

The Pennsylvania State University
The Graduate School
College of Earth and Mineral Sciences

ENSEMBLE DATA ASSIMILATION FOR THE ANALYSIS AND
PREDICTION OF MULTISCALE TROPICAL WEATHER SYSTEMS

A Dissertation in
Meteorology
by
Yue Ying

© 2018 Yue Ying

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Doctor of Philosophy

May 2018

The dissertation of Yue Ying was reviewed and approved* by the following:

Fuqing Zhang
Professor of Meteorology and Statistics
Dissertation Adviser
Chair of Committee

David J. Stensrud
Professor of Meteorology
Head of the Department of Meteorology and Atmospheric Science

Steven J. Greybush
Assistant Professor of Meteorology

Xiaofeng Liu
Assistant Professor of Civil Engineering

Jeffrey L. Anderson
Head of Data Assimilation Research Section, National Center for
Atmospheric Research
Special Member

Johannes Verlinde
Professor of Meteorology
Associate Head of the Department of Meteorology and Atmospheric Science

*Signatures are on file in the Graduate School

ABSTRACT

Tropical weather systems are important components of the global circulation that span a wide range of spatial and temporal scales. On the large-scale end of the spectrum, the Madden-Julian Oscillation (MJO) is found to be the dominant mode. Atmospheric wave motion due to Earth's rotation and gravity fills the spectrum from weeks to hours and from tens of thousands of kilometers to a few tens of kilometers. The thermally driven convective processes at smaller scales are chaotic in nature, which poses an intrinsic limit on the long-term predictability of tropical weather through coupling and scale interaction. This dissertation seeks to identify the predictability limits for tropical atmosphere, establishing an upper bound in expected prediction skill of these weather systems. Other scientific questions this dissertation answered are how much future satellite observations can improve the prediction skill, and how to design adaptive multiscale data assimilation methods that make better use of the available observations.

Using a convection-permitting numerical model, Weather Research and Forecasting (WRF), an MJO active phase during October 2011 is simulated. The practical predictability limit is estimated from an ensemble forecast with realistic initial and boundary condition uncertainties sampled from the operational global model forecasts. Predictability limit is reached when the ensemble spread is indistinguishable from random climatological draws. Results indicate predictability is scale dependent. There is a sharp transition from slow to fast error growth at the intermediate scales (~ 500 km), separating the more predictable large-scale components (~ 2 weeks) from the less predictable small-scale components (< 1 day). The intrinsic predictability limits, estimated by reducing the uncertainties to 1%, are > 2 weeks for larger scales and < 3 days for small

scales. An Observing System Simulation Experiment (OSSE) is conducted using the Ensemble Kalman Filter (EnKF) to evaluate the potential improvements in the prediction skill through assimilating current and future satellite observations. Results show that the currently available temperature, humidity profiles and wind vectors retrieved from infrared and microwave satellite sounder data can extend the skillful forecast lead time by as much as 4 days for the larger scales. With prospective improvement in resolution and complementary sampling strategies, the prediction skill can be further improved, especially for the smaller scales. These results shed lights on the need, design and cost-benefit analysis of future observing systems for better tropical weather prediction.

For ensemble filtering, covariance localization and inflation methods are required to account for sampling errors due to limited ensemble size and unrepresented model errors. Tuning the localization and inflation to achieve optimal filter performance is a laborious process, thus adaptive algorithms are much favored. In this dissertation, an adaptive covariance relaxation (ACR) method is proposed and tested in the Lorenz 40-variable system. The method is able to account for observations that are irregular in spatial and temporal distribution, which is typical for the tropics. In pursuit for an optimal localization method, the sensitivity of localization distance to ensemble size, model resolution, and observing network are comprehensively tested in a multiscale quasi-geostrophic (QG) model. The best localization distance is found dependent to the dominant scale of a system, which motivates the implementation of a multiscale localization for tropical weather. Some behavior related to nonlocal and irregular observation are also documented in this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES	viii
LIST OF TABLES	xiv
ACKNOWLEDGEMENTS	xv
Chapter 1 Overview of the challenges in the analysis and prediction of multiscale tropical weather systems	1
1.1 Introduction to multiscale tropical weather systems.....	1
1.2 Numerical simulation of tropical weather systems.....	2
1.3 Assessment of the predictability limits of the tropical weather systems	3
1.4 Evaluation of observation impact from data assimilation.....	6
1.5 Development of adaptive multiscale data assimilation methods	7
Chapter 2 Practical and intrinsic predictability of multi-scale weather and convectively-coupled equatorial waves during the active phase of an MJO	11
2.1 Introduction.....	11
2.2 Experimental design.....	12
2.2.1 Model configuration and the control simulation.....	12
2.2.2 Ensemble simulation.....	13
2.2.3 Predictability metric.....	14
2.3 Overview of the control simulation	16
2.4 Limit of practical predictability	18
2.5 Limit of intrinsic predictability	21
2.6 Predictability of CCEWs at different scales	23
2.7 Summary	25

Chapter 3 Potentials in improving predictability of multiscale tropical weather systems evaluated through ensemble assimilation of simulated satellite-based observations	41
3.1 Introduction.....	41
3.2 Experimental design.....	43
3.2.1 The verifying truth simulation	43
3.2.2 Data assimilation method: EnKF	45
3.3 Observing networks	47
3.3.1 Synthetic ATOVS retrieval products	47
3.3.2 Synthetic GPSRO observations	49
3.3.3 Synthetic AMV wind	50
3.3.4 Synthetic Meteosat-7 infrared brightness temperature	52
3.3.5 Synthetic CYGNSS surface wind speed	53
3.4 Scale- and variable-dependent improvements in prediction skill	54
3.5 Relative impact from different observing networks	57
3.6 Observation impact on CCEWs.....	60
3.7 Summary	62
Chapter 4 An adaptive covariance relaxation method for ensemble data assimilation	87
4.1 Introduction.....	87
4.2 Methodology	91
4.2.1. Ensemble Kalman filter	91
4.2.2. Calculation of inflation factor according to innovation statistics.....	92
4.2.3. Anderson’s adaptive covariance inflation (ACI) method ...	93
4.2.4. Adaptive covariance relaxation (ACR).....	95
4.3 Numerical experiment design	97
4.4 Results.....	99

4.4.1. Performance with sampling error due to limited ensemble size	99
4.4.2. Performance with an imperfect forecast model	101
4.4.3. Spatial and temporal behavior of inflation parameters	103
4.5 Summary	105
Chapter 5 On the selection of localization radius in ensemble filtering for multiscale quasi-geostrophic dynamics.....	115
5.1 Introduction.....	115
5.2 Experimental design.....	119
5.2.1 Two-layer quasi-geostrophic model	119
5.2.2 Ensemble filter	121
5.2.3 Observing network and assimilation experiments	122
5.3 Scale-dependent best localization radius	126
5.4 Sensitivity experiments.....	129
5.4.1 Model resolution	129
5.4.2 Sampling error due to limited-size ensemble.....	132
5.4.3 Accuracy and density of the observing network.....	134
5.5 Summary	136
Chapter 6 Concluding Remarks	152
6.1 Predictability of multiscale tropical weather systems.....	152
6.2 Potentials in assimilating satellite observations.....	154
6.3 Adaptive multiscale data assimilation methods.....	157
Bibliography	161

LIST OF FIGURES

- Figure 1.1. The spatial and temporal scales of multiscale tropical weather systems, including the Madden-Julian Oscillation (MJO), equatorial Rossby (ER) and Kelvin waves, mixed-Rossby-gravity (MRG) waves, tropical cyclones (TC), inertia-gravity (IG) waves, tropical squall lines, and (semi) diurnal convective systems. 10
- Figure 2.1 Horizontal maps of 5-day accumulated precipitation (mm) on October 20 (a, d), October 28 (b, e), and November 5 (c, f). The results are compared between (a)-(c) TRMM observation and (d)-(f) WRF control simulation. 29
- Figure 2.2. Longitude-time plots of 850-hPa zonal wind (color shadings from -20 to 20 m s^{-1}) and precipitation (black contours of 15 mm day^{-1}) averaged over 5°S - 5°N . The results are compared between (a) TRMM precipitation, ERA-Interim zonal wind, and (b) WRF control simulation. The white grid and numbers to the right indicate the observed phase of the Oct 2011 MJO (according to Fig. 6 from Johnson and Ciesielski 2013). The precipitation and wind fields are plotted at 3-h intervals, except for ERA wind that is at 6-h intervals. 30
- Figure 2.3. Wheeler-Kiladis space-time spectra of precipitation from (a, b) TRMM observation and (c, d) WRF control simulation averaged over 15°S - 15°N that is (a, c) symmetric and (b, d) anti-symmetric about the equator. Signal strengths from 1.1 to 2 are shown in shadings. The solid curves correspond to dispersion relations for dry equatorial waves with equivalent depths of 15 m. The zonal wavenumber and time frequency are labelled with corresponding zonal wavelength and time period, respectively, positive (negative) wavelength indicates eastward (westward) propagating signals. 31
- Figure 2.4. Longitude-time plot of RM-DTE (m s^{-1} , shadings) at 850 hPa and the precipitation from control simulation (black contour of 15 mm day^{-1}) averaged over 5°S - 5°N . The RM-DTE is the square root of ensemble-averaged DTE between the perturbed ensemble simulations and the control simulation from October 18 to November 2 (shown as $t = 0$ to 15 days)..... 32
- Figure 2.5. Reference energy spectra, $R(k)$, averaged over the 15 days (black lines) and error energy spectra, $E(k)$ (color-coded with simulation time $t = 0$ to 15 days), for (a) kinetic energy, (b) temperature, (c) specific humidity, and (d) precipitation. The shown spectrum is averaged over the vertical levels. 33
- Figure 2.6. Time series of spectral energy averaged at large (L; $l > 2000 \text{ km}$, black), intermediate (M; $200 < l < 2000 \text{ km}$, blue), and small (S; $l < 200 \text{ km}$, red) scales for (a) kinetic energy, (b) temperature, (c) specific humidity, and (d) precipitation. The thick lines show the error energy, while thin lines show the reference energy. 34
- Figure 2.7. Predictability limits for kinetic energy (black), temperature (blue), specific humidity (red), and precipitation (green) plotted as a function of zonal wavenumber (labeled as wavelength). Thick (thin) lines are practical (intrinsic) predictability limits. The limits are defined as the time it takes for the error energy (100% error for

- practical limit and 1% error for intrinsic limit) to reach reference energy. The limits are calculated for each member and the ensemble average is plotted. Smoothing is applied across wavelength to remove some noise for better visualization. 35
- Figure 2.8. Time series of spectral error energy integrated within L (a, d), M (b, e), and S (c, f) scales for (a)-(c) kinetic energy and (d)-(f) precipitation. The reference error energy is shown as black lines, the blue (red) lines show the error energy from the 100% (1%) error case, and the gray lines show the case with errors only in the initial condition..... 36
- Figure 2.9. Error growth time (days) plotted as a function of zonal wavenumber (labeled as wavelength) for kinetic energy (black), temperature (blue), specific humidity (red), and precipitation (green). The error growth time is defined as the time it takes for 1% initial error to grow and reach the 100% initial error. The growth time is calculated for each member and the ensemble average is plotted. 37
- Figure 2.10. Similar to Figs. 2.3c and 2.3d, but with zonal wavenumber and time frequency shown both in log scale. Space-time filtering windows are shown as black boxes for equatorial Rossby wave (ER), Kelvin wave, mixed-Rossby-gravity wave (MRG), $n = 1$ WIG wave at intermediate scale (WIG_M) and at large scale (WIG_L), $n = 0$ EIG at large scale (EIG_L) and $n = 1$ EIG at intermediate scale (EIG_M). 38
- Figure 2.11. Longitude-time plots (a)-(e) of precipitation (color shadings from -10 to 10 mm day^{-1} every 1 mm day^{-1} , zero shown in white) filtered for (a) Rossby, (b) Kelvin, (c) MRG, (d) WIG_L, and (e) EIG_L waves and averaged over $0-5^\circ\text{N}$. The second row (f)-(j) shows spaghetti plots of the contours highlighted in black in the first row, each color corresponds to a member from the perturbed ensemble simulation (100% error case). The third row (k)-(o) is similar to the second row but showing the ensemble simulation with error energy reduced to 1%. 39
- Figure 2.12. Similar to Fig. 2.11 but showing the two intermediate-scale waves, WIG_M and EIG_M. Only 5 members are shown in the spaghetti plots, and the longitude-time plots are zoomed in on $60^\circ-80^\circ\text{E}$ and the first 5 days of the simulation. 40
- Figure 3.1. Horizontal map of 15-day averaged daily precipitation (mm day^{-1}) shown in the computational domain with ERA-Interim as initial and boundary condition. The OSSE domain is shown in black 70
- Figure 3.2. Wheeler-Kiladis space-time spectra for the truth simulation: (a) symmetric and (b) asymmetric component about the equator. Signal strength from 1.1 to 2 is shown in gray shadings for precipitation and red contours for 850-hPa zonal wind. Dispersion relations for dry waves are shown with equivalent depth of 25 m. Horizontal (vertical) axis is zonal wavenumber (time frequency) but labeled with zonal wavelength (time period). 71
- Figure 3.3. Longitude-time Hovmöller diagrams of precipitation (black contours from 2 to 10 mm day^{-1} every 2 mm day^{-1}) and 850-hPa zonal wind (color shadings; m s^{-1}) averaged over $0-5^\circ\text{N}$ from the truth simulation and filtered for each CCEW: (a)

- MJO, (b) ER, (c) Kelvin, (d) MRG, (e) large-scale WIG, and (f) intermediate-scale WIG waves..... 72
- Figure 3.4. (a) Vertical observation error profiles from ATOVS (solid) and GPSRO (dashed) for T (red) and Q (blue); (b) Horizontal map of observation location during a 3-h window for ATOVS (solid black dots), ATOVS with reduced resolution (black cross), GPSRO at COSMIC2 resolution (red dots) and GPSRO at CubeSat resolution (blue circles); (c) AMV observation count (per 50-hPa vertical layer) during a 3-h window; (d) Horizontal map of AMV observation location color-coded with observation pressure height (hPa), black dots show location of CYGNSS wind observations. 73
- Figure 3.5. Temporally and vertically averaged spectra for (a) kinetic energy ($\text{m}^2 \text{s}^{-2}$), (b) vertical motion ($\text{m}^2 \text{s}^{-2}$), (c) temperature (K^2), (d) specific humidity ($\text{g}^2 \text{kg}^{-2}$), (e) rain+cloud water mixing ratio, and (f) ice+snow+graupel mixing ratios ($\text{g}^2 \text{kg}^{-2}$). Gray lines show spectra of the truth signal; red lines show spectra of the CNTL prior mean; black lines show spectra of CNTL error (difference between CNTL prior mean and the truth); and green lines show spectra of CNTL ensemble variance. 74
- Figure 3.6. Longitude-time Hovmöller diagrams of precipitation (mm day^{-1}) averaged over 0 to 5°N latitudes from (a) truth simulation, (b) NoDA ensemble mean, and (c) CNTL posterior ensemble mean. The precipitation fields are filtered for $l > 200$ km. ... 75
- Figure 3.7. (a)-(c) Snapshots of 850-hPa streamlines and 3-h accumulated precipitation (shadings; mm), and (d)-(f) 850-hPa relative humidity (%) after 5.5 days of cycling (valid at 12UTC, 21 Oct). Results are compared for (a, d) truth simulation, (b, e) NoDA ensemble mean, and (c, f) CNTL posterior ensemble mean. All fields are filtered for $l > 200$ km. The left half of the OSSE domain is shown (10°S to 10°N , 55° to 75°E). 76
- Figure 3.8. Times series of domain-averaged (a) root-mean difference kinetic energy (m s^{-1}), and RMSEs of (b) vertical motion (m s^{-1}), (c) temperature (K), (d) specific humidity (g kg^{-1}), (e) rain+cloud water mixing ratio (g kg^{-1}), and (f) ice+snow+graupel mixing ratio (g kg^{-1}). The errors are also averaged over the vertical column and filtered for $l > 200$ km. Black lines show error from CNTL posterior ensemble mean and gray lines from NoDA ensemble mean. Colored lines show error evolution from ensemble forecasts initiated at day 5, 7, 9, 11, 13, and 15. All ensemble means are calculated from 10 members. 77
- Figure 3.9. Vertical profiles of (a) RM-DKE (m s^{-1}), and RMSEs of (b) vertical motion (m s^{-1}), (c) temperature (K), (d) specific humidity (g kg^{-1}), (e) rain+cloud water mixing ratio (g kg^{-1}), and (f) ice+snow+graupel mixing ratio (g kg^{-1}). The errors are averaged over day 5 to 18 and filtered for $l > 200$ km. Black lines show error from CNTL posterior ensemble mean and gray lines from NoDA ensemble mean. Colored lines show error at forecast lead time of 12 hours, 1, 2, and 4 days. All ensemble means are calculated from 10 members. 78
- Figure 3.10. Similar to Fig. 3.9 but showing the temporally and vertically averaged error energy spectra. The spectral errors are in variance units, e.g. $\text{m}^2 \text{s}^{-2}$ for DKE and K^2

