whereby it combines with the baroclinic cyclone resulting from Tokage. Coincident with the baroclinic reintensification is a large increase in the ensemble variance near to and downstream of the transitioning cyclone; the standard deviation of 48-hour WRF EnKF SLP and 500 hPa height forecasts north of 40°N exceeds 7 hPa and 40 m, respectively, thus the WRF EnKF ensemble contains diverging solutions of how the transitioning cyclone and downstream state will evolve.

The variability in forecasts of Tokage's evolution is further demonstrated by examining the ensemble's cyclone track and intensity forecasts. WRF EnKF ensemble forecasts are able to replicate the TC track and intensity until the cyclone makes landfall in Japan on 06 UTC 20 October; beyond hour 24, larger differences between ensemble members emerge (Fig. 5.3a). Most of the WRF ensemble forecasts have the TC moving to the left of the best track position; however, the low percentage in track forecasts at any particular grid point indicate significant variability among members. A majority of the 90 ensemble members contain an intensifying sub-975 hPa cyclone after 48 hours, while a few members have a weakening cyclone, but that is still 10 hPa lower than the best track (Fig. 5.4a). Moreover, the corresponding-time GFS forecast also has a deepening baroclinic cyclone during this period, thus the additional observations used by the GFS forecast did not lead to a better forecast of this ET.

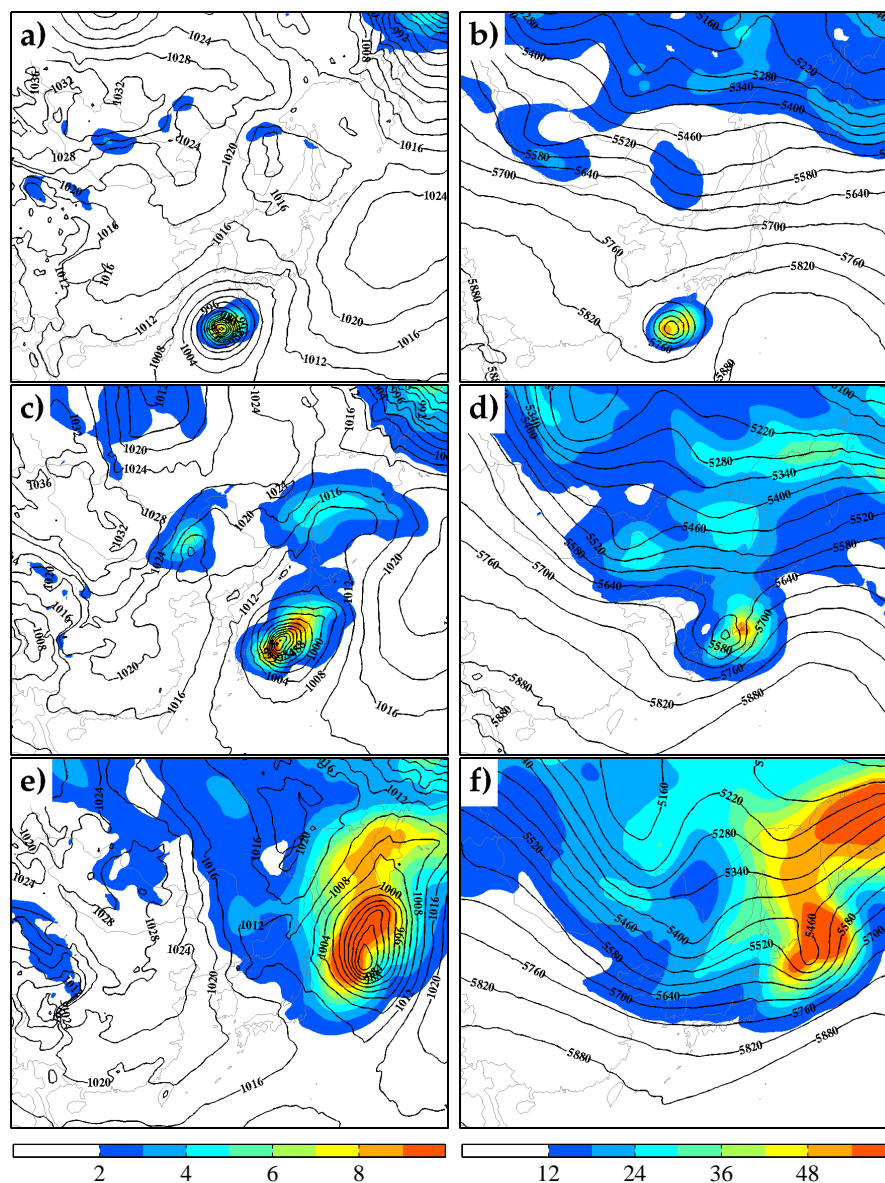


Figure 5.5: 00-hour (top row), 24-hour (middle row), and 48-hour (bottom row) WRF EnKF forecast of sea-level pressure (left column, units hPa) and 500 hPa height (right column, units m) for the ensemble member whose 48-hour forecast of SLP is closest to the ensemble-mean field (contours) initialized 12 UTC 19 October 2004 (Tokage). The shading denotes the ensemble standard deviation.

The diversity of ensemble solutions suggest that Tokage’s transition forecast has high initial condition sensitivity; therefore, this chapter will mainly focus on this forecast initialization time.

5.3.2 *Nabi*

Whereas the forecast of Tokage’s transition contained significant variability among the ensemble members, WRF EnKF forecasts of Nabi’s transition are characterized by less uncertainty (Fig. 5.6). WRF EnKF forecasts initialized at 00 UTC 6 September have Nabi moving poleward into the Sea of Japan and deepening into a 964 hPa cyclone on 00 UTC 8 September when the TC remnants phase with a slow-moving 500 hPa trough, shown in Fig. 5.6b along the northern border of Siberia. Unlike the Tokage transition forecast, the region of large ensemble spread in this 48-hour forecast is confined to the cyclone itself, thus Nabi’s transition does not introduce as much uncertainty to the downstream forecast. The WRF EnKF forecasts have the TC moving 200 km to the west of the best track position, with little variability among the ensemble members (Fig. 5.3b). Furthermore, all 90 ensemble members show Nabi weakening during the first 12 hours of the forecast while the storm is over land; however, similar to the corresponding GFS forecast, the ensemble-mean cyclone intensity decreases by 10 hPa during the final 36 hours of the forecast. The 48-hour ensemble-mean minimum SLP forecast is 28 hPa lower than the best track intensity estimate obtained from the GFS analysis on 00 UTC 8 September (Fig. 5.4b).

5.3.3 *Kirogi*

Although most of the ensemble members capture the evolution of Kirogi’s transition, 48-hour WRF EnKF forecasts have large variability in the TC position (Fig. 5.7). For the 48 hours following 00 UTC 17 October, the WRF EnKF forecasts show Kirogi slowly moving to the northeast in response to a synoptic-scale trough to its north. At forecast hour zero, the ensemble standard deviation in SLP is approximately 6 hPa

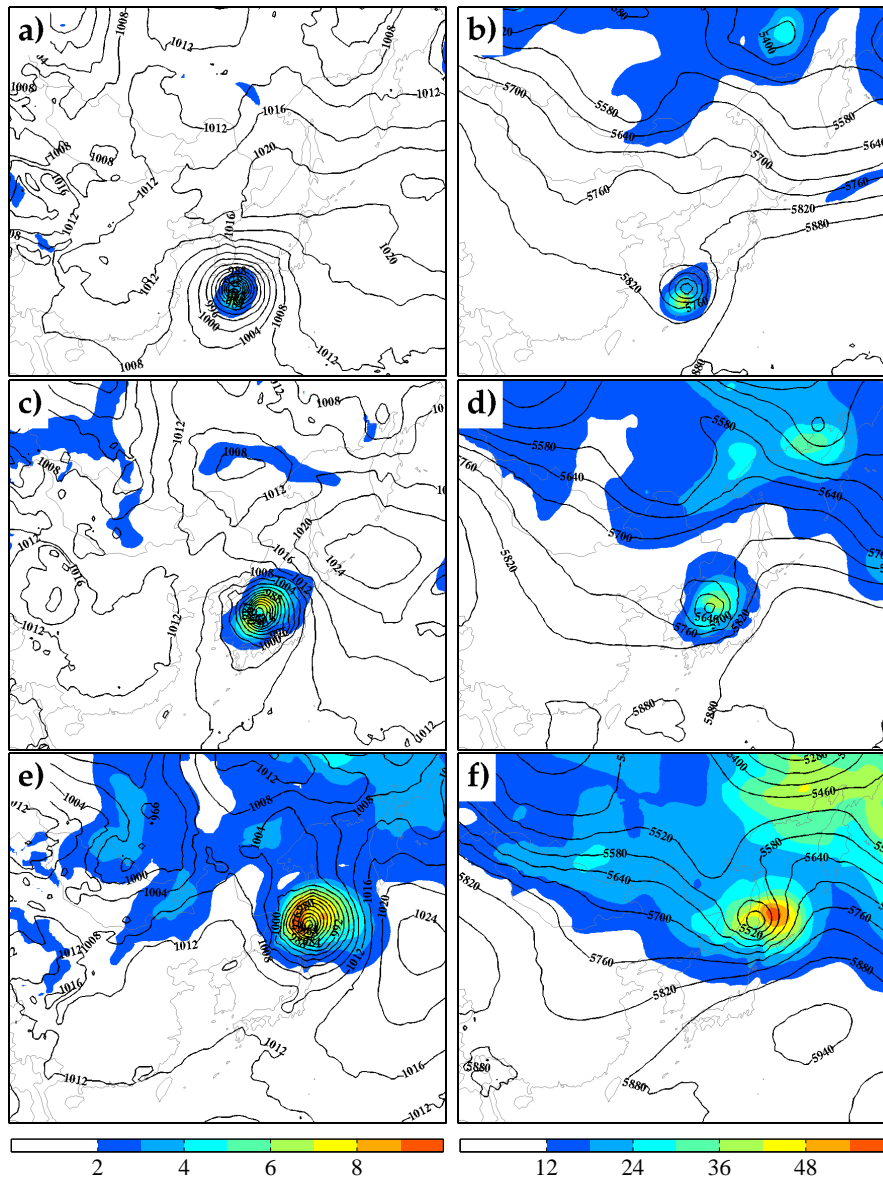
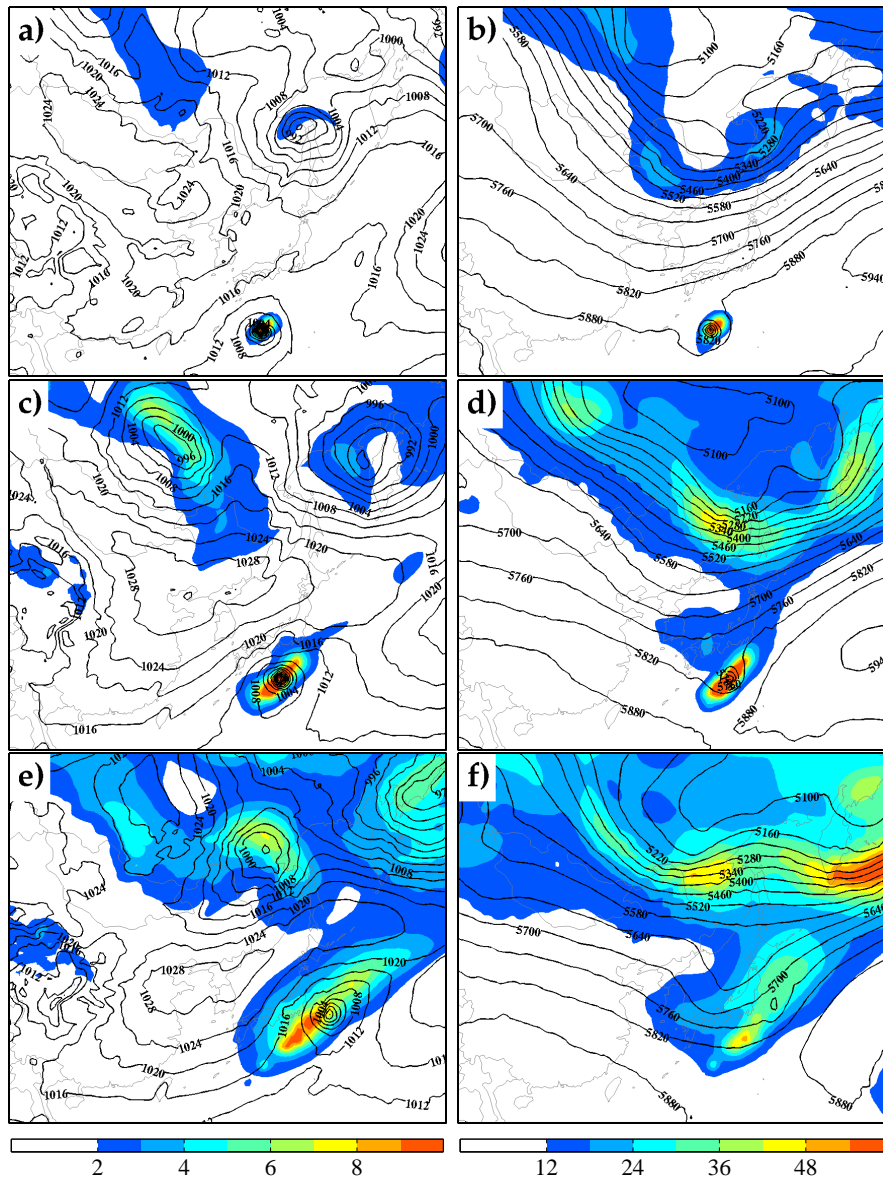


Figure 5.6: As in Fig. 5.5, but for the forecast initialized 00 UTC 6 September 2005 (Nabi).

and is confined to the TC, but at longer lead times, a region of ensemble spread in excess of 6 hPa grows parallel to the eastern coast of Japan (Fig. 5.7e). Kirogi does not have a chance to intensify into a baroclinic storm because the main trough is downstream of the TC as the storm enters the mid-latitudes, thus the storm is quickly sheared apart. Figure 5.3c shows that the region of large SLP variance in Fig. 5.7e reflects the uncertainty in the northeast movement of the TC during this 48 hour period. Whereas the best track position falls within the region characterized by 40% of the ensemble members passing within 90 km, some of the ensemble members have Kirogi's remnants 700 km to the northeast of the position for the 48-hour forecast valid 00 UTC 19 October. Although the ensemble-mean analysis of cyclone minimum SLP on 00 UTC 17 October is 30 hPa higher than the best track estimate, all of the ensemble members capture the TC weakening during the last 24 hours of the forecast (Fig. 5.4c). The GFS forecast shows Kirogi moving slower than the best track estimate; after 48-hours, the TC position is 400 km to the southwest of the GFS analysis position.

5.3.4 *Lupit*

WRF EnKF forecasts initialized 12 UTC 30 November accurately simulate the track, but not necessary the intensity of Lupit during its transition. Figure 5.8 shows that at forecast hour zero, Lupit is located 1300 km to the south of Japan, while a series of mid-latitude cyclones are to its north over Japan. Over the next 48-hours, the ensemble standard deviation for SLP associated with the cyclone increases from 4 hPa to 8 hPa as the cyclone moves to the northeast and transitions into a 980 hPa baroclinic storm to the east of Japan. The reintensification occurs in response to a weak upper-level trough located along the southern border of Mongolia and China in Fig. 5.8b. The two cyclones north of Japan combine into one 975 hPa storm at the southern tip of the Kamchatka peninsula by 12 UTC 1 December as the short-wave trough originally near the eastern China-Siberia border in Fig. 5.8b moves east.



With the exception of the first 12-hours, the best track position falls within the region of maximum forecast track probability determined from WRF EnKF forecasts (Fig. 5.3d). Although the ensemble-mean analysis of minimum SLP is 30 hPa higher than the best track estimate, the best-track intensity estimate falls within the span of the ensemble over the final 24 hours of the forecast when the storm is in its re-intensification phase (Fig. 5.4d). Furthermore, the GFS forecast of cyclone track and intensity initialized 12 UTC 30 November are also comparable to the ensemble-mean WRF EnKF forecast.

5.4 Forecast Sensitivities

The following section describes the sensitivity of transitioning TC forecasts to the initial conditions, which is determined by applying ensemble sensitivity methods to the WRF EnKF forecasts. This section seeks to understand where small errors in the initial conditions will have a large impact on the ET forecast. The verification results presented in the previous section indicate that the WRF EnKF forecasts contain large intensity errors, thus the first metric considered is the cyclone minimum SLP. An ensemble of minimum SLP forecast metric values is computed by finding the grid point in each ensemble member with the lowest SLP in the vicinity of the TC. The second metric considered is the RMS error in SLP forecasts within 800 km of the best track position of the cyclone, which will respond to both TC position and intensity errors. Similar to section 4.5, the RMS error within the circle is determined by verifying each ensemble member's SLP forecast within the 800 km radius circle against the appropriate time ensemble-mean analysis. It may also seem appropriate to determine the sensitivity of cyclone position forecast errors to the initial conditions; however, preliminary calculations showed potential problems and is thus not shown. For this metric, an ensemble member with a cyclone that is 200 km to the west of the verification position has the same position error as a member with a cyclone that is 200 km to the east. As a consequence, the linear regression coefficient between

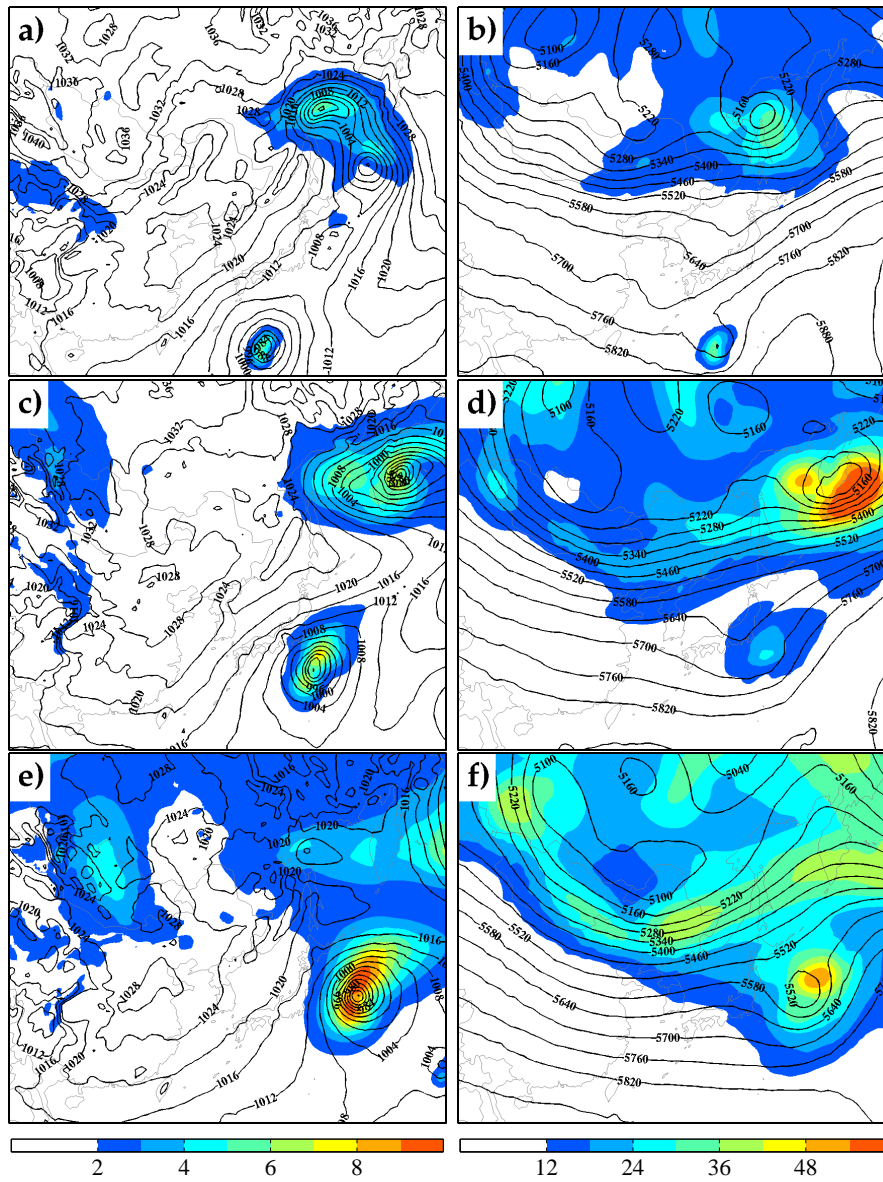


Figure 5.8: As in Fig. 5.5, but for the forecast initialized 12 UTC 30 November 2003 (Lupit).

cyclone position error and the analysis grid points, which are used to compute the initial condition sensitivity, often shows no meaningful relationship. The remainder of this section focuses primarily on the 48-hour forecast starting 12 UTC 19 October 2004 (the onset of Tokage’s transition) since WRF EnKF and GFS forecasts initialized at this time are characterized by large errors and ensemble spread.

5.4.1 *Cyclone Minimum SLP*

Figure 5.9 shows the sensitivity of Tokage’s minimum SLP forecast at various times to the analysis of SLP at the onset of transition. Similar to section 4.2, the sensitivity values at each grid point are multiplied by the analysis standard deviation at that grid point to allow for a quantitative comparison among fields and forecast hours. Grid points for which the sensitivity value is not statistically significant at the 95% confidence level, determined via (4.1), are set to zero. Regions of positive (negative) sensitivity indicate that increasing the analysis SLP at that grid point will lead to an increase (decrease) in the cyclone minimum SLP at the appropriate forecast hour. For the 12-hour minimum SLP forecast (Fig. 5.9a), a one standard deviation increase (decrease) in the analysis SLP near the center of the TC can lead to a 2.5 hPa increase (decrease) in the cyclone minimum SLP, thus short-term TC intensity forecasts have memory of the analysis intensity. At longer forecast lead times (36, 48 hours), the maximum sensitivity pattern is oriented in a dipole centered on the TC; increasing (decreasing) the southwest (northeast) by one standard deviation, achieved by shifting the storm to the northeast, leads to a 3 hPa increase in the cyclone minimum SLP (Fig. 5.9c-d). This sensitivity pattern indicates that the post-transition cyclone minimum SLP depends on the analysis TC position, rather than the analysis cyclone intensity. This sensitivity pattern is consistent with the Klein et al. (2000) climatology of western Pacific ET that showed no meaningful correlation between the cyclone intensity prior to ET and the cyclone intensity after ET. Both the 36 and 48-hour forecasts also have patterns of sensitivity in the mid-latitudes that have comparable

magnitude to those near the TC itself; one standard deviation change to the SLP field in between a cyclone, anti-cyclone, cyclone chain from southern Mongolia to the Siberian coast can lead to a 2.5 hPa change in the minimum SLP forecast. This sensitivity pattern implies that an eastward shift in the mid-latitude flow features is associated with a weaker cyclone.

Sensitivity patterns for 500 hPa height provide greater insight into how small initial condition errors in the mid-latitudes can effect the transition of Tokage (Fig. 5.10). For the 12 and 24-hour forecasts of Tokage's minimum SLP (Fig. 5.10a, b), the largest sensitivity values are associated with the reflection of Tokage in the 500 hPa flow; increasing (decreasing) the height in the center of the cyclone by one standard deviation can lead to a 3 hPa increase (decrease) in the forecast minimum SLP. In contrast, the sensitivity pattern for 36 and 48-hour forecasts is maximized in dipole patterns near Tokage's upper-level signature, and two shortwave troughs over Mongolia and Siberia (Fig. 5.10c-d). Moreover, a one standard deviation change to the 500 hPa height field in the regions of large sensitivity in the mid-latitudes can lead to the same change in the 48-hour minimum SLP forecast as a one standard deviation change in the 500 hPa height near the TC, thus initial condition errors in the mid-latitudes are of equal importance to initial condition errors near the TC.

