

ABSTRACT

Title of Document: AN EVALUATION OF HYBRID
VARIATIONAL-ENSEMBLE DATA
ASSIMILATION FOR THE NCEP GFS

Daryl Timothy Kleist, Ph.D., 2012

Directed By: Professor Kayo Ide, Department of Atmospheric
and Oceanic Science

Several variants of hybrid data assimilation algorithms have been developed and tested within recent years, particularly for numerical weather prediction (NWP). The hybrid algorithms are designed to combine the strengths of variational and ensemble-based techniques while at the same time attempting to mitigate their weaknesses. One such variational-based algorithm is under development for use with the National Centers for Environmental Prediction's (NCEP) global forecast system (GFS) model. In this work, we attempt to better understand the impact of utilizing a hybrid scheme on the quality of analyses and subsequent forecasts, as well as explore alternative extensions to make better use of the ensemble information within the variational solver.

A series of Observing System Simulation Experiments (OSSEs) are carried out. It is demonstrated that analysis and subsequent forecast errors are generally reduced in a 3D-hybrid scheme relative to 3DVAR. Several variational-based 4D

extensions are proposed and tested, including the use of a variety of dynamic constraints. A simple approach for hybridizing the 4D-ensemble with a time-invariant contribution is proposed and tested. The 4D variants are shown to be superior to the 3D-hybrid, with positive contributions from static **B** as well as the dynamic constraint formulations. It is clear from both the 3D and 4D experiments that more sophisticated methods for dealing with inflation and localization in the ensemble update are needed even within the hybrid paradigm. Lastly, a method for applying piecewise scale-dependent weights is proposed and successfully tested.

The 3D OSSE-based results are also compared with results from an experiment using real observations to corroborate the findings. It is found that in general, most of the results are comparable, though the positive impact in the real system is more consistent and impressive.

AN EVALUATION OF HYBRID VARIATIONAL-ENSEMBLE DATA
ASSIMILATION FOR THE NCEP GFS

By

Daryl Timothy Kleist

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Kayo Ide, Chair
Professor Eugenia Kalnay
Professor Takemasa Miyoshi
Professor Brian Hunt, Dean's Representative
Dr. John Derber

© Copyright by
Daryl Timothy Kleist
2012

Acknowledgements

I first thank my primary advisor, Prof. Kayo Ide, for her incredible patience, flexibility, and ability to provide productive feedback during this work. Without her guidance and support, I never would have been successful during my time at the University of Maryland. I am also forever grateful to my supervisors at the Environmental Modeling Center, particularly John Derber, Stephen Lord, and Bill Lapenta for their support and understanding while I pursued my studies. I also thank Dr. John Derber, Prof. Eugenia Kalnay, Prof. Takemasa Miyoshi, and Prof. Brian Hunt for taking the time to participate on my committee and provide valuable feedback and suggestions. I am especially thankful to Jeff Whitaker and Dave Parrish for their very fruitful collaboration.

The nature run data was kindly provided by the European Center for Medium Range Weather Forecasting. I am especially grateful to Erik Andersson and Michiko Masutani for their assistance. Thanks also to Ron Errico and Nikke Privé for kindly providing the simulated observations and providing guidance and assistance.

Last, but not least, I thank all of my family and many friends for providing constant support and encouragement.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iii
List of Acronyms and Abbreviations.....	v
List of Tables.....	vi
List of Figures.....	vii
Chapter 1: Introduction.....	1
1.1 Background.....	1
1.2 Thesis Objectives and Outline.....	7
Chapter 2: System Description and 3D-Hybrid OSSE results.....	10
2.1 Introduction.....	10
2.2 NCEP Hybrid Data Assimilation.....	13
2.2.1 System Description.....	13
2.2.2 Single Observation Example.....	16
2.3 Joint OSSE.....	17
2.4 Experiment Design.....	19
2.4.1 3DVAR.....	22
2.4.2 3D-Hybrid.....	23
2.5 3DVAR Results.....	28
2.6 3D-Hybrid Results.....	32
2.6.1 Analysis Comparison.....	32
2.6.2 Forecast Impact.....	39
2.6.3 Follow on 3D-Hybrid Experiments.....	42
2.7 Summary and Conclusions.....	51
Chapter 3: OSSE Experiments with 4D-Ensemble-Var and Hybrid Variants.....	55
3.1 Introduction.....	55
3.2 4D-Ensemble-Var.....	58
3.2.1 GSI-Based Hybrid Extension.....	58
3.2.2 4D-Ensemble-Var Hybrid.....	60
3.2.3 Single Observation Example.....	61
3.3 Constraint Formulations.....	66
3.3.1 Incremental Normal Mode Constraint.....	66
3.3.2 Weak Constraint Digital Filter.....	69
3.3.3 Impact on Single Analysis.....	70
3.4 Experiment Results.....	74
3.4.1 Impact of 4D Ensemble.....	74
3.4.2 Constraint Experiments.....	80
3.4.3 Forecast Impact.....	93
3.5 Summary and Conclusions.....	97
Chapter 4: Scale-Dependent Weighting.....	102
4.1 Introduction.....	102
4.2 Scale-dependent weighting in GSI-based hybrid.....	104

4.3 Experimental Results	109
4.3.1 Single Analysis Test Case.....	109
4.3.2 Cycling OSSE Experiment	112
4.4 Summary and Discussion.....	115
Chapter 5: Real Observation 3D-Hybrid Experiment.....	117
5.1 Introduction.....	117
5.2 Experimental configuration	118
5.3 Forecast Impact.....	121
5.4 Conclusions.....	130
Chapter 6: Summary and Conclusions.....	133
6.1 OSSE-based Hybrid Experiments.....	134
6.2 Prospects for Real NWP System.....	138
6.3 Future Work	139
Bibliography	143

List of Acronyms and Abbreviations

3DVAR	3-Dimensional Variational
4DVAR	4-Dimensional Variational
4DENSV	4-Dimensional Ensemble Variational
AC	Anomaly Correlation
AIRS	Atmospheric Infrared Sounder
AMSU	Advances Microwave Sounder Unit
AMV	Atmospheric Motion Vector
COMB	Combined normal mode and digital filter initialization
CRTM	Community Radiative Transfer Model
DFI	Digital Filter Initialization
ECMWF	European Centre for Medium-range Weather Forecasts
EMC	Environmental Modeling Center
EnKF	Ensemble Kalman Filter
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FGAT	First Guess at Appropriate Time
GDAS	Global Data Assimilation System
GFS	Global Forecast System
GSI	Gridpoint Statistical Interpolation
GOES	Geostationary Operational Environmental Satellite
H-4DENSV	Hybrid 4-Dimensional Ensemble Variational
H-4DVAR_AD	Hybrid 4-Dimensional Variational (with Adjoint)
HIRS	High Resolution Infrared Radiation Sounder
JCSDA	Joint Center for Satellite Data Assimilation
JMA	Japan Meteorological Agency
LETKF	Local Ensemble Transform Kalman Filter
MSU	Microwave Sounder Unit
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanographic and Atmospheric Administration
NMI	Normal Mode Initialization
NWP	Numerical Weather Prediction
OSE	Observing System Experiment
OSSE	Observing System Simulation Experiment
PIBAL	Pilot Balloon
RMS	Root Mean Square
RMSE	Root Mean Square Error
SSMI	Special Sensor Microwave Imager
TLNMC	Tangent Linear Normal Mode Constraint
VAD	Velocity Azimuth Display

List of Tables

Table 2.1: Description of various 3D (var and hybrid) OSSE-based experiments, as well as relevant figures and equations.	29
Table 3.1: Description of various 4D (and hybrid) single observation experiments and relevant equations for Figs. 3.1 and 3.2.	62
Table 3.2: The root mean square sum of the incremental (spectral) tendencies (total and gravity mode) as well as the ratio (gravity mode / total) for the eight vertical modes kept as part of the TLNMC for the single analysis valid at 06 UTC 15 July 2010 using various 4DENSX configurations.	74
Table 3.3: Description of various (hybrid) 4D-ensemble-var OSSE-based experiments and relevant figures showing comparison of analysis errors relative to the control run for each.	75

List of Figures

Figure 2.1: Model level 15 (~850 hPa) ensemble spread (K, upper left) as well as the 850 hPa temperature background (contours, other three panels) and analysis increment (shaded, 1K interval) resulting from the from the assimilation of a single 850 hPa temperature observation (location denoted with dot) given a 1K deviation from the model guess and 1K observation error for a 3DVAR (upper right), 3DENSV (lower left) and 3DHYB (50% weighting, lower right) valid at 00 UTC 12 September 2008. 17

Figure 2.2: Spatial distribution of simulated satellite radiance observations available for assimilation valid at 00 UTC 24 July 2005 from the Joint OSSE for AMSUA and MSU (upper left), AIRS and HIRS (upper right), AMSUB (lower left) and GOES Sounder (lower right) from AQUA (dark green), NOAA-14 (brown), NOAA-15 (red), NOAA-16 (blue), NOAA-17 (purple), GOES-12 (orange), and GOES-10 (light green). 20

Figure 2.3: As in Fig. 2.2, but for ‘conventional observation’ including surface (upper left, land meter [green], ships & buoys [red], and moored buoy [blue], rawinsonde (upper right), AMVs (middle left, from GOES [red, blue], EUMETSAT [green, brown], and JMA [orange, aqua]), aircraft (middle right), VAD winds (blue), lidar wind profilers (red) and pibal (green) [lower left], and SSMI derived surface wind speeds (lower right). 21

Figure 2.4: Schematic showing how the hybrid variational-ensemble paradigm operates for a given cycle using a three member ensemble. The yellow denotes the ensemble (potentially lower resolution) component whereas the green represents the high resolution deterministic component. The re-centering procedure (blue) replaces the EnKF analysis mean with the hybrid analysis. 25

Figure 2.5: Horizontal localization scale (km) utilized in the GSI hybrid (Gaussian de-correlation length scale) derived from a database of ensemble forecasts (using stream function). 26

Figure 2.6: Time mean zonally-averaged standard deviation of the 3DVAR analysis increment for a real observation experiment (left) and OSSE experiment (right) for zonal wind (top, $m s^{-1}$), temperature (middle, K), and surface pressure (bottom, hPa). The real observations case covers the period 00 UTC 01 August 2006 through 18 UTC 30 August 2006, whereas the OSSE experiment covers the simulated period spanning 00 UTC 01 August 2005 through 18 UTC 30 August 2005. 30

Figure 2.7: Time mean zonally-averaged standard deviation of the analysis error the 3DVAR (left) and 3DHYB experiment (right) for zonal wind (top, $m s^{-1}$), temperature (middle, K), and specific humidity (bottom, $g kg^{-1}$), covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005. 33

Figure 2.8: Difference in the time mean zonally-averaged standard deviation of the analysis error (3DHYB-3DVAR) for zonal wind (top, $m s^{-1}$), temperature (middle, K), and specific humidity (bottom, $g kg^{-1}$), covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005. 34

Figure 2.9: Time-averaged standard deviation of the 300 hPa zonal wind analysis error (ms^{-1}) for the 3DVAR (top) and 3DHYB (experiments) for the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005, as well as the difference (3DHYB-3DVAR, bottom).	36
Figure 2.10: Time series of the standard deviation of the 500 hPa zonal wind (top, m s^{-1}) and 850 hPa temperature (bottom, K) background (solid) and analysis (dashed) errors for the 3DVAR (green) and 3DHYB (red) experiments for the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.	38
Figure 2.11: Time-averaged root mean square geopotential height errors (m) for forecasts from the 00 UTC analysis in the 3DVAR experiment as a function of lead time for the Northern Hemisphere (upper left) and Southern Hemisphere (lower left) verified against the ECMWF nature run for forecasts verifying between 27 July 2005 and 01 September 2005. The difference between two experiments (3DHYB-3DVAR) for the Northern Hemisphere (upper right) and Southern Hemisphere (lower right) is also plotted.	40
Figure 2.12: Time-averaged 500 hPa anomaly correlation (upper panels) for the Northern Hemisphere (left) and Southern Hemisphere (right) for the 3DVAR (green) and 3DHYB (red) experiments for forecasts from the 00 UTC analyses as a function of lead time as well as the difference (3DHYB-3DVAR, lower panels). The 95% confidence threshold for a significance test (derived from a standard t-test) is also plotted in the lower panels.	41
Figure 2.13: As in Fig. 2.11, but for the vector wind RMSE in the tropics.....	42
Figure 2.14: Time mean zonally-averaged standard deviation of the zonal wind 06-hr forecast (background) error for the 3DVAR (upper left) and 3DHYB (upper right) experiments, as well as the NMC-method based static error estimate (lower left) and 3DHYB 06-hr ensemble spread (lower right) covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.....	43
Figure 2.15: As in Fig. 2.8, but for the difference between 3DENSV-3DVAR.....	45
Figure 2.16: As in Figure 2.14 (lower right, from original 3DHYB experiment), but for the experiment 3DHYB_RS (reduced spread).	47
Figure 2.17: As in Fig. 2.8 (which had comparison 3DHYB-3DVAR), but for the difference between 3DHYB_RS-3DHYB.	48
Figure 2.18: As in Fig. 2.8 and Fig. 2.17, but for the difference between 3DHYB_RS-3DVAR.....	50
Figure 3.1: 500 hPa Temperature (shaded, K) and vector wind (vectors, m s^{-1}) analysis increment resulting from the assimilation of a single 500 hPa temperature observation (-2K innovation, 1K error) taken -3h from analysis time (location of observation denoted with black dot) for 4DVAR (upper left), H-4DVAR_AD (lower left), 4DENSV (upper right), and H-4DENSV (lower right). The reference arrow is representative of 1 m s^{-1}	64
Figure 3.2: As in Fig. 3.1, except for the analysis increment valid at the beginning (t-3h, top), middle (t=0h, middle), and end (t+3h, bottom) of the assimilation window for the H-4DVAR_AD (left) and H-4DENSV (right) cases.....	65
Figure 3.3: Power spectra of the divergence increment for a single analysis valid at 06 UTC 15 July 2010 on model level 35 (top, approximately 200 hPa) and 14 (bottom, approximately 850 hPa) utilizing 4DENSV (non-hybrid) with no	

constraints (red), tangent linear normal mode constraint (blue), weak constraint digital filter (purple), and combined normal mode and weak constraint (green).	72
Figure 3.4: Observation penalty (top) and reduction of the gradient norm (bottom) by iteration for an analysis valid at 06 UTC 15 July 2010 utilizing 4DENSV (non-hybrid) with no constraints (red), tangent linear normal mode constraint (blue), weak constraint digital filter (purple), and combined normal mode and weak constraint (green).	73
Figure 3.5: Difference in the time-averaged (August) zonal mean standard deviation of the analysis error for the 4DENSV experiment relative to 3DENSV for zonal wind (top, ms^{-1}), temperature (middle, K), and specific humidity (bottom, g kg^{-1}), covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.	77
Figure 3.7: As in Fig. 3.6, but for the H-4DENSV_NMI minus H-4DENSV experiment errors.	81
Figure 3.8: Time series showing the global mean surface pressure error (hPa) for the background (top, solid) and analysis (bottom, dashed) from the H-4DENSV (red) and H-4DENSV_NMI (blue) experiments.	82
Figure 3.9: As in Fig. 3.7, but for the H-4DENSV_DFI minus H-4DENSV experiment errors.	84
Figure 3.10: The root mean square sum of the incremental spectral tendencies (top) for the total tendency (4DENSV-red dashed, 4DENSV_COMB-blue dashed) and gravity mode tendency (4DENSV-red solid, 4DENSV_COMB-blue solid) for the eight vertical modes kept as part of the TLNMC for the single analysis valid at 06 UTC 15 July as a function of observation bin (analysis relative time). Also plotted is the ratio (bottom, gravity mode/total tendencies) for the 4DENSV (red) and 4DENSV_COMB (blue) experiment.	86
Figure 3.11: As in Fig. 3.9 but for the H-4DENSV_COMB minus H-4DENSV experiment errors.	88
Figure 3.12: As in Fig. 3.11 but for the H-4DENSV_COMB minus 3DHYP experiment errors.	89
Figure 3.13: Time averaged analysis error standard deviation for 300 hPa zonal wind (ms^{-1}) from the 3DHYP (top) and H-4DENSV_COMB (middle), as well as the difference (bottom, H-4DENSV_COMB minus 3DHYP) for the simulated August period.	91
Figure 3.14: Analysis time-mean zonally averaged precipitable water (kg m^{-2} , upper left) and 06-hr forecasted precipitation (kg m^{-2} , upper right) for the simulated August period for the 3DENSV (blue, dashed), 4DENSV (orange, dashed), 3DHYP (red, solid) H-4DENSV_COMB (blue, solid) experiments. The differences are shown on the bottom between the 4DENSV minus 3DENSV (blue) and H-4DENSV_COMB minus 3DHYP (red) for precipitable water (kg m^{-2} , lower left) and precipitation forecast (kg m^{-2} , lower right).	92
Figure 3.15: As in Fig. 3.12 but for the H-4DENSV_COMB minus 3DVAR experiment errors.	94
Figure 3.16: The average geopotential height root mean square error by lead time for the 3DVAR control (green), 3DHYP (red) and H-4DENSV_COMB (blue)	

experiments verifying daily between the simulated 27 July and 01 September for the northern hemisphere (left) and southern hemisphere (right) at 1000 hPa (top) and 500 hPa (bottom). The difference between the experiments is shown on the bottom section of each panel for the 3DVAR-3DHYB (green) and H-4DENSV_COMB-3DHYB (blue) along with the 95% confidence derived from a t-test. Forecasts for each experiment were initialized at 00 UTC and verified against the ECMWF nature run. 96

Figure 3.17: As in Fig. 3.16, but for the vector wind root mean square errors in the tropics at 200 hPa (left) and 850 hPa (right). 97

Figure 4.1: Power spectra for zonal wind (top, kinetic energy m^2s^{-2}) and temperature (bottom, potential energy, K^2) for the nature run forecast (brown), 3DHYB 06-hr forecast error (red), and 06-hr EnKF perturbations from the 3DHYB (aqua) averaged from 06 UTC 08 August through 12 UTC 09 August. 105

Figure 4.2: Power spectra of the analysis increment from a 3DVAR (green) and dual-resolution 3DHYB (red) experiment valid at 06 UTC 15 July 2010 for 200 hPa divergence (upper left), 500 hPa vorticity (upper right), 850 hPa temperature (lower left) and surface pressure (lower right). 108

Figure 4.3: The scale-dependent weights, β_f^{-1} (black) and β_e^{-1} (grey) for the SD1 (left) and SD2 (right) single analysis experiments. 110

Figure 4.4: As in Fig. 4.2, but for the 3DHYB (red), SD1 (purple), and SD2 (light blue) experiments. 111

Figure 4.5: As in Fig. 4.2, but for the 3DHYB_RSSD cycling experiment. 113

Figure 4.6: Difference in the time-averaged (August) zonal mean standard deviation of the analysis error for the 3DHYB_RSSD experiment relative to 3DHYB_RS for zonal wind (top, $m s^{-1}$), temperature (middle, K), and specific humidity (bottom, $g kg^{-1}$). 114

Figure 5.1: Diagram of the cycling procedure used the operational NCEP GFS/GDAS and experiments in Chapter 5. Solid arrows represent the short term (09-h) forecast from the GDAS analysis (two errors per cycle represent a single forecast), whereas dashed arrows represent the 15-day forecast from the early (GFS) cycle. Green (red) colors represent those components that are (not) performed in the 3DVAR-R and 3DHYB-R experiments. Black arrows represent both the higher resolution deterministic and low resolution ensemble forecasts in the case of the hybrid. 120

Figure 5.2: Time-averaged root mean square geopotential height errors (m) for forecasts from the 00 UTC analysis in the 3DVAR-R experiment as a function of lead time for the Northern Hemisphere (upper left) and Southern Hemisphere (lower left) verified against the 3DVAR-R analyses for forecasts verifying between 05 August 2010 and 31 October 2010. The difference between two experiments ($[3DHYB-R]-[3DVAR-R]$) for the Northern Hemisphere (upper right) and Southern Hemisphere (lower right) is also plotted, where 3DHYB-R forecasts are verified against 3DHYB-R analyses. 122

Figure 5.3: The average geopotential height root mean square error by lead time for the 3DVAR-R (green) and 3DHYB-R (red) experiments verifying daily between 05 August 2010 and 31 October 2010 for the northern hemisphere (left) and southern hemisphere (right) at 1000 hPa (top) and 500 hPa (bottom). The

difference between the experiments is shown on the bottom section of each panel for the [3DHYB-R]-[3DVAR-R] along with the 95% significance. Forecasts for each experiment were initialized at 00 UTC and verified against analyses from their own experiment. 124

Figure 5.4: As in Fig. 5.4, but for anomaly correlation instead of RMSE..... 125

Figure 5.5: As in Fig. 5.2, but for vector wind RMSE instead of geopotential height. 126

Figure 5.6: Time-averaged wind speed bias (left) and vector wind RMSE (right) of 3DVAR-R (green) and 3DHYB-R (red) forecast with lead times of 24-h (top) and 48-h (bottom) as compared to rawinsonde and aircraft measurements in the tropics for the period covering 05 August 2010 through 31 October 2010. 127

Figure 5.7: The average tropical cyclone track error (km) for 3DVAR-R (green) and 3DHYB-R (red) forecasts for the period covering 10 August 2010 through 31 October 2010. The average track error was computed from a homogeneous sample of cases for storms in the Atlantic, East Pacific, and West Pacific basins. The number of cases for each lead time is identified below each forecast hour. Error bars indicate the 5th and 95th percentile of a re-sampled block bootstrap distribution. 129

Chapter 1: Introduction

1.1 Background

Operational Numerical Weather Prediction (NWP) centers such as the National Centers for Environmental Prediction (NCEP) have been using variational data assimilation techniques (Derber 1989) for initializing NWP model forecast for decades (Parrish and Derber 1992; Rabier et al. 2000; Lorenc et al. 2000; Kleist et al. 2009b). Variational methods are used to create a best estimate of the initial state for NWP models at a particular time by combining information from observations as well as a background, usually a short term NWP forecast from the same model. This best estimate is attained through the minimization of an objective function that includes measures of the weighted distance of an analysis from the observations and background. A key component to this procedure is the specification of the weights, or specifically, the error characteristics of both the observations and (model produced) background state.

Typically in these variational methods, the background error covariance matrix is estimated a-priori through the use of lagged forecast pairs, a free run of the NWP model, ensemble of forecasts, or by utilizing a comparison of model forecasts to observations (a comprehensive summary can be found in Bannister 2008a,b). The background error covariance matrix is typically kept static, meaning the analysis is unaware of flow-dependent errors (though some implicit flow-dependence is achieved in 4DVAR through the dynamics of the model by linearly propagating the increment throughout the assimilation window). For state of the art NWP models, the

error covariance matrix is prohibitively large and complex, and in particular, it is difficult to prescribe the multivariate aspects (for example, relating water vapor to the other dynamic variables).

Another class of algorithms based on the Kalman filter solution to the data assimilation problem has been developed in an attempt to avoid some of the weaknesses and difficulties with variational schemes. Ensemble Kalman filter (EnKF) data assimilation systems utilize fully flow-dependent background (and analysis) error covariances estimated from an ensemble of forecasts (e.g., Evensen 1994; Houtekamer and Mitchell 1998; Whitaker and Hamill 2002). In addition to being fully flow-dependent, the ensemble used to represent the background error covariance contains information about how each variable (at every location) is correlated with all of the other variables in the model. Two studies have demonstrated success in using an EnKF to assimilate real observations into the NCEP GFS (Szunyogh et al. 2008; Whitaker et al. 2008). The main disadvantage of the EnKF algorithms results from the fact that an ensemble of limited size is used to sample the background error covariance. This results in the need for various ad-hoc procedures such as localization and inflation to avoid filter divergence.

Fundamentally, EnKF and variational algorithms are solving for the same problem in slightly different manners. In fact, it is not difficult to reformulate one of the methods within the context of the other (given certain assumptions and considerations). For example, the extended Kalman filter can be formulated within a variational framework quite easily. We start with some nonlinear numerical model, M , that can propagate information forward in time (t):

$$\mathbf{x}^b_{t+1} = M(\mathbf{x}^a_t) \quad (1.1)$$

where \mathbf{x}^b and \mathbf{x}^a are the model state vector for the background (forecast of some length) and analysis respectively. Here, the $t+1$ represents relative cycle time, as the model propagates the analysis state forward in time to create the background state for the next cycle (i.e., the $t+1$ can represent 1-hr, 3-hr, 6-hr, etc. depending on the cycling frequency). The tangent linear model \mathbf{M} of nonlinear model M , and its adjoint, \mathbf{M}^T , can then be used to evolve the error covariances of the analysis (\mathbf{A}_{KF}) and background (\mathbf{B}_{KF}) forward in time,

$$\mathbf{B}_{\text{KF}} = \mathbf{M}\mathbf{A}_{\text{KF}}\mathbf{M}^T + \mathbf{Q} \quad (1.2)$$

where \mathbf{Q} represents the known model error. A new initial condition (analysis) can be found by updating the forecast background) state through the assimilation of observations

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K} (\mathbf{y}^o - \mathbf{H}\mathbf{x}^b) \quad (1.3)$$

where \mathbf{y}^o are the observations, \mathbf{H} the operator that translates the information from model state to observation space, and \mathbf{K} the Kalman gain matrix. This gain matrix is formulated to combine the observation operator with the known error covariances of the background and observations (\mathbf{R}):

$$\mathbf{K} = \mathbf{B}_{\text{KF}}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}_{\text{KF}}\mathbf{H}^T)^{-1}. \quad (1.4)$$

The analysis error covariance associated with the updated \mathbf{x}^a is then defined as an update to the background error covariance:

$$\mathbf{A}_{\text{KF}} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}_{\text{KF}} \quad (1.5)$$

representing a reduction of the uncertainty in the error covariance. Instead of solving for the \mathbf{x}^a in an optimal interpolation (OI) fashion as in equation 1.3, a variational cost function can be formulated using the same \mathbf{B} :

$$J_{\text{KF}}(\mathbf{x}') = \frac{1}{2}(\mathbf{x}')^T \mathbf{B}_{\text{KF}}^{-1}(\mathbf{x}') + \frac{1}{2}(\mathbf{H}\mathbf{x}' - \mathbf{y}')^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{x}' - \mathbf{y}'). \quad (1.6)$$

Here, the incremental form is used for simplicity and for direct comparison with derivations in other chapters, where the analysis increment and observation residuals (or innovations) are defined as

$$\mathbf{x}' = \mathbf{x}^a - \mathbf{x}^b \quad (1.7)$$

$$\mathbf{y}' = \mathbf{y}^o - \mathbf{H}\mathbf{x}^b. \quad (1.8)$$

The analysis increment (\mathbf{x}') is obtained by minimizing the cost function given by equation 1.6. The \mathbf{B}_{KF} is the same time-evolving background error covariance matrix that is derived from a propagation of the \mathbf{A}_{KF} , which is defined as the inversion Hessian of $J_{\text{KF}}(\mathbf{x}')$:

$$\mathbf{A}_{\text{KF}}^{-1} = \mathbf{B}_{\text{KF}}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad (1.9)$$

representing enhancement of accuracy.

For NWP applications, the size of \mathbf{x} is quite large ($\mathcal{O}(10^7)$), making explicit calculation of \mathbf{B}_{KF} and \mathbf{A}_{KF} (particularly the time propagation) practically impossible. The EnKF algorithms are designed to sample \mathbf{A} and \mathbf{B} through the use of an ensemble:

$$\mathbf{A}_{\text{EN}} = E \left[\left(\mathbf{x}_e^a - \mathbf{x}^{true} \right) \left(\mathbf{x}_e^a - \mathbf{x}^{true} \right)^T \right] \quad (1.10)$$

$$\mathbf{B}_{\text{EN}} = E \left[\left(\mathbf{x}_e^b - \mathbf{x}^{true} \right) \left(\mathbf{x}_e^b - \mathbf{x}^{true} \right)^T \right] \quad (1.11)$$

where \mathbf{x}_e now schematically denotes the ensemble of state vector states, \mathbf{x}^{true} is the known true state, and E the expectation. In reality, the true state is never really known, so the covariance matrices are approximated with ensemble perturbations (from an ensemble of size N), taking the background error covariance as an example:

$$\mathbf{B}_{EN} \cong \frac{1}{N-1} \sum_{n=1}^N \left((\mathbf{x}_e^b)^n - \bar{\mathbf{x}} \right) \left((\mathbf{x}_e^b)^n - \bar{\mathbf{x}} \right)^T \quad (1.12)$$

where the over bar is now representative of the ensemble mean. Assuming one had access to such a representation of the background error, it could easily be used within the variational framework as well just as in equation 1.6.

Variational assimilation algorithms do in fact have their own approximation to the background error covariance. This results in a very minor modification to the incremental 3D cost function from equation 1.6:

$$J_{3DVAR}(\mathbf{x}') = \frac{1}{2} (\mathbf{x}')^T \mathbf{B}_f^{-1} (\mathbf{x}') + \frac{1}{2} (\mathbf{H}\mathbf{x}' - \mathbf{y}')^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}' - \mathbf{y}'). \quad (1.13)$$

Here, a generic \mathbf{B}_f replaces the fully flow-dependent, time-evolved \mathbf{B}_{KF} from the extended Kalman filter framework. As mentioned in the introduction, this \mathbf{B}_f is typically calculated offline, kept static, and simplified through various approximations. The logical next step is to combine such information with an ensemble based estimate within the variational framework. This so-called hybrid method (Hamill and Snyder 2000; Lorenc 2003; Zupanski 2005) attempts to combine the advantages of both variational and EnKF methods while at the same time tries to ameliorate their deficiencies. Several studies have demonstrated the potential for hybrid methods using both simple models as well as full NWP models (Hamill and

Snyder 2000; Wang et al. 2007b, 2009; Buehner et al. 2010a, b). Some of the potential advantages of hybrid schemes over their stand-alone variational or EnKF counterparts include:

- Capability to incorporate fully flow-dependent, multivariate covariances within the variational scheme
- Supplementing an ensemble estimate with a thoroughly tested climatological **B**
- Ease of applying dynamic and physical constraints on the solution
- Implicit, physical-space localization, potentially advantageous for measurements such as satellite radiances
- Easy transition to operational NWP applications within the already established variational framework.

Results thus far have been so promising that several operational centers have implemented (UK Met office, Clayton et al. 2012) or are pursuing (NCEP, U.S. Navy, Environment Canada) hybrid methods for operational applications. Lastly, although a variety of (slightly different) hybrid algorithms have been developed and tested, it has been shown that many of them are theoretically equivalent (Wang et al. 2007a).

Although the term hybrid has come to mean different things in the literature, here it will refer to any configuration that actually blends an ensemble-based estimate for the background error covariance with a climatological estimate all within the variational framework. It is of course possible to hybridize an EnKF update through the use of a static estimate for **B**, but that is beyond the scope of this study. Any

variational-based configuration that uses only an ensemble perturbation representation of a covariance without blending with some static estimate will be referred to as “ensemble-var”.

Along the lines of 4D extensions to the EnKF algorithm (Hunt et al. 2004), several 4D variants to the use of ensembles within the variational framework have been proposed and developed as cheaper alternatives to the costly 4DVAR algorithm (Liu et al. 2008; Tian et al. 2008; Tian et al. 2011; Buehner et al. 2010a). Some of these algorithms are not truly hybrid within the context used in this work. However, the extension to supplement with a full-rank, climatological \mathbf{B} is fairly straight forward to implement. Because of the reduced computational resources required, such an algorithm could be very attractive for an operational NWP center, if it can be shown to be as effective as an algorithm such as 4DVAR.

1.2 Thesis Objectives and Outline

A hybrid data assimilation algorithm for use with the NCEP GFS model is explored in this study. Progress is already underway at NCEP to extend the operational assimilation algorithm, the GSI, to incorporate ensemble perturbations for use as a hybridization of the static error covariance. This work focuses on expanding upon this development, and tries to incorporate new features and algorithms to enhance the development. An initial operational GFS implementation of a dual-resolution hybrid scheme is planned for May 2012, and forms the building block which will be expanded upon in this work. We aim to answer several pertinent questions:

- Does the inclusion of ensemble covariances in a hybrid fashion reduce analysis errors?
- What are the key components or parameters yielding the best results within the dual resolution hybrid framework (localization, weighting, inflation, etc.)?
- What is the impact of expanding the 3D-hybrid to a 4D-ensemble based (and hybrid) extension? Can dynamical constraints improve the quality of analysis in the 4D context?
- Can ensemble information be incorporated into the hybrid in a better manner than has been utilized to this point (with single, global weighting parameters between the static and ensemble covariances)?

To answer some of these questions, a series of experiments utilizing an Observing System Simulation Experiment (OSSE) are carried out. The main advantage of this type of controlled experimentation is the access to the truth. Chapter 2 describes the details of the various components to the system, including the nature run and simulated observations as well as results from various 3D-Hybrid experiments using the NCEP GFS. An extension of the hybrid to include 4D ensemble perturbations is described in Chapter 3. This includes a novel approach to incorporating a hybridization of the 4D-ensemble-var algorithm, as well as proposals for a variety of dynamic constraints. Various 4D extensions are tested using the same simulated environment as is described in Chapter 2. This is then followed by Chapter 4 which investigates the use of scale-dependent weighting within a hybrid scheme. Experiments using real observations are then carried out in Chapter 5, in an effort to show the forecast impact of using a 3D-Hybrid instead of 3DVAR algorithm for

initializing a pseudo-operational version of the GFS model. This is then followed by the final chapter which provides a summary, as well as a variety of issues and ideas for future work and further development.

Chapter 2: System Description and 3D-Hybrid OSSE results

2.1 Introduction

To combine the advantages of the ensemble and variational methods while at the same time attempting to minimize their weaknesses, hybrid assimilation methods have been proposed and developed (Hamill and Snyder 2000; Lorenc 2003; Zupanski 2005). Typically, the proposed hybrid methods have utilized the variational method for the purposes of calculating the analysis increment (though, it is possible that one could utilize the EnKF framework and supplement the ensemble with a static background error covariance). Many of the (slightly different) algorithms have been shown to be theoretically equivalent, whether using a combined covariance through brute force or through a variational-based control variable method (Wang et al. 2007).

In the hybrid methods that utilize the variational framework, the flow-dependent, ensemble-based covariances are added to the cost function through the extended control variable method (Lorenc 2003; Buehner 2010a, b):

$$J(\mathbf{x}'_f, \boldsymbol{\alpha}) = \beta_f \frac{1}{2} (\mathbf{x}'_f)^T \mathbf{B}_f^{-1} (\mathbf{x}'_f) + \beta_e \frac{1}{2} \sum_{n=1}^N (\boldsymbol{\alpha}^n)^T \mathbf{L}^{-1} (\boldsymbol{\alpha}^n) + \frac{1}{2} (\mathbf{H}\mathbf{x}'_t - \mathbf{y}')^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}'_t - \mathbf{y}') \quad (2.1)$$

where \mathbf{B}_f is the static background error covariance, \mathbf{R} the observation error covariance, \mathbf{H} the observation operators, \mathbf{y}' the observation residuals (equation 1.8), $\boldsymbol{\alpha}^n$ the ensemble control variable for each member with \mathbf{L} defined as its covariance matrix (just as described in Lorenc 2003; Wang 2010). Here, n is the index for the

ensemble members for the total ensemble size N . This is equivalent to replacing \mathbf{B}_{KF} from equation 1.6 and \mathbf{B}_f from equation 1.13 with a (weighted) linear combination of \mathbf{B}_{EN} (equation 1.12) and \mathbf{B}_f (Hamill and Snyder 2000). The analysis increment, \mathbf{x}' , is obtained by minimizing the cost function (equation 2.1) and can be partitioned into two components:

$$\mathbf{x}'_t = \mathbf{x}'_f + \sum_{n=1}^N \left(\boldsymbol{\alpha}^n \circ \mathbf{x}_e^n \right). \quad (2.2)$$

The total analysis increment (\mathbf{x}'_t) is the sum of that which is derived from the static error covariance (\mathbf{x}'_f) and that which is derived from the ensemble perturbations (\mathbf{x}_e^n) as prescribed by the control variable ($\boldsymbol{\alpha}$). There is a single ensemble control variable for all variables, so for the single-resolution case where the ensemble is the same resolution as the analysis, the dimension of $\boldsymbol{\alpha}$ is equal to that of the 3D analysis grid. However, the additional computational cost associated with the addition of the hybrid control variable method is much smaller than adding N 3D variables to the control vector due to the simplified formulation of \mathbf{L} . Further computational savings can be achieved through the use of a dual-resolution framework, where the ensemble perturbations are applied using a lower resolution ensemble grid.

There are tuning parameters, β_f and β_e , that control how heavily to weight the static and ensemble contributions respectively. Historically, care has been taken to ensure that the sum of the inverses of the weighting parameters is equal to one. This is done in an effort to ‘preserve’ the effective error variances and use this parameter simply to assign the weight given to static and ensemble estimates. As an example, utilizing a β_e^{-1} parameter setting of 0.75 (thereby making $\beta_f^{-1} = 0.25$) results in the

hybrid analysis relying 75% on the ensemble and 25% on the static estimate for the background error covariance.

Various studies have demonstrated that the hybrid algorithm can in fact improve upon variational or ensemble-based algorithms on their own (Hamill and Snyder 2000; Wang et al. 2007b). A particularly thorough comparison that was described in Buehner et al. (2010 a, b) showed that a hybrid 4D algorithm improved upon the Canadian operational 4DVAR system. The advantages gained over the variational standalone systems come from the improved specification of the fully flow-dependent background error covariance with better multivariate definitions through the use of the ensemble. Additionally, hybrid systems have the ability to improve upon the standalone ensemble systems by linearly evolving the covariance in the full dimension space of the NWP model (in the case of 4DVAR hybrid with adjoint) and by applying the localization in physical space (whereas most EnKF algorithms for large NWP applications perform localization in observation space). Physical space localization is particularly helpful for the assimilation of those observations which are not point measurements, but instead integrated quantities such as satellite brightness temperatures (Campbell et al. 2010). This is in part because physical space localization avoids the need to explicitly assign a vertical location for such observations. It has also been suggested that the use of hybrid algorithms can be particularly useful for small ensemble sizes (Wang et al. 2007b, 2009).

Motivated by the successes of others in applying variational-based hybrid algorithms to the NWP data assimilation problem, a hybrid algorithm has been developed for the operational variational assimilation system at NCEP. This chapter

describes several experiments that utilize simulated observations to evaluate and demonstrate the impact of hybrid DA on the quality of analyses and subsequent forecasts using an observing system simulation experiment (OSSE). Section 2.2 describes the components of the NCEP data assimilation system including the implementation of the hybrid algorithm. Section 2.3 then provides details regarding the Joint OSSE project and simulated observations. This is followed by section 2.4 which provides a description of the various experimental components. Sections 2.5 and 2.6 describe experimental results from assimilating the simulated observations using 3DVAR and 3D-Hybrid configurations. Lastly, a summary and motivation for the subsequent chapters then follows.

2.2 NCEP Hybrid Data Assimilation

2.2.1 System Description

The gridpoint statistical interpolation (GSI) is a physical-space based variational analysis scheme (Wu et al. 2002; Kleist et al. 2009b) that has been made operational for several NCEP applications including the global data assimilation system (GDAS) and initialization of the global forecast system (GFS) model, used to produce global medium range deterministic forecast guidance (as well as boundary conditions for other applications). Although a variety of minimization algorithms exist within the code, a preconditioned (full background error covariance \mathbf{B}) double conjugate gradient solver (Derber and Rosati 1989) is the default choice within the GSI. Recursive filters are utilized (Purser et al. 2003 a, b) to model some of the off-diagonal components of the background error covariance matrix. Although the GSI

does have a 4DVAR capability embedded within it, it has yet to be exercised within the operational applications at NCEP. This is mostly due to the lack of development of the necessary tangent-linear and adjoint model codes as well as the significant computational cost associated with 4DVAR. Several dynamic and physical constraint options have been developed such as the penalty terms for unphysical tracers, incremental normal mode initialization, conservation of global mean dry mass, and a weak constraint digital filter for 4D applications (some of these are discussed in greater detail in section 3.3). The GSI that is run routinely as part of the GDAS uses a static estimate for the background error covariance which is supplemented with a tendency-based algorithm to apply flow dependent variances. In this algorithm, the static variances undergo a simple rescaling based on the 6-hr tendency in the model forecast valid for the assimilation window. The variances are increased (decreased) where the model tendencies are relatively large (small). An example of the reweighting procedure can be found in Figure 5 of Saha et al. (2010). The reweighting procedure is fairly simplistic and does not address the off-diagonal (stretching) nor multivariate deficiencies in the static error covariance estimate. More detailed information regarding the GSI can be obtained from the Developmental Testbed Center user's guide available online at <http://www.dtcenter.org/com-GSI/users/index.php>.

Following a similar procedure that is outlined in Wang (2010) as first proposed by Lorenc (2003), the hybrid method is implemented into the GSI using an extended control variable (as in equation 2.1)¹. Initially, the hybrid option was only developed for use with the GFS model, but has since been expanded to other

¹ Control variable method originally implemented by David Parrish and Daryl Kleist of NCEP/EMC.

applications (such as the NCEP regional and hurricane models). The hybrid option in GSI includes a dual-resolution capability, where the ensemble perturbations being used can be at lower resolution than the deterministic forecast and subsequent analysis. This also reduces the additional cost of the ensemble control variable over standard 3DVAR since the dimension of α is then reduced to the ensemble grid (with some work necessary to interpolate between analysis and ensemble grids). This is important for high resolution operational applications, where it is not affordable to run an ensemble at full resolution. As with standard 3DVAR in GSI for global applications, another option within the hybrid allows for an incremental tangent linear normal mode (strong) constraint to be applied (Kleist et al. 2009a) to the total analysis increment (the sum of the components from static and ensemble error covariances) or to the static contribution only.

The localization of the ensemble based covariances is handled through the specification of the error covariance matrix (\mathbf{L} , equation 2.1) for the ensemble control variable. The function \mathbf{L} is assumed to be Gaussian and explicitly local with unit amplitude. In the GSI for global applications, this function is applied by utilizing spectral correlation functions for the horizontal and a recursive filter in the vertical instead of utilizing quasi-Gaussian functions with compact support as is typically done for the EnKF. The GSI-specific implementation of the hybrid allows for a level-dependent specification of the decorrelation length scales used in \mathbf{L} (one each for the horizontal and vertical). Since there is (currently) only a single control variable (dimensioned to be the size of the 3D ensemble grid) for the entire ensemble perturbation matrix, the effective localization is the same for all analysis variables.

