Growth of Errors and Uncertainties in Medium Range Ensemble Forecasts of U.S. East Coast Cool Season Extratropical Cyclones

A Dissertation Presented

by

Minghua Zheng

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Marine and Atmospheric Sciences

Stony Brook University

December 2016

Stony Brook University

The Graduate School

Minghua Zheng

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Edmund K.M. Chang – Dissertation Advisor
Professor, School of Marine and Atmospheric Sciences

Sultan Hameed – Chairperson of Defense
Professor, School of Marine and Atmospheric Sciences

Brian A. Colle – Committee Member
Professor, School of Marine and Atmospheric Sciences

Minghua Zhang – Committee Member
Professor, School of Marine and Atmospheric Sciences

Sharanya J. Majumdar – External Committee Member
Professor, Department of Atmospheric Sciences
RSMAS, University of Miami

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Growth of errors and uncertainties in medium range ensemble forecasts of U.S. East Coast cool season extratropical cyclones

by

Minghua Zheng

Doctor of Philosophy

in

Marine and Atmospheric Sciences

Stony Brook University

2016

Cool-season extratropical cyclones near the U.S. East Coast often have significant impacts on the safety, health, environment and economy of this most densely populated region. For example, the "January 2015 blizzard" caused thousands of flights cancellations, travel bans enacted in five states, and two related deaths. Hence it is of vital importance to forecast these high-impact winter storm events as accurately as possible by numerical weather prediction (NWP), including in the medium-range (3-6 days). Ensemble forecasts are appealing to operational forecasters when forecasting such events because they can provide an envelope of likely solutions to serve user communities. However, it is generally accepted that ensemble outputs are not used efficiently in NWS operations mainly due to the lack of simple and quantitative tools to communicate forecast uncertainties and ensemble verification to assess model errors and biases. Ensemble sensitivity analysis (ESA), which employs a linear correlation

and regression between a chosen forecast metric and the forecast state vector, can be used to analyze the forecast uncertainty development for both short- (1–2 days) and medium-range forecasts. The application of ESA to a high-impact winter storm in December 2010 demonstrated that the sensitivity signals based on different forecast metrics (the EOF PCs, the MSLP run cycle differences, and the short-range forecast errors) are robust. In particular, the ESA based on the leading two EOF PCs can separate sensitive regions associated with cyclone amplitude and intensity uncertainties, respectively. The sensitivity signals were verified using the leave-one-out cross validation (LOOCV) method based on a multi-model ensemble from CMC, ECMWF, and NCEP.

The climatology of ensemble sensitivities for the leading two EOF PCs based on 3-day and 6-day forecasts of historical cyclone cases was presented. It was found that the EOF1 pattern often represents the intensity variations while the EOF2 pattern represents the track variations along west-southwest and east-northeast direction. For PC1, the upper-level trough associated with the East Coast cyclone and its downstream ridge are important to the forecast uncertainty in cyclone strength. The initial differences in forecasting the ridge along the west coast of North America impact the EOF1 pattern most. For PC2, it was shown that the shift of the tri-polar structure–the East Coast trough and its adjacent ridges–is most significantly related to the cyclone track forecasts.

The EOF/fuzzy clustering tool was applied to diagnose the scenarios in operational ensemble forecast of East Coast winter storms. It was shown that the clustering method could efficiently separate the forecast scenarios associated with East Coast storms based on the 90-member multi-model ensemble. A scenario-based ensemble verification method has been proposed and applied it to examine the capability of different EPSs in capturing the analysis scenarios for historical East Coast cyclone cases at lead times of 1−9 days. The results suggest that the NCEP model performs better in short-range forecasts in capturing the analysis scenario although it is under-dispersed. The ECMWF ensemble shows the best performance in the medium range. The CMC model is found to show the smallest percentage of members in the analysis group and a relatively high missing rate, suggesting that it is less reliable regarding capturing the analysis scenario when compared with the other two EPSs. A combination of NCEP and CMC models has been found to reduce the missing rate and improve the error-spread skill in medium- to extended-range forecasts.

By utilizing the scenario analysis, whether the ensemble mean from the multi-model ensemble or each individual model is really better than other subsets of an ensemble forecast has also been analyzed. It was found that in the majority of cases, the analysis does not lie within Group EM in the multi-mode ensemble. Meanwhile, the quadrant statistics suggest that the ECMWF model misses the analysis direction in a majority of past storms although it shows a slightly higher chance to be in the analysis quadrant in the medium range than the other two EPSs.

Based on the orthogonal features of the EOF patterns, the model errors for 1−6-day forecasts have been decomposed for the leading two EOF patterns. The results for error decomposition show that the NCEP model tends to better represent both EOF1 and EOF2 patterns by showing less intensity and displacement errors during 1−3 days. The ECMWF model is found to have the smallest errors in both EOF1 and EOF2 patterns during 4−6 days. The CMC model shows moderate errors for days 1−2 and the largest errors for days 3−6. We have also found that East Coast cyclones in the ECMWF forecast tend to be towards the southwest of the other two models in representing the EOF2 pattern, which is associated with the southwest-northeast shifting of the cyclone. This result suggests that ECMWF model may have a tendency to show a closer-to-shore solution in forecasting East Coast winter storms

The downstream impacts of Rossby wave packets (RWPs) on the predictability of winter storms are investigated to explore the source of ensemble uncertainties. The composited RWPA anomalies show that there are enhanced RWPs propagating across the Pacific in both large-error and large-spread cases over the verification regions. There are also indications that the errors might propagate with a speed comparable with the group velocity of RWPs. Based on the composite results as well as our observations of the operation daily RWPA, a conceptual model of errors/uncertainty development associated with RWPs has been proposed to serve as a practical tool to understand the evolution of forecast errors and uncertainties associated with the coherent RWPs originating from upstream as far as western Pacific. It suggests that the central and the leading regions of the RWP are the preferable regions for large errors/uncertainties to grow and develop. The errors and spread in a case study for a coherent RWP fit the conceptual model well. The ESA is also performed for this case study and the corresponding sensitivities also qualitatively fit the conceptual model.

To investigate the mechanism of how RWPs affect the downstream predictability, the forecasts errors associated with the downstream development have been investigated under the framework of eddy kinetic energy (EKE) budget based on the same RWP case. The results show that the errors in the total advection term contribute significantly to the errors in the local EKE tendency especially over the eastern U.S. and western Atlantic region. The errors in the ageostrophic flux term play a significant role in the initial development of the EKE errors. A comparison between two ensemble members shows that the ensemble member that has less errors better resolved the initial EKE center with less AGFD errors, leading to a much better forecast of the EKE tendency over the downstream areas (eastern U.S. and western Atlantic).

This work shed lights on improving the understanding of winter storm predictability and NWP model bias and provides new perspectives to communicate forecast uncertainty in predicting cool-season HIW events.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Adv	advective flux
AdvD	advective flux divergence
AGF	ageostrophic flux
AGFD	ageostrophic flux divergence
CMC	Canadian Meteorological Center
ECMWF	European Center for Medium-range Weather Forecasts
EOF	Empirical Orthogonal Function
ESA	ensemble sensitivity analysis
EPS	ensemble prediction system
ETKF	Ensemble Transform Kalman Filter
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
IC	initial condition
MSLP	mean sea level pressure
NAM	North American Mesoscale
NCEP	National Centers for Environmental Prediction
RWP	Rossby wave packet
RWPA	Rossby wave packet amplitude
TIGGE	the Observing System Research and Predictability

	Experiment (THORPEX) Interactive Grand Global Ensemble
Z500	geopotential height at 500 hPa

Acknowledgements

First and foremost, I'd like to gratefully acknowledge my dissertation advisor, Dr. Edmund Chang, for his advisement, support, and encouragement. When I was deciding between two graduate school offers years ago, my MS advisor, Dr. Zuntao Fu at Peking University, told me "*You will learn a lot from Dr. Chang*". Taking his suggestion seriously, I chose to continue my Ph.D. study at SoMAS, and had a great opportunity to join Dr. Edmund Chang's group. After several years, I could say it was the wisest decision I have ever made. He has been teaching me the knowledge in my research field. I used to take one advanced dynamic course he taught, and it proved to be the best dynamical course I have attended. He has been no only providing advice to solve the problems I met in my research but also strengthening my ability to solve the scientific problems independently, and gradually I have learned his way to think of science questions. Besides, he is a role model with very responsible personality and is known to be a scientist of integrity. I am in particular appreciative of all the opportunities he has given me to attend conferences and workshops, so I was able to improve my presentation and communication skills and participated in the atmospheric community. Every rehearsal he attended and every comment he gave for my PPT/papers/dissertation are greatly appreciated. Because there are intelligent and responsible people like him, atmospheric sciences are making progress.

I would also like to acknowledge Dr. Brian Colle for his guidance, support, and help. He provided a lot of useful comments and suggestions to my research and my paper writing. I have benefited a lot from his ideas and discussions. In particular, his operational experience helped me to interpret the ensemble sensitivity and fuzzy clustering results. I have also learned a lot from him on how to communicate science. Special thanks to him for helping me during my job search.

He and Dr. Edmund Chang are advisors who really care about the future of students. It has been a true privilege for me to work with them both.

I greatly thank my Ph.D. committee: Dr. Sultan Hameed, Dr. Minghua Zhang, and Dr. Sharan Majumdar at RSMAS, University of Miami. Dr. Sultan Hameed taught me the statistical methods on his class, which were very useful for my dissertation. He has always encouraged me to do good research and to be more confident. Dr. Minghua Zhang provided valuable advice and useful comments on my dissertation. Dr. Sharan Majumdar offered me specific and constructive feedback on this dissertation, which I believe will help me to complete our papers.

I also thank Dr. Ping Liu and Mark Lang at SoMAS for their technical support. Dr. Ping Liu helped us to recover our disks and has always been helpful in solving Linux problems. Mark Lang helped me fix my PC and print my posters. I am also grateful to my group for communicating the research and discussions. Special thanks to Carol Dovi, Gina Gartin and Christina Fink for the paperwork and reminding me all the deadlines. Thank all my friends at SoMAS for their friendship.

Finally, I thank my husband for his love, patience, and discussions. I am also grateful to my mother-in-law for taking care of our baby during the past year. Thank my parents for their endless love and support. Special thanks to my baby girl. She is the source of my greatest joy and happiness.

Chapter 1 Introduction

1.1 Background and motivations

Extratropical cyclones are dominant feature of the mid-latitudes, as their passage is associated with the day-to-day variability in weather and responsible for a substantial portion of mid-latitude climate. Thus cyclone activity represents an important measure of the state of the atmosphere, and plays important roles in global heat transport, energy redistributions, and regional hydrological cycle. The east coast of the U.S. and the surrounding ocean is one of the regions favorable for cool-season cyclone activity in the Northern Hemisphere. Intense extratropical cyclones along the coast often have large socioeconomic impacts on transportation (e.g., road, aviation, and marine), human health, and property with their strong winds and heavy precipitation. For example, on 26−28 December 2010, one major extratropical cyclone impacted the northeastern U.S., resulting in blizzard conditions across the region, including the New York metropolitan area, New Jersey, and portions of New England (Zheng et al. 2013). The northern part of New Jersey received around 82 cm (32 in.) of snow, while snowfall in excess of 51 cm (20 in.) fell in many places, which, in conjunction with 22−27 m s^{-1} winds, crippled the transportation system from New Jersey to New England during the busy traffic period just after Christmas. A more recent example is the "January 2015 blizzard" that occurred on 26−28 January 2015 (Zheng et al. 2016), which impacted the northeastern U.S., resulting in snowfall accumulations of 30 cm (12 in.) to 91 cm (36 in.) over the central part of Massachusetts, and blizzard conditions were prevalent from Long Island to southern and eastern New England (Winkler 2015). This storm caused thousands of flight cancellations, travel bans enacted in five states, and two related deaths. Considering the high population density of the eastern U.S., accurate forecasts of these storms are very important to reduce their impact, including both human and economic losses.

The skill of numerical weather prediction models in forecasting winter storms has varied in different lead times and regions. Some storms, such as the Superstorm of 1993 along the east coast of the U.S., have been relatively well forecasted several days in advance (Uccellini et al. 1995). In contrast, operational models have poorly predicted other cyclones 1−2 days in advance, such as the 25 January 2000 cyclone event along the North Carolina coast (Zhang et al. 2002), and the aforementioned 26–28 December 2010 blizzard event along the east coast (Zheng et al. 2013). Several poorly predicted cyclone events over the U.S. west coast were found to result from the poor initial conditions (ICs) over the relatively data-sparse upstream Pacific Ocean (McMurdie and Mass 2004). The performance of operational models in forecasting these storms has also varied. For example, Charles and Colle (2009a) found that for the short-term forecast the North American Mesoscale (NAM) model had a larger error in cyclone central pressure than the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) model in forecasting cool season cyclones for the eastern U.S. and western Atlantic Ocean during 2002–07. Meanwhile, Charles and Colle (2009b) found that the 15-member Short Range Ensemble Forecast (SREF) means for both cyclone position and central pressure on average have smaller errors than its subgroups and the NAM model in many regions, but not the NCEP GFS for many forecast times.

The forecast skills are associated with the predictability in the atmospheric flow. Previous studies on predictability are often associated with two types of predictability: one is the intrinsic predictability defined as "the extent to which prediction is possible if an optimal procedure is used" (Lorenz 1963; Lorenz 1969; Lorenz 1996), which places an upper bound upon the atmosphere's predictability; the other is practical predictability defined as "the extent to which we ourselves are able to predict by the best-known procedures, either currently or in the foreseeable future" (Lorenz 1982; Lorenz 1996), which is the actually achievable forecast skill in a given imperfect NWP, or the lower bound of the predictability. Lorenz (1982) proposed an approach to quantify the upper and lower bounds of predictability. The difference between upper and lower bounds was employed to determine how much predictions can be improved. Following this approach, Froude et al. (2013) revisited the predictability of large-scale atmospheric flow. Their results again confirmed the upper limit of two weeks (Lorenz 1963) still applied to the forecasts from the updated ECMWF model. They also pointed out that the forecast skill of the ECMWF system at that time was close to the maximum skill for a deterministic forecast without further reducing the errors in the initial conditions. However, the large differences in the potential predictability curves for the ensemble forecasts indicated that ensemble system has substantial potential in supporting the deterministic system. The above results suggest that ensemble forecasting is an area with enormous potential with respect to the practical predictability especially during the medium and extended range. Given the huge societal impact cyclones can bring to the middle latitudes, it is important to assess the skill of ensemble systems to predict winter storms in order to improve the forecasting models.

The predictability of extratropical cyclones, like other mid-latitude weather systems, has an upper limit, thus prediction of the sufficiently distant future is impossible by any method, unless the present conditions are known exactly. Within the upper limit, the forecast skills rely on a number of factors, including the initial state affected by observing system and data assimilation methods, the forecasting models, and physical parameterizations of unresolved scales. Therefore, the forecast errors and uncertainty for extratropical cyclones result from the forecast model itself (e.g., resolution, physics) and the ICs. Ensemble approach has been demonstrated to improve forecast skills in general when compared with a single-model (deterministic) approach (Toth and Kalnay 1993; Tracton and Kalnay 1993; Molteni et al. 1996; Buizza 1997) because of a variety of ICs, as well as physical parameterizations or models (Toth and Kalnay 1993; Molteni et al. 1996; Buizza 1999). Several recent studies have demonstrated the value of probabilistic information for medium-range forecasts of severe weather events, including the high-impact winter storms (Hewson et al. 2014; Matsueda and Nakazawa 2014; Swinbank et al. 2016). For example, Hewson et al. (2014) investigated two high-impact windstorms affecting northwestern Europe from 28 October 2013 (Christian) and 5 December 2013 (Xaver). For both storms, the probabilistic metrics provided an indication of extreme wind gusts 5−6 days in advance, although the finer details regarding the timing and strength of Christian were not well predicted.

There have been a number of studies to investigate the forecast errors and uncertainties of extratropical cyclone along the east coast of U.S. and the reasons (e.g., model deficiency; rapid error growth; sparse observational data; flow patterns) for them in both deterministic and ensemble models. Silberberg and Bosart (1982) analyzed the cyclone errors over North America and the adjacent oceans during the 1978−1979 cool season in the National Meteorological Center (NMC) Limited Finite Mesh model, and found that the cyclones over western Atlantic Ocean and southeastern U.S. were underpredicted. Corresponding 1000−500 hPa thickness showed a cold error. They emphasized the importance of IC errors over the Pacific in degrading the downstream skills of 1−2-day forecasts for five cyclone cases. Grumm et al. (1992) examined the overall systematic cyclone forecast errors of the Nested Grid Model (NGM) and found that the model cyclones tended to be too deep and moved too slowly in winter mainly due to the model's inability to fill cyclone properly where it was observed to fill. They also indicated that the NGM model had difficulties in simulating orographic effects. Langland et al. (2002) investigated the short- and medium-range (24–96-h) forecasts of the 24–25 January 2000 U.S. east coast snowstorm using the U.S. Navy global forecast model and an adjoint system. They concluded that the substantial loss of predictive skill in the 72- and 96-h forecasts were due to the rapid growth and propagation of small IC errors over a large upstream area, including part of the eastern Pacific and western and central North America. The low predictability regimes during the 22 January-2 February 2000 were characterized by enhanced ridging in the average 500-hPa height field over western North America. Zhang et al. (2002) found that the 1–2-day forecast errors for the same snowstorm case were associated with the rapid growth of error at scales below 500 km in association with moist processes. They also found that the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5) better resolved the precipitation patterns for the storm with increasing horizontal resolution. Elmore et al. (2006) found that the initialization error of upper-level short-wave troughs along the west coast of the U.S. tends to persist with the short wave as it crosses the continental U.S. . Large negative bias observed in the Eta and EtaKF models over the southeast U.S. was due to an error in the longwave radiation scheme interacting with water vapor and clouds.

Previous studies have suggested that the upstream forecast errors over the Pacific impact medium-range predictions associated with cyclone events over the eastern U.S. (Langland et al. 2002; Hakim 2005; Chang 2005; Elmore et al. 2006; Sellwood et al. 2008; Swanson and Roebber 2008; Majumdar et al. 2010; Zheng et al. 2013). However, most of these studies were limited to a small number of case studies, and used old versions of deterministic models or ensemble models for a specific short forecast period. Specifically, there is a lack of comprehensive studies of the source of the forecast errors/uncertainties and the characteristics of them in the current generation of ensemble models. Issues such as the most preferable regions associated with East Coast cyclone intensity and position uncertainties, as well as the growth and propagation speed of forecast uncertainties in the multi-model ensemble, are critical to improve the understanding of the East Coast dynamical relations, to help forecasters to better use ensemble outputs, to server as a guidance for deploying adaptive observations, and to improve ensemble models. Meanwhile, there are also needs to evaluate the performance of operational models in forecasting the high-impact East Coast storms while the ensemble verifications have been very difficult due to the extra dimension of ensemble member.

In order to help fill these needs and improve our knowledge of the predictability of winter storms, this dissertation has three objectives. The first objective is to investigate the development of uncertainties in cyclone intensity and position in the medium range within a multi-model ensemble by employing an ensemble sensitivity approach, which quantifies the sensitivity of a forecast metric at the verification time to the ICs or the ensemble forecast at an earlier forecast time. More details about this approach will be presented later. Quantification of the ensemble sensitivity enables the identification of the spread of the metrics among the ensemble members and the association of the metrics with the characteristics of the ICs or earlier forecasts. The climatologically preferred areas of ensemble sensitivity will be examined. The growth and propagation speed of the ensemble sensitivity will be statistically explored. Understanding of the ensemble sensitivity enables the forecasters to judge which ensemble members are more likely than others after new observations become available or the forecasting cycle is updated or forecaster's personal experiences are considered. It is often flow dependent. The goodness of ensemble sensitivity signals will also be explored by using the leave-one-out cross validation (LOOCV) method.

Ensemble sensitivity provides an overall linear relation between the forecast metrics and the initial conditions. However, the development scenarios may vary especially in a combination of multiple ensemble models. In the operational forecasting process, forecasters need to anticipate and evaluate the potential of the worst case scenario. Given a large dataset of many ensemble members, effective tools to extract the ensemble information and provide different development scenarios of a potential storm will benefit both the forecasters and the decision makers. Meanwhile, systematic evaluations of model biases with respect to winter storm forecast scenarios could be employed as guidance to benefit the forecasters and model developers.

Therefore, the second objective of this dissertation is to apply a fuzzy clustering tool to operational ensemble forecasts for East Coast cyclones to separate scenarios related with cyclone intensity and position uncertainties. Building on the successful applications of the fuzzy clustering tool to case studies and operational daily ensemble forecasts, we have proposed a scenario-based ensemble verification approach to evaluate the performance of the different models in simulating East Coast winter storms. Important ensemble properties, such as position and location biases, as well as the error-spread relation, will also be investigated.

In our study of the ensemble sensitivity signals of extratropical cyclones, we have found that the ensemble sensitivity signals, which are closely associated with forecast uncertainty/error development, resemble the development and propagation of Rossby wave packets (RWPs). Other studies have also suggested that extratropical cyclones are often associated with an RWP in the middle and upper troposphere. Orlanski and Sheldon (1995) proposed an energetic framework of downstream baroclinic evolution, and examined the trough formation over the east coast of North America associated with the “Blizzard of 93”. They concluded that the formation of the trough was primarily due to downstream propagation of Rossby wave energy from disturbances over the northeast Pacific. Moreover, they pointed out that under environmental conditions in which downstream development leads to packets of eddies with a life span much longer than individual eddies; the detection of such packets and energy transfer processes may impact the longer range predictability of such systems. Chang and Yu (1999) showed that there were downstream-developing wave trains propagating from the eastern end of the Pacific storm track across North America toward the Atlantic, suggesting that wave developments near the entrance of the Atlantic storm track are influenced by waves from the Pacific. Hakim (2003) showed that the Atlantic storm track is often seeded by wave packets that originate over the western North Pacific Ocean by studying a sample of North Atlantic storm track events. Forecast errors also tend to develop wave-packet-like structures (Hakim 2005) and propagate rapidly downstream with the group velocity (at least as fast as the upper-level flow). Model errors during the “surprise” January 2000 East Coast snowstorm also propagated with an RWP from eastern Pacific (Langland et al. 2002). Moreover, when a coherent RWP was present and the flow was zonal, distinctive targeted regions, which are associated with maximum forecast error reduction within the verification regions over U.S., could be traced upstream to near Japan at lead times of 4–7 days (Majumdar et al. 2010).

Thus, the third goal of this dissertation is to investigate the connection between RWPs and the forecast uncertainties and errors of winter weather over eastern U.S. and western Atlantic Ocean in the GFS and GEFS forecasts. The role of downstream development of RWPs in the growth and propagation of forecast errors will be explored. This part of work will help to improve the understanding of the source of forecast errors and uncertainties and diagnose the rapid downstream error propagation.

The remainder of this chapter will proceed as follows. An overview of East Coast cyclone climatology and dynamics will be presented in section 1.2. A brief introduction of numerical weather prediction and ensemble forecasting will be given in section 1.3. Section 1.4 will give a short review of the application of ensemble forecasting to East Coast cyclones. Section 1.5 will discuss the useful tools to examine the forecast sensitivity to initial conditions. Section 1.6 will present the detailed science questions that this dissertation will try to answer. The outline of this dissertation will be given in section 1.7.

1.2 East Coast cyclone climatology and dynamics

The U.S. East Coast and its adjacent ocean is a favorable region for cool-season cyclogenesis mainly due to the strong thermal and moisture contrasts between land and ocean. There have been numerous studies that documented the principal features of East Coast storms with respect to their development and classification. The studies by Austin (1941) and Miller (1946) laid the groundwork for the climatology and dynamics of the East Coast winter storms. Miller (1946) classified the East Coast storms into two basic categories according to their formation mechanisms. Type A cyclones often emerge as cyclone waves along cold fronts, typically over the southern U.S. or the Gulf of Mexico. The cold fronts are often of Polar or Arctic origin, separating cold continental air mass from maritime tropical Gulf or Atlantic air mass. The developing surface lows track northeastward out of the Gulf of Mexico, in some cases paralleling the Atlantic coastline and intensifying further. Miller Type A cyclogenesis is a common occurrence in regions where cold air outbreaks occur frequently, but it is most often observed along east coasts in the cool seasons. Miller Type B cyclones originate near the coastline to the southeast of a preexisting cyclone over the region of the Great Lakes. The preexisting cyclone or the primary cyclone occludes west of the Appalachians. As it dissipates in the Ohio Valley, a secondary low begins to develop along the warm front of the older cyclone. Typically, the warm front extends to the coastal waters where it connects with shallow boundaries that form locally near the coast. The secondary low and associated fronts separate the shallow wedge of cold air east of the Appalachians, usually as a result of cold-air damming, from the warmer air over the Atlantic Ocean. This type is not as common in other parts of the world and owes its frequency to the peculiar topography of North America. The Miller classification scheme remains to this day a useful means of cataloging many cyclogenesis events according to the surface features. Reitan (1974) estimated the most frequent paths of cyclones during the time period of 1951 to 1970. For the cool seasons, they distinguished the paths for storms over the northeastern U.S. as from the west, from the southwest, from the southeast and over the western Atlantic oceans. Zishka and Smith (1980) studied the climatology of cyclones and anticyclones over North America and the surrounding oceans for January and July, 1950–77. Three preferred cyclone tracks for January were found to impact the eastern U.S. and western Atlantic. Two begin in Alberta and Colorado, continue east-southeast and northeast respectively, and then merge over the Great Lakes before taking a more northward track into Canada. The third generally begins off the Virginia-North Carolina coast, although some cyclones originate as far south as the Gulf of Mexico. This track often moves parallel to the eastern seaboard of North America and then turns northward toward Greenland. Kocin and Uccellini (1990) showed that most of their major East Coast snowstorms originate to the south of 35°N. Since the average position of the Gulf Stream veers sharply away from the coast near 35°N, the interaction of cold air mass over the coastal plain and the warm Gulf Stream waters might promote explosive cyclogenesis to the south of 35°N. They also found that strong anticyclones often precede East Coast winter storms.

Mather et al. (1964) analyzed the temporal climatology of eastern U.S. cyclones by investigating records of coastal storms and water damage included in weather summaries, newspapers, and periodicals. They found that coastal storms of moderate or severe intensity affected the New Jersey and New York coast on average once every 1.4 yr. Davis and Dolan (1993) examined the synoptic characteristics of East Coast winter storms including wave data as a criterion. They defined storms producing deep-water waves greater than 1.6 m at Cape Hatteras, North Carolina from 1943 to 1984 as coastal storms. They found that the most dangerous storms are cyclones which originate either over Florida or north of Cuba from October to April, traveling northward, and are blocked by a stagnating anticyclone over Near England or the North Atlantic. They also found that coastal storm annual frequencies declined from the mid-1960s through the mid-1970s and then increased through 1984, but the frequency of potentially damaging storms had increased since 1965. Hirsch et al. (2001) developed the climatology of East Coast winter storms using an automated procedure for the NCEP-NCAR reanalysis dataset (1948, 1951–97). They included a strong wind threshold (>10.3m/s for at least 6 h) in their criteria for defining the storms. The East Coast winter storms were found to be relatively intense compared with other storms over the continental U.S. On average, 11.8 East Coast winter storms were found per season (Oct–Apr) and storms count peaks from December to February. They also found that the inter-annual variations are associated with ENSO, with the El Nino favoring the East Coast winter storms. Bernhardt and DeGaetano (2012) investigated the speeds of historical cool-season extratropical cyclones along the U.S. east coast during the period from 1951 to 2006.

The average storm speed was found to be 13.8 m s^{-1} with stronger storms generally moving faster than weaker storms. They also concluded that the coexistence of a strong El Nino and the negative NAO phase favors a slower propagation speed of the cyclones. Booth et al. (2015) analyzed the wintertime high-wind events in the northeastern U.S. associated with extratropical cyclones for 1979‒2012. Booth et al. (2015) analyzed the wintertime high-wind events in the northeastern U.S. associated with extratropical cyclones for 1979□2012. They showed that the storms associated with multi-station high-wind events are most likely to approach the northeastern U.S. from the southwest based on the frequency of the tracks in each of the pathways. They concluded that the extratropical cyclones that are associated with the strongest wind events often occur over land instead of the Atlantic Ocean, suggesting that they are not The same as those that cause storm surge (i.e., Nor'easters, Davis and Dolan 1993).

To sum up, cyclones affecting the eastern U.S. generally propagate along four tracks: Alberta Clippers that form over western Canada and track first east-southeastward across southern Canada before turning northward, Colorado lows that form over the southwest of the U.S. and track northeastward towards the Great Lakes, Gulf lows that form over the Gulf of Mexico that track either northward along the Mississippi-Ohio river valley or east-northeastward across southeastern U.S. before turning north-northeastward along the east coast, and the Hatteras lows that form off the Carolinas and track north-northeastward along the coast (Colle et al. 2015). Most of the severe winter weather event (e.g., high winds, heavy snowfall, coastal storm surge) over the eastern U.S. are associated with them.

There exists a rich literature that studied the detailed dynamical mechanisms associated with East Coast winter storms. The early work on extratropical cyclone self-development mechanism by Sutcliffle (1947) and Petterssen (1956) emphasized the positive feedback mechanism in which an approaching upper-level wave to a lower baroclinic zone acts to intensify surface features, while the pattern of thermal advection associated with the low-level circulation intensifies the upper-level features. Shapiro and Keyser (1990) updated the Norwegian frontal-cyclone model (Bjerknes 1919; Bjerknes and Solberg 1922) by using data from various observational and numerical modeling studies. The Shapiro-Keyser conceptual model includes four stages of the life cycle of a marine extratropical frontal cyclone: it starts as an incipient frontal cyclone, developing toward a frontal fracture near the cyclone center and becoming a frontal T-bone and bent-back warm front, and ultimately takes on a warm-core seclusion. The lifecycle of many Northeast cool-season cyclones takes place off the east coast of the U.S., and thus evolve in a similar manner as that proposed by Shapiro and Keyser (1990). Orlanski and Sheldon (1995) proposed another conceptual model of "downstream baroclinic evolution" from an energetics standpoint based on the results from case studies and idealized model experiments. They considered the transfer of kinetic energy between individual synoptic-scale systems due to an ageostrophic geopotential flux and divided the system into three stages. In *stage* 1, a pre-existing disturbance well upstream of an incipient trough weakens as it radiates energy via ageostrophic geopotential fluxes through a downstream ridge. The convergence of these fluxes downstream generates a new energy center on the western side of the trough. In *Stage 2*, this new energy center grows robustly, at first due to the convergence of these fluxes, and later by baroclinic conversion of eddy available potential energy (cold air sinks into the base of the trough). As the center matures, it begins to radiate energy via geopotential fluxes through the trough to the downstream side of the trough, initiating yet another energy center. In *Stage 3*, this new energy center continues to grow owing to the baroclinic conversion from eddy available potential energy to eddy kinetic energy (warm air rise), while that on the western side of the trough decays because of a decreasing supply of energy through fluxes from the older upstream system as well as its own export of energy downstream. As the eastern energy center matures, it may radiate energy further through the downstream ridge, and the sequence begins to repeat. The development of the trough associated with the U.S. "Blizzard of 93" fitted the conceptual picture of downstream baroclinic evolution quite well, with geopotential fluxes playing a critical role.

Uccellini and Kocin (1987) investigated the role of vertical transverse jet streak circulations in East Coast snowstorm events. They showed a link between the configuration of the surface and upper-level features, demonstrating that ageostrophic circulations associated with coupled jet streaks can connect the upper-level troughs and jet streaks to the orientation of surface cyclones, temperature advection patterns, moisture transport and the vertical motion pattern needed to produce heavy precipitation in East Coast snowstorms. A typical scheme for an East Coast snowstorm often develops both indirect and direct transverse circulations that span the entire troposphere in the exit and entrance regions of the southern and northern jet streaks, respectively. The advection of cold Canadian air southward in the lower branch of the direct circulation across the northeastern U.S. maintains the cold lower-tropospheric temperatures needed for snowfall along the East Coast. Meanwhile, the development of low-level jet streak (LLJ) within the lower branch of an indirect circulation can significantly increase the moisture transport into the region of heavy snowfall, such as for the Presidents' Day storm in 1979 (Uccellini et al., 1984). The LLJ also contributed to the decreasing sea level pressure that marked the secondary cyclogenesis along the coast for that case.

Kocin and Uccellini (2004) later described the evolution of surface and upper-level features associated with thirty northeastern U.S. snowstorms. For the typical jet streak circulation patterns common during Northeast snowstorms, the coupling between the thermally direct transverse circulation in the entrance region of the northern jet streak and thermally indirect transverse circulation in the exit area of the southern jet streak serves as enhanced forcing for surface cyclogenesis. Lupo et al. (1992) and Rolfson and Smith (1996) showed that cyclone development takes place during cyclonic vorticity advection and vertically integrated horizontal warm air advection, often accompanied by latent heat release. The above features usually lead to an increase in surface geostrophic vorticity and a decrease in central MSLP. Warm air advection will be especially effective if the cyclone lies downstream from a stratospheric warm pool associated with a stratospheric-level trough, while cyclonic vorticity advection usually occurs in association with an upstream mid- to upper-level trough and a jet streak (Sanders 1986).

As for the intense cyclones over eastern U.S., meteorologists employ “explosive cyclogenesis” to describe the rapid deepening of a low pressure center during cyclone formation. A meteorological bomb was defined by Sanders and Gyakum (1980) as an extratropical surface cyclone whose central pressure dropped an average of at least 1 hPa h^{-1} for 24 h at 60°N, normalized geostrophically at other latitudes. The “self development” mechanism is considered to be closely associated with the most intense cyclones (Sutcliffe 1947; Petterssen 1956; and Kocin and Uccellini 2004). Kocin and Uccellini (2004) noted that the most intense cyclones tend to be associated with the evolution of a 500-hPa trough into a closed vortex accompanied by self-development features, while the longest duration cyclones consist of a closed, or cutoff, 500-hPa circulation prior to cyclogenesis. Some studies also emphasized the importance of trough mergers, which is defined as the amalgamation of two separate 500-hPa vorticity centers, or shortwaves, into one coherent system (Gaza and Bosart 1990; Hakim et al. 1995, 1996; and Strahl and Smith 2001). Gaza and Bosart (1990) showed that trough mergers can lead to the development of a negatively tilted (northwest-southeast) upper-level trough, enhanced thermal and differential vorticity advection, and ultimately intense surface cyclogenesis using 21 merger events across North America during 1978−1985. The phasing of the northern and the southern stream shortwaves was observed in numerous historical cases such as the Cleveland "Superbomb" (Salmon and Smith 1980; Hakim et al. 1995, 1996). A southern stream shortwave and its embedded vorticity maximum that phases with a northern stream shortwave will produce more favorable dynamics for rapid surface cyclogenesis than a southern stream shortwave that damps the northern stream shortwave or ejects around its periphery. Winters and Martin (2016) suggested that certain trough merger cases may be simultaneously characterized by polar/subtropical jet superposition, which plays a primary role in the development of a surface cyclone during 18–20 December 2009 Mid-Atlantic Blizzard.

1.3 Numerical weather predictions and ensemble forecasts

Current numerical weather prediction models basically employ ensemble forecasting, the idea of which dates back to Lorenz (1963). Lorenz chose three ordinary differential equations to represent a convective process and found that all numerical solutions of the equations turned out to be unstable. When applying the instability of the non-periodic flow to the atmosphere, it indicates that prediction of the sufficiently distant future is impossible by any method, unless the present conditions are known accurately. Lorenz (1969) further concluded that the atmosphere is chaotically unstable with respect to perturbations of even small amplitudes and impacts as seemingly trivial as a butterfly flapping its wings will potentially determine whether or not a tornado forms in another part of the world. Since the atmospheric flow is sensitive to small perturbations, weather forecast skill is associated with the proper specification of the conditions used to initialize the model forecast equations. Judd and Smith (2001) further demonstrated that even under the ideal conditions of a perfect model and infinite past observations, it is impossible to make exact state estimations of a deterministic nonlinear system due to the uncertainty in the observations. Consistent with the noisy observations there is a set of states indistinguishable from the true state. Their result implies that an accurate forecast must be based on a probability density function (PDF) on the indistinguishable states. Accordingly, probabilistic approaches that recognize the unavoidable uncertainty of the true atmospheric state were gradually developed and applied to weather forecasting.

As indicated previously, numerical weather prediction is an initial-value problem: given an estimate of the present state of the atmosphere, the model simulates (forecasts) its evolution. Thus, either deterministic or probabilistic model requires a proper determination of the initial condition. To tackle the problem that the atmosphere is sensitively dependent to initial conditions, ensemble forecasts were proposed by launching a set of forecasts from different initial states instead of a single deterministic forecast from a single state (Leith 1974, Molteni et al. 1996). These studies showed that even a small ensemble size yields adequate approximation thereby providing a sufficiently accurate and computationally feasible approximation to the primitive equations. The accuracy of ensemble forecasting depends on whether initial members are drawn from the PDF that best approximated the uncertainty of the initial time state. Thus, the application of Monte Carlo ensemble forecasting requires an estimate of the current atmospheric state and its associated uncertainty. Variational assimilation (e.g., Talagrand and Courtier 1987) or Kalman filter approach (e.g. Welch and Bishop 2001; Houtekamer and Mitchell 1998) have been employed to approximate the state of current atmosphere and its associated uncertainty. Subsequently, the estimate of the current atmospheric state and its associated uncertainty are used to provide the initial conditions for the ensemble forecasts that yield the flow dependent output statistics.

Once an ensemble is initialized by one of the initialization methods (such as random perturbation, bred vector, or singular vector techniques), each ensemble member is used as the initial conditions for an NWP model run. Assuming that the initial ensemble correctly samples the initial uncertainty distribution, each individual ensemble forecast state defines an equally likely realization of the future atmospheric state and the joint distribution of ensemble forecasts defines the forecast uncertainty distribution (e.g. Kalnay 2003).

The value of ensemble forecasting and the related probabilistic information for high-impact weather events has been demonstrated in some recent studies using TIGGE data (Hewson et al. 2014; Matsueda and Nakazawa 2015; Swinbank et al. 2016). Matsueda and Nakazawa (2015) have developed a prototype suite of ensemble-based early warning products for severe weather events, using both single-model (ECMWF, JMA, NCEP, and Met Office) and multimodel grand ensembles. Several case studies (Swinbank et al. 2016) have demonstrated the ability of the products to successfully predict severe weather events, including the Russian heatwave in 2010, the 2010 Pakistan floods, and Hurricane Sandy in 2012. They showed that the grand ensemble provides more reliable forecasts than single-center ensembles, particularly with respect to strong wind speeds and severe temperature, aiding the advance detection of severe weather events to help mitigate the associated catastrophic damage.

1.4 Ensemble forecasts of East Coast cyclones

Cyclone forecasts using ensemble modeling have also been investigated in a number of studies (Du et al. 1997; Stensrud et al. 1999; Froude 2007; Charles and Colle 2009b; Froude 2009; Froude 2010). Froude et al. (2007) investigated the cyclone tracks in the 50-member ECMWF ensemble prediction system (EPS) and 10-member NCEP EPSs between 6 January and

5 April 2005 using an objective feature tracking methodology to identify and track cyclones along the forecast trajectories. They found that the ECMWF ensemble on average had a higher forecast skill than the NCEP ensemble for cyclones in the NH while in the SH the NCEP ensemble had smaller errors. Both EPSs show a higher level of forecast skill for the cyclone positon than their intensity. The propagation speed of cyclones is generally too slow in the ECMWF EPS. Both ECMWF ensemble mean and the best ensemble member have greater accuracy than control forecast for both the position and intensity of the cyclones although the ECMWF ensemble is underdispersed. Froude (2010) further analyzed the predictions of extratropical cyclones in the NH by nine EPSs from the TIGGE archive between 1 February 2008 and 31 July 2008. They again showed that the ECMWF ensemble has a higher predictive skill for all aforementioned cyclone properties. However, the ECMWF model consistently over-predicts cyclone intensity, although the bias is small. The JMA, UKMO, NCEP, and CMC have 1 day less skill for the position of cyclones throughout the forecast lead times. The NCEP model has larger errors for cyclone intensity than for position. It was also found that cyclones in all EPSs propagate too slowly.

As for extratropical cyclones over the eastern U.S. and the surrounding ocean, Charles and Colle (2009b) comprehensively verified the strengths and positions of storms around North America and the adjacent oceans within the SREF system of NCEP during 2004–07. They found that the SREF has slightly more accuracy over the eastern U.S. and western Atlantic Ocean than the western U.S. for the cyclone central pressure. The 15-member SREF mean for both cyclone position and central pressure on average has a smaller error than its subgroups and the NAM model in many regions, but not the GFS for many forecast times. The SREF probabilities are fairly reliable, although it is overconfident at higher probabilities in all regions. Colle and

Charles (2011) further examined the spatial distribution and evolution of sea level pressure (SLP) forecast errors across North America and its adjacent oceans in the NCEP GFS model in short- to medium-range. The forecast cyclones during 72−120 h are too weak on average by 2-3 hPa near the U.S. east coast. These cyclones move too fast in the GFS model at the medium range while being too slow and too far west for the short range.

Several studies have investigated the forecast errors and uncertainties related to high-impact East Coast winter storm cases in ensemble forecasts. For example, Zhang et al. (2002) showed that the 1−2-day forecast errors for the 24−25 January 2002 snowstorm case are associated with the rapid growth of errors at scales below 500 km associated with moist processes. They initialized the MM5 with several different initial conditions and found that an IC ensemble significantly improved the forecast. Novak and Colle (2011) compared the forecast uncertainty of mesoscale snowband formation and evolution based on predictions from a 16-member multimodel ensemble at 12-km grid horizontal spacing for the 25 December 2002, 12 February 2006, and 14 February 2007 northeastern U.S. snowstorms. They showed that even for very short range forecast considerable uncertainty in the timing and location of band formation and subsequent evolution existed. Among the three cases, large (small) initial differences in the upper-level PV and MSLP of the incipient cyclone were associated with large (small) differences in predicting snowband locations, which suggests that case-to-case differences in predictability may be related to the quality of the ICs. They also noted the complexity of the initial flow may also be a discriminator. Zheng et al. (2013) investigated the ECMWF ensemble forecast of the 26−28 December 2010 winter storm and found that the differences in initializing the short-wave trough over Texas led to the forecast jump with respect to cyclone position at the verification time.

Overall, there has been a lack of systematic studies on ensemble forecasts of East Coast storms in the literature mainly due to the difficulty for the research community in obtaining ensemble data and conduct numerical experiments with sufficient members. Ensemble data has one more dimension than other data, making it even harder to process and verify. However, given the prevalence of storms over the U.S. East Coast and the strong connections between the extratropical cyclones and extreme weather, a comprehensive understanding of the growth of errors and uncertainties in ensemble forecasts are necessary to improve the forecast accuracy and models.

1.5 Tools to examine forecast sensitivity to initial conditions

A small error in a meteorologically active region can be more significant than a major error in an inactive region for a short-term forecast. 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 assimilate additional observations in that area. Several objective techniques, such as singular vectors, adjoint sensitivity, Ensemble Transform Kalman Filter (ETKF) and ensemble sensitivity, have been used to quantify how small changes to the initial conditions will affect a forecast metric. We will review them in this section.

1.5.1 Singular vector, adjoint sensitivity and ETKF analysis

Singular vectors are a set of orthogonal structures that have different growth rates over a finite forecast time period for a given forecast metric, where the leading singular vectors represent the directions of fastest error growth in the forecast model optimized for a specific verification time (Molteni and Palmer 1993; Buizza and Palmer 1995; Molteni et al. 1996; Buizza et al. 1997; Palmer et al. 1998). Ensemble forecasting needs initial perturbations which grow with the evolution of atmospheric flow. The optimal perturbations are achieved from the singular vector decomposition of a propagator, which maps small perturbations from initial time to a later time. The structures of singular vectors depend on the choice of forecast metric. Palmer et al. (1998) showed that initial time singular vectors based on the total energy norm can represent the analysis-error variance better than the entropy and streamfunction variance, thus making it the norm of choice for ensemble forecasting. Buizza and Palmer (1995) have identified three preferred regions for singular vector growth: the East Asian/west Pacific sector, the northeast U.S./western Atlantic sector, and subtropical North Africa. The first two are familiar regions of cyclogenesis, corresponding to areas where high-pass transient variances are maximized. The ensemble forecasting systems at ECMWF has been based on linear combinations of initial-time singular vectors (Molteni and Palmer1993, Molteni et al. 1996). The singular vectors have also been used to determine the targeting regions for adaptive observations (Buizza and Montani 1998; Cardinali and Buizza 2003). 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 observational network. Buizza and Montani (1998) used the singular vectors with maximum energy at the final time inside a verification region to identify the target area where adaptive observations should be taken, at an earlier time, to reduce the forecast error over the verification area. Their results suggest that both the singular vectors and the target area are not only sensitive to the choice of the verification region, but also to the choice of the trajectory along which the singular vectors are evolved. The results based on the first 4 or 10 singular vectors are very similar. They also found that the root-mean-square errors over the verification area could be reduced by up to 13% by adding targeted observations.

Several authors have also applied adjoint technique to determine the initial condition sensitivity of mid-latitude cyclogenesis. In order to calculate adjoint sensitivity, the ensemble mean initial condition is first integrated forward with the nonlinear model; this trajectory is considered as the basic state for the adjoint sensitivity calculations. Next, the adjoint model is integrated backward in time from the gradient of the forecast metric at the forecast time. The adjoint sensitivity can be used to infer the importance of various synoptic-scale features of the cyclone's environment to the intensity of that cyclone. 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. Langland et al. (1995) used the adjoint model to examine the sensitivity of an idealized dry extratropical cyclogenesis simulation to perturbations of predictive variables and parameters during the cyclone life cycle. They found that the largest sensitivity for both temperature and wind perturbations is located between 600 and 900 hPa within the baroclinic zone above the developing cyclone. Zou et al. (1998) showed that five-day forecasts of an eastern U.S. 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 U.S., the vertically integrated adjoint 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.

ETKF is another technique to infer the relation between ICs and a forecast metric over a given verification area and provide targeting guidance (Bishop et al. 2001; Majumdar et al. 2002).

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 at analysis time. Rather than describing how the forecast metric value responds to initial condition errors, this technique shows how the forecast error variance within a given verification region at a verification time responds to assimilating observations at an analysis time. 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 a linear, increasing statistical relationship between the ETKF predicted signal variance and the variance of NCEP signals produced during the 2000 Winter Storm Reconnaissance (WSR) program at both the targeting and verification times. The ETKF method was used in NOAA operations to plan WSR missions, and proved to be an effective technique for short-range forecasts (Majumdar et al. 2001; Majumdar et al. 2002). Szunyogh et al. (2000, 2002) reported that in 70% of the WSR cases, the forecast errors are reduced, though the average improvement in the forecasts is only about 10%, or a 12 hour gain in forecast skill. Frequently, the optimal target area for forecasts is a single region where strong mid- and upper-tropospheric winds exist or a region of baroclinic instability (Petersen et al. 2007). Sellwood et al. (2008) investigated the capability of the ETKF approach to predict the impact of WSR dropwindsonde data released over northeastern Pacific Ocean on medium-range forecasts downstream. Employing a 51-member ECMWF ensemble, they concluded that the ETKF is capable of discriminating between observation locations that are effective or ineffective for 3–6-day NCEP GFS forecasts of 200-hPa winds within a verification region based on Rossby wave dispersion, if the flow was predominantly zonal. The ETKF performed poorly in blocked flows. Majumdar et al. (2010) used the ETKF to investigate medium-range forecasts for 20 cases of potential high-impact weather over the continental U.S.

using a 145-member ensemble comprising perturbations from NCEP, ECMWF, and the Canadian Meteorological Centre (CMC). They showed that when a coherent RWP was present and the flow was zonal, distinctive targeted regions determined by ETKF, which are associated with maximum forecast error reduction within the verification regions, could be traced upstream to near Japan at lead times of 4–7 days.

1.5.2 Ensemble sensitivity analysis

Both singular vector and adjoint sensitivity involve the linearization of the forecast model, and thus they may not work well in highly non-linear flow. Also, both methodologies are computationally expensive. Besides, the ETKF is only perfect when linear dynamics hold and the model and error covariance specification are perfect. An alternative and inexpensive approach is thus necessary to quantify the relation between a forecast and initial conditions.

A computational inexpensive ensemble error-covariance approach to sensitivity analysis was proposed by Zhang (2005) to investigate the dynamics and structure of mesoscale systems in the intensive extratropical cyclogenesis event that occurred on 24–25 January 2000. Zhang et al. (2006) tested the significance and effectiveness of the error covariance estimated from the ensemble forecasts in an EnKF data assimilation system based on MM5 for this event. They found that the EnKF was most effective in reducing larger-scale errors but less effective in reducing errors at smaller, marginally resolvable scales. Hawblitzel et al. (2007) employed the ensemble correlation on the base of error-covariance study to examine the dynamical relations and predictability of the mesoscale convective vortex (MCV) of 10–13 June 2003 through ensemble forecasting. The above work demonstrated the value of covariability techniques in inferring the dynamical linkage and couplings of initial state vector and forecast state vector in mesoscale systems.

Torn and Hakim (2008) further developed ensemble sensitivity analysis (ESA) as a practical co-variability technique, which uses both analysis and forecast ensemble data from an ensemble forecasting 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 straightforward and inexpensive. Ancell and Hakim (2007) showed that ensemble sensitivity is related to the adjoint sensitivity through 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 provided an optimal strategy for observation targeting because it incorporates information on the analysis error, observation error, dynamical error growth and data assimilation.

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. Chang et al. (2013) applied the ensemble sensitivity to medium range forecasts for two explosive Pacific cyclone and found the ensemble sensitivity can be very effective in predicting cyclone central pressure and position differences in ensemble members. Ensemble sensitivity has been used in a number of studies to infer the dynamical relation between a forecast metric and the state variables for weather systems of different scales (Hakim and Torn 2008; Torn 2010; Matsueda et al. 2011; Chang et al. 2013; Zheng et al. 2013; McMurdie and Ancell 2014; Bednarczyk and Ancell 2015). Whereas the above studies mainly explored ensemble sensitivities for case studies, this dissertation will apply this technique to a combined ensemble of operational models from the TIGGE archive to illustrate how ensemble sensitivities can be used to compute climatological sensitivities of the intensity and position forecast for East Coast cyclone. Meanwhile, there has been only a limited study to verify the robustness of ESA signals. The initial conditions derived based on ESA are used to perform perturbed ensemble experiment in a few case studies (Torn and Hakim 2009; Chang et al. 2013). While their results from the perturbed model experiments have verified the ESA relations in the corresponding case studies, it could be computationally very expensive and inefficient to perform ensemble experiments to verify a large number of cases in practice. Thus this study will use an alternative method (LOOCV) to verify the robustness of the ESA signals in ensemble members.

1.6 Science questions

One of the benefits of an ensemble forecast has when compared with a deterministic forecast is the additional probabilistic information provided by the ensemble envelope. This can provide a measure of the predictability of the atmosphere at a particular location and time. Large ensemble spread often indicates that different weather developments are possible. Forecasters have increasing ensemble guidance available, but ensemble data is often not used effectively since (1) ensembles have not been comprehensively verified and evaluated; (2) ensemble under-dispersion and biases limit ensemble skill; (3) forecasters lack tools to understand the origin of ensemble spread and errors in real time (Novak et al. 2008).

Though there is an increase in the availability of ensemble output from different centers, there is a lack of comprehensive studies on the model climatology of ensemble forecast uncertainty and error growth in medium-range forecasts for extratropical cyclones over eastern

U.S. and western Atlantic Ocean. Meanwhile, there are only limited studies on the evaluations of the performance of different models for predicting winter storms. These studies are only for a short, specific time period and mainly based on matching cyclone tracks, a process in which many ensemble members are not verified because a match could not be made. It is necessary to find an alternative way to evaluate cyclones and compare the results between different models as well as between the matching Lagrange method and other methods (e.g. Eulerian perspective). The role of large-scale flow regimes (e.g. RWPs) needs to be determined in improving the understanding of medium-range forecast of winter storms.

This dissertation will address the following science questions:

1) Can an ensemble sensitivity approach provide robust indications of the upstream uncertainties related with cyclone amplitude and position separately out to the medium range?

2) How to verify the goodness of ensemble sensitivity signals associated with the change of forecast metrics?

3) What is the model climatology (e.g., geographic preference) of ensemble sensitivity regarding to the cyclone intensity and position for medium range forecast in a multi-model ensemble from ECMWF, NCEP and CMC?

4) Can the fuzzy clustering method efficiently separate forecast scenarios associated with extratropical cyclones and thus condense the useful information from the large ensemble?

5) How can fuzzy clustering method be used as an ensemble verification tool for the prediction of winter storms?

6) What are the systematic errors and biases in forecasting winter storms in different models in

all significant storm cases?

7) What are the benefits of using multi-model ensembles over single model ensembles?

8) How do upper-level Rossby wave packets affect the development of forecast errors and uncertainties?

9) Under what flow regimes, the analyses are outside of the multi-model ensembles?

1.7 Outline of the dissertation

The motivational questions in section 1.6 will be explored in this dissertation. The outline of this dissertation is as follows. Chapter 2 will describe the data and methodologies used. It will also discuss how cyclone cases for the climatology of ensemble sensitivity study and evaluation of extratropical cyclone forecast in the multi-model ensemble from the three centers were selected. The application of ESA to medium-range forecast and its model climatology and verification using LOOCV will be presented in Chapter 3. The application of fuzzy clustering to separate ensemble forecast scenarios in multi-model ensembles will be introduced in Chapter 4. Chapter 4 will also present the evaluations of the different models in predicting East Coast winter storms using a scenario-based ensemble verification approach. The linkage between RWPs and forecast errors/uncertainties of East Coast winter storms will be discussed in Chapter 5. A concluding summary is given in Chapter 6.

This page is left to be blank intentionally

Chapter 2 Data and Methodologies

The focus of this work is the investigation of ensemble forecasts for extratropical cyclone events over the eastern U.S. and the surrounding oceans under the application of different ensemble tools. The ensemble data and techniques employed for this dissertation are described in this section. This work also involves cyclone cases for calculating model statistics and bias/errors; therefore how the cases were selected will be briefly presented.

2.1 Data

2.1.1 Deterministic and ensemble data: TIGGE

The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) project was established since 2005 to support a range of THORPEX research activities by providing operational ensemble forecast data to the international research community (Bougeault et al. 2010; Swinbank et al. 2016). One of the objectives of the project was to enhance collaboration on the development of EPSs between universities and operational centers by increasing the availability of ensemble data for research. Since 2008, ten operational weather forecasting centers, including EPSs from Australia, Brazil, Canada, China, ECMWF, Japan, Korean (KMA), Meteo France, U.K. (UKMet), and U.S. (NCEP), have been delivering near-real-time ensemble forecasting data to three TIGGE data archives located at the ECMWF, the National Center for Atmospheric Research (NCAR), and the Chinese Meteorological Administration (CMA). During a post processing at these archive centers, the individual ensemble forecasts are converted to a common grid/data format and are interpolated onto the same pressure levels. Meanwhile, the control forecast and deterministic forecast data from each of the 10 meteorological centers are also post-processed the same way and archived as part of the TIGGE data. After two days, these products are freely available for research and educational purpose, which form a valuable dataset to support research on predictability and dynamical processes and the development of ensemble-based forecast product (Swinbank et al. 2016). One particular benefit for research is that the individual ensemble forecasts can be arbitrarily combined to form a multimodel ensemble. This multimodel ensemble then provides the basis for various investigations, from single case studies to predictability investigations for the whole period. Standard meteorological parameters, such as the geopotential height, MSLP, temperature, etc., are available from all EPSs. Such a wealth of data from operational NWP systems, which has not previously been accessible to the research community, provides new opportunities to advance our knowledge of the predictability of specific weather systems like the high-impact East Coast winter storms.

This study will mainly analyze the ensemble forecasts in cool seasons (November to March in the following year) from 2007/08 to 2014/15 from three centers retrieved from the TIGGE archive, namely CMC, NCEP, and ECMWF. We chose these three because they are the ensembles used most by NWS forecasters. This multi-model ensemble facilitates exploration of uncertainties in NWP due to differences in initial uncertainties, model physics and parameterization methods among the different EPS. The three EPSs comprise a large set of ensemble forecast with 90 members in 12 h forecast steps and interpolated onto 1° latitude x 1° longitude grid. Note that the control and deterministic forecasts are not included in the multi-model ensemble. The details of updates and configurations of the three EPSs from 2007 to 2015 are listed in Table 2.1. Two major updates of GEFS model occurred in February 2010 and February 2012. CMC model had two major updates during August 2011 and November 2014. ECMWF model experienced three major updates in September 2009, January 2010, and November 2013, respectively.

Models / Configuration	GEFS	CMC	ECMWF
Model version	Nov 2007-Feb 2010: v.8 .0 Feb 2010-Feb 2012: v.9.0 Feb 2012-Mar 2015: v.10.0	Nov 2007-Aug 2011: CMC global ensemble model (GEM) Aug 2011-Feb 2013 GEPS2.0.2 Feb 2013-Dec 2013 GEPS 3.0.0 Dec 2013-Nov 2014 GEPS3.1.0 Nov 2014-Mar 2015: GEPS4.0.0	Nov2007-Mar 2008 Cycle 32r3 Mar 2008-Sep 2008 Cycle 33r1 Sep 2008-Sep 2009 Cycle 33r2 Sep 2009-Jan 2010 Cycle 35r3 Jan 2010-Mar 2015 Cycle 36r1
Initial uncertainty	Breeding Vector method with Ensemble Transform and Rescaling (BV-ETR)	Dec 2007-Aug 2011: Ensemble Kalman Filter Assimilation (EnKF); 96 members Aug 2011-Nov 2014: EnKF, 192 mem Nov 2014-Mar 2015: EnKF: 256 members	Nov 2007-Nov 2013: Singular Vectors (SVs) Nov 2013-Mar 2015" SVs+Perturbations from the Ensembles of Data Assimilation (EDA)
Model uncertainty	Nov 2007–Feb 2010: None Mar 2010–Mar 2015: Stochastic total tendency perturbation (STTP)	STTP, SKEB backscattering schemes	Nov 2007-Sep 2009: no stochastic physics Sep 2009-Mar 2015: Revisited Stochastic Physics STTP SKEB backscattering schemes
Data Assimilation	GSI 3D-Var	Nov 2007-Nov 2014: 4D-Var Nov 2014-Mar 2015: Coupling with 4D-En VAR	4D-Var
Resolution	Nov 2007-Feb 2010: T126 L28 Feb 2010-Feb 2012: T190 L28 Feb 2012-Mar 2015: T254 L42 (0-8) T190 L42 (8-16)	Nov 2007-Aug 2011: 100km(horizontal)/L28 (vertical, to 10hPa)/45 min (time step) Aug 2011-Feb 2013 66 km/L40 (to 2hPa)/30 min Feb 2013-Dec 2013 66km/L40(to 2hPa)/20min Dec 2013-Nov 2014:	Nov 2007-Jan 2010 T399(0-10day), T255(10-15/32day) Jan 2010-Nov 2013 T639L62 (0-10day) T319L62 (10-15day) Nov 2013-Mar 2015 T639L91 (0-10day), T319L91(10-15/32day)

		66 km/L40 (to 1.78hPa)/20min Nov 2014-Mar 2015 50km/L48(to 1.45hPa)/ 15 min	
Forecast length (days)	0-16	0-16	0-15, 0-32
Ensemble members	20+1	20+1	50+1
Daily frequency (UTC)	0000, 0600, 1200, 1800	0000, 1200	0000, 1200
Output Time Steps	6-hourly: 000h to 384h	6-hourly: 000h to 384h	6 hourly: 000h to 384h

Table 2.1: The Evolution of NCEP GEFS, CMC, and ECMWF model configurations during November 2007 and March 2015. Note the date in red means major updates.

2.1.2 Variables, forecasting, and analysis data used for each chapter

Here, I will summarize the data used for the main result of each chapter in this dissertation.

For the ensemble sensitivity calculations in Chapter 3, the following forecasting data and analysis data are used:

i. Fifty member ensemble forecasts of MSLP and Z500 from the ECMWF for the 2010 December case study, three runs are used with the initialization time at 0000UTC 22 December 2010, 0000 UTC 24 December 2010, and 0000 UTC 25 December 2010, respectively.

ii. A multi-model ensemble is used to calculate the climatology for ensemble sensitivity associated with cyclone intensity and position forecasts as well as to verify sensitivity signals using LOOCV. The details of the multi-model ensembles with 90 members from CMC, NCEP, and ECMWF are as follows:

- Forecast length used: 3 days and 6 days

- Forecast variables: MSLP, Z500
- Initialization time: 0000 or 1200 UTC
- Time period: 102 unique cyclone cases during 2007/08–2014/15 cool seasons
- Resolution: 1° longitude x 1° latitude

Note that the MSLP and Z500 are used to represent the surface and upper-level features, respectively.

iii. NCEP operational analysis data for MSLP and Z500

For the fuzzy clustering calculations in Chapter 4, the following forecasting data and analysis data are used:

i. Ninety member multi-model ensemble from the CMC, NCEP and ECMWF centers. MSLP , Z500, total precipitation for the 2015 January blizzard case, two runs with the initialization time at 1200UTC 21 January 2015 and 1200 UTC 24 January 2015.

ii. The 90-member multi-model ensembles from CMC, NCEP, and ECMWF are used for ensemble verification. The details of the multi-model ensembles are as follows:

- Forecast length used: 1–9-day forecast
- Forecast variables: MSLP, Z500
- Initialization time: 0000 or 1200 UTC
- Time period: cyclone cases (the case number will be given in Table 4.3 of Chapter 4) in the analysis during 2007/08–2014/15 cool seasons
- Resolution: 1° longitude x 1° latitude

iii. NCEP operational analysis data for MSLP and Z500

iv. Climate Prediction Center (CPC) unified gauge-based daily pre precipitation data over CONUS during 26−28 January 2015 on a 0.25° longitude x 0.25° latitude grid

To study the predictability of winter large-error events over eastern U.S. and western Atlantic Ocean associated with RWPs, the following data at 2.5°longitude x2.5°latitude grid are used in Chapter 5:

i. GFS deterministic 7-day forecast of Z300 during 2007/08 and 2013/14 cool seasons.

ii. NCEP operational analysis data for U wind and V wind at 300 hPa level to calculate RWP amplitude

iii. GEFS 20-member ensemble data for MSLP, U300, V300, and Z300 to calculate ensemble spread and ensemble wave packet amplitude

By studying the large set of TIGGE ensemble data, ensemble sensitivity analysis can make inference about the dynamical linkage between forecast metrics and antecedent forecasts at earlier forecast times or initial conditions. Different forecast scenarios can be evaluated using fuzzy clustering tool, which will shed light on understanding the development and propagation of forecast uncertainties. Meanwhile, it will provide a means to investigate the physical and dynamical processes responsible for the variability among EPS members and to connect these mechanisms to the predictability of the investigated cyclone events. In addition, large-error cases selected from the deterministic forecast data set will be investigated, to determine detailed major dynamical features affecting large forecast errors during the dropout cases. These techniques will be discussed further below.

2.2 Winter storm case selection

As mentioned in previous subsection, cyclone cases will be selected to study their predictability in ensemble forecasts. This section will briefly describe the criteria to select those cases.

2.2.1 Analyzed cyclone cases

To calculate ensemble sensitivity and evaluate the performance of different operational models in forecasting East Coast cyclones, we first tracked cyclones across two primary regions (Figure 2.1): region 1 and region 2. Region 1 is a bigger region including part of the central U.S.as well as the East Coast of the U.S.; region 2 is a relatively smaller region including the East Coast and part of the western Atlantic. The cyclone cases include verified significant cyclones with the minimum central pressure less than 1005 hPa crossing these two regions determined by the tracking scheme developed by Hodges (Hodges 1994, 1995, 1999). A minimum pressure of 1005 hPa is chosen because a climatological study of East Coast winter study by (Hirsch 2001) showed that over two third of the cyclones had minimum pressure less than 1005 hPa using NCEP/NCAR reanalysis data during 1951-1997. A choice of 1005 hPa as a threshold for minimum pressure will include most of the significant cyclones.

A total of 130 (120) cyclone tracks were found for region 1 (2). Among them, 63(82) cases are classified by NOAA WPC as significant storms, which "Impact large population areas, or major transportation systems, or otherwise make a significant impact upon the nation's or a region's commerce" (Reference: http://www.wpc.ncep.noaa.gov/winter_storm_summaries/winter_storm_summaries.shtml).

To compute the ensemble sensitivity as well as the LOOCV, we only use one verification time to avoid redundant count of cyclones. For each cyclone track, the verification time is selected using the time point with the maximum past 12 h MSLP change within each region. Therefore, 130 (120) cyclone cases for region 1 (2) were pre-selected. However, there are missing ensemble data in TIGGE archive. We finally used the 3-day and 6-day ensemble forecast from 115 (102) cyclone cases for region 1 (2) to compute the ensemble sensitivity climatology and LOOCV.

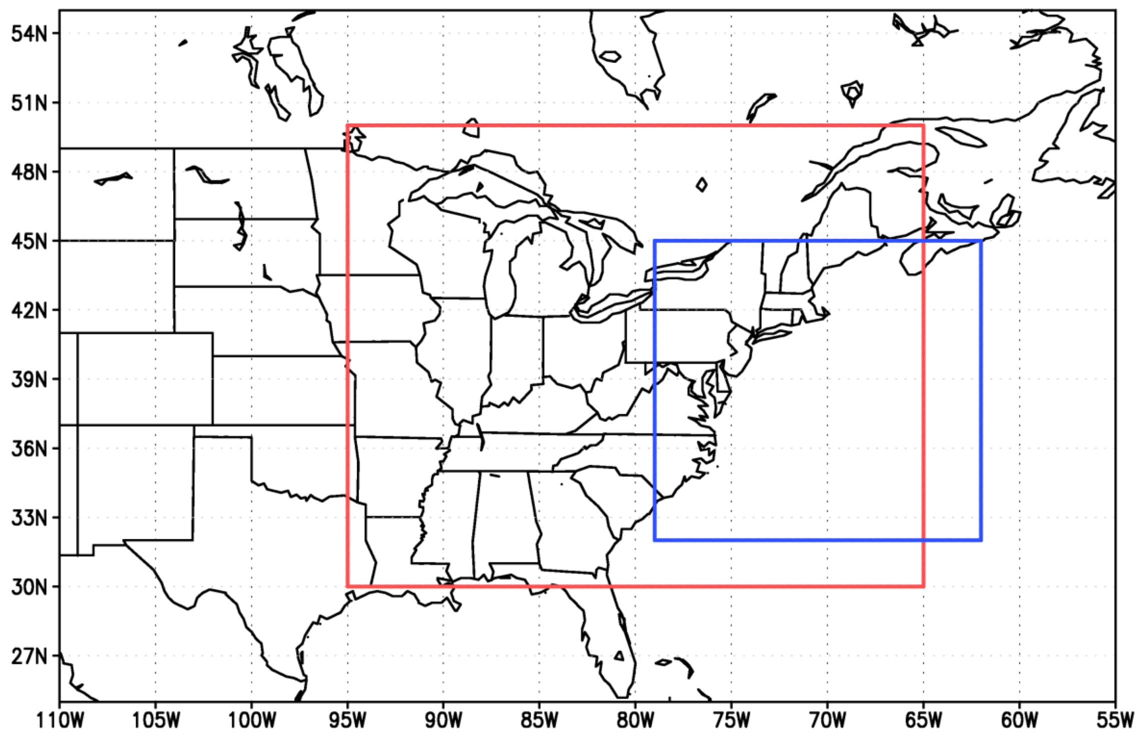


Figure 2.1: Region 1 (red box) and region 2 (blue box). The red box includes an area from 95°W to 65°W, and 30°N to 50°N; and the blue box includes an area from 79°W to 68°W, and 32°N to 45°N.

To perform the model evaluations for cyclone forecasts, the same 130 (120) cyclone tracks were selected. To increase sample size for model evaluations, one or two verification times were included as long as they are within the verified region. I evaluated the 1–9-day forecasts for the observed cyclones. Again, due to the missing members for some cases in TIGGE archive, I have chosen 158 to 185 cases (Table 4.3) for the scenario-based model evaluation calculations. I have also tested similar calculations using fewer cases (102 to 135) for 3- and 6-day forecasts; the results are very consistent.

2.2.2 Predicted cyclone cases

To evaluate model bias in forecasting cyclones, we also selected forecast cyclone cases for different forecast times. The NCEP deterministic forecasts from 2007/08 to 2014/15 cool seasons are chosen to track forecast cyclones. If the forecast cyclone at a forecast time of day N (N=1,...,9) crossed the verification region with a central pressure less than 1005 hPa, this cyclone will be a candidate for forecast cyclone case on day N forecast. If the initialization times of two candidate cases for the same verification time are very close (no more than one day), only one of them will be selected based on the criterion of its cyclone position closer to the center of verification region. The number of cases will be summarized in Table 4.4. Note that the evaluation of predicted cyclone cases is a complementary calculation for observed cyclone cases.

2.2.3 Dropout cases selection

In Chapter 5, large error cases or dropout cases will be investigated to explore the potential impacts of RWPs on the predictability of winter weather. Dropout cases are selected based on the NCEP deterministic forecast of geopotential height at 300 hPa (Z300). Root-mean-square errors of 7-day forecast for Z300 over a verification region (Figure 2.2) were calculated. The cases with the RMSE greater than 1.5 standard deviations are chosen as dropout cases. When two cases are very close (< 2day), only one case will be kept to avoid redundant cases. As a comparison, the small RMSE cases (< Mean−1.1 standard deviations) are also selected to represent good forecasts. Eighty-three dropouts and 85 small-error cases were selected.

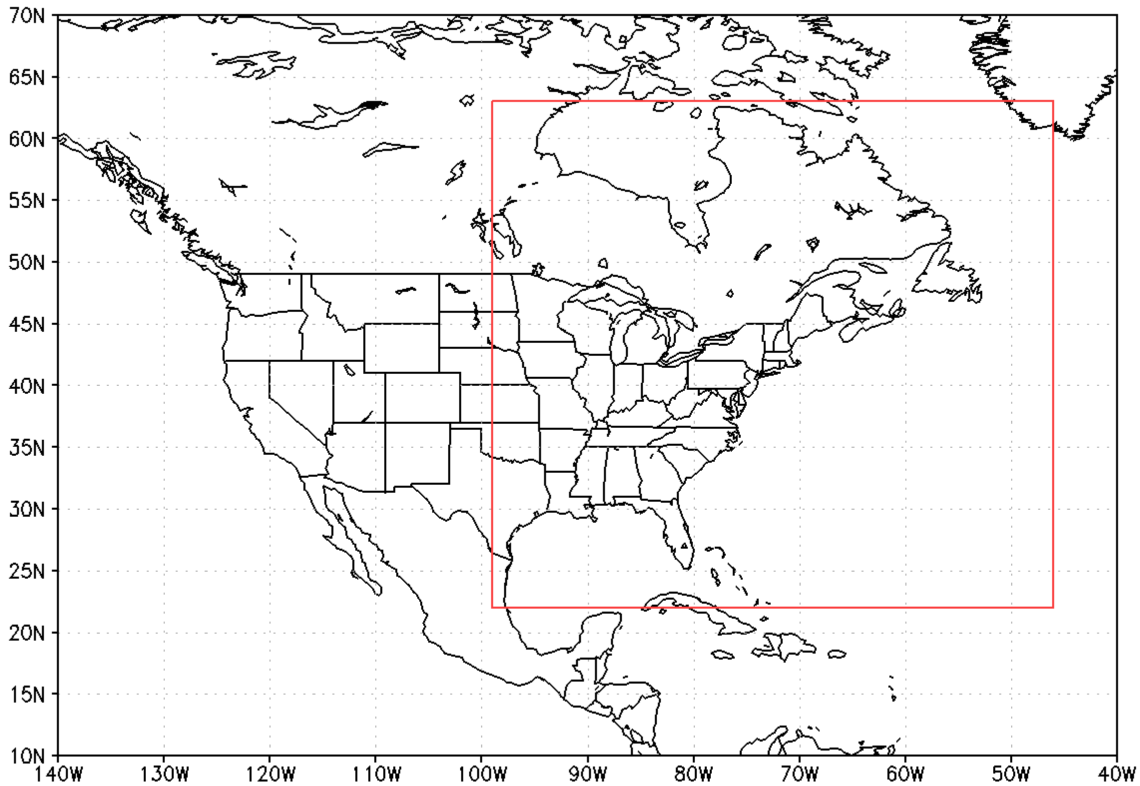


Figure 2.2: Region (red box) for calculating RMSE and selecting large error cases. The red box includes an area from 99°W to 47°W, and 22°N to 63°N.

2.3 Empirical Orthogonal Function analysis

The Empirical Orthogonal Function (EOF) analysis is a statistical method to condense the information of a large dataset without losing valuable information expressed by the original dataset. It was introduced to atmospheric sciences by Lorenz (1956) in the 1950s. Since then, it has become a popular method in atmospheric sciences to examine the variability of meteorological data due to its simplicity and analytic derivation (Hannachi et al. 2004; Hannachi et al. 2007; Wilks 2011). In terms of ensemble forecast, there exist regions of low forecast uncertainty where most members show nearly identical fields as well as regions of high forecast uncertainty where large disagreements in individual members appear. The EOF method can summarize the information or the dominant differences provided by the individual members while removing the redundant information among them. Basically, the EOF analysis defines a new set of variables that permits the expression of dominant parts of the variance contained in the original data set at one investigation time, but has a lower dimensionality. This new set of variables are determined by identifying a new set of vectors $\mathbf{e_i}$ onto which the original data matrix $\mathbf{X}$ explains the maximum variability, thus they can best describe the synoptic patterns in the ensemble members.

To be more specific, here the EOF analysis is calculated to determine the dominant patterns of variations in ensemble MSLP forecasts over a verification region at an investigation time. Following Hannachi et al. (2004, 2007) and Keller (2012), the formulation of EOF analysis will be summarized as below.

Assume the original MSLP data over a verification box at a verification time is a two-dimensional $M \times P$ matrix $\mathbf{X}$, where M and P represents the number of ensemble members and grid points, respectively. The value of MSLP for an ensemble member i at a grid point j is noted x_{ij} for $i=1,\ldots, M$ and $j=1,\ldots, P$.

If we denote by $\bar{x}_{\cdot j}$ the ensemble mean of the MSLP at the jth grid point, i.e.

$$\bar{x}_{\cdot j} = \frac{1}{M}\sum_{k=1}^{M} x_{kj} \tag{2.1}$$

Then the ensemble mean of all grids are defined by

$$\overline{\mathbf{x}} = (\bar{x}_{\cdot 1,\ldots,}\bar{x}_{\cdot p}) \tag{2.2}$$

The anomaly or the deviation of an ensemble member i from the ensemble mean at the jth grid point is defined by:

$$x'_{ij} = x_{ij} - \bar{x}_{\cdot j} \tag{2.3}$$

or in matrix form:

$$\mathbf{X}' = \mathbf{X} - \mathbf{I}\bar{\mathbf{x}}, \tag{2.4}$$

here **I** is the column vector containing M ones.

The anomaly matrix **X′** can be expressed as:

$$\mathbf{X}' = \begin{pmatrix} x'_{11} \; x'_{12} & \cdots & x'_{1P} \\ x'_{21} \; x'_{22} & \cdots & x'_{2P} \\ \vdots & & \vdots \\ x'_{M1} & \cdots & x'_{MP} \end{pmatrix}. \tag{2.5}$$

The covariance matrix is then defined by:

$$\boldsymbol{\Sigma} = \frac{1}{M-1}\mathbf{X}'^{T}\mathbf{X}', \tag{2.6}$$

which contains the covariance between any pair of grid points. The aim of an EOF analysis is to find the linear combination of all the variables that explains the maximum variance. That is to find a direction $\mathbf{e}=(e_1,e_2,\ldots,e_P)^T$ such that **X′e** has maximum variability. The variance of **X′e** is

$$\mathrm{var}\,(\mathbf{X}'\mathbf{e}) = \frac{1}{M-1}\left\|\mathbf{X}'\mathbf{e}\right\|^2 = \frac{1}{M-1}(\mathbf{X}'\mathbf{e})^T(\mathbf{X}'\mathbf{e}) = \mathbf{e}^T\boldsymbol{\Sigma}\mathbf{e}. \tag{2.7}$$

Under the constraint that **e** has unit length, the maximization-variance problem yields:

$$\max_{\mathbf{e}}(\mathbf{e}^T\boldsymbol{\Sigma}\mathbf{e}), \mathrm{s.t.}\,\mathbf{e}^T\mathbf{e} = 1 \tag{2.8}$$

The solution to eq. (2.8) is to solve the eigenvalue problem:

$$\boldsymbol{\Sigma}\mathbf{e} = \lambda\mathbf{e}, \tag{2.9}$$

where **e** and λ are referred to as the eigenvector and its corresponding eigenvalue of the covariance matrix **Σ**. The kth eigenvalue λ_k for the kth eigenvector gives a measure of the explained variance by e_k with k=1,…,P. The explained variance percentage can be defined as:

$$\frac{100\lambda_k}{\sum_{k=1}^{P}\lambda_k}\% \tag{2.10}$$

The projection of the anomaly field **X'** onto the kth EOF $\mathbf{e}_k$, i.e., $\boldsymbol{\alpha}_k = \mathbf{X}'\mathbf{e}_k$ is the kth principal components (PCs) as expressed by:

$$\alpha_{ik} = \sum_{j=1}^{P} x'_{ij} e_{kj} \tag{2.11}$$

where $i=1,\ldots,M$, $k=1,\ldots,P$.

The normalized PCs of ensemble members for the leading two EOF patterns will be used as forecast metrics to calculate ensemble sensitivity in chapter 3 and the basis to perform fuzzy clustering analysis in chapter 4. They represent the projections of each of the ensemble members onto the two dominant EOF patterns. Since the PCs are normalized to have unit variance, the EOF patterns carry the magnitude and dimensions of MSLP anomalies (Chang et al. 2013). As will be shown later, we find that typical EOF patterns represent two types of uncertainties in the forecast: amplitude and position uncertainties, consistent with the results of Gombos et al. (2012). We will also show in Chapter 4 that for East Coast cyclone cases, in the majority of cases most of the forecast error projects onto the two leading EOFs, justifying our focus on the two leading EOFs.

2.4 Sensitivity analysis

To relate the forecast uncertainty of a forecast at investigation time with the initial conditions or earlier forecasts, ensemble sensitivity analysis will be applied to the multi-model ensemble forecast of MSLP. This section briefly introduces the ensemble sensitivity method. Since ETKF method is also an efficient sensitivity analysis approach and will be used to compare the results from ensemble sensitivity in a case study, a brief introduction of the ETKF method will also be given here.

<u>2.4.1 Ensemble sensitivity method</u>

Ensemble sensitivity analysis can quantify the relationship between either initial conditions or forecasts at earlier forecast times and the forecasts at verification time, thus it can improve our understanding of the development and propagation of forecast uncertainties.

The calculation of ensemble sensitivity follows that of Torn and Hakim (2009). Given an ensemble with M members, the sensitivity of any forecast metric J to a state variable X_i can be expressed as

$$"Sensitivity" = \frac{Cov(\mathbf{J}, \mathbf{X_i})}{\sqrt{Var(\mathbf{X_i})}}; \tag{2.12}$$

here $\mathbf{J}$ and $\mathbf{X_i}$ are the $1 \times M$ ensemble estimates of the forecast metric J and the *i*th state variable X_i, respectively. Cov denotes the covariance between $\mathbf{J}$ and $\mathbf{X_i}$ across the ensemble, and Var is the variance. The "Sensitivity" defined by equation (2.12) carries the dimension of the forecast metric and describes the change in the forecast metric corresponding to one standard deviation change in the uncertainties of the selected state variable.

Figure 2.3 shows one illustrative example to interpret the ensemble sensitivity. The regression slope, which is the covariance between two variables (J and X_i) divided by the variance of X_i, is 11.3 Pa m^{-1}, which means a change of 1 m Z500 corresponds to a change of 11.3 Pa minimum pressure. Since the standard deviation of Z500 (or X_i in eq. 2.12) is 107.65 m, the ensemble sensitivity, which equals to the regression times the standard deviation based on eq. 2.12, is 107.65 m x 11.3 Pa m^{-1}, or 12.18 hPa. One standard deviation change of Z500 at the grid point of interest is associated with 12.18 hPa change of minimum pressure. Thus, the sensitivity using eq. 2.12 keeps the unit of forecast metric.