First, consider the trough initially located over Siberia that is associated with large initial condition sensitivity; decreasing (increasing) the heights to the east (west) of this feature by one standard deviation, achieved by moving the trough to the east, leads to a 2 hPa increase in the 48-hour forecast of minimum SLP. Recall from the previous section that the evolution of the 1010 hPa cyclone initially along the eastern China-Siberia border (cf., Fig. 5.5a), which intensifies depending on how this trough phases with it. Moving this trough to the east allows the 1010 hPa cyclone to undergo cyclogenesis due to the favorable interaction between these two features. It is hypothesized that this cyclone intensifies at the expense of the baroclinic potential energy in the area east of Japan; therefore, when Tokage attempts to undergo baroclinic cy-

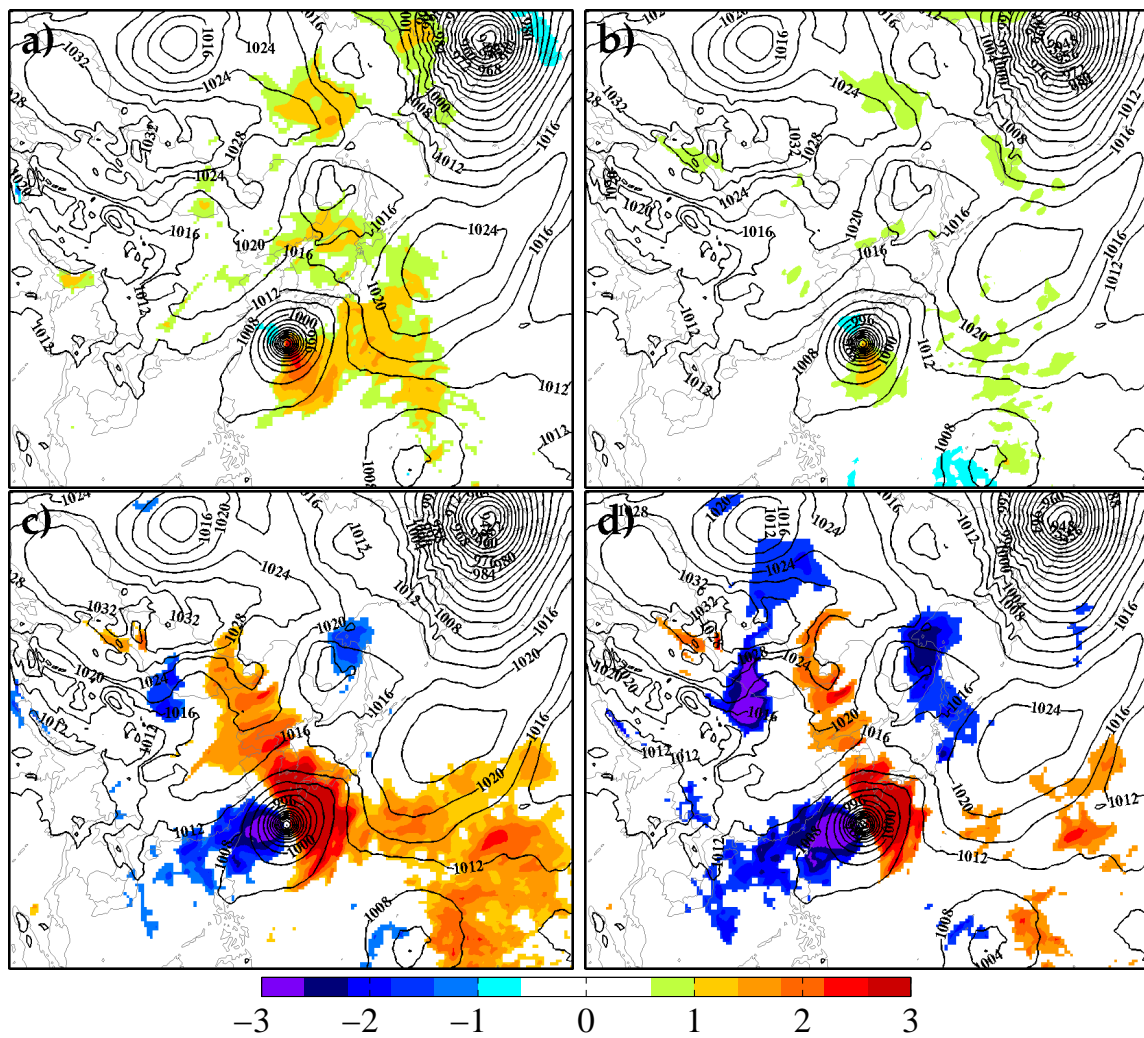


Figure 5.9: Sensitivity of the (a) 12, (b) 24, (c) 36, and (d) 48-hour forecast of Tokage's minimum sea-level pressure to the analysis of SLP times the analysis standard deviation (shading, hPa) for the forecast initialized 12 UTC 19 October 2004 (Tokage). The contours show the ensemble-mean analysis of SLP (hPa).

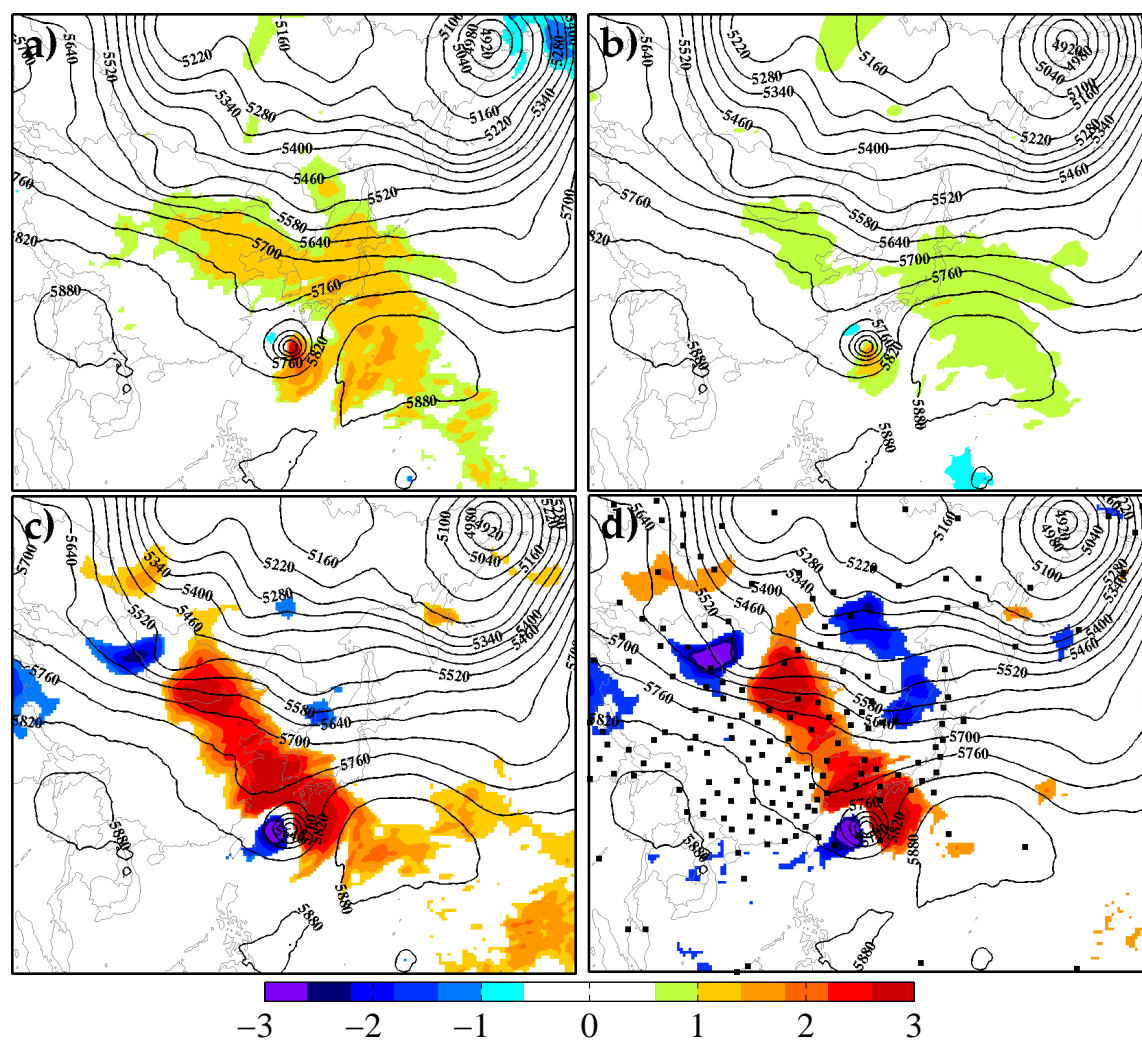


Figure 5.10: As in Fig. 5.9, but for the analysis of 500 hPa height. The squares in panel d denote the position of the rawinsonde stations available at the forecast initialization time. Units for the 500 hPa height field are m.

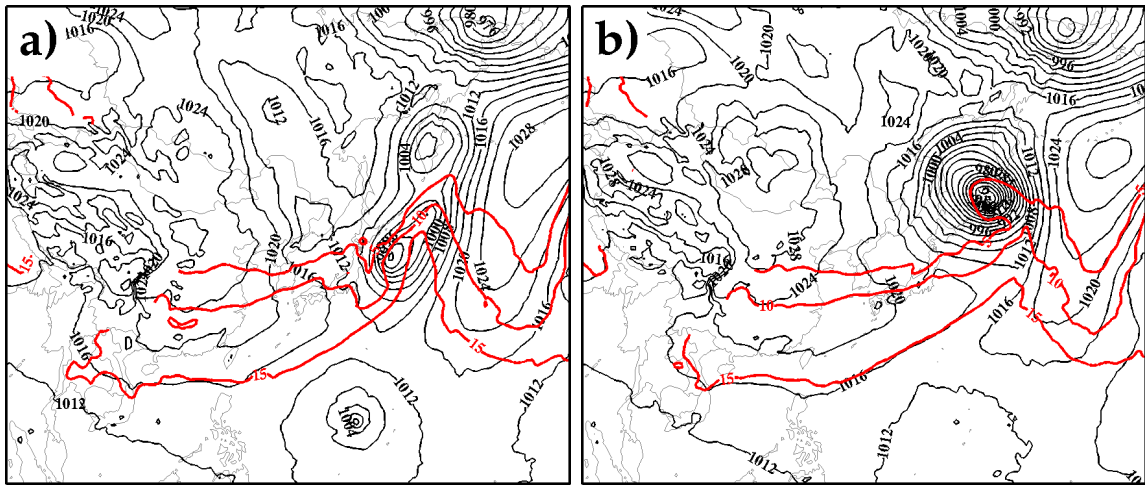


Figure 5.11: 48-hour SLP (black contours, units hPa) and 850 hPa temperature (red contours between 5 and 15°C) forecast for the WRF EnKF ensemble member with a 48-hour cyclone minimum SLP forecast of (a) 984 hPa, and (b) 952 hPa initialized 12 UTC 19 October 2004 (Tokage).

clogenesis after moving across Japan, the horizontal temperature gradient, and thus the available potential energy will be less. This hypothesis is supported by comparing the 48-hour forecast of SLP and 850 hPa temperature for the ensemble member that has Tokage becoming a 955 hPa cyclone, with the member that has Tokage becoming a 985 hPa cyclone (Fig. 5.11). In the weak-cyclone member's forecast (Fig. 5.11a), Tokage's remnants are moving toward the meridionally-elongated thermal ridge associated with the downstream cyclone, and is not conducive for baroclinic cyclogenesis to occur. In contrast, the strong-cyclone member has a more zonally-aligned horizontal temperature gradient that may be more beneficial to Tokage's transition (Fig. 5.11b).

The sensitivity dipole for the trough originally over central Mongolia has the opposite sign of the Siberian trough; increasing (decreasing) the heights to the east (west) of the trough by one standard deviation leads to a 3 hPa increase in the 48-hour forecast of Tokage's minimum SLP. As described in the previous section, this trough phases with Tokage's remnants during the forecast, thus shifting this trough to the

west will delay the interaction of these two features, thus the minimum SLP of Tokage will be higher. It is somewhat curious that the areas of large sensitivity for this case coincides with regions of few rawinsonde observations, denoted by the black squares in Fig. 5.10d. This result suggests that the absence of rawinsonde observations in these regions may have adversely affected the forecast of Tokage's transition. This idea will be further explored in section 5.5 where new initial conditions are produced based on the ensemble sensitivity structures.

Sensitivity patterns for the 48-hour minimum SLP forecast for Nabi to the analysis of 500 hPa height further suggest that changes to the mid-latitude analysis can be equally as important as changes to the TC (Fig. 5.12a). For the forecast initialized at 00 UTC 6 September, the 48-hour minimum SLP forecast is most sensitive to the 500 hPa height field at the base of an upper-level trough along the northern border of China and Mongolia. Increasing (decreasing) the geopotential heights by one standard deviation, which in turn will amplify (weaken) the trough, is associated with a 1.5 hPa decrease in the cyclone's minimum SLP 48 hours later. Recall from section 5.3 that this trough phases with Nabi's remnants during the forecast; therefore, this figure is suggestive that a stronger trough leads to a more intense baroclinic cyclone. A more highly amplified trough would have a larger horizontal vorticity gradient, which in turn could provide more vertical motion above the cyclone and greater height falls (e.g., Trenberth 1978). Unlike the Tokage forecast, the region of largest sensitivity is characterized by many rawinsonde observations; therefore, it might be expected that the analysis errors in that region should be smaller.

Whereas the previous two forecasts show large sensitivity to short-wave troughs in the mid-latitude flow, the 48-hour WRF EnKF forecast of Kirogi's minimum SLP is sensitive to the large-scale circulation pattern over eastern Asia, rather than individual short-wave troughs (Fig. 5.12b). The largest sensitivity values are in a dipole pattern surrounding the upper-level trough associated with Kirogi; increasing (decreasing) the 500 hPa heights to the south (north) by one standard deviation leads to a 8 hPa

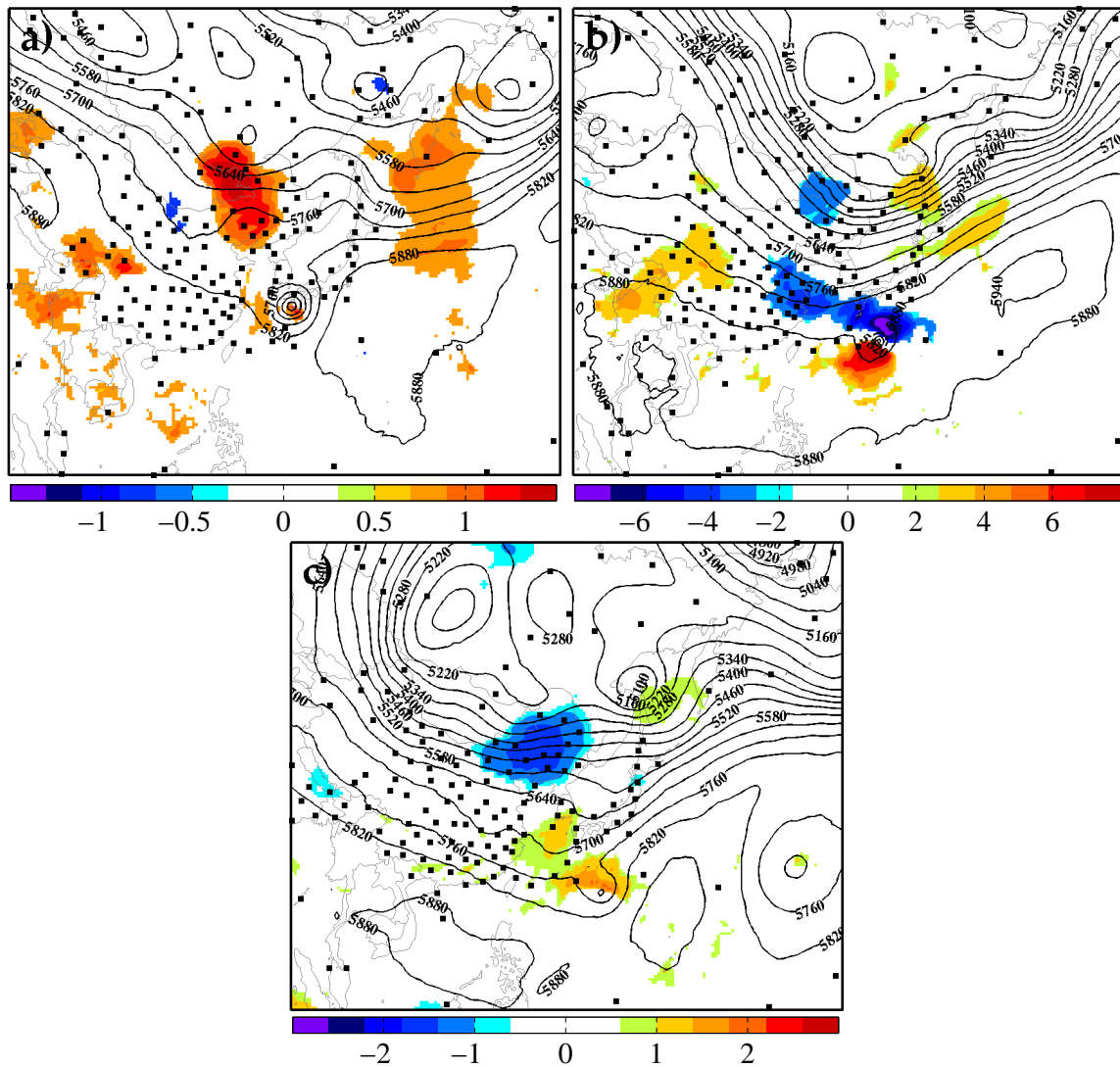


Figure 5.12: As in Fig. 5.10d, but for the forecast initialized (a) 00 UTC 6 September 2005 (Nabi), (b) 00 UTC 17 October 2005 (Kirogi), and (c) 12 UTC 30 November 2003 (Lupit).

increase in the minimum SLP forecast. This pattern may reflect the effect of shear on Kirogi's minimum SLP; when the TC is located further north at 00 UTC 17 October, it will encounter the jet sooner and rapidly weaken. Furthermore, the region of high sensitivity surrounding the TC is characterized by few rawinsondes, thus TC intensity forecasts could have benefited from additional reconnaissance data near the storm. Multiple areas of high sensitivity exist in the mid-latitudes on either side of the synoptic-scale trough centered near Korea. Increasing (decreasing) the heights to the east (west) of the longwave trough over Asia by one standard deviation leads to a 3 hPa increase in the 48-hour forecast of the cyclone's minimum SLP. This sensitivity pattern suggests that when the synoptic-scale trough is further west in the analysis, Kirogi moves directly into the jet on the eastern side of this trough and thus will quickly weaken in the forecast.

In contrast to the other ET forecast, the forecast of Lupit's transition initialized at 12 UTC 30 November displays minimal sensitivity to the trough that phases with TC remnants. Figure 5.12c shows that the largest sensitivity values are associated with the reflection of Lupit in the 500 hPa height field and a shortwave trough in eastern China. Increasing (decreasing) the 500 hPa heights near Lupit or decreasing (increasing) the heights in the base of the Chinese trough by one standard deviation will lead to a 2 hPa increase in the 48-hour forecast of cyclone minimum SLP. Recall that the upper-level trough characterized by large sensitivity values phases with the cyclone to the north of Japan, rather than Lupit, thus the reintensification of Lupit as a baroclinic cyclone is sensitive to the evolution of the cyclone to the north. Much like the Tokage forecast, it is hypothesized that the northern cyclone intensifies at the expense of the horizontal temperature gradients in the region where Lupit undergoes transition.

5.4.2 RMS error in SLP forecasts

The remainder of this section considers the initial condition sensitivities for the RMS error in the SLP field surrounding the TC, which are shown to correspond well with the regions of high initial condition sensitivity for the cyclone minimum SLP. For brevity, only the sensitivity to 500 hPa height is shown here in Fig. 5.13. For this metric, positive (negative) sensitivity values indicate that decreasing (increasing) the 500 hPa height analysis at that grid point will lead to a decrease in the RMS error within the thick black circle shown in Fig. 5.13; this circle is centered on the cyclone best track position and has radius 800 km. The RMS error in 12 and 24-hour SLP forecasts shows maximum sensitivity to the initial position of Tokage and the height field to the north of the storm; increasing (decreasing) the geopotential height to the east (west) of Tokage by one standard deviation can lead to a 1 hPa decrease in the RMS error in the circle. At longer lead times (36, 48 hours, Fig. 5.13c, d), the RMS error in SLP has maximum sensitivity to both the tropical cyclone and the previously described mid-latitude troughs (cf. Fig. 5.10d). The RMS error in SLP within the circle can be reduced by 1.5 hPa by decreasing (increasing) the heights to the east (west) of the Siberian trough by one standard deviation and/or by decreasing (increasing) the heights to the west (east) of the Mongolian trough by one standard deviation. Recall that the WRF EnKF 48-hour forecasts have a ET cyclone that is deeper than the best track estimate. The combination of Fig. 5.10d and Fig. 5.13d indicate that reducing the RMS error in SLP forecasts surrounding the cyclone would also lead to weaker cyclone. The potential of using these sensitivity structures to apply *a posteriori* diagnostic corrections to the initial conditions that will reduce the error in SLP forecasts will be explored in the next section.

Regions of high initial condition sensitivity for the RMS error in 48-hour forecasts for the remaining cases also located near the areas of large initial condition sensitivity for 48-hour cyclone minimum SLP forecasts. Fig. 5.14a shows that for the forecast of

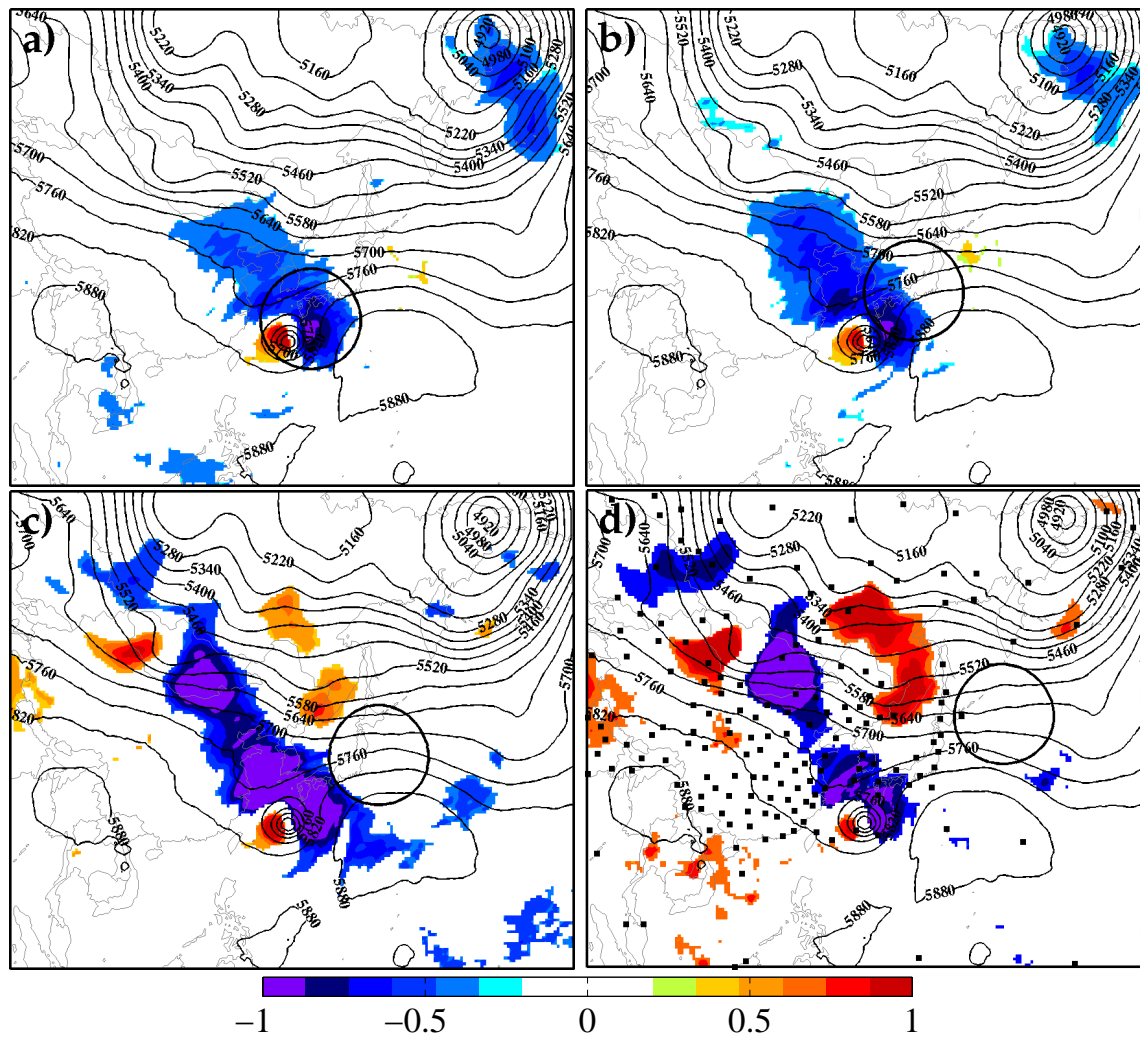


Figure 5.13: As in Fig. 5.10, but for the RMS error in SLP forecasts within 800 km of the best-track position of the cyclone, denoted by the black circle in each panel. Shading units are hPa.

