

**APPLYING FORECAST TRACK AND INTENSITY  
DIAGNOSTICS TO HIGH-IMPACT NORTHEAST WINTER  
STORMS**

by

Tomer Burg

A Thesis

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

College of Arts and Sciences

Department of Atmospheric and Environmental Sciences

2019

## ABSTRACT

A conventional forecasting notion is that as lead time decreases, numerical weather prediction models exhibit a leftward (i.e., west) trend in the forecast position of low-pressure systems along the East Coast of the U.S. This left trend, which may turn seemingly weak ocean cyclones into high-impact weather events for the Northeast U.S., is attributed to various potential causes, such as variability in upstream shortwave troughs, or the representation of latent heat release in the NWP models downstream of the trough associated with the incipient cyclone. This study seeks to address whether this rule of thumb holds any significant merit, and to examine a long-term climatology of Northeast U.S. cold season cyclones from a forecast skill and error perspective.

A climatology of ensemble forecasts of high-impact Northeast winter storms initialized from 0 to 5-day lead times was constructed using the Global Ensemble Forecast System (GEFS) Reforecast version 2, from 1 November 1985 through 31 December 2015. Cases included in this climatology were those identified within 750 km of the  $40^{\circ}\text{N}/70^{\circ}\text{W}$  benchmark from an ERA-I based cyclone track dataset created by Sprenger et al. (2017), but also could be tracked with CFSR data using a 925-hPa area-averaged vorticity maxima and height minima-based tracking algorithm. The verification of the ensemble forecasts at 0 through 5-day lead times are computed against the CFSR verification track, and the analysis quantifies the GEFS climatological forecast track and intensity errors and biases.

A total of 517 cyclones were identified, with the CFSR verification tracks consistent with prior climatologies of East Coast cyclones. When applying the tracking algorithm to the ensemble forecasts, an underdispersive bias was found in the GEFS for both position and

intensity forecasts, in addition to the GEFS exhibiting predominantly along-track position variability. The climatology results did not find a systematic right of track bias; for most of the reforecast lead times, cyclones tended to exhibit a slightly left of track bias on average, while the 12–66-hour lead time range shows a prominent slow bias for most cases. The forecasts exhibit a negative correlation between across-track bias and intensity bias, peaking at day 3 forecast lead time, which motivated assessing synoptic composites of left vs. right of track bias and weak vs. strong intensity bias. More amplified flow over the United States was found in the right of track and weak bias composites compared to the left of track and strong bias composites, in addition to a stronger downstream polar jet streak and upstream subtropical jet streak relative to the mean cyclone location. Assessing the synoptic composite differences of small vs. large across-track variability found a tendency for higher mean sea level pressure and upper level heights across western North America in the small across-track variability composite, which along with a significantly more negative temperature anomaly in the eastern two thirds of the U.S. suggests a higher likelihood of cold air outbreaks than large across-track variability.

## ACKNOWLEDGEMENTS

I would like to thank Dr. Andrea Lang, Dr. Ryan Torn and Dr. Kristen Corbosiero for their advisement on this thesis. Their invaluable input, as well as the different perspectives they offered from their respective areas of expertise, exposed me to approaches that I had previously not thought of and expanded my skill set. I would also like to thank my focal points from the National Weather Service, Neil Stuart, Joe Dellicarpini and Justin Arnott. Neil Stuart provided useful insight into the methodology and results from the perspective of an operational forecaster, which was useful in identifying ways to make the results of this project more applicable for the research to operations transition.

Appreciation is greatly extended to the faculty, staff and students at the Department of Atmospheric and Environmental Sciences at the University at Albany who helped to make my graduate school experience thus far a memorable one. The faculty provided top quality education throughout both my undergraduate and graduate years at Albany, and the staff has always helped with great enthusiasm, especially Kevin Tyle who helped out tremendously with technical challenges. I owe deep appreciation to some of the graduate students who helped me with various aspects, including Nick Schiraldi for helping me learn python; Philippe Papin, Alicia Bentley, and Hannah Attard for providing useful code and data that was instrumental in completing this thesis; and fellow graduate students in the department, especially Matthew Campbell, Massey Bartolini and Brendan Wallace for being supportive during the process of writing this thesis.

I would also like to thank my family, and my pet rabbit Bella, for their encouragement and support as I worked on this thesis. I would especially like to acknowledge my uncle Arnon

and my grandparents Tami and Moshe who were always there to encourage me during the writing of this thesis, Uri who passed away during my undergraduate studies, and Arza who unfortunately passed days before I began graduate school, as I would not have made it this far in my academic pursuits had it not been for their encouragement and support.

# CONTENTS

ABSTRACT . . . . .	ii
ACKNOWLEDGEMENTS . . . . .	iv
1. Introduction . . . . .	1
1.1 Motivation . . . . .	1
1.2 Literature Review . . . . .	2
1.2.1 East Coast Cyclones and Snowstorms: Overview, Climatology and Dynamics . . . . .	2
1.2.2 Numerical Modeling of Northeast ECs: Deterministic and Ensemble Approaches . . . . .	5
1.2.3 Objective Cyclone Identification and Tracking . . . . .	8
1.3 Research Goals and Hypotheses . . . . .	12
2. Data and Methodology . . . . .	15
2.1 Data . . . . .	15
2.2 Case Selection . . . . .	16
2.2.1 Case identification in Reanalysis Data . . . . .	16
2.2.2 Case identification in Ensemble Forecasts . . . . .	17
2.3 Cyclone Tracking Algorithm . . . . .	18
2.3.1 Selection of Variables and Pressure Level . . . . .	18
2.3.2 Cyclone Identification and Tracking . . . . .	20
2.3.3 Matching Ensembles Tracks to Cyclones . . . . .	22
2.4 Ensemble Forecast Diagnostics and Verification . . . . .	23
2.5 Composite Analyses . . . . .	24
3. Climatology . . . . .	31
3.1 Climatology of All Candidate Cyclones . . . . .	31
3.2 Ensemble Forecast Skill and Calibration . . . . .	32
3.3 Climatology of Cyclones Partitioned by Forecast Skill . . . . .	34
3.4 Climatology Discussion . . . . .	36

4. Composites . . . . .	49
4.1 Ensemble Bias Composites . . . . .	50
4.1.1 Left vs. Right of Track Bias . . . . .	50
4.1.2 Weak vs. Strong Bias . . . . .	51
4.2 Ensemble Variability Composites . . . . .	53
4.2.1 Across Track Ensemble Spread . . . . .	53
4.2.2 Along Track Ensemble Spread . . . . .	56
4.3 Composites Discussion . . . . .	57
5. Conclusions and Future Work . . . . .	66
5.1 Conclusions . . . . .	66
5.2 Future Work . . . . .	68

# 1. Introduction

## 1.1 Motivation

Cold season extratropical cyclones (ECs) are a common phenomenon in the United States (U.S.) East Coast. These storms can produce wide-ranging socioeconomic impacts including heavy snow, damaging wind, heavy rain and coastal flooding. There are many well-documented cases of high impact snowstorms in the Northeast U.S., including intense ECs that produced upwards of 50 cm (20 in) of snow, as in the Superstorm of 1993 (Kocin et al. 1995) and the Presidents' Day 1979 cyclone (Bosart 1981). Other cases associated with relatively weak ECs in terms of minimum sea level pressure but nonetheless produced heavy snowfall have also occurred, including the President's Day storm of 2003 (Kocin and Uccellini 2004a). Snowfall rate, time of day and preparation are contributing factors that allow even climatologically minor snow accumulation events to cause major societal disruptions. For example, the 14–16 November 2018 snowstorm in New York City produced heavy snow rates upwards of 4 cm hr<sup>-1</sup> during the middle of rush hour, which brought traffic to a standstill with cars stranded for many hours despite only 16.2 cm (6.4 in) of snow. Better identification of the potential for these events and understanding our ability to forecast them with sufficient forecast lead time is a motivating factor for this thesis.

While the ability to skillfully forecast Northeast ECs and associated snowfall accumulations has improved substantially over the last several decades, there remain high impact cases with deficiencies in forecast skill within the 48-hour to 96-hour lead time forecast range from the event. Notable cases include 25–27 December 2010 (Zheng et al. 2013), in which operational numerical weather prediction (NWP) model guidance simulated the cyclone too weak and too far to the east, and 25–27 January 2015 (Greybush et al. 2017), in which the operational model guidance simulated the cyclone too far to the west. Operational forecasters often utilize a rule of thumb stating that the forecast position of a Northeast EC will

trend left (i.e., closer to the coast) with decreasing forecast lead time. The implication of this rule of thumb is that this left trend leads to higher impact events than originally anticipated. While not well documented in literature, this rule of thumb appears to be based on a hypothesis that there is a systematic bias in NWP models to underrepresent the role of diabatic processes associated with convection in the incipient stage of the cyclone’s life cycle, thus failing to adequately simulate the downstream upper-tropospheric PV erosion associated with latent heat release and accordingly the amplitude of the EC. This rule of thumb motivates the question of whether such a systematic bias can be identified from a reforecast ensemble perspective, as well as differences between cases that exhibit different types of error and ensemble spread, which this thesis aims to assess with the goal of improving situational awareness for operational forecasting of these Northeast ECs.

## 1.2 Literature Review

This literature review contains three sections. Section 1.2.1 describes synoptic and mesoscale processes associated with cyclogenesis and snowfall. Section 1.2.2 provides an overview of NWP modeling of Northeast ECs and their biases and improvement over time, and Section 1.2.3 is a review of objective midlatitude cyclone identification and tracking algorithms.

### *1.2.1 East Coast Cyclones and Snowstorms: Overview, Climatology and Dynamics*

From both climatology and process perspectives, extratropical cyclogenesis in North America have been extensively studied in literature (e.g., Miller 1946; Uccellini 1990; Bentley 2018). At its most simplistic element, extratropical cyclogenesis is associated with baroclinic instability, as potential energy associated with low-tropospheric baroclinicity is converted into kinetic energy (Eady 1949). Early developments in simplifying the equations of motion into a quasigeostrophic (QG) system (Sutcliffe 1947; Charney 1948) allowed for an improved understanding of synoptic-scale dynamic forcing associated with cyclogenesis and upward vertical motion. Further application of QG theory by and later modifications by Trenberth

(1978) streamlined the usefulness of these equations for operations by highlighting the diagnostic relationship between omega and the advection of geostrophic vorticity by the thermal wind. This was also shown by Sutcliffe (1947) via the development theorem, in which the positive advection of thermal vorticity by the thermal wind contributes to a tendency for 1000-hPa convergence and amplification of the cyclone.

Cyclogenesis can also be assessed from a potential vorticity (PV) framework as well. Assuming a balanced flow, Davis and Emanuel (1991) applied an inversion technique to PV to describe surface cyclogenesis as a mutual interaction between a surface PV anomaly (SPVA) and an upstream upper-tropospheric PV anomaly (UPVA). As the UPVA approaches the downstream SPVA and its associated baroclinic zone, its circulation infringes on the baroclinicity and induces a southerly flow, with low-tropospheric warm air advection strengthening the SPVA and in turn inducing northerly flow advecting larger values of mean PV into the UPVA. This positive feedback process continues until warm seclusion occurs once the SPVA and UPVA become vertically aligned. Given that cyclogenesis is frequent near eastern North America (Sanders and Gyakum 1980), understanding the processes associated at the cyclogenesis stage, particularly along the coastal baroclinic zone, is important for analyzing these cyclones and identifying cyclogenesis objectively.

While ECs can occur in many parts of the world, the frequency of rapid cyclogenesis is favored in specific geographic regions, for example, off the eastern coasts of Asia and North America (Sanders and Gyakum 1980; Allen et al. 2010). Rapid cyclogenesis is more likely to occur in these regions due to enhanced low-tropospheric baroclinicity compared to over land, and/or latent heat release associated with higher moisture content over the warm ocean currents (Sanders and Gyakum 1980; Nuss and Anthes 1987; Reed et al. 1988). The combination of large baroclinicity and ample moisture acts to amplify the positive feedback cycle described by Davis and Emanuel (1991). A comprehensive study of Northeast U.S. snowstorms by Kocin and Uccellini (2004b) found that a majority of the highest impact snowstorms to affect the Northeast region were associated with intense cyclones compared to

climatology, generally below 990-hPa. In addition, a subset of the cyclones analyzed by Kocin and Uccellini (2004b) had rapidly deepened off the East Coast. The relationship between high impact weather and rapid cyclogenesis near the Northeast highlights the importance of understanding the ability to resolve and predict rapid cyclogenesis for forecasting Northeast snowstorms.

Despite variability in specific case environments in the high impact snowstorms analyzed by Kocin and Uccellini (2004b), compositing the synoptic-scale patterns during and preceding these high-impact cases shows an antecedent upper-tropospheric ridge over western North America with an amplifying downstream trough over eastern North America. Another feature of importance in the development of high-impact Northeast snowstorms is the presence of jet streak coupling near the Eastern U.S. In this composite, the poleward entrance region of the northern jet streak is near the Northeast, and associated with low-tropospheric cold air advection, while a subtropical jet streak is located to the south with cyclogenesis often occurring near the poleward exit region of this jet streak. Jet streak coupling is also often associated with rapid cyclogenesis, implying strengthened upper-tropospheric ageostrophic divergence and a deepening of the surface cyclone.

Mesoscale factors also have an important role in the intensity and location of heavy snowfall in Northeast snowstorms. Most mesoscale snow bands embedded within an EC tend to occur within the northwestern quadrant of the cyclone (Novak et al. 2004). The snow band formation and maintenance is the result of an increase in low-to-mid tropospheric frontogenesis (Novak et al. 2010). Stronger kinematic frontogenetical forcing, latent heat release associated with precipitation, and reduced conditional stability above the front in association with differential horizontal temperature advection contribute to the increased frontogenesis. Snow bands associated with Northeast snowstorms were partitioned into four categories: pivoting, quasi-stationary, laterally translating, and hybrid (Kenyon 2013). The climatological analysis of these four categories revealed that snow bands are typically associated with rapidly intensifying ECs. However, the analysis also found cases of synoptically weak low

pressures associated with a snow band resulting in heavy snowfall, such as the 30 March 2003 snowstorm (Evans 2006) which resembles the quasi-stationary band composite, and the Presidents' Day 2003 storm (Kocin and Uccellini 2004a), thus it should be recognized that not all Northeast heavy snowfall events are associated with rapidly intensifying cyclones.

### *1.2.2 Numerical Modeling of Northeast ECs: Deterministic and Ensemble Approaches*

Rapid cyclogenesis was often underforecast in the early days of NWP modeling through the 1970s (Sanders and Gyakum 1980). One such notable case was the 1979 Presidents' Day Storm (PDS). During the PDS, a strong anticyclone in the Northeast U.S. developed upstream of a strong baroclinic zone. Heavy snow developed in association with low-tropospheric warm air advection and isentropic ascent in the vicinity of the baroclinicity. The initial round of precipitation was accompanied by a weak surface cyclone over Kentucky. A subsequent secondary cyclone developed off the Mid Atlantic coast as a tropopause polar vortex (TPV) infringed upon the aforementioned baroclinic zone, inducing the development of this secondary cyclone which underwent rapid deepening (Bosart 1981; Uccellini et al. 1984; Uccellini et al. 1985). The operational NWP guidance available at the time of the PDS failed to simulate the rapid cyclogenesis of the secondary low. The major shortcoming in the operational forecasting of the PDS event spurred the need for research to improve understanding and forecasting of rapid cyclogenesis. When full physics were applied to simulations of this case, incorporating boundary layer fluxes and latent heat release, the model had a significantly better representation of the deepening rate of the cyclone. The improved representation of the boundary layer and diabatic processes led to a more intense and farther left of track cyclone than originally forecast (Nuss and Anthes 1987; Uccellini 1990).

The role of diabatic processes on surface cyclogenesis has been the subject of research since the 1950s (Reed et al. 1988). Studies such as Chen and Dell'osso (1987) and Reed et al. (1988) focusing on East Asian and North Atlantic cyclogenesis, respectively, found that latent heating had a large impact on the deepening rate of ECs. Condensation heating was

found to account for 40%-50% of the deepening for cases studied by Reed et al. (1988), with simulations that included latent heating having produced a deeper cyclones than those that contained full physics but removed the effects of latent heating. These studies highlighted the importance of properly accounting for diabatic processes in accurately simulating rapid cyclogenesis. Subsequently, additional studies evaluating operational NWP models available at that time found a systematic underforecasting of deepening rates of cyclones, especially over water, in addition to a slow motion bias (Grumm and Siebers 1989; Smith and Mullen 1993). It was hypothesized by Grumm and Siebers (1989) that the underdeepening bias over the ocean may be associated with lack of data as well as a systematic model cold bias.

Since the aforementioned studies, significant strides have been made in NWP modeling, including much higher horizontal and vertical model resolutions and parameterization of convective and cloud microphysical processes, and in sufficiently high model resolutions the ability to explicitly resolve cloud microphysics. As such, this set of same systematic biases that were evident in NWP models in the 1970s may not be applicable in modern NWP models. Nonetheless, there have been recent cases with short-term poor forecasts of EC track and intensity, such as the 24–25 January 2000 cyclone (Zhang et al. 2002), 26–27 December 2010 (Zheng et al. 2013), and 26–27 January 2015 (Greybush et al. 2017). In addition, improved forecast skill of EC track and intensity does not always correlate to improved snowfall forecasts, especially with regards to simulating mesoscale snow bands within the larger scale cyclone (Greybush et al. 2017).

A recent analysis of systematic NWP model biases in ECs in North America composited operational Global Forecast System (GFS) and North American Mesoscale (NAM) model forecasts for the 2002–2006 cool seasons, and found substantial spatial variability in model biases (Charles and Colle 2009). When focusing on the West Atlantic region off the East Coast and assessing position biases from a latitude and longitude coordinate framework, a westward EC track bias was evident in the NAM and slight westward track bias was calculated for the GFS in the 42–60 hour forecast lead time range. Colle and Charles (2011)

found further variability in biases as a function of lead time, with the GFS exhibiting a northeast cyclone position bias in the 102–120 hour lead time range, and a slight northwest cyclone position bias in the 36–48 hour lead time range. Cases where the GFS overdeepened the surface cyclone at 96-hour lead time also exhibited a northeast position bias, while cases where the GFS underdeepened the cyclone had a south position bias (Colle and Charles 2011).

Ensemble forecasting arose from the notion that random, small-scale perturbations in the initial conditions could have a large impact on the subsequent forecast evolution, that served as an early basis for chaos theory (Lorenz 1963). This concept served as a motivation for ensemble-based forecasting, where imposed perturbations in the model initial conditions and model physics are an attempt to represent the uncertainty and error in observations (Molteni et al. 1996). The ensemble of forecasts would theoretically produce a dispersive distribution of possible forecast outcomes, which should contain the eventual true observed solution.

More recent studies have applied an ensemble forecast approach to assess the predictability of East Coast ECs. While initial studies relied on using the adjoint of a forecast model to assess the sensitivity of the model forecast to initial conditions, Torn and Hakim (2008) demonstrated that an ensemble forecast approach can be used to link forecast uncertainty with initial conditions. This ensemble sensitivity method was applied to the 26–27 December 2010 snowstorm, which most operational models resolved as too weak and too far east, using the operational 50-member European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble (Zheng et al. 2013). The resulting analysis showed that medium-range uncertainty associated with the intensity of the cyclone could be traced back to the amplitude of an upstream wave packet, while along-track uncertainty was associated with the amplitude of the precursor cyclone and the position of an antecedent trough over the Gulf of Mexico. EC climatology research by Korfe and Colle (2018) compared ensemble forecasts for U.S. East Coast cyclones from the Global Ensemble Forecast System (GEFS), Canadian

Meteorological Centre (CMC), and the European Centre for Medium-Range Weather Forecasts (ECMWF). This study found a general bias to underdeepen the more intense ECs in the medium range, with a slow (negative along track) bias for most of the forecast range and a left of track bias from 120 to 144 hour lead times.

The studies on forecast skill climatology of East Coast ECs focus on similar, albeit slightly different, geographic regions and time frames. For many of these studies, the use of operational ensemble forecast data is a limiting factor to the time range of case selection. Often times, operational model upgrades lead to different ensemble versions which tend to exhibit different systematic forecast biases. Additionally, there have been comparatively few studies in literature pertaining to the synoptic-scale patterns associated with categorically different EC biases and errors. This thesis seeks to bridge this gap in literature by utilizing a reforecast dataset to create a longer term climatology of ECs and assessing the synoptic scale patterns associated with extremes in various error types.

### *1.2.3 Objective Cyclone Identification and Tracking*

The subject of analyzing trends and biases in EC forecasts brings forth the question of how to objectively identify and track ECs in a spatially gridded dataset. Tropical cyclones can be easily tracked due to their relatively slow motion and a coherent trackable vortex. In addition, there are authoritative documents that include official track and intensity data for observed tropical cyclones, such as IBTRACS (Knapp et al. 2010). In contrast, tracking midlatitude cyclones is more complicated. The literature lacks a single widely-accepted objective definition of a cyclone. ECs are also characterized by large vertical tilts, variability in cyclone horizontal symmetry and vertical depth, and ECs can merge and/or phase (Neu et al. 2013).

The current suite of cyclone identification and tracking algorithms utilize different variables to identify cyclones at consecutive time steps. The algorithms often use mean sea level pressure (Wernli and Schwerz 2006), 1000-hPa geopotential height minima (Blender

and Schubert 2000), 925-hPa area-averaged cyclonic relative vorticity (Bentley 2018), 850-hPa vorticity maxima (Hodges 1994; Sinclair 1997), or a combination of MSLP and vorticity (Murray and Simmonds 1991; Hewson and Titley 2010). These algorithms are primarily based on identifying and tracking features on single horizontal surfaces, which while generally sufficient, fails to account for the vertical structure of cyclones which can vary significantly throughout their life cycle. Some algorithms do account for vertical structure; Lim and Simmonds (2007) applied a pseudo-vertical identification and tracking criteria to represent the depth and tilt of the cyclone.

