UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A Comparison of Hybrid-4DEnVar and Hourly Hybrid-3DEnVar Assimilation of Tail Doppler Radar Observations on the Prediction of Rapidly Evolving Hurricanes

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

BENJAMIN DAVIS
Norman, Oklahoma
2019

A Comparison of Hybrid-4DEnVar and Hourly Hybrid-3DEnVar Assimilation of Tail Doppler Radar Observations on the Prediction of Rapidly Evolving Hurricanes

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

Dr. Xuguang Wang, Chair

Dr. Steven Cavallo

Dr. Aaron Johnson

Dr. Naoko Sakaeda

Acknowledgments

I would like to thank my advisor Dr. Xuguang Wang for her guidance and support throughout the project and the rest of my committee for their support and feedback. I would also like to thank the members of my research group, particularly Xu Lu, Sam Degelia, and Hristo Chipilski for their time discussing and providing feedback on my research and writing. Additionally, I would like to thank the professors and staff that have taught classes and assisted in navigating graduate school.

Thank you to my family and friends for supporting me during my time working on this degree. Thanks to Sara Wugofski, Stephanie Edwards, Andrew Berrington, Sean Ersnt, Kelton Halbert, CJ Sayre, Mitchell Kelleher, and all my other friends in Norman and around the country for making these last few years enjoyable. This would not have been possible with out you.

Table of Contents

List of Tables

List of Figures

Abstract

Beginning in the summer of 2017 the National Weather Service upgraded the operational Hurricane WRF (HWRF) to a continuously cycled Gridpoint Statistical Interpolation (GSI) based, 6-hourly, Hybrid 3-dimensional (3D) Ensemble-Variational (EnVar) data assimilation (DA) system. The Hybrid-3DEnVar system assumes the background error covariance is constant throughout the 6 hour DA window and is therefore unable to account for temporal evolution. During rapidly evolving conditions this assumption can cause analyses to be sub-optimal which can lead to degraded forecast performance. Furthermore, in a cycled DA system this problem may be more pronounced due to the accumulation of these errors over successive cycles. The first goal of this thesis is to evaluate methods to improve analysis and forecast accuracy compared to the 6-hourly Hybrid-3DEnVar system by accounting for the evolution of background error covariance.

Two methods are considered that are expected to produce improved analyses and forecasts compared to the 6-Hourly Hybrid 3DEnVar in HWRF by accounting for evolution of the background error covariance: 1) 4DEnVar with 6-hourly frequency (hereafter 4DEnVar), and 2) 3DEnVar with hourly frequency (hourly-3DEnVar). In 4DEnVar, the temporal evolution of error covariances is incorporated through the use of 4-dimensional ensemble perturbations with implicit linearity assumption during the minimizations. However, nonlinear error growth within a 6-hour window can still pose difficulties. Hourly-3DEnVar assimilates observations in one hour windows instead of one longer six hour window, allowing the error covariances to change each hour. However, frequent interruption of the model could introduce additional instabilities. The second goal of this study is to evaluate the performance of the hourly-3DEnVar and 4DEnVar hybrid data assimilation systems relative to each other in the HWRF model. It is expected that the two systems will produce similar results. While previous work has been done to evaluate different data assimilation frequencies and methods on thunderstorms, little work has been done comparing these two methods in the case of a hurricane.

Experiments are conducted for DA cycles that tail Doppler radar (TDR) data are available to be assimilated during Hurricane Edouard (2014) and Hurricane Irma (2017). Edouard became a category 3 hurricane while Irma became a category 5 storm. These cases are chosen due to the amount of TDR data available during rapid or near-rapid intensification and weakening phases of these storms. TDR data allows for abundant data in the inner core of the hurricane, whereas conventional observations may be sparse in this region particularly when the storm is over open oceans.

A baseline run is performed using 6-hourly Hybrid-3DEnVar. 4DEnVar and hourly-3DEnVar experiments are run for each DA cycle where TDR data are available. All experiments start from the same background produced by the control run, except consecutive DA cycles where TDR data are available, in which case the first cycle uses the background from the control and then is continuously cycled. Analyses and forecasts are verified against best track, TCVitals, satellite, stepped frequency microwave radiometer (SFMR), and TDR observations in order to explain differences seen in these systems.

Both the 6-Hourly 4DEnVar and the 1-Hourly 3DEnVar experiments produce better analyses and forecasts than 6-Hourly 3DEnVar by most metrics, however the advantages are limited to early forecast lead times. 6-Hourly 4DEnVar and 1-Hourly 3DEnVar perform similarly for most verification metrics, but early 1-Hourly 3DEnVar wind forecasts are degraded due to spindown. Additionally, 1-Hourly 3DEnVar is associated with larger computational cost than 4DEnVar.

Chapter 1

Introduction

Tropical Cyclones (TCs) can cause large losses of life and billions of dollars in damage. For example, recent category 5 hurricanes Irma (2017), Maria (2017), and Michael (2018) each caused more than $50 billion in damage and several dozen direct deaths with hundreds of injuries and indirect deaths. Despite being weaker on the Saffir-Simpson scale at landfall, Harvey (2017), and Florence (2018) produced additional significant impacts through widespread heavy rain and inland flooding after stalling near the coast. Summaries can be found at (https://www.nhc.noaa.gov/data/tcr). One way to reduce the significant risk to life and property is through improving numerical predictions of hurricanes. For example, if rapid intensification (RI) can be more confidently forecast in advance the decision to evacuate could be made sooner. While forecasts can be improved through several avenues; this study focuses on improving the forecasts of hurricanes by applying advanced data assimilation (DA) techniques.

Early studies used various vortex initialization methods to initialize hurricane forecasts when lacking inner-core observations (Kurihara et al. 1990, 1993, 1995, 1998; Thu and Krishnamurti 1992; Bender et al. 1993; Zou and Xiao 2000; Liu et al. 2000, 2006; Pu and Braun 2001; Tallapragada et al. 2014). Although these methods improve forecast skill, they may not produce a realistic storm (e.g. Bogusing). Further improvement has been shown through the use of ensemble-based DA methods such as the Ensemble Kalman Filter (EnKF) (Torn and Hakim 2009; Zhang et al. 2009; Li and Liu 2009; Hamill et al. 2011; Wang 2011; Zhang et al. 2011; Aksoy et al. 2012, 2013; Weng and Zhang 2012; Dong and Xue 2013; Poterjoy and Zhang 2014; Poterjoy et al. 2014). Such techniques allow the use of flow dependent characteristics during DA and allow for dynamic and thermodynamic consistency in the DA analysis for hurricane initialization.

Ensemble-Variational (EnVar) DA methods have been proposed as a way to further advance the quality of analyses and subsequent forecasts produced by DA (Hamill and Snyder 2000; Lorenc 2003; Etherton and Bishop 2004; Wang et al. 2007b, 2013a; Wang 2010, 2011; Li et al. 2012; Schwartz et al. 2013, 2015; Wang and Lei 2014; Li et al. 2015; Lu et al. 2017a,b). These techniques typically incorporate ensemble error covariances into the variational framework (Wang et al. 2013c). EnVar DA methods have been studied and implemented on both global and regional scales with promising results for both hurricane and non-hurricane applications (Buehner 2005; Wang et al. 2008a,b, 2013c; Buehner et al. 2010a,b; Bishop and Hodyss 2011; Wang 2011; Li et al. 2012; Zhang and Zhang 2012; Clayton et al. 2013; Wang et al. 2013b; Kutty and Wang 2015; Kleist and Ide 2015b,a; Lu et al. 2017a,b).

Ensemble-based DA methods, including EnVar, can provide better hurricane track forecasts than variational methods which use flow-independent static covariance (Wang et al. 2007a, 2009, 2013a; Buehner et al. 2010b; Wang 2011; Hamill et al. 2011; Zhang and Zhang 2012; Poterjoy and Zhang 2014; Wang and Lei 2014). Although intensity forecasts have presented more challenges than track forecasts (Rogers et al. 2013), some studies have shown there is improvement after assimilating inner-core observations (Pu et al. 2009; Li et al. 2012; Weng and Zhang 2012; Lu et al. 2017b). When the background is rapidly evolving during the DA time window (Wang and Lei 2014; Lu et al. 2017b) the stationary background error statistics commonly used by the three-dimensional (3D) EnVar may not be accurate, particularly for data that is temporally distant from the analysis time. For example, in the most commonly used 6-hourly 3DEnVar DA, the 6-hour long DA time window could introduce considerable errors in the analysis with some observations found near the edges of the time window (e.g. about 3 hours away from the analysis time; Wang and Lei 2014). Wang and Lei (2014) demonstrated using Hurricane Daniel (2010) that the 3DEnVar increment can be nearly the opposite of expected when the background is evolving rapidly. Such errors would be a significant problem if the storm were to undergo RI.

The RI process is sensitive to the inner-core structure (Leighton et al. 2018) and is often missed or weaker in model forecasts than observed (Pu et al. 2009). Therefore, a more temporally flow-dependent four-dimensional (4D) EnVar or 3DEnVar with a shorter time window might be helpful in improving RI forecasts (Lu et al. 2017a,b).

One method to account for rapidly-evolving background error is 4DEnVar, which uses background error covariance at the time of the observation rather than the analysis time. Several studies have shown that 4D methods perform better than their 3D counterparts for the same DA window length (Hunt et al. 2004, 2007; Zhang et al. 2011; Wang et al. 2013a,b; Wang and Lei 2014; Lu et al. 2017b; Zhang and Pu 2019). One reason for this result is seen in Wang and Lei (2014). They found that 4DEnVar produced an increment similar to that from assimilating the observation near its observation time and propagating it forward in the model. However, Wang et al. (2013a) found improvements in 4D EnKF compared to 3D EnKF were limited to early forecast times when assimilating Doppler radar data in a convective thunderstorm. Additionally, Gauthier et al. (2007) found that even when the 4D analysis does not notably improve, there may still be improvements in the forecast, such as with the implementation of 4DVar by Meteorological Service of Canada in 2005. While no improvement was seen in analyses when compared to radiosonde data, innovation and forecast statistics improved with the use of 4DVar.

Another method to account for a rapidly-evolving background error covariance is to reduce the length of the DA window. Shorter DA windows reduce temporal errors between the observations and background, which may help to reduce errors associated with storm location in addition to the changing error covariance structures. The shorter DA cycles also prevent the filtering of dynamical imbalances in the model with periods of 2-4 hours that are typically filtered out in 6 hour cycles (Huang and Lynch 1993). However, frequent model interruptions may introduce a shock to the model at each analysis time (Wang et al. 2013b; Houtekamer and Zhang 2016). As this shock can result in a degraded analysis and forecast, care must be taken in selecting the length of DA window for optimal results. Furthermore,

frequent assimilation is computationally expensive, largely due to the increased I/O cost of current DA systems.

Short DA cycles are often used when assimilating high-frequency data such as radar or surface data (Hu and Xue 2007; Dowell and Wicker 2009; Lu et al. 2017a), generally with positive results. Several studies found that using a DA window shorter than the traditional 6-Hour DA window showed improved forecasts (Lei and Anderson 2014; Tong et al. 2016; Lu et al. 2017a), although there is some debate over the optimal length. For example Lei and Anderson (2014) found that assimilating surface pressure at 1-hour intervals produced better analyses than longer 3- and 6-hour intervals, while Tong et al. (2016) found that assimilating radar data at 3-hour intervals produced better analyses than 1-hourly 3DVar. Both studies used the same amount of data per cycle, resulting in more data being assimilated overall for the same time period in configurations with shorter cycles. Lu et al. (2017a) found that 1-hour DA in an HWRF based system provided better forecasts than longer 6-hour DA windows.

While either 4D techniques or shorter 3D cycles can account for rapidly evolving background error covariances and temporal errors that are introduced by observations that are near the edges of the DA window, there are different advantages and disadvantages associated with each. Shorter 3D DA cycles can better account for non-linear background error evolution compared to 4D DA cycles. Evensen (2018) compared a 5 minute 3DEnKF and a 10-minute 4DEnKF for a linear case, finding that the two are equivalent. However, it is unclear how the two methods compare in a case with non-linear error evolution such as a hurricane. As mentioned previously, frequent interruptions to the model can result in an increased imbalance in 3D compared to 4D systems. Despite these limitations, Wang et al. (2013b) found that some shorter DA cycle lengths can be used for a 3D EnKF and still perform better than 4D EnKF with a longer window. However, the imbalance problem results in a limit to the frequency of DA. Fertig et al. (2007) found that for an EnKF it may be

necessary to have more frequent DA in some rapidly changing situations to capture short term or rapidly evolving features.

This thesis aims to compare 1-Hourly 3DEnVar with 6-Hourly 4DEnVar in rapidly evolving hurricanes using inner core DA. Furthermore, both systems are compared with 6-Hourly Hybrid-3DEnVar to quantify the improvements provided by accounting for the evolution of the background error covariance. Both 4DEnVar and 1-Hourly 3DEnVar experiments are conducted using background error covariance with 1-hour frequency compared to the 6-hour frequency of the standard 3DEnVar. This allows the evolution to be captured, even when the background error covariance is changing rapidly. However, to best of the author's knowledge, limited studies have compared these methods, especially in the case of inner-core hurricane DA. While Lu et al. (2017b) found improved results when applying the 4DEnVar over a 6-hour 3DEnVar for Hurricane Edouard (2014), they did not explore the potential advantages of an hourly 3DEnVar. It is expected that Hourly 3DEnVar should also improve upon the 6 Hourly 3DEnVar due to the ability to account for background error evolution but is unknown how it compares to 4DEnVar.

This study utilizes the Hurricane WRF (HWRF). HWRF is a regional hurricane model developed and maintained since 2002 by the Environmental Modeling Center (EMC), the Geophysical Fluid Dynamics Laboratory (GFDL), and the University of Rhode Island (URI) that covers the North Atlantic and Northeast Pacific Basins (Tallapragada et al. 2014; Biswas et al. 2017). The Hybrid DA systems developed by the OU Multiscale Data Assimilation (MAP) Lab and NOAA (Lu et al. 2017b) are used and further extended to a 1-Hourly 3DEnVar. The 6-Hourly 3DEnVar system differs from the operational HWRF primarily in the use of the directed moving nest strategy, with 6-Hourly 4DEnVar and 1-Hourly 3DEnVar differing further in their use of 4DEnVar and 1 Hour DA windows, respectively. The remainder of this thesis is organized as follows: Chapter 2 describes the DA configurations, Chapter 3 describes the model and experiment designs, Chapter 4 describes the results from

experiments with Edouard (2014), Chapter 5 describes the results from experiments with Irma (2017), and chapter 6 discusses these results.

Chapter 2

System Design

A Gridpoint Statistical Interpolation (GSI) based hybrid EnKF-Var DA system for HWRF is used in this study. The system follows Lu et al. (2017b) for 6-Hourly 3DEnVar and 6-hourly 4DEnVar experiments, with necessary changes made to extend to 1-Hourly 3DEnVar. The complete system description can be found in Lu et al. (2017b). The system is described here mirroring Lu et al. (2017b) in subsections 2.1-2.4, and a brief description of the configuration of each DA method is given in subsections 2.5-2.7.

2.1 General Overview

The GSI based hybrid EnKF-Var DA system used in this study utilizes a 40-member ensemble and a single control member. The control member is updated within a variational framework using the error covariances calculated from the ensemble, and provides the analysis from which a 120-hour free forecast is initialized. In the first cycle for each hurricane, the ensemble and control member are initialized using the ensemble analysis from the National Centers for Evironmental Predictions (NCEP) operational GFS hybrid DA system (Wang et al. 2013c). Ensemble storm centers are relocated using the TCVitals database (Liu et al. 2000, 2006) after 6 hours, while the control is relocated and modified at 3, 6, and 9h. Updated storm centers are determined by assimilating the TCVitals storm center location using an EnSRF (section 2.4). The modified ensemble and control member are then used as the background for the first DA cycle.

The Hybrid DA process is shown in figure (2.1) and consists of 4 steps:

1. The relocated background for the control member is updated by the dual-resolution GSI augmented control vector (GSI-ACV) with the relocated ensemble background covariance.

2. The relocated ensemble is updated using an EnKF.

3. The ensemble is recentered such that the ensemble mean matches the control analysis.

4. The outermost domain is replaced with the GFS Ensemble and control 27 or 18 km grids for all members.

Forecasts for the 40-member ensemble and the control member are initialized from the analysis until a time as required by the specific DA configuration (described in the following sections). A directed moving nest strategy is employed to prevent non-overlapping domains in the storm following nests. The directed moving nest prescribes the location for the center of the moving nests' domain, which is the same for all members, instead of adopting the HWRF moving nest strategy. The forecast produced by the model using the directed moving nest strategy is nearly indistinguishable from that produced using the HWRF moving nest strategy during the first 6 hours (Lu et al. 2017b), providing benefit during DA without increasing model errors. The directed moving nest strategy is fully described in section 2d of Lu et al. (2017b). Vortex relocation and modification are performed on the forecasts to be used as backgrounds for the next cycle.

A 120-hour free forecast is initialized from the control analysis. Since the primary benefit of the directed moving nest strategy is forcing overlapping domains for EnVar DA, the free forecast does not use the directed moving nest strategy, instead using the original HWRF vortex-following strategy. As a result, this forecast is independent of the control run.

2.2 GSI-ACV

The GSI-ACV system updates the relocated control background and is further detailed here following Wang and Lei (2014) and Lu et al. (2017b). For Dual-resolution 4DEnVar, the analysis increment at time t is defined as

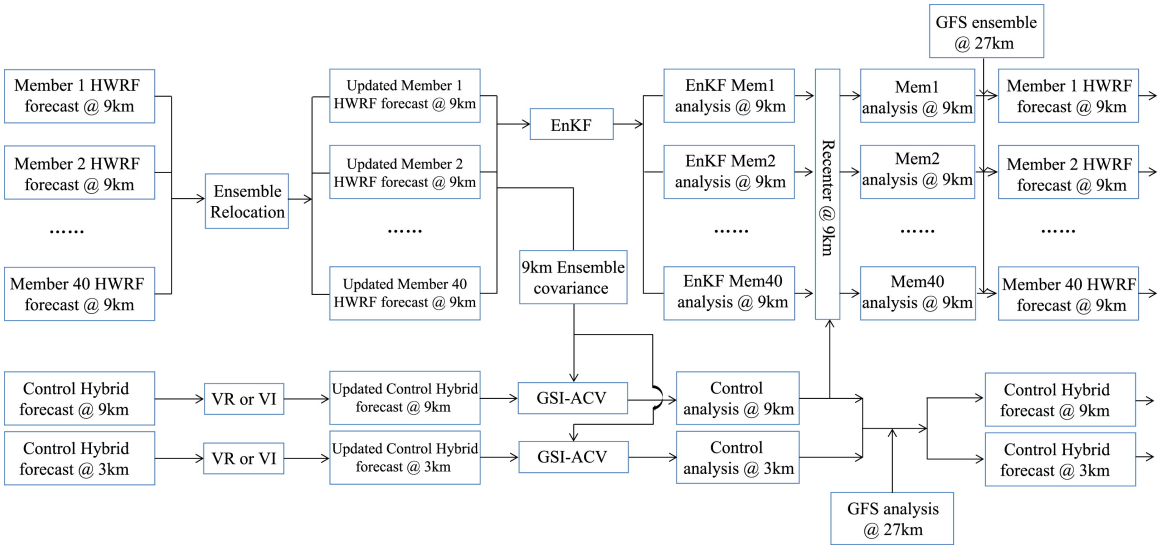


Figure 2.1: Flowchart of the GSI Based DA System for HWRF [Adapted from Wang et al. (2013c) and Lu et al. (2017b)]

$$\boldsymbol{x}'_t = \boldsymbol{D} \sum_{k=1}^{K} [\boldsymbol{a}_k \circ (\boldsymbol{x}^e_k)_t] \tag{2.1}$$

where $\boldsymbol{D}$ is an operator mapping coarse ensemble model fields to the finer model grid, $\boldsymbol{a}_k$ is the augmented control vectors for the k$^{\text{th}}$ ensemble member, $(\boldsymbol{x}^e_k)_t$ is the k$^{\text{th}}$ ensemble perturbation normalized by $\sqrt{(K-1)}$ at time t, with K being the ensemble size and $\circ$ is the Schur product.The 4DEnVar cost function is unchanged from single resolution 4DEnVar (Wang and Lei 2014):

$$\boldsymbol{J}(\boldsymbol{x}'_1, \boldsymbol{a}) = \frac{1}{2}(\boldsymbol{a})^T \boldsymbol{A}^{-1}(\boldsymbol{a}) + \frac{1}{2}\sum_{t=1}^{L}(\boldsymbol{y}^{o\prime}_t - \boldsymbol{H}_t \boldsymbol{x}'_t)^T \boldsymbol{R}^{-1}(\boldsymbol{y}^{o\prime}_t - \boldsymbol{H}_t \boldsymbol{x}'_t) \tag{2.2}$$

where $\boldsymbol{a}$ is $< \boldsymbol{a}_1, \boldsymbol{a}_, \ldots, \boldsymbol{a}_k >$ for k=1,K; $\boldsymbol{A}$ is the matrix defining the localization to the ensemble covariance, L is the length of the DA window, $\boldsymbol{R}_t$, $\boldsymbol{y}^{o\prime}_t$, and $\boldsymbol{H}_t$ are the observation error covariance, innovation vector, and linearized observation operator valid at time t, respectively. The localization matrix $\boldsymbol{A}$ is defined following Wang et al. (2013c) and Lu et al. (2017a,b).