Nabi's transition initialized 00 UTC 6 September, the error in 48-hour SLP forecast is most sensitive to the 500 hPa heights south of the short-wave trough along the Siberia-Mongolia border; increasing (decreasing) the heights in this trough by one standard deviation is associated with an 0.8 hPa decrease (increase) in the RMS error in SLP and is consistent with the change required to increase Nabi's 48-hour minimum SLP forecast. Initial condition sensitivities for the RMS error in Kirogi's forecast are largest in a dipole pattern surrounding Kirogi's position (Fig. 5.14b). Increasing (decreasing) the height to the south (north) by one standard deviation can reduce the RMS error in 48-hour SLP forecasts by 1 hPa. Furthermore, the RMS error in this forecast is also reduced by decreasing (increasing) the heights to the west (east) of the longwave trough over Asia. In contrast to the other three cases, the RMS error in 48-hour SLP forecast for Lupit shows minimal sensitivity to the initial conditions; one standard deviation changes to the 500 hPa height field would lead to a less than 0.2 hPa reduction in SLP error. This result is not surprising since the WRF EnKF ensemble forecasts of Lupit's transition have a more accurate forecast of cyclone track, as compared to the other storms considered here.

5.5 *Perturbed Initial Condition Experiments*

In the following section, the ensemble sensitivity structures identified above will be used to apply diagnostic corrections to the initial conditions to produce a desired effect in the 48-hour forecast of cyclone minimum SLP and the RMS error in 48-hour SLP forecasts. Moreover, the following experiments determine how better initial conditions can reduce the error in the WRF EnKF forecasts explored here. Since the WRF model contains errors in its formulation, the initial conditions that give the best forecast are not necessarily the initial conditions that are the best estimate of the state. These perturbed initial conditions are integrated forward using the non-linear model whereby the resulting forecast metric value is compared to the change predicted by ensemble sensitivities.

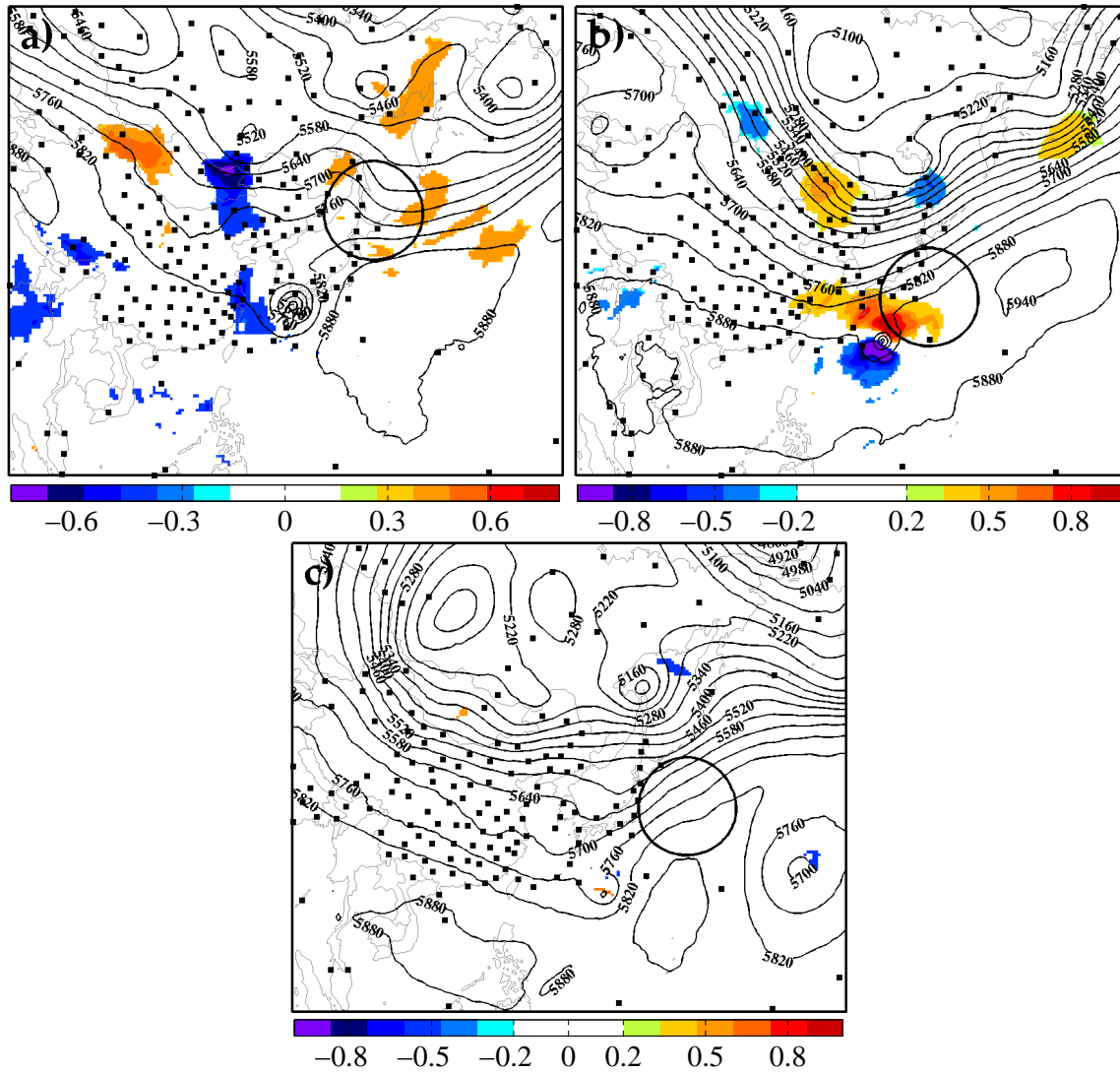


Figure 5.14: As in Fig. 5.12, but for the RMS error in SLP forecasts within 800 km of the best-track position of the cyclone, denoted by the black circle in each panel. Shading units are hPa.

For each of the ET forecasts described here, diagnostic corrections are computed and applied to one ensemble member's initial conditions using the following procedure. Since the sensitivity values apply to small changes about the ensemble-mean forecast metric value, the control forecast for these experiments is the ensemble member whose 48-hour forecast metric value (either the minimum SLP or RMS error in SLP) is closest to the ensemble-mean. Ensemble-mean analyses are not used here because it has smoothed out versions of the features in the initial conditions, especially near the TC. A perturbed version of each element of the initial condition state vector (x^p) is found via

$$\mathbf{x}^p = \mathbf{x}^a + \frac{\partial \mathbf{x}^a}{\partial J} \alpha, \quad (5.1)$$

where for an element of the state vector j ,

$$\frac{\partial x_j^a}{\partial J} = \frac{\text{cov}(\mathbf{X}_j^a, \mathbf{J})}{\text{var}(\mathbf{J})}, \quad (5.2)$$

x^a analysis state variable for the ensemble member closest to the mean, \mathbf{X}_j^a is a vector of the analysis ensemble's estimate of the j th analysis state variable, \mathbf{J} is the ensemble estimate of the 48-hour forecast metric and α is the desired change in the forecast metric. The equation for $\frac{\partial x_j^a}{\partial J}$ represents linear regression coefficient where the independent variable is the desired change in the forecast metric and the dependent variable is the perturbation to the analysis state vector required to realize that change. The perturbed initial conditions are integrated forward 48 hours using the WRF model whereby the forecast metric value is computed and compared to the control forecast metric value. This process can be repeated for several values of α to determine the range for which the ensemble predictions of the forecast metric change are valid. The procedure described here is similar to experiments performed by Rabier et al. (1996), Zou et al. (1998) and Langland et al. (2002), who used singular vector and adjoint-based sensitivities to adjust the initial conditions in a manner that would

reduce the error in mid-latitude cyclone forecasts. The results of perturbing the initial conditions to improve 48-hour cyclone minimum SLP forecasts is presented first before moving on to similar experiments for the RMS error in 48-hour SLP forecasts.

Figure 5.15 shows the control SLP and 500 hPa height analysis and the perturbation that must be added to each field in order to increase the 48-hour cyclone minimum SLP forecast by 18 hPa (the error in the 48-hour cyclone minimum SLP forecast for this member). The largest magnitude perturbations are located near Tokage; to raise the minimum SLP by 18 hPa, the SLP should be decreased (increased) on the southwest (northeast) side of Tokage by 12 hPa. Moreover, reducing the intensity of the cyclone 48-hours later also requires increasing the SLP near the downstream Asian cyclone by 6 hPa and on the southern side of the 948 hPa cyclone in the Bering Sea by 10 hPa. These initial condition changes are roughly two-times the SLP analysis standard deviation at these grid points. The perturbed initial conditions for 500 hPa height are 100 m larger (smaller) than the control analysis to the northeast (southwest) of Tokage (Fig. 5.15b). Furthermore, the geopotential height on the eastern side of the Siberian (Mongolian) trough is 30 m lower (50 m higher) than the control forecast, which will achieve an eastward (westward) shift to this features. Similar to the SLP perturbations, the 500 hPa height changes are also twice the analysis standard deviation in these locations.

After 48 hours, the differences between the control and perturbed initial condition forecasts have increased significantly and are primarily located near the transitioning cyclone. The 48-hour SLP forecast for the perturbed member is up to 40 hPa higher than the control forecast at the center of the cyclone and up to 25 hPa lower on the northern and southern side of the forecast ET cyclone (Fig. 5.15c). This tripole difference pattern indicates that the perturbed initial condition forecast has an ET cyclone that is further south, and also contains a separate baroclinic cyclone over the Kamchatka peninsula. In the perturbed initial condition forecast, the previously described weak cyclone over far eastern China (cf. Fig. 5.5a) does not combine with

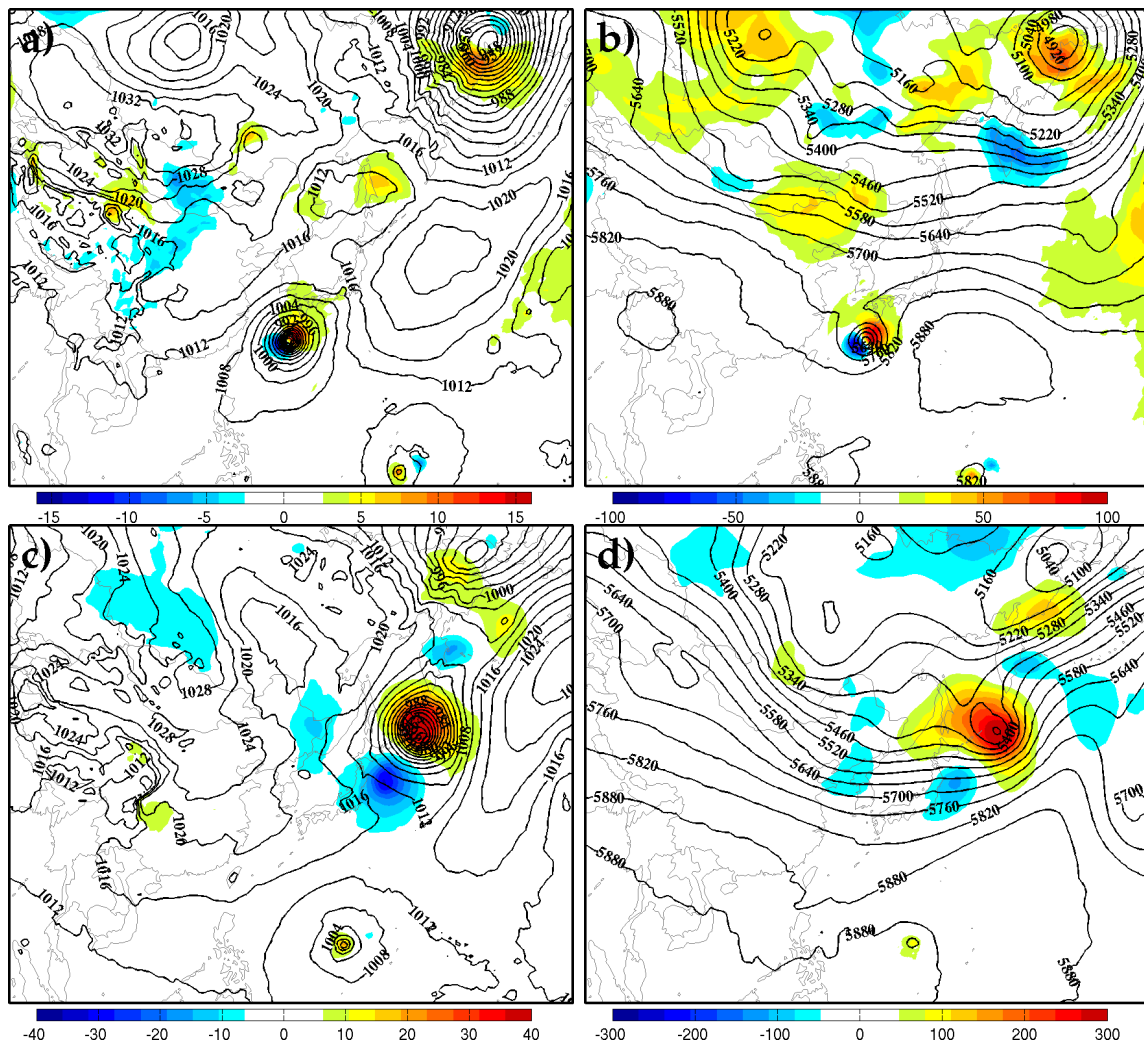


Figure 5.15: Difference between the perturbed and control analysis of (a) SLP (units hPa), and (b) 500 hPa height (units m) for the forecast initialized at 12 UTC 19 October 2004 (shading). The control analysis is given by the solid lines. The perturbed initial condition is constructed such that the 48-hour cyclone minimum SLP forecast is 18 hPa higher than the control forecast. (c) and (d) as in (a) and (b), but for the 48-hour forecast of SLP and 500 hPa height, respectively.

Tokage's remnants. Consistent with a weaker baroclinic cyclone, the 500 hPa heights near the upper-level trough associated with the Tokage's remnants are 300 m higher than the control forecast (Fig. 5.15d). In addition, there are also regions of alternating negative-positive differences downstream of the ET cyclone, co-located with ridge and troughs in the control forecast. This distribution of forecast differences implies that the downstream flow in the perturbed initial condition forecast is less amplified than the control forecast. The 48-hour perturbed initial condition minimum SLP forecast is 11 hPa higher than the control forecast value, but is smaller than the ensemble prediction of 18 hPa change.

Repeating these perturbed initial condition experiments with various values of α and comparing the resulting 48-hour minimum SLP forecast with the control simulation shows that between -5 hPa and 10 hPa, there is good agreement between the predicted change in the minimum SLP forecast and the actual difference in WRF solutions (Fig 5.16a). Beyond these values, the ensemble prediction falls below the main diagonal, which indicates that the WRF response is less than the prediction and linear perturbation dynamics are violated. As a consequence, it is not possible to use ensemble sensitivities to produce an initial condition change that will eliminate the 18 hPa error in the 48-hour cyclone minimum SLP forecasts without applying an iterative procedure.

Perturbed initial condition forecasts are generated for the three other ET forecasts explored in this study using the procedure described above for the Tokage forecast. In each forecast, the maximum magnitude of α is the error in the 48-hour ensemble-mean minimum SLP forecast. Perturbed initial condition forecasts for Nabi show good agreement between the predicted and actual changes between -8 hPa and 8 hPa; however for values greater than 8 hPa, the predicted change in the minimum SLP is smaller than the actual difference obtained for the forecast (Fig. 5.16b). The significant asymmetry between positive and negative perturbations indicates that at most Nabi can be deepened by 10 hPa, whereas it could be weakened by more than

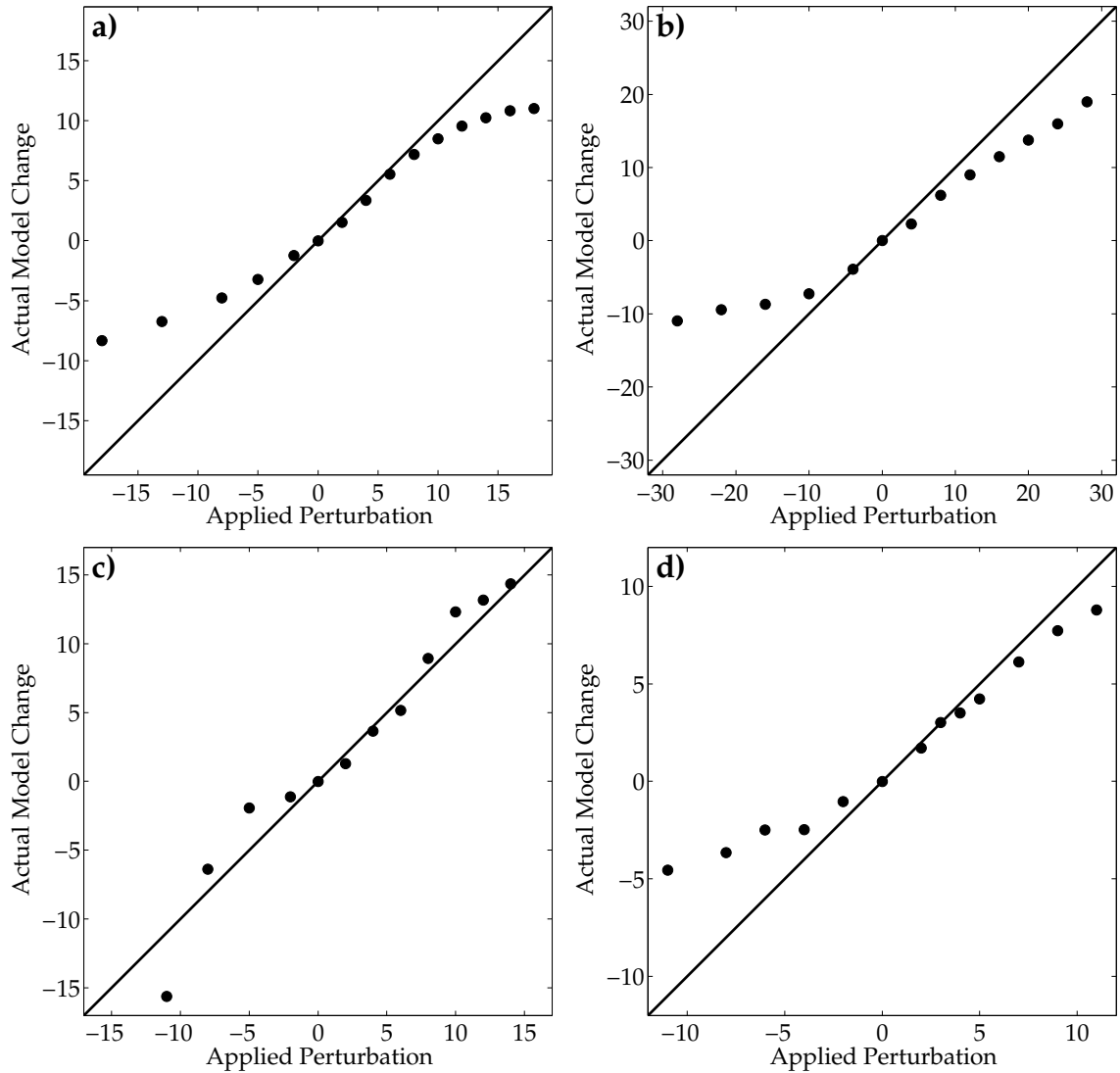


Figure 5.16: 48-hour cyclone minimum SLP differences as determined by perturbed integrations of the WRF model (ordinate) against the differences predicted by ensemble sensitivity analysis (abscissa) for forecasts initialized on (a) 12 UTC 19 October 2004 (Tokage), (b) 00 UTC 6 September (Nabi), (c) 00 UTC 17 October 2005 (Kirogi), and (d) 12 UTC 30 November 2003 (Lupit). The initial conditions for the perturbed integrations are created using ensemble statistics to obtain a desired change in the 48-hour forecast of cyclone minimum SLP. The solid line indicates perfect agreement between the predicted and WRF model integrations. Units are hPa.

20 hPa. Ensemble predictions of how initial condition changes affect the 48-hour minimum SLP forecast for Kirogi are in good agreement with the actual change obtained from the model over much of the range explored here, although for large negative perturbations (stronger cyclone), the actual change is much larger than the predicted change (Fig. 5.16c). This occurs because when large negative perturbations are applied, the cyclone is shifted far enough to the south so that is not affected by the mid-latitude trough to the north. For the Lupit forecast, the difference between the predicted and actual change is generally less than 1 hPa between -4 hPa and 6 hPa, with a large asymmetry between the positive and negative perturbations; positive perturbations follow linear dynamics in the ensemble sensitivities, whereas negative perturbations do not (Fig. 5.16d).

Ensemble sensitivities are also used to generate initial conditions via (5.1) that will decrease the RMS error in 48-hour SLP forecasts within 800 km of the cyclone. The procedure used for these experiments is similar to what is used for the perturbed 48-hour cyclone minimum SLP forecasts, except that J is the RMS error in 48-hour SLP forecasts and the control forecast is the ensemble member whose RMS error in the 48-hour SLP forecast is closest to the ensemble mean value. Recall from section 5.4 that decreasing J , which is akin to negative values of α , correspond to lower errors in SLP forecasts. Figures 5.17a and b show the initial condition perturbation required to reduce the RMS error in the Tokage SLP forecast by 11 hPa (the error in the 48-hour ensemble-mean forecast). The largest differences for SLP are in a dipole pattern surrounding Tokage; the perturbed initial condition has SLP values that are 10 hPa lower (higher) on the west (east) side of the cyclone. Moreover, there is a tripole of 3 hPa negative-positive-negative differences from Mongolia to the Asian coast, which leads to an eastward shift of these mid-latitude features. For 500 hPa height, differences of up to 60 m exist near the upper-level reflection of Tokage, while a 30 m increase (decrease) in 500 hPa heights to the west (east) of the Siberian trough and a 30 m decrease (increase) in the heights to the west (east) of the Mongolian trough

are also present.

Similar to the perturbed initial condition forecasts for the 48-hour cyclone minimum SLP, differences between this perturbed initial condition experiment and the control forecast are maximized near the forecast cyclone. The perturbed initial condition forecast has SLP values that are up to 30 hPa higher in the center of the forecast cyclone, with negative differences of up to 15 hPa on the northern and southern sides (Fig. 5.17c). This distribution implies that the perturbed initial condition forecast does not reintensify Tokage's remnants as much as the control forecast and the cyclone would be further south. At 500 hPa, height differences of up to 150 m coincide with the upper-level trough associated with the ET cyclone; however, there are also regions of negative and positive differences downstream of this trough that coincide with a ridge and trough in the 500 hPa heights for the control forecast (Fig. 5.17d). This distribution implies that the forecast characterized by lower SLP errors has less amplitude in the downstream flow. The RMS error in the perturbed initial condition forecast is 6 hPa smaller than the control forecast, thus systematic changes to the initial conditions can reduce the error in SLP forecasts by 60%.