There are advantages and disadvantages to using a cyclone identification algorithm based on either vorticity maxima or pressure minima. The reasoning for selecting either variable can depend on the specific aspects of a cyclone are most relevant. Hodges et al. (2003) notes vorticity-based algorithms place a larger emphasis on the wind field and high-frequency Synoptic scale, while pressure-based algorithms emphasize the mass field and low-frequency synoptic scale. Existing literature presents multiple reasons for why each variable should or should not be used; Pinto et al. (2005) notes that algorithms searching for pressure minima overestimate deep, mature cyclones while failing to detect cyclolysis or cyclogenesis, which a vorticity-based algorithm would be better able to identify.

The second part of an algorithm after the identification of a cyclone is to construct a cyclone track. As defined by Neu et al. (2013), "a cyclone track consists of a series of cyclones identified in sequential time steps at adjacent locations, which are deemed to represent the same physical feature in reality." Tracking cyclones is sensitive to the temporal resolution of the data used. Most standard modern reanalysis data uses 6-hour increments, during which an average midlatitude cyclone can traverse upwards of 660 km assuming an average motion of  $110 \text{ km h}^{-1}$  (Blender and Schubert 2000; Neu et al. 2013). Identifying a cyclone track at its first time step is more difficult since there is no motion vector data from a pre-existing track. Current methods in tracking algorithms apply climatological motion vectors (Murray and Simmonds 1991; Sinclair 1997), or use the initial position of the cyclone as a first guess

location for its next location (Wernli and Schwierz 2006). For cyclones that have existed beyond one time step, methods apply a combination of climatological cyclone motion and extrapolated motion vector to calculate a probability of the likelihood of a match between two cyclones in different time increments (Murray and Simmonds 1991; Sinclair 1997). These algorithms utilize a reduced linear continuation of the track from a previous time step to account for climatological deceleration of a cyclone with time (Wernli and Schwierz 2006), or applying an approximated steering flow-based extrapolation (Hewson and Titley 2010).

The varying methodologies of each cyclone tracking algorithm has its advantages and disadvantages. A comparison of numerous algorithms was conducted by the Intercomparison of Mid Latitude Storm Diagnostics (IMILAST; Neu et al. 2013) project, applied to the same reanalysis data and the same spatial and temporal resolutions. The IMILSAT results show general agreement in tracks of intense, long-duration cyclones, but also show notable differences including a large spread in the lifespan of short lived cyclones. The algorithms implement different magnitudes of spatial smoothing which effectively reduces the horizontal resolution of the data and thus can impact the number of cyclones identified (Sinclair 1997; Blender and Schubert 2000). The spatial smoothing also has the effect of relaxing the magnitude of mesoscale features such as vorticity maxima. The smoothing is intended to reduce the signature from frontal boundaries that may have larger magnitudes of vorticity than the center of the cyclone, thus focusing the algorithm on the larger scale circulation associated with the synoptic scale cyclone. The smoothing could incorrectly combine multiple separate features within close proximity into the same feature (Sinclair 1997). Though smoothed data may miss potentially trackable features, unsmoothed data may be oversensitive to mesoscale variability. This highlights the importance of the horizontal resolution of the data used for cyclone identification (Blender and Schubert 2000).

Objective cyclone tracking algorithms are subject to similar limitations and errors regardless of methodology. External factors such as input data are such factors; the rapid storm motion of midlatitude cyclones, in addition to rapid acceleration or deceleration of

these ECs, can result in the tracking algorithm splitting a cyclone track into two or more separate cyclones when upon visual inspection these features could be considered the same physical entity. Alternatively, the presence of two cyclones in close proximity may result in the temporal merging of two seemingly separate physical entities as the same cyclone (Neu et al. 2013). Using smaller time increments can increased confidence that the same feature is being tracked between one time step and the next one.

Another area of weakness in cyclone tracking algorithms is with cyclogenesis and cyclolysis portions of a cyclone's life cycle, meaning algorithms can begin or end the cyclone track at much different times (Neu et al. 2013). The asynchronous cyclogenesis or cyclolysis times can be attributed to the degree of horizontal smoothing, or utilizing different pressure levels for tracking. The variation in vertical tilt and depth of the cyclone can be a function of its life cycle stage, but also vary based on the variables used for tracking. While most conventional algorithms exhibit weaknesses with the aforementioned scenarios, Hanley and Caballero (2012) accounts for cyclone mergers and splits as well as multicenter lows by using MSLP contours to identify the presence of multicentre cyclones (MCCs), which it found are common in intense cyclones. This algorithm, while an overall improvement, nonetheless has limitations with about 20% of connections between cyclones deemed as spurious.

The identification and tracking of cyclones near the U.S. East Coast presents additional complexities due to regional topography and phenomena, such as secondary redevelopment (Kocin and Uccellini 1990). In secondary redevelopment a primary cyclone, represented by a MSLP minimum approaching the East Coast from west of the Appalachian Mountains, can weaken west of the mountains as a secondary, initially weak, surface cyclone develops near a coastal baroclinic zone. Such cyclones are often referred to as "Miller-B" type cyclones (Miller 1946). These cyclones are characterized by cold air damming east of the Appalachian Mountains associated with high lower-tropospheric stability and northeasterly surface flow blocked by the terrain. These low-level characteristics enhance baroclinicity near the coastal region and facilitate a second area of surface cyclogenesis while the antecedent surface low

west of the Appalachians decays. These cyclones are known to be problematic for tracking algorithms because cyclones at their incipient stage are typically shallow and based in the lower-troposphere. Depending on the pressure level selected for the tracking algorithm and the distance matching criteria between time steps, tracking algorithms may be inconsistent with keeping this as the same cyclone or splitting it into two cyclones during the transfer process. Other cases where both primary and secondary low pressures are active at the same time present challenges regarding which cyclone can be considered the dominant feature, especially if they remain within close proximity to one another.

### 1.3 Research Goals and Hypotheses

This thesis will examine the predictability and prediction skill of Northeast ECs from an ensemble reforecast perspective. The perceived existence of a systematic right of track bias in forecasts of Northeast ECs will be investigated with the goal to improve the understanding of biases and errors of forecasts of Northeast ECs and help increase situational awareness. Specifically, from an ensemble reforecast dataset, cyclones with 1) right vs. left of track biases, 2) weak vs. strong biases, and 3) along vs. across track variability will be compared. The key research goals for this thesis, and their accompanying hypotheses that will be addressed in this thesis, are as follows:

1. Investigate whether there is a systematic right of track bias in the short-to-medium range.
  - Hypothesis 1: The null hypothesis is that there is no systematic right of track bias for all cyclones. This hypothesis is motivated by the forecaster rule of thumb suggesting a directional bias with decreasing lead time, in addition to prior studies (Korfe and Colle 2018; Bentley 2018) that show no systematic right of track bias for ordinary cyclones. A slight right of track bias exists for more intense cyclones, similarly to prior research by Bentley (2018).

2. Assess the spatial distributions and temporal trends in track and intensity errors for forecasts of Northeast ECs.

- Hypothesis 2: There is a correlation between an underdeepening bias and a right of track bias in the GEFS forecast. This hypothesis is motivated by numerical case studies of the PDS event showing an underdeepening bias associated with a right of track bias (Nuss and Anthes 1987; Uccellini 1990). More intense cyclones are associated with stronger downstream warm air advection, isentropic ascent and latent heat release, thus shortening the downstream half wavelength and resulting in stronger downstream PV erosion and downstream ridge amplification. The logic is that the aforementioned set of processes can reconfigure the flow diabatically to change the storm motion, and this set of processes is not well resolved by models at longer lead times.

3. Identify upstream and local synoptic-to-large scale patterns for ECs characterized by large and small ensemble mean track and intensity biases.

- Hypothesis 3: Cases with a weak or right of track bias are associated with more meridional flow over North America relative to a strong or left of track bias, respectively. This is motivated by the same reasoning used in hypothesis 2, but as assessed from a synoptic composite perspective. The logic motivating this hypothesis is that the meridional flow over North America is conducive to preconditioning the flow with more subtropical moisture.

4. Identify upstream and local synoptic-to-large scale patterns for ECs characterized by large and small along and across track ensemble mean position spread.

- Hypothesis 4: Cases with large across track variability are associated with more downstream ridging relative to small across track variability. The logic motivating this hypothesis is that larger across track variability is associated with an ampli-

fying upper level flow downstream of the cyclone, with a shorter half wavelength relative to small across track variability cases.

- Hypothesis 5: Cases with large along track variability are associated with a stronger jet streak over eastern North America relative to small along track variability. It is hypothesized that cases with larger along track variability are entering a region with a strong upper tropospheric jet aligned parallel to the cyclone motion vector.

Chapter 2 will focus on the data and methodology used to construct a climatology of Northeast ECs and detail the cyclone identification and tracking algorithm applied in this thesis. Chapter 3 will show and analyze the climatology of Northeast ECs as well as their intensity and track biases, followed by Chapter 4 analyzing the results from different composites of cases with varying ensemble error and spread. Chapter 5 will conclude the results from the preceding two chapters while offering suggestions for future work relating to the transfer of research to operations.

## 2. Data and Methodology

The purpose of this thesis is to first identify cyclones that tracked close to the Northeast U.S., then quantify several aspects of the predictability of these storms. The analysis specifically considers cool-season (Nov–Mar) extratropical cyclones that can be tracked within both reanalysis and ensemble forecast frameworks. When identifying cyclones with a potential to impact the major population centers in the Northeast, operational meteorologists in the National Weather Service often reference cyclones that track near the coordinate located at the 40°N and 70°W “benchmark”. From here on, the term "benchmark" refers to the 40°N and 70°W coordinate.

### 2.1 Data

Three sets of data were primarily used for this study: a reanalysis cyclone track dataset, reanalysis data, and reforecast data. This analysis uses a dataset of MSLP minimum tracks from the 70-km resolution European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim; Dee et al. 2011) that was generated by Sprenger et al. (2017) (referred to as Sprenger from here on). The Sprenger dataset was generated with the cyclone identification and tracking algorithm developed by Wernli and Schwerz (2006). This data was used for case selection, as further elaborated in section 2.2.1.

Ensemble forecasts were retrieved from the 1.0° horizontal resolution, 11-member Global Ensemble Forecast System (GEFS) Reforecast version 2 (Hamill et al. 2013). The GEFS Reforecast is available from December 1984 to present day and is based on the 2012 version of the GEFS, ensuring a consistent version of the model over a long term hindcast period. The forecast verification is based on the 0.5° horizontal resolution NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) dataset. Calculations using this gridded data were done using Python and a variety of libraries, including MetPy (May et al. 2008 - 2017) and GeoPy.

## 2.2 Case Selection

### 2.2.1 Case identification in Reanalysis Data

A preliminary list of cyclone cases was created using the Sprenger dataset. From the original list containing all global cyclone data, several filters to identify candidate cyclones that were likely to have affected the Northeast U.S. during the October–March time frame included retaining cyclones that (a) existed for a minimum of 48 hours, (b) tracked within a 750-km radius of the benchmark for a period of 12 or more hours, and (c) had a minimum lifetime MSLP below 1016 hPa and a minimum MSLP within the 750-km domain surrounding the benchmark below 1005 hPa. These criteria were selected to ensure that sufficiently intense and long-lived cyclones were identified. This set of cyclones are likely to be trackable with an objective tracking algorithm in an ensemble reforecast dataset. The selection criteria was applied from 1 November 1985, since the GEFS Reforecast is available from beginning 1 December 1984, through 31 December 2015 when the Sprenger dataset terminates. The case list yielded a preliminary total of 601 candidate cases.

There are some high-impact cases that exhibit multiple MSLP minima and are not identified as a single cyclone in the Sprenger dataset, and thus fail to satisfy the criteria above. To ensure most such high-impact events were included in this climatology, the preliminary list of events was compared against events classified in the Northeast Snowfall Impact Scale (NESIS; Kocin and Uccellini 2004a), which ranks NE snowstorms as a function of the spatial extent of snow, amount of snow, and the population affected by the snowstorm. One prominent case that failed to meet the preliminary case criteria was the 6–8 January 1996 blizzard, which ranks as a category 5 "*extreme*" storm and has the second highest NESIS value in the dataset. This case was manually added to the case list for a total of 602 candidate cases.

After identifying the preliminary cases, the next step was to ensure that these cyclones were trackable within the CFSR. This step implemented the objective cyclone tracking algorithm discussed in section 2.3 to the CFSR, with the CFSR horizontal resolution coarsened

to a 1.0° grid using bilinear interpolation to maintain consistency between the resolution of the CFSR and GEFS, and compared the cyclone tracks obtained from the CFSR to those from the Sprenger dataset. Cyclones between the two datasets were considered a match if:

- A track identified using the CFSR is within 500-km of the track from the equivalent cyclone identified in the Sprenger dataset
- The CFSR and Sprenger tracks are within 500-km of each other within a 12-hour centered window of when the Sprenger track was closest to the benchmark
- The CFSR and Sprenger cyclones both lasted for 24 or more hours

The criteria above filtered the number of cases down to 517 candidate cases.

### *2.2.2 Case identification in Ensemble Forecasts*

For all identified cases in the CFSR, the tracking algorithm was subsequently applied to the GEFS forecast data at several lead times to obtain cyclone tracks for each ensemble member. The time at which the CFSR track was closest to the benchmark will be referred to from here on as the "time of peak impact", and is used as a reference time to determine the time corresponding to forecast initialization and lead times, as well as verification. For some cyclone track calculations that require information from the time steps preceding and following the time of peak impact (e.g., storm motion vectors), a separate "verification time" was obtained by adjusting the time of peak impact forward or back by 6 hours in cases where the time of peak impact coincided with the first or last time identified in the CFSR track, respectively. This adjustment procedure is represented visually in Figure 2.1.

The GEFS Reforecast dataset is only initialized once daily, at 0000 UTC, in contrast with operational model guidance which is initialized four times daily. This motivates the need to find an initialization time that most closely matches to the day 0 lead time. The aforementioned verification time was used as the reference to determine which forecast initialization corresponded to certain lead time groups. This procedure is represented visually

in Figure 2.2. For cases that verified at 0000 UTC, the 0000 UTC run from that day was chosen as the day 0 lead time. Following a similar analysis conducted by Torn (2017), for cases that verified at 0600 UTC (1800 UTC), the closest initialization time is 6 hours prior (after). Handling 1200 UTC verification times is more complex, as Torn (2017) notes it is unclear whether the initialization time 12 hours prior or after should be used. While Torn (2017) discarded cases with a 1200 UTC verification time, this study retains them for the purposes of retaining all useful data. To ensure that the hour 0 lead time forecast is as close as possible to the time of cyclogenesis, if the CFSR track has a cyclone position 12 hours prior to the verification time, then the initialization time used is 12 hours prior, otherwise the initialization time used is 12 hours after.

### 2.3 Cyclone Tracking Algorithm

There are many complexities in tracking Northeast cyclones. These complexities are the result of discrepancies in methodologies that lead to differences in cyclone tracks between cyclone tracking algorithms (e.g., section 1.2.3). Since this research is focused on understanding the predictability of Northeast cyclones, an objective cyclone tracking algorithm was developed to account for these complexities of Northeast cyclone tracking. This algorithm was subsequently applied to both reanalysis and ensemble forecast perspectives.

#### 2.3.1 Selection of Variables and Pressure Level

As discussed in section 1.2.3, tracking algorithms are sensitive to the temporal and spatial resolution of the dataset, variable used for tracking (e.g., mean sea level pressure, relative vorticity, center of circulation, geopotential height), and pressure level (if applicable), among other factors. This section describes why a combination of 925-hPa area-averaged relative vorticity maxima and geopotential height minima were utilized in this algorithm to identify cyclones in gridded datasets. 925-hPa was selected (as opposed to MSLP) to account for secondary redevelopment of a surface cyclone, as discussed in section 1.2.3. The discrepancy between the primary and secondary low tends to be maximized near the surface, while the

secondary low lags in development at 850-hPa<sup>1</sup>.

Comparing these redeveloping, or Miller-B, cyclones in the ensemble forecast and verification process raises several questions, including whether the whole cyclonic circulation should be considered as one cyclone or two separate features. How to consider these cyclones from a forecast perspective can differ among cases. The considerations are a function of the location and timing of the redevelopment depicted in the ensemble member, and the pressure level used for tracking.

In ensemble forecasts, variability in the location and timing of the secondary cyclogenesis frequently associated with East Coast cyclones also presents an issue of inducing a large ensemble position spread. If there is a broad synoptic scale circulation associated with a cyclone with a weak MSLP gradient that is similar among ensemble members, the location of the absolute MSLP minimum can range in location from just west of the Appalachians to just off the Mid Atlantic coast. The ensemble cyclone positions in these cases will generally bifurcate into two clusters. This induces an artificial spread in cyclone position that is not necessarily indicative of a similar spread in mid-upper tropospheric forcing, or cyclone impacts.

One method of mitigating this superficial spread in cyclone position is by using a mass centroid approach to determine cyclone location. The centroid approach is frequently utilized to identify tropical cyclone centers in high resolution data, and can similarly be applied to a midlatitude cyclone with a broad pressure gradient and multiple MSLP minima similar in magnitudes. Here the methodology of Nguyen et al. (2014) is applied to the 925-hPa geopotential height field in the proximity of the absolute geopotential height minimum. A caveat for this approach exists with cyclones with large asymmetries in the pressure gradient, where the pressure gradient is often weaker on the south and east quadrants where pressure is lower in association with frontal boundaries. This caveat leads to a frequent bias with this approach for midlatitude cyclones with a single minimum and strong pressure gradient,

---

<sup>1</sup> The tracking algorithm was also implemented using the 850-hPa level, but secondary cyclogenesis was detected much later than at 925-hPa, justifying the use of the latter.

which shifts the position of the centroid south and east of its true location.

In addition to the height centroid method, applying horizontal area-averaging to the vorticity field is another way of mitigating the potential for artificial ensemble spread. As referenced in section 1.2.3, numerous tracking algorithms applied area-averaging and/or spatial smoothing to the field of interest (e.g., vorticity, MSLP). This effectively reduces the horizontal resolution of the data, thus reducing mesoscale features not associated with the synoptic scale cyclone, such as frontal boundaries. The area-averaging allows the tracking algorithm to focus on tracking the synoptic-scale cyclone. While a useful approach in the short term, it is unavoidable to retain at least some ensemble position variability especially at longer lead times due to larger ensemble spread in the location of the transfer of the surface cyclone. This larger spread can result in some variability in whether ensemble members meet the criteria for retaining the same cyclone as opposed to initiating a new cyclone off the coast, depending on several factors such as the distance that the cyclone traveled in a single time step. Nonetheless, the ensemble forecast position spread using this method is reduced compared to an approach that does not apply horizontal smoothing or area-averaging.

### *2.3.2 Cyclone Identification and Tracking*

The cyclone tracking algorithm developed for this thesis uses 925-hPa area-averaged relative vorticity and geopotential height as well as Python specific functionality. An idealized example of this is illustrated in Figure 2.3. To identify cyclones, the algorithm first calculates 700-km area-averaged 925-hPa relative vorticity and applies a 2-sigma Gaussian filter to eliminate mesoscale and shorter wavelength features, thus retaining the synoptic-scale vorticity on the order of 700-km or larger, a methodology similar to Sinclair (1997). In order to identify area-averaged vorticity maxima greater than  $1.5 \times 10^{-5} \text{ s}^{-1}$ , Python's SciPy maximum filter function (`scipy.ndimage.filters.maximum_filter`) is applied using a  $10^\circ$  window. For every such identified maximum, a geopotential height minimum centroid is computed, following the methodology of Nguyen et al. (2014), provided that there is a

geopotential height minimum within a 750-km radius of the 925-hPa area-averaged vorticity maximum.

Next, the algorithm attempts to match currently identified cyclones to those from the previous time step. A simple solution used in other tracking algorithms is to match the closest cyclone from the previous time step with the current cyclone to create a track (Wernli and Schwierz 2006). For this algorithm, the weighted proximity of both the identified vorticity maxima and height minima from the cyclones in the previous time step were compared to the current cyclone, and the cyclone where both the vorticity maxima and height minima locations were closest was matched. If no cyclones are identified in the previous time step within a 750-km radius of the current cyclone, it is considered to be a new cyclone.

For the second time step and beyond, Figure 2.4 provides an illustration of the steps below that describe how the tracking portion of the algorithm functions.

- Step A: For all identified cyclones in the current time step, the closest cyclone from the previous time step is chosen as a potential candidate cyclone for matching. If there is such a cyclone within a 600-km radius of the current cyclone location given a 6-hour time interval in the data (e.g., Blender and Schubert (2000)), step A(i) is performed. If not, then step B is performed.
- Step A(i): An extrapolated motion vector is computed from the previous two, or three if available, time steps of the candidate cyclone, and if the current cyclone is within 400 km of the extrapolated vector then the two cyclones are matched to create a track. If this condition fails, or if there is no cyclone in the previous time step within a 600-km radius of the current cyclone, then step B is performed.
- Step B: If there are (a) no cyclones at the current time step within a 1000-km radius of this cyclone, and (b) a cyclone in the previous time step within 750 km of the current cyclone, the two cyclones are matched to create a track. If this condition fails, step C is performed.

- Step C: If this cyclone’s position is within 60 degrees of the extrapolated motion vector of the previous cyclone, then the two cyclones are matched to create a track. Otherwise, this process is repeated for all cyclones from the last time step within a 750-km radius of this cyclone, and if no successful match is found, then this cyclone is considered to be a new cyclone.

The algorithm was applied for all candidate cyclones, and was applied to track all cyclones in the North America domain beginning from 24 hours prior to when the cyclone was first identified in the Sprenger dataset. In order to focus on tracking cyclones near the U.S. East Coast and to avoid complexities associated with cyclolysis or cyclone mergers towards southern Greenland and Iceland, the algorithm was cut off 36 hours past the time of peak impact.