- for temperature. The spectrum are shown with respect to global horizontal wavenumber, $k = kx^2 + ky^2$, but labeled with its corresponding wavelength ($l = k - 1$). 79
- Figure 3.11. (a)-(d) Snapshots of channel-2 Tb (K), and (e)-(h) column integrated ice (gray), snow (red) and graupel (blue) mixing ratios (g kg^{-1}) after 5.5 days of cycling (valid at 12UTC, 21 Oct) from (a, e) truth simulation, (b, f) NoDA ensemble mean, (c, g) CNTL posterior ensemble mean, and (d, h) +Met7 posterior ensemble mean. 80
- Figure 3.12. (Top row) Time series, (center row) vertical profiles, and (bottom row) spectra of domain-averaged analysis errors in (columns from left to right) kinetic energy, temperature, specific humidity, and summed error from ice, snow, and graupel mixing ratios. Errors are shown as RMSE (RM-DKE for kinetic energy), except that the spectra show error variance. The time series are vertically averaged and filtered for $l > 200$ km; vertical profiles are averaged over time and filtered for $l > 200$ km; spectra are averaged temporally and vertically. Results from NoDA, CNTL, +Met7, +CYGNSS and NoPofile are compared. 81
- Figure 3.13. Same as Fig. 3.12, but showing results from NoDA, NoProfiles, +ATOVS, +ATOVScoarse, +GPSRO, and +GPSROdense. 82
- Figure 3.14. Same as Fig. 3.12, but showing results from NoProfiles, +ATOVS, +ATOVScoarse, +ATOVSthinV, and +ATOVSerror1.5. 83
- Figure 3.15. Same as Fig. 3.12, but showing results from NoProfiles, +GPSROdenseN, +GPSRON, +GPSROdense, and +GPSRO. 84
- Figure 3.16. Longitude-time Hovmöller diagrams of 850-hPa zonal wind (color shadings; m s^{-1}) averaged over $0-5^\circ\text{N}$ latitudes from (top row) truth simulation, (center row) NoDA ensemble mean, and (bottom row) CNTL posterior ensemble mean. The fields are filtered for each CCEW shown in columns from left to right. 85
- Figure 3.17. Longitude-time Hovmöller diagrams of RM-DTE with respect to the truth (m s^{-1}) averaged over $0-5^\circ\text{N}$ latitudes and vertical levels for (a)-(f) NoDA ensemble mean and (g)-(l) CNTL posterior ensemble mean. A space-time filter is applied to u , v and T for each CCEW (shown from left to right) before calculating the RM-DTE. 86
- Figure 4.1. Analysis RMSE (top row) and consistency ratio (bottom row) for the fully-observed cases with different degrees of sampling errors introduced by changing the ensemble size (N). Left (right) column shows result without (with) localization ($\text{ROI} = 10$). The gray line shows the observation noise level. The colored lines show inflation parameters for RTPS with α ranging from 0 to 1, the cross markers (no inflation) is equivalent to $\alpha = 0$. Triangle and square markers correspond to the ACI methods with $\sigma\lambda^2 = 0.1$ and 1, respectively. Circle markers correspond to the ACR method. Note that we show more data points in the transition region ranging from the filter being stable to diverging. The transition region is different for the cases without localization ($N = 15-40$) and with localization ($N = 4-10$). 109

- Figure 4.2. Analysis RMSE and consistency ratio for the half-observed cases ($N = 10$ with localization). RMSE_land (RMSE_ocean) is calculated using the observed (unobserved) variables along with the truth, while the consistency ratio is calculated in observation space (observed variables only). The legend is the same as Fig. 4.1. 110
- Figure 4.3. Analysis RMSE (left column) and consistency ratio (right column) for the fully-observed cases ($N = 40$ without localization) with different degree of model errors introduced by varying parameters F , a and d . Within each panel, F varies from 5 to 8. For each row, a different combination of a and d values are used. The legend is the same as Fig. 4.1..... 111
- Figure 4.4. Same as Fig. 4.3, but showing half-observed cases ($N = 10$ with localization). ... 112
- Figure 4.5. The spatial structure of inflation parameters for the half-observed cases ($N = 10$ with localization) with two types of severe error sources: (a) sampling error with $N = 5$ and (b) model error with $F = 5$. The colored lines show inflation parameters for RTPS with α ranging from 0 to 1. Gray triangle, square and cross markers correspond to ACI methods with $\sigma\lambda^2 = 0.1, 1$ and 10 , respectively..... 113
- Figure 4.6. Time series of the inflation factor (λ) from ACR method for half-observed cases ($N = 10$ with localization) with (a) perfect model and (b) model error $F = 5$. Only the first and last 200 steps are shown. Blue, red and black lines correspond to $\tau = 1, 10$ and 100 , respectively..... 114
- Figure 5.1. Snapshots of (a) θ and (b) u in color shadings and state variable ψ in black contours (solid lines for positive and dotted lines for negative values). The ψ , θ , and u variables are nondimensionalized. The scaling of the QG model is described in section 5.2.1. 140
- Figure 5.2. (a) Spectra of error kinetic energy from ensemble mean (color-coded with forecast time) and the time-averaged reference kinetic energy from the truth (black) during an ensemble spin-up period of 1.5 time units. (b) Spatially averaged mean absolute correlation (MAC) between θ and ψ plotted as a function of horizontal distance (number of grid points) also color-coded with forecast time as in (a). (c) Same as (b) but showing correlation between u and ψ 141
- Figure 5.3. Time series of θ MSE from CNTL filtered for the (a) L, (b) M and (c) S scales. Results are shown for errors from free ensemble (NoDA; black), observation (gray), and from analysis mean using ROI = 8, 16, 32, and 64 (colored)..... 142
- Figure 5.4. θ error spectra from (a) CNTL, (b) M_Scale, and (c) S_Scale. Results are shown for errors from a free ensemble (NoDA; black), observation (gray), and from analysis mean using ROI = 8, 16, 32, and 64 (colored). The CNTL NoDA error is shown as dotted lines in (b) and (c) for reference. 143
- Figure 5.5. (a) θ RMSE plotted with respect to ROI from CNTL (black), M_Scale (blue) and S_Scale (red); (b) Same as (a) but showing θ RMSE from CNTL filtered for the L (black), M (blue) and S (red) scales. The filled circles indicate the ROI with lowest analysis error..... 144

- Figure 5.6. MAC between θ and ψ filtered for L (dotted) and S (solid) scales estimated by the prior CNTL ensemble using ROI = 8, 16, 24, 32, and 64. The MACs are averaged over space and time and plotted as a function of horizontal distance (number of grid points). 145
- Figure 5.7. θ error spectrum of NoDA (black), observation (gray), and analysis ensemble mean using ROI = 8, 16, 32, and 64 (colored) from (a) LowRes, (b) LowRes_Model, (c) LowRes2, and (d) LowRes2_Model experiments. The CNTL NoDA error spectrum is shown as dotted lines for reference. The ROI is in physical distance units (number of grid points from CNTL model), e.g. ROI = 8 means 8 grid points in CNTL, which corresponds to 4 grid points in LowRes, and 2.67 grid points in LowRes2. 146
- Figure 5.8. θ RMSE filtered for (a) L and (b) M scales plotted with respect to ROI. Results are compared for CNTL, LowRes, LowRes_Model, LowRes2 and LowRes2_Model. 147
- Figure 5.9. θ RMSE filtered for (a) L, (b) M and (c) S scales plotted with respect to ROI compared for the $N = 16$, $N = 32$, $N = 64$ (CNTL), $N = 256$ and $N = 1024$ cases. Filled circles indicate the minimum error corresponding to the best ROI. (d)-(f) Same as (a)-(c) but showing root mean error kinetic energy from the cases assimilating u and v 148
- Figure 5.10. (a) MAC between θ and ψ from the posterior ensemble with changing ensemble size $N = 16, 32, 64, 256, 1024$ using the best ROI. (b) Same as (a) but showing MAC between u and ψ from the cases assimilating u and v . (c) Time evolution of MAC between θ and ψ from ensemble forecasts initialized with posterior from the $N = 64$ case, the lines are color-coded with forecast time $t = 0$ to $t = 1.5$ every 0.05 time units. (d) same as (c) but showing MAC between u and ψ from the cases assimilating u and v 149
- Figure 5.11. Error spectra of free ensemble mean (NoDA; black), observation (gray), and analysis mean using ROI = 8, 16, 32, and 64 (colored) from (a) ObsSparse, (b) ObsErrorX3, (c) ObsDense, and (d) ObsError/3 cases. 150
- Figure 5.12. θ RMSE filtered for (a) L, (b) M and (c) S scales plotted with respect to ROI for the cases with changing observing networks: ObsSparse, ObsDense, ObsErrorX3 and ObsError/3. Filled circle markers indicate the minimum error corresponding to the best ROI. (d)-(f) Same as (a)-(c) but showing results from using irregular observing networks. 151

LIST OF TABLES

Table 2.1. Averaged pattern correlation between member (perturbed run) and the control (unperturbed run) for zonal wind, temperature, specific humidity and precipitation associated with each CCEW mode. The improvement in pattern correlation from the 100% error case to the 1% error case is indicated by an arrow.	28
Table 3.1. Space-time filtering parameters for each tropical weather system (CCEW type).	65
Table 3.2. Observed variable, their associated uncertainty (observation error), spatial resolution and approximate data count within a 3-h window of each tested observing network.....	66
Table 3.3. Description of assimilated observing networks and their benefits (improvement in prediction skills in variables, scales and weather systems) found in each experiment.....	67
Table 3.4. Root-mean-square errors with respect to the truth simulation for u (m s^{-1}), T (K), Q (g kg^{-1}), and precipitation (mm day^{-1}) filtered for CCEW types described in Table 3.1. The errors are averaged over latitude ($0\text{-}5^{\circ}\text{N}$) and time (17-day period). Pattern correlations with the truth are shown in parentheses. Results are compared between NoDA ensemble mean and CNTL posterior ensemble mean.....	68
Table 3.5. Similar to Table 3.4, but comparing results from NoDA, NoProfile, +ATOVScoarse and +ATOVS posterior ensemble mean from the first 10 days and for WIG waves.	69
Table 4.1. Analysis RMSEs using no inflation, ACR, best-tuned RTPS and ACI methods (columns) for different error sources (rows) including sampling error and model error. The fully-observed case without localization is shown (corresponding to Figs. 4.1a and 4.3a).....	108
Table 5.1. Description of model and filter parameters used in control and sensitivity experiments. In each experiment, a range of ROIs is tested and the ROI that minimizes analysis RMSE is determined.....	139

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude to my adviser, Fuqing Zhang, for his endless support during my time as a graduate student. I deeply admire him for his passion for science and patience and optimism that encouraged me to carry on my research during difficult times. I want to thank my committee members, Steven Greybush, David Stensrud, Jeffrey Anderson, and Xiaofeng Liu, for their valuable suggestions and career advises. I am glad I had the privilege to work among so many passionate and devoted researchers in Fuqing's diverse research group. I could not have completed all my thesis research without the thought-provoking discussions. I also want to thank my wife, Dandan, whose love and company kept me going for the past years, and made those years so memorable and enjoyable.

The dissertation research received the funding support from National Science Foundation (NSF), National Center for Atmospheric Research (NCAR), and China Scholarship Council (CSC). Any opinions, findings, and conclusions or recommendations expressed in this dissertation are those of the author and do not necessarily reflect the views of the NSF, NCAR, or CSC.

Chapter 1

Overview of the challenges in the analysis and prediction of multiscale tropical weather systems

1.1 Introduction to multiscale tropical weather systems

Tropical atmosphere consists of weather systems spanning a wide range of spatial and temporal scales. Figure 1.1 illustrates the ranges of spatial and temporal scales associated with different kinds of tropical weather systems. At the planetary scale, the Madden-Julian Oscillation (MJO) is found to be the dominant mode of intraseasonal variability with typical periods of 20-100 days (Madden and Julian 1971, 1972; Zhang 2005). The active phase of an MJO is characterized by enhanced deep convection and intense precipitation that propagates eastward at a speed around 5 m s^{-1} . Within the MJO envelope, a wide variety of convectively-coupled equatorial waves (CCEWs) reside, including equatorial Rossby (ER), Kelvin, mixed-Rossby-gravity (MRG), and inertia-gravity (IG) waves.. A comprehensive review of observations and dynamics of CCEWs was provided by Kiladis et al. (2009). The equatorial Rossby waves are large-scale “cyclone pairs” that propagate westward at a speed around 4.5 m s^{-1} (Kiladis et al. 2009). The equatorial Kelvin waves are also known as the super-cloud clusters that propagates eastward at a speed of $15\text{-}20 \text{ m s}^{-1}$ (Nakazawa 1988; Dunkerton and Crum 1995). The MRG waves propagate westward at a speed of $15\text{-}20 \text{ m s}^{-1}$ and have the potential to

develop into tropical cyclones (Takayabu and Nitta 1993; Dickinson and Molinari 2002). The IG waves span a wide range of spatiotemporal scales, ranging from smaller-scale diurnal variations to larger-scale 2-day waves (Haertel and Kiladis 2004). At the sub-diurnal scales, convective systems such as the tropical squall lines can be found.

1.2 Numerical simulation of tropical weather systems

Simultaneously resolving systems at different scales has been a challenge for numerical simulation, because the computational domain needs to be large enough to capture the large-scale components as well as having small enough grid spacing to resolve the small-scale components. Huge computational cost is required to fulfill this dual requirement of domain size and resolution before it is feasible to study the interaction among systems at different scales. Recent development in high-resolution simulation of the tropical weather start to show better agreement with observation than previous coarse-resolution global models, indicating the importance of resolving smaller-scale components that are likely coupled with the larger-scale ones. For example, Wang et al. (2015) conducted a convective-permitting (9-km resolution) simulation of the October and November 2011 MJO events observed during the Dynamics of MJO (DYNAMO; C. Zhang et al. 2013) field campaign, using the Weather Research and Forecasting (WRF) model. They showed promising results that, thanks to the use of reanalysis data for model initialization and a nudging method to keep model solution from deviating too far away from large-scale observed dynamic and thermodynamic structures, the simulated MJO propagation, precipitation episodes, and moist static

energy budget of the MJO events all match relatively well with observations. However, there are clearly room for improvement in terms of finer-scale details of the simulated MJO events. The equatorial Kelvin wave and inertia-gravity waves have a lot of mismatches with the observation due to lack of constraints at smaller scales. At these scales, bias and errors in parameterization also pose some challenges in correctly representing the physical processes that are not well resolved.

1.3 Assessment of the predictability limits of the tropical weather systems

Lorenz (1963) first discovered the chaotic behavior of a dynamic system due to sensitivity to the initial condition, which pose a finite limit to the predictability of this system. Small errors in the initial condition may grow and eventually render the long-term predictions to have no skill as comparing to random draws from climatology. The concept of atmospheric predictability can be grossly categorized into intrinsic versus practical predictability (Lorenz 1996, Melhauser and Zhang 2012). Intrinsic predictability refers to the ability to predict given nearly-perfect representation of the dynamical system (by a forecast model) and nearly-perfect initial/boundary conditions, an inherent limit due to the chaotic nature of the atmosphere (Lorenz 1963, 1969; Zhang et al. 2003, 2007; Sun and Zhang 2016). Practical predictability, sometimes also referred to as the prediction skill, is the ability to predict given realistic uncertainties in both the forecast model and initial and boundary conditions (Lorenz 1982, 1996; Zhang et al. 2002, 2006) that can be both large at present. The limit of practical predictability can potentially be extended through the use of more accurate initial conditions (resulting from better data assimilation

methods and/or observations) and/or better forecast models (better model physics, numerics, and/or resolution).

Previous studies suggest that the predictability of mid-latitude weather systems can be intrinsically limited due to the chaotic nature of moist convection and the rapid upscale error growth, as shown for winter cyclones (Zhang et al. 2003, 2007; Sun and Zhang 2016) and summertime continental mesoscale convective systems (e.g., Bei and Zhang 2007; Melhauser and Zhang 2012; Selz and Craig 2015; Zhang et al. 2016). For tropical weather, tropical cyclones are also found to have similar limits in predictability due to moist convection (e.g., Zhang and Sippel 2009; Tao and Zhang 2015; Judt et al. 2016).

How predictable are the multiscale tropical weather systems? Is the predictability of CCEWs very limited due to coupling with moist convections in the tropics? Does each component of the multiscale tropical weather system have a unique predictability limit related to its scale (Bei and Zhang 2014)? At the planetary scale, several studies have documented their estimation of predictability of MJOs from global model simulations. Neena et al. (2014a, b) estimated from global circulation model simulations that MJO has the potential predictability of 35~45 days and the current prediction skill is 20~30 days. Hamill and Kiladis (2014) evaluated the current Global Ensemble Forecast System (GEFS) MJO simulations and found that the simulated MJO propagation speed is too slow and precipitation is overestimated. Miyakawa et al. (2014) estimated from a global cloud-resolving model that the current prediction skill for MJO is 27 days. The predictability is found to be dependent on MJO phases. Simulations initialized from active phases appear to have better predictability (Waliser et al. 2003; Nasuno et al. 2013;

Neena et al. 2014a,b). The models tend to generate too much precipitation during the suppressed phase, thus have difficulties predicting the correct timing of MJO initiation. Ling et al. (2014) suggested that MJO predictability may differ greatly when considering global or local scales and during different MJO phases. Such scale dependency also results in a more limited predictability estimated from higher-resolution simulation than coarser-resolution global models (Miyakawa et al. 2014). More recently, Zhang et al. (2017) also suggested that initiation of this MJO event may have resulted from a global circumnavigating signal, which can be potentially predictable at the planetary scale several weeks in advance.