The process of generating perturbed initial conditions to reduce the RMS error in 48-hour SLP forecasts is repeated using the above procedure for various values of α and all of the ET forecasts studied here. The maximum value of α for each initialization time is the RMS error in the 48-hour ensemble-mean SLP forecast. Comparisons of the predicted change in the forecast metric to the actual difference in model forecasts are shown in Fig. 5.18. With the exception of Lupit, altering the initial conditions based on ensemble sensitivities can reduce the RMS error in SLP by up to 60%. Furthermore, there is considerable agreement between the predicted reduction in SLP error estimated from ensemble statistics and the actual change obtained from the model forecasts for values of α up to 5 hPa, 3 hPa, and 2 hPa for the Tokage, Nabi and Kirogi forecasts, respectively. For larger magnitude perturbations, adjusting the initial conditions to reduce the RMS error has little impact, and in the case of Kirogi,

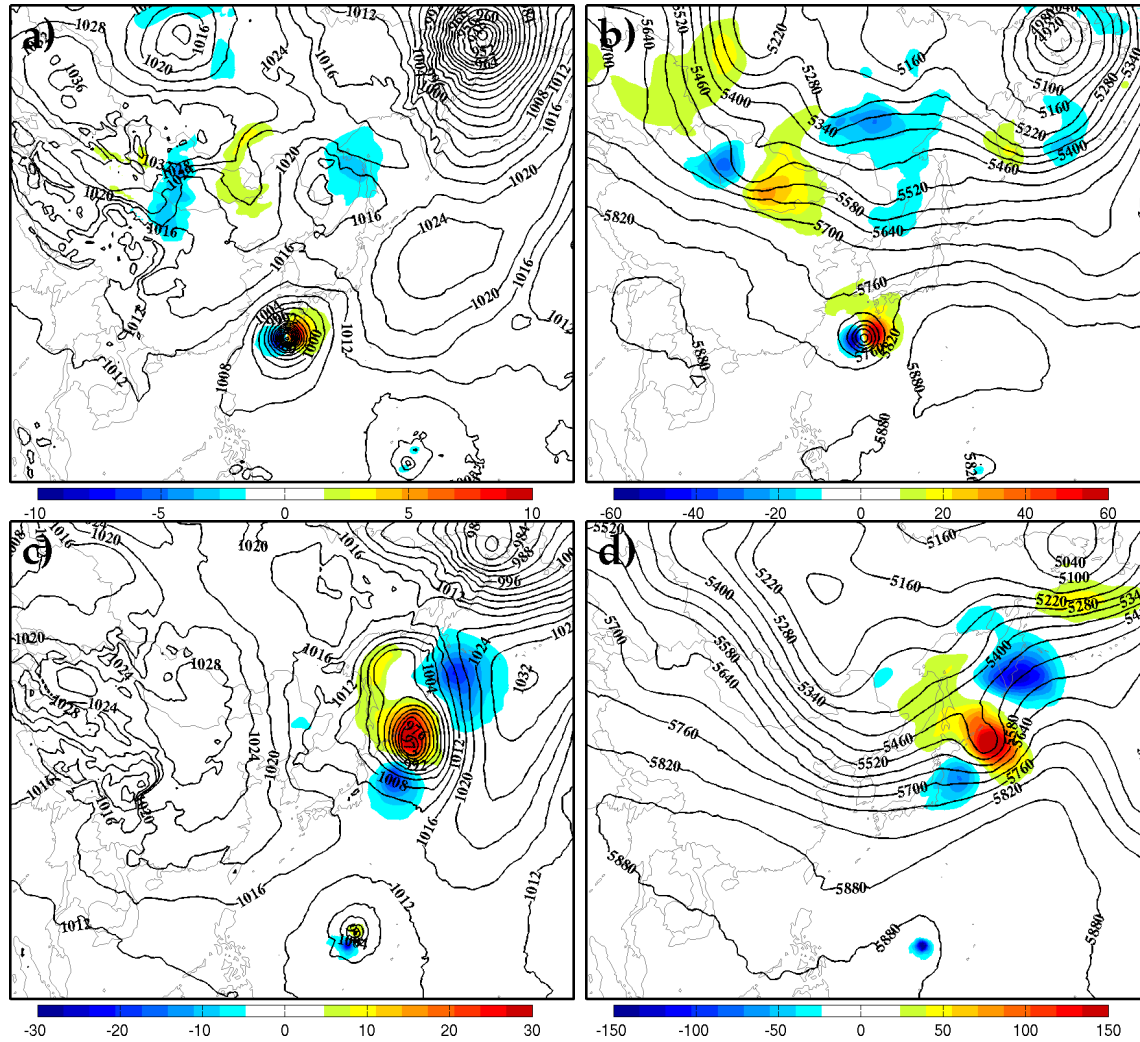


Figure 5.17: Difference between the perturbed and control analysis of (a) SLP (units hPa), and (b) 500 hPa height (units m) for the forecast initialized at 12 UTC 19 October 2004 (shading). The control analysis is given by the solid lines. The perturbed initial condition is constructed to decrease the RMS error in 48-hour SLP forecasts within the solid circle in Fig. 5.13d by 11 hPa. (c) and (d) as in (a) and (b), but for the 48-hour forecast of SLP and 500 hPa height respectively.

can actually lead to an increase in the error. The exception to these results is the Lupit transition forecast; altering the initial conditions does not reduce the RMS error in SLP beyond the ensemble-mean value of 5 hPa. This result is obtained because for this case, the RMS error in 48-hour SLP forecasts showed little sensitivity to the initial conditions (cf., Fig. 5.14c).

5.6 *Observation Impact*

This section uses ensemble sensitivities to determine how observation assimilation at the initialization time affects the WRF EnKF 48-hour minimum SLP forecast. In these experiments, the expected value and spread of the minimum SLP forecast for the case where observations are assimilated at hour 0 is compared to the control case where observations are not assimilated at hour 0. This section proceeds by identifying the observations that lead to the largest changes in the minimum SLP metric during each of the forecasts explored here, and then describes the results for several forecast experiments whereby different sets of observations are assimilated at hour 0 and the resulting analysis ensemble is integrated forward 48 hours.

The observations that produce the largest change in each cyclone's 48-hour minimum SLP forecast are identified using the following procedure. Consider all of the available observations at the ET forecast initialization time. The sensitivity of the 48-hour cyclone minimum SLP forecast to the model estimate of each observation can be computed using (2.23) and tested for statistical significance at the 99% confidence level using (4.1). If the sensitivity value is deemed to be statistically significant, the hypothetical change in the expected value and spread of this forecast metric is calculated from (2.25) and (2.28), respectively. Next, the observation that will produce the largest reduction in forecast metric variance is identified and assimilated serially using an EnKF. In addition, the change in the forecast metric values are updated in the same manner as the analysis state variables. After assimilating this observation, the process of computing the hypothetical change in the forecast metric due to an individ-

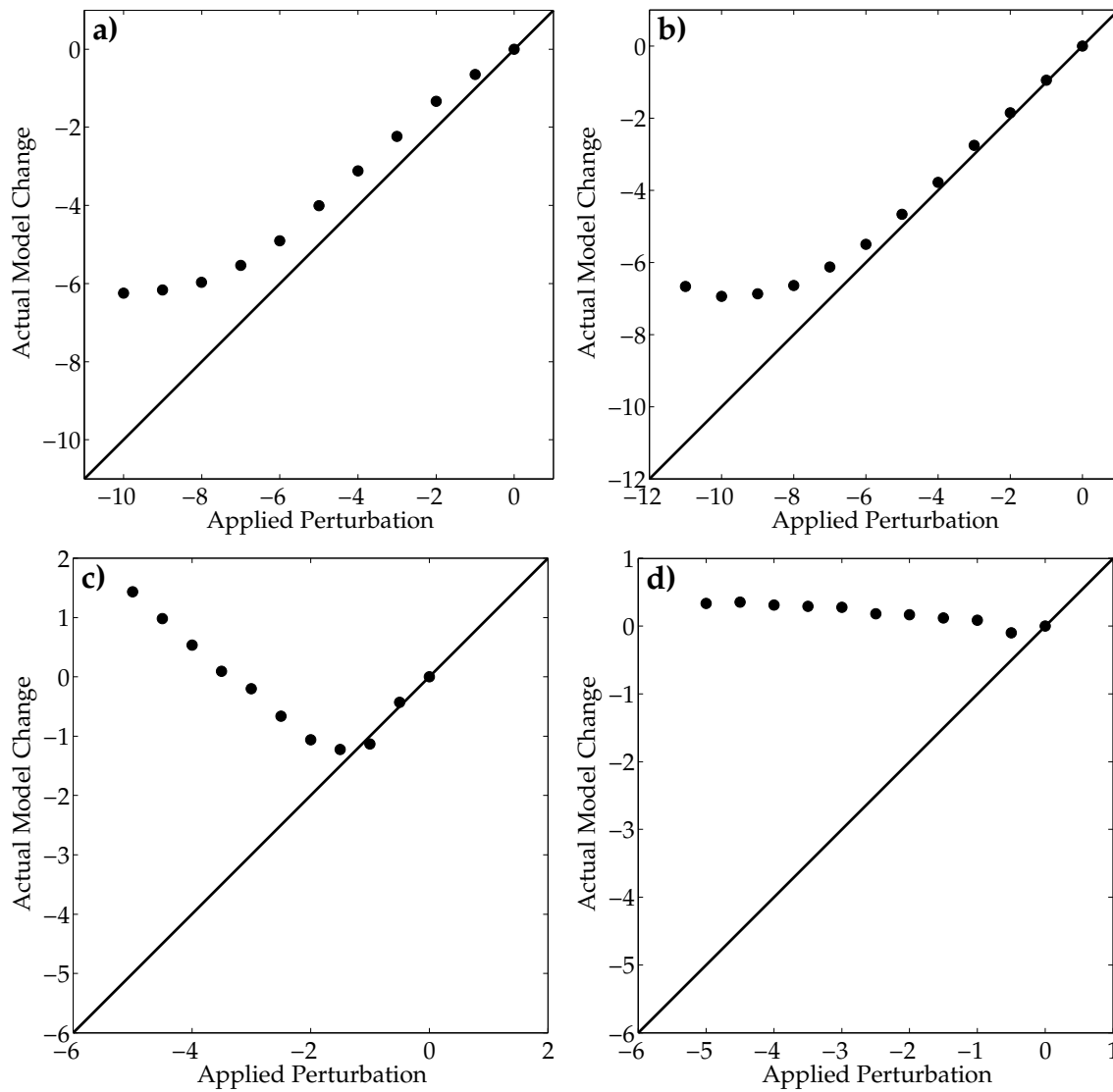


Figure 5.18: As in Fig. 5.16, but for the RMS error in 48-hour SLP forecasts within 800 km of the best track cyclone position, denoted by the appropriate circle in Figs. 5.13d and 5.14

Table 5.4: Change in the expected value (spread) in the 48-hour cyclone minimum SLP forecast due to the assimilating different sets of observations at hour 0. The control forecast ensemble does not assimilate any observations at hour 0. Units are hPa.

Tropical Cyclone	Tokage	Nabi	Kirogi	Lupit
Number of Sig. Obs.	32	42	32	39
All Observations	-0.4 (-5.6)	-3.0 (-2.6)	-2.9 (-9.1)	-3.0 (-3.9)
Best Obs. Only	2.2 (-2.0)	-2.6 (-1.1)	6.8 (-5.0)	-3.9 (-2.5)
No Best Obs.	-0.5 (-5.6)	-2.0 (-2.6)	0.0 (-8.7)	-2.9 (-3.9)
No Best Platforms	0.0 (-5.5)	-2.2 (-2.6)	-0.1 (-9.2)	-4.6 (-3.9)
Random Obs.	0.1 (-0.5)	-0.7 (-0.4)	0.1 (-0.1)	-0.9 (-0.6)

ual observation is repeated and the observation that will produce the largest reduction in forecast metric variance is identified and assimilated. This procedure continues until all statistically significant observations are exhausted. Table 5.4 shows the number of statistically significant observations identified for each ET forecast; approximately 40 observations are considered in each case. It should be noted that these results are highly dependent on the order of assimilation since nearby observations can produce similar updates to the model state vector, and thus changes to the forecast metric. If the change to the forecast metric due to previously assimilated observations are not accounted for, this technique will over-estimate the effect of individual observations because the observation data is correlated.

Figure 5.19 shows the estimated change in the expected value of the 48-hour cyclone minimum SLP forecast by the identified statistically significant observations. Each of the symbols represents one observation from an individual platform; for example, a square symbol could represent a rawinsonde zonal wind observation at one pressure level. In all cases, the observations that produce the largest changes in the expected value of the cyclone minimum SLP are associated with rawinsondes in

China and Japan, while cloud winds, ACARS and surface observations typically produce smaller changes. This result is related to the observation errors, information content in the observation data and sensitivity of the forecast metric to the initial conditions; rawinsonde observations have smaller wind errors than ACARS and cloud wind observations and provide vertical profiles of the atmosphere near the important mid-latitude features. Furthermore, observations that produce the largest change in the cyclone minimum SLP are near regions of large initial condition sensitivity shown in Figs. 5.10 and 5.12.

For the Tokage forecast, the Ulaan-Baator, Mongolia rawinsonde profile, which is located near the region of large 500 hPa height sensitivity, leads to a 4 hPa increase in the cyclone minimum SLP, while a few of the surrounding rawinsondes generally produce changes of up to 2 hPa (Fig. 5.19a). In contrast to the Tokage forecast, Nabi's 48-hour minimum SLP forecast is relatively unchanged by observation assimilation; the Ulaan-Baator profile identified during Tokage case and the observation of cyclone best track position, which will affect the nearby shortwave trough and TC respectively, each lead to a 2 hPa reduction in minimum SLP (Fig. 5.19b). The largest changes to Kirogi's minimum SLP forecast result from the Minamidaitojima rawinsonde profile near the TC and the Changchun, China rawinsonde north of Korea (Fig. 5.19c). These observations would modify the environment surrounding the tropical cyclone and synoptic-scale trough, respectively and each produce a 4 hPa change in the minimum SLP. Furthermore the estimated increase in the cyclone minimum SLP due to the best track position observation is 3 hPa. Two different rawinsonde observations near the TC lead to significant changes in Lupit's 48-hour minimum SLP forecast; the Yonago rawinsonde profile in southern Japan and Naze island station are associated with a 4 hPa increase and 3 hPa decrease in the metric, respectively. Moreover, the Dulan rawinsonde over central China, which is near the trough that phases with Lupit in the forecast, is associated with a 4 hPa decrease in minimum SLP.

Although several observations in Fig. 5.19 are shown to produce large changes

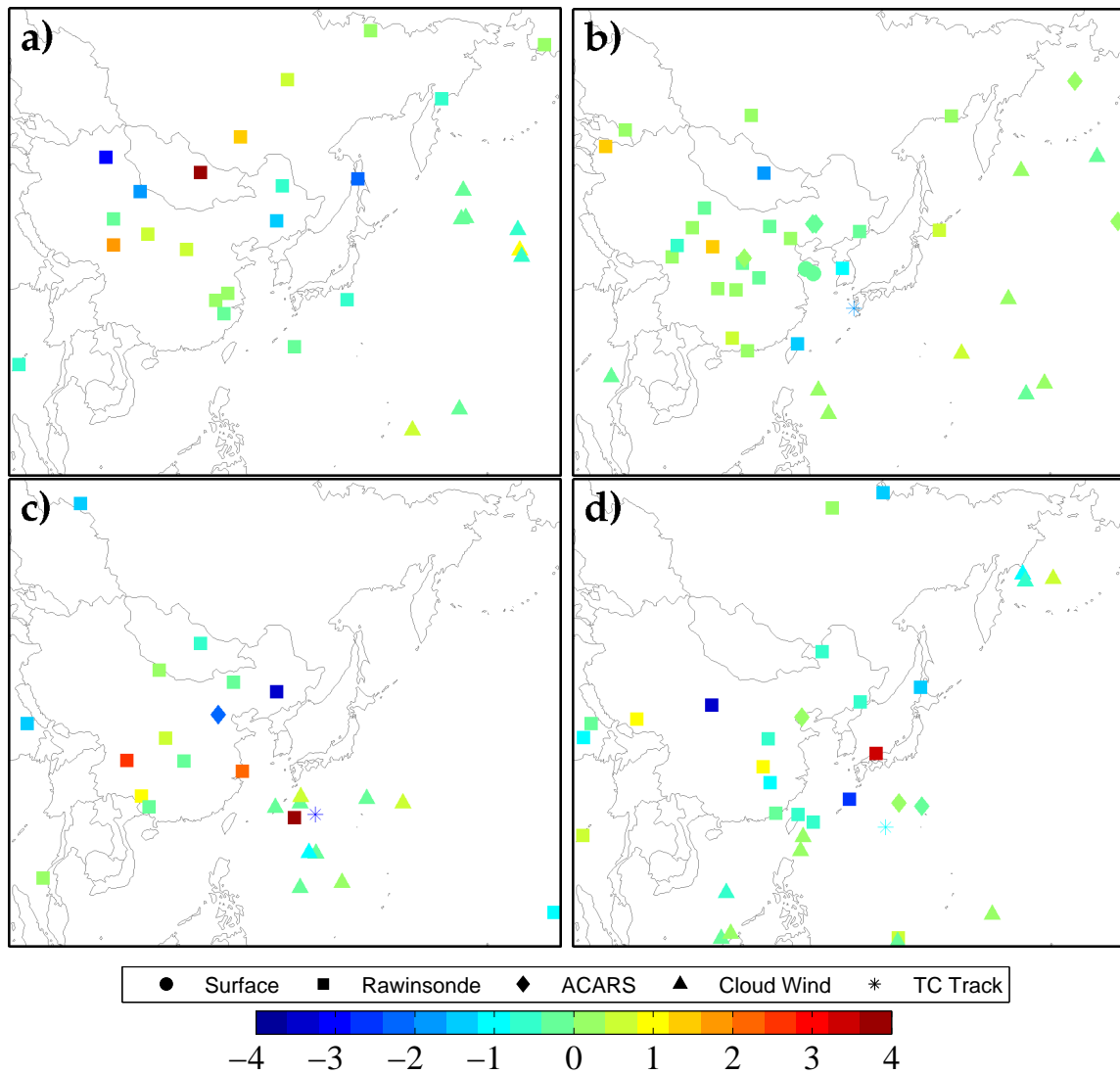


Figure 5.19: Change in the expected value of the 48-hour forecast of cyclone minimum central pressure valid (a) 12 UTC 21 October 2004 (Tokage), (b) 00 UTC 8 September 2005 (Nabi), (c) 00 UTC 19 October 2005 (Kirogi), and (d) 12 UTC 2 December 2003 (Lupit) due to the assimilation of statistically significant observations at the analysis time as estimated from ensemble statistics (shading). The shape of the symbol indicates the observation platform. Units are hPa.

to the expected value of the minimum SLP forecast, Fig. 5.20 shows that up to five observations are associated with large reductions in this forecast metric’s standard deviation. While there is generally good correspondence between the observations that are associated with large changes in the expected value and spread, there are a few observations that only lead to a large change in one of the two quantities. This can occur because the change in the expected value depends partly on the observation innovation ($\mathbf{y} - \mathcal{H}(\mathbf{x}^b)$), while the variance change does not. The Ulaan-Baator, Mongolia rawinsonde is associated with a 3.5 hPa reduction in the spread of Tokage’s minimum SLP forecast; however, the Minamidaitojima rawinsonde south of Japan and Chita, Siberia station produce a 2.5 hPa and 2.0 hPa decrease, respectively (Fig. 5.20a). In contrast to the other cases, observation assimilation for the Nabi forecast is characterized by smaller reductions in the forecast metric spread (Fig. 5.20b). Only the Dulan profile in central China has a metric spread reduction that exceeds 2 hPa. For Kirogi’s forecast, both the Minamidaitojima rawinsonde west of the TC and the best track position observation each reduce the spread in TC minimum SLP by 4.0 hPa (Fig. 5.20c). Furthermore, the rawinsonde observations in southeastern China and the Changchun station in northeastern China are associated with a 2 hPa and 1.5 hPa reduction in spread, respectively. Three observations lead to large reductions in minimum SLP spread during Lupit’s forecast; the Yonago profile in Japan and Dulan profile in central China, and best track position observation lead to a 2.7 hPa, 1.2 hPa, and 1.2 hPa reduction, respectively, while the remaining observations are associated with less than 1 hPa reductions (Fig. 5.20d).

The value of the identified observations to 48-hour ET forecasts is assessed by assimilating different sets of observations at forecast hour 0, and comparing the 48-hour forecasts that result from the resulting initial conditions. Observations for each experiment are assimilated serially using the procedure described at the beginning of this chapter, but without covariance inflation. After data assimilation is completed, the resulting analysis ensemble is integrated forward 48 hours and the expected value and

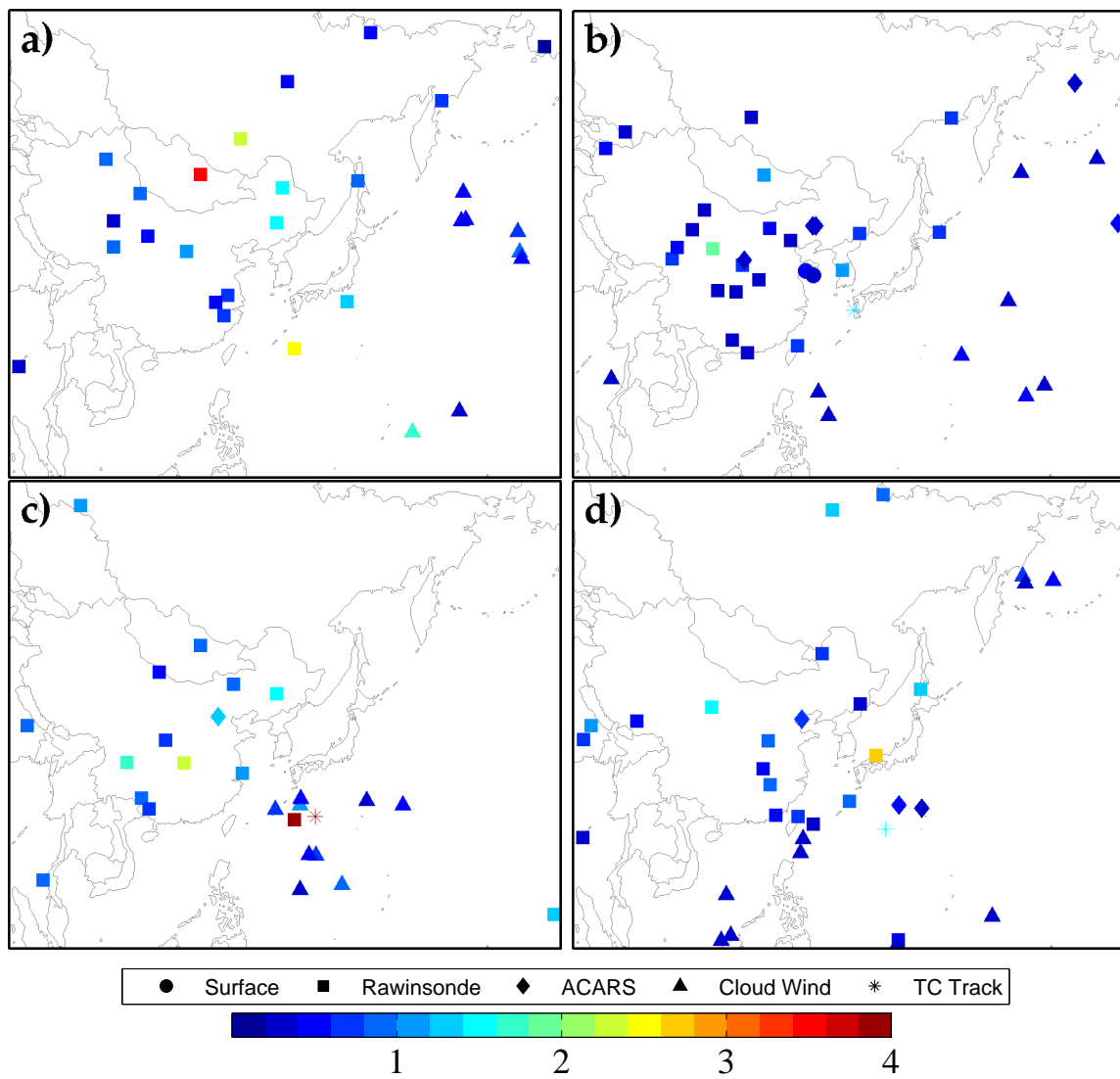


Figure 5.20: As in Fig. 5.19, but for the spread in the 48-hour forecast of cyclone minimum SLP.

spread of the cyclone minimum SLP is compared with the control ensemble forecasts where no observations are assimilated at hour 0. The “All” observation set considers all available observations at the forecast initialization time ($\approx 12,000$). Two of the five experiments assimilate orders of magnitude fewer observations. The “Best” experiment only assimilates the ≈ 40 observations identified at the beginning of this section that produce large changes in the metric, while the “Random” set randomly selects an equal number of observations from the full observation set. Whereas the “Best” experiment describes how a small number of observations can change the minimum SLP forecast, the “Random” experiment is used to determine whether the same change can be obtained by randomly assimilating the same number of observations. Finally, two additional observation sets are considered that withhold the “Best” observations. The “No Best” experiment assimilates all available observations except for the ≈ 40 statistically significant observations; however, other observations from the same platform are still considered. In contrast, the “No Platform” set contains all observations except for the observation platforms that contain the statistically significant observations. For example, consider a rawinsonde station that has a statistically significant temperature observation at 500 hPa. In the “No Best” experiment, only the 500 hPa temperature observation will be withheld; however, in the “No Platform” experiment, the entire rawinsonde profile would be removed.

Comparison of the Tokage “All” experiment to the other four ensemble forecasts indicate that a small sample of observations can lead to a large change in the cyclone minimum SLP (Table 5.4). Assimilation of all observations leads to a 0.4 hPa decrease in the expected value and a 5.6 hPa decrease in the spread of the cyclone minimum SLP as compared to not assimilating any observations. When the 32 observations associated with large changes in cyclone minimum SLP are assimilated (“Best”), the expected value (spread) in 48-hour minimum SLP forecasts are 2.2 hPa higher (2.0 hPa lower) than the control forecast where no observations are assimilated, thus 36% of the reduction in forecast metric variance can be attributed to

assimilating 32 observations. The “Random” experiment confirms that the “Best” observations are important; the reduction in the spread of the minimum SLP is 5% (25%) of the change found in the “All” (“Best”) experiments. Removing the statistically significant observations has minimal impact on the minimum SLP forecast; the change in the expected value and spread of the metric for the “All” and “No Best” experiments are nearly identical. When the 32 important observations are removed, assimilating the remaining observations with flow-dependent error statistics makes up for the absence of the “Best” data because this observation data is highly correlated. Moreover, this result suggests that the observations in a single rawinsonde profile contain redundant information since removing one observation from a number of rawinsondes has little impact on the metric. In contrast, when the *entire* Ulaan-Baator, Mongolia rawinsonde profile is removed in the “No Best Platform” experiment, the expected value of the 48-hour minimum SLP forecast is the same as for the forecast where no observations are assimilated, thus assimilating some information from this rawinsonde is important to the Tokage forecast.