The primary difference between single resolution and dual resolution 4DEnVar is the use of the operator $\boldsymbol{D}$, which projects the coarser-resolution ensemble onto the finer resolution control grid. If $\boldsymbol{D}$ is the identity matrix the equation simplifies to the single resolution formula. When only a single time is considered equations (2.1) and (2.2) describe 3DEnVar since 4DEnVar is a temporal extension of 3DEnVar (Wang and Lei 2014). Both the 4DEnVar and 3DEnVar versions are used in this study.

2.3 EnKF

An EnKF is used to update the ensemble members during the second step in the DA system. Following Lu et al. (2017a,b) the EnKF utilizes an ensemble square root filter (EnSRF) (Whitaker and Hamill 2002) The EnKF code is designed for use with HWRF (Lu et al. 2017a,b) and uses the observations preprocessing, quality control, and forward operators provided by GSI. Horizontal and vertical localization cutoffs are similar to those used in the GSI-ACV. Finally, the Relaxation to prior spread (RTPS) multiplicative inflation algorithm developed by Whitaker and Hamill (2002) is adopted and the inflation parameter is set to 0.9 following Lu et al. (2017b).

2.4 Vortex Relocation and Modification

For 6-hour forecasts the average storm location error is 15-40 km (Trahan and Sparling 2012; Tong et al. 2014). Location errors of this magnitude can significantly degrade the analysis when assimilating inner-core data through ensemble-based DA using a gaussian error assumption (Chen and Snyder 2007; Yang et al. 2013; Lu et al. 2017a,b). Therefore, this system adopts vortex relocation (VR) for both the ensemble and control in addition to the vortex modification for the control. The storm center locations are determined using the EnSRF method from Whitaker and Hamill (2002) for a single-variable problem. The EnSRF is a three-step process 1) update the ensemble mean (eq 2.3), 2) update the ensemble

perturbations (eq 2.4), 3) combine the results of 1) and 2) to compute the final update (eq 2.5). The ensemble mean is updated by

$$\bar{\boldsymbol{x}}^a = \bar{\boldsymbol{x}}^b + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{H}\bar{\boldsymbol{x}}^b) \tag{2.3}$$

where $\bar{\boldsymbol{x}}^a$ is the updated ensemble mean location, $\bar{\boldsymbol{x}}^b$ is the prior ensemble mean location, $\boldsymbol{K}$ is the Kalman gain, $\boldsymbol{y}$ is the TCVitals observation, and $\boldsymbol{H}$ is the unit observation operator. $\boldsymbol{K}$ is given by $\boldsymbol{P}^b\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{P}^b\boldsymbol{H}^T + \boldsymbol{R})^{-1}$, where $\boldsymbol{P}^b$ is the location error variance in the ensemble, and $\boldsymbol{R}$ is the error variance of the TCVitals location, estimated to be 10 km following Trahan and Sparling (2012).

Ensemble location perturbations are updated by:

$$\boldsymbol{x}_k'^a = (\boldsymbol{I} - \tilde{\boldsymbol{K}}\boldsymbol{H})x_k'^b \tag{2.4}$$

where $\boldsymbol{x}_k'^a$ is the ensemble analysis location perturbations, $\boldsymbol{I}$ is the identity matrix, $\tilde{\boldsymbol{K}}$ is the reduced Kalman game computed as $\tilde{\boldsymbol{K}} = [1 + \sqrt{\frac{\boldsymbol{R}}{(\boldsymbol{H}\boldsymbol{P}^b\boldsymbol{H}^T)+\boldsymbol{R}}}]^{-1}\boldsymbol{K}$, $\boldsymbol{x}_k'^b$ is the prior ensemble location perturbations. Equations (2.3) and (2.4) are combined to produce the final analysis of storm center locations used for VR:

$$\boldsymbol{x}_k^a = \bar{\boldsymbol{x}}^a + \boldsymbol{x}_k'^a \tag{2.5}$$

where $\boldsymbol{x}_k^a$ is the updated ensemble of locations.

Once updated ensemble storm locations are calculated, vortex relocation procedures are used based on the HWRF VR procedure (Liu et al. 2000, 2006; Tallapragada et al. 2014; Biswas et al. 2017) to relocate both the control and ensemble backgrounds. The vortex from the HWRF forecast is removed and placed in the location determined by equation (2.5). Vortex modification (VM) is then applied to control vortices by adjusting the size and intensity before placing the vortex back in the environment. Because negative impacts resulting from the interaction between GSI and VM have been seen (Zhou et al. 2015), the VM method is not used when inner core data (e.g. TDR data) are available.

2.5 6-Hourly 3DEnVar

To prepare for the next cycle, a 9-hour forecast is initialized from the previous cycles' control analysis, while a 6-hour forecast is initalized from each ensemble analysis. Due to the 6 hour length of the DA window, First Guess at Appropriate time (FGAT) (Trahan and Sparling 2012) is used during the GSI-ACV step of the DA, thus requiring backgrounds at hours 3, 6, and 9. These backgrounds are updated using VR and VM prior to DA. The analysis produced by the 6-Hourly 3DEnVar system is valid at the center of the 6-hour DA window with observations of +/- 3 hours from analysis time being assimilated.

2.6 6-Hourly 4DEnVar

For 6-Hourly 4DEnVar, a 9 hour forecast is initialized from the previous control and ensemble analyses. VR is employed to update the background prior to the GSI-ACV step at 3, 4, 5, 6, 7, 8, and 9 hours. The dual-resolution 4DEnVar data assimilation used for the GSI-ACV step in the analysis. The analysis is produced at the center of the 6 hour assimilation window.

2.7 1-Hourly 3DEnVar

Each 6 hour DA window is broken into 7 cycles with an analysis valid each hour, including the beginning and end of the mission. As the hourly DA has 7 cycles corresponding to a single 6-hour DA window, the term mission is used in discussing 1-Hourly 3DEnVar to refer to a set of cycles matching a 6 hourly DA window in 6-Hourly 3DEnVar and 6-Hourly 4DEnVar. The first and last cycles of each mission only consider data from the 30 minutes after and 30 minutes before the analysis time, respectively, with the remaining 5 cycles considering data both in the 30 minutes before and after the analysis time. When continuously cycled, the final analysis of the previous mission is used as the background for the first cycle of the next mission, effectively creating a single 1-hour cycle that spans

two missions. This is done so that the data included in each mission is consistent between experiments. When not continuously cycled, a 3-hour forecast is initialized from the previous cycle, with VR being used. Before the cycles 2 through 7 in each mission, a 1-hour forecast is initialized from the previous analysis for each ensemble member and the control member. A 3DEnVar GSI-ACV update is performed on the control member as in the 6-hourly 3DEnVar, except only using data that occurs during the 30-60 minute cycle as described above. Similarly, the ensemble is updated using a 3DEnKF using only the data in the shorter cycle. For times between the standard 6-hourly output for the GFS ensemble interpolated ensemble data is used to initialize the outermost domain.

The free forecast is only launched from the final analysis of each mission, placing the initialization time of the free forecast 3 hours later than that of the 6 Hourly 3DEnVar and 6 Hourly 4DEnVar. As described above, this ensures that the same data is assimilated during each mission for each experiment.

Chapter 3

Case Details and Experiment Design

3.1 HWRF Configuration

The HWRF model has been developed by the Environmental Modeling Center (EMC) with the Geophycisal Fluid Dynamics Laboratory (GFDL) and the University of Rhode Island (URI) since 2002 (Tallapragada et al. 2014; Biswas et al. 2017). The 2014 configuration is used in this study following Lu et al. (2017b) for experiments with Hurricane Edouard (2014) and the 2017 configuration is used for experiments with Irma (2017). For Edouard (2014) a two-way triple nested domain with horizontal grid spacing of 0.18°/0.06°/0.02° (approximately 27/9/3 km) is used, with 61 vertical levels. The model top is 2 hPa following Lu et al. (2017a,b). The outermost domain uses 216 x 432 horizontal grid points, the intermediate domain uses 232 x 454 horizontal grid points, and the innermost domain uses 181 x 322 (figure 3.1). For Experiments with Hurricane Irma (2017) a triple nested domain with horizontal grid spacing of 0.135°/0.045°/0.015°(approximately 18/6/2 km) is used, with 75 vertical levels, with the model top at 10 hPa. The outermost domain uses 348 x 516 horizontal gridpoints, the intermediate domain uses 265 x 532 horizontal grid points and the innermost domain uses 235 x 472 horizontal grid points. The physics parameters follow the 2014 (Edouard) (Tallapragada et al. 2014) and 2017 (Irma) (Biswas et al. 2017) operational HWRF (Table 3.1), except that the ocean coupling is turned off.

3.2 Observational Data

Conventional observations, clear-sky radiances from satellites, and sattelite derived winds (Tallapragada et al. 2014; Biswas et al. 2017) as well as radial velocities from TDR are

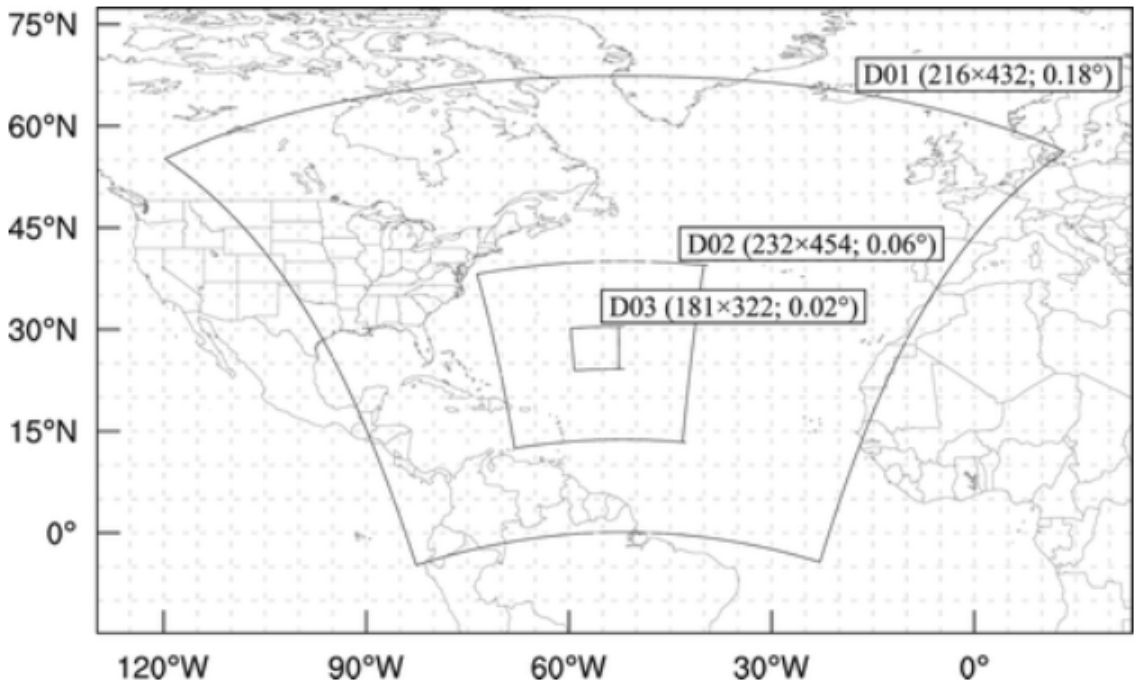


Figure 3.1: Example domain configuration for Edouard (2014) [Adapted from Lu et al. (2017b)]

Table 3.1: Table of model physics used for Edouard (2014) and Irma (2017).

Physics	2014	2017
Microphysics	Ferrier	Ferrier-Aligo
Cumulus	Simplified Arakawa-Schubert (SAS)	Scale Aware Simplified Arakawa-Schubert (SASAS)
Surface Layer	HWRF Surface Layer	HWRF Surface Layer
Land-surface Model	GFDL slab scheme	Noah LSM
Planetary Boundary Layer	Non-local	Non-local Hybrid Eddy-Diffusivity Mass-Flux (Hybrid EDMF)
Radiation	Eta Longwave and Shortwave	RRTMG Longwave and Shortwave

assimilated each cycle (table 3.2). While satellite radiances are only assimilated on the intermediate domain, conventional observations, satellite derived winds, and radial velocities (when available) are assimilated on both the intermediate and innermost domains.

Radial velocities are recorded by Tail Doppler Radar (TDR) mounted on the NOAA P-3 aircraft. Descriptions of the processing of TDR data can be found in Gamache et al. (2015) and Lu et al. (2017a)

3.3 Case Description

This study uses Hurricane Edouard (2014) and Hurricane Irma (2017) to evaluate the performance of 1-Hourly 3DEnVar and 6-Hourly 4DEnVar in the case of a rapidly evolving

Table 3.2: Data types assimilated.

Data Type		**Domain (d01)**	**Domain (d02)**	Domain (d03)
Conventional Observations	Radiosondes	No observations are assimilated	Y	Y
	Dropwindsondes		Y	Y
	Aircraft Reports		Y	Y
	Surface Ship and Buoy Observations		Y	Y
	Surface Observations over Land		Y	Y
	Pibal Winds		Y	Y
	Wind Profilers		Y	Y
	Radar-derived Velocity Azimuth Display Winds		Y	Y
	WindSat Scatterometer Winds		Y	Y
	Integrated Precipitable Water Derived from the Global Positioning System		Y	Y
Tail Doppler Radar Observations			Y	Y
Satellite Derived Winds			Y	Y
Satellite radiances	IR		Y	N
	MW		Y	N

hurricane when abundant data is available in the inner core. These cases are selected due to the amount of inner-core TDR data available during rapid or near-rapid intensification and weakening. A description of each hurricane, the data assimilated, and the experiments conducted are provided in this section.

3.3.1 Edouard

Hurricane Edouard developed 720 n mi west of the Cape Verde Islands on 1200 UTC 11 September 2014. Edouard peaked in intensity at 105 kts at 1200 UTC 16 September 2014 before immediately weakening due to an eyewall replacement cycle (`https://www.nhc.noaa.gov/data/tcr/AL062014_Edouard.pdf`). Edouard then began a northward and then northeastward motion during this weakening phase, accelerating ahead of a midlatitude trough. Finally, Edouard transitioned into a post-tropical cyclone on 19 September, before the remnant low was absorbed into a frontal system on 21 September.

This study focuses on inner core assimilation of TDR data, therefore experiments are performed only for cycles in which TDR data is available (Figure 3.2. These cycles are the 16th , 17th, 21st, 24th, and 25th cycles valid at 1200 UTC 15 October, 1800 UTC 15 October, 1800 UTC 16 October, 1200 UTC 17 October, and 1800 UTC 17 October (Table 3.3).

3.3.2 Irma

Hurricane Irma developed into a TC around 0000 UTC 30 August about 120 n mi west-southwest of Sao Vicente in the Cabo Verde Islands, before intensifying to a tropical storm just 6 hours later. Irma quickly became a hurricane by 0600 UTC 31 August, and later a major hurricane at 0000 UTC 1 September. Due to eyewall replacement cycles and dry air intrusion (`https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf`), Irma fluctuated between category 2 and 3 for the next 3 days. Irma then rapidly intensified from on 4 and 5 of September, reaching maximum intensity of 155 kts around 1800 UTC 5 September.

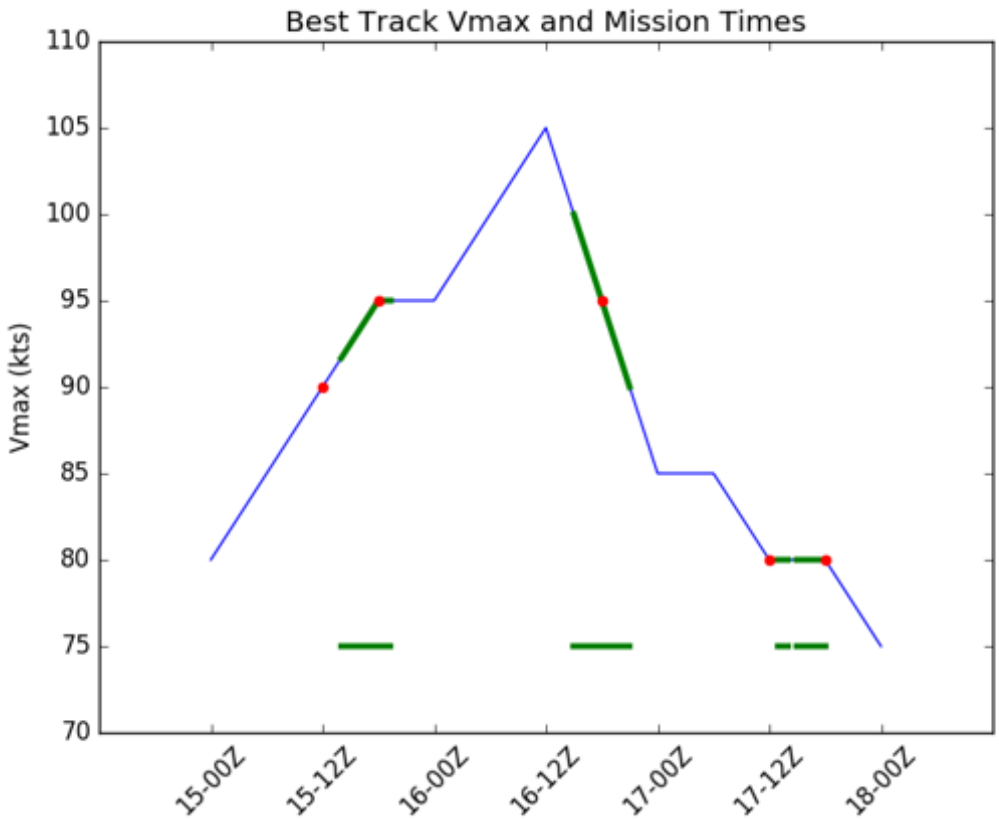


Figure 3.2: Edouard (2014) Best track maximum wind speed (Vmax; blue) with Tail Doppler Radar (TDR) data availability overlaid (green). Center of missions marked with red dot.

Table 3.3: Table of Edouard (2014) and Irma (2017) missions used. Time provided is the center of the 6-hour data assimilation (DA) window. Mission number is the indicates which DA cycle in the baseline 6-Hourly 3DEnVar system the mission corresponds with.

Date	**Mission**	**Nickname**
Edouard (2014)		
1200 UTC 15 September 2014	16	E16
1800 UTC 15 September 2014	17	E17
1800 UTC 16 September 2014	21	E21
1200 UTC 17 September 2014	24	E24
1800 UTC 17 September 2014	25	E25
Irma (2017)		
0000 UTC 04 September 2017	18	I18
0600 UTC 04 September 2017	19	I19
1200 UTC 04 September 2017	20	I20
1200 UTC 05 September 2017	24	I24
0000 UTC 06 September 2017	26	I26

Irma made landfall in Barbuda on 6 September as a category 5 hurricane. Further landfalls occurred on St. Martin, the British Virgin Islands, the Bahamas, Cuba, and Florida over the next several days. Irma finally became a remnant low by 0600 UTC 12 September, which dissipated 30 hours later.