To the best of my knowledge, the predictability of multiscale tropical weather beyond tropical cyclones and MJOs is rather underexplored. Using a then-operational global prediction system under a perfect model assumption, Reynolds et al. (1994) found that the internal error growth rate in the tropics is several times slower than that in mid-latitudes while the external error growth rate due to model deficiencies is considerably larger. Using a global convection-permitting aqua-planet model configured with different resolutions, Mapes et al. (2008) revealed that predictability of tropical weather can be potentially limited by error growth from mid-latitude moist baroclinic systems. Although there is usually an average predictability for a system, the actual predictability is quite variable and depends on the state of the system itself (Kalnay 2003). The growth rate of errors also depends on scales and variables (Bei and Zhang 2014). To further characterize the error growth associated with each component of the multiscale tropical weather system, a systematic assessment of predictability is performed in chapter 2 using ensemble techniques.

1.4 Evaluation of observation impact from data assimilation

The simulated MJO events in Wang et al. (2015) agrees well with observation at the large scales, while their smaller-scale details are not yet matching exactly with the observation. To further improve the simulation for systems at smaller scales, the data assimilation system should be tuned for optimal performance for a certain weather system of interest. In operational centers, the design of an ensemble assimilation and prediction system usually involves the consideration of temporal and spatial scale of the weather system, the availability of observations, as well as computational cost. For weather systems at smaller scales, denser observations should be assimilated at a higher temporal frequency comparing to weather systems at larger scales. The presence of multiple components at different scales in the complex tropical weather system poses some challenge in the design of a unified data assimilation scheme for this system. The larger and smaller scale components are expected to require different configurations.

While the ensemble simulations performed in chapter 2 provide estimates of the predictability of tropical weather at different scales, a follow-up Observing System Simulation Experiment (OSSE) will be performed in chapter 3 to evaluate the potentials in assimilating current and future satellite observations for improving the predictability, given that there are room for improvement, especially for components at smaller scales. An OSSE first runs a model simulation and considers it as the truth, then generate synthetic observations from the truth, and performs data assimilation trials using these synthetic observations and evaluate the performance of data assimilation. Unlike the real-data scenario where the truth is unknown, the perfect knowledge of the truth allows a

more thorough error statistics to be calculated in an OSSE scenario. Without the complication from quality control and cross calibration between model and observation required for the real-data scenario, an OSSE using perfect model assumption can help one focusing on the performance of data assimilation itself.

However, the caveat of OSSEs is that the results can be too optimistic when model errors and errors associated with the preparing and quality control process of observations are not accounted for. The findings from an idealized OSSE should always be complemented by further experiments in the real-data scenarios.

1.5 Development of adaptive multiscale data assimilation methods

Data assimilation is the process of incorporating observations to improve the accuracy of model initial conditions. Among the popular choices of data assimilation methods are the variational approach (e.g. 4DVar) and the ensemble-based filtering methods. Ensemble Kalman filter (EnKF; Evensen 1994) is a Monte Carlo approximation of the extended Kalman filter, using an ensemble of model realizations to sample the probability distribution. Observations are used to update both the ensemble mean and ensemble perturbations so that the analysis ensemble characterizes the mean and variance of the posterior probability distribution as given by the Bayesian theorem. Then, the model propagates the analysis ensemble forward in time to provide prior distribution for the next assimilation cycle. Through ensemble forecast, the flow-dependent background error covariance estimated from the ensemble accounts for the errors of the day, and thus ensure optimal use of local observation to adjust the model state. The EnKF algorithm

used in the OSSE in chapter 2 follows the ensemble square root filter (EnSRF; Whitaker and Hamill 2002) formulation, which also uses a serial assimilation strategy (Anderson 2003).

Although EnKF has the advantage of providing flow-dependent background error, the use of a limited-size ensemble to sample the high-dimensional state space does pose a lot of challenges. Sampling errors in the ensemble-estimated background error covariance due to small ensemble size requires the covariance to be localized (Hamill et al. 2001). The use of an imperfect model may cause model uncertainties to be unrepresented in the background error covariance, and thus requires inflation of the ensemble spread to prevent catastrophic filter divergence (Mitchell and Houtekamer 2000; Anderson and Anderson 1999; Zhang et al. 2004; Whitaker and Hamill 2012).

Tuning of the covariance localization and inflation is very costly, therefore adaptive methods were introduced to estimate the amount of covariance inflation according to innovation statistics (Wang and Bishop 2003; Anderson 2007, 2009; Li et al. 2009; Miyoshi 2011), and to estimate an optimal localization function (Anderson 2007, 2012; Bishop and Hodyss 2007, 2009; Anderson and Lei 2013; Zhen and Zhang 2014; Flowerdew 2015).

Given the often-limited observation availability in the tropics, a well-tuned data assimilation method is of vital importance in order to make the best use of the available observations. The tropical observing network is highly variable in space and time distribution, which is a quite challenging scenario for adaptive algorithms to operate in. An ideal adaptive covariance inflation method should be able to adjust the amount of inflation according to the density and quality of observations. One such method, called

Adaptive Covariance Relaxation (ACR) is introduced in chapter 4, and tested in the Lorenz (1996) 40-variable model framework.

Equally challenging is the choice of an optimal localization function for the tropical observing network. Due to the multiscale nature of the tropical weather, it is difficult to specify one localization function that is optimal for all components at different scales. It is necessary to use a multiscale localization scheme (Zhang et al. 2009; Miyoshi and Kondo 2013; Li et al. 2015; Buehner and Shlyayeva 2015). However, the theory of optimal localization is complicated by the fact that localization function does not only depend on the dominating scale of a system, it is also influenced by ensemble size and observing network properties. Previously proposed adaptive localization methods (e.g. Zhen and Zhang 2014) are not necessarily comprehensive and independent to the particular dynamical system and observing network. To extend the optimal localization theory to the scenario where multiple scales are present in a system, a sensitivity experiment is performed in chapter 5 using a quasi-geostrophic (QG) model, a simplified yet realistic atmospheric system. The experiment explores the behavior of best-performing localization distance in response to changes in model dynamics, resolution, ensemble size, and observation density and accuracy, which provide insights on the development of a scale-aware adaptive localization method.

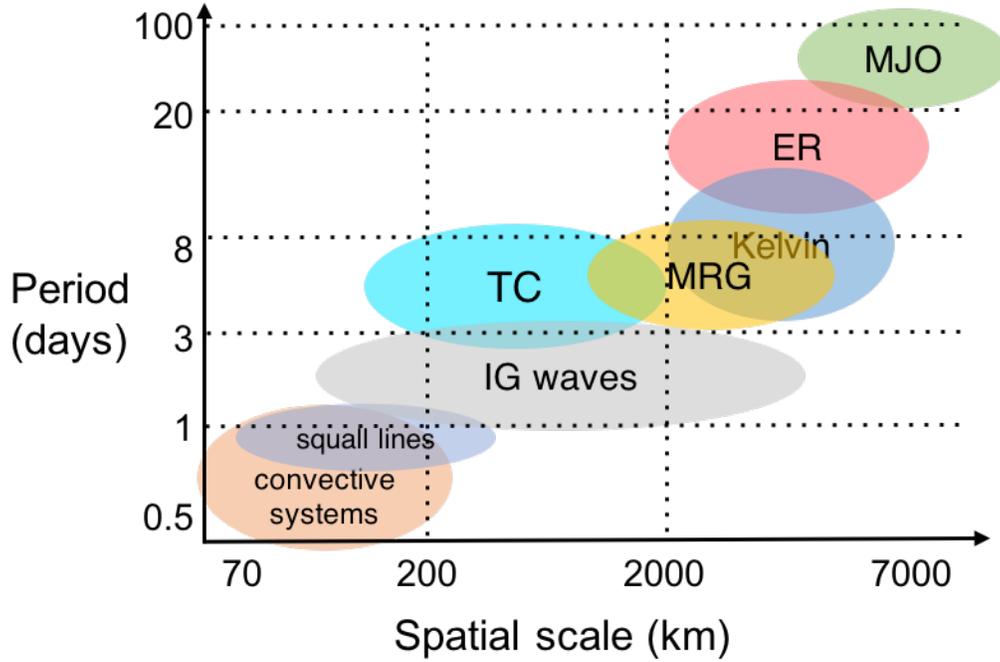


Figure 1.1. The spatial and temporal scales of multiscale tropical weather systems, including the Madden-Julian Oscillation (MJO), equatorial Rossby (ER) and Kelvin waves, mixed-Rossby-gravity (MRG) waves, tropical cyclones (TC), inertia-gravity (IG) waves, tropical squall lines, and (semi) diurnal convective systems.

Chapter 2

Practical and intrinsic predictability of multi-scale weather and convectively-coupled equatorial waves during the active phase of an MJO

2.1 Introduction

In 2011, the Dynamics of MJO (DYNAMO) field campaign is conducted over the Indian Ocean to gather more observations, to advance physical understanding of MJOs, and ultimately to improve MJO prediction (Zhang et al. 2013; Yoneyama et al. 2013). Two moderate to strong MJO events occurred during October and November 2011, respectively, were well campaigned by the field campaign observations, as documented in Johnson and Ciesielski (2013). Recently, Wang et al. (2015) conducted a successful simulation of these MJOs using the Weather Research and Forecasting (WRF) model at the convection-permitting resolution with a grid spacing of 9 km. They show that the model is capable of reproducing most of the observed MJO features, including its eastward propagation, dynamical structure and the overall rainfall pattern and magnitude. Sensitivity experiments with the same regional WRF configuration in Zhang et al. (2017) subsequently demonstrated the crucial importance of the global circumnavigating mode in the MJO initiation and propagation. These results motivate the use of such a model as a proxy of the tropical atmosphere for studying its predictability.

In this chapter, a systematic investigation of both the practical and intrinsic limits of multiscale predictability of tropical weather and CCEWs is conducted through a series of unprecedented convection-permitting regional-scale ensemble simulations. The resulting predictability estimates provide a benchmark for the future investigation with improved modeling systems. The model configuration and design of ensemble simulation are described in section 2.2. In section 2.3, an overview is given for the simulated features of CCEWs. The practical predictability is estimated in section 2.4, followed by an intrinsic predictability assessment in section 2.5. In section 2.6, the error growths are analyzed for CCEWs to illustrate their distinct predictability limits. Section 2.7 summarizes the findings.

2.2 Experimental design

2.2.1 Model configuration and the control simulation

In this study, the WRF model version 3.4.1 (Skamarock et al. 2008) is employed to conduct simulations. The model configuration is similar to that described in section 2a of Wang et al. (2015). The computational domain covers the equatorial Indian Ocean and part of the Maritime Continent (20°S to 20°N, 50° to 120°E). The model grid is 445×778 with 9 km spacing, and it has 45 vertical levels with 9 levels in the lowest 1 km and a model top at 20 hPa. The initial condition (IC) and lateral boundary condition (LBC) are specified by the ERA-Interim reanalysis data (Dee et al. 2011). The sea surface temperature (SST) for the lower boundary condition is updated every 6 h according to the ERA-Interim data.

The WRF Double-Moment (WDM) scheme (Lim and Hong 2010) is used to parameterize cloud physics with modifications to the shape parameters and terminal velocity of snow. Both shortwave and longwave radiation are treated with the CAM scheme (Collins et al. 2004). Surface processes are represented with the unified Noah land surface model (Chen and Dudhia 2001) with variable surface skin temperature (Zeng and Beljaars 2005). Subgrid-scale turbulent eddy mixing is parameterized using the Yonsei University (YSU) PBL scheme (Hong et al. 2006). No cumulus parameterization is used, and organized convective motion is explicitly represented by the 9-km model grid. A control simulation of the MJO active phase is initialized at 00Z 12 October and integrated for 31 days that ends on November 12, 2011. Section 3 will provide an overview of the simulated period and validation of model simulation with observations.

2.2.2 Ensemble simulation

A pair of 20-member ensembles are conducted to study the practical versus intrinsic predictability of the tropical weather systems and CCEWs. The first ensemble simulation designed to examine the practical predictability limits starts from October 18 to November 2 (corresponding to the MJO phases 1 to 3). The IC and LBC ensemble perturbations are sampled from the operational European Center for Medium-Range Weather Forecasts (ECMWF) global ensemble forecasts archived in The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE). The TIGGE archives 15-day global forecasts of horizontal winds, temperature, geopotential height, and specific humidity at 8 pressure levels (1000, 925, 850, 700, 500, 300, 250, and 200 hPa) at 12-h intervals. The then-operational ECMWF

ensemble has a horizontal resolution of 32 km (T639) for the first 10 days, and 63 km (T319) from day 10 to 15. Ensemble perturbations from the first 20 TIGGE members are interpolated to the 9-km WRF model grid and to 6-h intervals in time, then added to the control IC and LBC generated from ERA-Interim data. Since the TIGGE ensemble forecasts are valid at the same time as the IC and LBC being perturbed, the ensemble perturbations are physically consistent with the flow-dependent realistic uncertainties of the unperturbed model atmosphere. With the global model uncertainties downscaled to the regional model at the 9-km resolution, the ensemble forecasts designed herein will provide a realistic estimate of the practical predictability of the tropical weather and CCEWs during the MJO active phase under a perfect model assumption.

The intrinsic predictability is estimated from another set of 20-member ensemble simulations with the IC and LBC perturbation uncertainties reduced to 1 % in terms of error energy (or 10% error magnitude) comparing to the ensemble described above, which is a level of accuracy that is unlikely achievable in the foreseeable future. In other words, we seek the upper bound in prediction using a perfect model under nearly perfect initial and boundary conditions following Lorenz (1996).

2.2.3 Predictability metric

The limit in predictability will be quantified in terms of relative magnitude between the reference and error spectral energy, $R(k)$ and $E(k)$, defined respectively as

$$R(k) = \sum_{k_x^2 + k_y^2 = k^2} (\hat{x})_{k_x, k_y}^2, \quad (2.1)$$

and

$$E(k) = \sum_{k_x^2 + k_y^2 = k^2} \frac{1}{N-1} \sum_{i=1}^N (\hat{x}'_i)_{k_x, k_y}^2, \quad (2.2)$$

where x denotes the variable in consideration, $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ is the ensemble mean, $x'_i = x_i - \bar{x}$ is the ensemble perturbation, subscript $i = 1, 2, \dots, N$ indexes the ensemble member, the hat denotes the Fourier transformation in two horizontal dimensions and the subscripts k_x and k_y are the zonal and meridional wavenumbers, respectively. Following Bei and Zhang (2014), the reference and error energy are decomposed into spectral components (i.e. a function of global wavenumber k) and calculated separately for each model variable to demonstrate the scale- and variable-dependency in predictability. The variables of interest in this study are the u - and v -component wind, temperature, specific humidity, and precipitation. For the sake of simplicity, the u , v wind are combined as kinetic energy (KE), i.e. $\frac{1}{2}(u^2 + v^2)$. The reference and error KE are calculated by first evaluating (2.1) and (2.2) for u , v and then taking their average.

Bei and Zhang (2014) used only one perturbed simulation and its difference with the unperturbed (control) simulation to measure the error energy. Only one realization of forecast error is available, therefore, the predictability estimate is less robust. More ideally, one can use an ensemble to sample the forecast error and provide a much more robust estimate for predictability. In their predictability study on tropical cyclone intensity, Judt et al. (2016) defined an error energy as the averaged squared differences between two members from a 20-member ensemble, and the reference energy was

defined as the averaged energy from each member. The predictability limit was defined as the forecast time at which error saturates, i.e. the error energy becomes close enough to the reference energy. One caveat of such predictability limit definition is that the error growth usually slows down as approaching saturation so that the exact saturation time is difficult to evaluate due to this asymptotic behavior.

In this study, a slightly different definition is used. The error energy (noise) is defined as the ensemble variance and the reference energy (signal) is defined as the energy associated with the ensemble mean. As errors grow, the reference energy from ensemble mean will decrease due to the smoothing among ensemble members. The predictability limit is defined as the forecast time at which error energy reaches and exceeds the reference energy (signal-to-noise ratio drops below 1).

2.3 Overview of the control simulation

The control simulation conducted in this study is mostly consistent with the control experiment results from Wang et al. (2015), except that the simulation starts from a later time and analysis nudging is not performed. The active MJO phase features the eastward propagation of large-scale organized convection and precipitation. Figure 2.1 compares the 5-day accumulated precipitation from the control simulation to the Tropical Rainfall Measuring Mission (TRMM) observations. During the simulated period, the precipitation center moves across the Indian Ocean (phases 1 to 3) and to the Maritime Continent (phases 4 to 5). The propagation of the MJO is not at a constant speed. Phase 2 takes much longer time (10 days) than the following phases. Consistent with Wang et al.

(2015), the eastward propagation of the simulated precipitation agrees with the observation as shown in the Hovmöller diagrams in Fig. 2.2. At 850 hPa, westerlies (easterlies) are found west (east) to the precipitation center, which is the typical large-scale MJO flow pattern.

At smaller scales, the simulated precipitation shows inertia-gravity (IG) wave signals over the Maritime Continent (90° to 120°E). The propagation direction of these IG waves follows the prevailing zonal wind, i.e. westward-propagating IG (WIG) waves in easterly wind and vice versa. At least some of these waves and precipitation patterns over the Maritime Continent are likely forced by the thermal diurnal cycles associated with the land mass (Mapes et al. 2003; Love et al. 2011). The precipitation over the Indian Ocean (50° to 90°E) is organized into several 2-day episodes which are modulated by the phase of the MJO and several episodes of westward-propagating equatorial Rossby waves and eastward-propagating equatorial Kelvin waves (hereafter referred to as Rossby and Kelvin waves for simplicity). The model simulation of these finer scale features is less accurate than the MJO signal itself as compared to the observations. There are generally mismatches in timing of the IG waves over the Indian Ocean.