### *2.3.3 Matching Ensembles Tracks to Cyclones*

Once cyclone tracks have been identified in the CFSR dataset, the next step is to track these cyclones in the GEFS Reforecast dataset, as described in section 2.2. The tracking algorithm is applied to all cyclones identified in the forecast dataset from 0 through 5 day forecast lead time. Considering cyclones in a forecast framework necessitates a method to match ensemble tracks to the CFSR verification reference cyclone. Since forecast position error grows with forecast lead time, sometimes non-linearly (Korfe and Colle 2018), a search radius with varying size as a function of lead time was applied following the methodology of Bentley (2018). The cyclone matching search radius starts with 400 km at 0-hour lead time and expands by 200 km every 24 hours to reach 1400 km at 120-hour (5-day) lead time. To match the ensemble track to the verification CFSR cyclone, the ensemble track must fall within this search radius at any point within a 12-hour window centered at the verification time. If multiple ensemble tracks occur within this radius, the closest ensemble track is selected as a match to the CFSR verification cyclone track.

## 2.4 Ensemble Forecast Diagnostics and Verification

One of the main goals of this thesis is to assess the spatial distribution and temporal trends of ensemble forecasts for Northeast ECs, as well as to determine whether a systematic right of track bias exists (section 1.3). To investigate the ensemble track and intensity spread and error, the mean ensemble standard deviation (STDEV) and mean absolute error (MAE) for all cases were calculated.

To evaluate whether the GEFS forecasts of cyclone position and intensity are well calibrated, the ratio of the mean STDEV to the mean bias-corrected root mean square error (RMSE) was calculated following the methodology described in Murphy (1988). The STDEV is defined as follows:

$$STDEV(t) = \sqrt{\frac{1}{M(t)} \sum_{m=1}^{M(t)} \sigma_m^2(t)}$$

$$\sigma_m^2(t) = \frac{1}{N-1} \sum_{n=1}^N (x_n(t) - \bar{x}_m(t))^2$$

Where  $M(t)$  is the total number of cases at lead time  $t$ ,  $\sigma_m^2(t)$  is the variance of case  $m$  at lead time  $t$ ,  $N$  is the total number of ensemble members at lead time  $t$ ,  $x_n(t)$  is the forecast value from an individual ensemble member  $n$  at forecast lead time  $t$ , and  $\bar{x}_m(t)$  is the ensemble mean value at forecast lead time  $t$ . The bias-corrected root mean square error (RMSE) is defined as follows:

$$RMSE(t) = \sqrt{\frac{1}{M(t)} \sum_{m=1}^{M(t)} (\bar{x}_m(t) - b(t) - o_n(t))^2}$$

$$b(t) = \frac{1}{M(t)} \sum_{m=1}^{M(t)} (\bar{x}_m(t) - o(t))$$

Where  $b(t)$  is the bias (e.g., systematic error) of the ensemble forecast climatology at forecast lead time  $t$ , and  $o(t)$  is the verification CFSR value valid at the same time as

the ensemble mean forecast  $\bar{x}_m(t)$ . For ensemble position forecasts, the bias was calculated by partitioning the bias into latitudinal and longitudinal components, and removing the component biases from the ensemble mean forecast at time  $t$ . This calculation will provide information on whether the GEFS Reforecast ensemble is well calibrated. A well calibrated ensemble will have a STDEV-to-RMSE ratio equal to one, where the STDEV equals the RMSE. A ratio less than one, with a STDEV statistically significantly less than the RMSE, indicates an underdispersive bias – meaning the forecast verified outside of the ensemble spread. A ratio greater than one, with a STDEV statistically significantly larger than the RMSE, indicates an overdispersive bias – meaning the verification lies inside the ensemble spread (Murphy 1988).

In addition to STDEV, ensemble variability can be further analyzed from a position variability relative coordinate to ascertain whether most of the ensemble forecast position variability is in the along-track or across-track direction. To do so, the ensemble forecast position ellipse methodology utilized in Hamill et al. (2011) was employed, and is represented in Figure 2.5. To calculate this ellipse, eigenvectors of the displacement of ensemble forecast cyclone positions relative to the ensemble mean are computed, with the largest eigenvector (S1) denoting the direction of greatest position variability and thus the ellipse major axis. The angle between eigenvector S1 and the ensemble mean motion vector (MV) were subsequently compared. If the angle is zero, with MV parallel to S1, variability is predominantly in the along-track direction. If the angle is  $90^\circ$ , with MV normal to S1, variability is predominantly in the across-track direction.

## 2.5 Composite Analyses

The upstream and local synoptic-to-large scale patterns for ECs associated with various ensemble mean spread and errors are considered in hypotheses 3–5 and are assessed using a composite analysis approach. These composites were created using CFSR data, but cases are assigned to groups based on different bias and spread categories. Given the correlation

in intensity and across-track biases discussed in section 3.3, composites were created with CFSR data based on ensemble mean biases and spreads calculated for forecasts at day 3 lead time. These day 3 lead time biases and spreads are computed over a 12-hour period centered at the time of peak impact, or essentially when the CFSR verification cyclone track was at its closest proximity to the benchmark location, to ensure temporal continuity in the observed biases or ensemble spreads over the 12-hour time frame. These composites were further filtered to cyclones that occurred within a specific radius of the benchmark, varying by composite group, to prevent large positional variability. To assess the hypotheses stated in section 1.3, composites were created as follows:

- **Left vs. Right of Track Bias:** Cases where the day 3 lead time ensemble mean exhibited an across-track error corresponding to either left of track (LOT) or right of track (ROT) greater than 25 km compared to the CFSR storm motion vector, and consistently exhibited that error for the 12-hour period centered at verification time, were selected as candidate cases. The closest proximity of the CFSR verification cyclone to the benchmark must be 250 km or less for these cases to be selected. These thresholds yielded 49 LOT cases and 53 ROT cases.
- **Weak vs. Strong Bias:** Cases where the day 3 lead time ensemble mean exhibited a weak (WB) or strong (SB) 925-hPa geopotential height bias compared to the CFSR verification value, and consistently exhibited that error for the 12-hour period centered at verification time, were selected as candidate cases. The closest proximity of the CFSR verification cyclone to the benchmark must be 250 km or less for these cases to be selected. These thresholds yielded 63 WB cases and 44 SB cases.
- **Small vs. Large Across Track Variability:** Cases where the day 3 lead time ensemble position spread in the across track direction relative to the ensemble mean motion vector was in the lower quartile (small variability; SV) or the upper quartile (large variability; LV) of the GEFS climatological across-track variability, with that spread

consistent for the 12-hour period centered at verification time, were selected as candidate cases. The closest proximity of the CFSR verification cyclone to the benchmark must be 275 km or less for these cases to be selected. These thresholds yielded 36 SV and 36 LV cases.

- **Small vs. Large Along Track Variability:** Cases where the day 3 lead time ensemble position spread in the along track direction relative to the ensemble mean motion vector was in the lower quartile (small variability; SV-L) or the upper quartile (large variability; LV-L) of the GEFS climatological along-track variability, with that spread consistent for the 12-hour period centered at verification time, were selected as candidate cases. The closest proximity of the CFSR verification cyclone to the benchmark must be 275 km or less for these cases to be selected. These thresholds yielded 36 SV-L and 36 LV-L cases.

For each bias group, composites were computed for MSLP, 500-hPa geopotential height anomaly relative to the 1981-2010 CFSR climatological value at the time of each case, and 300-hPa wind speed. Composite differences were subsequently calculated and tested for statistical significance at the 95% confidence level using a 1000 resample bootstrap test following the methodology of Rios-Berrios et al. (2016).

Selecting Verification Time	
Case 1	<p>Time of peak impact is the <u>first</u> time step in the CFSR</p> <p>Verification time = TOPI + 6 hours</p>
Case 2	<p>Time of peak impact is the <u>last</u> time step in the CFSR</p> <p>Verification time = TOPI - 6 hours</p>
Case 3	<p>Time of peak impact is <u>neither</u> the last nor time step in the CFSR</p> <p>Verification time = TOPI</p>

Figure 2.1: Procedure followed for selecting verification time using the time of peak impact (TOPI) for each case. Green circles represent (top) TOPI and (bottom) verification time, or in case 3 both TOPI and verification time.

Selecting Day 0 Initialization					
Verification time	Day 0 initialization				
0000 UTC	Verification Time (0000 UTC this day)				
0600 UTC	Verification Time - 6 hours (0000 UTC this day)				
1200 UTC	<table border="1"> <tr> <td><i>If cyclone track existed at verification time - 12 hours:</i></td> <td><i>If the left condition is false:</i></td> </tr> <tr> <td>Verification time - 12 hours (0000 UTC this day)</td> <td>Verification time + 12 hours (0000 UTC next day)</td> </tr> </table>	<i>If cyclone track existed at verification time - 12 hours:</i>	<i>If the left condition is false:</i>	Verification time - 12 hours (0000 UTC this day)	Verification time + 12 hours (0000 UTC next day)
<i>If cyclone track existed at verification time - 12 hours:</i>	<i>If the left condition is false:</i>				
Verification time - 12 hours (0000 UTC this day)	Verification time + 12 hours (0000 UTC next day)				
1800 UTC	Verification Time + 6 hours (0000 UTC next day)				

Figure 2.2: Procedure followed for selecting the day 0 lead time initialization based on the case verification time.

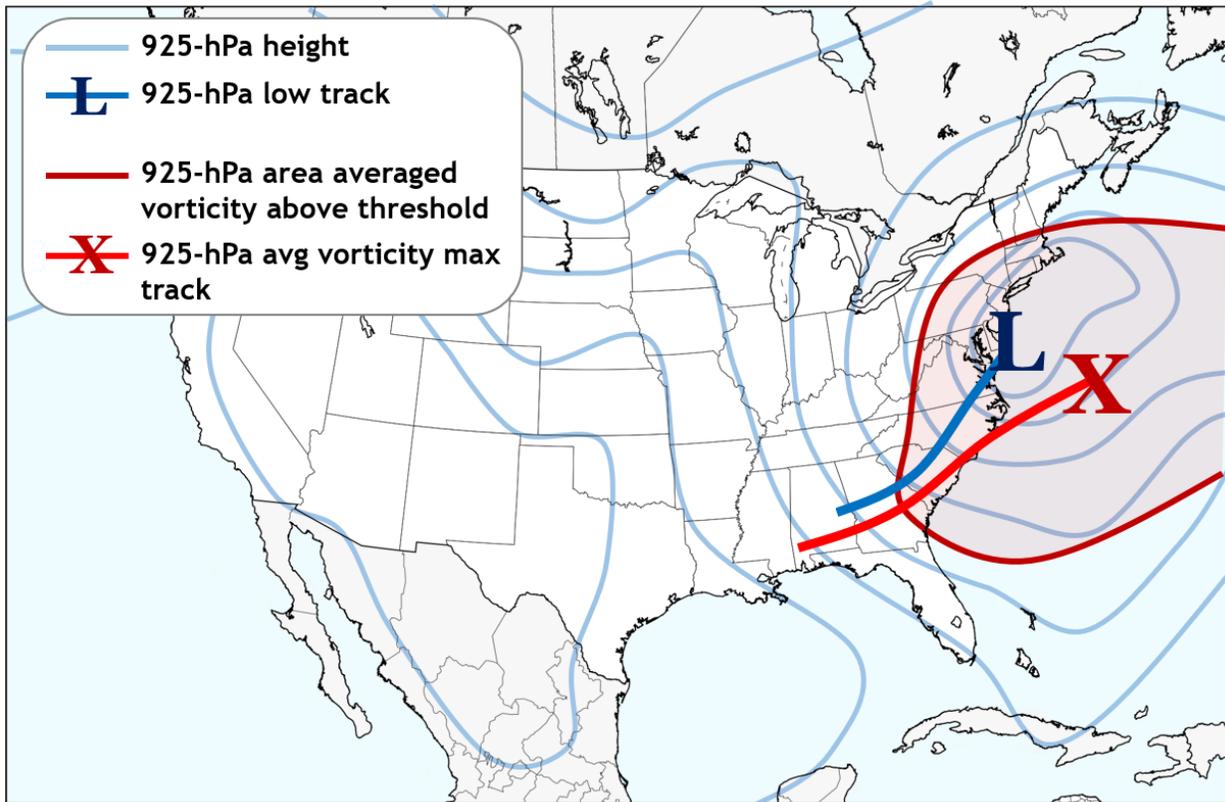


Figure 2.3: Idealized illustration of the cyclone tracking algorithm. The 925-hPa geopotential height field is contoured in light blue contours, with the dark blue line representing the 925-hPa height minimum track. The area encompassed by the area-averaged 925-hPa vorticity exceeding the minimum threshold is denoted by the dark red contour, with the area-averaged vorticity maximum track in the bold red line.

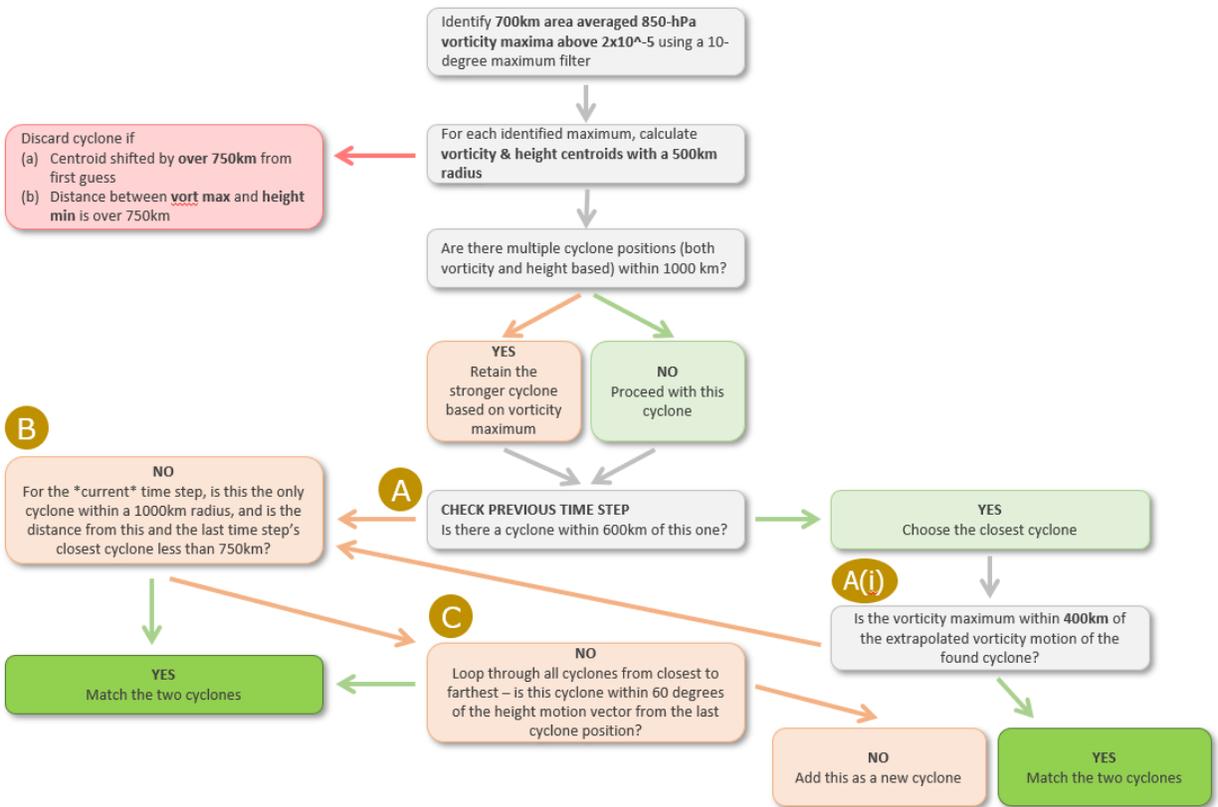


Figure 2.4: Procedure followed for the cyclone tracking algorithm from the second time step and beyond of an existing cyclone.

Sample ensemble member cyclone positions  
valid at time step  $t=1$

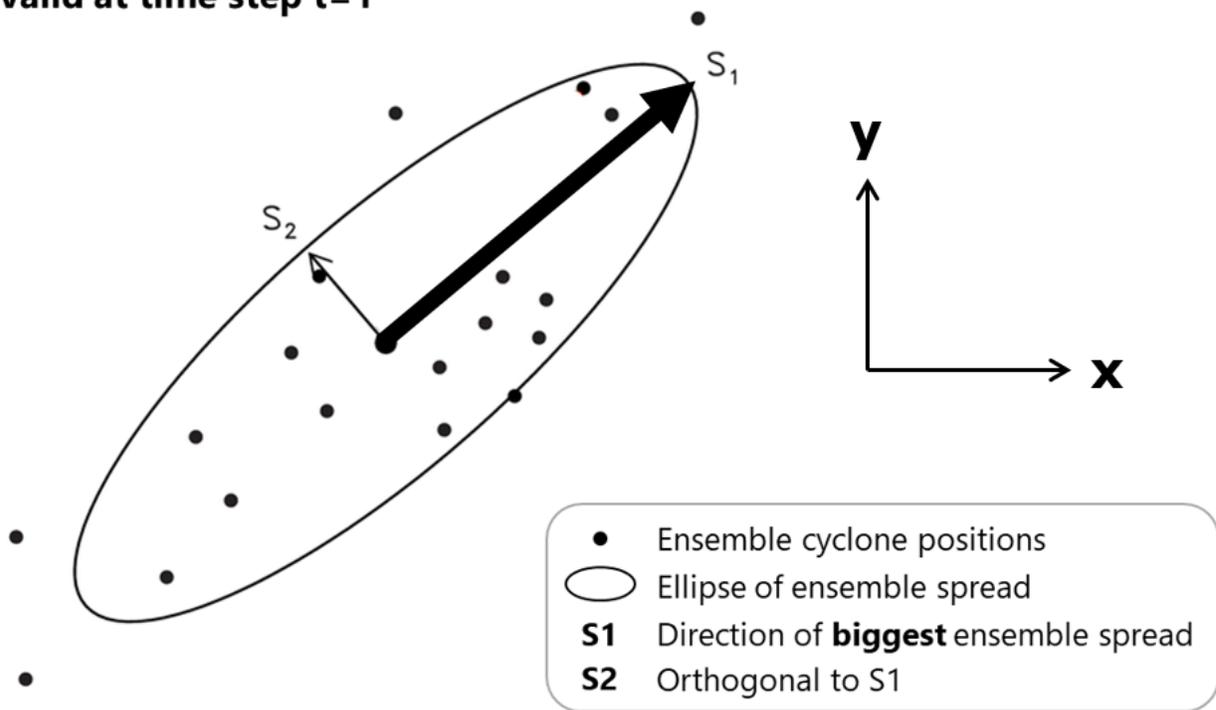


Figure 2.5: Forecast position ensemble spread ellipse, adapted and modified from the original figure in Hamill et al. (2011). Individual dots represent ensemble member positions, the circle represents the bivariate normal distribution fit to contain 90% of the data,  $S_1$  represents the eigenvector along the direction of greatest ensemble spread, and eigenvector  $S_2$  is normal to  $S_1$ .

### 3. Climatology

#### 3.1 Climatology of All Candidate Cyclones

Figure 3.1 shows the identified CFSR cyclone tracks for all cases in the October–March 1985–2015 period using the cyclone tracking algorithm described in section 2.3. The majority of the candidate cyclones can be considered coastal lows and progress from southwest to northeast. Generally, these cyclones are first identified near the Southern U.S., the northern Gulf of Mexico or the Southeast U.S. coast, with the more intense cyclones concentrated over the western Atlantic Ocean and in higher latitudes. There is an additional smaller subset of northwest to southeast progressing cyclones identified over southern Canada that are generally weaker than the coastal low subset. The density of CFSR cyclone tracks, calculated as the number of cyclone tracks within a 350-km radius normalized by the total number of cyclones, is shown in Figure 3.2a. The cyclone track density shown in Figure 3.2a differs from climatologies of all of North American cyclone tracks, with the present subset exhibiting a slight westward shift of the maximum east of New England and the absence of additional maxima in other regions of North America, but nonetheless exhibits a similar maximum in cyclone track density just north of the Gulf Stream southwest of Nova Scotia. The westward shift is likely a result of not including cyclones that track too far south to satisfy the criteria for candidate cases listed in section 2.2.1.

The cyclogenesis location density (Fig. 3.2b) is calculated based on the locations where cyclones are first identified by the tracking algorithm. For the purposes of this thesis, "cyclogenesis" will be defined from here on as the location where the cyclone was first identified in the tracking algorithm, with this caveat considered given that the same cyclone may have a different cyclogenesis location and time when calculated using a different tracking algorithm. There are two distinct maxima in cyclogenesis location density: one over the western Atlantic Ocean west of the benchmark, and a second over the South Mississippi

Valley. A third less distinct maximum exists over southern Manitoba. The first region of cyclogenesis near the benchmark is likely associated with coastal cyclogenesis events in the Gulf Stream region as well as secondary redevelopment of cyclones originating farther inland. The second region of cyclogenesis in the Southern U.S. is likely associated with cyclogenesis downstream of upper-tropospheric troughs emerging from the Rockies, and redevelopment of lee cyclones along the climatological baroclinic zone in the region (Bentley et al. 2019).

### 3.2 Ensemble Forecast Skill and Calibration

The objective tracking algorithm described in section 2.3 was applied to all candidate cyclones in the GEFS Reforecast dataset. Especially at longer lead times, uncertainties in the cyclone evolution, motion and intensity may result in ensemble members that do not have an identified cyclone matched to the CFSR verification track data. This requires an analysis of the average number of ensemble members detected for each cyclone to obtain a minimum threshold of ensemble members for subsequent analyses. Averaged over all cases, the average number of ensemble members that contain a cyclone matched to a CFSR track plotted by lead time, is shown in Fig. 3.3a. On average, there are 10–11 members of all 11 possible members that resolved the cyclone through 54 hours. This number decreases to 9 members by 90 hours and 8–9 members by 114 hours. Fig. 3.3b shows the percent of cases by lead time in which the GFS forecast contains  $\geq 7$  or  $\geq 11$  ensemble members of all 11 ensemble members that resolved the case. While the number of cases with all 11 members containing a matching cyclone decreases to below 40% by 102-hour lead time, the number of cases with  $\geq 7$  members remains at above 80% through the entire 120-hour range. With the higher percentage of cases containing 7 or more members considered, the subsequent results will only incorporate an evaluation of the GEFS reforecast data corresponding to cases where  $\geq 7$  ensemble members were identified.