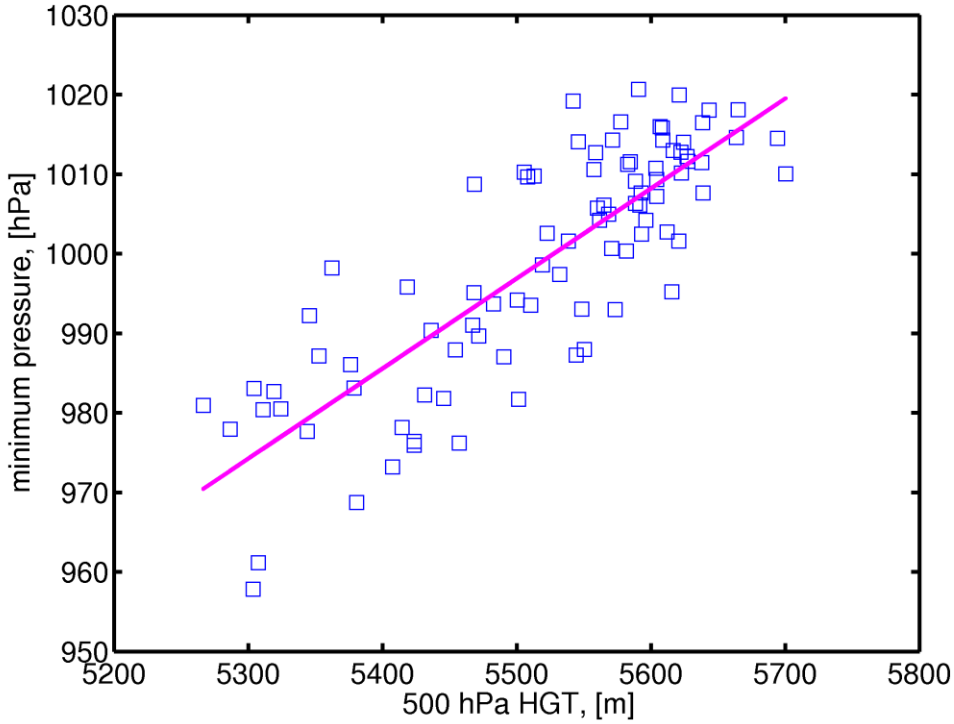


Figure 2.3: One illustrative example to interpret the ensemble sensitivity. The abscissas and ordinates are Z500 at one grid point within the verification region and the minimum pressure forecast from 90-member ensemble within the same verification region.

A forecast metric using EOF analysis is defined on a box over selected model grids to calculate the dominant EOF patterns in the ensemble MSLP forecast at a particular forecast hour. The EOF patterns basically represent the dominant forecast uncertainties within the box. This approach was used rather than identifying forecast metric for cyclone parameters directly, since the cyclones in each ensemble member with similar parameters (minimum pressure, longitude, or latitude) may not have a similar structure (Gombos et al. 2012; Chang et al. 2013). The corresponding PCs of each of the leading EOF patterns (which are the projection of the EOF pattern onto the difference between each ensemble member and the ensemble mean) can be employed as forecast metric (J). Therefore, using the PCs of an EOF as the forecast metric, ensemble sensitivity analysis can objectively evaluate the relationship between that uncertainty pattern and the state variables at initial time or other forecast times.

In our application, the PCs are dimensionless and normalized to unit variance, while the EOF patterns carry the dimension of the forecast variable, and the amplitude of the patterns reflects the square root of the amount of variance explained by each EOF. Thus, the forecast metric employed here has unit variance and is non-dimensional. To make this point more explicit, Chang et al. (2013) redefined "sensitivity" by dividing the RHS of equation (2.12) by the standard deviation of **J**, as follows:

$$"Sensitivity" = \frac{Cov(\mathbf{J}, \mathbf{X_i})}{\sqrt{Var(\mathbf{X_i})}\sqrt{Var(\mathbf{J})}}; \tag{2.13}$$

thus "sensitivity" in eq. 2.13 is dimensionless and is equal to the correlation between the ensemble estimates of J and X_i. Basically, "sensitivity" corresponds to a weighted average of the state vectors of the ensemble members; with the weight (J) reflecting how much the difference between each ensemble member and the ensemble mean resembles the EOF pattern.

With the knowledge of sensitivity signals within different ensemble forecast cycles, the forecasters can have better understanding of what kinds of antecedent atmospheric features are associated with a particular synoptic system at the verification time. The development and propagation of forecast uncertainties can be intuitively illustrated by examining the time evolution of sensitivity signals.

2.4.2 LOOCV: the verification of ensemble sensitivity signals

The accuracy of sensitivity regions identified by ensemble sensitivity can be evaluated by the leave-one-out cross validation method (Wilks 2011, Gombos et al. 2012). Cross validation is often employed to validate different ensemble forecasts by reserving an independent verification data set. Cross validation simulates prediction for future, unknown data by repeating the entire fitting procedure on data subsets, and then examining the predictions made for the data portions left out of each subset. A frequently used procedure is known as leave-one-out cross validation (LOOCV), in which the fitting procedure is repeated *n* times, each time with a sample of size *n*-1, because one of the predictand observations and its corresponding predictor set are left out in each replication of the fitting process. The result is *n* slightly different prediction equations.

The cross-validation estimate of the prediction mean squared error (MSE) is computed by forecasting each omitted observation using the equation developed from the remaining *N*-1 data values, computing the squared difference between the prediction and predictand for each of these equations, and averaging the *N* squared differences. Thus, LOOCV uses all *N* observations of the predictand to estimate the prediction MSE in a way that allows each observation to be treated one at a time, as independent data.

When applying the LOOCV to this dissertation, it will be performed by 1) removing a single ensemble member (member of interest, $m_{p,\ 1<p<,n_{ens}}$) from the ensemble (n_{ens}); 2) calculating the ensemble sensitivity for the leading two EOF PCs at verification time t_v using the remaining n_{ens}-1 members; 3) calculating the leading n_x EOF PCs of the state vector $\boldsymbol{X}$ over the selected sensitivity region at another time (t_i, $i \leq v$), which is the predictor ${}_{n_x}\tilde{X}_{n_{ens}-1}$; and choosing one of the leading two EOF PCs (${}_1J_{n_{ens}-1}$ and ${}_2J_{n_{ens}-1}$) at verification time as the predictand; 4) computing ensemble regression based on the predictor (${}_{n_x}\tilde{X}_{n_{ens}-1}$) and predictand (${}_1J^1_{n_{ens}-1}$ or ${}_1J^2_{n_{ens}-1}$); 5) projecting the forecast anomaly of member of interest at t_i to the leading n_x EOF patterns to the same sensitivity region and recoding the corresponding coefficient (${}_{n_x}\tilde{X}_1'$); and projecting the forecast anomaly for the same member at t_v to the leading two EOF patterns at t_v and recording the corresponding coefficient (${}_1J^1_1$ or ${}_1J^2_1$); 6) applying the ensemble regression relation to the predictor ${}_{n_x}\tilde{X}_1'$ and calculate the corresponding predictand (${}_1\bar{J}_1^1$ or ${}_1\bar{J}_1^2$) for the member of interest, and then 7) comparing the ensemble regression-estimated predictand (${}_1\bar{J}_1^1$ or ${}_1\bar{J}_1^2$) to the actual value (${}_1J_1^1$ or ${}_1J_1^2$) and computing the ratio r_p of the estimated value and the actual value of J. After repeating the above 7 steps for each ensemble member, the median of the r_p together with the correlation ρ between the predicted and the actual value for J will be measures of the robustness of the ensemble sensitivity signals. This method has been demonstrated to be effective in quantifying how well the ensemble sensitivity signals can predict the change in the forecast metric J (Gombos et al. 2012; Gombos and Hoffman. 2013). Chapter 3 will give more details of its application in verifying ensemble sensitivity signals.

2.4.3 An introduction to the ETKF method

The ETKF technique (Bishop et al 2001; Majumdar et al. 2002) was used to identify the potential upstream sensitive areas, where the assimilations of targeted observations are expected to maximize the improvement of the forecast over a verification area. In other words, given an ensemble run with a specified initialization time t_i, the ETKF uses the ensemble perturbations relative to the ensemble mean to predict the change of forecast error covariance within an investigation region at verification time t_v due to the assimilation of adaptive observations at time t_a ($t_i<t_a<t_v$). The impact of the targeted observations is expressed by using a difference total energy metric, or the variance of the "signal". It is predicted and mapped as a composite "summary map" that depicts sensitive areas for sampling (Majumdar et al. 2002, Majumdar et al. 2010).

The process for calculating ETKF is as follows:

Step 1: define the error covariance for perturbations at initial time:

$$\mathbf{P}^i(t \mid H^i) = \mathbf{Z}^i(t \mid H^i)\mathbf{Z}^{i\mathbf{T}}(t \mid H^i), \tag{2.14}$$

wherein $\mathbf{Z}^{i}(t|H^{i})$ is the forecast perturbation matrix, and H^{i} is the matrix operator.

Step 2: calculate error covariance matrix for routine observation network at analysis time by

$$\begin{aligned}\mathbf{P}^r(t_a \mid H^r) &= \mathbf{P}^i(t_a \mid H^i) - \mathbf{P}^i(t_a \mid H^i)H^{r\mathbf{T}}[H^r\mathbf{P}^i(t_a \mid H^i)H^{r\mathbf{T}} + \mathbf{R}^r]^{-1}H^r\mathbf{P}^i(t_a \mid H^i)\\ &= \mathbf{Z}^i(t_a \mid H^i)\mathbf{T}^r\mathbf{T}^{r\mathbf{T}}\mathbf{Z}^{i\mathbf{T}}(t_a \mid H^i)\end{aligned} \tag{2.15}$$

wherein H^{r} and $\mathbf{R}^{r}$ are the “routine” operator and error covariance matrix, and $\mathbf{Z}^{i}(t_a|H^{i})$ and $\mathbf{T}^{r}$ are the raw ensemble perturbation at analysis time and linear transformation matrix.

Step 3: using serial assimilation theory, calculate the error covariance update due to the qth deployment of targeted observations at analysis time:

$$\mathbf{P}^{r+q}(t_a \mid H^{r+q}) = \mathbf{P}^r(t_a \mid H^r) - \mathbf{P}^r(t_a \mid H^r)H^{q\mathbf{T}}[H^q\mathbf{P}^r(t_a \mid H^r)H^{q\mathbf{T}} + \mathbf{R}^q]^{-1}H^q\mathbf{P}^r(t_a \mid H^r) \tag{2.16}$$

wherein H^{q} and $\mathbf{R}^{q}$ are the qth adaptive operator and error covariance matrix.

Step 4: the forecast error covariance at time t_v can be expressed as:

$$\begin{aligned}&\mathbf{P}^{r+q}(t_v \mid H^{r+q}) = \mathbf{P}^r(t_v \mid H^r) - \mathbf{S}^q(t_v \mid H^q), \text{where}\\ &\mathbf{S}^q(t_v \mid H^q) = \mathbf{Z}^i(t_v \mid H^r)\mathbf{Z}^{i\mathbf{T}}(t_a \mid H^r)H^{q\mathbf{T}}[H^q\mathbf{P}^r(t_a \mid H^r)H^{q\mathbf{T}} + R^q]^{-1}H^q\mathbf{Z}^i(t_a \mid H^r)\mathbf{Z}^{iT}(t_v \mid H^r)\end{aligned} \tag{2.17}$$

$\mathbf{S}^{q}(t_v|H^{q})$ is the “signal covariance”. Following Majumdar et al. (2010), the diagonal of $\mathbf{S}^{q}(t_v|H^{q})$ localized within the verification region can be plotted as a function of the qth targeted observation on the ETKF guidance. The “pseudo sounding” of (u,v,T) at 850, 500, and 200 pressure levels are used as the adaptive observations, which are sampled at the model grid point at 1° resolution. Therefore, the summary map of the normalized signal represents the reduction in forecast variance within the fixed verification region. The regions with largest values represent the sensitive areas contributing most to the error variance reduction. For more details of this method, the reader is referred to papers by Majumdar et al. (2001, 2002), Langland (2005), and Majumdar et al. (2010, 2011). The ETKF code used for the case study in Chapter 3 was provided by Dr. Sharan Majumdar from University of Miami.

2.5 Fuzzy clustering method

To quantify the variability within an ensemble of forecasts and recognize different scenarios among different operational models, a Fuzzy cluster analysis will be applied to group ensemble members with similar forecast scenarios. The clusters of forecast scenarios will be used to understand different forecast scenarios among models and to evaluate models. Harr et al. (2008) suggested that once the EOF analysis on the ensemble within the region of interest was completed, the first and second principal components (PCs) for the collection of all ensemble members can be used as input to a Fuzzy clustering routine. To start the iterative procedure, a previously specified number of clusters or initial guess was randomly placed in the EOF PC1-PC2 phase space. Each ensemble member denoted by the pair of PCs is then assigned to the nearest group center. New centers are computed by minimizing an objective function that represents the distance from each point to each new cluster center. Each point is examined again relative to the updated cluster centers. If no points can be reassigned because they lie closer to another center, the iterations stop. Each member is assigned a weight value that identifies their relative strength of membership to their cluster (Harr et al. 2008). For a point k, the weight associated with the *i*th cluster is defined as

$$w_{i,k} = \frac{1}{\sum_{j=1}^{C} (\frac{d_{i,k}}{d_{j,k}})^{2/(q-1)}}, \tag{2.18}$$

such that $d_{i,k}$ is the distance between point k and the centroid of cluster i, and $d_{j,k}$ is the distance between point k and the other cluster centers j. The fuzziness coefficient q determines the level of cluster fuzziness. A large q results in smaller memberships $w_{i,k}$ and hence, fuzzier clusters. In the limit q=1, the memberships converge to 0 or 1, which implies a crisp partitioning. In the absence of experimentation or domain knowledge, q is commonly set to 2. Harr et al. (2008) set q to be 1.5. A total of C clusters can be calculated using equation (2.18) and the above procedure.

In this dissertation, fuzzy clustering will be applied to the PCs corresponding to the leading two EOF patterns. Conceptually, there should be an optimal number of population clusters associated with different synoptic-scale patterns. However, as pointed out by Harr et al. (2008), it is often very difficult to determine this number objectively. Keller et al. (2011) tested a range of the cluster number C from 2 to 8 and found that 6 was the suitable cluster number for their study. We have also tested 2–8-cluster solutions for both 3- and 6-day forecasts of the historical cyclone cases and found that 5 clusters tend to be optimal in most cases. One way to determine the optimal number of clusters is to find the stable solution with the highest adjusted Rand index (Yeung and Ruzzo 2001) based on 100 clustering results using random seeding points. When two or more cluster results are stable, the largest number of clusters is chosen. One reason is that we would like the clustering procedure to produce a group that can represent the ensemble mean, which is often hypothesized to be a good estimation of the truth (Du and Zhou 2011). Our experience suggests that a group clustered around the ensemble mean is more likely for larger number of clusters. Note that the ensemble mean (the origin on a PC1-PC2 coordinate) is used in clustering.

When examining past cases, after we have grouped forecast ensemble members based on their EOF PCs at VT, the difference between the analysis field and the ensemble mean at VT can be projected onto the leading EOF patterns and hence occupies one point on the EOF PC1-PC2 space just like one extra ensemble member, and this can be used to verify the scenarios in the ensemble forecasts (see Chapter 4).

The clusters of forecast scenarios, together with the ensemble sensitivity signals, will provide a more complete picture of the variability of forecast scenarios among ensemble members. We will also apply them to assess how the different NWP models perform in capturing the real scenario.

2.6 Rossby wave packet: detection and tracking

The relation between RWPs and large error cases will be investigated in chapter 5. The first step is to compute RWP amplitude, which is a representation of wave packet intensity. To investigate more details regarding to coherent RWP, the tracking of RWPs is necessary. So, an objective way to track RWPs will also be briefly introduced here.

2.6.1 Rossby wave packet amplitude

The calculation of Rossby wave packet amplitude (RWPA) follows Zimin et al. (2006), which involves the horizontal wind at 300 hPa on 2.5°x2.5° horizontal resolution from NCEP/NAR reanalysis. Zimin et al. (2006) demonstrate that when using a streamline, instead of a latitude circle line, the shape of the RWPA envelope is smoother and more contiguous in a non-zonal flow regime.

Given an atmospheric variable $\phi(x,y)$, and the basic flow with zonal components $u(x,y)$ and meridional components $v(x,y)$, the following algorithm summarizes how to calculate RWPA based on Zimin et al. (2003, 2006).

Step 1: At each grid point (x_0,y_0), a piecewise-linear approximation of a stream line is defined by (u,v) in the neighborhood of (x_0,y_0). If x, y are measured in units of longitude and latitude, respectively, the point (x_1,y_1) that lies a distance δ in the direction of the streamline from (x_0,y_0) is

$$
\begin{aligned}
x_1 &= x_0 + \frac{\delta}{\cos y_0}\frac{u(x_0,y_0)}{\sqrt{v(x_0,y_0)^2+u(x_0,y_0)^2}} \quad \text{and} \\
y_1 &= y_0 + \delta\frac{v(x_0,y_0)}{\sqrt{v(x_0,y_0)^2+u(x_0,y_0)^2}}.
\end{aligned}
\tag{2.19}
$$

Here δ has the units of latitude while cos y_0 compensates for the decreasing zonal distance between grid points near poles. Other points, (x_2, y_2), (x_3, y_3), … , (x_N, y_N), can be determined iteratively using the same formula with shifted indices. On the opposite direction, points (x_{-1}, y_{-1}),(x_{-2}, y_{-2}), …, (x_{-N}, y_{-N}) can be calculated similarly, working backward:

$$
\begin{aligned}
x_{-j-1} &= x_{-j} - \frac{\delta}{\cos y_{-j}}\frac{u(x_{-j},y_{-j})}{\sqrt{v(x_{-j},y_{-j})^2+u(x_{-j},y_{-j})^2}} \quad \text{and} \\
y_{-j-1} &= y_{-j} - \delta\frac{v(x_{-j},y_{-j})}{\sqrt{v(x_{-j},y_{-j})^2+u(x_{-j},y_{-j})^2}}.
\end{aligned}
\tag{2.20}
$$

Where, N is chose to keep $N\delta$ roughly the length of a latitude circle.

Step 2: With known coordinates (x_j,y_j), $j=-N, \ldots, +N$ for points, the atmospheric variable of interest is then interpolated onto each such point (x_j,y_j). The $\phi(x_j,y_j)$ is then localized using a Gaussian filter function centered at $j=0$:

$$\bar{\phi}(x_j, y_j) = \phi(x_j, y_j)\exp(-\alpha^2 \frac{j^2}{N^2}). \tag{2.21}$$

The 1/α is chosen to be roughly the length of the wave packet, as a fraction of the length of the latitude circle.

Step 3: Replace ϕ (x_j,y_j) with $\bar{\phi}$ (x_j,y_j) and define $\phi(s)$ given $s=j\delta$ for $|s|\leq\delta N$.

Step 4: The Fourier transform of the variable $\phi(s)$ is computed by

$$\hat{\phi}_k = \frac{1}{2N}\sum_{1}^{2N}\phi(\frac{\pi}{N}l)e^{-\pi ikl/N}, \text{ where } k=-N+1,\ldots, N \tag{2.22}$$

Step 5: The inverse Fourier transform is calculated with respect to a selected band (0<kmin≤k≤kmax) of the positive wavenumber half of the Fourier spectrum:

$$w(\frac{\pi l}{N}) = 2\sum_{k=k_{\min}}^{k_{\max}}\hat{\phi}_k e^{\pi ikl/N}. \tag{2.23}$$

Step 6: The wave packet envelope is computed by

$$A(2\pi/N) = |w(\pi l/N)| \tag{2.24}$$

For each original point (x_0, y_0), the above steps are repeated to compute wave packet amplitude A(x0, y0).

2.6.2 An objective way to track RWPs

Souders et al. (2014a, b) proposed a feature-based RWP tracking approach. The method used a hybrid approach to track RWPs by combining a point-based object identification method and a feature-based tracking approach together. This tracking approach is applied here to track RWPs in NCEP/NCAR analysis from 2007 to 2014 cool seasons (November to March).

The tracking approach starts by selecting significant local maxima in WPA ($\geq$14 m s^{-1}). The object is then defined by setting the minimum size to be 40 grid points at 2.5° resolution. The algorithm tracks the objects by searching for significant overlap across a single time interval within a search range (20°N/S, 30°W, and 90°E). Two objects in consecutive time steps will be given the same track ID if the overlap between them covers at least 50% of the area in either object. When multiple objects are related to one existing RWP at the same time, or the merging/splitting happens, the more intense object keeps the existing track ID while the smaller/weaker object will be given a new track ID. Readers are referred to Souders et al. (2014a, b) for more details of this tracking approach. Their results have been shown to have good agreement for most of the characteristics of RWPs with other studies (Grazzini and Vitart 2015).

2.7 Chapter summary

This chapter summarized the data and methodology used throughout this dissertation. The TIGGE data provides a great opportunity to investigate the predictability of high-impact East Coast winter storms. By using different ensemble tools, i.e., ensemble sensitivity, EOF/fuzzy clustering, the model statistics and model bias/errors will be computed and presented. Results from ensemble sensitivity and Fuzzy clustering analyses will assist forecasters to better understand and appreciate forecast uncertainty, while the evaluation of model performance will provide insights to improve forecast accuracy and models. The investigation of predictability associated with RWPs explores the role that large-scale flow regime plays in the forecasts of downstream cyclones.

This page is left to be blank intentionally.

Chapter 3 Ensemble sensitivity and its applications to ensemble forecasting

As previously mentioned, ensemble sensitivity analysis (ESA) is employed to link the ensemble forecast uncertainties with the initial conditions or earlier forecasts, and thus it provides a guidance for adaptive observing strategies as well as interpreting daily ensemble runs (Ancell and Hakim 2007; Hakim and Torn 2008; Torn and Hakim 2009). ESA uses a linear correlation and regression between the chosen forecast metric at verification time and the state vector at the initial or earlier forecast times from the ensemble members to derive the sensitivity of any forecast metric to the state vector.

Since the linear assumption for this sensitivity is expected to work better for short-range forecasts, ESA has mainly been employed for short range (1−3-day) forecasts. However, previous studies have shown that linearly derived sensitivity can still have value out to the medium range. For instance, Chang et al. (2013) showed that ESA signals for a forecast could be tracked backward at least 6 days when studying the evolution of two extreme extratropical cyclones over the Pacific. In this work, we will apply ESA for both short- and medium-range ensemble forecasts by summarizing three ways to assess the uncertainties in the forecasts of the intensity and track of extratropical cyclones. The verification of ensemble sensitivity signals using LOOCV will be given. The model climatology for ensemble sensitivity associated with cyclone intensity and position will also be presented.

Section 3.1 of this chapter will briefly introduce the background of ensemble sensitivity. Section 3.2 discusses the formulation of ESA and different ways of applying it. In Section 3.3, the 2010 December East Coast winter storm will be investigated to illustrate the procedure of applying ESA to diagnose forecast uncertainty with different forecast metrics/state vectors.

Section 3.4 discusses the verification of ESA using LOOCV. The comparison of ESA with ETKF method in a case study will be presented in section 3.5. Section 3.6 will present the climatology of ESA based on historical winter storms. Section 3.7 will briefly discuss the linkage between sensitivity signals and RWPs. The chapter summary will be given in section 3.8.

3.1 Ensemble sensitivity analysis background

Ensemble sensitivity was developed by Hakim and Torn (2008) and further explored in a couple of follow-up papers (Ancell and Hakim 2007; Torn and Hakim 2008). Basically, a scalar forecast metric (a response function *J*) at the verification time is linearly regressed onto a model state variable within an ensemble of forecast at either the verification time or other earlier forecast time (including the initial time). The slope of these linear regressions at a given time is defined as the ensemble sensitivity, which is an $N \times 1$ vector ($\partial J / \partial X$) with X representing the model state with state-space dimension N. While a univariate linear regression within a complex multivariate system would seemingly have its limitations, ensemble sensitivities have been demonstrated to possess a deep dynamical meaning in a number of studies for different applications. Ancell and Hakim (2007) demonstrated that the ensemble sensitivity is proportional to the projection of the analysis-error covariance onto the adjoint-sensitivity field. The sensitivity results from adjoint and ensemble sensitivity were compared in a winter Pacific Northwest cyclone for a forecast metric of near surface pressure at a point in the state of Washington. Adjoint sensitivity patterns mainly revealed lower-tropospheric structures that were more localized with small magnitudes and had a strong upshear tilt with height. In contrast, ensemble sensitivity patterns were mostly troposphere-deep structures that represented more synoptic scale with larger magnitudes and tilted modestly with height. Ensemble sensitivities can more efficiently pick out coherent features in the flow (e.g., upper-level troughs; mid-level temperature gradients) that are dynamically relevant to the forecast metric as compared with the adjoint sensitivity method.

ESA has been successfully applied to a variety of ensemble forecast problems. Torn and Hakim (2009) employed ensemble sensitivity to study two extratropical transition (ET) events associated with two typhoons in the western Pacific Ocean. The forecasts of minimum SLP were highly sensitive to the tropical cyclone (TC) position as well as the upstream midlatitude troughs that interacted with the TC during ET. The variability of the latter impacted the timing of the interaction with the TC, which affected its extratropical transition and ultimately the magnitude of forecast errors at later times. Torn (2010) also applied ESA to evaluate the impacts of initial condition errors on the amplitude and position forecast of a strong African Easterly Waves (AEW) in 2006. Short-term forecasts of wave strength were found to be most sensitive to the thermodynamic profile near the AEW and the initial mid-troposphere circulation strength whereas longer-range forecasts were most sensitive to the thermodynamic environment (potential temperature or θ_e). The AEW longitude forecasts were more sensitive to the downstream thermodynamic environment (θ_e) which suggests the area the wave moves into matters more than the initial circulation.

A series of studies have applied ESA to very short term (≤6 hours) wind speed forecasts for wind farms (Zack et al. 2010a,b) in California and along the Washington-Oregon border. Results of Zack et al. (2010a) showed that sensitivity tended to be mainly localized around the response function but still revealed specific areas upwind around a mountain pass, indicating the importance of diurnal mesoscale flows in Tehachapi Pass over southwest California. In contrast, Zack et al. (2010b) showed that the Mid-Columbia Basin region over Washington-Oregon border was affected by larger scale flow regimes and resulted in less defined and weaker sensitivity patterns. Several other studies applied ESA to convection (Hanley et al. 2013; Bednarczyk and Ancell 2015; Torn and Romine 2015). For example, Bednarczyk and Ancell (2015) showed that ESA could be successfully applied to high-resolution forecast of convection, with its most important features related to the synoptic scale flow such as the positioning of upper-level low and the low-level thermodynamic characteristics of the air mass.

Instead of describing the slope of the linear regression, another form of ensemble sensitivity using correlation evaluates how strong the linear relationship is between extended and early forecasts by dividing the covariance of J and X by the standard deviation of each variable (Sippel and Zhang 2010; Schumacher 2011; Chang et al. 2013). Schumacher (2011) showed that both sensitivity and correlation analysis highlighted the synoptic-scale features in the Z500 field, indicating that the strength variability in a closed anticyclone in the southern plains and a ridge in the southwestern U.S. controlled the maintenance and movement of a warm-core vortex that brought heavy rain to the southern plains. Sippel and Zhang (2010) used the ensemble correlation to study the predictability of a rapidly intensifying hurricane within an EnKF ensemble. The growth in ensemble spread of the storm was found to be related to the variations in midlevel moisture and low-level convective instability as well as the strength of a nearby front.

Ancell and McMurdie (2014) applied ESA to explore the predictability of North Pacific land-falling cyclones. Predictability was defined as the ensemble spread of surface pressure at the verification time, where large spread indicated low predictability. They found that storms that were deepening and tracked from the southwest exhibited the largest ensemble IC sensitivity and highest ensemble spread. The slow-moving storms from the northwest and ending south of 40°N exhibited higher predictability no matter whether they were decaying or deepening. Cyclones with large ensemble spread and low sensitivity were mature cyclones whose low predictability likely resulted from large IC spread instead of large perturbation growth. Their results have highlighted particular cyclone characteristics and synoptic situations those are associated with low predictability and can potentially be utilized to improve forecasts through improved observational coverage.

While the aforementioned studies generally employed ESA to examine the predictability at the short range (0–3 days), Chang et al. (2013) used ensemble correlation analysis in the medium range to study two Pacific extratropical cyclones. The initial conditions derived based on ESA are used to perform perturbed ensemble experiment. Their results suggested that while the development of these cyclones may be highly nonlinear, error growth during their development may still be quasi-linear in nature out to the medium range. The ensemble sensitivity analyses can show coherent sensitivity patterns in the medium range. Meanwhile, the initial condition perturbations derived based on ESA do succeed in modifying cyclone evolution, with the caveats that PCs derived from EOF analyses of forecast SLP variations should be used as forecast metrics instead of cyclone parameters themselves, and that the changes achieved in the perturbed ensembles have amplitudes smaller than those predicted by the ensemble sensitivity analyses. Zheng et al. (2013) further extended the sensitivity study to the operational medium-range forecasts and summarized three ways for applying ESA to diagnose the important features relevant to both cyclone track and intensity forecasts.

Since ESA has been shown to be successful in identifying important features relating to the predictability of weather systems at different time and spatial scales, it is employed in this study to systematically investigate the forecast uncertainty of high-impact winter storms and to establish a model climatology of ensemble sensitivity with respect to the intensity and position of winter storms. While most previous applications of ensemble sensitivity have been to short-range forecasts, this study will mainly discuss the value of ESA to medium-range forecast in global models. The robustness of ensemble sensitivity will be explored by using LOOCV.

3.2 Formulating ensemble sensitivity

The calculation of ensemble sensitivity follows that of Torn and Hakim (2009). The initial formulation of ensemble sensitivity here is a repeat of equation 2.12. Given an ensemble with M members, the sensitivity of any forecast metric J to a state variable X_i can be expressed as

$$\text{"}Sensitivity\text{"} = \frac{Cov(\mathbf{J}, \mathbf{X_i})}{\sqrt{Var(\mathbf{X_i})}}; \tag{3.1}$$

here $\mathbf{J}$ and $\mathbf{X_i}$ are the 1×M ensemble estimates of the forecast metric J and the ith state variable X_i, respectively. Cov denotes the covariance between J and X_i across the ensemble; and Var is the variance. The "Sensitivity" defined by equation (3.1) carries the dimension of the forecast metric J and describes the change in the forecast metric corresponding to one standard deviation change in the uncertainties of the selected state variable X_i.

Ensemble sensitivity analysis can be applied in three different ways based on the choice of forecast metrics: sensitivity using the EOF approach, sensitivity using run cycle difference, and forward sensitivity regression using short-range forecast errors. I will explain the details of these three formulations in the following subsections.

3.2.1 Ensemble sensitivity using the EOF approach

To perform the ensemble sensitivity, we first define the forecast metric using an EOF analysis (Hannachi et al. 2007; Wilks 2011). In this study, the EOF analysis is carried out on perturbations across an ensemble of forecasts at the verification time, i.e. perturbations defined as deviations from the ensemble mean. A box around the forecast cyclone is defined on the model grid to calculate the dominant EOF patterns in the ensemble MSLP forecast at a particular forecast hour. The EOF patterns basically represent the dominant forecast uncertainties within the box. This approach was used rather than identifying forecast metric for cyclone parameters directly, since the cyclones in each ensemble member with similar parameters (minimum pressure, longitude, or latitude) may not have a similar structure (Gombos et al. 2012; Chang et al. 2013). The corresponding PCs of each of the leading EOF patterns (which are the projection of the difference between each ensemble member and the ensemble mean onto the EOF patterns) can be employed as the forecast metrics (J). If the approach is applied with a cyclone within the specified box, the EOF patterns typically represent the amplitude and position uncertainties of the cyclone (Gombos et al. 2012; Chang et al. 2013). Therefore, using the PCs of an EOF as the forecast metric, ESA can objectively evaluate the relationship between that uncertainty pattern and the state variables at initial time or other forecast times. Sometimes the dominant patterns represent a combination of both amplitude and position uncertainties in the forecasts, so the ensemble sensitivity should be interpreted accordingly.

In the applications to this dissertation, the PCs are dimensionless and normalized to unit variance, while the EOF patterns carry the dimension of the forecast variable, and the amplitude of the patterns reflects the square root of the amount of variance explained by each EOF. In other words, the forecast metric employed here has unit variance and is non-dimensional. Recall the definition of correlation:

$$"Correlation" = \frac{Cov(\mathbf{J}, \mathbf{X_i})}{\sqrt{Var(\mathbf{X_i})}\sqrt{Var(\mathbf{J})}}; \tag{3.2}$$

the expression for “sensitivity” in (3.1) using metrics with unit variance is identical to the correlation between the ensemble estimates of J and X_i. In the discussions below, this definition will be used for “sensitivity” (see also Chang et al. 2013).

3.2.2 Sensitivity using run cycle differences

The second application of ESA is to study the shift in the model forecast of a particular feature (e.g. cyclone) between two close forecast cycles. Forecasters often refer to this as the “d(model)/dt” of the forecast cycles. This shift in the forecast pattern P can be defined by the difference in the ensemble mean forecasts at the same valid time between two forecasts initialized at two different times. The anomaly pattern Y_m, which is the difference between the mth ensemble member and the ensemble mean at the forecast valid time, can be projected onto the shift (forecast difference) pattern P as follows:

$$J_m^{dif} = \frac{\sum_{j=1}^{N} P_j Y_{mj}}{\sum_{j=1}^{N} P_j^2}; \tag{3.3}$$

where N is the number of total grid points over the boxed verification region, and $m = 1, \ldots, M$ identifies the ensemble member. This alternative forecast metric J_m^{dif} in equation (3.3) can be used to calculate the sensitivity (correlation) in equation (3.2) to objectively determine what may have caused the shift in the forecast between two forecast cycles.

3.2.3 Forward sensitivity regression using short-range forecast errors

The previous two approaches calculate the sensitivity backward in time, but the ESA can also be applied to study the forward impact of a short-term error early in the forecast on some later forecast hour. First, the difference pattern Q is calculated between the analysis valid at T_0

and the short-term forecast initialized at (T_0-ΔT) and also valid at T_0. This difference is The same as the forecast errors at time T_0 but with opposite signs. Then the pattern Z_m, which is the difference between the *m*th ensemble member and the ensemble mean at time T_0, for the ensemble initialized at time (T_0-ΔT), can be projected onto the pattern Q, and the projection coefficient is defined as the forecast metric, as follows:

$$J_m^{fwd} = \frac{\sum_{j=1}^{N} Q_j Z_{mj}}{\sum_{j=1}^{N} Q_j^2}; \tag{3.4}$$

here N is again the total number of grid points over the region we are interested in at time T_0. Finally, the projection coefficient J_m^{fwd} (m = 1, …, M) can be regressed onto the forecasts of cyclone state variables at analysis time T_0 as well as each future forecast time and thus used to evaluate the impact of the short-range error on the subsequent forecast evolution of the cyclone. The regression can be interpreted as another form of sensitivity, which is called "forward ensemble regression" here and expressed as

$$\text{"}Forward\ Ensemble\ Regression\text{"} = \alpha \frac{Cov(\mathbf{J}, \mathbf{X_i})}{\sqrt{Var(\mathbf{J})}}; \tag{3.5}$$

"forward ensemble regression" in (3.5) has the dimension of the state variable X_i. α in (3.5) is a scaling factor, which is defined such that the amplitudes of Q and the forward ensemble regression at time T_0 are the same. Both Z_m in equation (3.4) and X_i in equation (3.5) are calculated using the earlier forecast cycle since it has larger variations among ensemble members and thus give less noisy signals especially at earlier times. The ensemble sensitivity or regression in equation (3.5) is complementary to the usual way of ESA and can assess how forecast errors at earlier times affect the forecasts of later times within the same forecast cycle.

3.3 A case study applying ensemble sensitivity to an East Coast winter storm

This section will briefly illustrate how ensemble sensitivity can be applied to diagnose ensemble forecasts of East Coast winter storms by using one high-impact storm in December 2010. Interested readers are referred to Zheng et al. (2013) for more details of this case study.

On 26–28 December 2010, a powerful winter storm impacted the Northeast of the U.S., resulting in blizzard conditions across the region, including the New York metropolitan area, New Jersey, and portions of New England. The northern part of New Jersey received around 82 cm (32 inches) of snow, while snowfall total in excess of 51 cm (20 inches) fell in many places of New York City, which, in conjunction with 22 to 27 m s^{-1} winds, crippled the transportation from New Jersey to New England during the busy traffic period just after Christmas day.

Though the forecasts based on the ECMWF model indicated a potential East Coast snowstorm a week in advance, the ensemble mean position of the storm was too far east (Figures.3.1a-b, e-f). The cyclone track shifted back and forth in the forecasts before 0000 UTC 24 December. The NCEP model correctly forecasted the cyclone position half a day in advance of the ECMWF model. Finally, at 0000 UTC 25 December, operational models converged on the storm track and indicated the New York City-Boston corridor was in line with the potential for the heaviest snow and strong winds associated with a rapidly deepening surface low (Figures 3.1d, h). This consensus for this snowstorm event emerged only 36 to 48 h before the onset (1200 UTC 26 December to 0000 UTC 27 December) of the heaviest snows.

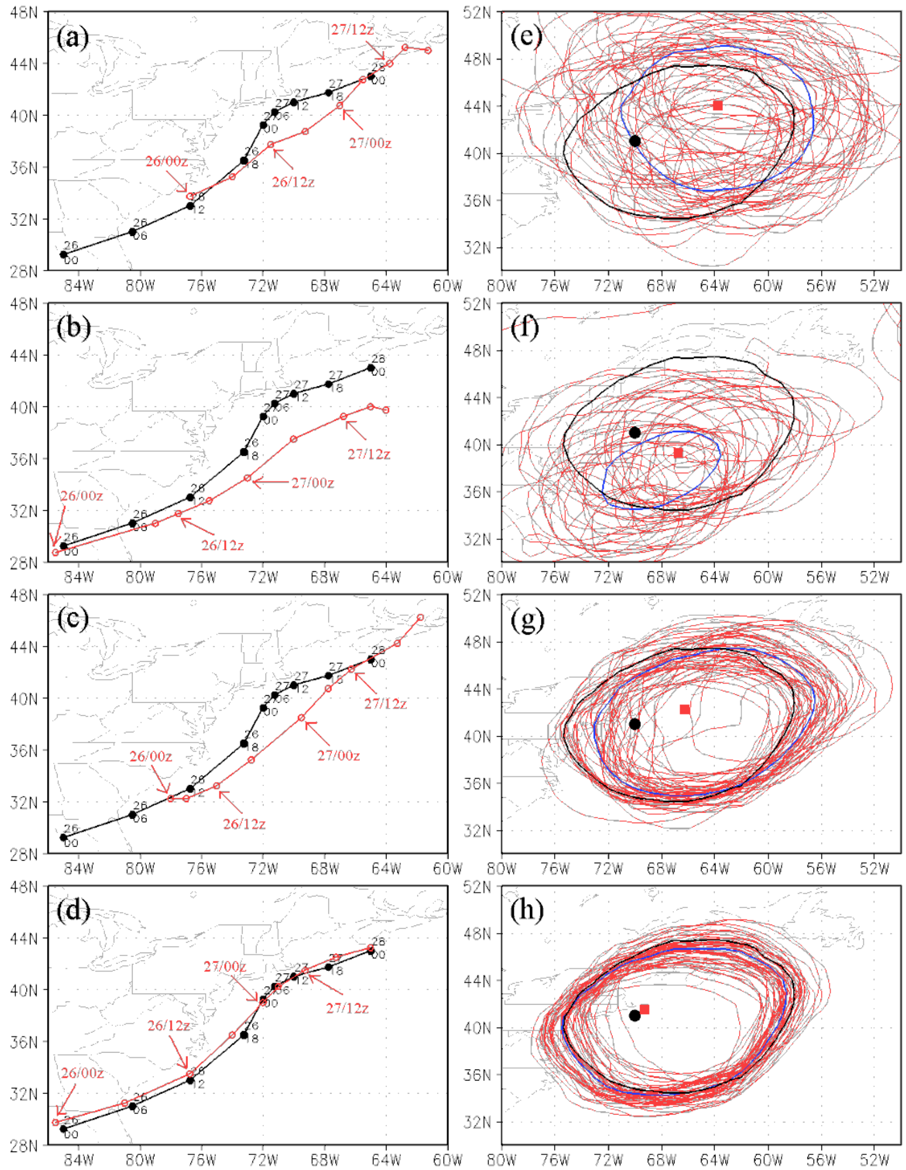


Figure 3.1: (left, a-d) Surface cyclone tracks starting from 0000 UTC 26 December to 1200 UTC 28 December initialized at 1200 UTC 20, 0000 UTC 22, 0000 UTC 24, and 0000 UTC 25 December 2010; (right, e-h) ensemble forecasts of 996 hPa MSLP for 50 ECMWF (red) members valid at 1200 UTC 27 December initialized at 1200 UTC 20, 0000 UTC 22, 0000 UTC 24, 0000 UTC 25 December 2010; black and blue thick line represent analysis and ensemble mean of 996 hPa MSLP contour; red and black dots represent ensemble mean positions of surface cyclone for ECMWF and analysis, respectively.

All the forecast models had difficulties in forecasting this event, but the primary model used in this case study is the ECMWF EPS. The ECMWF data was retrieved from the THORPEX TIGGE archive. The ECMWF EPS was chosen because of its relatively large number of members (50 members) and its superior performance compared to other EPSs based on several verification methods (Park et al. 2008). The initial conditions for the EPS are created by adding to the operational ECMWF analysis perturbations using the singular vector technique that produce the fastest energy growth during the first two days of the forecast period. The separation between the ensemble members is further enhanced by the use of stochastic physics. The ECMWF analysis, which is generated by the high-resolution deterministic model (T1279/N640, ~16-km horizontal grid spacing) and interpolated onto 1° latitude × 1° longitude grid, is used to compare with the ensemble forecasts. The four ECMWF ensemble forecasts used were initialized at 7, 5.5, 3.5 and 2.5 days before the verification time (1200 UTC 27 December 2010). The main forecasting parameters used are MSLP, Z500, and meridional velocity at 300 hPa pressure level.

3.3.1 Synoptic overview of the case

Figures 3.2 and 3.3 showed the evolution of the MSLP with 2-m temperature and the geopotential height with absolute vorticity at 500 hPa in the ECMWF analysis. At 0000 UTC 25 December (Figure 3.2a), there was a surface frontal boundary from southern Texas extending eastward along the Gulf Coast. At 500 hPa, there were a trough and an absolute vorticity maximum over central Texas and another short-wave trough over the central Plains (Figure 3.3a). Surface cyclogenesis occurred at 1200 UTC 25 December (Figure 3.2b) over southeastern Louisiana with a central pressure of 1013 hPa. The northern and southern stream upper-level troughs merged between 12Z/25 and 00Z/26 (Figures. 3.3b, c). The surface cyclone over the Gulf of Mexico deepened and moved first eastward across northern Florida, then northward to the east coast of South Carolina at 12Z/26 (Figures 3.2c, d), while the upper-level trough deepened and was situated over the southeast U.S. (Figure 3.3d). The surface

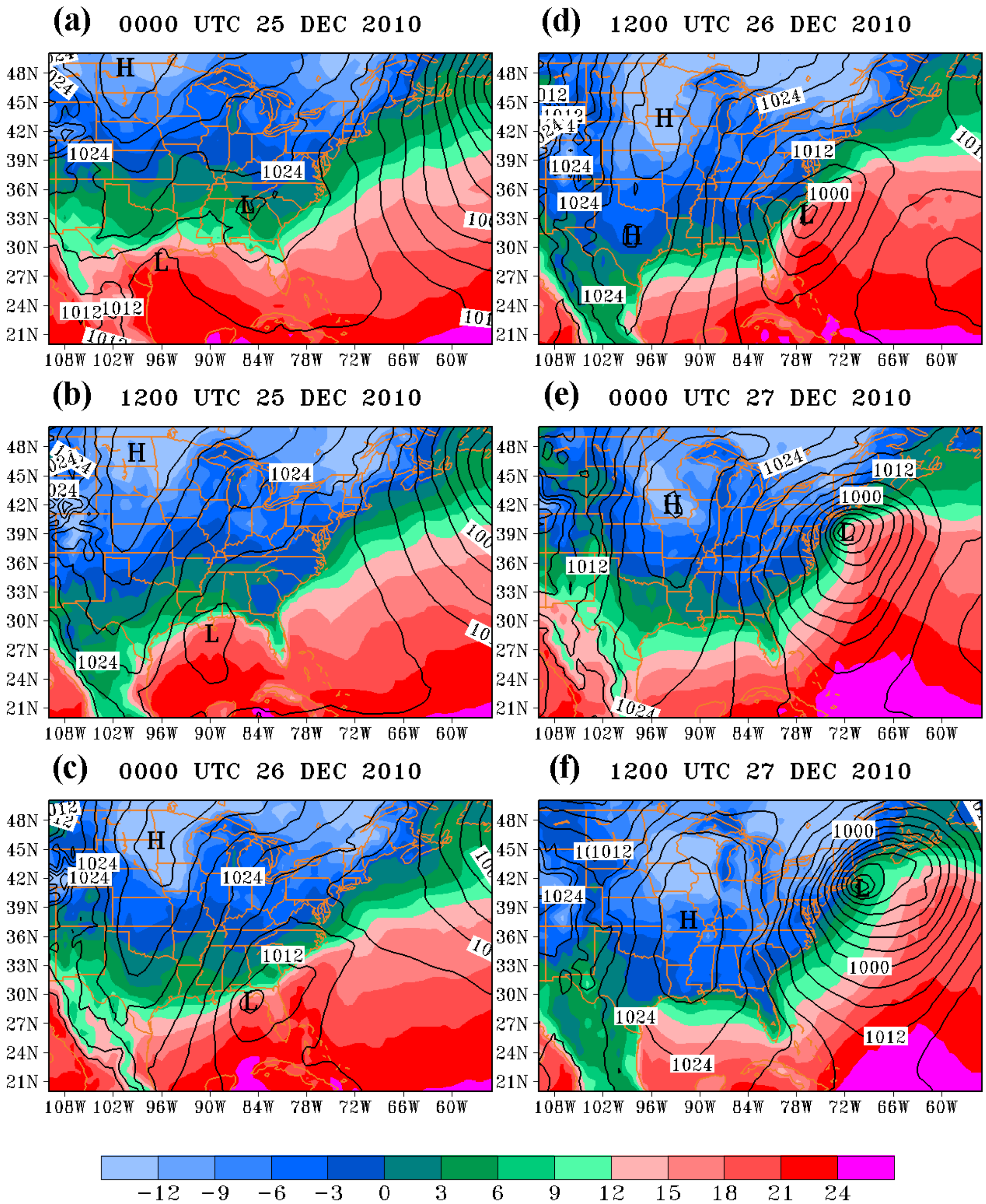


Figure 3.2: (a-f) Surface weather charts: MSLP (black contours, [hPa]), 2-meter air temperature (Shading, [°C]) starting from 0000 UTC 25 December to 1200 UTC 27 December 2010. Contour interval: 4 hPa.

cyclone tracked northeast of Cape Hatteras from 06Z/26 to 18Z/26 (Figure 3.1). Meanwhile, the central pressure dropped from 1002 hPa to 986 hPa (not shown). After 18Z/26, the main low quickly tracked more northward along the mid-Atlantic Coast and then northeastward along the New England Coast. Between 18Z/26 and 00Z/27 the central pressure dropped 12 hPa (not shown).

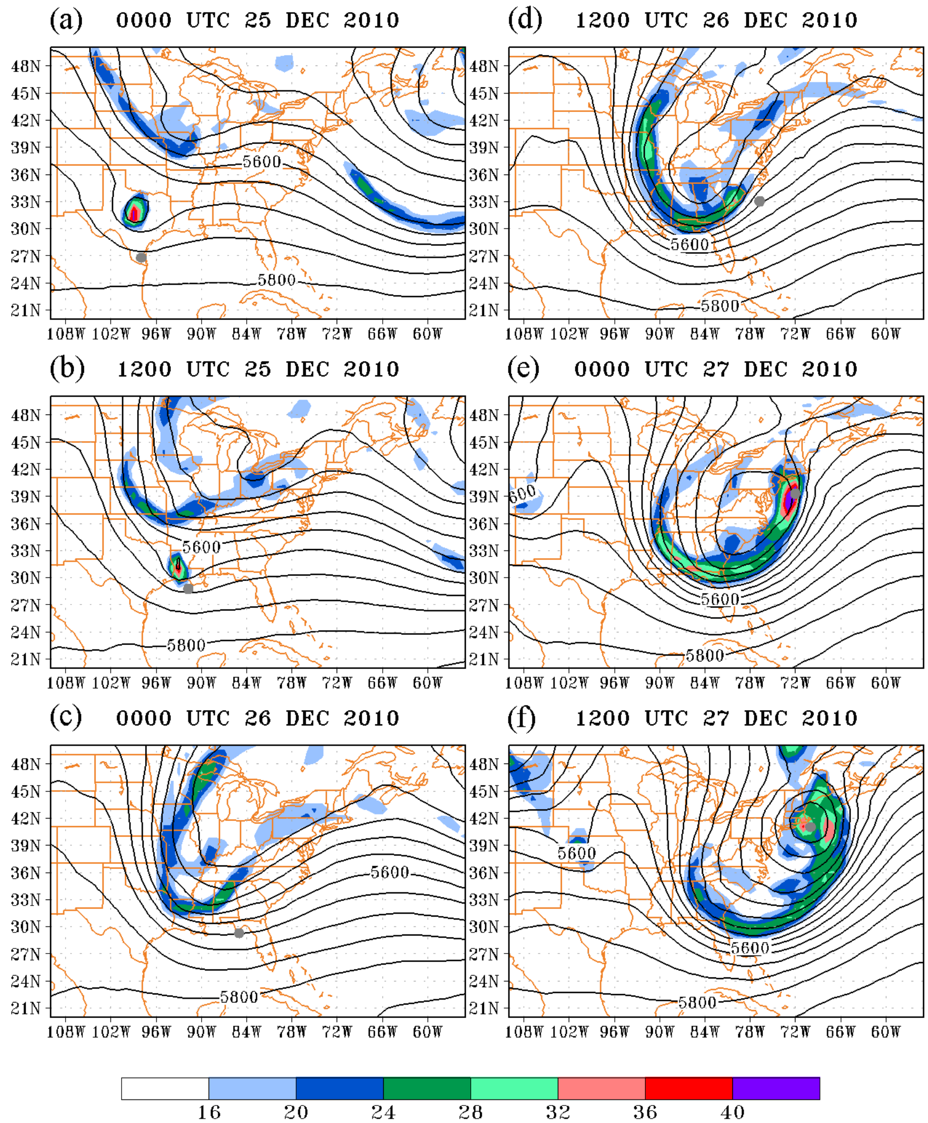


Figure 3.3: (a-f) Geopotential height (thick black contour, [m]) and absolute vorticity [×10-5 s-1] at 500 hPa level starting from 0000 UTC 25 December to 1200 UTC 27 December 2010. Contour interval: 50 m. The grey dot in each panel denotes the corresponding surface cyclone position.

By 00Z/27, the intense low pressure system was centered ~160 km to the southeast of Long Island, New York, with a central pressure of 976 hPa (Figure 3.2e). The positions of the surface cyclone and the 500 hPa trough show strong vertical phase tilt (Figures 3.3c, d, e), indicating that between 26/00z and 27/00z, there was a strongly baroclinic situation under which the surface low could strengthen rapidly. The associated 500 hPa trough rotated northeastward along the East coast between 00Z/27 and 12Z/27 (Figures 3.3e, f). By 12Z/27, the surface low has moved to near Cape Cod of Massachusetts (Figure 3.3f). From 00Z/27 to 18Z/27, the surface low continued moving along the coast to eastern New England, with the central pressure dropping from 976 hPa to 962 hPa.

3.3.2 Ensemble sensitivity results

ESA is applied in three different ways based on the choice of forecast metrics as discussed in section 3.2. First, we define forecast metric (J) using the PCs of each of the leading EOF patterns over a region around the forecast cyclone at a fixed investigation time. The MSLP and Z500 at different forecast time are chosen as the state vectors. The calculations of ensemble sensitivity are based on equation (3.1). Given 50 ensemble members, a "sensitivity" value of 0.28 is statistically significant at the 95% confidence level based on a two-tailed t-test. The second application of ESA here is to study the shift in the model forecast of cyclone between two close forecast cycles by employing equations (3.2)-(3.3). The third application is to study the forward impact of a short-term error early in the forecast on some later forecast hour. Equations (3.4) and (3.5) are used to evaluate the impact of short-range error on the subsequent forecast evolution of the cyclone.

3.3.2.1 Ensemble sensitivity using the EOF approach based on a 5.5-day forecast

As discussed in section 3.2, the EOF method can determine the dominant patterns of forecast variability that explain most of the variance within the ensemble spread. For the 5.5-day

MSLP forecast initialized at 0000 UTC 22 December, the ensemble mean cyclone is centered at (66.75°W, 39.25°N) and the maximum of the ensemble variance is ~300 km to the northwest of the cyclone center (Figure 3.4a). Figures 3.4b, c show the leading two EOF patterns for this 5.5-day MSLP forecast. Note that the sign of each EOF is arbitrary; the values of the PC just take the opposite sign if the opposite sign of the EOF is taken. The first EOF (EOF1) is a monopole pattern centered slightly north of the ensemble mean position of the surface cyclone (Figure 3.4b). Therefore, one can interpret the EOF1 pattern as a deeper and a slight northward shift of the ensemble predicted cyclone on day 5.5. Meanwhile, EOF2 is a dipole pattern that has a negative center west-southwest of the ensemble mean cyclone position and a positive center to the east-northeast (Figure 3.4c), representing more of the east-west position uncertainty on day 5.5. This pattern suggests the west-southwestward (as well as east-northeastward) shifted cyclone in a subset of ensemble members at the verification time.

To study the relation between each of these two forecast uncertainty patterns on day 5.5 and the upstream initial condition and forecasts, the PC of each EOF pattern was used as the forecast metric, and the sensitivity in equation (3.1) was calculated using Z500 as the state vector at different forecast/analysis times. Positive (negative) sensitivity indicates areas that an increase (decrease) in geopotential height is associated with an increase in the PCs and therefore an enhancement of each of the two EOF patterns.

The sensitivity for PC1 (PC for EOF1 pattern) at the forecast valid time, 0 h or the day 5.5 forecast (Figure 3.5a), shows a large negative sensitivity (< -0.9) near the center of a trough (<540 dm) along the eastern U.S. coast. Meanwhile, there is large positive sensitivity (> 0.7) that spreads from the upper-level ridge located over the western North Atlantic northward and westward towards the ridge upstream of the trough. The distribution of the sensitivity suggests that a deepening of the upper level trough and amplification of its upstream and downstream ridges corresponds to an increase in PC1 or an enhancement of EOF1 pattern, which is the strengthening of the surface cyclone at the verification time.

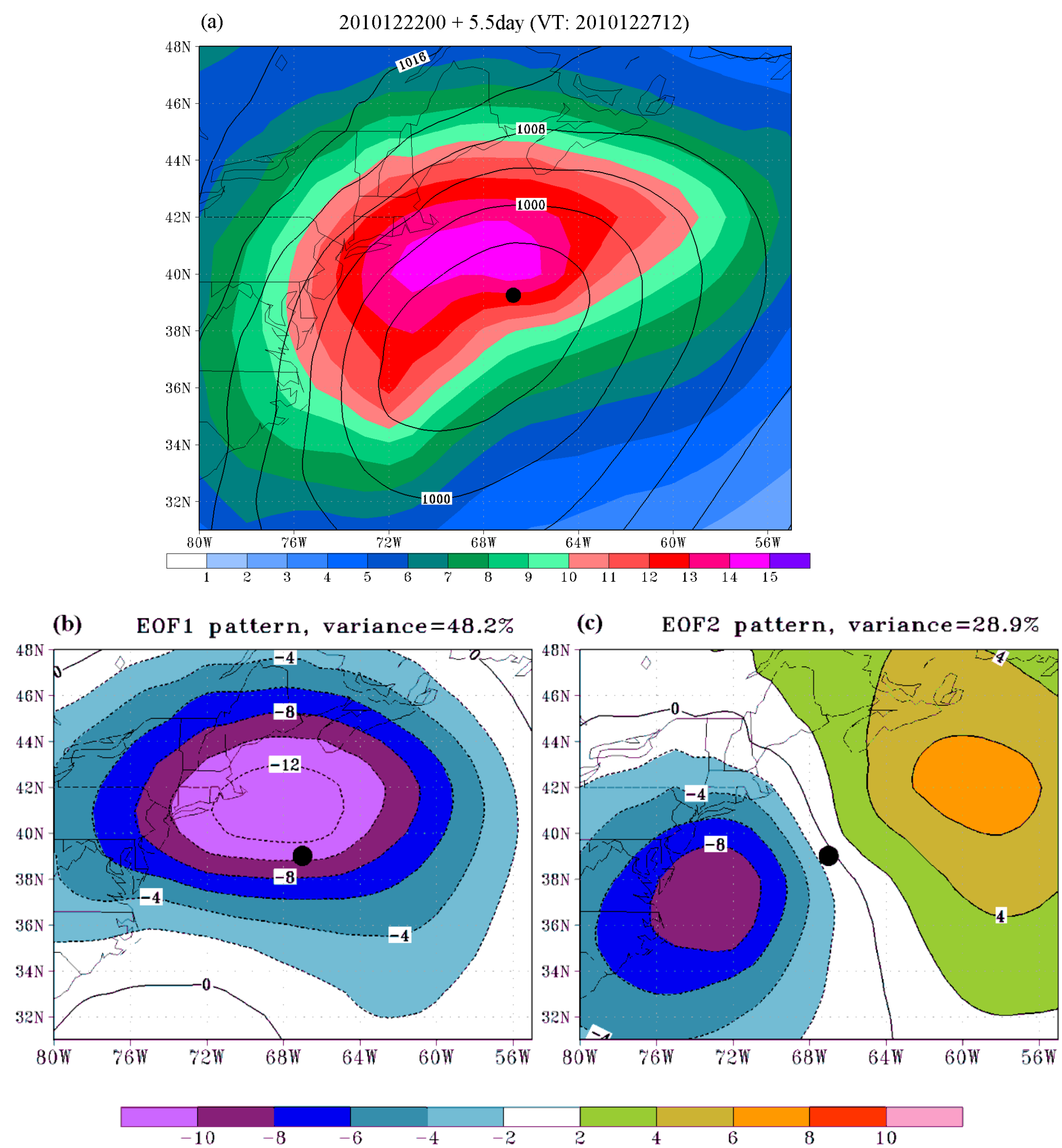


Figure 3.4: (a) Ensemble mean MSLP (contours, [hPa]) and spread (shading, [hPa]); (b) EOF 1 MSLP pattern, unit: [hPa]; (c) EOF 2 MSLP pattern, unit: [hPa]. Valid time (VT): 1200 UTC 27 December2010; Initial time (IT): 0000 UTC 22 December 2010.

At -36 h (day 4 forecast, Figure 3.5b), the sensitivities weaken and shift westward along with the synoptic weather systems. There are two minima (< -0.7) for the negative sensitivity, one northwest of the Great Lakes, and the other centered near Kentucky and Tennessee with the eastern U.S. trough. The positive sensitivity has two maxima, one near the northern part of the upstream ridge over central and northwestern Canada, and the other with the downstream ridge over the western Atlantic. At -60 h (day 3 forecast, Figure 3.5c), there are two short wave troughs, one over the northern Great Plains and another over Texas and Oklahoma. The negative sensitivity (< -0.7) is still centered within the northern stream trough, while the largest positive sensitivity is located downstream within the weak ridge along the East Coast. This pair of sensitivity regions indicates that the intensification of the northern trough and its downstream ridge are closely associated with the amplification of the surface cyclone at valid time.

At even earlier times (-84h to -108h, or day 2 to day 1 forecast, Figures 3.5d, e), the sensitivity areas shift westward, with some indications of upstream sensitivity developing over the northeastern North Pacific, western Canada, and southern Alaska, and this oscillating pattern of sensitivity signals resembles the structure of a Rossby wave train.

The sensitivity using the PCs of EOF2 is shown in the right panels in Figure 3.5. At valid time, 0h (Figure 3.5f), there is negative sensitivity (< -0.5) located over the base of the trough over the southeast U.S. coast. Meanwhile, there is a positive sensitivity (> 0.9) area located just west of the downstream western Atlantic ridge. The distribution of sensitivity suggests that the west-southwest shift and strengthening of the East Coast upper level trough and its downstream ridge correspond to an enhancement of EOF2 pattern, which is the west-southwest shift of the potential cyclone at forecast valid time.

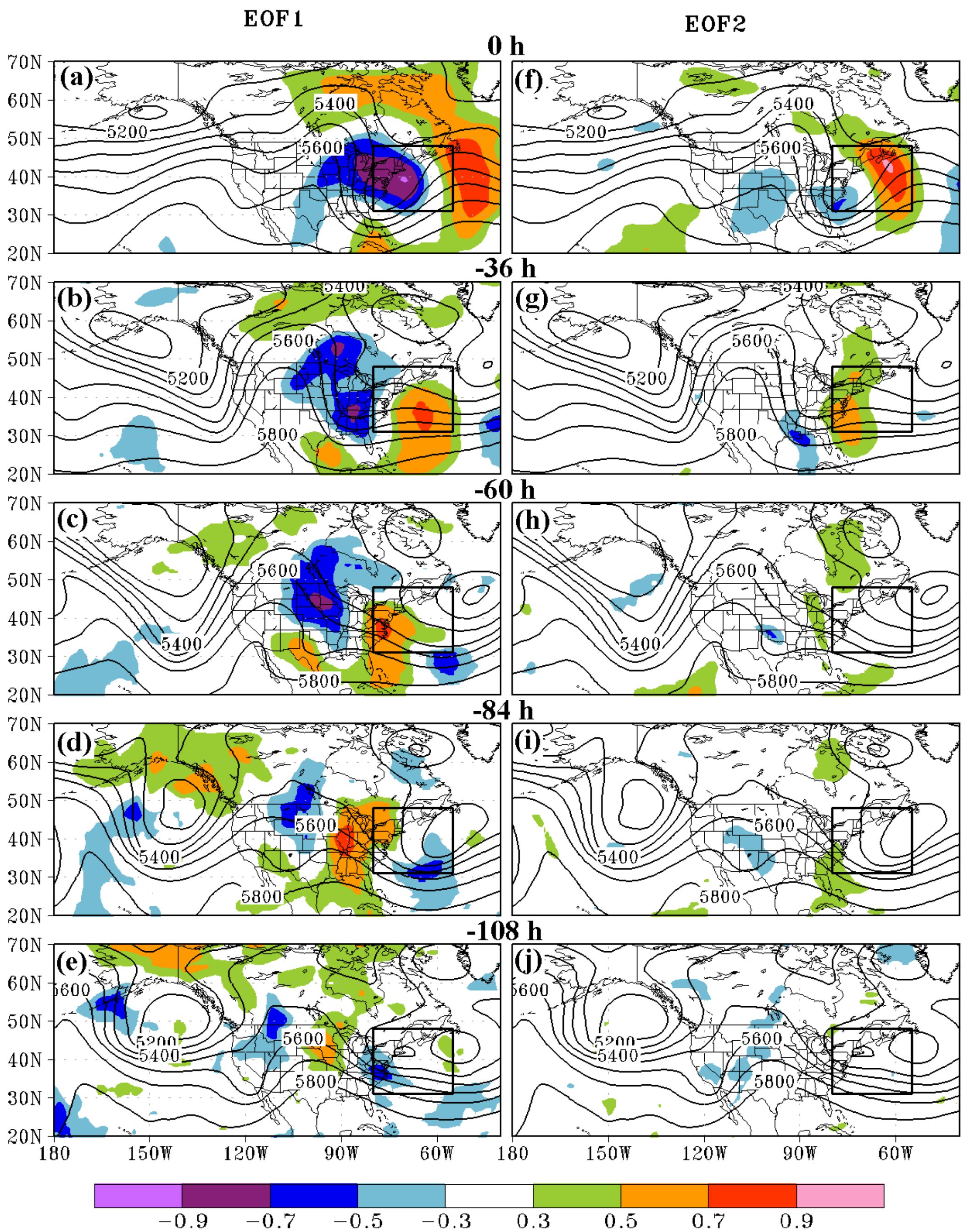


Figure 3.5: (left, a-e) Sensitivity (shading) of EOF PC1 to Z500; (right, f-j) Sensitivity (shading) of EOF PC2 to geopotential height. Contours are ensemble mean of Z500. VT (0 h): 1200 UTC 27 December 2010; IT: 0000 UTC 22 December 2010.

When going backward in time (Figures 3.5g-j), the sensitivity regions shift westward with the corresponding weather systems and are reduced in amplitude. The positive sensitivity with the western Atlantic upper-level ridge decreases rapidly and splits into broad but weak pieces (<0.5). However, the negative sensitivity maintains its central value of ~-0.6 at -36 h and it is still less than -0.5 even at -60 h (Figure 3.5h). Even at earlier times (-84 h and -108h; Figures 3.5i, j), there is still negative sensitivity around the southern stream trough, suggesting this negative sensitivity with this southern Plains short-wave trough is the most consistent and robust signal in affecting the west-southwest shift of the day-5.5 surface cyclone. Thus, accurate predictions of the southern short wave trough is very important for the potential cyclone position, i.e., onshore versus offshore solution, at the valid time. We will highlight similar results for this short-wave trough in the following two subsections.

3.3.2.2 Sensitivity using run cycle differences

There were inconsistent short-term forecast predictions by the ECMWF ensemble for this case. For example, the ECMWF ensemble run cycles before 0000 UTC 25 December had the mean storm position valid at 1200 UTC 27 December ~550 km to the east of the U.S. East coast (Figure 3.1), but starting with the 0000 UTC 25 December run cycle the ECMWF ensemble positioned the cyclone within ~300 km of the coast (Figure 3.1). To determine what may have caused the sudden change in the forecast cyclone position, the difference between the ensemble MSLP forecasts valid at 1200 UTC 27 December, from 0000 UTC 24 December and 0000 UTC 25 December (84-h and 60-h forecast), was computed.

Figure 3.6 shows the ensemble mean and spread of MSLP for the two forecasts valid at 1200 UTC 27 December as well as the difference between them. For the 0000 UTC 24

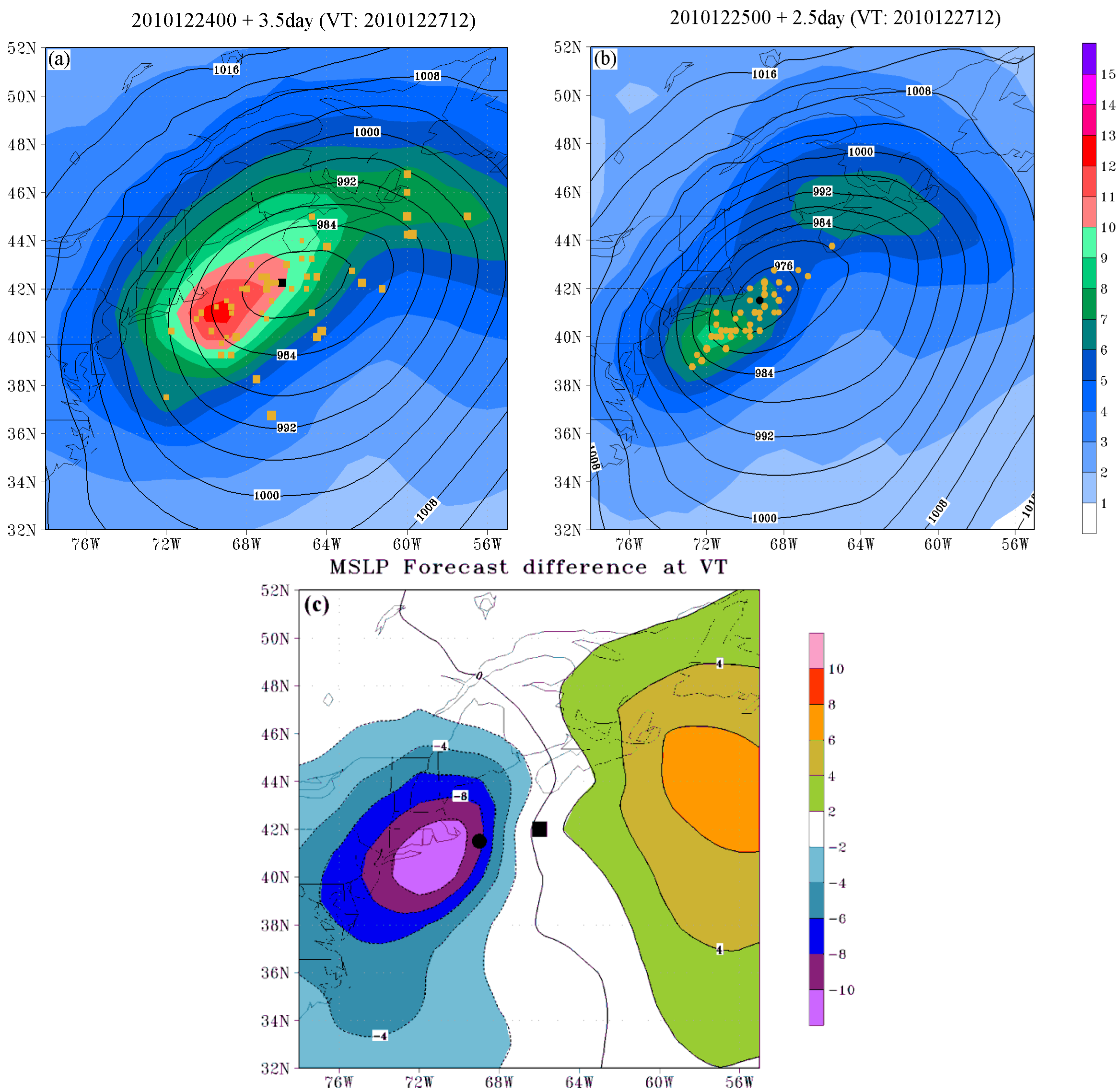


Figure 3.6: (a) Ensemble mean MSLP (contours, [hPa]) and spread (shading, [hPa]) for forecast initialized at 0000 UTC 24 DEC 2010; (b) Ensemble mean MSLP (contours, [hPa]) and spread (shading, [hPa]) for forecast initialized at 0000 UTC 25 DEC 2010; (c) Difference between ensemble mean MSLP initialized at 0000 UTC 25 DEC 2010 and that initialized at 0000 UTC 24 DEC 2010 (shading, [hPa]). VT: 1200 UTC 27 2010. Solid black circle is the ensemble mean position of surface cyclone forecast by cycle 0000 UTC DEC 2010 and solid black square is that forecast by cycle 0000 UTC DEC 2010. In (a) and (b), dark yellow squares/dots represent cyclone positions forecast by 50 individual ensemble members and their sizes represent cyclone central pressure relative to ensemble mean central pressure with larger (smaller) size representing weaker (stronger) cyclone.

December forecast (Figure 3.6a), the surface cyclone at the valid time is centered at (42.25 °N, 66.25 °W) with a 978 hPa minimum pressure, while the cyclone forecast for the 0000 UTC 25

December cycle (Figure 3.6b) is centered ~250 km to the west-southwest with a minimum pressure of 974 hPa. Therefore, the forecast difference in the MSLP valid at 1200 UTC 27 December shows a dipole pattern, with a negative difference to the west-southwest and the positive difference to the east-northeast over the western North Atlantic (Figure 3.6c).

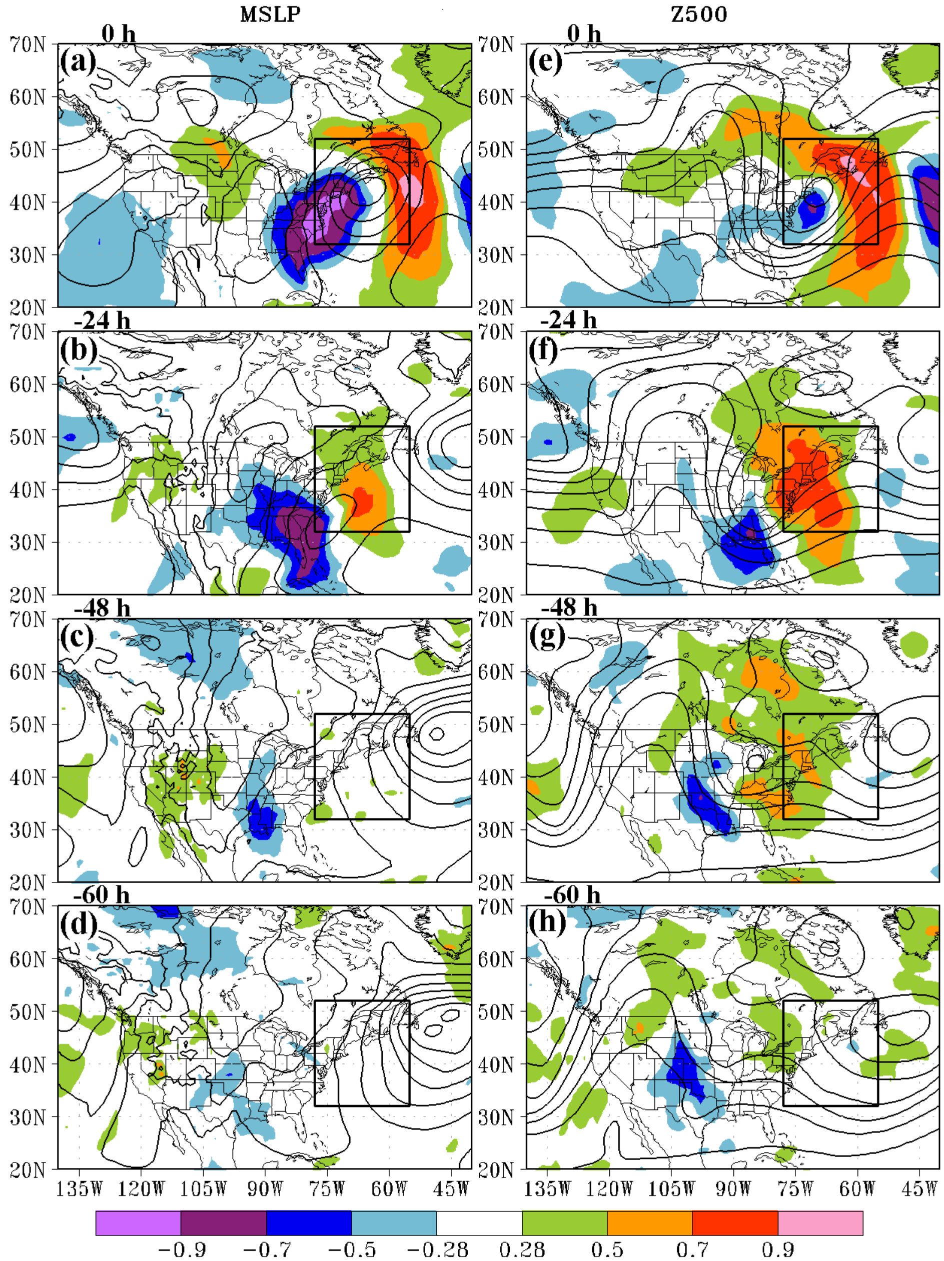


Figure 3.7: Sensitivity (shading) of projection coefficients on forecast jump pattern to MSLP (a-d, left) and Z500 (e-h, right). Contours are ensemble mean of MSLP and geopotential height, respectively. VT (0 h): 1200 UTC 27 DEC 2010; IT: 0000 UTC 24 DEC 2010.

To determine the features that may have caused this run cycle shift, the sensitivity was computed using equations (3.1) and (3.3). Since the earlier cycle initialized at 0000 UTC 24 December has larger spread (Figure 3.6a) and results based on these two run cycles are consistent, only the results for the earlier run will be shown. The forecast metric J for equation (3.3) is the projection coefficient of each member's anomaly over the boxed region (32°N-52°N, 78°W-55°W) onto the forecast shift pattern (Figure 3.6c) over the same region at forecast valid time.

Figure 3.7 shows the sensitivity of the projection coefficient to MSLP and 500 hPa Z. Significant positive (negative) sensitivity (absolute value > 0.28) indicates areas that an increase (decrease) of MSLP/Z500 is associated with the enhancement of the cyclone difference pattern in Figure 3.6c. At forecast valid time, the sensitivity displays a dipole pattern at the surface (Fig. 3.7a), which resembles the forecast difference pattern in Figure 3.6c. Meanwhile, there is large negative sensitivity over the southwestern part of the upper level low along the Northeast U.S. coast and positive sensitivity just west of the downstream ridge (Figure 3.7e). The sensitivity shifts westward with the major synoptic systems going backward in time. At -24 h, or the 60 h forecast (Figures 3.7b, f), the surface negative and positive sensitivities are still located at southwestern and northeastern parts of the surface cyclone, respectively, while at upper levels there is a large negative sensitivity over the southern base of the East Central U.S. trough and a positive sensitivity region to its northeast. At -48 h, or 36 h forecast (Figures 3.7c, g), surface negative sensitivity (< -0.5) is centered over Louisiana, while for the upper levels there is negative sensitivity along the U.S. Gulf coast and positive sensitivity over the eastern U.S.

Meanwhile, a weak negative sensitivity region can be seen in both surface and 500 hPa levels over parts of northwest Canada, indicating the fall of surface pressure and Z500 may also contribute to the shift in the cyclone forecast. At -60 h, or 24 h forecast (Figures 3.7d, h), which is also the initial time for the 0000 UTC 25 December forecast cycle, negative and weak sensitivity pieces are distributed over the southern Great Plains at the surface, which is associated with a negative sensitivity northwest of the short wave trough over Texas. In addition, some negative sensitivity is located over northern Canada. Results shown in Figure 3.7 suggest that the short wave over Texas may be the most important feature contributing to the forecast cyclone difference between the two run cycles.

3.3.2.3. Forward sensitivity regression using short-range forecast errors

To further confirm the impact of the short wave trough over the southern Great Plains at early times (from 0000 UTC 24 December to 0000 UTC 25 December) on the surface cyclone forecast at 1200 UTC 27 December, the forward ensemble regression was computed forward in time using equations (3.4) and (3.5). This analysis used the difference between the 24-h forecast initialized at 0000 UTC 24 December and analysis at 0000 UTC 25 December. Figure 3.8 Shows the differences in ensemble mean MSLP/Z500 between the two runs.

To assess the impact of the forecast error at 0000 UTC 25 December on forecasts of MSLP/Z at later times within the cycle, the forward ensemble regression on MSLP/geopotential height is computed based on equations 3.4 and 3.5, in which Z_m is each member's Z500 anomaly to ensemble mean within the red box (box1) over Texas in Figure 3.8h; Q is the negative ensemble mean forecast error of Z500 within the same box (Figures 3.8h and 3.9); *J* is the projection coefficient of Z_m onto Q, and X_i is either the ensemble MSLP or Z500. Note that the forward ensemble regression shown in Figure 3.10 is computed based only on the ensemble initialized on 0000 UTC 24 December, and the ensemble initialized on 0000 UTC 25 December is only used to derive the initial error pattern (Figure 3.9b).