2.2.2 Single Observation Example

To demonstrate that the implementation of the hybrid algorithm has been done correctly in the GSI, an experiment that assimilates only a single observation is performed. The observation chosen for this test is a single temperature observation for a particular case, and assigned a deviation of one degree from the background (a 6-hr GFS forecast) and an observation error of one degree. The observation is chosen to reside in an area of a large local gradient in temperature in an attempt to demonstrate the importance of the flow-dependent part of the background error specification (Fig. 2.1). Three GSI runs are performed using this single observation, including a standard 3DVAR run with static error covariance estimate, and two hybrid assimilation runs utilizing β_e^{-1} specifications of 0.5 and 1.0 to demonstrate the case of equal and full ensemble weighting, respectively. For the hybrid runs, an 80 member ensemble (from an ensemble of forecasts initialized using an offline EnKF) is used to prescribe the background error. A comparison of the experiments reveals that the resultant increment from the hybrid options is heavily stretched along the temperature gradient from southwest to northeast (Fig. 2.1). To contrast, the increment from standard 3DVAR is smaller in amplitude, larger in scale, and nearly isotropic, spread equally with little knowledge of the background temperature gradient. The hybrid run that utilizes the 0.5 weighting qualitatively appears to be a smoothed out version of the run that utilized a pure ensemble specification of the background error. Although not shown, the analysis increment that results from an EnKF run using the same ensemble and observation looks qualitatively and quantitatively similar to the hybrid run that utilized the 1.0 weighting.

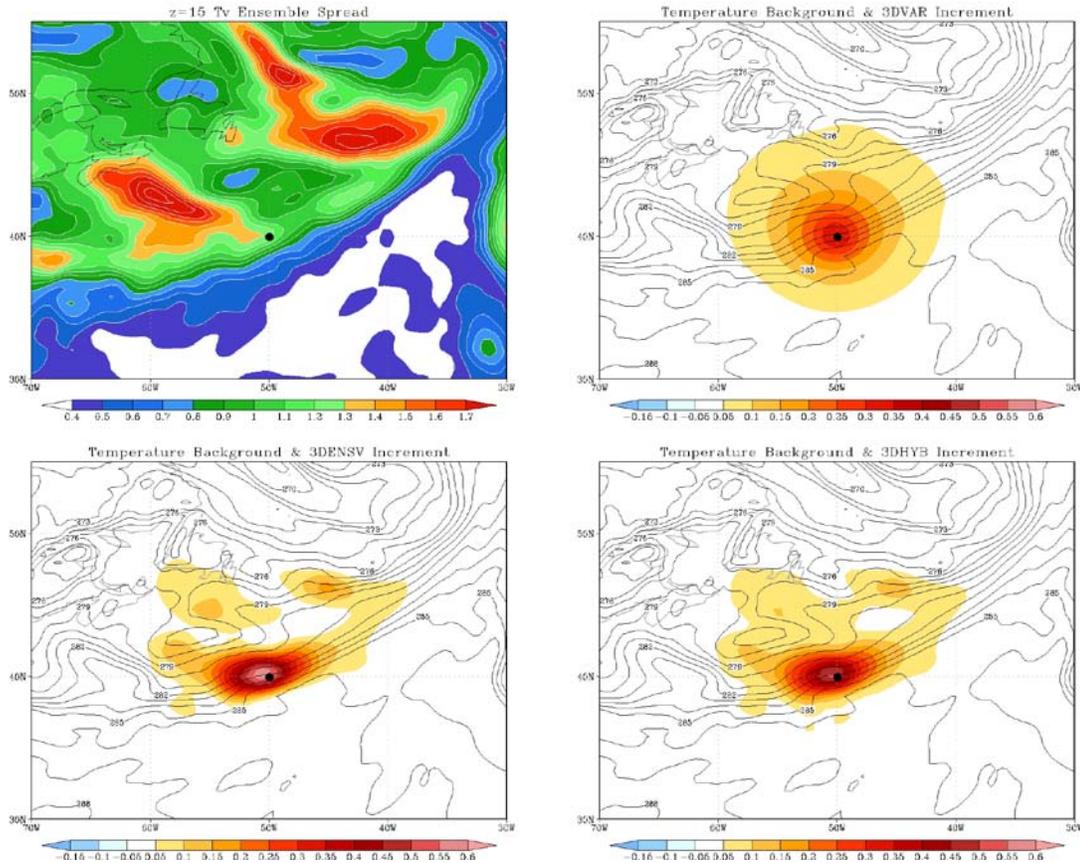


Figure 2.1: Model level 15 (~850 hPa) ensemble spread (K, upper left) as well as the 850 hPa temperature background (contours, other three panels) and analysis increment (shaded, 1K interval) resulting from the assimilation of a single 850 hPa temperature observation (location denoted with dot) given a 1K deviation from the model guess and 1K observation error for a 3DVAR (upper right), 3DENSF (lower left) and 3DHYS (50% weighting, lower right) valid at 00 UTC 12 September 2008.

2.3 Joint OSSE

An OSSE is typically designed to be able to investigate the potential impact of a future observing system (Masutani et al. 2007, 2010). However, OSSE experiments can also be utilized to investigate various aspects of a data assimilation system such as analysis error (Errico et al. 2007). In an OSSE for atmospheric NWP applications, a reference state (nature run), is first generated by making a climate, uninterrupted

free run using a NWP model (using state-dependent boundary conditions such as observed sea surface temperatures, if desired). This free run is then considered to be the true state. Simulated observations are generated by extracting the appropriate information from the nature run and adding realistic errors. The simulated observations can then be used by a data assimilation system to assess their impact on analysis and forecast accuracy. For an OSSE to be useful, it is critical to ensure that the nature run is a suitable representation of the real atmosphere. In order to achieve realistic results, it is important to use a model within the data assimilation system that is different than that which was used to generate the nature run (if the same model is used for both, the so-called identical-twin experiment, model error goes unaccounted for).

An internationally collaborative effort called the Joint OSSE has formed over the past several years. A nature run for use within the Joint OSSE community has been generated by the European Center for Medium Range Weather prediction (ECMWF), by completing a 13-month forecast using cycle 30r1 of their Integrated Forecast System model with T511 horizontal resolution and 91 vertical levels (Andersson and Masutani 2010). The initial condition for the run was 12 UTC on 01 May 2005, though since this was a free run of the forecast model, the starting date is not very relevant. The nature run was carefully evaluated to ensure it was realistic in terms of general climatology, storm tracks, as well as clouds and precipitation (Reale et al. 2007, McCarty et al. 2012). From this nature run, scientists at NCEP and NASA have simulated observations that were operationally available in 2005, including ‘conventional’ (radiosonde, surface, aircraft, satellite derived atmospheric

motion vectors, wind profiler, ship and buoy, scatterometer-based surface winds, etc.) and satellite microwave and infrared brightness temperature (including HIRS, AMSUA, AMSUB, MHS, AIRS on the simulated polar orbiters of NOAA 14, 15, 16, 17 as well as AQUA, in addition to the GOES sounder radiances from simulated geostationary satellites GOES 10 and 12). The simulated observations have been assimilated into a NWP model and gone through initial validation to ensure their usefulness (Errico et al. 2012). Both the nature run and simulated observations have been made available to the research community for the purpose of running OSSE experiments.

A subset of calibrated observations covering the simulated period from 01 July through 31 August was generated by GMAO and made available. Observations were only simulated to correspond to the NCEP GDAS (late cut-off) cycle (example of observations available for a single cycle shown in Figs. 2.2 and 2.3). The errors that were generated and added to the simulated observations from the nature run were calibrated in an attempt to match observation impacts from the real system, evaluated with a series of data denial (observing system experiment) runs (Privé, personal communication)².

2.4 Experiment Design

To test the impact of the various components and aspects of including a hybrid variational-ensemble component to the system, it is necessary to first produce a fully-cycled 3DVAR (non-hybrid) run utilizing simulated observations from the

² Simulated observations kindly provided by Nikki Privé, NASA GMAO.

Joint OSSE as described in section 2.3. The model used for the assimilation experiments is a degraded resolution version of the operational NCEP GFS that became operational in May 2011. The version of the GFS utilized in this study is a T382 spectral model with 64 hybrid sigma-pressure vertical levels. A description of the GFS model version 9.0.1 is available from the NCEP Environmental Modeling Center (EMC), <http://www.emc.ncep.noaa.gov/GFS/doc.php>. No tuning is done to the physical parameterization schemes for this work. The same version of the model is used for both data assimilation cycling and longer (7 day) free forecasts.

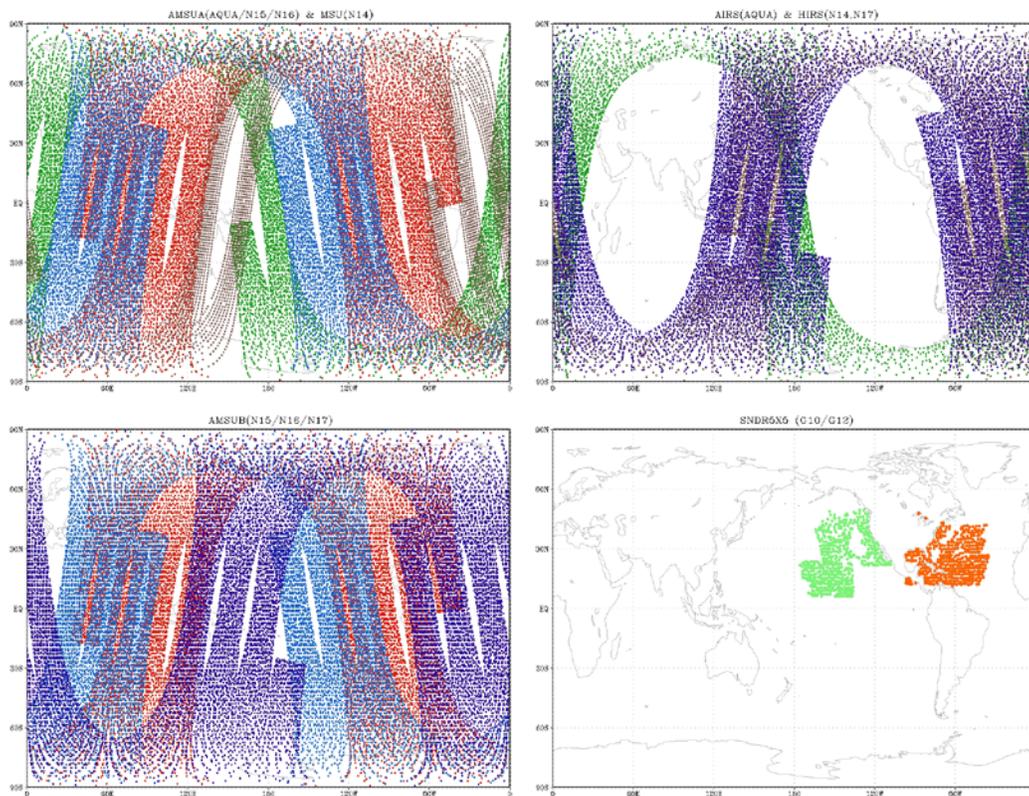


Figure 2.2: Spatial distribution of simulated satellite radiance observations available for assimilation valid at 00 UTC 24 July 2005 from the Joint OSSE for AMSUA and MSU (upper left), AIRS and HIRS (upper right), AMSUB (lower left) and GOES Sounder (lower right) from AQUA (dark green), NOAA-14 (brown), NOAA-15 (red), NOAA-16 (blue), NOAA-17 (purple), GOES-12 (orange), and GOES-10 (light green).

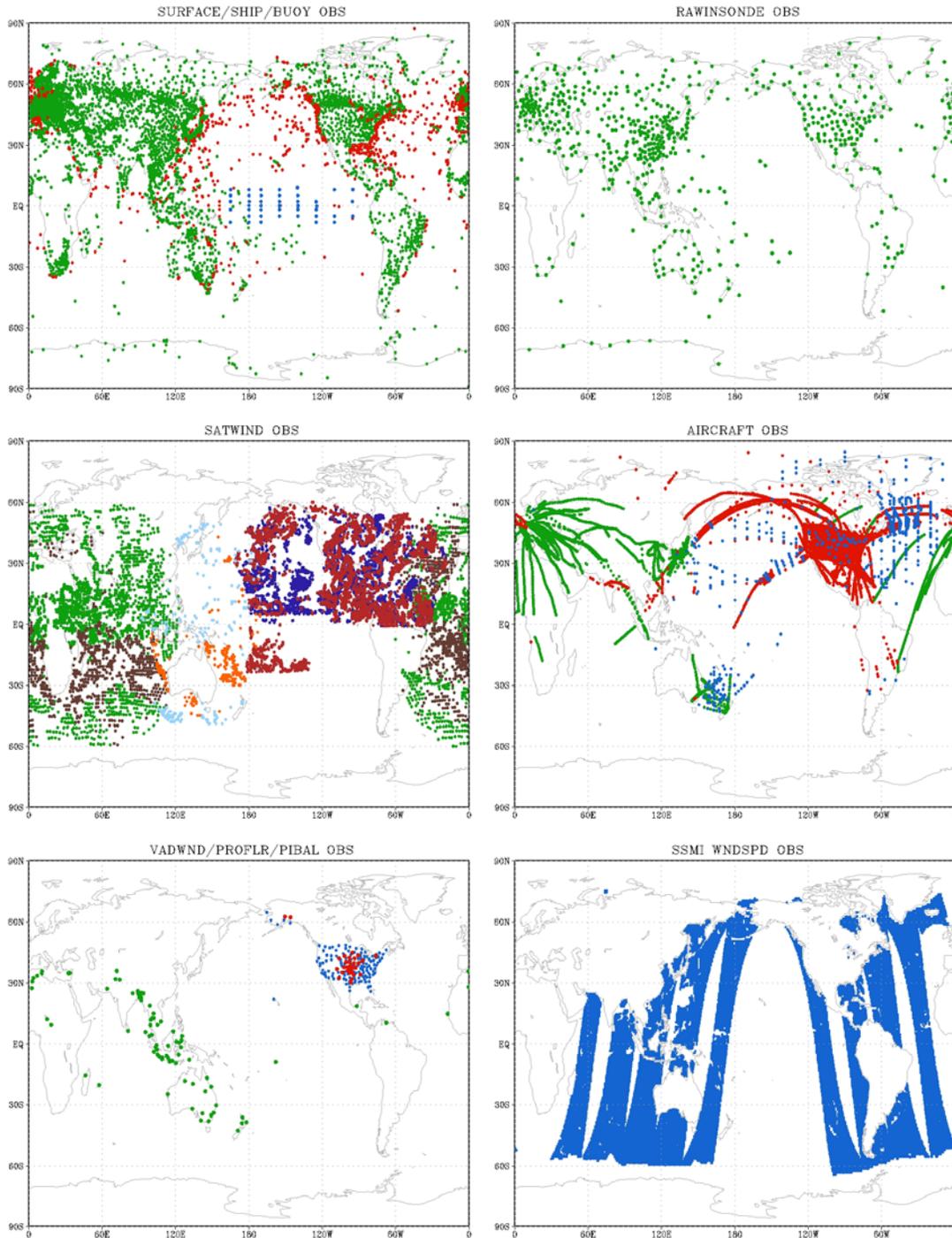


Figure 2.3: As in Fig. 2.2, but for ‘conventional observation’ including surface (upper left, land meter [green], ships & buoys [red], and moored buoy [blue]), rawinsonde (upper right), AMVs (middle left, from GOES [red, blue], EUMETSAT [green, brown], and JMA [orange, aqua]), aircraft (middle right), VAD winds (blue), lidar wind profilers (red) and pibal (green) [lower left], and SSMI derived surface wind speeds (lower right).

2.4.1 3DVAR

The 3DVAR control experiment is configured to mimic as closely as possible an operational configuration, in the hope that the OSSE-based results will match the real system. The hybrid-ready version of the GSI is utilized in the control run for proper comparison. The static background error estimate for this version of the model is derived using the so-called NMC method (Parrish and Derber 1992), extracting statistics from 24-hr and 48-hr lagged forecast pairs (and the same tuning parameters that are used in the operational GSI). The incremental normal mode constraint (TLNMC, Kleist et al. 2009a), global mean dry mass weak constraint, and flow-dependent variance reweighting are all utilized. Radiative transfer calculations for assimilation of satellite radiances within the GSI are performed using the JCSDA Community Radiative Transfer Model (CRTM, Han et al. 2006). This particular version of the CRTM is slightly different than the version that was used to generate the simulated brightness temperatures from the nature run. Since the simulated observations were only generated for the GDAS (late cut-off) cycle, no early data cut-off (GFS) cycles are performed. Long forecasts are carried out once per day at 00 UTC from the analysis generated by assimilating the GDAS simulated observations.

The GSI/GFS is cycled through the period for which the simulated observations were made available, 01 July through 31 August. An initial condition for 01 July is spun up by assimilating the Joint OSSE simulated observations in a separate, offline, low-resolution experiment (that assimilated observations for the simulated June period). The spun-up 01 July initial condition is utilized for the starting point for all experiments (including those that utilize various hybrid

configurations). From this initial condition, an additional two weeks of the experiment is ignored to allow for proper spin-up to this experimental configuration.

2.4.2 3D-Hybrid

A hybrid experiment (hereafter referred to as 3DHYB) is carried out that utilizes an ensemble of forecasts which are initialized each cycle with a serial square root filter form of an EnKF (Whitaker and Hamill 2002). The version of the EnKF update used in the experiment utilizes the GSI for all of the observation operators and applies both additive and multiplicative inflation after the ensemble update step. The multiplicative inflation factor (ρ) is proportional to the posterior standard deviation reduction from the prior through the assimilation of observations:

$$\rho = \omega \left(\frac{s^a - s^b}{s^a} \right) \quad (2.3)$$

where s is the sample standard deviation for each component to the analysis control vector:

$$s = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2. \quad (2.4)$$

The analysis perturbations $\{x_n^a\}$ are post multiplied by this factor, ρ . The amplitudes are controlled through the global tuning parameter ω . It can be seen that the inflation factor is allowed to be different for each variable, vertical level, and horizontal grid point and in general will be larger where observations are dense. The additive inflation extracts perturbations from a database of lagged forecast pairs (24-hr and 48-hr) within a window about the valid analysis time (to ensure the perturbations are

appropriate for the given season), and the final analysis post-perturbation (\mathbf{x}^{ap}) becomes:

$$\mathbf{x}_n^{\text{ap}} = \mathbf{x}_n^{\text{a}} + \kappa(\mathbf{x}'_r). \quad (2.5)$$

Each ensemble member, n , is assigned a different quasi-random perturbation vector, \mathbf{x}'_r . The lagged forecast database contains cases spanning an entire year (4 per day associated with the data assimilation cycling frequency). There is a single amplitude parameter (κ) to rescale the perturbations that are extracted from the lagged forecast pairs. A more detailed description of the inflation procedure can be found in Whitaker and Hamill (2012). The 3DHYB experiment utilizes the default parameter settings of $\omega=0.85$ and $\kappa=0.32$.

To maintain synergy between the ensemble and hybrid (deterministic) analyses, the ensemble mean is replaced every cycle using a re-centering procedure that ensures the ensemble of analysis perturbations are always centered about the (presumably better) hybrid analysis (schematic of the hybrid procedure can be seen in Fig. 2.4). The same localization parameters used for the hybrid (control variable in physical space) are used in the EnKF (observation space). A level dependent specification for the horizontal decorrelation length scales that is derived from an ensemble of forecasts (similar to Pannecoucke et al. 2008) is used (Fig. 2.5). A single value of 0.5 scale heights is used for the vertical localization in the hybrid. A conversion factor is applied in the EnKF to convert between the two localization definitions (distance to zero within the Gaspari and Cohn (1999) framework versus the Gaussian e-folding distance for the hybrid).

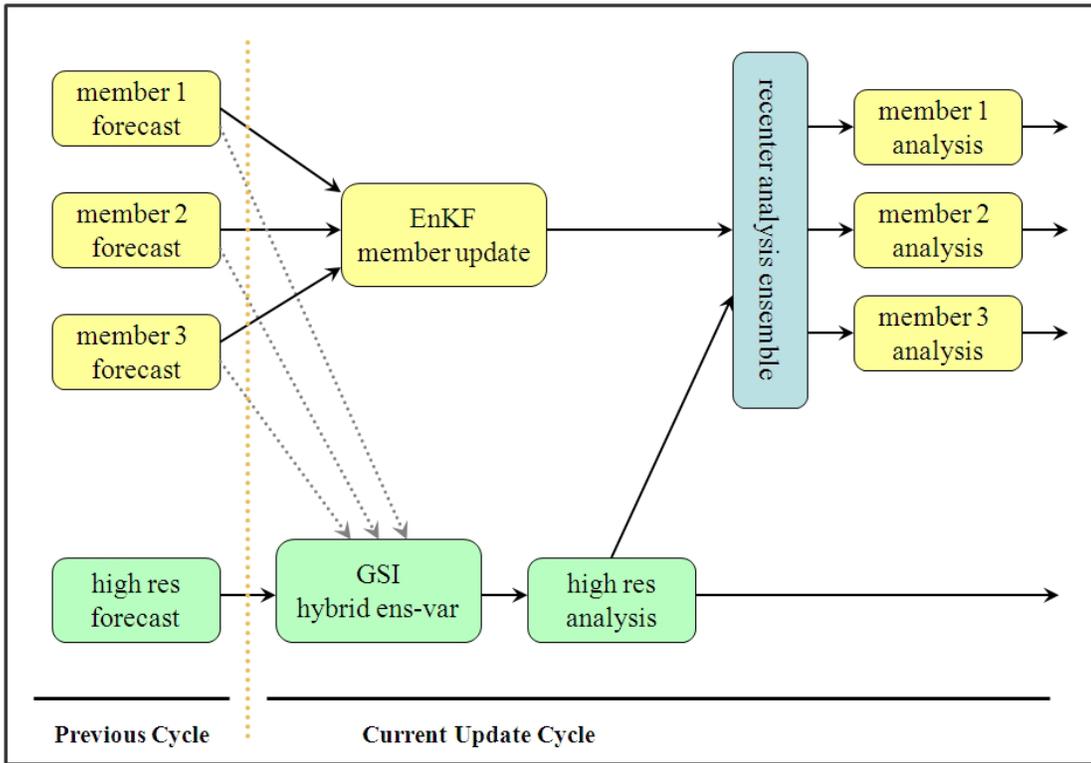


Figure 2.4: Schematic showing how the hybrid variational-ensemble paradigm operates for a given cycle using a three member ensemble. The yellow denotes the ensemble (potentially lower resolution) component whereas the green represents the high resolution deterministic component. The re-centering procedure (blue) replaces the EnKF analysis mean with the hybrid analysis.

The ensemble used in the experiment contains 80 members that are run at T254 spatial resolution, but maintains the same specification of (64) sigma-pressure vertical levels using an identical version of the GFS model (physical parameterizations are not changed despite the degraded resolution). The 3DHYB experiment uses parameter settings of $\beta_f^{-1}=0.25$ and $\beta_e^{-1}=0.75$, so the static contribution to the analysis increment is chosen to be 25% (and therefore, relying 75% on the ensemble covariances).

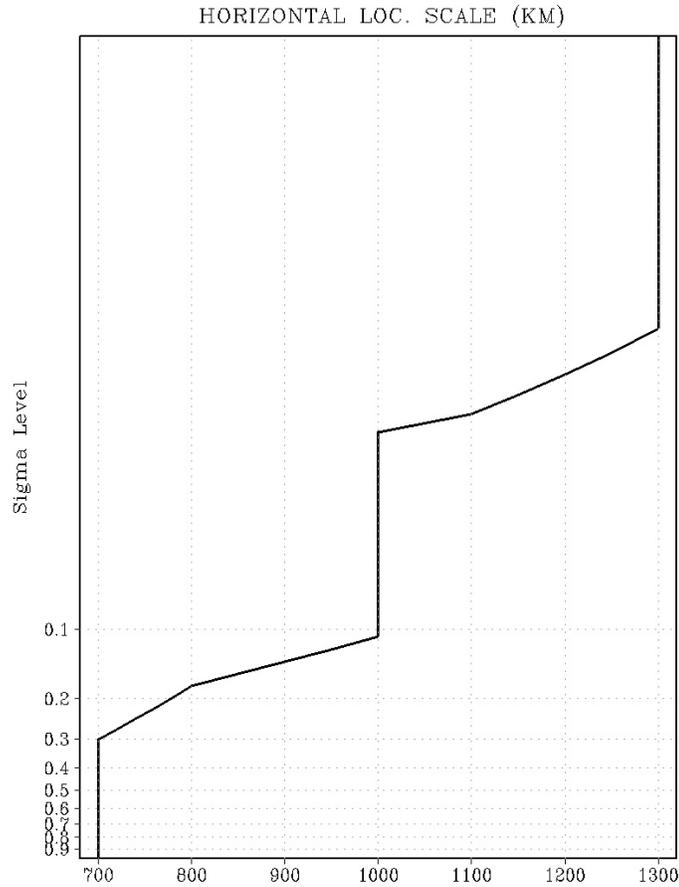


Figure 2.5: Horizontal localization scale (km) utilized in the GSI hybrid (Gaussian de-correlation length scale) derived from a database of ensemble forecasts (using stream function).

The initial ensemble is generated by taking the same initial condition for 01 July that is utilized in the control experiment and performing the additive inflation procedure as in equation 2.5 (with 100% larger κ). As with the control, the first two weeks of the experiment are ignored to allow for the analysis (and ensemble) to spin up. The initial condition for the high resolution deterministic component to the 3DHYP experiment is identical to that from the 3DVAR control. The hybrid experiment selects simulated observations to assimilate from the exact same set of files that the 3DVAR control utilizes, though the exact data selection and counts are

allowed to differ for any given cycled based on GSI internal quality control checks (gross error, as well as decisions made by the variational quality control procedure within the minimization).

The observations that are selected and assimilated as part of the EnKF update are not identical to the subset used within the hybrid update. The data selection and quality control decisions for the EnKF are driven by the (low resolution) guess from the ensemble mean instead of the high resolution deterministic guess. The data selection for this particular EnKF, while utilizing the GSI for the observation operators, is done in this way as a practical means to ensure that identical observations are chosen for all of the ensemble members. If the GSI-based observer is allowed to operate on each ensemble member background independently without further constraint, a different set of quality control and thinning decisions will be made for each member. Additionally, a coarser mesh (225 km) is used for the satellite thinning procedure for the ensemble relative to the high resolution deterministic update (which uses a 150 km mesh). This results in fewer observations being assimilated into the lower resolution ensemble update. Since the EnKF used in this study is a serial, square root filter, the reduced number of satellite observations assimilated helps improve efficiency without degrading the quality of the ensemble update.

The analysis increment derived from the hybrid deterministic update utilizes the tangent linear normal mode constraint (Kleist et al. 2009a) on the total analysis increment. Applying the constraint in this manner maintains consistency in terms of the balance and noise characteristics of the analysis increment between the 3DVAR

and 3DHYP experiments. The total analysis increment (as in equation 2.2) that is used in the observation term in (2.1) then becomes

$$\mathbf{x}'_t = \mathbf{C} \left[\mathbf{x}'_f + \sum_{n=1}^N \left(\boldsymbol{\alpha}^{n \circ} \mathbf{x}_e^n \right) \right] \quad (2.6)$$

where \mathbf{C} represents the forward linear normal mode constraint operator (exactly as described in Kleist et al. 2009a). In addition to the consistency between the increments within the 3DVAR and 3DHYP paradigms, the use of such a constraint can act to help ameliorate potential imbalance and noise introduced through the localization of the control variable as well as sampling error inherent to the ensemble covariances. An alternative would be to filter the ensemble perturbations themselves, though this would be computationally expensive relative to the application of the constraint over a single instance of the total analysis increment. A summary of all of the OSSE-based experiments performed for this chapter can be found in Table 2.1.

2.5 3DVAR Results

As an initial validation of the experimental configuration, an evaluation of the zonal mean square root of the variance in the analysis increments from the 3DVAR experiment is carried out as in Errico et al. (2007). The zonal wind increment in the OSSE-based 3DVAR control exhibits two local maxima (Fig. 2.6b), associated with the extratropical jets and consistent with the background error variance specification. Two secondary maxima are also noted in the near surface extratropics, with larger amplitude in the southern hemisphere consistent with the season for which the observations are assimilated.

<i>Experiment</i>	<i>Control</i>	<i>Description</i>	<i>Relevant Figures</i>	<i>Relevant Equations</i>
3DVAR	--	3DVAR Control $\beta_e^{-1}=0.0$	Fig. 2.6	Eq. 2.1
3DHYB	3DVAR	3D-Hybrid (TLNMC) $\beta_f^{-1}=0.25, \beta_e^{-1}=0.75$	Fig. 2.7- Fig. 2.13	Eq. 2.1 Eq. 2.6
3DENS	3DVAR 3DHYB	3D-Ensemble-Var No TLNMC $\beta_f^{-1}=0.0, \beta_e^{-1}=1.0$	Fig. 2.15	Eq. 2.1
3DHYB_RS	3DHYB	3DHYB + Reduced Inflation	Fig. 3.9	Eq. 2.1 Eq. 2.3- Eq. 2.6

Table 2.1: Description of various 3D (var and hybrid) OSSE-based experiments, as well as relevant figures and equations.

The increments from the OSSE-based 3DVAR experiment are then compared in a qualitative sense to a system that assimilated real observations over a similar time period. To perform this comparison, analysis increments are extracted from an experiment that utilizes a T382 version of the GFS model as well as the GSI for assimilation (Kleist et al. 2009b)³. One has to keep in mind that the atmospheric state (truth) is different in this comparison, since the OSSE-based experiment utilized a model free run as a baseline. The zonal wind increment in the real system exhibits similar structure to that in the OSSE-based run, with maxima associated with the extratropical jets and a secondary maximum in the near-surface southern hemisphere extratropics (Fig. 2.6a). Interestingly, the magnitude of the low-level maxima is

³ Experiment was part of a pre-implementation run to test the impact of replacing a spectral analysis with the GSI as part of the GDAS/GFS system.

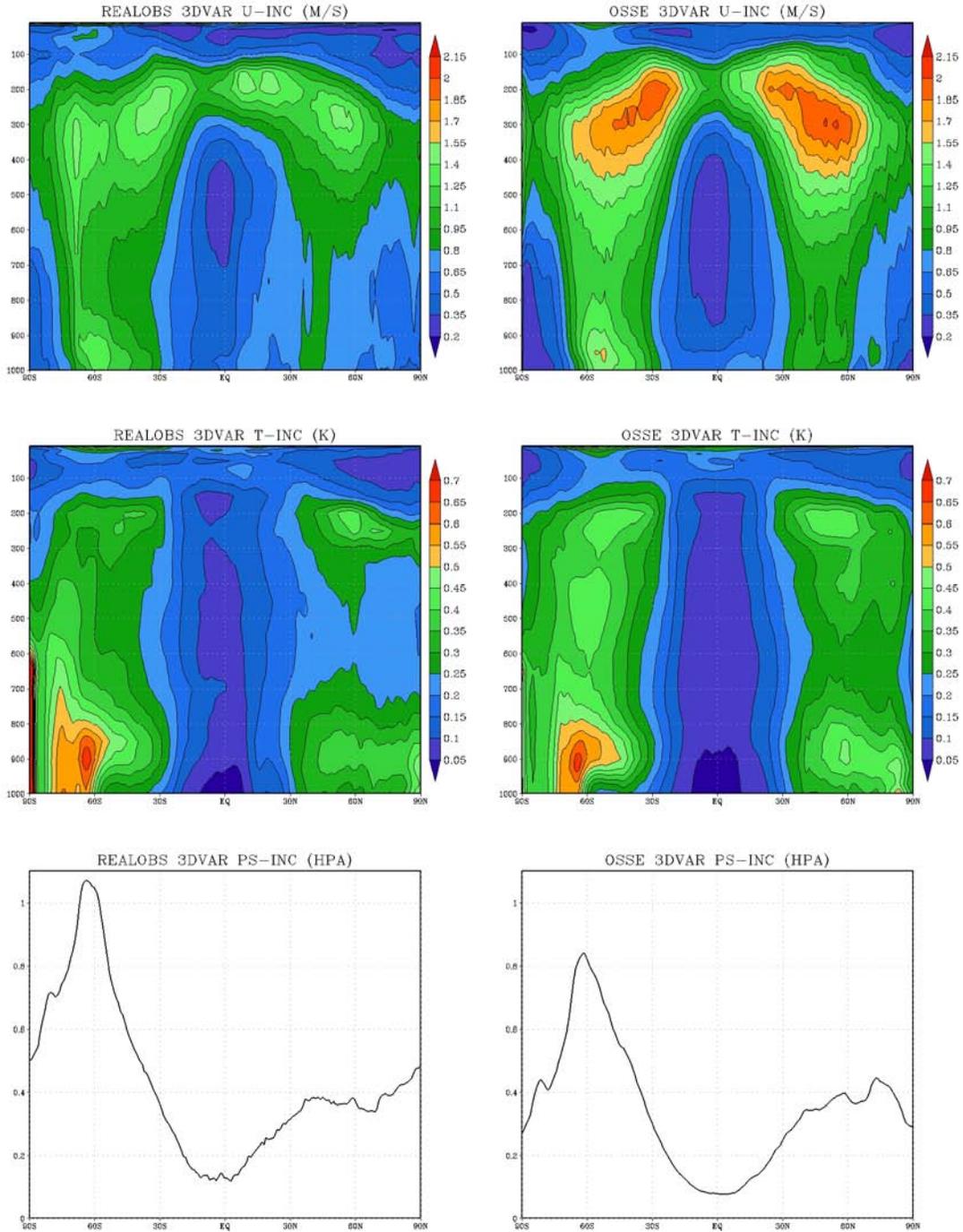


Figure 2.6: Time mean zonally-averaged standard deviation of the 3DVAR analysis increment for a real observation experiment (left) and OSSE experiment (right) for zonal wind (top, m s^{-1}), temperature (middle, K), and surface pressure (bottom, hPa). The real observations case covers the period 00 UTC 01 August 2006 through 18 UTC 30 August 2006, whereas the OSSE experiment covers the simulated period spanning 00 UTC 01 August 2005 through 18 UTC 30 August 2005.

similar between the OSSE and real observation experiments, with larger discrepancies in amplitude associated with the jet level maxima. The real system also exhibits two local maxima in each hemisphere at jet level, one associated with the tropics/subtropics, and a second (lower) maximum associated with the mid-latitude jet. The OSSE-based control, on the other hand, does not exhibit these two distinct features, and instead has a much broader (across latitude) structure.

A comparison of other variables shows that the OSSE and real observations are quite similar in terms of incremental statistics. For example, the temperature increment exhibits very similar structure and amplitude between the two experiments (Fig 2.6 c,d). Both experiments exhibit distinct maxima in the temperature increments poleward of thirty degrees in both hemispheres, with separate local maxima near the surface and near jet level. Other than an odd feature in the lower troposphere right over the South Pole in the real observation experiment, the OSSE compares remarkably well. Similarly, the surface pressure analysis increment is very similar between the two runs (Fig. 2.6 e,f), with the OSSE-based experiment exhibiting smaller amplitude. The statistics described here are similar to the findings in Errico et al. (2007) and Errico et al. (2012) where the largest discrepancies between the OSSE-based and real observation systems seem to be associated with the wind increments.

It is to be expected that there will be some quantitative discrepancies between the real analysis system and the OSSE based experiment (Errico et al. 2007). However, the focus of this work is not the impact of particular observations or the evaluation of real analysis error, and instead is the relative (potential) improvement in

the quality of analysis resulting in algorithmic changes such as the introduction of the hybrid. In this sense, the only expectation is that one can produce something that at least behaves similarly to the real system. Of course, one would hope that the OSSE and real system are close enough so that the findings of the OSSE-based experimentation would translate directly to the real system.

2.6 3D-Hybrid Results

2.6.1 Analysis Comparison

The analysis (and forecast) quality for the 3DVAR control and 3DHYP experiments are compared by computing errors relative to the ECMWF provided nature run. For this purpose, all fields are interpolated to a common grid (one degree regular latitude/longitude grid in the horizontal, pressure surfaces for the vertical). The analysis error structure is quite similar between the 3DVAR and 3DHYP runs (Fig. 2.7), exhibiting zonal wind error local maxima in the upper troposphere associated with jets, large temperature errors in the lower troposphere poleward of sixty degrees south, and lower tropospheric maxima in specific humidity errors. Despite the obvious similarities, several differences immediately stand out. A plot of the difference between the 3DHYP experiment and 3DVAR control reveals that analysis errors are generally reduced in the hybrid system (Fig. 2.8). Notably, there appears to be a significant reduction in the wind errors, particularly in the locations where the largest errors exist and near the model top. The reduction in analysis temperature errors is not as impressive, consistent with the fact the observing system is made up predominantly of mass-type observations (satellite radiances).

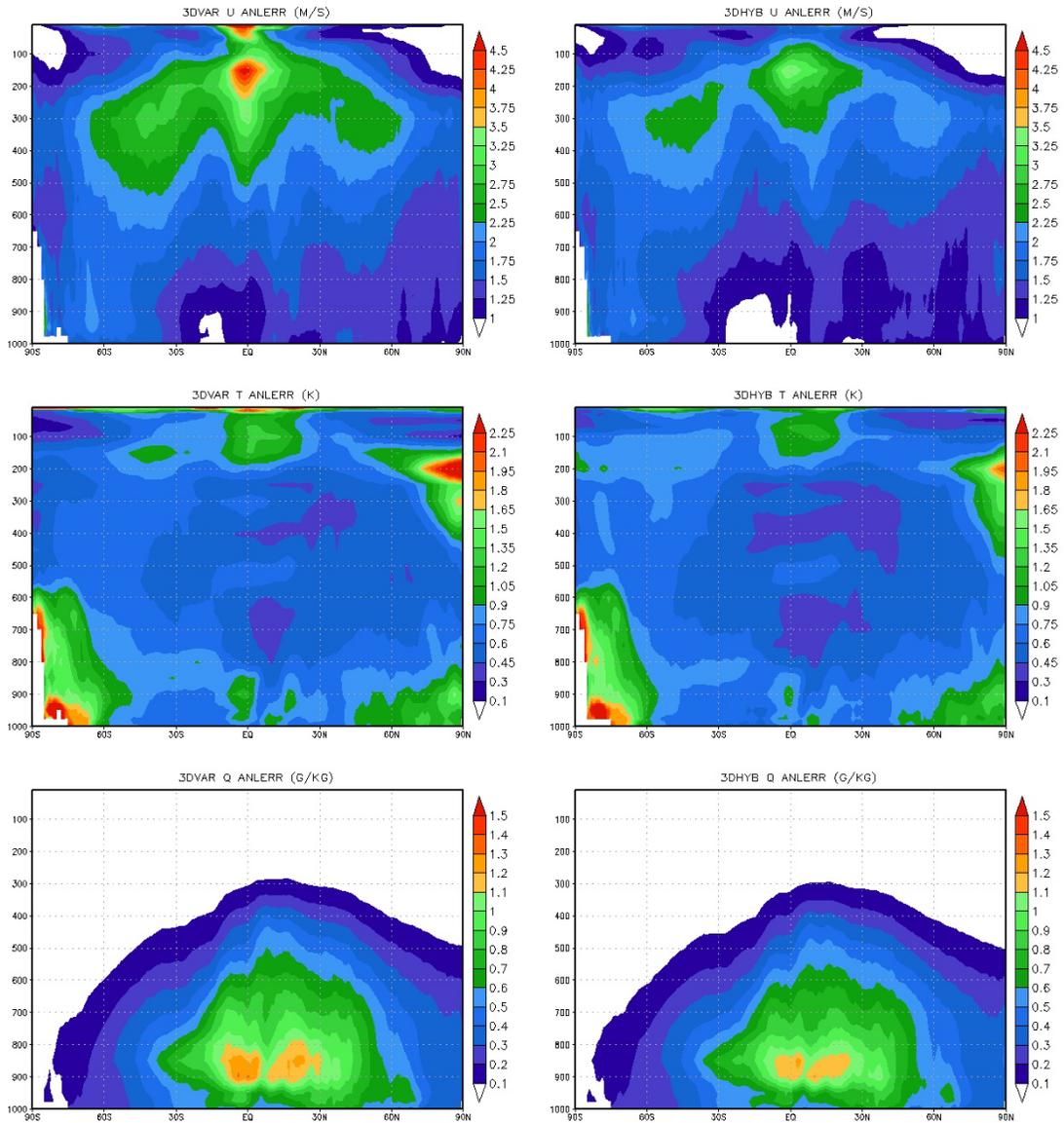


Figure 2.7: Time mean zonally-averaged standard deviation of the analysis error the 3DVAR (left) and 3DHYP experiment (right) for zonal wind (top, m s^{-1}), temperature (middle, K), and specific humidity (bottom, g kg^{-1}), covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.