These ensemble forecast position tracks are used to evaluate the skill in forecasting cyclone position and intensity as quantified by 925-hPa geopotential height. Figure 3.4

shows the mean ensemble standard deviation (STDEV) and mean absolute error (MAE) for both forecast cyclone position and intensity. The position standard deviation (Fig. 3.4a) increases exponentially in the first 84 hours with a transition to a linear increase thereafter. The average position STDEV is 59 km at 24-hour lead time, 197 km at 72-hour lead time, and 472 km at 126-hour lead time. The ensemble mean error (Fig. 3.4b) increases linearly with lead time, with an average of 97 km at 24-hour lead time, 253 km at 72-hour lead time, and 454 km at 126-hour lead time. The cyclone intensity forecast STDEV (Fig. 3.4c) increases somewhat non-linearly with lead time, with an average of 6.8 m at 0-hour lead time, 27.4 m at 72-hour lead time, and 56.7 m at 126-hour lead time. The MAE of the cyclone intensity (Fig. 3.4d) increases nearly linearly, with an average of 1.9 m at 0-hour lead time, 36.1 m at 72-hour lead time, and 57.7 m at 126-hour lead time.

To evaluate whether the GEFS forecasts of cyclone position and intensity are well calibrated, the ensemble mean standard deviation (STDEV) and the mean bias-corrected root mean square error (RMSE) are computed following the methodology described in section 2.4, averaged over all cases for all initializations and all forecast hours where seven or more ensemble members were identified. As elaborated in section 2.4, a STDEV-to-RMSE ratio of one indicates a well calibrated ensemble, a STDEV below the RMSE indicates an underdispersive bias, and a STDEV above the RMSE indicates an overdispersive forecast. The resulting RMSE and STDEV for position and intensity forecasts are shown in Figure 3.5. The forecast position results (Fig. 3.5a) suggest a slight overdispersive forecast at initialization, which may be a result of the initial condition perturbation applied to the GEFS. However, there is a consistent underdispersive position bias from 12-hour to 126-hour lead time, which is significant at the 95% confidence level when using a two-sided t-test for the null hypothesis that these two samples have identical averages. The intensity forecast (Fig. 3.5b) is similarly overdispersive at initialization, and is significantly (at 95% confidence level) underdispersive from 6-hour lead time through the end of the analysis period at 126-hour lead time.

Ensemble variability can be further assessed utilizing a position variability relative coordinate, as described in section 2.4. The ensemble position ellipse methodology from Hamill et al. (2011) was applied to all cases at all forecast initializations, for all forecast hours where seven or more ensemble members contained cyclones matched to the verification cyclone. The vector along the direction of greatest ensemble spread was compared to the ensemble mean motion vector to assess whether variability is predominantly in the along or across track direction. The angle between the two vectors was calculated, with the distribution of the angles as a function of lead time shown in Figure 3.6. An angle of zero indicates predominantly along-track variability, while an angle of  $90^\circ$  indicates predominantly across-track variability. The majority of the cases exhibit mainly along-track variability, with the median angle between  $25^\circ$  and  $30^\circ$  from 0 through 5 day lead times.

The ratio of the S1 to S2 eigenvector (defined in Fig. 2.5) can be used to further assess whether there is large preferred direction variability. The resulting distribution of ellipse ratios for all cases across 0 through 120 hour forecast lead times shown in Figure 3.7. Most cases have a ratio between 1 and 3, with a smaller subset of cases having a ratio between 3 and 4. Compared to an analysis of ensemble position ellipse ratios applied to tropical cyclones (Elless 2015), there is a larger frequency of extreme preferred direction variability within this climatology of ECs (i.e., ratio above 3) compared to a climatology of tropical cyclone forecasts, implying that there often is preferred direction variability.

### **3.3 Climatology of Cyclones Partitioned by Forecast Skill**

The forecast position error of the ensemble mean cyclone is examined with respect to the CFSR storm motion vector, at multiple forecast lead times (Fig. 3.8). The majority of the cases exhibit minimal error at initialization (Fig. 3.8a). A small number of cases exhibit larger position errors, as large as 100 km. While some larger position errors at initialization are to be expected given that initial condition perturbation was applied to the GEFS (Hamill et al. 2013), some cases where a broad area of low pressure with a weak

925-hPa height gradient may also inevitably lead to larger uncertainty regarding the initial position of the cyclone in each ensemble member. By 24-hour lead time, the forecast position error grows as represented by the frequency distribution of the maximum shifting towards a slow and left track bias (Fig. 3.8b). This slow and left of track bias is also apparent in the 48-hour forecast data, before becoming less discernible by the 72-hour forecast. Longer lead times towards 5-day lead time (not shown) show a more prominent left of track bias. This near-term left of track and slow bias in the GEFS data is consistent with previous results by Bentley (2018), as well as the slow bias found in an examination of North American cyclones in the operational GEFS by Korfe and Colle (2018). The short-term left bias is statistically significant from zero across-track bias at the 95% confidence level from 24-36 hour and 48-66 hour lead times. The slow bias is significant from 0 through 66 hour lead times. These biases suggest that for an overall climatology of cyclones, there is no systematic right of track bias for Northeast ECs.

Bentley (2018) found a right of track bias beyond forecast hour lead time of 156 hours for extreme weather events, described by a category of anomalously intense cyclones that lasted for a sufficiently long duration. While this thesis does not consider such long lead times, a similar comparison of position error for all cyclones compared to the top 20th percentile of cyclones as ranked by intensity, defined here as magnitude of 925-hPa area-averaged vorticity, is presented in Figure 3.9. All cyclones and intense cyclones alike show a significant (at the 95% confidence level) slow and left of track bias through 36-hour lead time. While all cyclones continue to exhibit a slow and left bias through 78-hour lead time, the left bias is no longer significant for intense cyclones beyond 36-hour lead time, with a slow bias evident through 54 hours. The all cyclone group exhibits a left of track bias consistently throughout the forecast period, especially towards 120 hours, while the intense cyclones have no significant bias at longer lead times in neither the along nor across track direction. The results presented here do not show a systematic right of track bias for Northeast ECs. It is possible that such a bias may be evident at longer lead times not shown here, but captured

by Bentley (2018).

Despite the significant short term left of track mean bias, it can be subjectively assessed that the number of cases with a right of track bias especially prior to 72-hour lead time is not negligible compared to the number of cases with a left of track bias (Fig. 3.8b,c). The relatively similar number of right vs. left of track bias cases was used to test Hypothesis 2 (section 1.3) stating that a correlation exists between an underdeepening and right of track bias. Figure 3.10 shows the Pearson correlation coefficient (PCC) between across-track error and intensity error as a function of forecast lead time. There is a consistent negative PCC across all lead times, which can be interpreted as a correlation between an underdeepening and right of track bias, and similarly overdeepening and left of track bias in an ensemble mean forecast. This correlation is significant from one through five day lead times at the 95% confidence level using the Pearson  $r$ , Spearman  $r$ , and Kendall Tau significance tests (Wilks 2011). The relationship peaks at day 3 lead time (Fig. 3.11) with  $r = -0.45$  before decreasing to  $r = -0.33$  by 5-day lead time (Fig. 3.11).

### 3.4 Climatology Discussion

Section 3.1 reviewed a climatology of all CFSR cyclone tracks from the 1985–2015 period (Fig. 3.1), as well as cyclone track density (Fig. 3.2a) and cyclogenesis density (Fig. 3.2a). The CFSR cyclone track distribution is consistent with prior North America extratropical cyclone (EC) climatology studies (e.g., Bentley et al. (2019)). The overall regions of cyclogenesis are similar to the cyclogenesis density plot of ordinary cyclones from Bentley et al. (2019). The differences occur near the coastal Atlantic where the maximum is shifted poleward in this climatology, likely associated with the selection of 925-hPa vs. surface levels used to calculate the cyclone tracks. The second maximum in the southern U.S. is shifted equatorward compared to Bentley et al. (2019), likely due to failing to account for cyclones that track too far north to satisfy the criteria for candidate cyclones. The cyclogenesis region associated with lee cyclogenesis in Bentley et al. (2019) is absent in this

climatology, likely indicating that the majority of lee cyclones track too far north and/or undergo cyclolysis too early to satisfy the criteria for candidate cyclones.

Sections 3.2 and 3.3 reviewed whether the ensembles are well calibrated, and examined across-track biases for ensemble mean forecast position as a function of lead time to assess if a right of track bias exists. For both the all cyclone group and intense cyclones, no systematic right of track bias was identified. For all cyclones, a large left of track bias is apparent at long lead times, towards 120 hours (Fig. 3.9a). It is possible that this left bias may be a result of the type of cases that exhibit a right of track (and accordingly weak bias) containing ensemble members with cyclones that were too weak to be identified by the tracking algorithm, while the cyclones with a left of track and strong bias were well identified (section 3.3). This situation would lead to an undersampling of right of track cases at longer lead times which would manifest as an overall left of track bias averaged over all cases. This undersampling hypothesis has not been tested in this thesis and may be a subject of future work.

Section 3.3 further found a negative correlation between track directional biases and intensity biases, peaking at day 3 lead time before decreasing towards longer lead times. It is speculated that the peak in correlation at day 3 lead time may be a result of ensemble uncertainty mostly focused on shortwave troughs serving as precursors to cyclogenesis. These upper-level features can lead to relatively minor track or intensity variability while the longer wavelength regime is generally well modeled. This may not be the case at longer forecast lead times beyond day 3, where additional uncertainties such as the type of cyclone evolution and long wavelength regime break down the correlation between track and intensity biases. The decrease in the correlation beyond day 3 lead time may also be associated with the average time of an upper-tropospheric trough resulting in the incipient East Coast cyclogenesis to traverse North America over the radiosonde sampling network, with upper-level signatures 72 hours preceding Northeast U.S. snowstorms often over western North America (Kocin and Uccellini 2004b). However, this interpretation is speculative and these hypotheses have

not been tested as part of this thesis.

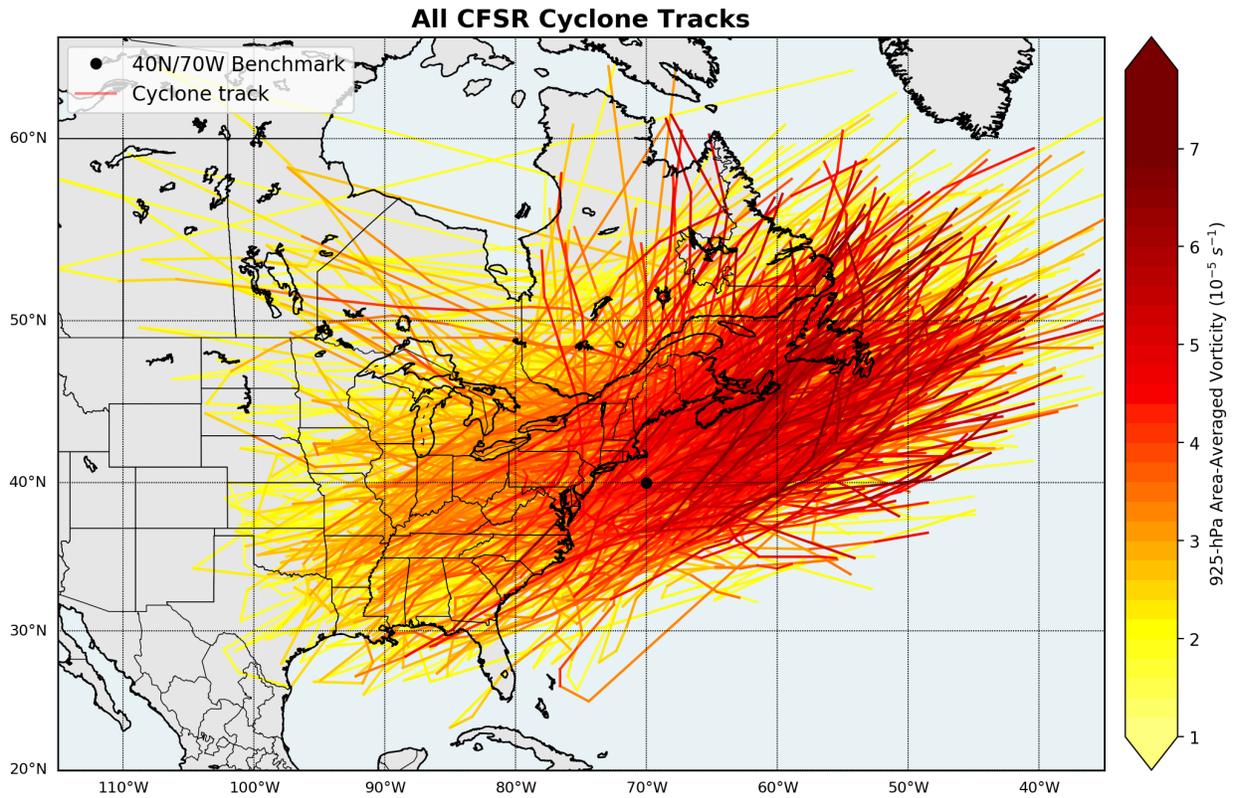


Figure 3.1: Tracks of all objectively identified cyclone tracks from the CFSR, shaded by their 925-hPa area-averaged vorticity amplitude in  $10^{-5} s^{-1}$ .

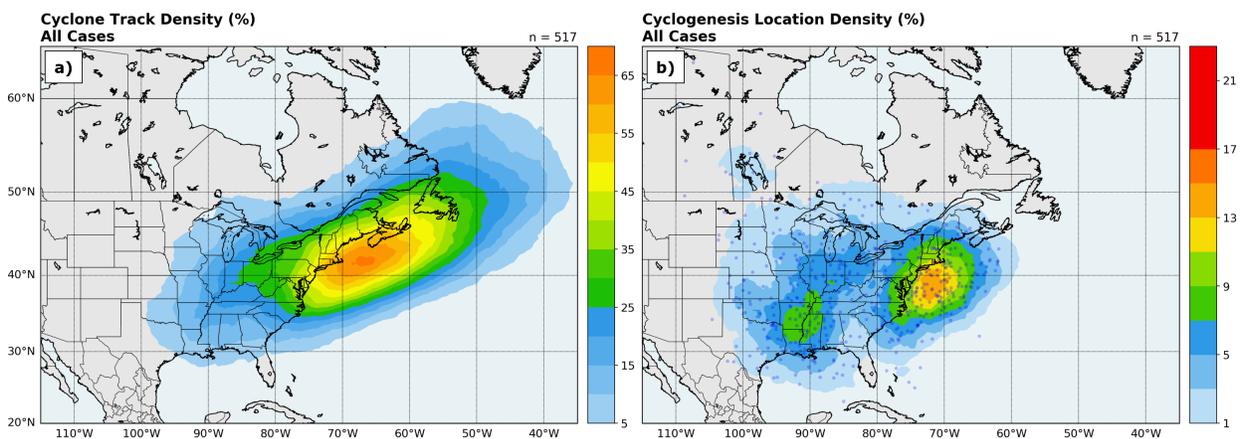


Figure 3.2: Density of (a) all cyclone tracks and (b) location of cyclone when first objectively identified, shaded by the percentage of cyclones within 350 km of a grid point normalized by the total number of cyclones.

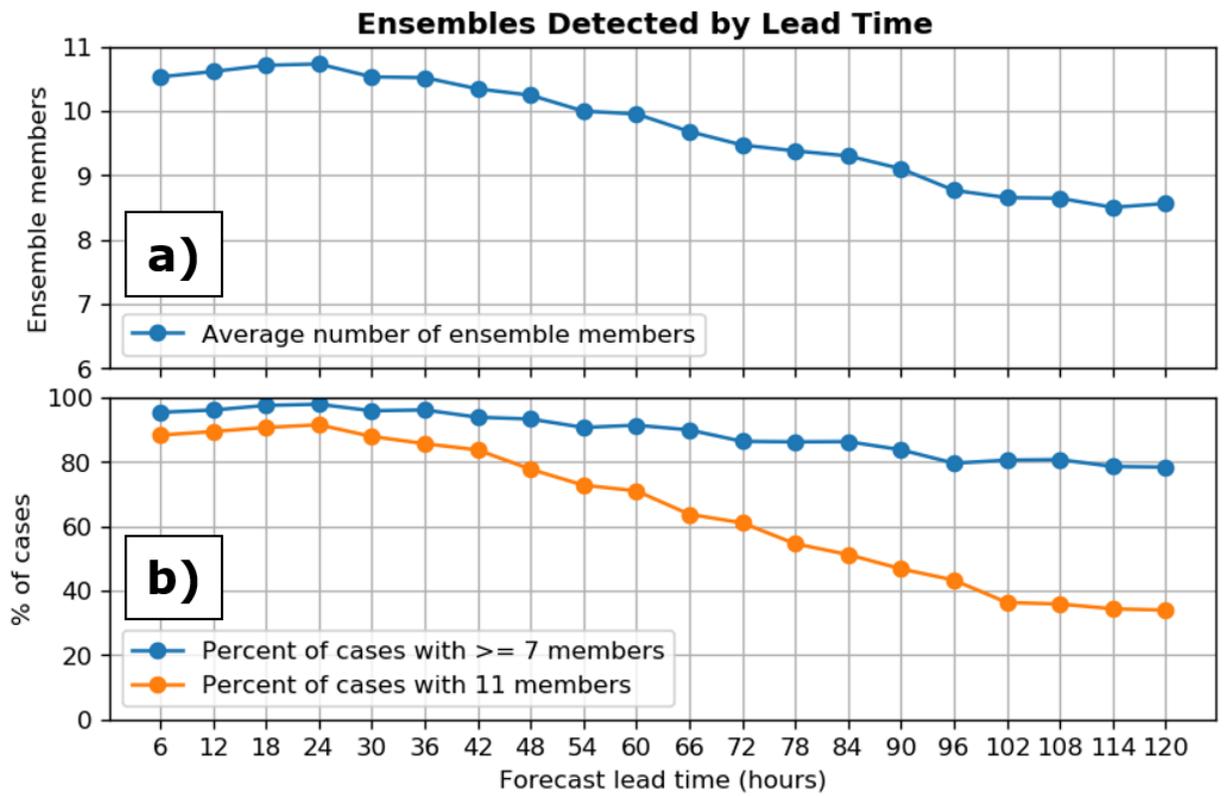


Figure 3.3: (a) Average number of ensemble members identified for each case and (b) percent of cases with all 11 members or 7 or more members, as a function of forecast lead time.

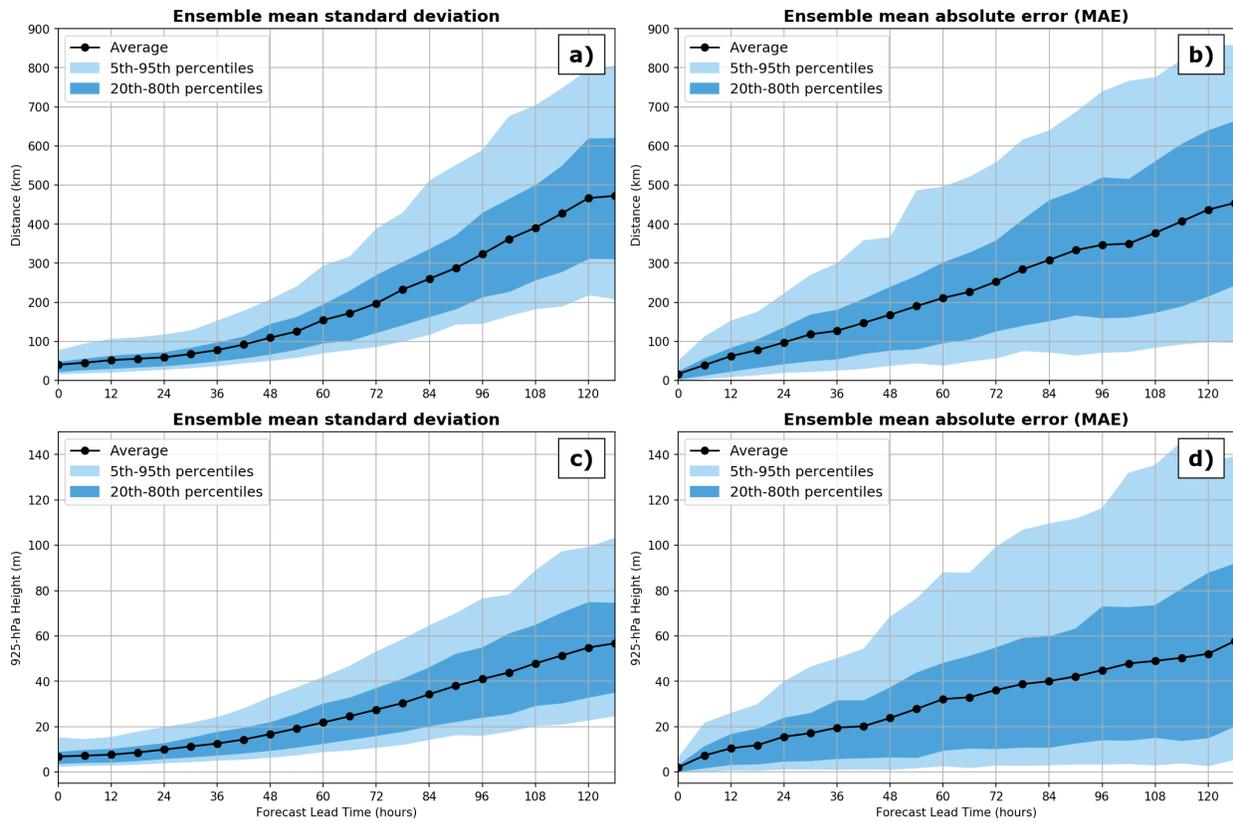


Figure 3.4: Ensemble forecast position (a) mean standard deviation in km and (b) mean absolute error in km, as a function of forecast lead time. Panels (c) and (d) are the same as (a) and (b) but for ensemble forecast intensity intensity. The 5th-95th percentile range is shaded in light blue, with the 20th-80th percentile range shaded in darker blue.