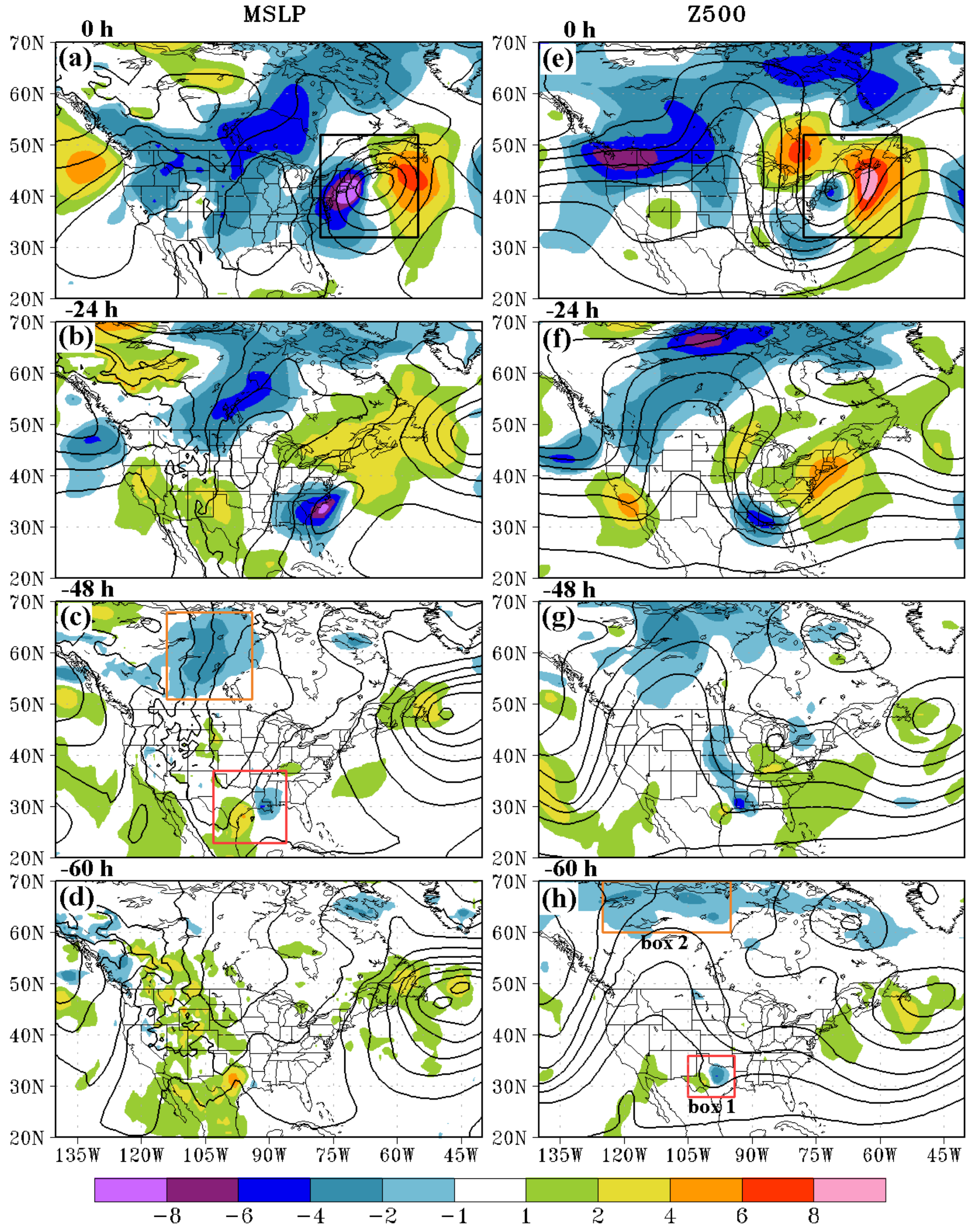


Figure 3.8: Difference between ensemble mean MSLP (a-d, shading, [hPa]) or Z500 (e-h, shading, [dm]) initialized at 0000 UTC 25 DEC 2010 and that initialized at 0000 UTC 24 DEC 2010. VT (0 h): 1200 UTC 27 DEC 2010. Contours are ensemble mean MSLP and Z500

initialized at 0000 UTC 24 DEC 2010, respectively. Boxes 1 and 2 will be used in forward ensemble regression calculations.

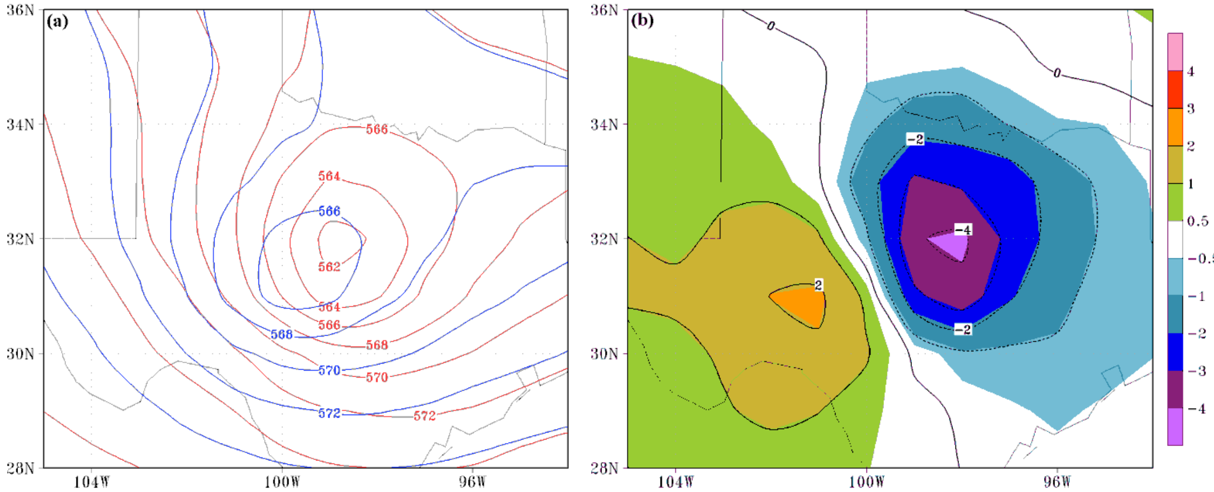


Figure 3.9: (a) 500 hPa Geopotential height analysis at 0000 UTC 25 DEC 2010 (red contours, [dm]) and 24-hr forecast of 500 hPa geopotential height initialized at 0000 UTC 24 DEC 2010 (blue contours, [dm]); (b) difference (shading) of Z between analysis [red contours in (a)] and 24-h forecast [blue contours in (a)], unit: [dm]. Note, the region is the solid red box (box 1) in Figure 3.8h.

The left and right panels in Figure 3.10 show the forward ensemble regression for MSLP and Z500, respectively, which can be considered to be the weighted mean of ensemble members that have anomalies resembling the chosen forecast error pattern (Figure 3.9b). At the initial time, -60h or 0000 UTC 25 December, there is a dipole pattern with the short wave trough over Texas (Figure 3.10), which is consistent with the forecast error over that region (Figure 3.10h), while the MSLP regression shows negative value over eastern Texas (Figure 3.10d). Subsequently, both the MSLP and Z500 sensitivity develop and propagate northeastward with the major synoptic systems. At -48 h (Figure 3.10c), a dipole forms on the surface over the southwestern Gulf coast. At 500 hPa (Figure 3.10g), the dipole over Texas is more enhanced with the negative part spreading to the northwest of the trough.

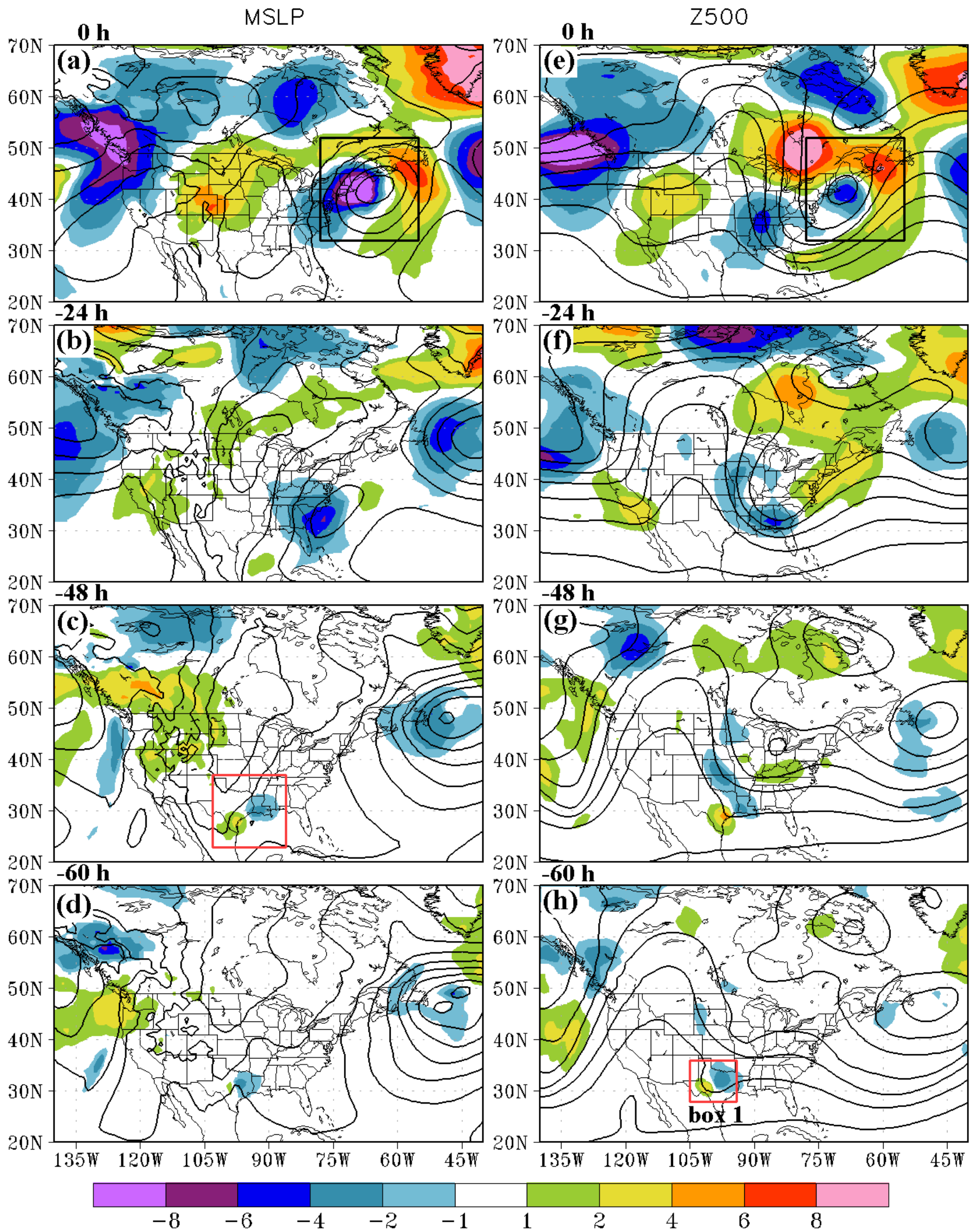


Figure 3.10: Sensitivity regression (shading) of projection coefficients on forecast error pattern in Fig.3.9b over box 1 to MSLP (a-d; [hPa]) and Z500 (e-h; [dm]). Contours are ensemble mean of MSLP and Z500, respectively. VT (0 h): 1200 UTC 27 DEC 2010; IT: 0000 UTC 24 DEC 2010.

From -24h (1200 UTC 26 December) to the valid time (1200 UTC 27 December) in Figures 3.10b and a, the surface negative part of the dipole over the U.S. Gulf coast develops and moves northeastward. During this process, a downstream positive anomaly develops to the east of New England. Meanwhile, at 500 hPa, the negative forward ensemble regression has two minima, with one over the upper level low and the other southwest of the East Coast trough, while the positive forward ensemble regression is first located over the mid-Atlantic coast and then over Nova Scotia (Figures 3.10f, e).

The regressed MSLP and Z500 at the valid time (1200 UTC 27 December) display a shift and intensification of both the surface cyclone and the upper level trough (Figures 3.10a, e). These patterns show strong resemblances to the actual forecast differences (Figures 3.8a, e) between the two forecast cycles, with pattern correlations within the valid box of 0.82 and 0.61 for MSLP and Z500, respectively. These results suggest that the initial time forecast error at 0000 UTC 25 December over the southern stream trough led to a majority of the forecast uncertainties of the surface cyclone and upper level trough positions as well as their amplitudes.

The forward ensemble regression was also calculated using the forecast difference in Z500 over northern Canada (Figure 3.8h, box 2), where the geopotential height in the analysis at 0000 UTC 25 December is lower for the northern part of the ridge from northern Canada to the central U.S. The regression result based on the error within box 2 over northern Canada showed that the initial forecast error spreads eastward across northern Canada and amplifies gradually at both the surface and upper levels (not shown). At -48 h, a positive signal starts developing at 500 hPa to the southeast of this negative signal. Over the next 48 h, this positive signal spreads and amplifies both at the upper level and at the surface. At valid time, the regressed patterns resemble the forecast difference at higher latitudes (north of ~45°N), and contribute to the positive part of the MSLP dipole over the northeastern portion of the boxed region at the forecast valid time (Figure 3.8a).

The above case study illustrated the 3 ways that ESA can be employed in interpreting ensemble forecast uncertainties. The readers are referred to Zheng et al. (2013) for discussions regarding to the utility of ensemble sensitivity analysis in operational ensemble forecasting.

3.4 The verification of ESA signals using LOOCV

The sensitivity calculated using ESA in this study shows the covarying relation between a grid point and the forecast metrics (e.g. EOF PCs). However, it is hard to verify the sensitivity signals and determine how robust they are in modifying the forecast metrics. As an extension of ESA, Gombos et al. (2012) proposed to use multivariate ensemble regression (ER) to investigate the relation between a perturbed domain and a forecast metric. Basically, they use the leading n_p EOF patterns over a prescribed area to predict the leading n_y EOF patterns of a forecast field. They used the ER method to relate the 1000-h potential vorticity forecast for Supertyphoon Separt (2007) with the Z500 field from the Japan Meteorological Agency's 50-member ensemble. Their results clearly showed that the sensitive areas determined by ensemble regression are quite similar with the areas using ensemble correlation. Their work also suggested that the LOOCV can be used to estimate the error of ER.

Given the similarity between the ER and ESA, we will use the LOOCV method to verify the sensitivity areas. The LOOCV will be employed the way as introduced in Chapter 2 to illustrate the robustness of ensemble sensitivity (or ER) signals for multi-model ensemble forecasting in the medium range. The 3- and 6-day forecasts are used to represent the medium range ensemble forecasts in the 90-member multi-model ensemble for East Coast cyclone forecasts. The forecast metrics (*J*) are the PCs corresponding to each of the leading two EOF patterns of ensemble MSLP over the east coast of U.S. and western Atlantic (Region 2, Figure 2.1).

The LOOCV employed here begins by removing one member from the ensemble and define this member as the member of interest m_i. The remaining 89 members are used to calculate the ensemble sensitivity of the forecast metric to the state vectors. The region with maximum area-averaged absolute ensemble sensitivity at earlier forecast (or analysis) time is chosen to perform EOF analysis of Z500. The area of the sensitive area is defined to be a 56°longitude x 36°latitude box at the verification region. Because of the dispersive nature of mid-latitude weather systems, the box is increasing by 4°longitude and latitude per day when going back in time, and reaches 80°longitude x 60°latitude on -6 day and 68°longitude x 48°latitude on -3 day. Note that the sensitive region is not very sensitive to the choice of longitudes/latitudes as long as it contains the largest ensemble sensitivities and is within a >30°longitude and <100°longitude bands. The resultant leading 9 EOF PCs over the sensitive region are used as predictors. Nine PCs are chosen because in the test cases, >10 leading EOF PCs often lead to the overfitting problem while a smaller number of EOF PCs cannot explain enough ensemble variance (<80%). The leading two EOF PCs for the 89 members are computed over the verification region (Region 2, Figure 2.1) at the verification time. Either EOF PC1 or PC2 is chosen as the predictand. The 89-member predictor and predictand are used to establish an ensemble regression relation ($L_{M\text{-}i}$). The corresponding linear regression model is used to predict the value of forecast metric for the member of interest based on the projection of its anomaly onto the leading 9 EOF patterns over the sensitive area. The predicted value for forecast metric J'_i, is compared with the original value J_i, which is a projection of the anomaly for the member of interest to the corresponding EOF pattern over the verification region at the verified time. The above procedure is repeated for each ensemble member, each time leaving out a different member. Finally a predictand forecast scalar vector J'_{1x90}, which can be compared with original vector J_{1x90} used as the forecast metric.

By employing LOOCV, the goodness of ensemble sensitivity can be estimated via the correlation coefficient or the median of the ratio (Wilks 2011) between the ESA-computed predictand field J'_{1x90} and the actual ensemble predictand field J_{1x90}. The ability of ESA to correctly forecast the forecast metric will represent its ability to derive the dynamical relation between the initial perturbed conditions and the forecast anomaly pattern.

Figure 3.11 illustrate one example for the ensemble correlation and median ratio between the predicted and actual forecast metric using EOF PC1. Note that the sensitive areas for each time step are most likely different. The correlation varies from 0.73 to 0.98 while the median ratio varies from 0.69 to 1.02. Both the correlation and median ratio decrease with the increasing lead time with respect to the VT due to the decreasing contemporaneousness between the forecast metrics and the state vectors. The high correlation and high ratio suggest that the ESA is highly capable of capturing the majority of the gross features associated with the leading EOF pattern in this example.

Figure 3.12 compares the cross validated correlations for PC1 and PC2 using another example. For this case, the correlation drops quickly from forecast times of 36 h to 30 h (with lead times of 36 h to 42 h relative to the VT, respectively). Also, the correlation for PC2 is overall smaller than that for PC1, suggesting the ensemble sensitivity relation based on PC1 is more robust in this example.

Figure 3.13 shows the correlation for 3-day forecasts of all 102 cyclone cases. The correlation for both EOF PC1 and PC2 decrease with reduced forecast hours (or increasing lead time relative to VT), suggesting the ESA relations worsen closer to the initial time, which is consistent with Gombos et al. (2012). However, statistically significant correlations are found in most cases (75th percentile for PC1, and >50th percentile for PC2) even at initialization time (a lead time of 72 h), demonstrating the ability of ESA to forecast the gross relation between the state vector and the forecast metrics using PCs. Meanwhile, the same plot for the median ratio at each forecast time (not shown) showed consistent result, with median values >0.2 for PC1, and >0.15 for PC2 at initialization time, showing the forecasted value of forecast metrics in most cases have the same sign with the actual value though the amplitude can be smaller close to initialization time.

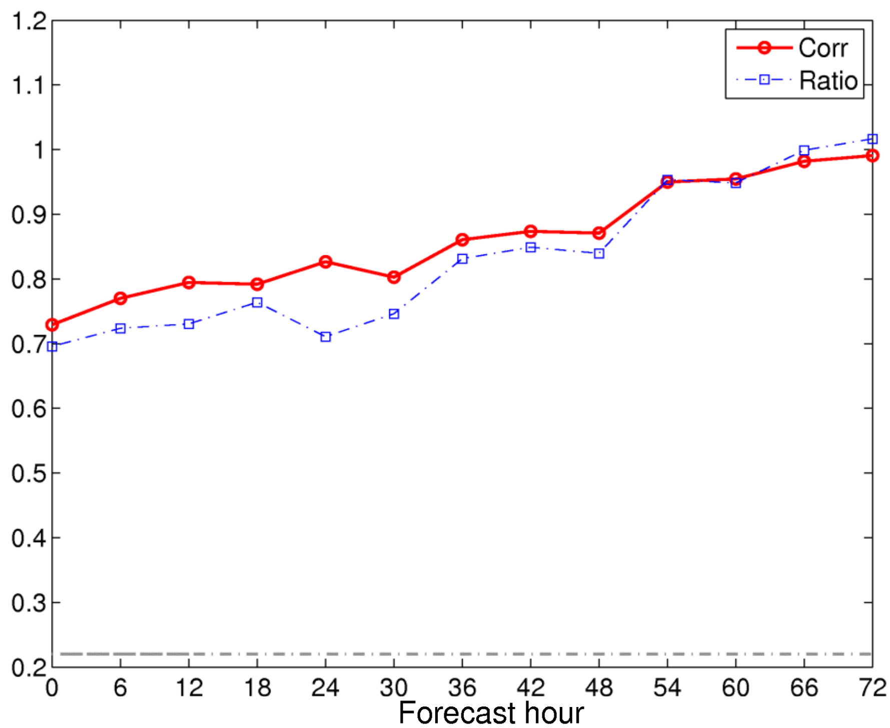


Figure 3.11: One example comparing leave-one-out cross validated correlation (red dotted solid) and median ratio (blue squared dashed) of predicted EOF PC1 and actual EOF PC1 for 3-day forecast initialized at 1200 Dec 09 2008. VT is the 72 h on the plot. Note that the x-axis is the forecast hours, not the lead times for the LOOCV calculations. The lead time can be inferred by subtracting the forecast hour from VT (see text).

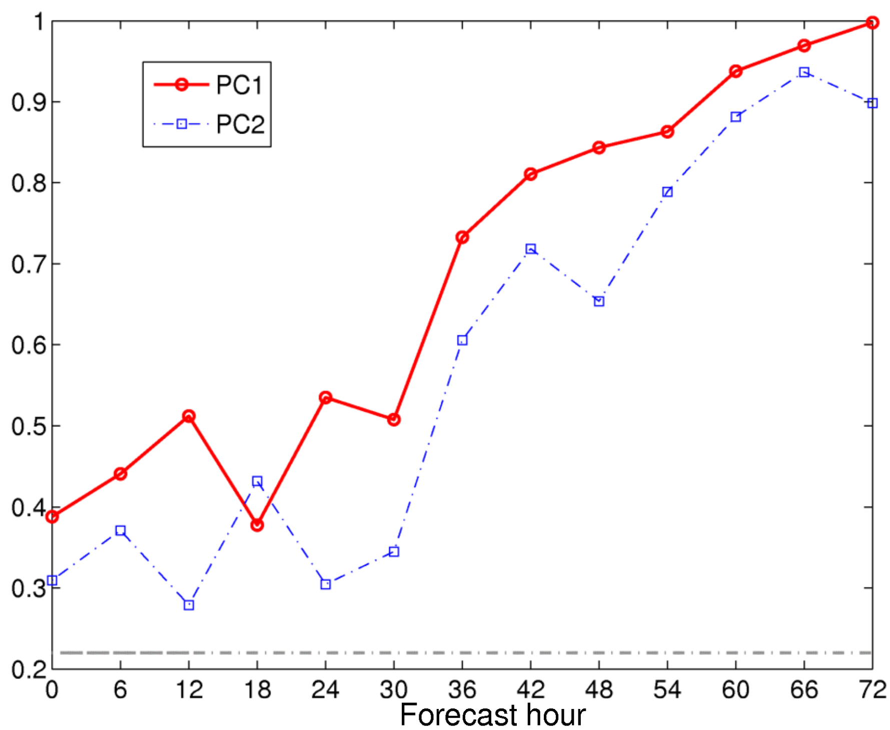


Figure 3.12: Leave-one-out cross validated correlation using EOF PC1 and PC2 for 3-day forecast initialized at 1200 Dec 24 2010. VT is 72 h on the plot. Note that the x-axis is the forecast hours, not the lead times for the LOOCV calculations. The lead ti me can be inferred by subtracting the forecast hour from VT (see text).

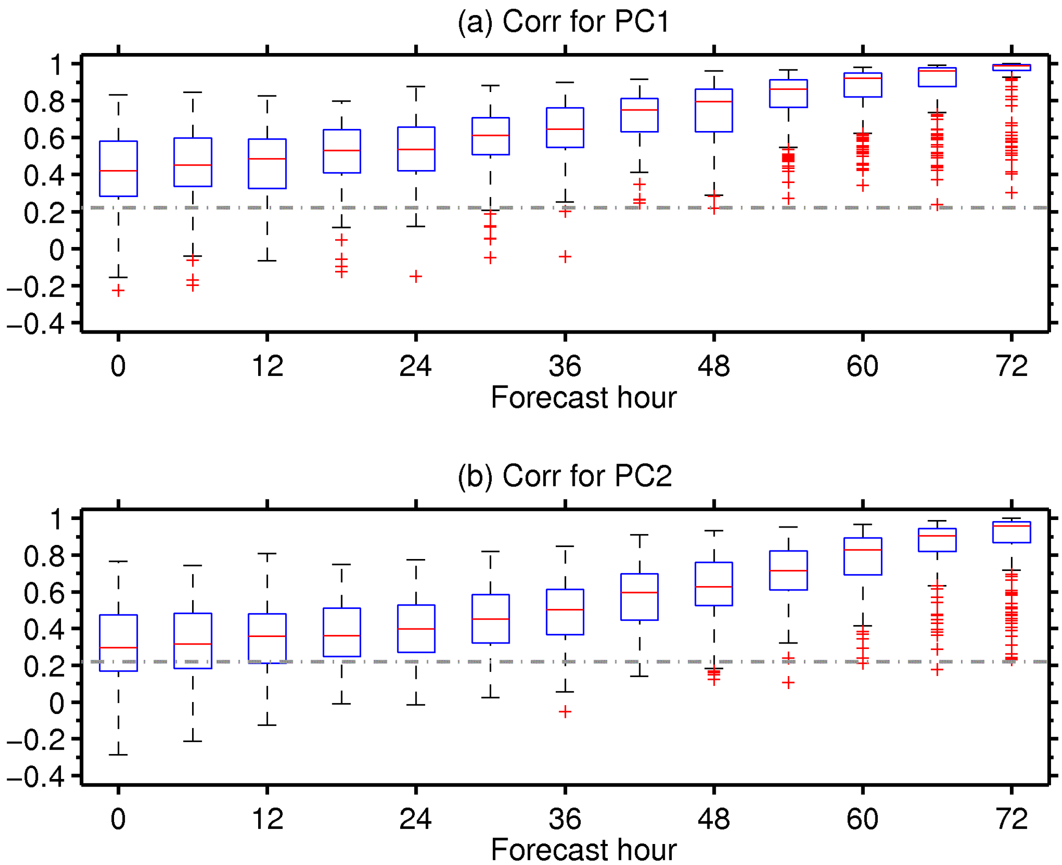


Figure 3.13: Boxplot of leave-one-out cross validated correlation for PC1 (a), and PC2 (b) using 3-day forecast for all 102 verified cyclone cases over region 2. VT is 72 h on the plot. Note that the x-axis is the forecast hours, not the lead times for the LOOCV calculations. The lead time can be inferred by subtracting the forecast hour from VT (see text).

The LOOCV has also been calculated using 6-day forecasts (Figure 3.14). The spread of correlation becomes large prior to 96-h forecast (a lead time of 48 h relative to VT) and the 75th percentile drops to below the significant line at 30 h (a lead time of 114 h) for PC1 and 60 h (a lead time of 84 h) for PC2. Nevertheless, the median correlations are above the significant line through the entire 6-day forecast for PC1, and between day 1 (a lead time of 5 days) and day 6 (VT) for PC2. The median ratios are greater than 0.15 through the 6-day forecast for PC1, and are positive for PC2 through the 6-day forecast.

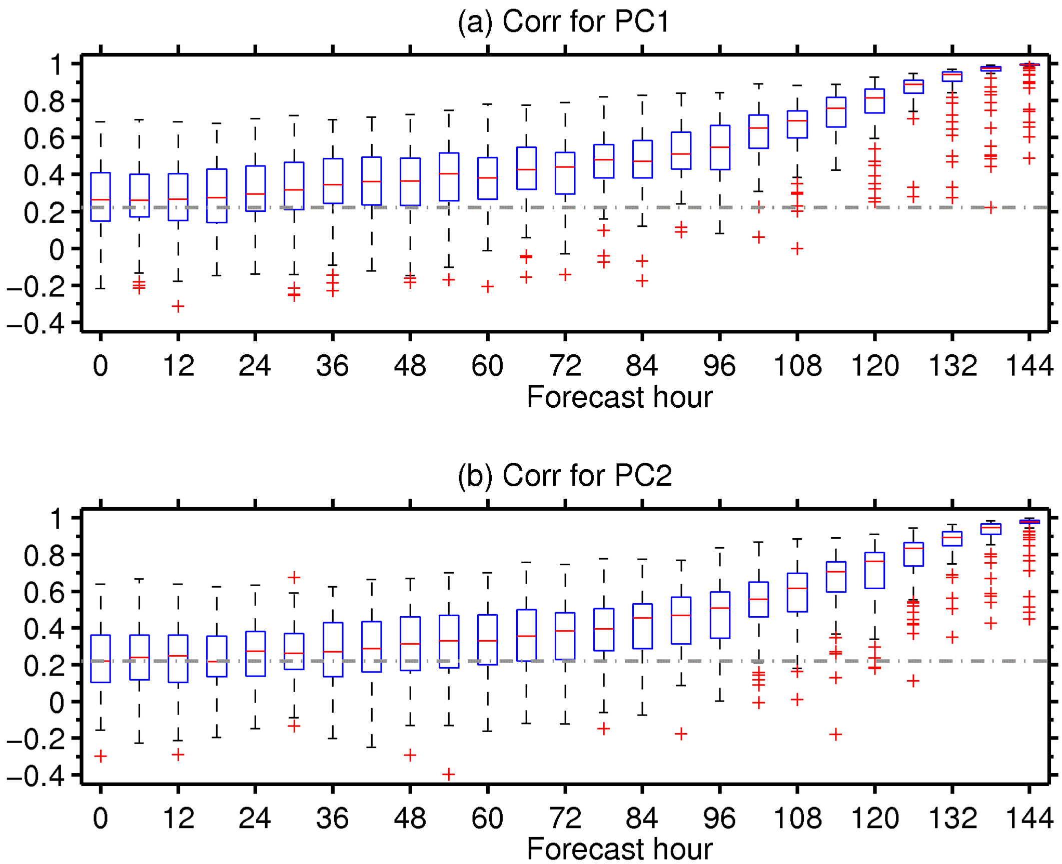


Figure 3.14: The same as Figure 3.13 but for 6-day forecast. VT is 144 h on the plot. Note that the x-axis is the forecast hours, not the lead times for the LOOCV calculations. The lead time can be inferred by subtracting the forecast hour from VT (see text).

Both 3-day and 6-day LOOCV results demonstrated the forecast skill of ESA in the medium range is significant for most of the cyclone cases, although the relations worsen close to the initialization time. Note that the LOOCV here only validates the goodness of using sensitivity to predict the forecast of an independent member of the ensemble. When applying sensitivity to forecast the analyzed field, we expect that the forecast skills will decrease. However, some preliminary results suggest that it can forecast significant correlation and positive ratios for PC1 at least back to 18 h (a lead time of 54 h) for 3-day forecasts, and back to 36 h (a lead time of 108 h) for 6-day forecasts.

3.5 An alternative way of verifying the robustness of medium range ESA signals

Ancell and Hakim (2007) has derived the relation between ETKF and ESA and suggested that in the appropriate norm, ESA is equivalent to the ETKF, but does not require the determination of a transformation matrix, and also allows straightforward inclusion of statistical confidence measures to address sampling error. The difference between two is that ETKF involves the determination of the transformation matrix. As an alternate way to quantify the relation between the forecast metric and the initial conditions, the application of ESA can be compared with ETKF signals as a complementary way to verify sensitivity signals. Therefore, we have selected a few cases to compare the ensemble sensitivity analysis and ETKF method in quantifying the distribution and development of sensitive regions at the initial time and earlier forecast steps associated with forecast uncertainties and errors at the verification time. Here the results for one case, corresponding to the case 15 discussed in Majumdar et al. (2010), will be presented.

The ensemble mean shows a clear cyclone in 850-hPa at verification time (not shown) using a 10-day forecast initialized at 0000 UTC 15 Feb; accordingly, the EOF1 dipole pattern represents a southwestern shift of this cyclone (not shown), which explains 44.2% forecast variance of 850Z.

Ensemble sensitivity of EOF PC1 to 500-hPa Z at different forecast times is shown in Figure 3.15. The shades denote ensemble sensitivity signals, while the contours display the ensemble mean of 500-hPa Z. Here we only show four forecast times: 0 h, -36 h, -72 h and -108 h (Figures 3.15a-d). At -108 h, there are a series of sensitivity signals starting from western India (marked by “X”), crossing south or central Asia (“C+”, “C-1” and “C-2”) and Japan (“B+”) into the north of Bering Sea (“B-1”). At -72 h, sensitivity signals “C+”, “C-2” and “B+” move eastward along synoptic systems. However, signals “X” stays in place. Signals “C-1” and “B-1” fade out. Meanwhile, a few downstream signals (i.e., “B-2”, “A+” and “A-”) are generated. At -36 h, most sensitivity signals move eastward along synoptic systems. However, signals “X” and part of “C+” have stayed around the same place. Upstream signals (i.e., “C+”, “C-2”, and “B+”) either remain the same amplitude or weaken, while downstream signals (i.e., “A+”, “B-” and “A-”) intensify. The easternmost downstream signal “A-” has developed most. At the same time, new signals (“D+” and “D-”) downstream of signal “A-” are generated. At 0 h (verification time), the strongest signals “A-” and “D+” are within the central and eastern U.S., which is associated with the major trough over central U.S. Downstream signal “D-” has also got strengthened. Upstream signals either disappear (“C-2” and “A+”) or become much weaker (“B+” and “B-”). Signals “X” and “C+” remains more or less the same places.

Figure 3.16 shows the ETKF signals, which represents the predicted reduction of wind forecast error variance due to the adaptive observation of u,v,T at three levels (850, 500, and 200 hPa). The largest value suggests the best place for extra observation because these observations will help reducing forecast error most. The black contours are for the ensemble mean of 500-hPa Z. To compare with ensemble sensitivity signals, we will also only show four forecast steps: 0 h (verification time), -36 h, -72 h and -108 h (Figures 3.16a-d). Since ETKF uses an energy norm as forecast metric, the ETKF signals only display positive values. At -108 h, there are four larger signals from west to east marked by “X”, “C”, “B” and “Y”, which are generally consistent with ensemble signals in Figure 3.15d except for signal “Y”. At -72 h, except for the above four signals, there is another large signal (marked by “A”) downstream of signal “B”, which also matches with sensitivity signal “A+” in Figure 3.16c. However, the strong ETKF signal “Y” again doesn’t exist in ensemble sensitivity plots in Figure 3.15c. At -36 h, ETKF signal “X” is sort of stationary. Upstream signals “C” and “B” become weaker, while downstream signal “A” is the strongest signal at this time. There is a new signal “D” downstream of signal “A”. The ETKF signals are generally consistent with the sensitivity signals in Figure 3.16b, except that signal “Y”, which doesn’t exist in the sensitivity plots, has moved to the east of the verification region. At 0 h, upstream ETKF signals “C” and “B” fade out. The most robust signal is the combination of signals “A” and “D” within verification region, which is consistent with the sensitivity signals in Figure 3.15a. However, “D-” signal in Figure 3.15a is not present in the ETKF signals plot.

Based on the results from ensemble sensitivity and ETKF methods, ETKF signals/sensitivity signals (“D”/ “D+”, “A”/ “A+”, “A-”) that develop and propagate into verification region can be shown using both methods. The important feature of downstream development (the upstream signals become weaker while the downstream signals get strengthened) can also be seen in both kinds of plots. Some stationary signals (i.e., “X”) can also be determined using both methods. However, these two methods are different at some forecast times in showing remote signals. For instance, ETKF method suggests a signal “Y” close to verification region at -108 h, -72 h and -32 h, while the sensitivity method doesn’t show this signal. Initial condition experiment is required to distinguish significant signals from spurious signals. One thing worth nothing is that since ESA provides phase information of signals, it can clearly show downstream development related with wave packet while ETKF may sometimes either miss or not show clearly this important feature. Note that the case study here only qualitatively compares the two signals. Further studies need to be done to objectively compare these two methods. The related numerical experiment will also be necessary to further study the robustness of signals identified by these two methods.

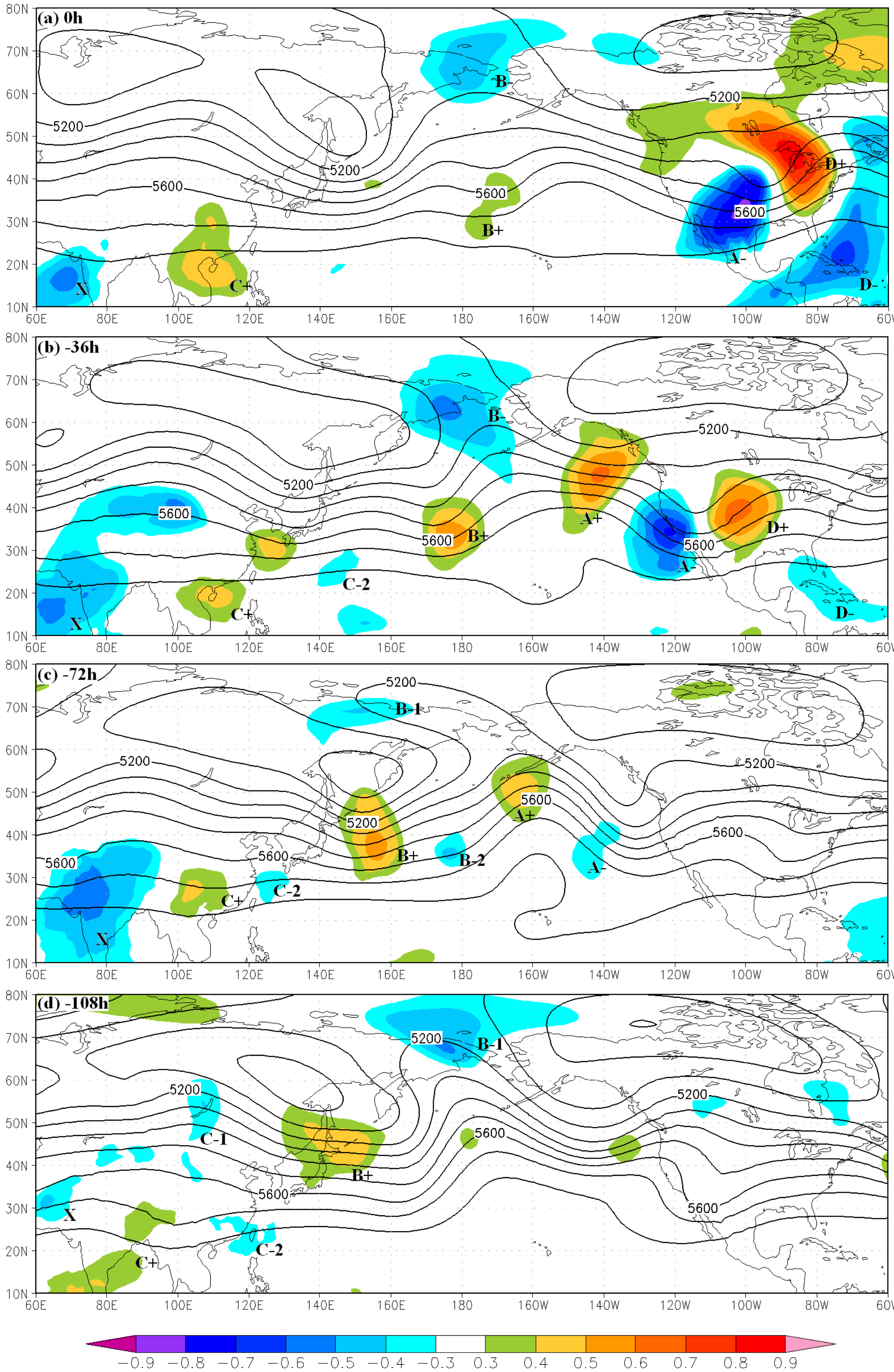

(a) 0h
(b) -36h
(c) -72h
(d) -108h
80N
70N
60N
50N
40N
30N
20N
10N
60E
80E
100E
120E
140E
160E
180
160W
140W
120W
100W
80W
60W
5200
5600
A+
A-
B+
B-
B-1
B-2
C+
C-1
C-2
D+
D-
X
−0.9 −0.8 −0.7 −0.6 −0.5 −0.4 −0.3 0.3 0.4 0.5 0.6 0.7 0.8 0.9

Figure 3.15: Ensemble sensitivity (shades)of EOF PC 1 to 500 hPa Z at verification time (a), -36 h (b), -72 h (c) and -108 h (d). Ensemble forecasts are based on 50 ECMWF members initialized at 0000 UTC 15 FEB 2007. Verification time: 0000 UTC 25 FEB 2007.

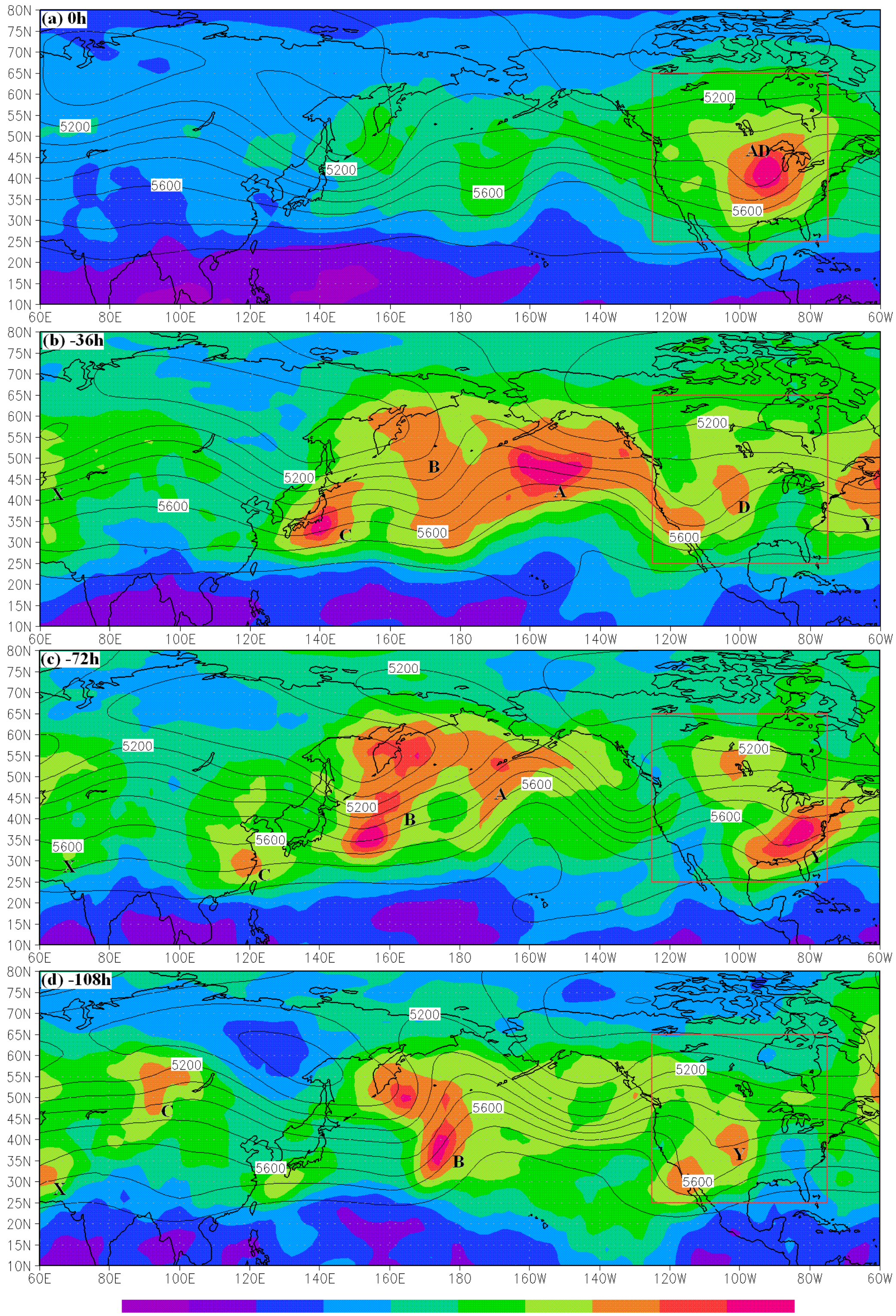


Figure 3.16: ETKF reduction of wind forecast error variance due to adaptive obs of (u,v,T) at850/500/200hPa for verification time (a), -36 h (b), -72 h (c) and -108 h (d). Ensemble forecasts are based on 50 ECMWF members initialized at 0000 UTC 15 FEB 2007. Verification time: 0000 UTC 25 FEB 2007. Orange frame shows verification region.

3.6 Climatology of ensemble sensitivity

Ensemble sensitivity provides a way to investigate the sensitive regions where the perturbations can have significant impacts on the selected forecast metrics. In Section 3.4 it has been demonstrated by LOOCV that the sensitive areas are overall effective in modifying the forecast areas. In this section, we will examine the climatology of ensemble sensitivity signals and investigate different evolutions of sensitivity for different development scenarios of cyclones.

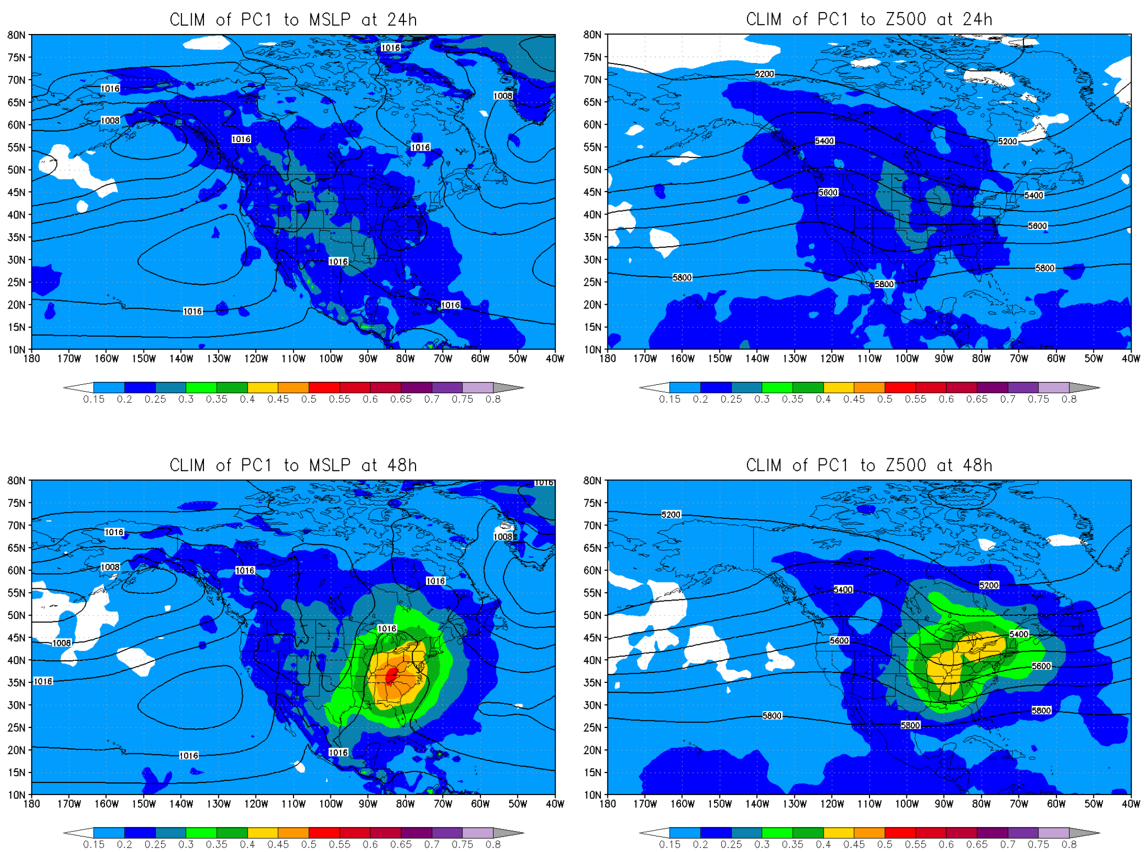
The ensemble data used is a combination of the NCEP, CMC and ECMWF ensemble data from the TIGGE archive. Cyclone cases (102) are defined as minimum pressure <1005 hPa in GFS analysis at verification time over a verification box (32°N-45°N, 78°W-62°W). We chose 3-day and 6-day MSLP and Z500 forecast to represent medium range forecast. The methods include ESA using EOF PCs of MSLP as forecast metrics (equation (3.1)), composite analysis, and Extended EOF method (Wilks 2011).

Sensitivity analysis has proven to efficiently determine the region most significantly associated with a forecast metric (Torn and Hakim 2008; Ancell and Hakim 2007; Chang et al. 2013). Garcies and Homar (2009) applied ESA based on the ECMWF ERA-40 data and derived climatologically sensitivity regions for Mediterranean intense cyclones. This work follows our previous studies on applying ensemble sensitivity analysis to the operational ensemble, and derives the model climatological sensitivity regions for U.S. East Coast cyclones to provide insights into the forecast uncertainty of winter storms in medium range forecast.

Figure 3.17 shows the averaged absolute sensitivity of EOF PC1 of MSLP to MSLP and Z500 based on 3-day forecast for all cyclone cases. At 24h, the surface sensitivity maxima are associated with the eastern and the southeastern side of a surface high over the Rocky Mountains. Meanwhile, the upper-level sensitivities are mainly over the trough downstream of the ridge over the west coast of the North America. The initial condition uncertainty in MSLP could be partly due to the mountain topography. At 48 h, the sensitivity to MSLP increased from <0.3 at 24h to 0.5 on the surface, and tended to be maximized with the development of the surface low pressure area over the Ohio Valley, as well as the upper level trough over the central part of the U.S. From 48 h to 72 h, the composite surface low developed and moved northeastward across the South-Atlantic coast. It had a minimum pressure <1004hPa and centered at (40°N, 69°W). The maximum surface sensitivity followed the surface cyclone center and reached 0.8 at 72 h. The sensitivity for Z500 followed the aforementioned trough as well as its downstream ridge from the west and reached 0.6 at 72 h. Therefore, the sensitivity at verification time (72 h) is more like a monopole centered with the averaged cyclone, suggesting the dominant uncertainty pattern for EOF1 sensitivity is associated with the cyclone intensity uncertainty. The Z500 characteristics suggest that practically in all cases the corresponding upper-level trough and its downstream ridge are important to the forecast uncertainty in cyclone intensity, which originates from the ridge along western coast of North America.

Figure 3.18 shows the averaged sensitivity for EOF PC2. At 24 h, the averaged sensitivities are smaller than using EOF PC1.The upper-level sensitivity is associated with the central part of the ridge over the west coast of North America and its downstream trough. At

48h, the surface sensitivity showed two maxima associated with the low area over the Ohio Valley, with one over its southwest and the other to its northeast. This indicated the sensitivity could represent the shifting of this surface low system. Meanwhile, the upper-level sensitivity also showed two maxima, one over the western Ohio Valley and the other over the east of the Great Lakes. From 48h to 72h, the surface sensitivity developed and continued to have two maxima, one over the west-southwest and the other over the east-northeast of the surface cyclone. At the upper level, the sensitivity formed one center over the east of the East Coast trough. The surface and upper-level structure suggested that the sensitivity for EOF2 is associated with the uncertainty in cyclone position and its upper-level trough position.



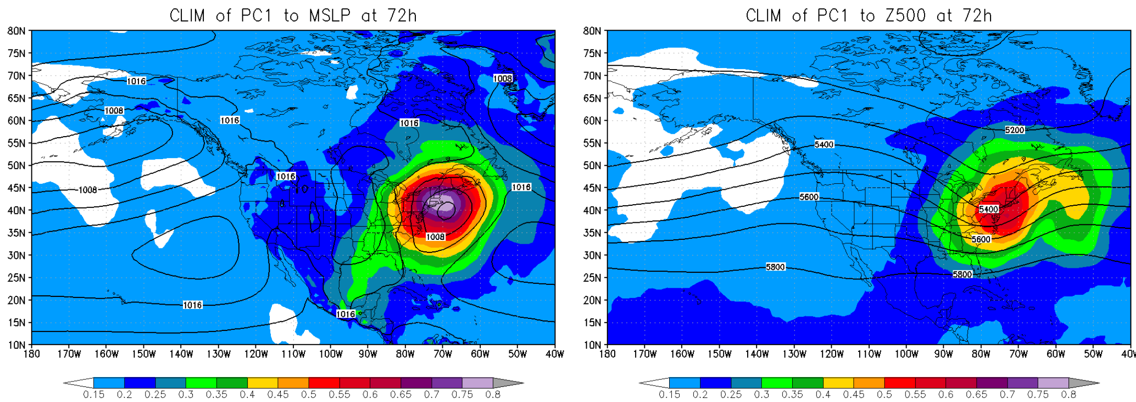
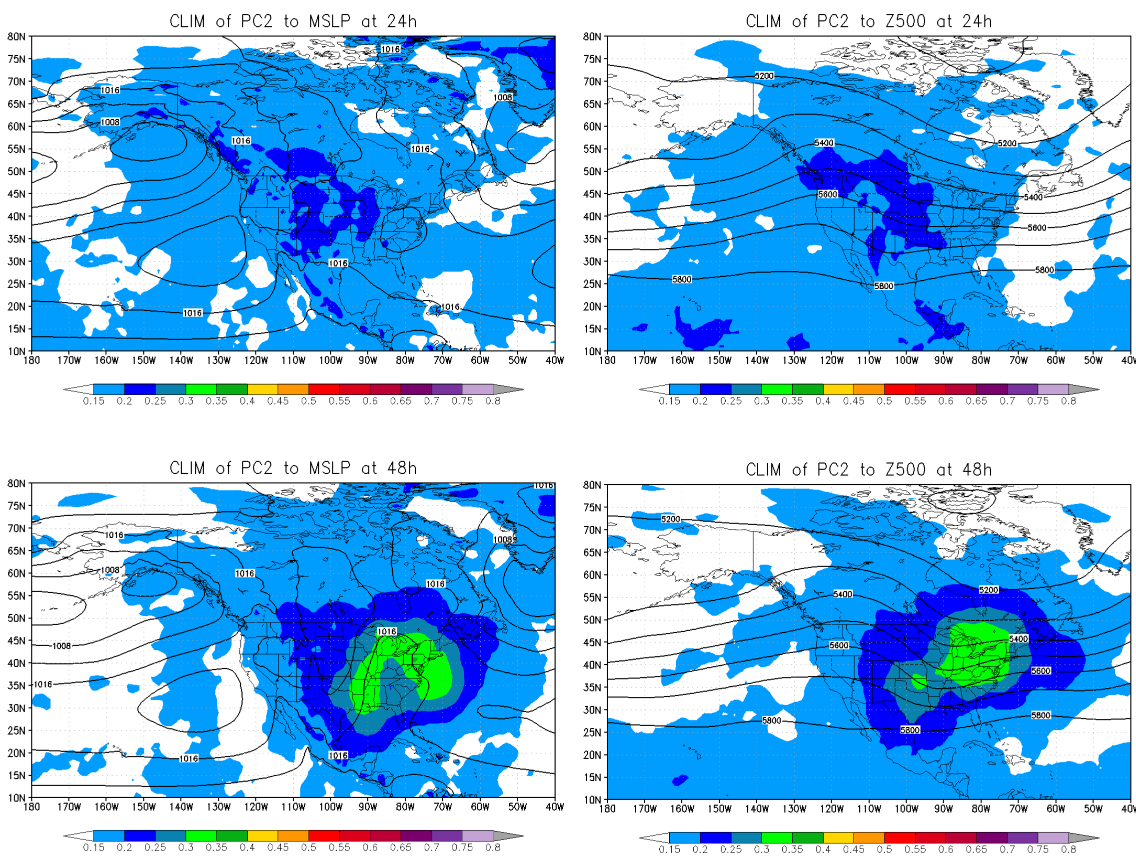


Figure 3.17: Mean of absolute ensemble sensitivity using EOF PC1 for MSLP (left, shaded) and Z500 (right, shaded) over 102 cyclone cases, and ensemble mean MSLP (left, black contours) and Z500 (right, black contours) at different lead times. The verification time is 72h.



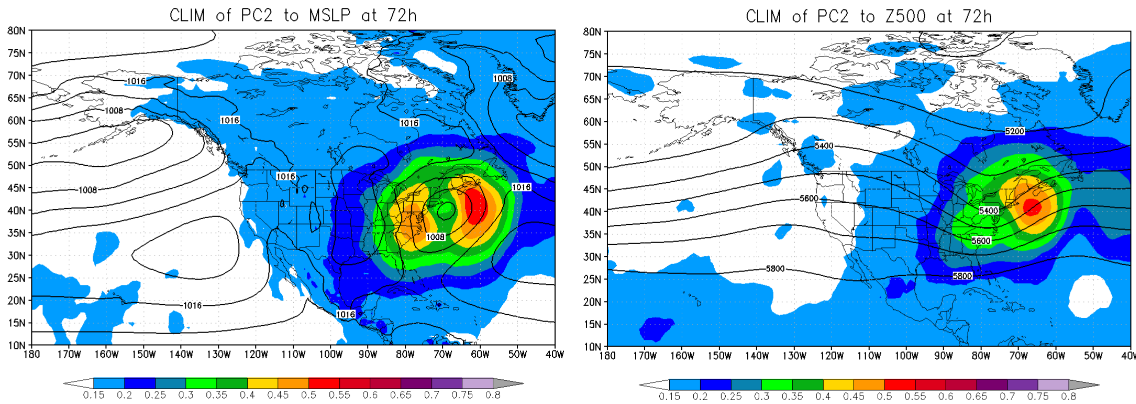
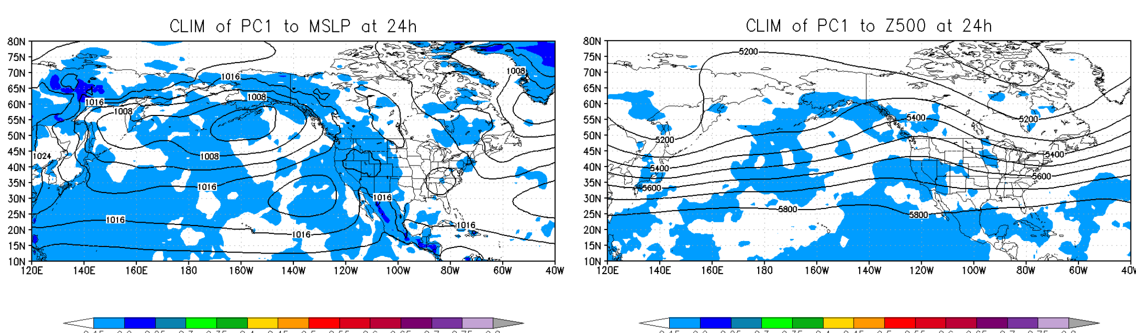


Figure 3.18: Mean of absolute ensemble sensitivity using EOF PC2 for MSLP (left, shaded) and Z500 (right, shaded) over 102 cyclone cases and ensemble mean MSLP (left, black contours) and Z500 (right, black contours) at different forecast times. The verification time is 72h.

The climatological sensitivity for 6-day forecast is also examined. Figure 3.19 shows the 6-day averaged absolute forecast sensitivity using EOF PC1 of MSLP on day 6 (144 h). Before 96 h, the ensemble sensitivity was mainly over the North and East Pacific oceans, associated with the Aleutian low over the northeastern Pacific. From 96 h to 144 h, the surface sensitivity was associated with the cyclone originating from the U.S. South, increasing from <0.3 at 96 h to >0.8 at 144 h. The surface sensitivity was mainly a monopole at 144 h, indicating that it was associated with the cyclone intensity uncertainty. Meanwhile, the upper-level sensitivity showed a dipole around the East Coast trough and its downstream ridge at 144 h, again suggesting the strength of this trough and its downstream ridge have an impact on the surface cyclone intensity forecast.



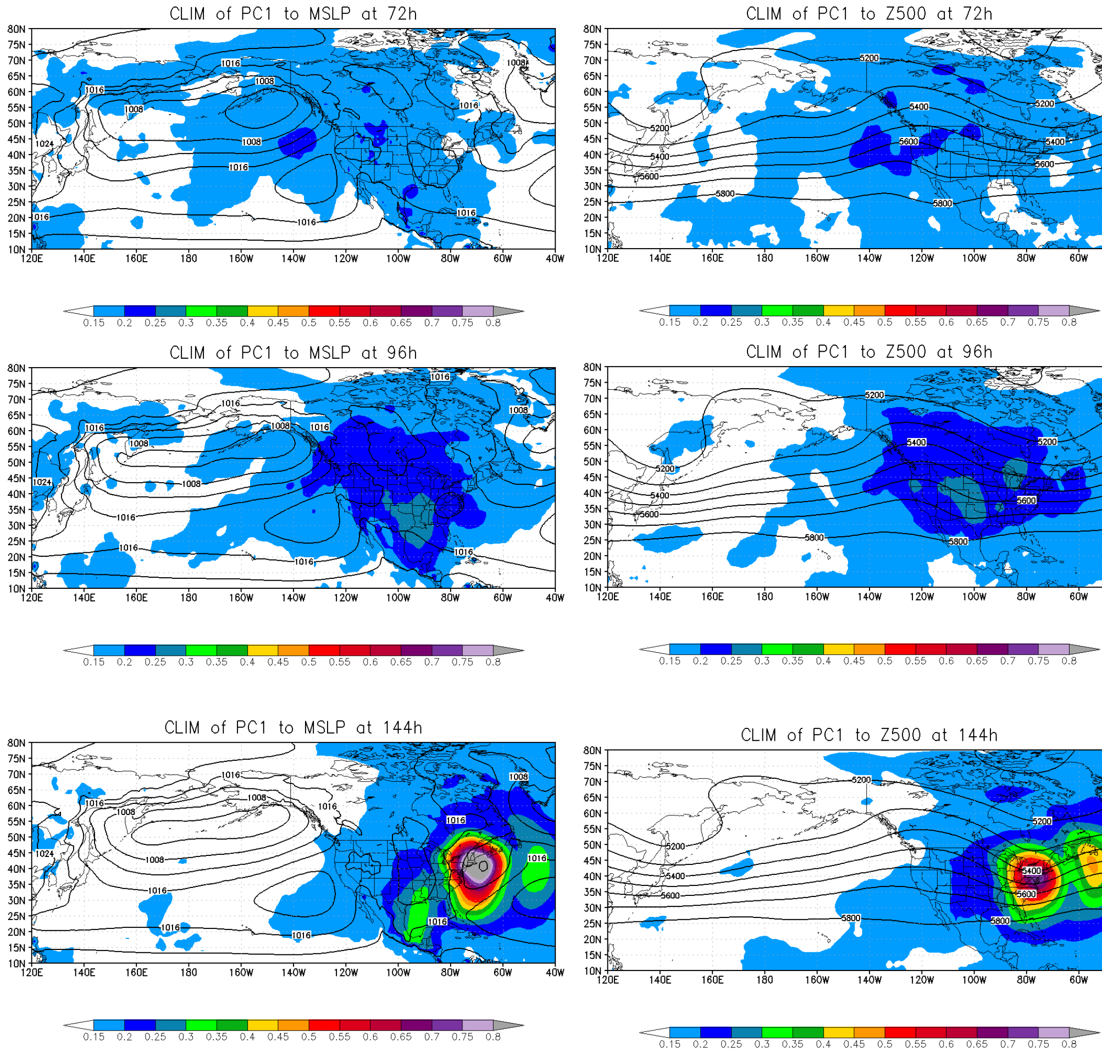


Figure 3.19: The same as Figure 3.17 but for 6-day forecast with the verification times at 144h.

The ensemble sensitivity for EOF PC2 (Figure 3.20) shows scattered signals at 24 h associated with short wave troughs over central Pacific at both surface and upper levels. These sensitivity signals seemed to propagate eastward across the west coast of the U.S. at 72 h. From 96 h to 144 h, the surface sensitivity signals showed a dipole associated with the surface cyclone, suggesting the EOF2 pattern is on average associated with the forecast uncertainty in cyclone positon. Meanwhile, the upper-level sensitivity showed a triple-center structure, suggesting the shift of the East Coast trough and its adjacent systems could impact the cyclone position forecast uncertainty.

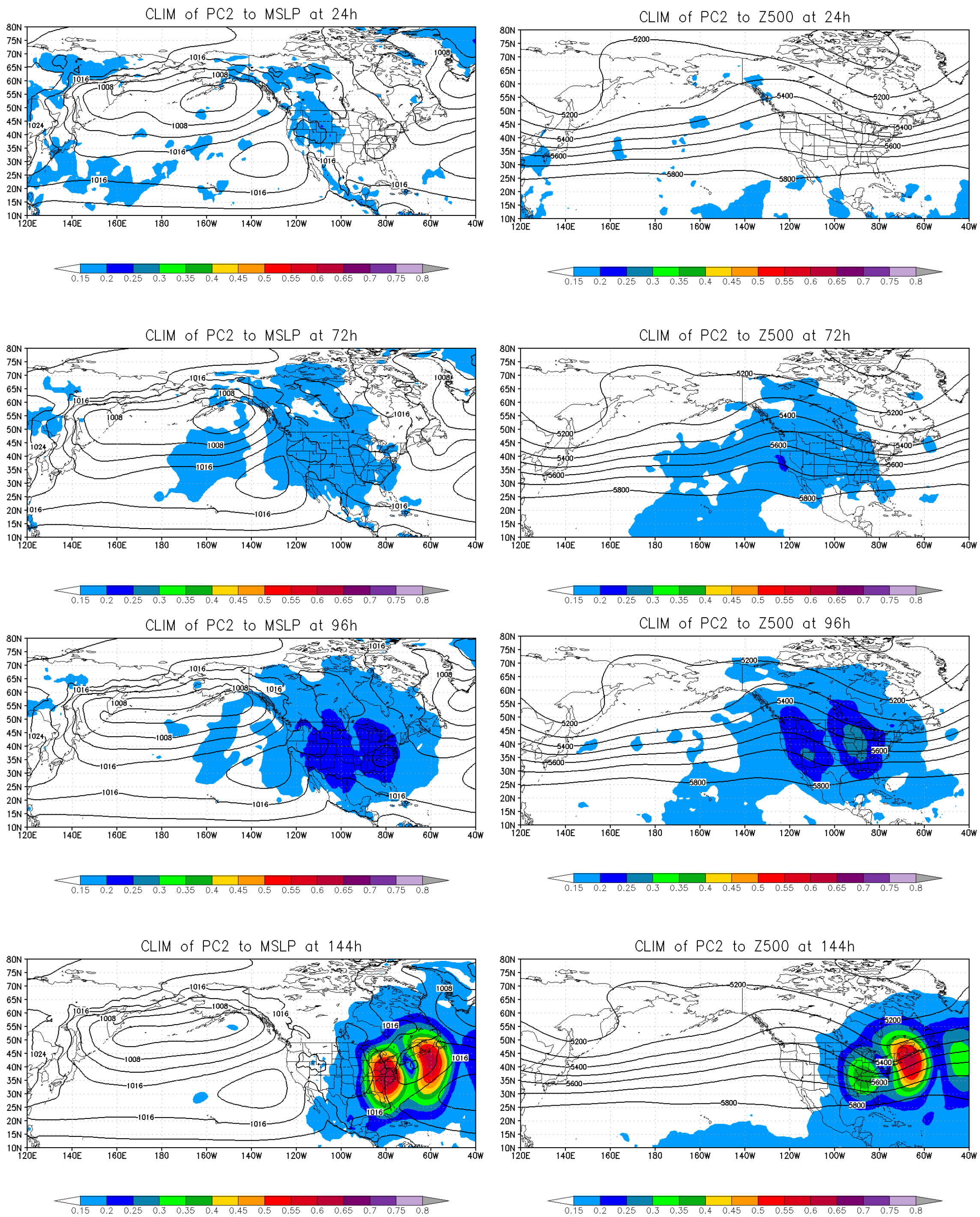


Figure 3.20: The same as Figure 3.19 but for EOF PC2.

To investigate the propagation of medium-range ensemble sensitivity signals, the absolute sensitivity to Z500 was averaged from 30°N to 60°N and plotted against forecast time using 6-day forecast as a Hovmoller plot (Hovmoller 1949). At verification time, the sensitivity for PC1 was maximized around 77.5°W (282.5°E on Figure 3.21). It spreads westward and decreases in amplitude when going backward with time. On day 1, the sensitivity value dropped to 0.15, and showed two weaker centers: one is around 175°W and the other around 122.5W. So the propagation speed could range between ~10° and ~20° longitude day^{-1}, which is typically a bit faster than the synoptic weather system and slower than the group speed of Rossby wave packets. Note that this result is an average over cases and latitudes, indicating the propagation speed can vary with different cyclone case.

The Hovmoller plot of sensitivity for PC2 is shown in Figure 3.22. When compared with Figure 3.21, the sensitivity for PC2 overall is weaker than for PC1, which dropped to 0.15 on day 2. Another difference is that the averaged sensitivity on day 6 has three centers, which is consistent with its triple-center structure in Figure 3.20, again indicating the shift of Z500 trough and its adjacent systems. Until day 3.5, the division of centers is still clear to observe. The sensitivity center shifted from 70°W on day 6 to around 130°W on day 2 (or 162.5°W on day 2.5); thus the propagation speed is ~15° longitude day^{-1}.

To investigate the horizontal scale of ensemble sensitivity, the sensitivity at a central longitude is used for each forecast day to correlate with the sensitivity at all longitudes. Figure 3.23 showed the correlation for each central longitude on each day. A 0.5 horizontal line is also plotted as a reference line to denote where the correlation drops to half. The width at which reference line crosses the sensitivity correlation curve is used to roughly represent the coherent sensitivity signal centered at the central longitude. On day 6 and day 5 (black line and green line), the sensitivity scale is around 15.5°longitudes. However, it increases to ~20°longitudes on day 4, ~38.9°longitude on day 3, ~44.4°longitude on day 2, and ~61.1° longitude on day 1. The increase in width is due to the dispersive nature of the middle latitude waves.

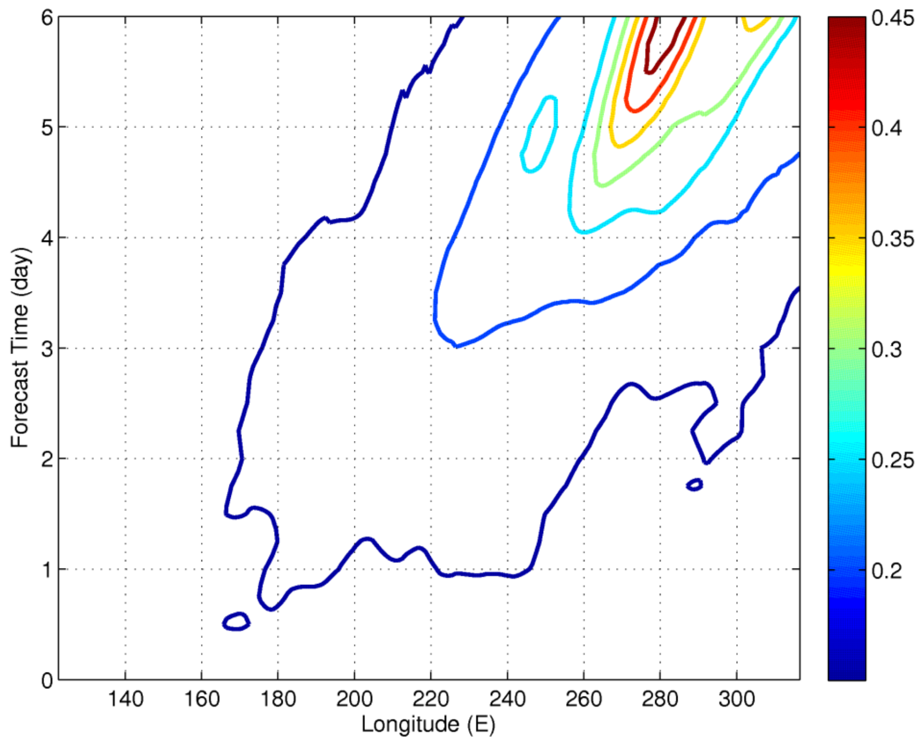


Figure 3.21: ESA using PC1 averaged in the mid-latitudes from day 0 to day 6.

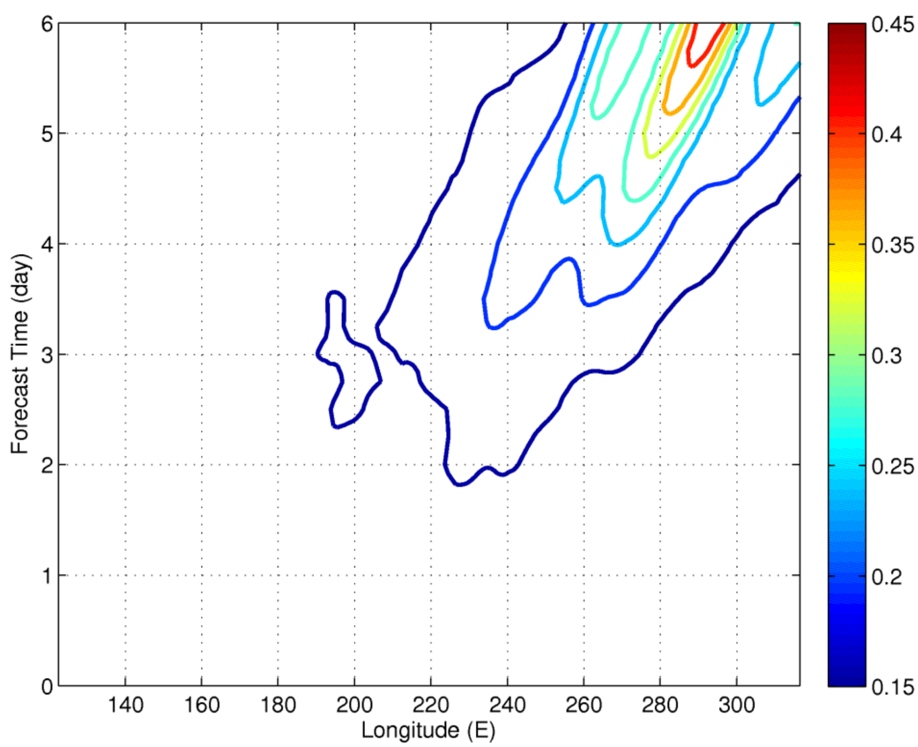


Figure 3.22: The same as Figure 3.21 but for PC2.

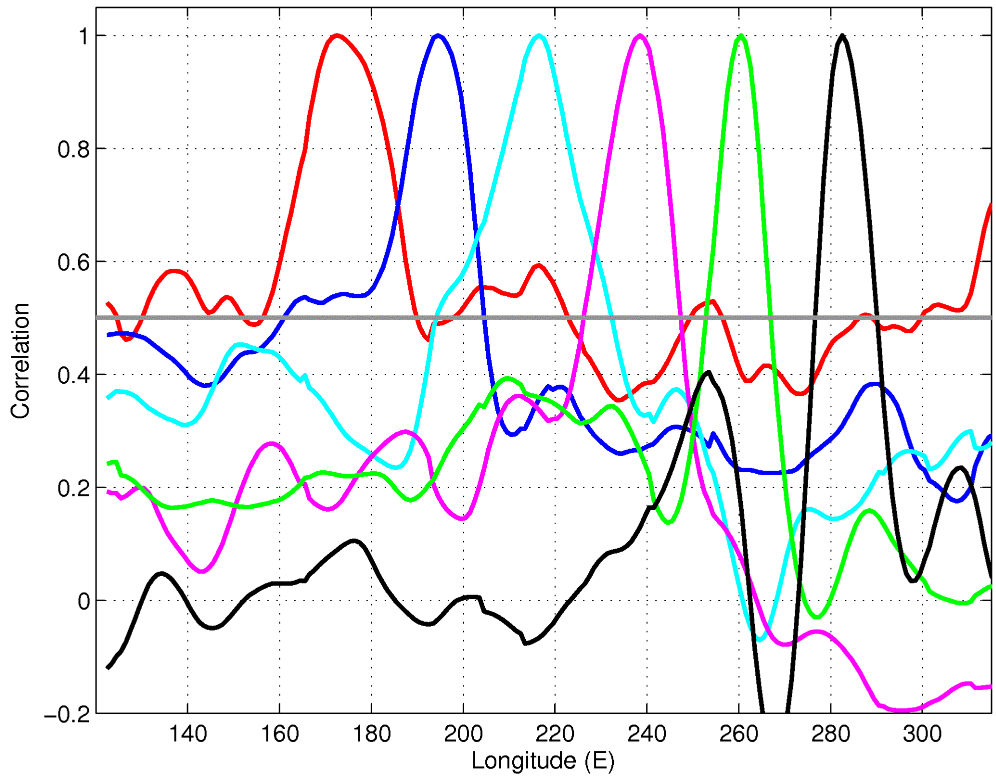


Figure 3.23: Correlation of sensitivity for PC1 at the central longitude of each day with sensitivity average over other longitude. Red, blue, cyan, magenta, green and blue represent day 1 to day 6 average. The central longitudes are from 170°E to 282.5°E with an increase of 22.5° longitude.

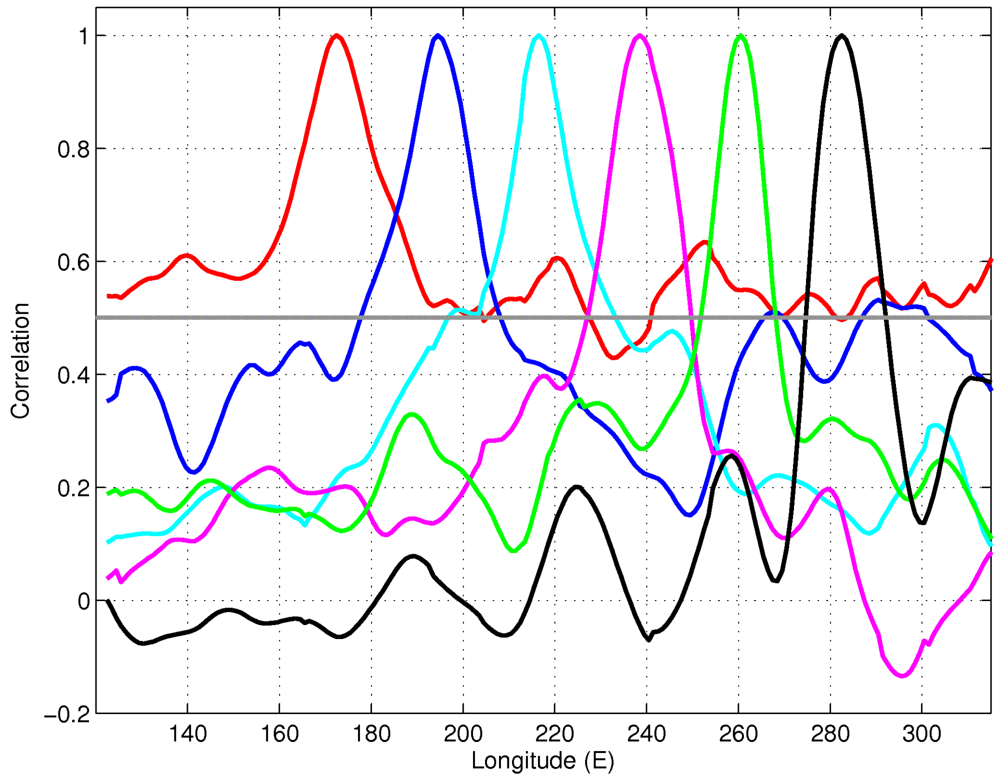


Figure 3.24: The same as Figure 3.23 but for sensitivity using PC2.

The width for sensitivity using PC2 (Black line, Figure 3.24) is 19.5° longitude at verification time, which is a bit larger than using PC1 and consistent with the triple-center pattern. It decreases a little bit on day 5 to 19° longitude, then increases to 24.4° longitude on day 4, 33.3° longitude on day 3 and day 2, and >88.5° longitude on day 1. Note that the correlations for day 1 are very high, which may be due to the spurious sensitivities close to initial times. Nevertheless, the increasing with decreasing forecast time does indicate the dispersive features of midlatitude atmospheric waves.

3.7 Possible link to RWPs via. Downstream development mechanism

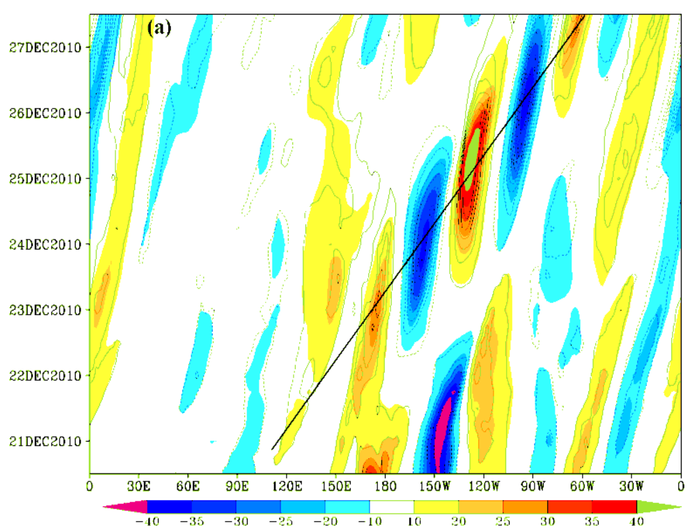
As discussed in Section 3.3, in the sensitivity plots using EOF PC1 based on 5.5-day forecast in December 2010 case, some indications of the wave packet like behavior can be seen in the sensitivity signals (Figure 3.5a-h). Chang (1993) showed that RWPs can be better diagnosed using the meridional wind perturbations rather than the geopotential height, because the variations in geopotential height are dominated by low frequency variability. To examine whether the ensemble sensitivity signals for medium range forecast are associated with RWPs, the ECMWF ensemble forecast starting on 1200 UTC 20 December was used to produce a Hovmoller diagram (longitude-time plot) of the 300-hPa meridional wind averaged between 30°N and 60°N (Figure 3.25a). A clear wave packet can be seen propagating eastward from the central North Pacific on 21 December to the western Atlantic on 27 December, with a group speed of around 29°longitude per day. The wave packet can be found in both the analysis (black lines) and the ensemble forecast (shades).