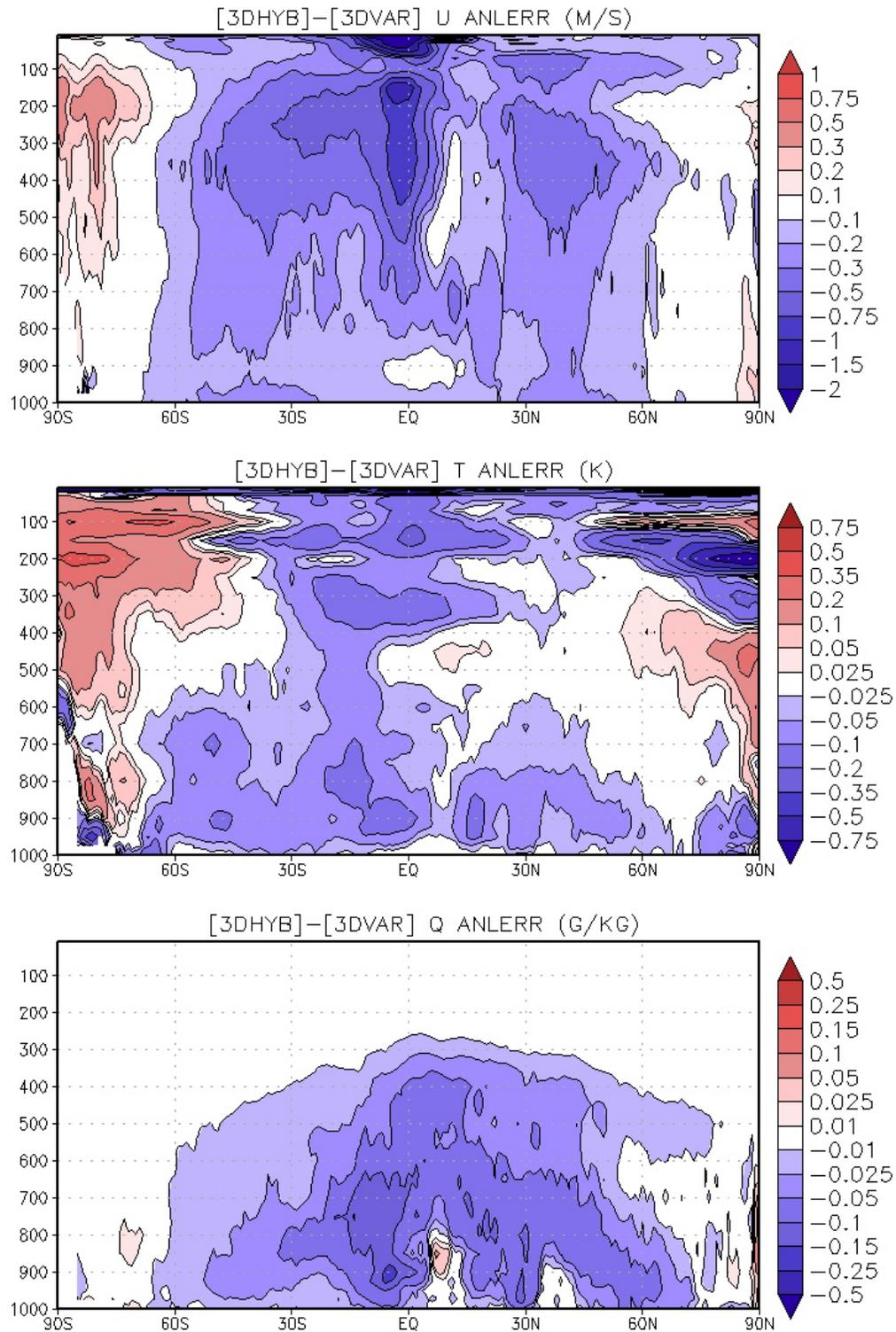


Figure 2.8: Difference in the time mean zonally-averaged standard deviation of the analysis error ($3DHYB-3DVAR$) for zonal wind (top, $m\ s^{-1}$), temperature (middle, K), and specific humidity (bottom, $g\ kg^{-1}$), covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.

The reduction in the specific humidity errors in the 3DHYB run (Fig. 2.8c) is consistent with the wind error reduction, with near uniform error reduction in the regions associated with the largest errors. The large reduction in analysis errors for wind and humidity, but not temperature, is suggestive that the multivariate aspects of the analysis increments in the hybrid play a critical role in improving the quality of analysis. It is noteworthy that the 3DHYB system did not result in uniform error reduction, with a noticeable increase in the analysis errors relative to the 3DVAR control for temperature and wind between 50 hPa and 400 hPa poleward of sixty degrees south (Fig. 2.8 a,b). This is discussed in more detail in Section 2.6.3, as follow-on hybrid experiments were carried out to attempt to investigate the reasons for this degradation.

A comparison of the 300 hPa zonal wind analysis error standard deviation between the 3DHYB and 3DVAR experiments reveals that the biggest impact from the hybrid system can be found in the tropics (Fig. 2.9), where the 3DVAR control appears to have the most issues. In particular, a large local maximum in error over the Indian Ocean region, associated with deep convection, is significantly reduced in the 3DHYB experiment. The difference in analysis errors (Fig. 2.9c) reveals that the 3DHYB is almost uniformly better than the 3DVAR control for the simulated August period, with small scale exceptions located just off the equator north of South America, and subtle hints of degradation poleward of sixty south (as previously revealed in the zonal mean plots). Interestingly, the errors in the 3DVAR control are quite large over the Southern Ocean as expected but quite small over the central and north Pacific. This is likely an artifact of the errors that were added to the simulated

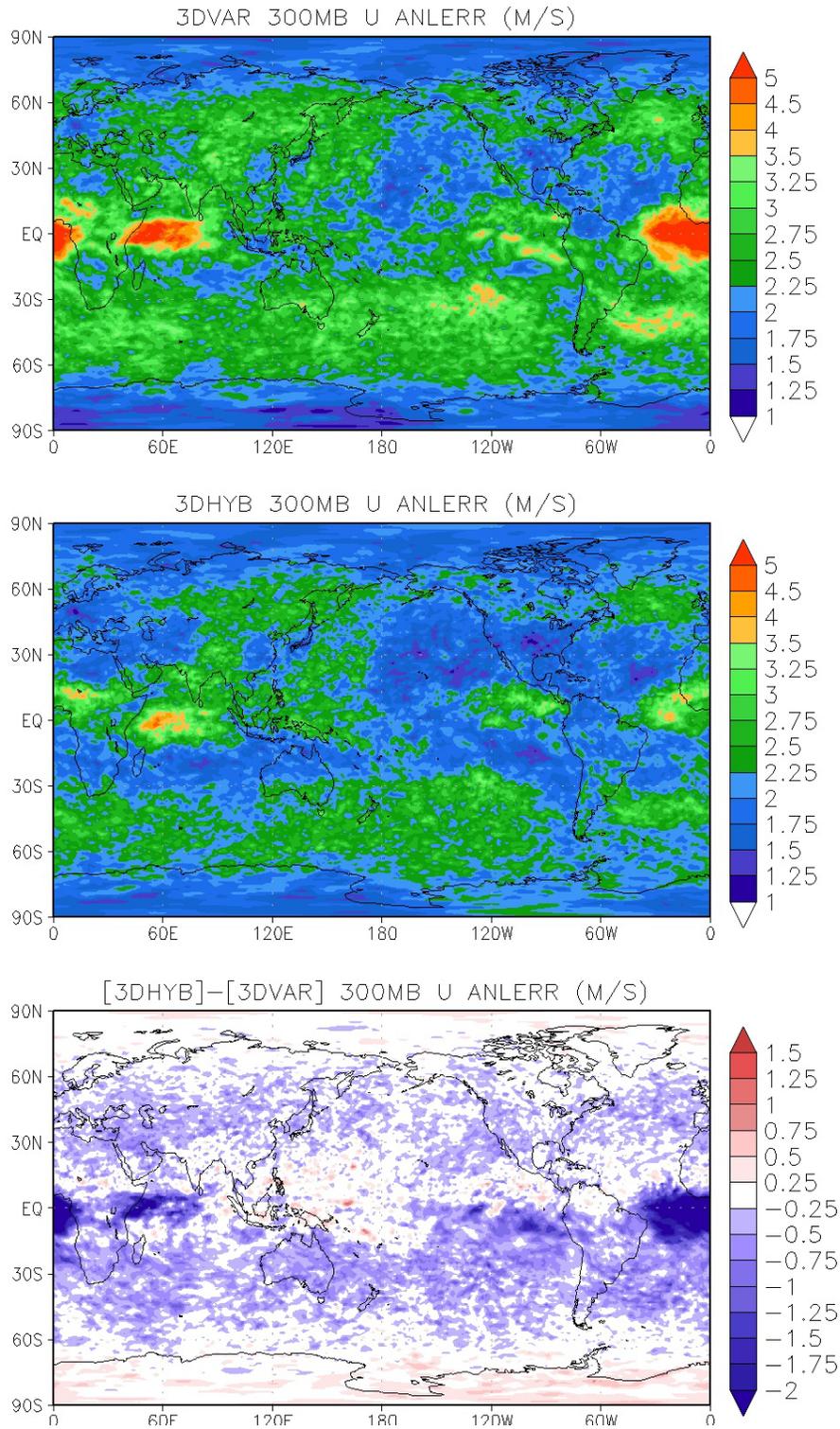


Figure 2.9: Time-averaged standard deviation of the 300 hPa zonal wind analysis error (ms^{-1}) for the 3DVAR (top) and 3DHYP (experiments) for the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005, as well as the difference (3DHYP-3DVAR, bottom).

observations and trying to match the impact of each observing platform relative to realistic OSEs (Privé, personal communication).

Some interesting behavior in both the 3DVAR and 3DHYB systems is revealed by the time series of the analysis and background error standard deviations (as in Fig. 2.10). In general, the analysis error is smaller than the background error for any given cycle, though there are rare exceptions. For mid-tropospheric zonal wind, the background errors for the 3DHYB experiment are significantly smaller than the analysis errors from the 3DVAR control. Interestingly, the magnitude of the difference between the analysis errors and background errors for 500 hPa zonal wind is comparable, despite the hybrid exhibiting significantly smaller errors. For 850 hPa temperature, the behavior is quite different as the difference between the background and analysis error is much more subtle with a clear signal by cycle. The temperature errors are a minimum for the 00 UTC analysis cycle each day in both the 3DVAR and 3DHYB. A much more subtle secondary minimum in analysis error is associated with the 12 UTC cycle. The smaller errors at 00 and 12 UTC can be attributed to the addition of the radiosonde network for those times. The zonal wind errors appear to have a bit of cyclical nature to them as well, but the pattern is much more difficult to ascertain amongst the much larger day to day variability relative to the temperature errors. Consistent with previous discussion, the temperature errors between the 3DVAR and 3DHYB experiments are much more similar to each other and exhibit much less variability, attributable to the fact that the temperature analysis errors are anchored substantially through the assimilation of satellite radiances.

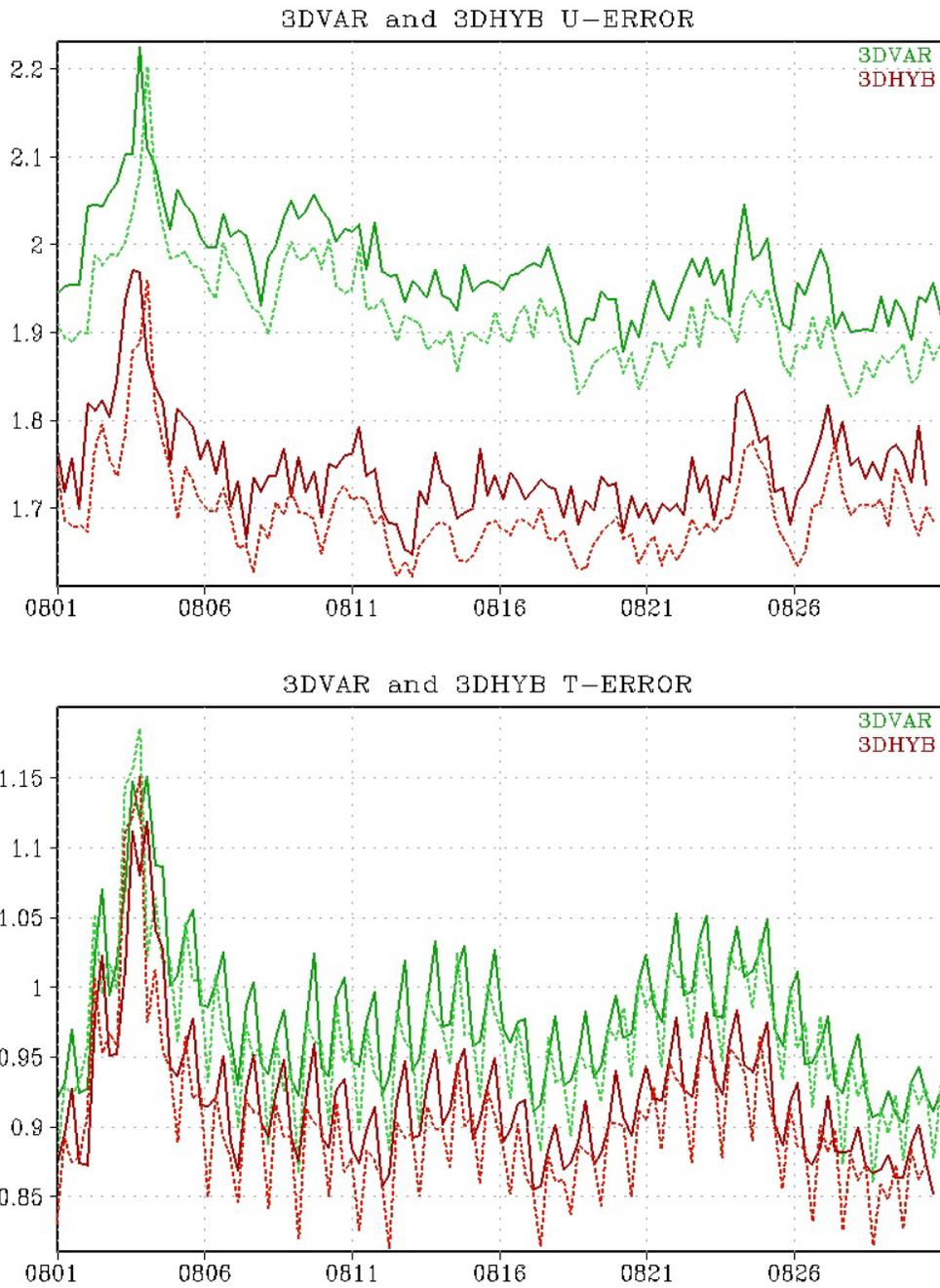


Figure 2.10: Time series of the standard deviation of the 500 hPa zonal wind (top, m s^{-1}) and 850 hPa temperature (bottom, K) background (solid) and analysis (dashed) errors for the 3DVAR (green) and 3DHYB (red) experiments for the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.

2.6.2 Forecast Impact

It is clear that the analysis quality in the 3DHYB experiment is superior relative to that from the 3DVAR control. However, for operational NWP, it is imperative that the improved analyses also translate to improved forecasts. To test the impact of the hybrid in this framework, forecasts are initialized once per day at simulated 00 UTC, and run out to 7.5 days using the same forecast model that was used for data assimilation cycling, starting after the two weeks spin-up. Forecasts from the 3DVAR and 3DHYB runs are then verified by comparing with the ECMWF nature run (instead of doing verification relative to its own analysis as is standard practice in operational NWP).

The forecasts generated by starting from the 3DVAR control analyses exhibit behavior similar to the real system, at least in terms of geopotential height errors. The errors generally increase significantly with time, especially at jet level and near the top of the model (Fig. 2.11). The forecasts in the 3DHYB experiment do not seem to have much of an impact for short lead times. However, the hybrid analysis results in substantial improvement for longer lead times in the Southern Hemisphere troposphere and in the upper stratosphere globally. The skill in the Northern Hemisphere troposphere appears very similar between the 3DVAR and 3DHYB. A die-off curve showing the 500 hPa anomaly correlation (a WMO established standard verification metric for operational centers, see Wilks 2006 for a description) shows a consistent result, with the 3DHYB having a significant impact on the Southern Hemisphere skill. As was the case for the RMSE, the anomaly correlation shows a somewhat positive impact for short lead times and neutral impact at days 5 and 6 in

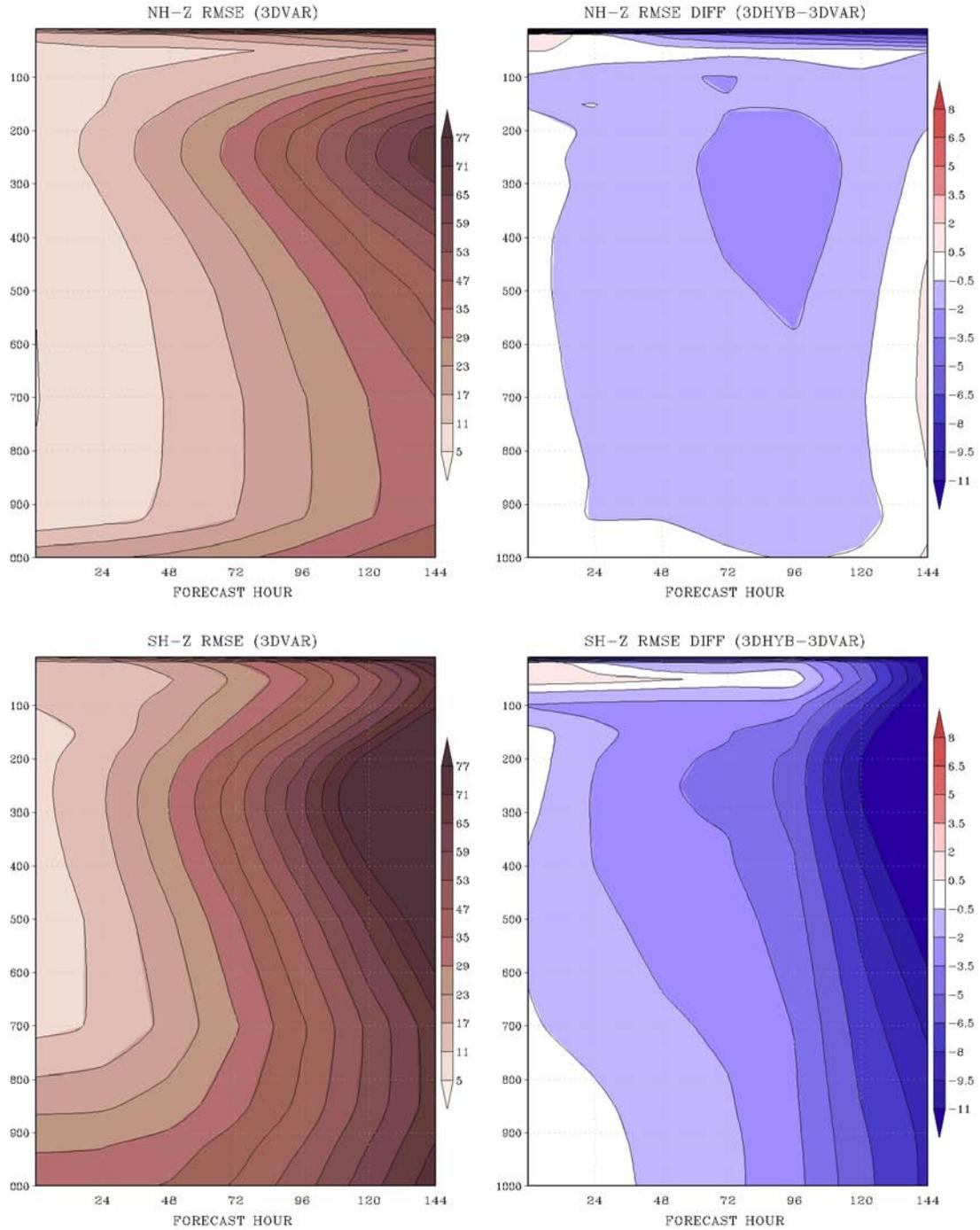


Figure 2.11: Time-averaged root mean square geopotential height errors (m) for forecasts from the 00 UTC analysis in the 3DVAR experiment as a function of lead time for the Northern Hemisphere (upper left) and Southern Hemisphere (lower left) verified against the ECMWF nature run for forecasts verifying between 27 July 2005 and 01 September 2005. The difference between two experiments (3DHYB-3DVAR) for the Northern Hemisphere (upper right) and Southern Hemisphere (lower right) is also plotted.

the Northern Hemisphere (Fig. 2.12). Lastly, just as for the background and analysis errors for zonal wind, the 3DHYP experiment results in significantly improved forecast errors for vector wind in the tropics for all levels and all lead times (Fig. 2.13).

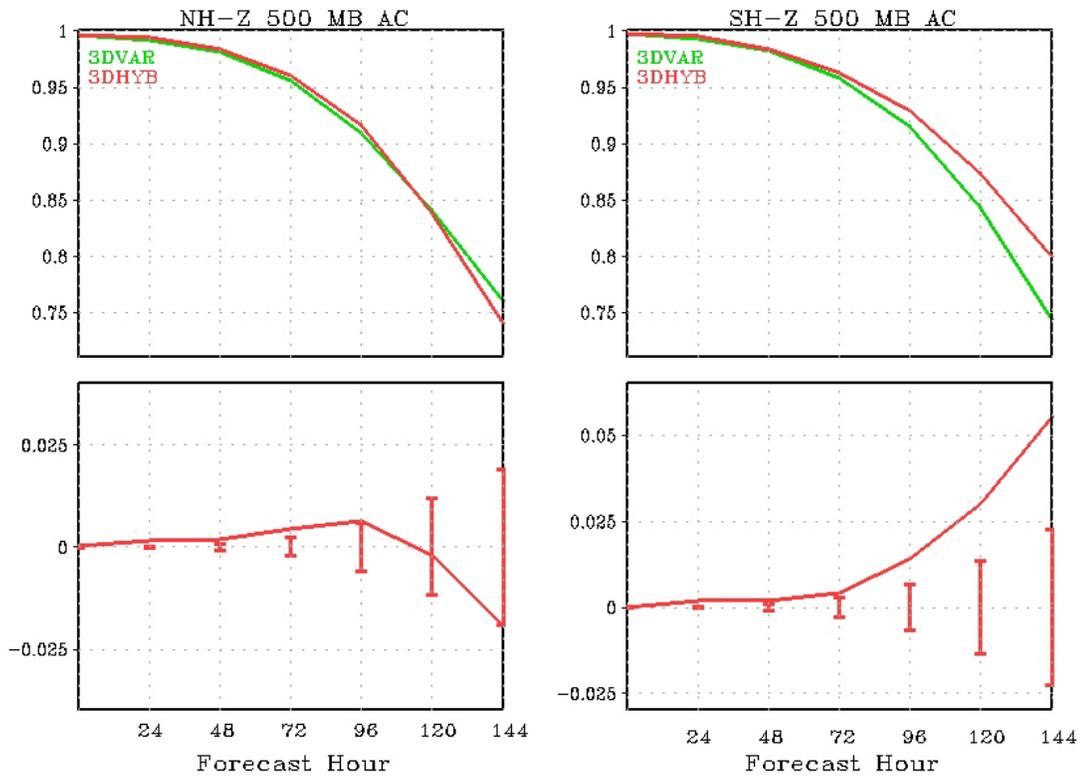


Figure 2.12: Time-averaged 500 hPa anomaly correlation (upper panels) for the Northern Hemisphere (left) and Southern Hemisphere (right) for the 3DVAR (green) and 3DHYP (red) experiments for forecasts from the 00 UTC analyses as a function of lead time as well as the difference (3DHYP-3DVAR, lower panels). The 95% confidence threshold for a significance test (derived from a standard t-test) is also plotted in the lower panels.

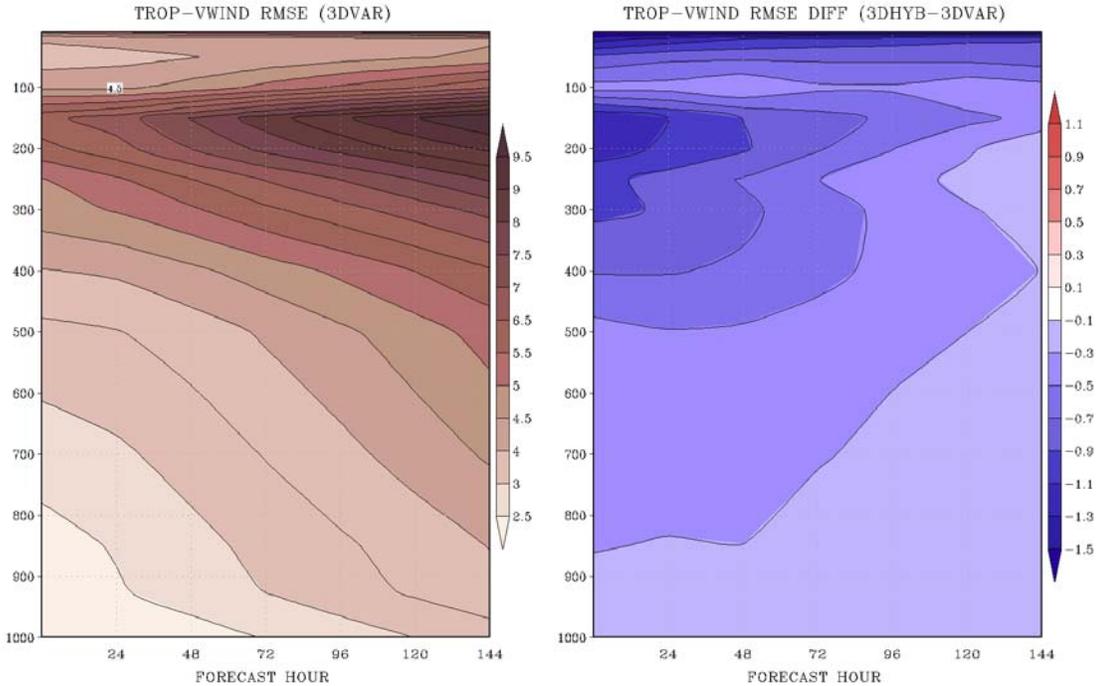


Figure 2.13: As in Fig. 2.11, but for the vector wind RMSE in the tropics.

2.6.3 Follow on 3D-Hybrid Experiments

Although the analysis errors are almost uniformly reduced in the hybrid experiment, it is interesting that there appear to be isolated regions and metrics for which the hybrid system actually results in a degraded analysis. This seems especially true for wind and temperature in a layer between 100 and 400 hPa over the south polar cap (Fig. 2.8). Given that everything else is kept the same, it seems intuitive that the ensemble used within the hybrid framework must have issues in these regions to result in a worse analysis. It is clear by looking at the average spread for zonal wind that the ensemble used within the hybrid experiment is over-dispersive. Compared to the actual 6-hr (background) forecast error, the over-inflated ensemble spread is most egregious in the same region where the hybrid resulted in increased

analysis errors (Fig. 2.14). The average amplitude of the ensemble based background error is in general agreement with the static estimate derived from the NMC-method, except for the large discrepancy between 500 and 150 hPa over the southern latitudes.

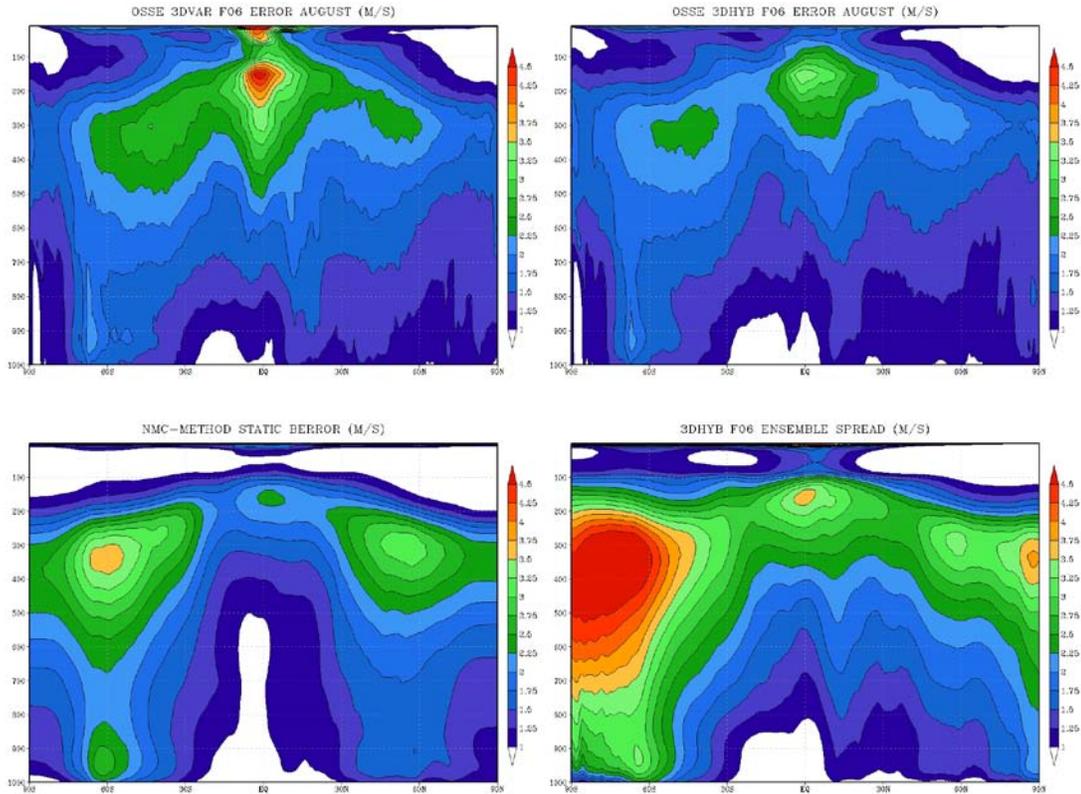


Figure 2.14: Time mean zonally-averaged standard deviation of the zonal wind 06-hr forecast (background) error for the 3DVAR (upper left) and 3DHYB (upper right) experiments, as well as the NMC-method based static error estimate (lower left) and 3DHYB 06-hr ensemble spread (lower right) covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.

Additional 3D hybrid experiments are designed to further investigate these small regions where the hybrid seems to degrade the quality of analysis. The first additional experiment (3DENSV) is similar to the original 3DHYB experiment,

except that it does not utilize a static background error contribution by setting $\beta_e^{-1}=1$ and $\beta_f^{-1}=0$ in equation 2.1 (nor does it use the incremental normal mode constraint, so that the impact of using the ensemble covariances directly can be investigated). This experiment is analogous to running a high resolution analysis of 3D-EnKF using a low resolution forecast ensemble but in a variational framework. The 3DENSV experiment is designed to ascertain whether or not the ensemble is really to blame for the degradation seen over the southern latitude upper troposphere and lower stratosphere. All other settings (including inflation and localization parameters) are kept the same. A comparison of the analysis error differences reveals that, as expected, the same problem areas from the 3DHYB system (Fig. 2.8) show up again in the 3DENSV experiment (Fig. 2.15). Even more interesting is the fact that the errors in the southern latitude upper troposphere are even worse in the 3DENSV experiment. This is consistent with the ensemble spread being much too large in these regions, resulting in the analysis drawing too closely to the (imperfect) observations. It is also noticeable that there is a much larger mix of positive and negative impact in the 3DENSV case, whereas the result was almost uniformly positive in the 3DHYB case. The 3DENSV experiment does significantly reduce the tropical zonal wind errors just as in the 3DHYB case. It is clear that there is a positive contribution from the static error covariance in the hybrid as well as the incremental normal mode initialization, over the 3DENSV case, particularly in regions where the ensemble is over-dispersive.

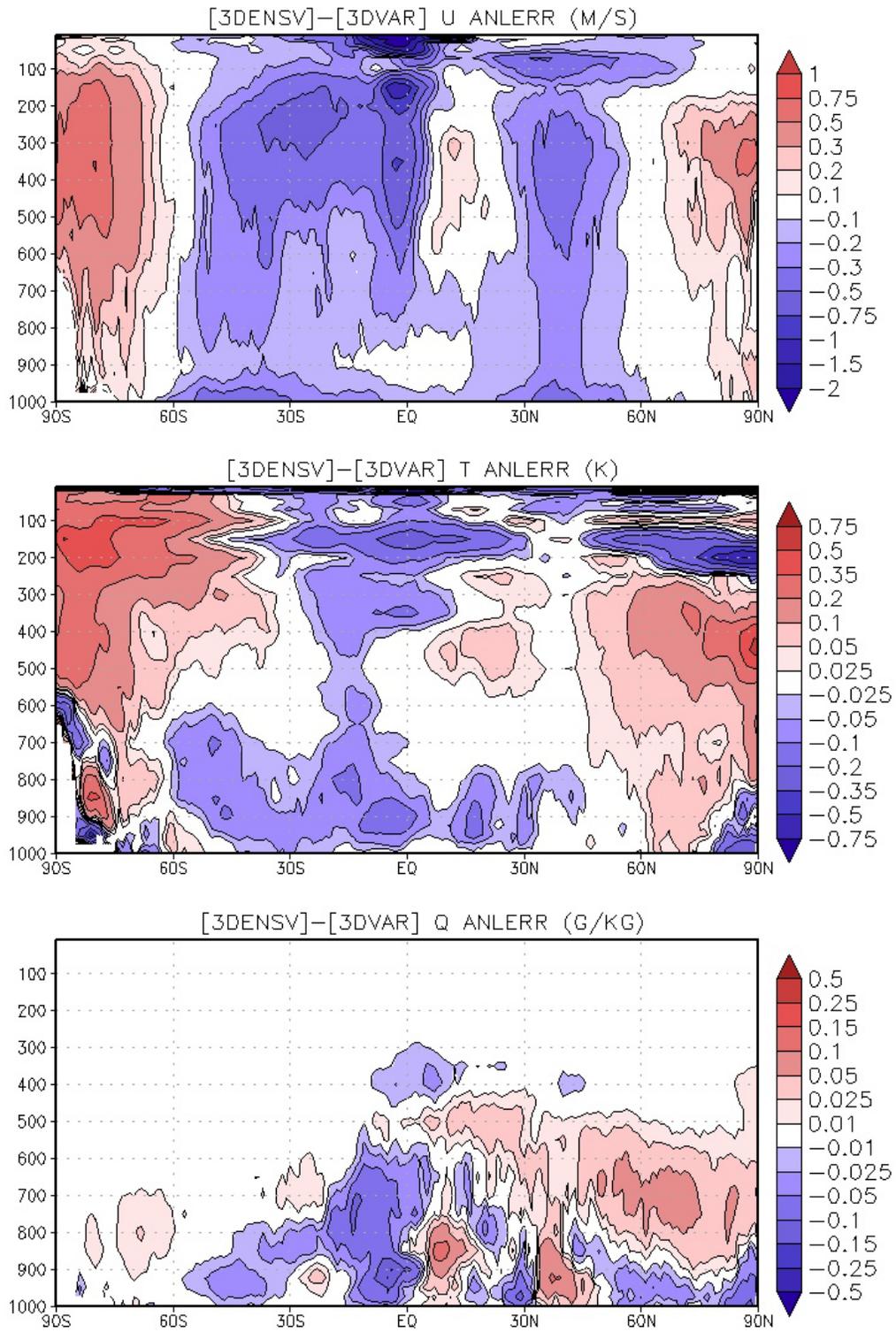


Figure 2.15: As in Fig. 2.8, but for the difference between 3DENSV-3DVAR

As a result of sampling error and to avoid filter divergence, additive and multiplicative inflation are applied to the ensemble members after the EnKF update. The amount of inflation that is applied is a key component to controlling how much spread exists within the background ensemble. The inflation parameters that are used in the 3DHYB and 3DENSV experiments are chosen to match those utilized in a real EnKF (or hybrid) system, such as the one in development and testing for the operational NCEP GDAS. Given that those particular parameter settings result in an ensemble with too much spread (in certain regions and variables) and keeping in mind the results from the 3DHYB and 3DENSV experiments, an additional experiment (3DHYB_RS) is carried out that utilizes smaller inflation parameters (see Section 2.4.2) to reduce the spread in the background ensemble. The multiplicative inflation parameter (ω) is reduced from 0.85 to 0.7 globally and the amplitude factor (κ) for the additive perturbations is reduced by 35% globally. These parameters are tuned in an attempt to better match the ensemble spread (for multiple variables) with the actual background errors of the system in a global, general sense.

In the 3DHYB_RS experiment, everything is identical to the 3DHYB experiment (including localization, as well as weights given to the static and ensemble contributions) except for the inflation parameters. The resultant average standard deviation of the ensemble spread for zonal wind can be seen in Fig. 2.16. In the 3DHYB_RS experiment, the amplitude of the ensemble-based error standard deviation is in better agreement with the actual 6-hr forecast error. In fact, there are regions where the ensemble appears to have spread that is too small, such as in the tropical troposphere as well as much of the stratosphere. The inflation parameters are

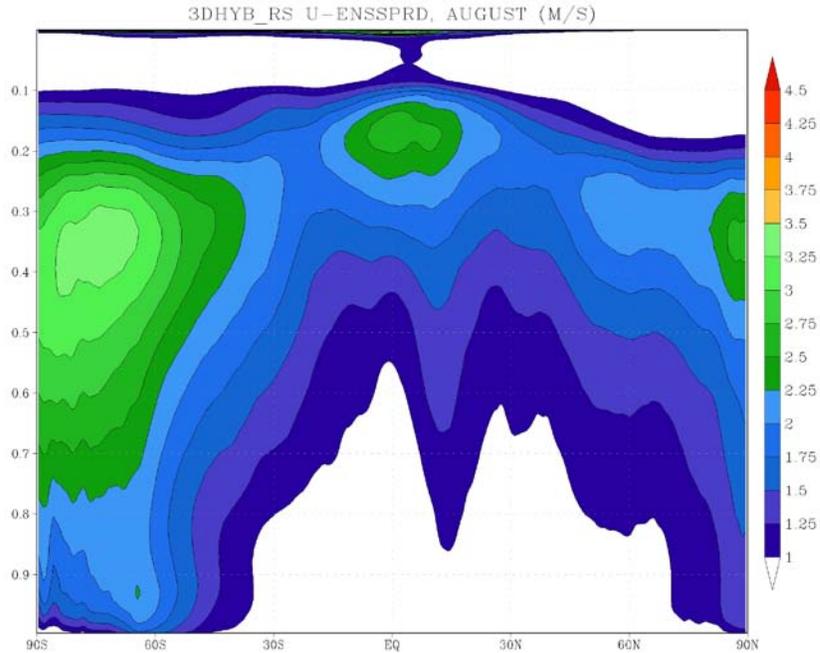


Figure 2.16: As in Figure 2.14 (lower right, from original 3DHYB experiment), but for the experiment 3DHYB_RS (reduced spread).

global numbers making this type of experiment a tuning exercise. Again, the motivation is not to match the ensemble spread and error relative to the nature run as closely as possible, but instead to test the impact of smaller ensemble spread on the analysis error in the hybrid paradigm.

A comparison of the time averaged zonal mean analysis error reveals the analyses in the 3DHYB_RS experiment exhibit smaller errors for most variables and most levels (Fig. 2.17). As designed, the 3DHYB_RS experiment does have a positive impact on the analysis in the southern latitude upper troposphere by reducing the ensemble spread in those regions to more realistic levels. However, this has come at the expense of an increase in the temperature errors for some levels in the upper

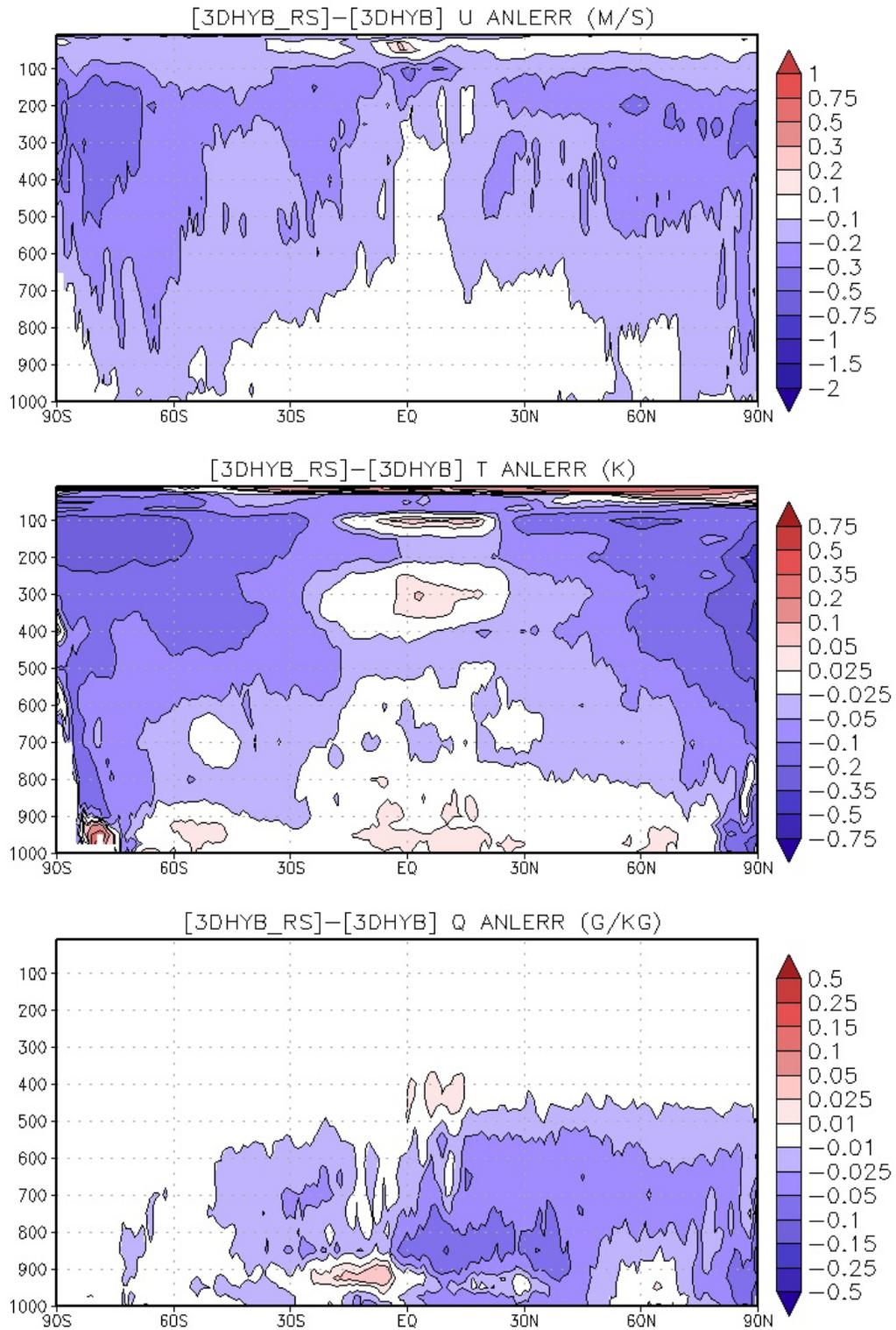


Figure 2.17: As in Fig. 2.8 (which had comparison 3DHYB-3DVAR), but for the difference between 3DHYB_RS-3DHYB.

stratosphere. The reduction in zonal wind errors is of slightly larger amplitude in the 3DHYB_RS experiment than the original 3DHYB experiment and expands over larger areas of the globe. A comparison of the 3DHYB_RS experiment with the original 3DVAR run reveals remarkable analysis error reduction for nearly all variables and all levels (Fig. 2.18). Almost all of the problem areas from the original 3DHYB run (Fig. 2.8) are eliminated, with the exception of some small regions for temperature, again over the South Pole in the upper troposphere (Fig. 2.18). The amplitude of the original error increase for temperatures in these regions (3DHYB over 3DVAR) has been significantly reduced in 3DHYB_RS. Furthermore, the spatial coverage of the reduction in analysis errors is much greater in 3DHYB_RS compared to 3DHYB.