Repeating these observation impact experiments for the remaining ET forecasts confirms the relative importance of the “Best” observations for 48-hour minimum SLP forecasts. The reduction in the spread of the 48-hour cyclone minimum SLP forecast when the “Best” observations are assimilated is 42%, 55% and 64% of the change obtained when all observations are assimilated for the Nabi, Kirogi and Lupit forecast times, respectively. Furthermore, randomly selecting observations from the full observation set is associated with smaller reductions in forecast metric spread as compared to the “Best” experiment. Whereas the changes in the expected value of Nabi’s cyclone minimum SLP for the “All”, “Best”, “No Best” and “No Best Platform” observation experiments are similar, there is minimal change to the expected value of minimum SLP when the most important observations are removed from the Kirogi forecast. In contrast to the other three forecasts, the most important observations to Kirogi’s minimum SLP forecast are in regions characterized by few *in situ*

observations and high forecast sensitivity (cf., Fig 5.12b), thus the remaining observations cannot compensate for their absence. In the case of Lupit, the change in the expected value for the “All” experiment is similar to “No Best” experiment; however, when the entire platform is removed (“No Platform”), the 48-hour minimum SLP forecast is 1.5 hPa lower than what is found when all observations are assimilated, thus the “Best” observation platforms limit the reintensification of the cyclone.

5.7 Using Ensemble Statistics to Understand the Dynamics of Downstream Ridge Development

For the remainder of this chapter, the statistics from the short-term forecast ensemble generated by the WRF EnKF system are used to explore the dynamical mechanisms responsible for the amplification of the downstream flow during the onset of ET. In order to obtain accurate downstream forecasts during ET, it is important to be able to understand how ET impacts the downstream state. The upper-level flow will be described using potential vorticity (PV) since this quantity is conserved in the absence of diabatic and frictional processes. The relationship between the upper-tropospheric PV and several factors related to the transitioning TC, such as different regions of precipitation, is found by computing the linear regression between the ensemble estimates of these fields. The remainder of this section proceeds by determining the mechanisms responsible for the downstream ridging that occurred with Tokage’s transition before a similar analysis is performed for Nabi and Lupit’s transition. Recall that Kirogi moved into a region that was unfavorable for baroclinic cyclogenesis, and as a result, did not lead to the amplification of the mid-latitude flow.

5.7.1 Tokage

Figure 5.21 shows the ensemble-mean six-hour forecast of SLP, precipitation, 850 hPa temperature, 850 hPa frontogenesis, and the upper-level flow, represented by the 2 potential vorticity unit (PVU) contour on the 250 hPa pressure surface, each twelve

hours prior to and during the transition of Tokage. Over this 24 hour period, the upper-level flow downstream of Tokage amplifies as the TC moves into the mid-latitudes south of Japan. Further, there are two primary regions of precipitation, and thus latent heating, near Tokage: a quasi-circular region near the center of the TC, with precipitation rates in excess of 45 mm (6 h)^{-1} , and another west-east oriented region to the north and east of the storm with rates around 20 mm (6 h)^{-1} . Whereas the former precipitation area is associated with the TC, the latter region is co-located with the warm front to the northeast of the TC. On 00 UTC 19 October, the TC is located well to the south of this baroclinic zone; however, 24 hours later, Tokage becomes embedded within the 850 hPa thermal ridge and the two areas of precipitation become indistinguishable. The southerly winds in between the tropical cyclone and the sub-tropical high to its east advect tropical values of heat and moisture into the mid-latitudes and is responsible for the frontogenesis in excess of $1 \text{ K (100 km)}^{-1} (3 \text{ h})^{-1}$ along this baroclinic zone northeast of the TC. Assuming that the 2 PVU contour follows the mid-latitude flow, the horizontal component of the relative vorticity vector, which is proportional to the vertical wind shear, is directed to the northeast in the vicinity of the TC, thus either the precipitation associated with the TC inner core, or the frontogenesis could be responsible for the amplification of the downstream ridge.

The relative importance of each precipitation region is determined from the ensemble statistics by evaluating how changes in horizontally-averaged precipitation affect the 250 hPa PV at each grid point (Fig. 5.22). This figure and several others like it are created using the following procedure. For each ensemble member, the six-hour accumulated precipitation associated with the baroclinic zone (TC) is determined by computing the area-average precipitation within the box (circle) shown in Fig. 5.22. Next the ensemble estimates of the area-average precipitation are normalized by dividing the ensemble estimates by the ensemble standard deviation. Finally, the change in the 250 hPa PV at each horizontal grid point associated with a one

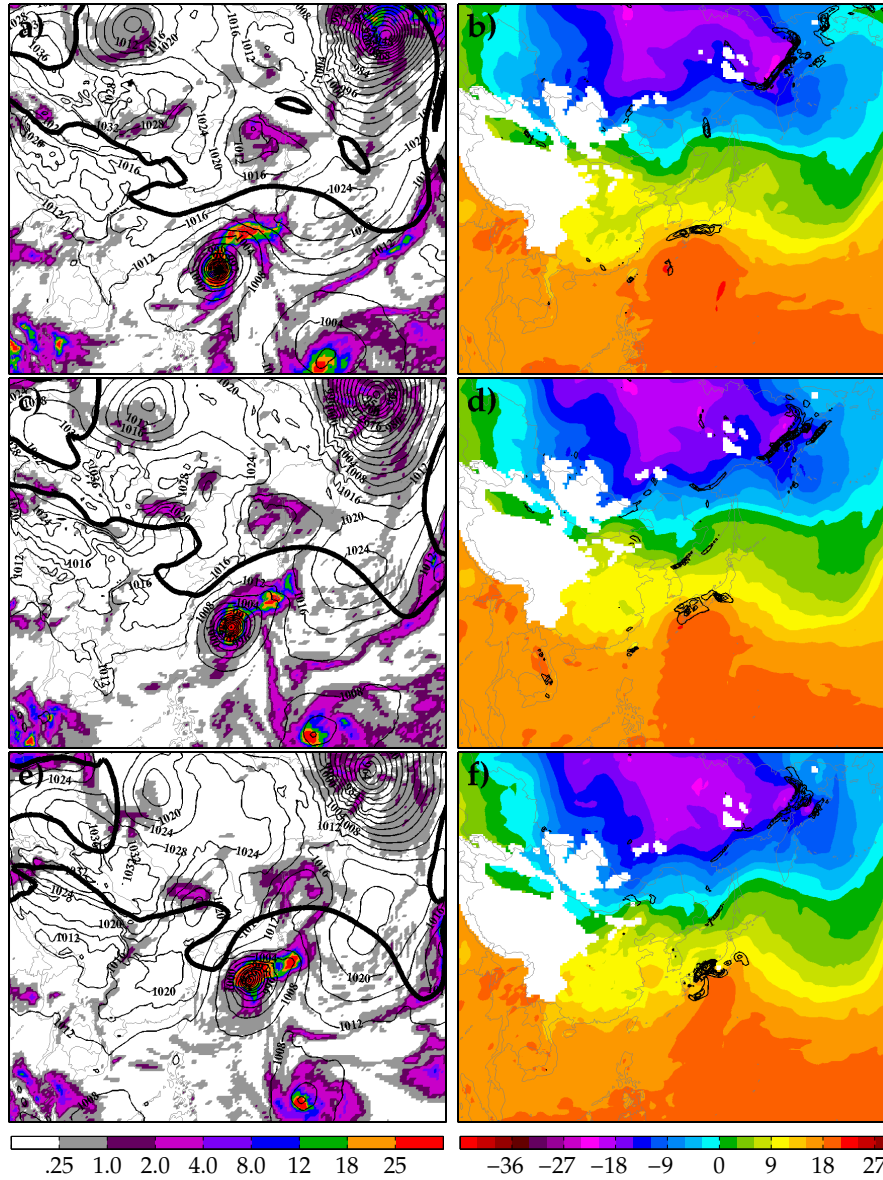


Figure 5.21: (left column) Six-hour ensemble mean forecast of accumulated precipitation (shading, mm (6 h)^{-1}), sea-level pressure (thin contours, hPa), and the 2 potential vorticity unit contour on the 250 hPa surface (thick contour, $1 \text{ PVU} = 10^{-6} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$) valid 06 UTC 19 October 2004 (top row), 18 UTC 19 October 2004 (middle row) and 06 UTC 20 October 2004 (bottom row). The right column shows the corresponding time 850 hPa temperature (shading, $^{\circ}\text{C}$) and 850 hPa frontogenesis contoured for values greater than $1 \text{ K (100 km)}^{-1} (3 \text{ h})^{-1}$ every $1 \text{ K (100 km)}^{-1} (3 \text{ h})^{-1}$.

standard deviation change in each area of precipitation ($\frac{\partial \mathbf{x}^a}{\partial J}$) is computed using (5.2), where J is the normalized average precipitation in either the box or circle and \mathbf{x}^a is the six-hour forecast of 250 hPa PV. Similar to Fig 5.9, the sensitivity at each grid point is tested for statistical significance at the 95% confidence level using (4.1); grid points that do not pass the test are set to zero. Regions characterized by warm (cold) colors on this figure show where a one standard deviation increase in the respective area of precipitation is associated with an increase (decrease) in the 250 hPa PV.

On 06 UTC 19 October, increasing the six-hour forecast of warm-front precipitation by one standard deviation is associated with a 0.25 PVU decrease in the six-hour 250 hPa PV forecast within the developing ridge north of Japan, while increasing the precipitation in the inner core does not seem to relate to the 250 hPa PV in the ridge (Fig. 5.22a, b). Decreasing the PV to the north of Japan achieves an amplification of the ridge, thus the statistics of the ensemble indicate that the development of the ridge at this time is tied to the latent heating along the baroclinic zone. Twelve (24) hours later, a one standard deviation increase to the warm-front precipitation is associated with a 0.4 PVU (0.6 PVU) reduction in the ridge PV, thus the precipitation has a larger impact as the storm moves poleward. Moreover, one standard deviation changes to the tropical cyclone precipitation during these same times are associated with a 0.25 and 0.6 PVU reduction, respectively. These results suggest that when the tropical cyclone is located well south of the mid-latitude baroclinic zone, the warm-front precipitation is primarily responsible for the initial upper-tropospheric ridge amplification; however, when the TC moves into the lower-tropospheric thermal ridge, both the TC and frontal precipitation are equally important in amplifying the upper-level flow, possibly because it is difficult to distinguish between the two areas of precipitation.

Two factors can contribute to the amount of precipitation, and thus the amount of latent heating, associated with the baroclinic zone: the strength of the vertical circulation along the baroclinic zone, as measured by the magnitude frontogenesis near

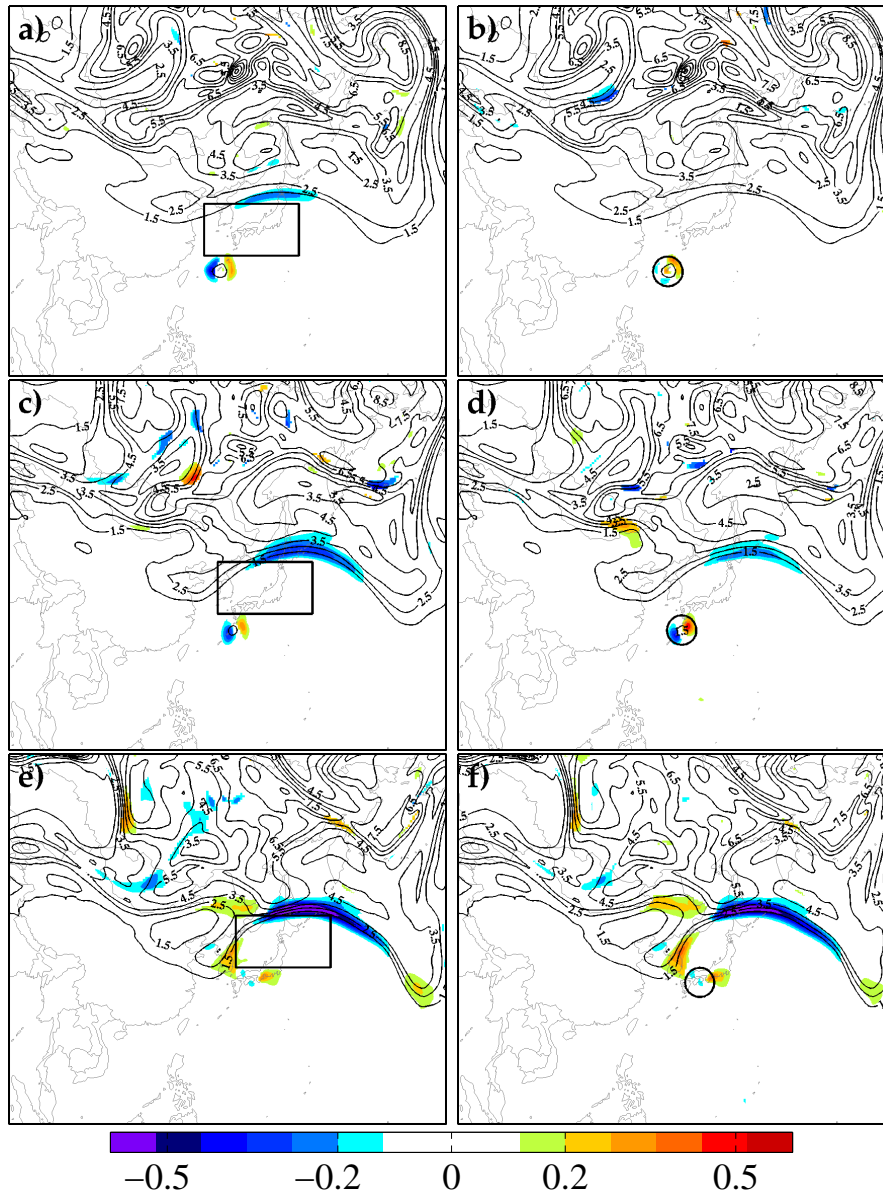


Figure 5.22: Change in the six-hour forecast of 250 hPa PV associated with a one standard deviation change in the six-hour forecast of the area-average precipitation within the box (left column), and the average precipitation within the circle (right column) valid 06 UTC 19 October 2004 (top row), 18 UTC 19 October 2004 (middle row), and 06 UTC 20 October 2004 (bottom row). The box encloses the precipitation associated with the warm-front, while the circle encloses the tropical cyclone precipitation. Contours are the six-hour ensemble-mean forecast of 250 hPa PV. Units are PVU.

the warm front, and the amount of water vapor available to precipitate, which is replenished by the poleward flux of moisture south of the baroclinic zone. To determine how the magnitude of each factor contributes to the destruction of PV, the change in the six-hour forecast of 250 hPa PV associated with a one standard deviation change in the horizontally-averaged six-hour frontogenesis and moisture flux forecast is computed from the ensemble statistics in a manner similar to what is done above for the average precipitation (Fig. 5.23). At 06 UTC 19 October, increasing (decreasing) the frontogenesis or moisture flux in the respective boxes by one standard deviation is associated with a 0.15 PVU and 0.30 PVU decrease (increase), respectively, in the PV within the ridge (Fig. 5.23a, b), while 12 hours later, the same change leads to a 0.35 PVU decrease (increase) in PV (Fig. 5.23c, d). In addition, a one standard deviation increases to the moisture flux are also associated with increasing (decreasing) the 250 hPa PV to the east (west) of the circular PV region associated with the TC. This result suggests that the moisture flux is larger when the TC is further east because the southerly winds are stronger.

Both the 850 hPa frontogenesis and 850 hPa moisture flux are a function of the wind field at that level, and therefore, it is possible that the magnitude of these two factors could be related. To evaluate which of these processes is a better predictor of the magnitude of the PV destruction within the ridge, multi-variate linear regression (e.g., Wilks 2005, section 6.2.8) of the form $y = a_1x_1 + a_2x_2$ is performed where the independent variables are the normalized frontogenesis (x_1) and normalized moisture flux (x_2) in the respective boxes shown in Fig. 5.23, and the dependent variable (y) is the horizontally-averaged 250 hPa PV within the dashed box shown in Fig 5.23b and d, which serves as a proxy for the ridge amplitude. Table 5.5 displays the regression coefficients and correlation between the predictors for each time considered. Although the magnitudes vary based on the time considered, the larger regression coefficient for 850 hPa moisture flux indicates that a one standard deviation change to the moisture flux leads to a larger change in the average 250 hPa PV within the box. As suspected,

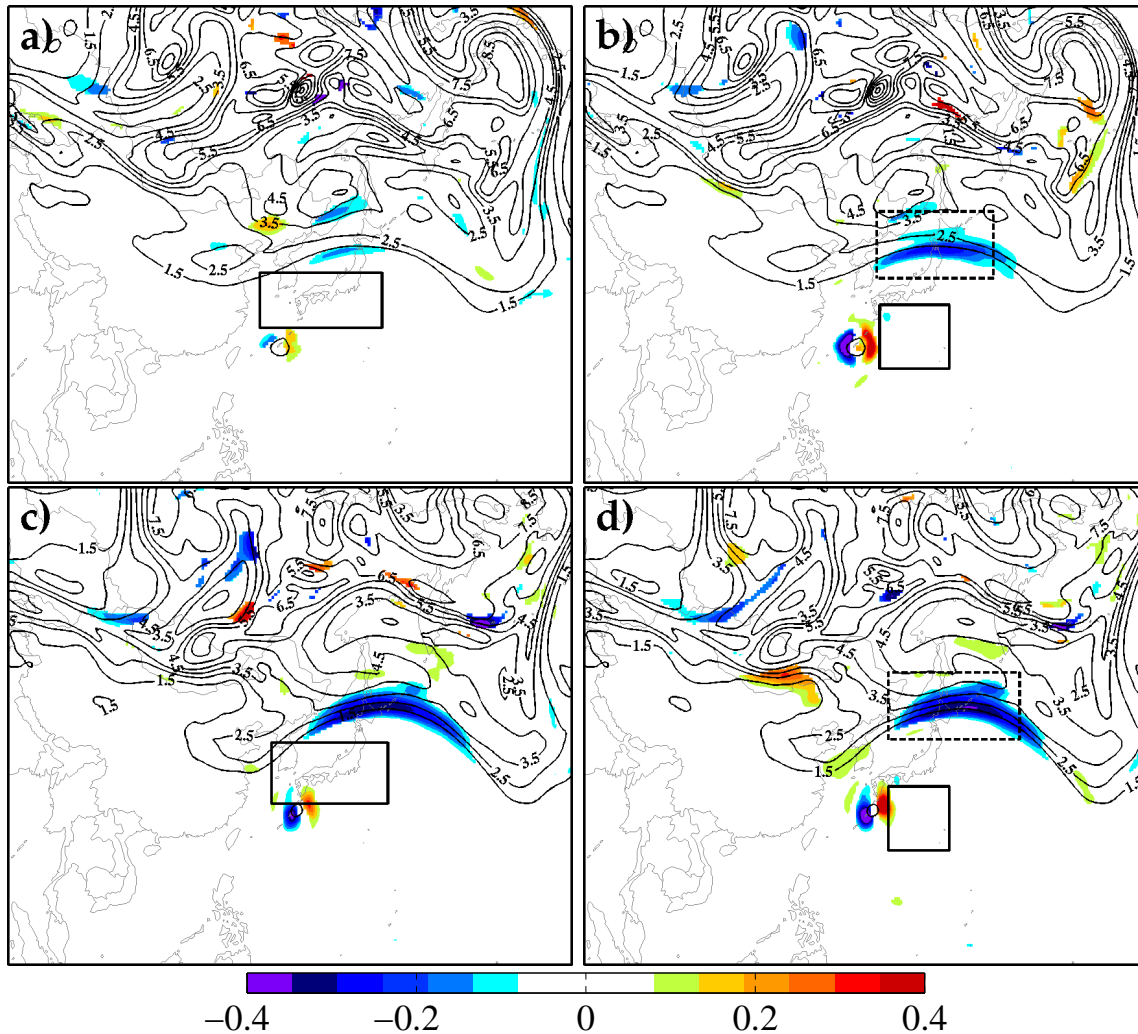


Figure 5.23: Change in the six-hour forecast of 250 hPa PV associated with a one standard deviation increase in the six-hour forecast of the average 850 hPa frontogenesis within solid box in the left column, and the average 850 hPa moisture flux within solid box in the right column valid 06 UTC 19 October 2004 (top row), and 18 UTC 19 October 2004 (bottom row). The dashed box in panels b and d denote the region of 250 hPa PV used for the multi-variate regression calculations displayed in Table 5.5. Contours are the six-hour ensemble-mean forecast of 250 hPa PV. Units are PVU.

Table 5.5: Multi-variate linear regression coefficients for 850 hPa frontogenesis (a_1) and 850 hPa water vapor flux (a_2), which are the independent variables, and the dependent variable is the average 250 hPa PV in the dashed boxes in Figs. 5.23, 5.27, and 5.31 for the Tokage, Nabi, and Lupit forecasts, respectively. $r_{1,2}$ and R^2 are the correlation between the predictors and the fraction of variance explained, respectively.

Forecast valid time	a_1	a_2	$r_{1,2}$	R^2
Tokage				
06 UTC 19 October 2004	0.001 PVU	-0.072 PVU	0.40	0.41
12 UTC 19 October 2004	0.0005 PVU	-0.093 PVU	0.24	0.47
18 UTC 19 October 2004	-0.046 PVU	-0.066 PVU	0.56	0.54
Nabi				
12 UTC 6 September 2005	-0.042 PVU	-0.054 PVU	0.50	0.55
18 UTC 6 September 2005	-0.025 PVU	-0.069 PVU	0.41	0.55
Lupit				
18 UTC 29 December 2003	-0.0007 PVU	-0.038 PVU	0.27	0.39
00 UTC 30 December 2003	-0.015 PVU	-0.091 PVU	0.07	0.54

the magnitude of the frontogenesis and moisture flux are correlated above 0.25, thus there is redundant information in these predictors. Nevertheless, these two variables explain approximately 50% of the variance in the average 250 hPa PV. The moisture flux is a function of both the winds and water-vapor mixing ratio, thus the multi-variate regression described above is repeated where the independent variables are now the normalized average wind, and normalized average water vapor in the moisture flux box shown in Fig. 5.23d. The regression coefficients for meridional wind and water-vapor at 12 UTC 19 October are -0.11 and -0.03, respectively; therefore, the evolution of the downstream ridge is linked to the lower-tropospheric wind speed.

The relationship between the moisture flux and 250 hPa PV is confirmed by generating perturbed initial conditions that will increase the analysis water vapor flux at

00 UTC 19 October. This perturbed forecast is integrated forward 12-hours whereby the difference between the 12-hour forecast of 250 hPa PV within the dashed box in Fig. 5.23b is compared to the control forecast and the ensemble-based prediction. These perturbed initial condition experiments are created similar to those described in section 5.5, but with the following differences. The perturbed initial condition is generated via (5.1), where J is the horizontally-averaged analysis 850 hPa water vapor flux in the solid box in Fig. 5.23b, and the control analysis is the ensemble member whose average 850 hPa moisture flux is closest to the ensemble-mean value. In addition, the largest value of α for these experiments is three times the ensemble standard deviation of the analysis moisture flux.

Figure 5.24 shows that, over the range of values tested here, the 12-hour forecast of the average 250 hPa PV within the dashed box in Fig. 5.23d is linearly related to the horizontally-averaged 850 hPa moisture flux. A 0.03 m s^{-1} increase (decrease) in water vapor flux leads to a 0.28 PVU decrease (increase) in the 12-hour average 250 hPa PV forecast. Moreover, the ensemble prediction of how the PV is altered by a change in the analysis moisture flux is in good agreement with the actual difference obtained from the model, which lends confidence to the hypothesis that the poleward transport of tropical moisture and the resulting latent heating along the baroclinic zone are associated with downstream ridging during ET.