For comparisons, a Hovmoller diagram of the averaged ensemble sensitivity signals computed based on EOF PC1 of meridional wind over a boxed region corresponding to the surface cyclone (Figure 3.26, the magenta box) is shown in Figure 3.25b. The ensemble sensitivity has a wave-packet-like structure, which can be traced back to the central Pacific on day -6 (21 December). Even though the sensitivity signals are not coincident with the Rossby wave packet regarding the structure and phases, both show clear wave packet behavior and comparable group speed.

Apart from the above similarity, the sensitivity signals also exhibit clear characteristics of downstream development (Chang 1993), which can be seen not only in Figure 3.25b but also in Figure 3.26, which shows the sensitivity of EOF PC1 of meridional wind uncertainties over the boxed region to the meridional wind field. At -84 h, there is a small piece of positive sensitivity A+ located at (160°W, 36°N) and negative sensitivity A- at (142°W, 40°N). Another even smaller positive signal B+ is around (125°W, 45°N). All signals propagate eastward with time. At -60 h, the upstream signal A+ keeps the same intensity, but A- and B+ have strengthened.

New signals B-, C+ and D+ are generated over (98°W, 47°N), (98°W, 58°N) and (80°W, 44°N). At -36 h, the A+ signal becomes weaker; the A- remains the same intensity; while the B+, B-, C+, and D+ have further strengthened. At -12 h, the upstream A+ decays to pieces and A- becomes much weaker, the B+, B-, and C+ have strengthened moderately; the downstream signal D+ gets much stronger. Overall, this evolution resembles the behavior of an RWP's downstream development.

Considering the close link between ensemble sensitivity signals and RWPs, we speculate that the growth and propagation characteristics of the forecast uncertainties and errors are associated with the growth and propagation of Rossby wave packets with the presence of an RWP. Chapter 5 will discuss more on the influence of Rossby wave packets on forecast errors/uncertainties.



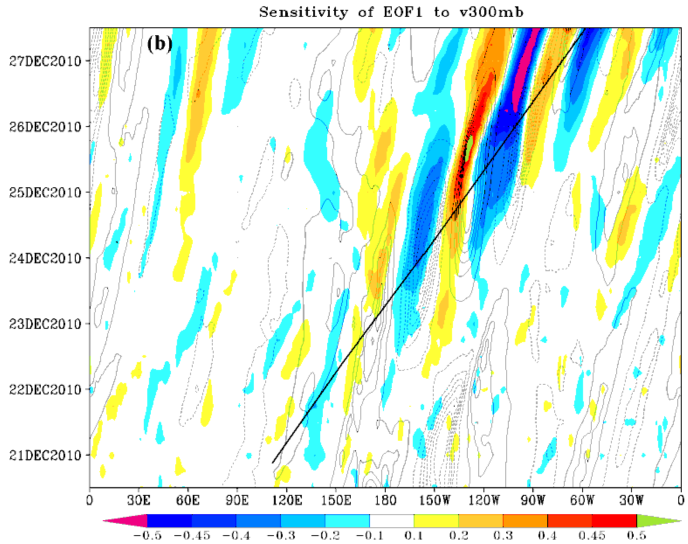
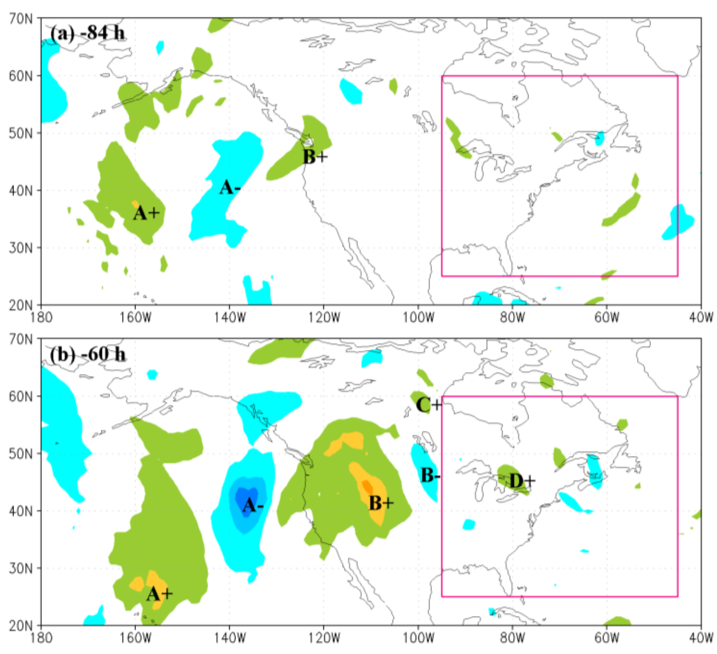


Figure 3.25: (a) Hovmoller diagram of ensemble mean forecast of meridional wind (shading, [m s-1]); (b) sensitivity signals (shading) averaged between 30°N and 60°N; IT: 1200 UTC 20 DEC 2010; sensitivity forecast metric is PC1 of the meridinal wind over a boxed region (magenta box in Fig.3.26). Black contours are for Hovmoller diagram of analyzed meridional wind averaged over 30°N to 60°N with an interval of 10 m s^{-1}. Thick black line shows the movement of analyzed Rossby wave packets.



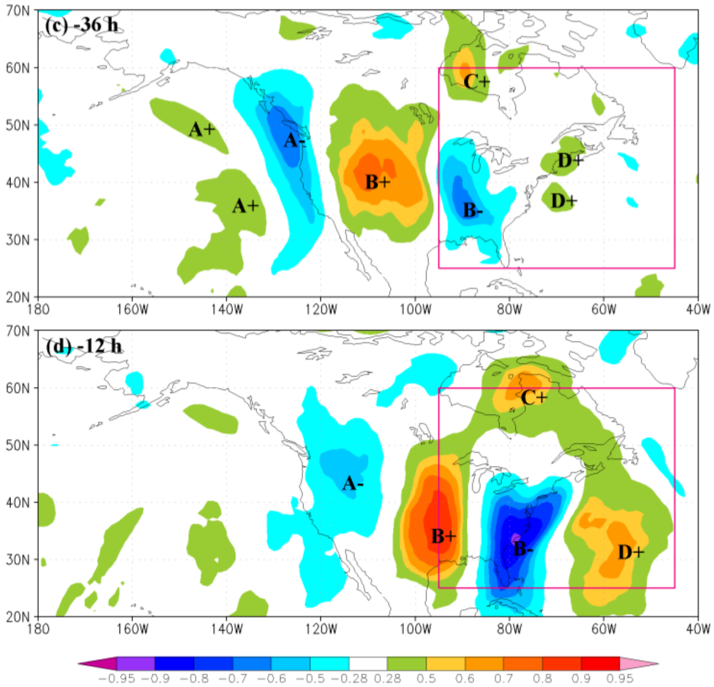


Figure 3.26: Sensitivity (shading) of EOF PC1 to meridional wind from -84 h to -12 h (a-d). The magenta Box shows the verification region for EOF pattern 1. IT: 1200 UTC 20 DEC 2010; VT (0 h): 1200 UTC 27 DEC 2010.

3.8 Chapter summary

In this chapter, the application of ESA to interpret medium-range forecast uncertainties of East Coast cyclone based on a multi-model ensemble introduced in Chapter 2 has been presented.

ESA has been applied in three different ways to diagnose a high-impact storm in December 2010 winter snowstorm based on the choice of forecast metrics: sensitivity using the EOF approach, sensitivity using run cycle difference, and forward sensitivity regression using short-range forecast errors. The 5.5-day forecast cyclone intensity uncertainty in the ECMWF ensemble was associated with the trough and ridge systems over the northeastern Pacific and central U.S., respectively, while the track uncertainty is associated with a short-wave trough over the southern Great Plains. Sensitivity based on the ensemble mean MSLP difference between two run cycles also suggests that the track's shift between the two cycles was linked to the initial errors in the short-wave trough over the southern Great Plains. The forward sensitivity approach based on short-range forecast errors further confirms that the short-term error associated with the short-wave trough over the southern plains developed and contributed significantly to the shift in cyclone position between two cycles.

The LOOCV is used to verify the goodness of ESA. Both 3-day and 6-day LOOCV results demonstrated the forecast skill of ESA in the medium range is significant for most of the cyclone cases, although the relations worsen as the forecast time decreases. These results show that the sensitive areas are overall effective in modifying the forecast areas. The sensitivity using the leading two EOF PCs can be quite reliable in medium-range.

The model climatology of ESA is displayed in section 3.6. It is shown that the ensemble sensitivities at the verification time for both 3-day and 6-day forecast are associated with the cyclone intensity forecast for PC1 and the cyclone position shift for PC2. For PC1, the Z500 characteristics suggest that practically in all cases the corresponding upper-level trough and its downstream ridge are important to the forecast uncertainty in cyclone intensity, which originates from the ridge along western coast of North America. In contrast, the upper level sensitivity for PC2 showed a triple-center structure, suggesting the shift of the East Coast trough and its adjacent systems could impact the cyclone position forecast uncertainty.

In the case study, a coherent RWP originated from the central North Pacific 6 days before this snowstorm event. The sensitivity signals behave like a wave packet and exhibit the same group velocity of ~29° longitude per day, indicating that RWPs may have also amplified uncertainty in both the cyclone amplitude and track forecast. The Hovmoller plot of ensemble sensitivity climatology also suggested that the propagation speed of sensitivities for both PC1 and PC2 are a bit faster than the phase speed of individual weather systems, indicating the propagating RWPs could have an impact on the evolution of ensemble sensitivity signals.

The model climatology of ensemble sensitivities provides guidance for operational forecasters to interpret the preferable regions for forecast uncertainty of significant East Coast cyclone events. Though the exact location of maximum sensitivity varies case by case, it could still benefit the forecasters to build a conceptual model of forecast uncertainty development. More importantly, the ensemble sensitivity climatology highlights the regions where the model disagreement will impact the forecast most, and a further exploration of the reasons for the formation of the sensitive areas might help modelers to diagnose the models' deficiency in forecasting East Coast winter storms.

This page is left to be blank intentionally.

Chapter 4 Fuzzy clustering and evaluations of multi-model ensemble

The ensemble sensitivity tool discussed in Chapter 3 provides an overall linear relation to diagnose forecast uncertainty related to a particular feature. It works better in quasi-linear flow. However, given the chaotic and nonlinear nature of the atmosphere, the linear relation represented by ensemble sensitivity sometimes may miss part of the important information in the ensemble forecast. Meanwhile, the detailed development scenarios among an ensemble can also be important especially for high-impact East Coast winter storm events. Given a large dataset from an ensemble, effective tools to extract the ensemble information in a form that can provide the most important development scenarios could benefit the operational forecasters and decision makers. The development of different scenarios can be due to various ICs in different ensemble members. An investigation of these different development scenarios will also benefit modelers when diagnosing model deficiency in order to improve forecast skill.

To quickly extract important information from a large ensemble and diagnose forecast uncertainty, this section will apply an EOF/fuzzy clustering tool to operational ensemble forecasts of East Coast winter storms. Forecast scenarios will be separated by dividing a multi-model ensemble (NCEP+CMC+ECMWF) into *N* clusters, with each cluster representing one scenario. A scenario-based ensemble verification method will also be proposed to diagnose the cyclone forecasts in these three EPSs as a novel way to evaluate ensemble performance.

Section 4.1 will briefly introduce some relevant previous studies on the EOF/fuzzy clustering method. Zheng et al. (2016) applies the EOF/fuzzy clustering method to an extratropical cyclone case, which brought high-impact weather to the U.S. East Coast on 26–28 January 2015. Section 4.3 will compare the cluster results using different forecast metrics for the aforementioned case. Section 4.4 introduces the procedure to apply EOF/fuzzy clustering to verify ensemble forecasting. Section 4.5 presents the scenario-based verification of the three EPSs in forecasting East Coast winter storms. Section 4.6 investigates the outlier cases determined from the EOF PC phase space. Section 4.7 provides the chapter summary.

4.1 A brief introduction to fuzzy clustering method in ensemble studies

Cluster analysis can be used to diagnose the variability among ensemble members for the short-range (0.5-2.5 day, Johnson et al. 2011), medium-range (Ferranti and Corti 2011) and extended-range (beyond 10 days; Palmer et al. 1990). Brill et al. (2015) introduced a new divisive clustering algorithm to medium-range forecasting based on the one-dimensional discrete Fourier transformation. The application of this method can provide meteorologically coherent pictures of outcomes not captured sharply by ensemble means or deterministic model output. However, this approach often excludes around a dozen of the combined 70 ensemble members by not including them in any cluster. Besides, this method reduces the geopotential data to one spatial dimension in large scale, which may miss a lot of two-dimensional information in complex weather systems. To take full advantage of the ensemble product and provide consistent guidance for National Centers and Weather Forecast Offices (WFOs), the U.S. National Weather Service (NWS) has developed a National Blend of Models (NBM) by postprocessing and combining (blending) guidance from models and EPSs from multiple centers (Gilbert et al. 2016; Tew et al. 2016). To get the best possible science, it is important to document the variability among different EPSs, and to evaluate the benefit of combining models especially in the sense of providing a sufficient variety of forecast scenarios.

Harr et al. (2008) introduced the fuzzy clustering method to identify groupings of forecast scenarios during the extratropical transitioning (ET) of tropical cyclones within a collection of ensemble members. Keller et al. (2011) compared the forecast scenarios associated with 10 ET cases in the TIGGE (Bougeault et al. 2010) data by applying fuzzy clustering analysis and found that some EPSs are confined to a few scenarios while others contribute to almost all scenarios. Other studies (e.g. Grams et al. 2011) have also shown that fuzzy clustering is a suitable diagnostic method to detect the physical processes associated with different weather systems. In this study, the fuzzy clustering method of Harr et al. (2008) is applied to winter storms cases over the U.S. East Coast to separate forecast scenarios in an objective and efficient way and to assess ensemble forecast output, with a goal of providing guidance to forecasters to improve the understanding of scenarios as well as interpreting model biases.

4.2 Applying fuzzy clustering to winter storm forecast in a case study: 26-28 January 2015 winter storm event

The mid-Atlantic to Northeast major winter storm during January 26─28, 2015 is used to illustrate the application of fuzzy clustering. This storm impacted the northeastern U.S., resulting in snowfall accumulations of 30 cm (12 in.) to 91 cm (36 in.) over the central part of Massachusetts, and blizzard conditions were prevalent from Long Island to southern and eastern New England (Winkler 2015).

4.2.1 Synoptic overview

At 1200 UTC 25 January, a surface low pressure system was near the Iowa and Missouri border with an associated upper-level short wave trough approaching the Ohio Valley (Figure 4.1a). The upper-level cyclonic potential vorticity (PV) anomaly (2 PVU dynamic tropopause < 305K) was located over the central part of the Great Plains. During the following 12 h, the surface low progressed eastward along the Kentucky─Tennessee border and weakened as it approached a pre-existing ridge, as indicated by the warm potential temperature (PT) on the tropopause (Figure 4.1b). At 1200 UTC 26 January (Figure 4.1c), the short wave trough and the associated surface low moved eastward across the central Appalachians. A new surface cyclone formed off the east coast of North Carolina at 1800 UTC 26 January, between the left exit of an upper level jet stream over the southeastern U.S. and the right entrance of a downstream jet east of New England (not shown).

Between 1200 UTC 26 January and 0000 UTC 27 January, the upper-level trough became negatively tilted as it approached the East Coast, and the surface low intensified just off the mid-Atlantic coast (Figure 4.1d). The upper-level PV maximum was over the northwest of the trough axis. Hence there was positive PV advection near the trough axis off the mid-Atlantic coast as well as the corresponding surface low pressure system. During the following 12 h (Figure 4.1e), a closed 500 hPa low developed and strengthened off eastern Long Island. The surface low continued tracking north-northeastward to the western Atlantic east of Long Island, and deepened rapidly from 992 hPa at 0000 UTC 27 to 980 hPa at 1200 UTC 27. Over the 24-h period between 12 UTC 26 and 12 UTC 27, heavy snow fell over southeastern New England with maximum precipitation near Boston (Figure 4.2a).

At 0000 UTC 28 January (Figure 4.1f), the 500 hPa low was situated over Cape Cod and became neutrally tilted while the surface cyclone was centered east of Boston with a minimum pressure of 985 hPa (Figure 4.1f). An occluded front extended northeastward from the surface low into the Atlantic Ocean (not shown). The surface low weakened as the upper level closed low started to fill through 1200 UTC 28 January and snowfall intensity diminished to light and scattered nature over the Northeast (not shown).

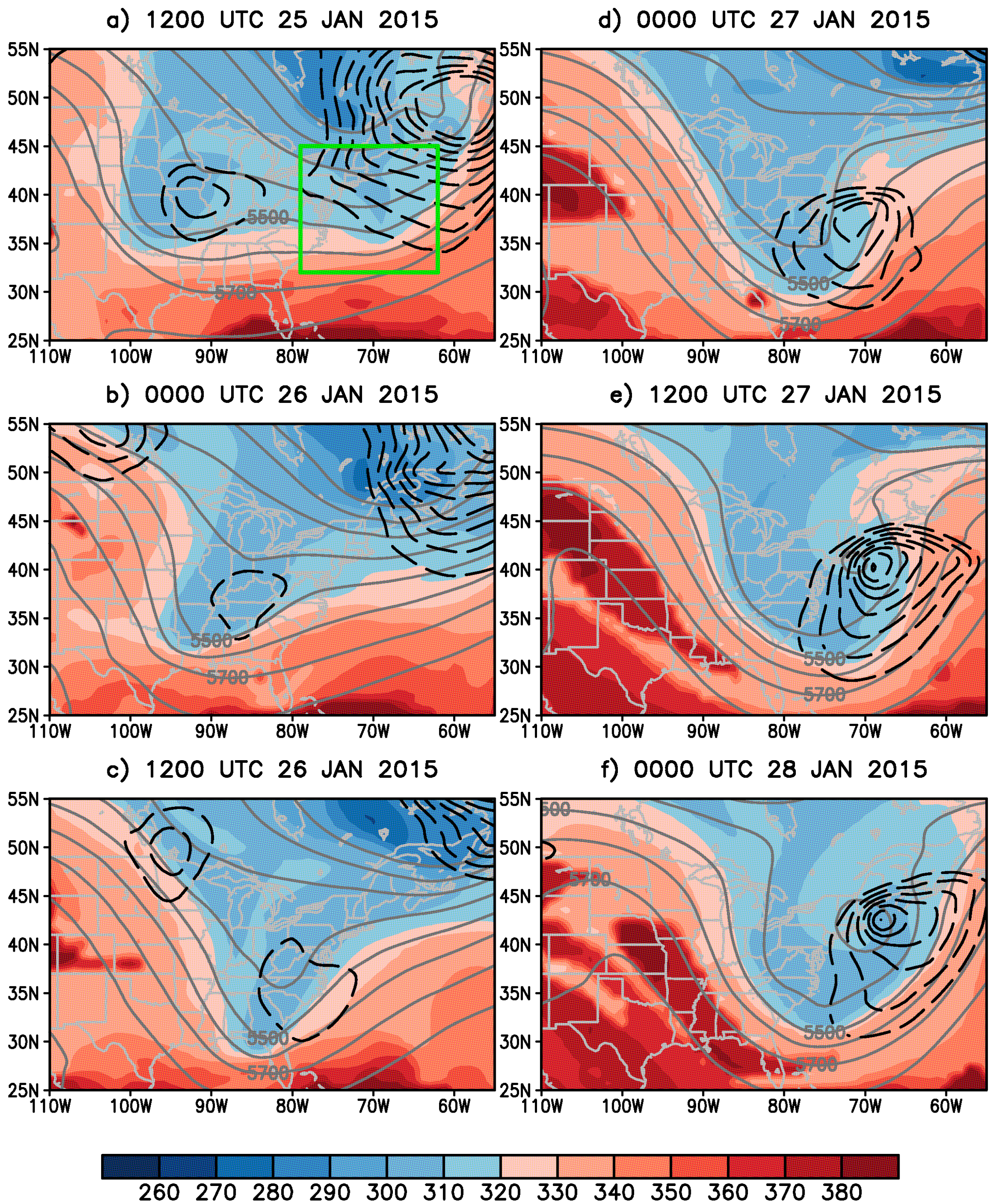


Figure 4.1: (a)–(f) Z500 (gray solid contour), MSLP (black dashed contour) and potential temperature at 2PVU dynamical tropopause (shading, [K]) from 1200 UTC 25 January to 0000 UTC 28 January 2015. Contour interval for geopotential height is 100 m, and the 5700 and 5500 m lines are labeled. The MSLP contours start from 1008 hPa and decrease at 4 hPa intervals. The green solid box in (a) represents the verification region at valid time (1200 UTC 27 January 2015).

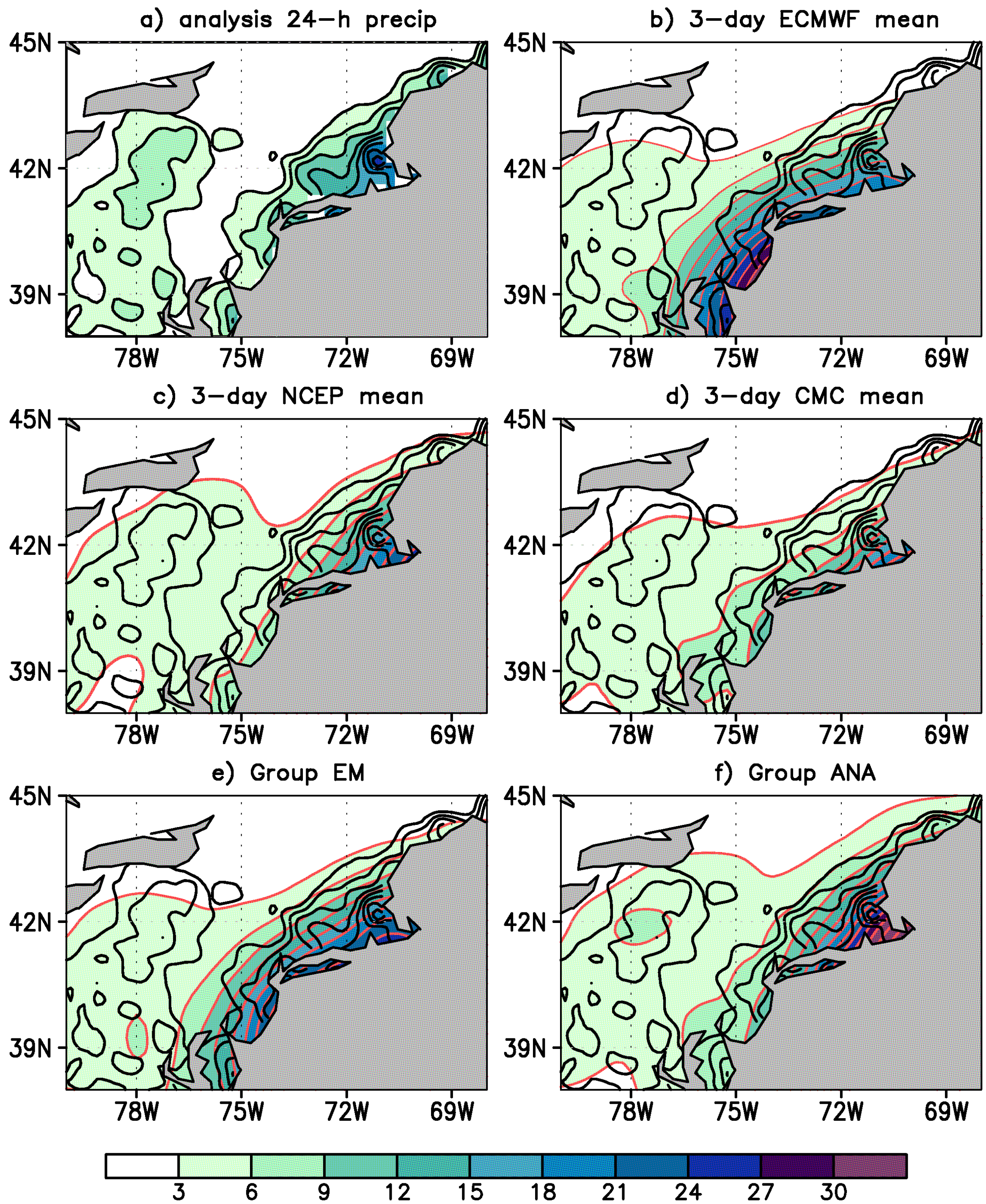


Figure 4.2: (a) The CPC 24-h precipitation (shading and black contour), (b)–(d) 3-day ensemble mean forecasts (shading and red contour) from three models, and (e)–(f) the means (shading and red contour) from Group EM and Group ANA (or Group 2) for 3-day m ulti-model ensemble. The black contours in (b)–(f) are the same with the contours in (a) to represent the observed precipitation. Unit: [mm]. Time period is from 1200 UTC 26 January to 1200 UTC 27 January 2015.

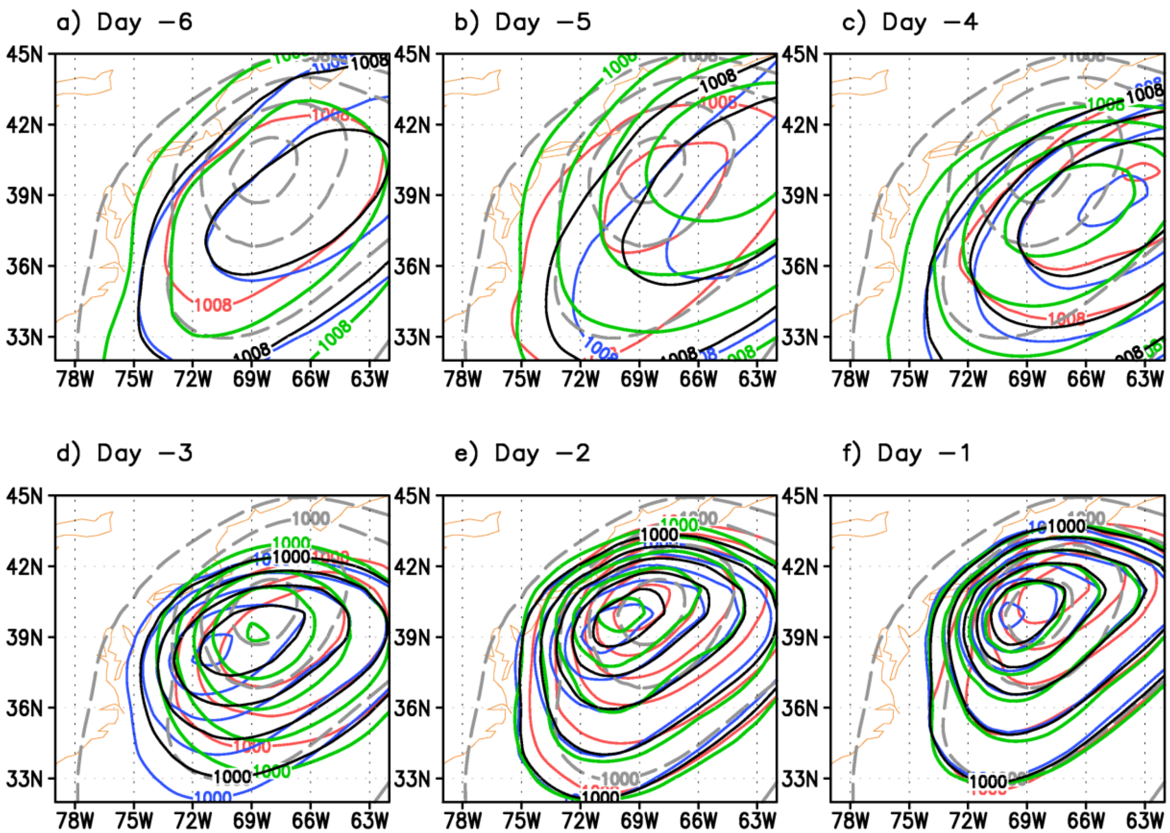


Figure 4.3: (a)–(f) Ensemble mean forecasts of MSLP from a lead time of 6 days to 1 days. The red, green and blue contours are the ensemble means from the CMC, NCEP, and ECMWF model, respectively. Black and gray are the multi-model mean and analysis. Contours for the ensemble means in (a)–(c) range from 1008 to 1000 hPa, and in (d)–(f) range from 1000 to 980 hPa with a contour level of 4 hPa.

4.2.2 Ensemble forecasts at different lead times

Figure 4.3 shows the ensemble means from each of the models, the multi-model mean, and the analysis verifying at 1200 UTC 27 January 2015 for different lead times. Between days -6 and -4, all 3 EPSs forecasted a mean cyclone that was east or southeast of the analyzed cyclone. Between days -5 and -4, the NCEP and ECMWF mean cyclone shifted west-southwestward. This trend continued between days -4 and -3, and during that 24-h the CMC forecasted cyclone also shifted westward. From days -3 to -1, the ECMWF mean cyclone was the deepest and furthest west, while the CMC mean cyclone was the weakest and furthest east, with the NCEP mean cyclone located between the two. Overall, for the short-range forecasts the multi-model ensemble mean indicated a closer to shore cyclone than the analysis, which led the forecasters at NWS to predict a major snowstorm and issue a blizzard warning for parts of seven states including New York City (Figure 4.3), which turned out to receive far less snow than expected (Fanning 2015; Flegenheimer 2015).

In the following subsection, we will focus on the ensemble forecasts at the forecast time of 3 days to illustrate how to apply the EOF/fuzzy clustering method to separate the forecast scenarios in the medium range associated with this blizzard case.

4.2.3 Steps for applications to operational forecast

The fuzzy clustering method has been introduced in Section 2.5. The steps to apply the EOF/fuzzy clustering method are as follows:

STEP 1: compute the variability pattern among the multi-model ensemble using the EOF method

Given a set of ensemble runs, an EOF analysis is first performed to find out the leading variability patterns or forecast uncertainty patterns among the ensemble members.

Day 3 ensemble mean forecast (Figures 4.4a-b) predicts a cyclone centered at (39.0°N, 69.5°W) with a minimum pressure of 985 hPa, ~120 km to the south-southwest of the analyzed cyclone. The maximum ensemble variance is ~400 km to the west-southwest of the ensemble mean cyclone center (Figure 4.4a). The spaghetti plot of the 996 hPa contour lines (Figure 4.4b) suggests that there is large forecast uncertainty regarding the position of the cyclone.

Here, we perform EOF calculations on the 90-member 3-day MSLP forecast anomaly relative to the ensemble mean. The two leading EOF patterns for this 3-day MSLP forecast explain 42.9% and 28.7% of the variance, respectively. Positive (negative) EOF1 (Figure 4.4c) pattern is a monopole centered ~260 km west of the ensemble mean cyclone, representing a deeper (weaker) as well as a westward (eastward) shifted cyclone. Positive (negative) EOF2 (Figure 4.4d) pattern is a dipole, representing the northeastward (southwestward) shift of the cyclone, indicating that there is a subset of ensemble members shifted to the northeastward (as well as southwestward) at VT.

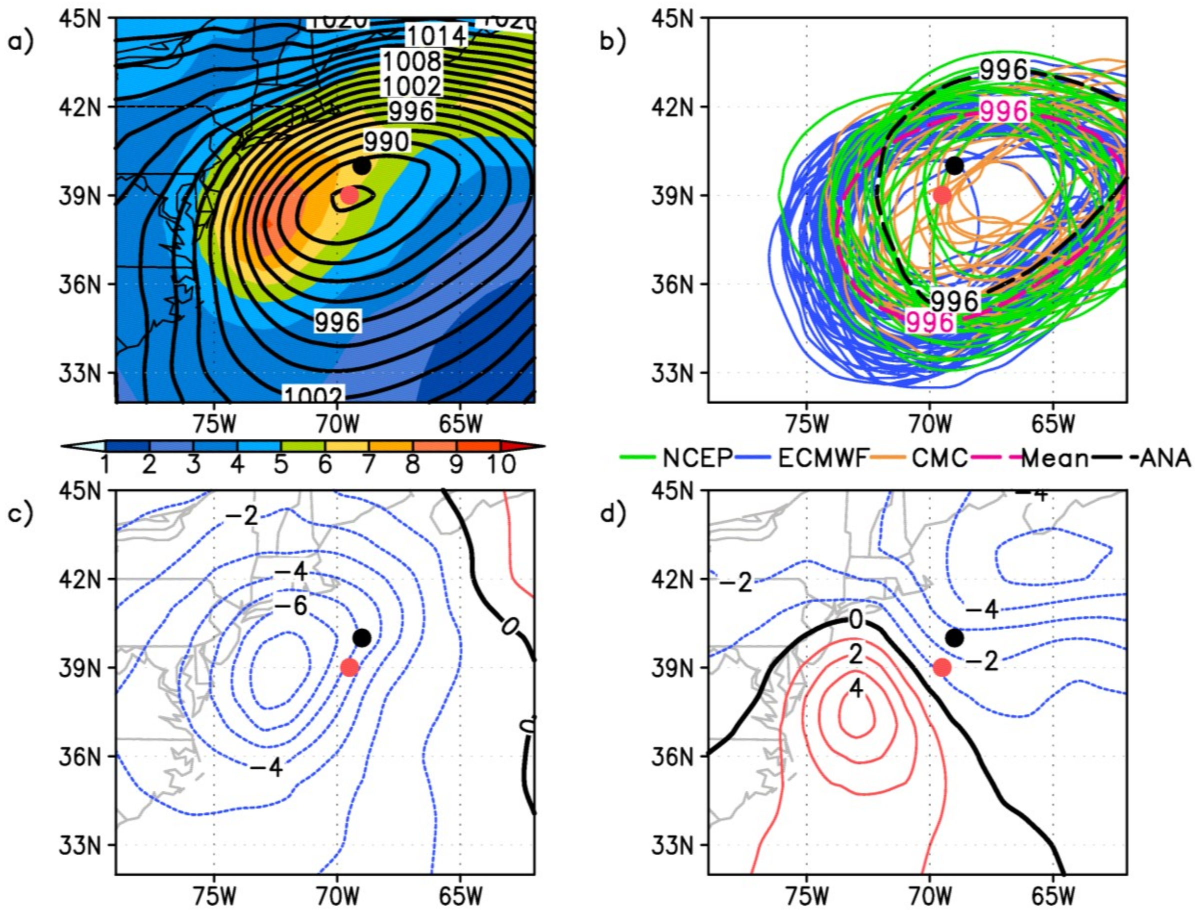


Figure 4.4: (a) MSLP ensemble mean (contours, [hPa]) and spread (shading, [hPa]), (b) spaghetti plots of 996 hPa contour for 90 multi-model ensemble members (blue are for the ECMWF members; green are for the NCEP members; and orange are for the CMC members) with the dashed magenta lines and black lines to be the ensemble mean and the analysis. (c) EOF1 MSLP pattern (contours, [hPa]), and (d) EOF2 MSLP pattern (contours, [hPa]). The VT is 1200 UTC 27 January 2015; IT: 1200 UTC 24 January 2015. The black/red dot in each panel denotes the analyzed/ensemble mean position of the surface cyclone at VT.

STEP2: Determine forecast scenarios based on the leading two EOF PCs using fuzzy

clustering method

With the corresponding EOF PCs for the ensemble forecast, fuzzy clustering can group them based on each pair of PC of the 90 members.

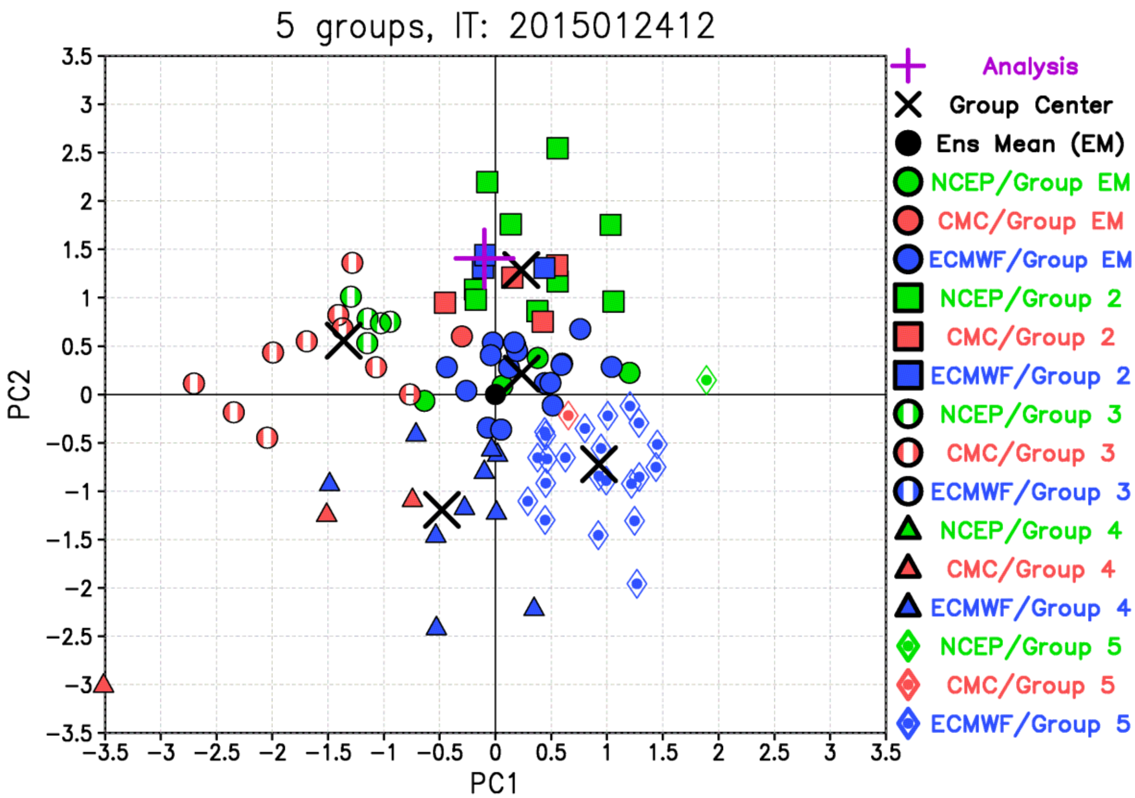


Figure 4.5: The 5 clusters divided using fuzzy clustering method on the PC1-PC2 space from the 90 ensemble members for 3-day forecast. The VT is 1200 UTC 27 January 2015, and IT is 1200 UTC 24 January 2015.

Figure 4.5 shows the 5 groups based on the fuzzy clustering method. Group EM is centered close to the ensemble mean represented by the origin on the PC1-PC2 coordinate. Group 2 represents positive EOF2 characteristics–the northeastward shift of the cyclone position relative to the ensemble mean (Figure 4.4d). Group 3 represents mainly a negative EOF1 pattern, with the cyclone in this group anticipated to be weaker and more eastward than the ensemble mean. Group 4 has mainly a weak negative EOF1 anomaly and a negative EOF2 anomaly, suggesting a cyclone weaker and more southwestward relative to the ensemble mean. One of the CMC member in this group is at (-3.5, -3.0), seemingly an outlier with respect to the rest of the ensemble. Group 5 indicates a combination of positive EOF1 and negative EOF2 anomalies. Therefore, the cyclone in this group should be more west-southwestward (more onshore) and deeper than the ensemble mean.

To sum up, groups 2 and 4 represent the largest positive and negative EOF2 pattern, corresponding to more northeastward and southwestward shift, respectively. Meanwhile, groups 3 and 5 represent the largest negative and positive EOF1 pattern, corresponding to weaker/off-shore and deeper/on-shore cyclone scenarios, respectively.

STEP3: Plot cluster means for different parameters

With the partitions of five clusters, the cluster means can be plotted to show the differences between the clusters.

Figure 4.6 shows the group means of the 5 clusters (partitioned based on the EOF PCs on day 3) from day 1 to day 3 using the 1008 hPa MSLP and 5400 m Z500 contour lines. On day 1 (Figures 4.6a-b), the surface cyclone in all groups seem to be similar, but the cyclone for Group 2 is more intense than that of the other 4 groups. The Z500 short wave trough shows more separations among the groups, with that of Group 2 being the strongest and most southeastward one. On day 2 (Figures 4.6c-d), the surface cyclone as well as the upper level trough in Group 2 are still the most eastward among the 5 groups (Figures 4.6c-d). Group 5, on the other hand, has the cyclone being the most westward. Group 3 shows the weakest cyclone with the center splitting into two (Figure 4.6c).

On day 3 (VT), Group 2 has both the surface cyclone and its associated upper level trough being ~120 km more northeastward than that of Group EM (Figures 4.6e-f), which is consistent with the earlier forecast time steps. Group 5 has the cyclone deeper than that of Group

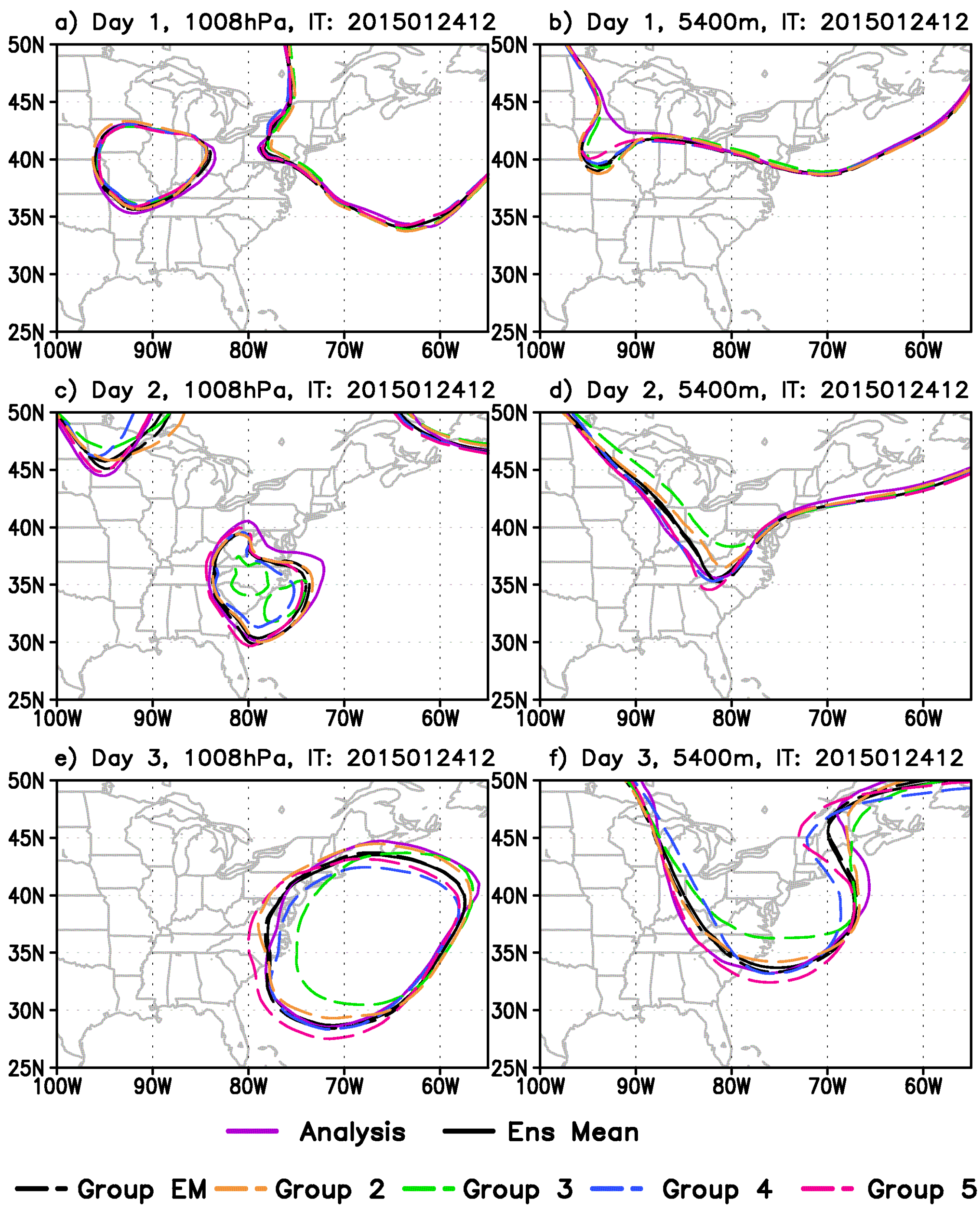


Figure 4.6: A summary of 5 cluster means using 1008hPa MSLP contour line (a, c, and e) and 5400m Z500 contour line (b, d, and f) for 3-day forecast. The VT is 1200 UTC 27 January 2015; IT, 1200 UTC 24 January 2015.

EM and is ~320 km more west-southwestward (Figure 4.6e). Its associated trough is the deepest one among the five groups and extends more west-southwestward than the ensemble mean

(Figure 4.6f). In contrast, both the surface cyclone and the upper level trough in Groups 3 and 4 (Figures 4.6e-f) are weaker than that of Group EM. The cyclone of Group 3 continues to be the weakest one with the center ~320 km more northeastward than the ensemble mean. That of Group 4 is weaker by 3 hPa and ~170 km more southwestward than the ensemble mean. The synoptic features represented by both surface and upper level variables again match the interpretations of clusters of Figure 4.5 in the previous step.

STEP4: Plot spaghetti plot for each cluster to show cluster details

Spaghetti plots can provide more details of a variable from an ensemble forecast. If a forecaster sees a scenario he/she is interested in, spaghetti plots for that cluster can provide more details of each cluster member and often include most of the similar members corresponding to that cluster mean scenario.

Figure 4.7 shows the spaghetti plots of the 1000 hPa MSLP contour lines for the 5 groups aforementioned. The ECMWF members are mainly contributing to groups EM, 4 and 5 (Figures 4.7a, d, and e), forecasting a more close-to-shore cyclone. However, 3 ECMWF members also contribute to Group 2, together with 9 NCEP and 2 CMC members, forecasting a deeper and more east-northeastward scenario. Half of the CMC members contribute to Group 3, together with 5 NCEP members, forecasting the weakest and more northeastward cyclone (Figure 4.7c). The spaghetti plot confirms that the cluster method is able to find the members with similar forecast scenario not only from the same EPS but also different models. Figures 4.2b-d shows that the 24-h total precipitation in ECMWF mean is closer to shore than the other two models, with the heaviest precipitation regions including New York City and New Jersey. Figures 4.2e-f compares the mean precipitation for Group EM and Group 2, suggesting the heavy precipitation in Group 2 is more northeastward and closer to the analysis than Group EM. Figure 4.2 again demonstrates the consistency of the cluster scenarios for different important weather elements.

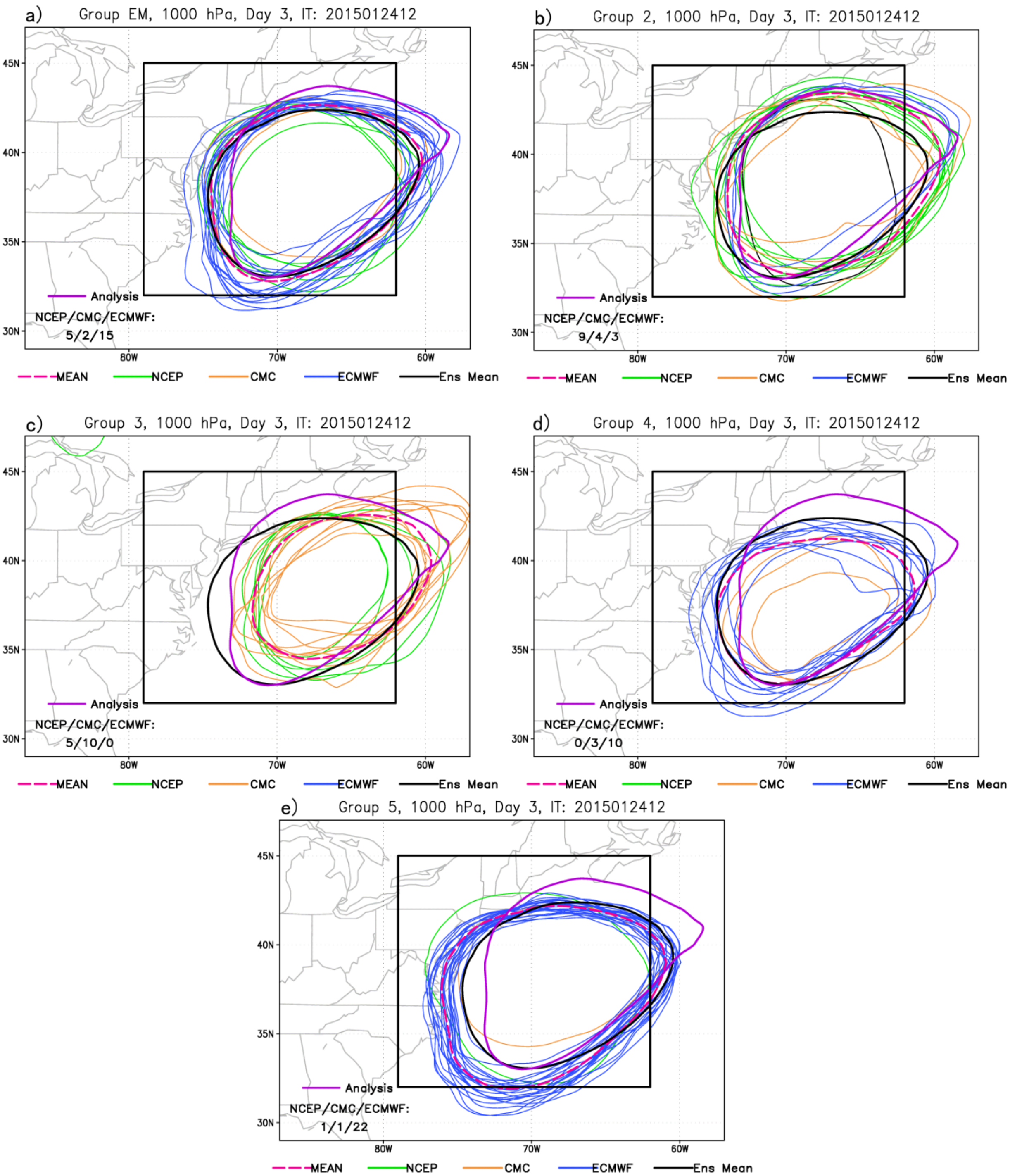


Figure 4.7: Spaghetti plot for the 5 groups (a)–(e) using 1000 hPa MSLP contour line. Red, green and blue contour lines are from CMC, NCEP and ECMWF members, respectively. The dashed magenta line is for the group mean. The black and purple lines are the multi-model mean and the analysis in all panels. The VT is 1200 UTC 27 January 2015; IT: 1200 UTC 24 January 2015.

4.3 Comparing different forecast metrics for clustering

One important question about applying fuzzy clustering to ensemble forecast is how sensitive the results depend on the field used to conduct the clustering. One objective tool to verify cluster results from two different fields is the Rand index or RI (Rand 1971). As suggested by Keller et al. (2011), this index is very effective in determining whether the cluster solution for one forecast variable separates the members in the same way that the cluster solution for another forecast variable does. By using the RI to investigate the clusters between TIGGE ensemble and TIGGE-without-EC ensemble, they found that the RI ranges from 0.72 to 0.92 in the ten ET cases by using Z500 as the forecast variable. They indicated that >0.89 would indicate high agreement and <0.65 would indicate low agreement between two clustering approaches. Here the RI is applied to compare clusters using different variables.

Here, we use a simple example to illustrate the computation of the RI. If we apply the fuzzy clustering approach to two distinct forecast variables (e.g. using MSLP and 24-h total precipitation as variables for EOF- and cluster- analysis, respectively) with, for example, N=6 members, we will most likely get two distinct solutions Y and Y' on how to group the ensemble members. Assuming that the clustering approach divides the members into two clusters, $Y_{1,2}$ for using the first variable and $Y'_{1,2}$ for using the other variable. Assume the N=6 ensemble members are assigned to the clusters as follows:

Y_1=1,2,6 Y_2=3,4,5; and Y_1'=1,2,3,6 Y_2'=4,5.

To compute the RI, we first count the number of members that are assigned together or separately in the two cluster solutions and that are mixed, i.e., assigned together in one, but separately in the other solution. For the example stated above (adapted from Rand, 1971), the assignment for each pair out of 15 pairs of members is shown in Table 4.1.

Member-pair	Together in both	Separate in both	Mixed
1&2	*		
1&3			*
1&4		*	
1&5		*	
1&6	*		
2&3			*
2&4		*	
2&5		*	
2&6	*		
3&4			*
3&5			*
3&6			*
4&5	*		
4&6		*	
5&6		*	
Total	4	6	5

Table 4.1: Member-pairs assignment

Therefore, ten pairs of members are assigned similarly in both solutions (either together in both or separate in both), while the five remaining pairs are assigned differently. The RI is simply the ratio of the amount of member-pairs assigned to the same solution in both to the total possible member-pairs for assignment in the two solutions. In summary, this can be written as (Adapted from Rand, 1971)

$$RI(Y,Y') = \frac{\sum_{j=2}^{N}\sum_{i=1}^{j}\gamma_{ij}}{\binom{N}{2}} \tag{4.1}$$

where N represents the total number of ensemble members, and γ_{ij} ($i<j$) is defined to be:

$$\gamma_{ij} = \begin{cases} 1 & \text{If members } i \text{ and } j \text{ are assigned together in both} \\ 1 & \text{If members } i \text{ and } j \text{ are assigned separated in both} \\ 0 & \text{If members } i \text{ and } j \text{ are together in on but separated in the other} \end{cases}$$

Thus, RI ranges from 0 to 1. One represents two identical cluster solutions while 0 represents that none of the member-pairs are assigned in the same manner in the two solutions. In the above example, RI equals to 0.67, which would be referred to as a rather moderate agreement.

The cluster results for the 2015 January winter storm case (Figures. 4.5-4.7) discussed above are based on the MSLP forecast. However, it is important to verify and compare the cluster solutions with those based on other forecast variables. Here we employ RI to verify the cluster results using the MSLP, 24-h total precipitation (24-h TP), Z500, and 850-hPa moisture flux (850MF) as variables for EOF/fuzzy clustering, respectively. These three variables are chosen because together they represent the basic thermodynamic parameters to represent three vertical levels. Table 4.2 compares the RI for clustering using the four variables.

RI	MSLP	24-h TP	850MF	Z500
MSLP		0.756	0.713	0.659
24-h TP	0.756		0.725	0.663
850MF	0.713	0.725		0.680
Z500	0.659	0.663	0.680	

Table 4.2: The RI between cluster results of 2015 January winter storm case using MSLP, 24-h total precipitation, 850-hPa moisture flux and Z500 for EOF/fuzzy clustering, respectively.

The member assignments of any two cluster solutions out of the four forecast variables are compared in the above table. This gives an indication of whether two cluster solutions are mainly the same, or strong differences exist in the assignment. With a Rand Index RI=0.756, the cluster solution of the MSLP and 24-h TP have the highest agreement in member assignments. In contrast, the cluster solution of MSLP has only rather a moderate agreement with that of Z500 (RI=0.659). Overall, the MSLP has good agreement with the total precipitation and 850 hPa moisture flux, indicating the MSLP scenarios are consistent with one of the most important forecast parameters: accumulated precipitation. Z500 has the lowest RI with the other parameters, suggesting the scenarios based on upper-level geopotential height are only consistent with precipitation scenarios in a moderate way. Note that the RI values are lower than the results from Keller et al. (2011), which is expected because Keller (2011) calculated RI using the same variable (Z500). Their purpose was to examine the clusters in TIGGE ensemble with or without the ECMWF model. Since we are using RI to validate the clusters using different parameters, the RI values would be expected to be a bit lower owing to the differences in two parameters. Nevertheless, even the lowest RI value (0.659) can still be considered as moderate agreement. The RI results confirm that the clusters using MSLP overall agree well with clusters using other parameters.

4.4 Applying fuzzy clustering to ensemble verification of historical winter storm cases

In sections 4.2 and 4.3 we have discussed the operational application of EOF/fuzzy clustering approach to a winter storm case study and compared the cluster solutions using different variables. When the analysis is available, the forecast scenarios in the multi-model ensemble can be verified by projecting the analysis onto the EOF PCs coordinate. The scenarios closest to the analysis can be used to verify the ensemble forecast. This section discusses how to apply the EOF/fuzzy clustering method to ensemble verification of historical winter storm cases.

4.4.1 Analysis scenario definition

To perform scenario-based ensemble verification of EPSs in predicting East Coast winter storms, the definition of the analysis scenario is of paramount importance. In this study, our cluster analysis is based on EOF PC space; therefore, the analysis scenario is also defined based on the projection (von Storch and Zwiers, 1999) of the analysis onto the leading EOFs. To be more specific, the analysis anomaly (A') relative to the ensemble mean at the verification time is projected onto the leading two EOF patterns (E_1 and E_2) by using the following equation:

$$\alpha_i = \frac{\mathrm{cov}(A', E_i)}{\mathrm{var}(E_i)}, i = 1, 2 \qquad (4.3)$$

where α_i stands for the projection coefficient of the analysis anomaly onto the EOF 1 or 2 patterns. This is based on the property that the EOF patterns are all orthogonal to each other. Therefore the verifying analysis at the verification time is translated onto the EOF PC1-PC2 phase space by adding the pair of projection coefficient α_i ($i=1, 2$) to the PC1-PC2 scatter plot. The point occupied by α_1 and α_2 represents the analysis on the PC1-PC2 coordinate. The cluster with the center having the shortest distance to the analysis point is considered to represent the analysis more closely than the remaining clusters; in other words, we define that cluster as "Group ANA" or the "analysis scenario". One clarification here is the analysis point is not included in the clustering procedure to prevent modifications of the cluster assignments.

Based on the above definition, for the 2015 January case, Group 2 in Figure 4.5 is the analysis group. As can be seen from Figures 4.6-7, Group 2 (or Group ANA) is the closest cluster to the analysis (purple solid line in Figures 4.6-7), which forecasts a cyclone more north-northeastward than the ensemble mean. The average precipitation corresponding to Group 2 is also more north-northeastward than that of the multi-model ensemble mean (Figures 4.2e-f). The heaviest precipitation in Group 2 is much closer to the observed than the other clusters as well as individual model means (Figure 4.2). In another case study for the 2010 December snowstorm examined in Chapter 3, we found that Group 3 was assigned as the analysis group for a 3.5 day forecast initialized at 0000 UTC 24 December 2010, which is the closest to the analysis contours in MSLP and Z500 over the Eastern U.S. and western Atlantic at verification time among the 5 groups. In both cases, Group ANA is distinct from Group EM, that is, the analyses do not lie within the space occupied by the ensemble mean groups.

Here, the analysis group has been defined based on the projection of the analysis onto the PC1-PC2 space. Two more conventional metrics, root-mean-squared error (RMSE) and pattern correlation coefficient (Corr), are used to verify that the analysis group defined this way is indeed the closest to the analysis.

Figures 4.8 and 4.9 show the RMSE and Corr of Z500/MSLP from IT to VT for the 2015 January winter storm (Case 1) and 2010 December case (Case 2), respectively. At VT (72 h for case 1 and 84 h for case 2), the analysis group (Group 2 for case 1 and Group 3 for case 2) has smaller RMSE (Figures. 4.8a-b and 4.9a-b) and higher pattern correlation (Figures. 4.8c-d and 4.9c-d) than the other four groups. In addition, it also has smaller RMSE and higher Corr than the multi-model mean as well as each individual model mean. The superiority of the analysis group is significant at the 95% level based on 100 random sampling of groups.

The superiority of the analysis group is not only limited to the VT. The time range of its robustness can also be seen from Figures 4.8-9. For the 2015 case (Figure 4.8), the analysis group is significantly better from 42h (VT-30h, Figures. 4.8b,d) to 84h (VT+12h, not shown), spanning over 40h. For the 2010 December case (Fig. 4.9), the analysis group is significantly better from 30h (VT-54h, Figs. 4.9b,d) to 120h (VT+36h, not shown), spanning over 3.5 days. Therefore, the time range of the dominant benefit of using the analysis group varies case by case, but it can be robust spanning as long as over 3.5 days. Nevertheless, Figures 4.8-9 show that at early forecast times, other groups are closer to the analysis than the analysis group, indicating that one cannot easily pick a "best" forecast group based on its agreement with observations during early forecast hours, consistent with the findings of Smith (2001) and Orrell et al. (2001).

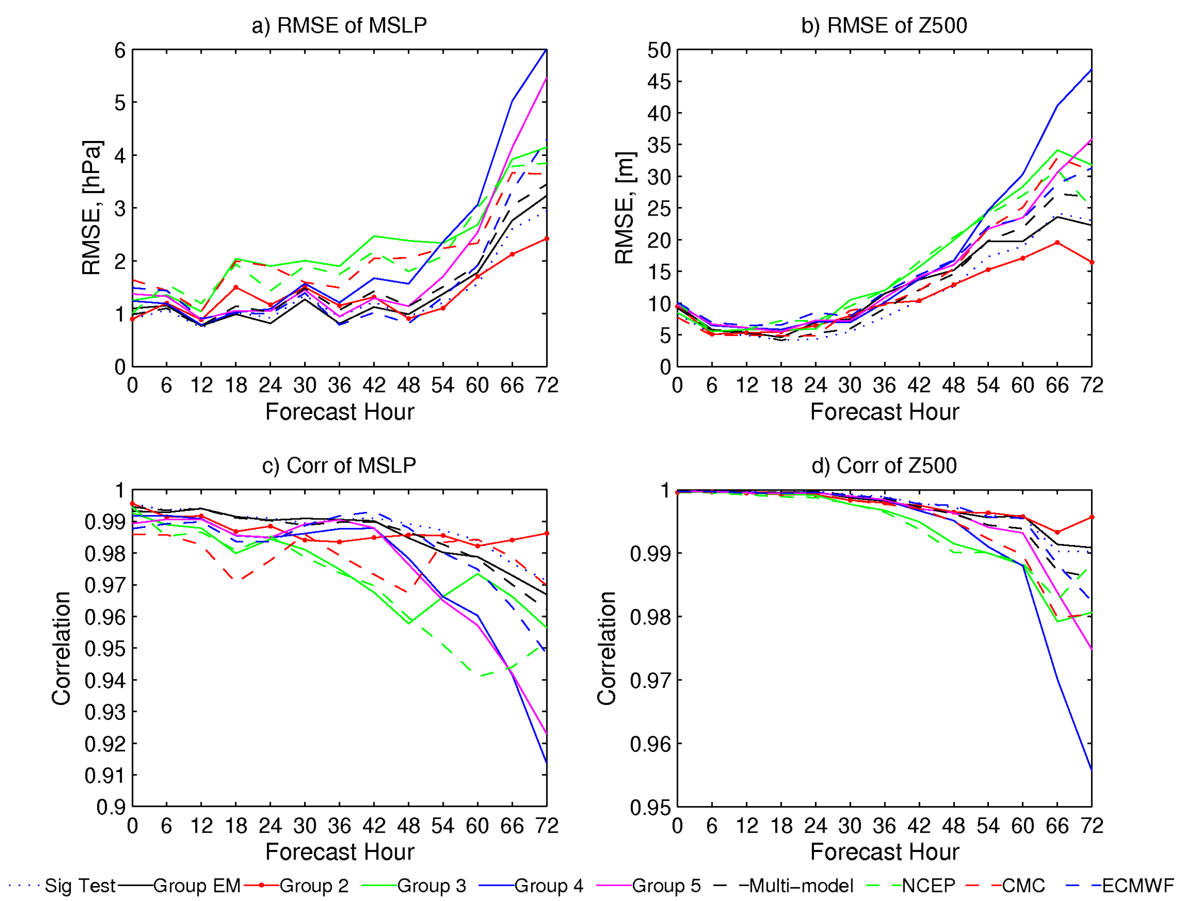


Figure 4.8: RMSE of MSLP (a) and Z500 (b) for each cluster mean and model means at different forecast times; Correlation of MSLP (c) and Z500 (d) between each mean and the analysis at different forecast times. The clusters are selected based on the EOF PC s at 72 h forecast initialized at 1200 UTC 24 January 2015. The black, green, blue and purple solid lines are for

Group EM, 3, 4 and 5, respectively. The red dotted solid line is for Group 2 (analysis group). The black dashed line is for multi-model ensemble mean. The green, red and blue dashed lines are for the NCEP, CMC and ECMWF model means. The blue dot is the 95% significant line based on 100 randomly selected clusters. The verification area is a box (24 longitude x 20 latitude) moving with the cyclone (for MSLP) and the corresponding upper level trough (for Z500) system.

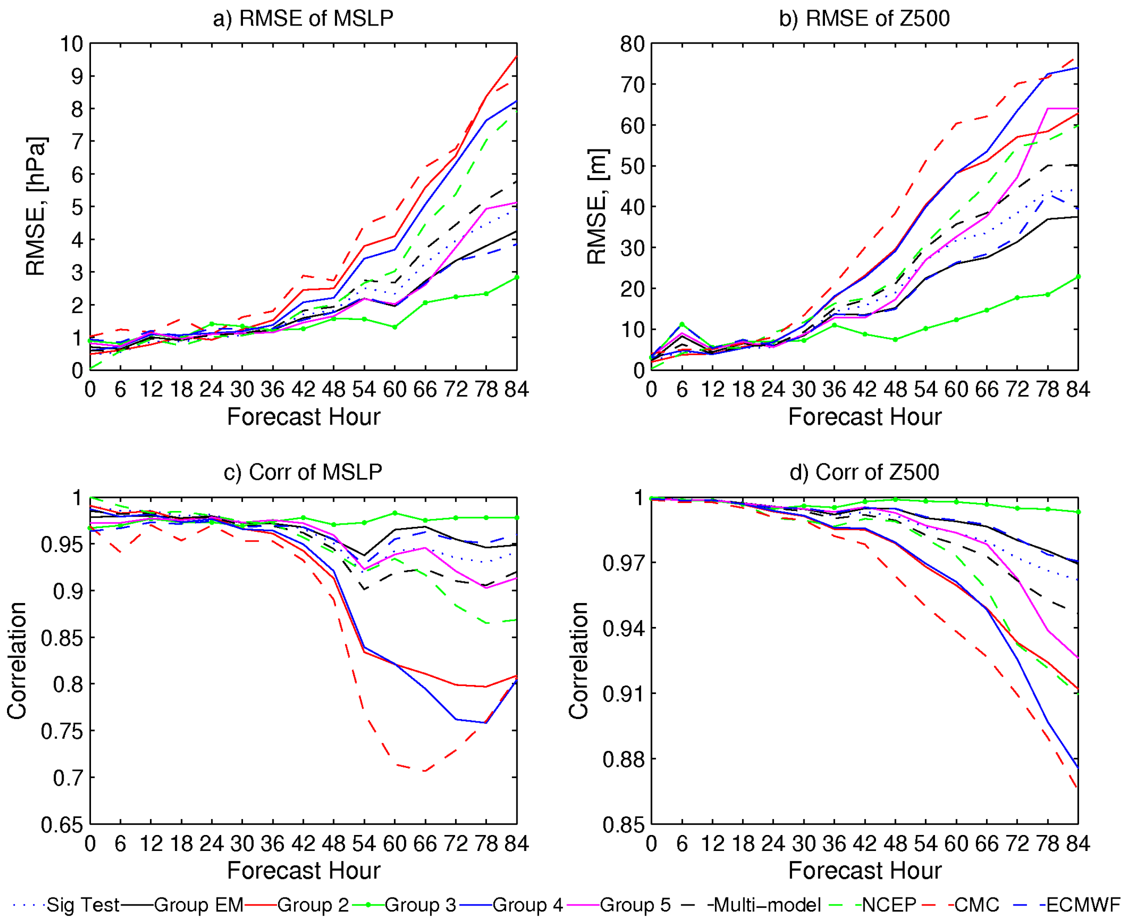


Figure 4.9: The same as Figure 4.8 but for Case 2 initialized at 0000 UTC 24 December 2010 and verified at 1200 UTC 27 December 2010. Group 3 (green dotted solid line) is the analysis group.

The verifications using RMSE and Corr confirm that the analysis group has relatively lower errors and higher pattern correlations among the 5 groups over a time range spanning the VT. The selection of the analysis group is hence reliable to be applied to operational post-processing and model calibration. The result also indicates that by combining members from different models, an optimal cluster can be more valuable than both multi-model ensemble mean and individual model means.

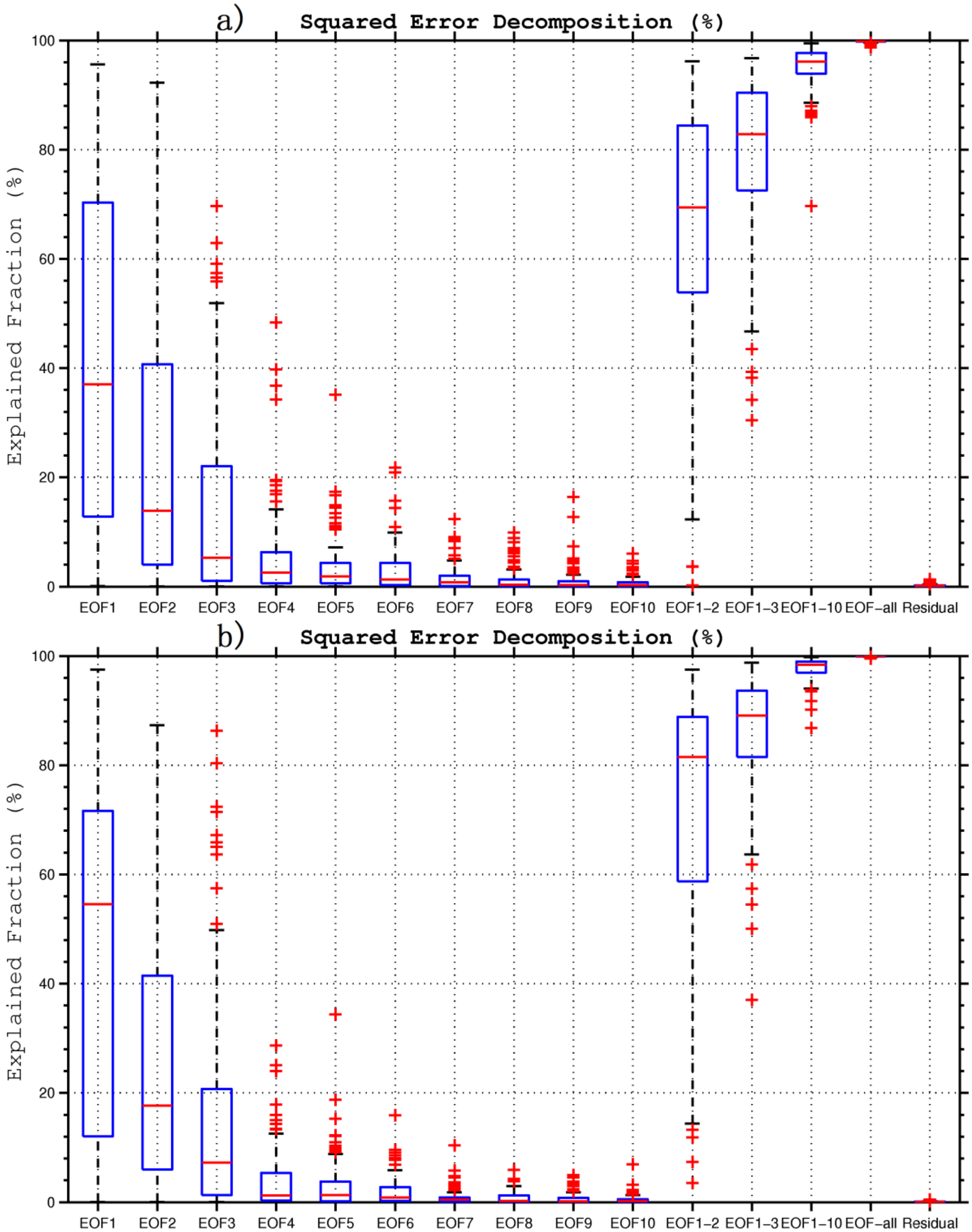


Figure 4.10: The explained squared error fraction by EOF1 to EOF10 patterns respectively, as well as the accumulated fraction by EOF1-2, EOF1-3, EOF1-10, all 90 EOF patterns, and the residual for (a) day 3 forecast and (b) day 6 forecast.

The projection of the analysis onto the PC space and the definition of the analysis group can be used as a verification tool in ensemble verification. An assumption of this application is that the forecast errors do project primarily onto the leading two EOF patterns. In other words, the leading two EOF patterns explain the dominant forecast errors. To examine whether this holds true in most cases, the fraction of squared error (error here is defined as ensemble mean minus the analysis) explained by each EOF pattern is plotted in Figure 4.10 based on day 3 and day 6 forecasts for 111 cyclone cases over region 2. For day 3 forecasts, the median fraction of explained squared error by the leading two EOF patterns is around 70%, suggesting that over two thirds of the squared forecast error can be explained by the leading two EOF patterns. This value reaches 81% for day 6, suggesting that most of the forecast errors do project on the leading two patterns. Note that there are a few outlier cases for day 3 and day 6 with the leading two EOFs explaining very small amount of squared errors. Nevertheless, in most cases most of the forecast errors can be explained by EOF1 and 2 patterns.

4.4.2 Outside-of-envelope (OOE) cases

When analyzing the ensemble forecasts, there can be cases in which the analysis is out of the multi-model ensemble envelope. Figure 4.11 is an example comparing cases with the analysis in and out of the forecast envelope. The analysis is located within the cluster on the PC1-PC2 space for the case shown in Figure 4.11a; in contrast, the analysis is clearly out of the region enclosed by the dashed line in Figure 4.11b. Quantitatively, the outlier cases or OOE cases are defined by the following criteria: 1) the analysis is outside the boundary defined by the line segments joining the vertices on the PC1-PC2 coordinate; 2) the distance between the analysis and the closest member on the PC space is larger than the average distance between any two members plus 1 standard deviation. Among the cases we examined, there are 4 outlier cases for 3-day forecast, 16 for 6-day forecast, and 19 for 9-day forecast. For most of the statistics discussed in sections 4.5 and 4.6 which are dependent on the existence of an analysis group, we have excluded the corresponding OOE cases for forecast at each lead time considering that the outlier cases don't really have an analysis group and those cases should not be included in the related-statistics. However, the OOE cases are still included in calculating error/spread relation statistics (sections 4.5.3 and 4.5.4) since these statistics do not directly depend on the existence of an analysis group when evaluating the multi-model ensemble to avoid biasing the results. More details will be given in sections 4.5 and 4.6.

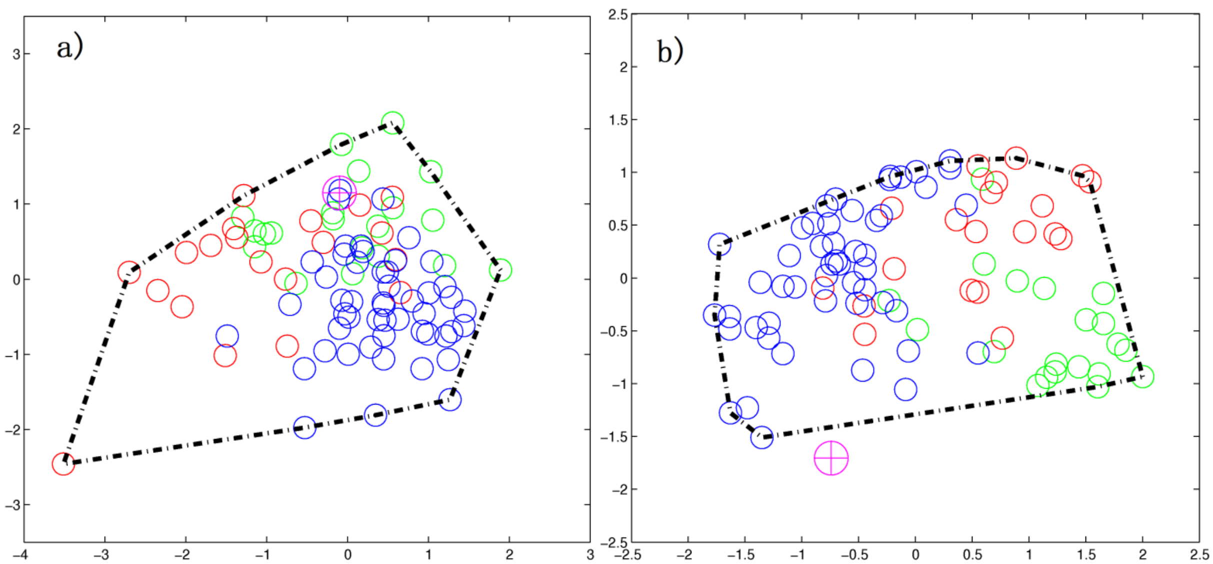


Figure 4.11: In the envelope (left, a) and outside-of-envelope (right, b) examples. Green, red, and blue open circles represent members from the NCEP, CMC, and ECMWF models, respectively. Black dashed represents the outside envelope of the multi-model ens emble. Magenta circle with a plus sign denotes the analysis point. Left (a) is for 3-day MSLP forecast initialized at UTC1200 January 24 2015; while right (b) is for 3-day MSLP forecast initialized at UTC1200 December 09 2008.

Given that the OOE cases are selected based on PC1-PC2 space, it is important to confirm that the case selection criteria make sense on the corresponding physical atmospheric field. Figure 4.12 shows the spaghetti plot for the OOE case shown in Figure 4.11b. It is clear that the analysis (black dashed line) is more southwestward than most of the ensemble contours;

except for three members from the ECMWF model. However, even the closest three members are much deeper and more west-northwestward than the analysis. Figure 4.12b depicts the corresponding group means for the five-cluster solution, which clearly shows that the analysis is quite distinct from the mean of any of the five clusters. Therefore, the definition of the OOE case for this case appears reasonable in the physical space. Apart from this case, we have also manually examined the other OOE cases determined based on the PC1-PC2 phase space, and confirmed that most of these cases are real outliers based on the physical field. Therefore, our definition of OOE cases is robust and can be used for investigating the statistics of outlier cases. More details concerning these cases will be presented in Section 4.6.

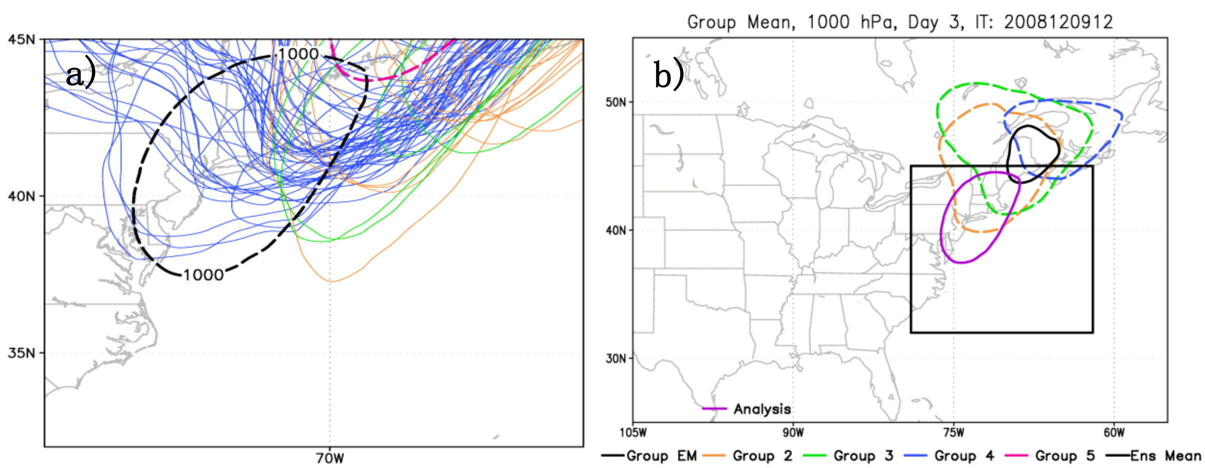


Figure 4.12: (left, a) Same spaghetti plot as Figure 4.4b except for 1000 hPa contour line and initialized at 1200UTC December 09 2008; (right, b) same cluster mean plot as Figure 4.7e except for 1000 hPa contour line with initial time at 1200UTC December 09 2008.