This is an interesting result for real NWP applications, where it is impossible to tune the ensemble spread to match exactly the real forecast error standard deviation. Given that care has to be taken to avoid filter divergence, one can take the cautious approach and ensure that the ensemble spread is large enough. It is encouraging that the hybrid does seem to overcome *some* of the problems that come with using an over-spread ensemble. However, this evidence suggests that an ensemble spread that is closer to the real background error will result in an improved analysis. Perhaps through adaptive inflation methods (Anderson 2009; Miyoshi 2011), or through a hybridization of the background error covariance used in the EnKF as well, some of these issues can be mitigated.

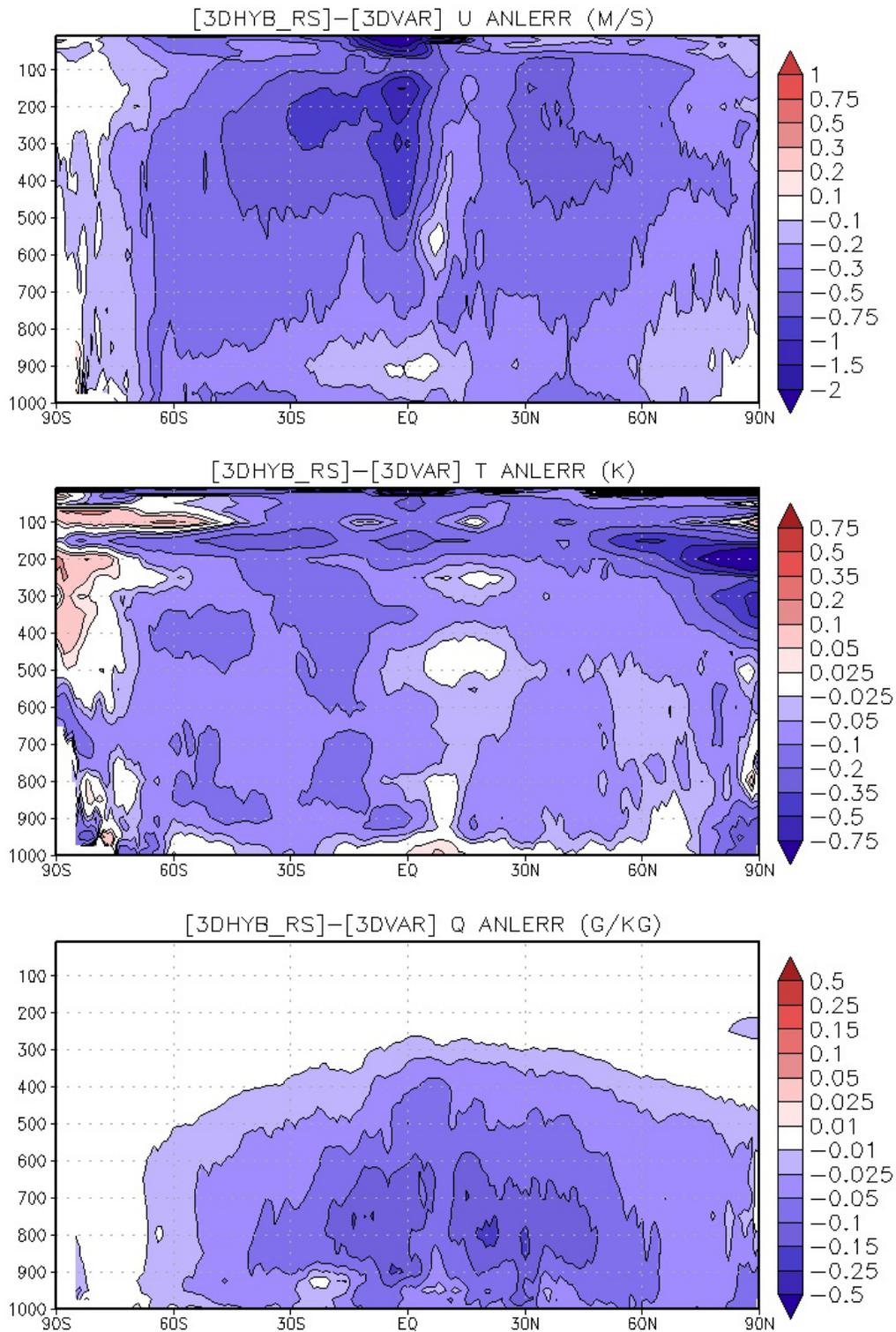


Figure 2.18: As in Fig. 2.8 and Fig. 2.17, but for the difference between 3DHYB_RS-3DVAR.

2.7 Summary and Conclusions

Hybrid data assimilation algorithms that attempt to combine the strengths of variational and ensemble algorithms while attempting to minimize their weaknesses have become popular within the data assimilation community. This is true at NCEP where efforts are underway to develop and test a hybrid system for the operational application. For an operational center, a hybrid algorithm allows for an easier transition toward using ensemble-based, flow-dependent error covariances without a complete paradigm shift by incorporating ensemble perturbations into well-established variational algorithms. There are advantages to maintaining the variational framework beyond the practical aspects, such as supplementing the ensemble estimate with a static error covariance, the application of already established dynamic constraints, and localization in physical space.

Although impact tests with real observations have corroborated that hybrid algorithms can prove superior to their stand-alone variational and ensemble counterparts (Hamill and Snyder 2000; Wang et al. 2007b, 2009; Buehner et al. 2010 a,b), OSSEs provide a platform for which to evaluate the characteristics of the analysis error since the truth is known. With this in mind, a series of experiments are carried out to evaluate the impact of a hybrid data assimilation algorithm relative to 3DVAR control through the assimilation of observations generated from an ECMWF-produced nature run. The errors that are added to the simulated observations were done in such a manner that their impact on the analysis and forecast skill within the OSSE framework matched their counterparts from real observations OSEs. Using the NCEP GFS model, 3DVAR and 3DHYP experiments are carried out over a

simulated two month period in a framework that closely resembles a real world NWP application.

Consistent with previous studies, the quality of analyses and forecasts from the hybrid experiment are generally superior to those derived from the 3DVAR control. In particular, the analysis errors for wind and moisture are substantially reduced, while the impact on the quality of temperature analysis was much smaller. Interestingly, the spatial distribution of analysis errors is qualitatively similar to the background errors (compare Figs. 2.7 and 2.14, for example). Surprisingly, it was found that the analysis error in the control for some metrics is substantially greater than the background error (6-hr forecast) in the hybrid. The improved analyses are found to yield improved forecast skill for most variables and most lead times.

A few problem areas are identified in the hybrid analyses, where the introduction of the ensemble based covariances actually leads to an increase in the analysis error, such as the case for temperature and winds in the southern latitude upper troposphere. A comparison of the time mean ensemble spread with the actual 6-hr forecast error reveals that the ensemble that is utilized in the hybrid experiment has too much spread in the same regions that were identified to be problematic. The EnKF is not drawing as closely to the simulated observations as is seen for real observation experiments, resulting in larger analysis spread (not shown). Of course, this implies that the analysis ensemble requires less inflation. Furthermore, the default set of parameters were tuned from a real observation system at *higher resolution*. System growth rates also play a role in how much one needs to inflate an initial ensemble, as a model configuration with faster perturbation growth rates does

not require as large initial perturbations to achieve a certain level of spread after some time interval compared to a model with slower perturbation growth rates.

Two additional experiments, 3DENSV and 3DHYB_RS are then carried out to further demonstrate that the large spread in the ensemble is in fact the cause of the degradation in those regions. The first experiment turns off the static error contribution as well as the normal mode constraint, to test the impact of using only the ensemble-based covariance within the 3D framework. The same problem areas show up in this experiment (along with other regions), demonstrating the impact of the ensemble spread being too large. Interestingly, the 3DHYB experiment seems to help quite a bit for most of the problem areas in the 3DENSV experiment. Given that the ensemble spread is largely controllable through inflation parameters, a final experiment is carried out to test the impact of an ensemble with more realistic spread (relative to the actual forecast errors) on the hybrid system. It is found that this experiment generally resembles the original hybrid experiment but without the problem areas. For winds, the reduced spread experiment exhibits even smaller errors.

It is encouraging that the forecast impact seen from the hybrid within this framework is somewhat consistent with results that have been achieved for more realistic systems (that assimilate real observations). This system then provides a framework for which to experiment with ways to improve the hybrid system, with the expectation that gains achieved in this framework will translate to the real NWP and DA systems. A natural extension to the hybrid that has already been developed and tested is to utilize 4D ensemble perturbations. As a product of a very fruitful

collaboration with NASA GMAO, the same code that was used for these experiments already has a 4DVAR capability. These two things combined allow for the application of a 4D ensemble-based method utilizing the variational framework, and without the need for a tangent linear and corresponding adjoint model (similar to Liu et al. 2008; Buehner et al. 2010 a,b). This same OSSE framework can be utilized to test the impact of such an extension to the hybrid algorithm and is the subject of Chapter 3.

Chapter 3: OSSE Experiments with 4D-Ensemble-Var and Hybrid Variants

3.1 Introduction

Four-dimensional variational data assimilation (4DVAR) techniques that use tangent linear (Lewis and Derber 1985; Courtier et al. 1994) or linear perturbation models (Rawlins et al. 2007) and their corresponding adjoints have been shown to be powerful natural extensions to the 3DVAR technique. In fact, 4DVAR is the method of choice for initialization of deterministic NWP applications at many operational centers (Rabier et al. 2000; Rosmond and Xu 2006; Gauthier et al. 2007; Rawlins et al. 2007). One attractive feature of 4DVAR is that a dynamic model is used to help impose temporal smoothness and physical constraints. Additionally, 4DVAR allows for the simultaneous assimilation of asynchronous observations throughout a window at their appropriate times by producing a 4D analysis trajectory (Lorenc and Rawlins 2005). This is in contrast with the so-called 3DVAR-FGAT method (Rabier et al. 1998; Lawless 2010), which employs time interpolation to compute innovations at the appropriate time but only solves for a solution at a single time (typically at the center of a window). The major drawbacks to the 4DVAR technique are the computational cost (since the method involves propagating a model forward and backward in an iterative manner) as well as the complications related to developing and maintaining linearized forecast models and their corresponding adjoints.

Much like the 3DVAR technique, 4DVAR methods typically assume a static error covariance (though, valid at the beginning of the assimilation window instead of

the middle). This background error covariance is then implicitly evolved by the linear dynamic (and adjoint) model as part of the variational solver. This procedure does allow for some flow-dependence, though the quality of the initial background error covariance can still play a crucial role for short assimilation windows. Further, the use of a dynamic model to constrain the solution does help improve the multivariate aspects. Much like for the 3D case, the development and application of a hybrid ensemble-4DVAR technique, which utilizes ensemble based covariances for prescribing the background error covariance at the beginning of the assimilation window has been shown to be beneficial (Buehner et al. 2010a,b). The drawbacks of such a method are the same as those for 4DVAR, namely computational cost and the need for linearized and adjoint models.

Along the lines of the 4D ensemble-based techniques (Hunt et al. 2004) such as the 4D-LETKF (Hunt et al. 2007), several methods expanding on the idea introduced by Lorenc (2003) have recently been proposed to utilize 4D ensemble perturbations within a variational framework to help solve for a 4D-analysis increment without the need for a tangent linear and adjoint model (Liu et al. 2008; Tian et al. 2008; Tian et al. 2011; Buehner et al. 2010a). Most of the methods in these previous works rely exclusively on an ensemble based error covariance, without combining information with a climatological estimate as has been done in many 3D-based hybrid studies. This exposes these particular techniques to the same weaknesses that have been documented in regards to the EnKF technique, in particular issues related sampling or lack of temporal smoothness. Formulating the problem in the variational framework allows one to take full advantage of the many

developments that have taken place over the years, such as dynamic constraints (Gauthier and Thépaut 2001; Kleist et al. 2009a) and variational bias correction (Derber and Wu 1998; Dee 2005).

In Chapter 2, it was demonstrated that including ensemble covariances in a variational-based hybrid algorithm yielded improvements in the quality of analyses and subsequent forecasts for the National Centers for Environment Prediction (NCEP) global forecast system (GFS) model in the context of an observation simulation system experiment (OSSE). The experiments that were conducted were all performed using 3DVAR and 3D-Hybrid variants, leaving significant room for improvement. Without access to a tangent linear and adjoint model, a natural extension of the 3D hybrid to include 4D ensemble perturbations (i.e., 4D-ensemble-var, just as proposed in Buehner et al. 2010a) is a logical next step for improving upon the previous work. This chapter focuses on such a 4D extension to the GSI-based hybrid. Section 3.2 describes the implementation of the 4D extension to the hybrid including a time-invariant static error covariance supplement. Section 3.3 then follows with a description of various OSSE-based experiments that demonstrate the impact of utilizing 4D-ensemble-var (and hybrid variants) relative to the 3DHYB experiments that were carried out in Chapter 2. Several experiments are carried out to demonstrate the impact of including a static error covariance in the dual-resolution 4D-ensemble-var paradigm, as well as show the impact of various dynamic constraints. A summary and motivation for future work then follows.

3.2 4D-Ensemble-Var

3.2.1 GSI-Based Hybrid Extension

Traditionally, incremental 4DVAR involves solving for the optimal solution (\mathbf{x}'_0) at the beginning of a time window, obtained by minimizing a cost function

$$J(\mathbf{x}'_0) = \frac{1}{2}(\mathbf{x}'_0)^T \mathbf{B}_f^{-1}(\mathbf{x}'_0) + \frac{1}{2} \sum_{k=1}^K (\mathbf{H}_k \mathbf{M}_k \mathbf{x}'_0 - \mathbf{y}'_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{M}_k \mathbf{x}'_0 - \mathbf{y}'_k). \quad (3.1)$$

For notation please refer to Chapter 2, but now a time-level index (k) is included for the initial condition valid at $t=0$ and asynchronous observations (for each k) for K -time levels. The adjoint of the linear model (the transpose of \mathbf{M}) is necessary in order to obtain the gradient in this formulation. This cost function can then be easily extended to a hybrid 4DVAR cost function by including a control variable ($\boldsymbol{\alpha}$) for the ensemble contribution to the analysis increment at the beginning of the window:

$$J(\mathbf{x}'_f, \boldsymbol{\alpha}) = \beta_f \frac{1}{2}(\mathbf{x}'_f)^T \mathbf{B}_f^{-1}(\mathbf{x}'_f) + \beta_e \frac{1}{2} \sum_{n=1}^N (\boldsymbol{\alpha}^n)^T \mathbf{L}^{-1}(\boldsymbol{\alpha}^n) + \frac{1}{2} \sum_{k=1}^K (\mathbf{H}_k \mathbf{M}_k \mathbf{x}'_0 - \mathbf{y}'_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{M}_k \mathbf{x}'_0 - \mathbf{y}'_k) \quad (3.2)$$

where, just as in the 3D hybrid case, \mathbf{x}'_0 becomes a linear combination of that which is derived from a static error covariance (\mathbf{x}'_f) and that which is derived from the ensemble perturbations (\mathbf{x}_e^n) as prescribed by the control variable ($\boldsymbol{\alpha}$) for an ensemble of size N :

$$\mathbf{x}'_0 = \mathbf{x}'_f + \sum_{n=1}^N (\boldsymbol{\alpha}^n \circ \mathbf{x}_e^n). \quad (3.3)$$

This hybrid configuration will hereafter be referred to as H-4DVAR_AD (with AD to denote the use of the adjoint [and forward linear] model within the minimization).

As pointed out by Buehner (2010a), this can be further manipulated to solve a similar 4D cost function without the need for the linear model and its adjoint by utilizing 4D ensemble perturbations (hereafter referred to as 4DENSV). Solving instead for the control variable in this so-called 4D-ensemble-var case, the cost function then becomes:

$$J(\boldsymbol{\alpha}) = \frac{1}{2} \sum_{n=1}^N (\boldsymbol{\alpha}^n)^T \mathbf{L}^{-1} (\boldsymbol{\alpha}^n) + \frac{1}{2} \sum_{k=1}^K (\mathbf{H}_k \mathbf{x}'_k - \mathbf{y}'_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{x}'_k - \mathbf{y}'_k) \quad (3.4)$$

where the linear dynamic model is replaced with 4D, nonlinear ensemble perturbations, so the increment valid at each of the observation time levels can be prescribed as

$$\mathbf{x}'_k = \sum_{n=1}^N \left(\boldsymbol{\alpha}^n \circ (\mathbf{x}_e)_k^n \right). \quad (3.5)$$

As in the 3DHYB case, \mathbf{L} denotes the error covariance for the ensemble control variable, specified to be unit amplitude and Gaussian in structure, thereby acting to enforce localization of the ensemble based error covariance. In this particular formulation, $\boldsymbol{\alpha}$ is assumed to be the same throughout the assimilation window, analogous to the weights in a 4D-LETKF without temporal localization.

The components necessary to solve the GSI using the formulation that is described by equations 3.4 and 3.5, 4DENSV, are already in place. The ensemble control variable for the hybrid is developed and implemented as part of the 3D-hybrid work described in Chapter 2. A 4DVAR capability within the GSI was previously developed through collaboration with colleagues at NASA GMAO (Todling and Trémolet, personal communication). A key component of the extension of 3DVAR to 4DVAR within the GSI was the addition of new observation handling features and

the inclusion of time binning. With these pieces in place, the 3D hybrid capability is extended to allow for 4D ensemble perturbations and a 4DENSV option.

3.2.2 4D-Ensemble-Var Hybrid

As was demonstrated in the 3D-hybrid context, the inclusion of a static (climatological) estimate for the background error covariance within the hybrid can prove beneficial. This seems especially true for dual-resolution applications or for those problems that require the utilization of a (relatively) small ensemble. As shown equation 3.2, this can be achieved in the hybrid H-4DVAR_AD context, where the ensemble is used only to supplement the background error covariance at the initial time. However, such a paradigm is computationally expensive since it has all the cost of a 4DVAR formulation (plus some), in addition to having to maintain and update the ensemble. A cost effective alternative does exist, if one is willing to allow for a time-invariant (3DVAR-like) contribution to the solution. By combining equations 3.2-3.5,

$$\begin{aligned}
 J(\mathbf{x}'_f, \boldsymbol{\alpha}) = & \beta_f \frac{1}{2} (\mathbf{x}'_f)^T \mathbf{B}_f^{-1} (\mathbf{x}'_f) + \beta_e \frac{1}{2} \sum_{n=1}^N (\boldsymbol{\alpha}^n)^T \mathbf{L}^{-1} (\boldsymbol{\alpha}^n) + \\
 & \frac{1}{2} \sum_{k=1}^K (\mathbf{H}_k \mathbf{x}'_k - \mathbf{y}'_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{x}'_k - \mathbf{y}'_k)
 \end{aligned} \tag{3.6}$$

to formulate a cost function that is a blend of a 4DENSV and 3DVAR-FGAT cost function. In this context, \mathbf{x}'_f is the analysis increment that is prescribed from a static, time-invariant error covariance estimate \mathbf{B}_f (i.e. the same \mathbf{B} that is used in 3DVAR or 4DVAR at the beginning of the window). The contribution of this part of the solution to the analysis increment at any given observation time is constant:

$$\mathbf{x}'_k = \mathbf{x}'_f + \sum_{n=1}^N \left(\boldsymbol{\alpha}^n \circ (\mathbf{x}_e)_k^n \right). \quad (3.7)$$

The ensemble contribution to the analysis increment at any given observation time is exactly like that for the 4DENSV case. In this instance, however, the (4D) ensemble and (3D) static contributions are weighted similarly to the 3D-hybrid case, with β_e and β_f respectively. By setting the extended control variable for the ensemble, $\boldsymbol{\alpha}^n=0$, this method is equivalent to 3DVAR-FGAT. The capability to run the 4DENSV option in a truly hybrid mode, by including a time-invariant static contribution to the increment as shown in 3.6 is implemented for use within the GSI-based hybrid (referred to as H-4DENSV from here on out, as not to be confused with the H-4DVAR_AD formulation). Given a configuration that utilizes a dual-resolution option, the static contribution is aimed at accomplishing two things: 1) filling the null space at higher frequencies that the low resolution ensemble cannot resolve and 2) anchoring the solution with a static error covariance (which can itself help further ameliorate potential sampling and localization issues). An H-4DVAR_AD option is also implemented into the code though is not the main focus of this study.

3.2.3 Single Observation Example

To demonstrate the impact of the 4DENSV (and H-4DENSV variant) algorithm, an experiment that assimilates only a single observation is performed. The observation chosen for this is a mid-tropospheric temperature observation taken at the beginning of a 6-hr assimilation window (three hours prior to the “update” analysis time), and is assigned a negative two degree departure from the background at that time (a 3-hr forecast) and one degree observation error. The observation is chosen to

reside upstream from the base of a shortwave trough, similar to the single observation experiments shown in Buehner et al. (2010a) for comparison. Several 4D experiments are then carried out that assimilate this single observation. A summary and description of the experiments is available in Table 3.1. The two hybrid cases that include a static and ensemble contribution to the increment are done so with $\beta_f^{-1}=0.25$ and $\beta_e^{-1}=0.75$. For the two 4DVAR variants, the tangent linear and adjoint model that are utilized in the inner loop are adaptations of a fairly simple model that has been under development at NCEP for the sole purpose of running 4DVAR. Although not yet ready for large problems or cycling due to its current computation inefficiencies, it is still quite useful for evaluating (low resolution) single observation experiments that require only a few iterations to minimize.

<i>Experiment</i>	<i>Description</i>	<i>Relevant Equations</i>
4DVAR	4DVAR	Eq. 3.1
H-4DVAR_AD	Hybrid ensemble-4DVAR (with adjoint) $\beta_f^{-1}=0.25, \beta_e^{-1}=0.75$	Eq. 3.2 Eq. 3.3
4DENS	4D-Ensemble-Var (no static B) $\beta_f^{-1}=0.0, \beta_e^{-1}=1.0$	Eq. 3.4 Eq. 3.5
H-4DENS	Hybrid 4D-Ensemble-Var $\beta_f^{-1}=0.25, \beta_e^{-1}=0.75$	Eq. 3.6 Eq. 3.7

Table 3.1: Description of various 4D (and hybrid) single observation experiments and relevant equations for Figs. 3.1 and 3.2.

The resultant analysis increment at the middle of the assimilation window (three hours after the observation is taken) for the various 4D configurations is shown in Fig. 3.1. All four experiments show the maximum increment downstream from where the actual observation was taken consistent with the northwesterly background flow. The 4DVAR experiment results in a spatially broad, quasi-Gaussian temperature increment (Fig. 3.1a). This is not terribly surprising given that only three hours elapses between the time that the observation was taken (beginning of the window) and the center of the window. The three experiments that utilized ensemble covariances exhibit a temperature increment that is stretched along the height gradient as would be expected (Fig. 3.1b-d). All of the experiments show a cyclonic wind response to the cold temperature observation and increment, with the ensemble/hybrid based experiments showing a stronger wind response than the 4DVAR case. It is clear that the 4DENSV case suffers from sampling (spurious correlations) more so than the two hybrid variants. This is not surprising either, in this case, given that a 40 member ensemble is utilized. However, adding a time-invariant, static contribution to the increment helps ameliorate these apparent issues substantially, without hurting the 4D nature of the increment. In fact, the H-4DENSV increment is qualitatively and quantitatively very similar to that from the H-4DVAR_AD experiment, despite the fact that one uses a fairly simple dynamic model while the other utilizes 4D, nonlinearly evolved ensemble perturbations.

The same single observation test can be utilized to further investigate how the algorithms handle the propagation of information through the assimilation window, by visualizing the analysis increment at various k -time levels. Recall that for the

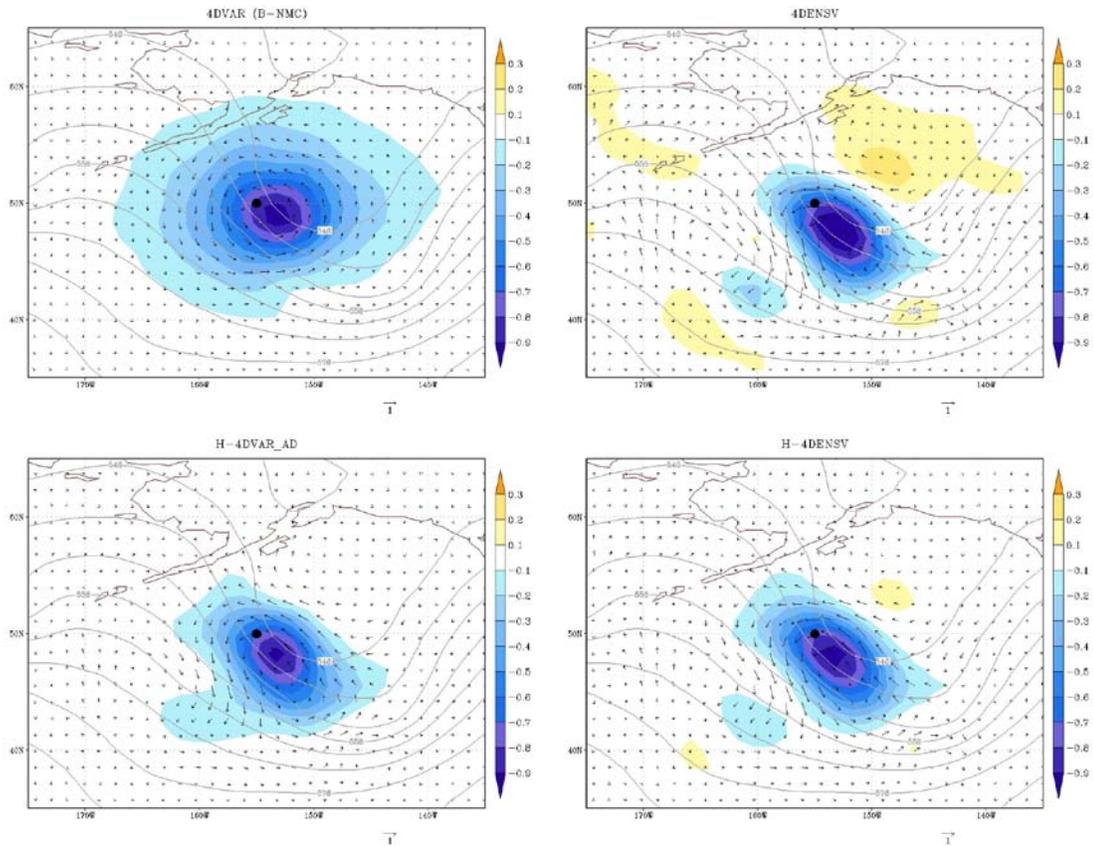


Figure 3.1: 500 hPa Temperature (shaded, K) and vector wind (vectors, m s^{-1}) analysis increment resulting from the assimilation of a single 500 hPa temperature observation (-2K innovation, 1K error) taken -3h from analysis time (location of observation denoted with black dot) for 4DVAR (upper left), H-4DVAR_AD (lower left), 4DENS (upper right), and H-4DENS (lower right). The reference arrow is representative of 1 m s^{-1} .

4DVAR cases, this involves an explicit propagation of the increment through the use of a (linear) dynamic model. For the 4DENS variants, the propagation of information is achieved implicitly through correlations contained within the 4D ensemble perturbations. The time evolution of the temperature and wind increments at 500 hPa in 3-hr intervals for the H-4DVAR_AD and H-4DENS cases is shown in Fig. 3.2. First note that the solution between the two is virtually identical at the beginning of the window (the time at which the observation was taken). This is

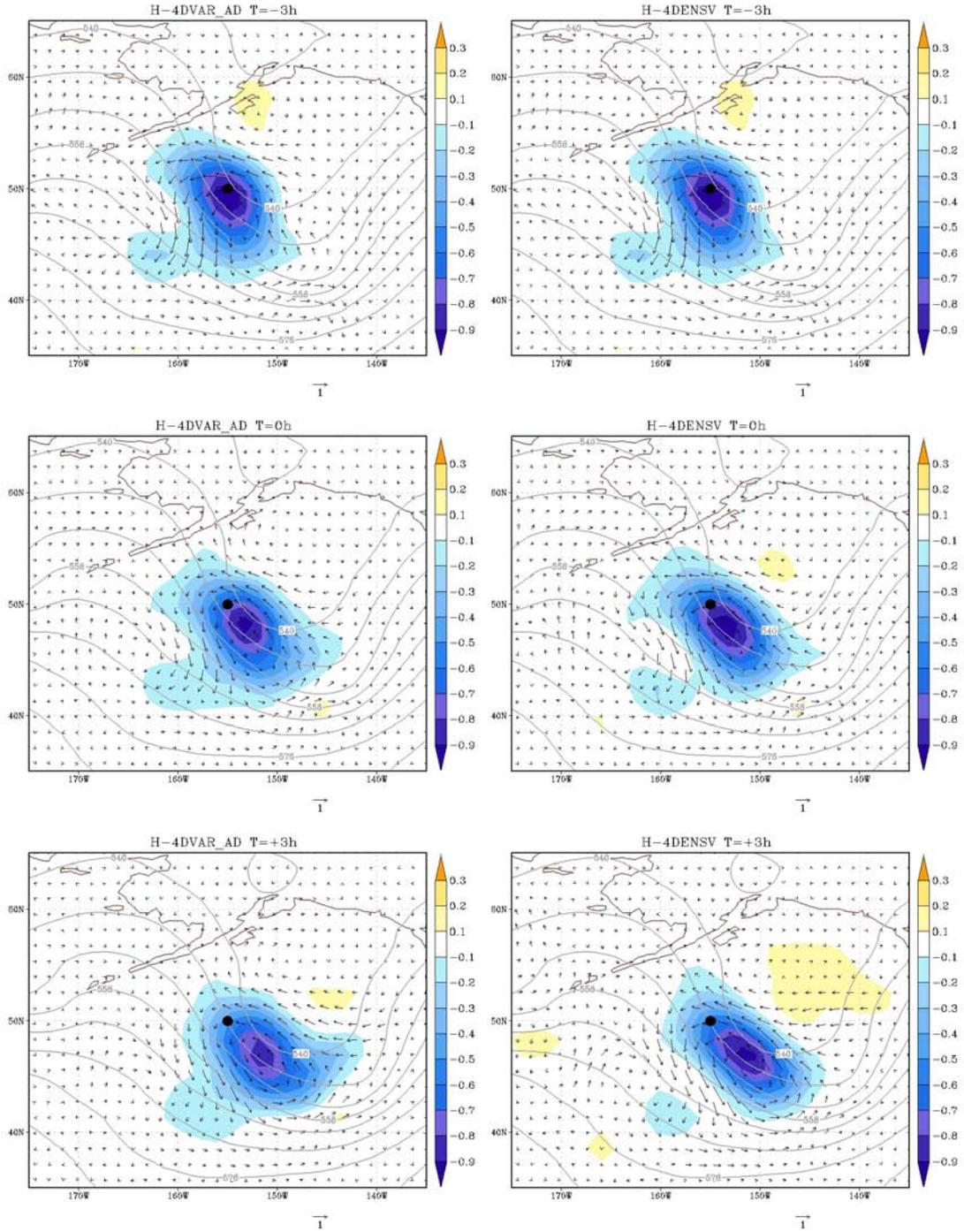


Figure 3.2: As in Fig. 3.1, except for the analysis increment valid at the beginning ($t-3h$, top), middle ($t=0h$, middle), and end ($t+3h$, bottom) of the assimilation window for the H-4DVAR_AD (left) and H-4DENSIV (right) cases.

expected given that the same ensemble perturbations, relative weights, and static error covariance (since it has not yet been evolved by the dynamic model in the 4DVAR case) are identical between the two experiments. This is not the case for observations taken at other times in the assimilation window (not shown), where for the H-4DVAR_AD case, the effective “static” error covariance is evolved forward in time consistent with the linear dynamic model. For the other times in the six hour assimilation window (middle and end) the temperature and wind increment between the two hybrid cases is strikingly similar both qualitatively and quantitatively. By the end of the window, the H-4DENSV case develops more small amplitude features away from the location of the original observation and has a slightly tighter structure aligned with the height contours. The H-4DVAR_AD case, on the other hand, has slightly broader features at the end of the window. It is quite encouraging that the (computationally cheaper) ensemble method can mimic quite nicely the evolution of the increment within the assimilation window. Although not shown, other single observation experiments for other variables and levels reveal similar results.

3.3 Constraint Formulations

3.3.1 Incremental Normal Mode Constraint

The 4DENSV (and hybrid) option is implemented in such a way as to allow for the application of many of the standard features included in the 3D GSI such as variational quality control, variational satellite bias correction, the tangent linear normal mode constraint, and various weak constraints. However, the 4DENSV solution to the analysis problem requires special attention when it comes to the

application of such constraints given that the 4D aspect of the problem is obtained through a dot product of the weights and the ensemble perturbations in a discrete manner without the explicit use of a dynamic model. This introduces the possibility that it may be necessary to apply such constraints over all k -time levels to ensure consistency. For a dynamic constraint such as the tangent linear normal mode constraint, this can be prohibitively expensive if one were to use hourly time levels and a six hour assimilation window, as an example. Furthermore, the original implementation of the normal mode constraint for application within the 3D-Hybrid requires the filter to be applied to the total analysis increment (sum of static and ensemble contributions) without much flexibility (as in equation 2.1 from Chapter 2). For these reasons, several options related to the application of various constraints are explored and implemented into the analysis code.

Building on previous successes in applying the tangent linear normal mode constraint in 3DVAR (Kleist et al. 2009a) and 3DHYB (Chapter 2) applications, various possibilities for application of such a constraint in the 4DENS (and hybrid) context are developed and tested. First, the possibility to apply the constraint to only the static contribution to the increment is implemented:

$$\mathbf{x}'_k = \mathbf{C} \mathbf{x}'_f + \sum_{n=1}^N \left(\boldsymbol{\alpha}^n \circ (\mathbf{x}_e)_k^n \right). \quad (3.8)$$

This requires a single constraint operation per iteration, making the cost equivalent to 3D applications. Note that this capability can also be applied within the 3DHYB context, though previous results suggest the inclusion of the constraint to the total increment aids in ameliorating potential imbalances introduced by the ensemble based

increment. This formulation is motivated by the desire to allow for the use of the unfiltered ensemble covariance, while still filtering the static contribution to the increment.

In order to help mitigate imbalances that can be introduced through sampling and localization, the ability to apply the tangent linear normal mode constraint to the total increment is explored. To apply the incremental normal mode constraint to the total increment in the H-4DENSV (and 4DENSV) context, two possibilities are considered. The first is to simply apply the constraint over all k -time levels:

$$\mathbf{x}'_k = \mathbf{C}_k \left[\mathbf{x}'_f + \sum_{n=1}^N \left(\boldsymbol{\alpha}^n \circ (\mathbf{x}_e)_k^n \right) \right]. \quad (3.9)$$

Here, the balance operator \mathbf{C}_k is also denoted with a time level index, since the possibility exists for linearization about the background for each time level (since the constraint is in fact tangent-linear). However, for practical reasons (computational cost and memory), the linearization state is assumed to be the center of the window. Even still, the application of the constraint in this manner can be prohibitively expensive. For a six-hour window with hourly time levels (and hourly ensemble perturbations) the computational cost of the analysis goes up substantially. For this reason, the possibility to apply the full constraint only to the solution in the middle of the window is considered. The advantage to such a method is the increment that is applied to the background and used to restart the model can be filtered explicitly reducing spin-up and spin-down, without the considerable cost of having to filter all time levels. This of course introduces inconsistencies between the incremental

solution in the center of the window and the other time-levels, which could have undesirable consequences.

3.3.2 Weak Constraint Digital Filter

Digital filter weak constraints have been shown to be beneficial in many 4DVAR applications (Polavarapu et al. 2000; Gauthier and Thépaut 2001, Wee and Kuo 2004). A similar constraint has previously been implemented by colleagues at NASA GMAO for the 4DVAR extension of the GSI.³ The formulation of such a constraint involves the addition of a new penalty term (see Gustafson 1993 or Polavarapu et al. 2000 for a more detailed derivation):

$$J_{dfi} = \chi \left\langle \mathbf{x}_m - \mathbf{x}_m^i, \mathbf{x}_m - \mathbf{x}_m^i \right\rangle \quad (3.10)$$

where \mathbf{x}_m^i denotes a filtered or “initialized” state at time m (in the case of 4DENSF, assumed to be the center of the assimilation window), and χ a general weighting term (typically denoted α in the literature but avoided in this context to remove confusion with the ensemble control variable). Although previous derivations involve the tangent linear and adjoint models required to run 4DVAR, such a constraint can be utilized in the 4DENSF context based on the 4D increment that is prescribed from the ensemble perturbations. The filtered state is constructed from the 4D increment using the same filter coefficients (h) as in the typical formulation (Lynch and Huang 1992)

$$\mathbf{x}_m^i = \sum_{k=1}^K h_{k-m} \mathbf{x}_k^u \quad (3.11)$$

³ The JcDFI that was originally developed for 4DVAR applications was not implemented for the standard double conjugate gradient minimization algorithm that is utilized for the hybrid applications. Minor modifications were necessary to adapt the code for use in this context.

where \mathbf{x}_k^u denotes the unfiltered increment at each time level- k . The norm for the penalty function in this implementation is chosen to be a dry energy norm (though the capability does exist to use a moist energy norm). Such a constraint in the 4DENSV context potentially allows for noise control within the 4D increment without adding much computational cost. Additionally, it is now possible to explore the use of a combination of noise constraints (weak constraint digital filter and incremental normal mode) to improve the quality of analysis.

3.3.3 Impact on Single Analysis

To ensure that the constraints are all properly functioning for use with the new 4DENSV paradigm within GSI, a single analysis test case that assimilates real observations (including satellite radiances) within a 6-hr window valid at 06 UTC 15 July 2010 is performed. The background and ensemble files are generated from an offline experiment that utilizes the dual-resolution hybrid configuration, but at higher resolution than the OSSE-based cycled experiments described in Chapter 2. For practical reasons, this analysis test case is also run at the higher resolution (from the T574 GFS model, with an 80 member, T254 ensemble with hourly output). The analysis is run utilizing four separate configurations: 1) 4DENSV, 2) 4DENSV with tangent linear normal mode constraint on the increment over all time levels, 3) 4DENSV with weak constraint digital filter, and 4) 4DENSV with weak constraint digital filter and the tangent linear normal mode constraint on the center of the window only. The fourth configuration that utilizes both constraints is done with computational considerations in mind. In all four configurations, there is no static

contribution to the solution (i.e. $\beta_f^{-1}=0$). The desire is to have the normal mode constraint remove gravity mode tendencies in the middle of the window while allowing the weak constraint digital filter to maintain the connection to the other time levels through the procedure to get the filtered state, and penalize additional noise accordingly.

Consistent with previous work in the literature utilizing 4DVAR, the divergence increment for the experiments that ran with the weak constraint digital filter is damped across the entire spectrum (Fig. 3.3). Some interesting behavior is observed on the high spatial frequency part of the spectrum, most of which is an artifact of the dual-resolution aspect of the configuration (which is the subject of Chapter 4). As expected, the more the solution is constrained, the more difficult it becomes to draw to the observations (Fig. 3.4). The penalty reduction is greatest for the 4DENSV case with no extra constraints, followed by the experiment that utilizing the normal mode constraint and then the two experiments that utilized the weak constraint digital filter. In all four configurations, the minimization appears well behaved with a steady reduction of the gradient norm. By design, the tangent linear normal mode constraint has a significant impact on the incremental tendencies, particularly the gravity mode tendencies (Table 3.2). On the other hand, while the weak constraint digital filter reduces the total tendencies by a factor of three, it has very little impact on the amount of tendencies that project onto undesirable gravity modes. The configuration that utilizes the weak constraint digital filter and normal mode constraint in the center of the window appears to be the best compromise, by significantly reducing the total tendencies while also reducing the percentage that

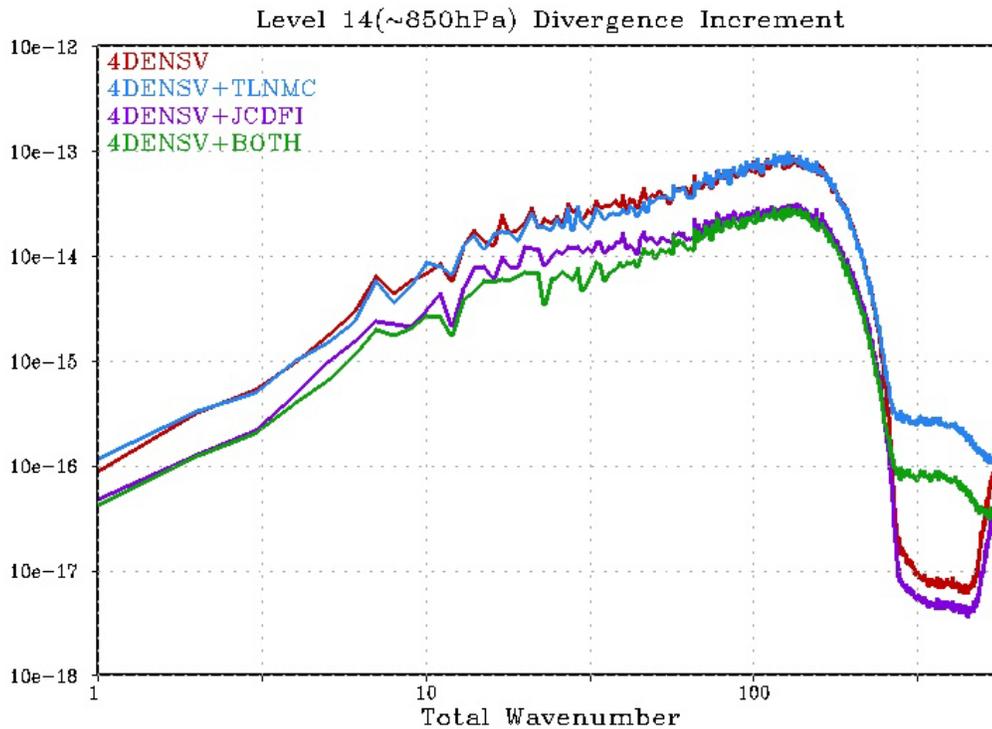
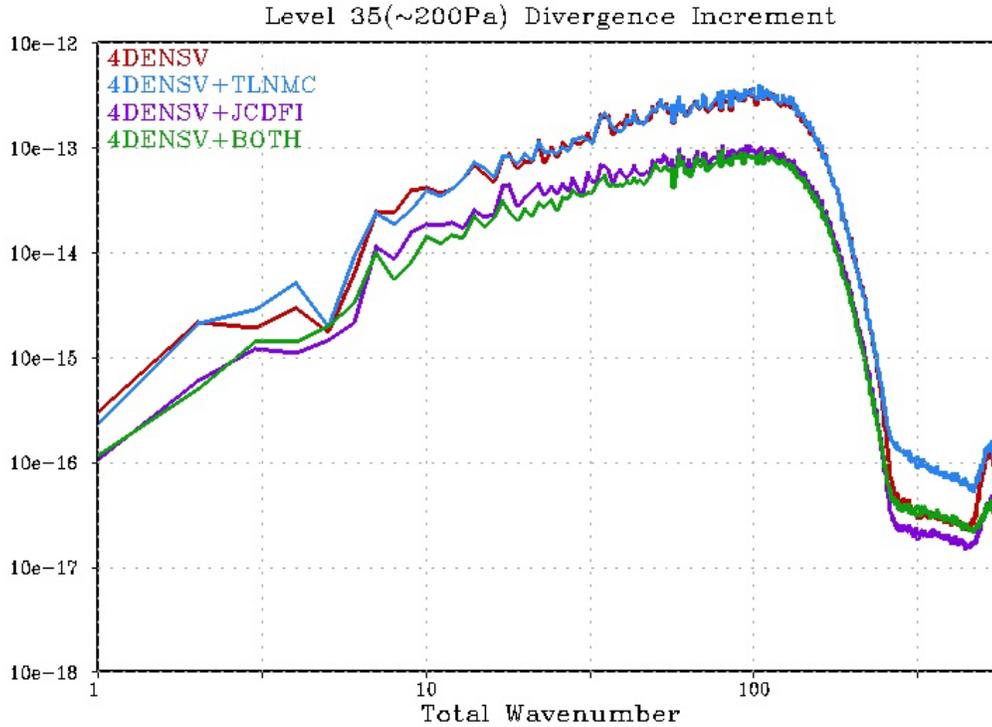


Figure 3.3: Power spectra of the divergence increment for a single analysis valid at 06 UTC 15 July 2010 on model level 35 (top, approximately 200 hPa) and 14 (bottom, approximately 850 hPa) utilizing 4DENSV (non-hybrid) with no constraints (red), tangent linear normal mode constraint (blue), weak constraint digital filter (purple), and combined normal mode and weak constraint (green).