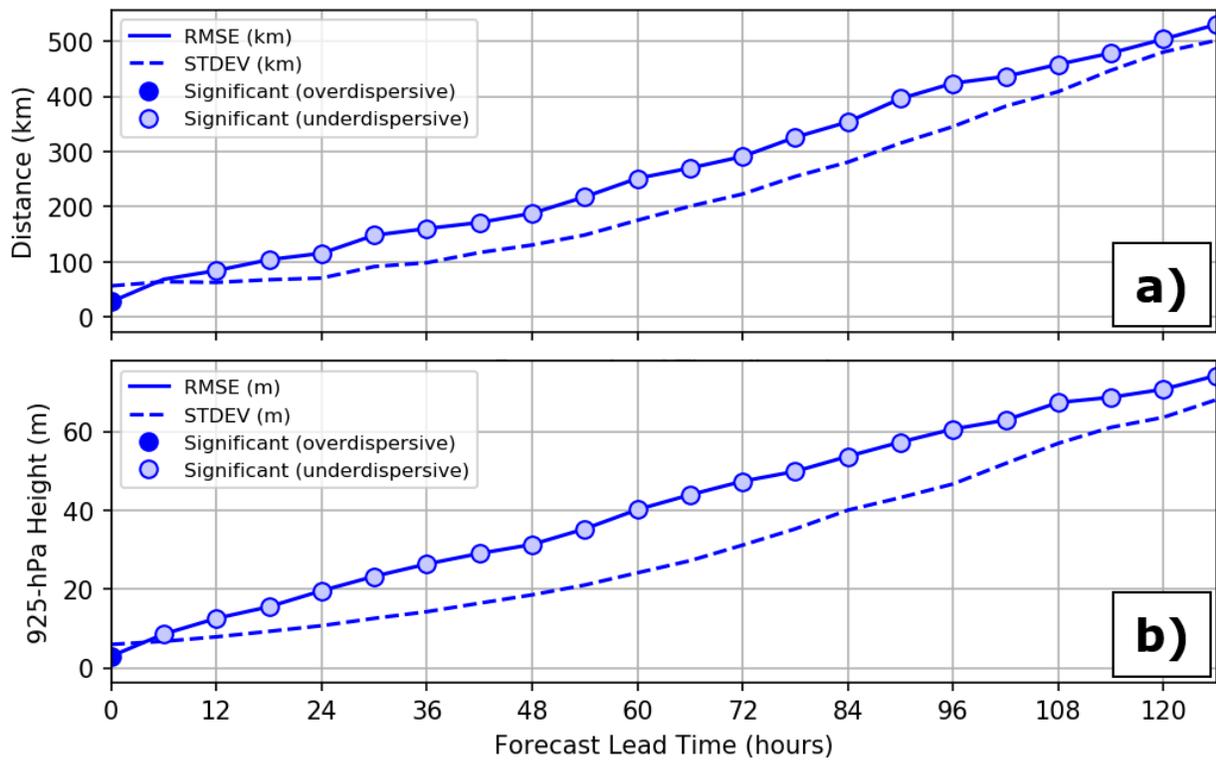


Figure 3.5: Bias-corrected RMSE (solid) and STDEV (dashed) for (a) position and (b) intensity, as a function of forecast lead time. Dots denote where the difference between the RMSE and STDEV is statistically significant at the 95% confidence level, with dark blue dots indicating an overdispersive bias and light blue dots indicating an underdispersive bias.

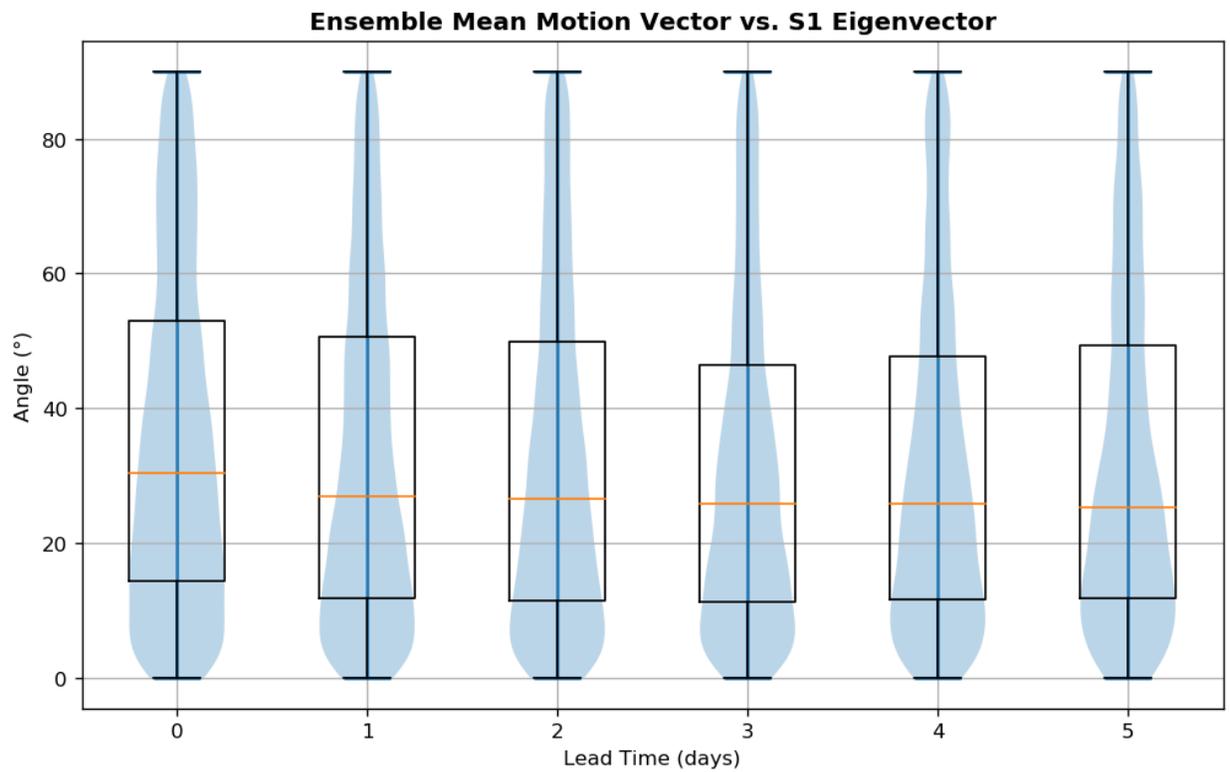


Figure 3.6: Violin plot distribution of the angle between ensemble mean motion vector and the S1 eigenvector representing the direction of greatest ensemble spread, binned by forecast lead time. Orange lines represent the median, and the box represents the interquartile range (IQR).

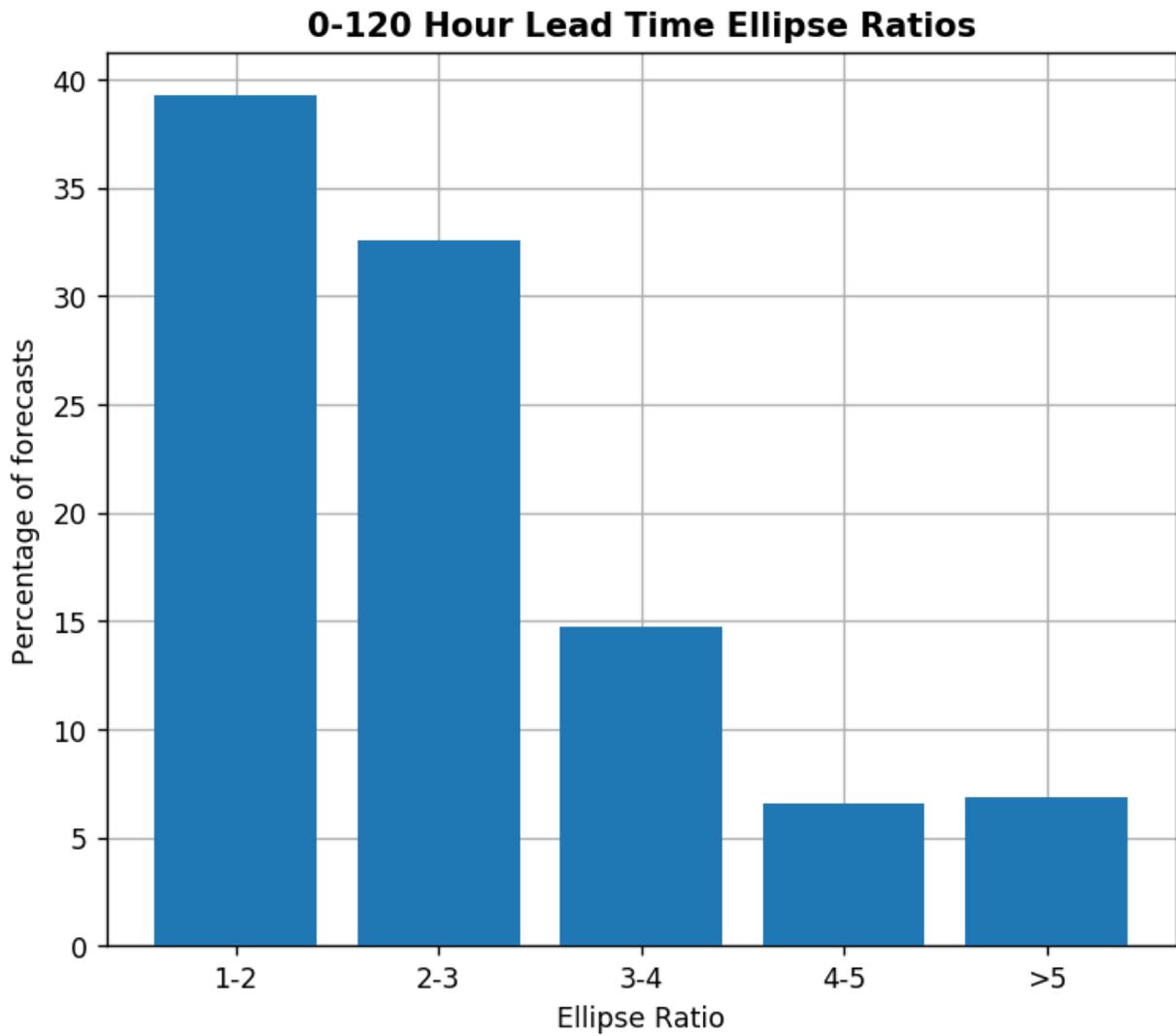


Figure 3.7: A histogram of the percentage of all forecasts from 0 through 120 hour lead times in each ellipse ratio category.

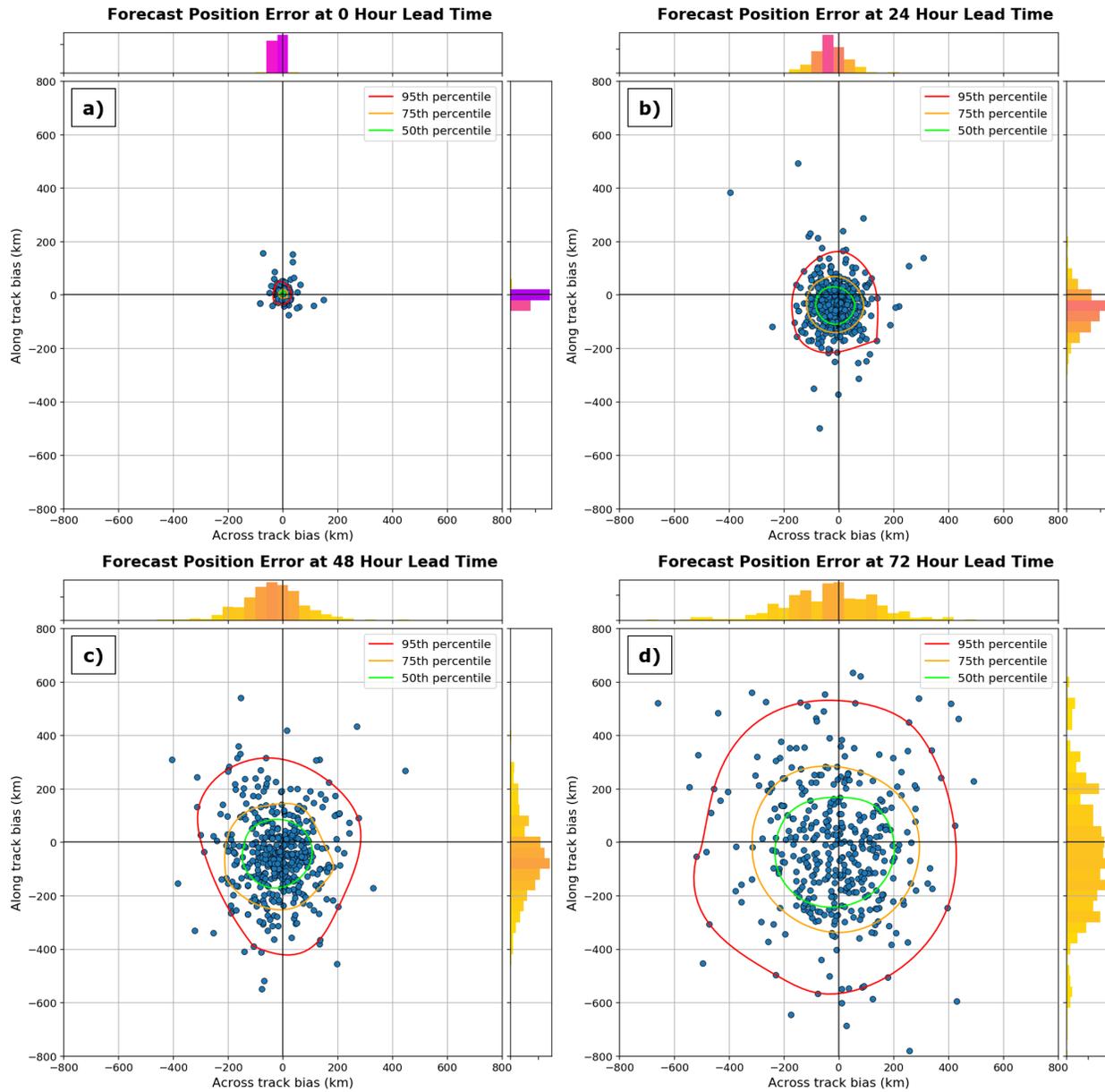


Figure 3.8: Ensemble mean forecast position error relative to the verification CFSR storm motion vector oriented north, in kilometers, valid at (a) 0 hour, (b) 24 hour, (c) 48 hour, and (d) 72 hour forecast lead times. Convex hulls representing the 50th, 75th and 95th percentiles are contoured, with 40 km binned histograms for along and across track error shaded by the frequency of cases in each bin, with higher values in red and purple.

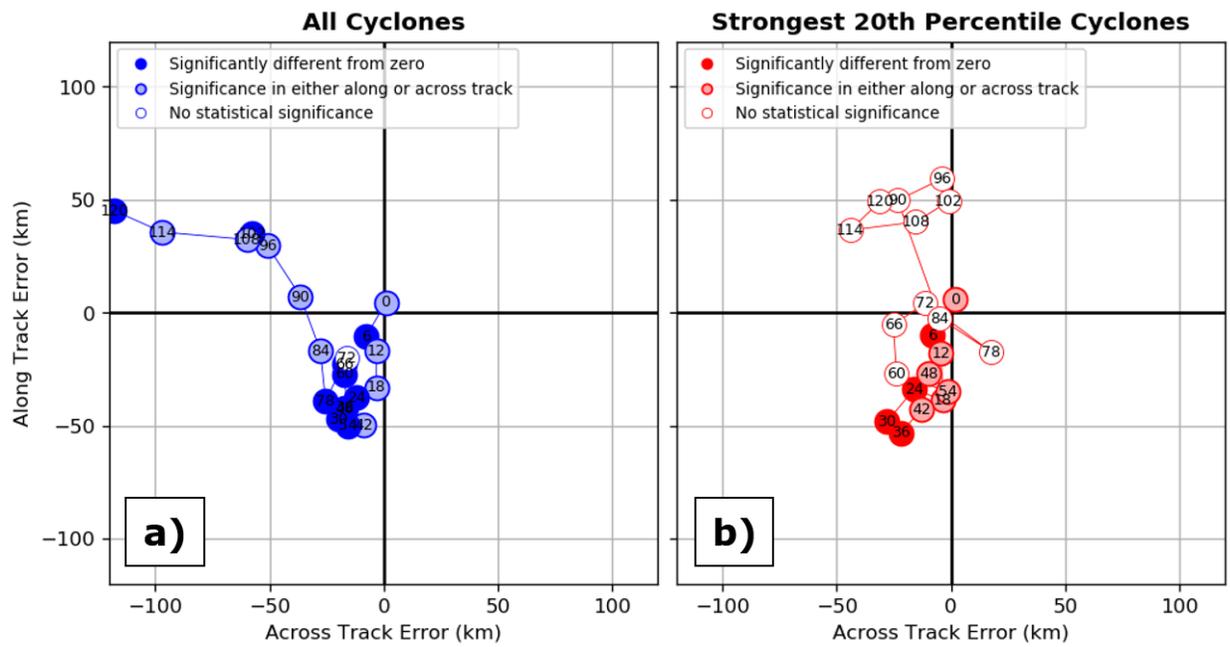


Figure 3.9: Ensemble mean forecast position error relative to the verification CFSR storm motion vector oriented north, in kilometers, for (a) all cyclones and (b) Strongest 20% percentile of cyclones. Numbers represent the forecast hour. Dark filled shading indicates that both along and across track biases are statistically significantly different from zero at the 95% confidence level, while lighter filled shading indicates either along or across track biases are significantly different from zero at the 95% confidence level.

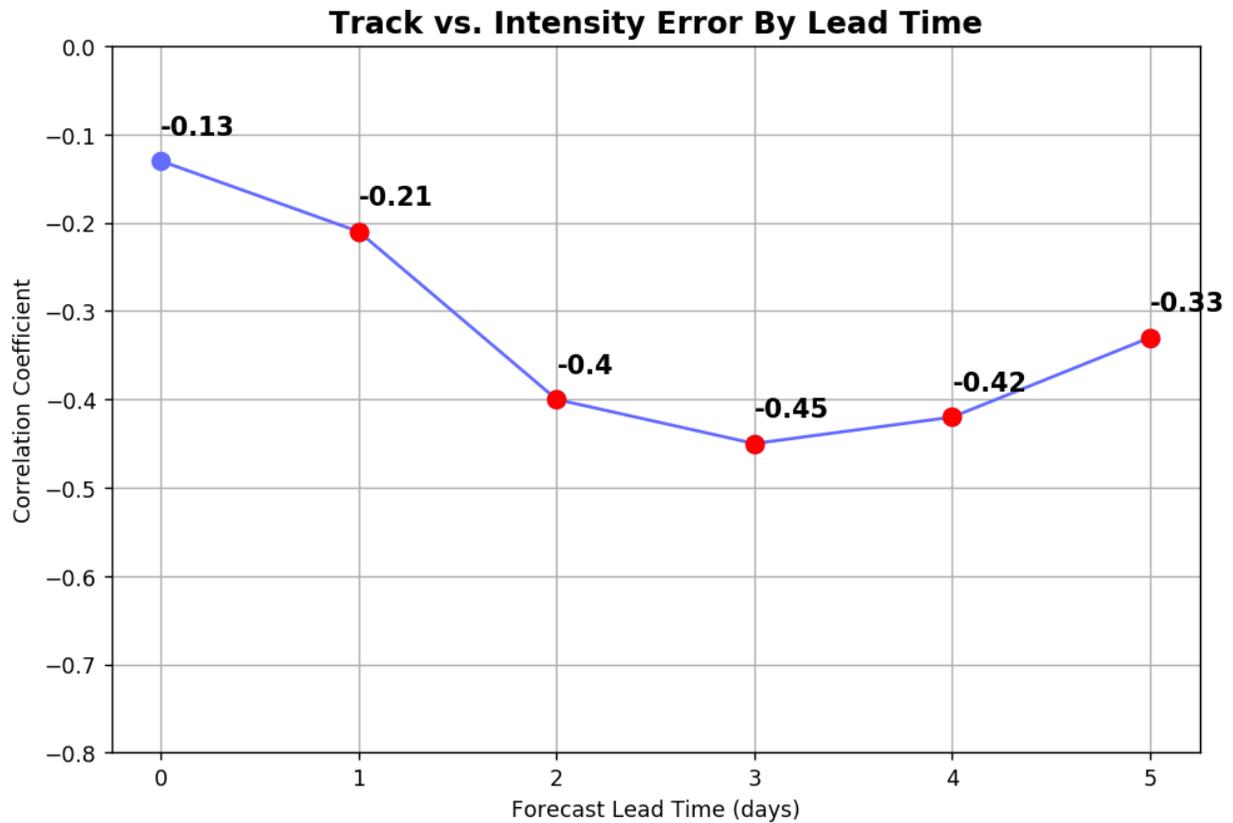


Figure 3.10: Pearson correlation coefficient between across-track error vs. intensity error as a function of forecast lead time. Dots in red indicate where the correlation coefficient is statistically significant at the 95% confidence level.