<u>4.4.3 Steps for applying clustering approach to verify a historical storm case</u>

To apply the fuzzy clustering approach to verify a past storm case, the clusters of the 90 ensemble members are first determined based on the steps listed in subsection 4.2.3. The analysis anomaly relative to the ensemble mean is then calculated and projected on to the leading EOF1 and EOF2 patterns, respectively. The corresponding projection coefficient pair α_i ($i=1, 2$) is plotted on the scatter plot of PC1 and PC2. According to the distance between analysis point and each ensemble member, whether the analysis is outside of the multi-model ensemble is first determined. If it is, then this case is referred as an OOE case; if not, the cluster with the center closest to the analysis point is defined as the analysis group (Group ANA) to represent the analysis scenario. The corresponding statistics associated with the analysis group can then be calculated, and these statistics will be discussed in the next section.

4.5 Evaluation of the multi-model ensemble

This section evaluates the capability of different EPSs in forecasting past East Coast winter storm cases. The cyclone cases are selected based on the NCEP operational analysis of MSLP with the minimum pressure less than 1005 hPa during the cool seasons from 2007/08 to 2014/15. The multi-model ensemble is verified over region 2 (Figure 2.1) at lead times of 1 day to 9 days with an interval of 1 day. The statistics for region 1 are also calculated for day 3, 6 and 9. The results are overall consistent. Though the number of verified cyclones at the same verification time should be the same, the number of cases are still different for different lead times due to missing data in the TIGGE archive. Table 4.3 shows the number of cases used as observed cyclone cases for each lead time.

Lead Time [day]	1	2	3	4	5	6	7	8	9
Case number	164	176	180	158	170	185	170	168	178

Table 4.3: The number of cyclone cases used for each forecast day.

Three types of scenario-based statistics are examined here: the average percentage of members that are in the analysis group for each model (section 4.5.1); the rate that each model misses the analysis group (section 4.5.1); and the frequency of the ensemble mean group capturing the analysis scenario (section 4.5.2). Note that when calculating the above scenario-based statistics, OOE cases are not included. Later in this section, the error-spread relation (section 4.5.3), and the model bias as well as its associated physical interpretation (section 4.5.4) will be explored using the PC metrics, and OOE cases are included in calculating these statistics.

4.5.1 Different EPS's ability in capturing the analysis scenario

As demonstrated in the previous section, the analysis group resembles the analysis most closely with the smallest distance between cluster center and the analysis. By employing this definition, we can find out the analysis scenario at verification time for each ensemble run for each past cyclone case. Thus, we have repeated this calculation for all the historical cyclone cases over the East Coast region (region 2). In this study the NCEP operational analysis is used. We have also tested similar calculations using the ECMWF analysis and found that the analyses from different centers are very similar and lie close together on the PC1-PC2 scatter plot. Hence the choice of the analysis will not affect the related statistics. Another thing worth noting is that the analysis could be out of the ensemble envelope in some cases as discussed in subsection 4.4.2, and these cases will be investigated separately in section 4.6. Unless specified otherwise, the OOE cases have been excluded from the calculation of the following statistics.

Figure 4.13 shows the average percentage of ensemble members contributing to the analysis group with respect to the total number of ensemble members for the NCEP, CMC and ECMWF over region 2 at lead times of 1−9-days. The percentage statistic is averaged over all cases with the OOE cases excluded. For short-range forecast (days 1−2) (Figure 4.13), on average, there are around 26% (5.2) of NCEP members contributing to the analysis group. The highest percentage of members among the three EPSs suggests the NCEP is more superior in capturing the analysis scenario in short-range forecast. In contrast, the CMC model has the lowest percentage (~18%, or 3.6 members) contributing to the analysis scenario among the three models. The average percentage of ECMWF model is in-between.

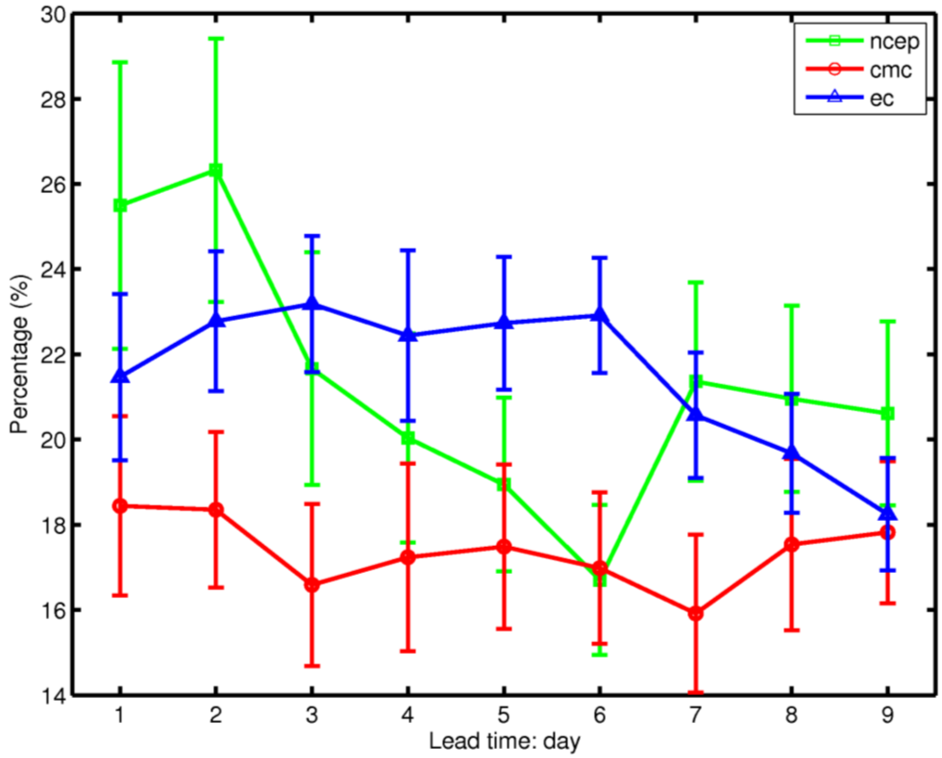


Figure 4.13: The average percentage of member numbers in the analysis group for each EPS. The NCEP, CMC, and ECMWF have a total of 20, 20, and 50 ensemble members, respectively. The percentage is calculated by counting the number of each model member in the analysis group and dividing it by the total member number of each model. The average percentage is calculated based on all the historical cases for each lead time. The vertical bars represent the 95% confidence level for each lead time.

For medium-range forecast (days 3–6), the ECMWF model has the highest percentage (~23%, or 11.5) of members assigned to the analysis group, suggesting that it does a very good job in capturing the analysis scenario in medium-range forecast. The CMC model still has the lowest percentage (~17%, 3.4), but they are not significantly lower than that for the NCEP model (except for day 3). The average percentage of the NCEP model has dropped from day 2 to day 6, suggesting its decreasing skill in the medium range in capturing the analysis scenario.

As for extended-range forecasts (days 7−9), the NCEP model's percentage (~21% or 4.2) has increased since day 6, and becomes the highest one among the three models. Meanwhile, the average percentage for the ECMWF model has decreased since day 6, but it is not significantly lower than the NCEP model. The average percentage for the CMC model is still the lowest one among three; however, it has slightly increased from day 6 to day 9. On day 9, it is not significantly different from the ECMWF model.

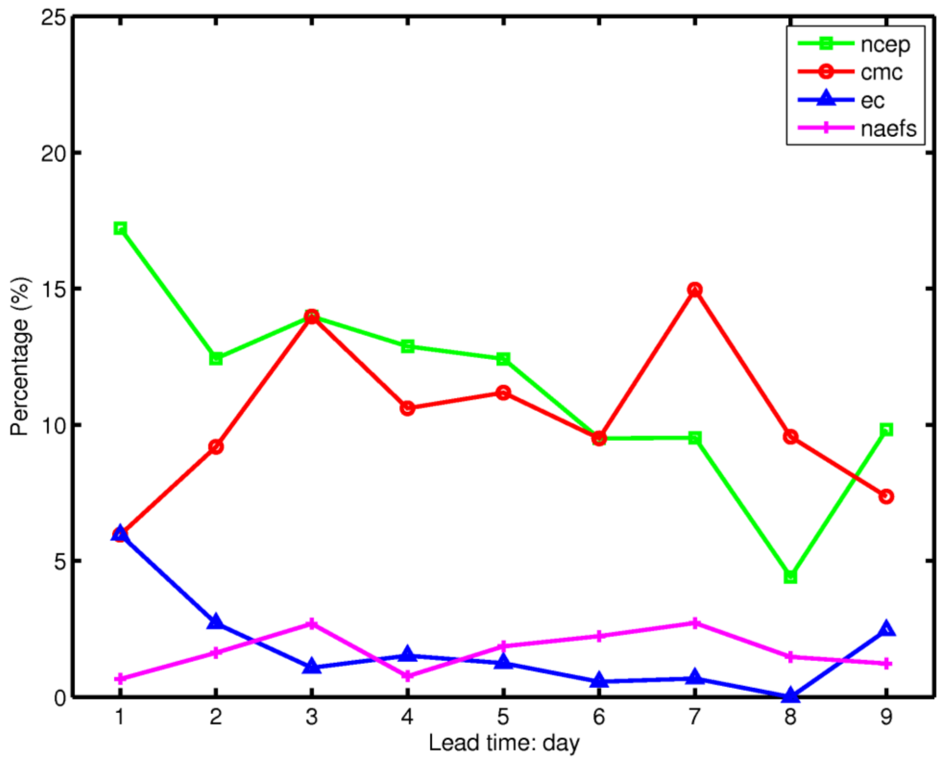


Figure 4.14: The percentage (missing rate) for each EPS and NAEFS that misses the analysis group based on all cases at each lead time. The percentage is calculated by counting the cases with zero member in the analysis group and dividing it by the total case number at each lead time.

In some cases, one or two EPS(s) can have zero number of members in the analysis group. In other words, the EPS(s) fails to predict that particular analysis scenario. We define a missing rate for each EPS to represent the fraction of cases that the EPS fails to predict the analysis scenario. Figure 4.14 shows the missing rate for each EPS as well as a combination of

NCEP and CMC model (NAEFS). For short-range forecasts (days 1−2), the NCEP model has the largest missing rate (>12%) among the three models. The ECMWF model has a missing rate value of ~6% on day 1, but it reduces to less than 3% on day 2. The CMC model also has a missing rate value of ~6% on day 1, but it increases to ~9% on day 2. During the medium-range (days 3−6), the NCEP and CMC models have comparable missing rates (9%-14%), which are much larger than that of the ECMWF model. For the extended range (days 7−9), the CMC has larger missing rate than the NCEP model (except for day 9) while the ECMWF model still has the smallest missing rate (<4%) among the three models. One thing worth noting is that the NAEFS has comparable missing rate as the ECMWF model. It even has the smallest missing rate for short-range, day 4 and day 9 among all groupings considered, suggesting one benefit of combining the NCEP and CMC ensembles is to significantly reduce the missing cases in forecasting winter storms.

To sum up, the percentage of members contributing to the analysis group together with the missing rate can quantify how well the individual EPS captures the analysis scenario. The NCEP model has the highest probability to be included in the analysis group for the short-range forecasts; however, it also has the highest missing rate. This indicates that NCEP model may have less forecast errors but also a smaller ensemble spread for short-range forecasts. The ECMWF ensemble shows the best performance in the medium range with the highest percentage contributing to the analysis group as well as the lowest missing rate among the three EPSs, suggesting its superiority in medium-range forecasts of East Coast storms. The CMC model overall shows the smallest percentage of members contributing to the analysis group and a relatively higher missing rate, suggesting that it is less reliable in terms of capturing the analysis scenario. However, it has lower missing rate than the NCEP model in the short range, indicating that it may have both large forecast error and ensemble envelope for short-range forecasts. The combination of the CMC and NCEP models can reduce the missing rate significantly, demonstrating the value of combining the two models.

4.5.2 Ensemble mean group versus other groups

In general, the ensemble mean forecast is often assumed to be closer to the truth than any of the individual forecasts comprising the ensemble (Leith 1974; Murphy 1988; Whitaker and Loughe 1998). In operational forecast, the ensemble mean is most widely used to represent the best available estimate of the future state of the atmosphere. However, most of the verification of ensemble mean errors are based on the correlation, the mean absolute errors or RMSE. The majority of verification methods are not intuitive and don't help improve the understanding of scenario development. The knowledge of correlation and MAE/RMSE for the ensemble mean (e.g. over the CONUS) is insufficient to be guidance for a forecaster to make a sensible forecast when expecting a severe winter storm. For example, in the 2015 January case, the ensemble mean was forecasting a cyclone closer to shore while a subset of ensemble members forecast a more northeastward cyclone. Under this situation, if the forecaster relied too heavily on the ensemble mean and ignored the other scenarios in the ensemble, a wrong forecast could be made. Therefore, a complementary and efficient verification of the ensemble mean is necessary to investigate whether the ensemble mean is really better than other subsets of an ensemble forecast.

Scenarios based on PC1-PC2 phase space provide a new perspective to verify the ensemble mean. Since the analysis group is a representation of the analysis scenario while the ensemble mean group (Group EM) is a representation of the ensemble mean scenario, it is easy to verify whether the ensemble mean group is similar to the analysis group. If the ensemble mean group includes the analysis scenario more often than the other scenario groups, it will suggest that the ensemble mean does have a better skill than the other subsets of the ensemble in terms of the capability of including real development scenarios.

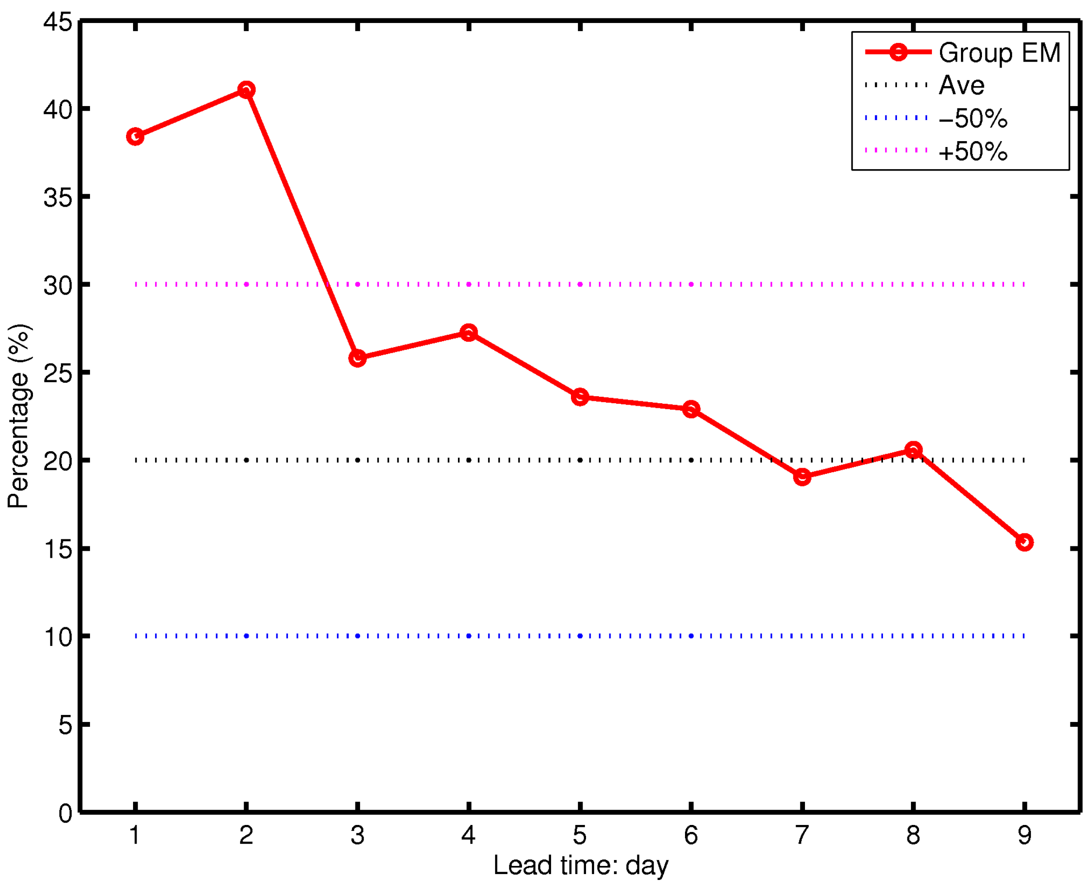


Figure 4.15: The fraction of cases (red line with dot) with the Group EM same with Group ANA for all cases at each lead time. For the five-cluster solution, each group is supposed to have 20% (black dot line) chance (average chance) to be similar with the Group ANA. The blue and magenta dashed lines represent 50% less and 50% more chance than the average chance to include analysis group, respectively.

Figure 4.15 shows the fraction of cases in which the ensemble mean group is The same as the analysis group. For short-range forecast (1−2-day), the ensemble mean group does include the analysis scenario more often (~40%) than the expected average chance (20%) of each group.

This indicates that the ensemble mean scenario is more reliable than the remaining ensemble groups. However, the percentage decreases beyond day 2, becoming only slightly higher than the average percentage in the medium range and even less than the expected average in the extended range. This suggests that the ensemble mean group has no advantage when compared with the other groups in medium- and extended-range forecasts. Even in the short-range forecast, there are still around 60% of cases with the analysis group being different from the ensemble mean group. Therefore, focusing on the ensemble mean too much could be misleading in most cases. The ensemble mean information is far from a sensible guidance for forecasters in operations.

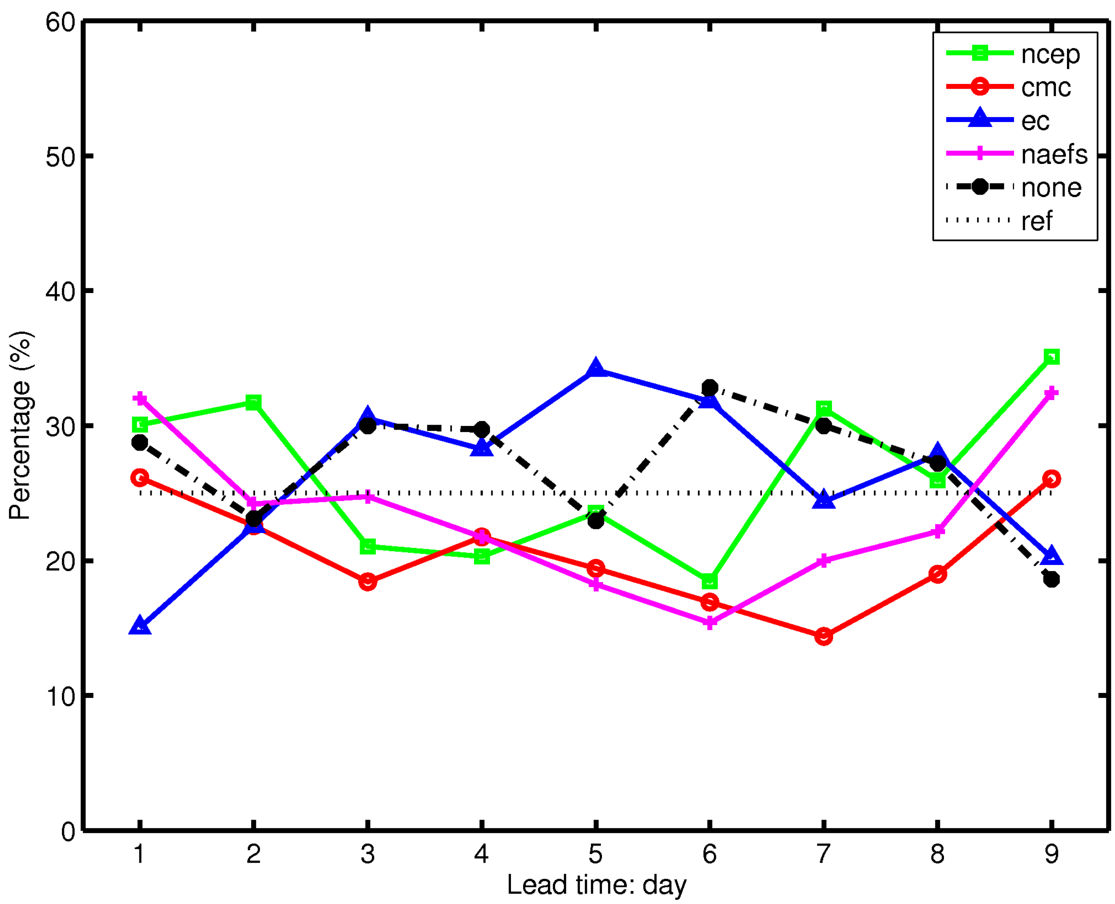


Figure 4.16: The percentage of cases in which the analysis falls into the same quadrant as NCEP, CMC, ECMWF, and NAEF means, as well as in a quadrant in which none of the three EPS means are in.

The above results demonstrate that in the majority of cases, the analysis does not lie within Group EM, similar to the aforementioned two case studies, suggesting that one should not put too much emphasis on the ensemble mean and ignore the other groups.

In operational forecasting, the individual model means are also used a lot, especially the ECMWF mean. Previous studies suggested that the ECMWF model is the “best” in global ensemble model (e.g., Buizza et al. 2005; Keller 2011). Operational forecasters have the tendency to hedge towards the ECMWF forecast direction when the forecast uncertainties are large. Figure 4.15 has shown that in the medium range the analysis scenario is not more likely to be in the Group EM. Here, we examine how often the analysis tends to be in the direction of individual model means on the PC1-PC2 space. We investigate this question by examining whether the projection of the analysis on the PC1-PC2 space is in the same quadrant as the different EPS means. Figure 4.16 shows the percentage of cases in which the analysis falls within the same quadrant as each EPS mean at each lead time. The percentage that the analysis is located outside of any of the EPS mean quadrants is also calculated for comparison. Since there are 4 quadrants, if everything is random, each EPS mean should have a probability of 1/4 (25%) to lie within the same quadrant as the analysis. As can be seen from Figure 4.16, the ECMWF model does show higher chance during the medium range to be in the analysis quadrant. However, even the highest percentage (for day 5) is no more than 35%. The NCEP and NAEFS show slightly higher chance for days 1–2 and day 9 to be in the analysis quadrant. One thing worth noting is that there are a moderate number (~20-30%) of cases in which the analysis lies in a quadrant in which none of the three EPS means fall. The quadrant statistics in Figure 4.16 suggest that although the ECMWF ensemble shows a slightly higher chance to be in the quadrant in which the analysis falls for the medium range, it also misses the analysis direction in around

2/3 of the cases. The above result has practical implications in operational forecast, suggesting that it is not a good practice for forecasters to hedge towards the ECMWF ensemble mean solution as we often see, even though the ECMWF might be the best ensemble. Together with the statistics for Group EM and Group ANA in Figure 4.15, these results indicate that all scenarios must be taken into consideration in the formulation of a forecast.

4.5.3 Error-spread relationship

One widely accepted measure of the utility of an EPS is the relationship between its forecast accuracy and ensemble spread. From Figures 4.13-14, we have seen the capability of different EPSs in capturing analysis scenarios. The performance of an EPS is largely associated with its error-spread relationship or whether the EPS could provide sufficient forecast variability in simulating East Coast storms. In previous studies, the relationship is often simplified by linearly correlating the absolute error of either the ensemble mean forecast or the deterministic forecast with the corresponding ensemble spread of a forecast variable (Kalnay and Dalcher 1987; Murphy 1988; Buizza 1997; Grimit and Mass 2005). However, there is a lack of studies to employ a complementary metric or to compare the error-spread skill for both multi-model and individual models. In this study, the error-spread relation in different individual models as well as the multi-model is revisited under the framework of EOF PCs. The metrics we are utilizing to do cluster analysis are EOF PCs, with the standard deviation normalized to be 1 for each case. To be consistent, we choose this metric to study the spread-error relation. Note that the OOE cases are included in calculating the following statistics in this subsection to avoid biasing the results.

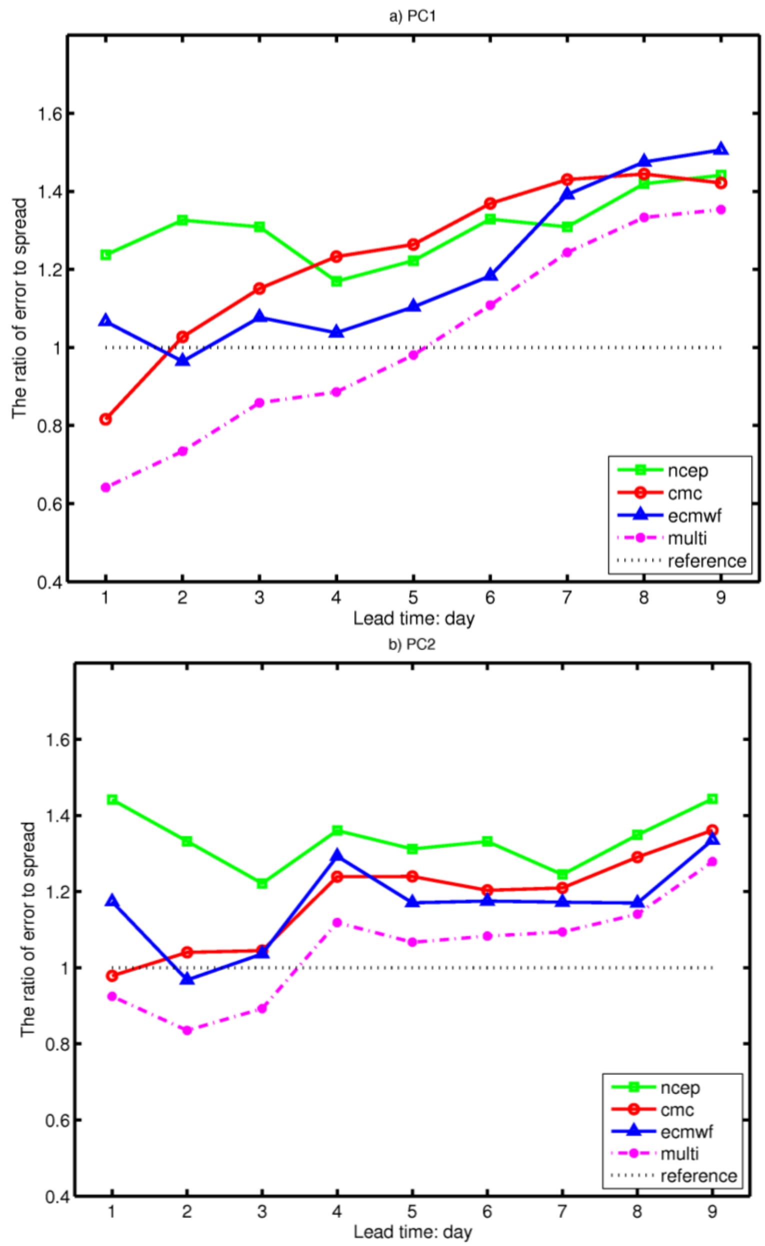


Figure 4.17: The ratio of error to the spread of (a) PC1, and (b) PC2at each lead time for each EPS. The error is calculated based on the distance of analysis relative to each EPS or multi-model mean; while the spread is calculated using all the EPS ensemble members relative to the ensemble mean of each EPS or mulit-model.

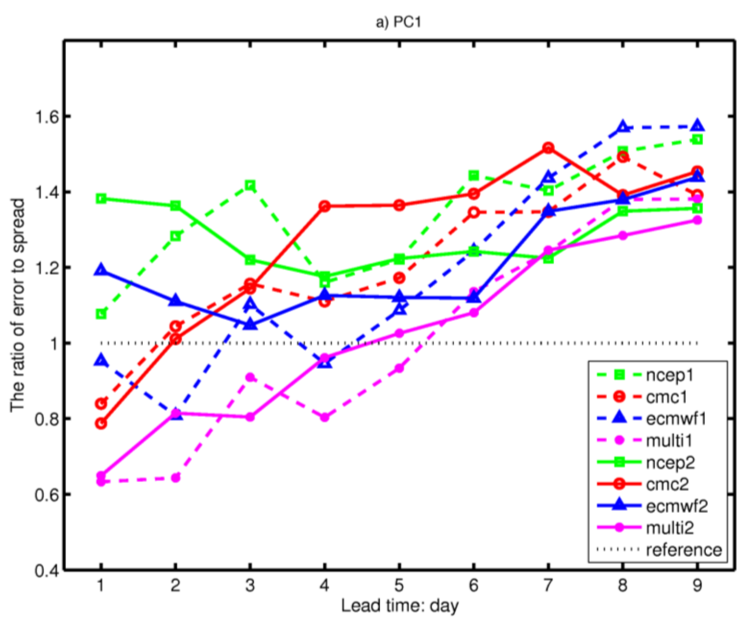
Figure 4.17 shows the error-spread ratios for the leading two EOF PCs based on the observed cyclone cases at each lead time for all three EPSs and the multi-model. The error-spread ratio is calculated by dividing the RMSE of each model mean by the standard deviation of all members with respect to the model mean for all cases. One represents the perfect relation.

The model is considered to be "under-dispersed" if the ratio is greater than 1; otherwise, it is "over-dispersed" if the ratio is less than 1. For short-range forecasts, both PC metrics suggest that the NCEP model is severely under-dispersed (>1.2), which partly explains the larger missing rate shown in Figure 4.14. The CMC model is slightly over-dispersed on day 1 for PC1 and close to the reference line on day 2 for PC1 and on days 1−2 for PC2. The ECMWF model shows slight under-dispersion on day 1 especially for PC2, but it is overall closer to the reference line than the other two models. For medium-range forecasts, both the NCEP and CMC models are more under-dispersed than the ECMWF model for PC1 metric. The NCEP model shows the highest under-dispersion in PC2 while the other two show comparable under-dispersion. As for extended-range forecasts, all three models are severely under-dispersed in PC1 with comparable error-spread ratios. The ECMWF model has less under-dispersion than the other two EPSs in PC2.

For PC1, the multi-model ensemble shows over-dispersion in short- to medium-range forecasts (days 1−4), a nearly perfect relation on day 5, and increasing under-dispersion after day 6. For PC2, it also shows over-dispersion for forecasts before day 4. Between day 4 and day 7, it is slightly under-dispersed. However, it becomes more under-dispersed after day 7. For the extended range, the error-spread skills for all ensembles converge towards too low of a spread, which is consistent with previous studies (e.g. Park et al. 2008). Overall, the multi-model ensemble shows a much lower under-dispersion for medium- and extended-range forecasts, suggesting the benefit of combining different EPSs in the medium and extended range.

To sum up, the NCEP model is severely under-dispersed in the short range for both PC metrics, suggesting the NCEP model may not have enough ensemble dispersion. Connecting this with its higher percentage of members contributing to the analysis scenario (Figure 4.13), a preliminary conclusion is that the NCEP model has less forecast errors but a narrow ensemble spread in the short-range. In contrast, the CMC model is slightly over-dispersed or close to reference line during the short range. Since it has lower chance to be included in the analysis group, the CMC model seems to have large forecast errors but also a broad ensemble. The ECMWF model has better error-spread relationship during the medium range, demonstrating its superior performance in the medium range which is consistent with Figures 4.13-14. Note that during the earlier forecast time (1–3 days), the differences in the error-spread skills for the three models are larger, which could be due to their differences in generating ensemble perturbations (Table 2.1). The ECMWF model has been using the singular vector method for perturbing initial conditions (Molteni and Palmer 1993; Buizza and Palmer 1995). Previous studies found that the ECMWF spread is small and grow fast during the short range (Trevisan et al. 2001; Park et al. 2008; Buizza et al. 2005), which partly explains why the ECMWF model is more underdispersed on day 1 than on day 2 in Figure 4.17. Overall, the best error-spread skill is obtained by the ECMWF ensemble for it has the best data assimilation/modeling components (Buizza et al. 2005). The NCEP ensemble model uses the breeding vector approach for generating perturbations. The basic assumption is that the fast-growing perturbations naturally develop in a data assimilation cycle and will continue to grow as short- and medium-range forecast errors (Toth and Kalnay 1993). Previous studies found that the NCEP model exhibits the lowest (and least realistic) perturbation growth and the smallest spread beyond the first day (Buizza et al. 2005), which perhaps explains the overall underdispersive characteristics for both short- and medium-range forecasts. The CMC EPS uses the operational ensemble Kalman filter (EnKF) to generate initial conditions (Houtekamer et al. 2009; Houtekamer et al. 2014). An ensemble of trajectory assimilates randomly perturbed observations, and uses different model versions. A number of ensemble analyses are generated, among which 20 representative members are chosen to initialize the global ensemble forecasts. Hagedorn et al. (2012) found that CMC EPS starts with a large spread but exhibits a serious mismatch in the growth of the spread and error, which grows with lead time. This partly explains the largest increase in the error to spread ratio of the CMC EPS when compared with the other two models. The relatively good performance of the CMC system beyond 7 days may be due to the use of both perturbed observations and models in that ensemble (Buizza et al. 2005).

On the other hand, the multi-model (NCEP+CMC+ECMWF) ensemble shows over-dispersion in short-range forecasts, suggesting a combination of the three models is over-dispersed. Beyond day 4, it is under-dispersed like the individual models. However, it is closer to the perfect value than any individual model is, suggesting the benefit of combining different EPSs to provide more forecast variability in the medium and extended range. The above results are consistent with the findings by Hagedorn et al. (2012) who analyzed using forecasts of 850-hPa temperature, 2-m temperature and 500 hPa geopotential in the extratropics.



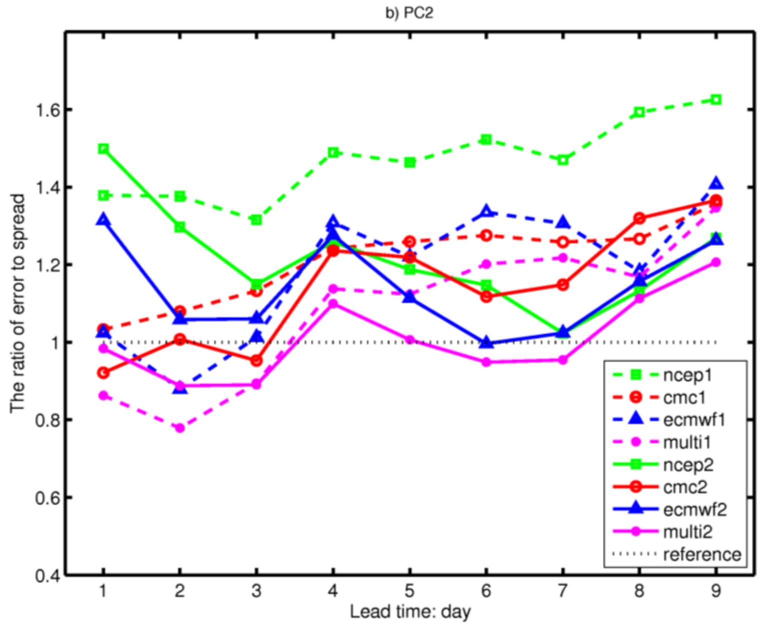


Figure 4.18: The same as Figure 4.17 but for two time periods during the examined time. Dashed lines with markers are for the cases during the first half period while the solid lines with markers are for the second half period.

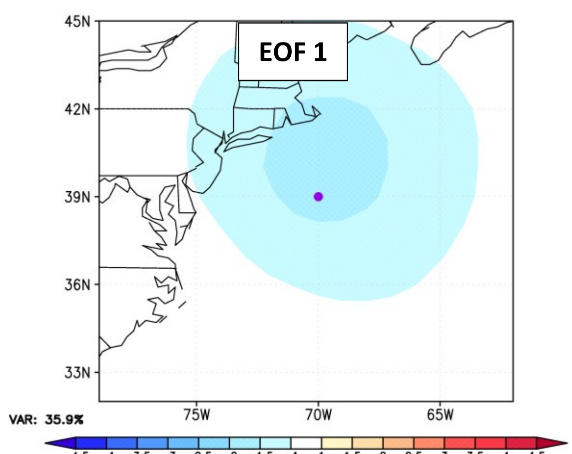
Since all three models experienced updates during the time period we examined, the error-spread ratios have been recalculated for two different periods: the first half (2007/08 to 2010/11) and the second half period (2011/12-2014/15), separately. Figure 4.18 shows the error-spread ratio for the two periods. For the PC1 metric, the error-spread ratio for short-range forecast doesn't show large differences between the two periods except that the NCEP model becomes even more under-dispersed. Beyond day3, the NCEP and ECMWF models show improved error-spread relation in the second period while the CMC becomes more under-dispersed. The multi-model showed slight improvement during the second period with the curve closer to the reference line except for day 3. For PC2, all three models as well as the multi-model show improvement beyond day 1 with the curves much closer to the reference line during the second period than during the first period. The NCEP model shows the largest improvement among the three models. Therefore, the error-spread relation in PC2 shows larger improvement in these two metrics. In the next subsection, we will show that PC2 is mainly associated with the southwest-northeast shifting of the cyclone position. The above results suggest that the models improved the error-spread relation in forecasting cyclone southwest-northeast track uncertainties especially for the NCEP model.

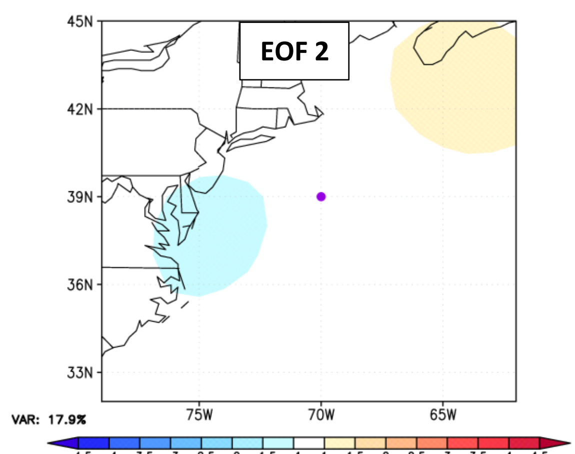
4.5.4 Physical interpretation of model errors based on PC metrics

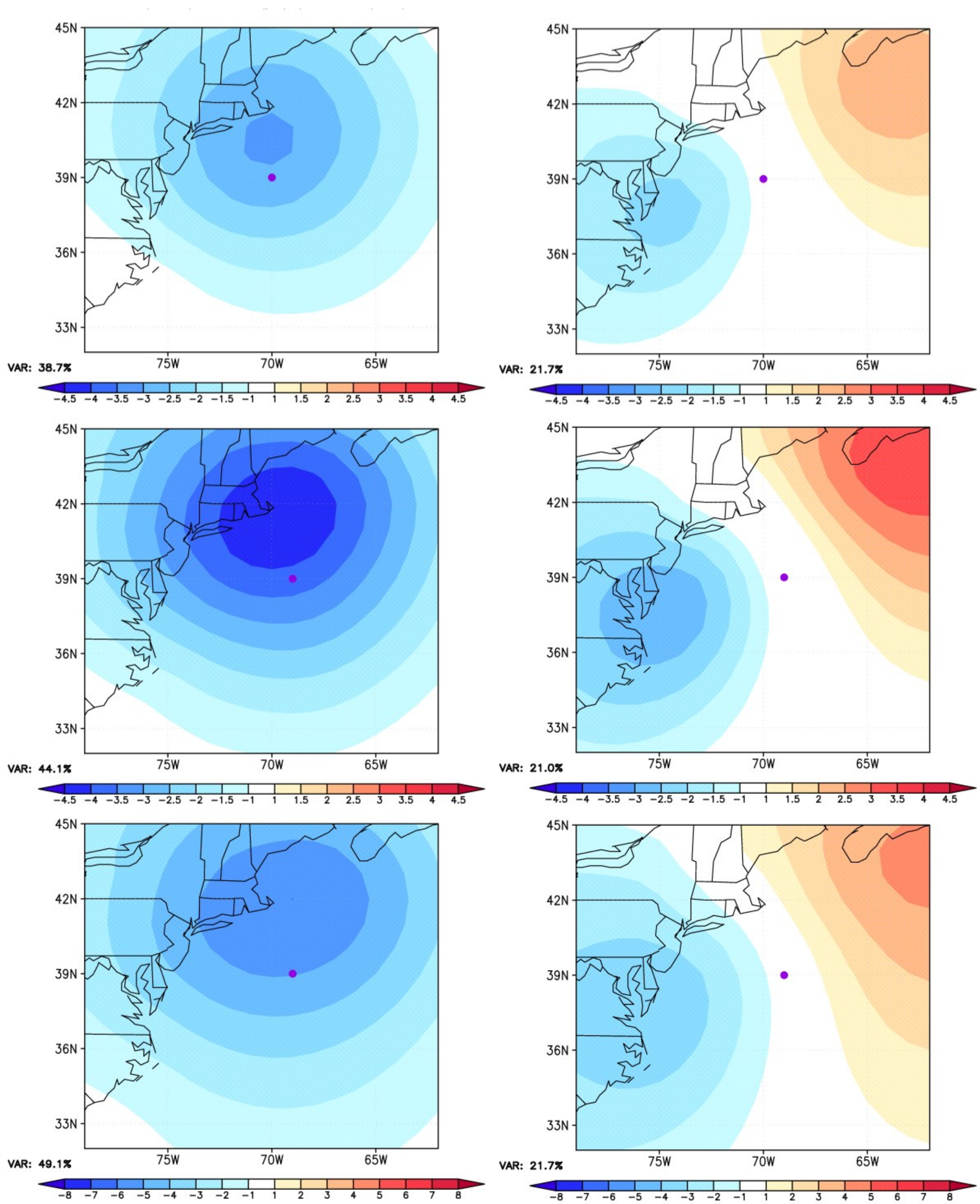
One benefit using EOF analysis is the orthogonal property of EOF patterns, which can also be used to perform the error decomposition. To decompose the forecast uncertainty in the multi-model ensembles, we first computed each member's anomaly with respect to the ensemble mean for each case, and repeated this calculation for all cases. An EOF analysis was then calculated on the combined anomaly for all members and all cases. The leading patterns represent the dominant forecast uncertainties in all cyclone forecasts.

Figure 4.19 shows the dominant uncertainty patterns in all ensemble members in forecasting the observed cyclone cases from day 1 to day 6 forecasts. The leading pattern (EOF1) tends to be a monopole at all forecast times, which is a bit northward (westward for day 6) of the ensemble mean cyclone (purple dot). This pattern suggests the main uncertainty in forecasting East Coast storms is the intensity together with a slight shift of the center in the north-south direction. The variance explained varies from 35.9% (for day 1) to 56.6% (for day 6). The second pattern (EOF2) is always a dipole orthogonal to the first pattern, which mainly represents the southwest and northeast shift in cyclone position. The explained forecast variance by the leading two patterns increases from 60.4% for day3 to 76.6% for day 6. Therefore, the leading two patterns represent most of the forecast uncertainties.

The systematic errors on each EOF direction can also be calculated. On the above EOF PC1-PC2 space, each member occupies one point for one case. When the analysis is projected onto the corresponding EOF1 and 2 patterns, it also occupies one point for each case. The difference in the abscissa or the ordinate between each member and the analysis reflects the relative error for that member in EOF PC1 or PC2. By averaging the differences between the members from one EPS and the analysis, the systematic error associated with EOF1/2 pattern corresponding to that EPS can be evaluated. Though the differences calculated in PC1 or PC2 is dimensionless, the EOF patterns have the dimension of MSLP. Each EOF pattern represents an anomaly with respect to the ensemble mean. Therefore, the PC1/PC2 errors can be converted back to the dimensional errors since they are associated with the EOF anomaly. The errors in cyclone forecasts can thus be decomposed into errors associated with the EOF1 or EOF2 pattern. Since each of the two patterns is two-dimensional, it can be further decomposed into cyclone intensity (minimum pressure) and displacement errors. The above decompositions of forecast errors associated with the leading two patterns are summarized in Figures 4.19-20 to provide a physical understanding of the model errors based on the EOF PC1-PC2 space.






45N
42N
39N
36N
33N
75W
70W
65W
VAR: 38.7%
VAR: 21.7%
−4.5 −4 −3.5 −3 −2.5 −2 −1.5 −1 1 1.5 2 2.5 3 3.5 4 4.5
VAR: 44.1%
VAR: 21.0%
VAR: 49.1%
VAR: 21.7%
−8 −7 −6 −5 −4 −3 −2 −1 1 2 3 4 5 6 7 8

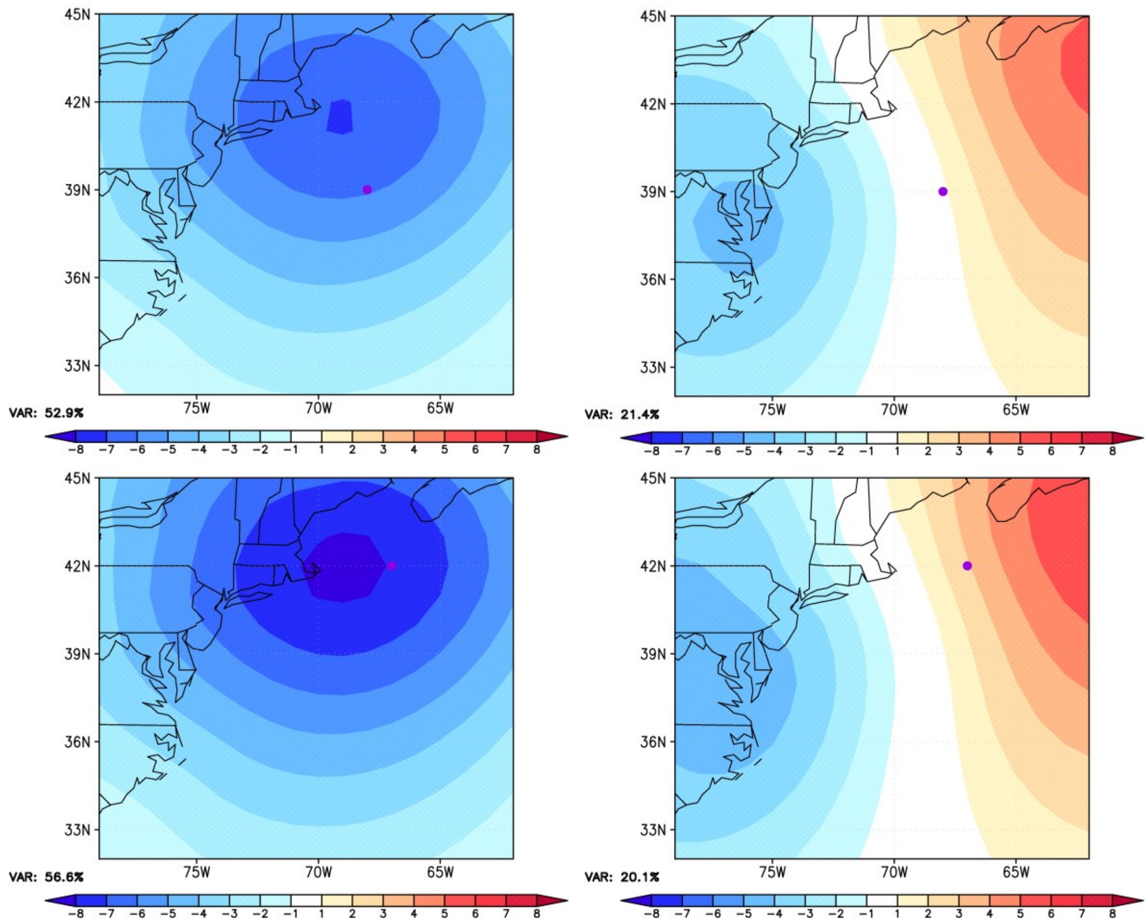


Figure 4.19: The first (left panel) and second (right panel) EOF pattern (shaded) in the multi-model ensemble for day 1 to day 6 forecast (top to bottom). The purple dot represents the ensemble mean cyclone position averaged over all cases for each lead time.

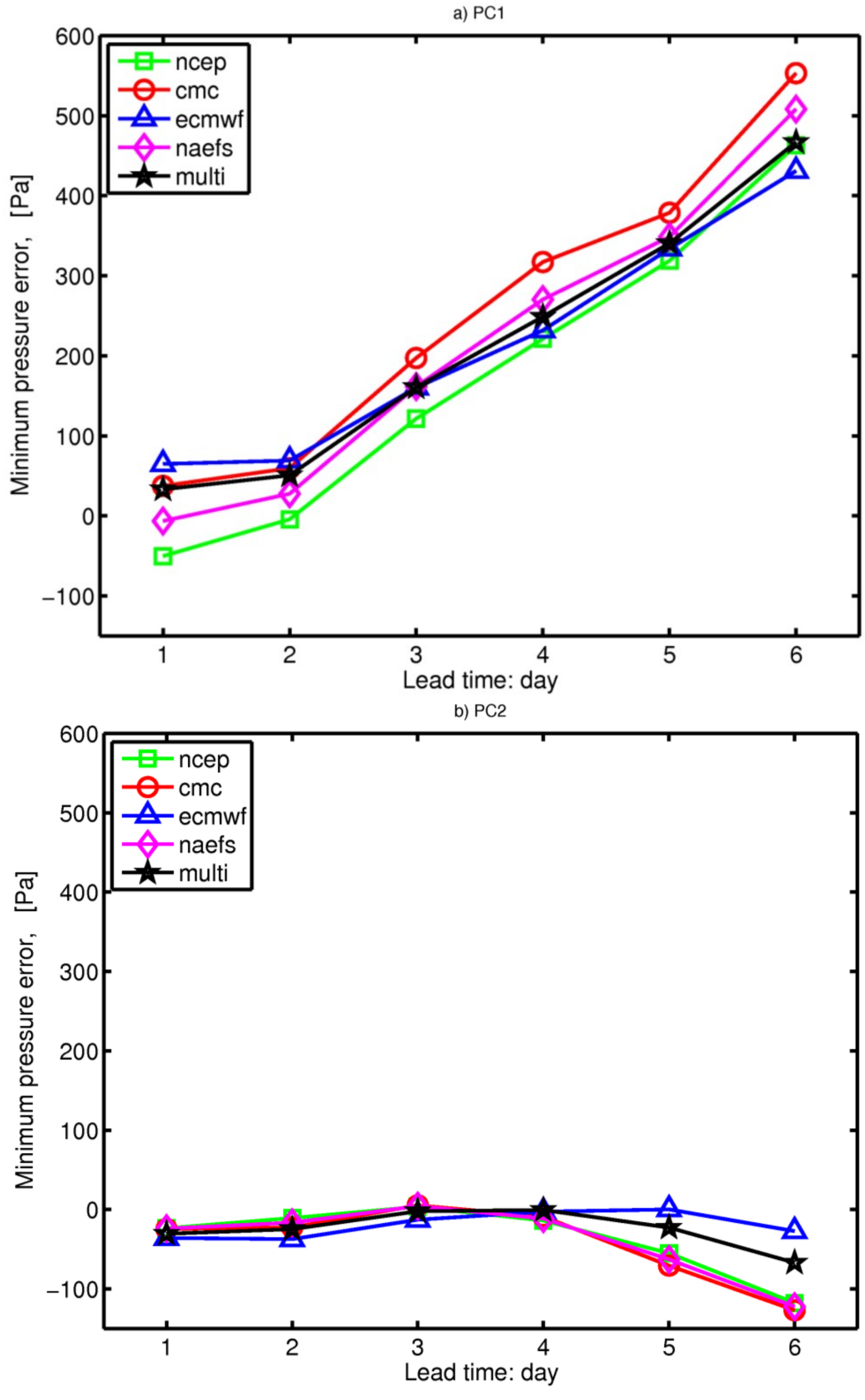
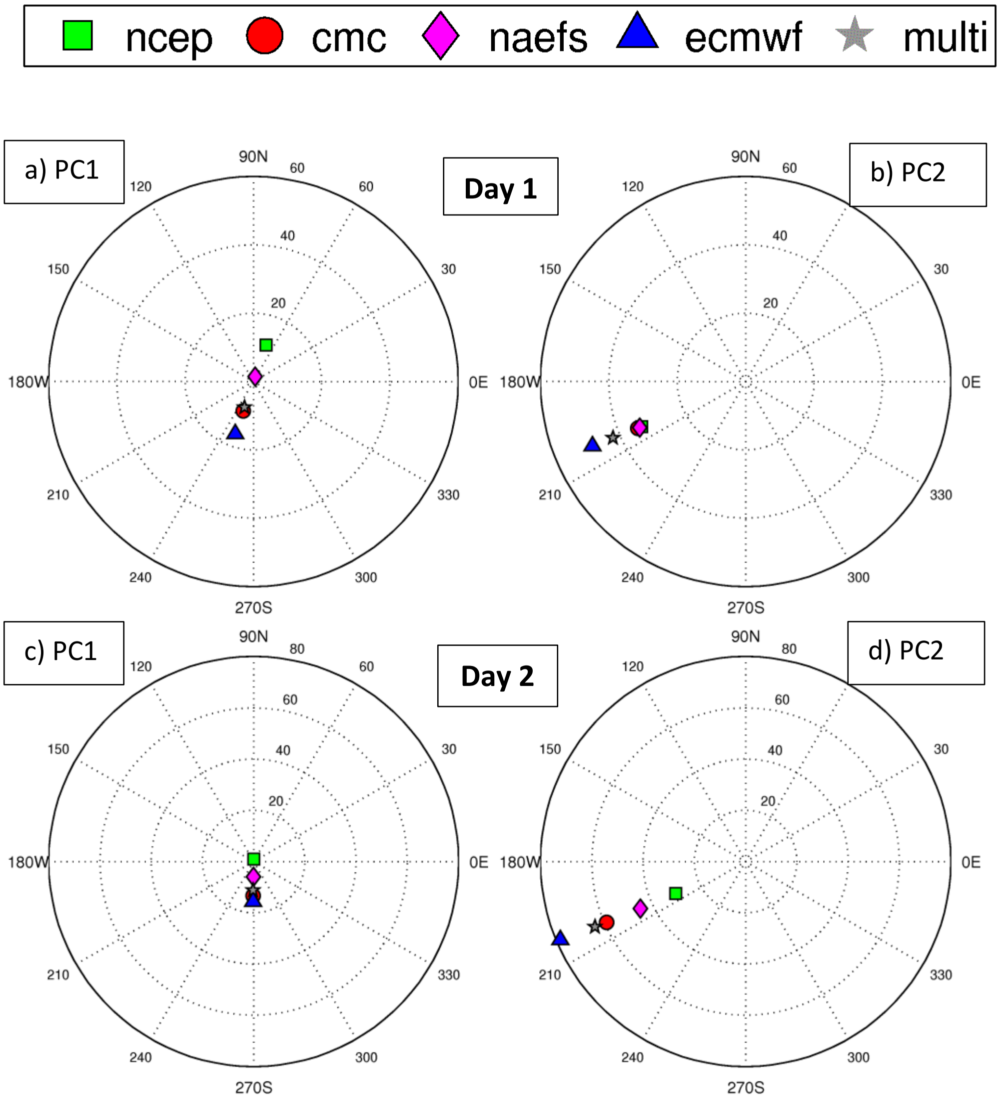
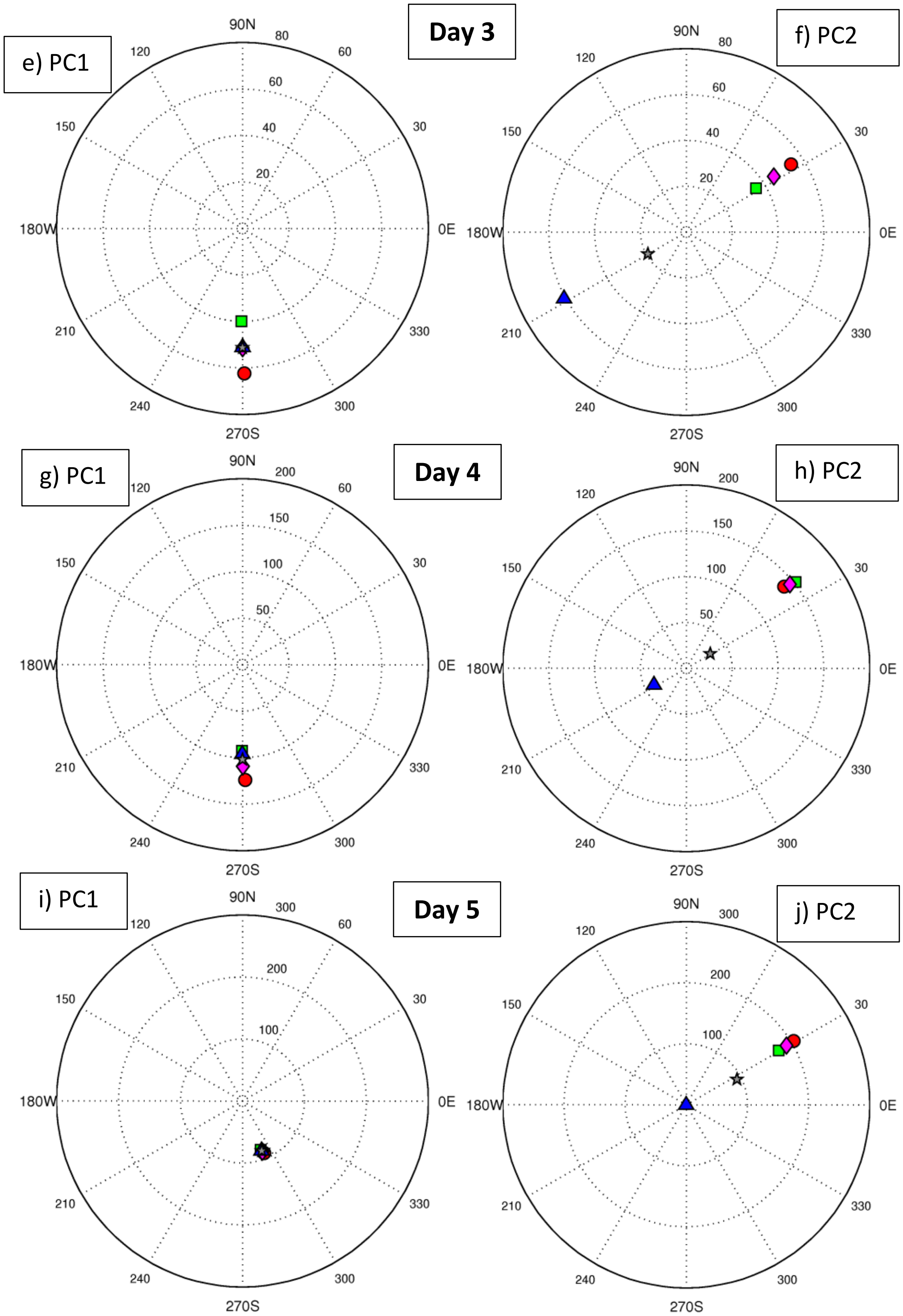


Figure 4.20: Mean minimum pressure error (unit: [Pa]) of each EPS and the combined model for PC1 (a) and PC2 (b) errors from day 1 to day 6 forecast.

Figure 4.20 illustrates the minimum pressure error for each EPS as well as the combined EPSs from day 1 to day 6 forecasts derived from each EOF PC errors. The minimum pressure errors for the EOF1 patterns are much larger than those for the EOF2 patterns, which is consistent with Figure 4.19. Figure 4.21 depicts the displacement errors for each EPS associated with the leading two patterns. For the errors associated with the EOF1 pattern, the ECMWF model shows slightly larger positive intensity errors and displacement error toward the south or south-southwest direction than the other models for day 1 and day 2 forecasts (Figures 4.20a, 4.21a, c). In contrast, the NCEP model has negative intensity errors and is located more toward the north than the other models. The CMC, NAEFS, and the three-model combined ensemble have the errors between the above two. For the medium range (3−6 days, Figure 4.20a, 4.21e, g, i, k), all models have positive intensity errors, suggesting an under-prediction of the observed East Coast winter storms. Among them, the CMC forecasts have the largest intensity and displacement errors toward the south (southeast for day 6) while the NCEP forecasts have the smallest intensity and displacement errors during days 3 and 5 and the ECMWF model has the smallest errors for day 6 forecasts. Throughout day 1 and day 3 forecasts, the NCEP model tends to be more northward than the other models, indicating their less southward bias.

As for the EOF2 pattern, the overall intensity errors for all EPSs (Figure 4.20b) are small and negative, which again confirms the second EOF pattern often represents the position shift uncertainty instead of intensity uncertainty. During days 1−3 (Figures 4.20b, 4.21b, d, f), the ECMWF has the largest displacement error toward the west-southwest while the NCEP model has the least errors. During day 4-6 (Figures 4.20b, 4.21h, j, l), the ECMWF model has the least intensity and displacement errors among all models. Note that when comparing with the other models, the ECMWF model is always to the west-southwestward direction, suggesting that when forecasting East Coast winter storm, one tendency of the ECMWF model is to forecast them too west-southwestward than the other two models. This could be partly due to the slow propagation in the ECMWF model as suggested by Froude et al. (2010).




Day 3
e) PC1
f) PC2
Day 4
g) PC1
h) PC2
Day 5
i) PC1
j) PC2
90N
270S
180W
0E
30
60
120
150
210
240
300
330
20
40
60
80
50
100
150
200
300

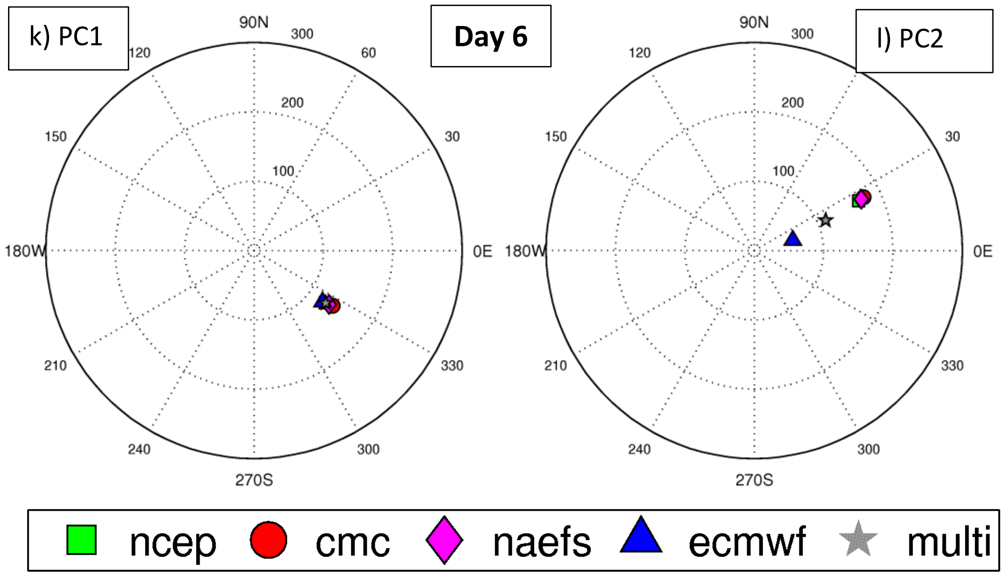


Figure 4.21: The displacement errors corresponding to averaged PC1 (left) or PC2 (right) errors for each EPS as well as the combined models for day 1–6-day forecast. The radius has unit of kilometer (km).

Here we want to emphasize that the forecast errors shown in Figures 4.20-21 only provide a qualitative understanding of the model errors. The simplification by extracting the leading two forecast uncertainty patterns is already an approximation in representing the forecast information. The mean errors of each EPS associated with the leading two patterns provide a gross tendency diagnosis of the biases present in the different EPSs.

Note that in calculating the forecast errors (as well as for the error/spread relation discussed in the previous subsection), all cyclone cases and ensemble members are included in the calculations, hence we believe that the evaluations conducted in this study can provide results that are complementary to those provided by cyclone matching (Froude 2010; Korfe 2016) that can directly verify forecast errors in intensity and displacement, but can only verify cases in which many ensemble members can be matched to the observed cyclone, as well as verifying only the ensemble members that can be matched, which represent a reducing fraction of all cyclone forecasts as the lead time is increased.

4.5.5 Evaluations based on predicted cyclones

To confirm the results of evaluations for the three EPSs using observed cyclone cases, similar statistics were calculated using the predicted cyclone cases based on the NCEP deterministic forecast. The number of cases used for different lead times are listed in table 4.4.

Lead Time [day]	1	2	3	4	5	6	7	8	9
Case number	140	143	145	119	120	119	123	110	100

Table 4.4: The number of predicted cyclone cases used for each forecast day.

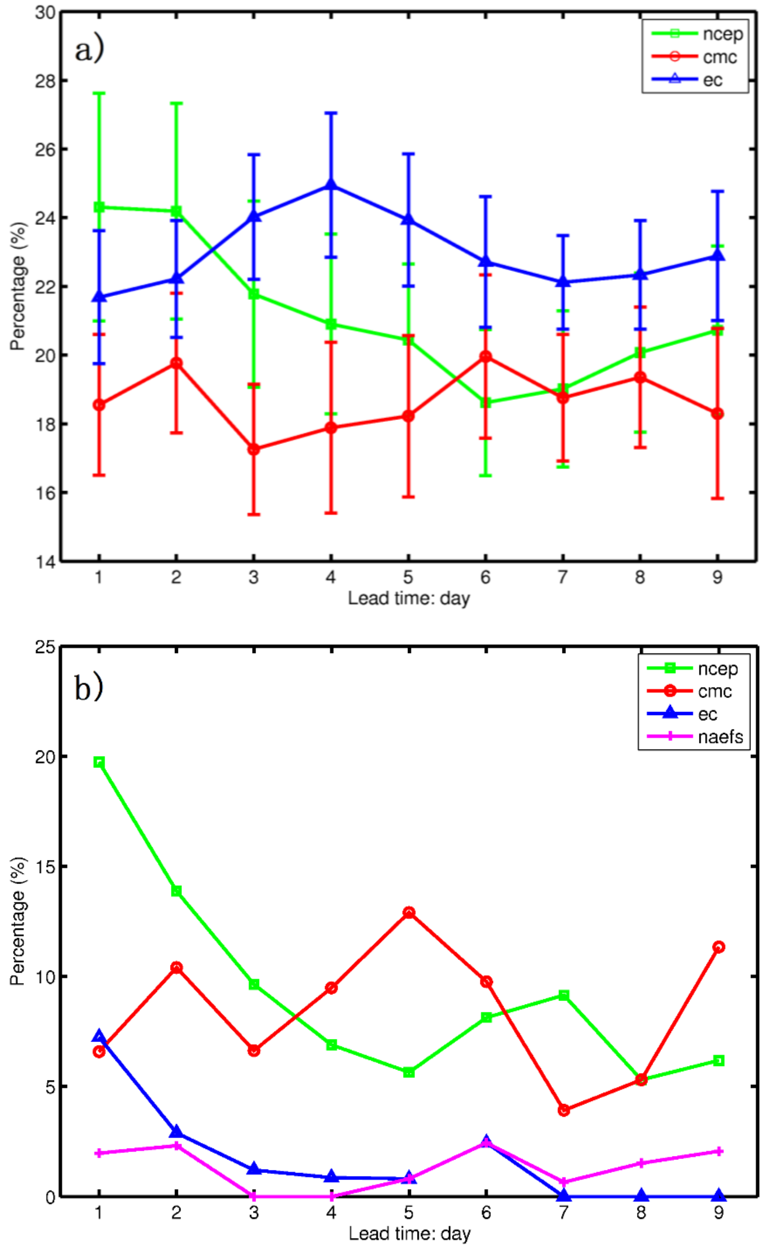


Figure 4.22: (a) The same as Figure 4.13 but for the predicted cyclone cases; (b) The same as Figure 4.14 but for the predicted cyclone cases.

Figure 4.22 shows the percentage of members contributing to the analysis group and the missing rate for each model. For the short range (days 1–2, Figure 4.22a), the NCEP members have the highest chance to be in the analysis group while the CMC members have the lowest chance. During the medium- and extended-range (Figure 4.22a), the ECMWF members have significantly larger chance in contributing to the analysis group. Overall the CMC members have the lowest chance to be included in the analysis group. Meanwhile, NCEP has the largest missing rate during days 1–3 while ECMWF shows much smaller missing rate during medium- and extended-range (Figure 4.22b). The combination of NCEP and CMC (NAEFS) significantly reduces the missing rate and even shows smaller rate during days 1–4 when compared to the ECMWF model (Figure 4.22b). Therefore, the scenario-based statistics overall are consistent with the statistics using the observed cyclone cases in section 4.5.1. The differences between the observed cases and predicted cases occur during day 7-9, which could be due to the large differences in sample size.

We have also calculated other statistics such as the group EM vs. Group ANA, the quadrant statistics, and the error-spread ratios using the predicted cyclone cases (not shown). The statistics are grossly consistent with results in subsections 4.5.2-3 especially during days 1–5, demonstrating the results for the evaluations in this study are robust for short- and medium-range forecasts.

4.6 Summary of outlier cases

While for an ensemble forecast, a rate of outliers, i.e., the analysis lying outside of the ensemble envelop, is expected, the outlier rates are of paramount importance to ensemble verification since they represent the poorest forecast skills in NWP. Since our study is based on the PC phase space, the outlier or OOE cases are also defined based on the PC1-PC2 space. As stated in section 4.4.2, the outlier cases or OOE cases are defined by the condition that the distance between the analysis and the closest member is significantly larger than the average distances between any two ensemble members.

The fraction of OOE cases for each lead time is shown in Figure 4.23. During the lead times of 1- to 3-day, there are only less than 3% of OOE cases, which is consistent with the overall over-dispersion characteristic of the multi-model ensemble (Figure 4.17). It increases to 5% on day 5 and reaches 8% for 6- to 7-day. The fraction of OOE increases to ~14% on day 8 and drops slightly on day 9. Overall, the OOE fraction increases with lead time in medium- to extended-range, which could be associated with the increasing error-spread under-dispersion (Figure 4.17).

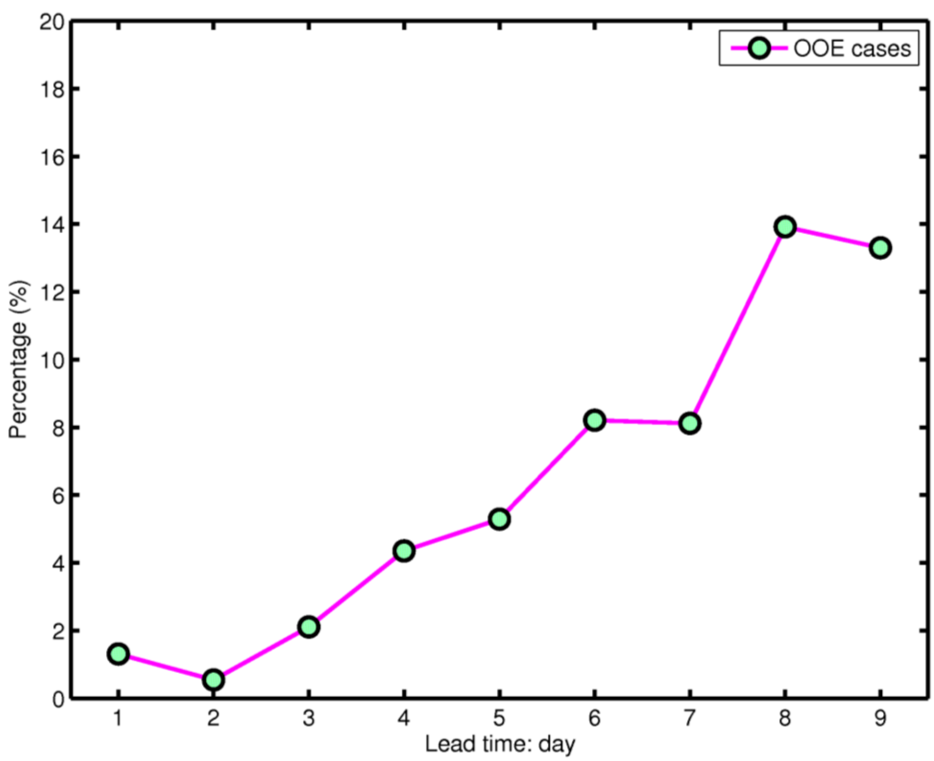


Figure 4.23: The fraction of OOE cases to all observed cyclone cases for each forecast lead time.

Table A4.1 in Appendix 1 listed all the OOE cases for 1−6-day lead times in this study, including their forecast lead time, initial/verification time, and the verified weather event. Figure 4.24 shows the counts of the OOE cases at all lead times for each cool season. It is shown that 25 out of 38 OOE cases happened during the first half of the examined years (before 20011/2012, Figure 4.24). Eighteen of them happened in 2008/2009 and 2009/2010 cool seasons. The improvements of model resolution, data assimilation method and physical parameterization schemes (Table 2.1) could have brought about the overall reduction of OOE cases in the second half of the examined years. However, since the ensemble runs are flow dependent, the OOE cases can be also associated with very poor predictability of some types of synoptic flow patterns while NWP models remain the same. Thus, we examined the details of the 13 OOE case for 1−4-day forecasts and the composited flow patterns for the 25 OOE cases for 6-day forecasts. The results are summarized in Appendix 1.

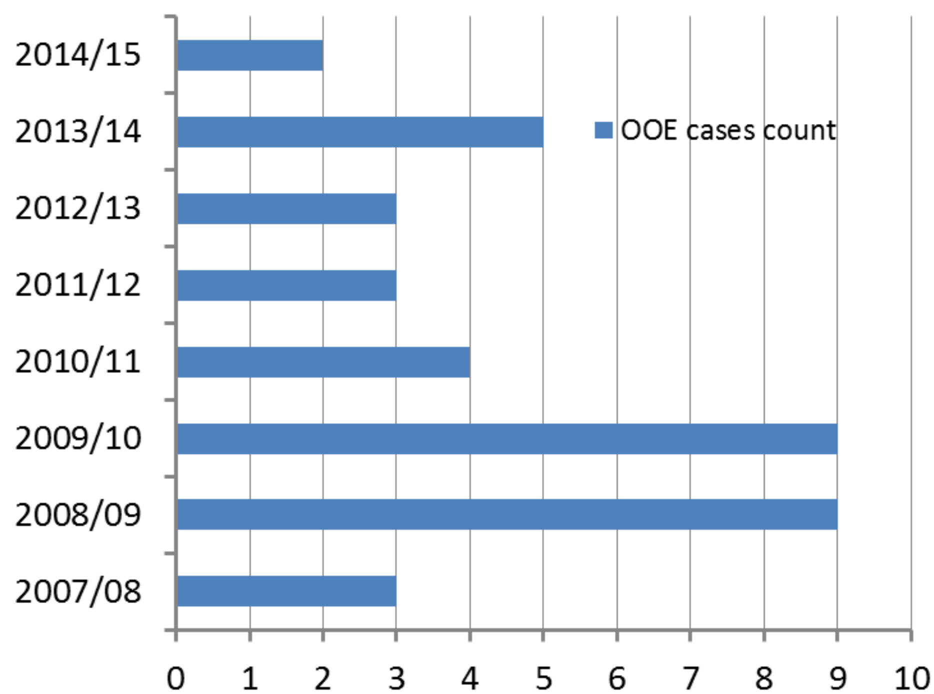


Figure 4.24: OOE cases count for each cool season based on all the OOE in table A4.1 in Appendix 1.

4.7 Chapter summary

This chapter has presented the application of a fuzzy clustering method to diagnose forecast scenarios in operations and to evaluate the capability of different EPSs in forecasting historical East Coast winter storm cases.

It has been shown that the clustering method can be applied to efficiently separate the scenarios in forecasting East Coast storms based on a 90-member multi-model ensemble. Application to the 26–28 January 2015 storm event demonstrates that the forecast scenarios determined by the fuzzy clustering method are well separated and consistent in different state variables (i.e., MSLP, Z500, and total precipitation). The robustness of performing cluster analysis based on MSLP has been also confirmed by computing a Rand Index between clusters using different variables.

Building upon the successful application of EOF/fuzzy clustering to operational cyclone cases, we have also proposed a scenario-based ensemble verification method by projecting the analysis onto the EOF PCs coordinate and finding the closest cluster to represent the analysis scenario. We have applied this tool to examine the capability of different models in capturing the analysis scenarios for historical East Coast cyclones. The NCEP model is found to perform better in the short range in terms of capturing the analysis scenario although it is under-dispersed. The ECMWF ensemble shows the best performance in the medium range, suggesting its superiority in medium-range forecasts of East Coast storms. The CMC model overall shows the smallest percentage of members contributing to the analysis group and a relatively high missing rate, suggesting that it is less reliable in terms of capturing the analysis scenario. One thing worth noting is that the combination of NCEP and CMC can significantly reduce the missing rate. In addition, the combination of the three models is found to improve the error-spread relation, demonstrating the benefit of combining models in reducing missing rate and alleviating the under-dispersion problem especially in the medium- to extended-range.

Since the ensemble mean of a multi-model or individual models (in particular the ECMWF model) is most widely used by the operational forecasters to represent the best available estimate of the future state of the atmosphere, whether the ensemble mean from the multi-model or individual models is really better than other subsets of an ensemble forecast has also been investigated based on the scenario analysis or the quadrant statistics. It is found that in the majority of cases, the analysis does not lie within Group EM. The quadrant statistics suggest that although the ECMWF ensemble shows a slightly higher chance to be in the quadrant in which the analysis falls for medium-range forecasts, it misses the analysis direction in around 2/3 of the past storm cases. The above results have practical implications in operational forecast, suggesting that: 1) one should not put too much emphasis on the ensemble mean and ignore the other groups in the multi-model ensemble; 2) it is not a good practice for forecasters to hedge towards the ECMWF ensemble mean solution as we often see, even though the ECMWF might be the best ensemble. All scenarios must be taken into consideration in the formulation of a forecast.

The forecast uncertainty patterns in forecasting winter storms based on the multi-model have also been investigated. For the 1−6-day forecasts, the first pattern is a monopole within the verification region, suggesting the dominant uncertainty is associated with the variation in forecasting cyclone intensity. The second pattern is a dipole along the west-east or southwest-northeast direction, representing the variation in forecasting cyclone positions. For the 1−3-day forecasts, the NCEP model tends to better represent both EOF1 and EOF2 patterns by showing less intensity and displacement errors. The ECMWF model has the largest bias in both EOF1 and EOF2 for days 1−2. For the 3−6 forecasts, the CMC model has the largest errors in representing both EOF1 and EOF2 patterns. Meanwhile, the ECMWF model is found to have the smallest errors in both EOF1 and EOF2 patterns during days 3−6. Therefore, as the acknowledged "best" ensemble among global EPSs, the ECMWF only starts to show superiority beyond day 3. One interesting point is that our results suggest that East Coast cyclones in the ECMWF forecasts tend to be towards the southwest of the other two models in representing EOF2 pattern, suggesting ECMWF model may have a tendency to show a closer-to-shore solution in forecasting East Coast winter storms.

This page is left to be blank intentionally

Chapter 5 Predictability associated with Rossby wave packets

In Chapter 3, we have found that ensemble sensitivity signals sometimes develop and propagate like an RWP with downstream development (Chang et al. 2013). Some OOE cases determined in Chapter 4 (see Appendix 1) are also associated with the presence of RWPs. Thus, another goal of this study is to find the connection between RWPs and the predictability associated with the rapid downstream error propagation. As early as the 1990s, RWPs in the middle and upper troposphere were considered to affect extratropical cyclones which often bring high-impact weather in winter. For instance, Orlanski and Sheldon (1995) proposed an energetic framework of downstream baroclinic evolution, and examined the trough formation over the east coast of North America associated with the “Blizzard of 93”. They concluded that the formation of the trough was primarily due to the downstream propagation of Rossby wave energy from disturbances over the northeast Pacific. Moreover, they pointed out that under environmental conditions in which the downstream development leads to packets of eddies with a life span much longer than individual eddies, the detection of such packets and energy transfer processes may impact the longer range predictability of such systems. Chang and Yu (1999) found that wave packets appear to be most coherent across southern Asia, extending all the way across the Pacific, into North America, across the mid-Atlantic, across North Africa, and back toward southern Asia. This band is referred as the primary wave guide for synoptic-scale wave packets. There are downstream-developing wave trains propagating from the eastern end of the Pacific storm track across North America toward the Atlantic, suggesting that wave developments near the entrance of the Atlantic storm track are influenced by waves from the Pacific. Hakim (2003) showed that the Atlantic storm track is often seeded by wave packets that originate over the western North Pacific by studying a sample of North Atlantic storm track events.