Eight missions have TDR data available, however only five of the missions are used. The remaining missions occurred during analyses near landfall, and the forecast statistics were dominated by intensity differences caused by small track differences. This is not representative of the case as a whole and are therefore excluded from this study. The cycles

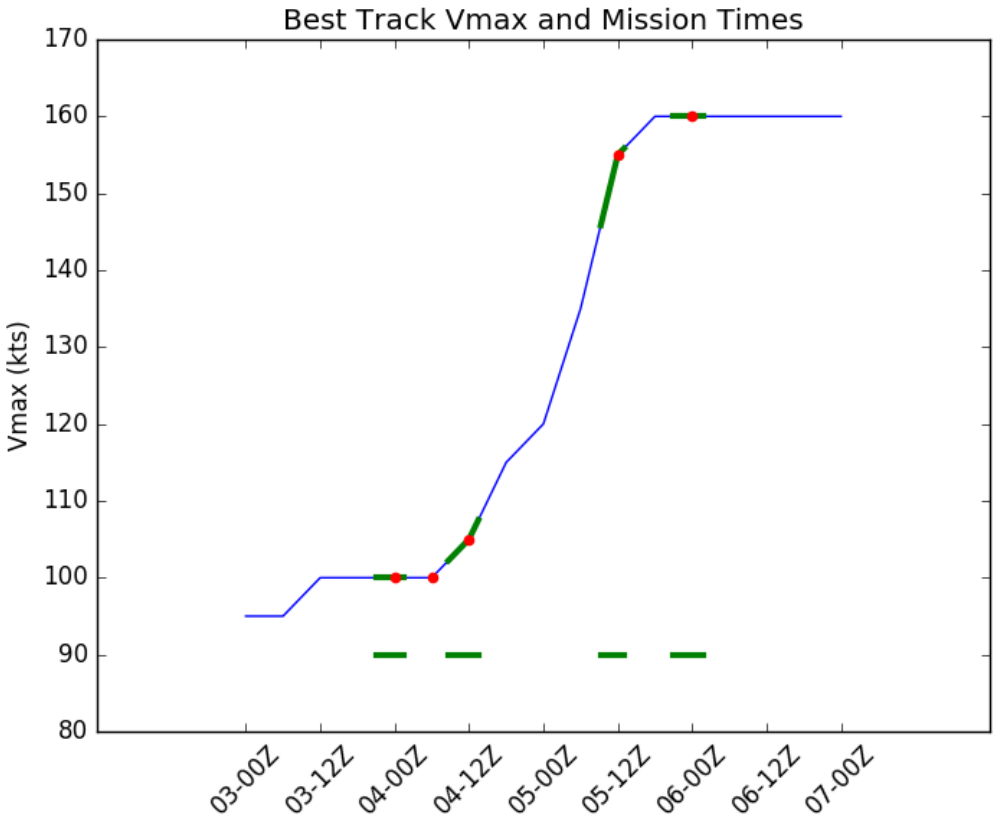


Figure 3.3: Irma (2017) Best track Vmax (blue) with TDR data availability overlaid (green). Center of missions marked with red dot.

used are the 18th, 19th, 20th, 24th, and 26th cycles valid at 0000 UTC 04 September, 0600 UTC 04 September, 1200 UTC 04 September, 1200 UTC 05 September, and 0000 UTC 06 September (Figure 3.3, Table 3.3).

3.4 Experiments

The baseline experiment 6H-3DEnVar uses the 6-Hourly 3DEnVar method as described in section (2.5). and is run for the entire storm to be used as the background for the other experiments. The model is continuously cycled starting on 1200 UTC 11 September for Edouard (2014), beginning when it developed into a tropical depression. For Irma (2017) continuous cycling began at 1200 UTC 30 August. A 6 hour spinup time is allowed prior to DA.

Table 3.4: Table of experiments.

Experiment Name	DA Window	Vortex Relocation Method	DA Method
6H-3DEnVar	6 hours	TCVitals	3DEnVar
6H-4DEnVar	6 hours	TCVitals	4DEnVar
1H-3DEnVar	1 hours	TCVitals	3DEnVar
6H-4DEnVar-sl	6 hours	Satellite	4DEnVar
1H-3DEnVar-sl	1 hours	Satellite	3DEnVar

Two experiments to test the hypothesis that 6-Hourly 4DEnVar and 1-Hourly 3DEnVar produce better analyses and forecasts than 6-Hourly 3DEnVar are conducted. The first is 6H-4DEnVar and the second is 1H-3DEnVar. The results of these two experiments will be compared to each other in order to determine if a difference exists between the two systems. 6H-4DEnVar uses the 6-Hourly 4DEnVar method as described in section (2.6).

It is only run for cycles in which TDR data is available. If no TDR data is available for the previous cycle the same background as 6H-3DEnVar is used, except that the relocation occurs every hour for hours 3-9 instead of just hours 3, 6, and 9. If TDR data is available for multiple cycles 6H-4DEnVar is continuously cycled, using the forecast from the previous 6H-4DEnVar cycle as the background, which is then relocated. 1H-3DEnVar uses the 1-Hourly 3DEnVar method described in section (2.7). Like 6H-4DEnVar, it is only run when TDR is available. If TDR data is available for consecutive missions 6H-4DEnVar is continuously cycled. When TDR data is not available for the previous mission, the 3-hour forecast from the previous cycles 6H-3DEnVar is used as the background for the first cycle, and undergoes VR prior to DA.

To investigate the impact of storm center location on the analysis and forecast two additional experiments are conducted. 6H-4DEnVar-sl and 1H-3DEnVar-sl are the same as 6H-4DEnVar and 1H-3DEnVar, respectively, except that during vortex relocation storm center locations are manually determined using satellite imagery. All members are relocated to this location instead of using TCVitals and EnSRF to determine the position of each member.

Chapter 4

Edouard Results

4.1 RMSE for Wind and Pressure

Maximum wind speed (Vmax) and minimum sea-level pressure (MSLP) forecasts are verified against the National Hurricane Center's (NHC) best track data for the five cases where TDR data is available. The root mean square error (RMSE) is calculated for each experiment at each forecast lead time to evaluate the typical error seen in the forecasts, with each experiment having 5 forecasts. An F-Test of equal variances is used to determine statistical significance of all RMSE based metrics and significance is indicated on figures. For Vmax and MSLP no times are significant at the 95% level due to a maximum sample size of 5. However, the results are still analyzed with emphasis on the need for a larger sample size. Both 1H-3DEnVar and 6H-4DEnVar produce smaller Vmax errors at 0-6 h than 6H-3DEnVar (Figure 4.1a). From forecast 12 h until 24 h, 6H-3DEnVar is similar to 6H-4DEnVar and has less Vmax error than 1H-3DEnVar. After 24 h, the Vmax RMSE is similar for all experiments. Multiple spuriously strong wind maxima in 6H-3DEnVar analyses (figure 4.1c) are responsible for the increased RMSE for 6H-3DEnVar during the first 6 hours of the forecast. Figure 4.1b shows 6H-3DEnVar has a larger RMSE for 0-12h than 1H-3DEnVar and 6H-4DEnVar before improving to produce the least error of all experiments for 24-48h. As is the case for Vmax, MSLP is spuriously strong in multiple cases resulting in large error for 6H-3DEnVar. Most forecasts cross the best track value for MSLP at some point in the first 48 hours (Figure 4.1d), with some overcorrecting, and even overcorrecting back to being too strong. As a result, the magnitude of the improvement is misleading during the 18-48h period, as the individual forecasts tend to have comparable error to other experiments when they are not crossing the best track value. Comparing

6H-4DEnVar and 1H-3DEnVar reveals initial Vmax RMSEs are similar but 12-24h 1H-3DEnVar Vmax RMSE is increased due to spindown that is occurring more frequently in 1H-3DEnVar. The tendency for spindown in 1H-3DEnVar is not seen in MSLP. 1H-3DEnVar has a larger MSLP RMSE in 0-48 h than 6H-4DEnVar due to a bias toward higher MSLP values.

Figure 4.1e shows the relationship between Vmax and MSLP. If the model is capturing the intensity evolution correctly the slope and the mean of each variable will be similar to best track. 6H-3DEnVar and 6H-4DEnVar have similar slopes to best track but each shows a bias along the slope, suggesting the strength in the model is biased but the relationship is reasonable. 1H-3DEnVar has a shallower slope than best track (significant at 83% level), suggesting the relationship between wind and pressure is too weak. This is primarily seen in the 0-30h analysis and forecast due to the spindown only occurring in Vmax. While 6H-3DEnVar and 6H-4DEnVar produce similar results, 1H-3DEnVar is worse in both slope and bias.

4.2 Structure Verification using TDR data

Calculation of a three-dimensional spatial correlation of the model wind and the TDR wind composite (figure 4.2) reveals that 6H-3DEnVar provides a worse wind analysis than other experiments. 6H-3DEnVar never produces the highest correlation and the mean of the correlation is about 4.7% smaller that of 1H-3DEnVar and 6H-4DEnVar. 6H-4DEnVar and 1H-3DEnVar produce similar quality wind analyses. 1H-3DEnVar provides the highest or second highest correlation in all cases. E21 and E24 show the largest difference between experiments, with 6H-4DEnVar having the highest correlation, 1H-3DEnVar the second highest, and 6H-3DEnVar is the lowest. Despite 6H-3DEnVar having the worst analyzed structure of the 3 experiments results were inconclusive when applied to forecast structure.

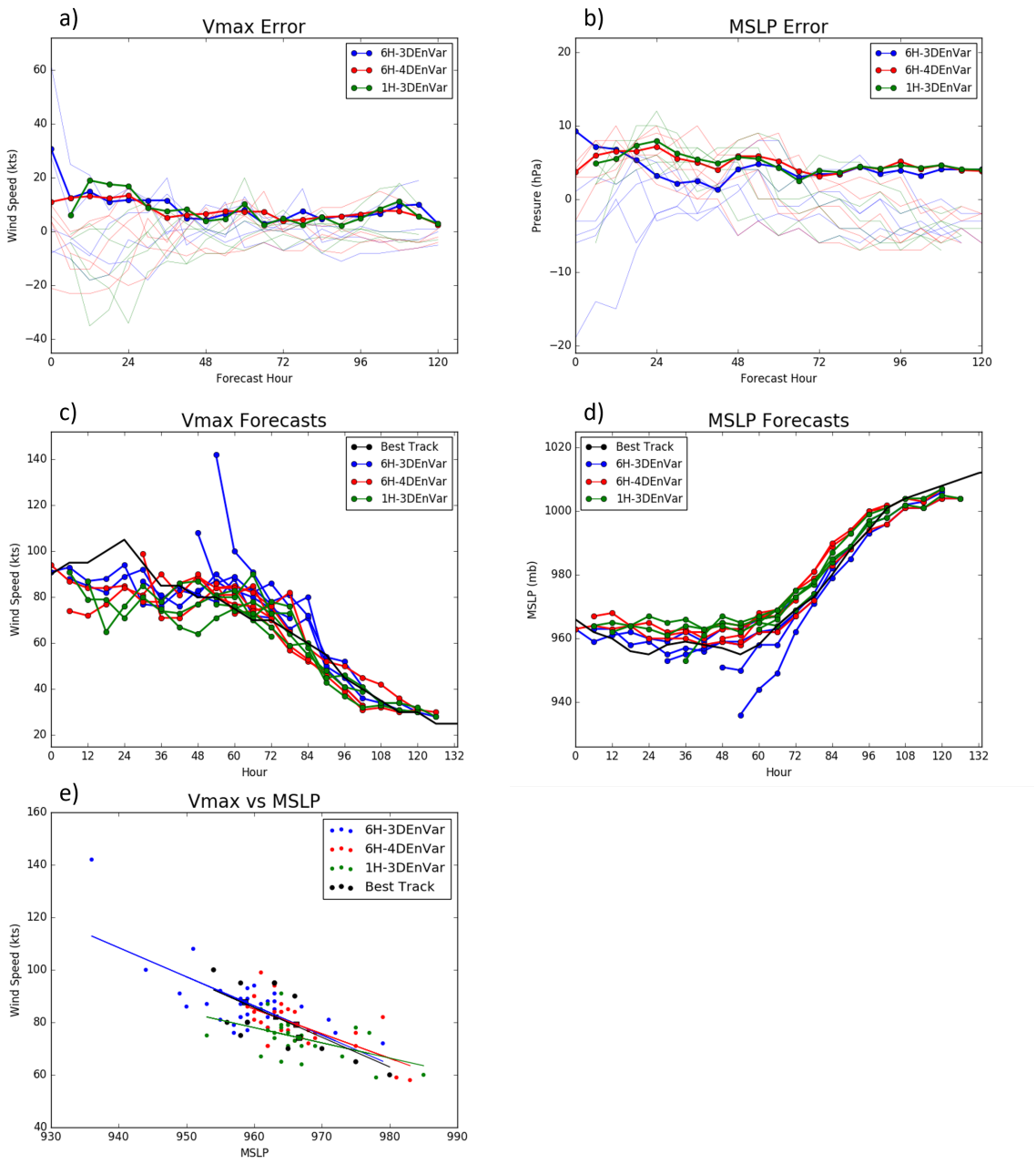
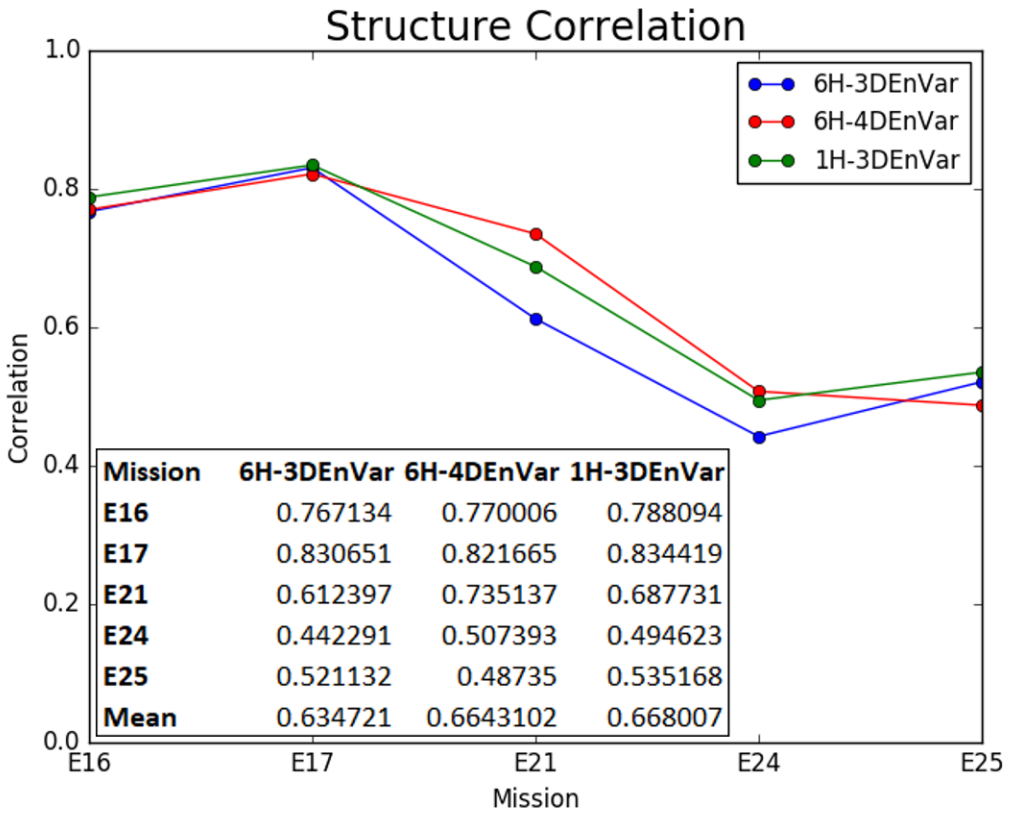


Figure 4.1: a) Vmax and b) MSLP RMSE. Thin lines are forecast minus best track for each individual forecast, thick lines are RMSE for each experiment. Individual c) Vmax and d) MSLP Forecasts plotted over Best track. Hour is hours from 1200 UTC September 15. Scatter plot of e) Vmax vs MSLP plotted over Best track. Red lines and dots are 6H-4DEnVar, blue lines are 6H-3DEnVar, green lines are 1H-3DEnVar, and black lines are best track. $p = 0.17$ for difference of 1H-3DEnVar and Best track slopes



Mission	6H-3DEnVar	6H-4DEnVar	1H-3DEnVar
E16	0.767134	0.770006	0.788094
E17	0.830651	0.821665	0.834419
E21	0.612397	0.735137	0.687731
E24	0.442291	0.507393	0.494623
E25	0.521132	0.48735	0.535168
Mean	0.634721	0.6643102	0.668007

Figure 4.2: 3-Dimensional spatial correlation of the model wind speed analysis with the TDR wind composite.

4.3 Structure verification vs TDR reflectivity

Model simulated reflectivity is compared to reflectivity from TDR. Due to the spinup time for hydrometeors in the model, only forecasts are compared. Only two missions observed the inner core near the valid time of forecasts. This discussion focuses on 6 hour forecast from valid at 1800 UTC 15 September (Figure 4.3) and the 18 hour forecast from valid at 1200 UTC 17 September (Figure 4.4). 6H-3DEnVar produces an eyewall that is stronger and broader than that of 6H-4DEnVar and 1H-3DEnVar in Figure (4.3), although observations show that reflectivity is too high in all experiments. Observations show a region with weak reflectivity counterclockwise from the primary band of precipitation in the eyewall, which 1H-3DEnVar fails to capture. Additionally, a rain shield is observed by TDR that is weaker than in all models, with 6H-3DEnVar and 6H-4DEnVar showing a less continuous but stronger shield than 1H-3DEnVar. Generally 1H-3DEnVar and 6H-4DEnVar produce a better forecast than 6H-3DEnVar. Comparing 6H-4DEnVar and 1H-3DEnVar reveals that 6H-4DEnVar simulates the inner core better with both the eyewall and the dry region present, however the outer regions are better represented by 1H-3DEnVar.