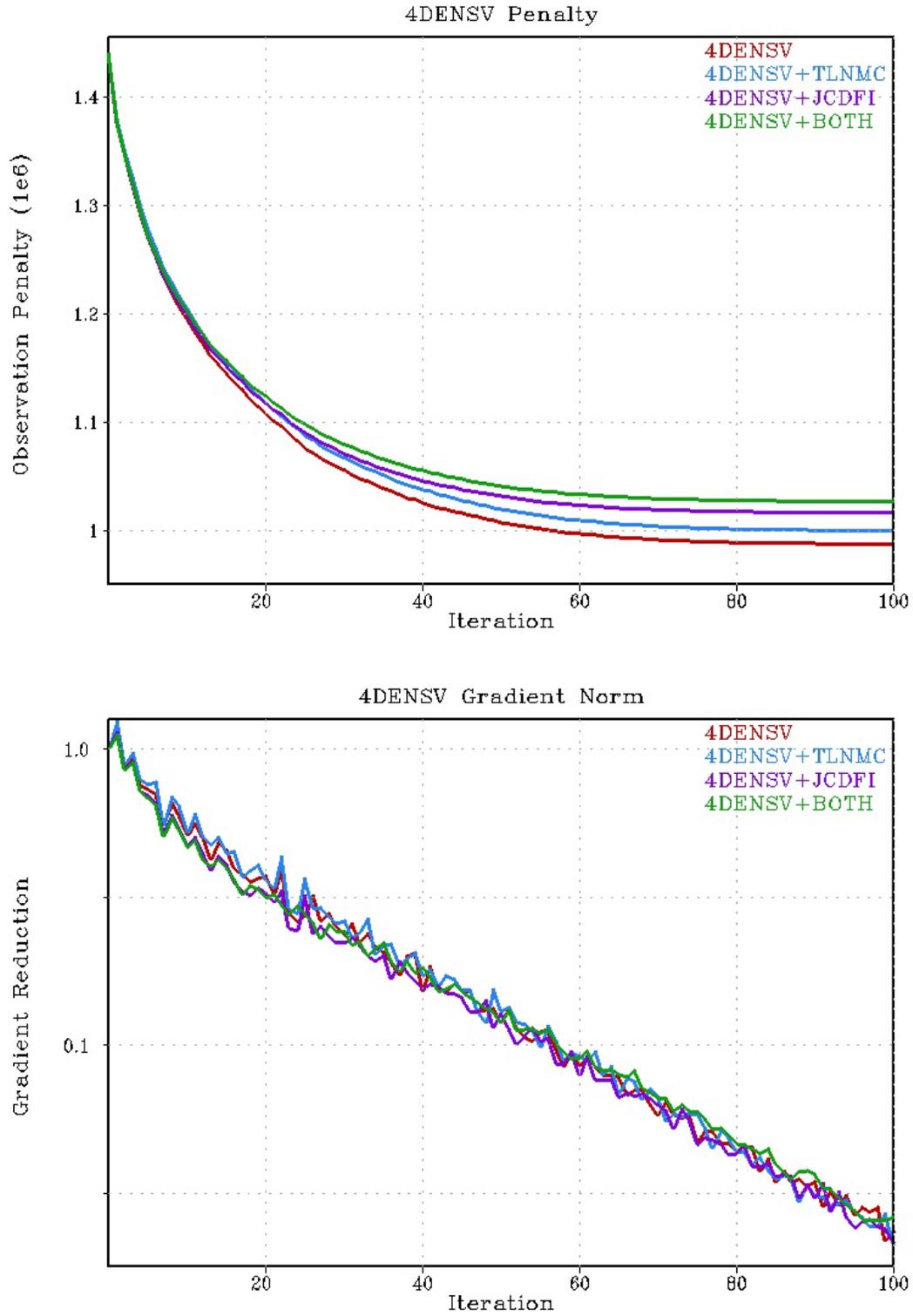


Figure 3.4: Observation penalty (top) and reduction of the gradient norm (bottom) by iteration for an analysis valid at 06 UTC 15 July 2010 utilizing 4DENSV (non-hybrid) with no constraints (red), tangent linear normal mode constraint (blue), weak constraint digital filter (purple), and combined normal mode and weak constraint (green).

<i>Experiment</i>	<i>Total Tendency</i>	<i>Gravity Mode Tendency</i>	<i>Ratio</i>
4DENSV	2.994×10^{-3}	2.884×10^{-3}	0.96
4DENSV+TLNMC	3.313×10^{-4}	2.029×10^{-4}	0.61
4DENSV+JCDFI	1.305×10^{-3}	1.279×10^{-3}	0.98
4DENSV+COMB	8.874×10^{-5}	6.608×10^{-5}	0.74
H-4DENSV	1.791×10^{-3}	1.597×10^{-3}	0.89

Table 3.2: The root mean square sum of the incremental (spectral) tendencies (total and gravity mode) as well as the ratio (gravity mode / total) for the eight vertical modes kept as part of the TLNMC for the single analysis valid at 06 UTC 15 July 2010 using various 4DENSV configurations.

project onto gravity modes, all at a significantly cheaper computational cost than running the normal mode constraint on all time levels. The H-4DENSV experiment is included in Table 3.2 to aid in the interpretation of results later in this chapter.

3.4 Experiment Results

3.4.1 Impact of 4D Ensemble

A series of experiments are designed to include various 4D-ensemble and hybrid components as an extension to the 3D runs that were described in Chapter 2. All of these new experiments utilize the same cold start initial condition for the control and ensemble, simulated observations from the Joint OSSE (Andersson and Masutani 2010), GFS model version and spatial resolutions (dual resolution mode), comparable versions for the GSI and EnKF codes, identical inflation parameters, and

cycling configuration for the data assimilation including the re-centering procedure. The experiments are designed in such a way as to infer which components of the hybrid and 4D extensions yield improvements (or degradations) to the quality of analyses, building upon the 3DVAR, 3DENSV, and 3DHYB runs that were included in Chapter 2. A listing and description of the five experiments can be found in Table 3.3.

<i>Experiment</i>	<i>Control</i>	<i>Description</i>	<i>Relevant Figures</i>	<i>Relevant Equations</i>
4DENSV	3DENSV	Pure 4D-Ensemble-Var (No Static \mathbf{B}), no TLNMC or JCDFI	Fig. 3.5	Eq. 3.4
H-4DENSV	4DENSV	4DENSV + $\beta_f^{-1}=0.25, \beta_e^{-1}=0.75$	Fig. 3.6	Eq. 3.6
H-4DENSV_NMI	H-4DENSV	H-4DENSV + TLNMC (all observation k -levels)	Fig. 3.7 Fig. 3.8	Eq. 3.6 Eq. 3.9
H-4DENSV_DFI	H-4DENSV	H-4DENSV + JCDFI ($\chi=10$)	Fig. 3.9	Eq. 3.6 Eq. 3.10
H-4DENSV_COMB	H-4DENSV 3DHYB	H-4DENSV + TLNMC (k =mid only) + JCDFI ($\chi=10$)	Fig. 3.11 Fig. 3.12	Eq. 3.6 Eq. 3.9 ($k=m$) Eq. 3.10

Table 3.3: Description of various (hybrid) 4D-ensemble-var OSSE-based experiments and relevant figures showing comparison of analysis errors relative to the control run for each.

The first experiment, 4DENS_V, is designed to test the impact of going simply from 3DENS_V to 4DENS_V. In this context, none of the constraints are invoked and the variational solver is utilized simply to initialize the high resolution deterministic component. Likewise, the solution in both the 3DENS_V and 4DENS_V experiments is exclusively reliant upon the ensemble to prescribe the background error covariance. The assimilation window is kept fixed at six hours, though the 4DENSE_V experiment utilizes hourly time levels (and therefore, hourly background and ensemble perturbations). The change in the analysis error for zonal wind, temperature, and specific humidity is shown in Fig. 3.5. Generally speaking, the analysis error is smaller in the 4DENS_V experiment, especially for upper tropospheric extratropical winds and temperature, and lower tropospheric water vapor. It appears that by going from 3D to 4D, the temperature analysis error has actually increased over the southern polar cap in the lower troposphere. Recall from Chapter 2 (Fig. 2.6) that this region (southern polar cap, below 700 hPa) had seemingly large analysis errors to begin with, at least partly an artifact of utilizing pressure level data on a common grid and exposing interpolation and extrapolation issues (done independently and in a different manner in the validation data set and post-processing of each of the data assimilation experiments). It is still interesting that this region is being exposed to exhibit an error increase in going to 4D, perhaps indicative of issues related to differences in topography between the nature run and model used in the experiments being exacerbated. Also of note is the slight increase in analysis error for specific humidity in the mid-tropospheric tropics. By in large, however, the impact in going from 3DENS_V to 4DENS_V is generally positive,

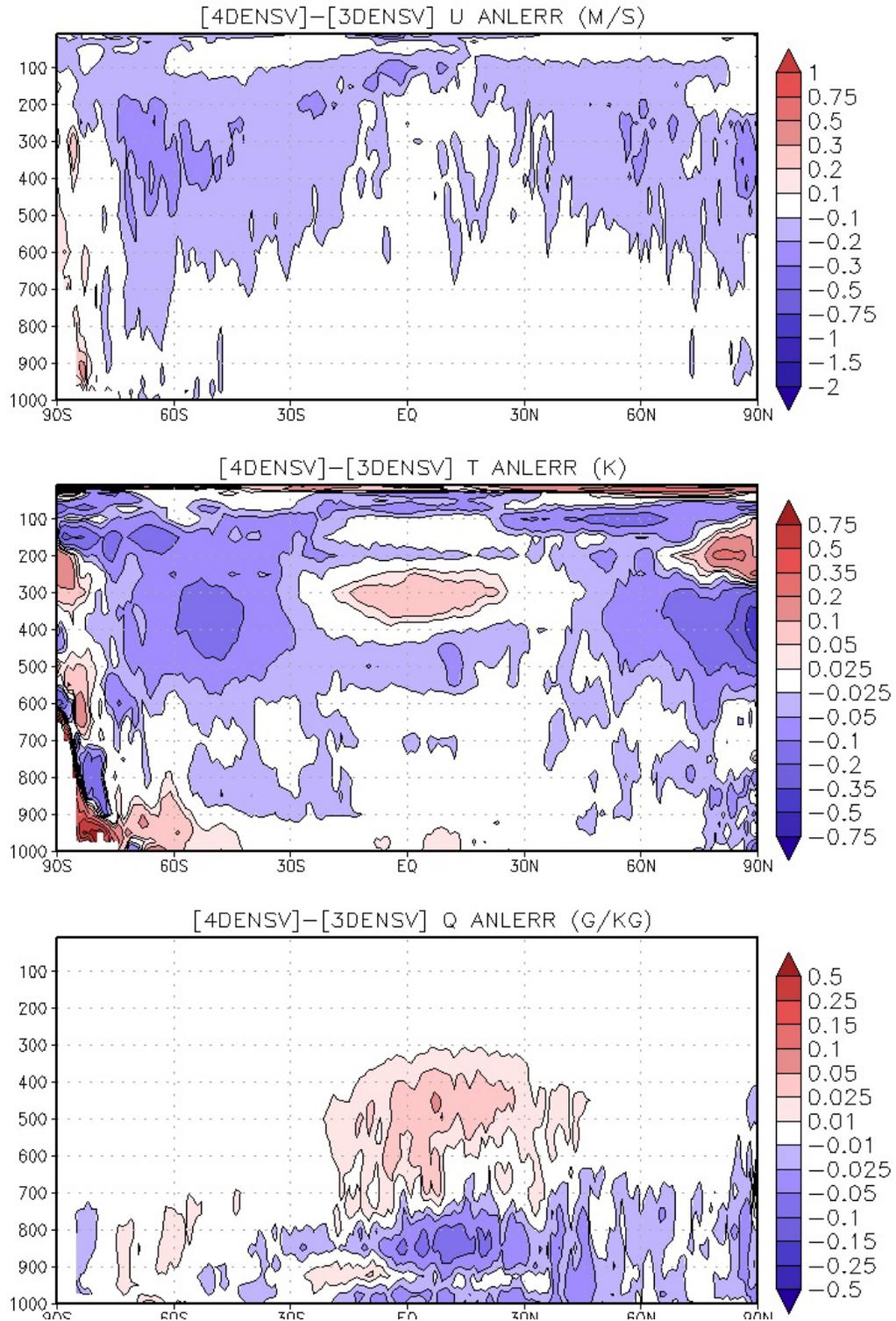


Figure 3.5: Difference in the time-averaged (August) zonal mean standard deviation of the analysis error for the 4DENSV experiment relative to 3DENSV for zonal wind (top, ms^{-1}), temperature (middle, K), and specific humidity (bottom, g kg^{-1}), covering the period spanning 00 UTC August 2005 through 18 UTC 30 August 2005.

consistent with the better use of observations distributed through the assimilation windows as was demonstrated with the single observation test.

By utilizing equations 3.6 and 3.7, an H-4DENS_V experiment is designed to demonstrate the impact of including a time invariant static contribution to the analysis increment. In this case, parameter settings similar to the 3DH_{YB} experiment from chapter 2 are chosen with $\beta_f^{-1}=0.25$ and $\beta_e^{-1}=0.75$, so the (time-invariant) static contribution to the analysis increment is chosen to be 25% (and therefore, relying 75% on the ensemble covariances). The tangent linear normal mode constraint is applied to the static component to the analysis increment solution only (and not to the total increment, see equations 3.8 and 3.9) in an effort to cleanly demonstrate the impact of the static contribution alone. Given that the default settings chosen for the inflation parameters in the ensemble update result in an over-spread ensemble, it should be expected that the hybrid also helps to reduce the effective error variances thereby resulting in a better analysis. A comparison of the analysis errors from this H-4DENS_V experiment with the 4DENS_V run show a nearly uniform reduction in the analysis error (Fig. 3.6), with the largest impacts evident in the extratropical wind and temperature fields, tropical upper tropospheric winds, and lower tropospheric specific humidity. The inclusion of the static **B** to the 4DENS_V solution has as big of an impact for some variables and levels as was found when going from 3DVAR to 3DH_{YB} (Fig. 2.8), though in different geographic regions (and more consistently).

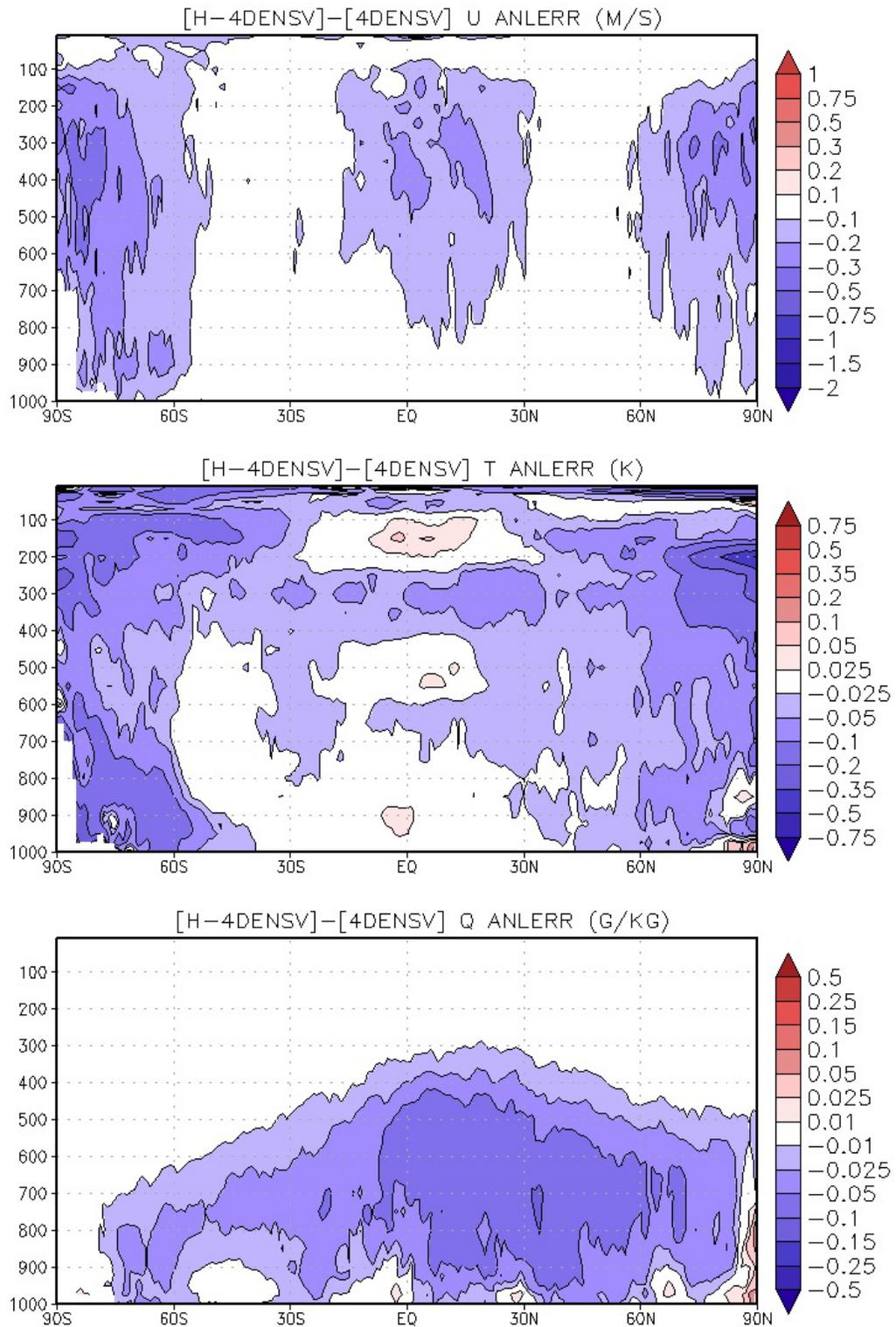


Figure 3.6: As in Fig. 3.5, but for the H-4DENSV minus 4DENSV experiment errors.

3.4.2 Constraint Experiments

To further improve the quality of analyses within the 4D-ensemble-var hybrid context, experiments are designed to explore the contribution of adding further constraints to the total increment, such as through the inclusion of the normal mode constraint and weak constraint digital filter. The first of such experiments is analogous to the original 3DHYB experiment that yielded such good results in Chapter 2, and applies the tangent linear normal mode constraint to the total analysis increment over all (hourly) k -time levels as in equation 3.9. It should be noted, again, that such a configuration is quite computationally expensive as this particular constraint requires the calculation of incremental tendencies as well as spectral to grid transforms, all within the iterative minimization scheme. However, the point here is to demonstrate the impact of such a constraint on the quality of analysis. This experiment, denoted H-4DENSV_NMI, is carried out in the hopes of improving upon the H-4DENSV result described in the previous section. However, the constraint appears to have very little (and localized) impact on reducing the analysis error (Fig. 3.7), confined to small areas poleward of 70S for temperature and wind. Similar to the findings in Kleist et al. (2009), the normal mode constraint does have an impact on reducing the background and analysis errors for surface pressure (Fig. 3.8). This is consistent with the removal of incremental tendencies that project onto gravity modes, as also demonstrated within the context of the single analysis investigation in Section 3.3.3.

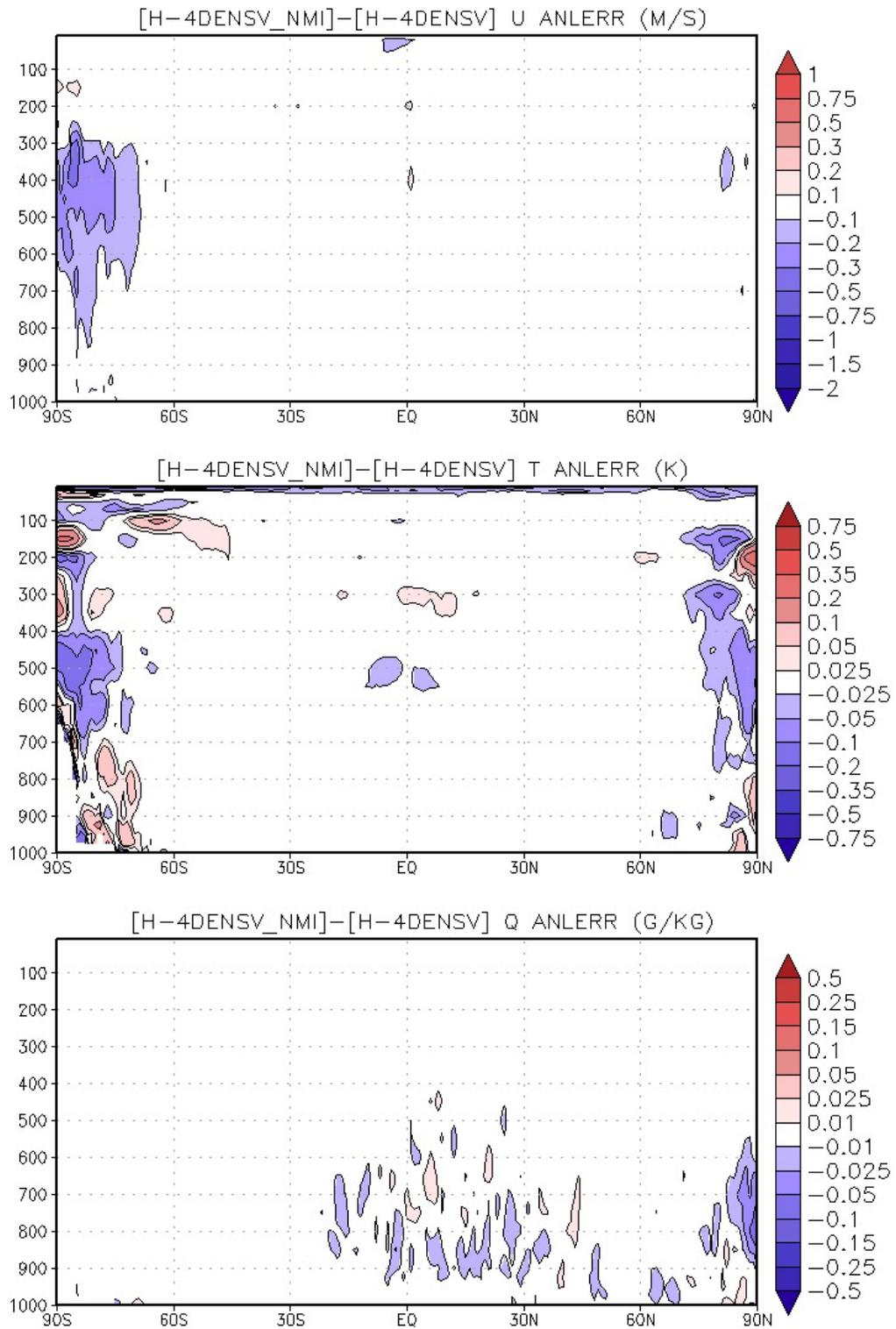


Figure 3.7: As in Fig. 3.6, but for the H-4DENSV_NMI minus H-4DENSV experiment errors.

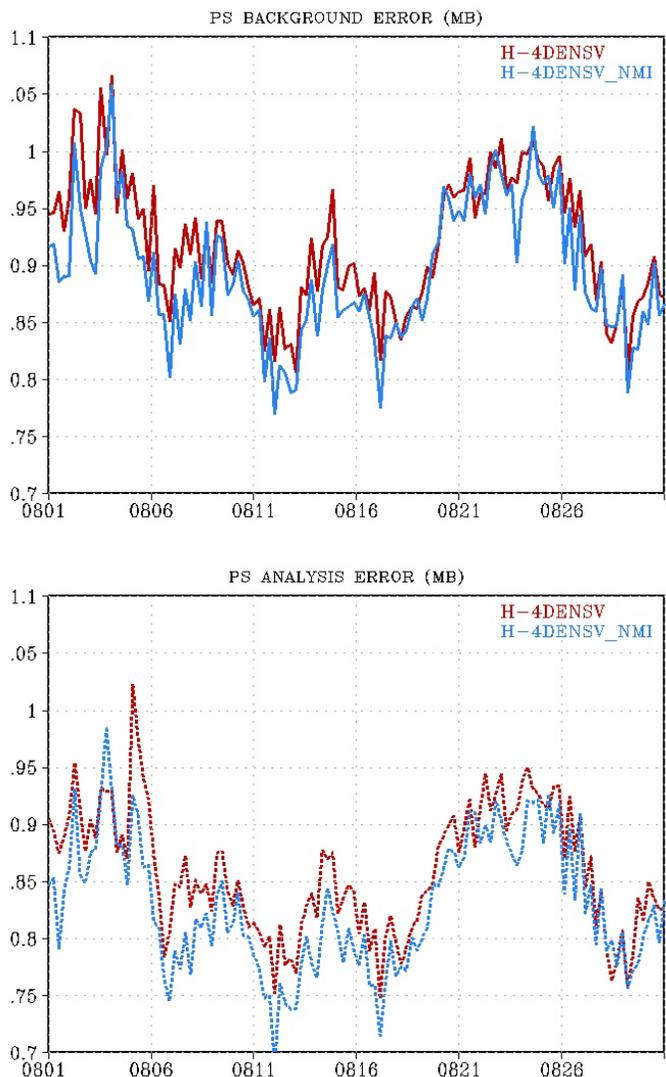


Figure 3.8: Time series showing the global mean surface pressure error (hPa) for the background (top, solid) and analysis (bottom, dashed) from the H-4DENSV (red) and H-4DENSV_NMI (blue) experiments.

Another experiment, called H-4DENSV_DFI is conducted to test the impact of including the weak constraint digital filter in place of the normal mode constraint in the 4DENSV context. Outside of the change in constraint terms, all of the other settings from the previous H-4DENSV experiments remain unchanged. The digital

filter term is chosen to be based on a dry energy norm and utilizes a weighting parameter (χ) of ten, based on previous findings and documentation in literature. The impact on reducing the analysis error is found to be larger than that for the normal mode constraint for temperature, wind, and specific humidity (Fig. 3.9). The digital filter constraint does not have an impact on the surface pressure background and analysis errors as was noted for the normal mode constraint (not shown). The improvement found in the specific humidity analysis error is surprising, given the choice of norm for the weak constraint penalty term (which does not include a moist component). This suggests that the constraint is improving the overall quality of analysis in addition to positively impacting moist and precipitation processes in the short-term forecasts.

It should be pointed out, however, that the impact of including the dynamic constraints on the 4DENSV formulation would likely be different in the absence of a static **B** contribution, which is utilized in H-4DENSV_NMI and H-4DENSV_DFI. One of the impacts that the static covariance has on the solution is to reduce the amplitude of the imbalance/noise. A comparison of the incremental tendencies for the previously discussed single analysis case (Table 3.2) shows that H-4DENSV has tendencies that are roughly twice as small in amplitude as compared to the pure 4DENSV case, with a much smaller projection onto gravity modes. Although the impact of the static **B** on the imbalance is not as large as either of the constraints, it is noteworthy to aid in the interpretation of results.

Building on the success of the H-4DENSV_NMI and H-4DENSV_DFI experiments and given that each of the two constraints appears to be contributing to

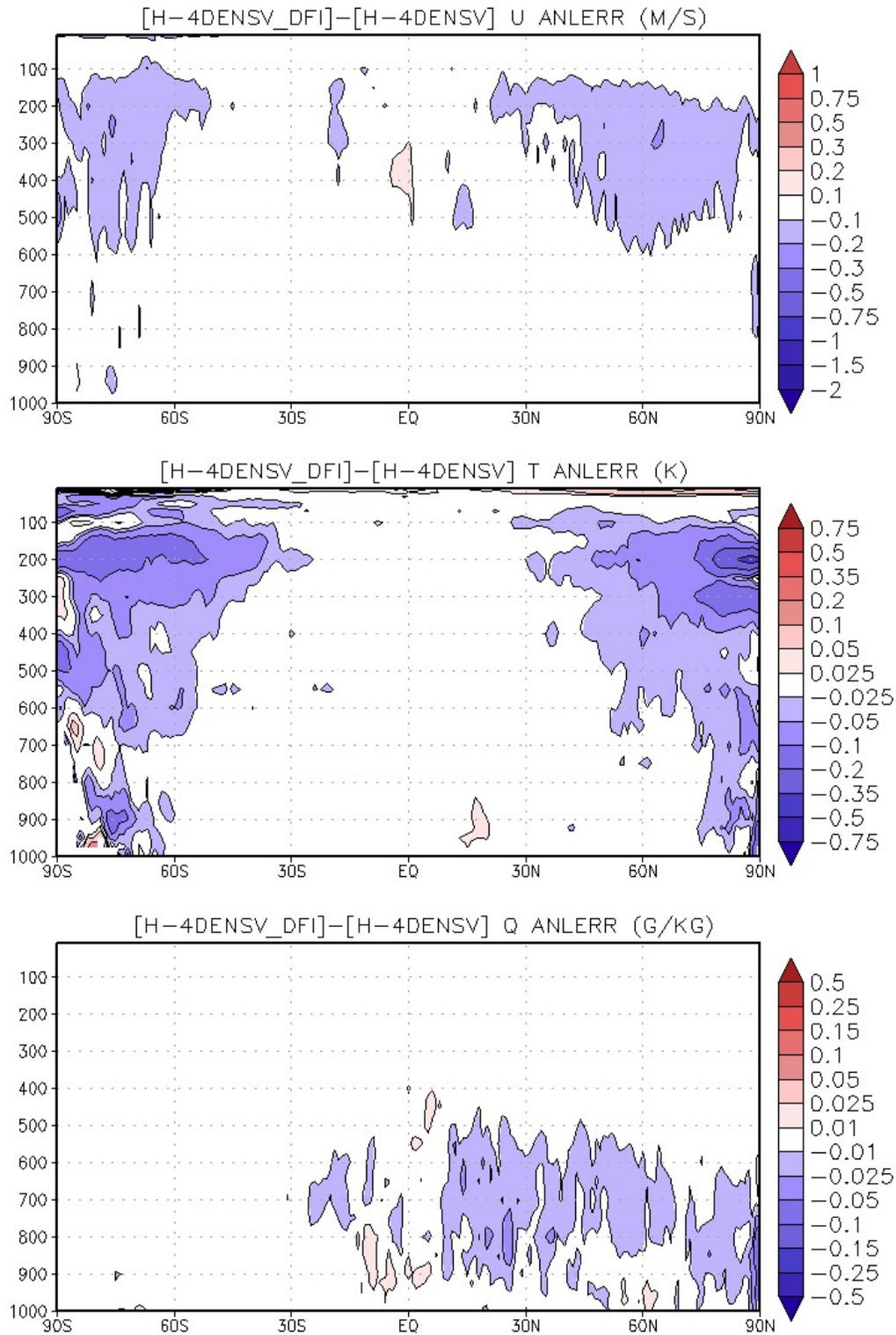


Figure 3.9: As in Fig. 3.7, but for the H-4DENSV_DFI minus H-4DENSV experiment errors.

improving the analysis through different means, a new experiment (H-4DENSV_COMB) is designed that utilizes a combined tangent linear normal mode constraint and weak constraint digital filter. In this experiment, the weak constraint digital filter parameter choices are identical to those utilized in H-4DENSV_DFI. However, the normal mode constraint is only applied to the center of the assimilation window, or in other words, to the actual increment that is added to the 6-hr forecast which is then passed on as the initial condition for the forecast model. Explicitly applying the constraint to this specific time level will help ameliorate spin up issues by forcing the parts of the increment that project onto high frequency modes to zero. The weak constraint digital filter uses information over all of the time levels, and as such, aids to ‘propagate’ the balance information to the time levels away from the center of the window according to the filter weights (the center of the window is the largest contributor to the filtered state). This is demonstrated by looking at the total and gravity mode tendencies from the single analysis case used in section 3.3, comparing the increments from a 4DENSV configuration and a configuration that utilizes both the weak constraint digital filter and normal mode constraint in the center of the assimilation window (Fig. 3.10). First, note that the tendencies are significantly larger for the analysis without any constraints, with a very large portion of the tendencies coming from those that project onto gravity modes. The tendencies in the analysis that utilizes the constraints have a distribution throughout the window similar to the inverse of the filter weights, with the smallest amplitude in the center of the window and then increasing further away. By design, the incremental tendencies at the center of the window have a much smaller projection onto gravity modes.

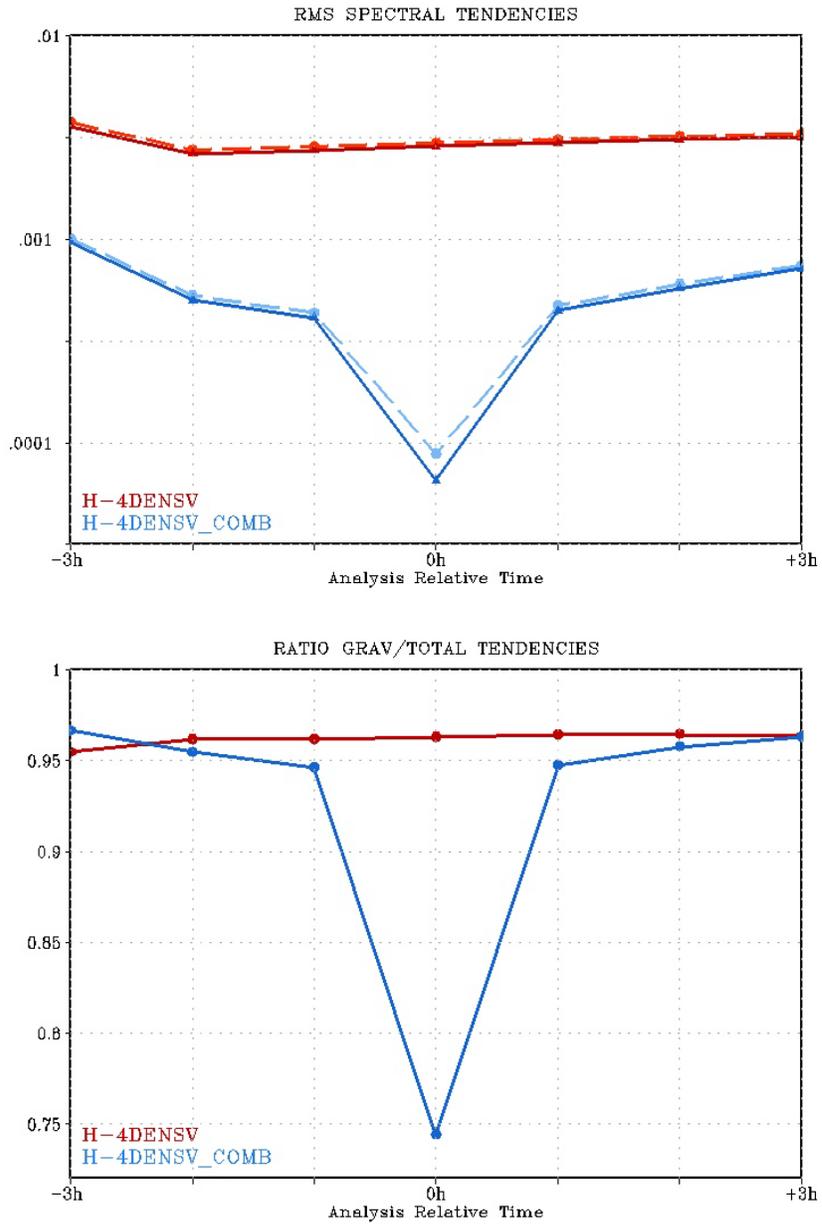


Figure 3.10: The root mean square sum of the incremental spectral tendencies (top) for the total tendency (4DENSV-red dashed, 4DENSV_COMB-blue dashed) and gravity mode tendency (4DENSV-red solid, 4DENSV_COMB-blue solid) for the eight vertical modes kept as part of the TLNMC for the single analysis valid at 06 UTC 15 July as a function of observation bin (analysis relative time). Also plotted is the ratio (bottom, gravity mode/total tendencies) for the 4DENS (red) and 4DENSV_COMB (blue) experiment.

The inclusion of both constraints in the H-4DENSV_COMB experiment has a positive impact on the quality of analysis (Fig. 3.11), in the same regions as was noted in the H-4DENSV_NMI and H-4DENSV_DFI experiments. Interestingly, the impact from including both constraints is slightly larger in amplitude than the combined impact of each of the constraints individually. This validates the notion that the two constraints are contributing positively to the analysis in quasi-independent manners (i.e., they are filtering different aspects of noise and imbalance). It is also interesting the constraints are contributing positively in the very same aspects as the static error covariance (comparing Figs. 3.6 and 3.11), resulting in significant improvements over pure 4DENSV when all three are utilized. This is quite encouraging for operational (high resolution) applications, given that the configuration utilized in the H-4DENSV_COMB experiment is only slightly more expensive than the 4DENSV case (less than 10% increase in computational time), which in itself is almost exactly twice as computationally expensive as the 3DENSV configuration (all of which are much cheaper than a comparable 4DVAR or H-4DVAR_AD would be).

A comparison of the analysis error from the H-4DENSV_COMB experiment with the 3DHYB experiment from Chapter 2 reveals that the inclusion of the 4D ensemble and constraint configuration results in significant reduction in analysis errors (Fig. 3.12). This is particularly true for zonal wind, where the error reduction is evident for nearly all levels and regions, with a notable lack of change in the tropical lower troposphere. Recall from Chapter 2 that the reduction in error was more evident in the tropics when going to the 3DHYB, with smaller error reduction

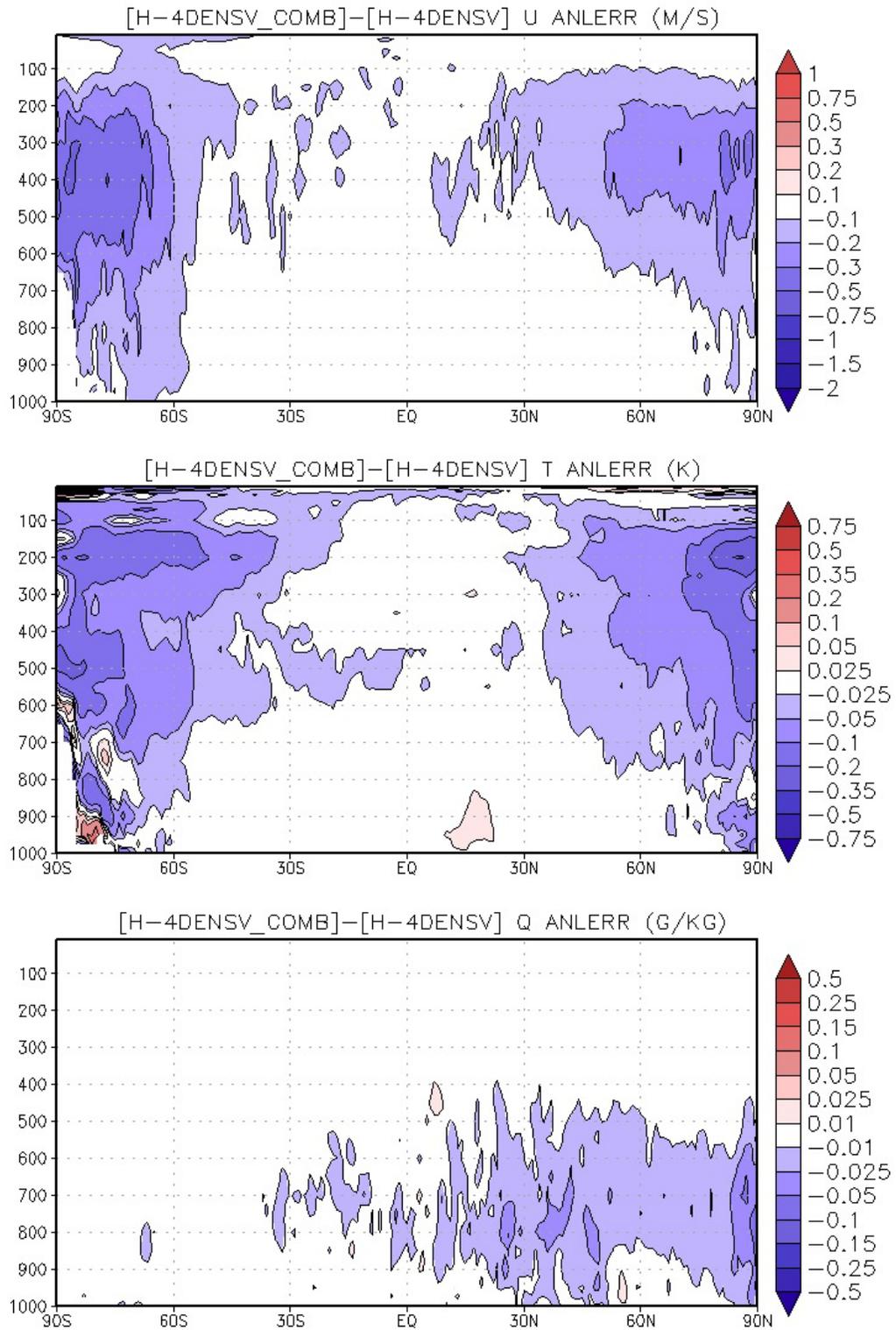


Figure 3.11: As in Fig. 3.9 but for the H-4DENSV_COMB minus H-4DENSV experiment errors.