5.7.2 *Nabi*

The downstream ridge amplification during Nabi's transition has characteristics that are similar to the Tokage transition; however, the response of mid-latitude flow to the transition and thus the downstream impact is less. Figure 5.25 shows two primary regions of precipitation during Nabi's transition, one associated with the TC inner core and another along a weak warm front to the northeast of the TC. Nabi's transition occurs during early September, thus the 18° isotherm at 850 hPa extends north into Japan and the meridional temperature gradient is smaller than what would be

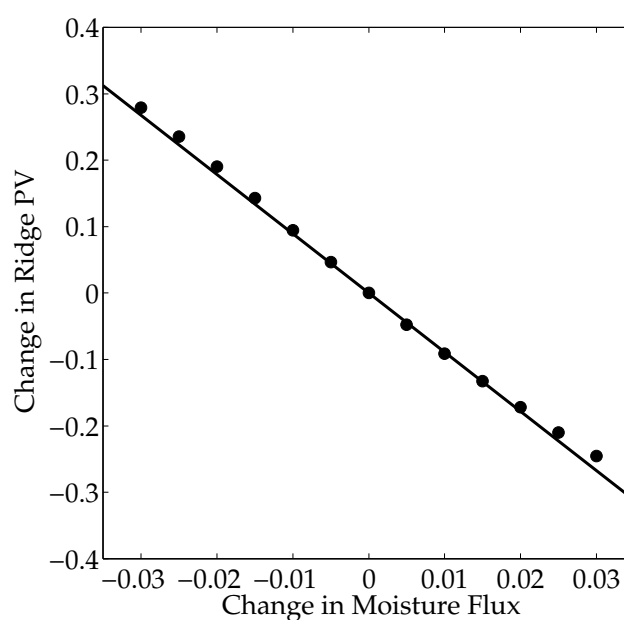


Figure 5.24: Change in the 12-hour average 250 hPa PV forecast within the dashed box in Fig. 5.23b (ordinate, units PVU) associated with adjusting the average 850 hPa moisture flux in the solid box in Fig. 5.23b (abscissa, units m s^{-1}) as determined by perturbed integrations of the WRF model for the forecast initialized on 00 UTC 19 October 2004. The solid line denotes the ensemble prediction of how adjusting the analysis moisture flux changes the 12-hour 250 hPa PV forecast.

obtained during October or November based on climatology. In comparison to the Tokage transition, the frontogenesis in this region is smaller magnitude and more disorganized.

Statistics from the six-hour ensemble forecasts valid 12 UTC and 18 UTC 6 September indicate that 250 hPa PV forecast distribution is sensitive to both the tropical cyclone and baroclinic zone precipitation (Fig. 5.26). This figure is computed using the procedure used to estimate the relative impact of different precipitation areas on the 250 hPa PV during Tokage's transition. The only differences relate to the regions of TC and baroclinic zone precipitation during Nabi's transition, which are shown in Fig. 5.26. Whereas increasing (decreasing) the baroclinic zone precipitation by one standard deviation yields a 0.25 PVU decrease in the 250 hPa PV within the developing ridge north of Japan, a similar change to the TC precipitation correlates with a 25% smaller change in the PV field. Comparable values are most likely obtained at both times because the baroclinic zone and TC precipitation are closer together than in the Tokage transition.

Figure 5.27 compares the change in six-hour 250 hPa PV forecasts associated with adjustments to the magnitude of the six-hour forecast of 850 hPa frontogenesis and 850 hPa water vapor flux. Increasing (decreasing) the magnitude of the average frontogenesis or water vapor flux in the respective boxes by one standard deviation at 12 UTC 6 September is associated with a 0.3 PVU decrease in PV within the ridge. Six-hours later, a one standard deviation change in the moisture flux (frontogenesis) relates to the same (0.2 PVU) reduction in PV in the ridge.

Using the procedure described for the Tokage case, multi-variate linear regression coefficients are computed for these times to determine the relative importance of the 850 hPa frontogenesis and moisture flux to the mid-latitude PV. For these calculations, the dependent variable is again the average 250 hPa PV within the dashed box in Fig. 5.27b, d, while the independent variables are the normalized 850 hPa frontogenesis and 850 hPa moisture flux in the solid boxes shown in Fig. 5.27. The regression

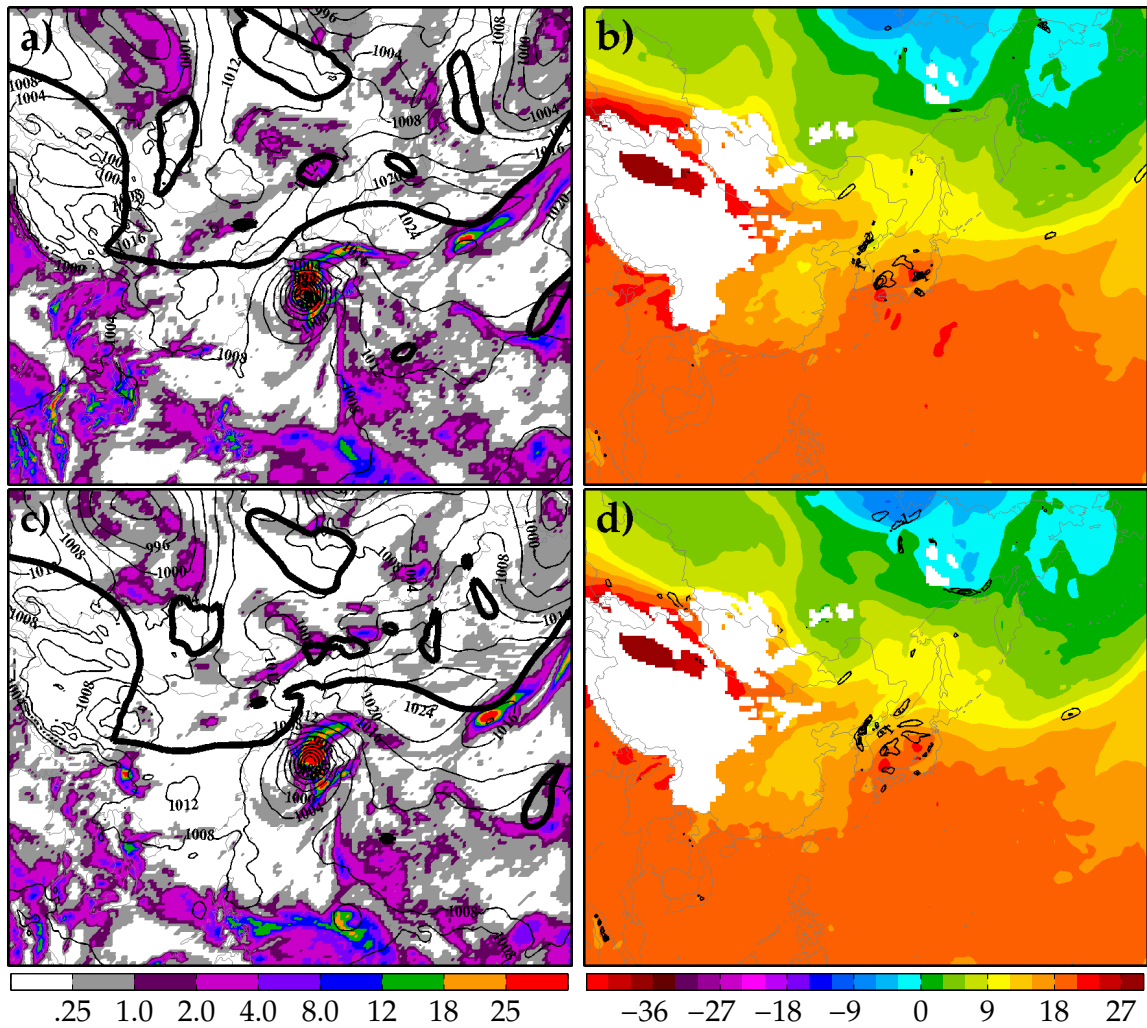


Figure 5.25: As in Fig. 5.21, but for the six-hour forecast valid 12 UTC 6 September 2005 (top row) and 18 UTC 6 September 2005 (bottom row).

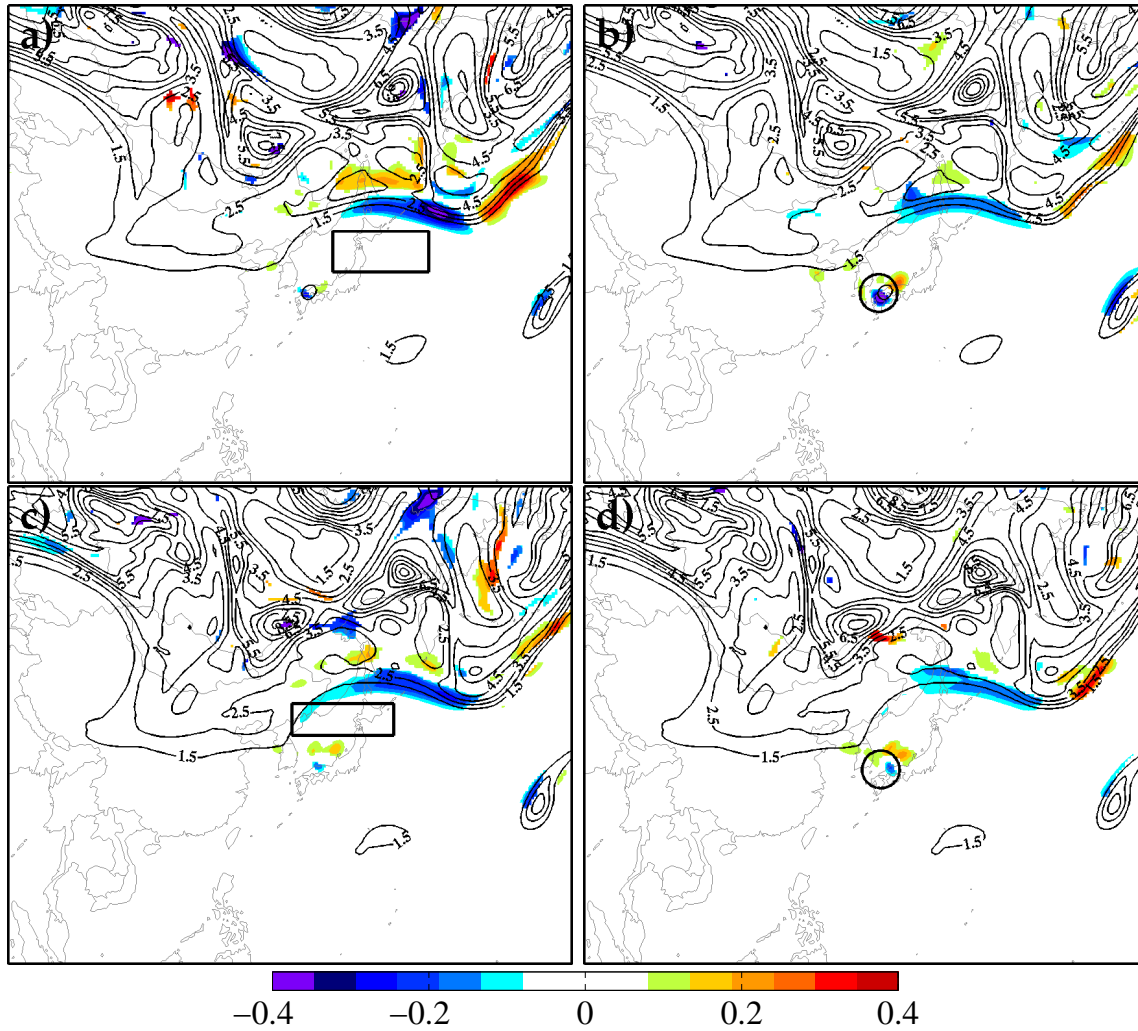


Figure 5.26: As in Fig. 5.22, but for the six-hour forecast valid 12 UTC 6 September 2005 (top row) and 18 UTC 6 September 2005 (bottom row).

coefficients for these forecasts indicate that one standard deviation changes to the moisture flux leads to a larger reduction of PV within the ridge (Table 5.5). Whereas a one standard deviation change in the 850 hPa moisture flux leads to a 0.05 PVU decrease in the average PV within the ridge, the same change to the frontogenesis leads to a 0.03 PVU decrease in the average PV. Moreover, these two predictors explain 55% of the variance in the average 250 hPa PV, but are still correlated above 0.40.

The impact of increasing the moisture flux to the east of Nabi is further evaluated by generating twelve-hour forecasts initialized at 06 UTC 6 September with perturbed initial conditions consistent with a larger (smaller) analysis moisture flux. The procedure used to produce the perturbed initial conditions for this time is similar to what is used for the Tokage forecast above, except that the average PV is evaluated within in the dashed box in Fig. 5.27d and the region of average moisture flux is given by the solid box in Fig. 5.27b. Figure 5.28 shows that, much like the Tokage forecast described previously, the average PV in the ridge is a linear function of the analysis moisture flux to the east of the TC; a three standard deviation increase (decrease) in the water vapor flux is associated with a 0.4 PVU decrease (increase) in the average 250 hPa PV within the box 12 hours later. Furthermore, the ensemble-based predictions of how the magnitude of the analysis water vapor flux impacts the 12-hour forecast of average PV nearly match the actual difference obtained from non-linear forecasts.

5.7.3 *Lupit*

Unlike the previous two cases, the downstream ridging that occurs during transition of Lupit is complicated by the presence of a series of mid-latitude cyclones to the north of the TC. Figure 5.29 shows that at 18 UTC 29 December and 00 UTC 30 December, the western Pacific Ocean is dominated by two regions of large precipitation; one associated with Lupit's core where the rate exceeds 35 mm (6 h)^{-1} , and another

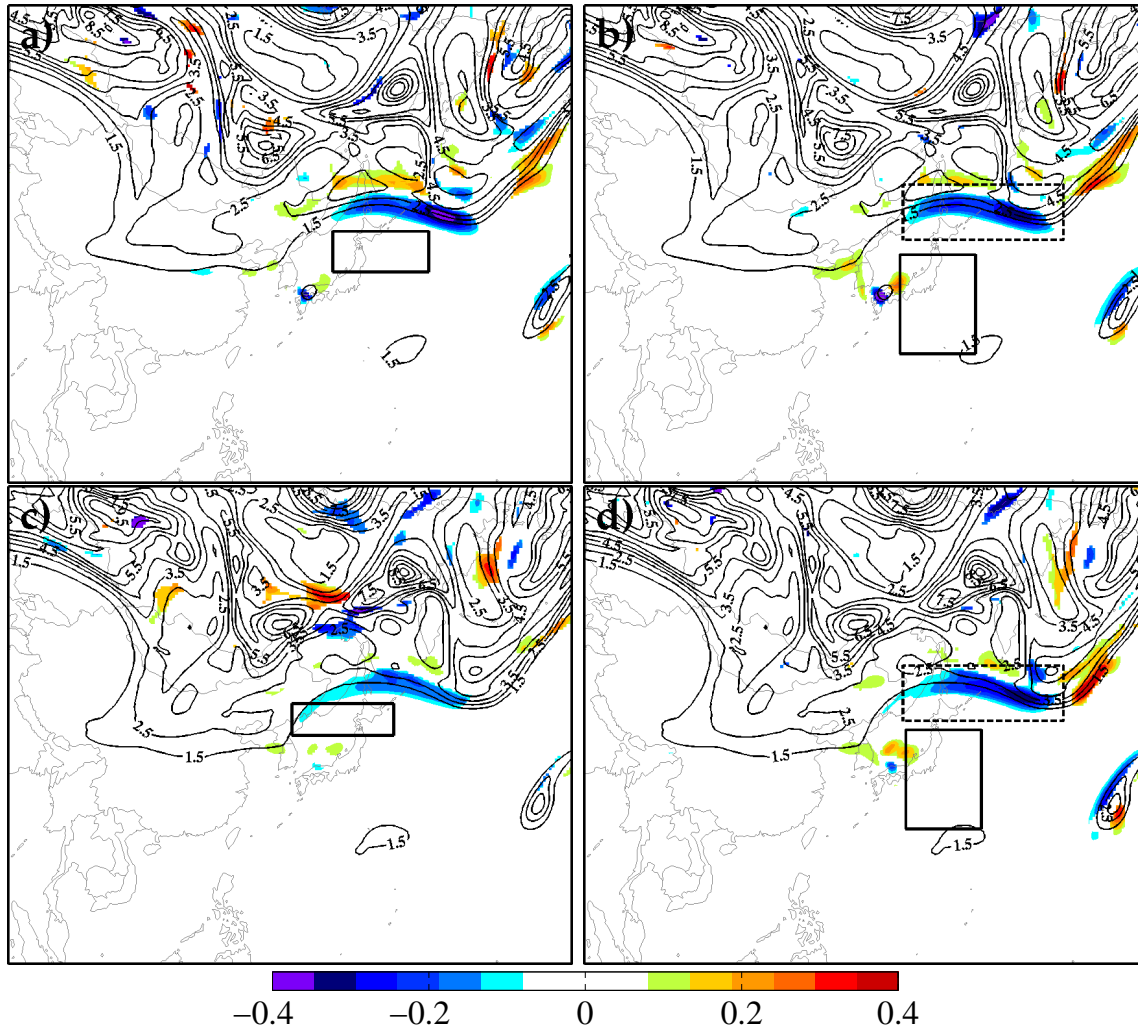


Figure 5.27: As in Fig. 5.23, but for the six-hour forecast valid 12 UTC 6 September 2005 (top row) and 18 UTC 6 September 2005 (bottom row).

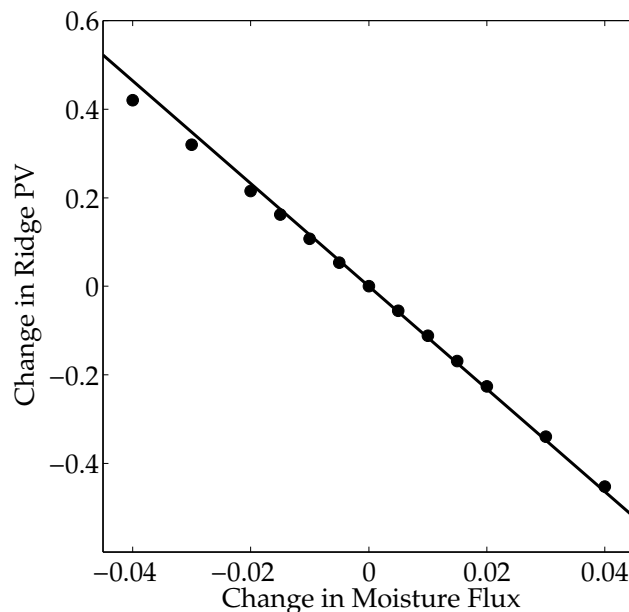


Figure 5.28: As in Fig. 5.24, but for the twelve-hour forecast valid 18 UTC 6 September 2005. The region used for the average 12-hour PV forecast (analysis 850 hPa moisture flux) is denoted by the dashed (solid) box in Fig. 5.27d (5.27b).

extending from the northern periphery of the TC along the north-south baroclinic zone to the northern coast of Japan with rates exceeding 20 mm (6 h)^{-1} . Furthermore, the baroclinic zone has two areas of larger precipitation rate that are co-located with larger frontogenesis values, one to the north of the TC where the TC winds are advecting the tropical air poleward and a second to the west of Japan where the polar airmass (850 hPa temperature $< -20^\circ\text{C}$) is being advected equatorward. The entire region to the south of the baroclinic zone is characterized by southerly winds that are transporting tropical heat and moisture poleward in between the surface high in the northwest Pacific and Lupit. Unlike the previous two cases, the ridge initially builds in response to the downstream cyclone over Japan, thus the TC transition is not solely responsible for the ridge; however, it will be shown that it contributes to its continued amplification. The procedures used to construct each of the following figures are the same as what is used to evaluate the Tokage and Nabi transitions,

except where noted.

The change in the six-hour forecast of 200 hPa PV associated with changes to the six-hour forecast of average baroclinic zone (box) and TC core precipitation (circle) is shown in Fig. 5.30. Since the cyclones north of Lupit have already lead to an amplified mid-latitude flow, 200 hPa PV is used to evaluate downstream ridge amplification for the Lupit transition because it better reflects the tropopause level near the TC. While the baroclinic zone precipitation box is larger than the previous two cases, for consistency, it encloses the entire area of precipitation from south of Japan to the Kamchatka peninsula. This figure shows that increasing (decreasing) the amount of precipitation in the baroclinic zone box by one standard deviation is associated with up to a 0.2 PVU decrease in the 200 hPa PV at the apex of the ridge on 18 UTC 29 November and a 0.3 PVU decrease in ridge PV on 00 UTC 30 November. In contrast, precipitation associated with the tropical cyclone is associated with up to 0.3 PVU and 0.2 PVU change in the PV along with western side of the ridge at 12 UTC and 18 UTC, respectively. This result suggests that while the latent heating associated with the baroclinic zone precipitation amplifies the ridge, the latent heat with the TC precipitation acts to increase the wavelength of the ridge, even though the TC is further from the jet.

Although it is not as clear as the Tokage and Nabi transition, the ensemble statistics indicate that the magnitude of the 850 hPa frontogenesis and 850 hPa moisture flux are good predictors of the downstream ridge amplitude. Figure 5.31 shows that whereas a one standard deviation increase (decrease) in the frontogenesis along the baroclinic zone at 18 UTC 29 November is associated with a 0.2 PVU decrease (increase) in the PV at the apex of the ridge west of Japan, six hours later the same change is associated with an 0.3 PVU increase (decrease) the 200 hPa PV near the tip of the Kamchatka peninsula. Adjusting the magnitude of the moisture flux south of Japan has an asymmetric impact on the PV distribution; a one standard deviation increase in the 850 hPa moisture flux is associated with a 0.4 (0.2) PVU decrease to

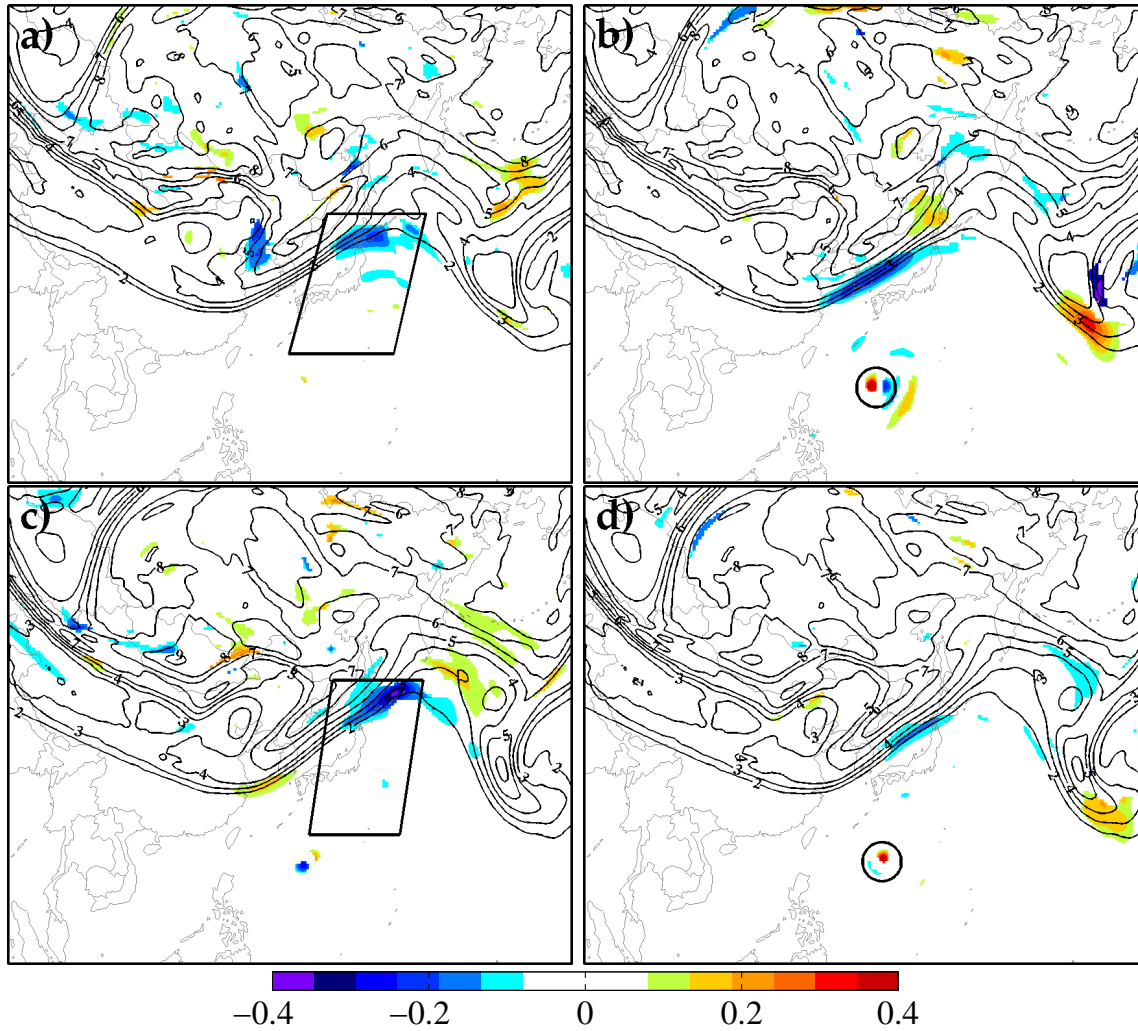


Figure 5.30: As in Fig. 5.22, but for the six-hour forecast valid 18 UTC 29 November 2003 (top row), 00 UTC 30 December 2003 (bottom row) and 200 hPa PV.

the west (east) side of the ridge at either time. This distribution suggests that when the moisture flux is larger, the extra latent heating is associated with an amplified ridge and an increase in the horizontal wavelength.