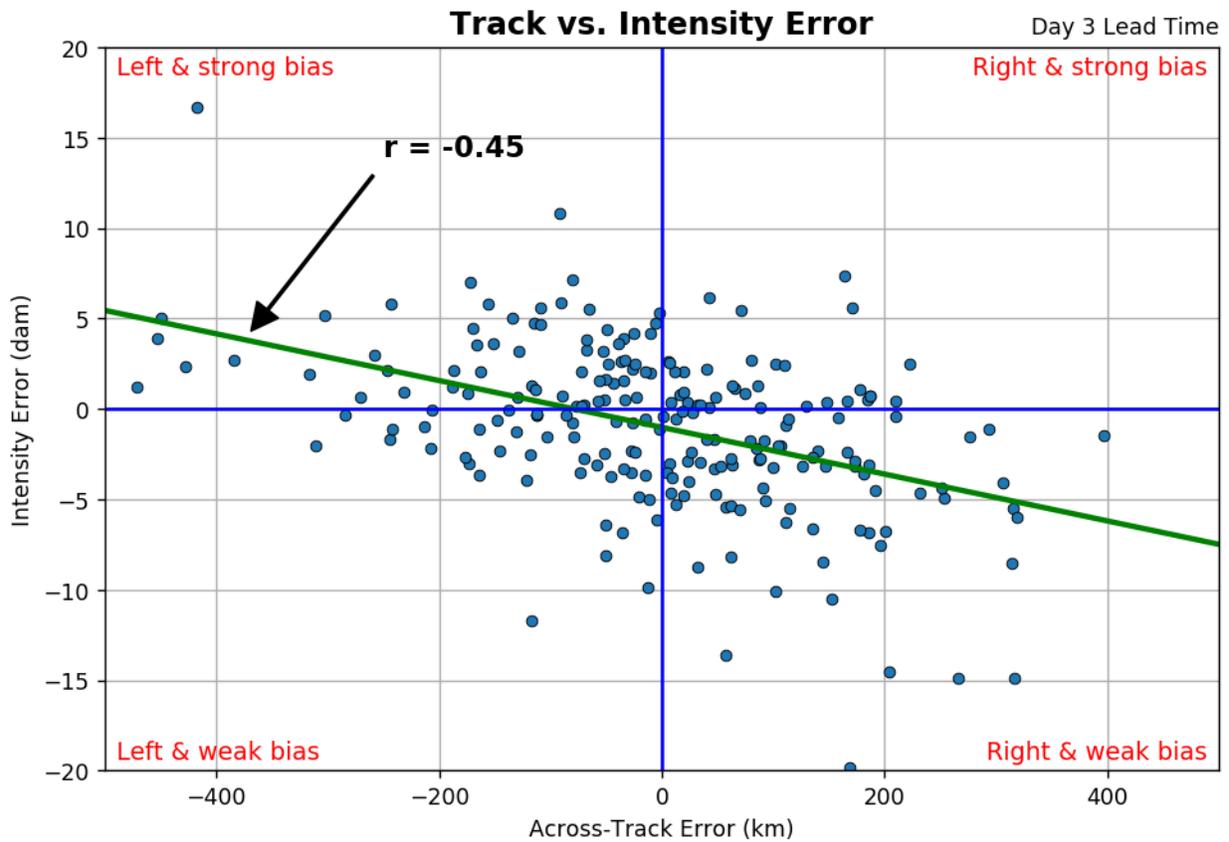


Figure 3.11: Across-track error (km) vs. intensity error (dam) for day 3 lead time forecasts. The linear regression between the two variables and the r-value of the Pearson correlation coefficient are labeled on the plot.

## 4. Composites

The previous chapter examined a climatology of verification cyclone tracks from the CFSR, obtained using the tracking algorithm referenced in section 2.3, as well as climatologies of ensemble forecast position spread, finding an underdispersive position bias, in addition to variability predominantly in the along-track direction with a smaller frequency of predominantly across-track variability. While no systematic right of track bias was found, there is a negative correlation between across-track bias and intensity bias, peaking at day 3 lead time (section 3.3). These results motivate this chapter's goals to further examine these biases and position variability groups from a synoptic perspective to assess the spatial differences in the flow configurations, compared to the features typically associated with Northeast U.S. snowstorms (Kocin and Uccellini 2004b).

To examine hypotheses 3-5 as stated in section 1.3, suggesting that there are synoptic-to-large scale flow configurations that are associated with systematic biases in the forecasts of Northeast ECs, this chapter considers composites of cyclones categorized by bias groups. These composites are created using CFSR verification centered at the time of peak impact, with cases selected based on their bias or position spread evident at day 3 lead time. The methodology for this chapter is further discussed in section 2.5. Section 4.1.1 considers left of track vs. right of track biases at day 3 lead time, and section 4.1.2 considers weak vs. strong intensity biases at day 3 lead time. Motivated by the presence of a relatively small subset of cases exhibiting predominantly across-track variability, section 4.2.1 considers small vs. large ensemble forecast position spread in the across track direction, and section 4.2.2 considers small vs. large ensemble forecast position spread in the along track direction.

## 4.1 Ensemble Bias Composites

### 4.1.1 Left vs. Right of Track Bias

Figure 4.1 shows the CFSR composite differences between the day 3 lead time right of track (ROT) minus the left of track (LOT) biases for MSLP (Fig. 4.1a-c), 500-hPa geopotential height anomaly relative to climatology (Fig. 4.1d-f), and 300-hPa wind (Fig. 4.1g-i) at  $t_0 - 48$  h,  $t_0 - 24$  h and  $t_0$  lead times. Focusing on the MSLP composite at  $t_0$  (Fig. 4.1c), there is a region of significantly higher MSLP from the Rockies into the South Plains and the Gulf of Mexico in the ROT composite compared to the LOT composite. This difference originates over the Northwestern U.S. at  $t_0 - 48$  h (Fig. 4.1a) and propagates equatorward while encompassing most of the continental U.S. by  $t_0 - 24$  h (Fig. 4.1b). The ROT composite additionally has a region of significantly higher MSLP over the Northeast U.S. 24 hours preceding the closest approach of the cyclone to the benchmark, compared to the LOT composite.

The 500-hPa height anomaly field at the 0 hour lag time (Fig. 4.1f) shows a more meridional flow over North America in the ROT cases compared to the LOT cases. The difference in the flow is highlighted by a significantly deeper trough in the East Coast and a higher amplitude ridge in the West Coast (Fig. 4.1f). This difference can be traced back to  $t_0 - 48$  h with a significantly larger positive height anomaly over the Northwest U.S. (Fig. 4.1d), persisting over the West Coast through  $t_0$  while the downstream trough over the East Coast is deeper in the ROT cases from  $t_0 - 24$  h to  $t_0 + 6$  h (the latter not shown). There is an additional region of positive height anomaly downstream of the East Coast trough that is significantly larger in the ROT than LOT cases from  $t_0 - 24$  h (Fig. 4.1e) to  $t_0$ .

The evolution of the evolution of the 300-hPa composite difference with lag time (Fig. 4.1g-i) suggests that differences in the wind are associated with smaller scale variations compared to the other fields analyzed. At  $t_0 - 48$  h, there are few regions of significant differences aside from the larger magnitude in wind with the ROT cases relative to the LOT cases over the core of the jet stream. At  $t_0$ , a coupled jet streak signature is apparent in

the mean of both composites, with the ROT composite mean exhibiting a stronger and equatorward displaced subtropical jet streak, and a stronger downstream polar jet streak.

The composite differences representing higher MSLP in the ROT cases progress equatorward, and the differences indicate more meridional flow over North America in these cases, suggesting that northerly flow in the Central U.S. may be more enhanced in the ROT cases. Given that these cases occur during the cold season, there may be a higher frequency of cold air outbreaks (CAOs) over the central United States leading up to and during ROT cases. CAOs over the U.S. are generally associated with positive height anomalies over the western U.S. and negative height anomalies to the east, and positive MSLP anomalies over the western U.S. two days prior to the peak of the CAO (Konrad 1996). The composite 500-hPa height and MSLP fields in Figure 4.1 exhibit similarities to the aforementioned features associated with CAOs, with higher amplitude flow over North America implying an anomalous northerly wind with a more poleward source of air parcel trajectories and likely a higher probability of a colder source region. The higher MSLP in the Northeast U.S. at  $t_0 - 24$  h is also likely indicative of stronger antecedent anticyclones over the Northeast U.S., associated with cold air damming and thus retaining low-tropospheric cold air while enhancing low-tropospheric baroclinicity and perhaps aiding in cyclogenesis (Kocin and Uccellini 2004b).

#### *4.1.2 Weak vs. Strong Bias*

Hypothesis 2 expects a negative correlation between the magnitude of the across-track and intensity biases. The climatological analysis found a negative correlation between these two biases in section 3.3, thus verifying hypothesis 2. This section will test hypothesis 3, the composite difference between WB and SB cases will exhibit large-scale similarities to the composite difference between the ROT and LOT cases. Specifically, WB cases are expected to exhibit a more meridional flow over North America than SB cases.

Figure 4.2 shows the CFSR composite differences between the day 3 lead time weak

(WB) minus strong (SB) biases for MSLP (Fig. 4.2a-c), 500-hPa geopotential height anomaly relative to climatology (Fig. 4.2d-e), and 300-hPa wind (Fig. 4.2c-i) at  $t_0 - 48$  h,  $t_0 - 24$  h and  $t_0$ . The MSLP composite difference at  $t_0$  (Fig. 4.2c) shows higher MSLP associated with WB than SB cases in the Southeast U.S. and Gulf of Mexico regions. Compared with the corresponding composite difference for ROT minus LOT cases (Fig. 4.1c), this positive MSLP difference is confined to the South Plains and the Gulf of Mexico. Additionally, there is a region where MSLP is lower in WB than SB cases over the North Plains and southern Manitoba. The less expansive statistically significantly positive MSLP composite difference is similarly evident at  $t_0 - 24$  h and  $t_0 - 48$  h.

The MSLP values within and surrounding the composite MSLP minimum at the benchmark at  $t_0$  are lower for the WB than SB cases. This implies that WB cases include more intense cyclones than SB cases. Considering that the composite is not storm centered and there is position variability among the cases within the composite, this does not necessarily mean that WB cases are stronger than SB cases. However, an evaluation of the 925-hPa cyclone area-averaged vorticity for both the LOT vs. ROT cases and the SB vs. SB cases (Fig. 4.3) does suggest there are cyclone intensity differences. A comparison of the cyclone intensity suggests that both WB and ROT cases are on average slightly stronger than SB and ROT cases, respectively. The differences are statistically significant at the 95% confidence level using a difference of means test.

The WB composite is associated with a lower 500-hPa height anomaly within the East Coast trough than SB cases at  $t_0$  (Fig. 4.2f). Unlike the ROT minus LOT composite difference (Fig. 4.1f), however, the ridge over western North America is not significantly more amplified in the WB composite than the SB composite at  $t_0$ . The composite difference patterns over North America propagate downstream with time, with a positive difference in height anomaly in the WB cases compared to the SB cases over the Northwest U.S. from  $t_0 - 84$  h (not shown) through  $t_0 - 48$  h (Fig. 4.2d), and a negative difference in height anomaly with the WB cases compared to SB cases associated with the East Coast trough from  $t_0 -$

24 h through  $t_0$  (Fig. 4.2e,f). While not shown, there is a subsequent positive difference in height anomaly in WB cases compared to SB cases with the downstream ridge from  $t_0 + 12$  h through  $t_0 + 36$  h.

The 300-hPa wind composite at  $t_0$  (Fig. 4.2i) shows that the polar jet streak downstream of the cyclone is displaced poleward, and the subtropical jet streak upstream of the cyclone is displaced equatorward, in the WB composite relative to the SB composite. While there is no difference in wind speed in the polar jet streak, the subtropical jet streak is stronger in the WB than SB composite. This difference is similarly evident at  $t_0 - 24$  h (Fig. 4.2h). This is different than the difference composite for ROT minus LOT cases (Fig. 4.1i), in which the downstream jet streak was associated with wind speed differences between both composites over the core of both subtropical and polar jet streaks.

Hypothesis 3 expected weak bias cases to be associated with a more meridional flow than strong bias cases. At a single lag time, there are no coinciding positive (negative) 500-hPa height anomaly differences in the western (eastern) U.S. However, the previously noted downstream propagation of the 500-hPa height anomaly differences between the WB and SB composites suggests that when considering the temporal evolution of the 500-hPa height anomaly field in the three day time period centered on the event, there is a more meridional flow over North America in the WB composite compared to the SB composite. Thus, hypothesis 3 is verified.

## 4.2 Ensemble Variability Composites

### 4.2.1 Across Track Ensemble Spread

This section examines hypothesis 4, which expects that cases with large across track variability may be associated with more downstream ridging and a weak downstream polar jet. Figure 4.4 shows the CFSR composite differences between the day 3 lead time of cases with small across-track variability (SV) minus large across-track variability (LV) for MSLP (Fig. 4.4a-c), 500-hPa geopotential height anomaly relative to climatology (Fig. 4.4d-f), and

300-hPa wind (Fig. 4.4g-i) at  $t_0 - 48$  h,  $t_0 - 24$  h and  $t_0$ . The MSLP composite at  $t_0$  (Fig. 4.4c) shows an expansive area of higher MSLP in the SV composite than in the LV composite from the central U.S. poleward into western Canada and the Northwest Territories. This difference is persistent in that region through  $t_0 - 84$  h (not shown). There is an additional equatorward and westward shift in the position of the Icelandic low at  $t_0 - 24$  h in the SV composite compared to the LV composite. At upper levels, no significant difference was found with the 300-hPa wind field near the coupled jet streak in the eastern U.S. at  $t_0$  (Fig. 4.4i). This is in contrast to prior composites of LOT vs. ROT biases and weak vs. strong biases, which found differences with either the subtropical or polar jet streaks.

Hypothesis 4 asserts that large across track variability is associated with a more amplified downstream ridge than small across track variability. Contrary to the expectation stated in hypothesis 4, there are no significant differences in the 500-hPa height field between SV and LV cases near the East Coast in the three days prior to  $t_0$  (Fig. 4.4d-f). Differences appear after  $t_0 + 12$  h (not shown) with lower heights over the Northeast U.S. and Southeast Canada. The composite does not suggest an amplified downstream ridge, and these results do not support hypothesis 4.

The persistence of the central U.S. signature of higher MSLP composite over several days is indicative of colder temperature anomalies preceding and during the SV cases. This suggestion is supported by the SV minus LV composite of 850-hPa temperature anomaly (Fig. 4.5). SV cases are associated with widespread colder temperature anomalies across the United States relative to the LV cases, with actual temperatures below climatological normals across the eastern U.S. into the Gulf of Mexico (not shown). An additional feature that is particularly notable in this composite is the persistence of higher 500-hPa height anomalies across most of northern and western Canada in the SV composite relative to the LV composite throughout the  $t_0 - 84$  h to  $t_0 + 24$  h range. When considered in combination with lower heights equatorward of these more positive height anomalies in Canada, this couplet implies a weaker westerly zonal wind over the western U.S. in the SV composite

relative to the WV composite. This is supported by the analysis of weaker composite 300-hPa wind for the SV case at  $t_0 - 48$  h.

The persistence of a western North American ridge anomaly throughout the 5-day analysis period in the SV composite resembles a positive Pacific-North America (PNA) teleconnection and a negative Eastern Pacific Oscillation (EPO) teleconnection pattern. A positive PNA pattern has been correlated with cold North American patterns (Harnik et al. 2016). Similarly, a negative EPO pattern was linked to the anomalously cold winter of 2013–2014 in the United States, which was associated with an upstream upper level block near Alaska (Marinaro et al. 2015). In addition, negative North Atlantic Oscillation (NAO) and negative Arctic Oscillation (AO) teleconnections have been linked to North American cold air outbreaks (Walsh et al. 2001). A working hypothesis is that the SV cases are associated with a more positive PNA and a more negative EPO that exhibit an amplified Alaskan ridge and cold central U.S. Similarly, the large scale patterns over the eastern U.S. suggest that SV cases are associated with a more negative NAO and AO than the LV cases, which would indicate a higher likelihood of cold air outbreaks over the United States. To test these hypotheses, teleconnection indices were compared between the SV and LV composites (Fig. 4.6), including the PNA, EPO, NAO, and AO. Both SV and LV composites are associated with a +PNA on average, but the SV cases have a larger +PNA index 1 and 2 days prior to the event. The signal is more prominent with the EPO, as LV cases associated with a slightly positive EPO preceding the event while SV cases exhibit a neutral EPO, with the difference significant from 1 to 4 days prior to the event. The SV cases also exhibit a more negative AO and a more positive NAO compared to the LV cases. The overall implication is that SV cases are on average associated with a higher likelihood of cold air outbreaks, with the preceding PNA more positive and EPO more negative over western North America than LV cases. However, a negative EPO pattern was not found for the average of SV cases.

### 4.2.2 Along Track Ensemble Spread

Figure 4.7 shows the CFSR composite differences between the day 3 lead time small along-track variability (SV-L) minus large along-track variability (LV-L) for MSLP (Fig. 4.7a-c), 500-hPa geopotential height anomaly relative to climatology (Fig. 4.7d-f), and 300-hPa wind (Fig. 4.7g-i) at  $t_0 - 48$  h,  $t_0 - 24$  h and  $t_0$ . In contrast with the across-track ensemble spread, which showed little composite differences near the East Coast and downstream of the cyclone, this composite difference plot shows a temporally and spatially consistent area of lower MSLP (Fig. 4.7a-c) and 500-hPa height anomaly (Fig. 4.7d-f) over the northwest Atlantic Ocean downstream of the cyclone from  $t_0 - 72$  h to  $t_0 + 24$  h (the latter not shown). This difference in the Atlantic Ocean suggests that LV-L cases are associated with a stronger downstream ridge and higher MSLP downstream of the cyclone than LV-L cases. The SV-L cases are also associated with higher 500-hPa heights over much of Canada in the period preceding the event (Fig. 4.7d-f). Both of these areas suggest that LV-L cases are associated with a more meridional flow across the United States, and lower geopotential heights across central Canada, than the SV-L cases. This difference pattern is consistent with the SV and LV composite difference, but is displaced eastward compared to Figure 4.4d-f, with large along-track ensemble spread being associated with higher 500-hPa heights upstream over the U.S. West Coast.

Hypothesis 5 asserts a stronger downstream jet streak with LV-L cases relative to SV-L cases. Given the composite differences in the 500-hPa field, it would be expected that there would be a difference in the 300-hPa wind field as well, specifically regarding the latitude of the jet stream. Fig. 4.7i shows that at  $t_0$ , the downstream polar jet streak is both stronger and displaced poleward in the LV-L composite than the SV-L composite, and these differences are statistically significant. This result is consistent with the stronger 500-hPa geopotential height gradient in the LV-L composite relative to the SV-L composite downstream of the cyclone. Thus, hypothesis 5 is considered to be true.

### 4.3 Composites Discussion

Section 4.1.1 analyzed the composite differences between left of track (LOT) and right of track (ROT) day 3 lead time bias cases. It was found that the ROT composite is associated with a more meridional flow over North America than the ROT composite, along with higher MSLP over the Central and Western U.S. in addition to the Northeast U.S. preceding the event. This signature is likely indicative of a higher frequency of cold air outbreaks over the Central and Eastern U.S. associated with ROT cases. At 300-hPa, the ROT composite mean exhibits a stronger and equatorward displaced subtropical jet streak with a stronger downstream polar jet streak. This signature is indicative of a strong ageostrophic wind component in the equatorward entrance region of the jet streak downstream of the cyclone. The jet signature suggests stronger tropospheric upward vertical motion with the ROT than LOT cases, which may be indicative of heavier precipitation and stronger latent heat release downstream of the cyclone to further amplify the downstream ridge. This hypothesis has not been tested in this thesis and may be explored further in future work.

Section 4.1.2 analyzed the composite differences between weak (WB) and strong (SB) day 3 lead time bias cases. The difference field suggested wave packet-like downstream propagation of the 500-hPa height differences, with significant differences originating from a higher amplitude Western U.S. ridge in the WB composite and translating downstream to a deeper East U.S. trough and a subsequent deeper ridge downstream of the cyclone. Prior research has noted similarities in downstream error propagation to the long distance propagation of a Rossby wave packet (Zheng et al. 2013; Torn 2017). However, it is unclear if the signal shown here is in fact associated with a wave packet. The 300-hPa composite mean shows a coupled jet structure in the East U.S., with the WB composite showing the downstream polar jet streak displaced poleward while the subtropical jet streak is displaced equatorward. This may suggest that troughs associated with WB cases have a higher overall amplitude than those associated with SB cases, although this hypothesis has not been tested in this thesis.

When comparing ensemble variability in the across-track direction (section 4.2.1), cases with small across-track variability (SV) were found to be associated with higher MSLP and 500-hPa heights over much of central and western Canada compared to large across-track variability (LV) cases in the 2–3 days leading up to the event. There was a particularly strong signature of higher MSLP over the central U.S. associated with SV cases, which was associated with more negative 850-hPa temperature anomalies across the eastern two thirds of the United States. Hypothesis 4, asserting that large across track variability is associated with a more amplified downstream ridge and a weaker downstream polar jet streak, was not supported by the composite difference analysis. However, further testing of this hypothesis such as feature ridge axis tracking would likely be required to assert this with high confidence.

Analyzing ensemble variability in the along-track direction (section 4.2.2) showed that cases with large along-track variability (SV-L) are associated with a stronger and poleward shifted jet streak, as well as stronger downstream ridging over the western Atlantic Ocean, than those with small along-track variability (LV-L). The stronger downstream jet streak suggests that LV-L cases may be stronger and move faster than SV-L cases, with the uncertainty in the magnitude of forcing for propagation a potential cause of the large variability, which may be explored more in future work.