Forecast errors also tend to develop wave packet-like structures (Hakim 2005), and propagate rapidly downstream with the group velocity (at least as fast as the upper level flow). Model errors during the "Surprise" January 2000 East Coast snowstorm also propagated with an RWP from the Pacific (Langland et al. 2002). Moreover, when a coherent RWP is present and the flow is zonal, distinctive targeted regions, which are associated with maximum forecast error reduction within verification regions, can be traced upstream to near Japan at lead times of 4−7 days (Majumdar et al. 2010). Rodwell et al. 2013 investigated the forecast busts of the ECMWF day-6 forecasts over a European verification box and found that the busts are associated with a Rossby wave train that crossed the Pacific, suggesting the RWPs might enhance the medium-range forecast errors and the ensemble spread over the verification region. In this work, we hypothesize that model errors are generally enhanced when wave packets are present upstream. To be more specific, we hypothesize the RWPs originating from the Pacific Ocean can degrade the predictability of the eastern North America and western Atlantic in the medium range.

We start with the RWP envelope characteristics of large forecast error (dropout) cases in Section 5.1. In Section 5.2, the corresponding RWP anomaly in large ensemble spread cases will be presented. As a complementary perspective, Section 5.3 proposes a conceptual model about the errors/uncertainties relative to a coherent RWP and examines the corresponding forecaster errors and uncertainties under a coherent RWP case. Ensemble sensitivity based on the same coherent RWP will be examined in Section 5.4. The mechanism of downstream development for ensemble errors/uncertainties will be summarized in Section 5.5. Section 5.6 concludes this chapter.

5.1 Dropout cases and RWP

Dropout cases here refer to large forecast error cases over a prescribed verification region (Figure 2.2) for 300 hPa geopotential height day-7 forecasts from the NCEP GFS deterministic forecasts. Figure 5.1 shows the time series for the dropout cases used in this section. Subsection 5.1.1 will explore the RWP anomaly in the dropout cases. As a comparison, subsection 5.1.2 will investigate the RWP anomaly in good forecast cases.

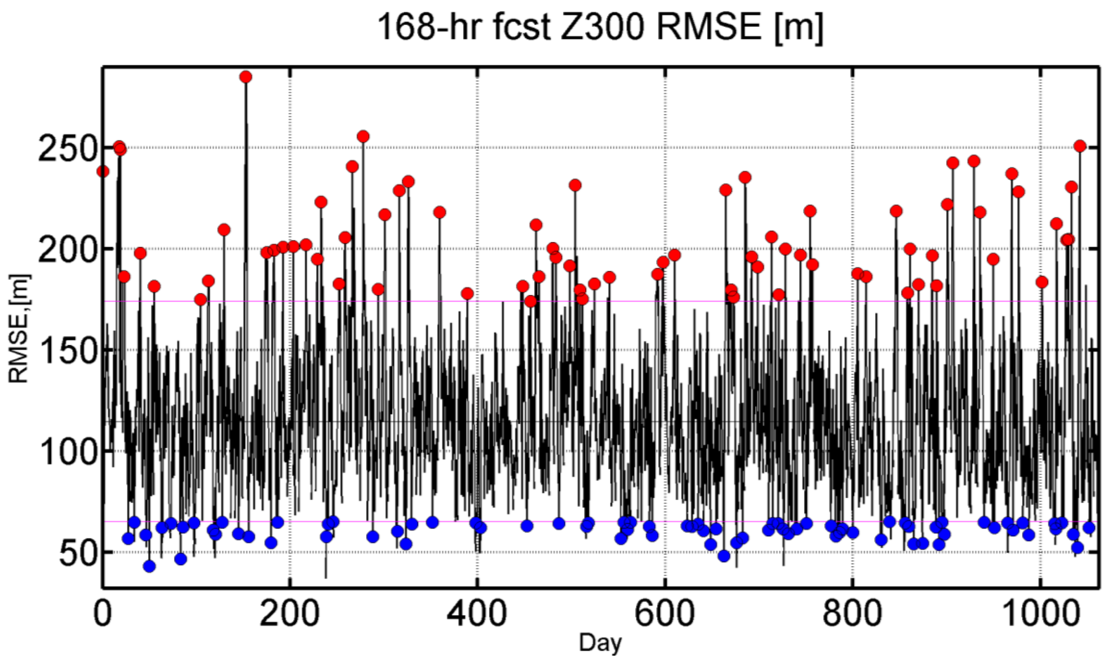


Figure 5.1: The time series of168-h forecast RMSE of Z300 over VR for 5 cool seasons from 2007/08 to 2012/13. Unit: [m]. Red dots and blue dots represent 83 large error cases and 85 small error cases, respectively.

<u>5.1.1 RWP amplitude anomaly for Dropout cases</u>

RWP amplitude (RWPA) anomaly here is the average anomaly relative to the climatology of RWP amplitude during 2007/2008 and 2012/2013 cool seasons (Figure 5.2). The dropout cases represent the cases with the largest forecast errors. By compositing the corresponding RWPA anomaly, the corresponding RWP features can be examined in the dropout cases.

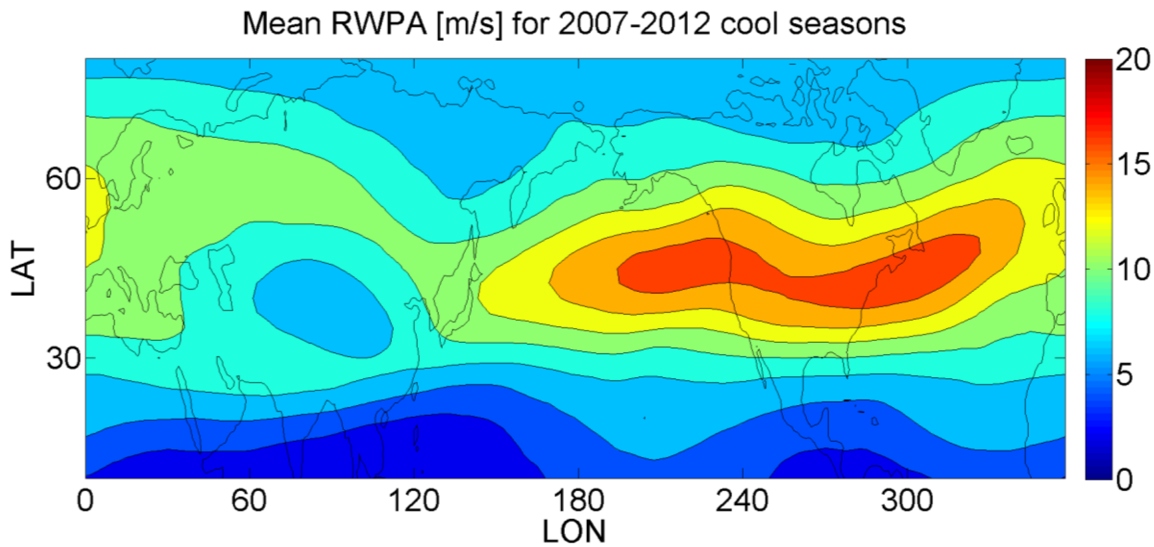


Figure 5.2: Climatology of RWPA over the 2007-2012 cool seasons. Unit: [m/s].

Figure 5.3 shows the time evolution of the absolute error anomaly relative to its climatology based on the day-7 large error cases. At a lead time of 1 day, the forecast errors are very small and distributed randomly within the plotting area. Starting from day 2, significant positive errors begin to develop over the Gulf of Alaska, northwest of the U.S., and southwest of Canada. There is also a small but significant positive error over northern Japan. On day 3, the error maxima are over the Gulf of Alaska and northwestern U.S., respectively. There are also positive errors east of Kamchatka but they are not significant. The maximum error grows from ~20 m for day 3 to >30 m for day 4. There are three error maxima on day 4: one east to the Kamchatka, the second across northwestern coast of the U.S., and the third one over western U.S. The eastern two pieces of errors grow rapidly from day 4 to day 5 and merge into one center west of the verification regions. The aforementioned first error maximum east of the Kamchatka has moved to Gulf of Alaska on day 5. The forecast errors grow from ~60 m on day 5 to 100 m on day 6. The error maximum is centered slightly southwest of the center of the verification region on day 6. Significant large errors cover most of the continental U.S. and its adjacent sea east of the coast. On day 7 (verification time), the error has increased to >140m and is centered over northeastern U.S.

Figure 5.4 shows the corresponding RWPA anomaly. The 95% significance contour for positive error in Figure 5.3 is also plotted to show where the significant errors are. On day 1, there is a significant RWPA positive anomaly over eastern Asia. Another positive anomaly is over northwestern U.S. On day 2, the RWPA anomaly over eastern Asia on day 1 has moved to western Pacific. Another positive anomaly is now over the northern plains of the U.S. On day 3, the first RWPA anomaly has moved to Central Pacific with the center east of the date line. Large errors tend to develop around its center and northern parts. There is another large error with the center over northwestern U.S, seemingly corresponding to the leading edge of the RWPA anomaly. The second RWPA anomaly has decreased radically. On day 4, the RWPA anomaly is centered over northeastern Pacific, with the leading edge crossing the northwest coast of the U.S. Two large error maxima are over the leading edge of the positive RWPA anomaly and west to the positive anomaly center, respectively. On day 5, the RWPA anomaly has split into two centers, with the eastern one centered over the southern plains of the U.S., and the western one over the west coast of North America. The largest errors tend to develop over its downstream over the Great Lakes. During day 6 and day 7, the RWPA anomaly moves southeastward, with the center over the Great Lakes on day 7. The largest error is centered over its leading edge.

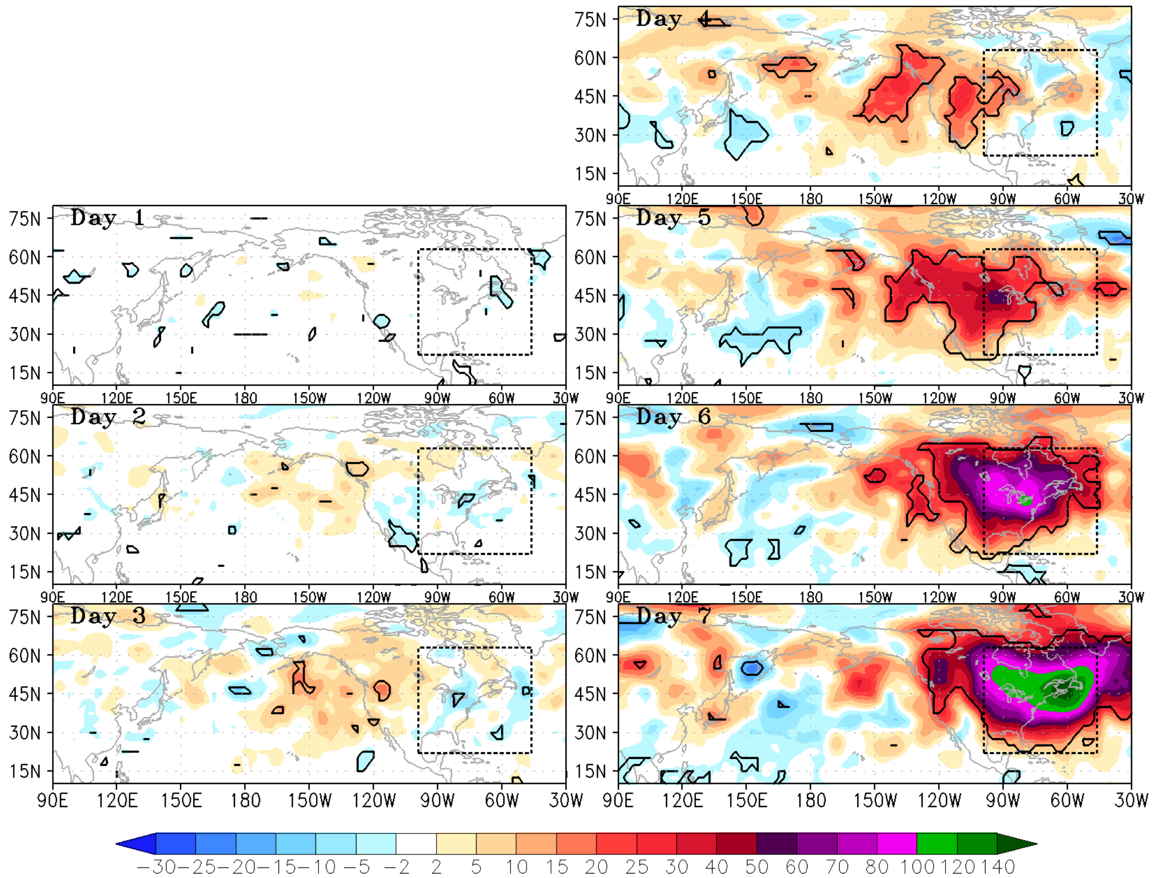


Figure 5.3: Composited forecast absolute error anomaly of Z300 conditional on 77 drop-out cases. Unit: [m]. Black line: significant values at 95% confidence level using t-test. Day 1-7: 1-7 day forecast times. Day 7 is also the forecast verification time.

To sum up, the forecast errors tend to develop and propagate from the central and eastern Pacific, suggesting the errors are associated with the propagation of large scale flow. For the forecast dropout cases, there was coherent RWPA positive anomaly starting from eastern Asia, crossing central and eastern Pacific, and entering the North American continent. Large forecast errors tend to develop with the RWPA anomaly when the RWPA anomaly is small. However, when the RWPA anomaly becomes large in amplitude and area, the largest errors tend to develop over its leading edge. The trailing region of RWPA is the second favorite region for error growth when a mature RWPA starts to cross the west coast of North America. America.

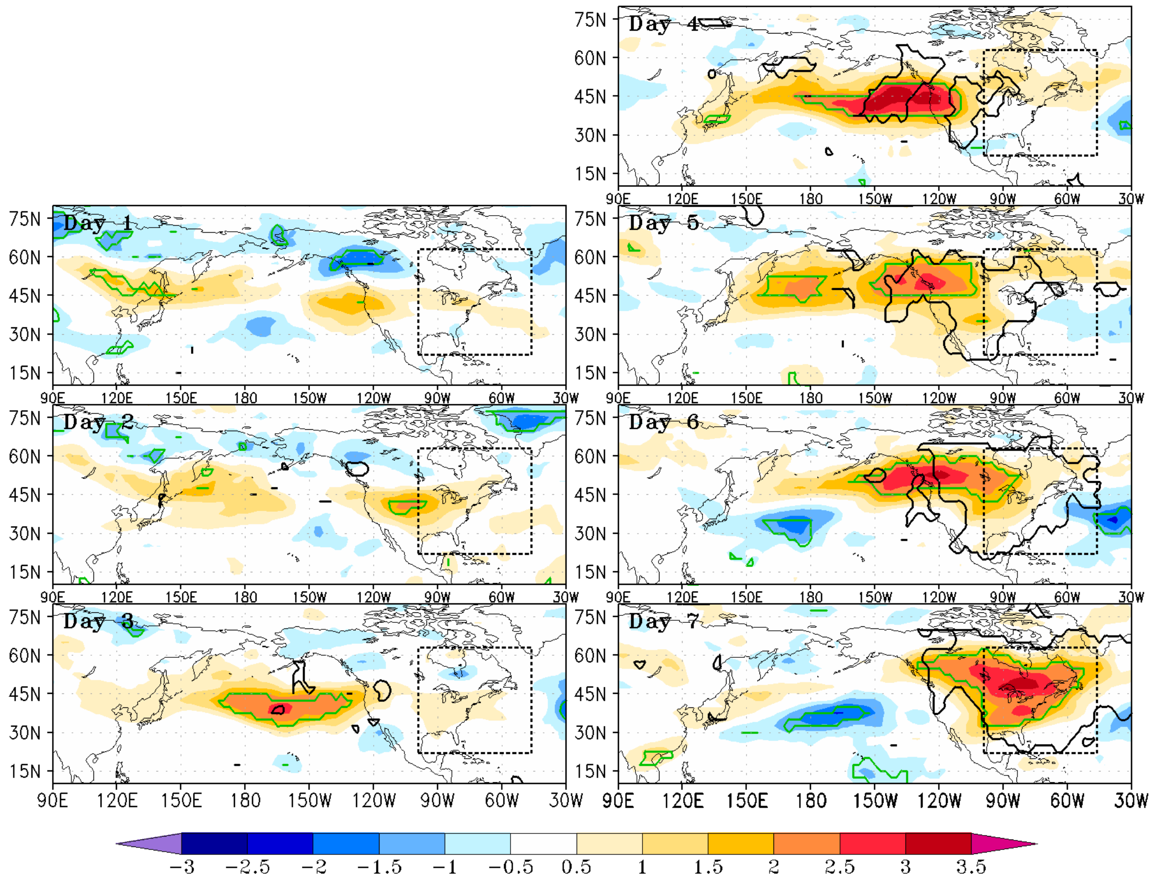


Figure 5.4: Composited RWPA anomaly (relative to climatology) conditional on 77 drop-out cases. Unit: [m/s]. Green line: significant values at 95% confidence level using t-test. Black contour: significant large absolute error region in Figure 5.3.

<u>5.1.2 RWPA anomaly for good-forecast cases</u>

From Figures 5.3-4, we have seen that composite large error cases over eastern U.S. and western Atlantic are associated with positive RWPA anomalies from upstream. By comparison, the RWPA anomalies in good-forecast cases have also been investigated in this subsection.

Figure 5.5 shows the time evolution of the absolute error anomaly based on the day 7 small error cases. During days 1−3, the forecast error anomalies are very small and mainly smaller than the climatological errors. Starting from day 4, an isolated piece of small error anomaly (negative values) is just crossing the southwestern coast of the Canada. On day 5, the significantly smaller errors are mainly over the western/southwestern coast of the U.S. The significantly smaller errors develop over the whole North America on day 6 and shifts towards the western Atlantic on day 7. The error anomaly shows an over 70 m reduction over the northeastern U.S.

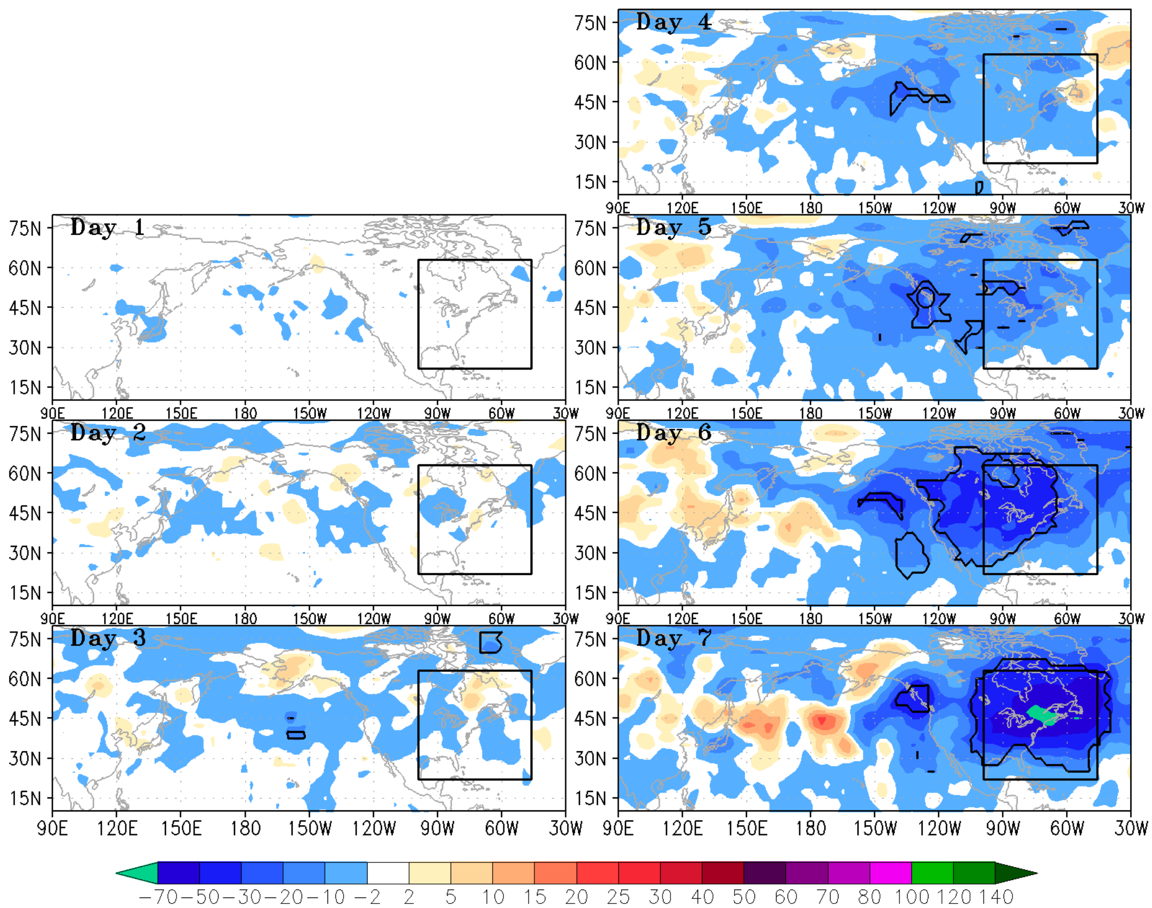


Figure 5.5: The same as Figure 5.3 but for the good-forecast cases. Note the black contours represent the significantly smaller errors relative to the climatological errors.

The corresponding RWPA anomaly is shown in Figure 5.6. The 95% significant contour for negative error anomalies in Figure 5.5 is plotted to show where the significantly small errors are. On day 1 and day 2, there are very weak positive RWPA anomalies over the western Pacific. The positive RWPA anomalies dissipate on day 3. A negative RWPA anomaly is over the west coast of the Canada, corresponding to the small error anomaly there on day 4. On day 5, the above negative RWPA anomaly becomes elongated and spans the middle latitudes (~35°N-55°N) of North America. The center of the RWPA anomaly moves to the east coast of the U.S. and western Atlantic on days 6–7. Meanwhile, there exists weakly negative RWPA anomaly over most regions of eastern North America.

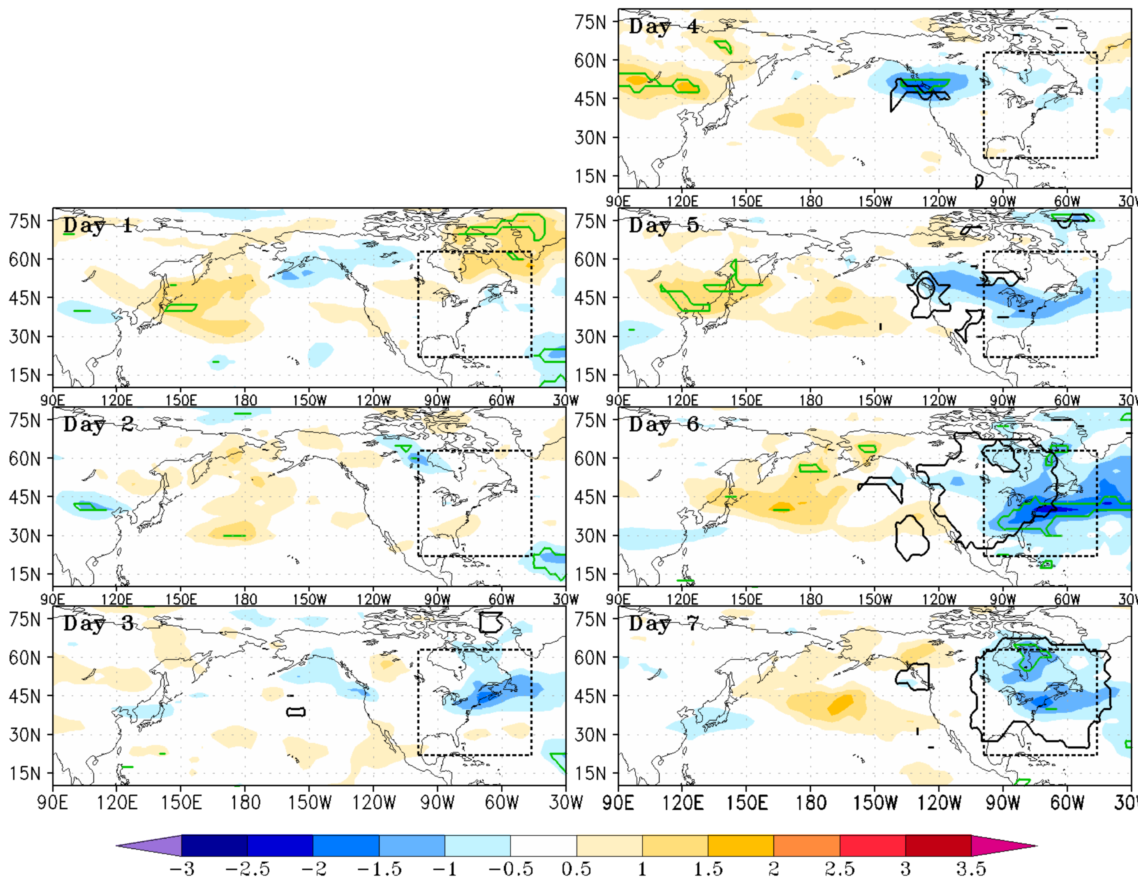


Figure 5.6: The same as Figure 5.4 but for the good-forecast cases. Note that black contours show the significantly smaller errors relative to the climatological errors.

Therefore, a weakly negative RWPA anomaly develops from the west coast of Canada and is seemingly associated with the decreasing absolute errors over the eastern North America. When compared with the results for the large-error cases, the negative RWPA anomalies are weaker and more localized, indicating that the weakening of the RWP is linked to the small errors over the eastern North America and western Atlantic.

5.2 Large spread cases and RWP

Ensemble spread is used in some studies (Grimit and Mass 2007; Du and Zhou 2011; McMurdie and Ancell 2014) to represent the predictability, with large ensemble spread indicating poor predictability. Therefore, we used 70 large ensemble spread cases based on the 7-day forecast of Z300 from 2007/2008 to 2012/2013 cool seasons over the same verification region in section 5.1. The large-spread cases are the cases with area-averaged spread >1.5 standard deviation.

Figure 5.7 shows the evolution of the composite spread anomaly from day 1 to day 7. Note the dashed box in each panel is the verification region. On day 1, the average spread is quite small everywhere. On day 2, large positive spread anomalies are mainly over the western North Pacific. On day 3, significantly larger spread anomalies are over the North Pacific. One anomaly is over the north of the Bering Sea, and another small positive anomaly is centered on (155°W, 35°N). There is also a third large spread anomaly spreading from northeastern Pacific to the northwest of the Great Lakes. On day 4, the largest spread anomaly has moved to the northeastern Pacific and the west of North America with two centers over the southwestern coastal area of Canada and northern Great Plains. From day 5, both positive spread centers grow and expand eastward. The west coast one is still centered over western Canada, whereas the downstream one has moved to the Midwest of the U.S. Note that the downstream anomaly developed more than the upstream one north of the Rocky Mountains. On day 6, the spread anomaly over the west coast has moved eastward and started to merge with the downstream one over the Great Lakes. Meanwhile, the downstream one has significantly increased from 25 m to 35 m. On day 7, the largest spread anomaly is within the verification region, with the center spreading from south of the Great Lakes to the Northeast of the U.S. Note that there is another center just over the Alaska Peninsula.

The composited RWPA anomaly together with the largest normalized spread over the whole region is shown in Figure 5.8. A positive RWP anomaly is over northwestern North Pacific on day 1. Another positive RWP anomaly is over its southeast spreading from East Pacific to the west coast of the U.S. The largest spread develops between them over the North Pacific. On day 2, the positive RWP anomaly to the west has moved to the south of the Bering Sea and increased to 4 m s^{-1}, whereas the eastern one remains over eastern Pacific between 35°N and 40°N. Significant larger positive spread anomalies are around the centers of the two anomalies. The above two RWP positive anomalies merge and strengthen on day 3 with the stronger center at (40°N, 140°W). Note that the largest spread is around the northern RWP center over the Gulf of Alaska. On day 4, the merged positive RWP anomaly has crossed the west coast of North America with the center over the northern part of the Rockies. The largest spread is over the west coast of Canada. On day 5, the positive RWPA center has weakened and shifted to the west of the Great Lakes. The largest spread is still over the southwest coast of Canada; meanwhile, another positive spread anomaly is over the west of the Great Lakes. Hence the large spread anomalies develop over the trailing region as well as the central region of the positive RWPA. On day 6, the RWPA anomaly continues weakening and moves to the East Coast of the U.S. The largest positive errors follow the RWPA anomaly center. Note that from day 5 to day 6, the largest spread center has jumped from the west coast of Canada to the Great Lakes, indicating that downstream development may have also occurred in the development of the forecast uncertainties over eastern U.S. On day 7, the RWPA anomaly intensifies slightly over the Northeast of the U.S. The large positive spread anomaly is with the RWPA positive anomaly. Meanwhile, a weak positive spread anomaly is over the Gulf of Alaska, which is the trailing region of the RWPA anomaly.

From Figures 5.7-5.8 we can see that the large-spread cases are also associated with enhanced RWPA propagating eastward from the Pacific. The centers of the large spread anomaly seem to shift from upstream over northeastern Pacific to the downstream over eastern and northeastern U.S. from day 4 to day 6 when the enhanced RWPA reaches the west coast of North America and enters the central North America. These results suggest that the leading edge of an RWP is a preferred region for the growth of both errors and uncertainties.

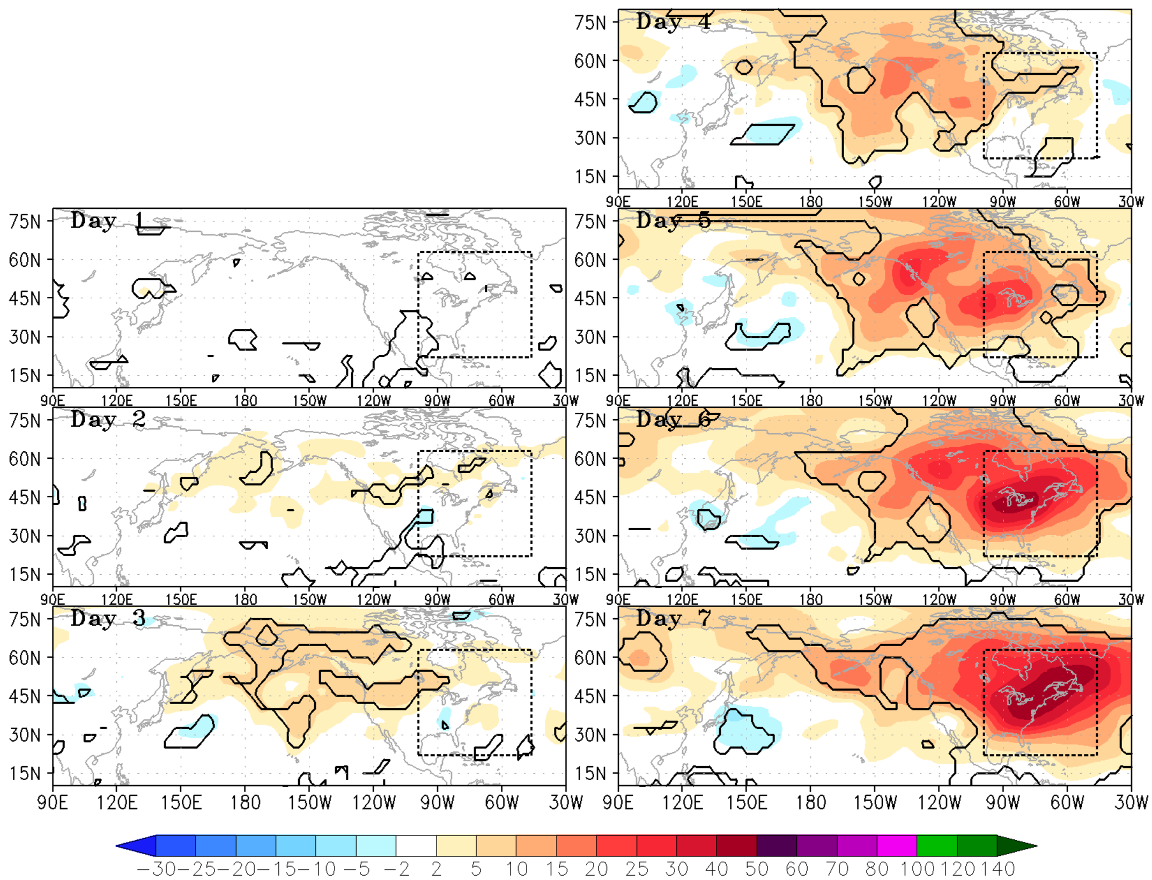


Figure 5.7: Composited ensemble forecast spread anomaly of Z300 conditional on 70 large spread cases. Unit: [m]. Black line: significant values at 95% confidence level using *t*-test.

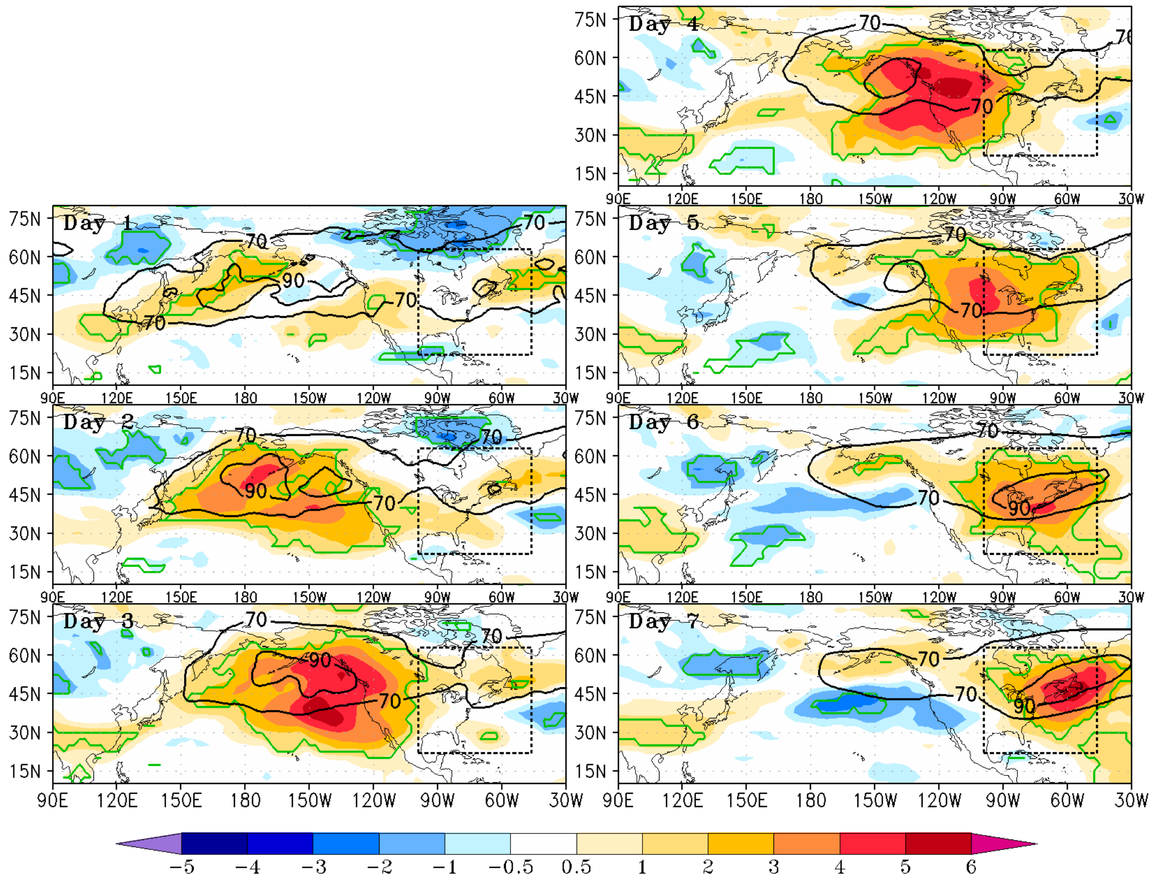


Figure 5.8: Composited RWPA anomaly (shaded, [m/s]) and the normalized ensemble spread (black contour, [%]) in 70 large spread cases.

5.3 Errors and uncertainties in coherent RWPs

5.3.1 A conceptual model of errors/uncertainty development associated with RWPs

Previous studies of both individual cases and a large number of cases have suggested that RWPs can serve as energy sources of downstream disturbances (Chang 1993; Orlanski and Sheldon 1995; Chang 1999; Chang and Yu 1999; Chang 2000; and Hakim 2003). Hakim (2005) further demonstrated that midlatitude forecast errors can develop and propagate similarly to RWPs. Chang et al. (2013) proposed that forecast uncertainties resemble the behavior of the downstream development of RWPs. Meanwhile, the impacts of targeted observations also propagate downstream in the form of an RWP (Szunyogh et al. 2000; Szunyogh et al. 2002; Majumdar et al. 2010).

From the studies of both large-error and large-spread cases for the eastern North America and western Atlantic discussed in the previous sections, we have found that both large errors and spread are associated with enhanced RWPs from upstream as far as eastern Asia or western Pacific. Building on the previous studies as well as the interpretation of operational daily ensemble RWPs, we propose a conceptual model for error and forecast uncertainty development as illustrated in Figure 5.9.

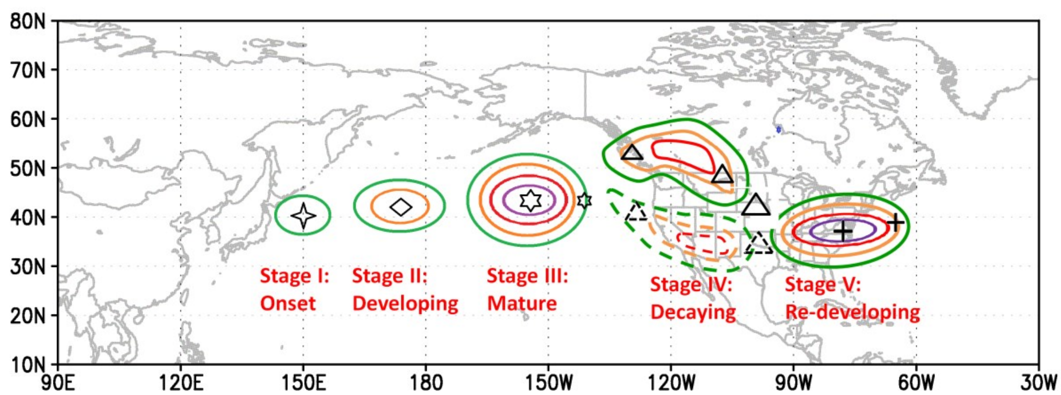


Figure 5.9: A conceptual model of the errors/uncertainties relative to the different stages of an RWP. Green, orange, red, and purple solid contours represents the RWP amplitude with the green being the least value and purple the largest value. The dashed contours represent an alternate process between two adjacent stages of RWP. The area enclosed by the green contour represents one RWP object. The black four-pointed star, diamond, six pointed star, triangle (dashed triangle), and plus sign, denote the possible regions for large errors/uncertainties associated with each RWP object◇Note that the positions of RWP objects just denote an approximation of the preferred RWP location, instead of the exact locations of an RWP.

Figure 5.9 shows an RWP forming over western Pacific (or central/eastern Pacific), and propagating into continental North America. We define the main stages of the RWP as follows: the onset of the RWP (stage I), the developing stage (stage II), the mature stage (stage III), the decaying stage (stage IV) and the re-developing stage (stage V). An RPW can form in different ways (Souders et al. 2014b): it may begin independently from a synoptic eddy; it may form when upstream energy "seeds" the entrance region of a storm track with a strong waveguide in place, or it may split off from a parent RWP either at its leading or trailing edges. The most likely genesis locations for RWPs in the NH are found to be at 140°E−170°W and 80°W−60°W (Souders et al. 2014b). During the onset stage, the likely large errors or forecast uncertainties tend to follow the RWP center when there is not a pre-existing upstream RWP. The RWP often intensifies over the central and eastern North Pacific. During the developing stage, the largest forecast errors/uncertainties are often over the center as well as the leading edge of the RWP. The RWP that impacted the predictability of eastern U.S. often reaches a peak or mature stage (Stage III) around eastern North Pacific or the west coast of North America. During this stage the largest errors/uncertainties are mainly over its center. Meanwhile, there often exist significant large errors/uncertainties over its leading edge. The RWP starts to decay when it crosses the mountainous regions (the Rocky Mountains) over western North America. The flow often splits into two streams: the northern stream and the southern stream. Hence the RWP crosses the mountain by turning either to its north (solid contours over the west coast of North America Figure 5.9) or south (dashed contours, Figure 5.9). From the composite results in sections 5.1-2, the northern track of the RWP over the west coast of North America is more closely associated with both large error and uncertainty cases. The northern track is considered to be the primary track to impact the predictability of eastern North America and western Atlantic. Henceforth, the southern track is considered to be the secondary track for RWPs to impact the forecasts of downstream regions. No matter which track the RWP takes, the forecast errors/uncertainties are often over its trailing region and leading edges. Both Souders et al. (2014b) and Glatt and Wirth (2014) suggested that western-central U.S. (west of 80°W) is a preferred region for RWP decay. The decaying stage (stage IV) is thus defined when the RWP crosses the Rocky Mountains and reaches its minimum intensity over western North America. During the decaying stage, the largest forecaster errors/uncertainties can develop further downstream of the leading edge of RWP. East Coast of the U.S. and the surrounding ocean is another preferred region (80°W-60°W) for RWP formation (Souders et al. 2014b). Therefore, a re-developing stage (stage V) is defined when the RWP starts to strengthen again over the east coast of the U.S. and western Atlantic. During this stage, the largest errors/uncertainties are over the central and leading regions of the RWP. Note that RWP during the re-developing stage can be actually a new RWP seeded by the decaying RWP. However, no matter whether a new RWP is triggered over the East Coast and the western Atlantic region, the largest errors and uncertainties often increases significantly around the previous decaying RWP center and the leading regions.

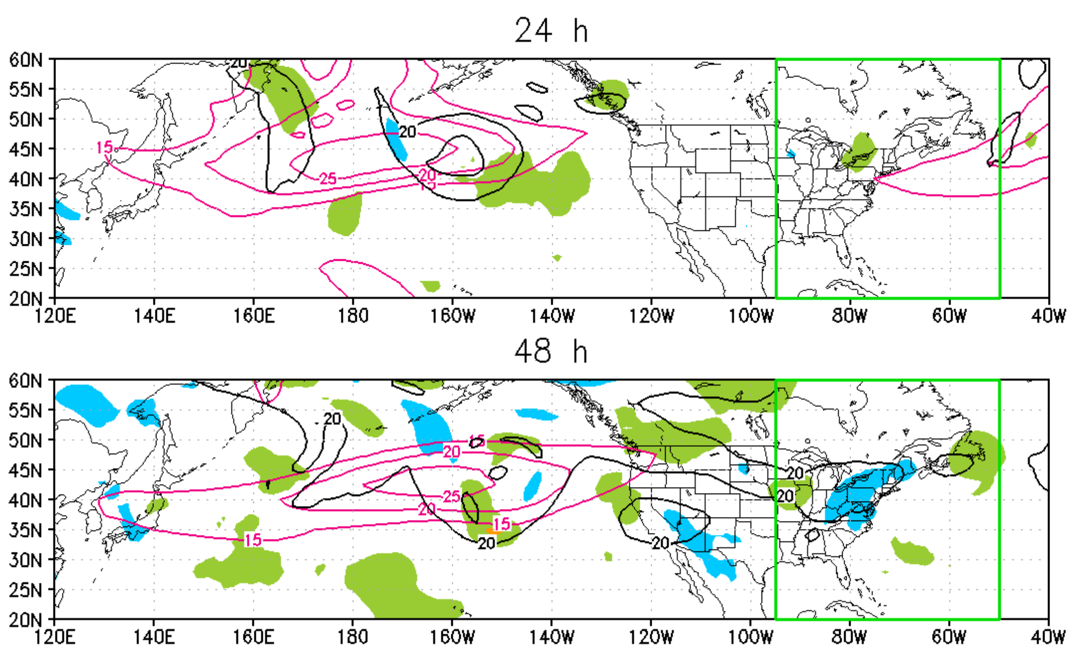
Previous studies (Lee and Held 1993; Orlanski and Chang 1993; Swanson and Pierrehumbert 1994) used numerical experiments to demonstrated that nonlinearity modifies the wave characteristics substantially, with the leading edge of the linear wave packet developing into a localized nonlinear wave packet, and the trailing end somehow dissipating, leaving just a localized downstream developing wave packet propagating with the speed of the leading edge as predicted by the linear theory. Chang (1993) showed that over the Pacific storm track regions, baroclinic waves exhibit the distinct characteristics of downstream development and occur in wave packets that propagate with group velocities much faster than the phase speeds of individual waves. RWP movement can be largely explained by the divergence of the fluxes of local eddy kinetic energy and ageostrophic geopotential (Chang 2001). Szunyogh et al. (2000; 2002) employed the ETKF method to investigate the impact of targeted observations and found that the ETKF signals propagate like upper-tropospheric wave packets generated by baroclinic energy conversion. Meanwhile, the largest error reduction was observed at the leading edge of

ETKF signals. Hence it is reasonable to hypothesize that the large amplitude of the growing waves over the leading edge regions could impact the growth and distribution of error/uncertainties especially downstream of RWPs. For example, the center of the largest errors/uncertainties can jump from the center of an RWP to the downstream regions when the center of eddy activity is shifting toward the downstream side during the mature and decaying stages (Chang 1993; Chang 2001). Szunyogh et al. (2002) found that the peaks of forecast improvement over the eastern U.S. followed the peaks of improvement over the western half with a 24-h delay, which is the time it takes for a wave packet to travel from the western to the eastern half of the country. Their results also showed the eastward "jump" of the signal maxima (See Figure 8 in Szunyogh et al. 2002), suggesting that the shift of the ETKF signals appear with the strong energy conversions. Extreme RWPs are often formed more commonly (more than 70%, Souders et al. 2014) in the North Pacific. The average propagating speed is around 25-30 m s^{-1} (Glatt and Wirth 2014) or a bit slower (15-25 m s^{-1}, Souders et al. 2014b), and the average duration is 5.8 days. Hence it is common for a coherent RWP originating over western or central Pacific to propagate across western North America and reach eastern North America within 4-7 days. The conceptual model shown in Figure 5.9 should be a useful model to illustrate how RWPs can modulate the development of forecast errors/uncertainties, thereby the medium-range predictability of East Coast high-impact weather event. One caveat for using this conceptual model is that the observed RWP objects in Figure 5.9 are not necessarily consistent with the exact locations and the relative amplitudes. Nevertheless, they represent a practical tool to understand the evolution of forecast errors/uncertainties associated with the coherent RWPs originating from upstream as far as western Pacific.

5.3.2 A case study of Error/forecast uncertainty growth with respect to a propagating RWP from western Pacific

To illustrate how the growth of errors and uncertainties in operational runs fit the conceptual model in Figure 5.9, a typical RWP case propagating from western Pacific is investigated in Figures 5.10-11

Figure 5.10 shows the time evolution of the ensemble mean RWPA as well as the forecast errors/uncertainties from day 1 to day 5 (verification time). On day 1 (24h), the RWP is over the central Pacific between 35°N and 50°N. There is large spread over its leading edge centered at (160°W, 43°N). The largest errors are also found over the leading edge between 155°W and 135°W. On day 2 (48 h), the RWP has moved eastward by ~15°longitudes. The largest spread is located east of the RWPA center. Meanwhile, the forecast errors have increased to over 60 m just southeast to RWPA center. Downstream of the RWP, there is a large spread center around the southwest coast of the U.S.



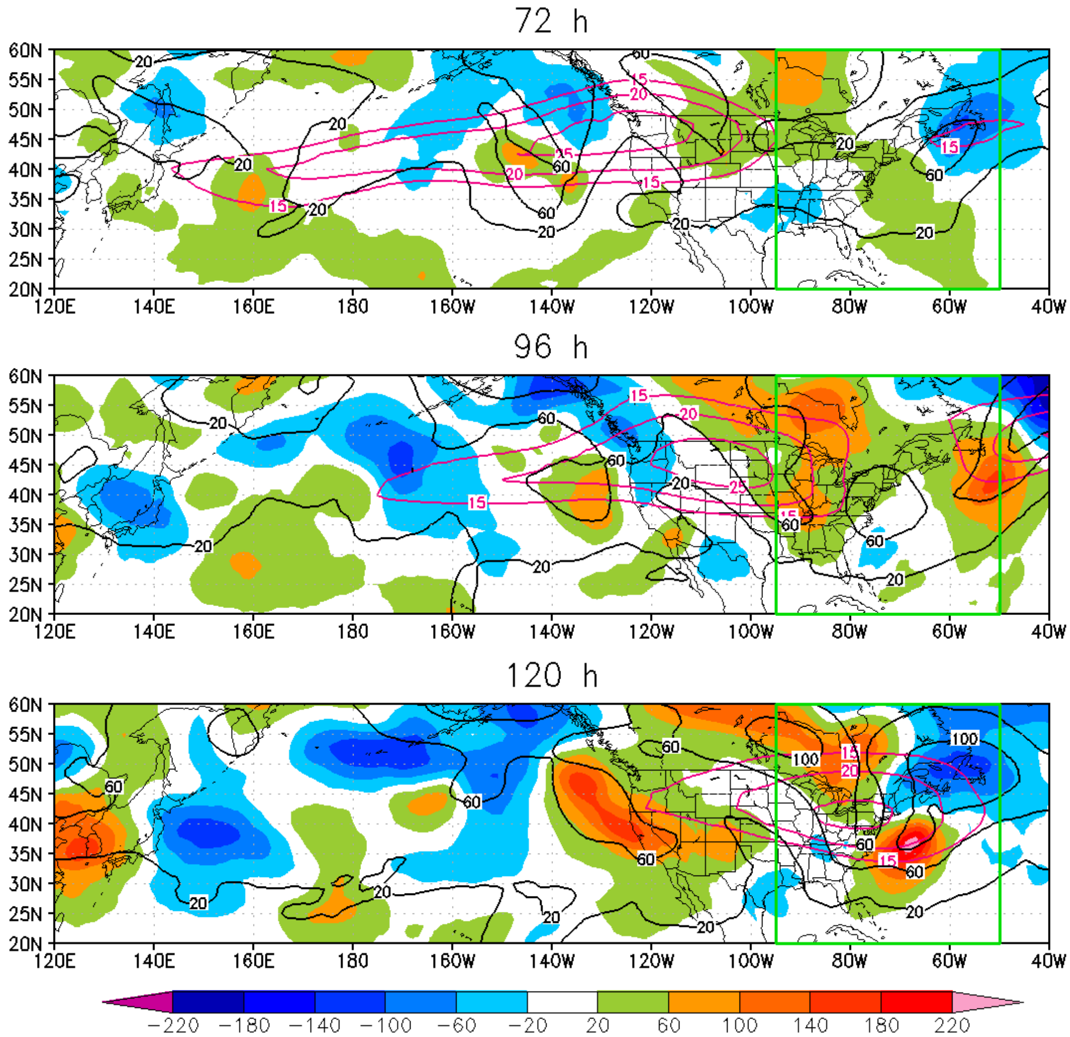


Figure 5.10: The ensemble mean RWPA (magenta solid contours, [m/s]), ensemble spread of Z300 (black contours,[m]), and the ensemble mean Z300 errors (shaded, [m]) of a 20-member GEFS run from 24h to 120h forecasts initialized at 1200 UTC March 21, 2014. The green box is the verification box later used to compute ensemble sensitivity.

On day 3 (72 h), the RWP center has shifted to the west coast of North America, leaving a longitudinally elongated trailing region which spans almost the entire Pacific. Taking both intensity and area into account, the RWP reaches its maximum on day 3. The largest errors/spread grows rapidly just west to RWP center as well as its northeastern quadrant. In other words, the largest errors and forecast uncertainties start to develop over both the leading edge as well as the trailing area. On day 4 (96 h), the RWP center has crossed the west coast of North

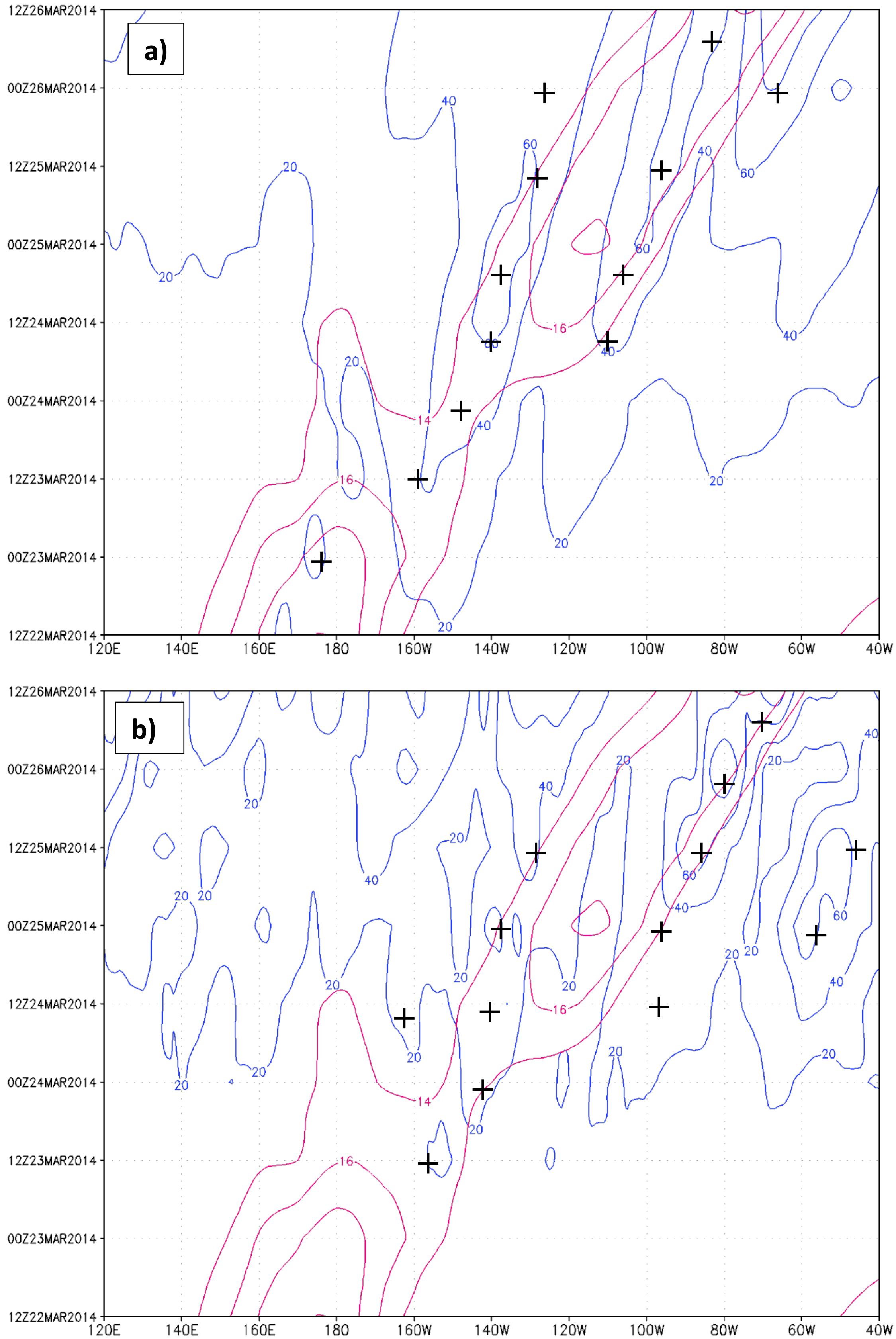

a)
12Z26MAR2014
00Z26MAR2014
12Z25MAR2014
00Z25MAR2014
12Z24MAR2014
00Z24MAR2014
12Z23MAR2014
00Z23MAR2014
12Z22MAR2014
120E
140E
160E
180
160W
140W
120W
100W
80W
60W
40W
b)

Figure 5.11: The Hovmoller plot of ensemble mean RWPA (magenta solid contours, [m/s]), and the ensemble spread of Z300 (a, blue contours,[m]) averaged between 30°N and 55°N, and the ensemble mean Z300 absolute errors (b, blue contours, [m]) of a 20-member GEFS run from 24 h to 120 h forecasts initialized at 1200 UTC March 21, 2014. Black plus signs represents the maxima of spread or errors at each forecast day.

America and weakened a bit. Large errors and spread are over its leading regions. A new large spread center forms over the East Coast of the U.S. On day 5 (120 h), the RWP is centered over the Mid-Atlantic. The largest errors and spread are spreading from its center towards the east and the northeast. The magnitude of the positive errors of Z300 over the western Atlantic has increased rapidly to >220 m from a negative error of ~40m on the previous day. The spread has increased rapidly from ~50 m to ~140 m over the east and northeast of U.S.

Figure 5.11 shows the corresponding Hovmoller diagram for both RWPA and the errors/spread. It is clear that the largest errors and spread tend to be mainly over the east of RWPA centers from day 1 (12Z 22 Mar 2014) to day 5 (12Z 26 Mar 2014). In other words, the errors and spread exhibit clear downstream development characteristics associated with the wave packet. However, there are also large errors/spread over the trailing regions of RWP between day 3 and day 4. This time is the transition time when the wave packet transitions from the mature stage to the decaying stage with the errors/uncertainties developing over both its trailing end and its leading edge, which qualitatively fits the northern track scenario displayed in Figure over both its trailing end and its leading edge, which qualitatively fits the northern track scenario displayed in Figure 5.9.

5.4 ESA of forecast metric to RWP

As discussed in Chapter 3, ESA can reveal the co-varying relation between a forecast metric and a state vector. The most sensitive areas are the critical regions associated with the

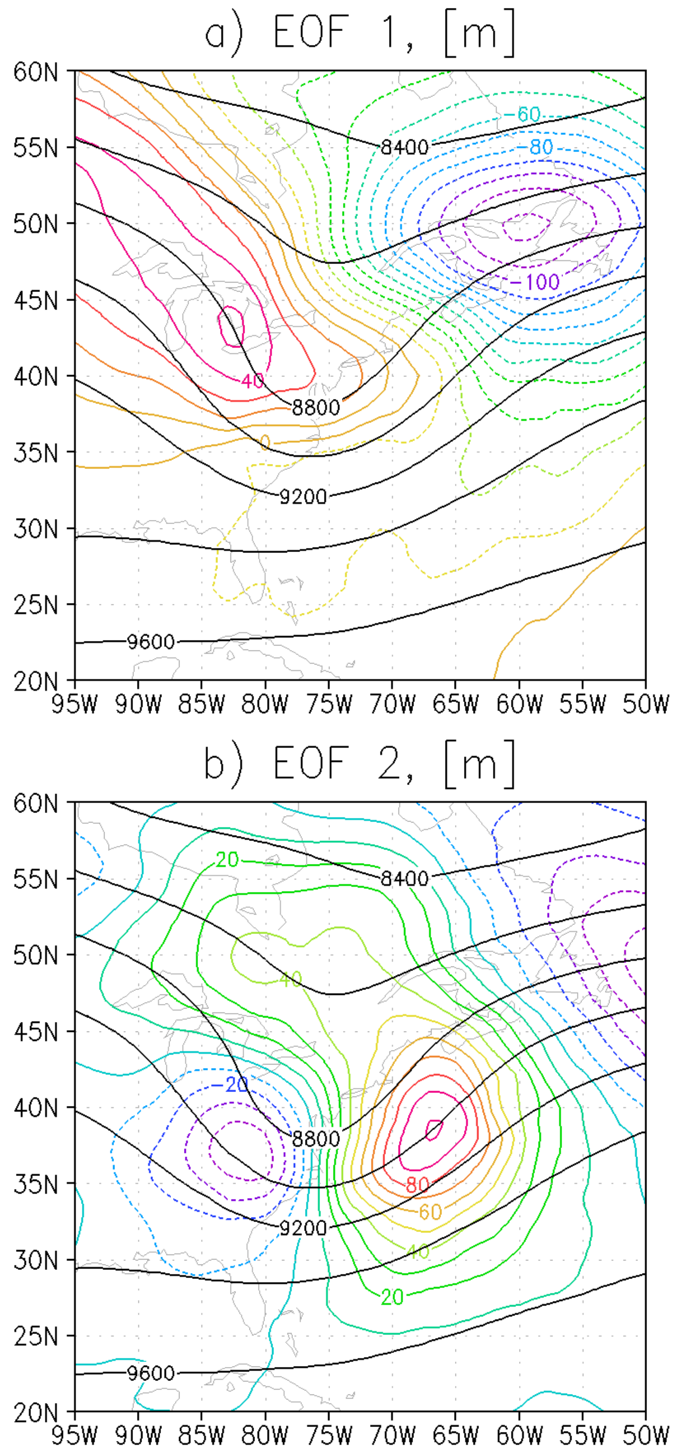


Figure 5.12: Ensemble mean of Z300 (black solid contour), with EOF1 (a, colored contour) and EOF2 (b, colored contour) of Z300 patterns based on the 20-member GEFS ensemble of Z300 initialized at 1200 UTC March 21, 2014 and valid at 1200 UTC March 26, 2014.

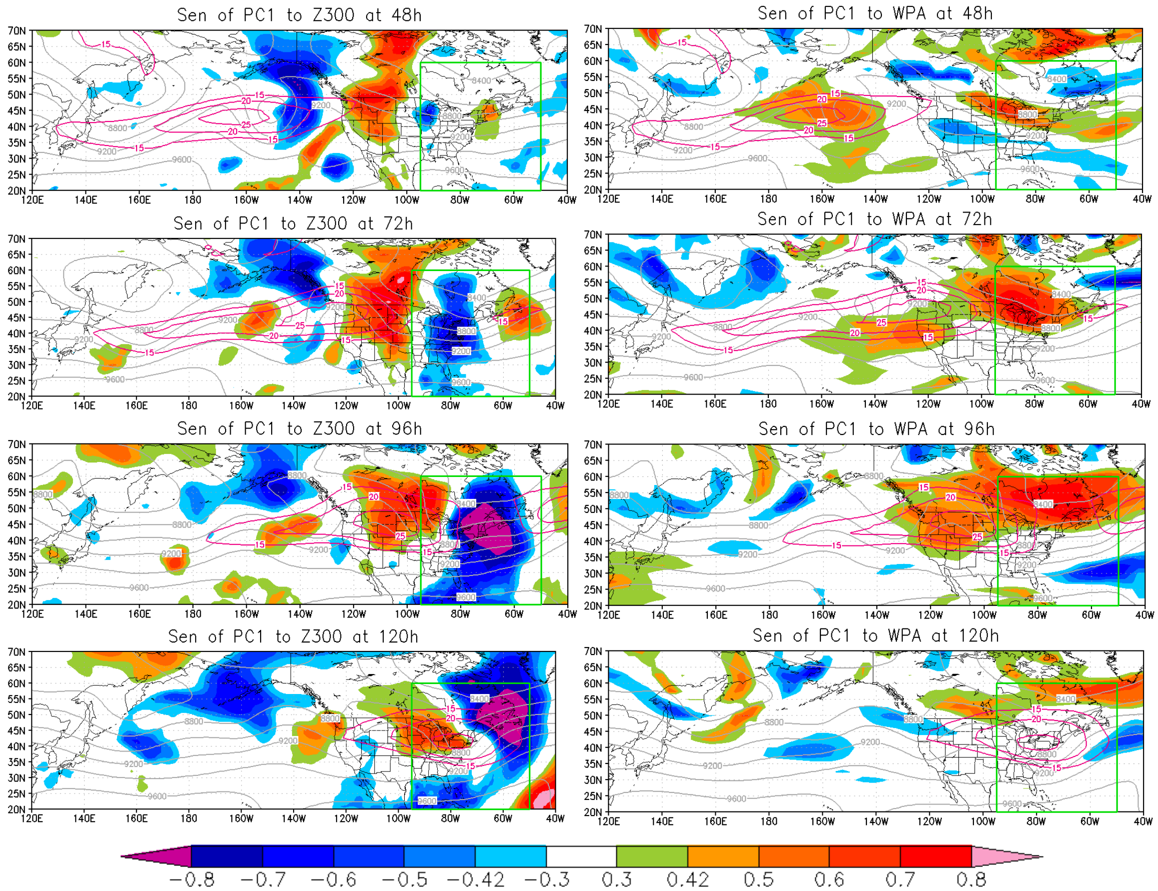


Figure 5.13: Sensitivity of EOF PC1 of Z300 to Z300 (left, shaded) and WPA (right, shaded) from 24h (day 1) to 120h (day 5) forecast time. The ensemble is the 20-member GEFS run initialized at 1200 UTC March 21, 2014. The EOF PC1 is calculated based on the 120 h or day 5 forecast over the green box. Magenta contours: the ensemble mean WPA; Gray contours: ensemble mean Z300.

variation in forecast metric corresponding to a verification time over a specified area. Sensitivity can be considered as an alternate way to represent forecast uncertainties in an ensemble run. In this section, we use ESA to investigate the sensitivity of a forecast metric to an RWP, which can be considered as an alternative way to examine the development of forecast uncertainty relative to RWP.

The same case discussed in subsection 5.3.2 (Figures 5.10-11) is used here. The leading two EOF PCs of Z300 are used as forecast metrics while the state vector is the RWPA anomaly in different members with respect to the ensemble mean. To compare the results, the Z300 is also employed as a state vector to calculate sensitivity.

The 5-day GEFS ensemble run is chosen and the EOF analysis is performed to find the dominant pattern of the Z300 forecast. The ensemble mean of Z300 shows a trough over the East Coast (Figure 5.12). Positive EOF 1 pattern shows a positive anomaly southwest of the trough and a negative anomaly over its north and northeast (Figure 5.12a). Hence EOF1 pattern represents a north-northeastward shifting and strengthening of the trough. EOF2 pattern shows a negative anomaly just west of the trough and strong positive anomaly over the east and the north (Figure 5.12b), indicating a slightly westward shifted and overall weaker trough.

Figures 5.13 and 5.14 show the ensemble sensitivity lon-lat plot and the hovmoller diagram using both Z300 and WPA as state vectors. Both sensitivities using Z300 and WPA as state vectors shows that the Z300 forecast over the verification regions are closely associated with the propagating RWP from western Pacific. The sensitivities using Z300 as state vector are overall stronger, which is expected for the forecast metric is based on the Z300 ensemble. Sensitivity using Z300 also clearly showed downstream development features. ESA using WPA suggested that the trough intensification and northeastward shifting at verification time are most significantly associated with the RWP intensity on day 1 and day 2, and the northeastward expansion on day 3 and day 4. Both sensitivities suggested that overall the leading edges downstream of the RWP center tend to develop the largest ensemble sensitivity. Nevertheless, west of the central regions between day 2 and day 4 also exhibits significant ensemble sensitivity.

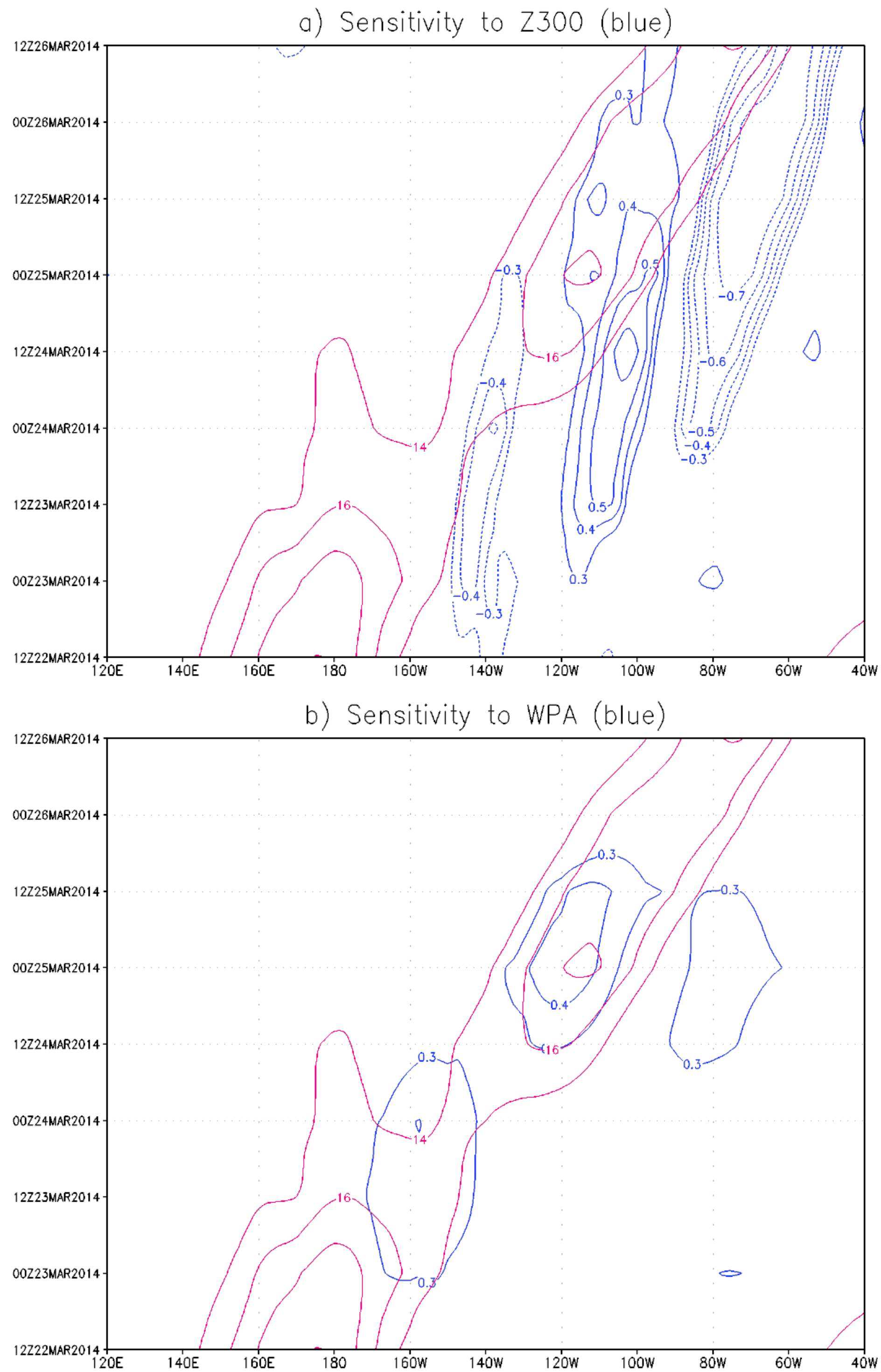


Figure 5.14: Hovmoller plot of sensitivity to Z300 (a, blue contours) and to WPA (b, blue contours) averaged between 30°N and 55°N. The magenta contours are for RWPA averaged over the same longitudinal bands.

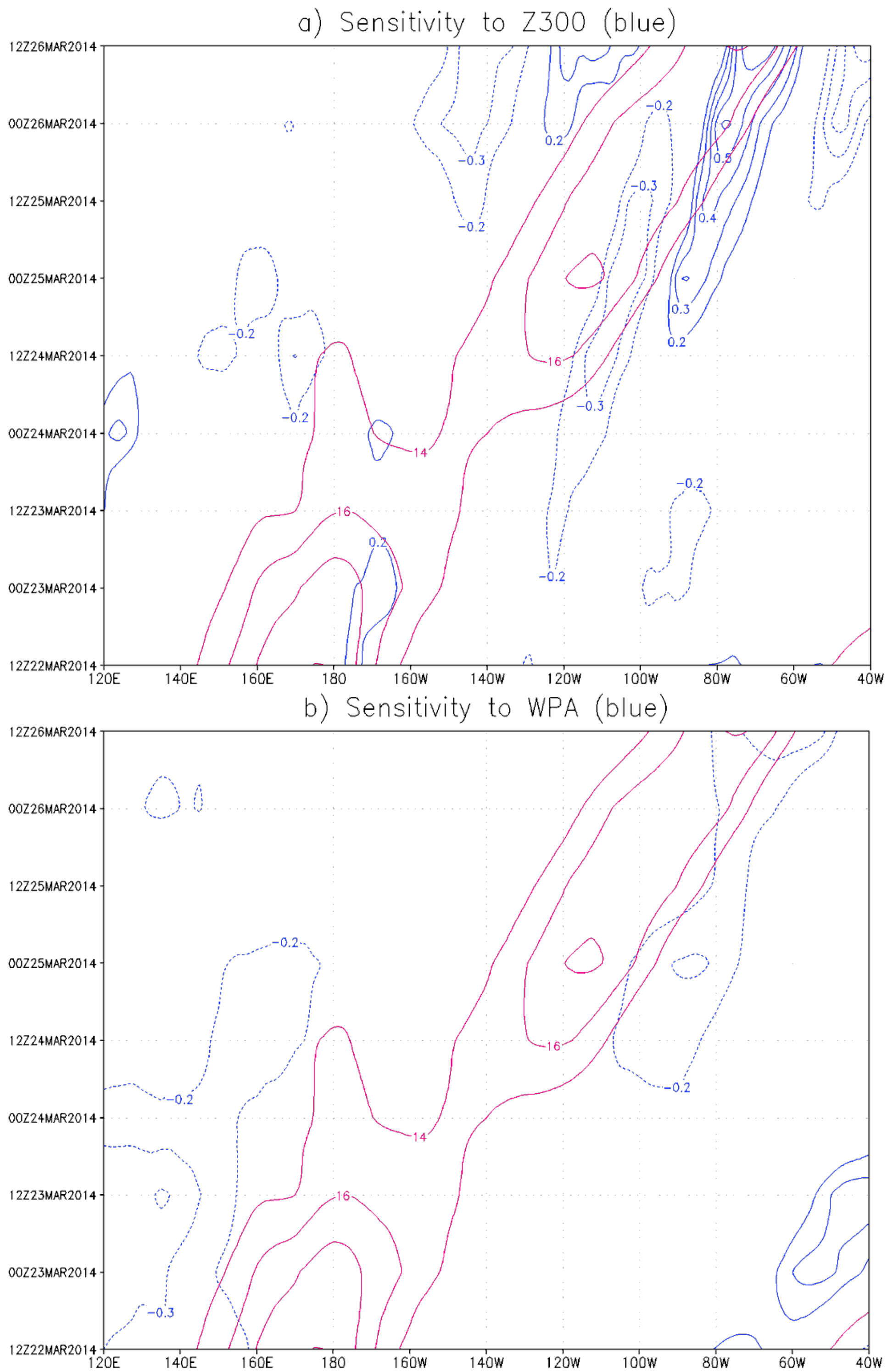


Figure 5.15: The same as Figure 5.14 but for sensitivity using EOF PC2.

Sensitivity of EOF PC2 to Z300 shows that the weakening and westward-shifting of the trough and its downstream ridge is associated with a weaker wave group both upstream and downstream at verification time (Figure 5.15a). The sensitivity to RWPA (Figure 5.15b) is consistent, suggesting the weakening and slightly southwestward shifting RWP is associated with the EOF2 pattern. When comparing with sensitivity for EOF1, both sensitivities for the EOF2 pattern are weaker in amplitude. However, the most robust signals which finally propagate into the verification region at verification time are over the leading edges.

The ESA for the case study confirms that the forecast uncertainties over the verification region are most closely linked with the central to leading part of the RWP. The sensitivity develops and propagates with the RWP leading edges, exhibiting the characteristics of downstream development. These results suggest that RWPs propagating across the North Pacific have a pervasive impact on the forecast skill over eastern U.S. and western Atlantic.

5.5 Downstream development of forecast errors

In the previous sections we have shown the importance of downstream development in the time evolution of forecast errors and uncertainties. The conceptual model in Figure 5.9 was proposed based on the composited results as well as the daily ensemble WPA interpretations. Orlanski and Katzfey (1991) derived the eddy kinetic energy (EKE) budget equations and showed that the local kinetic energy fluxes between adjacent baroclinic systems are critical to cyclone growth and downstream development. Chang (1993) employed the EKE equation to build a comprehensive framework of downstream development theory of RWPs. Since we hypothesize that the large forecast errors and uncertainties are associated with the downstream development of RWPs, the EKE budget will be used to diagnose whether the error increases are mainly due to the downstream process.

5.5.1 The eddy kinetic energy budget

In order to derive the EKE budget equation, the atmospheric flow is first separated into a time mean (usually monthly mean) part and a perturbation part. The horizontal velocity, vertical velocity and the geopotential can be divided as follows:

$$\mathbf{v} = \mathbf{v}_m + \mathbf{v}' \tag{5.1}$$

$$\phi = \phi_m + \phi' \tag{5.2}$$

$$\omega = \omega' \tag{5.3}$$

Where v is the horizontal and ω is the vertical velocity; ϕ represents the geopotential. The subscript $_m$ denotes the time-mean part and the prime describes the perturbations.

The eddy part of the kinetic energy (EKE) can be defined as:

$$K_e = (u'^2 + v'^2)/2 \, . \tag{5.4}$$

The equation governing the local tendency of the EKE can be derived from the horizontal momentum equations as well as the continuity equation and be written as follows (see Appendix 2 and Orlanski and Katzfey 1991):

$$\frac{\partial K_e}{\partial t} = -(\mathbf{v}'\cdot\nabla_p \phi') - \nabla_p \cdot (\mathbf{v}K_e) - \frac{\partial(\omega' K_e)}{\partial p} - (\mathbf{v}'\cdot(\mathbf{v}_3' \cdot \nabla_p \mathbf{v}_m)) + (\mathbf{v}'\cdot(\overline{\mathbf{v}_3' \cdot \nabla_p \mathbf{v}'})) - Residual$$

(5.5)

Note that the **v** in the above equation represents the total horizontal wind. The variables in the equations are defined in equations 5.1-4.