However, the model simulation has a relatively good representation of the spectral modes of the multiscale CCEWs. Figure 2.3 compares the Wheeler and Kiladis (1999) space-time spectra (WK spectrum) of precipitation between the observation and the control simulation. Comparing to the TRMM observation, the WRF simulation captures most of the CCEW modes although with errors in their signal strengths. The zonal wavelengths and time periods of the simulated waves agree with the observation, except that Kelvin waves are propagating more slowly and the large-scale eastward propagating

IG (EIG) waves are missing the 2.5-day period in the simulation. The signal strengths of diurnal WIG, MRG, and Rossby waves are weaker in the simulation than in the observation, while Kelvin wave and MJO signal strengths are comparable to the observation. The large-scale 2-day WIG wave signals are simulated stronger than the observation. The WK spectra are calculated for a single MJO active phase during the simulated one month period, which is a relatively short sample size. Therefore, the large-scale low-frequency wave signals are expected to have some errors due to sampling noises. The errors in model and the specified initial and boundary conditions also cause the simulated WK spectra to differ from observation. Despite of these discrepancies, the WRF simulation provides a reasonable representation of the observed CCEW and well serves the need as a control simulation.

2.4 Limit of practical predictability

To estimate the practical predictability (prediction skill) of the tropical multi-scale weather and CCEWs, the error growth during the first ensemble simulation is investigated in this section. Figure 2.4 shows the longitude-time Hovmöller diagram of root mean difference total energy (RM-DTE; as defined in Melhauser and Zhang 2012), which is a combined measure of errors in horizontal winds and temperature. As forecast time progresses, the overall error increases due to the realistic uncertainties from the LBC. Larger RM-DTE is located near the region of stronger precipitation. There are apparent westward-propagating streaks of RM-DTE that are related to the CCEWs over the entire domain. The errors associated with the IG waves over the Maritime Continent

have diurnal maxima following the peak precipitation, and there is no obvious trend during the two week period. On the other hand, the three successive moist phases of Kelvin waves over the Indian Ocean has increasingly larger errors. Such error growth behavior indicates a flow dependency in predictability which is more limited in the areas of precipitative systems.

Figure 2.5 show the spectra of reference and error energy, $E(k)$ and $R(k)$, for KE, temperature, specific humidity, and precipitation. Let $l = k^{-1}$ be the horizontal wavelength, the spectra are plotted as a function of k but labelled with its corresponding l values. The time-averaged reference KE spectrum (black line in Fig. 2.5a) has a $-5/3$ slope at small scales, and transitions to a steeper -3 slope at around $l = 500$ km toward larger scales. According to Lorenz (1969) and Rotunno and Snyder (2008), there will be a predictability limit for the small scales due to its shallower spectral slope. The temperature and humidity spectra (Figs. 2.5b and 2.5c) have a similar shape compared to the KE spectrum, while the precipitation spectrum (Fig. 2.5d) is much shallower. Error energy spectral evolution can be viewed from the colored lines with time progresses from blue to red. At $t = 0$, the initial error KE from the ECMWF analysis is about one order of magnitudes smaller than the reference energy at large scales, and becomes comparable with reference energy at intermediate scales around $l = 500$ km. The initial error is artificially too small at small scales, because the ECMWF analyses are archived at a relatively coarse resolution (32 km) thus not fully resolving $l < 200$ km. The error growth is more or less linear at larger scales, while the small-scale error saturates almost immediately. At small scales, the error energy exceeds the time-averaged reference

energy, indicating the loss of predictability. Compared to other variables, precipitation has a wider range of scales with lost predictability after several days of forecast.

To further illustrate the time evolution of errors at different scales, the reference and error energy are averaged over three arbitrarily selected scale ranges: large ($l > 2000$ km), intermediate ($200 < l < 2000$ km), and small ($l < 200$ km), and their time series are shown in Fig. 2.6. The errors in KE, temperature and humidity have similar multi-stage growth behavior as described in Zhang et al. (2007). Small-scale errors grow the fastest in the first 12 hours, then after saturation they stay at relatively the same level. Large- and intermediate-scale errors grow slowly during the whole simulation period, and the large-scale error never reached reference energy for KE. The predictability limit not only depends on the rate of error growth, but is also complicated by the variations in reference energy. For temperature, the reference energy has a clear diurnal cycle at both large and small scales. The large-scale reference energy also appear to be modulated by the low-frequency waves, the humidity reference energy has three minima that are associated with the precipitation episodes over the Indian Ocean. For precipitation, the small- and intermediate-scale error energy exceeds their corresponding reference energy very early in the simulation, and the large-scale error energy is comparable to the reference after 4 days of simulation but the exact time for loss of predictability is uncertain due to the constantly varying reference energy.

Figure 2.7 plots the estimated practical predictability limits (thick lines) as a function of horizontal wavelength. At large scale, the KE and temperature have practical predictability limits up to 15 days, followed by the specific humidity that has a limit of 8 days, and practical predictability of precipitation is limited to only 3 days. The

predictability of all variables dropped significantly across the intermediate scale. The practical predictability is limited to less than 12 hours for KE, temperature, and specific humidity for scales $l < 200$ km, and $l < 800$ km for precipitation.

2.5 Limit of intrinsic predictability

In this section, the intrinsic predictability limit is identified by investigating the second ensemble simulation with IC and LBC error energy reduced to 1%, a level of accuracy that is unlikely attainable in the foreseeable future. Figure 2.8 shows the time series of error energy of this new ensemble (red) for KE and precipitation at three scales and compares their error evolution to the original ensemble with 100% error (blue). When error energy is reduced to 1%, the large- and intermediate-scale error KE (Figs. 2.8a and 2.8b) still grows at a similar rate as the 100% error case during most of the simulation period, although there is some indication of increased error growth rate during the first 3 days. By the end of simulation, error KE from the 1% error case remains an order of magnitude lower than that from the 100% error case. This indicates that the intrinsic predictability for these scales are likely achievable beyond the 15-day simulation period, since the predictability horizon can be extended by reducing the IC and LBC errors.

On the other hand, the small-scale error KE (Fig. 2.8c) grows much more rapidly and the reduced error only delay the loss of predictability by about 1 day. This behavior is as expected accordingly to Rotunno and Snyder (2008) because of the $-5/3$ power law of small-scale KE. The intermediate-scale KE error growth rate is higher at the beginning

of the simulation, because part of the intermediate scale range has a shallower KE spectrum ($200 < l < 500$ km). The temperature and humidity error energy results for the 1% error case are similar to KE (not shown), and precipitation (Figs. 2.8d~f) also displays similar behavior but its range of scales that have limited predictability is much wider than other variables. The intrinsic predictability limits are also plotted as thin lines in Fig. 2.7 as a function of horizontal wavenumber. For the $-5/3$ power law range ($l < 500$ km), theories predict that predictability will be limited. The results are consistent with theory that reducing the IC and LBC error perturbations to 10% does not increase the predictability limit 10 times. The intrinsic predictability limit for KE is about 10 days at $l = 500$ km and decreases to < 1 day at small scales. The same predictability limit is true for other variables such as temperature and humidity except for precipitation which has more limited intrinsic predictability.

For regional models, the specification of LBC is nontrivial for the accuracy of simulation. To evaluate the relative importance of IC and LBC, an extra set of ensemble simulation is conducted with perturbations (from 100% error) only added to the IC, and the results are plotted in Fig. 2.8 as gray lines (IC error only). For large- and intermediate-scale KE, the LBC error contributes a lot more to the overall error growth, without LBC error, the error energy remains at similar IC error level throughout the simulation. However, for small-scale KE, a correct LBC does not help to reduce the initial error growth; it only slightly reduces errors later in the simulation when they are already saturated. For precipitation, the large-scale error energy is reduced by specifying a correct LBC, while the intermediate- and small-scale errors are not significantly reduced before saturation. The intermediate-scale precipitation error is only occasionally

lowered with the correct LBC, indicating a mixed influence from boundary and local region.

Similar to the definition of error doubling time (Lorenz 1969), an “error growth time” is defined as the time in which the initial error energy grows 2 orders of magnitude (from 1% to 100% initial error energy). Figure 2.9 plots the error growth time as a function of horizontal wavenumber. For $l > 500$ km, the error growth time exceeds 15 days for KE and temperature, and ~ 10 days for humidity. The error growth rate rapidly increases as l decrease from 500 km to 100 km, for $l < 100$ km the error growth time is well below 6 hours. The precipitation error growth rate is much higher than other variables for the large and intermediate scales. However, its error growth rate increases more smoothly across scale, unlike other variables that have a rapid growth rate boost within a narrow scale band.

2.6 Predictability of CCEWs at different scales

In the previous sections, the practical and intrinsic predictability limits are estimated for different model variables and for different spatial scales. To identify the underlying processes for the error growth at different spatial and temporal scales, a Wheeler-Kiladis space-time spectral analysis is conducted in this section to extract CCEWs from the simulation, and study the predictability associated with each wave mode.

Figure 2.10 shows the WK spectra for precipitation from the control simulation. Both the zonal wavenumber and time frequency axes are shown in log scale. The spatial

scale is again separated into large (L), intermediate (M), and small (S) scales similar to previous sections, the specified temporal and spatial scale windows for the CCEWs are indicated with black boxes. Along with the MJO signal, the Rossby ($n = 1$ ER), Kelvin, MRG, 2-day WIG ($n = 1$ WIG_L), and EIG ($n = 0$ EIG_L) waves all reside in the large scale. At intermediate scales, the $n = 1$ IG waves (WIG_M and EIG_M) with diurnal to semidiurnal period are the dominant wave modes. The small scale has no clearly identifiable wave signals. Precipitation signals related to small-scale moist convection spread throughout the small-scale spectrum, and according to the previous sections its predictability is intrinsically limited to < 1 day.

A space-time band-pass filter is applied to the precipitation field to extract each wave mode according to its period and wavelength window as shown in Fig. 2.10. Since the perturbed ensemble forecast is performed only for a shorter 15-day period, the control simulation is used to fill in the missing days before filtering. Figures 2.11a~e show longitude-time plots for the filtered precipitation associated with large-scale waves. Spaghetti plots of a selected precipitation contour (in black) among members from the first ensemble (100% error) are shown in Figs. 2.11f~j, and they are compared to the 1% error cases in Figs. 2.11k~o. For the ensemble with 100% error sampled from the ECMWF forecasts, the precipitation contours diverge and become out of phase. However, when the error energy is reduced to 1%, the contours remain in phase among members throughout the whole simulation. The same analysis is performed for the intermediate-scale IG waves and results are shown in Fig. 2.12. In contrast to large-scale wave results, the error reduction does not bring the contours in phase and large displacement errors still exist for intermediate-scale waves. These results agree with the

findings from previous section that the estimated large-scale predictability for precipitation can potentially be extended from 3 days to ~15 days if errors are reduced to 1%, while for intermediate scales its predictability remains intrinsically limited (Fig. 2.7).

Table 2.1 list the averaged pattern correlations between perturbed and unperturbed simulations as a more quantitative measure of CCEW phase errors. A correlation of 1 indicates that two waves are perfectly in phase, while zero correlation indicates two waves completely out of phase. For precipitation, the large-scale CCEWs all show that a significant improvement in wave phase (from ~0.6 to ~0.8 correlation) is possible, which is contrasted by the intermediate-scale IG waves that stay out of phase (correlation < 0.3) even with reduced error. For other variables, results from the previous section show that the practical predictability for KE, temperature and humidity is much less limited than precipitation at large and intermediate scale. Therefore, the zonal wind, temperature and humidity phases associated with Rossby, Kelvin, and MRG waves can potentially be improved to almost perfect (correlation > 0.9) with reduced error. The intermediate-scale predictability for zonal wind, temperature and humidity is more limited, but less so than precipitation.

2.7 Summary

In this study, the October 2011 MJO active phase is simulated using the WRF model with similar configuration as Wang et al. (2015). The control simulation is initialized with ERA-Interim data. The model faithfully reproduced most of the large-scale features of the MJO and CCEWs. Twenty perturbed simulations are conducted for

the 15-day period from October 18 to November 2 to estimate the practical predictability of the multi-scale tropical weather. The IC and LBC perturbations are sampled from the ECMWF global ensemble forecasts from the TIGGE archive. Predictability limit is defined as the time in which error energy (ensemble variance) reaches/exceeds the reference energy (energy associated with the ensemble mean field). Intrinsic predictability is identified by another set of perturbed simulations with the IC and LBC error energy reduced to 1%. Two-dimensional spectral decomposition is applied to the error and reference energy to reveal the horizontal scale dependency in predictability. The predictability limits are calculated separately for kinetic energy, temperature, specific humidity, and precipitation at each scale. Findings from this study are summarized as follows.

- 1) The practical predictability is scale- and variable-dependent. For large scale ($l > 2000$ km) horizontal winds and temperature, the practical predictability limit is ~ 15 days; and for humidity the limit is ~ 8 days. The predictability rapidly drops across the intermediate scale ($200 < l < 2000$ km), and for small scales ($l < 200$ km) their practical predictability is limited to < 12 hours. Precipitation has more limited predictability than other variables, its large-scale practical predictability limit is only 3 days and drops to < 12 hours for the smaller scales.
- 2) Intrinsic predictability limits for horizontal winds, temperature and humidity are > 10 days for scales larger than 500 km. At these larger scales, the practical predictability horizon can be well extended by reducing errors in IC and LBC. However, for scales smaller than 500 km, the intrinsic predictability limit decreases, at < 100 km scales the limit is < 1 day, which is likely related to a shallower $-5/3$

power law in KE spectrum. For precipitation, a wider range of its smaller scales has intrinsically limited predictability comparing to other variables.

- 3) Error growth rate is low at large scales and high at small scales, the increase takes place in a very narrow scale range (100 ~ 500 km) for horizontal winds, temperature and humidity. On the other hand, precipitation error growth rate increases more smoothly across scales.
- 4) Large-scale CCEWs, i.e. Rossby, Kelvin, MRG and the 2-day IG waves, have a predictability that can be potentially improved by reducing the IC and LBC errors. With errors reduced to 1%, the zonal wind, temperature, and humidity associated with large-scale waves can be improved to almost perfectly in phase. However, for the intermediate scale, the diurnal and semidiurnal IG waves have a predictability that is more intrinsically limited.

Table 2.1. Averaged pattern correlation between member (perturbed run) and the control (unperturbed run) for zonal wind, temperature, specific humidity and precipitation associated with each CCEW mode. The improvement in pattern correlation from the 100% error case to the 1% error case is indicated by an arrow.

CCEW mode	zonal wind	temperature	specific humidity	precipitation
Rossby	0.84 → 0.99	0.85 → 0.99	0.72 → 0.97	0.75 → 0.88
Kelvin	0.77 → 0.97	0.82 → 0.97	0.67 → 0.92	0.65 → 0.85
MRG	0.75 → 0.95	0.64 → 0.94	0.65 → 0.94	0.64 → 0.86
WIG_L	0.51 → 0.86	0.51 → 0.86	0.46 → 0.81	0.45 → 0.74
EIG_L	0.51 → 0.85	0.64 → 0.92	0.49 → 0.83	0.46 → 0.70
WIG_M	0.27 → 0.43	0.38 → 0.50	0.27 → 0.44	0.17 → 0.28
EIG_M	0.16 → 0.27	0.33 → 0.45	0.17 → 0.32	0.09 → 0.17

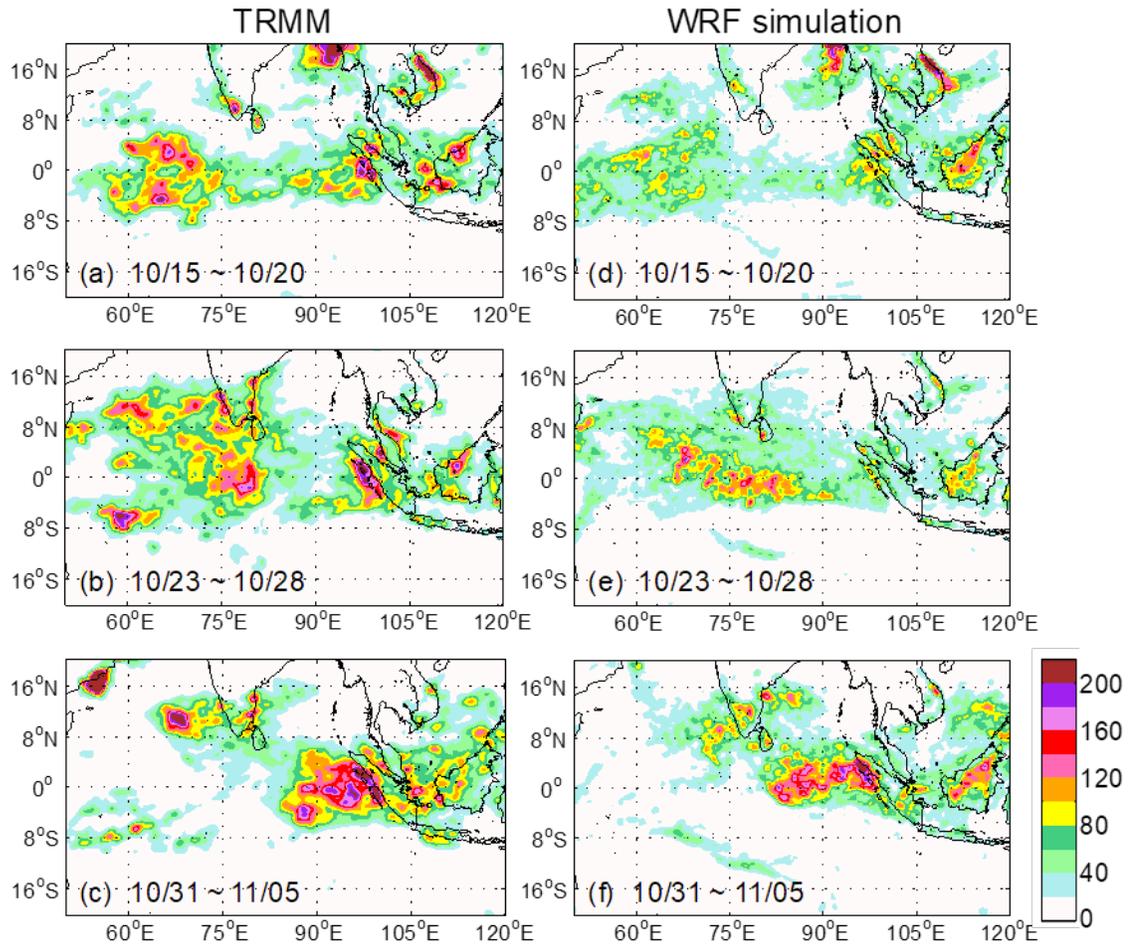


Figure 2.1 Horizontal maps of 5-day accumulated precipitation (mm) on October 20 (a, d), October 28 (b, e), and November 5 (c, f). The results are compared between (a)-(c) TRMM observation and (d)-(f) WRF control simulation.