The relative impact of the frontogenesis and moisture flux is determined by performing a multi-variate regression similar to what is done in the Tokage and Nabi forecasts, where the region of average 200 hPa PV used in this calculation is denoted by the dashed box in Fig. 5.31b and d and the regression coefficients are listed in Table 5.5. Whereas a one standard deviation change in the water vapor flux is associated with a 0.038 and 0.091 change in the average PV at 18 UTC 29 November and 00 UTC 30 November, respectively, a similar perturbation to the average frontogenesis produces a much smaller difference. In contrast to the other cases, the correlation between frontogenesis and moisture flux is smaller because the two predictor regions are further apart.

Finally, initial condition perturbations, consistent with changing the magnitude of the analysis moisture flux in the box on 12 UTC 29 November, are constructed using the procedure described for the two other ET events (Fig. 5.32), where the region of average 200 hPa PV is shown in Fig. 5.31d and the analysis moisture flux by the solid box in Fig. 5.31b. Over the range of values tested here (\pm three standard deviations in the analysis moisture flux), the 12-hour forecast of PV within the dashed box in Figs. 5.31d responds linearly to the initial moisture flux; however unlike the previous two cases, the ensemble prediction is generally less than the actual change in the average PV obtained from the non-linear forecast. Whereas the ensemble sensitivity predicts that a 0.03 m s^{-1} increase (decrease) in moisture flux leads to a 0.32 PVU decrease (increase) in the 12-hour forecast of the average 200 PV in the dashed box, the model runs indicate that the actual change is 0.5 PVU. It is not clear why the ensemble prediction is inaccurate for this case, though one possibility is that the initial condition perturbation that changes the analysis moisture flux will also lead to adjustments to nearby cyclones.

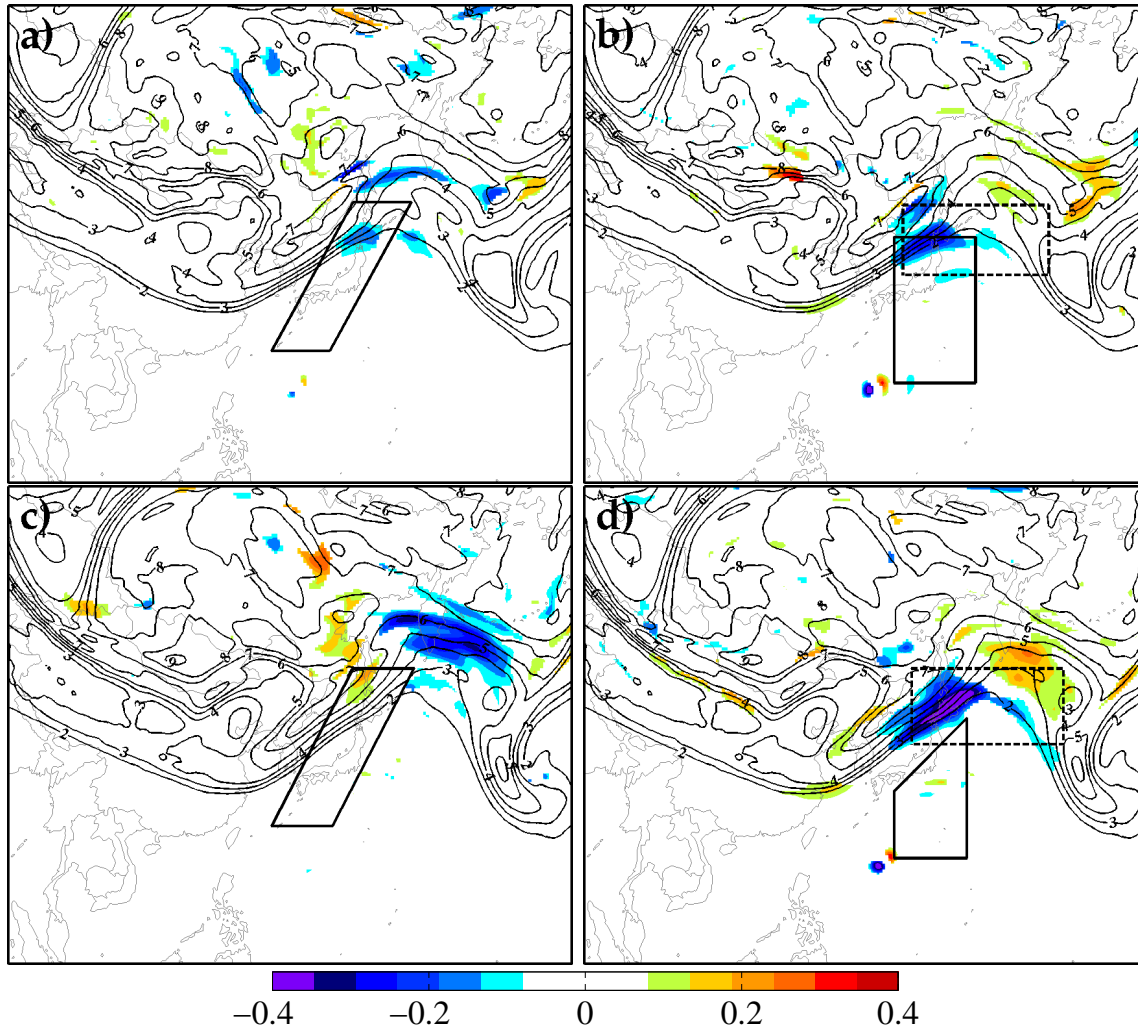


Figure 5.31: As in Fig. 5.23, but for the six-hour forecast valid 18 UTC 29 November 2003 (top row), 00 UTC 30 December 2003 (bottom row) and 200 hPa PV.

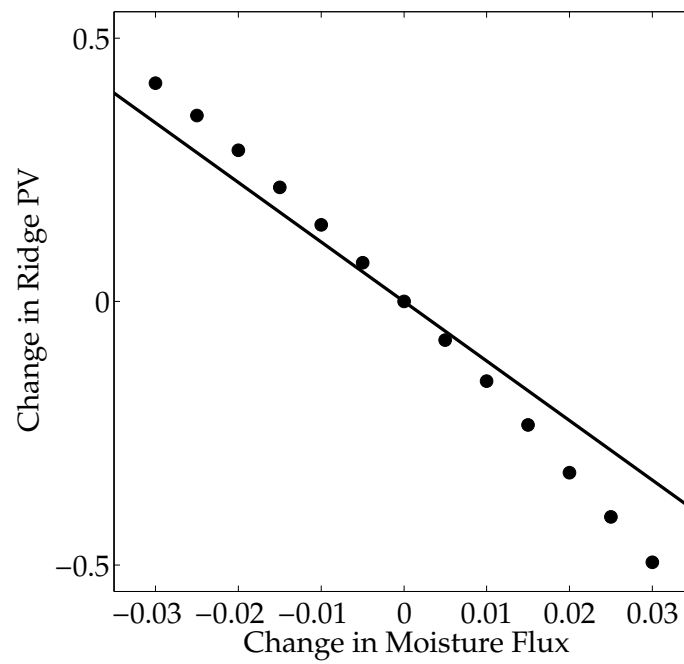


Figure 5.32: As in Fig. 5.24, but for the twelve-hour forecast valid 00 UTC 30 November 2003. The region used for the average 12-hour PV forecast (analysis 850 hPa moisture flux) is denoted by the dashed (solid) box in Fig. 5.31d (5.31b).

Chapter 6

SUMMARY AND CONCLUSIONS

The chaotic nature of the atmosphere allows small deficiencies in the initial conditions of NWP models to grow into large errors during the subsequent forecast. As a consequence, several techniques have been independently developed, such as sensitivity analysis, data assimilation and ensemble forecasting, to determine regions where small initial condition changes will have a large impact on a forecast and generate the best analysis of the atmosphere by combining a short-term forecast with observations. This dissertation describes two applications of ensemble sensitivity analysis, which when combined with ensemble data assimilation, provide a consistent framework for evaluating the predictability and dynamics of the atmosphere. These ensemble techniques are applied to weather systems in the northeast Pacific Ocean and to western Pacific extratropical transition events, both of which are characterized by forecast errors resulting from inaccurate initial conditions.

The benefit of assimilating observations with flow-dependent error statistics in a region of sparse *in situ* data is determined by evaluating output from a pseudo-operational, limited-area EnKF system. Whereas a large fraction of observations used by operational forecasting systems is satellite radiance data, the UW EnKF system only considers *in situ* data from surface stations, rawinsondes, ACARS and cloud motion vectors. In addition to providing an ensemble of analyses and short-term forecasts for observation assimilation each six hours, the UW EnKF system produces 90 independent forecasts out to 24 hours at 00 UTC and 12 UTC. Ensemble forecasts from this system and deterministic output from operational global models on the same domain are verified against rawinsonde and surface observations over a

two-year period starting 1 January 2005.

Although the UW EnKF system assimilates a small fraction of the observations used by other operational centers, the RMS error in UW EnKF forecasts of wind and temperature is similar to those yielded by the NOGAPS and UKMO forecasts and slightly larger than GFS and CMC forecasts in this region. In contrast, the bias in geopotential height forecasts from the UW EnKF system leads to errors that are slightly worse than other operational systems. Multiple reasons could exist for why UW EnKF forecasts have comparable errors, which include assimilating data using flow-dependent error statistics, using GFS lateral boundary conditions and the WRF model formulation. Two additional forecasts are produced during this same period to evaluate the relative contribution of each of these factors. The “no assimilation” forecast is a WRF ensemble member that is run for the two-year period without being updated with observations, while the WRF-GFS is a forecast initialized with the GFS analysis. The “no-assimilation” forecast, which defines the contribution from the lateral boundary conditions, has RMS errors that are up to 50% larger than six-hour UW EnKF forecasts. Moreover, WRF forecasts initialized from the GFS analysis yield errors that are comparable to GFS forecasts in this region, thus the WRF model is not responsible for the skill in the UW EnKF forecasts. As a consequence, the performance of the UW EnKF system must be due to performing data assimilation with an EnKF. Previous studies on the impact of various observation sets have shown that satellite radiance observations lead to large improvements to operational forecasts, especially in data sparse regions (e.g., Rabier et al. 2005). These results show that in regions of sparse *in situ* data, such as the ocean, assimilating observations with flow-dependent covariances can make up for not having a large number of observations.

Quite surprisingly, the error in 24-hour UW EnKF forecasts of dew-point temperature are equal to or smaller than all operational centers shown here, especially for rawinsonde stations along the North American coast. The region upstream of the coast is

characterized by many cloud motion vector and ACARS observations, but no direct observations of moisture. Whereas the fixed-error statistics used by other operational systems do not allow these wind and temperature observations to adjust the water vapor field (e.g., Parrish and Derber 1992, Rabier et al. 1998), the flow-dependent error statistics used an EnKF permit such modifications. Time-averaged correlations computed from UW EnKF data, indicate that on average, upper-tropospheric wind and mid-tropospheric temperature data lead to systematic corrections to the water vapor field. Previous studies (e.g., Ralph et al. 2004, Knippertz and Martin 2007) have shown that short-term global model forecasts of water vapor in regions of clouds are often characterized by large errors. The results presented here suggest that this problem may be reduced by assimilating other available observations, such as cloud winds, with flow-dependent error statistics.

Verification rank histograms for winds, temperature and moisture indicate that the forecast ensemble spread has skill in predicting the forecast error; however, for the mass field, the model biases are the dominant signal in this diagrams. Whereas the analysis mass field exhibits smaller bias, the mean difference between observations and the ensemble-mean forecast grows with increasing forecast hour, suggesting that observations can eliminate the bias for short-term forecasts, but cannot completely overcome it at longer lead times. Moreover, this low bias exists even when the WRF model is initialized with the GFS analysis, thus the loss of mass does not result from EnKF data assimilation, though it is possible that using covariance averaging to parameterize model error makes the problem worse at shorter lead times. Unlike the EnKF system described by Houtekamer et al. (2005), the spread in UW EnKF forecasts grows in tandem with the mean error during all forecast hours, thus the analysis perturbations project onto growing disturbances. Consistent error growth at all forecast hours most likely results from how the UW EnKF system parameterizes model error (covariance averaging), which will at least partially retain the growing structures from the previous six-hour forecast.

Similar to other mesoscale ensemble systems based on different initial conditions (e.g., Eckel and Mass 2005), forecasts of parameterized surface fields, such as 10 meter winds and 2 meter air temperature and dew-point temperature, are characterized by relatively constant errors at all forecast lead times and small growth in ensemble spread. The errors in these fields are related to the model's formulation of their evolution. Surface and boundary layer parameterization schemes have two different types of input: dynamical fields that are different for each ensemble member (i.e., low-level temperature and wind), and fixed fields that are the same for each member (i.e., solar constant, land surface type). The relative lack of ensemble spread for surface fields indicates that the invariant input parameters restrict error growth in the ensemble. Although the UW EnKF system did not attempt to assimilate surface wind, temperature or moisture data, the lack of variance in these fields would have caused the data assimilation scheme to under-weight the new information from observations. As a consequence, flow-dependent error statistics will not improve surface field forecasts unless the spread of these fields is boosted. Sutton et al. (2006) attempted to overcome the lack of variance in surface variables during periods of convection by adding uncertainty to soil moisture fields. Their results indicates that adding variance to the soil moisture showed limited improvement over using constant parameters for all ensemble members.

The large errors in surface field forecasts at all lead times indicate that model deficiencies, rather than initial condition errors are responsible for the mean error, and as a consequence, improvements to the initial conditions will not necessary lead to better forecasts. Model errors can be corrected by either improving the formulation of parameterization schemes, or by increasing the horizontal resolution, which would allow the model to resolve the local features that influence the surface wind and temperature fields. Another possible method of decreasing the model error in surface fields is objective parameter estimation, whereby the state vector is augmented with unknown parameters and thus can be adjusted by observation assimilation. Aksoy et

al. (2006) and Tong and Xue (2007) had some success improving forecasts of a sea-breeze circulation and supercell thunderstorm, respectively, when applying parameter estimation with an EnKF. The two year repository of UW EnKF forecasts offers an opportunity to retrospectively improve WRF surface forecasts through parameter estimation.

One method of improving UW EnKF forecasts is to determine where small changes to the initial conditions would have a large impact on subsequent forecasts and to assimilate observations in that region. This can be determined by applying the ensemble sensitivity technique described by Hakim and Torn (2007) and Ancell and Hakim (2007) to data drawn from the UW EnKF system during January–July 2005. Ensemble analyses and forecasts are used to determine locations of persistent sensitivity for selected forecast metrics near western North America. The skill of ensemble sensitivity analysis in predicting the change in a forecast metric mean, δJ , and variance, $\delta\sigma$, due to observation assimilation is also determined. Although this work focuses mainly on forecast metrics near the West Coast of North America, this technique is general and may be applied to any other scalar forecast metric.

Climatological sensitivity fields for 24-hour western WA SLP and precipitation forecasts are most frequently sensitive to the upstream mass field and to a lesser extent the temperature field. While a large fraction of the frequently sensitive region is observed by the fixed buoy network, the buoy closest to the sensitivity maximum was not functioning during the period and thus could have adversely affected western WA forecasts. Composite patterns for the most-sensitive forecasts indicate that the region of largest sensitivity for 24-hour western WA SLP and precipitation forecasts is approximately 1000 km west of the metric box and exhibits modest upshear tilt in the vertical. A one-standard-deviation change in the most-sensitive region of the analysis mass field would have a larger impact on either the SLP or precipitation metrics compared to a one standard deviation change in the most-sensitive region of the temperature field, which suggests that targeted buoy and ship SLP observations

have the largest potential impact on short-range, surface-based forecast metrics in this area.

The change in average SLP forecasts due to withholding a buoy from the region of frequent sensitivity is evaluated for the 30 most-sensitive cases of western WA SLP forecasts. Removing the buoy pressure observation from the assimilation process yields a perturbed forecast metric that is compared with the prediction from ensemble sensitivity analysis. For all forecast cycles, the ensemble-based estimate of the forecast metric change is in good agreement with the actual change obtained from perturbed non-linear model forecasts. These results indicate that a single SLP observation within the region of frequent sensitivity can change 24-hour western WA area-averaged SLP forecasts by up to 0.8 hPa and reduce the ensemble spread by up to 0.5 hPa.

Single-observation impact calculations are extended to estimate how a larger number of assimilated observations change 12-hour SLP forecasts in the western Washington region. Approximately 100 observations per analysis time will produce a statistically significant change in the forecast metric mean value at the 99% confidence level. As a consequence, the ensemble sensitivity approach attempts to predict the change in the forecast metric mean and variance with approximately 100 observations from the several thousand observations that are assimilated. An attractive attribute of this approach is that it can be applied “off line” to an existing dataset of ensemble analyses and forecasts without running the model or cycling a data assimilation system. Results show that the ensemble-based estimates provide a moderately accurate prediction of how observations change to the spread of the forecast metric, although agreement is worse than the case of a single observation. Similar results are found for a forecast metric that covers the west coast of North America, suggesting that the results are not limited to metrics covering small geographical areas.

Ensemble sensitivities can also be applied to test how observations impact the SLP forecast error in a manner similar to what is summarized above. For an error metric defined by the root-mean-square error in the box over western Washington state, re-

sults also show that ensemble sensitivity provides accurate estimates of the reduction in error standard deviation, and to a lesser extent the error expected value. Partitioning the error estimates by observation type indicates that surface observations, rather than ACARS or cloud-wind observations, are more likely to reduce the error for the areas described here. In particular, the fixed-position buoys 500 km from the North American coast have the largest impact because they are in the region of persistent sensitivity. It should be noted that the impact of these observation platforms can vary depending on the forecast metric, season, model, and particular observation set.

For the experiments and metrics considered here, it is evident that ensemble sensitivity analysis is more reliable in predicting changes in the spread (variance) of a forecast metric as compared to the mean-value. In the absence of significant non-linearity and model error, this difference in predictive ability may be understood by exploring how sampling error impacts the calculation of these quantities. The key quantity in (2.25) and (2.28) is the covariance between the forecast metric and the model estimate of an observation, $\mathbf{J}(\mathbf{H}\mathbf{x}^b)^T$, which affects the predicted changes in the metric mean and spread differently. If sampling error (ϵ_{cov}) for this covariance is unbiased ($E[\epsilon_{cov}] = 0$), then the predicted changes to the metric mean value will have zero mean, and the scatter about the line of perfect prediction will scale with the magnitude of the sampling error. In contrast, the error in the predicted change in the metric *spread* is proportional to the sampling error *variance* ($E[\epsilon_{cov}\epsilon_{cov}^T] \neq 0$), which will introduce a bias to estimates of $\delta\sigma$ even when the sampling error itself is unbiased (see Fig. 4.10b); the larger the sampling error, the greater the overprediction in the variance reduction. Using a confidence test on the forecast metric-observation covariance reduces the effect of sampling error, but also places limits on the number of observations that are included in the calculation. This technique works well until a point of diminishing return is reached when, for confidence levels approaching 100%, important observations are not considered and thus the ability to predict the change in the forecast metric is adversely affected. The results shown here with UW EnKF

data show that this point is reached around the 99% confidence level.

The results presented here suggest that ensemble sensitivity analysis provides an attractive alternative to adjoint sensitivity analysis. In addition, the results indicate that this technique may prove useful for observation thinning, where a large sample of observations is reduced to a set that is expected to produce the largest decrease in forecast-metric spread, while also producing a statistically significant change in the forecast-metric mean value. Unlike previously described thinning algorithms (e.g., Liu and Rabier 2002, Ochotta et al. 2005), it is possible to adapt the thinning algorithm to the forecast metric of interest. Similarly, ensemble sensitivity may be useful for observation targeting because the impact of a hypothetical observation on the forecast metric variance can be determined prior to knowing the observation value. Furthermore, this method is particularly attractive because, unlike other targeting methods, it takes the analysis-error statistics into account (Berliner et al. 1999, Langland 2005). Ensemble sensitivity analysis may also prove useful for selecting observations based on their predicted affect on forecast error variance in previous forecasts.

Given the skill of UW EnKF ensemble forecasts and the ability of ensemble sensitivity to estimate the impact of changes to the initial conditions, these methods are applied to the extratropical transition of western Pacific tropical cyclones, which are often characterized by large forecast errors. Forecasts of the transition of Typhoons Lupit (2003), Tokage (2004), Nabi (2005) and Kirogi (2005), which differ by their post-transition evolution, degree of predictability, and impact on the downstream state, are explored by cycling a WRF EnKF data assimilation system for the period prior to and during the transition of each cyclone. This EnKF system assimilates conventional *in situ* observations from surface stations, rawinsondes, ACARS, cloud wind observations and best track position data each six-hours. For the analysis time closest to the onset of transition, all 90 ensemble members are integrated forward 48 hours whereby the initial condition sensitivity and impact of assimilating observations is

evaluated. Although this WRF EnKF system only assimilates a small fraction of the observations available to the GFS, the ensemble-mean track and intensity forecasts have comparable skill.

Ensemble forecasts of the cyclone minimum SLP show large initial condition sensitivity to the features that are expected to be important during ET, namely the TC and the upper-level trough that phases with it. Whereas the region of largest sensitivity for short-term forecasts of cyclone minimum SLP are near the TC, at longer lead times, the areas of significant sensitivity are still near the TC, but are also located near mid-latitude features. Initial condition sensitivities for Tokage and Kirogi's (Lupit and Nabi's) minimum SLP forecast have larger (smaller) magnitude and are associated with the position (amplitude) of the mid-latitude troughs. Forecasts where the largest sensitivities are associated with the position of the trough are characterized by a stronger mid-latitude jet since small displacements to the trough's position can lead to a mismatch of the TC and trough during ET. In contrast, transition events for which the largest sensitivities are associated with the trough amplitude are associated with a weaker mid-latitude jet. For these forecasts, the model is likely to obtain the correct phasing between the cyclone and trough, thus the degree to which the tropical cyclone reintensifies as a mid-latitude cyclone depends on the amplitude of the trough. Furthermore, the sensitivity of the RMS error in the SLP field near the TC forecast is also large in the same regions as the initial condition sensitivity for minimum SLP forecasts, thus small errors in the TC or mid-latitude troughs can have a large impact on errors in the ET forecast.

Based on previous research (e.g., Rabier et al. 1996, Browning et al. 2000, Klein et al. 2002, Ritche and Elsberry 2003), it might have been expected that the largest sensitivity values would only be associated with the upper-level trough that phases with the TC remnants; however, these results show that the evolution of mid-latitude cyclones poleward of the TC that are not associated with the ET can have a large effect the ET forecast. The southerly winds on the east side of the TC advect tropical

air poleward into the mid-latitudes, and therefore increase the horizontal temperature gradient. Coincident with Tokage and Lupit's transition, mid-latitude cyclones initially poleward of the TC undergo cyclogenesis along the resulting temperature gradient. The results shown here suggest that when the TC moves into a region where another cyclone deepens, the ET cyclone is less likely to undergo baroclinic cyclogenesis. As a consequence, the errors in ET forecasts not only originate from how the TC interacts with a mid-latitude trough, but also the evolution of mid-latitude features in the area where the ET storm will reintensify as a baroclinic system.

The ability to improve forecasts of cyclone minimum SLP and the RMS error in SLP forecasts is demonstrated by employing ensemble sensitivity to produce diagnostic corrections to the initial conditions. Since the model is not perfect, these initial conditions may improve the forecast, but are not necessarily best analysis of the current state. The benefit of these perturbed initial condition forecasts is verified by integrating this forecast forward 48 hours and comparing the resulting forecast to the control where no perturbation is applied. For changes of up to 8 hPa, the ensemble prediction of how the initial condition change will affect the 48-hour minimum SLP forecast is in good agreement with the actual difference obtained from the non-linear forecast; however, for larger magnitude changes, non-linearities begin to dominate and the predictions are less accurate. With the exception of the Lupit forecast, diagnostic corrections to the initial conditions can reduce the error in 48-hour SLP forecasts by up to 50%, thus more accurate initial conditions, which can be obtained through additional observations, or improved model physics could lead to better forecasts of these events, though it is not clear which approach would be more beneficial.