Orlanski and Katzfey (1991) further demonstrated that the first term on the rhs in equation 5.5 associated with the pressure work can be further partitioned into two, one for the geostrophic eddy flux and the other for the ageostrophic eddy flux through the following derivation:

$$
\begin{aligned}
-(\mathbf{v}'\cdot\nabla_p\phi') &= -(\nabla_p\cdot(\mathbf{v}'\phi') - \phi'\cdot\nabla_p\cdot\mathbf{v}') \\
&= -\nabla_p\cdot(\mathbf{v}_a'\phi') - \nabla_p\cdot(\mathbf{v}_g'\phi') + \phi'\cdot\nabla_p\cdot\mathbf{v}_a') + \phi'\cdot\nabla_p\cdot\mathbf{v}_g') \\
&= -\nabla_p\cdot(\mathbf{v}_a'\phi') - \mathbf{v}_g'\cdot\nabla_p\phi' - \phi'\frac{\partial\omega'}{\partial p} + 0 \\
&= -\nabla_p\cdot(\mathbf{v}_a'\phi') - \frac{1}{f}\vec{\mathbf{k}}\times\nabla\phi'\cdot\nabla_p\phi' - \frac{\partial(\omega'\phi')}{\partial p} + \omega'\frac{\partial\phi'}{\partial p} \\
&= -\nabla_p\cdot(\mathbf{v}_a'\phi') - \omega'\alpha' - \frac{\partial(\omega'\phi')}{\partial p}
\end{aligned}
$$

Hence the EKE equation 5.5 can be summarized as follows:

$$
\frac{\partial K_e}{\partial t} = -\nabla_p\cdot(\mathbf{v}K_e + \mathbf{v}_a'\phi') - \omega'\alpha' - [\mathbf{v}'\cdot(\mathbf{v}_3'\cdot\nabla_p\mathbf{v}_m) - \mathbf{v}'\cdot\overline{(\mathbf{v}_3'\cdot\nabla_p\mathbf{v}')}] - [\frac{\partial(\omega' K_e)}{\partial p} + \frac{\partial(\omega'\phi')}{\partial p}] - Residual
$$

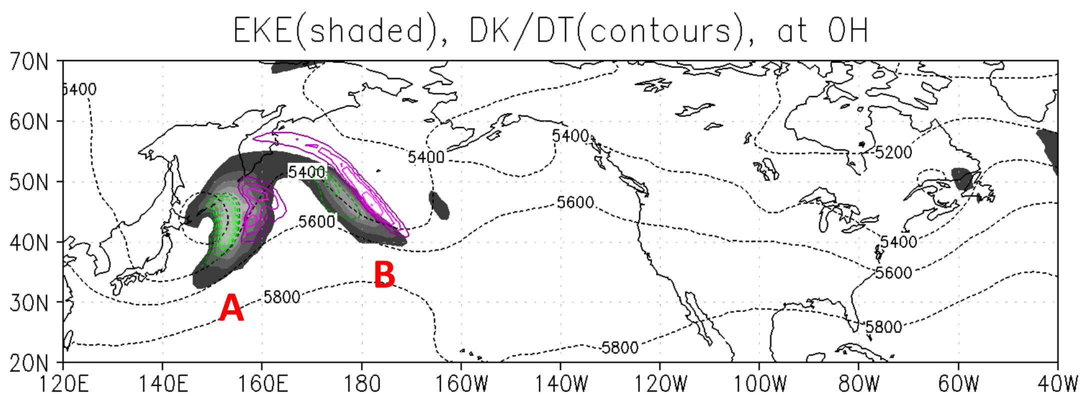
(5.6)

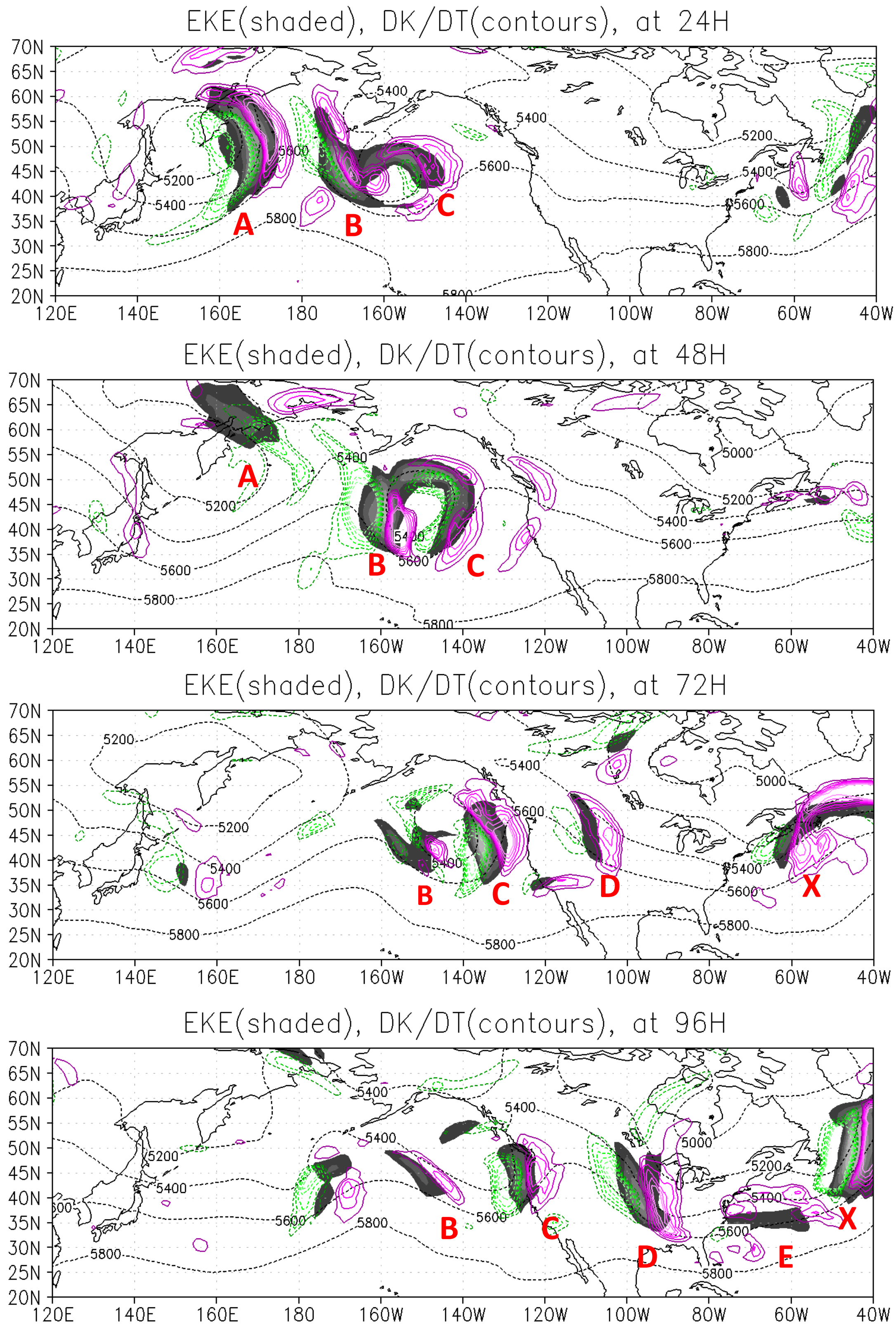
In equation 5.6, the first term on the rhs is the convergence of the total energy flux, which is the sum of advective flux (AdvF) and ageostrophic geopotential flux (AGF). The second term is the baroclinic conversion from eddy available potential energy. The third term is the barotropic conversion. The fourth term represents the convergence of the vertical advective flux of EKE and the vertical energy flux (the total energy flux entering the volume from the top and bottom if vertically integrated).

It has been shown in the literature that the ageostrophic geopotential fluxes (AGF) contribute significantly to the growth of downstream waves (e.g. Orlanski and Chang 1993). Chang and Orlanski (1994) showed that for nearly plane waves, the summation of the AGF and the AdvF normalized by the total eddy energy is exactly equal to the group velocity of RWPs, indicating the total energy flux in equation 5.6 is most closely associated with the downstream energy dispersion. Their results also showed that waves over the wintertime Pacific storm track region radiate a lot of energy toward the downstream direction, suggesting this energy can affect the weather and climate of the eastern Pacific and the continental North America. Since the integral of the total energy flux convergence can be used to diagnose downstream development between adjacent waves when integrating the local EKE tendency, we will use the vertical integral of the first term in the rhs of equation 5.6 as well as its individual components associated with the AGF and the AdvF to diagnose the development of errors and investigate how much it contributes to the EKE budget. Since the vertical velocity data (ω' in equations 5.6) is not available from the TIGGE archive and the focus of this study is on the terms associated with downstream development, the baroclinic term in equation 5.6 is not calculated for this work even though it may also be important.

5.5.2 The diagnostic of forecast errors under EKE framework

In order to examine how the errors develop and propagate under the EKE framework, the vertically integrated EKE as well as its local tendency (DK/DT) will be investigated firstly using the same 2014 March case discussed in sections 5.3-5.4.




EKE(shaded), DK/DT(contours), at 24H
70N 65N 60N 55N 50N 45N 40N 35N 30N 25N 20N
120E 140E 160E 180 160W 140W 120W 100W 80W 60W 40W
A
B
C
EKE(shaded), DK/DT(contours), at 48H
A
B
C
EKE(shaded), DK/DT(contours), at 72H
B
C
D
X
EKE(shaded), DK/DT(contours), at 96H
B
C
D
E
X

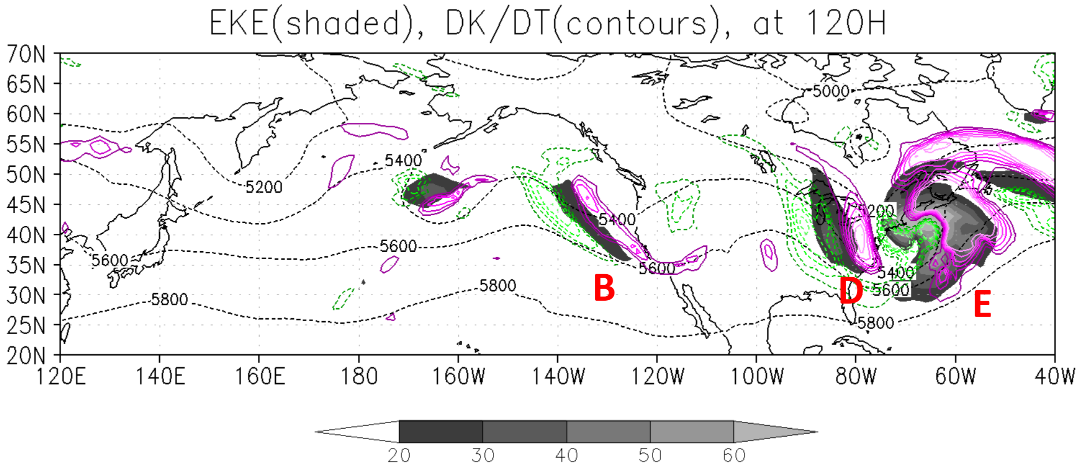


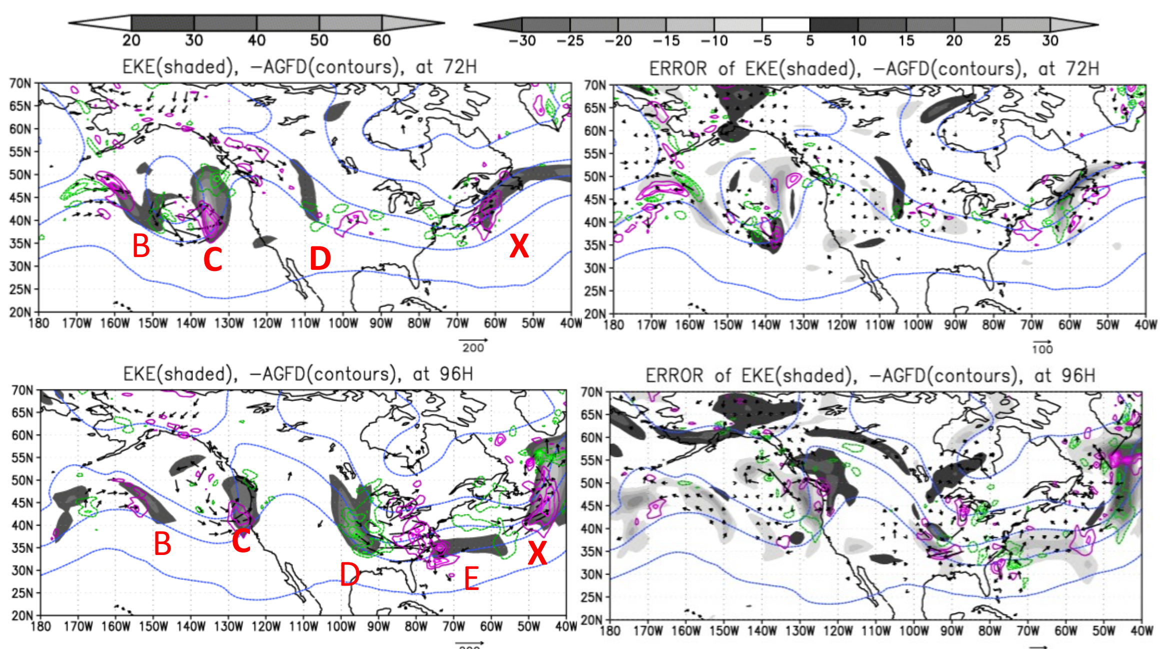
Figure 5.16: Vertically integrated EKE (gray shades 1x105 J m-2) and the local EKE tendency (colored contours, -70–-20, and 20–70 W m-2) from 0 h (day 0, 12Z March 21 2014) to 120 h (day 5, 12Z March 26 2014) in the NCEP analysis. Black dashed contours are Z500.

Figure 5.16 shows the time evolution of the EKE and its tendency (DK/DT). Two EKE centers (A and B) developed over the western and central Pacific, which were associated with the presence of a wave packet as discussed in sections 5.3-4. The EKE tendency suggests that the EKE often increases (positive DK/DT) over the leading regions of the EKE center and decreases (negative DK/DT) over the trailing areas, indicating predominantly eastward movement. By 24 h, the developing center B has grown in both amplitude and area. A new EKE center (C) formed downstream of center B. At 48 h, the upstream EKE center A weakened while the downstream center C had been growing. Center B maintained its intensity. Note that at this time there were two positive tendency centers just over the west coast of North America, indicating the EKE there will subsequently increase. Figure 5.13 has shown that at this time, the leading edge of the RWP reached the west coast of North America. The positive EKE tendency there is indicative of the dispersion of EKE energy from upstream. At 72 h, center C developed while the upstream center B weakened. A new center D was forming downstream of C associated with a trough lying northeast to the Rocky Mountains. Large positive EKE tendency was over the downstream side of center D. At 96 h, both centers B and C weakened while the center D developed over the

Northern Plains of U.S. Another new center E formed over the ocean east to the East Coast of the U.S. At 120 h, center C disappeared and B was very weak. Center D had moved to the Midwest of the U.S. and weakened a bit. The downstream center E, which was over the eastern side of an upper-level trough along the East Coast, had rapidly intensified. The positive tendency over its eastern and northeastern side indicates it will be further intensifying in the following time. Note that center E was also connected with a pre-existing western Atlantic RWP (center X).

Figure 5.16 shows that the EKE associated with the RWP started to increase from 48 h over the west coast of North America, impacting the continental U.S. during the following days, and reached the East Coast of the U.S. at 120 h. As can be seen from Figures 5.10 and 5.11, the forecast errors over the continental U.S. and the surrounding ocean started to grow by 72h. Previous studies suggested that the downstream development of a new EKE center is triggered by the convergence of the AGF (Chang 1993; Orlanski and Sheldon 1995). While the AdvF is often large, it acts mainly to redistribute energy from the rear to the front of the energy center and not between energy centers, as is the case for the AGF (Orlanski and Sheldon 1995). In terms of the EKE budget, Chang (1993) found that over weakly baroclinic regions, initially the convergence of the AGF dominates the growth of EKE with other terms playing only secondary roles. Therefore, we consider the AGF to be more representative of the initial downstream development, though AGF and the AdvF together are associated with the group velocity of RWPs (Chang and Orlanski 1994). Here we hypothesize that the errors in a new EKE center are triggered by the errors in AGF during the early stages while the AdvF becomes increasingly important during the later stages. Meanwhile, the errors in the baroclinic conversion term can be comparable with the AGFD term in the later stage over the regions with strong baroclinicity.

To examine the errors associated with downstream development, the forecast errors in EKE are calculated. Meanwhile, the errors in the convergence of the AdvF, the AGF, and the total advective flux are also computed separately to represent the contributions of each term to the errors in the local EKE tendency. Hence the errors associated with downstream development can be quantified. The control member from the NCEP operational model, and two ensemble members (member 4 and member 10) will be investigated. These two members are chosen because member 4 has larger MAE than other member while member 10 has smaller MAE over a verification box (80°W-50°W, 25°N-55°N) and better forecast the location and the intensity of the surface cyclone at 120 h than other members. Since the large errors over North America developed from 72 h, the errors from 72 h (day 3) to 120 h (day 5) will be shown.



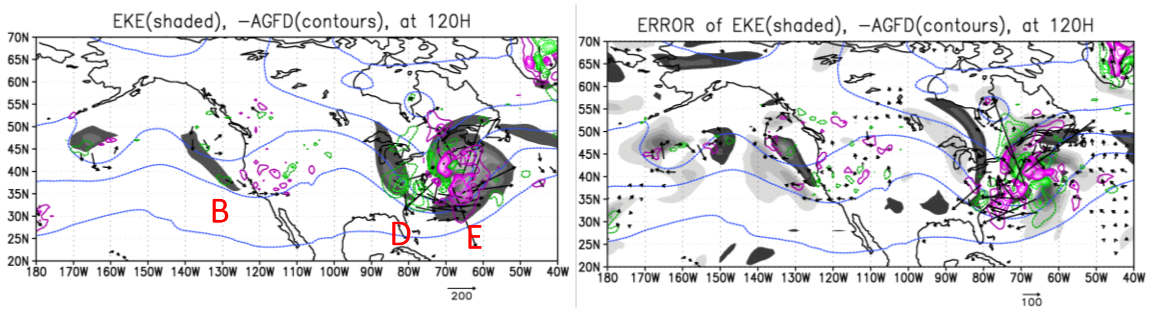


Figure 5.17: (left panels) The EKE (shaded) and the AGF (black arrows) and its convergence (colored contours, -AGFD); (right panels) the errors of EKE (shaded), AGF (black arrows) and its convergence (colored contours). The colored contours are starting from ±100 Wm^{-2}, with a 100 increase or decrease.

The vertically integrated EKE, the AGF, and the convergence of AGF are shown in the left panels of Figure 5.17 while the right panels show their errors in the control forecast. It is clear that strong AGF exhibits between two EKE centers, with the upstream center radiating energy to the downstream one. At 72 h, the EKE enter C radiated energy to its northeast across the west coast of North America, leading to a convergence of AGF around (125°W, 50°N). The newly formed center D received energy from its northwest through weak AGF. In the forecast, center C was a bit weaker and displaced southward. The northeastern side of C radiated more energy to its north instead of its east, leading to the divergence error around (125°W, 50°N). The new EKE center D was shifted north-northeastward over the northern Plains. At 96 h, center D was more north-northeastward in the forecast. In the analysis, there were very strong AGF transferring energy from D to the newly formed downstream center E over the southeast coast of the U.S., while the AGF was much weaker (opposite direction around 90°W, 35°N) in the forecast. The downstream center E was thus much weaker in the forecast. There also existed significant AGF transferring EKE from center D to eastern Great Lakes in the analysis while the AGF was also weaker in the forecast. At 120 h, the AGF errors between centers D and E were very large over the East Coast of the U.S., leading to the less energy transfer from D to E and excessive divergence (energy dissipation) over the southern part of E, which contributes to the negative EKE errors over western side of E and positive EKE errors over the east-southeastern side of D. Note that the AGF was anticlockwise within center E in the analysis while in the forecast the AGF had a clockwise errors, creating positive EKE errors around (65°W, 41°N). Overall, there was a qualitative correspondence between the AGF convergence (divergence) errors and the positive (negative) errors of two adjacent EKE centers, seemingly associated with the downstream development over regions with strong AGF.

Figure 5.18 shows the EKE, the AdvF, and the convergence of the AdvF. The relation between the EKE and advection term is straightforward as can be seen from the left panels in Figure 5.18. The AdvF follows the propagation of EKE, creating divergence over the rear of the energy center and convergence over the front of the center, which is consistent with the results from Orlanski and Shelton (1995). The errors of the advection term showed divergence error over the eastern (western) side of a negative (positive) EKE error and convergence error over the western (eastern) side of a negative (positive) EKE error. When compared with the AGFD from 96 h, the convergence/divergence errors caused by the advection term overall were smaller than that by the AGFD over eastern U.S. and western Atlantic, indicating the errors associated with the AGFD term are larger. Meanwhile, the errors of the advection term were mainly distributed within a same EKE center while the AGFD errors could be large during the regions between two adjacent energy centers, suggesting the AGFD can shift the error centers from upstream to downstream and thus initiate a downstream error center while the errors in advection term are more local. Nevertheless, together with the AGFD, they modulate the change of local EKE change and its errors.

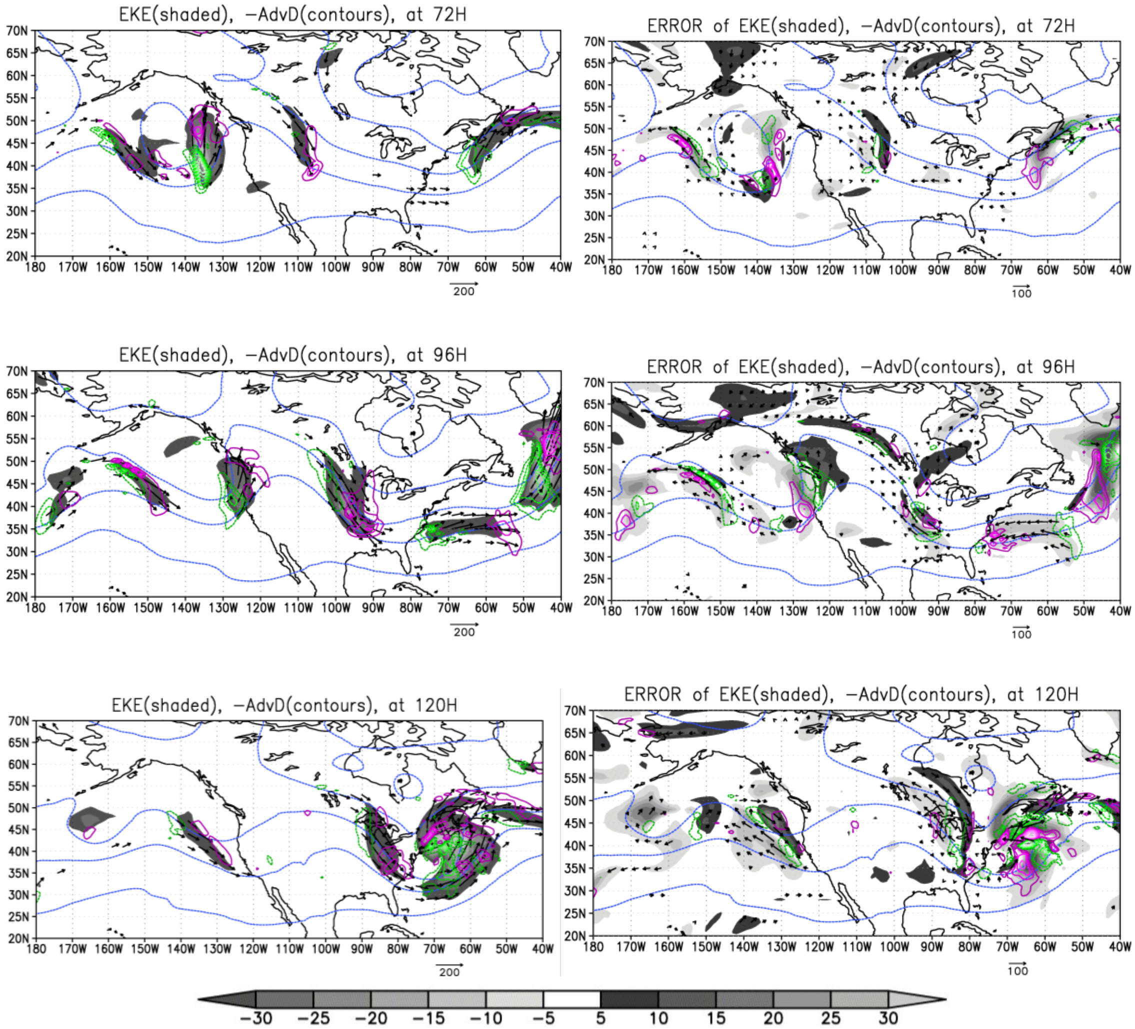
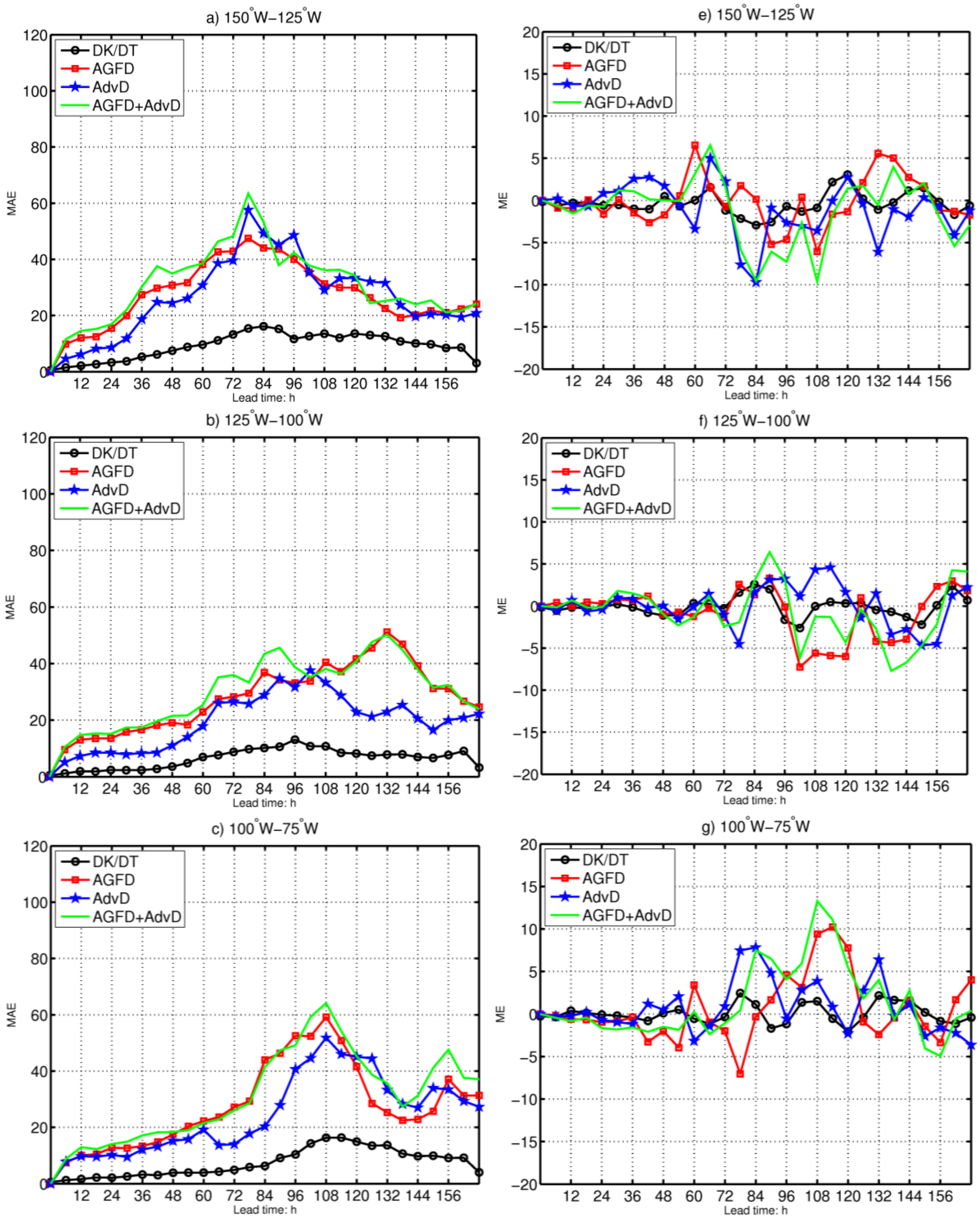


Figure 5.18: The same as Figure 5.17 but for the AdvF and its convergence (the advection term).


a) 150°W–125°W
b) 125°W–100°W
c) 100°W–75°W
e) 150°W–125°W
f) 125°W–100°W
g) 100°W–75°W
DK/DT
AGFD
AdvD
AGFD+AdvD
MAE
ME
Lead time: h

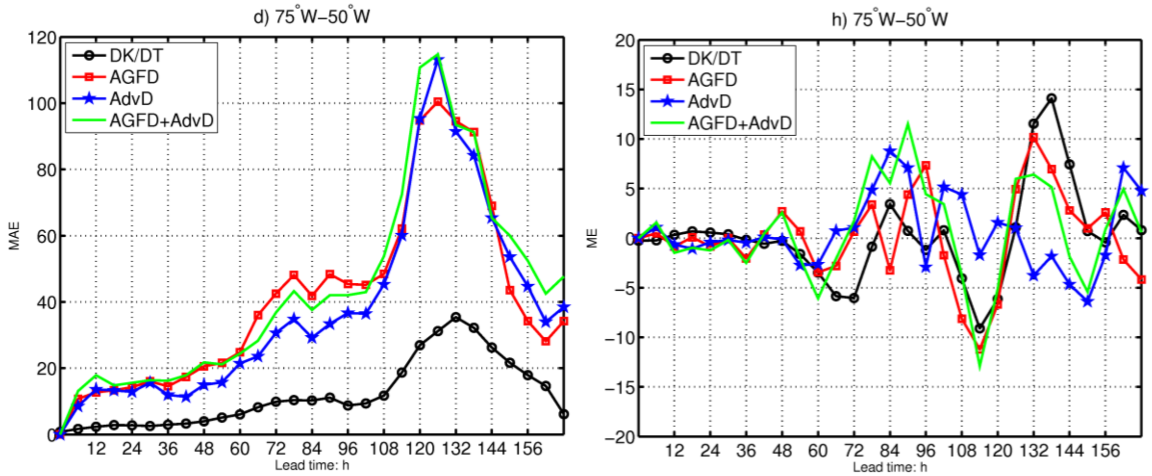


Figure 5.19: Averaged (a-d) MAE and (e-h) ME of the EKE tendency, -AGFD, -AdvD, and the total flux convergence for four regions (top to bottom: EP, 150°W-125°W; WN, 125°W-100°W; CE, 100°W-75°W; and WA, 75°W-50°W) between 30°N and 60°N. Unit: W m^{-2}.

Figure 5.19 shows the area averaged mean absolute error (MAE) as well as the mean error (ME) for the EKE tendency (DK/DT) and the two flux terms associated with the AGF and AdvF for four regions from west to east. For the MAE, the errors in DK/DT increases with the increase of the errors in the AdvFD term, which seem to lag (by ~6-24 hours) behind the growth of MAE in AGFD term. The peak of the MAE in DK/DT shifts with time to the east and grows rapidly between region CE (central and eastern North America, Figure 5.19c) and region WA (western Atlantic, Figure 5.19d). The peak value of MAE in DK/DT for a region seems to be influenced by the position of the aforementioned RWP relative to this region. For instance, the center of the RWP reaches the central longitude of region EP at 78 h (not shown), at which the MAE over region EP started to approach its peak value (Figure 5.19a). The amplitude of MAE in AGFD term is slightly larger than that of AdvD term when a region is located east of the center of the RWP, indicating that the errors in the AGFD term may contribute more to those in the DK/DT over the leading edge of the RWP or the further downstream areas of the leading edge.

A closer look at the mean errors (Figures 5.19e-h) show that the net mean errors in DK/DT over a region prior to the peak of MAE are grossly consistent with the net errors in the AGFD term, indicating that the initial errors in DK/DT might be triggered by the errors for the AGFD term when the region is downstream to the RWP, for example, between 24 h and 42 h over region EP in Figure 5.19e; and between 102 h and 120 h over region WA in Figure 5.19h. For region WA, the signs of the ME for the DK/DT were similar with the ME in the AGFD term between 108 h and 144 h (Figure 5.19h), indicating the errors in DK/DT are mainly attributable to the errors in the AGF convergence: the under-prediction between 102 h and 120 h and the over-prediction between 126 h and 156 h. Between 144 h and 168 h, errors in advection term contributed more to the errors in DK/DT over region WA. Note that the net errors in total advection can be much smaller or larger than those in DK/DT, such as during 66 h and 108 h for region EP in Figure 5.19e, and during 132 h and 144 h for region WA in Figure 5.19f, suggesting that other terms (e.g., the dissipation term, the barioclinic conversion term) were cancelling or enhancing the errors in total advection term.

Other than the control forecast, the EKE errors associated with two ensemble members (member 4 and 10) are also computed. Note that member 4 showed larger MAE among the ensemble than member 10 did over the verification box (80°W-50°W, 25°N-55°N) at 120 h. At 72 h (Figure 5.20), there was an AGF divergence error over the west coast around (125°W, 50°N) in member 4, suggesting less energy transferring from the EKE center C in Figure 5.16 to its northeast, leading to weaker amplitude and the east-northeastward displacement of the downstream center D. In contrast, there was an AGF convergence error over at same location in member 10, leading to excessive EKE errors over the northeastern side of Rocky Mountains. The overall errors over EKE center D was smaller in member 10 than in member 4. During 96-120 h

(Figure 5.20), the errors in the AGF downstream of center D were much larger in member 4 than those in member 10. Both members under-predicted the southern and eastern side of EKE center E. However, member 4 under-predicted the EKE much more than member 10 did. The errors in the AdvD term (Figure 5.21) for the two members followed the EKE errors. Member 4 had more advection errors than member 10 had at 72 h. Both members showed comparable errors in the advection term at 120 h.

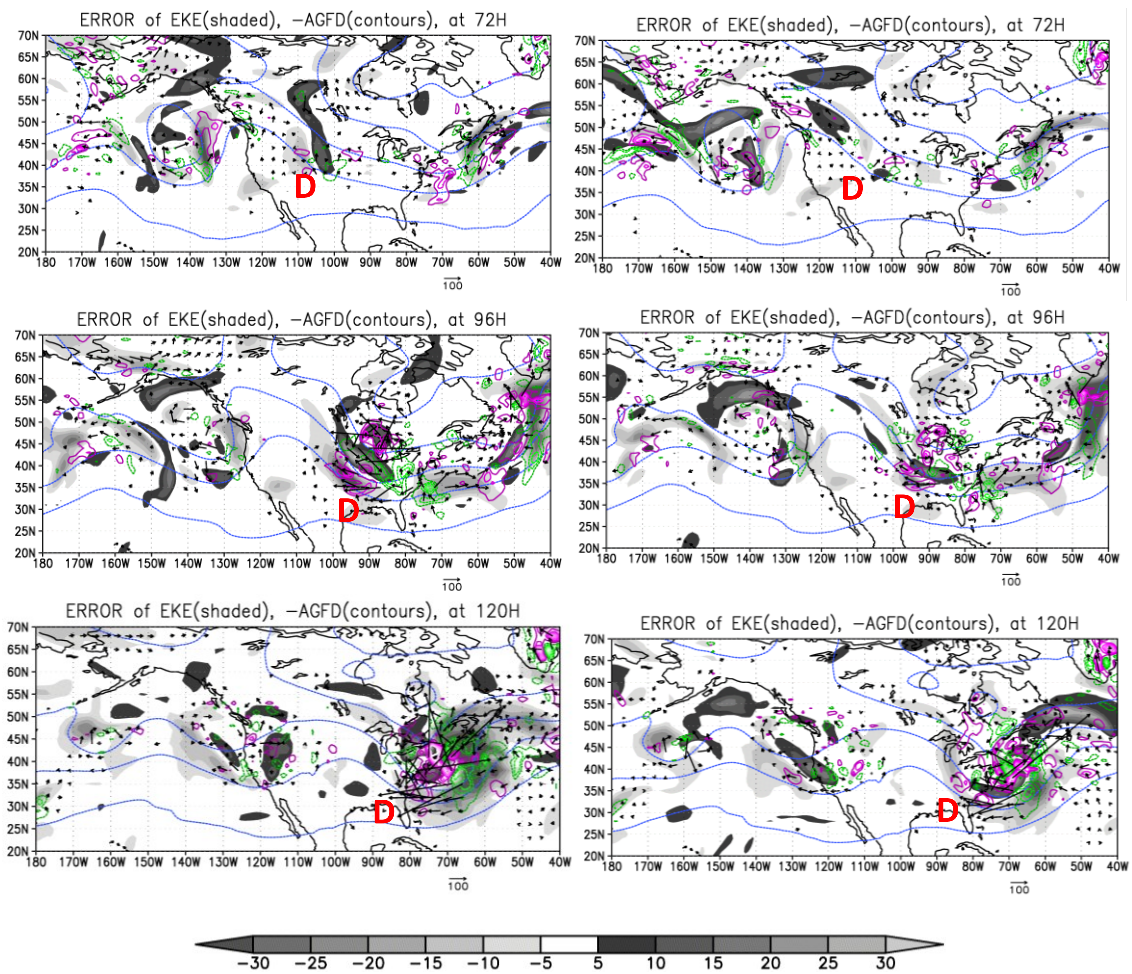


Figure 5.20: The same as right panels in Figure 5.17 but for member 4 (left panels) and member 10 (right panels) for lead times of 72 h, 96 h and 120 h.

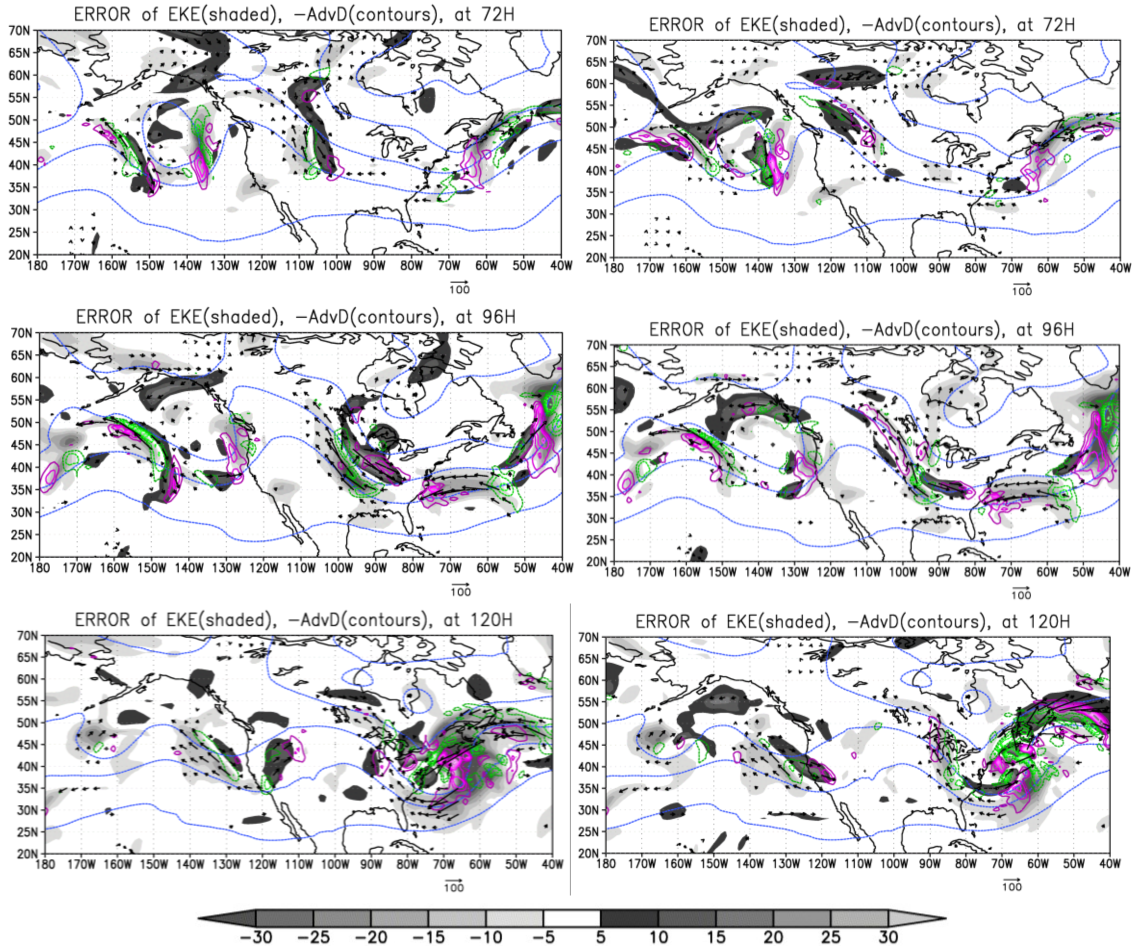
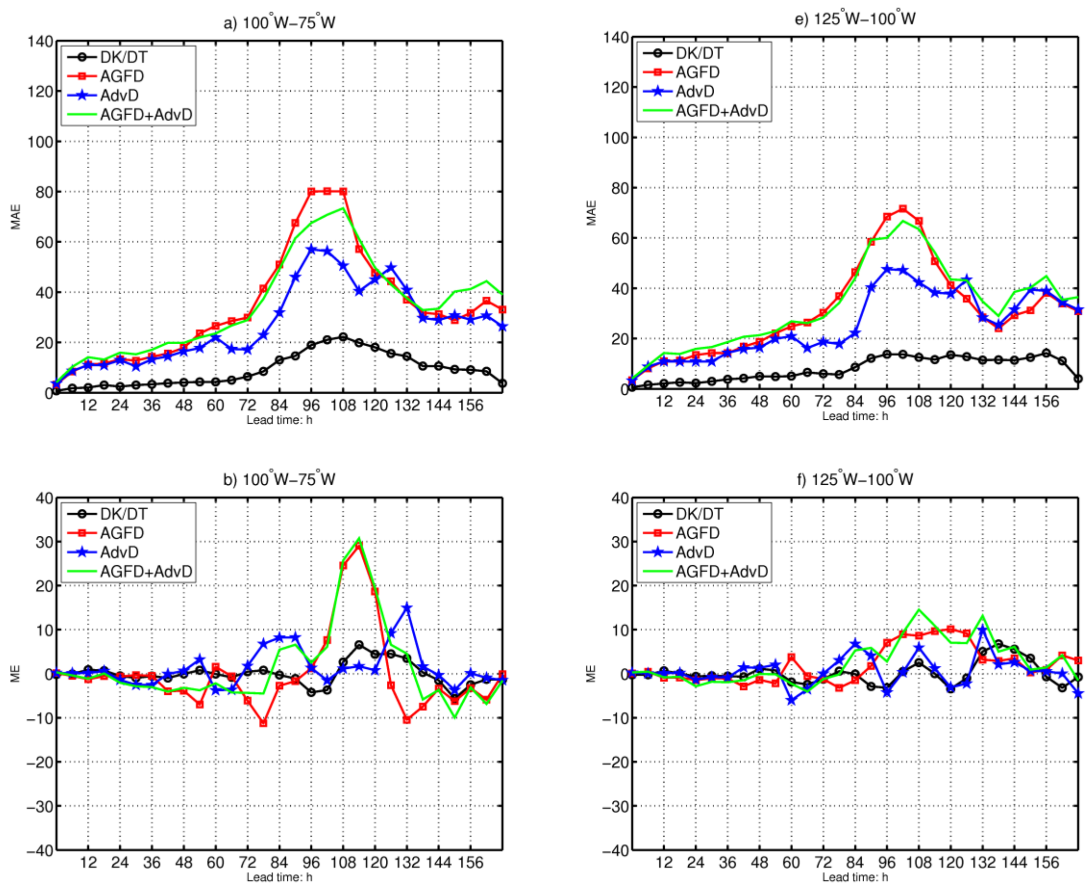


Figure 5.21: The same as Figure 5.20 but for the DK/DT errors associated with the AdvF term in member 4 (left panels) and member 10 (right panels).

Figure 5.22 compares the MAE and ME of the two members for regions CE and WA, respectively. For region CE (Figure 5.22), the EKE tendency errors peak around 108 h for member 4 and 96 h for member 10. The MAE in the AGFD term for both members was much larger than the MAE in the AdvF term during 54–120 h (Figures 5.22a, e), again suggesting that the AGFD term is more important in contributing to the errors in DK/DT over this region when it is over the central and leading regions of the RWP. However, the overpredictions of the AGFD term for member 4 were much larger than for member 10 especially during 96-120 h. Note that around the peak of DK/DT errors, the errors in the total advection were much larger than the DK/DT errors, suggesting that errors in other terms (e.g., baroclinic conversion term) might be fairly large and cancelling errors in the total flux terms especially in member 4 (Figure 5.22b). For region WA (Figures 5.22c, g), the MAE in DK/DT peaks at 132 h in member 4 and at 126 h in member 10. The DK/DT for member 4 was severely underpredicted in member 4 during 102-120 h and over-estimated between 126-150 h, which was mainly attributable to the errors in AGFD terms during 102-132 h and to the errors in AdvD term during 136-150 h (Figure 5.22h). Hence the total advective term dominate the large errors in DK/DT for member 4. For member 10, the DK/DT was only slightly under-estimated during 120-126 h and over-estimated during 132-150. Note that the reduction of the errors in AGFD might be the main reason for the larger underestimation errors associated with center E in member 4 prior to its mature stage (132 h).



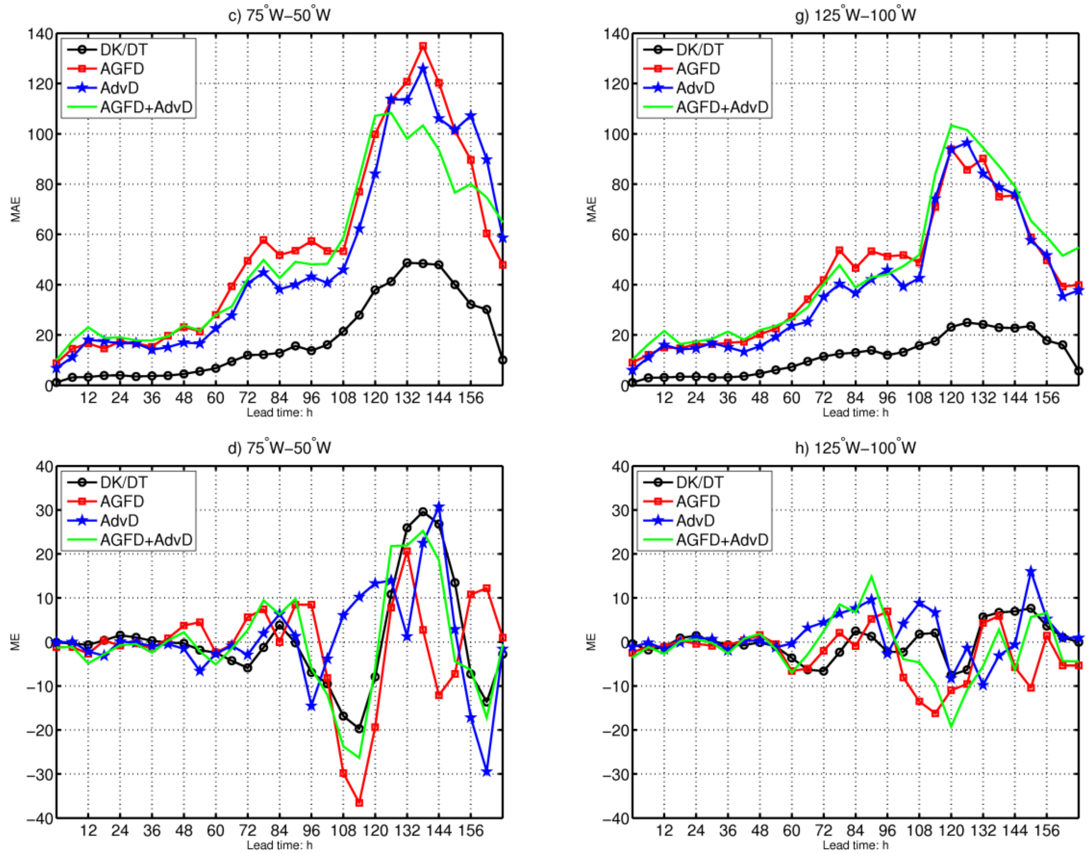


Figure 5.22: Averaged MAE (a, c, e, g) and ME (b, d, f, h) for member 4 (left, a-d) and member 10 (right, e-h) over region C (a, b, e, f) and D (c, d, g, h).

To sum up, the EKE errors increased with the AGFD and advection terms, in particular for the downstream regions over the eastern U.S. and western Atlantic Ocean, indicating that the two terms contribute significantly to the EKE local tendency. The comparisons among the three members show that the AGFD term is more important in modulating the EKE errors during the initial stages when there are significant AGF but small AdvF, in particular over the leading regions of a RWP. When the RWP propagates into northeastern Pacific, there are always change of shape over the west coast of North America. The directions of AGF can thus change rapidly, affecting the energy transfer between the RWP center and its leading areas. A proper representation of the AGF in models can modify the convergence or divergence areas and thus modulate the local EKE change when the convergence of the AGF is dominating the growth of a new EKE center. The initial locations of newly formed EKE centers are critical in impacting the downstream forecasts. Member 10 in this case study better resolved the initial EKE center (center D) with less AGFD errors, leading to a much better forecast of DK/DT over the downstream areas (eastern U.S. and western Atlantic). For instance, both the cyclone position and intensity of member 10 are closer to the analysis off the East Coast as can be seen from Figure 5.23. Nevertheless, the errors in the AGF and the AdvF terms together contribute significantly to the error growth of the local EKE center as well as the propagation of error peaks.

As can be seen from equation 5.6, the EKE tendency has other terms: the baroclinic term and the barotrophic terms. Previous studies suggested that the advection term, the AGFD term, and the baroclinic term are the dominant terms in EKE tendency budget (Chang 1993). However, the baroclinic term in a new EKE center only plays a secondary role in modifying the growth of the RWP during the initial stages. Due to the lack of the vertical velocity data (ω) in TIGGE archive, the baroclinic term was not calculated for this work. In this case, as well as many other coherent RWPs impacting the central and eastern U.S., an RWP often starts to decay when crossing the West Coast; the baroclinic term is not expected to contribute much in the local energy change over the central and west coast of the U.S. However, when there is strong baroclinicity over the East Coast and the western Atlantic Ocean, the baroclinic term could also contribute significantly to the error growth. We speculate that the rapid intensification of center E is also impacted by the baroclinicity over the southeastern U.S. during days 4-5. The averaged errors during 96-120 h for region CE showed that the other terms are also important other than the advective term (Figures 5.18 c, g, and 5.22a,b,e,f). However, our study confirms that the downstream development over the leading regions of a RWP crossing west coast of U.S.

provides an excessive error source to degrade the downstream forecast. In other words, when the thermal and moisture conditions are similar, the presence of an RWP from upstream can further degrade the forecast skills over the central and eastern U.S.

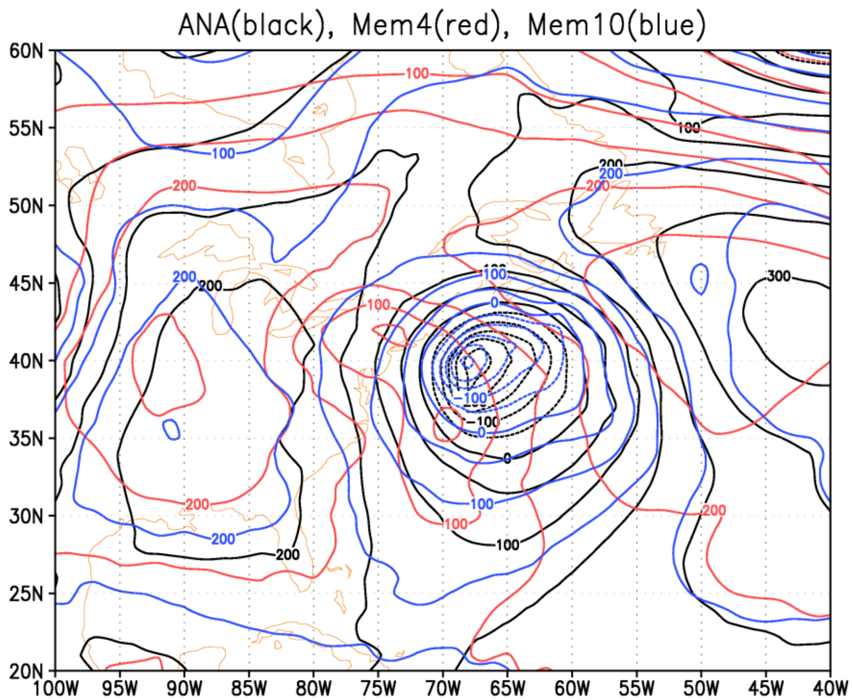


Figure 5.23: Geopotential height of 1000 hPa for analysis, member 4 and member 10 valid on day 5 (1200 March 26, 2014).

5.6 Chapter summary

This chapter has explored the impacts of RWPs on medium-range forecasts over eastern U.S. and western Atlantic in the cool season. Previous studies have suggested that RWPs exhibit downstream development characteristic (e.g. Chang 1993), with the co-existence of the decaying upstream waves and developing downstream new waves. A hypothesis was proposed that RWPs originating from Pacific can degrade the predictability over eastern U.S. and western Atlantic Ocean.

The RWPA anomalies in both large-error cases and large-spread cases over the verification region have been analyzed in sections 5.1 and 5.2. The composite results show that both the large-errors and large-spread cases are associated with enhanced RWPs from upstream from as far as western Pacific. The error maxima can shift from upstream of an RWPA anomaly to its downstream regions (see Figures 5.3-4, 5.7), indicating the errors might propagate with the speed comparable with the group velocity of RWPs.

Building upon the composite results as well as our observations of the operation daily RWPA, a conceptual model of errors/uncertainty development associated with RWPs has been proposed in subsection 5.3.1. The conceptual model shows that there are five stages of a RWP originating from the western or central Pacific and propagating into the continental North America. It suggests that large errors/uncertainties are mainly over the central and the leading edge of the RWP in most stages. One exception is the decay stage, during which the errors/uncertainties also develop over the trailing areas of the RWP. An investigation of the errors and ensemble spread in a typical RWP case confirms the validity of the conceptual model. ESA associated with the same case also demonstrate that the development of the sensitivities qualitatively fits the conceptual model.

In order to investigate the mechanism of how RWPs affect the downstream predictability, forecasts error growth associated with downstream development have also been investigated under the framework of eddy kinetic energy budget based on the same coherent RWP case in Section 5.5. The results show that the errors in the total advective term (AGF and AdvF) associated with downstream development contribute significantly to the errors in the local EKE tendency especially over the eastern U.S. and western Atlantic region. The AGF plays a significant role during the initial development of the EKE errors. A comparison between two ensemble members shows that the ensemble member (e.g. member 10 in the case study) that better resolved the initial EKE center with less AGFD errors subsequently had a much better forecast of the EKE tendency over the downstream areas (eastern U.S. and western Atlantic).

This page is left to be blank intentionally.

Chapter 6 Conclusions and future work

In this dissertation, we have explored the source and propagation of forecast errors and uncertainties in forecasting cool-season extratropical cyclones over the east coast of the U.S. and western Atlantic Ocean utilizing a multi-model ensemble from CMC, NCEP, and ECMWF. Ensemble sensitivity analysis has been employed to diagnose the upstream uncertainties associated with cyclone amplitude and position forecasts using the multi-model ensemble. The sensitivities determined by the ESA are verified by the LOOCV technique. In order to investigate the detailed evolution of different forecast scenarios in forecasting the high-impact winter storms, an EOF/fuzzy clustering tool has been applied to operational forecast to separate forecast scenarios in the multi-model ensemble. Building on the successful application of the EOF/fuzzy clustering method to cyclone case studies, a scenario-based ensemble verification method has been proposed to examine the capability of different models in capturing the analysis scenarios for historical East Coast winter storms. Meanwhile, the model errors from the three EPSs have been evaluated based on the leading EOF PC metrics. The role of the large scale RWPs originating from the Pacific Ocean in impacting medium-range forecasts over eastern North America and western Atlantic Ocean have been explored. The primary goal of the research presented here is to increase our understanding of the predictability of East Coast winter storms and apply novel techniques to reap the benefits of ensemble outputs and potentially improve medium-range forecasts.

6.1 Summary and conclusions

Because of the chaotic nature of the atmosphere, small differences in different model members at the initial times can be amplified or distorted at later forecast time steps. Thus the variations of a state variable at initial times may have a profound effect on medium range forecasts associated with a high-impact winter storm. Therefore, it is necessary to quantify the relation between the state variable at earlier forecast times and the forecast metric of interest in the medium range. Meanwhile, it is important to determine regions where small changes can be largely associated with the variations in forecasts over the verification region at a valid time, such as the intensity and track uncertainties of the cool-season storms over the eastern North America and western Atlantic Ocean in the medium range. In Section 3, the ESA was applied to diagnose the forecast uncertainties for cyclone amplitude and position forecasts using the multi-model ensemble. The application of ESA to a high-impact winter storm in December 2010 demonstrated that the sensitivity signals based on different forecast metrics (the EOF PCs, the MSLP run cycle differences, and the short-range forecast errors) are robust. In particular, the ESA based on the leading two EOF PCs can separate sensitive regions associated with cyclone amplitude and intensity uncertainties, respectively. The sensitivity signals were verified using the LOOCV method. Both 3-day and 6-day LOOCV results based on historical cyclone cases demonstrated that the forecast skills of ESA are fairly high in modifying the forecasts over the verification region. Thus, the sensitivity using the leading two EOF PCs can be quite reliable out to medium range.

Section 3.6 also presented the climatology of ESA for 3-day and 6-day forecasts based on the multi-model forecasts of historical cyclone cases. It was found that the EOF1 pattern often represents the intensity variations while the EOF2 pattern represents the track variations along west-southwest and east-northeast direction. For PC1, the upper-level trough associated with the East Coast cyclone and its downstream ridge are important to the forecast uncertainty in cyclone strength. The initial differences in forecasting the ridge along the west coast of North America impact the EOF1 pattern most. For PC2, it was shown that the shift of the tri-polar structure–the

East Coast trough and its adjacent ridges–is most significantly related to the cyclone track forecasts.

While the ESA tool discussed in Chapter 3 can provide an overall linear relation to diagnose forecast uncertainty related to a forecast metric, it may miss parts of the important information in the ensemble forecast given the nonlinear nature of the atmospheric flow. The detailed development scenarios in an ensemble can also be important in the operational forecast process and to improve the understanding of the predictability of winter storms. Meanwhile, effective tools to extract the ensemble information in a form that can provide the most important development scenarios could benefit the operational forecasters and decision makers given the increasing volume of ensemble data.

To quickly extract important information from a large ensemble and diagnose forecast uncertainty, Chapter 4 applied the EOF/fuzzy clustering tool to diagnose the scenarios in operational ensemble forecast of East Coast winter storms. It was shown that the clustering method can efficiently separate the forecast scenarios associated with East Coast storms based on the 90-member multi-model ensemble. The application of the EOF/fuzzy clustering tool to the 26-28 January 2015 storm event demonstrated that the forecast scenarios are well separated and consistent in different state variables such as the MSLP, Z500, and 24-h total precipitation.

Based on the successful application of the EOF/fuzzy clustering tool to case studies, we have proposed a scenario-based ensemble verification method and applied it to examine the capability of different EPSs in capturing the analysis scenarios for historical East Coast cyclone cases at lead times of 1–9 days (Section 4.5). The results suggest that the NCEP model performs better in short-range forecasts in capturing the analysis scenario although it is under-dispersed. The ECMWF ensemble shows the best performance in the medium range. The CMC model is found to show the smallest percentage of members in the analysis group and a relatively high missing rate, suggesting that it is less reliable in terms of capturing the analysis scenario when compared with the other two EPSs. A combination of NCEP and CMC models has been found to reduce the missing rate and improve the error-spread skill in medium- to extended-range forecasts.

By utilizing the scenario analysis, whether the ensemble mean from the multi-model ensemble or each individual models is really better than other subsets of an ensemble forecast has also been analyzed. It was found that in the majority of cases, the analysis does not lie within Group EM in the multi-mode ensemble. Meanwhile, the quadrant statistics suggest that the ECMWF model misses the analysis direction in around 2/3 of the past storms although it shows a slightly higher chance to be in the analysis quadrant in the medium range.

Based on the orthogonal features of the EOF patterns, the model errors for 1−6-day forecasts have been decomposed for the leading two EOF patterns. The results show that the NCEP model tends to better represent both EOF1 and EOF2 patterns by showing less intensity and displacement errors during 1−3 days. The ECMWF model is found to have the smallest errors in both EOF1 and EOF2 patterns during 3−6 days. The CMC model shows moderate errors for days 1−2 and the largest errors for days 3−6. We have also found that East Coast cyclones in the ECMWF forecast tend to be towards the southwest of the other two models in representing the EOF2 pattern, which is associated with the southwest-northeast shifting of the cyclone. This result suggests that ECMWF model may have a tendency to show a closer-to-shore solution in forecasting East Coast winter storms.

Motivated by the observation that the sensitivity signals sometimes develop and propagate like an RWP with downstream development (e.g., Figures 3.24-25), the relation between RWPs and the forecast errors/uncertainties in the GFS and GEFS models is investigated to explore the role of RWPs in the predictability of downstream regions over eastern U.S. and western Atlantic Ocean.

The composited RWPA anomalies show that there are enhanced RWPs propagating across the Pacific in both large-error and large-spread cases over the verification regions (sections 5.1-5.2. There are also indications that the errors might propagate with a speed comparable with the group velocity of RWPs.

Building upon the composite results as well as our observations of the operation daily RWPA, a conceptual model of errors/uncertainty development associated with RWPs has been proposed to serve as a practical tool to understand the evolution of forecast errors and uncertainties associated with the coherent RWPs originating from upstream as far as western Pacific. It suggests that the central and the leading regions of the RWP are the preferable regions for large errors/uncertainties to grow and develop. The errors and spread in a case study for a coherent RWP fits the conceptual model well. The ESA is also performed for this case study and the corresponding sensitivities also qualitatively fit the conceptual model.

In order to investigate the mechanism of how RWPs affect the downstream predictability, the forecasts errors associated with the downstream development have been investigated under the framework of eddy kinetic energy budget based on the same coherent RWP case in Section 5.5. Previous studies showed that the total advective term in the EKE budget equation is most closely related to the downstream development (Chang 1993; Orlanski and Shelton 1995). The results show that the errors in the total advective term contribute significantly to the errors in the local EKE tendency especially over the eastern U.S. and western Atlantic region. The errors in the AGF term plays a significant role in the initial development of the EKE errors. A comparison between two ensemble members shows that the ensemble member that has less errors better resolved the initial EKE center with less AGFD errors, leading to a much better forecast of the EKE tendency over the downstream areas (eastern U.S. and western Atlantic).

6.2 Implications

This dissertation applied ESA and clustering to investigate the source of forecast uncertainty and to separate the forecast scenarios for the medium range in forecasting cool-season cyclones over the East Coast of the U.S. Meanwhile, the three most widely used EPSs are comprehensively evaluated in terms of the scenario-based statistics, the quadrant statistics, the PC-based error-spread skills, and the EOF pattern-based intensity/displacement error decompositions. Finally how RWPs impact the predictability over the eastern U.S. was investigated. Consequently, a conceptual model was proposed and analyzed under the EKE framework. Results in this dissertation have important operational and scientific implications as summarized below.

i. The application of ESA in operational forecasting provides an efficient and intuitive way to help forecasters to track the source of the forecast uncertainty associated with an EOF pattern of interest (e.g. the close-to-shore or far-from-shore solutions). In past literature about ESA (e.g. Ancell and Hakim 2007; Torn and Hakim 2008; Zheng et al. 2013), the verification of the sensitivity signals had been a limitation. The main verification method of ESA has been based on initial-condition numerical experiments (Torn and Hakim 2009; Chang et al. 2013), which are computationally expensive and thus cannot be applied to a large number of cases. The verification using LOOCV method (Gombos and

Hoffman 2013) provides an efficient and robust way that allows us to verify the ensemble sensitivity signals for many cases. The results from the LOOCV demonstrated the goodness of the sensitivity signals in the multi-model ensemble forecasts we examined.

ii. The climatology of ESA for medium-range forecasts of East Coast cool-season storms can provide guidance to operational forecasters to help them identify and conceptualize the preferable regions for the development of forecast uncertainties during significant East Coast cyclone events. Meanwhile, the ESA climatology highlights the regions where model disagreement can impact the forecast most.

iii. The fuzzy clustering provides a complementary way to investigate the forecast uncertainties. Together with the ESA, they can help forecasters to make better use of the ensemble output. The ensemble verification based on scenario is a novel and efficient way to verify cyclone forecasts in the multi-model ensembles. Previous studies mainly employed the tracking and matching method to verify extratropical cyclones (Froude 2010), which only verify cases in which many ensemble members can be matched to the observed cyclone, as well as only the ensemble members that can be matched. Hence these verifications exclude a significant fraction of cyclone forecasts especially in the medium range. In this dissertation, all (most) of the cyclone cases and ensemble members are included in evaluating the model errors (calculating scenario-based statistics), hence we believe that the evaluations here can provide results that are complementary to the results from previous studies. Despite the difference in our method and the matching method, we have found consistent results with other studies (Froude 2010), for example the westward tendency in ECMWF model associated with the slow propagation speed of the forecast cyclones.

iv. We have shown that the ensemble mean group (the ECMWF mean) is not superior to other subsets of ensemble members in terms of the scenario (quadrant) statistics. Meanwhile, the quadrant statistics suggests that although the ECMWF ensemble shows a slightly higher chance to be in the quadrant in which the analysis falls for the medium range, it also misses the analysis direction in around 2/3 of the cases. A few clarifications were made about the benefit of ensemble mean. The above results have practical implications in operational forecast, suggesting that: 1) one should not put too much emphasis on the ensemble mean and ignore the other groups in the multi-model ensemble; 2) it is not a good practice for forecasters to hedge towards the ECMWF ensemble mean solution as we often see, even though the ECMWF might be the best ensemble. All scenarios must be taken into consideration in the formulation of a forecast.

v. The composite results based on large-error and large-spread cases suggest that the RWPs originating from Pacific Ocean and propagating into North America degrade the predictability over the eastern U.S. and western Atlantic Ocean. These results are consistent with the findings based on the forecast busts over Europe in Rodwell et al. (2013). However, other studies (Grazzini 2007; Grazzini and Vitart 2015) suggested that the atmospheric predictability may be enhanced with the presence of coherent RWPs. The differences might be due to the metrics used to represent RWPs as well as the different forecast lead times, for instance, the focus of Grazzini and Vitart (2015) is on the medium- and long-range forecasts (beyond day 8).

vi. The conceptual model proposed in Section 5 represent a practical tool to illustrate how RWPs can modulate the development of forecast errors/uncertainties, thereby the medium-range predictability of East Coast high-impact weather event. One caveat is that the observed RWP objects in Figure 5.9 are not necessarily similar with the exact locations and the relative amplitudes. Meanwhile, the EKE based budget is creatively applied in diagnosing downstream-development-related errors. It provides a physical interpretation of the mechanism for RWPs to impact the cyclone forecasts over the eastern U.S.

vii. Both ensemble sensitivity and the fuzzy clustering approaches have been applied to real-time ensemble forecasting out to medium range as part of our collaborations with the NOAA Environmental Modeling Center (EMC) group. The near real-time products and web interfaces of ESA and fuzzy clustering provide an efficient and intuitive way for operational forecasters to interpret the ensemble outputs.

6.3 Future work

The theme presented in this dissertation provides a framework for ongoing and future work. Some potential future research paths are listed as follows.

i. Further comparisons of ESA and other sensitivity analysis methods, such as the ETKF and adjoint sensitivity. We have shown in Section 3.5 that the ETKF signals are qualitatively consistent with the ESA signals in a case study. Quantitative comparisons between them using more cases are necessary to improve the understanding of the dynamical relation between a response function and state variables and to provide confident guidance for targeting and assimilating new observations.

ii. Perturbed initial condition experiments can be performed to differentiate significant sensitivity signals from spurious signals, as well as to validate the verifications based on the LOOCV method.

iii. To examine how differences in large scale flow patterns can impact coherent ESA signals. Preliminary suggested that the blocked events over eastern Pacific have an impact on the propagation of ESA. The future work will investigate the differences in ESA for the zonal and blocked flows.

iv. Given its simplicity and low computational cost, the ESA can be applied to other high-impact weather events. The impact-based response functions can be developed to help the decision makers and to benefit the end-users.

v. The OOE cases represent the cases with the poorest forecasts. Detailed case studies are needed to understand what factors might have led to the poor predictability. Numerical experiments can be designed to test those related factors either using different physics under a similar model or using the different models (e.g. WRF and CAM) with similar physics.

vi. The fuzzy clustering method and the scenario-based verifications can be applied to other high impact weather events, such as tropical cyclones and extremely heavy precipitation events.

vii. The model errors determined in Section 4.5 can be diagnosed using the same models but using the physics or data assimilation method from another model to test the reasons for the model errors.

viii. The relation between RWPs and the forecast dropout cases provides motivation that future work can be done to investigate in more details the relation between RWPs, flow patterns, and cyclogenesis.

ix. Some preliminary results based on case studies have suggested that the ensemble sensitivity using the metrics of total precipitation also show characteristics of RWPs. The impacts of RWPs on the predictability of large-scale precipitation events, in particular the shifting of the heaviest precipitation region, can be investigated to potentially improve the forecasts on the boundary of the heaviest precipitation region.

This page is left to be blank intentionally.

Appendix 1

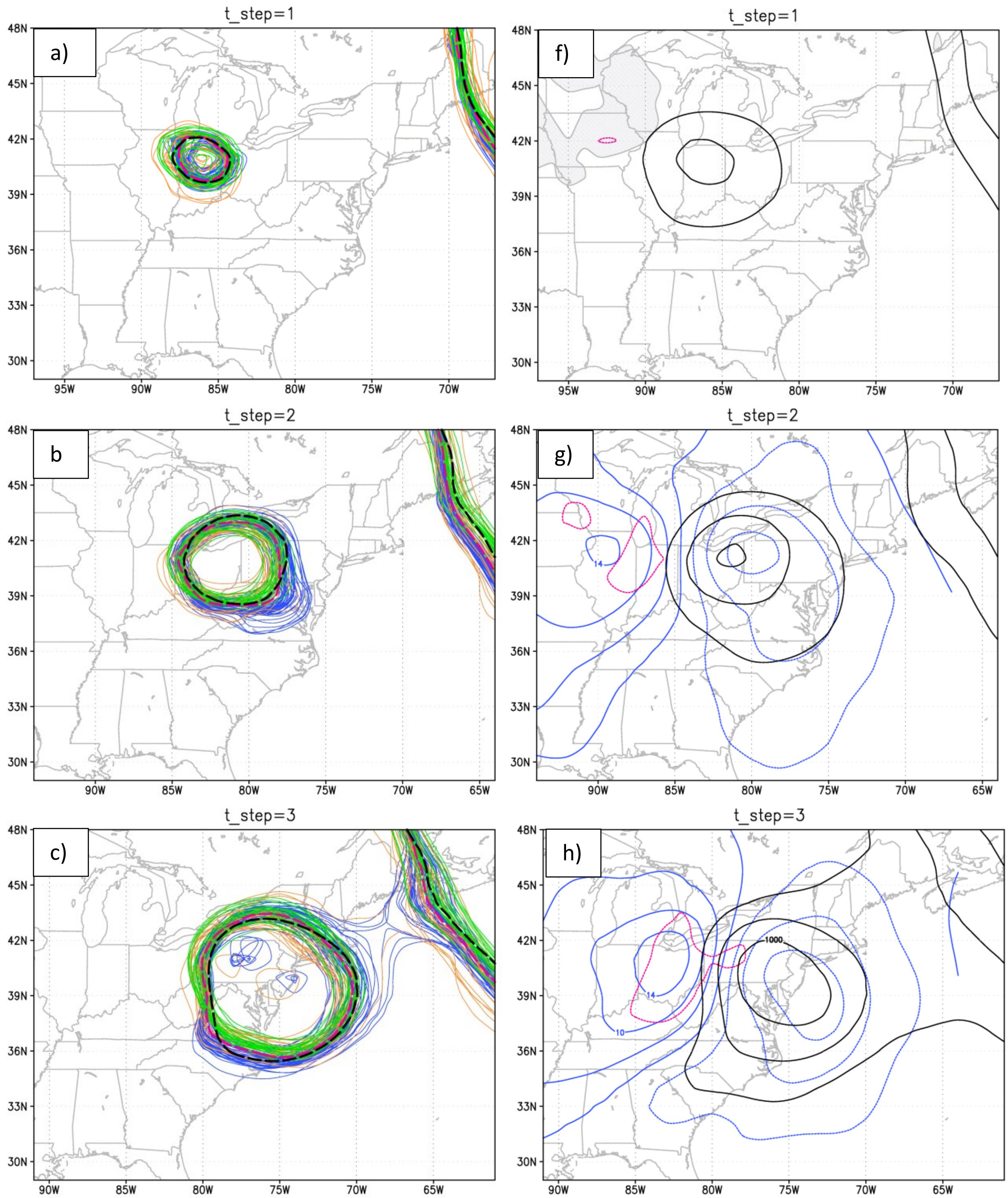
This appendix summarizes the OOE cases as listed in the Table A4.1 for Section 4.6.

Lead Time	Case	Initial Time	Verification time	Description of weather event
1 day	**1a**	**2008123100**	**2009010100**	**Northeastern U.S. winter Storm**
	1b	2015032600	2015032700	N/A
2 day	2a	2010012100	2010012300	N/A
3 day	3a	2008120912	2008121212	Eastern U.S. mixed precipitation event
	3b	2009012600	2009012900	Southern Plains/Ohio Valley/Northeast U.S. winter storm
	3c	2011010612	2011010912	N/A
	3d	**2014011912**	**2014012212**	**Midwest to Eastern U.S. winter storm**
4 day	4a	2007122300	2007122700	N/A
	4b	**2008122800**	**2009010100**	**Northeastern US winter Storm**
	4c	2009012500	2009012900	N/A
	4d	2012122600	2012123000	N/A
	4e	**2014011800**	**2014012200**	**Midwest to Eastern U.S. winter storm**
	4f	2014021212	2014021612	Northeast winter storm
5 day	5a	2008111100	2008111600	N/A
	5b	**2008122700**	**2009010100**	**Northeastern U.S. winter Storm**
	5c	2009121412	2009121912	Eastern U.S. winter storm
	5d	2009121500	2009122000	Eastern U.S. winter storm
	5e	2010020812	2010021312	Southern snowstorm
	5f	2010122212	2010122712	Eastern U.S. winter storm
	5g	2011120312	2011120812	N/A
	5h	2013020400	2013020900	Great Lakes and Northeast major winter storm
	5i	2013102900	2013110300	N/A
6 day	6a	2008020112	2008020712	Northeastern winter storm
	6b	2008031512	2008032112	N/A

	6c	2008120612	2008121212	Eastern U.S. mixed precipitation event
	6d	**2008122600**	**2009010100**	**Northeastern U.S. winter Storm**
	6e	2009110700	2009111300	N/A
	6f	2009121312	2009121912	Eastern U.S. winter storm
	6g	2009121400	2009122000	Eastern U.S. winter storm
	6h	2010022300	2010030100	N/A
	6i	2010022512	2010030312	N/A
	6j	2010110212	2010110812	N/A
	6k	2011011712	2011012312	N/A
	6l	2012020600	2012021200	N/A
	6m	2012021412	2012022012	N/A
	6n	2013020312	2013020912	Great Lakes and Northeast major winter storm
	6o	2014020812	2014021412	Southern Plains to East Coast winter storm
	6p	2015012500	2015013100	Northeastern U.S. winter storm

Table A4.1: List of OOE cases for 1–6-day lead times in this study, their forecast lead time, initial/verification time, and the verified weather event summarized by NOAA/WPC forecasters. N/A denotes that the case is not a significant large-scale storm event and not included in the NOAA WPC storm summary. Cases highlighted by bold and/or colored fonts represent events verified at the same valid time.

Figures A4.1-2 show the two OOE cases for 1 day forecasts. Case 1a had a cyclone over the northern Indiana with the central pressure less than 1005 hPa at initial time (Figures A4.1a, f). The cyclone moved east-southeastward with large negative pressure change ahead of its track during time step 1-3 (initial time to 12-h forecast; Figures A4.1a-c, and f-h). Meanwhile, the multi-model ensemble forecast errors show a negative center (~2 hPa) west of the analyzed cyclone center, suggesting the ensemble mean forecasts a cyclone a bit more westward and slightly deeper than the analyzed cyclone. The analyzed cyclone deepened rapidly from time step 3 to 4 (12 to 18 h forecasts; Figures A4.1h, i) with the 6-hr pressure drop of 18 hPa around the analyzed cyclone center at step 4 (18 h; Figure A4.1i). The ensemble spread shows an over 2-hPa increase, which is large when compared with the previous time steps. Forecast errors tend to form an asymmetric dipole around the analyzed cyclone, suggesting the model forecasts a cyclone more westward and slightly deeper than the analyzed cyclone. At time step 5 (day 1 forecast; Figure A4.1j), the analyzed cyclone continued to intensify with a central pressure of 982 hPa. The spaghetti plot (Figure A4.1e) clearly shows that the analyzed cyclone is more east-northeastward with respect to the overall envelope. The forecast errors grew quickly with the negative value reaching 6hPa. One thing worth noting is the forecast error centers overlapped with the largest spread change areas. Previous study suggested that the change of ensemble spread can be used to quantify the change of flow nonlinearity (Grimit and Mass 2007; Du and Zhou 2011). The large errors in this case seemed to be associated with the large nonlinear process during rapid cyclone development.


t_step=1
a)
f)
t_step=2
b
g)
t_step=3
c)
h)
48N
45N
42N
39N
36N
33N
30N
95W
90W
85W
80W
75W
70W
65W
14
10
1000

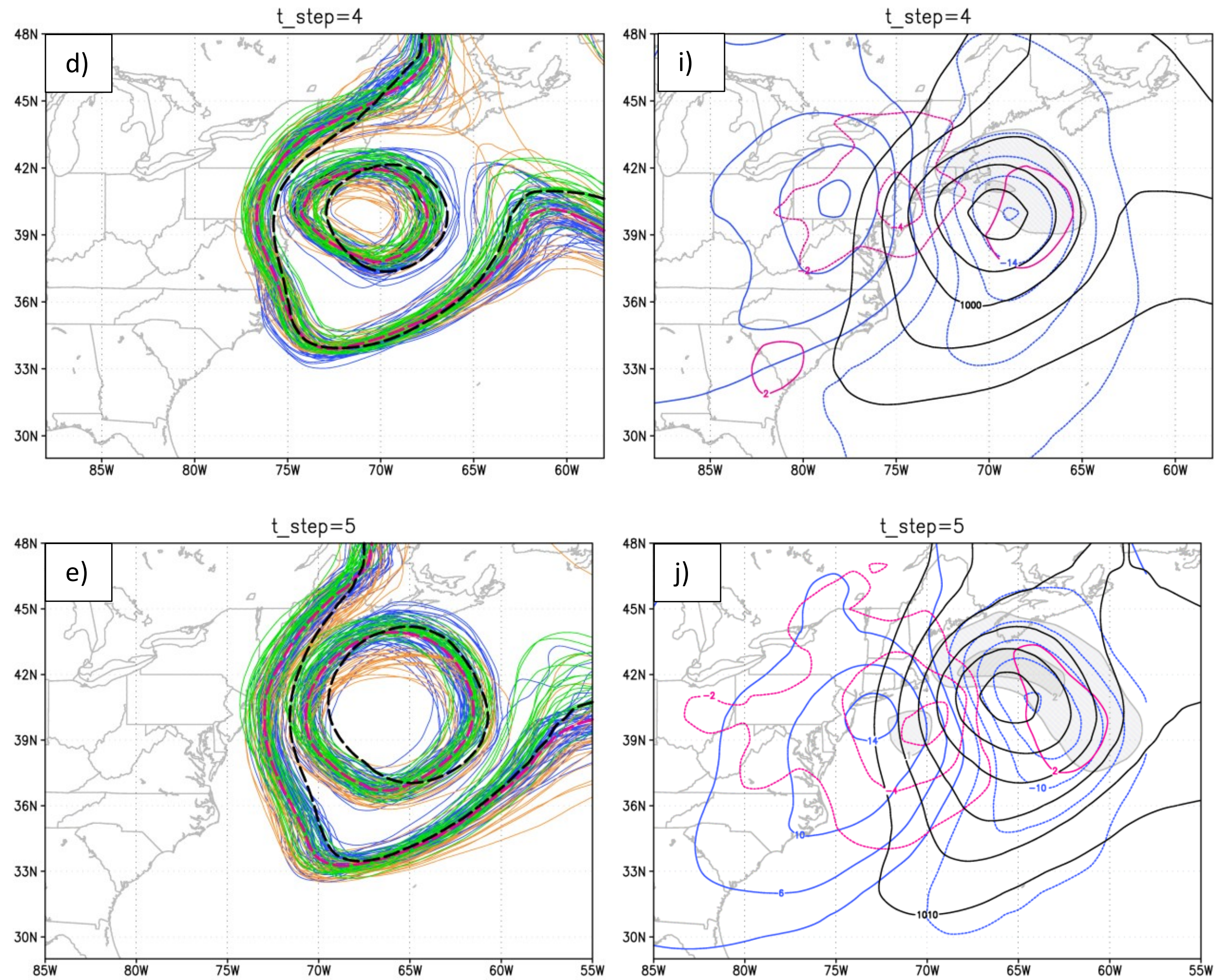
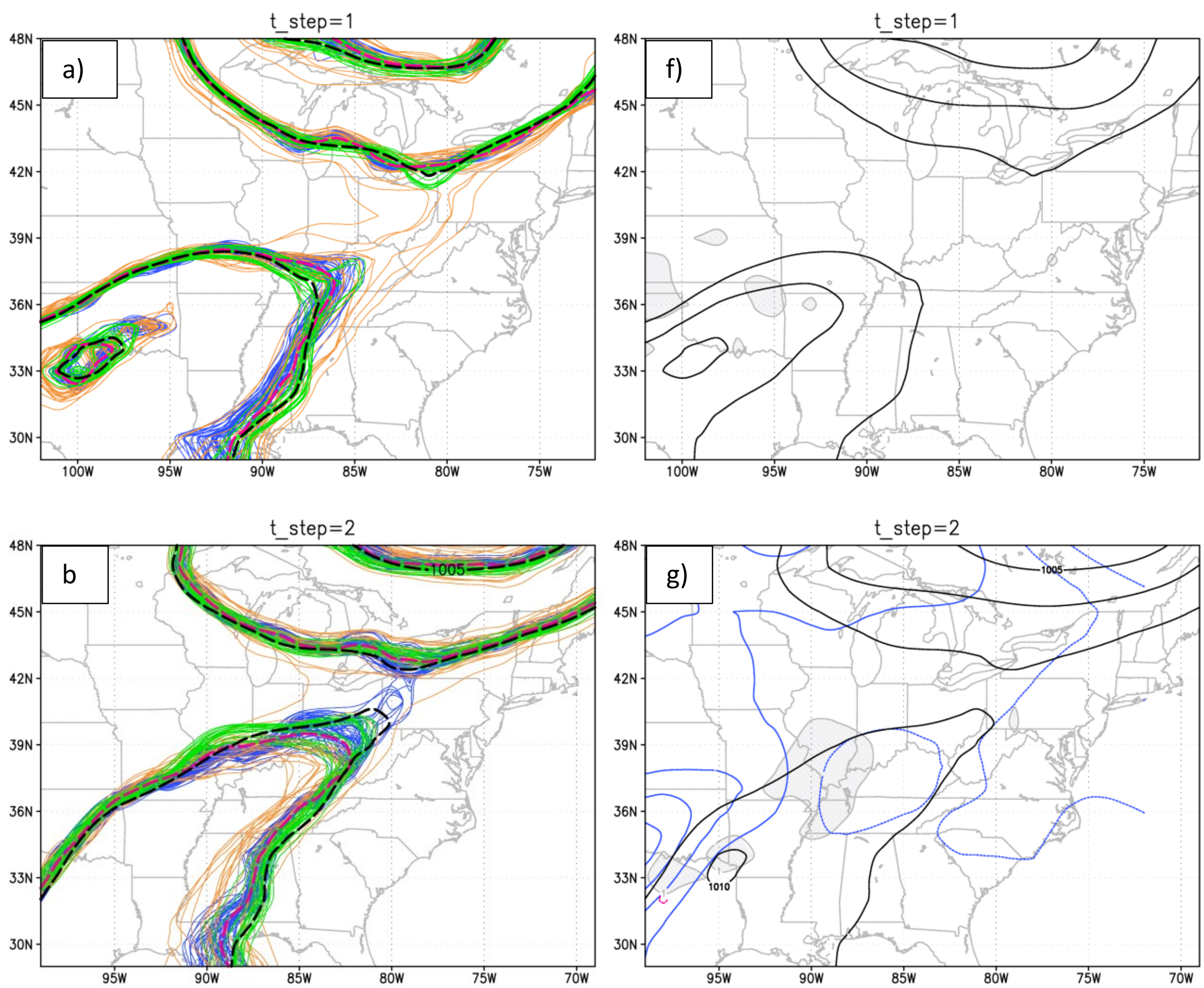


Figure A4.1: Day 1 OOE case: case 1a in table A4.1 spaghetti plot for 1005 and 995 hPa MSLP contour line (a-d, left) at different forecast step; and (e-h, right) ensemble mean forecast error (magenta dashed) with an interval of 2 hPa starting from ±2hPa, pressure change (blue contours) between two time steps with an interval of 4 hPa, spread change (gray shades and contour with an interval of 1 hPa), and analyzed MSLP contour (solid black lines). Unit for all plots: [hPa]. The time step is 6 hr.