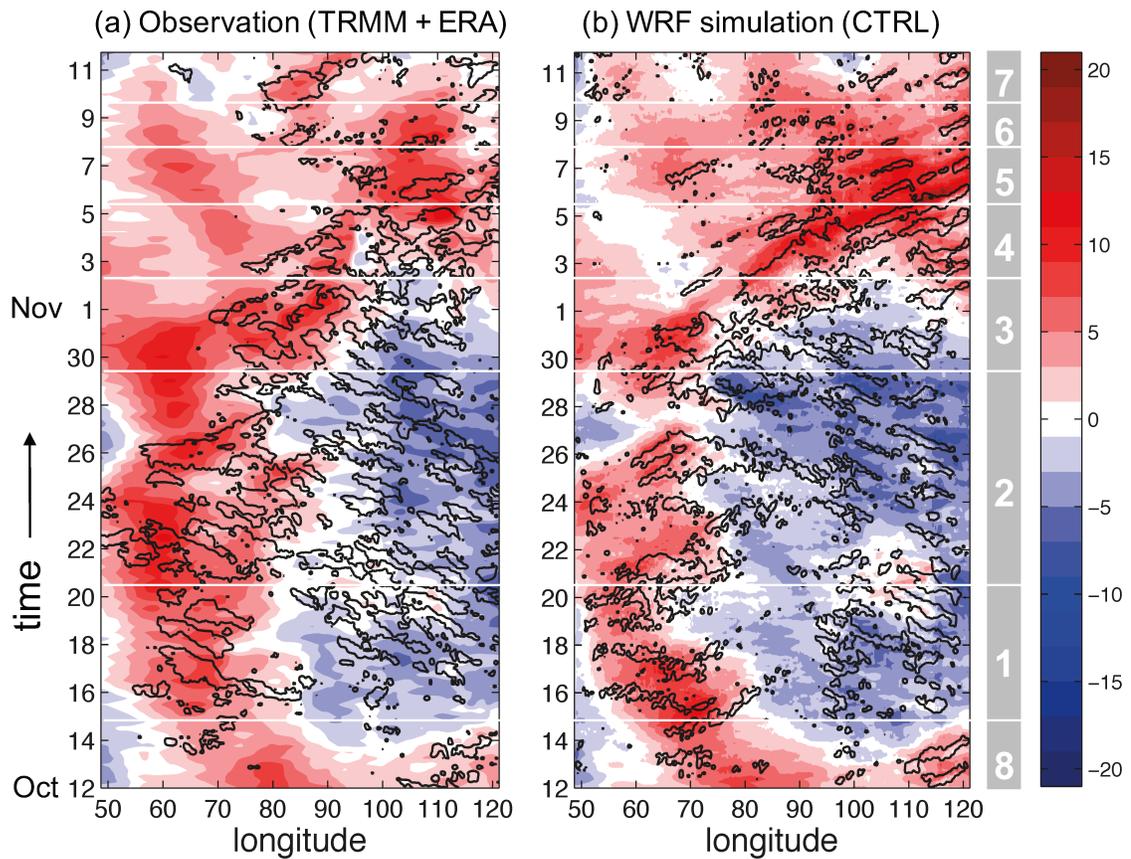


Figure 2.2. Longitude-time plots of 850-hPa zonal wind (color shadings from -20 to 20 m s^{-1}) and precipitation (black contours of 15 mm day^{-1}) averaged over 5°S - 5°N . The results are compared between (a) TRMM precipitation, ERA-Interim zonal wind, and (b) WRF control simulation. The white grid and numbers to the right indicate the observed phase of the Oct 2011 MJO (according to Fig. 6 from Johnson and Ciesielski 2013). The precipitation and wind fields are plotted at 3-h intervals, except for ERA wind that is at 6-h intervals.

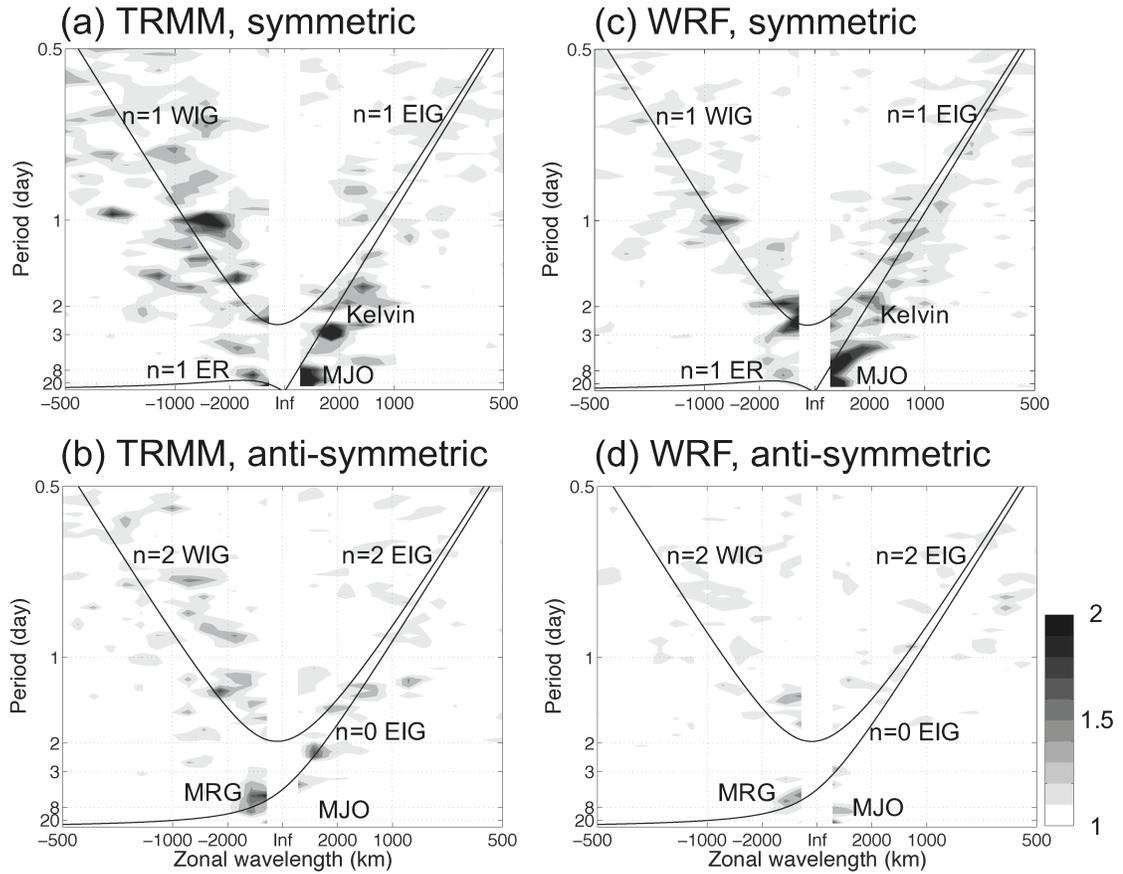


Figure 2.3. Wheeler-Kiladis space-time spectra of precipitation from (a, b) TRMM observation and (c, d) WRF control simulation averaged over 15°S - 15°N that is (a, c) symmetric and (b, d) anti-symmetric about the equator. Signal strengths from 1.1 to 2 are shown in shadings. The solid curves correspond to dispersion relations for dry equatorial waves with equivalent depths of 15 m. The zonal wavenumber and time frequency are labelled with corresponding zonal wavelength and time period, respectively, positive (negative) wavelength indicates eastward (westward) propagating signals.

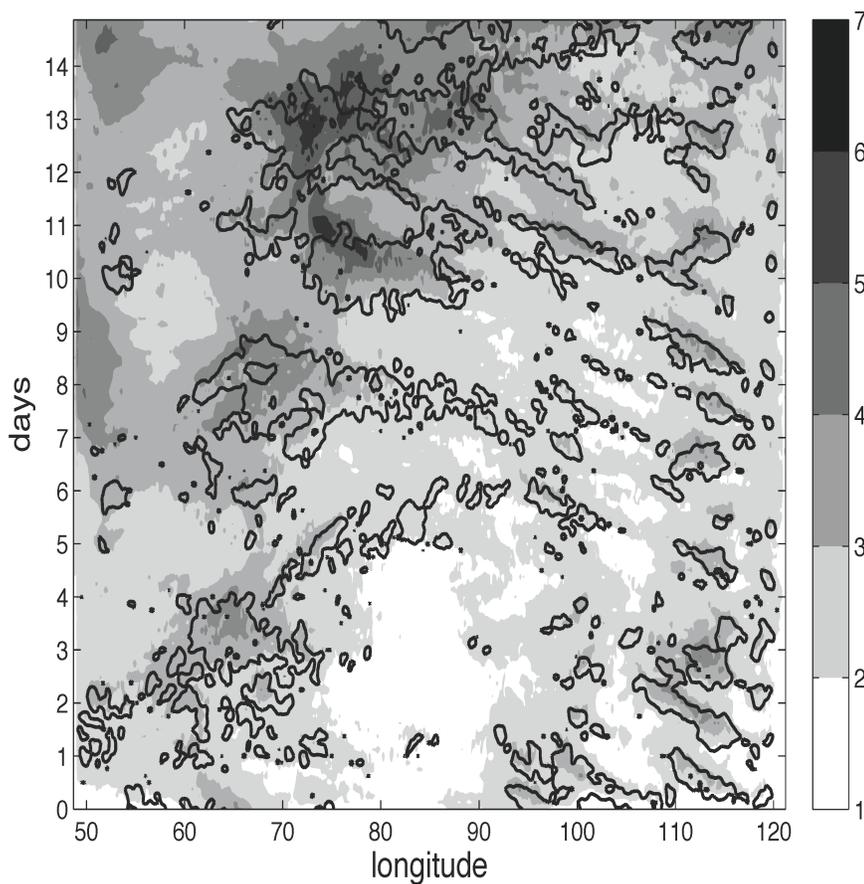


Figure 2.4. Longitude-time plot of RM-DTE (m s^{-1} , shadings) at 850 hPa and the precipitation from control simulation (black contour of 15 mm day^{-1}) averaged over 5°S - 5°N . The RM-DTE is the square root of ensemble-averaged DTE between the perturbed ensemble simulations and the control simulation from October 18 to November 2 (shown as $t = 0$ to 15 days).

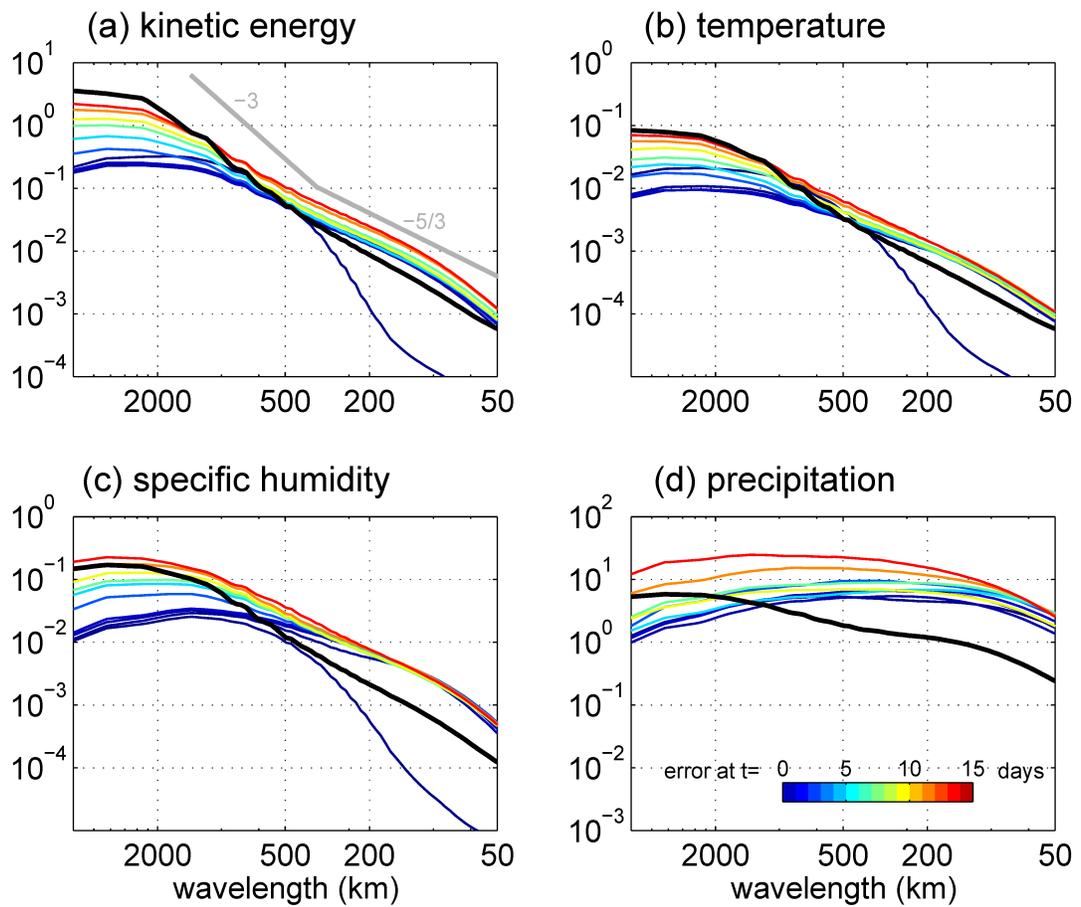


Figure 2.5. Reference energy spectra, $R(k)$, averaged over the 15 days (black lines) and error energy spectra, $E(k)$ (color-coded with simulation time $t = 0$ to 15 days), for (a) kinetic energy, (b) temperature, (c) specific humidity, and (d) precipitation. The shown spectrum is averaged over the vertical levels.

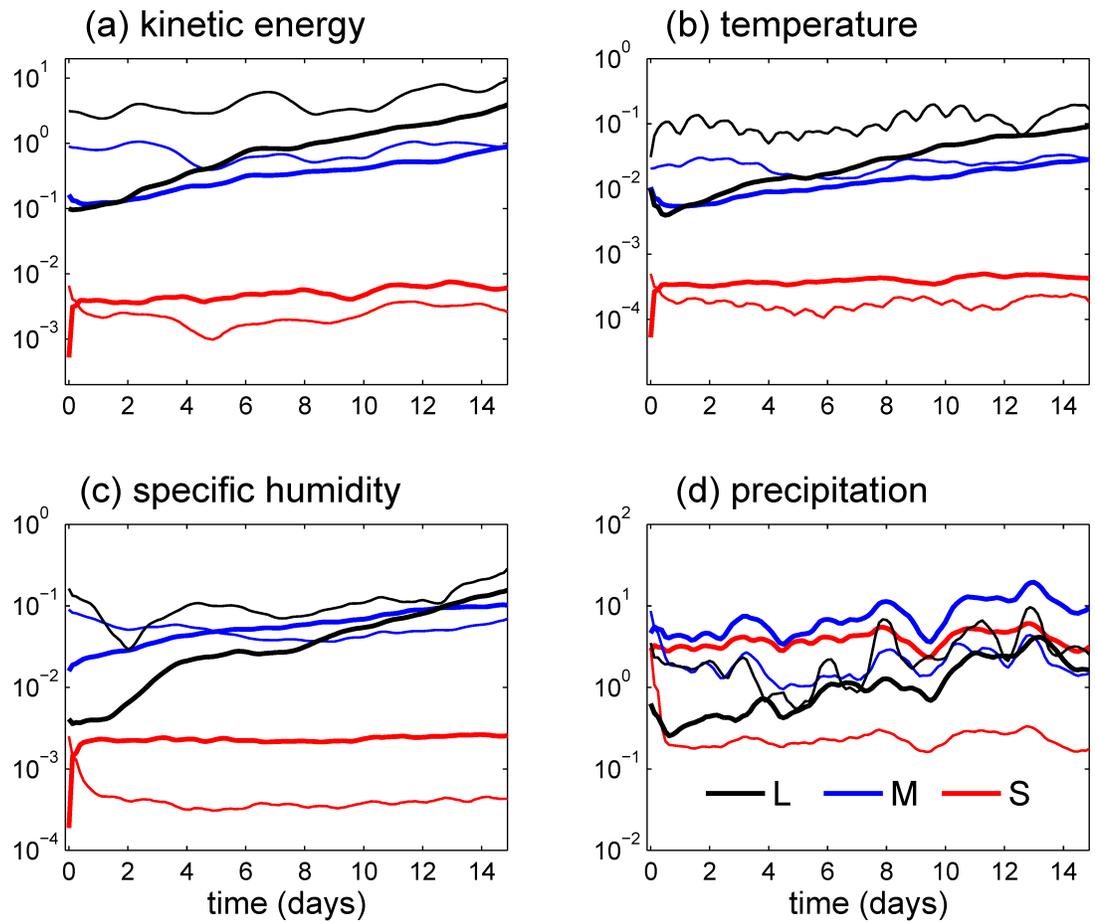


Figure 2.6. Time series of spectral energy averaged at large (L; $l > 2000$ km, black), intermediate (M; $200 < l < 2000$ km, blue), and small (S; $l < 200$ km, red) scales for (a) kinetic energy, (b) temperature, (c) specific humidity, and (d) precipitation. The thick lines show the error energy, while thin lines show the reference energy.