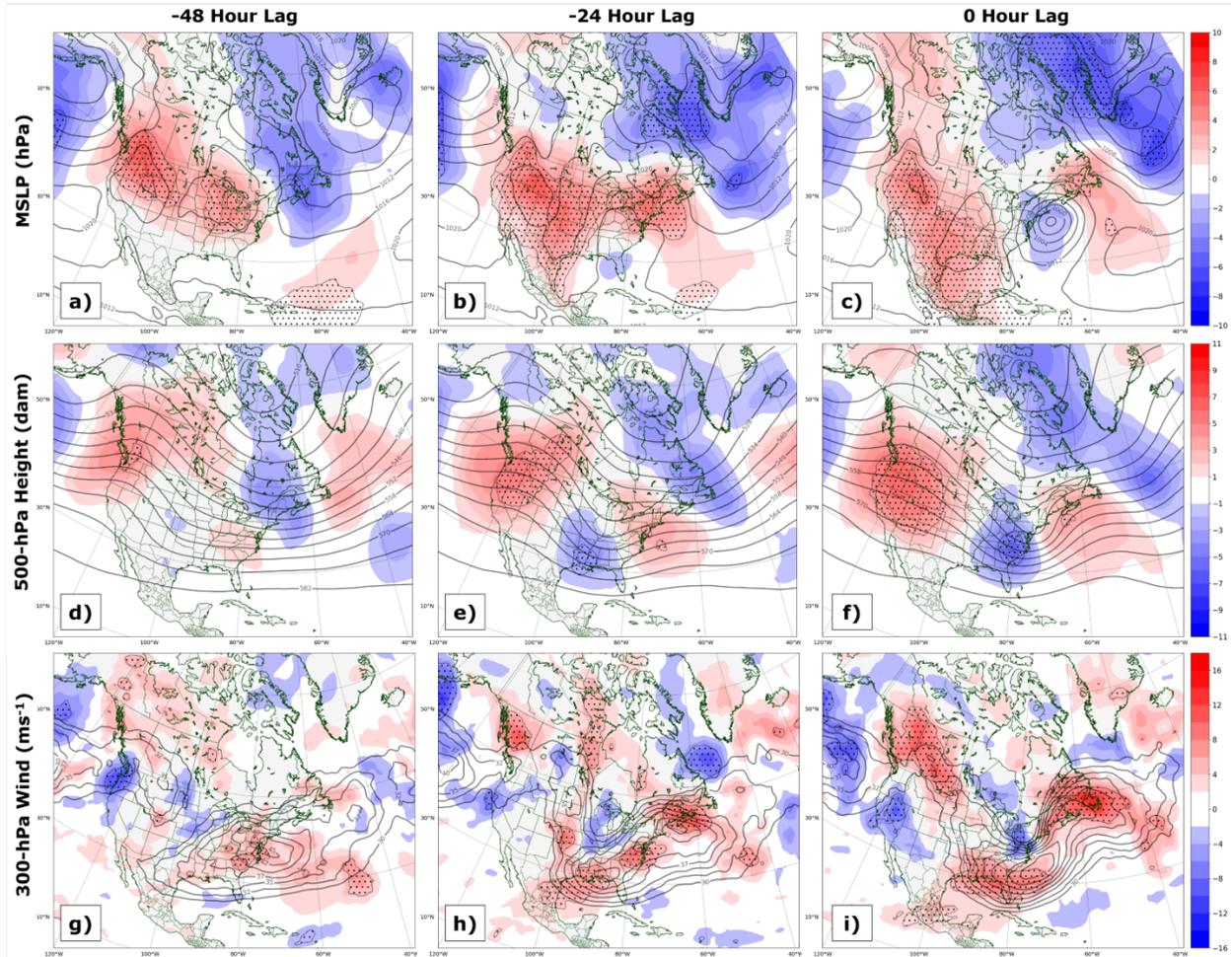


Figure 4.1: Difference plots of ROT minus LOT bias composites for (a,b,c) MSLP (hPa), (d,e,f) 500-hPa height anomaly (dam), and (g,h,i) 300-hPa wind ( $ms^{-1}$ ). The difference plots are presented for (a,d,g) -48 hour lag time, (b,e,h) -24 hour lag, and (c,f,i) 0 hour lag, with 0-hour lag composite centered at the time of peak impact. For each panel, the difference field is shaded, with warm colors indicating the ROT composite has a higher value than the LOT composite. The mean value of both ROT and LOT composites is contoured in black lines, and black stippling indicates statistical significance at the 95% confidence level.

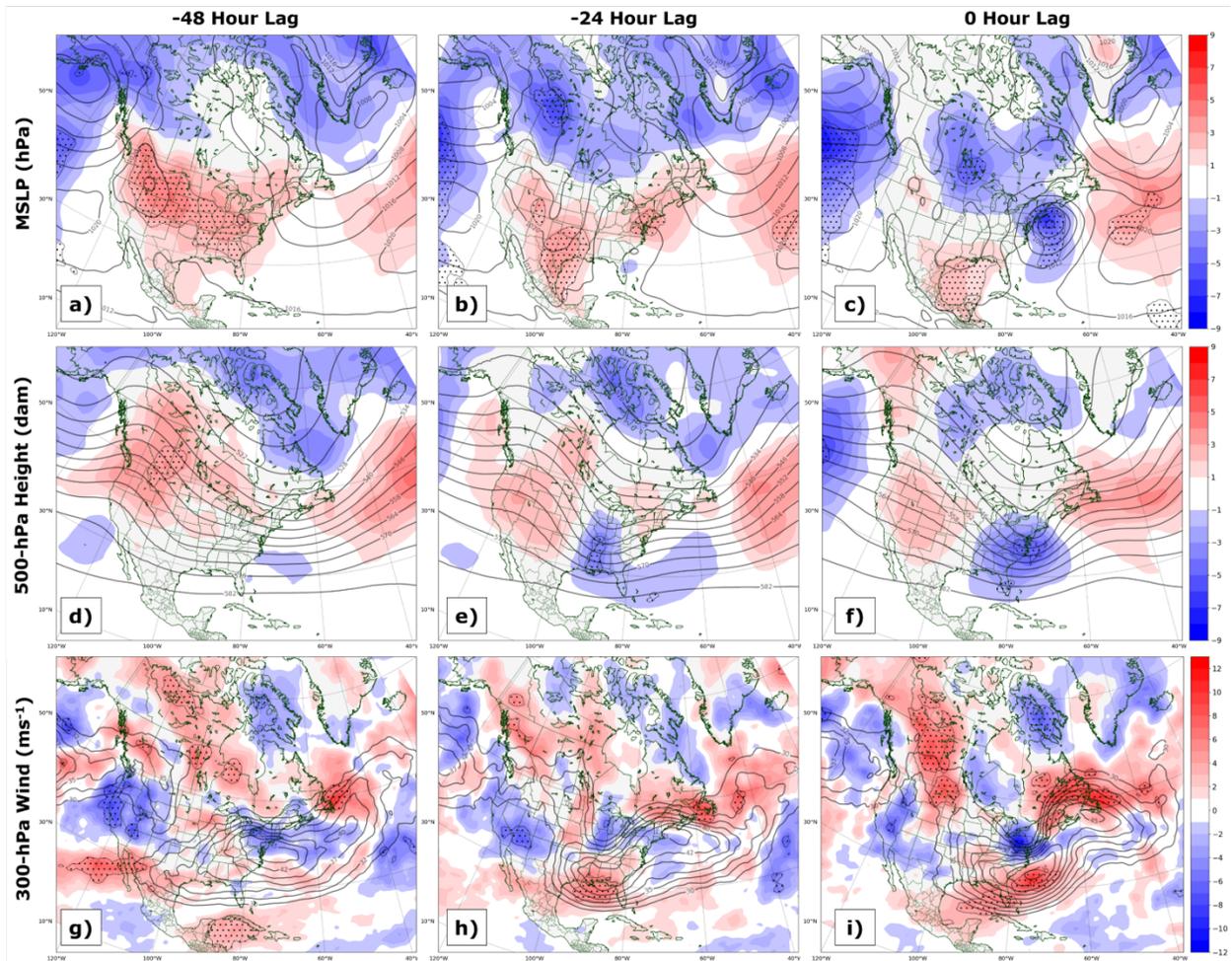


Figure 4.2: Same as Fig. 4.1, but for composite difference of WB minus SB cases.

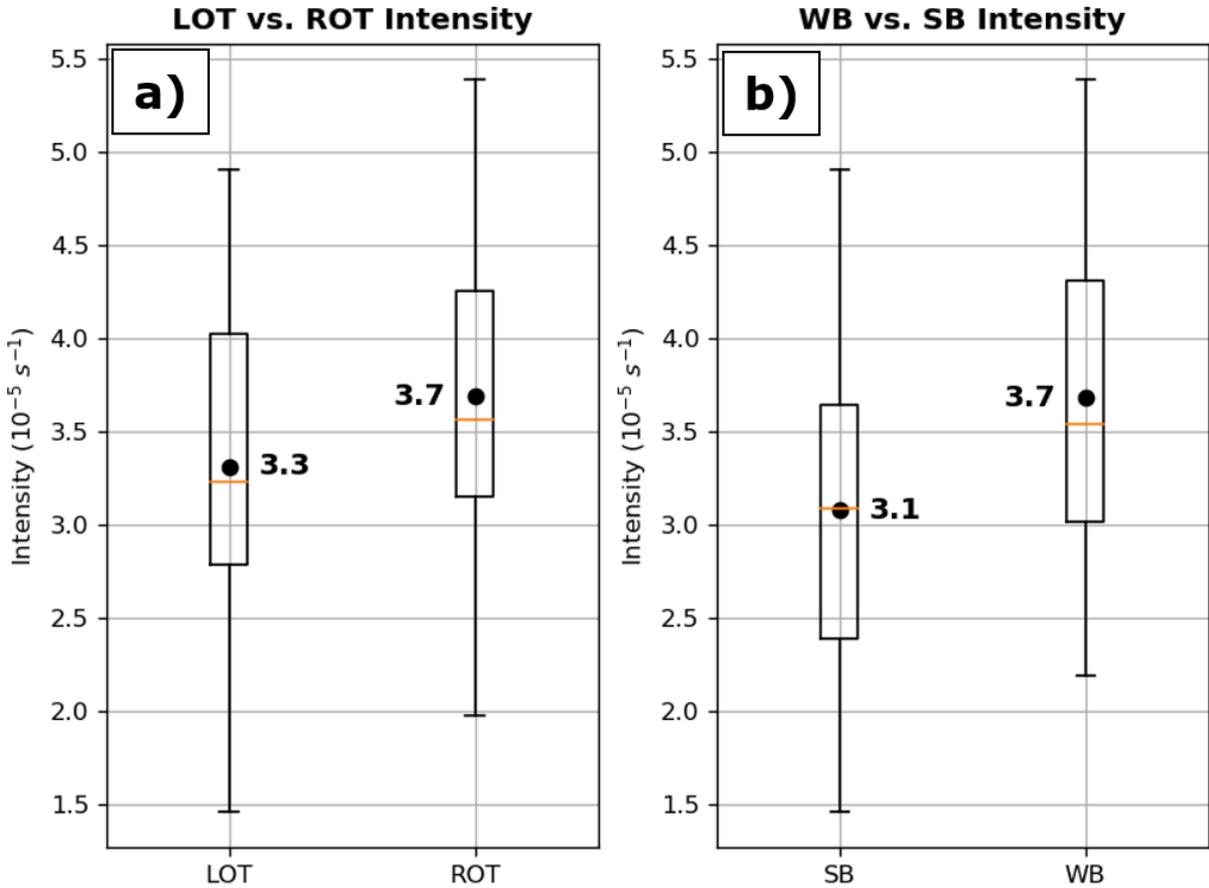


Figure 4.3: Distribution of CFSR observed 925-hPa area-averaged vorticity for (a) LOT (n=48) vs. ROT (n=45) cases and (b) WB (n=44) vs. SB (n=63) cases. Orange lines represent the median, black dots represent the average, with the box extending from the lower to upper quartile values. Whiskers denote the range of the data.

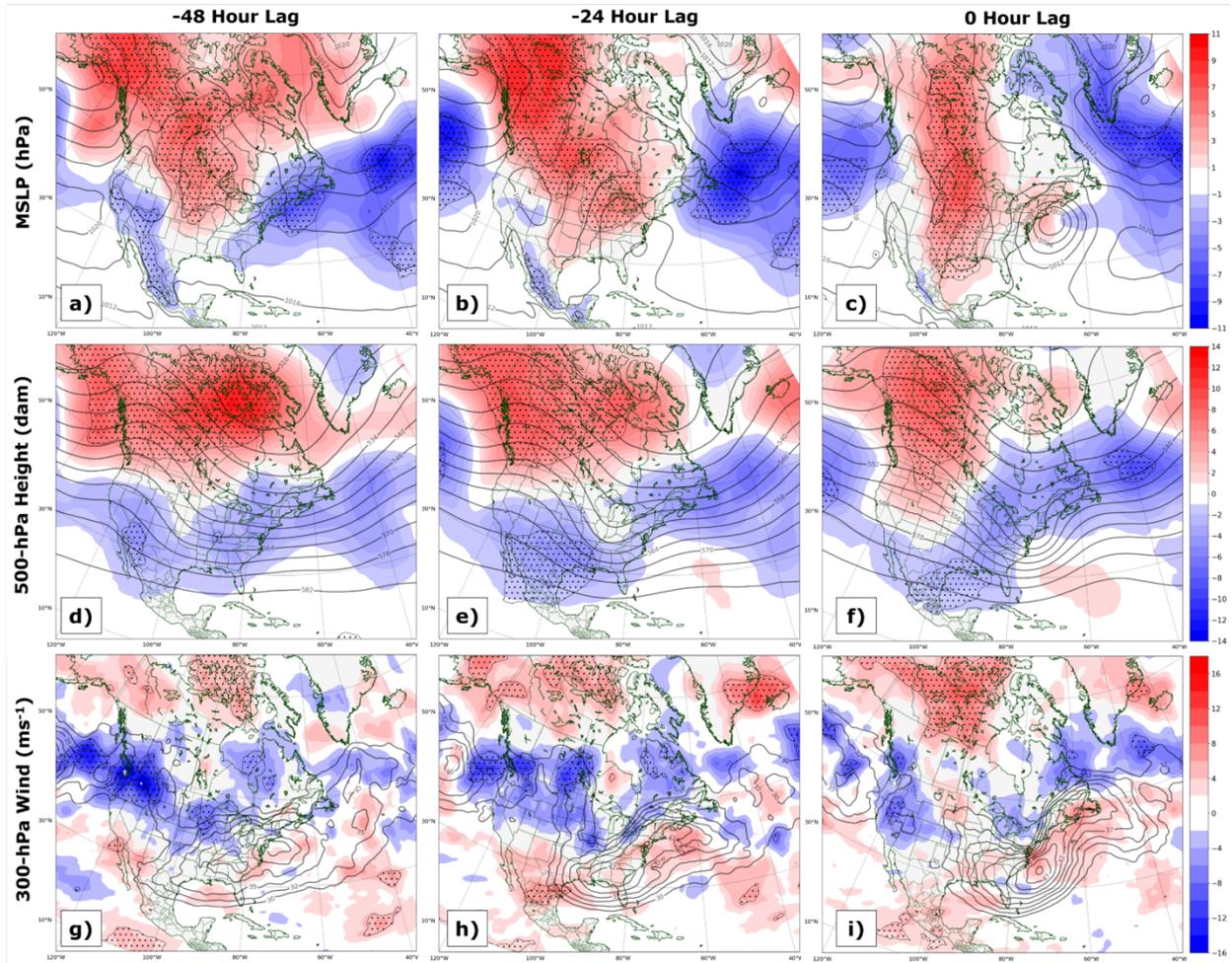


Figure 4.4: Same as Fig. 4.1, but for composite difference of SV minus LV across-track variability cases.

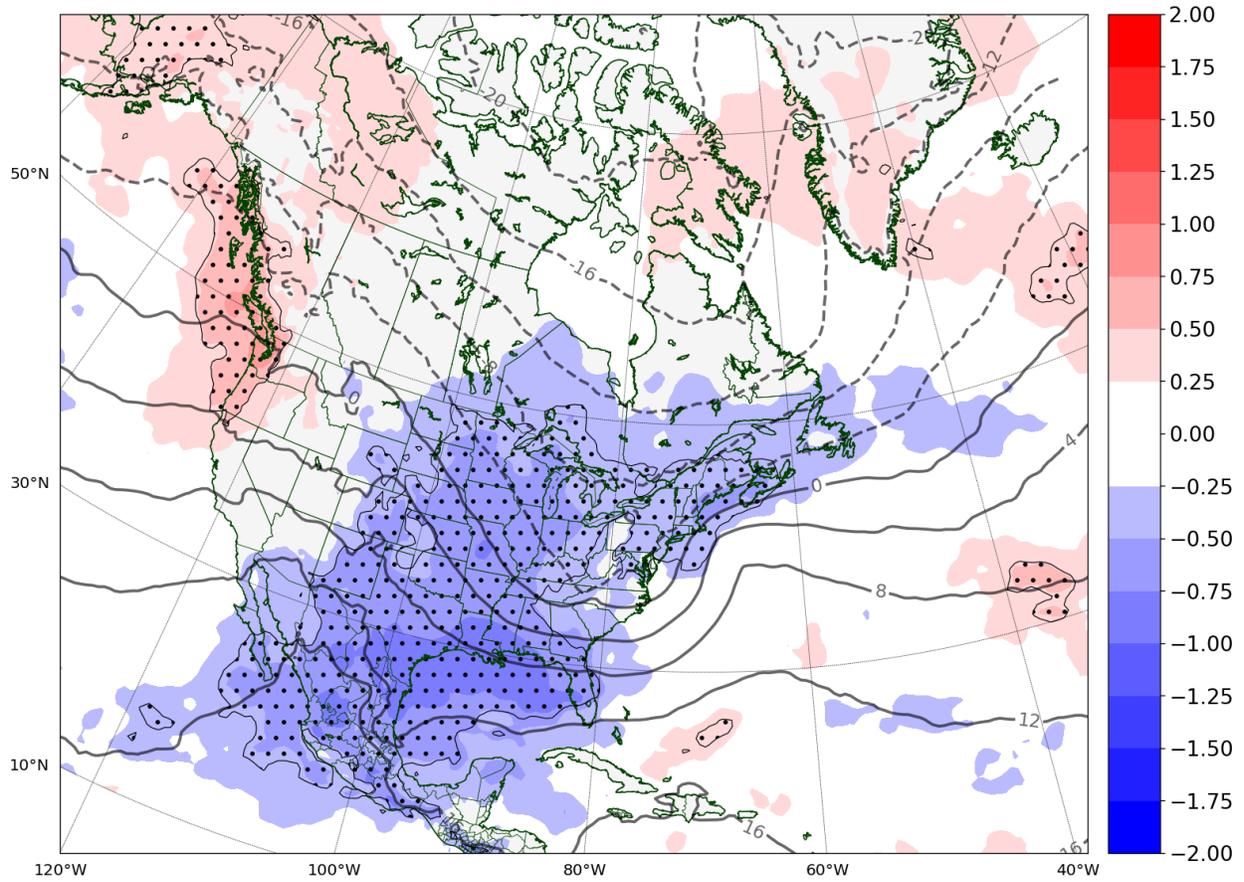


Figure 4.5: Composite difference of SV minus LV across-track variability cases for 850-hPa normalized temperature anomaly ( $\sigma$ ) at 0 hour lag time.

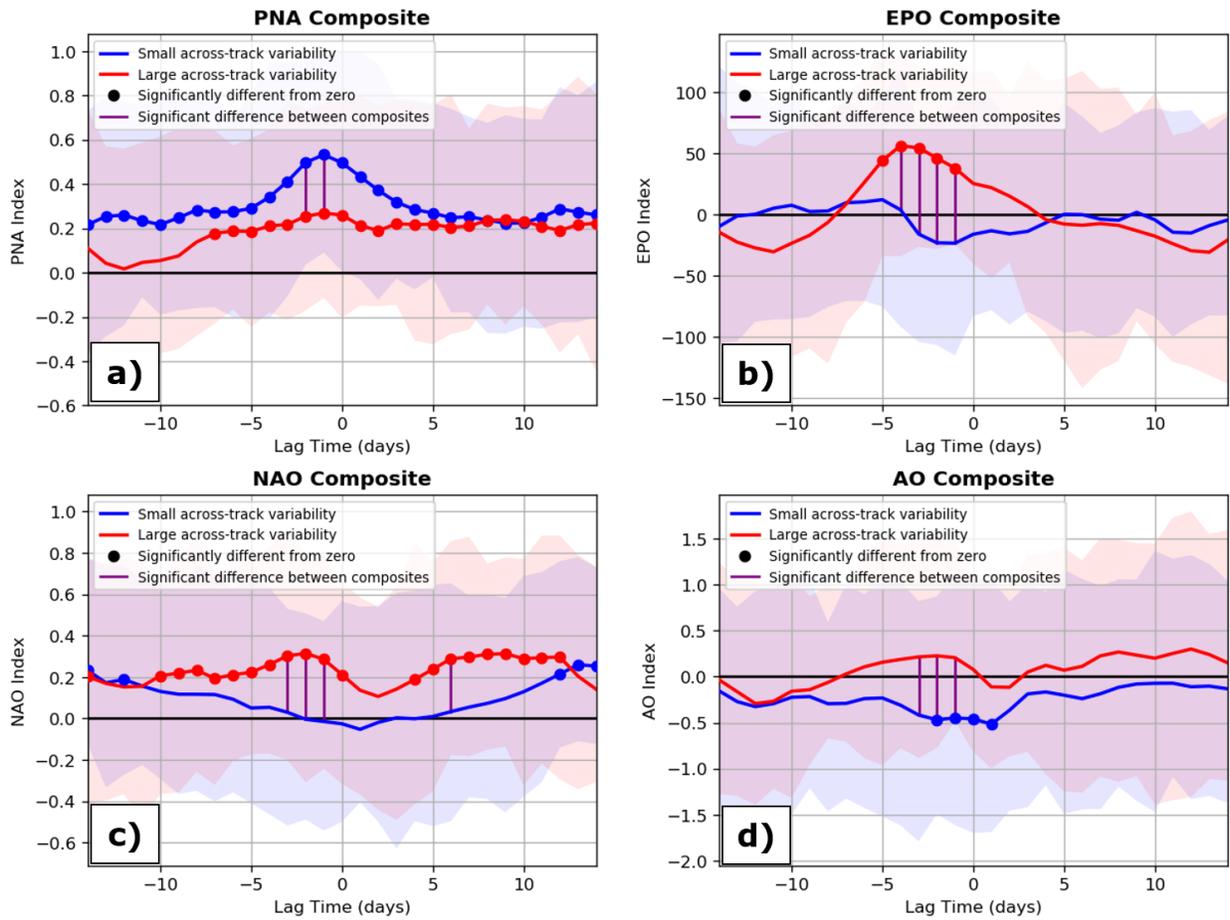


Figure 4.6: Teleconnection comparisons of (a) PNA, (b) EPO, (c) NAO, and (d) AO, for -14 to 14 day lag times centered at the time of peak impact. Blue lines denote small across-track variability composite with light blue shading denoting the Interquartile range (IQR), with the same for red denoting large across-track variability. Statistical significance is tested at the 95% confidence level.