Ensemble-based sensitivities are applied to identify how assimilating individual observations at the forecast initialization time affects the 48-hour minimum SLP forecasts shown here. This procedure identifies the ≈ 40 observations that produce a statistically significant change in the minimum SLP forecast. In general, rawinsonde observations near the regions of large initial condition sensitivity, which are located

near the TC and the important mid-latitude features, lead to the largest changes in the expected value and spread of the minimum SLP metric. Rawinsonde observations are important because they are the only source of observation data in the WRF EnKF system over Asia, and directly measure model state variables. For the Nabi and Lupit forecasts, the regions of high sensitivity and identified critical observations are in the dense Chinese rawinsonde network, thus additional observations are unlikely to make large corrections to the initial conditions. For the Tokage forecast, the largest sensitivities are in Mongolia, which is characterized by few rawinsonde observations and thus may have could contribute to the large uncertainty observed in the WRF EnKF forecasts. Moreover, when the most sensitive region is in a sparsely observed area, the forecast could benefit from targeted observations, such as reconnaissance flights, near the TC.

The actual change associated with assimilating these significant observations is determined by generating multiple sets of ensemble forecasts, which differ by the observation set assimilated at forecast hour zero. For each ET forecast studied here, the ≈ 40 important observations account for 50% of the reduction in forecast metric variance obtained when assimilating the 12,000 available observations. When these critical observations are withheld from the analysis, the change in the expected value and spread in the cyclone minimum SLP due to observations is the same as the case where all observations are assimilated because the information in the remaining observations is correlated and thus will produce a similar correction to the analysis state vector. The exception to this result is when the most important observations are in a region characterized by few observations; both the Kirogi and Tokage minimum SLP forecast are different when the important rawinsondes are withheld because there is a lack of surrounding observations to compensate for their absence. Moreover, it can be concluded from these results that when assimilating observations with flow-dependent error statistics, ET forecasts do not require a large quantity of data, but rather a few high quality observations in the locations of large initial condition sensitivity.

Tropical cyclones that deepen as baroclinic systems often lead to the rapid amplification of the mid-latitude flow, and as a consequence, produces an eastward-propagating wave packet that can have significant impact on North American forecasts. It is unclear from the literature what processes are responsible for the genesis of the downstream ridge; therefore the statistics from the short-term ensemble forecasts from the predictability experiments are used to explore how the upper-tropospheric potential vorticity (PV) is affected by latent heating associated with different regions of precipitation. Prior to the onset of transition, a majority of the PV destruction is associated with precipitation along the developing warm-front on the northeast side of the TC, and not the precipitation associated with the TC itself. This warm front forms in response to the large-scale southerly winds on the east side of the TC that advect tropical air poleward into the mid-latitudes. After the onset of transition, the TC is located within the thermal ridge of this baroclinic zone and it becomes difficult to distinguish between the two precipitation regions. As a consequence, the mid-latitude flow during the onset of ET responds to the synoptic-scale latent heating associated with the warm-front and to a lesser extent, the convection within the TC.

The amount of latent heat release within the warm front precipitation region depends on both the magnitude of the vertical circulation, measured by the average frontogenesis, and the lower-tropospheric moisture flux south of the baroclinic zone. Ensemble statistics indicate that the PV within the mid-latitude ridge responds to changes in either of these factors. Multi-variate linear regression where the frontogenesis and moisture flux are the independent variables and the average PV within the ridge is the dependent variable indicates that a one unit change moisture flux leads to a larger amount of PV destruction. This idea is confirmed by using the ensemble statistics to adjust the initial conditions such that there is more (less) water vapor flux at forecast hour zero, integrating the model forward, and comparing the resulting change in 12-hour PV forecasts to the predictions from ensemble statistics. Ensemble predictions of how adjustments to the analysis moisture flux changes the 12-hour

forecast of PV within the ridge are in good agreement with actual change obtained from the non-linear forecast.

As described previously, the distribution of lower-tropospheric water vapor in the global models is often characterized by significant errors, which combined with these results, suggest that these deficiencies could have a large impact on the downstream forecast. Given that UW EnKF dew-point temperature forecasts yield lower errors than global model forecasts and that transition usually occurs over regions of sparse *in situ* data, observations of lower-tropospheric winds and moisture assimilated with flow-dependent error statistics may improve downstream forecasts during ET. Future work may include testing this hypothesis during the upcoming THORPEX Pacific field campaign and through predictability experiments on larger domains that include North America.

BIBLIOGRAPHY

Aberson, S. D., 2003: Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Wea. Rev.*, **131**, 1613–1628.

Aksoy, A., F. Zhang, and J. W. Nielsen-Gammon, 2006: Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model. *Mon. Wea. Rev.*, **134**, 2951–2970.

Ancell, B. and G. J. Hakim, 2007: Comparing adjoint and ensemble sensitivity analysis. *Mon. Wea. Rev.*, **135**, Accepted.

Ancell, B. C. and C. F. Mass, 2006: Structure, growth rates, and tangent linear accuracy of adjoint sensitivities with respect to horizontal and vertical resolution. *Mon. Wea. Rev.*, **134**, 2971–2988.

Anderson, J. L., 1996: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, **9**, 1518–1530.

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

Anderson, J. L. and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.

Anwender, D., M. Leutbecher, S. Jones, and P. Harr, 2006: Sensitivity of ensemble forecasts of extratropical transition to initial perturbations targeted on the tropical cyclone. *27th Conference on Hurricanes and Tropical Meteorology*, American Meteorological Society, Monterey, CA.

Atallah, E. M. and L. F. Bosart, 2003: The extratropical transition and precipitation distribution of Hurricane Floyd (1999). *Mon. Wea. Rev.*, **131**, 1063–1081.

Atangan, J. F. and A. Preble, 2004: 2004 annual tropical cyclone report. Technical report, U.S. Naval Pacific Meteorology and Oceanography Center/ Joint Typhoon Warning Center, Pearl Harbor, HI.

- Atangan, J. F. and A. Preble, 2005: 2005 annual tropical cyclone report. Technical report, U.S. Naval Pacific Meteorology and Oceanography Center/ Joint Typhoon Warning Center, Pearl Harbor, HI.
- Barker, D. M., W. Huang, Y. R. Guo, A. J. Bourgeois, and Q. N. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897–914.
- Barkmeijer, J., M. Van Gijzen, and F. Bouttier, 1998: Singular vectors and the estimates of the analysis-error covariance matrix. *Quart. J. Roy. Meteor. Soc.*, **124**, 1695–1713.
- Bergot, T., 2001: Influence of the assimilation scheme on the efficiency of adaptive observations. *Quart. J. Roy. Meteor. Soc.*, **127**, 635–660.
- Berliner, M. L., Z.-Q. Lu, and C. Synder, 1999: Statistical design for adaptive weather observations. *J. Atmos. Sci.*, **56**, 2536–2552.
- Bierman, G. J., 1977: *Factorization Methods for Discrete Sequential Estimation*, volume 128 of *Mathematics in Science and Engineering*. Academic Press, 241 pp.
- 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.
- Bishop, C. H. and A. J. Thorpe, 1994: Potential vorticity and the electrostatics analogy: Quasigeostrophic theory. *Quart. J. Roy. Meteor. Soc.*, **120**, 713–731.
- Bormann, N., S. Saarinen, G. Kelly, and J.-N. Thepaut, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.*, **131**, 706–718.
- Bosart, L. F. and G. M. Lackmann, 1995: Postlandfall tropical cyclone reintensification in a weakly baroclinic environment: A case-study of Hurricane David (September 1979). *Mon. Wea. Rev.*, **123**, 3268–3291.
- Browning, K. A., A. J. Thorpe, A. Montani, D. Parsons, M. Griffiths, P. Panagi, and E. M. Dicks, 2000: Interactions of tropopause depressions with an extratropical cyclone and sensitivity of forecasts to analysis errors. *Mon. Wea. Rev.*, **128**, 2734–2755.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.

Cardinali, C. and R. Buizza, 2003: Forecast skill of targeted observations: A singular-vector-based diagnostic. *J. Atmos. Sci.*, **60**, 1927–1940.

Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081–3094.

Chang, E. K. M. and I. Orlanski, 1993: On the dynamics of a storm track. *J. Atmos. Sci.*, **50**, 999–1015.

Chang, E. K. M. and D. B. Yu, 1999: Characteristics of wave packets in the upper troposphere. Part I: Northern Hemisphere winter. *J. Atmos. Sci.*, **56**, 1708–1728.

Chen, Y. and C. Snyder, 2007: Assimilating vortex position with an ensemble Kalman filter. *Mon. Wea. Rev.*, **135**, Accepted.

Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. 2. Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113**, 1329–1347.

Dee, D. P. and A. M. DaSilva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295.

Dirren, S., R. D. Torn, and G. J. Hakim, 2007: A data assimilation case-study using a limited-area ensemble Kalman filter. *Mon. Wea. Rev.*, **135**, 1455–1473.

Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005.

Eckel, F. A. and C. F. Mass, 2005: Aspects of effective mesoscale, short-range ensemble forecasting. *Wea. Forecasting*, **20**, 328–350.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rodgers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational Eta model. *J. Geophys. Res.*, **108**, 8851.

- Epstein, E. S., 1969: Stochastic-dynamic prediction. *Tellus*, **21**, 739–759.
- Errico, R. M. and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU-NCAR mesoscale model. *Mon. Wea. Rev.*, **120**, 1644–1660.
- Evans, J. L., J. M. Arnott, and F. Chiaromonte, 2006: Evaluation of operational model cyclone structure forecasts during extratropical transition. *Mon. Wea. Rev.*, **134**, 3054–3072.
- Evans, J. L. and R. E. Hart, 2003: Objective indicators of the life cycle evolution of extratropical transition for Atlantic tropical cyclones. *Mon. Wea. Rev.*, **131**, 909–925.
- 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., 2003: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367.
- Ferreira, R. N. and W. H. Schubert, 1999: The role of tropical cyclones in the formation of tropical upper-tropospheric troughs. *J. Atmos. Sci.*, **56**, 2891–2907.
- Froude, L. S. R., L. Bengtsson, and K. I. Hodges, 2007: The predictability of extratropical storm tracks and the sensitivity of their prediction to the observing system. *Mon. Wea. Rev.*, **135**, 315–333.
- Furze, P. and A. Preble, 2003: 2003 annual tropical cyclone report. Technical report, U.S. Naval Pacific Meteorology and Oceanography Center/ Joint Typhoon Warning Center, Pearl Harbor, HI.
- Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.
- Gelaro, R., R. Buizza, T. N. Palmer, and E. Klinker, 1998: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, **55**, 1012–1037.
- Hakim, G. J., 2003: Developing wave packets in the North Pacific storm track. *Mon. Wea. Rev.*, **131**, 2824–2837.

Hakim, G. J. and R. D. Torn, 2007: Ensemble synoptic analysis. *Fred Sanders Monograph*, L. F. Bosart and H. B. Bluestein, eds., American Meteorological Society, In press.

Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550–560.

Hamill, T. M., 2005: Ensemble-based atmospheric data assimilation: A tutorial. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, eds., Cambridge Press, 124–156.

Hamill, T. M. and S. J. Colucci, 1997: Verification of Eta-RSM short-range ensemble forecasts. *Mon. Wea. Rev.*, **125**, 1312–1327.

Hamill, T. M. and C. Snyder, 2002: Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. *Mon. Wea. Rev.*, **130**, 1552–1572.

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

Harr, P. A. and R. L. Elsberry, 2000: Extratropical transition of tropical cyclones over the western north Pacific. Part I: Evolution of structural characteristics during the transition process. *Mon. Wea. Rev.*, **128**, 2613–2633.

Harrison, M. S. J., D. S. Richardson, K. Robertson, and A. Woodcock, 1995: Medium-range ensembles using both the ECMWF T63 and unified models—An initial report. Technical Report 153, UKMO, Bracknell, Berkshire, UK.

Hart, R. E., 2003: A cyclone phase space derived from thermal wind and thermal asymmetry. *Mon. Wea. Rev.*, **131**, 585–616.

Hart, R. E. and J. L. Evans, 2001: A climatology of the extratropical transition of Atlantic tropical cyclones. *J. Climate*, **14**, 546–564.

Henderson, J. M., G. M. Lackmann, and J. R. Gyakum, 1999: An analysis of Hurricane Opal's forecast track errors using quasi-geostrophic potential vorticity inversion. *Mon. Wea. Rev.*, **127**, 292–307.

Hoffman, R. N. and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, **35**, 100–118.

- Hollingsworth, A. and P. Lonnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data 1. The wind-field. *Tellus*, **38**, 111–136.
- Hong, S. Y., J. Dudhia, and S. H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120.
- Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946.
- 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., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with the ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620.
- Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189.
- Janjic, Z. I., 2002: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso model. Technical Report NCEP Office Note No. 437, National Centers for Environmental Prediction, Camp Springs, MD.
- Joly, A., K. A. Browning, P. Bessemoulin, J. P. Cammas, G. Caniaux, J. P. Chalon, S. A. Clough, R. Dirks, K. A. Emanuel, L. Eymard, R. Gall, T. D. Hewson, P. H. Hildebrand, D. Jorgensen, F. Lalaurette, R. H. Langland, Y. Lemaitre, P. Mascart, J. A. Moore, P. O. G. Persson, F. Roux, S. M. A., C. Snyder, Z. Toth, and R. M. Wakimoto, 1999: Overview of the field phase of the Fronts and Atlantic Storm-Track EXperiment (FASTEX) project. *Quart. J. Roy. Meteor. Soc.*, **125**, 3131–3163.
- Jones, S. C., P. A. Harr, J. Abraham, L. F. Bosart, P. J. Bowyer, J. L. Evans, D. L. Hanley, B. N. Hanstrum, R. E. Hart, F. Lalaurette, M. R. Sinclair, R. K. Smith, and C. Thorncroft, 2003: The extratropical transition of tropical cyclones: Forecast challenges, current understanding, and future directions. *Wea. Forecasting*, **18**, 1052–1092.

Kain, J. S. and J. M. Fritsch, 1990: A one-dimensional entraining detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802.

Kalman, R. E. and R. S. Bucy, 1961: New results in linear filtering and prediction theory. *Trans. of the AMSE-J. Basic Engin.*, **83D**, 35–45.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 364 pp.

Klein, P. M., P. A. Harr, and R. L. Elsberry, 2000: Extratropical transition of western North Pacific tropical cyclones: An overview and conceptual model of the transformation stage. *Wea. Forecasting*, **15**, 373–395.

Klein, P. M., P. A. Harr, and R. L. Elsberry, 2002: Extratropical transition of Western North Pacific tropical cyclones: Midlatitude and tropical cyclone contributions to reintensification. *Mon. Wea. Rev.*, **130**, 2240–2259.

Klinker, E., F. Rabier, and R. Gelaro, 1998: Estimation of key analysis errors using the adjoint technique. *Quart. J. Roy. Meteor. Soc.*, **124**, 1909–1933.

Knippertz, P. and J. E. Martin, 2007: A Pacific moisture conveyor belt and its relationship to significant precipitation event in the semiarid southwestern United States. *Wea. Forecasting*, **22**, 125–144.

Kurihara, Y. M., A. Bender, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev.*, **123**, 2791–2801.

Langland, R. H., 2005a: Observation impact during the North Atlantic TReC-2003. *Mon. Wea. Rev.*, **133**, 2297–2309.

Langland, R. H., 2005b: Issues in targeted observing. *Quart. J. Roy. Meteor. Soc.*, **131**, 3409–3425.

Langland, R. H. and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189–201.

Langland, R. H., M. A. Shapiro, and R. Gelaro, 2002: Initial condition sensitivity and error growth in forecasts of the 25 January 2000 east coast snowstorm. *Mon. Wea. Rev.*, **130**, 957–974.

Langland, R. H., Z. Toth, R. Gelaro, I. Szunyogh, M. A. Shapiro, S. J. Majumdar, R. E. Morss, G. D. Rohaly, C. Velden, N. Bond, and C. H. Bishop, 1999: The North Pacific Experiment (NORPEX-98): Targeted observations for improved North American weather forecasts. *Bull. Amer. Meteor. Soc.*, **80**, 1363–1384.

Leutbecher, M., A. Doerenbacher, F. Grazzini, and C. Cardinali, 2004: Planning of adaptive observations during the Atlantic THORPEX regional campaign 2003. Technical report, European Center for Medium Range Weather Forecasts, Reading, UK.

Liu, Z.-Q. and F. Rabier, 2002: The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J. Roy. Meteor. Soc.*, **128**, 1367–1386.

Lorenc, A. C., 1981: A global 3-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.

Lorenc, A. C., 2005: Why does 4D-Var beat 3D-Var? *Quart. J. Roy. Meteor. Soc.*, **131**, 3247–3257.

Lorenz, E. N., 1963: The predictability of hydrodynamic flow. *Trans. NY Acad. Sci. Series II*, **25**, 409–432.

Ma, S. H., H. Ritchie, J. Gyakum, J. Abraham, C. Fogarty, and R. McTaggart-Cowan, 2003: A study of the extratropical reintensification of former hurricane Earl using Canadian Meteorological Centre regional analyses and ensemble forecasts. *Mon. Wea. Rev.*, **131**, 1342–1359.

Majumdar, S. J., S. D. Aberson, C. H. Bishop, R. Buizza, M. S. Peng, and C. A. Reynolds, 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.*, **134**, 2354–2372.

Majumdar, S. J., C. H. Bishop, B. J. Etherton, I. Szunyogh, and Z. Toth, 2001: Can an ensemble transform Kalman filter predict the reduction in forecast-error variance produced by targeted observations? *Quart. J. Roy. Meteor. Soc.*, **127**, 2803–2820.

McMurdie, L. and C. Mass, 2004: Major numerical forecast failures in the northeast Pacific. *Wea. Forecasting*, **19**, 338–356.

McTaggart-Cowan, R., J. R. Gyakum, and M. K. Yau, 2001: Sensitivity testing of extratropical transitions using potential vorticity inversions to modify initial conditions: Hurricane Earl case study. *Mon. Wea. Rev.*, **129**, 1617–1636.

McTaggart-Cowan, R., J. R. Gyakum, and M. K. Yau, 2003: The influence of the downstream state on extratropical transition: Hurricane Earl (1998) case study. *Mon. Wea. Rev.*, **131**, 1910–1929.

Mitchell, H. L., C. Charette, C. Chouinard, and B. Brasnett, 1990: Revised interpolation statistics for the Canadian data assimilation procedure: Their derivation and application. *Mon. Wea. Rev.*, **118**, 1591–1614.

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

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

Morss, R. E. and K. A. Emanuel, 2002: Influence of added observations on analysis and forecast errors: Results from idealized systems. *Quart. J. Roy. Meteor. Soc.*, **128**, 285–321.

Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 89–125.

Nutter, P., M. Xue, and D. Stensrud, 2004: Application of lateral boundary condition perturbations to help restore dispersion in limited-area ensemble forecasts. *Mon. Wea. Rev.*, **132**, 2378–2390.

Ochotta, T., C. Gebhardt, D. Saupe, and W. Wergen, 2005: Adaptive thinning of atmospheric observations in data assimilation with vector quantization and filtering methods. *Quart. J. Roy. Meteor. Soc.*, **131**, 3427–3437.

Orlanski, I. and J. P. Sheldon, 1995: Stages in the energetics of baroclinic systems. *Tellus*, **47**, 606–628.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics and adaptive observations. *J. Atmos. Sci.*, **55**, 633–653.

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

Petterssen, S. and S. J. Smebye, 1971: On the development of extratropical storms. *Quart. J. Roy. Meteor. Soc.*, **97**, 457–482.

Potter, J., 1964: W matrix augmentation. Technical Report SGA 5-64, M. I. T. Instrumentation Laboratory Memo, Cambridge, MA.

Rabier, F., 2005: Overview of global data assimilation developments in numerical weather-prediction centres. *Quart. J. Roy. Meteor. Soc.*, **131**, 3215–3233.

Rabier, F., H. Jarvinen, E. Klinker, J. F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.*, **122**, 121–150.

Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Uden, J. Eyre, A. Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Quart. J. Roy. Meteor. Soc.*, **124**, 1809–1829.

Ralph, F. M., P. J. Nieman, and G. A. Wick, 2004: Satellite and CALJET aircraft observations of atmospheric rivers over the eastern North Pacific Ocean during the winter of 1997/98. *Mon. Wea. Rev.*, **132**, 1721–1745.

Ritchie, E. A. and R. L. Elsberry, 2003: Simulations of the extratropical transition of tropical cyclones: Contributions by the midlatitude upper-level trough to reintensification. *Mon. Wea. Rev.*, **131**, 2112–2128.

Sinclair, M. R., 2002: Extratropical transition of Southwest Pacific tropical cyclones. Part I: Climatology and mean structure changes. *Mon. Wea. Rev.*, **130**, 590–609.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF Version 2. Technical Report 468+STR, National Center for Atmospheric Research, Boulder, CO.

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

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

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

Stensrud, D. J. and N. Yussouf, 2003: Short-range ensemble predictions of 2-m temperature and dewpoint temperature over New England. *Mon. Wea. Rev.*, **131**, 2510–2524.

Sutton, C., T. M. Hamill, and T. T. Warner, 2006: Will perturbing soil moisture improve warm-season ensemble forecasts? A proof of concept. *Mon. Wea. Rev.*, **134**, 3174–3189.

Szunyogh, I., Z. Toth, R. E. Morss, S. J. Majumdar, B. J. Etherton, and B. C. H., 2000: The effect of targeted dropsonde observations during the 1999 winter storm reconnaissance program. *Mon. Wea. Rev.*, **128**, 3520–3537.

Szunyogh, I., Z. Toth, A. V. Zimin, S. J. Majumdar, and A. Persson, 2002: Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance program. *Mon. Wea. Rev.*, **130**, 1144–1165.

Talagrand, O., R. Vautard, and B. Strauss, 1997: Evaluation of probabilistic prediction systems. *Proceedings, ECMWF Workshop on Predictability*, ECMWF, Reading, UK, 1–25.

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

Tong, M. and M. Xue, 2007: Simultaneous estimation of microphysical parameters and atmospheric state with radar data and ensemble square-root Kalman filter. Part II: Parameter estimation experiments. *MWR*, **135**, Accepted.

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

Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at NMC: the generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.

Toth, Z., E. Kalnay, S. Tracton, R. Wobus, and J. Irwin, 1997: A synoptic evaluation of the NCEP ensemble. *Wea. Forecasting*, **12**, 140–153.

Toth, Z., I. Szunyogh, C. Bishop, S. Majumdar, R. Morss, J. Moskaitis, D. Reynolds, D. Weinbrenner, D. Michaud, N. Surgi, M. Ralph, J. Parrish, J. Talbot, J. Pavone, and S. Lord, 2002: Adaptive observations at NCEP: Past, present and future. *Symposium on observation, data assimilation and probabilistic prediction*, American Meteorological Society, Orlando, FL.

Trenberth, K. E., 1978: On the interpretation of the diagnostic quasi-geostrophic omega equation. *Mon. Wea. Rev.*, **106**, 131–137.

Velden, C., J. Daniels, D. Stettner, D. Santeck, J. Key, J. Dunion, K. Holmlund, G. Dengel, W. Bresky, and P. Menzel, 2005: Recent innovations in deriving tropospheric winds from meteorological satellites. *Bull. Amer. Meteor. Soc.*, **86**, 205–223.

Wandishin, M. S., S. L. Mullen, D. J. Stensrud, and H. E. Brooks, 2001: Evaluation of a short-range multimodel ensemble system. *Mon. Wea. Rev.*, **129**, 729–747.

Wei, M., Z. Toth, R. Wobus, Y. Zhu, C. H. Bishop, and X. Wang, 2006: Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus*, **58**, 28–44.

Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132**, 1190–1200.

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

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2007: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **135**, Accepted.

Wilks, D. S., 2005: *Statistical Methods in the Atmospheric Sciences*. Elsevier Academic Press, 648 pp.

Wu, C. C., P. H. Lin, S. Aberson, T. C. Yeh, W. P. Huang, K. H. Chou, J. S. Hong, G. C. Lu, C. T. Fong, K. C. Hsu, I. I. Lin, P. L. Lin, and C. H. Liu, 2005: Dropwindsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) - An overview. *Bull. Amer. Meteor. Soc.*, **86**, 787–790.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253.

Zou, X., Y.-H. Kuo, and S. Low-Nam, 1998: Medium-range prediction of an extratropical oceanic cyclone. *Mon. Wea. Rev.*, **126**, 2737–2763.

VITA

Ryan Torn was born in September 1979 in Menomonee Falls, Wisconsin. At an early age, he became interested in the weather, and by age 8, he drawing his own forecast maps. After graduating from Menomonee Falls High School in 1998, he enrolled in the University of Wisconsin-Madison, where he graduated with a degree in Atmospheric and Oceanic Science and Mathematics in 2002. During his time there, he was involved in research projects that explored the surface properties of atmospheric aerosols and using satellite data to determine land surface fluxes. Ryan started graduate school during Fall 2002 and began working on projects related to data assimilation and the atmospheric predictability with Greg Hakim. He graduated with his PhD from the University of Washington in 2007 and will join the faculty at the University of Albany in the Fall of 2008 after a one year post-doc position at the National Center for Atmospheric Research.