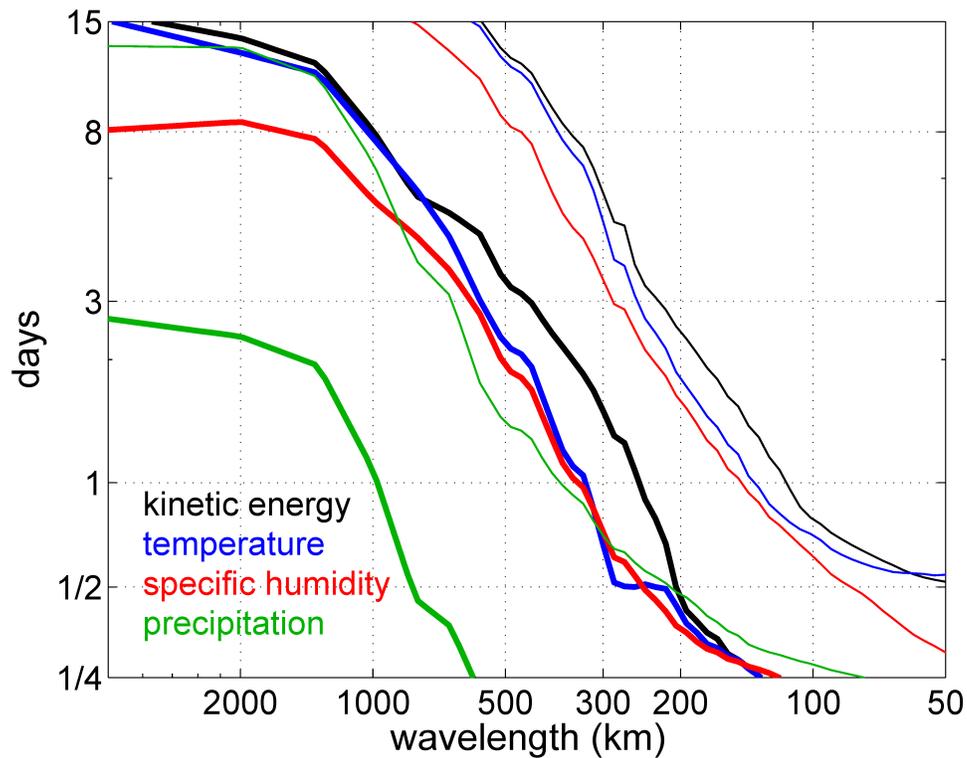


Figure 2.7. Predictability limits for kinetic energy (black), temperature (blue), specific humidity (red), and precipitation (green) plotted as a function of zonal wavenumber (labeled as wavelength). Thick (thin) lines are practical (intrinsic) predictability limits. The limits are defined as the time it takes for the error energy (100% error for practical limit and 1% error for intrinsic limit) to reach reference energy. The limits are calculated for each member and the ensemble average is plotted. Smoothing is applied across wavelength to remove some noise for better visualization.

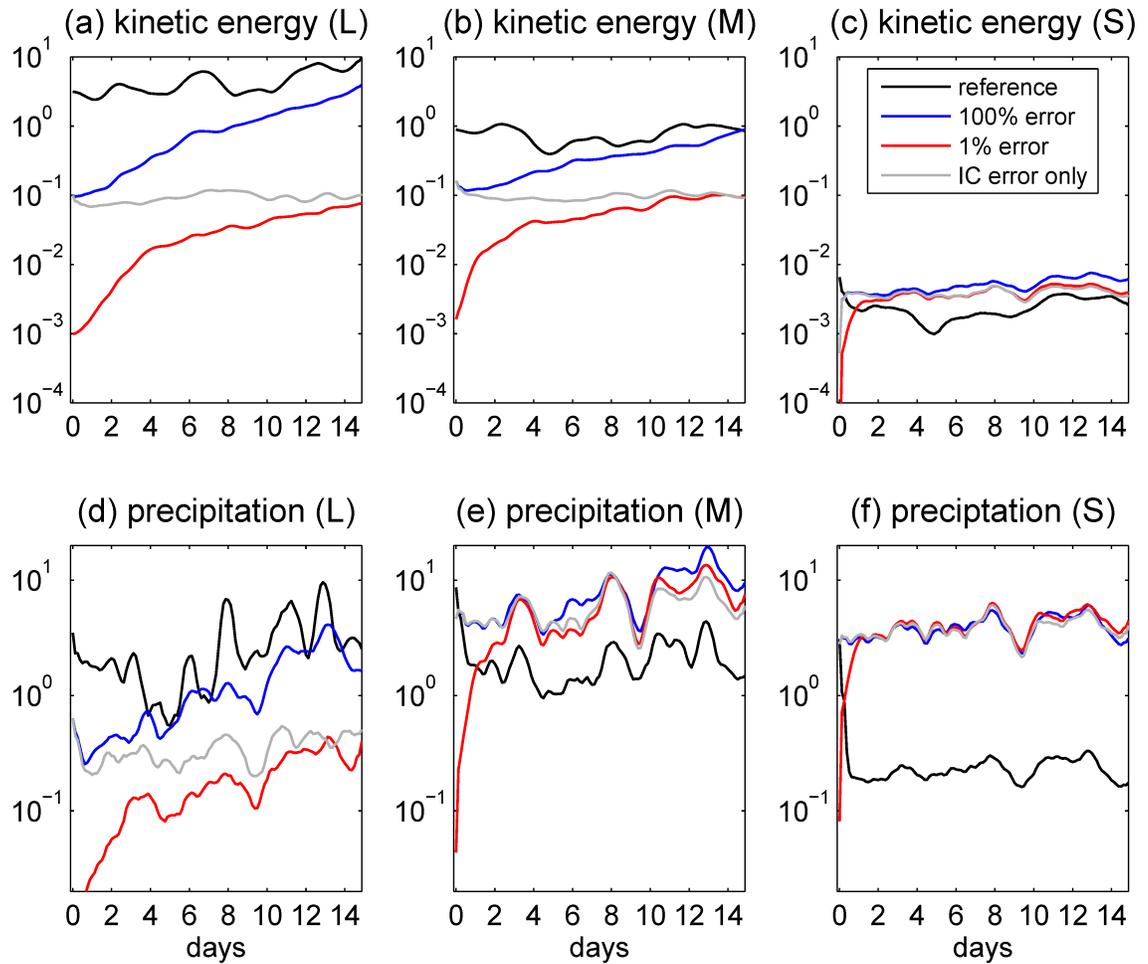


Figure 2.8. Time series of spectral error energy integrated within L (a, d), M (b, e), and S (c, f) scales for (a)-(c) kinetic energy and (d)-(f) precipitation. The reference error energy is shown as black lines, the blue (red) lines show the error energy from the 100% (1%) error case, and the gray lines show the case with errors only in the initial condition.

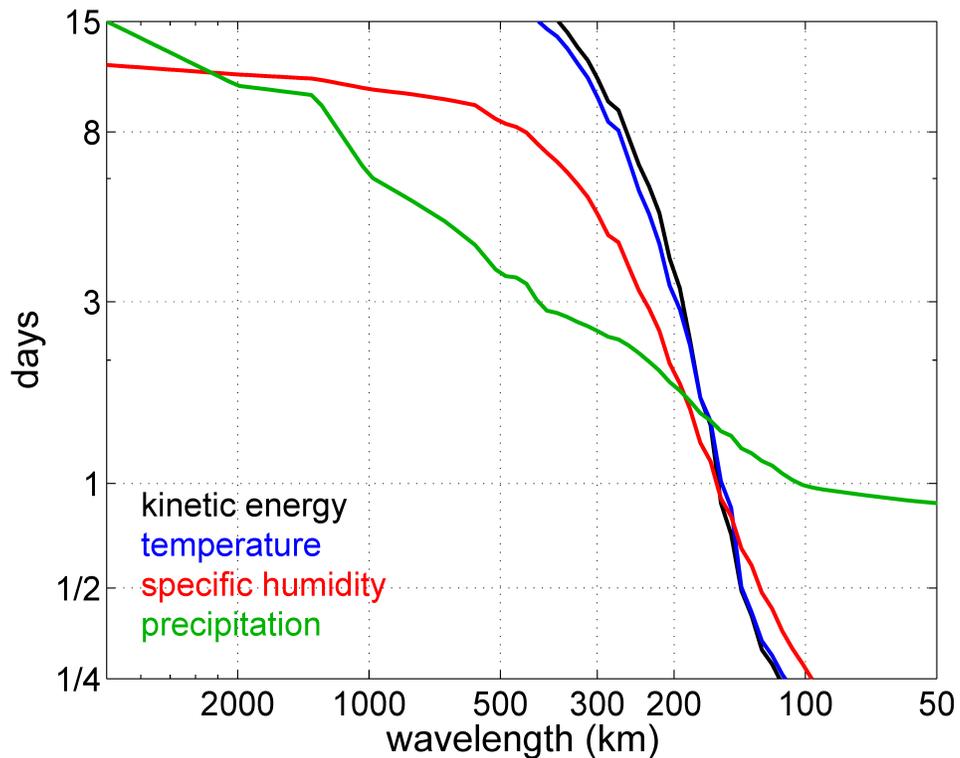


Figure 2.9. Error growth time (days) plotted as a function of zonal wavenumber (labeled as wavelength) for kinetic energy (black), temperature (blue), specific humidity (red), and precipitation (green). The error growth time is defined as the time it takes for 1% initial error to grow and reach the 100% initial error. The growth time is calculated for each member and the ensemble average is plotted.

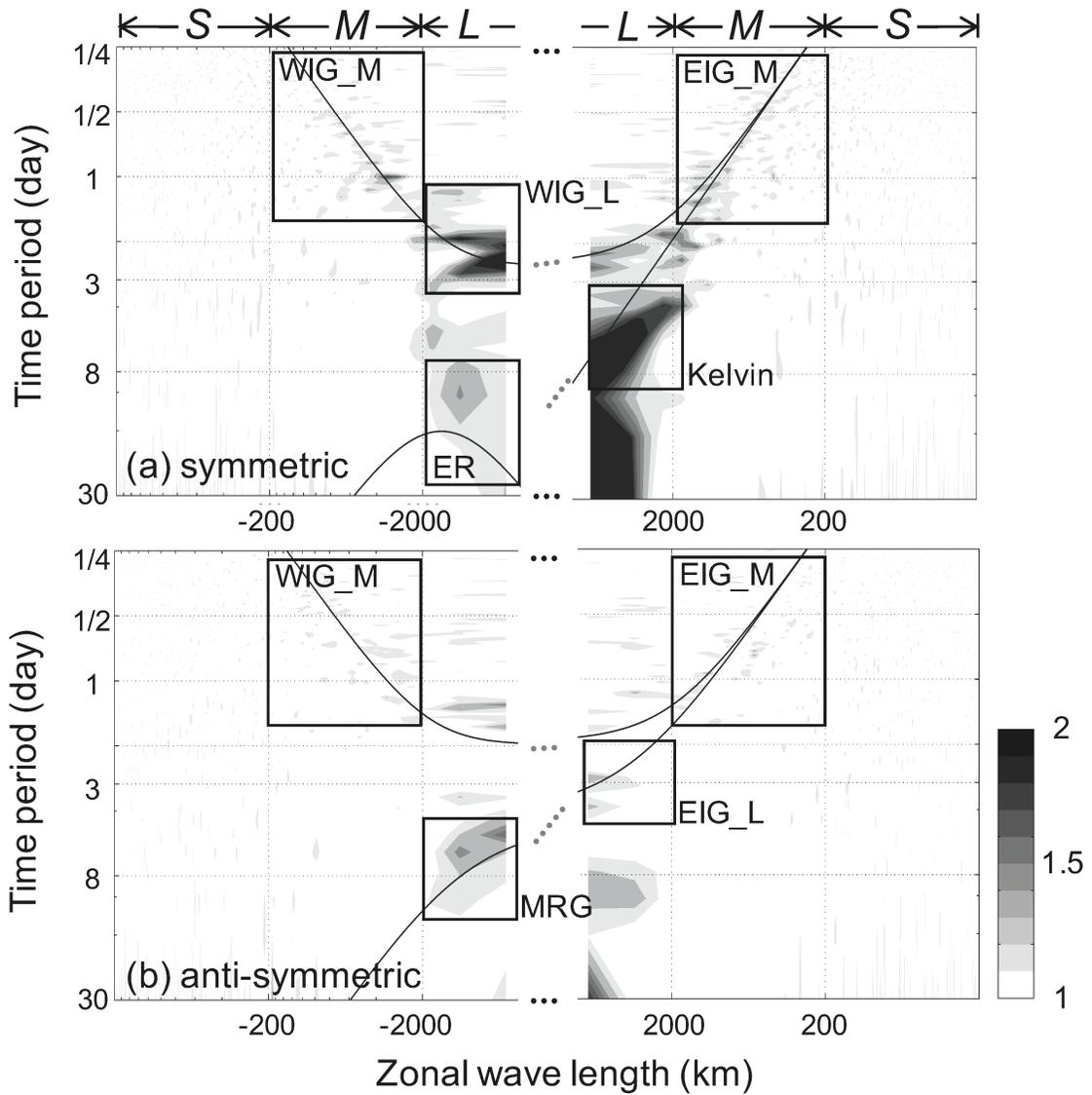


Figure 2.10. Similar to Figs. 2.3c and 2.3d, but with zonal wavenumber and time frequency shown both in log scale. Space-time filtering windows are shown as black boxes for equatorial Rossby wave (ER), Kelvin wave, mixed-Rossby-gravity wave (MRG), $n = 1$ WIG wave at intermediate scale (WIG_M) and at large scale (WIG_L), $n = 0$ EIG at large scale (EIG_L) and $n = 1$ EIG at intermediate scale (EIG_M).

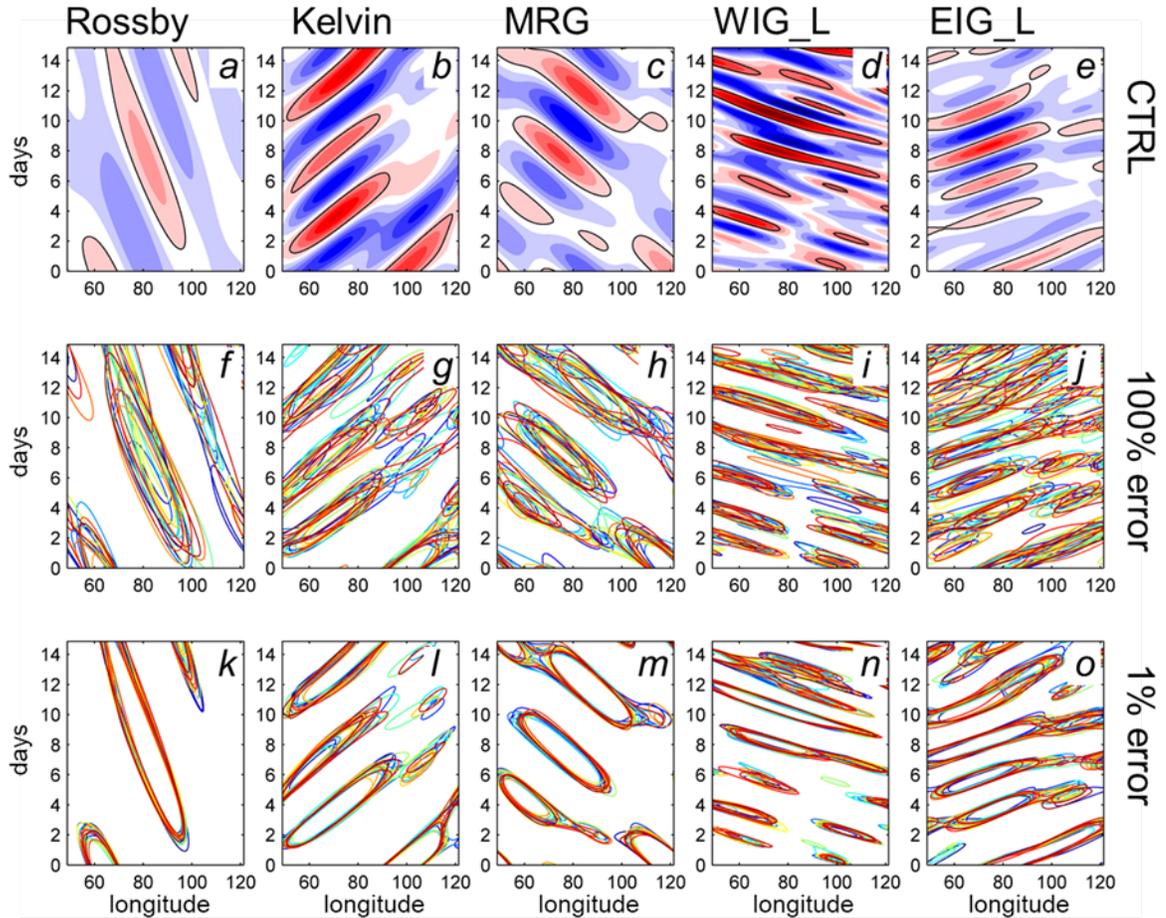


Figure 2.11. Longitude-time plots (a)-(e) of precipitation (color shadings from -10 to 10 mm day^{-1} every 1 mm day^{-1} , zero shown in white) filtered for (a) Rossby, (b) Kelvin, (c) MRG, (d) WIG_L, and (e) EIG_L waves and averaged over $0-5^\circ\text{N}$. The second row (f)-(j) shows spaghetti plots of the contours highlighted in black in the first row, each color corresponds to a member from the perturbed ensemble simulation (100% error case). The third row (k)-(o) is similar to the second row but showing the ensemble simulation with error energy reduced to 1%.