Case 1b (Figures A4.2a, f) shows an area of low pressure located at the surface in northern Texas at initial time. At time step 2 (6 h; Figures A4.2b, g), the low moved to its northeast and had weakened by 5 hPa. From time step 3 (12 h; Figures A4.2c, h), another low center formed over the south Appalachia within the broad low pressure area. Meanwhile, negative forecast errors of 2 hPa developed to the west of the analyzed cyclone. During time steps 4 and 5 (18 and 24 h; Figure A4.2d-e, i-j), the multi-model ensemble overall forecasts a more west-northwestward cyclone than the analyzed cyclone. This case shows that the forecast errors are mainly due to the misrepresentation of cyclone center position on the background of a broad low surface area in models. Although the multi-model has a small ensemble spread (1 hPa), the systematical northwest shifting error in the models still render it to be an OOE case. A further examination confirmed that there coexisted a coherent RWP from initial time (not shown), indicating that the RWP may have degraded the forecast skill of cyclone position in this case.



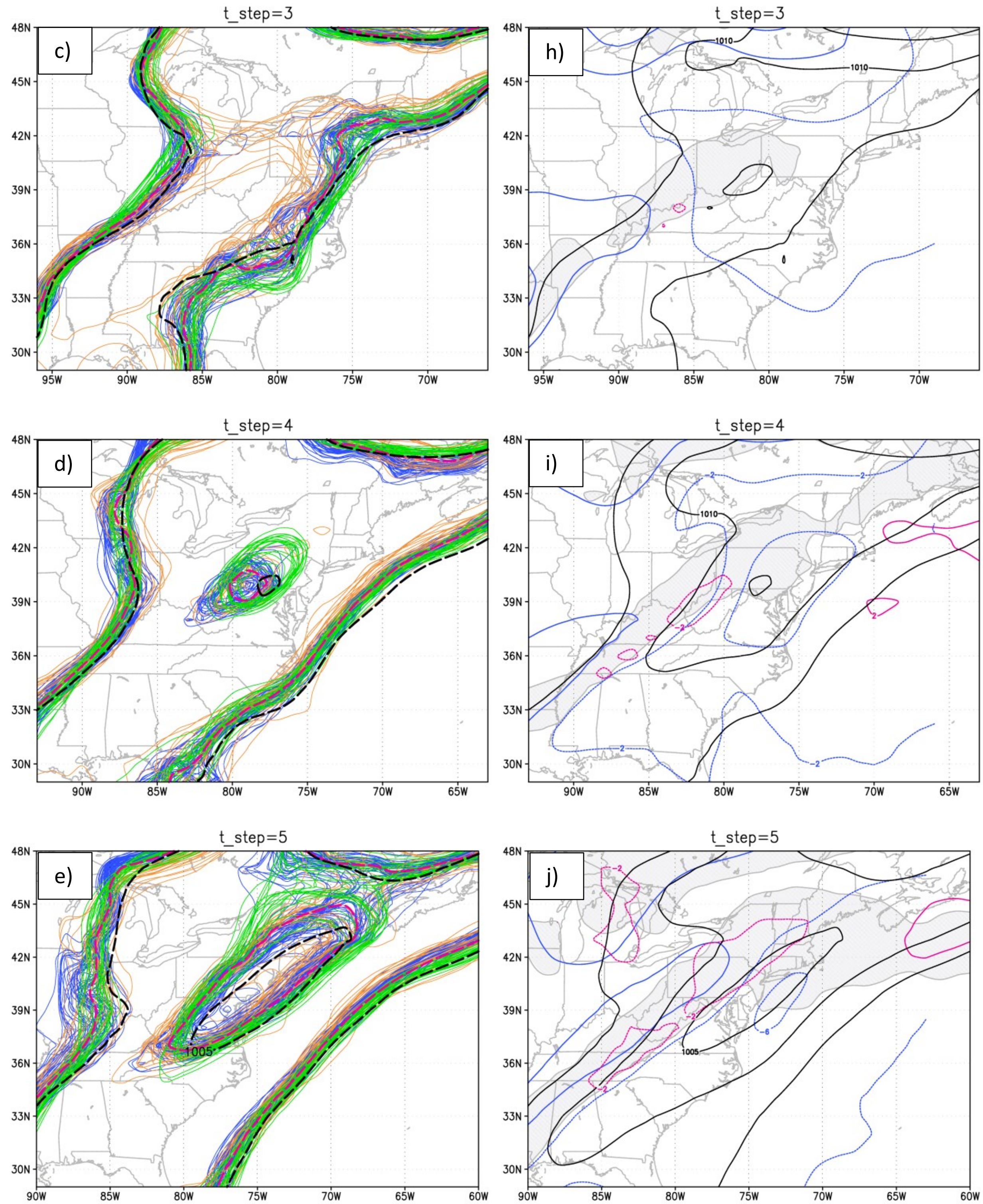
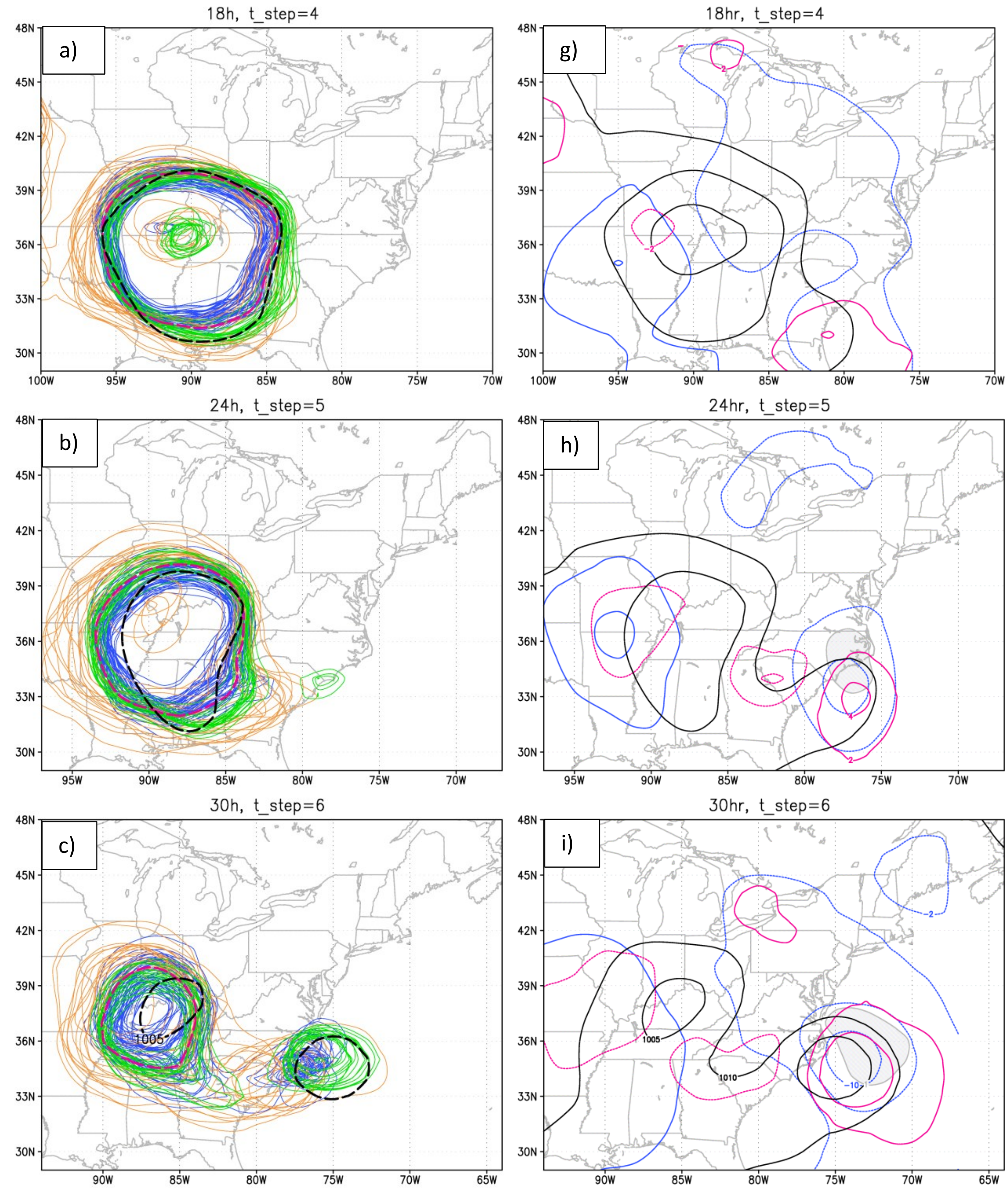


Figure A4.2: The same as Figure A4.1 except for case 1b in Table A4.1 and spaghetti plot using 1015 and 1005 hPa MSLP contour lines. The shades in f-j represent the spread instead of spread change.

The single OOE case for day 2 is shown in Figure A4.3, which represents a typical Miller type B cyclone. At time step 4 (18 h; Figures A4.3a, g), the primary cyclone was over the southwestern Ohio Valley with a minimum pressure of 999 hPa. The multi-model mean showed negative errors of -3 hPa northwest to the cyclone center, and positive errors of 4 hPa southeast of the Georgia coast. At time step 5 (24 h; Figures A4.3b, h), a new surface low area appeared off the South Carolina coast, where forecast errors increased to 4 hPa accompanying positive ensemble spread change. Models seemed to shift the second low too northwestward and forecasted the primary cyclone a bit deeper and more westward than the analysis. From time step 5 to 6 (24-30 h; Figures A4.3c, i), the secondary cyclone off the Mid-Atlantic coast developed rapidly with a pressure drop of ~12 hPa. The secondary low in the models was too weak and too westward while the primary low was a bit deeper and more westward than the analyzed low. Large positive errors are over the region of large spread increase. From time step 6 to 8 (30-42 h; Figures A4.3c-e; i-k), the cyclone continued to develop rapidly with large pressure drop. The forecast errors are mainly positive over the center and northeastern side of the analyzed cyclone accompanying with spread increase. Between 42 h and 48 h (Figures A4.3f, l), large negative errors developed over the western side of the cyclone together with a spread increase, indicating the westward shift became the dominant error.

In this case, the NWP models cannot predict the secondary low's position and amplitude correctly at its onset time. The corresponding errors developed rapidly with the rapid development of the secondary low. Previous studies suggested that rapid cyclone development is often the result of nonlinear interaction among different processes, such as strong upper-level forcing, the intrusion of stratospheric high PV air, latent heat release, surface energy fluxes from the ocean, as well as enhanced local baroclinicity from differential diabatic heating, although all these factors need not be present (Uccellini 1995; Wang and Rogers 2001) at the same time. The overlap of error maxima and positive spread change again indicate that the models tended to continue to provide a linear solution while the nonlinear processes lead to rapid development.


18h, t_step=4
a)
18hr, t_step=4
g)
24h, t_step=5
b)
24hr, t_step=5
h)
30h, t_step=6
c)
1005
30hr, t_step=6
i)
1005
1010
48N
45N
42N
39N
36N
33N
30N
100W
95W
90W
85W
80W
75W
70W
65W

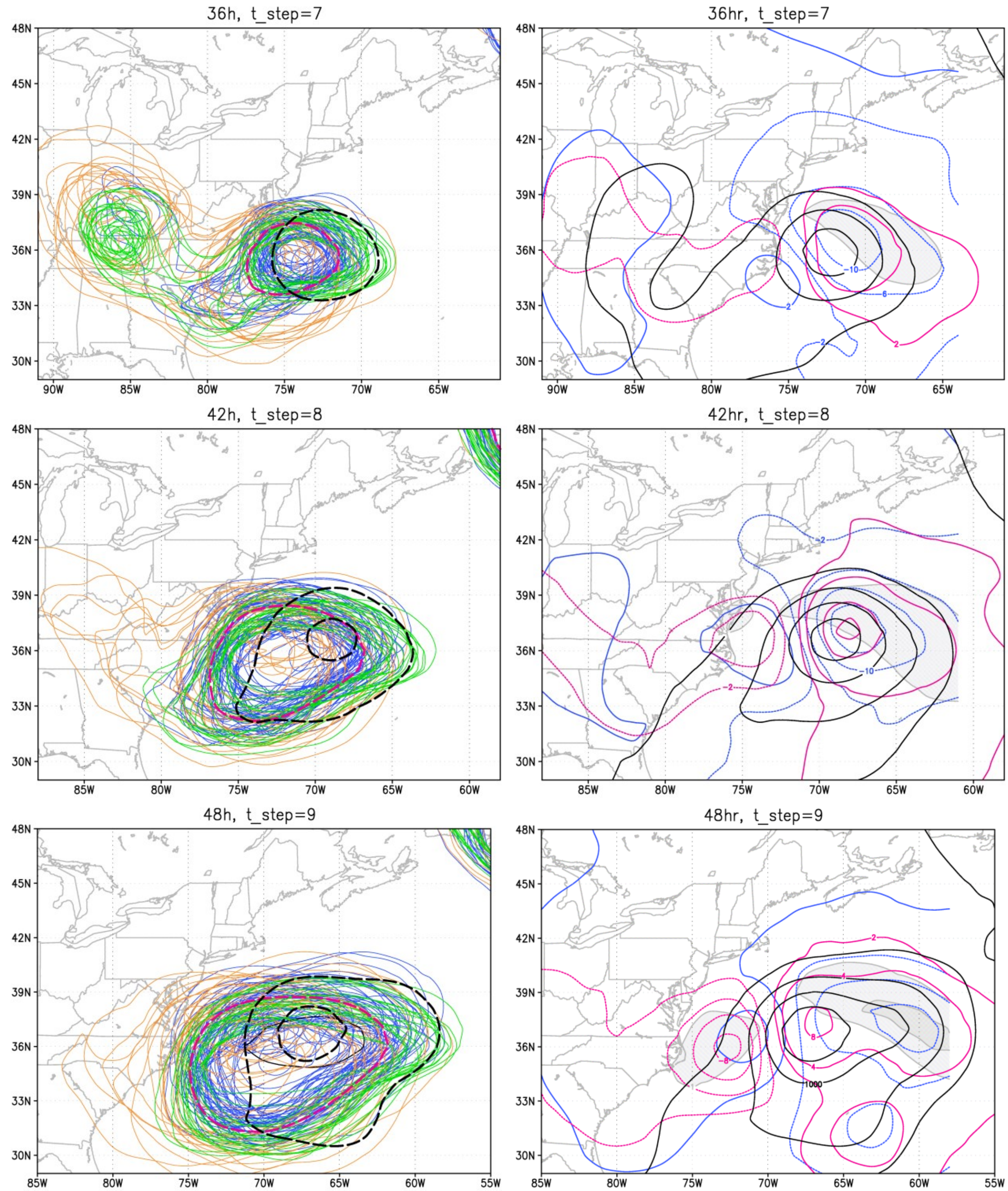


Figure A4.3: Day 2 OOE case: case 2a spaghetti plot for 1005 and 995 hPa MSLP contour line (a-f, left) at different forecast step; and (g-l, right) ensemble mean forecast error (magenta dashed), pressure change (blue contours) between two time steps, spread change (gray shades and contour), and analyzed MSLP contour (solid black lines). Unit for all plots: [hPa]. Each time step represents 6 hr.

As for the four 3-day OOE cases, case 3a is associated with a surface low tracking from the Gulf Coast across the Mid-Atlantic northward through southern New England and into Northern Maine (not shown). This case involved the re-development of the cyclone and strong nonlinearity ahead of the cyclone track, making the forecast extremely difficult. The cyclone in the model was around ~650km more northeastward than the analyzed cyclone was (Figure 4.12). This displacement error exceeds the double value of the typical displacement errors for 3-day forecasts for an East Coast storm in the GFS model, which is ~250 km in Colle and Charles (2011). Case 3a was also confirmed as a high-impact storm case by NOAA HPC forecasters.

Case 3b is a high-impact ice storm case originating from the southern Plains. The storm in the ensemble models was too weak and too slow. The errors are partly due to the rapid development of the cyclone as well as the increasing speed (~488 km (6 hr)$^{-1}$) of the storm.

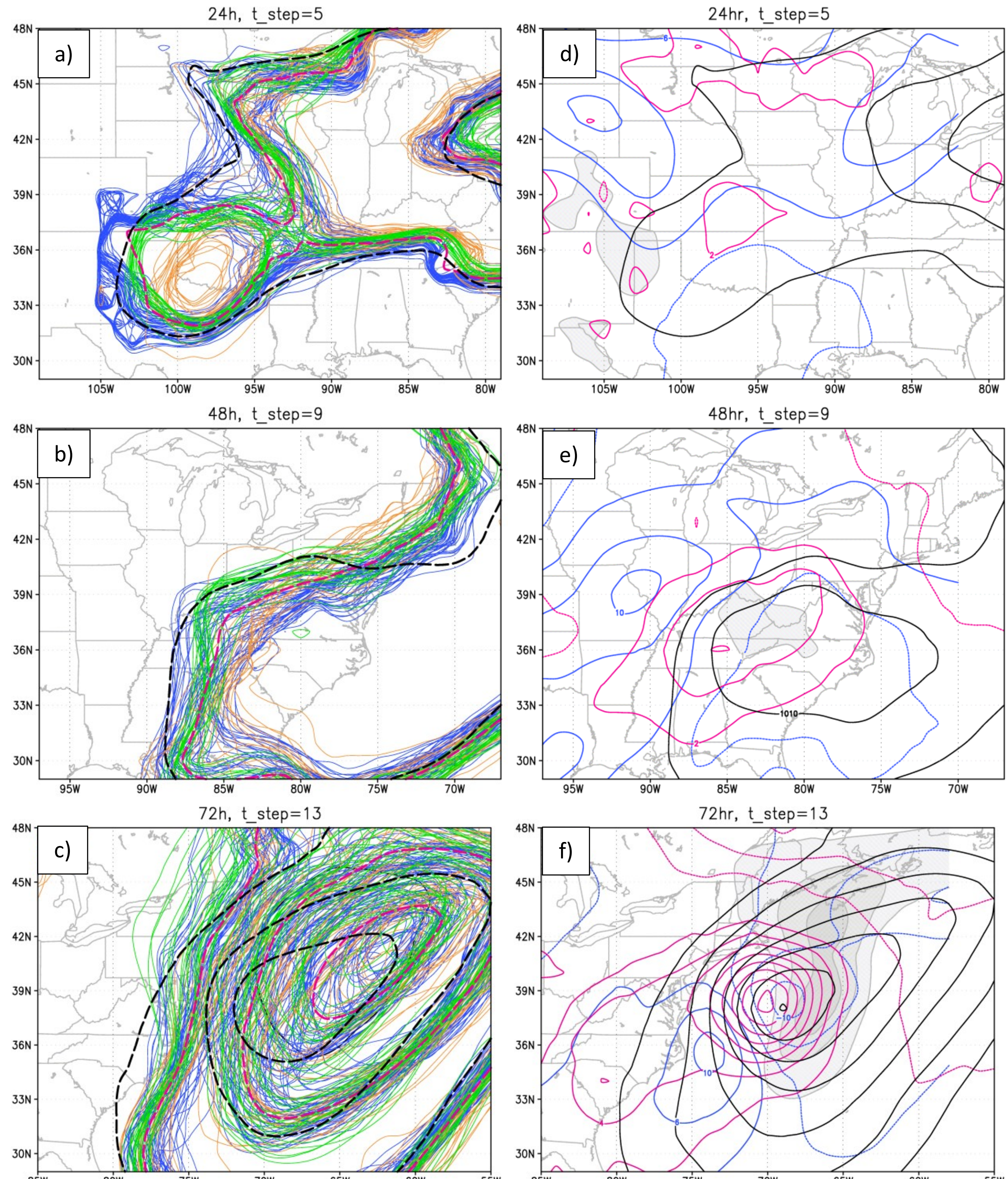
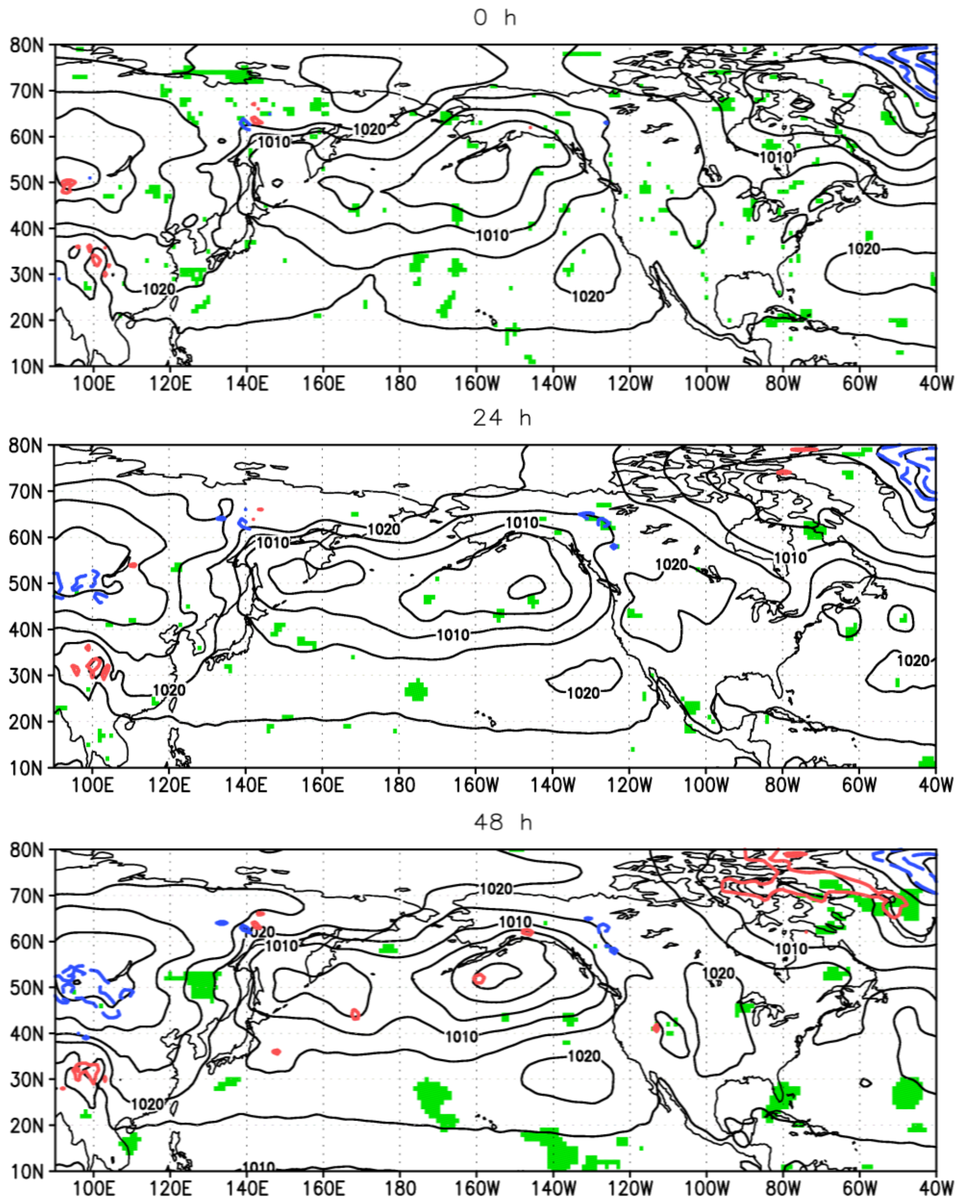


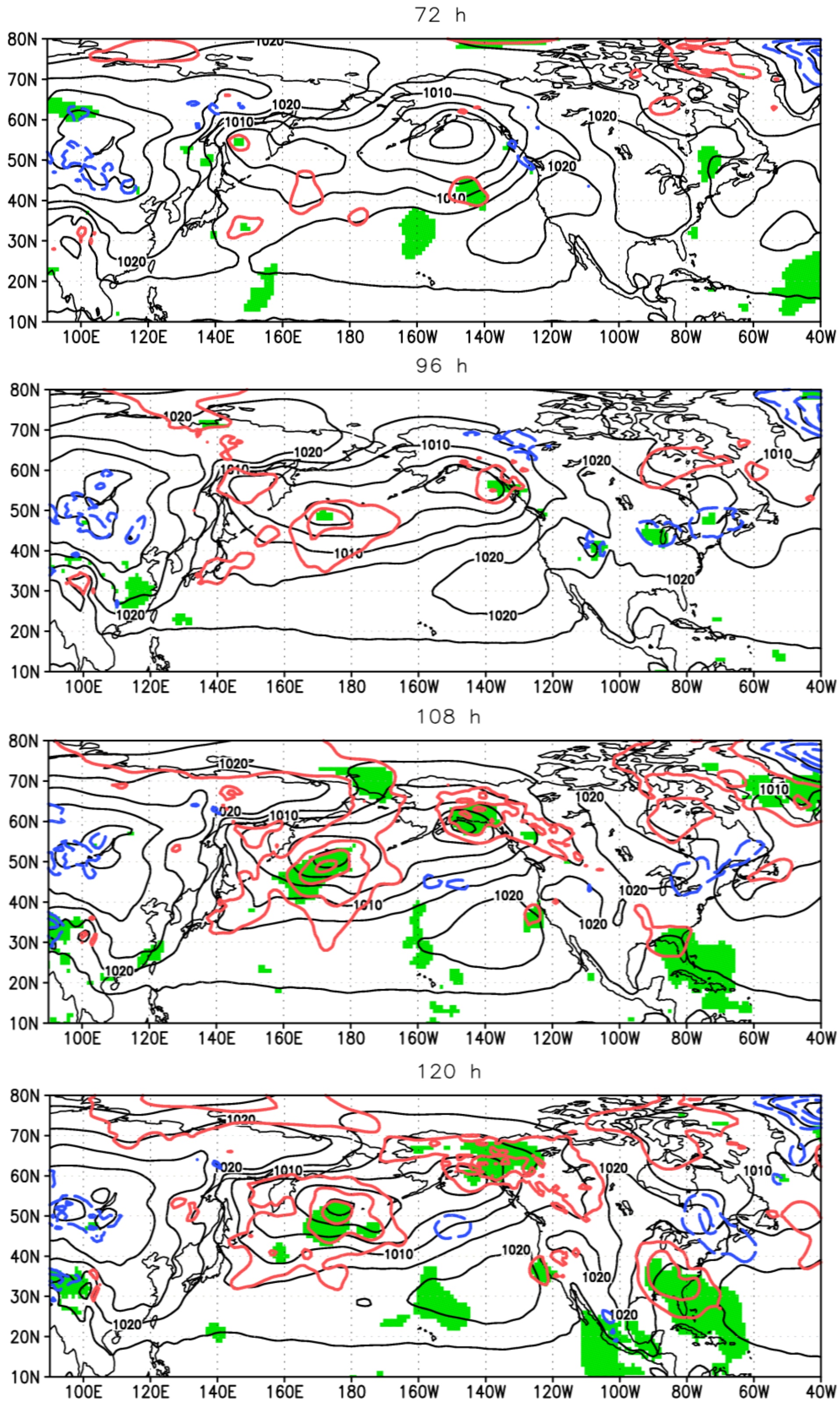
Figure A4.4: Day 3 OOE case: case 3d spaghetti plot for1015, 1005 and 995 hPa MSLP contour line (a-c, left) at different forecast step; and (d-f, right) ensemble mean forecast error (magenta dashed), pressure change (blue contours) between two time steps, spread change (gray shades and contour), and analyzed MSLP contour (solid black lines). Unit for all plots: [hPa]. The time step is 6 hr.

Both case 3c and case 3d are associated with a cyclone embedded in the remnants of a pre-existing cyclone. Figure A4.4 shows the synoptic development and error evolution of case 3d. The models under-predicted the cyclone at the onset stage of the storm, and finally forecasted a much weaker and more northeastward cyclone on day 3. Large positive forecast errors existed with the spread increase, again suggesting that nonlinearity is one of the reasons for the large forecast error.

To sum up, for days 1-4 (day 4 is not shown), the OOE cases are mainly associated with one or two of the following synoptic features: a) rapid cyclogenesis; b) cyclone development embedded in a pre-existing low; c) multiple cyclone centers over an elongated low pressure area; d) development of a secondary low; e) Rossby wave packet. One thing worth noting is that in most of the cases, error maxima tended to be over the areas with large spread increase, which may or may not be in the centers of the positive spread change. We hypothesize that the increasing nonlinearity and the models' misrepresentation of the nonlinear processes, such as latent heat release (Whitaker and Davis 1994; Zhang et al. 2002; Ahmadi-Givi et al. 2004; Moore et al. 2008), enhanced local baroclinicity from differential diabatic heating (Bosart 1981; Moore and Montgomery 2004; Boettcher and Wernli 2011), or strong upper-level forcing (Rogers and Bosart 1986; Smith et al. 1992) are the major reasons for the OOE cases. However, given the limited sample size here and the complexity of nonlinear flow representation, this point needs to be further investigated through well-designed numerical experiments.

Figures A4.5 and A4.6 show the composited MSLP and Z500 ensemble mean errors as well as the flow patterns for the 6-day OOE cases. From the initial time (0 h) to 24 h, there was a blocked upper-level ridge around the dateline in the mid-high latitudes (Figure A4.6). Another open ridge was just off the west coast of North America. Between the above two ridges, a surface low developed with the corresponding cut-off low at upper level off the southern Alaska coast. Meanwhile, there was another low west of the dateline at around 50°N corresponding to the cut-off low west of the blocked ridge. Note that although the high latitude flow was blocked, the flow was quite zonal south of 40°N from the west Pacific to the western U.S coast.




72 h
96 h
108 h
120 h
80N
70N
60N
50N
40N
30N
20N
10N
100E
120E
140E
160E
180
160W
140W
120W
100W
80W
60W
40W
1020
1010

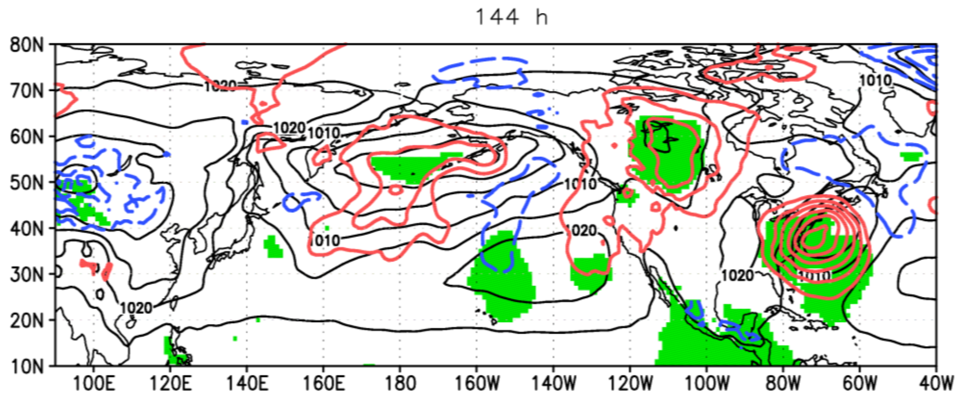
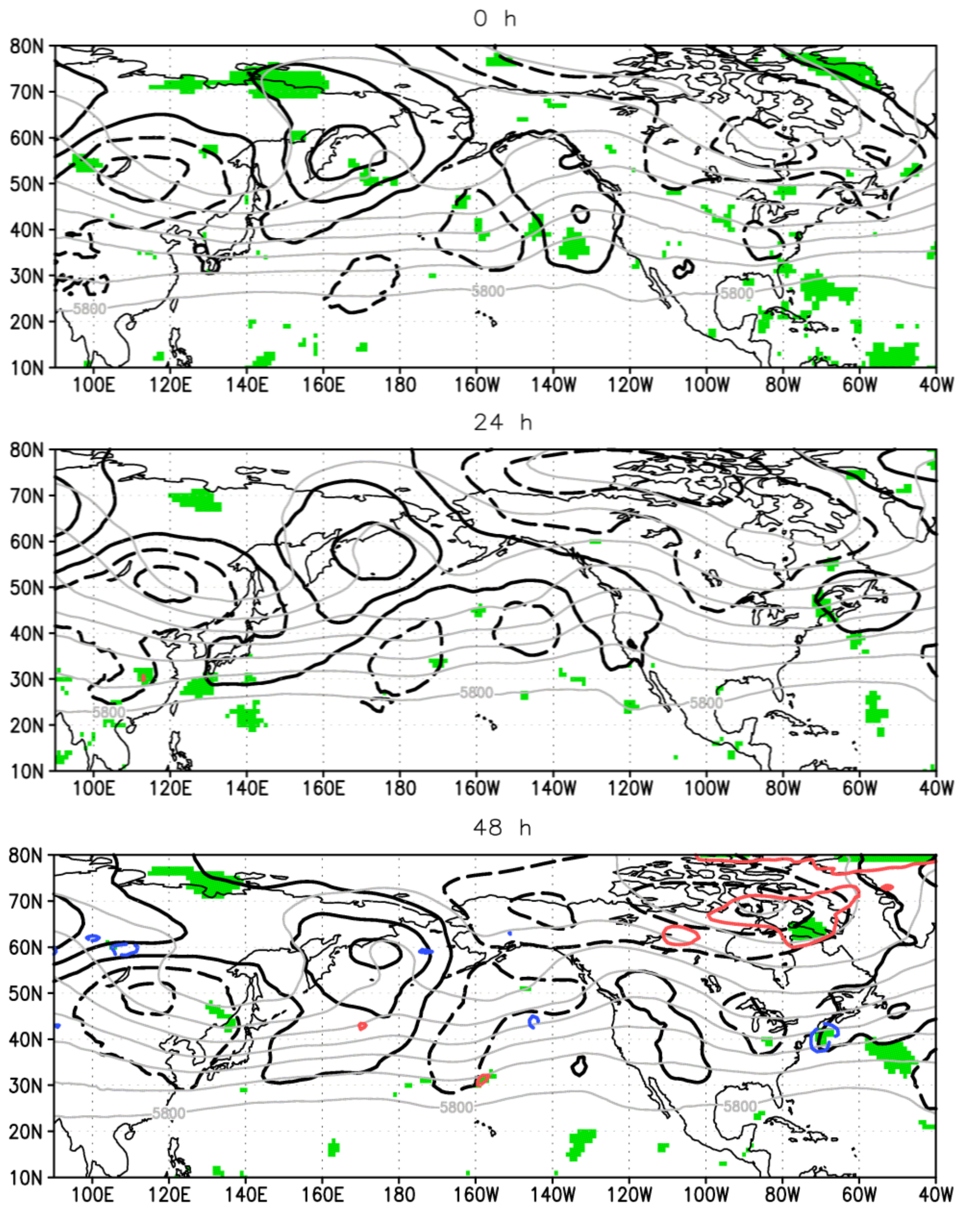
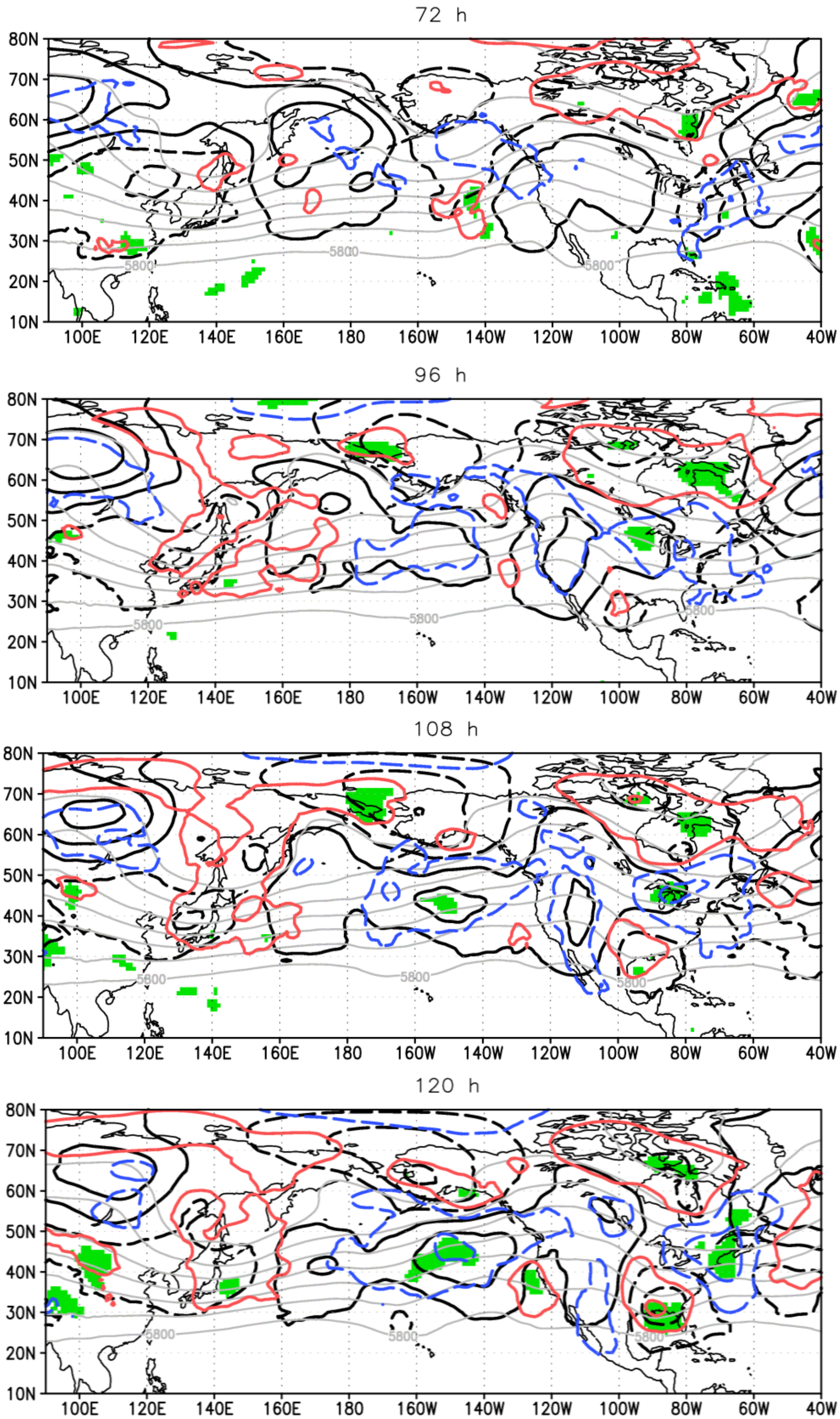


Figure A4.5: Composited MSLP forecast errors (red and blue contours) and MSLP analysis (black contour) for 6-day forecast OOE cases. The green shaded shows the significant values of errors different from the averaged errors of all cyclone cases. The MSLP contours have an interval of 5 hPa with 1010 and 1020 hPa lines labeled.

The amplitude of the blocked ridge over central Pacific decreased from 72 h (Figure A4.6). Forecast errors increased from 72 h at both the surface and upper levels associated with the southern side of the trough downstream of the blocked ridge and over the U.S. west coast ridge, indicating the models underestimated the short-wave trough and its downstream ridge over the southwest coast of the U.S. Between 96 h and 120 h, the flow pattern changed to be quite zonal over the central Pacific. The preexisting two surface lows over West and Northeast Pacific were underestimated by the ensemble. A wavy pattern of ridge-trough-ridge-trough developed at 96 h from 160°W to 100°W between 20° and 50°N, and was underestimated by the models at the upper level. Meanwhile, a weak surface low developed over the southern part of Mexico. From 108 h, the surface low area moved northeastward across the U.S. south coast, and intensified over the North Carolina coast at 132 h and moved to the east of the Mid-Atlantic coast. The surface low was underestimated significantly by over 12 hPa at the verification time. There was also some negative errors just ahead of the cyclone. At the upper level, the associated trough and its downstream ridge were underestimated by 100 m at the verification time, which can be directly linked with the weakening of the shortwave trough at 96 h over the southwest of U.S.;

The upstream errors can be further traced back to the errors in the wave group originating the south and downstream of the preexisting blocking ridge.




72 h
96 h
108 h
120 h
80N
70N
60N
50N
40N
30N
20N
10N
100E
120E
140E
160E
180
160W
140W
120W
100W
80W
60W
40W
5800

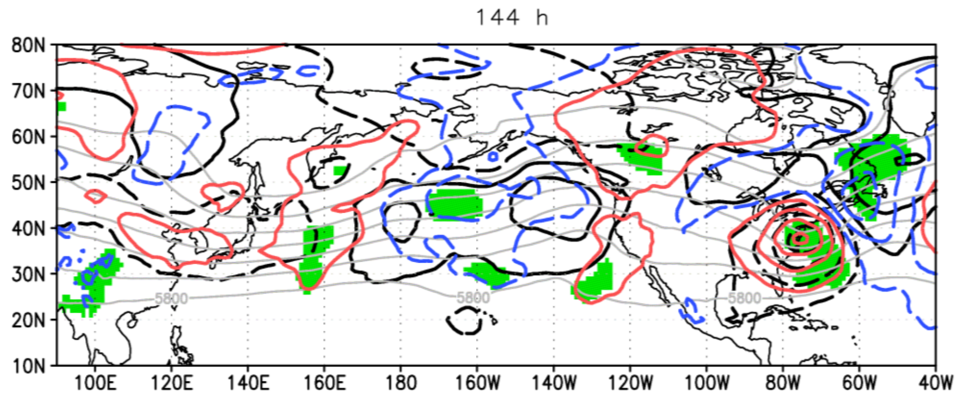


Figure A4.6: Composited Z500 forecast errors (red and blue contours), Z500 analysis (black contour) and anomaly to the climatology of Z500 for 6-day forecast OOE cases. The green shaded shows the significant values of errors different from the averaged errors of all cyclone cases. The climatology of Z500 is based on the mean of cool season (NDJFM) during 2007/8 and 2014/15. The errors starts from ±10m with an interval of 30 m; the Z500 anomaly starts from ±20 m with an interval of 40m; and the Z500 contours have an interval of 100 m with 5800m line labeled.

Therefore, the composite 6-day OOE cases seem to be related with a cyclone from the Gulf of Mexico and the upper-level southern stream short-wave trough. Both the surface and upper-level errors are associated with the under-estimation of the corresponding system. However, the MSLP errors are more local, while the upper-level errors are associated with a propagating wave group from East Pacific south of the 45°N. We speculate that the surface forecast errors could be partly due to the moisture conditions over the Gulf of Mexico, while the upper-level errors are associated with the downstream development of wave groups embedded on a blocking-to-zonal changing flow regime. Previous studies also suggested that the transition of large scale flow regimes can degrade the forecast skills over East Coast. For example, Frederiksen et al. (2004) found that the forecast errors tend to be larger when the blocks are developing or decaying and smaller for mature blocks. Langland et al. 2012 showed that the ACC for 120-h forecast of Z500 over NH is lower on average during the transition from negative Arctic Oscillation (AO) to positive phase in operational deterministic models. Archambault et al.

(2010) showed that the cool-season Northeast precipitation tends to be enhanced during negative-to-positive PNA and positive-to-negative NAO pattern. Therefore, our result for OOE fits the notion that the errors grow rapidly when dynamical development is rapidly changing (Frederiksen and Bell 1990; Frederiksen et al. 2004; Trevisan et al. 2001; Langland et al. 2002, 2012).

References

Ahmadi-Givi, F., Graig, G.C. and Plant, R.S., 2004. The dynamics of a midlatitude cyclone with very strong latent-heat release. *Quarterly Journal of the Royal Meteorological Society*, *130*(596), pp.295-323.

Ancell, B. and Hakim, G.J., 2007. Comparing adjoint-and ensemble-sensitivity analysis with applications to observation targeting. *Monthly Weather Review*, 135(12), pp.4117-4134.

Archambault, H.M., Keyser, D. and Bosart, L.F., 2010. Relationships between large-scale regime transitions and major cool-season precipitation events in the northeastern United States.*Monthly Weather Review*, *138*(9), pp.3454-3473.

Austin, J.M., 1941. Favorable conditions for cyclogenesis near the Atlantic coast. *Bulletin of the American Meteorological Society*, *22*, pp.270-272.

Barkmeijer, J., Bouttier, F. and Van Gijzen, M., 1998. Singular vectors and estimates of the analysis-error covariance metric. *Quarterly Journal of the Royal Meteorological Society*, *124*(549), pp.1695-1713.

Bednarczyk, C.N. and Ancell, B.C., 2015. Ensemble sensitivity analysis applied to a southern plains convective event. *Monthly Weather Review*,*143*(1), pp.230-249.

Bernhardt, J.E. and DeGaetano, A.T., 2012. Meteorological factors affecting the speed of movement and related impacts of extratropical cyclones along the U.S. east coast. *Natural hazards*, *61*(3), pp.1463-1472.

Bishop, C.H., Etherton, B.J. and Majumdar, S.J., 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly weather review*, *129*(3), pp.420-436.

Bjerknes, J., 1919. On the structure of moving cyclones. *Monthly Weather Review*, *47*(2), pp.95-99.

Bjerknes, J. and Solberg, H., 1922. Life cycle of cyclones and the polar front theory of atmospheric circulation. Grondahl.

Boettcher, M. and Wernli, H., 2011. Life cycle study of a diabatic Rossby wave as a precursor to rapid cyclogenesis in the North Atlantic-Dynamics and forecast performance. *Monthly Weather Review*, *139*(6), pp.1861-1878.

Booth, J.F., Rieder, H.E., Lee, D.E. and Kushnir, Y., 2015. The Paths of Extratropical Cyclones Associated with Wintertime High-Wind Events in the Northeastern United States*. *Journal of Applied Meteorology and Climatology*, *54*(9), pp.1871-1885.

Bosart, L. F., 1981: The Presidents' Day snowstorm of 18–19 February 1979: A subsynoptic-scale event. *Mon. Wea. Rev.*, **109,** 1542–1566.

Bougeault, P., Toth, Z., Bishop, C., Brown, B., Burridge, D., Chen, D.H., Ebert, B., Fuentes, M., Hamill, T.M., Mylne, K. and Nicolau, J., 2010. The THORPEX interactive grand global ensemble. *Bulletin of the American Meteorological Society*, *91*(8), p.1059.

Brill, K.F., Fracasso, A.R. and Bailey, C.M., 2015. Applying a divisive clustering algorithm to a large ensemble for medium-range forecasting at the weather prediction center. *Weather and Forecasting*, *30*(4), pp.873-891.

Buizza, R. and Palmer, T.N., 1995. The singular-vector structure of the atmospheric global circulation. *Journal of the Atmospheric Sciences*, *52*(9), pp.1434-1456.

Buizza, R., 1997. Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Monthly Weather Review*, *125*(1), pp.99-119.

Buizza, R. and Montani, A., 1999. Targeting observations using singular vectors. *Journal of the Atmospheric Sciences*, *56*(17), pp.2965-2985.

Buizza, R., Miller, M. and Palmer, T.N., 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *125*(560), pp.2887-2908.

Buizza, R., Houtekamer, P.L., Pellerin, G., Toth, Z., Zhu, Y. and Wei, M., 2005. A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Monthly Weather Review*,*133*(5), pp.1076-1097.

Cardinali, C. and Buizza, R., 2003. Forecast skill of targeted observations: A singular-vector-based diagnostic. *Journal of the atmospheric sciences*, *60*(16), pp.1927-1940.

Chang, E.K., 1993. Downstream development of baroclinic waves as inferred from regression analysis. *Journal of the atmospheric sciences*, *50*(13), pp.2038-2053.

Chang, E.K. and Orlanski, I., 1994. On energy flux and group velocity of waves in baroclinic flows. *Journal of the atmospheric sciences*, *51*(24), pp.3823-3828.

Chang, E.K. and Yu, D.B., 1999. Characteristics of wave packets in the upper troposphere. Part I: Northern Hemisphere winter. *Journal of the Atmospheric Sciences*, *56*(11), pp.1708-1728.

Chang, E.K., 1999. Characteristics of wave packets in the upper troposphere. Part II: Seasonal and hemispheric variations. *Journal of the atmospheric sciences*, *56*(11), pp.1729-1747.

Chang, E.K., 2005. The impact of wave packets propagating across Asia on Pacific cyclone development. *Monthly Weather Review*, *133*(7), pp.1998-2015.

Chang, E.K., Zheng, M. and Raeder, K., 2013. Medium-range ensemble sensitivity analysis of two extreme Pacific extratropical cyclones. *Monthly Weather Review*, *141*(1), pp.211-231.

Charles, M.E. and Colle, B.A., 2009a. Verification of extratropical cyclones within the NCEP operational models. Part I: Analysis errors and short-term NAM and GFS forecasts. *Weather and Forecasting*, *24*(5), pp.1173-1190.

Charles, M.E. and Colle, B.A., 2009b. Verification of extratropical cyclones within the NCEP operational models. Part II: The short-range ensemble forecast system. *Weather and Forecasting*, *24*(5), pp.1191-1214.

Colle, B.A. and Charles, M.E., 2011. Spatial distribution and evolution of extratropical cyclone errors over North America and its adjacent oceans in the NCEP global forecast system model.*Weather and Forecasting*, *26*(2), pp.129-149.

Colle, B.A., Booth, J.F. and Chang, E.K., 2015. A review of historical and future changes of extratropical cyclones and associated impacts along the U.S. east coast. *Current Climate Change Reports*, *1*(3), pp.125-143.

Davis, R.E. and Dolan, R., 1993. Nor'easters. *American Scientist*, *81*(5), pp.428-439.

Du, J., Mullen, S.L. and Sanders, F., 1997. Short-range ensemble forecasting of quantitative precipitation. *Monthly Weather Review*, *125*(10), pp.2427-2459.

Du, J. and Zhou, B., 2011. A dynamical performance-ranking method for predicting individual ensemble member performance and its application to ensemble averaging. *Monthly Weather Review*, *139*(10), pp.3284-3303.

Elmore, K.L., Schultz, D.M. and Baldwin, M.E., 2006. The behavior of synoptic-scale errors in the Eta model. *Monthly weather review*, *134*(11), pp.3355-3366.

Errico, R.M. and Vukicevic, T., 1992. Sensitivity Analysis Using an Adjoint of the PSU-NCAR mesoseale Model. *Monthly weather review*, *120*(8), pp.1644-1660.

Fanning, A., (2015): "Storm Summary Number 09 For Mid-Atlantic To Northeast Major Winter Storm". *Weather Prediction Center.* College Park, Maryland: National Centers for Environmental Prediction. [Available online at http://www.wpc.ncep.noaa.gov/winter_storm_summaries/2015/storm2/stormsum_9.html.]

Ferranti, L., and S. Corti. "New clustering products." *ECMWF Newsl* 127 (2011): 6-11.

Flegenheimer, Matt. "Leaders in New York and New Jersey Defend Shutdown for a Blizzard That Wasn't". *New York Times*. The New York Times Company. [Available online at http://www.nytimes.com/2015/01/28/nyregion/new-york-blizzard.html.]

Frederiksen, J.S. and Bell, R.C., 1990. North Atlantic blocking during January 1979: linear theory. *Quarterly Journal of the Royal Meteorological Society*, *116*(**496**), pp.1289-1313.

Frederiksen, J.S., Collier, M.A. and Watkins, A.B., 2004. Ensemble prediction of blocking regime transitions. *Tellus A*, *56*(5), pp.485-500.

Froude, L.S., Bengtsson, L. and Hodges, K.I., 2007. The predictability of extratropical storm tracks and the sensitivity of their prediction to the observing system. *Monthly weather review*, *135*(2), pp.315-333.

Froude, L.S., 2009. Regional differences in the prediction of extratropical cyclones by the ECMWF Ensemble Prediction System. *Monthly Weather Review*, *137*(3), pp.893-911.

Froude, L.S., 2010. TIGGE: Comparison of the prediction of Northern Hemisphere extratropical cyclones by different ensemble prediction systems. *Weather and Forecasting*, *25*(3), pp.819-836.

Froude, L.S., Bengtsson, L. and Hodges, K.I., 2013. Atmospheric predictability revisited. *Tellus A*,*65*.

Gaza, R.S. and Bosart, L.F., 1990. Trough-merger characteristics over North America. *Weather and Forecasting*, *5*(2), pp.314-331.

Garcies, L. and Homar, V., 2009. Ensemble sensitivities of the real atmosphere: application to Mediterranean intense cyclones. *Tellus A*, *61*(3), pp.394-406.

Gilbert, K.K., Myrick, D.T., Craven, J.P., Gagan, J.P., Hamill, T.M., Novak, D. R., Ruth, D.P., Settelmaier, J., Sieveking, J.E., and Veenhuis Jr, B., 2016: The National Blend of Global Models, Version One. 23rd Conf. on Probability and Statistics in the Atmospheric Sciences, New Orleans, LA, Amer. Meteor. Soc., 1.3.

Glatt, I. and Wirth, V., 2014. Identifying Rossby wave trains and quantifying their properties.*Quarterly Journal of the Royal Meteorological Society*, *140*(679), pp.384-396.

Gombos, D., Hoffman, R.N. and Hansen, J.A., 2012. Ensemble statistics for diagnosing dynamics: Tropical cyclone track forecast sensitivities revealed by ensemble regression. *Monthly Weather Review*, *140*(8), pp.2647-2669.

Gombos, D. and Hoffman, R.N., 2013. Ensemble-based exigent analysis. Part II: Using ensemble regression to estimate conditions antecedent to worst-case forecast damage scenarios. *Weather and Forecasting*, *28*(3), pp.557-569.

Grams, C.M., Wernli, H., Böttcher, M., Čampa, J., Corsmeier, U., Jones, S.C., Keller, J.H., Lenz, C.J. and Wiegand, L., 2011. The key role of diabatic processes in modifying the upper-tropospheric wave guide: a North Atlantic case-study. *Quarterly Journal of the Royal Meteorological Society*, *137*(661), pp.2174-2193.

Grazzini, F., 2007. Predictability of a large-scale flow conducive to extreme precipitation over the western Alps. *Meteorology and Atmospheric Physics*,*95*(3-4), pp.123-138.

Grazzini, F. and Vitart, F., 2015. Atmospheric predictability and Rossby wave packets. *Quarterly Journal of the Royal Meteorological Society*, *141*(692), pp.2793-2802.

Grimit, E.P. and Mass, C.F., 2007. Measuring the ensemble spread-error relationship with a probabilistic approach: Stochastic ensemble results. *Monthly weather review*, *135*(1), pp.203-221.

Grumm, R.H., Oravec, R.J. and Siebers, A.L., 1992. Systematic model forecast errors of surface cyclones in NMC's Nested-Grid Model, December 1988 through November 1990. *Weather and forecasting*, *7*(1), pp.65-87.

Hagedorn, R., Buizza, R., Hamill, T.M., Leutbecher, M. and Palmer, T.N., 2012. Comparing TIGGE multimodel forecasts with reforecast-calibrated ECMWF ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, *138*(668), pp.1814-1827.

Hakim, G.J., Bosart, L.F. and Keyser, D., 1995. The Ohio Valley wave-merger cyclogenesis event of 25-26 January 1978. Part I: Multiscale case study. *Monthly Weather Review*, *123*(9), pp.2663-2692.

Hakim, G.J., Keyser, D. and Bosart, L.F., 1996. The Ohio Valley wave-merger cyclogenesis event of 25-26 January 1978. Part II: Diagnosis using quasigeostrophic potential vorticity inversion. *Monthly weather review*, *124*(10), pp.2176-2205.

Hakim, G.J., 2003. Developing wave packets in the North Pacific storm track. *Monthly Weather Review*, *131*(11), pp.2824-2837.

Hakim, G.J., 2005. Vertical structure of midlatitude analysis and forecast errors. *Monthly Weather Review*, *133*(3), pp.567-578.

Hakim, G.J. and Torn, R.D., 2008. Ensemble synoptic analysis. In *Synoptic—Dynamic Meteorology and Weather Analysis and Forecasting* (pp. 147-161). American Meteorological Society.

Hanley, K.E., Kirshbaum, D.J., Roberts, N.M. and Leoncini, G., 2013. Sensitivities of a squall line over central Europe in a convective-scale ensemble. *Monthly Weather Review*, *141*(1), pp.112-133.

Hannachi, A., 2004. A primer for EOF analysis of climate data. *Department of Meteorology, University of Reading*, pp.1-33.

Hannachi, A., Jolliffe, I.T. and Stephenson, D.B., 2007. Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, *27*(9), pp.1119-1152.

Harr, P.A., Anwender, D. and Jones, S.C., 2008. Predictability associated with the downstream impacts of the extratropical transition of tropical cyclones: Methodology and a case study of Typhoon Nabi (2005). *Monthly Weather Review*, *136*(9), pp.3205-3225.

Hawblitzel, D.P., Zhang, F., Meng, Z. and Davis, C.A., 2007. Probabilistic evaluation of the dynamics and predictability of the mesoscale convective vortex of 10-13 June 2003. *Monthly weather review*, *135*(4), pp.1544-1563.

Hewson, L. Magnusson, O. Breivik, F. Prates, I. Tsonevsky and J. W. de Vries, 2014: Windstorms in northwest Europe in late 2013. *ECMWF Newsletter*, No. 139, ECMWF, Reading, United Kingdom, pp.22–28.

Hirsch, M.E., DeGaetano, A.T. and Colucci, S.J., 2001. An East Coast winter storm climatology. *Journal of Climate*, *14*(5), pp.882-899.

Hodges, K. I., 1994. A general-method for tracking analysis and its application to meteorological data. *Monthly Weather Review, 122* (**11**). pp. 2573-2586.

Hodges, K. I. (1995) Feature tracking on the unit-sphere. *Monthly Weather Review, 123* (**12**). pp. 3458-3465.

Hodges, K. I. (1999) Adaptive constraints for feature tracking. *Monthly Weather Review, 127* (**6**). pp. 1362-1373.

Houtekamer, P.L. and Mitchell, H.L., 1998. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, *126*(3), pp.796-811.

Houtekamer, P.L., Mitchell, H.L. and Deng, X., 2009. Model error representation in an operational ensemble Kalman filter. *Monthly Weather Review*, *137*(7), pp.2126-2143.

Houtekamer, P.L., He, B. and Mitchell, H.L., 2014. Parallel implementation of an ensemble Kalman filter. *Monthly Weather Review*, *142*(3), pp.1163-1182.

Hovmöller, E., 1949. The Trough-and-Ridge diagram. *Tellus*, *1*(2), pp.62-66.

Johnson, A., Wang, X., Kong, F. and Xue, M., 2011. Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous weather Testbed 2009 spring experiment. Part I: Development of the object-oriented cluster analysis method for precipitation fields. *Monthly weather review*,*139*(12), pp.3673-3693.

Judd, K. and Smith L.A., 2001. Indistinguishable states: I. Perfect model scenario. Physica D: Nonlinear Phenomena, Volume 151, Issues 2-4, pp.125-141, ISSN 0167-2789.Kalnay, E., 2003. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.

Keller, J.H., Jones, S.C., Evans, J.L. and Harr, P.A., 2011. Characteristics of the TIGGE multimodel ensemble prediction system in representing forecast variability associated with extratropical transition. *Geophysical Research Letters*, *38*(12).

Keller, J.H., 2012. *Diagnosing the downstream impact of extratropical transition using multimodel operational ensemble prediction systems* (Vol. 57). KIT Scientific Publishing.

Kocin, P.J. and Uccellini, L.W., 1990. Snowstorms along the northeastern coast of the United States, 1955 to 1985 (Vol. 22). *American Meteorological Society.*

Kocin, P. J., and L. W. Uccellini, 2004: Northeast Snowstorms (Volume I: Overview, Volume II: The Cases). Meteor. Monogr., No. 54, *American Meteorological Society*, 818 pp. Langland, R.H., Elsberry, R.L. and Errico, R.M., 1995. Evaluation of physical processes in an idealized extratropical cyclone using adjoint sensitivity. *Quarterly Journal of the Royal Meteorological Society*, *121*(526), pp.1349-1386.

Langland, R.H., Shapiro, M.A. and Gelaro, R., 2002. Initial condition sensitivity and error growth in forecasts of the 25 January 2000 East Coast snowstorm. *Monthly Weather Review*, *130*(4), pp.957-974.

Langland, Rolf H, 2005. Issues in targeted observing. *Quarterly Journal of the Royal Meteorological Society*, *131*(**613**), pp.3409-3425.

Langland, R.H. and Maue, R.N., 2012. Recent Northern Hemisphere mid-latitude medium-range deterministic forecast skill. *Tellus A*, *64*.

Lee, S. and Held, I.M., 1993. Baroclinic wave packets in models and observations. *Journal of the atmospheric sciences*, *50*(10), pp.1413-1428.

Leith, C.E., 1974. Theoretical skill of Monte Carlo forecasts. *Monthly Weather Review*, *102*(6), pp.409-418.

Lorenz, E.N., 1956. Empirical orthogonal functions and statistical weather prediction.

Lorenz, E.N., 1963. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, *20*(2), pp.130-141.

Lorenz, E.N., 1969. Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric sciences*, *26*(4), pp.636-646.

Lorenz, E.N., 1982. Atmospheric predictability experiments with a large numerical model. *Tellus*, *34*(6), pp.505-513.

Lorenz, E.N., 1996, September. Predictability: A problem partly solved. In *Proc. Seminar on predictability* , vol. 1, no. 1.

Lupo, A.R., PJ Smith, and P. Zwack, 1992: A diagnosis of the explosive development of two extratropical cyclones. *Mon. Wea. Rev*, *120*, pp.1490-1523.

Magnusson, L., Leutbecher, M. and Källén, E., 2008. Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system. *Monthly Weather Review*, *136*(11), pp.4092-4104.

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,** pp.2803–2820.

Majumdar, S.J., Bishop, C.H., Etherton, B.J. and Toth, Z., 2002. Adaptive sampling with the ensemble transform Kalman filter. Part II: Field program implementation. *Monthly Weather Review*, *130*(5), pp.1356-1369.

Majumdar, S.J., Sellwood, K.J., Hodyss, D., Toth, Z. and Song, Y., 2010. Characteristics of target areas selected by the Ensemble Transform Kalman Filter for medium-range forecasts of high-impact winter weather. *Monthly Weather Review*, *138*(7), pp.2803-2824.

Majumdar, S.J., Chen, S.G. and Wu, C.C., 2011. Characteristics of ensemble transform Kalman filter adaptive sampling guidance for tropical cyclones. *Quarterly Journal of the Royal Meteorological Society*, *137*(655), pp.503-520.

Mather, J.R., Adams III, H. and Yoshioka, G.A., 1964. Coastal storms of the eastern United States. *Journal of Applied Meteorology, 3*(6), pp.693-706.

Matsueda, M., Kyouda, M., Toth, Z., Tanaka, H.L. and Tsuyuki, T., 2011. Predictability of an atmospheric blocking event that occurred on 15 December 2005. *Monthly weather review*, *139*(8), pp.2455-2470.

Matsueda, M. and Nakazawa, T., 2015. Early warning products for severe weather events derived from operational medium-range ensemble forecasts.*Meteorological Applications*, *22*(2), pp.213-222.

McMurdie, L. and Mass, C., 2004. Major numerical forecast failures over the northeast Pacific. *Weather and forecasting*, *19*(2), pp.338-356.

McMurdie, L.A. and Ancell, B., 2014. Predictability characteristics of landfalling cyclones along the North American west coast. *Monthly Weather Review*, *142*(1), pp.301-319.

Miller, J.E., 1946. Cyclogenesis in the Atlantic coastal region of the United States. *Journal of Meteorology*, *3*(2), pp.31-44.

Molteni, F. and Palmer, T.N., 1993. Predictability and finite-time instability of the northern winter circulation. *Quarterly Journal of the Royal Meteorological Society*, *119*(510), pp.269-298.

Molteni, F., Buizza, R., Palmer, T.N. and Petroliagis, T., 1996. The ECMWF ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, *122*(529), pp.73-119.

Moore, R. W., and M. T. Montgomery, 2004: Reexamining the dynamics of short-scale, diabatic Rossby waves and their role in midlatitude moist cyclogenesis.

Murphy, J.M., 1988. The impact of ensemble forecasts on predictability.*Quarterly Journal of the Royal Meteorological Society*, *114*(480), pp.463-493.

Novak, D.R., Bright, D.R. and Brennan, M.J., 2008. Operational forecaster uncertainty needs and future roles. *Weather and Forecasting*, *23*(6), pp.1069-1084.

Novak, D.R. and Colle, B.A., 2012. Diagnosing snowband predictability using a multimodel ensemble system. *Weather and Forecasting*, *27*(3), pp.565-585.

Orlanski, I. and Katzfey, J., 1991. The life cycle of a cyclone wave in the Southern Hemisphere. Part I: Eddy energy budget. *Journal of the atmospheric sciences*, *48*(17), pp.1972-1998.

Orlanski, I. and Chang, E.K., 1993. Ageostrophic geopotential fluxes in downstream and upstream development of baroclinic waves. *Journal of the atmospheric sciences*, *50*(2), pp.212-225.

Orlanski, I. and Sheldon, J.P., 1995. Stages in the energetics of baroclinic systems. *Tellus A*, *47*(5), pp.605-628.

Orrell, D., Smith, L., Barkmeijer, J. and Palmer, T.N., 2001. Model error in weather forecasting. *Nonlinear processes in geophysics*, *8*(6), pp.357-371.

Palmer, T.N., Brankovic, C., Molteni, F., Tibaldi, S., Ferranti, L., Hollingsworth, A., Cubasch, U. and Klinker, E., 1990. The European Centre for Medium-range Weather Forecasts (ECMWF) program on extended-range prediction. *Bulletin of the American Meteorological Society*, *71*(9), pp.1317-1330.Palmer, T.N., Gelaro, R., Barkmeijer, J. and Buizza, R., 1998. Singular vectors, metrics, and adaptive observations. *Journal of the Atmospheric Sciences*, *55*(4), pp.633-653.

Park, Y.Y., Buizza, R. and Leutbecher, M., 2008. TIGGE: Preliminary results on comparing and combining ensembles. *Quarterly Journal of the Royal Meteorological Society*, 134(637), pp.2029-2050.

Petterssen, S., 1956. Motion and motion systems. McGraw-Hill.

Petersen, G.N., Majumdar, S.J. and Thorpe, A.J., 2007. The properties of sensitive area predictions based on the ensemble transform Kalman filter (ETKF). *Quarterly Journal of the Royal Meteorological Society*, *133*(**624**), pp.697-710.

Rand, W.M., 1971. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical association*, *66*(336), pp.846-850.

Reitan, C.H., 1974. Frequencies of cyclones and cyclogenesis for North America, 1951-1970. *Monthly Weather Review*, *102*(12), pp.861-868.

Rodwell, M.J., Magnusson, L., Bauer, P., Bechtold, P., Bonavita, M., Cardinali, C., Diamantakis, M., Earnshaw, P., Garcia-Mendez, A., Isaksen, L. and Källén, E., 2013. Characteristics of occasional poor medium-range weather forecasts for Europe. *Bulletin of the American Meteorological Society*, *94*(9), pp.1393-1405.

Rogers E, Bosart LF. 1986. An investigation of explosively deepening oceanic cyclones. *Monthly Weather Review* **114**: 702–718, DOI: 10.1175/1520-0493.

Rolfson, D.M. and Smith, P.J., 1996. A composite diagnosis of synoptic-scale extratropical cyclone development over the United States.

Salmon, E.M. and Smith, P.J., 1980. Focus on Forecasting: A synoptic analysis of the 25-26 January 1978 blizzard cyclone in the central United States. *Bulletin of the American Meteorological Society*, *61*(5), pp.453-453.

Sanders, F. and Gyakum, J.R., 1980. Synoptic-dynamic climatology of the “bomb”. *Monthly Weather Review*, *108*(10), pp.1589-1606.

Sanders, F., 1986. Explosive cyclogenesis in the west-central North Atlantic Ocean, 1981-84. Part I: Composite structure and mean behavior. *Monthly weather review*, *114*(10), pp.1781-1794.

Schumacher, R.S., 2011. Ensemble-Based Analysis of Factors Leading to the Development of a Multiday Warm-Season Heavy Rain Event*. *Monthly Weather Review*, *139*(9), pp.3016-3035.

Sellwood, K.J., Majumdar, S.J., Mapes, B.E. and Szunyogh, I., 2008. Predicting the influence of observations on medium-range forecasts of atmospheric flow. *Quarterly Journal of the Royal Meteorological Society*, *134*(637), pp.2011-2027.

Shapiro, M.A. and Keyser, D.A., 1990. Fronts, jet streams, and the tropopause. US Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, Wave Propagation Laboratory, pp.177-186.

Silberberg, S.R. and Bosart, L.F., 1982. An analysis of systematic cyclone errors in the NMC LFM-II model during the 1978-79 cool season. *Monthly Weather Review*, *110*(4), pp.254-271.

Sippel, J.A. and Zhang, F., 2010. Factors affecting the predictability of Hurricane Humberto (2007). *Journal of the Atmospheric Sciences*, *67*(6), pp.1759-1778.

Smith, P.J., Zwack, P. and Lupo, A.R., 1992. A diagnosis of the explosive development of two extratropical cyclones.

Smith, L.A., 2001. Disentangling uncertainty and error: On the predictability of nonlinear systems. In *Nonlinear dynamics and statistics* (pp. 31-64). Birkhäuser Boston.

Souders, M.B., Colle, B.A. and Chang, E.K., 2014a. A description and evaluation of an automated approach for feature-based tracking of Rossby wave packets. *Monthly Weather Review*, *142*(10), pp.3505-3527.

Souders, M.B., Colle, B.A. and Chang, E.K., 2014b. The climatology and characteristics of Rossby wave packets using a feature-based tracking technique. *Monthly Weather Review*, *142*(10), pp.3528-3548.

Stensrud, D.J., Brooks, H.E., Du, J., Tracton, M.S. and Rogers, E., 1999. Using ensembles for short-range forecasting. *Monthly Weather Review*, *127*(4), pp.433-446.

Strahl, J.L. and Smith, P.J., 2001. A diagnostic study of an explosively developing extratropical cyclone and an associated 500-hPa trough merger. *Monthly weather review*, *129*(9), pp.2310-2328.

Sutcliffe, R.C., 1947. A contribution to the problem of development.*Quarterly Journal of the Royal Meteorological Society*, *73*(317-318), pp.370-383.

Swanson, K. and Pierrehumbert, R.T., 1994. Nonlinear Wave Packet Evolution on a Baroclinicaily Unstable Jet. *Journal of the atmospheric sciences*, *51*(3), pp.384-396.

Swanson, K.L. and Roebber, P.J., 2008. The impact of analysis error on medium-range weather forecasts. *Monthly Weather Review*, *136*(9), pp.3425-3431.

Swinbank, R., Kyouda, M., Buchanan, P., Froude, L., Hamill, T.M., Hewson, T.D., Keller, J.H., Matsueda, M., Methven, J., Pappenberger, F. and Scheuerer, M., 2016. The TIGGE project and its achievements. *Bulletin of the American Meteorological Society*, *97*(1), pp.49-67.

Szunyogh, I., Toth, Z., Morss, R.E., Majumdar, S.J., Etherton, B.J. and Bishop, C.H., 2000. The effect of targeted dropsonde observations during the 1999 Winter Storm Reconnaissance Program. *Monthly weather review*,*128*(10), pp.3520-3537.

Szunyogh, I., Toth, Z., Zimin, A.V., Majumdar, S.J. and Persson, A., 2002. Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Monthly weather review*, *130*(5), pp.1144-1165.

Talagrand, O. and Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory.*Quarterly Journal of the Royal Meteorological Society*, *113*(478), pp.1311-1328.

M. A. Tew, A. Stern, A. Horvitz, K. Gilbert, and D. T. Myrick, 2016: National Blend of Models: Transformational Forecast Change by the National Weather Service, 2016, 4th Symp. on Building a Weather-Ready Nation, New Orleans, LA, Amer. Meteor. Soc., 6.2.

Torn, R.D. and Hakim, G.J., 2008. Ensemble-based sensitivity analysis. *Monthly Weather Review*, 136(2), pp.663-677.

Torn, R.D. and Hakim, G.J., 2009. Initial condition sensitivity of western Pacific extratropical transitions determined using ensemble-based sensitivity analysis. *Monthly Weather Review*, *137*(10), pp.3388-3406.

Torn, R.D., 2010. Ensemble-based sensitivity analysis applied to African easterly waves. *Weather and Forecasting*, *25*(1), pp.61-78.

Torn, R.D. and Romine, G.S., 2015. Sensitivity of central Oklahoma convection forecasts to upstream potential vorticity anomalies during two strongly forced cases during MPEX. *Monthly Weather Review*, *143*(10), pp.4064-4087.

Toth, Z. and Kalnay, E., 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, *74*(12), pp.2317-2330.

Tracton, M.S. and Kalnay, E., 1993. Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Weather and Forecasting*,*8*(3), pp.379-398.

Trevisan, A., Pancotti, F. and Molteni, F., 2001. Ensemble prediction in a model with flow regimes. *Quarterly Journal of the Royal Meteorological Society*, *127*(572), pp.343-358.

Uccellini, L.W., Kocin, P.J., Petersen, R.A., Wash, C.H. and Brill, K.F., 1984. The Presidents' Day cyclone of 18-19 February 1979: Synoptic overview and analysis of the subtropical jet streak influencing the pre-cyclogenetic period. *Monthly weather review*, *112*(1), pp.31-55.

Uccellini, L.W. and Kocin, P.J., 1987. The interaction of jet streak circulations during heavy snow events along the east coast of the United States. *Weather and Forecasting*, *2*(4), pp.289-308.

Uccellini, L.W., Kocin, P.J., Schneider, R.S., Stokols, P.M. and Dorr, R.A., 1995. Forecasting the 12–14 March 1993 superstorm. *Bulletin of the American Meteorological Society*, *76*(2), pp.183-199.

Von Storch, H., 1999. Misuses of statistical analysis in climate research. In *Analysis of Climate Variability* (pp. 11-26). Springer Berlin Heidelberg.

Wang, C-C., and J. C. Rogers, 2001: A composite study of explosive cyclogenesis in different sectors of the North Atlantic. Part I: Cyclone structure and evolution. Mon. Wea. Rev, **129,** 1481–1499.

Welch, G. and Bishop, G., 2001. An introduction to the kalman filter.*Proceedings of the Siggraph Course, Los Angeles*.

Whitaker, J.S. and Davis, C.A., 1994. Cyclogenesis in a saturated environment. *Journal of the atmospheric sciences*, *51*(6), pp.889-908.

Whitaker, J.S. and Loughe, A.F., 1998. The relationship between ensemble spread and ensemble mean skill. *Monthly weather review*, *126*(12), pp.3292-3302.

Wilks, D.S., 2011. *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.

Winkler, R.L., 2015. The importance of communicating uncertainties in forecasts: Overestimating the risks from winter storm Juno. *Risk Analysis*,35(3), pp.349-353.

Winters, A.C. and Martin, J.E., 2016. Synoptic and mesoscale processes supporting vertical superposition of the polar and subtropical jets in two contrasting cases. *Quarterly Journal of the Royal Meteorological Society*. 142(695), pp. 1133-1149

Yeung, K.Y. and Ruzzo, W.L., 2001. Details of the adjusted Rand index and clustering algorithms, supplement to the paper "An empirical study on principal component analysis for clustering gene expression data". *Bioinformatics*, *17*(9), pp.763-774.

Zack, J., Natenberg, E., Young, S., Van Knowe, G., Waight, K., Manobianco, J. and Kamath, C., 2010a. Application of ensemble sensitivity analysis to observation targeting for short-term wind speed forecasting in the Tehachapi region winter season. *LLNLTR-460956, October*.

Zack, J., Natenberg, E., Young, S., Van Knowe, G., Waight, K., Manobianco, J. and Kamath, C., 2010b. Application of ensemble sensitivity analysis to observation targeting for short term wind speed forecasting in the Washington-Oregon Region. Technical Report LLNL-TR-458086, 65 pp., Lawrence Livermore National Laboratory, Livermore, CA.

Zhang, F., Snyder, C. and Rotunno, R., 2002. Mesoscale predictability of the "surprise" snowstorm of 24-25 January 2000. *Monthly Weather Review*, *130*(6), pp.1617-1632.

Zhang, F., 2005. Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Monthly weather review*, *133*(10), pp.2876-2893.

Zheng, M., Chang, E.K. and Colle, B.A., 2013. Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Weather and Forecasting*, *28*(5), pp.1133-1156.

Zheng, M., Chang, E.K., Colle, B.A., Luo, Y. and Zhu, Y., 2016. Applying fuzzy clustering to a multi-Model ensemble for U.S. East Coast winter storms: scenario identification and forecast verification. *Weather and Forecasting*, submitted.

Zimin, A.V., Szunyogh, I., Patil, D.J., Hunt, B.R. and Ott, E., 2003. Extracting envelopes of Rossby wave packets. *Monthly weather review*,*131*(5), pp.1011-1017.

Zimin, A.V., Szunyogh, I., Hunt, B.R. and Ott, E., 2006. Extracting envelopes of nonzonally propagating Rossby wave packets. *Monthly weather review*, *134*(4), pp.1329-1333.Zishka, K.M. and Smith, P.J., 1980. The climatology of cyclones and anticyclones over North America and surrounding ocean environs for January and July, 1950-77. *Monthly Weather Review*, *108*(4), pp.387-401.Zou, X., Kuo, Y.H. and Low-Nam, S., 1998. Medium-range prediction of an extratropical oceanic cyclone: Impact of initial state. *Monthly weather review*, *126*(11), pp.2737-2763.