A double eyewall structure is observed on 1200 UTC 17 September (Figure 4.4). Both 6H-4DEnVar and 1H-3DEnVar simulate a weak secondary eyewall, however this structure is absent from 6H-3DEnVar. All 3 experiments show the dry region surrounding the eyewalls, however the radial distance of the rain band from the eyewall in 6H-3DEnVar is shorter than observed. Outside this band observations show a weaker second rainband that is present in all models. However, 6H-4DEnVar produces broad regions of $>$35 dBZ reflectivity that are not present in the observations. Comparing 1H-3DEnVar and 6H-4DEnVar, the outer eyewall wraps further around and is stronger than in observations in 1H-3DEnVar, while 6H-4DEnVar has a magnitude and size are similar to that of the observations. The dry region around the eyewall in 1H-3DEnVar is far broader than in 6H-4DEnVar, where both are already dryer than the observations. However, the outer rain band in 1H-3DEnVar is generally close to that observed, in contrast to the regions of $>$35 dBZ already mentioned

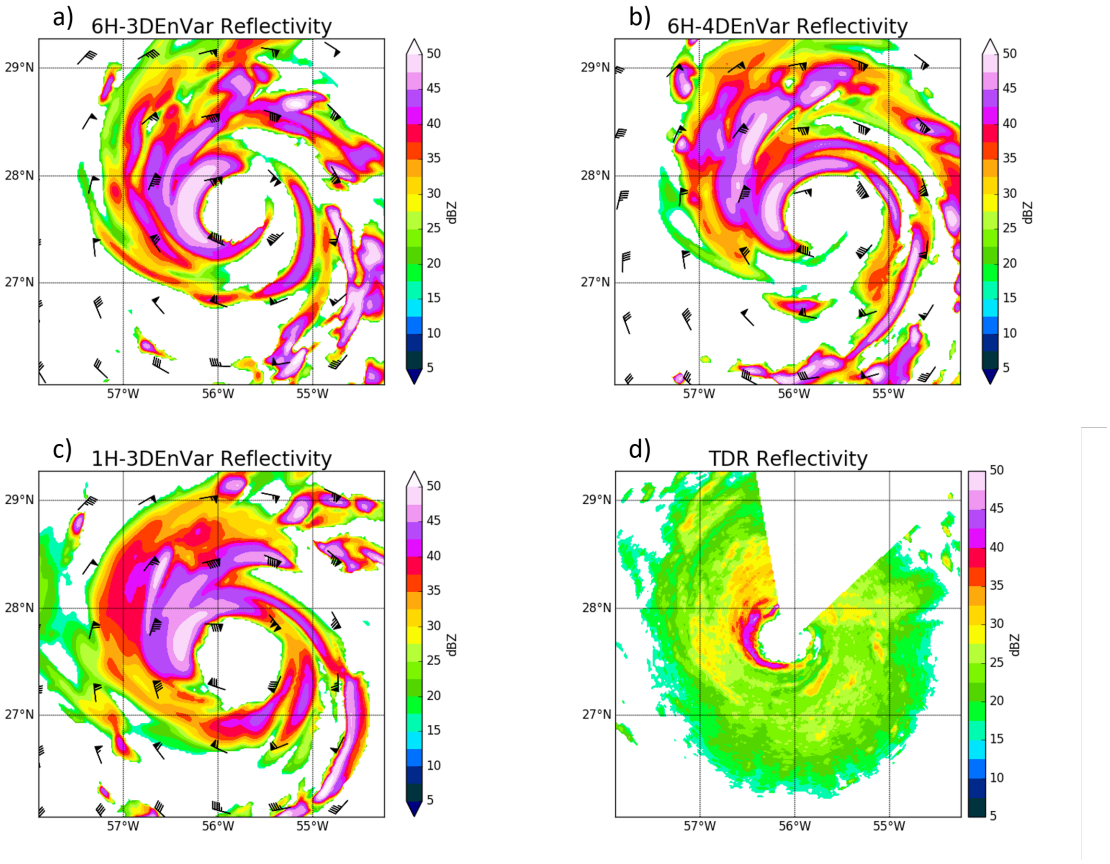


Figure 4.3: Simulated radar reflectivity valid 1800 UTC 15 September for a) 6H-3DEnVar, b) 6H-4DEnVar from forecast launched 1200 UTC 15 September, and c) 1H-3DEnVar from forecast launched at 1500 UTC 15 September (E16) and Observed and Observed TDR reflectivity valid at 1801 UTC 15 September 15

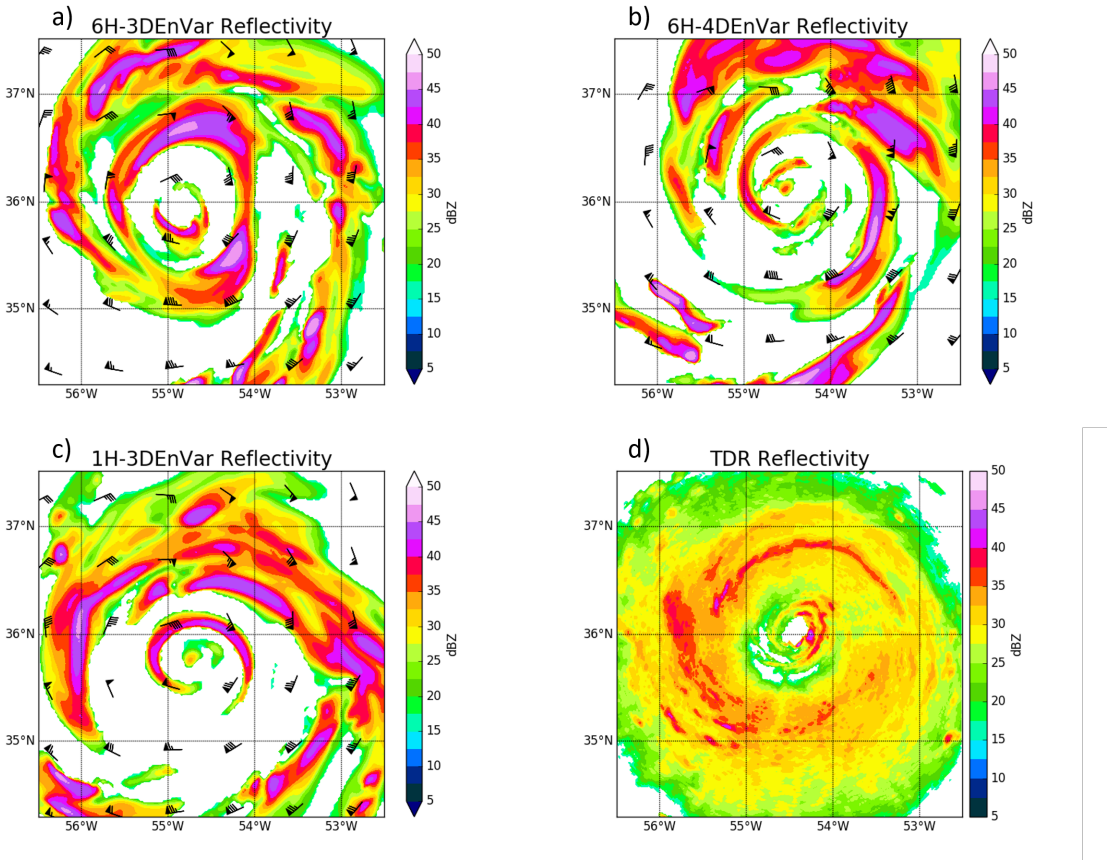


Figure 4.4: Simulated radar reflectivity valid 1200 UTC 17 September for a) 6H-3DEnVar, b) 6H-4DEnVar from forecast launched 1800 UTC 16 September and, c) 1H-3DEnVar from forecast launched at 2100 UTC 16 September (E21) d) and observed TDR reflectivity valid at 1312 17 UTC September

in the 6H-4DEnVar. In general both 6H-4DEnVar and 1H-3DEnVar produce better reflectivity forecasts than 6H-3DEnVar, with 6H-4DEnVar producing better inner core forecast, and 1H-3DEnVar producing better forecasts in the outer regions of the hurricane for both forecasts analyzed.

4.4 Verification against independent flight-level and SFMR data

In situ measurements made by the NOAA-P3 aircraft and the Stepped Frequency Microwave Radiometer (SFMR) provide independent measurements with which to verify the inner-core structure of the simulated TCs. SFMR wind speed and NOAA-P3 flight level (hereafter flight level) specific humidity (q) and temperature are used. RMSEs for analyses and model forecasts during each penetrating leg are calculated and the mean of the RMSEs for each leg in each mission are computed. Model output is recentered so the simulated TC center matches the observed TC location at the time of each leg, allowing for the direct comparison of the hurricane structure. The number of legs in each mission is listed in table 4.1. For forecasts only the first forecast time that SFMR and flight level data is available is verified, specifically E16's 6 hour forecast, E17's 24 hour forecast, E21's 18 hour forecast and E24's 6 hour forecast. No forecast is verified for E25 as there is no data available after analysis time.

SFMR wind speed verification of the analyses shows that 3D has a mean RMSE 40% larger than those of both 6H-4DEnVar and 1H-3DEnVar (Figure 4.5), and 6H-3DEnVar has a larger RMSE than both 1H-3DEnVar and 6H-4DEnVar for 4 of the 5 analyses. No notable difference in RMSE occurs between 1H-3DEnVar and 6H-4DEnVar. Figure (4.6) shows the increased error in 6H-3DEnVar can be attributed to a wider eye than observations with stronger wind maxima than other experiments. 1H-3DEnVar produces a storm with a properly sized eye, but 6H-4DEnVar captures the wind speed better except in the eyewall. For the forecast, 3 of 4 analyses show the 6H-4DEnVar and 1H-3DEnVar produce better results than 6H-3DEnVar (Figure 4.7). Similarly, the mean RMSE of all cases shows the

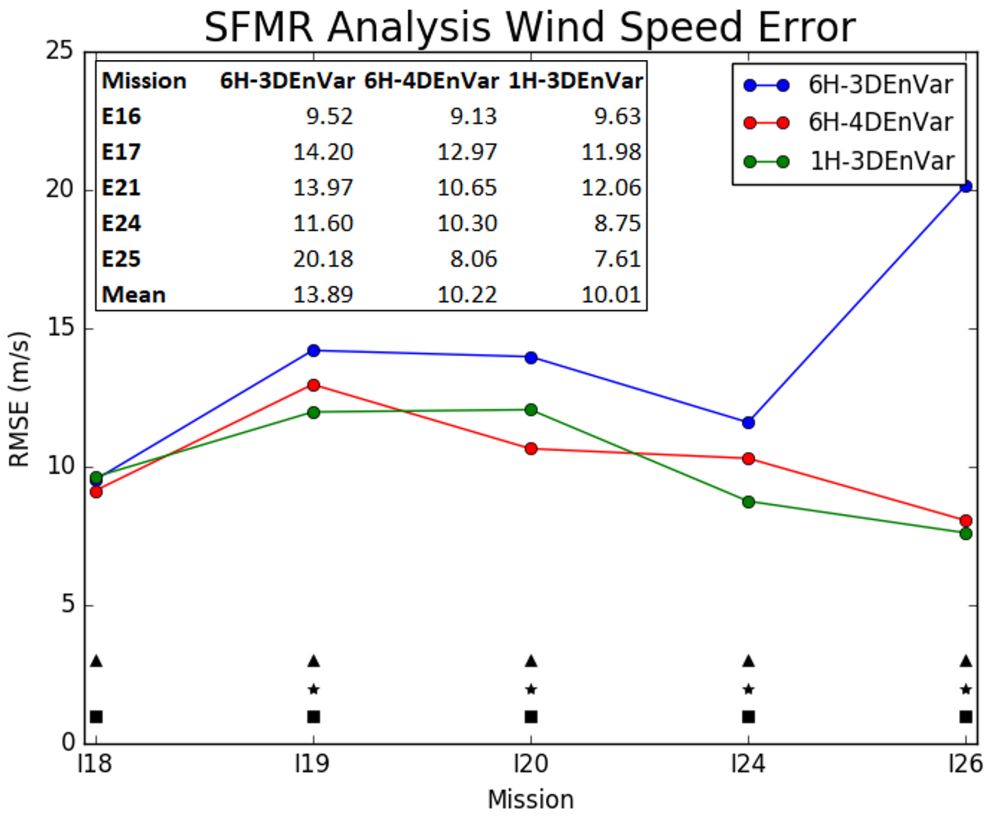


Figure 4.5: Wind speed Analysis error as verified using Stepped Frequency Microwave Radiometer (SFMR). The value for each experiment is combined RMSE of the penetrating legs for each mission. Black triangles indicate statistically significant difference between 6H-3DEnVar and 6H-4DEnVar at 95%level, black stars indicate statistically significant differences between 6H-3DEnVar and 1H-3DEnVar at 95% level, and black squares indicate statistically significant differences between 6H-4DEnVar and 1H-3DEnVar at 95% level.

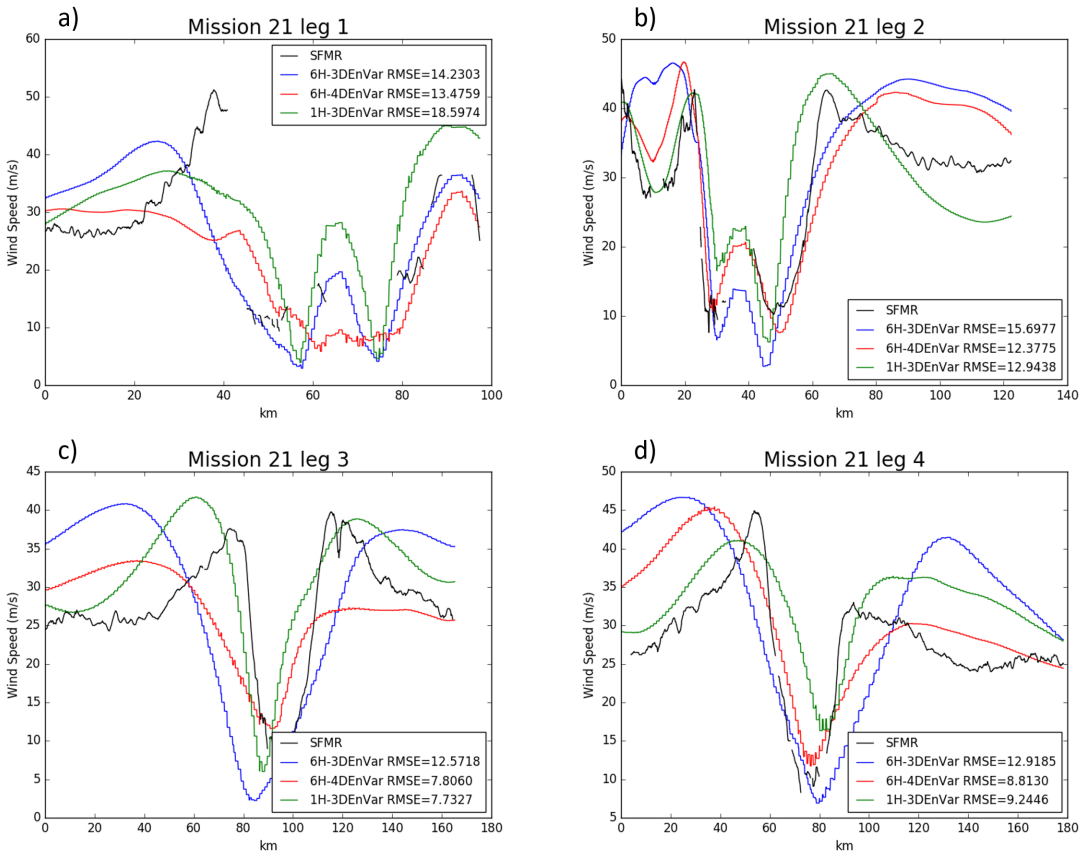


Figure 4.6: SFMR wind speed (black) and model surface wind speed along the SFMR flight track for each leg. Analysis valid at 1800 UTC 16 September 2014

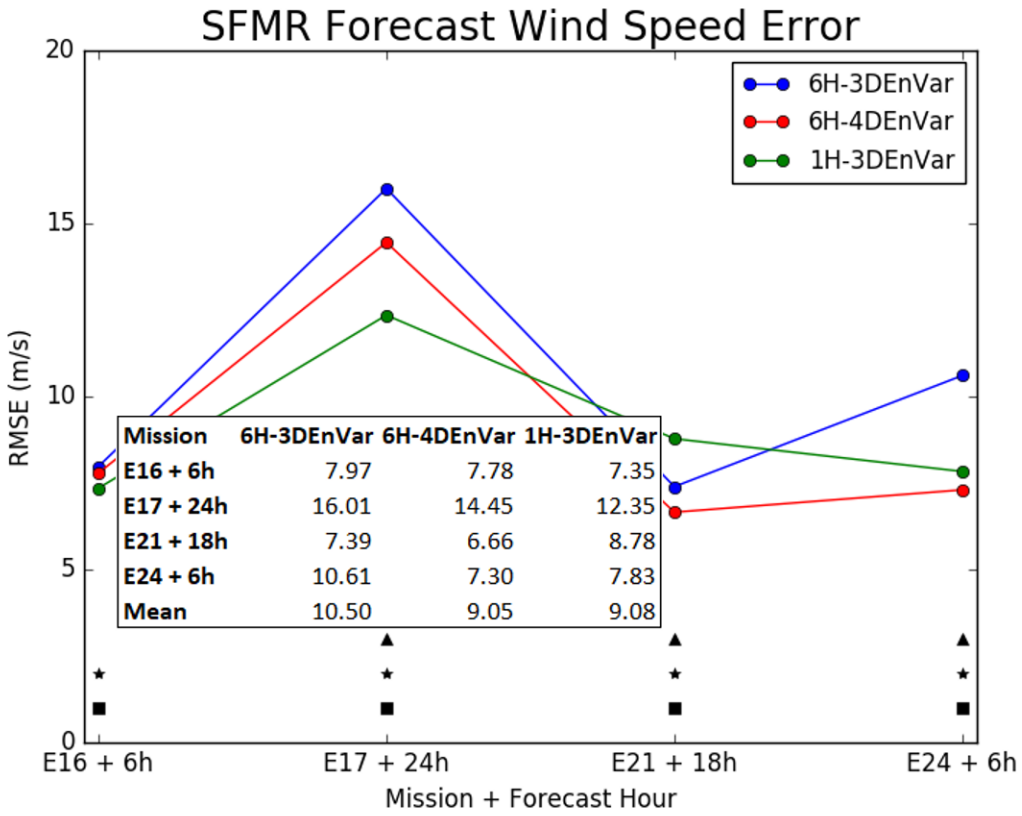
Table 4.1: Number of SFMR and flight level penetrating legs used for each mission. Dates and mission numbers as in table 3.3

.

Date	Mission	Number of Legs
Edouard (2014)		
1200 UTC 15 September 2014	16	5
1800 UTC 15 September 2014	17	2
1800 UTC 16 September 2014	21	4
1200 UTC 17 September 2014	24	2
1800 UTC 17 September 2014	25	1

largest RMSE for 6H-3DEnVar, although the difference is smaller than for the analyses. 6H-3DEnVar produced the largest wind speed of any experiment in all legs (Figure 4.8). Because all experiments result in similarly sized eyes, the primary differences are caused by the strength of the eyewall.

Flight level data only showed notable differences for analyses. In 4 of the 5 cases, as well as in the overall mean, 6H-3DEnVar produces a worse specific humidity analysis than 1H-3DEnVar and 6H-4DEnVar (Figure 4.9). Figure (4.10) shows that 6H-3DEnVar produces a specific humidity analysis that is more moist than other experiments and observations. There is not a notable difference between 1H-3DEnVar and 6H-4DEnVar. Model temperature analyses were also verified. While 6H-3DEnVar performed worse than 6H-4DEnVar in 4 of 5 missions, 1H-3DEnVar performed worse than 6H-3DEnVar in 4 missions (Figure 4.9). Similarly, 6H-4DEnVar performed better than 1H-3DEnVar in 4 missions. The mean RMSE for 6H-4DEnVar was smallest with 1H-3DEnVar being the largest. All experiments have a bias toward a horizontally large warm core, however 1H-3DEnVar



Mission	6H-3DEnVar	6H-4DEnVar	1H-3DEnVar
E16 + 6h	7.97	7.78	7.35
E17 + 24h	16.01	14.45	12.35
E21 + 18h	7.39	6.66	8.78
E24 + 6h	10.61	7.30	7.83
Mean	10.50	9.05	9.08

Figure 4.7: Wind speed forecast error as verified using Stepped Frequency Microwave Radiometer (SFMR). The value given for each experiment is the combined RMSE of the penetrating legs for each mission. Black triangles indicate statistically significant difference between 6H-3DEnVar and 6H-4DEnVar at 95%level, black stars indicate statistically significant differences between 6H-3DEnVar and 1H-3DEnVar at 95% level, and black squares indicate statistically significant differences between 6H-4DEnVar and 1H-3DEnVar at 95% level.