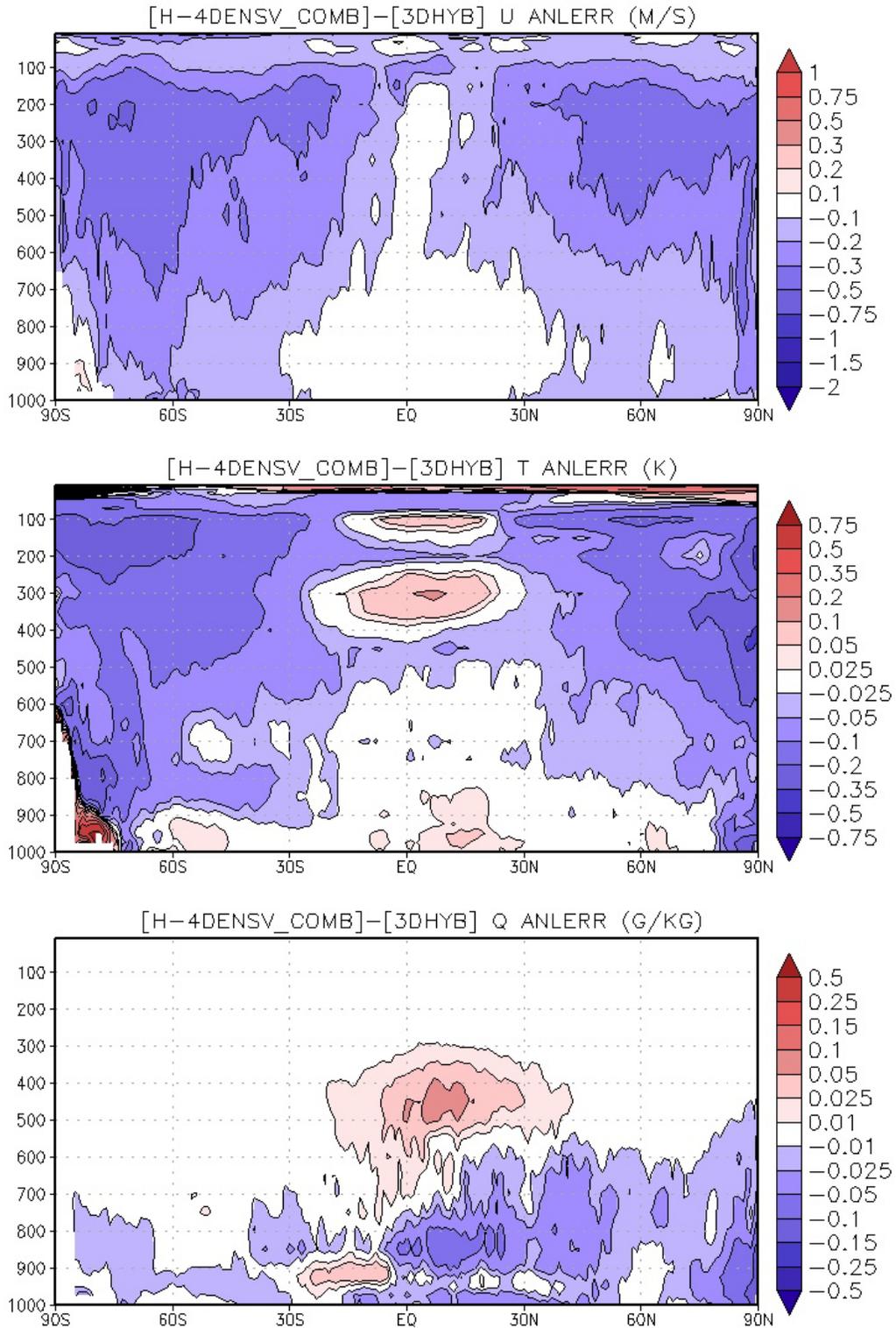


Figure 3.12: As in Fig. 3.11 but for the H-4DENSV_COMB minus 3DHYB experiment errors.

(and in fact, an error increase for some variables and levels) in the extratropics. The addition of the ensemble-based covariance seems to have the biggest impact in the tropics, whereas the extension to 4D has a bigger impact on the more dynamically active extratropics (particularly for temperatures and winds near jet level). Note that the amplitude in the error reduction is similar when going from 3D to 4D hybrid as compared to the reduction found when adding the hybrid in the 3D context over the 3DVAR control. This is further confirmed by visualizing the geographic distribution of the 300 hPa zonal wind analysis errors in the H-4DENSV_COMB and 3DHYB experiments as well as their differences (Fig 3.13). Although the local maxima in wind errors in several regions in the tropics associated with convection still exist, the amplitude is comparable to the 3DHYB results. The reduction in the errors outside of the tropics is significant.

The two problem areas that showed up in the comparison of the 4DENSV and 3DENSV experiments are present yet again, however: the low-level temperatures near the Antarctic ice shelf and the mid to upper tropospheric specific humidity in the tropics (Fig. 3.12). The increase in analysis error for specific humidity (and very subtly for temperature) in the upper tropospheric tropics is a bit puzzling, given the lower tropospheric errors are in fact reduced. A closer inspection of two sets of 3D to 4D experiments reveals that the 4D paradigms result in more precipitable water in the analysis (Fig. 3.14), most notably in the tropics. This subtle change in the amount of moisture in the analysis between the 3D and 4D paradigms is likely attributable to the treatment of the unphysical moisture (nonlinear) weak constraint terms used in the GSI. The penalty terms for negative water vapor and supersaturation involve global

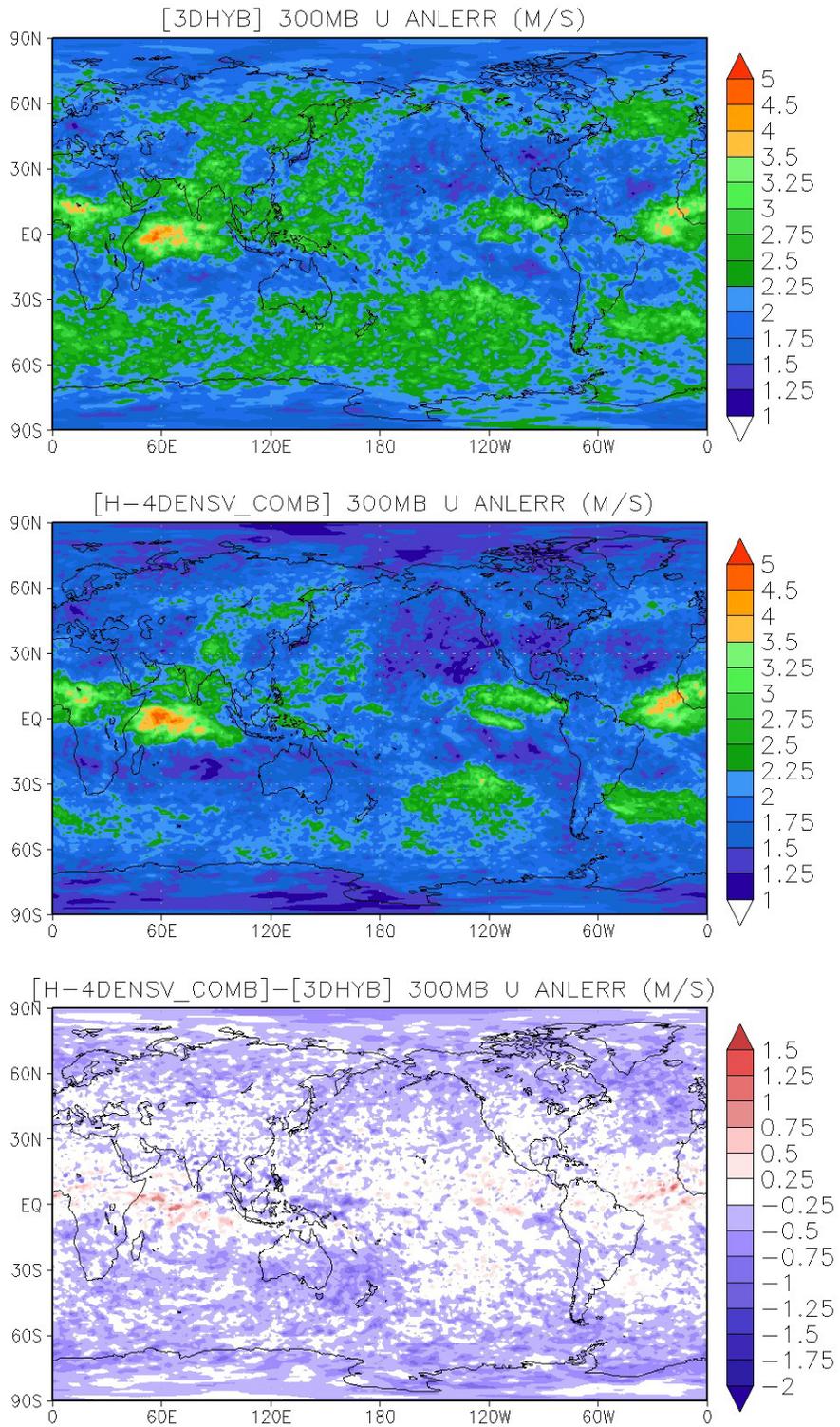


Figure 3.13: Time averaged analysis error standard deviation for 300 hPa zonal wind (ms^{-1}) from the 3DHYB (top) and H-4DENSV_COMB (middle), as well as the difference (bottom, H-4DENSV_COMB minus 3DHYB) for the simulated August period.

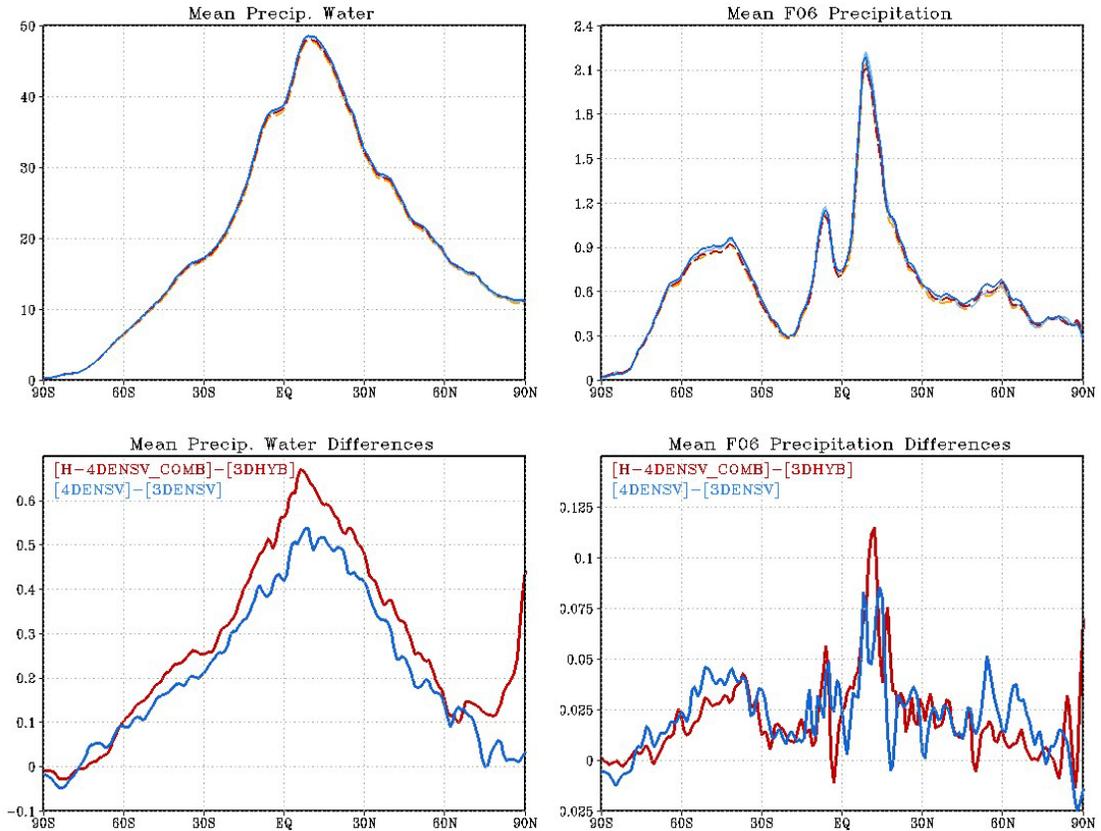


Figure 3.14: Analysis time-mean zonally averaged precipitable water (kg m^{-2} , upper left) and 06-hr forecasted precipitation (kg m^{-2} , upper right) for the simulated August period for the 3DENSV (blue, dashed), 4DENSV (orange, dashed), 3DHYB (red, solid) H-4DENSV_COMB (blue, solid) experiments. The differences are shown on the bottom between the 4DENSV minus 3DENSV (blue) and H-4DENSV_COMB minus 3DHYB (red) for precipitable water (kg m^{-2} , lower left) and precipitation forecast (kg m^{-2} , lower right).

sums of the amount of unphysical values. In the extension of the solution to the 4D hybrid, these penalty terms are also extended to include options to utilize the entire 4D increment as to ensure consistency between the different time levels (and therefore treatment of observations). This results in a significant change in the manner for which these (highly nonlinear) weak constraint terms operate. In this case, it seems that this change in weak constraint term (without further tuning of the

weighting parameters) in combination with the 4D use of the observations results in a slight increase in the amount of analyzed moisture.

All other things being equal (such as the forecast model, nature of balance in the initial conditions, etc.), an increase in available moisture will result in a corresponding increase in precipitation for a short term forecasts, as is the case here (Fig. 3.14). Given that the modeled tropical precipitation is dominated by convective parameterizations in both the nature run model and GFS, it is possible that such a distribution of decrease (lower level) and increase (upper level) in errors is associated with differences in the parameterized convection between the two models.

A comparison of the H-4DENSV_COMB experiment to the original 3DVAR control reveals a substantial reduction in analysis error for all variables and levels (Fig. 3.15). One exception is in the region that was noted to have ensemble spread (error variances) that is much too large compared to the actual background errors. It is expected that a re-run of the H-4DENSV_COMB experiment with reduced inflation parameters will result in analysis errors that are even further reduced, similar to the findings in the 3DHYB and 3DHYB_RS experiments. The slight increase in error for moisture in the upper tropospheric tropics that was found when going from 3D to 4D is masked by the substantial gains attributable to the addition of the ensemble based covariances.

3.4.3 Forecast Impact

Similar to Chapter 2, a forecast impact experiment is carried out utilizing 00 UTC analyses from the H-4DENSV_COMB configuration to initialize the GFS

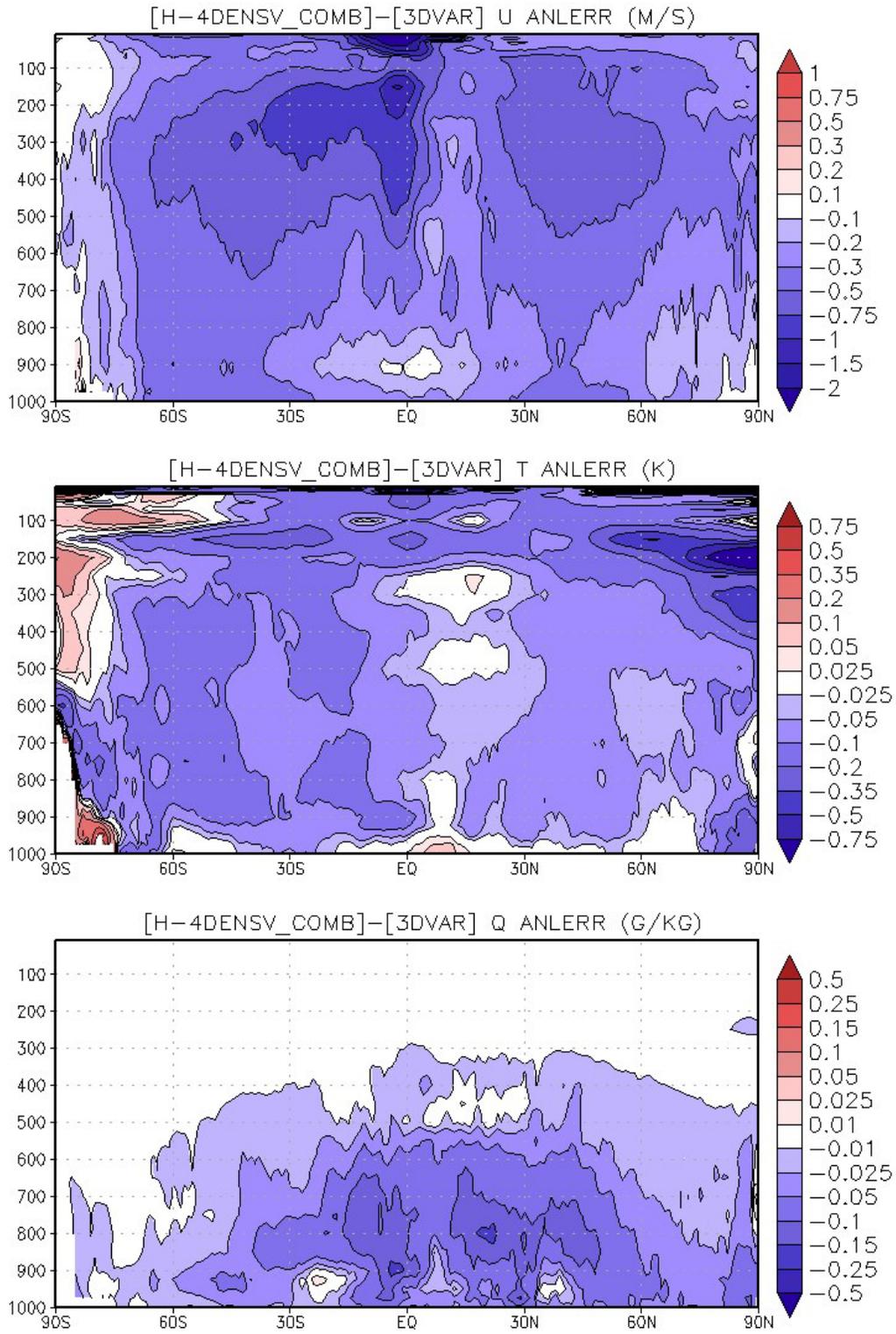


Figure 3.15: As in Fig. 3.12 but for the H-4DENSV_COMB minus 3DVAR experiment errors.

model for integration out to 7.5 days. The first two weeks of the experiment are ignored to account for spin-up. All verification is done relative to the ECMWF nature run on a common grid. Recall that it was found that forecasts initialized from the 3DHYB analyses were generally superior to those that were initialized from 3DVAR analyses. Given that the H-4DENSV_COMB experiment demonstrated reduced analysis errors relative to 3DHYB, it is reasonable to expect a similar reduction in forecast errors.

A comparison of the mean geopotential height errors in the extratropics reveals the forecasts initialized from the H-4DENSV_COMB analyses are superior to those from the 3DHYB experiment in the northern hemisphere, but slightly worse in the southern hemisphere (Fig. 3.16). The differences between the H-4DENSV_COMB and 3DHYB forecasts for short lead times for 500 hPa geopotential height are statistically significant, whereas the differences for other regions and levels are not. The apparent degradation to the southern hemisphere forecasts in the H-4DENSV_COMB experiment is quite interesting, given the strong evidence for a generally superior analysis. The H-4DENSV_COMB forecasts are superior to the original 3DVAR control for both regions, levels, and for all lead times as expected.

Similar to the extratropical geopotential heights, there is evidence that the H-4DENSV_COMB results in less skillful forecasts in terms of tropical vector wind root mean square errors (Fig. 3.17). It is a bit puzzling that a significant reduction in analysis error for the 200 hPa vector wind quickly disappears, within 24 hours, relative to the 3DHYB. This is possibly consistent with the previously noted issues regarding the increased convection in the tropics for all of the 4D-based experiments.

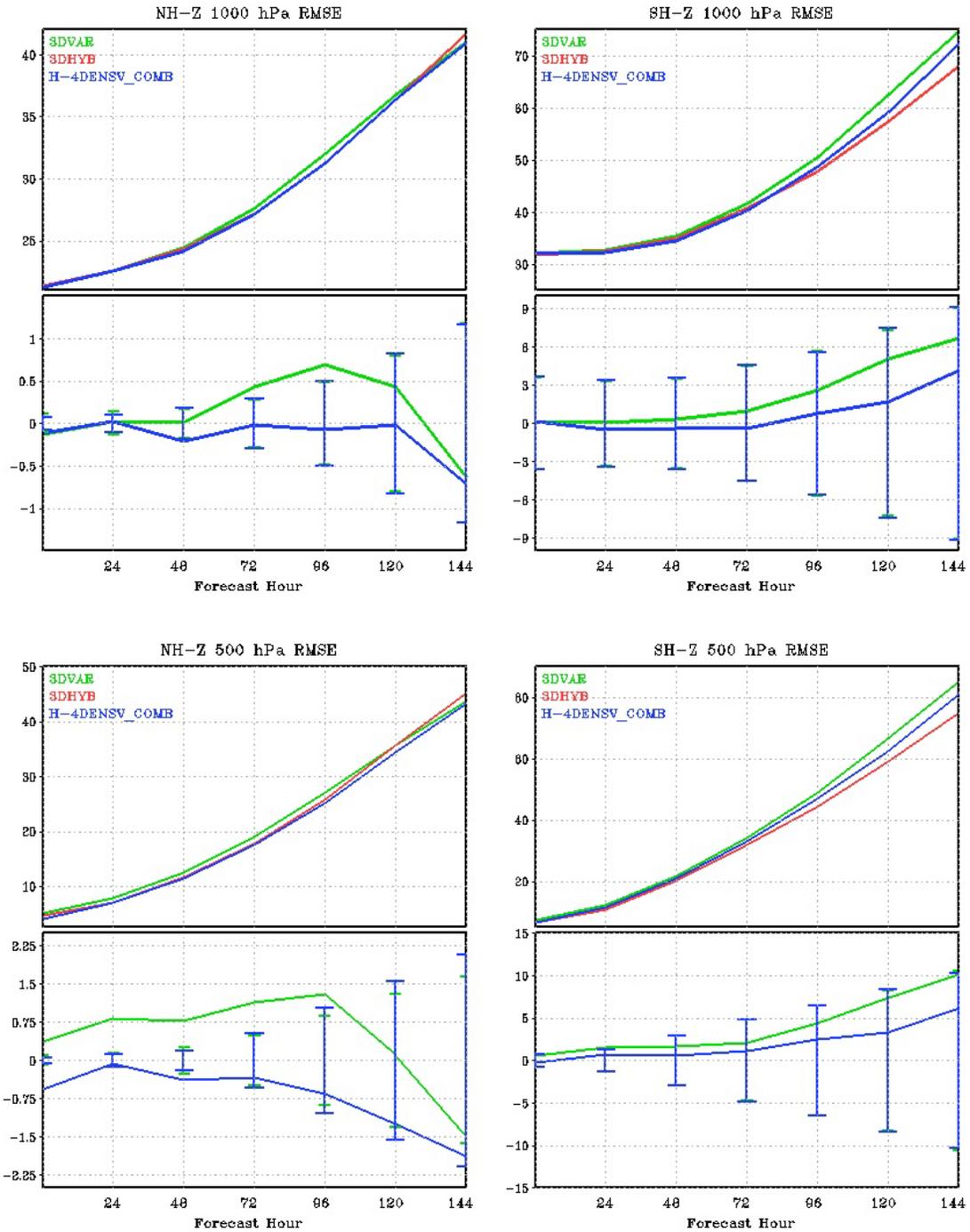


Figure 3.16: The average geopotential height root mean square error by lead time for the 3DVAR control (green), 3DHYB (red) and H-4DENSV_COMB (blue) experiments verifying daily between the simulated 27 July and 01 September for the northern hemisphere (left) and southern hemisphere (right) at 1000 hPa (top) and 500 hPa (bottom). The difference between the experiments is shown on the bottom section of each panel for the 3DVAR-3DHYB (green) and H-4DENSV_COMB-3DHYB (blue) along with the 95% confidence derived from a t-test. Forecasts for each experiment were initialized at 00 UTC and verified against the ECMWF nature run.

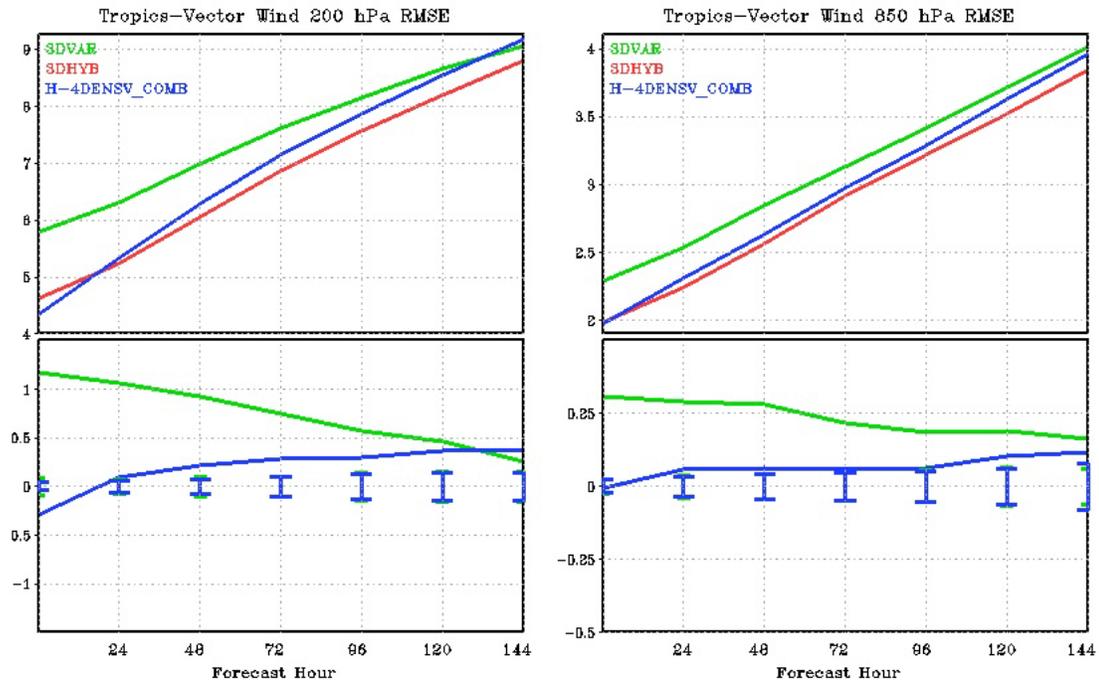


Figure 3.17: As in Fig. 3.16, but for the vector wind root mean square errors in the tropics at 200 hPa (left) and 850 hPa (right).

The degradation in the H-4DENSV_COMB experiment does appear to be statistically significant. It should be noted, however, that both the H-4DENSV_COMB and 3DHYB experiments are superior beyond a 95% confidence threshold relative to the 3DVAR control for all lead times out to 5 days.

3.5 Summary and Conclusions

An extension of the GSI-based hybrid variational ensemble algorithm to include 4D ensemble perturbations is proposed and implemented. The 4DENSV algorithm has several advantages relative to 4D-EnKF and 4DVAR algorithms that make it attractive for an operational center such as NCEP. Since 4DENSV does not

require the use of an additional dynamic model (i.e., the tangent linear or adjoint), it requires less in terms of computational resources than a traditional 4DVAR algorithm. Since it is based on a variational algorithm, it becomes quite trivial to supplement the ensemble-based (4D) covariance with some static estimate which can help ameliorate potential sampling issues, particularly for small ensemble sizes. Furthermore, such a configuration allows for easy implementation of a dual-resolution algorithm, solving for a high resolution increment using a low resolution ensemble and control variable (analogous to the interpolation of ensemble weights algorithm, Yang et al. 2009). Lastly, variational-based methods have the advantage of physical space localization (implicit through the “background error” for the control variable) which can be important for observations such as satellite radiances.

The ability to incorporate dynamic constraints on the solution within the variational framework is quite attractive. Some examples include tangent linear normal mode constraints, weak constraints on unphysical values for atmospheric tracers, and weak constraints on conservation of global mean dry mass. It is demonstrated that these constraints operate quite well within the 4DENSV framework. Along these lines, an adaptation of a weak constraint digital filter (typically used in 4DVAR applications) is proposed and implemented. A weak constraint digital filter allows for the control of noise within the 4D increment without much additional computational costs.

Several OSSE-based experiments are carried out to demonstrate that the 4D-ensemble-var algorithm and its hybrid variants contribute positively to the quality of analysis relative to the 3DENSV and 3DHYB experiments that were carried out in

Chapter 2. It is found that going from 3D to 4D ensemble perturbations reduces the analysis error for most variables and levels. The addition of a time-invariant static contribution to the analysis increment (i.e., hybridization) further reduces the analysis error. The dynamic constraints (tangent linear normal mode constraint and weak constraint digital filter) are found to have a much smaller impact (individually) on the reduction in analysis error, but both do contribute positively.

A follow-on experiment is designed to attempt to utilize both constraints, where the normal mode constraint is used to filter the solution only at the center of the window and the weak constraint digital filter is utilized to help control the noise throughout the 4D increment. Such a configuration is computationally inexpensive relative to applying the normal mode constraint over all time levels, making it an attractive option for realistic (and high resolution) applications. It is found that this combined constraint reduces the analysis error more than the sum of the individual constraint experiments do. The combined constraint has the largest impact on reducing the wind errors, particularly in the extratropics.

All of the 4D-based experiments have a larger impact, relative to their 3D counterparts, in the extratropics (particularly in the upper troposphere) than in the tropics. This is consistent with the fact that these regions are more dynamically active, where the propagation of information and the usage of the appropriate time levels within the observations are more important. The amplitude of the error reduction in going from the 3DHYB to H-4DENSV_COMB is similar to that when going from the 3DVAR control to 3DHYB. The combined impact then of the H-4DENSV_COMB relative to the original 3DVAR control is substantial for nearly

all variables and levels. The only real problem areas are related to the regions for which the ensemble was found to be significantly over-spread (particularly poleward of 70S). It is expected that a rerun of the H-4DENSV_COMB experiment with tuned inflation parameters (to reduce the ensemble spread, particularly in those regions) will result in even further improved analyses. Further investigation may be needed to sort out precisely why the 4D extensions result in greater precipitable water in the analysis, and therefore exciting more convection. One such set of experiments could involve a rerun of the 3DHYB and H-4DENSV experiments without the use of the weak constraint for unphysical tracers, since it is extremely difficult to force such a constraint to operate the same way in the 3D and 4D paradigms. Another thought may be to look more closely into the 4D nature of the simulated observations, since they were extracted from discrete (3 hourly) time levels from the nature run, using linear interpolation to fill in the gaps.

The interpretation of the forecast impact experiments from the H-4DENSV_COMB analyses is more difficult than was the case for the 3DHYB experiment (relative to the control). The H-4DENSV_COMB analysis does result in superior forecast skill in the northern hemisphere, but seemingly degrades the forecasts in the southern hemisphere and tropics relative to the 3DHYB forecasts. However, the differences in the southern hemisphere are not statistically significant, and the degradation in the tropics is likely related to the increased convective precipitation. Work is already underway to test the H-4DENSV_COMB configuration using real observations to test the forecast impact on the real system.

Although the GSI has a 4DVAR capability within it, the inefficiencies of the inner loop dynamic model make it unaffordable for running fully cycled experiments. Work is ongoing to make the dynamic model more efficient. Once it is ready and can be utilized for such an experiment, it will be interesting to compare the results of the H-4DENSEV experiments with actual 4DVAR (including H-4DVAR_AD). There is also additional work that can be done on the 4DENSEV algorithm itself, such as treating the outer loop more appropriately (along the lines of the ideas proposed in Yang et al. 2012). To this point, the 4DENSEV algorithm has been executed more like a 3DVAR implementation at least in terms of how the outer loop operates. Perhaps some of the noted issues can be improved through a quasi-outer loop (where the nonlinear control model is run as part of the outer loop but the perturbations are held fixed) or full outer loop (where both the control and ensemble forecasts are rerun).

Chapter 4: Scale-Dependent Weighting

4.1 Introduction

It has been demonstrated that ensembles can be effectively used to provide flow-dependent, multivariate estimates of analysis and background error covariances in ensemble (Evensen 1994; Houtekamer and Mitchell 1998; Whitaker and Hamill 2002; Hunt et al. 2004, 2007; and many others) and hybrid (Wang et al. 2007b; Buehner et al. 2010a,b) data assimilation algorithms. Ensembles of assimilations have also been shown to be beneficial in providing flow dependent error variances for variational systems (Kucukkaraca and Fisher 2006; Isaksen et al. 2007; Raynaud et al. 2009, 2011). However, due to computational considerations, compromises are often necessary in terms of the choice of ensemble resolution and/or size. The use of a finite sized ensembles, typically between $\mathcal{O}(10)$ and $\mathcal{O}(10^2)$, introduces the necessity to deal with sampling noise in order to effectively use the covariance (or variance) estimate.

The issue of spurious correlations over large spatial distances has been widely studied and addressed in the framework of localization within ensemble data assimilation algorithms (examples include but are not limited to Houtekamer and Mitchell 1998; Hamill et al. 2001; Bishop and Hodyss 2007; Hunt et al. 2007). Within the hybrid algorithm as described in Chapters 2 and 3, two separate sets of localization are utilized: one within the ensemble assimilation update (as in Whitaker and Hamill 2002) and the other implicitly in the variational-based hybrid through the

specification of the \mathbf{L} covariance matrix for the ensemble control variable (as in equations 2.1 and 3.6).

For the use of ensemble-based variances from limited-sized ensembles within variational algorithms, a variety of spatial filtering techniques have been proposed and successfully implemented (Raynaud et al. 2009; Bonavita et al. 2011). In the objective filtering work proposed by Reynaud et al. (2009), a filtering step is designed that attempts to maximize the signal to noise ratio of ensemble variances. As such, it requires a method for calculating the energy spectrum of the sampling noise in the ensemble-based variances. Several examples demonstrate that sampling noise dominates for high frequency scales of motions, whereas even a fairly small ensemble can capture the true error variance without much contamination from sampling (see. Figs. 2 and 3 in Reynaud et al. 2009; and Fig. 5 in Bonavita et al. 2011).

Within the hybrid algorithm up to this point, the ensemble perturbations have been utilized as-is when being blended with climatological estimates for \mathbf{B} , assuming that sampling error is constant for all variables, locations, and scales of phenomenon. Given that it has been demonstrated for a finite-sized ensemble that the variance estimates are dominated by sampling for small scales, there are clearly alternative methods for using the ensemble-based information more intelligently within the hybrid update, such as through a (piecewise) scale-dependent weighting between the ensemble and static contributions. One such idea was proposed in Lorenc (2003), where it was argued that it might be desirable to have the small scales determined more heavily by the traditional control variable (i.e., through the use of static \mathbf{B}),

while relying more on the ensemble for synoptic (large) scales. For the 3DHYB experiment carried out in Chapter 2, it is evident that the evolved EnKF perturbations used within the hybrid have too little power at the very largest scales, too much power for wave numbers between 25 and 100, and no information at all at high frequencies (Fig. 4.1). Interestingly, there is pretty good agreement between the power spectra within the EnKF approximated background error and the actual background error for the case shown. It is evident even from this simple example that the weighting parameter between the static and ensemble contributions should be expanded beyond a single value for all frequencies.

The rest of this chapter describes an algorithm for applying scale-dependent weights for the ensemble and static contributions to the solution within hybrid data assimilation. Section 4.2 describes further motivation and practical implementation of a scale-dependent scheme into the GSI-based hybrid. The results of including scale-dependent weighting in a 3D-hybrid experiment are provided in section 4.3. This is followed by a brief summary and motivation for future work.

4.2 Scale-dependent weighting in GSI-based hybrid

The extension to include scale-dependent weighting terms within the variational-based hybrid is straightforward through a minor modification of equation 2.1, if \mathbf{x} , \mathbf{B} , and \mathbf{L} are formulated in spectral space (superscript s denotes spectral or wave space):

$$J(\mathbf{x}'^s, \boldsymbol{\alpha}^s) = \boldsymbol{\beta}_f^s \frac{1}{2} (\mathbf{x}'_f^s)^T \mathbf{B}^{-1} (\mathbf{x}'_f^s) + \boldsymbol{\beta}_e^s \sum_{n=1}^N \frac{1}{2} [(\boldsymbol{\alpha}^n)^s]^T \mathbf{L}^{-1} [(\boldsymbol{\alpha}^n)^s] + \frac{1}{2} (\mathbf{H}\mathbf{x}'_t^s - \mathbf{y}')^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}'_t^s - \mathbf{y}') \quad (4.1)$$

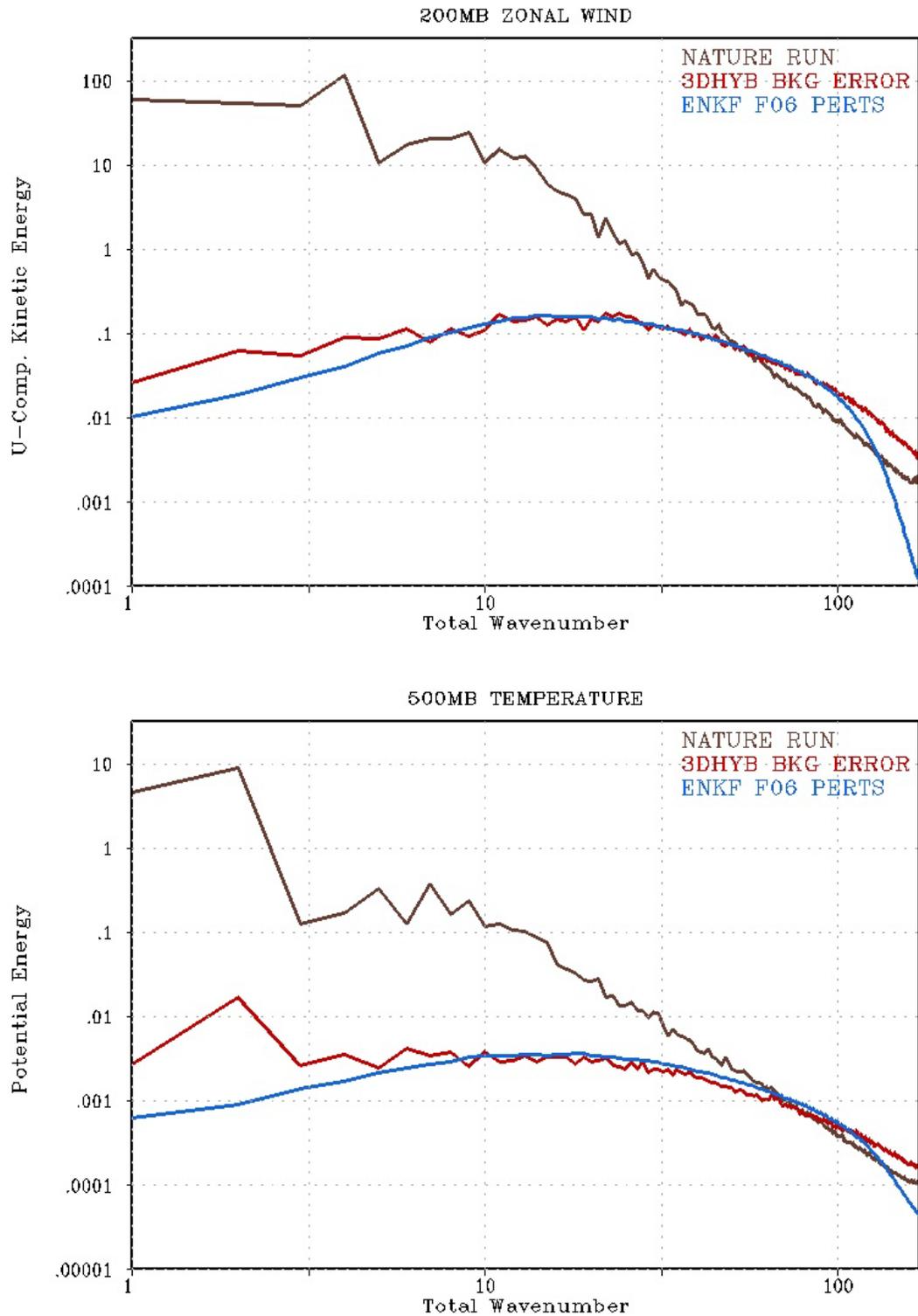


Figure 4.1: Power spectra for zonal wind (top, kinetic energy m^2s^{-2}) and temperature (bottom, potential energy, K^2) for the nature run forecast (brown), 3DHYB 06-hr forecast error (red), and 06-hr EnKF perturbations from the 3DHYB (aqua) averaged from 06 UTC 08 August through 12 UTC 09 August.

where, β_f^s and β_e^s are now functions of total wavenumber. In practice, however, the GSI is solved utilizing a preconditioned conjugate gradient algorithm. For this purpose, we rewrite the spectral cost function using two transformed variables (similar to Wang 2010):

$$\mathbf{z}^s = \beta_f^s \mathbf{B}^{-1} \mathbf{x}'_f{}^s \quad (4.2)$$

$$\left(\mathbf{v}^n\right)^s = \beta_e^s \mathbf{L}^{-1} \left(\boldsymbol{\alpha}^n\right)^s \quad (4.3)$$

resulting in

$$J\left(\mathbf{z}^s, \mathbf{v}^s\right) = \frac{1}{2} \left(\mathbf{x}'_f{}^s\right)^T \mathbf{B} \mathbf{z}^s + \frac{1}{2} \sum_{n=1}^N \left[\left(\boldsymbol{\alpha}^n\right)^s\right]^T \left[\left(\mathbf{v}^n\right)^s\right] + J_o \quad (4.4)$$

where J_o is representative of the observation penalty term. The control variables can then be represented as

$$\mathbf{x}'_f{}^s = \left(\beta_f^s\right)^{-1} \mathbf{B} \mathbf{z}^s \quad (4.5)$$

$$\left(\boldsymbol{\alpha}^n\right)^s = \left(\beta_e^s\right)^{-1} \mathbf{L} \left(\mathbf{v}^n\right)^s. \quad (4.6)$$

Conveniently, the inverses of \mathbf{B} and \mathbf{L} are no longer necessary. Some details regarding the minimization algorithm can be found in the GSI User's Guide (see Chapter 6, available online at <http://www.dtcenter.org/com-GSI/users/index.php>). The scale dependent weighting is implemented directly into equations 4.5 and 4.6. However, to do so requires an additional step since the solution is formulated in physical (not wave) space. By inserting a grid to spectral operator (\mathbf{S}) and its inverse, and operating directly on the \mathbf{z} and \mathbf{v} in physical space (dropping the s) we finally end up with

$$\mathbf{x}'_f = \mathbf{S}^{-1} \left[\left(\beta_f^s\right)^{-1} \mathbf{S} (\mathbf{B} \mathbf{z}) \right] \quad (4.7)$$

$$\boldsymbol{\alpha}^n = \mathbf{S}^{-1} \left[\left(\beta_e^s\right)^{-1} \mathbf{S} (\mathbf{L} \mathbf{v}^n) \right]. \quad (4.8)$$

The calculation of the total analysis increment (\mathbf{x}_t) as the sum of the static and ensemble contributions is exactly as in equation 2.2. Note that this derivation results in the scale dependence being applied directly to the control variable (and not the contribution to the increment) for the ensemble term.