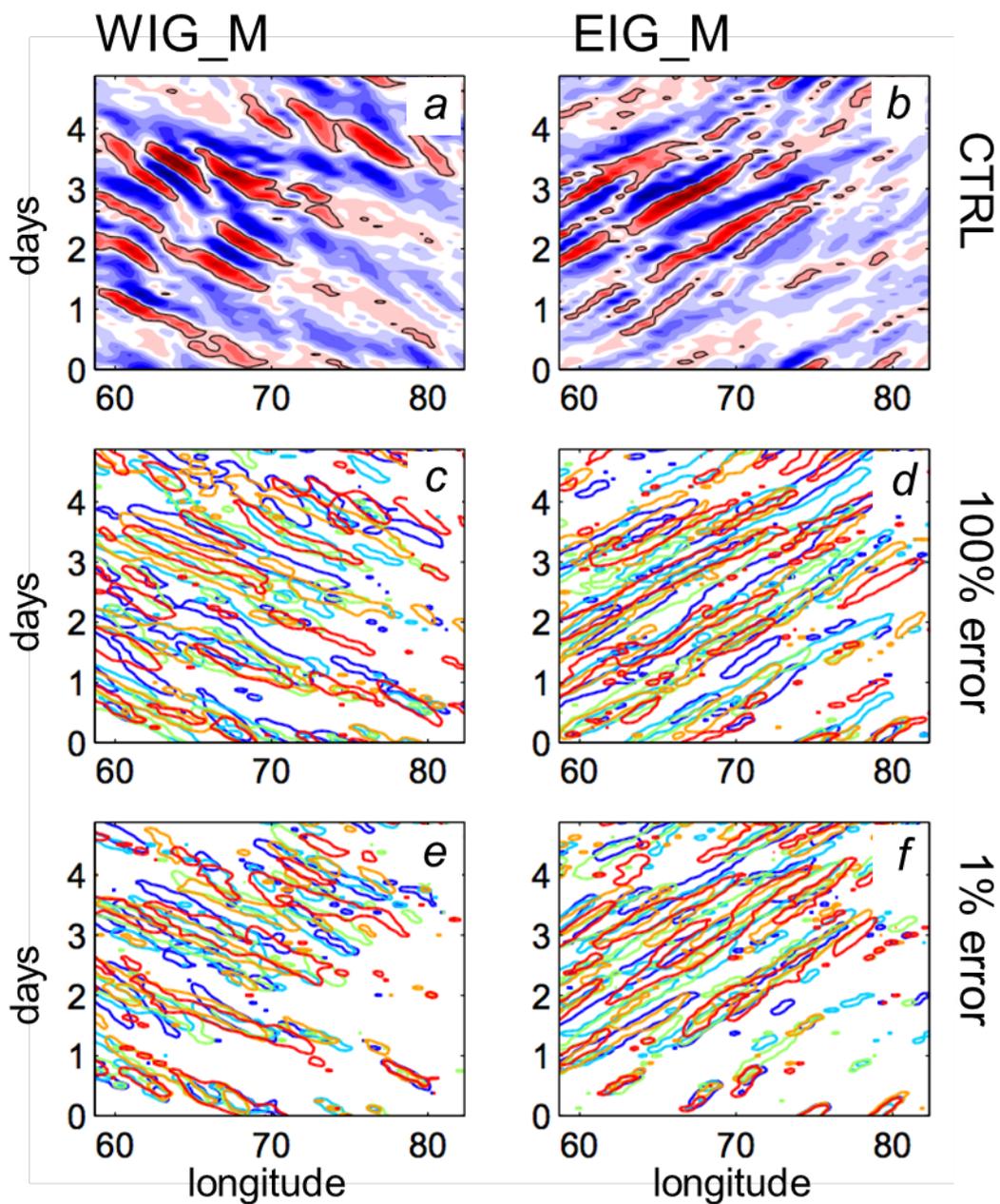


Figure 2.12. Similar to Fig. 2.11 but showing the two intermediate-scale waves, WIG_M and EIG_M. Only 5 members are shown in the spaghetti plots, and the longitude-time plots are zoomed in on 60°-80°E and the first 5 days of the simulation.

Chapter 3

Potentials in improving predictability of multiscale tropical weather systems evaluated through ensemble assimilation of simulated satellite-based observations

3.1 Introduction

Skillful modeling of the multiscale tropical weather systems is an important and challenging task for global numerical weather prediction. Lorenz (1963) first discovered that the atmospheric flow can be intrinsically unpredictable due to its chaotic nature even with near-perfect model and initial/boundary condition. Moist convective processes are identified as responsible for the intrinsic limit of predictability for moist baroclinic waves (Zhang et al. 2003, 2007; Sun and Zhang 2016), mesoscale convective systems (Bei and Zhang 2007; Melhauser and Zhang 2012; Selz and Craig 2015; Zhang et al. 2016), and tropical cyclones (Zhang and Sippel 2009; Tao and Zhang 2015). CCEWs are also susceptible to the moist convective processes that will limit their predictability. Despite such intrinsic limits in predictability, there is still room for improvement in the current prediction skill by assimilating better observations and improving the models. For example, Reynolds et al. (1994) found that the then-operational global prediction system has deficiencies in the tropics that lead to large error growth, although the internal error growth rate is actually slower in the tropics than in mid-latitudes.

Due to the lack of in-situ soundings with high resolution and good coverage, the tropics heavily rely on satellite remote sensing to provide observations for forecast systems. In this chapter, we evaluate the potentials in improving the predictability of tropical weather at different scales through assimilating current and future satellite observing networks. To the best of our knowledge, this study is the first satellite data assimilation experiment focusing on multiscale tropical weather systems beyond tropical cyclones. The MJO active phase simulation from the previous chapter is used as a test case, and an idealized Observing System Simulation Experiment (OSSE) is conducted under the assumption of perfect dynamic model and perfect observation error models used in data assimilation. For observing networks with different spatial density, several studies suggest that a shorter localization distance is required for denser observations to achieve better analyses (Dong et al. 2011; Kirchgessner et al. 2014; Periañez et al. 2014). In this study, the localization for each observing network is manually tuned to allow a relatively fair comparison of their impact.

This chapter is organized as follows. Section 3.2 describes the dynamic model, test case, and data assimilation method. Section 3.3 provides details of the satellite observing networks under investigation. The following sections will discuss results regarding scale- and variable-dependent predictability (section 3.4), the relative contribution from each observing network (section 3.5), and the specific improvement found for each CCEW (section 3.6). Section 3.7 will summarize the findings from this study.

3.2 Experimental design

3.2.1 The verifying truth simulation

A similar configuration of the Weather Research and Forecasting (WRF) model is adapted from the previous chapter to conduct a month-long simulation from 12 October to 12 November 2011. The model domain contains 777×444 horizontal grid points with 9 km spacing, and has 45 vertical levels with 9 in the lowest 1 km and a model top at 20 hPa. Cloud microphysics processes are represented by the WRF Double-Moment 6-class (WDM6) scheme (Lim and Hong 2010). No cumulus parameterization is used, and organized convections are explicitly represented by the 9-km model grid. Shortwave and longwave radiative processes are simulated using the New Goddard (Chou and Suarez 1999) and RRTMG (Iacono et al. 2008) schemes, respectively. Surface processes are simulated using the unified Noah land surface model (Chen and Dudhia 2001) with variable surface skin temperature (Zeng and Beljaars 2005). Subgrid-scale turbulent mixing is treated with the Yonsei University (YSU) boundary layer scheme (Hong et al. 2006). The initial conditions (IC) and lateral boundary conditions (LBC) are specified by the ERA-Interim reanalysis data (Dee et al. 2011). The sea surface temperature (SST) for the lower boundary conditions is updated every 6 h according to the National Centers for Environmental Prediction (NCEP) Final Operational Global Analysis (FNL data).

Figure 3.1 shows the computation domain that covers the precipitation associated with the active MJO phase. The simulation successfully captures the slow eastward propagation of the active phase as well as most of the CCEW modes. **Figure 3.2** shows the Wheeler and Kiladis (1999) space-time spectra for surface precipitation and 850-hPa

zonal wind. For precipitation, the dominant signals are the MJO, Kelvin and westward-propagating IG (WIG) waves. The WIG waves contain a hierarchy of spatial and temporal scales. For spatial scales larger than 2000 km, there is the “2-day wave”, while the diurnal and semi-diurnal waves have wavelengths slightly shorter than 2000 km. The Kelvin wave has a precipitation signal at a shorter wavelength than the wind signal. It is also evident that the propagation speed of the convectively coupled Kelvin wave is slower than the dry Kelvin waves. For ER and MRG waves, the precipitation signal is relatively weaker than their wind signal. The precipitation signal for ER wave also has a shorter wavelength than the wind signal, similar to the Kelvin waves. The signals related to eastward propagating IG waves are not as strong as their westward-propagating counterparts.

Due to the limitation of computation resources, the OSSE is conducted in a smaller domain with 378×222 grid points (10°S to 10°N , 55° to 90°E) during the 17-day period from 16 October to 2 November (phase 1 to 3 of the MJO). The location of this OSSE domain is shown in **Fig. 3.1**. The month-long large-domain simulation provides IC and LBC for the smaller OSSE domain. The OSSE domain has the same 45 vertical levels and 9-km grid spacing as the larger domain.

To extract a certain CCEW mode, a space-time band-pass filter is applied to the month-long large-domain model simulation. **Table 3.1** lists the filter parameters for each CCEW. **Figure 3.3** shows the filtered surface precipitation and 850-hPa zonal wind during the 17-day test period and within the OSSE domain. Hereafter, the time from 16 October to 2 November will be labeled as day 1 to 18. During the test period, the MJO precipitation signal slowly propagates across the entire OSSE domain from west to east.

The low-level westerly wind anomaly is trailing the precipitation peak. The ER wave features two strong westerly wind signals associated with the cyclonic Rossby gyres within the test domain and replaced by easterlies (anti-cyclonic gyre) later into the test period. The Kelvin wave has three successive phases propagating across the test domain during the 17 days, each one progresses more to the east due to modulation from the MJO. MRG waves have a strong wind signal but very weak precipitation signal during the test period. WIG waves consist of a wide range of spatial and temporal scales. Large-scale (>1000 km) WIG signals have a life cycle of about 2 days and propagate across the whole test domain, while intermediate-scale (500-2000 km) WIG signals are more transient in space and time. Both types of WIG waves appear to be modulated by the MJO envelope as well.

3.2.2 Data assimilation method: EnKF

The PSU WRF EnKF (Meng and Zhang 2008) is employed as data assimilation method. The EnKF utilizes an ensemble of model realizations to estimate flow-dependent error covariance, which is an approximation of the extended Kalman filter (Evensen 1994). The ensemble perturbations around the mean are updated with the square root algorithm from Whitaker and Hamill (2002). Covariance relaxation-to-prior-perturbation (Zhang et al. 2004) with $\alpha = 0.8$ inflates the ensemble variance to prevent filter divergence. Covariance localization is applied with the Gaspari and Cohn (1999) tapering function to remedy sampling noises in error covariance.

The OSSE is conducted under a perfect-model assumption, i.e. the forecast model and LBC are the same from the truth simulation. An ensemble of 60 members is created

one day before the data assimilation cycle begins. The initial ensemble perturbations are randomly sampled from the climatological error covariance matrix created by the NMC method (Parrish and Derber 1992). The ensemble mean is created by adding another perturbation to the truth. This perturbation is drawn from the same error covariance but made sure that its magnitude is close to the ensemble spread. The initial ensemble is run forward for one day to develop flow-dependent error covariance.

Synthetic observations are created by randomly perturbing the observed variable simulated from the truth with prescribed observation error. The perfect forward operator is used, and the correct observation errors matching the uncertainties in synthetic observations are specified in EnKF. The prescribed observation errors are unbiased and uncorrelated in this OSSE study. **Table 3.2** lists the prescribed observation error and resolution for each observing network, which will be discussed in more details in next section. **Table 3.3** provides a succinct list of experiments conducted in this paper. NoDA is a free ensemble run from the perturbed initial ensemble. CNTL is a benchmark case assimilating ATOVS temperature (T) and specific humidity (Q) profiles and AMV winds. The observations are assimilated every 3 h during the 17-day test period, which results in 136 data assimilation cycles in total. We will show that the EnKF spin-up period appears to be shorter than 4 days, therefore some of the sensitivity experiments are only conducted for the first 9 days. The localization cutoff distance, i.e. the radius of influence (ROI), is specified separately for each observing network in consideration of the respective observation density. A larger ROI is used for sparser networks, and some manual tuning with a few cycles are performed to reach a reasonable ROI. A sensitivity experiment that doubles the horizontal ROI in CNTL confirmed that the resulting change

in analysis error is minor. As pointed out by Zhen and Zhang (2014) and Lei and Whitaker (2017), the analysis sensitivity to localization is minor when approaching the best-performing ROI. Thus, we will not further discuss how the exact choice of ROI changes analysis accuracy in this paper.

3.3 Observing networks

3.3.1 Synthetic ATOVS retrieval products

The Advanced TIROS Operational Vertical Sounder (ATOVS) is a collection of polar-orbiting satellite instruments that are designed to sample the atmospheric T and Q at different vertical levels. These instruments currently include the High-Resolution Infrared Sounder (HIRS), the Advanced Microwave Sounding Unit-A (AMSU-A) and Microwave Humidity Sounder (MHS) on board NOAA-19, MetOp-A and MetOp-B satellites. Operational centers routinely produce atmospheric sounding products from the ATOVS instruments using sophisticated retrieval methods (Reale 2001), typically including cloud detection, radiative transfer, and other physical process-based algorithms. The T and Q profiles can be retrieved for both clear and cloudy atmospheric conditions. Although direct assimilation of ATOVS radiances may yield more realistic results, current regional-scale data assimilation systems (including the one employed here) cannot effectively and simultaneously assimilate all the ATOVS radiances, especially for cloud-affected microwave-channel radiances that are extremely sensitive to parameterizations in model microphysical processes and related representations of scattering properties. Therefore, we assimilate synthetic retrieved T and Q profiles with

regular spatiotemporal resolutions to evaluate the impact from these ATOVS radiances (similar to English et al. 2000), which also serves a baseline for other observing systems.

The spatial and temporal resolution of ATOVS profiles depend on both the availability of satellite instruments and the level of quality control during retrieval. The current profiles have a horizontal spacing of about 40 km. In the vertical, there are 40 layers from surface up to 0.1 hPa for T profiles, and 17 layers from surface up to 200 hPa for Q profiles. Miyoshi et al. (2013) demonstrated how non-orthogonal observation operators reduce the available information from observations. Due to the overlapping of weighting functions of radiances used in the retrieval process, the number of independent pieces of information from real ATOVS profiles is likely to be less than what assumed in this study, the impact of which will be examined with sensitivity experiments through reducing the information content in either the horizontal and/or vertical directions, and in time. The horizontal spacing is set to a coarser 90 km to account for missing data due to gaps between swaths and/or quality control processes. The current three-satellite constellation could sample the whole tropics 6 times every day. This observation interval can be shortened with an increase in number of participating satellites in the constellation. Since year 2015, with the launch of the Global Precipitation Measurement (GPM) mission (Hou et al. 2014), it has been estimated that there are enough satellites to achieve an average revisiting time of 1.5 h. Therefore a 3 h cycling period is reasonable for ATOVS in this study. **Figure 3.4a** shows the vertical profile of prescribed ATOVS observation errors in solid lines. Li et al. (2000) validated the ATOVS retrievals with respect to radiosonde observations, and found the accuracy of retrieved T is about 2 K and dewpoint temperature is 3-6 K (about 30% error in Q). The horizontal distribution of

ATOVS profiles is shown as black dots in **Fig. 3.4b**. The horizontal localization radius (ROI) is set to 400 km. The vertical ROI is set to 5 levels, which correspond to ~500 m near surface and increases to ~3000 m near tropopause.