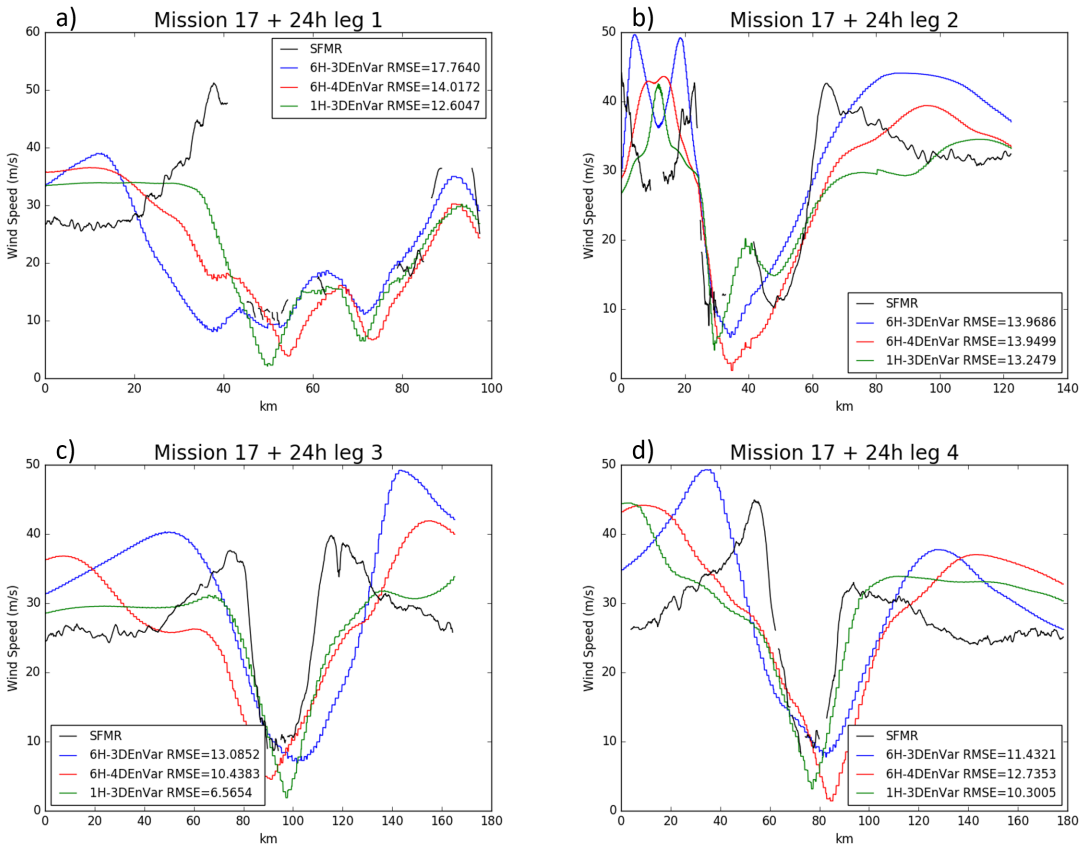


Figure 4.8: SFMR wind speed (black) and model surface wind speed along the SFMR flight track for each leg. Forecast valid at 1800 UTC 16 September 2014 from forecast launched 18 hours prior

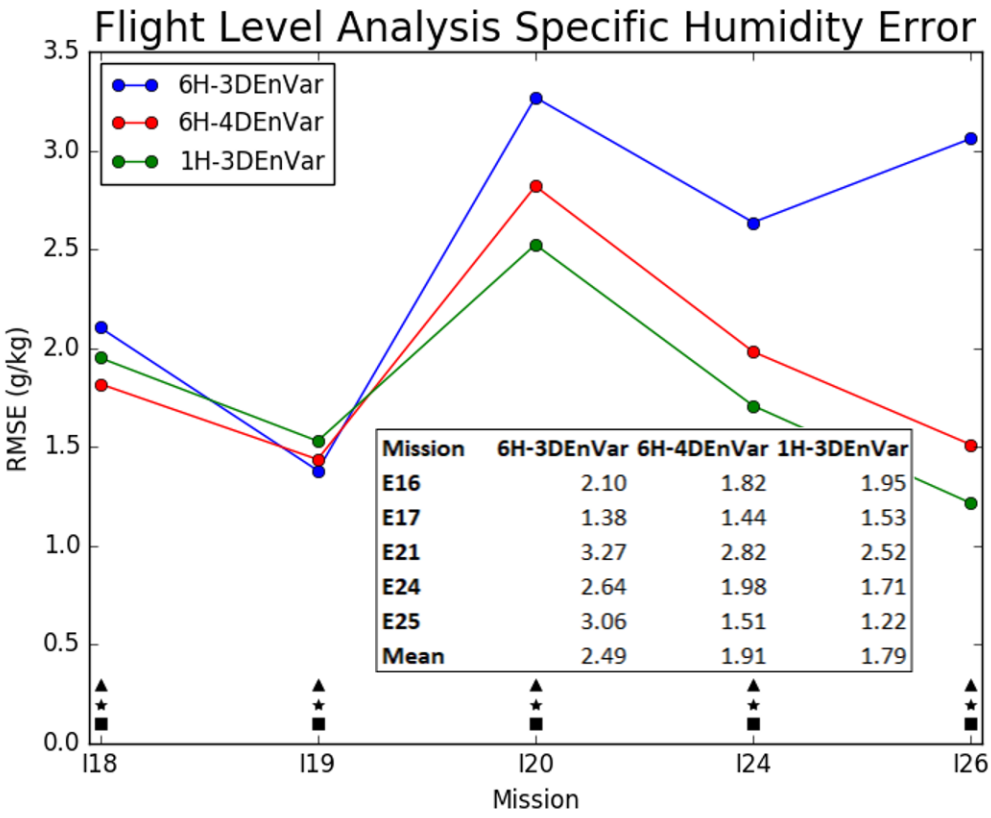


Figure 4.9: Specific humidity analysis error as verified using NOAA-P3 flight level data. The values for each experiment is the combined RMSE of the penetrating legs for each mission. Black triangles indicate statistically significant difference between 6H-3DEnVar and 6H-4DEnVar at 95%level, black stars indicate statistically significant differences between 6H-3DEnVar and 1H-3DEnVar at 95% level, and black squares indicate statistically significant differences between 6H-4DEnVar and 1H-3DEnVar at 95% level.

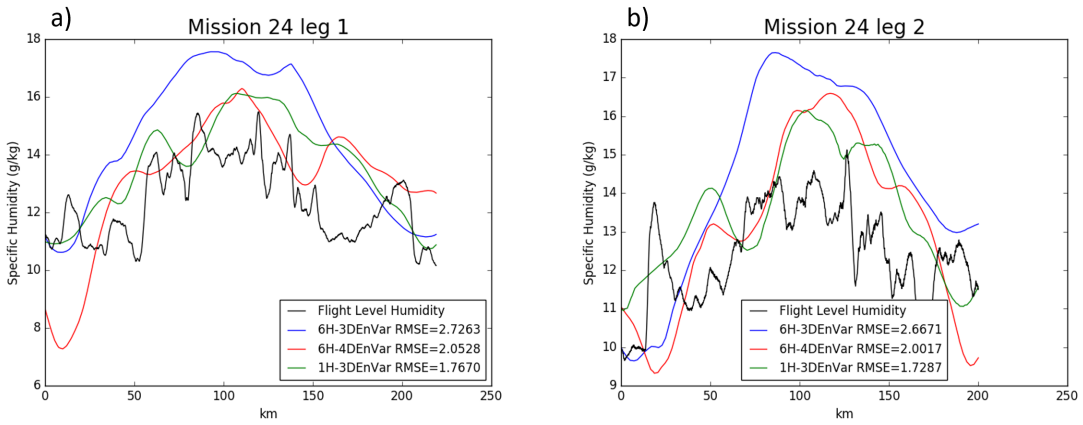


Figure 4.10: Flight level specific humidity (black) and model surface wind speed along the NOAA-P3 flight track for each leg. Analysis is valid at 1200 UTC 17 September 2014

tends to be consistently too warm near the center (Figure 4.10). Outside the inner core there is not a notable difference in many legs.

4.5 Diagnosis of the spindown issue in 1H-3DEnVar

Frequent spindown is observed in 1H-3DEnVar (figure 4.1c), causing a degradation of the Vmax verification early in forecasts. Spindown occurs when Vmax decreases greater than 5 m s^{-1} $(6\ h)^{-1}$ during the first 6-12 h of model integration (Lu and Wang 2019) with no such weakening occuring in observations. An investigation into the cause of the spindown is discussed in this section. First, model stability is evaluated due to the limitations discussed in chapter 1 pertaining to shortening the DA window. To further investigate the cause of the inferior performance of 1H-3DEnVar a representative case on 1800 UTC 15 September 2014 case is chosen (E17) due to the spindown in 1H-3DEnVar that is typical of all cases and more in-depth diagnostics are performed. This section provides results and a discussion of the investigation.

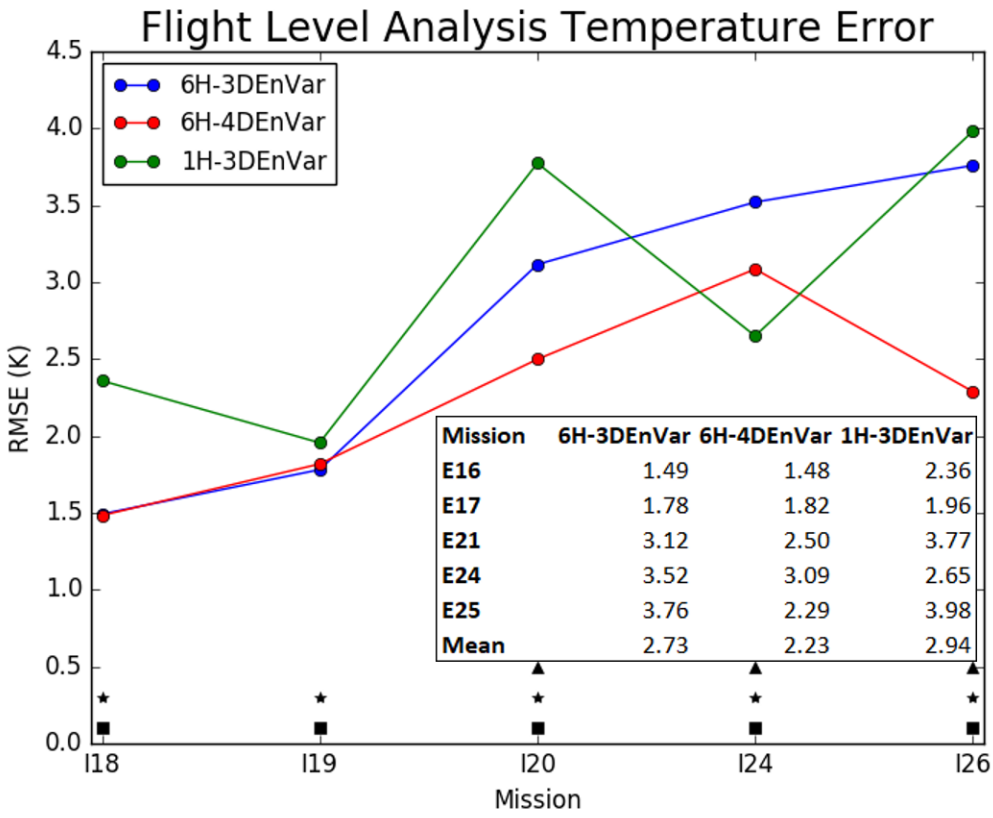


Figure 4.11: Temperature analysis error as verified using NOAA-P3 flight level data. The values for each experiment is the combined RMSE of the penetrating legs for each mission. Black triangles indicate statistically significant difference between 6H-3DEnVar and 6H-4DEnVar at 95%level, black stars indicate statistically significant differences between 6H-3DEnVar and 1H-3DEnVar at 95% level, and black squares indicate statistically significant differences between 6H-4DEnVar and 1H-3DEnVar at 95% level.

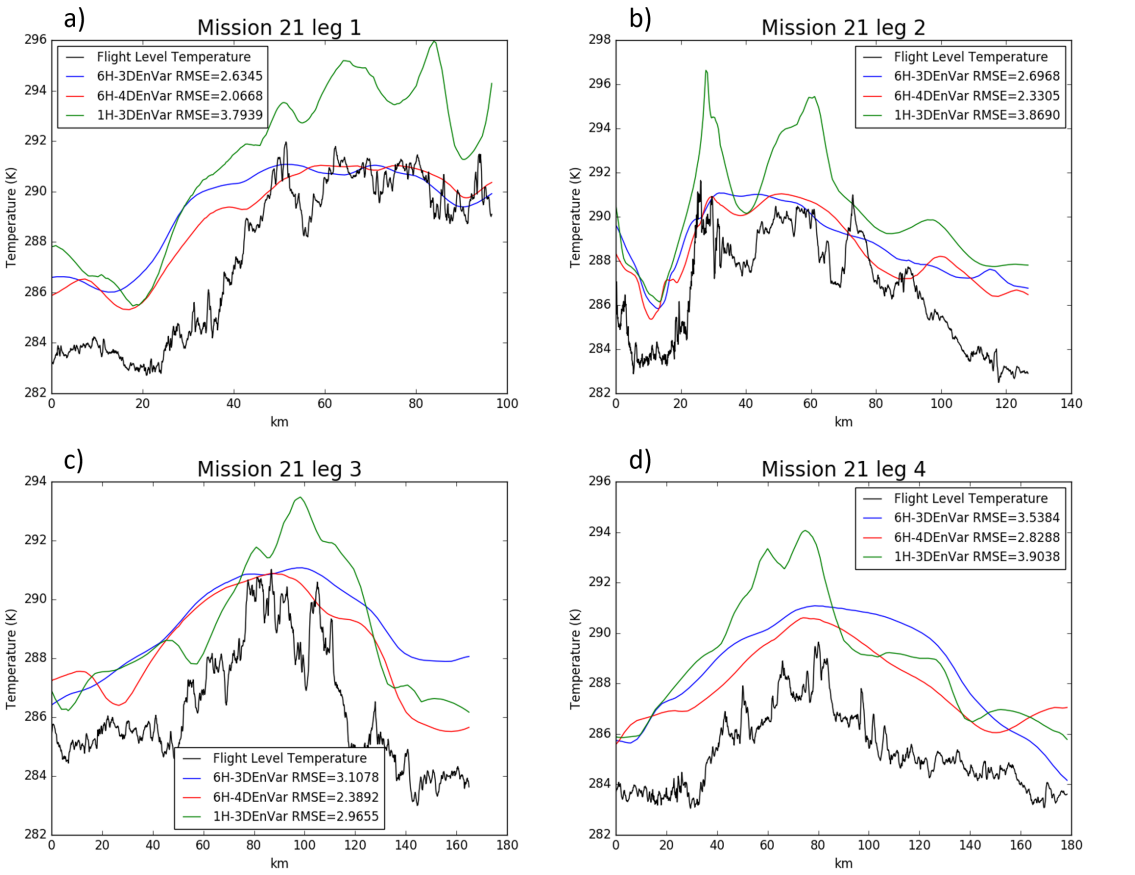


Figure 4.12: Flight level temperature (black) and model surface wind speed along the NOAA-P3 flight track for each leg. Analysis is valid at 1800 UTC 16 September 2014

4.5.1 Model Stability.

Mean absolute pressure tendency (Mdpdt) is calculated following eq (1) to diagnose model instability.

$$Mdpdt = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{dpdt_{ij}}{m*n}, \qquad dpdt = |\frac{p_t - p_{t-1}}{\Delta t}| \tag{4.1}$$

where p is pressure, t is time, m and n are the number of grid points along each axis in the subdomain being averaged over. Data assimilation can produce an analysis that is not dynamically balanced, resulting in a pressure wave that shows up as rings of high dpdt propagating outward from a source. These waves are present in this experiment (Figure 4.13). Averaging over the outermost domain reveals that dpdt in 1H-3DEnVar increases steadily over time (not shown), while the 6H-4DEnVar and 6H-3DEnVar do not show a similar increase. However, when averaging over the region corresponding to the innermost domain, the peak magnitude of dpdt is similar for all experiments. No growth is seen in 1H-3DEnVar and Mdpdt returns to similar baseline levels as 6H-3DEnVar and 6H-4DEnVar before each analysis. The wave propagates outside of the inner domain in 45 minutes (seen as a return to baseline values of Mdpdt in figure 4.13) therefore this instability is not impacting future DA.

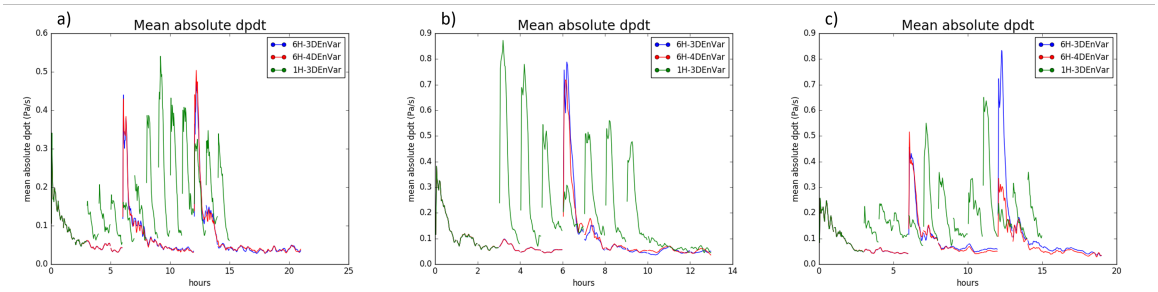


Figure 4.13: Mean absolute pressure tendency averaged over a box approximating the path of the inner domain during E16 and E17 (a), E21 (b), and E24 and E25 (c)

4.5.2 Moisture Increments and model spread

Two regions appear in the 1H-3DEnVar analyses during E17 where a negative moisture increment occurs in the analysis for several consecutive cycles (Figure 4.14) at 2 km above ground level (agl). The first is 3° northeast of Edouard where drying occurs in several consecutive cycles. The second is north and northwest of Edouard, which is not seen in 6H-4DEnVar and has a smaller magnitude of drying than the former region. The later region of dry air is advected into the inner core during the free forecast (not shown) intruding on and weakening convection leading to spindown. Comparing simulated satellite to satellite imagery confirms that the 1H-3DEnVar analysis is too dry, whereas the 6H-4DEnVar analysis is not nearly as dry (Figure 14.15).

A difference in storm location is observed between the model background and the satellite imagery (Figure 4.16a). Due to a strong wind gradient in the eyewall, radar radial velocity data located along this wind gradient near can result in large innovations during DA if the gradients are not aligned accurately. Large innovations from the GSI plotted on the model background in figure (4.16b) confirm that large innovations occur in regions along the axis of dislocation. This analysis shows that the largest innovations occur near the eyewall along the axis of dislocation between the background and observations, confirming that this is the cause of the large innovations. Due to the location in the storm and the dislocation of the eye the observations will result in a U innovation <0. Ensemble cross-covariance between a wind observation at a point representative of large innovations with specific humidity (Figure 4.16c) shows that regions of positive cross-covariances occur in regions of drying suggesting the drying is a result of this dislocation.

A reduction in storm location spread in the ensemble (Figure 4.17) occurs over time for 1H-3DEnVar leading to an underdispersive ensemble. Due to the EnSRF process used to determine the relocated centers the underdispersive ensemble results in less correction than expected in an optimal analysis. The 6H-4DEnVar ensemble is also underdispersive but the magnitude of the underdispersiveness and the ratio of the two lines is smaller. As

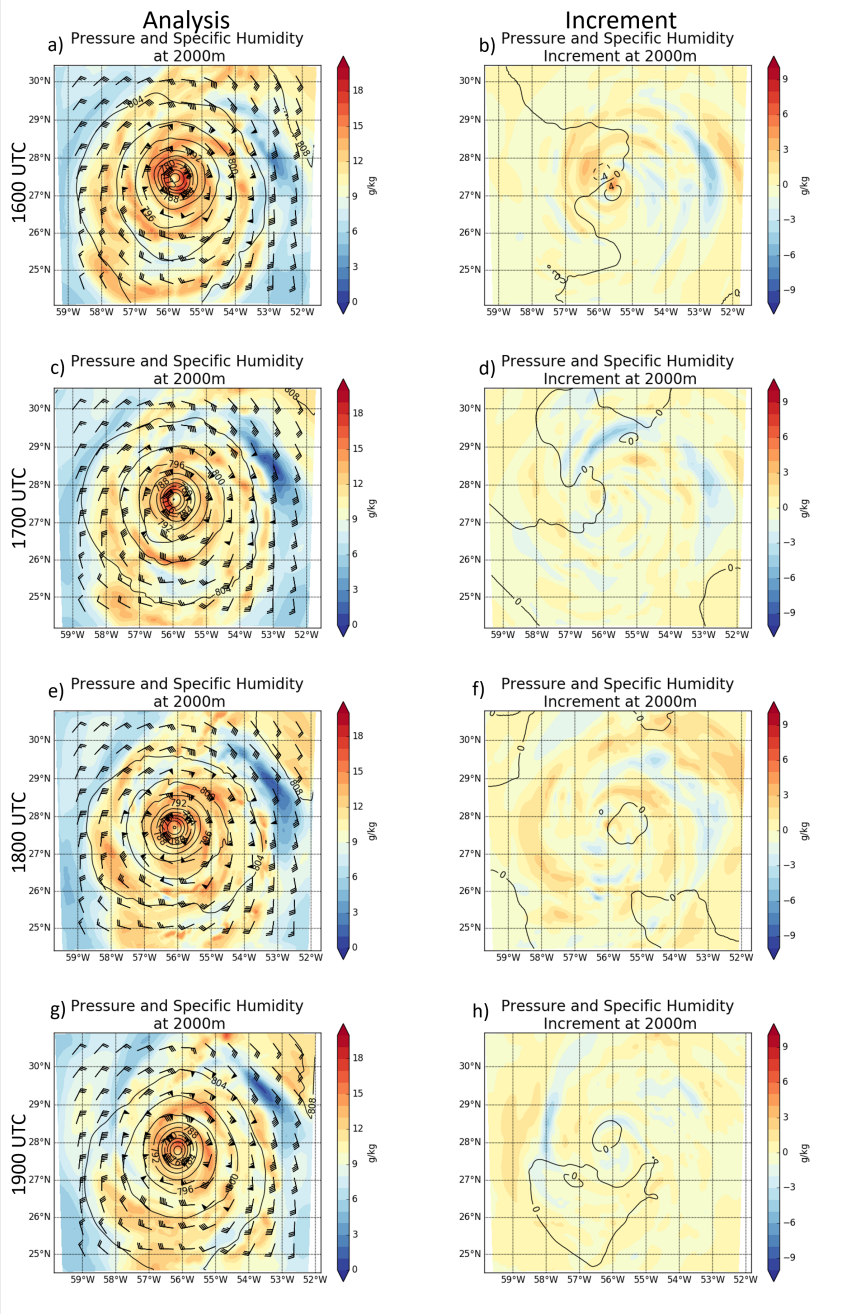


Figure 4.14: Left column: Specific humidity analysis (colors) for E17 1H-3DEnVar at a) 1600 UTC, c) 1700 UTC, e) 1800 UTC, and g) 1900 UTC with pressure (contours) every 4 hPa and wind barbs overlaid. Right column: Specific humidity (colors) and pressure (contours) increments for E17 1H-3DEnVar at b) 1600 UTC, d) 1700 UTC, f) 1800 UTC, and h) 1900 UTC. All figures are for a high of 2 km agl.