The use of a dual-resolution hybrid update step (where the ensemble is at lower spatial resolution, as in Chapters 2 and 3) introduces an interesting problem. To this point, only a single weighting parameter is utilized to control the relative weights for the entire spectrum, even though the ensemble and static contributions have information over a different range of scales (i.e. the ensemble has zero information for scales below the truncation). For frequencies that are higher than the truncation of the ensemble, the amplitude of the solution is going to be artificially damped since the ensemble perturbations cannot contribute anything meaningful and the static contribution is weighted to account for a pre-specified percentage across all scales.

To demonstrate the problem, a single analysis test case that assimilates real observations (including satellite radiances) within a six hour window valid at 06 UTC 15 July 2010 is performed (identical to the high resolution single analysis experiment in section 3.3.3). Two single cases are run from the same dataset, a 3DVAR control and dual-resolution 3DHYB (T574 deterministic control, 80 member T254 ensemble, with $\beta_f^{-1}=0.25$). For the dual-resolution hybrid configuration, interpolation (and the adjoint thereof) between the lower resolution (ensemble) grid and full analysis grid is necessary to translate the ensemble information to the analysis increment. This

results in spurious information being aliased to the highest of frequencies in the analysis increment for several of the fields (Fig. 4.2). For the 3DHYP increment, the truncation of the ensemble resolution is obvious, with a sharp cutoff at wave number 254, except for the spurious power at high frequencies generated from the interpolation. This further motivates the desire to have a capability to apply scale dependence, as a means for masking issues related to the interpolation of information between the lower resolution ensemble and high resolution hybrid increment.

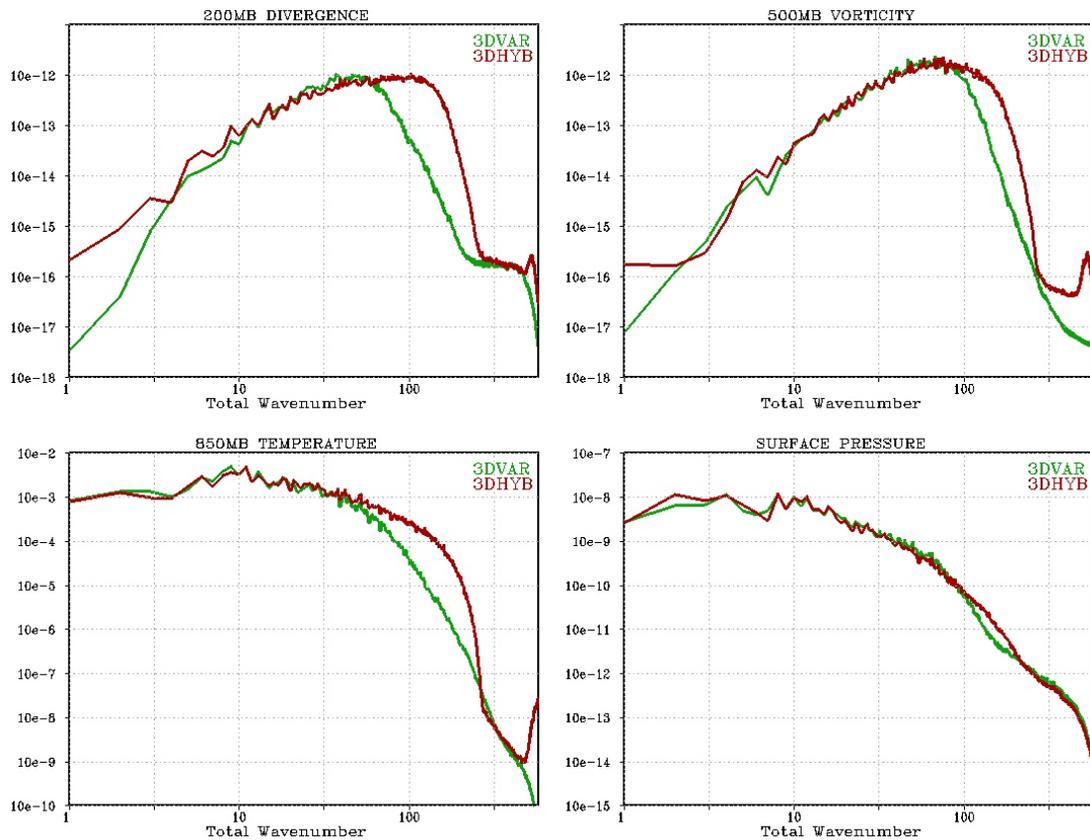


Figure 4.2: Power spectra of the analysis increment from a 3DVAR (green) and dual-resolution 3DHYP (red) experiment valid at 06 UTC 15 July 2010 for 200 hPa divergence (upper left), 500 hPa vorticity (upper right), 850 hPa temperature (lower left) and surface pressure (lower right).

The dual-resolution hybrid necessitates a slight reformulation for the application of the scale dependent weights for the ensemble term. Instead of applying the scale dependent weights to the ensemble control variable, the algorithm is modified:

$$\boldsymbol{\alpha}^n = \mathbf{L}\boldsymbol{\nu}^n \quad (4.9)$$

$$\mathbf{x}'_e = \mathbf{S}^{-1} \left[\left(\boldsymbol{\beta}_e^s \right)^{-1} \mathbf{S} \left(\sum_{n=1}^N \left(\boldsymbol{\alpha}^n \circ \mathbf{x}_e^n \right) \right) \right] \quad (4.10)$$

where \mathbf{x}_e is the solution for the ensemble contribution only. In this formulation, the scale dependent weights are applied to the transformed physical space variables directly for both the ensemble and static contributions to the increment (i.e. equation 4.7 is kept the same). One advantage of this formulation is that it allows for an explicit mask to be applied (i.e. zero) for wavelengths that the ensemble has no information.

4.3 Experimental Results

4.3.1 Single Analysis Test Case

To test that the scale dependent weighting has been implemented correctly, the previously used single analysis dual-resolution 3DHYB experiment is rerun with scale dependent weightings turned on. The weights that are utilized for the two scale dependent tests can be found in Fig. 4.3. In the first test (SD1), the exact same hybrid weights are used for wave numbers 0 through 254 ($\beta_f^{-1}=0.25$, $\beta_e^{-1}=0.75$), however, for wave numbers 255 through 574, the increment comes exclusively from the static contribution ($\beta_f^{-1}=1.0$, $\beta_e^{-1}=0.0$). This has the effect of explicitly zeroing out any

spurious contribution from the part of the spectrum that the ensemble has no information. Also, this allows the analysis to make appropriate changes at the highest frequencies according to the specification of **B**. A second test (SD2) is designed to attempt to exercise the scale-dependent code in a manner consistent with the assumed characteristics of the sampling error within the ensemble. The weighting terms are assigned values to rely heavily on the ensemble for the large scales, not at all for high frequencies, and a compromise or blend for wavenumbers in between (weights can be found in Fig. 4.3).

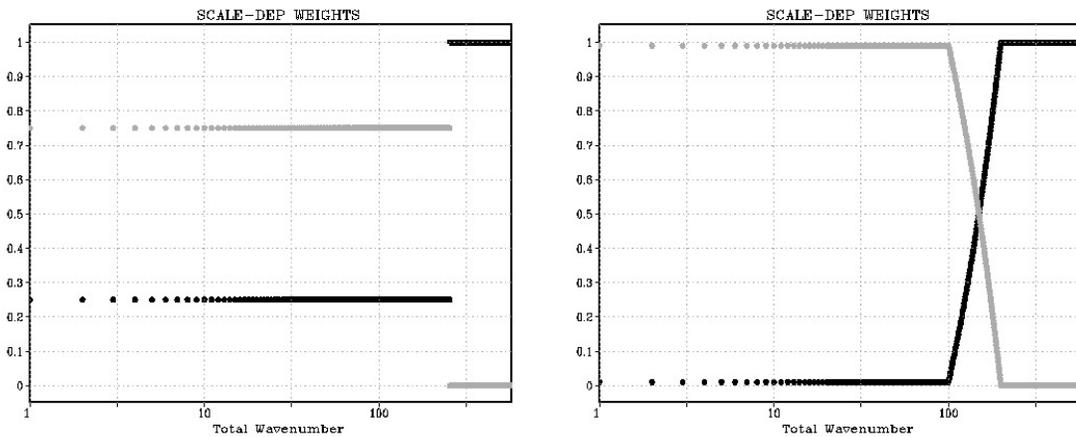


Figure 4.3: The scale-dependent weights, β_f^{-1} (black) and β_e^{-1} (grey) for the SD1 (left) and SD2 (right) single analysis experiments.

By design, both of the experiments SD1 and SD2 are able to mitigate the previously noted aliasing issues at the highest frequencies within the dual-resolution hybrid related to the interpolation between grids (Fig. 4.4). Relative to the 3DHYB, the resultant analysis increment in the SD1 experiment is very similar for all wave numbers within the resolution of the ensemble. The increment from SD2 exhibits

characteristics consistent with the designed weighting parameters, having the amplitude down-weighted from the 3DHYB toward the 3DVAR based solutions for the wave numbers ranging between 100 and 200. Since the power spectra for the surface pressure increment is very similar in the 3DVAR and 3DHYB test cases (Fig. 4.2) due to the tangent linear normal mode constraint filtering, the scale-dependent weights have very little impact (Fig. 4.4).

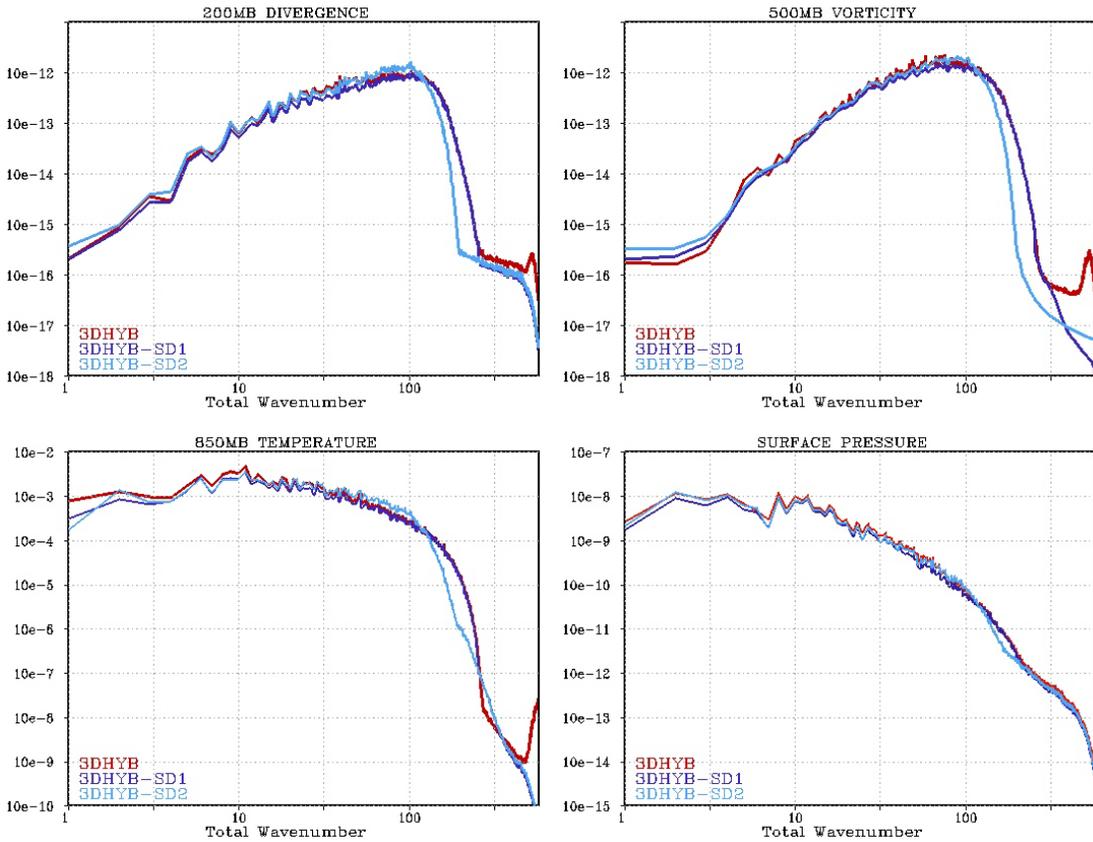


Figure 4.4: As in Fig. 4.2, but for the 3DHYB (red), SD1 (purple), and SD2 (light blue) experiments.

4.3.2 Cycling OSSE Experiment

Just as in previous chapters, a cycling experiment is carried out by assimilating the simulated observations from the Joint OSSE for the same period. The new experiment utilizes the 3DHYB_RS experiment from Chapter 2 as the control run, and tests the impact of adding scale dependence on the quality of analysis. The 3DHYB_RS experiment is chosen as the baseline given the superior performance relative to the original 3DHYB experiment, resulting from tuning the inflation parameters to give better ensemble spread. The choice of scale-dependent weights for this new experiment (3DHYB_RSSD) is illustrated in Fig. 4.5. This configuration results in a system that should behave exactly as the previous hybrid experiment (3DHYB_RS) for all wave numbers smaller than 120 (i.e. 75% ensemble, 25% static contribution). For all wave numbers greater than 170, only the static **B** contributes to the solution. The frequency band in between then contains a transition zone that goes from heavy ensemble to all static **B**. The biggest difference, then, between the 3DHYB_RS and 3DHYB_RSSD should be for the highest frequencies, where the original 3DHYB_RS experiment likely suffered from the aforementioned aliasing problem.

The inclusion of the scale dependent weighting, even though quite conservative, yields a reduction in the analysis errors relative to the 3DHYB_RS experiment (which itself was quite an improvement over the original 3DHYB) for zonal wind and temperature, particularly in the extratropical troposphere (Fig. 4.6). The scale dependence increases the analysis error for stratospheric temperatures and localized regions in the tropics for humidity. This provides evidence that the use of

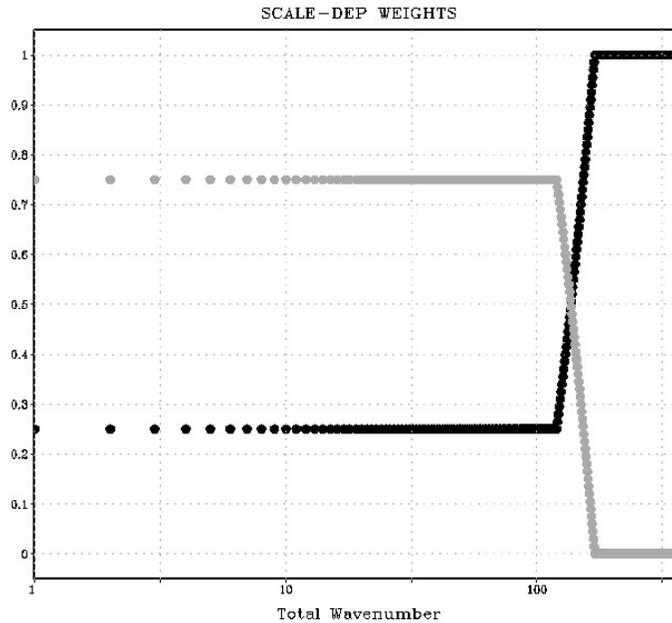


Figure 4.5: As in Fig. 4.2, but for the 3DHYB_RSSD cycling experiment.

the ensemble-based information needs to be made even more flexible beyond the specification of scale dependent global weights, and instead may need to include level- and/or variable-dependent weights. For example, the truncation for which the water vapor and cloud information from the ensemble may be more useful is likely at much smaller wavelengths than for temperature and winds. However, introducing variable specific scale-dependent weights introduces the possibility of destroying some of the balance between variables within the perturbations. In terms of applying level-dependent weights, to utilize only the very largest of scales from the ensemble in the stratosphere, for example, is straightforward to implement and will be one of many areas for future work. Ideally, an algorithm needs to be developed to utilize a fully adaptive, flow-dependent, evolving set of weights between the static and ensemble contributions.

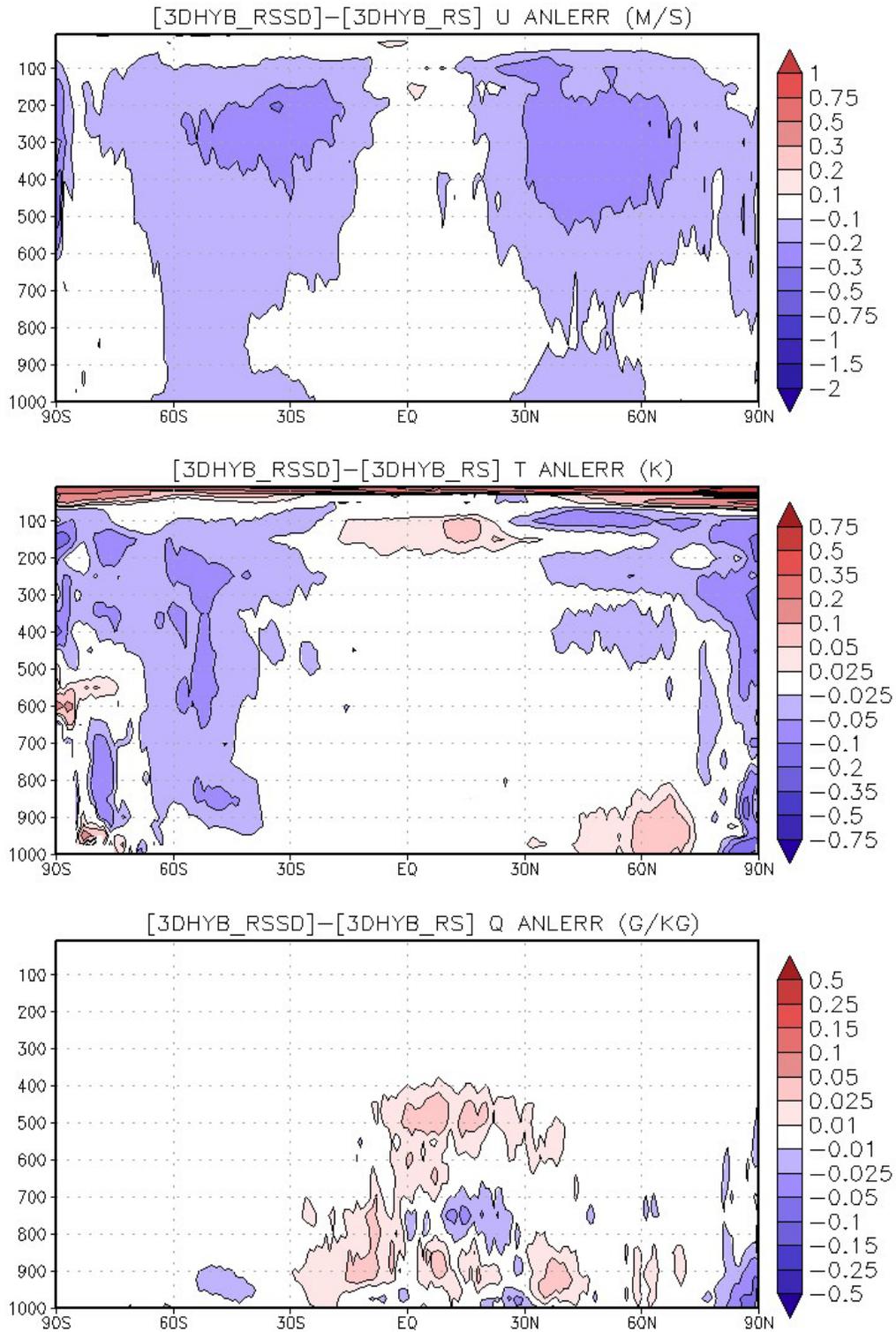


Figure 4.6: Difference in the time-averaged (August) zonal mean standard deviation of the analysis error for the 3DHYB_RSSD experiment relative to 3DHYB_RS for zonal wind (top, m s^{-1}), temperature (middle, K), and specific humidity (bottom, g kg^{-1}).

4.4 Summary and Discussion

Ensemble-based (and hybrid) methods are always going to be sensitive to compromises that need to be made related to resolution and ensemble size. Methods for dealing with sampling error in ensemble and variational algorithms have been explored in some detail through covariance localization and spatial filtering techniques. Along these lines, it remains an area of research to find better ways for utilizing ensemble information within hybrid algorithms. One such idea is to build weighting parameters between the static and ensemble covariance approximations that are a function of scale. The idea is motivated by the fact that the signal to noise ratio for limited size ensembles is much smaller at high frequencies than for small.

A simple method for applying scale dependent β -factors is derived and implemented for use within the GSI-based hybrid scheme. The impact of the scale-dependent scheme is demonstrated through single analysis and cycling experiments. The impact on the power spectrum of the analysis increment is as expected, demonstrating that the scale-dependent weights are implemented correctly. Despite the fact that the initial set of experiments utilized fairly conservative set of scale dependent weights, the net result was a general reduction in analysis errors in the OSSE-based cycled run, particularly in the extratropical troposphere.

The scheme is able to eliminate some artificial features at the highest frequencies that appear within the dual-resolution hybrid as a result of aliasing from the interpolation algorithm between the low and high resolution grids. The use of scale-dependent weights is not the only way to mitigate this issue. As was shown in the 4D context, the use of a weak constraint digital filter can help eliminate such

issues (Fig. 3.3). Alternative (higher order or smoothing) interpolation schemes could also be implemented as an alternative to the brute force scale-dependent weights method.

The increase in analysis errors for several variables and regions highlights the necessity for more work investigating how to make better use of the ensemble information within the hybrid. The weights should likely be expanded to also be level- and/or variable-dependent. However, doing so does increase the number of parameters that need to be specified and makes maintaining a model consistent balance more difficult. Building on previous research on adaptive inflation (Anderson 2009; Miyoshi 2011) and localization (Bishop and Hodyss 2007, 2009a, 2009b, 2011; Anderson 2012) for ensemble assimilation schemes, a method for quantitatively specifying adaptive weights (which could be functions of level, scale, or variable) is a direction for future research.

Chapter 5: Real Observation 3D-Hybrid Experiment

5.1 Introduction

It has been demonstrated in Chapters 2-4 that the inclusion of ensemble-based covariance estimates within the variational scheme in a hybrid manner yields improvements to the quality of analyses and subsequent forecasts within a controlled, OSSE framework where the truth is known. However, there is no guarantee that such a finding will translate directly to the real system where the true atmospheric state is never known. The nature run is produced by a discretized numerical model including parameterized physics resulting in many potential issues. The observations and their (largely unknown) associated errors are simulated based on discrete output from the nature run model, and therefore may not be truly representative of the quasi-continuous set of observations assimilated into the real NWP system.

Forecast impact experiments within the OSSE framework may be especially susceptible to the issue of unrepresentative results of the real system, given that it is likely that the nature run and experiment models are more like each other than they are to the real atmosphere. Although most of the results found in previous chapters were generally consistent, several lingering issues remain. For example, not all of the experiments that yielded a significant reduction in analysis errors were able to produce better forecasts. Recall that in the 3DHYB versus 3DVAR comparison, the results for the Northern Hemispheric forecasts were basically neutral, despite the fact that the 3DHYB exhibited smaller analysis errors than the 3DVAR counterpart. For any OSSE-based findings to be of practical use to an operational center such as

NCEP, they need to be validated to produce the same expected behavior for the real system.

To further validate the findings from Chapter 2, a set of analysis and forecast impact experiments analogous to 3DVAR and 3DHYB are designed and carried out using a real NWP system. Section 5.2 provides details regarding the experiment design. This is followed by a description of the impact of including the 3D-Hybrid on forecast skill relative to 3DVAR. A brief summary and motivation for future work is then presented.

5.2 Experimental configuration

Two experiments analogous to the OSSE-based 3DVAR and 3DHYB are carried out to test the impact of including the 3D-hybrid on forecast skill for a real NWP system. For these experiments, the NCEP GFS model that became operational in May 2011 is again utilized. However, for these experiments, the model is run at the operational T574 spectral resolution utilizing the same 64 hybrid (sigma-pressure) vertical levels that have been utilized for all other experiments.

The same version of the model is utilized for both cycling within the global data assimilation system (GDAS) as well as making the two-week deterministic forecasts. A 3DVAR control using real observations (3DVAR-R) is carried out spanning the period covering 15 July 2010 through 01 November 2010. The first two weeks of the experiment are ignored for spin-up. The data assimilation cycling mimics the procedure utilized in operations, where the long (15 day) forecasts are initialized from a cycle with an earlier data cut-off time (called GFS) to allow for

distribution of products as quickly as possible. A second assimilation is run each cycle with a later data cutoff time (called GDAS). This secondary, late cycle creates all of the necessary components for the following cycle. The analysis for both the GFS and GDAS cycles utilizes a short-term forecast from the previous GDAS cycle as the background (a more detailed explanation is provided in Kleist et al. 2009 a, b). To save on computational resources, the longer deterministic forecasts for the GFS cycle are only initialized for the 00 UTC cycle (see Fig. 5.1 for illustration of experiment components). Additionally, the forecasts are only carried out to 8 days, as opposed to 15 days as is done operationally by NCEP. As in the 3DVAR experiments from previous chapters, the 3DVAR-R utilizes a similar, hybrid-capable version of the GSI. The analysis is performed on the linear Gaussian grid that corresponds to T574, with 1152×578 grid points in the horizontal. Tuned estimates for background and observation errors are taken from the operational GDAS (see Kleist et al. 2009b, with additional information available from <http://www.emc.ncep.noaa.gov/gmb/gdas/>). All operationally available observations are assimilated for the period including rawinsondes, aircraft, surface pressure, wind profilers, ship and buoy, satellite derived AMVs, GPS radio occultation bending angle, as well as microwave and infrared satellite radiances (using the CRTM, Han et al. 2006). A variational bias correction algorithm (Derber and Wu 1998; Dee 2005) is applied to the radiance observations as is done operationally.

A second experiment is then run to test the impact of the hybrid. This new experiment, 3DHYP-R, starts from the same set of initial conditions as the

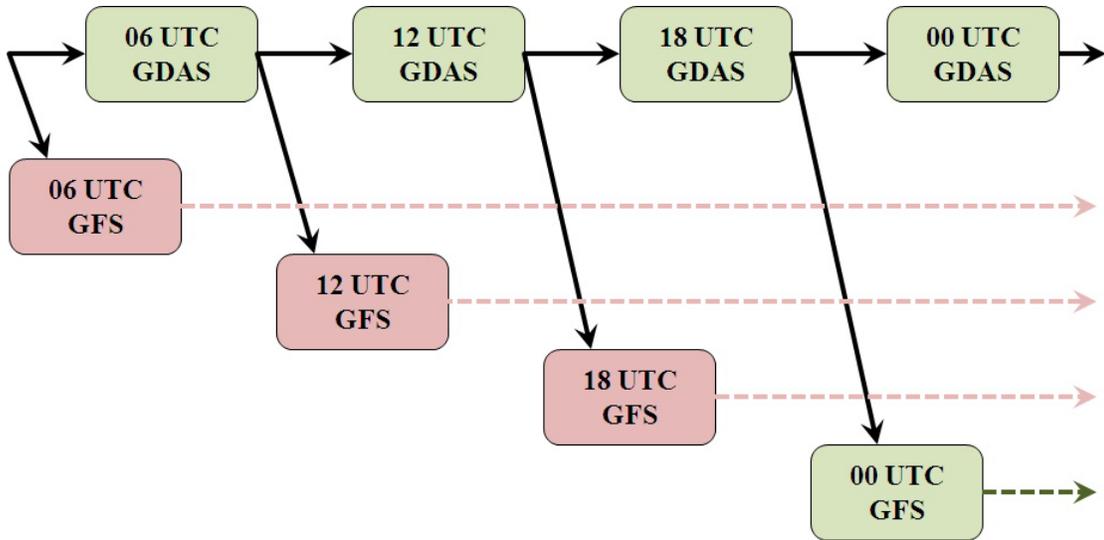


Figure 5.1: Diagram of the cycling procedure used the operational NCEP GFS/GDAS and experiments in Chapter 5. Solid arrows represent the short term (09-h) forecast from the GDAS analysis (two errors per cycle represent a single forecast), whereas dashed arrows represent the 15-day forecast from the early (GFS) cycle. Green (red) colors represent those components that are (not) performed in the 3DVAR-R and 3DHYB-R experiments. Black arrows represent both the higher resolution deterministic and low resolution ensemble forecasts in the case of the hybrid.

3DVAR-R and uses an identical configuration but adds the hybrid capability exactly as described in Chapter 2. However, since the deterministic model is at higher resolution than what was used in the OSSE-based studies, the resolution of the ensemble is similarly increased to T254L64. The ensemble utilized for the hybrid is an 80 member ensemble, updated with the same EnKF (serial square root filter, Whitaker and Hamill 2002) and inflation procedures as previous chapters. The same inflation parameters that were used in the OSSE-based 3DHYB are used here. Like previous experiments, the EnKF utilizes the GSI for all observations operators, and again uses separate data selection and quality control from the high resolution deterministic analysis (relying on the ensemble mean, as well as coarser thinning mesh). Although it has an algorithm to do so internally, the EnKF relies on the GSI-

hybrid for updating the variational bias correction coefficients. This ensures consistency in the treatment of observations between the GSI-hybrid and EnKF. The EnKF-based perturbations are also recentered about the hybrid analysis every cycle (Fig. 2.4). The EnKF is only run as part of the late cutoff (GDAS) cycle. The hybrid analysis for the early (GFS) cycle uses the ensemble forecasts from the previous GDAS cycle (Fig 5.1). The initial ensemble for the 3DHYB-R experiment comes from a set of an interpolated set of operational NCEP Global Ensemble Forecast System (GEFS) initial conditions valid 00 UTC 15 July 2010. The parameter settings for the GSI hybrid are $\beta_e^{-1}=0.75$ and $\beta_f^{-1}=0.25$.

5.3 Forecast Impact

The forecasts that are initialized from the 00 UTC GFS cycle from the 3DVAR-R and 3DHYB-R are evaluated using a variety of metrics utilized by many operational centers. All analysis grid-based metrics (RMSE, AC) are done by evaluating forecasts against analyses from each own's experiment (as opposed to an independent or consensus analysis. All forecasts valid within the time spanning 00 UTC 05 September and 00 UTC 31 October 31 are verified.

Consistent with the results from the OSSE-based experiments discussed in Chapter 2, the addition of the hybrid to the analysis system yields a significant reduction in the forecast error for geopotential height in the extratropics for all lead times and levels (Fig. 5.2). The error reduction shows up almost immediately in the forecasts and grows throughout in both hemispheres. For this particular case, the error reduction is greater in the Southern Hemisphere (which itself exhibits larger

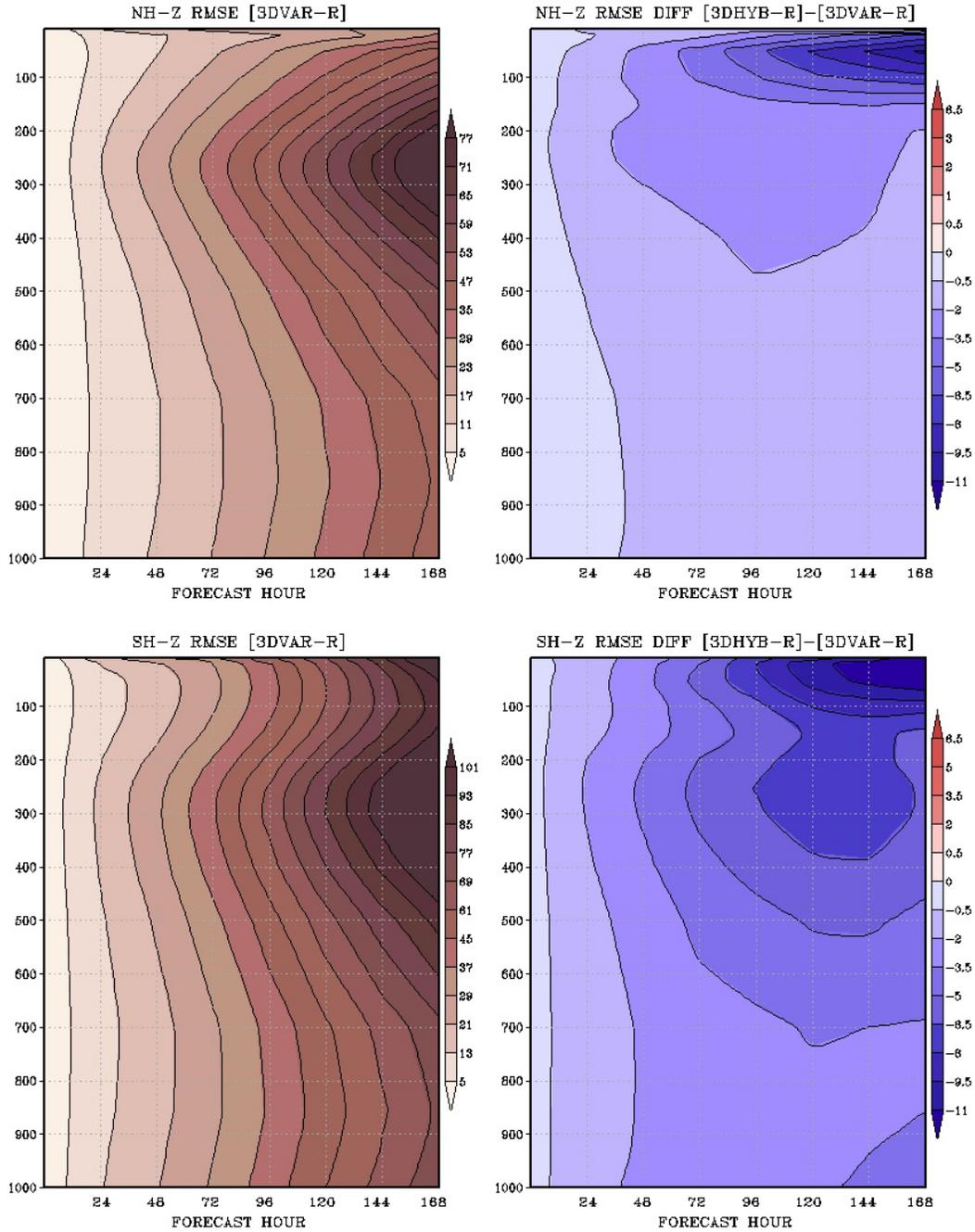


Figure 5.2: Time-averaged root mean square geopotential height errors (m) for forecasts from the 00 UTC analysis in the 3DVAR-R experiment as a function of lead time for the Northern Hemisphere (upper left) and Southern Hemisphere (lower left) verified against the 3DVAR-R analyses for forecasts verifying between 05 August 2010 and 31 October 2010. The difference between two experiments ($[3DHYB-R]-[3DVAR-R]$) for the Northern Hemisphere (upper right) and Southern Hemisphere (lower right) is also plotted, where 3DHYB-R forecasts are verified against 3DHYB-R analyses.

errors than the Northern Hemisphere). In both hemispheres, there is a substantial reduction in geopotential height errors for lead times greater than 4 days in the stratosphere. Focusing on the die off curves for 1000 hPa and 500 hPa, it is verified that most of the improvement is statistically significant at the 95% confidence level (Fig. 5.3). This is particularly true for the Northern Hemisphere, despite the fact that amplitude of the error reduction is smaller than the Southern Hemisphere. For the Southern Hemisphere, the error reduction is statistically significant for short lead times but on the edge of being significant for lead times beyond 4 days.

In terms of anomaly correlation, the results are quite similar to the findings based on geopotential RMSE. The 3DHYB-R is significantly more skillful in the Northern Hemisphere at both 1000 and 500 hPa (Fig. 5.4). For the Southern Hemisphere, the improvement in skill in terms of this metric is even more impressive, and now shown to be significantly significant all the way out to 7 day lead times. This type of consistent, statistically significant improvement is very difficult to achieve for these types of metrics. The results here are a bit in conflict with the findings found in Chapter 2, where the forecast skill improvement from the hybrid was not nearly as impressive in the Northern Hemisphere, attributable to the previously noted issues with using an OSSE for this type of investigation.

As impressive as the impact on extratropical height forecasts is exhibited to be through use of the hybrid, the OSSE-based experimentation provides evidence that the inclusion of ensemble based covariances has a larger impact in the tropics. A comparison of the tropical vector wind RMSE reveals that hybrid does seemingly

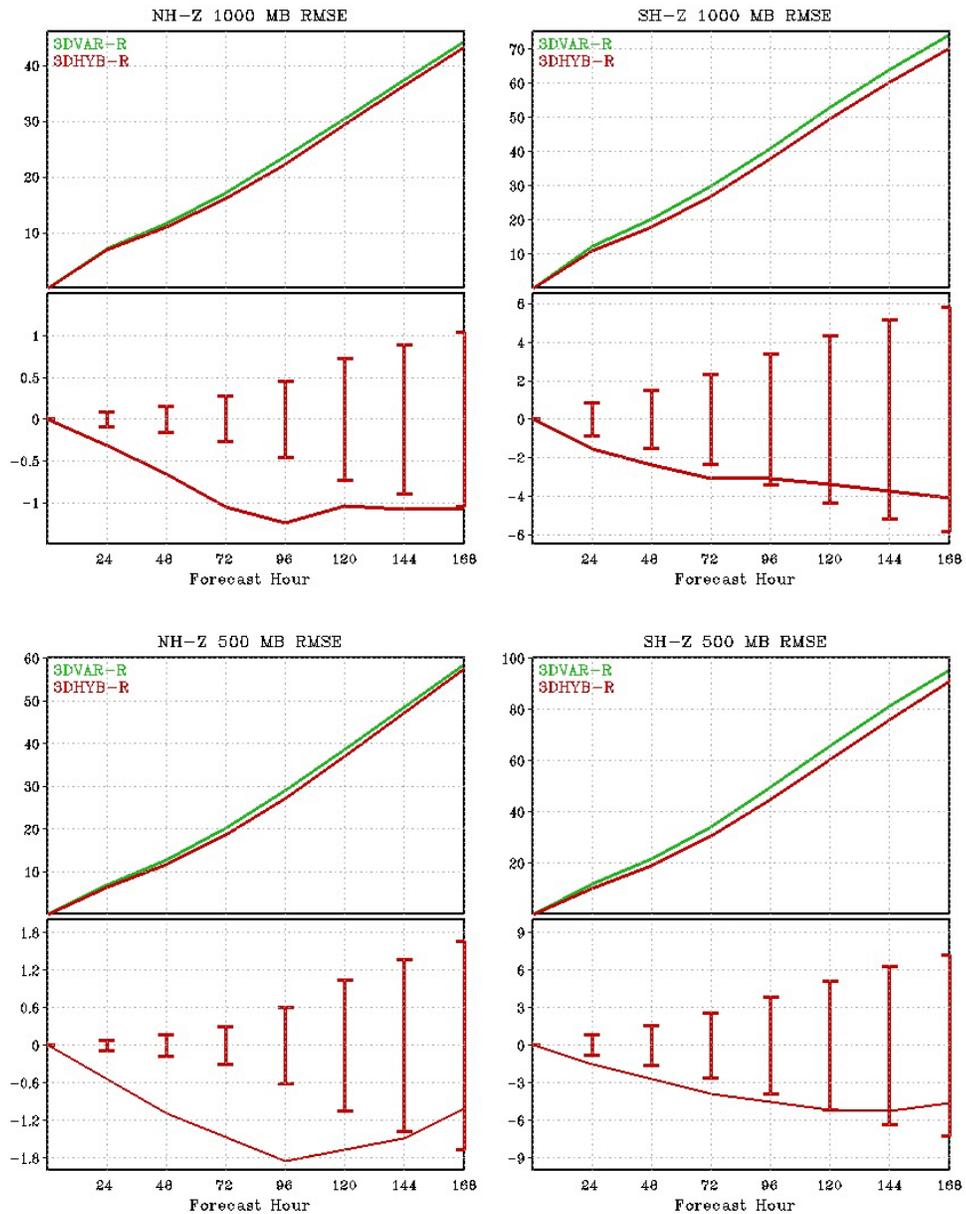


Figure 5.3: The average geopotential height root mean square error by lead time for the 3DVAR-R (green) and 3DHYP-R (red) experiments verifying daily between 05 August 2010 and 31 October 2010 for the northern hemisphere (left) and southern hemisphere (right) at 1000 hPa (top) and 500 hPa (bottom). The difference between the experiments is shown on the bottom section of each panel for the [3DHYP-R]-[3DVAR-R] along with the 95% significance. Forecasts for each experiment were initialized at 00 UTC and verified against analyses from their own experiment.