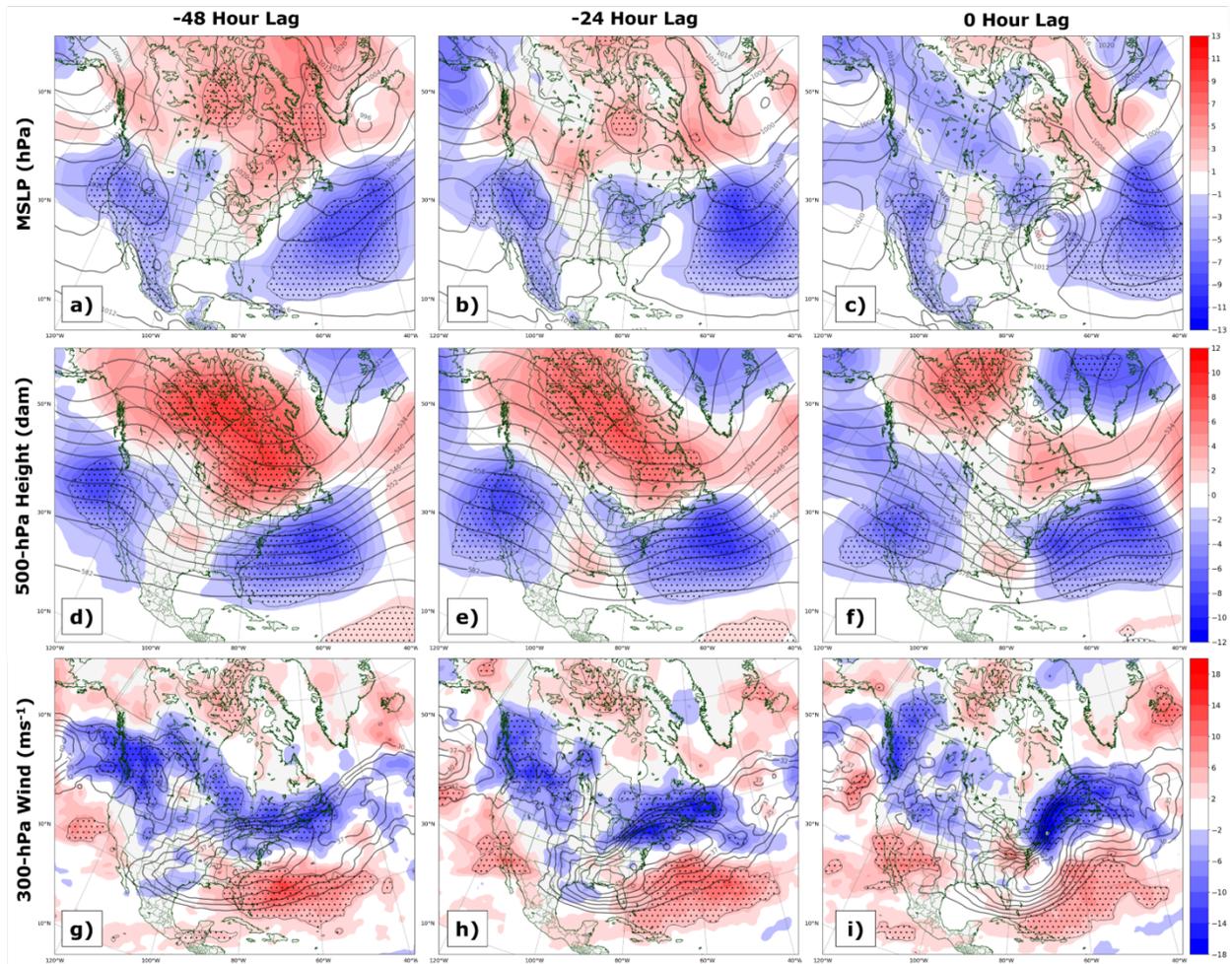


Figure 4.7: Same as Fig. 4.1, but for composite difference of SV minus LV along-track variability cases.

## 5. Conclusions and Future Work

### 5.1 Conclusions

Motivated by the need to increase situational awareness of high impact weather, this thesis examines the errors and biases in forecasts of cool-season Northeast ECs. The overarching goals of this thesis are to (a) examine whether a systematic right of track bias exists, (b) analyze the climatology of track and intensity errors from a GEFS Reforecast ensemble perspective for the Northeast U.S., and (c) compare synoptic patterns for cases that exhibit various ensemble mean errors and biases. First, objectively identified cyclone tracks were obtained for the verification CFSR data and forecast GEFS data using a tracking algorithm created for this research centered on Northeast ECs.

The available GEFS ensemble forecast data was used to assess the hypotheses put forth in section 1.3. A caveat to these results is that a systematic underdispersive bias was found for both position and intensity forecasts from hour 6 lead time onward, consistent with prior similar analyses on the operational GEFS (Froude 2010) and the GEFS Reforecast (Bentley 2018). Analyzing the spatial distributions and temporal trends in track and intensity errors for Northeast ECs revealed the following conclusions:

- There is no systematic right of track bias, both for all cyclones analyzed and for the most intense 20th percentile of cyclones (Fig. 3.9). This verifies hypothesis 1 regarding the existence of a systematic right of track bias, as stated in section 1.3.
- Short forecast lead times (i.e., 1 to 2 days prior to the event) exhibit a slow and left of track bias relative to the storm motion vector.
- Longer forecast lead times (i.e., 4 to 5 days prior to the event) exhibit a left of track bias for the category of all cyclones. Intense cyclones do not exhibit this medium range left of track bias.

- A negative correlation exists between forecast position and intensity errors for North-east ECs; specifically, weak bias cases are correlated with right of track bias cases, and strong bias cases are correlated with left of track bias cases. This correlation peaks at day 3 lead time before decreasing towards longer lead times. This verifies hypothesis 2 regarding an expected correlation between position and intensity biases.

The existence of a correlation between cyclone position and intensity biases at day 3 lead time motivated further analyses of cases with left of track (LOT) and right of track (ROT) biases, as well as weak (WB) and strong (SB) intensity biases. This analysis was done utilizing a synoptic composite approach, in order to compare the differences in the synoptic-to-large scale patterns between these different bias groups using CFSR data. These composites revealed the following conclusions:

- Both ROT and WB composites are associated with a more meridional flow over North America relative to LOT and SB cases, respectively, and higher antecedent MSLP within the Central U.S.
- ROT cases are associated with a stronger downstream polar jet streak and upstream subtropical jet streak relative to LOT cases, while in the WB cases the polar jet streak is displaced poleward and the subtropical jet streak is displaced equatorward relative to SB cases.
- These analyses verified hypothesis 3 regarding the existence of a correlation between WB and ROT cases and a more meridional flow over North America relative to SB and LOT cases.

To assess the climatology of ensemble variability, ellipses representing ensemble forecast position variability were calculated following the methodology of Hamill et al. (2011). Most cases exhibit position variability predominantly in the direction of the ensemble mean motion vector (i.e., along-track variability), with a much smaller subset of cases exhibiting position

variability predominantly normal to the direction of the ensemble mean motion vector (i.e., across-track variability). A composite approach to assess synoptic differences in cases that exhibit small vs. large ensemble position spread in either the along or across track direction revealed the following conclusions:

- Cases with small across-track variability (SV) are associated with a higher amplitude ridge and higher MSLP over western North America compared to the large across-track variability (LV) cases. There is no evidence of a stronger downstream ridge for LV cases compared to SV cases, thus rejecting hypothesis 4.
- SV cases are associated with more positive PNA, more negative EPO, and more negative NAO and AO relative to LV cases in the 1–4 days preceding the event. These result offer some prognostic value, and in conjunction with previous analyses suggest a higher likelihood of cold air outbreak over the eastern half of the U.S. leading up to and during a SV event compared to a LV event.
- Cases with small along-track variability (SV-L) are associated with a weaker downstream ridge over the western Atlantic Ocean and higher heights over Canada, as well as a weaker and equatorward shifted downstream jet streak, relative to large along-track variability (LV-L) cases. These results verify hypothesis 5.

## 5.2 Future Work

It should be recognized that the results obtained from this thesis, while providing value into the climatology of Northeast ECs and their forecast skill and biases, are sensitive to the tracking algorithm applied and the criteria used to match ensemble cyclones to the observed cyclone, especially at longer lead times. Different cyclone identification algorithms, especially those incorporating different criteria for cyclone identification as well as criteria for accounting for cyclone mergers or splitting, may have some impact on the duration of and number of cases. An area of future work may involve applying several different cyclone

tracking algorithms, perhaps including more sophisticated algorithms that consider the vertical structure of cyclones (e.g., Lim and Simmonds (2007)), to compare the climatology of GEFS errors and biases while using the same dataset. Additionally, the systematic biases presented in this work, and the synoptic patterns that correspond with these biases and errors, while similar to past literature, specifically are in reference to the version of the GEFS Reforecast used for this thesis (Hamill et al. 2013). Future upgrades to a new ensemble system, especially including a different dynamical core such as the upcoming Finite Volume Cubed-Sphere Dynamical Core (FV3) model, will likely lead to different systematic biases. In addition, while statistically significant differences were found between composites of intensity biases, position biases and position variability, the physical mechanisms that are associated with these differences were not investigated as part of this thesis, and serve as opportunities for future work.

This thesis is additionally part of a research-to-operations (R2O) project, with the goal of applying the results found in this thesis to operational forecasting in the National Weather Service (NWS). One such goal is to create a web-based tool to view ensemble forecasts of past high-impact Northeast ECs, and to apply the same methodology to real time operational GEFS forecasts to compare the forecast ensemble spread and type of variability to the climatology presented in this thesis.

## BIBLIOGRAPHY

- Allen, J. T., A. B. Pezza, and M. T. Black, 2010: Explosive Cyclogenesis: A Global Climatology Comparing Multiple Reanalyses. *J. Climate*, **23**, 6468–6484.
- Bentley, A., 2018: Extratropical Cyclones Leading to Extreme Weather Events over Central and Eastern North America. Ph.D. thesis, University at Albany, State University of New York.
- Bentley, A. M., L. F. Bosart, and D. Keyser, 2019: A Climatology of Extratropical Cyclones leading to Extreme Weather Events over central and eastern North America. *Mon. Wea. Rev.*, **147**, 1471–1490.
- Blender, R., and M. Schubert, 2000: Cyclone Tracking in Different Spatial and Temporal Resolutions. *Mon. Wea. Rev.*, **128**, 377–384.
- Bosart, L. F., 1981: The Presidents' Day Snowstorm of 18–19 February 1979: A Subsynoptic-Scale Event. *Mon. Wea. Rev.*, **109**, 1542–1566.
- Charles, M., and B. A. Colle, 2009: Verification of extratropical cyclones within cyclones within NCEP forecast models using an automated tracking algorithm: Part I: Comparison of the GFS and NAM models. *Wea. Forecasting*, **24**, 1173–1190.
- Charney, J. G., 1948: On the scale of atmospheric motions. *Geophys. Publ. Oslo*, **17 (2)**.
- Chen, S., and L. Dell'osso, 1987: A Numerical Case Study of East Asian Coastal Cyclogenesis. *Mon. Wea. Rev.*, **115**, 477–487.

- Colle, B. A., and M. E. Charles, 2011: Spatial distribution and evolution of extratropical cyclone errors over North America and its adjacent oceans in the NCEP Global Forecast System model. *Wea. Forecasting*, **26**, 129–149.
- Davis, C. A., and K. A. Emanuel, 1991: Potential Vorticity Diagnostics of Cyclogenesis. *Mon. Wea. Rev.*, **119** (8), 1929–1953.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597.
- Eady, E. T., 1949: Long waves and cyclone waves. *Tellus*, **1**, 33–52.
- Elless, T. J., 2015: Evaluating preferred direction tropical cyclone track variability in an operational global ensemble prediction system. M.S. thesis, Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York.
- Evans, M., 2006: An analysis of a frontogenetically forced early-spring snowstorm. *Bull. Amer. Meteor. Soc.*, **87**, 27–32.
- Froude, L. S., 2010: TIGGE: Comparison of the Prediction of Northern Hemisphere Extratropical Cyclones by Different Ensemble Prediction Systems. *Wea. Forecasting*, **25**, 819–836.
- Greybush, S. J., S. Saslo, and R. Grumm, 2017: Assessing the Ensemble Predictability of Precipitation Forecasts for the January 2015 and 2016 East Coast Winter Storms. *Wea. Forecasting*, **32**, 1057–1078.
- Grumm, R. H., and A. L. Siebers, 1989: Systematic Surface Cyclone Errors in NMC’s Nested Grid Model November 1988–January 1989. *Wea. Forecasting*, **4**, 246–252.

- Hamill, T. M., G. T. Bates, J. S. Whitaker, D. R. Murray, M. Fiorino, T. J. Galarneau, Y. Zhu, and W. Lapenta, 2013: NOAA's Second-Generation Global Medium-Range Ensemble Reforecast Dataset. *Bull. Amer. Meteor. Soc.*, **94**, 1553–1565.
- Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global Ensemble Predictions of 2009's Tropical Cyclones Initialized with an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **139**, 668–688.
- Hanley, J., and R. Caballero, 2012: Objective identification and tracking of multicentre cyclones in the era-interim reanalysis dataset. *Quarterly Journal of the Royal Meteorological Society*, **138 (664)**, 612–625.
- Harnik, N., G. Messori, R. Caballero, and S. B. Feldstein, 2016: The Circumglobal North American wave pattern and its relation to cold events in eastern North America. *Geophysical Research Letters*, **43 (20)**, 11,015–11,023.
- Hewson, T. D., and H. A. Titley, 2010: Objective identification, typing and tracking of the complete life-cycles of cyclonic features at high spatial resolution. *Meteorological Applications*, **17 (3)**, 355–381.
- Hodges, K. I., 1994: A General Method for Tracking Analysis and Its Application to Meteorological Data. *Mon. Wea. Rev.*, **122**, 2573–2586.
- Hodges, K. I., B. J. Hoskins, J. Boyle, and C. Thorncroft, 2003: A Comparison of Recent Reanalysis Datasets Using Objective Feature Tracking: Storm Tracks and Tropical Easterly Waves. *Mon. Wea. Rev.*, **131 (9)**, 2012–2037.
- Kenyon, J. S., 2013: The motion of mesoscale snowbands in northeast U.S. winter storms.

- M.S. thesis, Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York.
- Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Amer. Meteor. Soc.*, **91**, 363–376.
- Kocin, P. J., P. N. Schumacher, R. F. Morales, and L. W. Uccellini, 1995: Overview of the 12–14 March 1993 Superstorm. *Bull. Amer. Meteor. Soc.*, **76**, 165–182.
- Kocin, P. J., and L. W. Uccellini, 1990: *Snowstorms Along the Northeastern Coast of the United States: 1955 to 1985*. Meteor. Monogr.
- Kocin, P. J., and L. W. Uccellini, 2004a: A Snowfall Impact Scale Derived From Northeast Storm Snowfall Distributions. *Bull. Amer. Meteor. Soc.*, **85**, 177–194.
- Kocin, P. J., and L. W. Uccellini, 2004b: *Northeast Snowstorms. (Volume I: Overview, Volume II: The Cases)*, Vol. 54. Meteor. Monogr., Amer. Meteor. Soc.
- Konrad, C. E., 1996: Relationships between the Intensity of Cold-Air Outbreaks and the Evolution of Synoptic and Planetary-Scale Features over North America. *Mon. Wea. Rev.*, **124**, 1067–1083.
- Korfe, N. G., and B. A. Colle, 2018: Evaluation of Cool-Season Extratropical Cyclones in a Multimodel Ensemble for Eastern North America and the Western Atlantic Ocean. *Wea. Forecasting*, **33**, 109–127.
- Lim, E., and I. Simmonds, 2007: Southern Hemisphere Winter Extratropical Cyclone Char-

- acteristics and Vertical Organization Observed with the ERA-40 Data in 1979 – 2001. *J. Climate*, **20**, 2675–2690.
- Lorenz, E. N., 1963: Deterministic Nonperiodic Flow. *J. Atmos. Sci.*, **20**, 130–141.
- Marinaro, A., S. Hilberg, D. Changnon, and J. Angel, 2015: The North Pacific–Driven Severe Midwest Winter of 2013/14. *J. Appl. Meteor. Climatol.*, **54**, 2141–2151.
- May, R., S. Arms, P. Marsh, E. Bruning, and J. Leeman, 2008 - 2017: Metpy: A Python package for meteorological data. Boulder, Colorado, URL <https://github.com/Unidata/MetPy>, doi:10.5065/D6WW7G29.
- Miller, J. E., 1946: Cyclogenesis in the Atlantic Coastal Region of the United States. *J. Meteor.*, **3**, 31–44.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122** (529), 73–119.
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 463–493.
- Murray, R. J., and I. Simmonds, 1991: A numerical scheme for tracking cyclone centers from digital data. *Aust. Meteor. Mag.*, **39**, 155–166.
- Neu, U., and Coauthors, 2013: IMILAST: A Community Effort to Intercompare Extratropical Cyclone Detection and Tracking Algorithms. *Bull. Amer. Meteor. Soc.*, **94**, 529–547.
- Nguyen, L. T., J. Molinari, and D. Thomas, 2014: Evaluation of tropical cyclone center identification methods in numerical models. *Mon. Wea. Rev.*, **142**, 4326–4339.

- Novak, D. R., L. F. Bosart, D. Keyser, and J. S. Waldstreicher, 2004: An Observational Study of Cold Season–Banded Precipitation in Northeast U.S. Cyclones. *Wea. Forecasting*, **19**, 993–1010.
- Novak, D. R., B. A. Colle, and A. R. Aiyyer, 2010: Evolution of Mesoscale Precipitation Band Environments within the Comma Head of Northeast U.S. Cyclones. *Mon. Wea. Rev.*, **138**, 2354–2374.
- Nuss, W. A., and R. A. Anthes, 1987: A Numerical Investigation of Low-Level Processes in Rapid Cyclogenesis. *Mon. Wea. Rev.*, **115**, 2728–2743.
- Pinto, J. G., T. Spanghel, U. Ulbrich, and P. Speth, 2005: Sensitivities of a cyclone detection and tracking algorithm: individual tracks and climatology. *Meteorologische Zeitschrift*, **14 (6)**, 823–838.
- Reed, R. J., A. J. Simmons, M. D. Albright, and P. Uden, 1988: The role of latent heat release in explosive cyclogenesis: Three examples based on ECMWF operational forecasts. *Wea. Forecasting*, **3**, 217–229.
- Rios-Berrios, R., R. D. Torn, and C. A. Davis, 2016: An Ensemble Approach to Investigate Tropical Cyclone Intensification in Sheared Environments. Part I: Katia (2011). *J. Atmos. Sci.*, **73**, 71–93.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91 (8)**, 1015–1057.
- Sanders, F., and J. R. Gyakum, 1980: Synoptic-Dynamic Climatology of the “Bomb”. *Mon. Wea. Rev.*, **108**, 1589–1606.

- Sinclair, M. R., 1997: Objective Identification of Cyclones and Their Circulation Intensity, and Climatology. *Wea. Forecasting*, **12**, 595–612.
- Smith, B. B., and S. L. Mullen, 1993: An Evaluation of Sea Level Cyclone Forecasts Produced by NMC’s Nested-Grid Model and Global Spectral Model. *Wea. Forecasting*, **8**, 37–56.
- Sprenger, M., and Coauthors, 2017: Global Climatologies of Eulerian and Lagrangian Flow Features based on ERA-Interim. *Bull. Amer. Meteor. Soc.*, **98**, 1739–1748.
- Sutcliffe, R. C., 1947: A Contribution to the Problem of Development. *Quart. J. Roy.*, **73**, 370–383.
- Torn, R. D., 2017: A Comparison of the Downstream Predictability Associated with ET and Baroclinic Cyclones. *Mon. Wea. Rev.*, **145**, 4651–4672.
- Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677.
- Trenberth, K. E., 1978: On the interpretation of the diagnostic quasi-geostrophic omega equation. *Mon. Wea. Rev.*, **106**, 131–137.
- Uccellini, L. W., 1990: Processes contributing to the rapid development of extratropical cyclones. *Extratropical Cyclones: The Erik Palmén Memorial Volume*, C. W. Newton, and E. O. Holopainen, Eds., Amer. Meteor. Soc., 81–105.
- Uccellini, L. W., D. Keyser, K. Brill, and C. Wash, 1985: The Presidents’ Day Cyclone of 18–19 February 1979: Influence of Upstream Trough Amplification and Associated Tropopause Folding on Rapid Cyclogenesis. *Mon. Wea. Rev.*, **113**, 962–988.

- Uccellini, L. W., P. Kocin, R. Petersen, C. Wash, and K. Brill, 1984: The Presidents' Day Cyclone of 18–19 February 1979: Synoptic Overview and Analysis of the Subtropical Jet Streak Influencing the Pre-Cyclogenetic Period. *Mon. Wea. Rev.*, **112**, 31–55.
- Walsh, J. E., A. S. Phillips, D. H. Portis, and W. L. Chapman, 2001: Extreme Cold Outbreaks in the United States and Europe. *J. Climate*, **14**, 2642–2658.
- Wernli, H., and C. Schierz, 2006: Surface Cyclones in the ERA-40 Dataset (1958–2001). Part I: Novel Identification Method and Global Climatology. *J. Atmos. Sci.*, **63**, 2486–2507.
- Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*, International Geophysics, Vol. 100. 3rd ed., Academic Press.
- Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the “surprise” snowstorm of 24–25 January 2000. *Mon. Wea. Rev.*, **130**, 1617–1632.
- Zheng, M., E. K. Chang, and B. A. Colle, 2013: Ensemble Sensitivity Tools for Assessing Extratropical Cyclone Intensity and Track Predictability. *Wea. Forecasting*, **28**, 1133–1156.