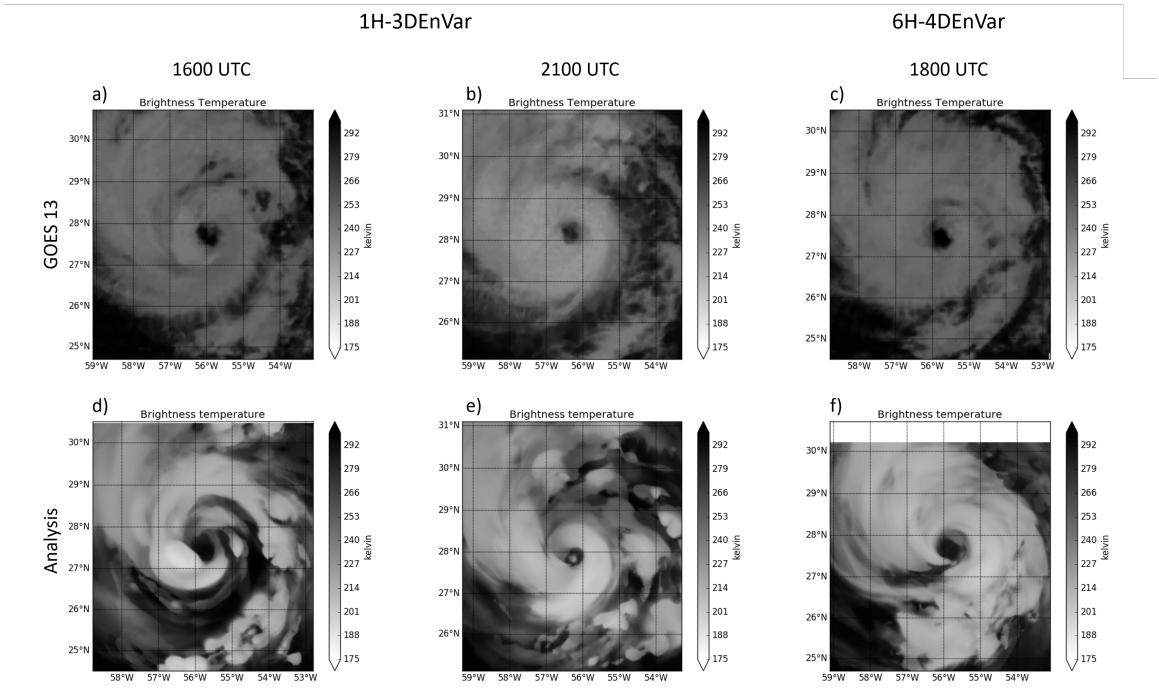


Figure 4.15: Observed satellite from GOES 13 in a), c) and e) at 1600 UTC, 2100 UTC, and 1800 UTC 15 September 2014 respectively, simulated brightness temperature from the 1H-3DEnVar analysis in d) and e) at 1600 UTC and 2100 UTC respectively, and f) 6H-4DEnVar analysis at 1800 UTC.

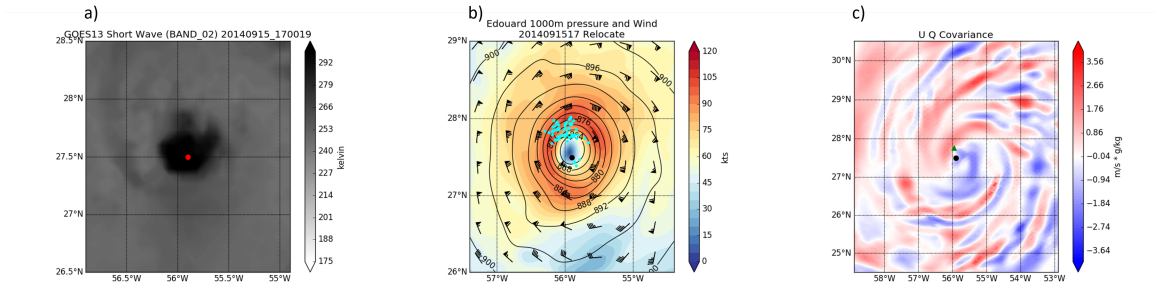


Figure 4.16: a) Satellite imagery with TCVital (red) overlaid. b) 1H-3DEnVar 1700 UTC wind analysis (colors and barbs) overlaid with wind speed increments >10m/s (blue dots) and TCVital (red). c) 1700 UTC 1H-3DEnVar Ensemble cross-covariance between U wind speed and specific humidity (Q) for a sample observation (green triangle) corresponding to the region of large innovations. TCVital overlaid (Black square)

a result, 6H-4DEnVar does not experience the same problems with poor location updates during the relocation step.

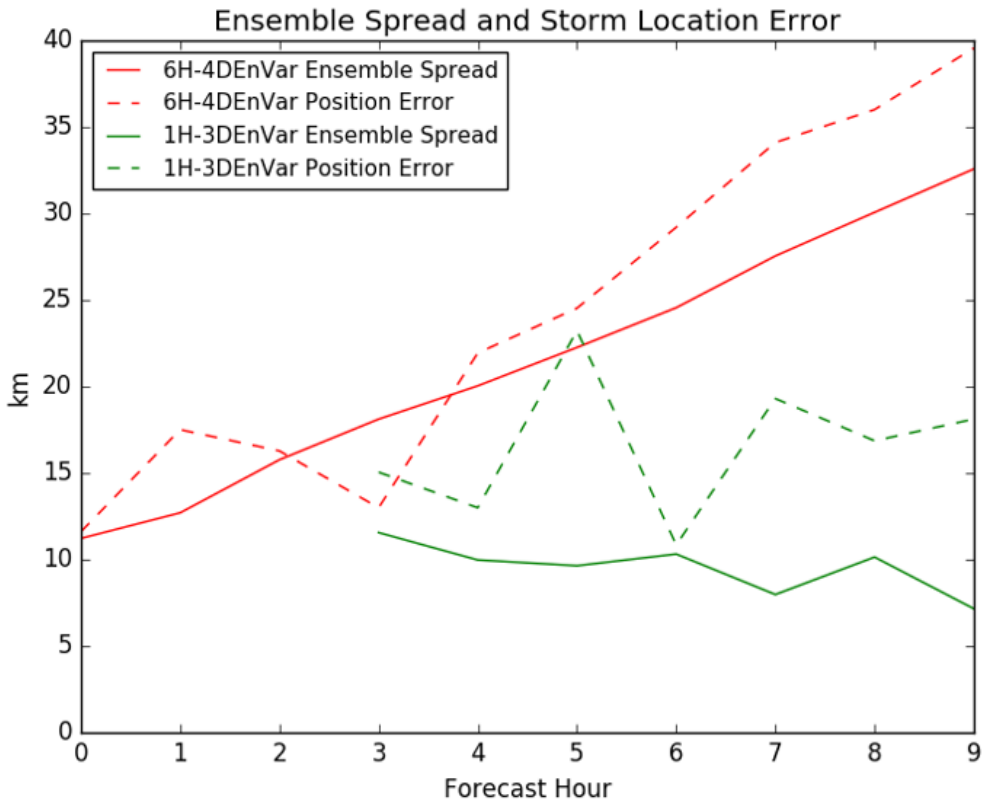


Figure 4.17: Ensemble spread (solid lines) and mean background position error (dotted lines) for 6H-4DEnVar (red) and 1H-3DEnVar (green) by forecast hour

To confirm that poor location updates are degrading the quality of forecast 1H-3DEnVar and 6H-4DEnVar are rerun using satellite derived storm centers for the relocated background instead of the EnSRF and TCVitals derived centers. Figure 4.18 shows that while some spindown still occurs in 1H-3DEnVar, the magnitude is reduced such that that error remains comparable to that of 6H-4DEnVar for 12-24h. This is consistent with the hypothesis that position error is causing the reduced performance of 1H-3DEnVar in this time

frame. No improvement (and even a slight degradation) is seen in 1H-3DEnVar MSLP forecast. However, a reduction in MSLP error is observed in 6H-4DEnVar for 6-48h. The improvement appears as be a reduction in pressure spindown, suggesting that the 6H-4DEnVar system could still see further improvement with improved VR and VM strategies.

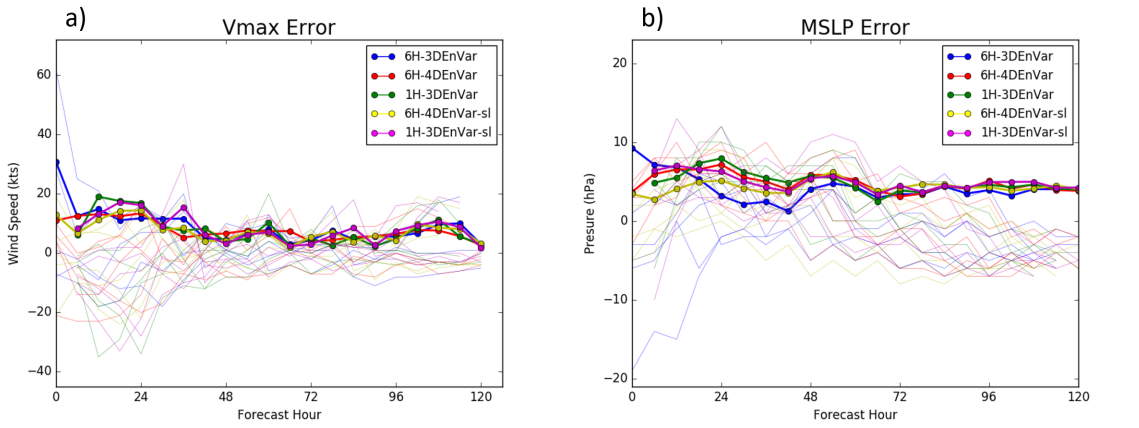


Figure 4.18: Vmax (a) and MSLP (b) RMSE (thick lines) and individual forecast errors (thin lines). Experiments with -sl use satellite derived locations during VR

Chapter 5

Irma Results

5.1 RMSE for wind and pressure

RMSE is calculated for Vmax (figure 5.1a) and MSLP (figure 5.1b) against best track data for Irma. The Vmax RMSE reveals that for early lead times 1H-3DEnVar has larger error than 6H-3DEnVar before becoming comparable for 24-36h. 6H-4DEnVar is comparable to 6H-3DEnVar for 0-48h. An improvement is seen in both 1H-3DEnVar and 6H-4DEnVar such that during the 48-90h time frame both experiments consistently exhibit smaller error than 6H-3DEnVar. However the error difference is small. During 0-12h, 1H-3DEnVar has a larger RMSE than 6H-4DEnVar. From 36-80h, 1H-3DEnVar has a smaller error than 6H-3DEnVar, with both experiments having similar RMSE's for remaining times. RMSE for MSLP reveals 1H-3DEnVar has smaller errors than 6H-3DEnVar for 12-54h but has larger errors for later times (figure 5.1b). MSLP error for 6H-4DEnVar are comparable to 6H-3DEnVar prior to 84h, but are typically worse at later times. Due to differences in storm tracks in individual free forecasts, 1H-3DEnVar and 6H-4DEnVar see degraded Vmax and MSLP RMSEs at longer forecast lead times. Large errors are introduced when an experiment produces a landfalling hurricane when none was present in the observations, or an experiment's forecast track stays over the ocean when the observations show the storm making landfall. Prior to 60h 1H-3DEnVar shows smaller error than 6H-4DEnVar, after 60h 1H-3DEnVar shows larger RMSE than 6H-4DEnVar.

Both 6H-4DEnVar and 6H-3DEnVar exhibit spindown in cases with the largest wind speeds (Figure 5.1c). While not shown in Figure 5.1c, 1H-3DEnVar spins down in the first 3 hours in 4 of 5 cases. During the peak intensity all experiments exhibit a weak bias. For MSLP, there is a consistent weak bias in all experiments during peak intensity. However, 1H-3DEnVar shows a smaller bias than either 6H-4DEnVar and 6H-3DEnVar, consistent

with the smaller RMSE. Comparing the relationship between pressure and Vmax (Figure 5.1e) reveals that the relationship between Vmax and RMSE is too weak in all experiments (significant at the 95% level for each experiment relative to best track) during the first 30 hours, as shown by the shallower slope than in the best track. All 3 experiments show a similar slope (The differences are not statistically significant at the 95% level), however 1H-3DEnVar is an outlier with stronger MSLP for a given wind speed than both 6H-3DEnVar and 6H-4DEnVar.

5.2 Structure Correlation

A three-dimentional structure correlation is calculated between the model wind analyses and the TDR wind speed (Figure 5.2). 1H-3DEnVar correlates 24% better to observations than both 6H-3DEnVar and 6H-4DEnVar and has the largest correlation in 4 of 5 cases. This increase in the mean of the correlation is largely due to I24, however, if this case is removed the correlations are similar with 1H-3DEnVar still being the largest. The difference between 6H-3DEnVar and 6H-4DEnVar is small.

5.3 Structure verification using TDR reflectivity

Simulated reflectivity is compared to TDR reflectivity at 1200 UTC September 04 and 0000 UTC September 06, analysis time for I20 and I26 respectively. The I19 6h forecast (figure 5.3) and the I24 12h forecast (figure 5.4) are compared at 3000m and 2000m respectively. Both 6H-4DEnVar and 1H-3DEnVar show the small inner eyewall accurately, where 6H-3DEnVar fails to produce one (Figure 5.3). The rain bands in the inner core are too strong in all experiments. 1H-3DEnVar produces precipitation in the eye that is not present in the observations. Outside the inner core, both 1H-3DEnVar and 6H-4DEnVar capture the rain bands better than 6H-3DEnVar. However, due to attenuation it is difficult to quantify the magnitude of the reflectivity produced by the model in these bands. At 0000 UTC

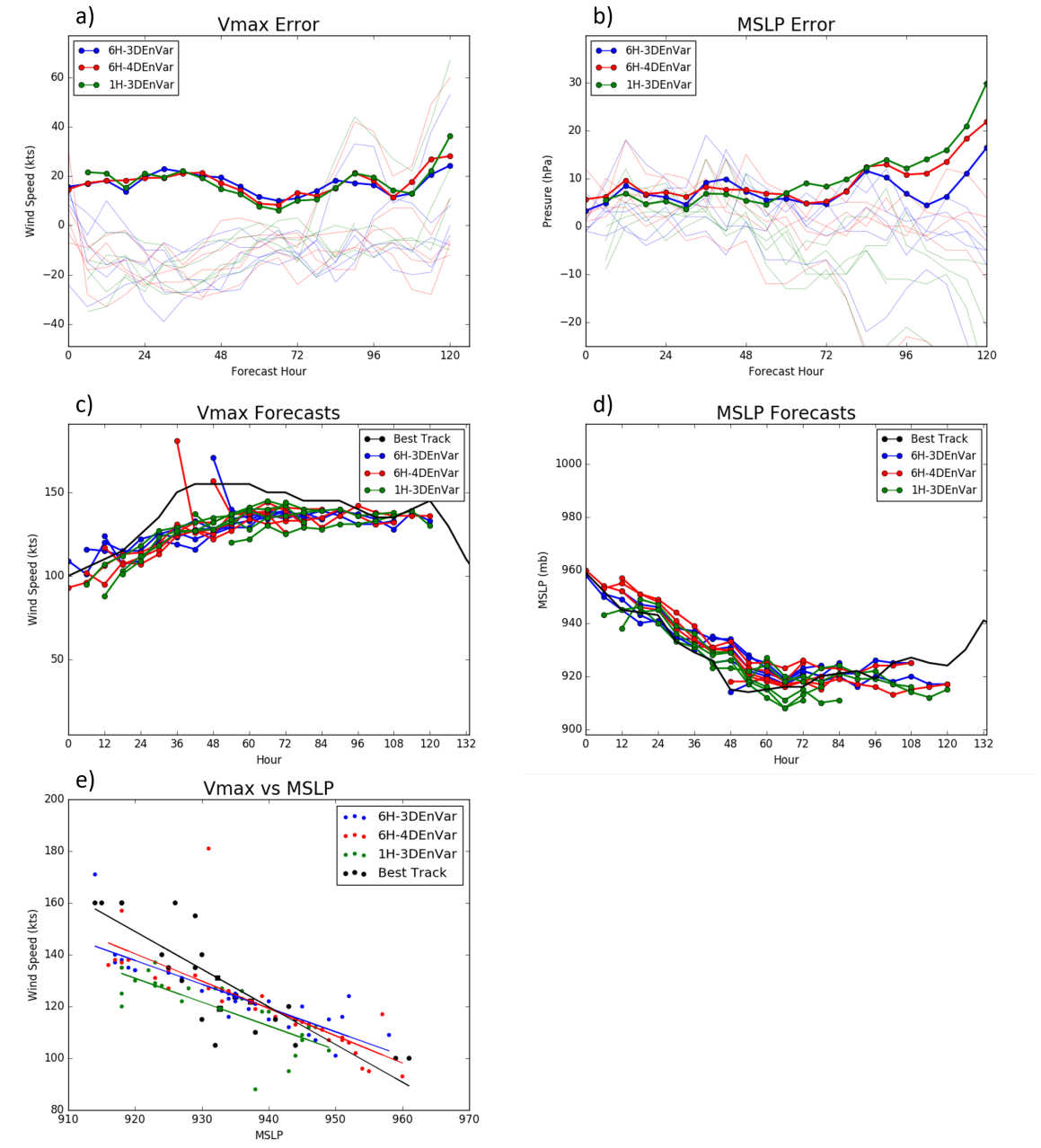
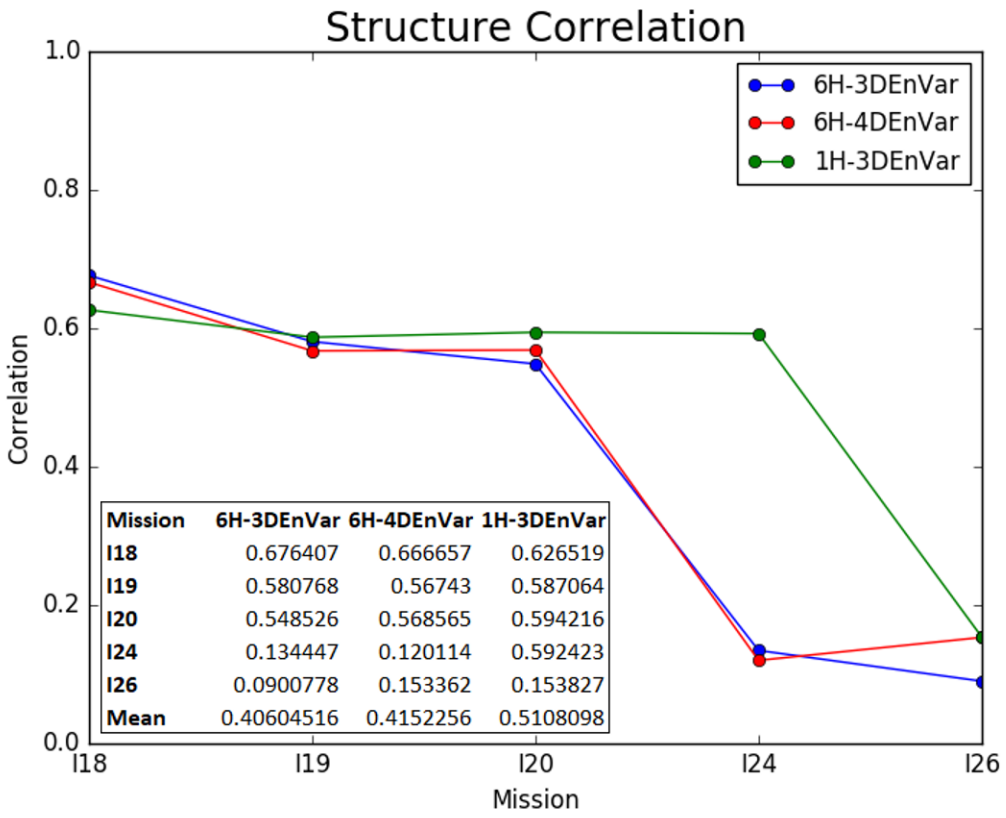


Figure 5.1: As in 4.1, but for Irma (2017), and all slopes are different than best track at the 95% significance level.