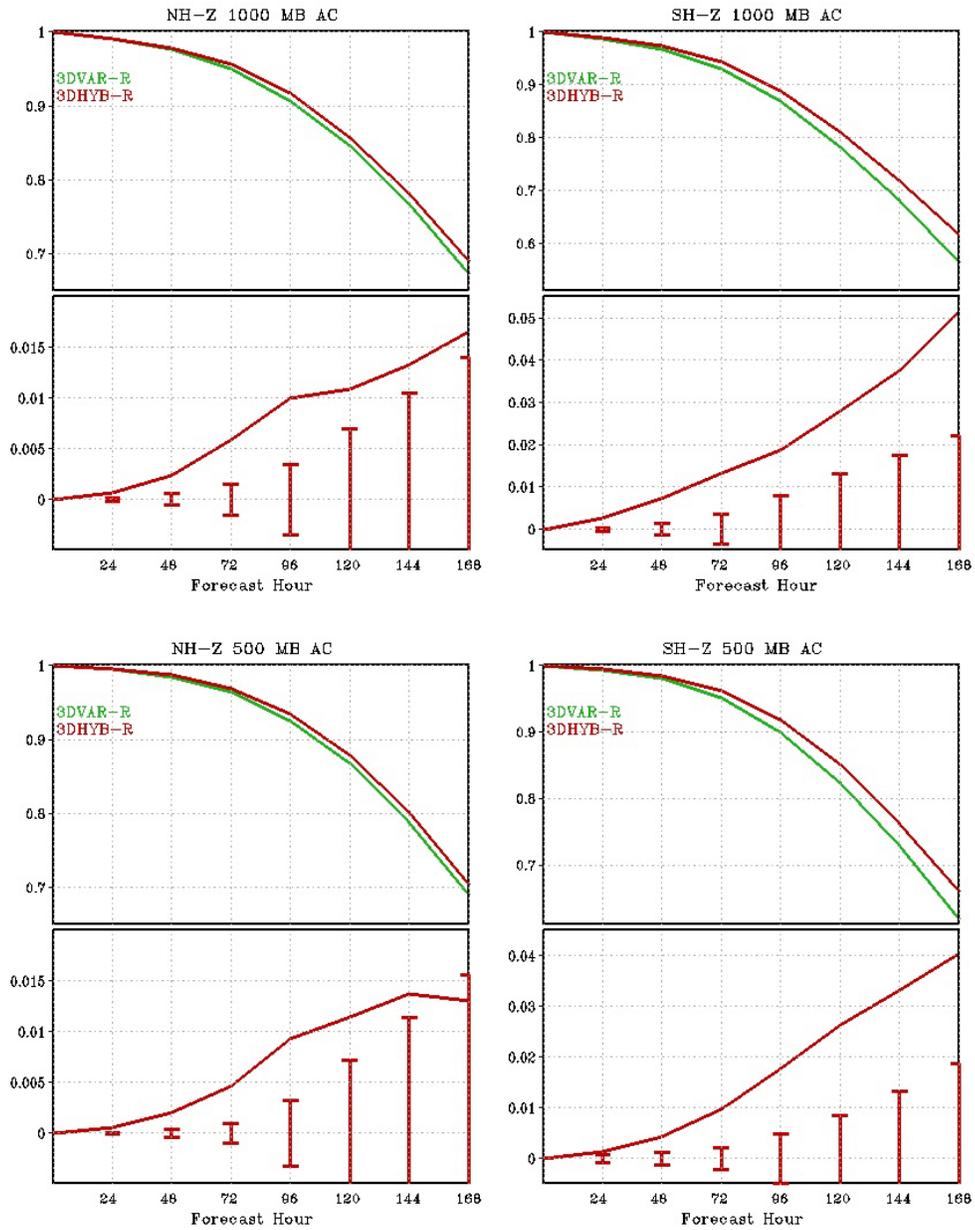


Figure 5.4: As in Fig. 5.4, but for anomaly correlation instead of RMSE.

reduce the forecast error (Fig. 5.5). However, the reduction in wind errors is not as impressive as those found for geopotential heights in the midlatitudes, except for the very substantial reduction noted in the tropical stratosphere. There is evidence in the analysis based verification that the hybrid is actually increasing the errors by a small margin for short lead times in the troposphere. However, a comparison of short-term forecasts to in situ wind observations in the tropics (rawinsonde and aircraft) shows that the 3DHYB-R forecasts are actually superior to the 3DVAR-R forecasts (Fig. 5.6), in direct conflict with the information from the analysis-based RMSE metric. Compared to the observations, both the RMSE and low-speed bias are improved in the 3DHYB-R, throughout the depth of the troposphere.

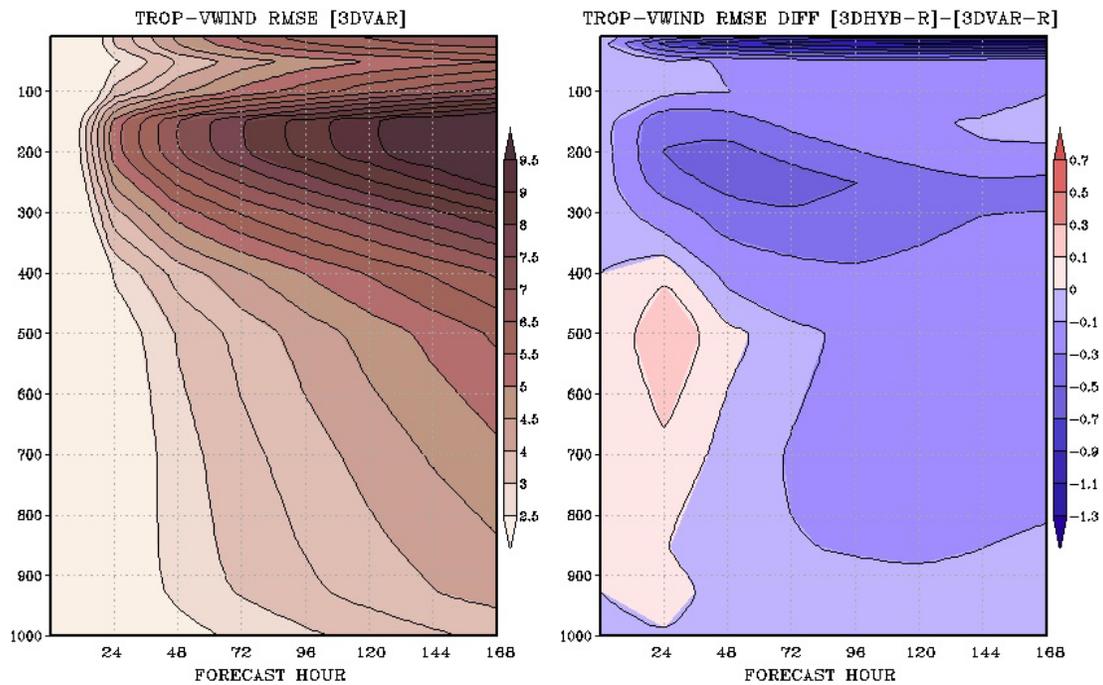


Figure 5.5: As in Fig. 5.2, but for vector wind RMSE instead of geopotential height.

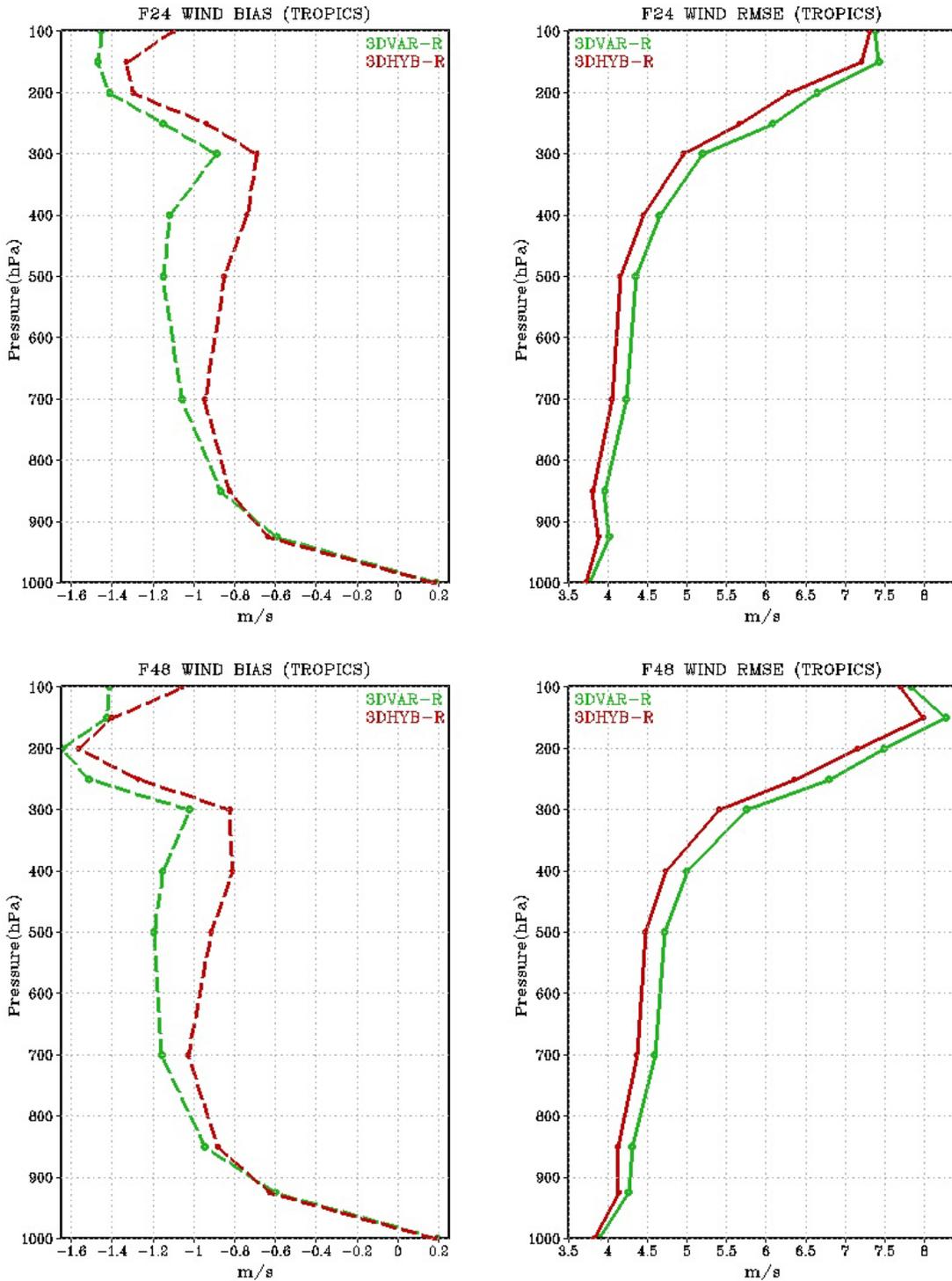


Figure 5.6: Time-averaged wind speed bias (left) and vector wind RMSE (right) of 3DVAR-R (green) and 3DHYB-R (red) forecast with lead times of 24-h (top) and 48-h (bottom) as compared to rawinsonde and aircraft measurements in the tropics for the period covering 05 August 2010 through 31 October 2010.

In terms of the conflicting verification, there is a region around 500 hPa where the analysis based metric shows that the 3DHYB-R forecasts are worse by some 0.2 ms^{-1} , whereas the observation based RMSE is improved in the 3DHYB-R by a similar amplitude for 24-hr forecasts. It is important to note that both metrics are using a “truth” that has errors associated it (be it analysis or observation error). Grid-based RMSE calculations can be misleading, as things can be inadvertently rewarded (penalized) for having a smoother (noisier) verifying analysis. This may have implications for some of the interpretation of the OSSE-based results and worthy of further consideration in future work.

A comparison of tropical cyclone forecasts reveals that the hybrid results in a sizeable reduction in track/position errors, particularly beyond 48-hr lead time (Fig. 5.7). The differences between the 3DVAR-R and 3DHYB-R forecasts are found to be at or near 95% significance for 48-hr through 96-hr by utilizing a paired block bootstrapping algorithm (Hamill et al. 2011a). These results are consistent with the findings for both EnKF and hybrid-based results presented in Hamill et al. (2011b). Given that these are global model forecasts at fairly coarse resolution for tropical cyclone dynamics, much of the gain can likely be attributable to the improved steering forecasts and not necessarily improvements to the initialization of the tropical cyclones themselves. In both the 3DVAR-R and 3DHYB-R experiments, the assimilation of tropical cyclone minimum sea level pressure is utilized (Kleist 2011), which is the one of the predominant mechanisms along with a relocation procedure for initializing the storms in the GFS (most observations are actually screened due to representativeness issues). The intensity bias is very similar between the two

experiments for all lead times including initialization (not shown), indicating that the hybrid is not simply providing better initial storms, thereby further suggesting that the improved steering flow in the tropics is critical.

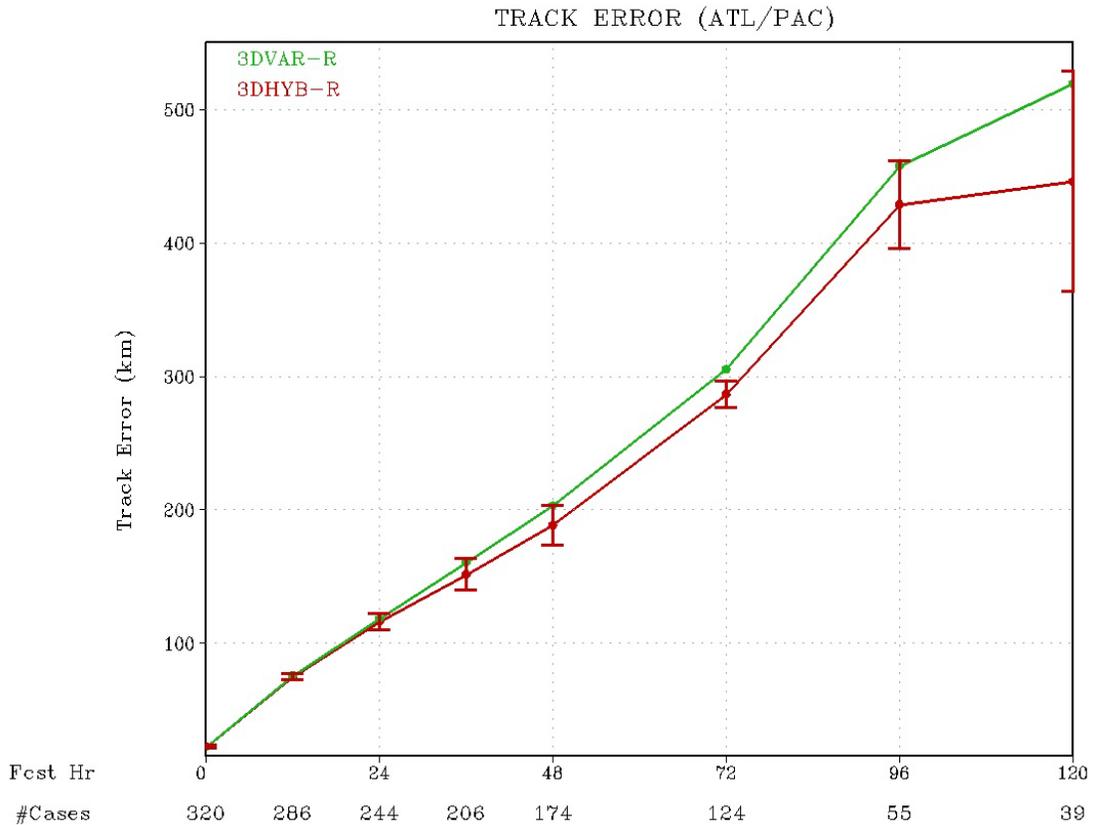


Figure 5.7: The average tropical cyclone track error (km) for 3DVAR-R (green) and 3DHYB-R (red) forecasts for the period covering 10 August 2010 through 31 October 2010. The average track error was computed from a homogeneous sample of cases for storms in the Atlantic, East Pacific, and West Pacific basins. The number of cases for each lead time is identified below each forecast hour. Error bars indicate the 5th and 95th percentile of a re-sampled block bootstrap distribution.

Although only small subset of verification metrics is presented here for the comparison of the 3DVAR-R and 3DHYB-R experiments, it is generally representative of other variables, levels, and metrics. These results are also consistent with the findings of pre-operational trials of the hybrid for the NCEP GFS/GDAS, scheduled to be implemented in spring 2012. A further evaluation of these experiments is ongoing, and will be the subject of another manuscript that will include some follow-on work and interpretation.

5.4 Conclusions

Experiments have been designed and carried out to test the impact of including a 3D-Hybrid scheme using a realistic (operational resolution) prototype system and real observations. The experiment is designed to mimic the operational GDAS/GFS as closely as possible including observational processing, cycle configuration (early cut-off for initialization of deterministic forecast), and post-processing of model output. Unlike the OSSE-based results, there is no known “truth” to validate the analysis error. Instead, focus is placed on validating the quality of forecasts initialized from the hybrid analyses versus those initialized using 3DVAR. The three month experiment covers the 2010 hurricane season.

It is immediately clear the hybrid analyses result in more skillful forecasts than those initialized using 3DVAR. Several metrics are considered, including analysis-based RMSE and AC, fits of forecasts to observations, and tropical cyclone track and intensity errors. The gain in skill appears for most variables, regions, and lead times. For most metrics, the skill difference is found to be statistically

significant at the 95% level. Although not focused on in this chapter, a sizeable reduction in errors in the stratospheric forecasts is also observed. A more thorough evaluation is ongoing and will be the subject of a future manuscript. Similar to the systematic OSSE-based runs discussed in Chapters 2 and 3, more sensitivity experiments are underway (though using real observations) in an attempt to ascertain which aspects of the hybrid yield the significant gains (background error, physical space localization, dynamic constraint terms in the variational framework), relative to standalone 3DVAR and 3D-EnKF counterparts.

Some of the results found in the real observation experiments are in conflict with the results from the OSSE-based experiments. Consider the forecast impact experiments carried out in Chapter 2 (3DVAR versus 3DHYB) with those carried out in this chapter (3DVAR-R versus 3DHYB-R). The experiments were designed in such a fashion as to be analogous (running through a similar season/time period, using the same forecast model, parameter settings, with the OSSE-based experiments using slightly lower resolution). The forecast impact results from the real observation experiments are much more impressive than the results found from the OSSE-based experiments, especially when considering the gain in skill from the hybrid in Northern Hemispheric geopotential height forecasts. It is still not entirely clear why this is the case, though it is possible that one of the main issues is related to the fact that the GFS is more like the nature run model than either is to the real atmosphere.

Given the discrepancy found between the OSSE-based and real observation forecast impact experiments, it will be interesting to see what happens when the H-4DENSV_COMB configuration is tested using real observations. There is clear

evidence that the analyses in such a configuration are superior to 3DHYB, though as documented in Chapter 3, this did not translate clearly to improved forecast skill (except in the Northern Hemisphere). Experiments of this sort are already underway, and NCEP is considering such a paradigm for a prototype GDAS configuration in lieu of traditional 4DVAR (with adjoint), given resource (computational and manpower) issues.

Chapter 6: Summary and Conclusions

Hybrid data assimilation algorithms (Hamill and Snyder 2000; Lorenc 2003; Zupanski 2005) have become popular, in a concerted effort to attempt to combine the advantages of variational- and ensemble-based methods while at the same time trying to minimize their weaknesses. Most methods, to this point, have tried to incorporate ensemble-based covariance estimates into variational algorithms. Usually, this ensemble based information is combined with some full-rank, climatological estimate (i.e. the standard \mathbf{B} utilized in variational algorithms). Many studies have found that hybrid based-algorithms improve upon the variational or ensemble-based algorithms on their own (Hamill and Snyder 2000; Wang et al. 2007b, Buehner et al. 2010 a,b).

Hybrid algorithms have many potential advantages over their stand alone counterparts. Ensemble-based covariance estimates contain fully flow-dependent and multivariate information, though typically suffer from sampling issues. Variational-based hybrid algorithms can easily apply localization in physical space, which can be particularly advantageous for the assimilation of observations that are integrated quantities such as satellite radiances (Campbell et al. 2010). Dynamic constraints are quite easy to implement within variational solvers, and as such, hybrid algorithms can help improve balance and noise issues for ensemble-based analysis increments. Lastly, it has been suggested that hybrid algorithms may be particularly useful for small ensemble sizes (Wang et al. 2007b, 2009).

The purpose of this research was to investigate the use of a hybrid data assimilation algorithm for use with the NCEP GFS model. The development and

testing of a prototype 3D variational-ensemble hybrid system for use with the operational system was already underway at the time this research began. This work was designed to learn more about the hybrid system already under development and to expand upon it with further innovation, testing under various configurations all with eventual implementation into the operational NCEP system in mind. There is a lot of research currently ongoing in regards to hybrid data assimilation within the earth science community (particularly for atmospheric and oceanic applications), but the work presented here has several novel aspects to it:

- Focus on systems and algorithmic developments that can eventually be transitioned to an operational framework. Experiments are designed to run at (or near) operational resolution, using a dual-resolution paradigm for the ensemble component.
- Use of a hybridized 4DENS_V algorithm, through application of a time-invariant static **B** supplement.
- Experimentation with various dynamical constraints within the hybrid-4DENS_V framework, including use of a weak constraint digital filter, incremental normal mode constraint, and combinations thereof.
- Proposal and experimentation with scale dependence within the hybrid “weights” (i.e. β -terms in the cost function).

6.1 OSSE-based Hybrid Experiments

To answer some of the questions proposed regarding the use of a hybrid algorithm, a series of OSSE-based experiments are carried out. A realistic set of

simulated observations (including reasonable errors) from the Joint OSSE nature run were kindly provided for use in such experiments by GMAO. These simulated observations, as well as the truth (nature run) provide the basis for a whole suite of experiments to test various aspects of the hybrid algorithm. An initial 3DVAR experiment was carried out using a T382L64 version of the NCEP GFS model, to validate the usability of the simulated observations and evaluate how realistic the experimental configuration was. The OSSE-based 3DVAR innovation statistics were found to be very similar to those extracted from a real observation experiment. This finding, combined with previous work on the validation of the Joint OSSE nature run (Reale et al. 2007; McCarty et al. 2012) and simulated observations (Errico et al. 2007; Errico et al. 2012) provided confidence in the experiment design.

A 3DHYB experiment was then carried out in the OSSE framework using a dual-resolution configuration, designed to mimic the hybrid prototype that was under development for implementation for use with the NCEP GDAS/GFS. It was found that the hybrid algorithm resulted in analyses that were generally superior to those from 3DVAR, especially in the tropics. For some variables, the background errors within the hybrid were smaller than the analysis errors from 3DVAR. It was noted, however, that the hybrid seemed to actually increase the analysis errors in small regions for temperature and wind in the southern latitude upper troposphere. A follow-on hybrid experiment with reduced inflation (3DHYB_RS) demonstrated that a hybrid system that used an ensemble with spread more representative of the real background error amplitude resulted in significantly improved analyses. This result was a bit surprising, at least in terms of the amplitude of the error reduction through a

simple tuning of inflation parameters within the ensemble update. This further motivates the need to develop better means for controlling sampling issues within the EnKF (through use adaptive inflation methods or hybridization of the EnKF itself).

The GSI-based hybrid was then extended to include 4D-ensemble perturbations, i.e. 4DENSEV (just as in Lorenc 2003, Buehner et al. 2010a). The use of 4D-ensemble based methods in place of traditional 4DVAR has become popular in the research community, given the fact that the method does not require the development of an adjoint model (Liu et al. 2008; Tian et al. 2008, 2011; Buehner et al. 2010a). The 4DENSEV algorithm is very similar to a 4D-LETKF (Hunt et al. 2007), except that the weights are solved for within the variational framework. A hybrid variant of the 4DENSEV algorithm was then proposed, that simply uses a time-invariant \mathbf{B} estimate to supplement the 4D-ensemble based increment. Various configurations of dynamic constraints are also proposed for use within the H-4DENSEV algorithm, including the tangent linear normal mode constraint, a weak constraint digital filter, and combined normal mode-digital filter constraint. The combined constraint is motivated by the exorbitant cost of running the normal mode constraint over all time levels, and simply applies it to the center of the window only (allow for the weak constraint digital filter to act upon the other time levels, “propagating” the information).

The OSSE based experiments demonstrated that the inclusion of 4D ensemble perturbations improves upon the use of 3D perturbations, and then, the addition of a static \mathbf{B} contribution improves the solution even further. Although the normal mode constraint and weak constraint digital filter seem to have a small additional impact on

the system, both do contribute positively. If such experiments had been done without the hybrid component, the impact may have been larger as static **B** does seem to help improve the balance to the solution (i.e. reducing the necessity for such constraints to begin with). Interestingly, the experiment with the combined constraint resulted in improvements that were larger than the combined impact of each of the individual constraint experiments. Of all the experiments tested in Chapter 3, the combined H-4DENS_V experiment (with hybrid) not only yields the best results, but does so without extreme additional computational cost over a 3D-hybrid paradigm.

Lastly, a method for using scale dependent weighting between the ensemble and static contributions to the solution within the GSI-based hybrid algorithm was proposed. From a practical standpoint, such an algorithm is not trivial to implement and does not come without cost. The impact of the algorithm on the power spectrum of the analysis increment was as expected (depending on choice of scale-dependent weights). Within the dual-resolution hybrid, interesting features were observed for the highest frequency part of the spectrum, an artifact of interpolation errors and aliasing between the ensemble and analysis grids. The scale dependent weighting was demonstrated to be one way to clean up such noise. Preliminary results from a cycling experiment were somewhat encouraging, with a conservative set of scale-dependent weights resulting in a reduction in analysis error for extratropical wind and temperature.

6.2 Prospects for Real NWP System

For all of the OSSE-based hybrid runs, forecast impact experiments were attempted with some limited success. The 3DHYB forecasts were demonstrated to be more skillful than the 3DVAR forecasts for many variables and levels, though the Northern Hemispheric scores were unimpressive despite the fact that the analysis errors were shown to be smaller. The forecasts from the 4D-experiments were even more difficult to interpret. Some improvement relative to the 3DHYB run was found, but generally the extension to 4D analyses resulted in larger forecasts errors (again, despite the smaller analysis errors). There may be an inherent limitation within OSSE studies in terms of usefulness for doing forecast impact evaluation.

A comparison of the OSSE-based 3D forecast impact results with those from experiments using real observations further raised this possibility. In the real observation experiments, the hybrid forecasts were clearly superior to the 3DVAR-based forecasts for nearly all metrics explored. This is despite the fact that the OSSE-based experiments were designed to mimic as closely as possible those for the real observation experiments (including similar versions of the model, parameter choices, etc.). Given this discrepancy, it will be interesting to see if the results from an H-4DENS experiment using real observations results in improvements (as the reduction in analysis error within the OSSE framework suggested) or not (as some of the forecast impact results seem to suggest).

Algorithms such as the 4DENS (and hybrid variants) are quite enticing for an operational center for many reasons. Such an algorithm would allow one to adopt a non-adjoint extension to 4D without having to overhaul existing infrastructure and

familiarity (say if one wanted to replace a 3DVAR with a 4D-EnKF). Further, with a 3D-hybrid set to become operational as part of the GDAS in spring 2012, an extension to include a 4D ensemble is the next logical step (again, given the lack of adjoint model and sufficient computational resources). The extension to 4D is in fact quite inexpensive (not free) at least relative to a true 4DVAR, without having to totally sacrifice resolution. Although a full (operational) resolution ensemble is not feasible, the static **B** contribution to the H-4DENSV (and 3DHYB for that matter) helps to fill the void for the parts of the spectrum that the ensemble simply cannot represent. A real observation follow-on experiment to those carried out in Chapter 5 is already underway to test the impact of an H-4DENSV_COMB algorithm in the real system.

6.3 Future Work

The use of scale-dependent weighting is not quite as mature as the 4DENSV hybrid extensions. In fact, as implemented, the application of scale dependent weights is more expensive than the 4D extension to the hybrid. In addition to the computational issues, further work is needed to better understand quantitatively just how to set the weights properly. This may include a further expansion to level-dependent specification of the weights. Beyond a simple scale-dependence, it seems reasonable that one should be able to develop fully flow-dependent, adaptive methods for deciding on the choice of weights between the static and ensemble contributions. Such a method could be analogous to methods that have already been developed within the ensemble data assimilation community for adaptive localization (Bishop

and Hodyss 2007, 2009a, 2009b, 2011; Anderson 2012) and inflation (Anderson 2009; Miyoshi 2011).

In all of the 4DENSV experiments carried out as part of this research project, the assumption has been made that the control variable (or ensemble weights) are valid throughout the assimilation window. For many regions and applications, this seems to be a reasonable assumption. However, there may be situations for which one may want to have the capability to apply time localization within the assimilation. This would be especially true if one were interested in extending the assimilation window to something longer than several hours. This may also be true for specific applications such as storm scale or hurricane initialization. To perform such time localization within the GSI-based hybrid as formulated previously, the control variable needs to be expanded to be a function of time level, so the cost function becomes:

$$J(\mathbf{x}'_f, \boldsymbol{\alpha}) = \beta_f \frac{1}{2} (\mathbf{x}'_f)^T \mathbf{B}_f^{-1} (\mathbf{x}'_f) + \frac{1}{2} \frac{\beta_e}{K} \sum_{n=1}^N \sum_{k=1}^K (\boldsymbol{\alpha}_k^n)^T \mathbf{L}^{-1} (\boldsymbol{\alpha}_k^n) + \frac{1}{2} \sum_{k=1}^K (\mathbf{H}_k \mathbf{x}'_k - \mathbf{y}'_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{x}'_k - \mathbf{y}'_k). \quad (6.1)$$

Now, some method for passing information between time levels, perhaps through a simple weighting procedure (w) is included

$$\boldsymbol{\alpha}_k^n = \sum_{j=1}^K w_k^j \boldsymbol{\alpha}_j^n \quad (6.2)$$

and the increment at each time level is prescribed by

$$\mathbf{x}'_k = \mathbf{x}'_f + \sum_{n=1}^N \left(\boldsymbol{\alpha}_k^n \circ (\mathbf{x}_e)_k^n \right). \quad (6.3)$$

Here, the same spatial correlation matrix, \mathbf{L} , is used for the control variable over all time levels. For the implementation in chapter 3, the w is simply set to one for all values of j and k , so the control variable at each grid point is assumed to be constant for all of the time levels. How precisely to prescribe equation 6.2, and specifically assign values for w (or some other method for information propagation/control between time levels) is unclear.

The current formulation of the hybrid 4DENS V only utilizes a time-invariant contribution to the increment from a static \mathbf{B} . However, there is the possibility of adding some time information to the “fixed” contribution to the increment. One such idea would be to utilize a first-order time extrapolation to the observation (FOTO) algorithm, taking advantage of the time tendencies of the state variables that are computed as part of the incremental normal mode constraint (Rančić et al. 2008). Such an extension would provide some synergy (at least in terms of geographic displacement) between the static and ensemble based increments, for those observations taken away from the center of the analysis window.

Lastly, several follow-on experiments are already planned within the OSSE framework. First and foremost, an experiment utilizing 4DVAR (with adjoint) will be carried out for comparison with the 4DENS V results. This will then be followed by a hybrid 4DVAR (with adjoint) experiment, to see if the use of the linear model is beneficial relative to the straight use of the nonlinear 4D ensemble within the 4DENS V algorithm. These experiments are still waiting for further development of a usable tangent linear and adjoint model for use within the inner loop for 4DVAR. Additionally, to help further understand the moisture related issues that were noted in

Chapter 3 when going to 4D, some of the experiments will be rerun but exclude the weak constraint on unphysical moisture.

Bibliography

- Andersson, E., and M. Masutani, 2010: Collaboration on observing system simulation experiments (Joint OSSE), ECMWF Newsletter No. 123, Spring 2010, 14-16. [Available online at <http://www.ecmwf.int/publications/newsletters/pdf/123.pdf>]
- Anderson, J. L., 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, **61**, 72-83.
- Anderson, J. L., 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, in press. [Available online at <http://dx.doi.org/10.1175/MWR-D-11-00013.1>]
- Bannister, R. N., 2008a: A review of forecast error covariance statistics in atmospheric variational data assimilation. Part I: Characteristics and measurements of forecast error covariances. *Quart. J. Roy. Meteor. Soc.*, **134**, 1951-1970.
- Bannister, R. N., 2008b: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quart. J. Roy. Meteor. Soc.*, **134**, 1971-1996.
- Bonavita, M., L. Raynaud, and L. Isaksen, 2011: Estimating background-error variances with the ECMWF ensemble of data assimilations system: some effects of ensemble size and day-to-day variability. *Quart. J. Roy. Meteor. Soc.*, **137**, 423-434.
- Bishop, C. H. and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 2029-2044.
- Bishop, C. H. and D. Hodyss, 2009a: Ensemble covariances adaptively localized with ECO-RAP. Part 1: tests on simple error models. *Tellus*, **61A**, 84-96.
- Bishop, C. H. and D. Hodyss, 2009b: Ensemble covariances adaptively localized with ECO-RAP. Part 2: a strategy for the atmosphere. *Tellus*, **61A**, 97-111.
- Bishop, C. H., and D. Hodyss, 2011: Adaptive Ensemble Covariance Localization in Ensemble 4D-VAR State Estimation. *Mon. Wea. Rev.*, **139**, 1241-1255
- 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. *Mon. Wea. Rev.*, **138**, 1550-1566.

- 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. *Mon. Wea. Rev.*, **138**, 1567-1586.
- Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138**, 282-290.
- Clayton, A. M., A. C. Lorenc and D. M. Barker, 2012: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quart. J. Roy. Meteor. Soc.*, submitted.
- Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367-1388.
- Dee, D. P., 2005: Bias and data assimilation, *Quart. J. Roy. Meteor. Soc.* **613**, 3323-3343.
- Derber, J. C., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.*, **117**, 2437-2446.
- Derber, J. C., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333-1347.
- Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287-2299.
- Errico, R.M., R. Yang, M. Masutani, M., and J. Woollen, 2007: Estimation of some characteristics of analysis error inferred from an observation system simulation experiment. *Meteorologische Zeitschrift*, **16**, 695-708.
- Errico, R.M., R. Yang, N. Privé, K.-S. Tai, R. Todling, M.E. Sienkiewicz, and J. Guo, 2012: Development and validation of observing system simulation experiments at NASA's Global Modeling and Assimilation Office. *Quart. J. Roy. Meteor. Soc.*, submitted.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99(C5)**, 10143-10162.
- Gaspari, G, and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723-757.

- Gauthier, P., and J.-N. Thépaut, 2001: Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.*, **128**, 2905-2919.
- Gauthier, P., M. Tanguay, S. Laroche, and S. Pellerin, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.*, **135**, 2339–2364.
- Gustafsson, N., 1993: Use of a digital filter as weak constraint in variational data assimilation. *Proc. Workshop on Variational Assimilation, with Special Emphasis on Three-Dimensional Aspects*, Reading, United Kingdom, ECMWF, 327–338.
- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776-2789.
- Hamill, T. M., J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011a: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668–688.
- Hamill, T.M., J. S. Whitaker, D. T. Kleist, M. Fiorino, S. G. Benjamin, 2011b: Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Mon. Wea. Rev.*, **139**, 3243–3247.
- Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM) – Version 1, *NOAA Technical Report #122*, 33pp.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.
- Hunt, B., E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Sauer, I. Szunyogh, J. A. Yorke, and A. V. Zimin, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273-277.
- Hunt, B. R., E. Kostelich, and I. Szunyogh 2007: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter, *Physica D*, **230**, 112-126.
- Isaksen I., M. Fisher, and J. Berner, 2007: Use of analysis ensembles in estimating flow-dependent background error variance. *Proc. ECMWF Workshop on Flow-dependent aspects of data assimilation*, Reading, United Kingdom, ECMWF, 65-86. [Available online at <http://www.ecmwf.int/publications/>].

- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009a: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137**, 1046-1060.
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. J. Lord, 2009b: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691-1705.
- Kleist, D. T., 2011: Assimilation of tropical cyclone advisory minimum sea level pressure in the NCEP Global Data Assimilation System. *Wea. Forecasting*, **26**, 1085–1091.
- Kucukkaraca E. and M. Fisher, 2006: Use of analysis ensembles in estimating flow-dependent background error variances. ECMWF Tech. Memorandum 492. [Available online at <http://www.ecmwf.int/publications/>].
- Lawless, A. S., 2010: A note on the analysis error associated with 3D-FGAT. *Quart. J. Roy. Meteor. Soc.*, **136**: 1094–1098.
- Lewis, J. M. and J. C. Derber, 1985: The use of adjoints equations to solve a variational adjustment problem with advective constraints. *Tellus*. **37A**, 309–322.
- Liu C., Q. Xiao, and B. Wang, 2008: An ensemble-based four dimensional variational data assimilation scheme. Part I: technique formulation and preliminary test. *Mon. Wea. Rev.*, **136**, 3363-3373.
- Lorenc, A. C., and Coauthors, 2000: The Met. Office global 3-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **126**, 2991-3012.
- Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183-3203.
- Lorenc, A. C., and F. Rawlins, 2005: Why does 4D-Var beat 3D-Var?. *Quart. J. Roy. Meteor. Soc.*, **131**, 3247-3257.
- Lynch, P., and X.Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019-1034.
- Masutani, M., and Coauthors, 2007: Progress in Joint OSSEs. AMS preprint for 18th conference on NWP, Park City, UT, 25-29 June 2007. [Available online at <http://ams.confex.com/ams/pdfpapers/124080.pdf>].

- Masutani, M., and Coauthors, 2010, Observing system simulation experiments at the National Centers for Environmental Prediction, *J. Geophys. Res.*, **115**, D07101, doi:10.1029/2009JD01258.
- McCarty, W., R. M. Errico, and R. Gelaro, 2012: Cloud coverage in the Joint OSSE nature run. *Mon. Wea. Rev.*, in press. [Available online at <http://dx.doi.org/10.1175/MWR-D-11-00131.1>].
- Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535.
- Pannekoucke, O., L. Berre, and G. Desroziers, 2008: Background error correlation length-scale estimates and their sampling statistics. *Quart. J. Roy. Meteor. Soc.*, **134**, 497–511.
- Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Polavarapu, S., M. Tanguay, L. Fillion, 2000: Four-dimensional variational data assimilation with digital filter initialization. *Mon. Wea. Rev.*, **128**, 2491–2510.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.
- Rabier, F., J.-N. Thépaut, and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quart. J. Roy. Meteor. Soc.*, **124**: 1861–1887.
- Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170.
- Rančić, M., J. C. Derber, D. Parrish, R. Treadon, and D. T. Kleist, 2008: The development of the first-order time extrapolation to the observation (FOTO) method and its application in the NCEP global data assimilation system. *Proc. 12th Symp. On Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, New Orleans, LA, Amer.

Meteor. Soc., J6.1. [Available online at <http://ams.confex.com/ams/pdfpapers/131816.pdf>.]

- Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global 4-Dimensional data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362.
- Raynaud, L., L. Berre, and G. Desroziers, 2009: Objective filtering of ensemble-based background error-variances. *Quart. J. Roy. Meteor. Soc.*, **135**, 1177-1199.
- Raynaud, L., L. Berre, and G. Desroziers, 2011: An extended specification of flow-dependent background –error variances in the Meteo-France global 4D-Var system. *Quart. J. Roy. Meteor. Soc.*, **137**, 607-619.
- Reale O., J. Terry, M. Masutani, E. Andersson, L.-P. Riishojgaard, J.-C. Jusem, 2007: Preliminary evaluation of the European Centre for Medium-range Weather Forecasts (ECMWF) nature run over the tropical Atlantic and African monsoon region. *Geophys. Res. Lett.*, **34**, L22810, doi:10.1029/2007GL031640.
- Rosmond, T., and L. Xu, 2006: Development of NAVDAS-AR: nonlinear formulation and outer loop tests. *Tellus*, **58A**, 45-59.
- Saha, S., and Coauthors, 2010: The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015-1057.
- Szunyogh, I., E.J. Kostelich, G. Gyarmati, E. Kalnay, B.R. Hunt, E. Ott, E. Satterfield, and J.A. Yorke, 2008: A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus*, **60A**, 113-130.
- Tian, X., Z. Xie, and A. Dai, 2008: An ensemble-based explicit four-dimensional variational assimilation method. *J. Geophys. Res.*, **113**, D21124, doi: 10.1029/2008JD010358.
- Tian, X., Z. Xie, and Q. Sun, 2011: A POD-based ensemble four-dimensional variational assimilation method. *Tellus.*, **63A**, 805-816.
- Wang, X., C. Snyder, and T. M. Hamill, 2007a: On the theoretical equivalence of differently proposed ensemble/3D-Var hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222-227.
- Wang, X., T. M. Hamill, J. S. Whitaker and C. H. Bishop, 2007b: A comparison of hybrid ensemble transform Kalman filter-OI and ensemble square-root filter analysis schemes. *Mon. Wea. Rev.*, **135**, 1055-1076.

- Wang, X., T. M. Hamill, J. S. Whitaker, C. H. Bishop, 2009: A Comparison of the hybrid and EnSRF analysis schemes in the presence of model errors due to unresolved scales. *Mon. Wea. Rev.*, **137**, 3219–3232
- Wang, X., 2010: Incorporating ensemble covariance in the Gridpoint Statistical Interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990–2995.
- Wee, T.-K., Y.-H. Kuo, 2004: Impact of a Digital Filter as a Weak Constraint in MM5 4DVAR: An Observing System Simulation Experiment. *Mon. Wea. Rev.*, **132**, 543–559.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences* (2nd Edition), Academic Press, 627 pages.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.
- Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482.
- Whitaker, J. S. and T. M. Hamill, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, in press. [Available online at <http://dx.doi.org/10.1175/MWR-D-11-00276.1>].
- Wu, W.-S., D. F. Parrish, and R. J. Purser, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.
- Yang, S.-C., E. Kalnay, B. Hunt, and N.E. Bowler, 2009: Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 251–262.
- Yang, S.-C., E. Kalnay, and B. Hunt, 2012: Handling nonlinearity in an Ensemble Kalman Filter: experiments with the three variable Lorenz model. *Mon. Wea. Rev.*, in press.
- Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.*, **133**, 1710–1726.