Mission	6H-3DEnVar	6H-4DEnVar	1H-3DEnVar
I18	0.676407	0.666657	0.626519
I19	0.580768	0.56743	0.587064
I20	0.548526	0.568565	0.594216
I24	0.134447	0.120114	0.592423
I26	0.0900778	0.153362	0.153827
Mean	0.40604516	0.4152256	0.5108098

Figure 5.2: As in 4.2 but for Irma (2017)

September 16 observations show two distinct eyewalls (Figure 5.4). In 6H-4DEnVar the outer band appears to be two spiraling bands emanating from the inner band. 6H-3DEnVar and 1H-3DEnVar more correctly resolve the outer eyewall as a separate feature from the inner eyewall, although neither produces a fully closed ring as in the observations. Both 6H-4DEnVar and 1H-3DEnVar resolve the band of convection to the north of the inner eyewall better than the 6H-3DEnVar as it is absent in 6H-3DEnVar. All models place the storm further east than the observations as well, with 1H-3DEnVar being the furthest east.

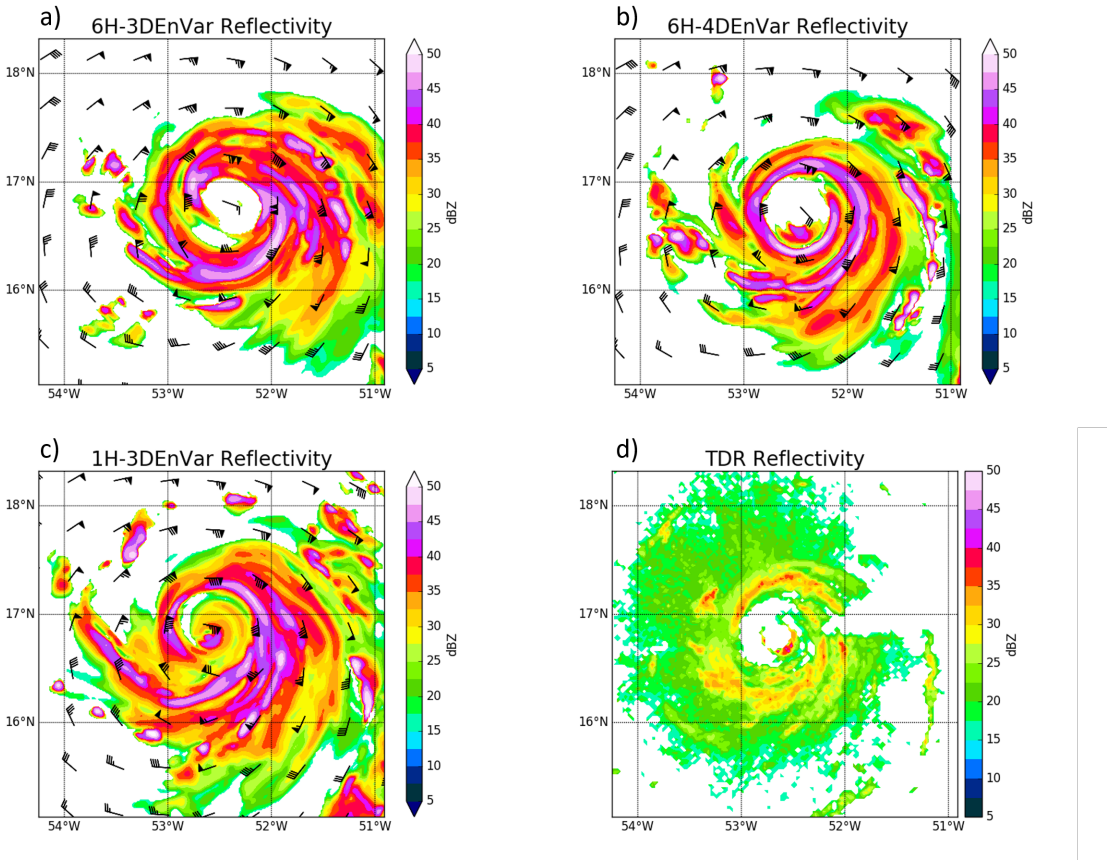


Figure 5.3: Simulated radar reflectivity valid 1200 UTC 04 September for a) 6H-3DEnVar, b) 6H-4DEnVar from forecast initialized 0600 UTC 04 September and, c) 1H-3DEnVar from forecast initialized at 0900 UTC 04 September (I19) and Observed and Observed TDR reflectivity valid at 1142 UTC 04 September.

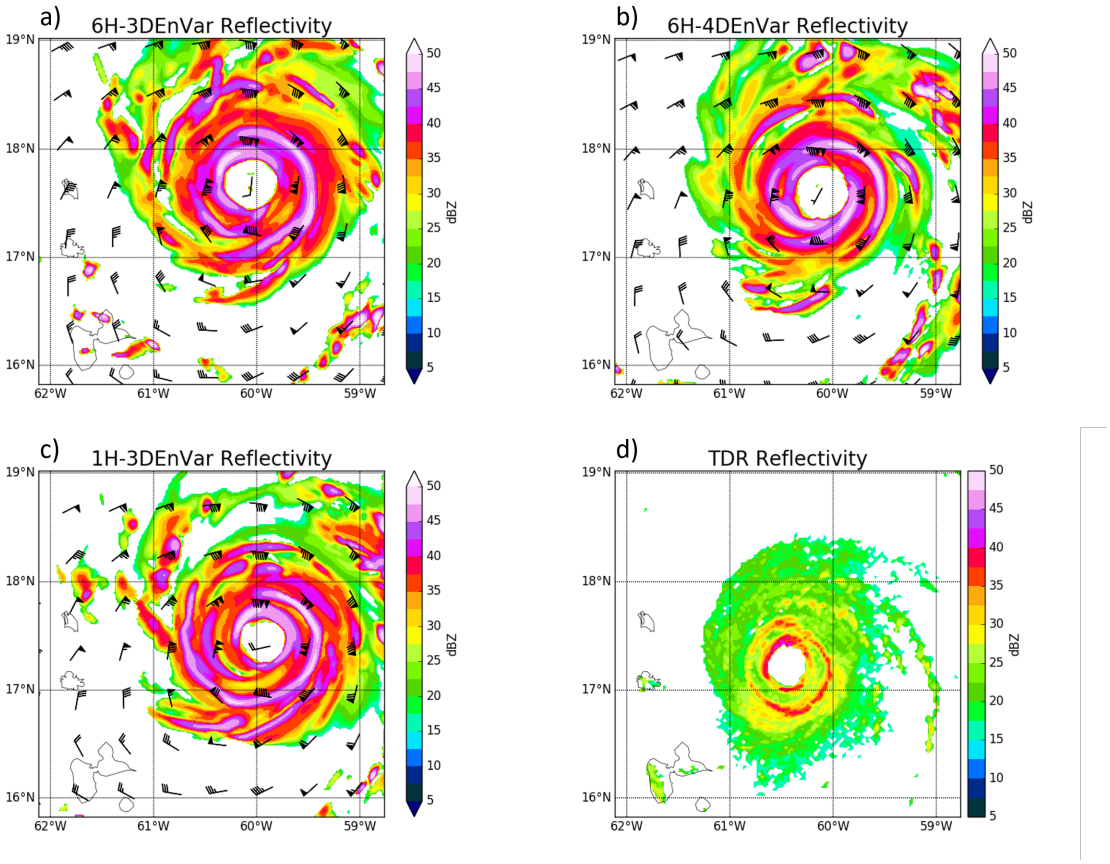


Figure 5.4: Simulated radar reflectivity valid 0000 UTC 06 September for a) 6H-3DEnVar, b) 6H-4DEnVar from forecast initialized 1200 UTC 05 September and, c) 1H-3DEnVar from forecast launched at 1500 UTC 05 September (I24) and Observed and Observed TDR reflectivity valid at 2359 UTC 05 September

5.4 Satellite location

As was done for Edouard in section 4.5.2, experiments were conducted using satellite derived locations instead of the TCVitals and EnSRF derived locations for VR. Both 1H-3DEnVar-sl and 6H-4DEnVar-sl show improvements over 1H-3DEnVar and 6H-4DEnVar for Vmax for 0-24h (Figure 5.1a). Similarly, 6H-4DEnVar-sl improves upon the 6H-3DEnVar results after 66h. Both 1H-3DEnVar-sl and 6H-4DEnVar-sl exhibit smaller RMSEs for much of the forecast than 6H-3DEnVar and provide comparable results to each other for most times after 30h. For MSLP (Figure 5.1b), 1H-3DEnVar-sl improves upon 1H-3DEnVar errors primarily after 54h, becoming comparable to 6H-4DEnVar errors for this time frame. 6H-4DEnVar-sl further improves upon errors seen in 6H-4DEnVar for all forecast hours. Both 1H-3DEnVar-sl and 6H-3DEnVar-sl show smaller errors than 6H-3DEnVar in 0-48, and are comparable or better than 6H-3DEnVar from 48h to 84h. After this time 6H-4DEnVar-sl and 1H-3DEnVar-sl have larger RMSE's than 6H-3DEnVar. However, the improvements compared to 1H-3DEnVar and 6H-4DEnVar are largest during this time.

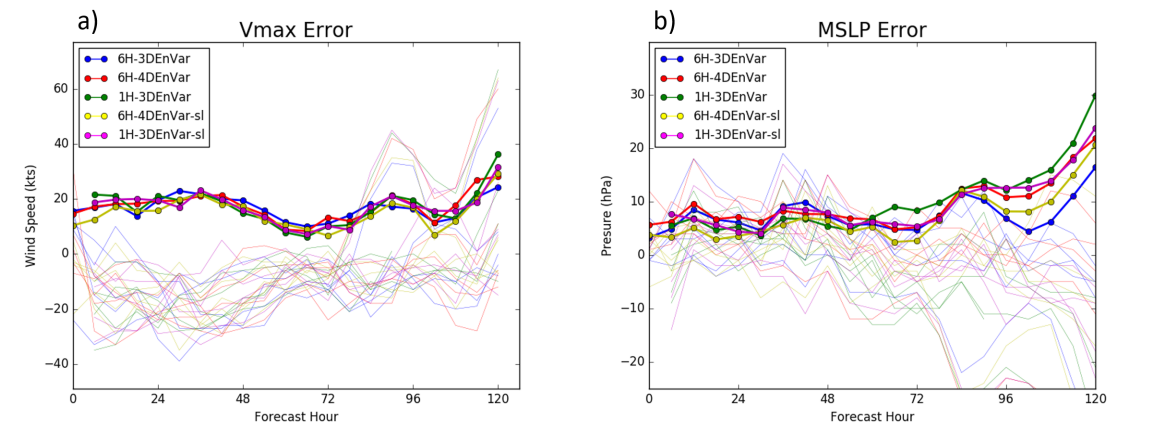


Figure 5.5: As in 4.18, but for Irma

Chapter 6

Discussion and Conclusion

A GSI-Based Ensemble-Variational Data Assimilation system for HWRF is tested to compare the four-dimensional and one-hourly three-dimensional implementations. Experiments are conducted using Edouard (2014) and Irma (2017) to evaluate whether accounting for the evolution of background error covariance in rapidly evolving hurricanes with inner core data can improve analyses and forecasts in HWRF. Specifically 1-Hourly 3DEnVar and 6-Hourly 4DEnVar are proposed as alternatives to the 6-Hourly Hybrid-3DEnVar system that uses stationary covariances over the 6-hour DA window. Furthermore, 1-Hourly Hybrid-3DEnVar and 6-Hourly Hybrid-4DEnVar are compared to determine if one system produces better analyses and forecasts than the other.

Between several experiments, 6H-3DEnVar is seen to produce poor analyses. Edouard (2014) 6H-3DEnVar shows a large RMSE in Vmax and MSLP as well as SFMR wind speed and flight level specific humidity. Structure correlation shows a similar result but with a small difference in magnitude for Edouard (2014). In Edouard, flight level temperature shows that 6H-3DEnVar is worse than 6H-4DEnVar but not 1H-3DEnVar. These results suggest that both 6H-4DEnVar and 1H-3DEnVar produce better analyses than 6H-3DEnVar, supporting the hypothesis that accounting for background evolution when the background is evolving rapidly improves the analysis.

Comparing 6H-4DEnVar and 1H-3DEnVar, Vmax and MSLP RMSE are the only metrics that show a notable difference at analysis time. 6H-4DEnVar analysis MSLP verifies better than 1H-3DEnVar, however for Vmax 1H-3DEnVar has a smaller RMSE for Edouard (2014). The 1H-3DEnVar forecast experiences spindown in nearly every cycle, so the improved analysis is not representative of the early forecast.

During the forecast, simulated reflectivity and SFMR wind speed suggest that 6H-3DEnVar typically does not perform as well as 1H-3DEnVar and 6H-4DEnVar, especially when a double eyewall is present for reflectivity. For Vmax the poor analysis does not necessarily produce a poorer forecast than 6H-4DEnVar or 1H-3DEnVar, where any degradation does not last beyond hour 12. Instead, the spindown in 1H-3DEnVar results in 1H-3DEnVar producing the worst Vmax forecast for the first 12-18 hours. For Edouard (2014) MSLP RMSE shows that early 6H-3DEnVar forecasts take longer to adjust than in Vmax, but from hours 24-48 the 6H-3DEnVar MSLP forecast produces the smallest RMSE. This is in part due to a combination of a bias toward low MSLP in the analysis correcting to a better forecast before strengthening again to degrade the forecast. Overall, most differences between 6H-3DEnVar and the other experiments is within the first 24-48 hours, with most suggesting that 6H-3DEnVar is the worst in this time frame, aside from MSLP during the second half of this window.

1H-3DEnVar exhibits a poor relationship between MSLP and Vmax during the Analysis through the first 30 hours of the forecast in Edouard (2014). 6H-3DEnVar and 6H-4DEnVar show a similar slope to best track, with both experiments having a bias along the slope. Additionally, 1H-3DEnVar has a larger bias than 6H-4DEnVar and 6H-3DEnVar. These issues appear to be related to the spindown issue seen in the 1H-3DEnVar forecast where Vmax weakens, and there is no notable corresponding increase in MSLP. This spindown was shown to be a result of poorly predicted storm locations in the relocated background. A strong wind gradient exists in the transition from eye to eyewall, and a slight dislocation between the observations and the background can result in large innovations and a violation of gaussian assumptions. This problem is further exacerbated by an underdispersive 1H-3DEnVar ensemble with regard to storm center locations. Since the vortex modification procedure uses an EnSRF to determine the location of the storm in the relocated background, storms location may not be relocated far enough toward the TCVitals observation.

Regular TCVitals observations are available every 6 hours. However, 6H-4DEnVar and 1H-3DEnVar perform vortex relocation every hour. A linear interpolation is used to provide an observation in between regular TCVitals observations. Given the results suggesting that the background location impacts results, in some cases using a more advanced non-linear interpolation method may improve results. In cases where the storm is propagating non-linearly over the 6 hour window additional error is introduced in the linearly interpolated TCVitals. Barring an increase in the temporal frequency of TCVitals observations, the quality of the TCVitals for times other than 0, 6, 12, and 18 UTC will be dependent on the quality of the interpolation.

Experiments with Irma (2017) show that relative Vmax RMSEs between experiments are similar to Edouard (2014). The primary differences are 6H-3DEnVar not showing as large of error in the analysis, and 1H-3DEnVar having increased RMSE starting at hour 3, instead of 9. For MSLP RMSE 6H-3DEnVar has the largest error instead of the smallest as is seen in the Edouard results, suggesting that the small error in Edouard is not representative. MSLP-Vmax relationship is similar to that seen in Edouard, however 6H-3DEnVar and 6H-4DEnVar have poor slopes similar to 1H-3DEnVar. Radar forecast verification also results in similar conclusions to Edouard, with 6H-4DEnVar and 1H-3DEnVar resolving inner core structure better than 6H-3DEnVar. Structure correlation for Irma suggests that 1H-3DEnVar produces the best results. However, this is due to a single case, and the correlations are similar for the remaining cases. Comparing results for individual missions reveals that for most metrics 4 of the 5 missions produce 6H-3DEnVar analyses that are worse than both 6H-4DEnVar and 1H-3DEnVar. While most results for Irma (2017) are similar to those for Edouard (2014), there are some differences. To fully understand the differences more cases are needed, as some results may be due to the small sample sizes used in this study.

As mentioned in chapter 1 decreasing the length of the DA cycle increases computational costs, as does using 4D DA. As such, 1H-3DEnVar is the most computationally

expensive, with 6H-4DEnVar being the second most expensive, and 6H-3DEnVar the least expensive. The lesser computational cost is the only consistent advantage seen in 6H-3DEnVar compared to the other experiments. The 6H-4DEnVar and 1H-3DEnVar experiments perform similarly in most metrics, but there is a greater computational cost associated with 1H-3DEnVar due to the large number of cycles. Additional problems were presented in 1H-3DEnVar related to choosing the exact setup of the experiments. Because of the 1-hour DA cycles in 1H-3DEnVar, several difficulties were presented regarding the timing of the cycles and the free forecast. If the free forecast is launched at the same time as with the 6-hour DA, any data valid in the second half of the 6-hour DA window will not be assimilated, thus providing a 3 hour mismatch in the data assimilated between free forecasts. If the 6-hour DA window is divided into six 1-hour windows then the analyses will be valid at HH:30, leaving a 30 minute difference in the launch of the free forecast between experiments, as well as not being consistent with the operational HH:00 analyses. In order to remain consistent with the HH:00 convention, the window that observations fall into between free forecasts in 1H-3DEnVar would be shifted by 30 minutes compared to the 6-hour DA window in 6H-3DEnVar and 6H-4DEnVar. In order to alleviate this problem, the DA cycles that span two DA windows are split into two smaller 30-minute cycles, but no forecast or VR/VM is performed between them. This allows the same data to be assimilated for each mission. However, it does present the negative affect of having the free forecast launch at the end of the mission, thus three hours later than the 6H-3DEnVar and 6H-4DEnVar free forecasts.

In summary, both 6H-4DEnVar and 1H-3DEnVar perform better than 6H-3DEnVar by most metrics in both the analysis and early forecast but become similar after about 48 hours. Most differences between 6H-4DEnVar and 1H-3DEnVar are small except for 1) Flight level temperature, 2) spindown, 3) computational costs. However, it is uncertain if the first two differences could be eliminated with further development. Given the improvements seen using satellite derived locations, hourly TCVitals observations could provide further

improvement. Additionally, inflating the ensemble of storm location perturbations prior to the EnSRF step in VR may reduce the affects of the underdispersive ensemble in 1H-3DEnVar. Further experimentation with more casesand development is necessary to answer these questions.

Bibliography

Aksoy, A., S. D. Aberson, T. Vukicevic, K. J. Sellwood, S. Lorsolo, and X. Zhang, 2013: Assimilation of High-Resolution Tropical Cyclone Observations with an Ensemble Kalman Filter Using NOAA/AOML/HRD's HEDAS: Evaluation of the 2008-11 Vortex-Scale Analyses. *Monthly Weather Review*, **141 (6)**, 1842–1865, doi:10.1175/mwr-d-12-00194.1.

Aksoy, A., S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Aberson, and F. Zhang, 2012: The HWRF Hurricane Ensemble Data Assimilation System (HEDAS) for High-Resolution Data: The Impact of Airborne Doppler Radar Observations in an OSSE. *Monthly Weather Review*, **140 (6)**, 1843–1862, doi:10.1175/MWR-D-11-00212.1.

Bender, M. A., R. J. Ross, R. E. Tuleya, and Y. Kurihara, 1993: Improvements in tropical cyclone track and intensity forecasts using the GFDL initialization system. *Monthly Weather Review*, **121**, 2046–2061.

Bishop, C. H., and D. Hodyss, 2011: Adaptive Ensemble Covariance Localization in Ensemble 4D-VAR State Estimation. *Monthly Weather Review*, **139 (4)**, 1241–1255, doi:10.1175/2010mwr3403.1.

Biswas, M., and Coauthors, 2017: Hurricane Weather Research and Forecasting (HWRF) Model: 2017 Scientific Documentation. Tech. rep., Developmental Testbed Center, 105 pp. URL https://dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRFv3.9a_ScientificDoc.pdf.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quarterly Journal of the Royal Meteorological Society*, **131 (607)**, 1013–1043, doi:10.1256/qj.04.15.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments. *Monthly Weather Review*, **138 (5)**, 1550–1566, doi:10.1175/2009mwr3157.1.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010b: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part II: One-Month Experiments with Real Observations. *Monthly Weather Review*, **138 (5)**, 1567–1586, doi:10.1175/2009mwr3158.1.

Chen, Y., and C. Snyder, 2007: Assimilating Vortex Position with an Ensemble Kalman Filter. *Monthly Weather Review*, **135 (5)**, 1828–1845, doi:10.1175/mwr3351.1.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, **139 (675)**, 1445–1461, doi:10.1002/qj.2054.

Dong, J., and M. Xue, 2013: Assimilation of radial velocity and reflectivity data from coastal WSR-88D radars using an ensemble Kalman filter for the analysis and forecast of landfalling hurricane Ike (2008). *Quarterly Journal of the Royal Meteorological Society*, **139 (671)**, 467–487, doi:10.1002/qj.1970.

Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *Journal of Atmospheric and Oceanic Technology*, **26 (5)**, 911–927, doi:10.1175/2008JTECHA1156.1.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of Hybrid Ensemble/3DVAR Analysis Schemes to Model Error and Ensemble Covariance Error. *Monthly Weather Review*, **132 (5)**, 1065–1080, doi:10.1175/1520-0493(2004)132⟨1065:rohdas⟩2.0.co;2.

Evensen, G., 2018: Analysis of iterative ensemble smoothers for solving inverse problems. *Computational Geosciences*, **22 (3)**, 885–908, doi:10.1007/s10596-018-9731-y.

Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D Ensemble Kalman Filter: Perfect model simulations with Lorenz-96. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **59 (1)**, 96–100, doi:10.1111/j.1600-0870.2006.00205.x.

Gamache, J. F., J. Franklin, N. Surgi, and Q. Liu, 2015: Real-Time Dissemination of Hurricane Wind Fields Determined from Airborne Doppler Radar Data Real-Time Dissemination of Hurricane Wind Fields Determined. Tech. rep., NOAA, 38 pp. URL http://www.nhc.noaa.gov/jht/2003-2005reports/DOPLRgamache_JHTfinalreport.pdf.

Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Monthly Weather Review*, **135 (6)**, 2339–2354, doi:10.1175/mwr3394.1.

Hamill, T. M., and C. Snyder, 2000: A Hybrid Ensemble Kalman Filter-3D Variational Analysis Scheme. *Monthly Weather Review*, **128 (8)**, 2905–2919, doi:10.1175/1520-0493(2000)128⟨2905:ahekfv⟩2.0.co;2.

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. *Monthly Weather Review*, **139 (2)**, 668–688, doi:10.1175/2010mwr3456.1.

Houtekamer, P. L., and F. Zhang, 2016: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, **144 (12)**, 4489–4532, doi:10.1175/MWR-D-15-0440.1, URL http://journals.ametsoc.org/doi/10.1175/MWR-D-15-0440.1.

Hu, M., and M. Xue, 2007: Impact of Configurations of Rapid Intermittent Assimilation of WSR-88D Radar Data for the 8 May 2003 Oklahoma City Tornadic Thunderstorm Case. *Monthly Weather Review*, **135 (2)**, 507–525, doi:10.1175/mwr3313.1.

Huang, X.-Y., and P. Lynch, 1993: Diabatic Digital-Filtering Initialization: Application to the HIRLAM Model. *Monthly Weather Review*, **121 (2)**, 589–603, doi:10.1175/1520-0493(1993)121⟨0589:ddfiat⟩2.0.co;2.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, **230 (1-2)**, 112–126, doi:10.1016/j.physd.2006.11.008, 0511236v2.

Hunt, B. R., and Coauthors, 2004: Four-dimensional ensemble Kalman filtering. *Tellus, Series A: Dynamic Meteorology and Oceanography*, **56 (4)**, 273–277, doi:10.1111/j.1600-0870.2004.00066.x.

Kleist, D. T., and K. Ide, 2015a: An OSSE-Based Evaluation of Hybrid Variational-Ensemble Data Assimilation for the NCEP GFS. Part I: System Description and 3D-Hybrid Results. *Monthly Weather Review*, **143 (2)**, 433–451, doi:10.1175/mwr-d-13-00351.1.

Kleist, D. T., and K. Ide, 2015b: An OSSE-Based Evaluation of Hybrid Variational Ensemble Data Assimilation for the NCEP GFS. Part II: 4DEnVar and Hybrid Variants. *Monthly Weather Review*, **143 (2)**, 452–470, doi:10.1175/mwr-d-13-00350.1.

Kurihara, Y., M. A. Bender, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Monthly Weather Review*, **121**, 2030–2045.

Kurihara, Y., M. A. Bender, R. E. Tuleya, and R. J. Ross, 1990: Prediction Experiments of Hurricane Gloria (1985) Using a Multiply Nested Movable Mesh Model. *Monthly Weather Review*, **118 (10)**, 2185–2198, doi:10.1175/1520-0493(1990)118⟨2185:peohgu⟩2.0.co;2.

Kurihara, Y., M. A. Bender, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Monthly Weather Review*, **123**, 2791–2891.

Kurihara, Y., R. E. Tuleya, and M. A. Bender, 1998: The GFDL Hurricane Prediction System and Its Performance in the 1995 Hurricane Season. *Monthly Weather Review*, **126 (5)**, 1306–1322, doi:10.1175/1520-0493(1998)126⟨1306:tghpsa⟩2.0.co;2.

Kutty, G., and X. Wang, 2015: A Comparison of the Impacts of Radiosonde and AMSU Radiance Observations in GSI Based 3DEnsVar and 3DVar Data Assimilation Systems for NCEP GFS. *Advances in Meteorology*, **2015**, 1–17, doi:10.1155/2015/280546.

Lei, L., and J. L. Anderson, 2014: Impacts of Frequent Assimilation of Surface Pressure Observations on Atmospheric Analyses. *Monthly Weather Review*, **142 (12)**, 4477–4483, doi:10.1175/mwr-d-14-00097.1.

Leighton, H., S. Gopalakrishnan, J. A. Zhang, R. F. Rogers, Z. Zhang, and V. Tallapragada, 2018: Azimuthal Distribution of Deep Convection, Environmental Factors, and Tropical Cyclone Rapid Intensification: A Perspective from HWRF Ensemble Forecasts

of Hurricane Edouard (2014). *Journal of the Atmospheric Sciences*, **75 (1)**, 275–295, doi:10.1175/jas-d-17-0171.1.

Li, J., and H. Liu, 2009: Improved hurricane track and intensity forecast using single field-of-view advanced IR sounding measurements. *Geophysical Research Letters*, **36 (11)**, 9–12, doi:10.1029/2009GL038285.

Li, X., J. Ming, M. Xue, Y. Wang, and K. Zhao, 2015: Implementation of a dynamic equations constraint based on the stead state momentum equations within the WRF hybrid ensemble-3DVar data assimilation system and test with radar T-TREC wind assimilation for tropical Cyclone Chanthu (2010). *Journal of Geophysical Research : Atmospheres*, **120 (1)**, 4017–4039, doi:10.1002/2014JD022706.

Li, Y., X. Wang, and M. Xue, 2012: Assimilation of Radar Radial Velocity Data with the WRF Hybrid Ensemble-3DVAR System for the Prediction of Hurricane Ike (2008). *Monthly Weather Review*, **140 (11)**, 3507–3524, doi:10.1175/mwr-d-12-00043.1.

Liu, Q., S. Lord, N. Surgi, Y. Zhu, R. Wobus, Z. Toth, and T. Marchok, 2006: Hurricane Relocation in Global Ensemble Forecast System. *27th Conference on Hurricanes and Tropical Meteorology*, Monterey, CA, P5.13, URL https://ams.confex.com/ams/pdfpapers/108503.pdf.

Liu, Q., T. Marchok, H.-l. Pan, M. Bender, and S. Lord, 2000: Improvements in Hurricane Initialization and Forecasting at NCEP with Global and Regional (GFDL) models. *NCEP Office Note*, 472.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP - A comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, **129 (595 PART B)**, 3183–3203, doi:10.1256/qj.02.132.

Lu, X., and X. Wang, 2019: Improving Hurricane Analyses and Predictions with TCI, IFEX Field Campaign Observations, and CIMSS AMVs Using the Advanced Hybrid Data Assimilation system for HWRF. Part I: What is missing to capture the rapid intensification of Hurricane Patricia (2015) when HWRF is already initialized with a more realistic analysis? *Monthly Weather Review*, **(2015)**, MWR–D–18–0202.1, doi:10.1175/MWR-D-18-0202.1, URL http://journals.ametsoc.org/doi/10.1175/MWR-D-18-0202.1.

Lu, X., X. Wang, Y. Li, M. Tong, and X. Ma, 2017a: GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction: impact of various error covariances for airborne radar observation assimilation. *Quarterly Journal of the Royal Meteorological Society*, **143 (702)**, 223–239, doi:10.1002/qj.2914.

Lu, X., X. Wang, M. Tong, and V. Tallapragada, 2017b: GSI-Based, Continuously Cycled, Dual-Resolution Hybrid Ensemble-Variational Data Assimilation System for HWRF: System Description and Experiments with Edouard (2014). *Monthly Weather Review*, **145 (12)**, 4877–4898, doi:10.1175/mwr-d-17-0068.1.

Poterjoy, J., and F. Zhang, 2014: Intercomparison and Coupling of Ensemble and Four-Dimensional Variational Data Assimilation Methods for the Analysis and Forecasting of Hurricane Karl (2010). *Monthly Weather Review*, **142 (9)**, 3347–3364, doi:10.1175/mwr-d-13-00394.1.

Poterjoy, J., F. Zhang, and Y. Weng, 2014: The Effects of Sampling Errors on the EnKF Assimilation of Inner-Core Hurricane Observations. *Monthly Weather Review*, **142 (4)**, 1609–1630, doi:10.1175/mwr-d-13-00305.1.

Pu, Z., X. Li, and E. J. Zipser, 2009: Diagnosis of the Initial and Forecast Errors in the Numerical Simulation of the Rapid Intensification of Hurricane Emily (2005). *Weather and Forecasting*, **24 (5)**, 1236–1251, doi:10.1175/2009waf2222195.1.

Pu, Z.-X., and S. A. Braun, 2001: Evaluation of Bogus Vortex Techniques with Four-Dimensional Variational Data Assimilation. *Monthly Weather Review*, **129 (8)**, 2023–2039, doi:10.1175/1520-0493(2001)129⟨2023:eobvtw⟩2.0.co;2.

Rogers, R., and Coauthors, 2013: NOAA'S Hurricane Intensity Forecasting Experiment: A Progress Report. *Bulletin of the American Meteorological Society*, **94 (6)**, 859–882, doi:10.1175/bams-d-12-00089.1.

Schwartz, C. S., Z. Liu, and X.-Y. Huang, 2015: Sensitivity of Limited-Area Hybrid Variational-Ensemble Analyses and Forecasts to Ensemble Perturbation Resolution. *Monthly Weather Review*, **143 (9)**, 3454–3477, doi:10.1175/mwr-d-14-00259.1.

Schwartz, C. S., Z. Liu, X.-Y. Huang, Y.-H. Kuo, and C.-T. Fong, 2013: Comparing Limited-Area 3DVAR and Hybrid Variational-Ensemble Data Assimilation Methods for Typhoon Track Forecasts: Sensitivity to Outer Loops and Vortex Relocation. *Monthly Weather Review*, **141 (12)**, 4350–4372, doi:10.1175/mwr-d-13-00028.1.

Tallapragada, V., S. G. Gopalakrishnan, Q. Liu, and T. P. Marchok, 2014: Hurricane Weather Research and Forecasting (HWRF) model: 2014 scientific documentation. Tech. Rep. September, Developmental Testbed Center, 1–105 pp. URL http://www.dtcenter.org/HurrWRF/users/docs/scientific_documents/HWRFScientificDocumentation_August2011.pdf.

Thu, T. V., and T. N. Krishnamurti, 1992: Vortex initialization for typhoon track prediction. *Meteorology and Atmospheric Physics*, **47 (2-4)**, 117–126, doi:10.1007/BF01025612.

Tong, M., V. Tallapragada, E. Liu, W. Wang, C. Kieu, Q. Liu, and B. Zhan, 2014: Impact of Assimilating Aircraft Reconnaissance Observations in Operational HWRF. *2014 HFIP annual meeting*, URL http://www.hfip.org/events/annual_meeting_nov_2014/wed/15_Tong_2014_HFIP_annual_meeting.pdf.

Tong, W., G. Li, J. Sun, X. Tang, and Y. Zhang, 2016: Design Strategies of an Hourly Update 3DVAR Data Assimilation System for Improved Convective Forecasting. *Weather and Forecasting*, **31 (5)**, 1673–1695, doi:10.1175/waf-d-16-0041.1.

Torn, R. D., and G. J. Hakim, 2009: Ensemble Data Assimilation Applied to RAINEX Observations of Hurricane Katrina (2005). *Monthly Weather Review*, **137 (9)**, 2817–2829, doi:10.1175/2009mwr2656.1.

Trahan, S., and L. Sparling, 2012: An Analysis of NCEP Tropical Cyclone Vitals and Potential Effects on Forecasting Models. *Weather and Forecasting*, **27 (3)**, 744–756, doi:10.1175/waf-d-11-00063.1.

Wang, S., M. Xue, and J. Min, 2013a: A four-dimensional asynchronous ensemble square-root filter (4DEnSRF) algorithm and tests with simulated radar data. *Quarterly Journal of the Royal Meteorological Society*, **139 (672)**, 805–819, doi:10.1002/qj.1987, URL http://doi.wiley.com/10.1002/qj.1987.

Wang, S., M. Xue, A. D. Schenkman, and J. Min, 2013b: An iterative ensemble square root filter and tests with simulated radar data for storm-scale data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **139 (676)**, 1888–1903, doi:10.1002/qj.2077.

Wang, X., 2010: Incorporating Ensemble Covariance in the Gridpoint Statistical Interpolation Variational Minimization: A Mathematical Framework. *Monthly Weather Review*, **138 (7)**, 2990–2995, doi:10.1175/2010mwr3245.1.

Wang, X., 2011: Application of the WRF Hybrid ETKF-3DVAR Data Assimilation System for Hurricane Track Forecasts. *Weather and Forecasting*, **26 (6)**, 868–884, doi:10.1175/waf-d-10-05058.1.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A Hybrid ETKF-3DVAR Data Assimilation Scheme for the WRF Model. Part I: Observing System Simulation Experiment. *Monthly Weather Review*, **136 (12)**, 5116–5131, doi:10.1175/2008MWR2444.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2444.1.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008b: A Hybrid ETKF-3DVAR Data Assimilation Scheme for the WRF Model. Part II: Real Observation Experiments. *Monthly Weather Review*, **136 (12)**, 5132–5147, doi:10.1175/2008MWR2445.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2445.1.

Wang, X., T. M. Hamill, J. S. Whitaker, and C. H. Bishop, 2007a: A Comparison of Hybrid Ensemble Transform Kalman Filter-Optimum Interpolation and Ensemble Square Root Filter Analysis Schemes. *Monthly Weather Review*, **135 (3)**, 1055–1076, doi:10.1175/mwr3307.1.

Wang, X., T. M. Hamill, J. S. Whitaker, and C. H. Bishop, 2009: A Comparison of the Hybrid and EnSRF Analysis Schemes in the Presence of Model Errors due to Unresolved Scales. *Monthly Weather Review*, **137 (10)**, 3219–3232, doi:10.1175/2009mwr2923.1.

Wang, X., and T. Lei, 2014: GSI-Based Four-Dimensional Ensemble-Variational (4DEnsVar) Data Assimilation: Formulation and Single-Resolution Experiments with

Real Data for NCEP Global Forecast System. *Monthly Weather Review*, **142 (9)**, 3303–3325, doi:10.1175/MWR-D-13-00303.1, URL http://journals.ametsoc.org/doi/10.1175/MWR-D-13-00303.1.

Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013c: GSI 3DVar-Based Ensemble-Variational Hybrid Data Assimilation for NCEP Global Forecast System: Single-Resolution Experiments. *Monthly Weather Review*, **141 (11)**, 4098–4117, doi:10.1175/mwr-d-12-00141.1.

Wang, X., C. Snyder, and T. M. Hamill, 2007b: On the Theoretical Equivalence of Differently Proposed Ensemble-3DVAR Hybrid Analysis Schemes. *Monthly Weather Review*, **135 (1)**, 222–227, doi:10.1175/MWR3282.1, URL http://journals.ametsoc.org/doi/abs/10.1175/MWR3282.1.

Weng, Y., and F. Zhang, 2012: Assimilating Airborne Doppler Radar Observations with an Ensemble Kalman Filter for Convection-Permitting Hurricane Initialization and Prediction: Katrina (2005). *Monthly Weather Review*, **140 (3)**, 841–859, doi:10.1175/2011mwr3602.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, **130 (7)**, 1913–1924, doi:10.1175/1520-0493(2002)130⟨1913:edawpo⟩2.0.co;2.

Yang, S.-C., K.-J. Lin, T. Miyoshi, and E. Kalnay, 2013: Improving the spin-up of regional EnKF for typhoon assimilation and forecasting with Typhoon Sinlaku (2008). *Tellus A: Dynamic Meteorology and Oceanography*, **65 (1)**, 20 804, doi:10.3402/tellusa.v65i0.20804.

Zhang, F., Y. Weng, J. F. Gamache, and F. D. Marks, 2011: Performance of convection-permitting hurricane initialization and prediction during 2008-2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophysical Research Letters*, **38 (15)**, 2–7, doi:10.1029/2011GL048469.

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-Resolving Hurricane Initialization and Prediction through Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter. *Monthly Weather Review*, **137 (7)**, 2105–2125, doi:10.1175/2009mwr2645.1.

Zhang, M., and F. Zhang, 2012: E4DVar: Coupling an Ensemble Kalman Filter with Four-Dimensional Variational Data Assimilation in a Limited-Area Weather Prediction Model. *Monthly Weather Review*, **140 (2)**, 587–600, doi:10.1175/mwr-d-11-00023.1.

Zhang, S., and Z. Pu, 2019: Numerical Simulation of Rapid Weakening of Hurricane Joaquin with Assimilation of High-Definition Sounding System Dropsondes during the Tropical Cyclone Intensity Experiment: Comparison of Three- and Four-Dimensional Ensemble-Variational Data Assimilation. *Weather and Forecasting*, **34 (3)**, 521–538, doi:10.1175/waf-d-18-0151.1.

Zhou, C., H. Shao, and B. Ligia, 2015: Applications of the GSI-Hybrid Data Assimilation for High-Resolution Tropical Storm Forecasts: tackling the intensity spindown issue in 2014 HWRF. *16th WRF Users Workshop*, Boulder, CO, URL https://dtcenter.org/eval/data_assim/publications/GSI-Hybridat2015WRFUsersWorkshop.v2_poster.pdf.

Zou, X., and Q. Xiao, 2000: Studies on the Initialization and Simulation of a Mature Hurricane Using a Variational Bogus Data Assimilation Scheme. *Journal of the Atmospheric Sciences*, **57 (6)**, 836–860, doi:10.1175/1520-0469(2000)057⟨0836:sotias⟩2.0.co;2.