3.3.2 Synthetic GPSRO observations

The Global Positioning System Radio Occultation (GPSRO) samples the atmosphere with limb radio signals transmitted by the GPS satellites and received by low-Earth-orbit (LEO) satellites. Atmospheric T and Q can be retrieved from the refraction of GPS signals. Since radio signals are not influenced by particulates, the retrieval can be performed for both clear and cloudy weather conditions. The limb sounding geometry provides a higher vertical resolution than the passive nadir sounders (Kursinski et al. 1997; Yunck et al. 2009). The achievable vertical spacing is ~100 m for the low troposphere and reduce to 1 km into stratosphere for both temperature and humidity. The horizontal density of retrieved profiles depends on the number of LEO satellites in operation. For the second Constellation Observation System for Meteorology, Ionosphere and Climate (COSMIC2) mission (Cook et al. 2011), 12 LEO satellites will be launched that provide more than 8000 profiles globally within a 3 h window. The retrieval products also have better accuracy than ATOVS. Wang et al. (2013) validated the GPSRO profiles with global radiosondes and found that T error is 1.72 K and Q error is 0.67 g kg^{-1} . **Figure 3.4a** shows the GPSRO error profiles as dashed lines in comparison to ATOVS error profiles.

In this study, synthetic GPSRO profile location is randomly specified. At COSMIC2 density, there are about 20 profiles in the OSSE domain. A much higher

horizontal density may be achieved with the future launch of CubeSat (Mannucci et al. 2010). Assuming a 300-satellite CubeSat constellation, there will be 25 times more profiles in the domain. **Figure 3.4b** compares the horizontal distribution of GPSRO at COSMIC2 and CubeSat density to the ATOVS profile grid. The horizontal ROI is set to 1600 km for GPSRO profiles at COSMIC2 density, while a shorter 400 km is used for denser CubeSat GPSRO profiles. Vertical ROI is set to 5 levels, the same as for the ATOVS profiles.

To compare the assimilation of retrieved T and Q profiles with the direct assimilation of refractivity (N) data, synthetic N profiles are generated at the same location of the T and Q profiles according to the following equation (Smith and Weintraub 1953).

$$N = 77.6 \frac{p}{T} + 3.73 \times 10^5 \frac{e}{T^2}, \quad (3.1)$$

where p is pressure, T is temperature and e is water vapor pressure. According to Kursinski et al. (1997), the observation errors for N is $\sim 0.2\%$ above 500 hPa and increases to $\sim 1\%$ at lower levels. Since the observation operator is nonlinear, a shorter localization distance is specified. The horizontal ROI is 600 km for coarser network and 200 km for denser network. Vertical ROI is set to 3 levels.

3.3.3 Synthetic AMV wind

The Atmospheric Motion Vector (AMV) is derived from a sequence of geostationary satellite images by tracking features such as cloud edges and water vapor

gradients (Nieman et al. 1997). A cross-correlation tracking algorithm locates sharp gradients in raw satellite images, and their displacement vectors are retrieved as wind observations. Heights are assigned to each AMV according to its observed brightness temperature. Operational centers now produce routine AMV wind products from the Geostationary Operational Environment Satellite (GOES) images. NCEP GFS assimilates AMVs in real time every 6 h with relatively strict quality control processes. CIMSS prepares hourly AMV products with more detailed coverage over tropical regions (Velden et al. 1997). A “rapid scan” mode can be switched on to achieve even higher temporal frequency (every 15 minutes). Velden and Bedka (2009) performed a careful comparison of large volumes of AMV data with collocated rawinsonde wind profiles, and estimated that AMV observation errors are about 5 m s^{-1} and the height assignment is the dominant error source (up to 70%).

To generate synthetic AMV observations, instead of simulating the cloud detection and tracking algorithms, the real AMV observation locations from GTS dataset are directly used to interpolate wind from truth simulation. Such treatment does not guarantee the synthetic AMVs to be located at the simulated clouds, especially for cloud features at smaller scales. The AMVs collocated with real data may be located in the simulated clear-air regions. At small scales, such biased coverage will introduce flow dependence in results, and the results may be different from a real-case scenario. However, for larger scales, the results are less influenced since simulated cloud clusters agrees well with observations. More important is that such procedure yields synthetic AMVs that match the spatial resolution and distribution of the current real-world data. **Figure 3.4c** shows the number of AMVs at each vertical layer. Most of the AMVs are

located at upper levels (~ 200 hPa), and only a small fraction is at low levels (~ 750 hPa). At upper levels, the number of AMVs is comparable to ATOVS profiles (~ 800). **Figure 3.4d** shows the horizontal distribution of AMVs. The horizontal ROI is set to 400 km. The vertical coverage is limited to where the cloud or vapor gradient is located, therefore a larger vertical ROI of 15 model levels is used.

3.3.4 Synthetic Meteosat-7 infrared brightness temperature

Among the first-generation geostationary satellites, Meteosat-7 (Met7) has coverage to the Indian Ocean. It provides full-disk brightness temperature (Tb) every 30 minutes from two infrared channels and a visible channel. The horizontal observation spacing is 5 km. Starting from 2017, Met7 will be replaced by the second-generation Meteosat-8, which has more channels and increased resolution (similar to GOES-R that achieves a spatial grid spacing of 2 km every 15 minutes). Although clear-air satellite Tb is routinely assimilated in operational centers, the direct assimilation of cloud-affected Tb is still quite challenging. Recent development in methods that remedy the non-Gaussian observation errors makes the assimilation of all-sky Tb much more effective (Geer and Bauer 2011; Tavolato and Isaksen 2015; Zhang et al. 2016; Minamide and Zhang 2017). The Tb from infrared channels are easier to assimilate than those from microwave channels thanks to less sensitivity to scattering processes that are not well represented in current microphysics schemes and radiative transfer models. In this study, the water vapor band Tb (channel 3) is assimilated directly using the CRTM as forward operator. The thermal infrared window (channel 2) is preserved for verification purposes. The Adaptive Observation Error Inflation (AOEI; Minamide and Zhang 2017) is applied to

reduce the negative effect from clear/cloudy-air representativeness errors. Horizontal ROI is set to 30 km. The vertical height of a Tb observation is specified according to its cloudy/clear air condition. A cloudy Tb is assigned to 250 hPa and a clear-air Tb to 400 hPa (peaks of their weighting function). The vertical ROI is set to 25 levels around the Tb observation at its specified height so that the localization function is similar to its weighting function.

3.3.5 Synthetic CYGNSS surface wind speed

The Cyclone Global Navigation Satellite System (CYGNSS) is a constellation of 8 micro-satellites that receive both direct and reflected signals from GPS satellites (Ruf et al. 2016). The direct signals pinpoint the location of a CYGNSS observation, and the reflected signals carries information of the roughness of ocean surface, from which wind speed is retrieved. The relatively low inclination orbits of the microsattellites are designed to provide excellent coverage for the tropics. The observation footprint is 25 km wide with a mean and median revisit time of 7.2 and 2.8 h, respectively. The observation error is about 2 m s^{-1} or 10% for a wide dynamic range of wind speed. The synthetic CYGNSS observation is generated at 45 km grid spacing every 3 h. Instead of locating the observations along the realistic orbits that are horizontally inhomogeneous, a slightly coarser uniform grid is used for simplicity. The CYGNSS observation grid is shown as black dots in **Fig. 3.4d**. The horizontal ROI is set to 200 km for the CYGNSS observation and their vertical ROI is 15 model levels from surface up to ~ 700 hPa.

3.4 Scale- and variable-dependent improvements in prediction skill

In this subsection, the benchmark CNTL experiment is compared to NoDA to evaluate the impact from assimilating synthetic ATOVS and AMV observations. First, to assess the EnKF performance in CNTL, the two-dimensional spectra of error in prior ensemble mean are compared to the ensemble spread in **Fig. 3.5**. For most variable and scales, the ensemble spread well represents the prior error, and no filter divergence occurred. **Figure 3.5** also shows the spectra from the CNTL prior ensemble mean and truth signals as references. The two reference spectra match at the large scales, however, the ensemble mean spectral power are lower than the truth at smaller scales. This drop in spectral power is due to the smoothing of small-scale features that are dislocated among the members, which to a certain extent reflects the more limited predictability for small scales, where error grows and displacing these features at a higher rate than the data assimilation cycles can constrain. In chapter 2, the predictability limits are characterized by comparing the ensemble spread spectra to the ensemble mean reference spectra. Going from large to small scales, at a given time the ensemble spread reaches and exceeds the signal level from the reference (signal-to-noise ratio smaller than 1), indicating loss of predictability. Notice that hydrometeors and vertical motion (w) reach predictability limits at larger scales than temperature (T), specific humidity (Q) and winds. At even smaller scales, when signal-to-noise ratio is dropped significantly below 1, the error spectra match the truth signal (error saturation). Scales smaller than 200 km appear to have saturated hydrometeor and w errors and wind, T , Q errors approaching their predictability limits at 3 h in the prior ensemble. Therefore, the following diagnostics will

filter out these smaller scales for a clearer demonstration of observation impact. For the sake of simplicity, “NoDA” will refer to NoDA ensemble mean and “CNTL” to CNTL posterior ensemble mean hereafter.

Figure 3.6 compares the longitude-time plots of surface precipitation from NoDA and CNTL to the truth. The NoDA mean precipitation signal is much weaker than the truth, although it still captures the precipitation events around day 6, 11, and 14-17. On the other hand, the CNTL analysis precipitation better agrees with the truth than NoDA. The amplitude, location, timing, and propagation of the precipitation signals are much improved by assimilating the observations. Snapshots of low-level horizontal wind and relative humidity fields from NoDA and CNTL just prior to the precipitation event at day 6 are compared to the truth in **Fig. 3.7**. The truth low-level flow forms a convergence line and an anti-cyclonic shear around the precipitation and a cyclonic gyre is located northeast to the precipitation. Although NoDA has the perfect lateral boundary conditions by the current OSSE design, the convergence line and precipitation peaks are almost missing by the ensemble simulation without data assimilation. EnKF assimilation of satellite observations in CNTL recovers the flow pattern, although the location of precipitation is still slightly but noticeably different from the truth. Similarly, CNTL has a better representation of atmospheric moisture near the line of precipitation than NoDA as shown from the relative humidity fields. Note that CNTL has a weaker precipitation signal due to the averaging among members that have precipitation at different locations. Although small-scale signals ($l < 200$ km) are filtered out, this effect is still present.

Figure 3.8 shows time series of the domain-averaged root mean square errors (RMSEs) with respect to the truth from NoDA, CNTL and the forecasts from CNTL.

Zonal wind (u) and meridional wind (v) errors are combined and shown as root-mean difference kinetic energy (RM-DKE). Difference kinetic energy (Zhang et al. 2002) is defined as $(u'^2 + v'^2)/2$, where u' and v' are the errors in u and v , respectively. The forecasts are used to evaluate how long in time the prediction skill can be retained. The posterior ensemble mean is lack of small-scale signals, which is unrealistic and not necessarily in physical balance. Therefore, an ensemble forecast is required. Schwartz et al. (2014) suggested that subensembles of 20-30 members will have comparable forecast skills with a full 50-member ensemble. In this study, a small 10-member ensemble forecast is performed, and forecast errors are calculated from the mean of ensemble forecasts. The directly observed variables (u , v , T , and Q) all show evident error reduction in CNTL from NoDA. The hydrometeor mixing ratios and w are also improved slightly from the better flow pattern and thermodynamic structures. The horizontal wind, T , Q forecast errors grow linearly for the first day and more flow-dependent afterwards as errors approach the NoDA level. Hydrometeor and w error growth is flow-dependent almost right from the beginning. The apparent error peaks at day 6, 11, and 14-17 are associated with the precipitation episodes.

Figure 3.9 shows vertical error profiles from NoDA, CNTL, and the forecasts and **Fig. 3.10** shows their spectra. The forecast errors at certain lead times are averaged over 10 forecast runs during the test period (5 forecasts starting from 00 UTC on day 5, 7, ..., 13, and 5 forecasts starting 12 UTC). Errors peak at different vertical levels for each variable. Horizontal wind, w , T , and solid hydrometeor (ice, snow and graupel) errors have peaks at upper levels that associate with cloud tops. Deep convective clouds are likely responsible for the liquid hydrometeor (rain and cloud) errors at mid-level, while

shallow low-level clouds likely induce errors in horizontal wind, T , and Q around 900 hPa. Their spectral error distribution confirms that improvement from NoDA to CNTL is mostly at scales larger than 200 km. The forecast errors saturate faster for smaller scales. For 200-500 km scales, error saturates after 2 days; for scales larger than 1000 km, errors have not reached saturation after 4 days. The hydrometeors and w have faster small-scale error growth than other variables. These scale- and variable-dependent forecast error growth is consistent with the findings from Bei and Zhang (2014) and chapter 2.

3.5 Relative impact from different observing networks

The +Met7 experiment assimilates Met7 channel-3 Tb in addition to CNTL. **Figures 3.11a-d** shows a snapshot of its resulting channel-2 Tb as an independent verification in comparison to other experiments. In CNTL, the assimilation of ATOVS and AMV recovers the large-scale precipitating cloud cluster from NoDA already. However, its small-scale details of convective clouds are still erroneous comparing to the truth. The +Met7 result has a better agreement with the truth for the location and strength of cloud clusters (low Tb) than CNTL. The horizontal distribution of ice, snow and graupel, to which infrared Tb is most sensitive, are shown in **Figs. 3.11e-h**. Although not perfect comparing to the truth, the +Met7 hydrometeors are improved over CNTL.

The +CYGNSS experiment assimilates CYGNSS wind speed in addition to CNTL, and another experiment NoProfile assimilates both Met7-Tb and CYGNSS in addition to CNTL but excludes ATOVS profiles. **Figure 3.12** plots the error time series, vertical profiles, and spectra from these experiments for comparison. The +Met7 can

effectively reduce errors in ice, snow, and graupel mixing ratios (**Fig. 3.12d**), but not so much for rain and cloud (not shown). Horizontal wind, T , and Q are also slightly improved by assimilating Met7-Tb. Since Met7-Tb has a much higher horizontal resolution than ATOVS profiles, it can potentially improve the predictability for scales smaller than 200 km. The summed error spectra of solid hydrometeors (**Fig. 3.12l**) show that +Met7 indeed reduces error at small scales. However, in this study, the 3 h cycling period is too long to constrain errors at these scales before their saturation. Potentially, assimilating Met7-Tb at hourly intervals and with a higher resolution model will better utilize its positive impact. The single channel Tb assimilated is also not providing sufficient vertical profile information. Comparing NoProfile to CNTL, it is evident that ATOVS profiles better constrains horizontal wind, T , and Q . The ATOVS profiles combines the information from multiple satellite Tb images that are sensitive to different vertical heights, effectively having better vertical resolution than a single-channel Tb. The future development in all-sky infrared and microwave Tb assimilation will allow a fairer comparison between direct Tb assimilation and assimilating retrieved profiles. In addition to AMV wind, CYGNSS observations provide extra wind information for the lowest levels, reducing horizontal wind, T , and Q errors from surface up to 850 hPa. This improvement is mostly found for larger scales due to its moderate resolution. In NoProfile, the denial of ATOVS profiles causes a larger error in mid-level horizontal winds, however, their errors at upper and low levels remain small with the assimilation of AMV and CYGNSS (**Fig. 3.12e**).

To further test the sensitivity to resolution of the retrieval T and Q profiles from either ATOVS or GPSRO, several additional experiments are conducted assimilating

retrieved profiles in addition to NoProfile. +ATOVS assimilates the ATOVS profiles at their current resolution, while +ATOVScoarse assimilates them with temporal and horizontal resolution reduced by half. +GPSRO assimilates GPSRO profiles at COSMIC2 resolution, while +GPSROdense assimilates a denser version from the hypothetical populated CubeSat constellation. Note that the localization distance is increased for coarse networks according to the increase in observation intervals. **Figure 3.13** compares their resulting error time series, vertical profiles and spectra. When reduced-resolution ATOVS profiles are assimilated, their impact is also reduced for all variables and scales. The COSMIC2 GPSRO profiles bring an impact that is approaching but from time to time less than the reduced-resolution ATOVS profiles. The random location of these profiles is a disadvantage over the uniformly spaced ATOVS profiles in continuous sampling of convective systems of interest. This issue is likely resolved with an increase in observation density. As shown in +GPSROdense, the CubeSat constellation with 25 times more data will yield a retrieval product dense enough to surpass the performance of the current network of ATOVS profiles.

In +ATOVS, the ATOVS profiles are assumed to have ~40 independent pieces of information for T in the vertical, which likely overestimates the information content and may yield an analysis that is too optimistic. To at least partially address this concern, two additional experiments are conducted. +ATOVSthinV assimilates ATOVS profiles with vertical levels thinned to 10 levels for T and 4 levels for Q , and +ATOVSerror1.5 assimilates ATOVS profiles with vertical resolution unchanged but errors for both T and Q inflated to 150%. Note that for +ATOVSthinV, the vertical ROI is enlarged to 20 levels as the observations are thinned in the vertical. **Figure 3.14** shows that reduction in

vertical resolution/accuracy does not considerably degrade the analysis accuracy. The increase in analysis error in these experiments is much smaller comparing to +ATOVScoarse with reduced horizontal and temporal resolutions. This also indicates that the vertical information content estimated to contain ~ 10 independent pieces of information is likely able to capture the key vertical modes in the tropics.

Figure 3.15 compares the direct assimilation of GPSRO N profiles (+GPSRON and +GPSROdenseN) with assimilation of retrieved T and Q profiles (+GPSRO and +GPSROdense). The overall analysis error is similar for both assimilation strategies. The nonlinear observation operator results in a shorter decorrelation length scale, which requires the localization distance to be shorter to achieve similar performance. The nonlinearity also causes longer filter spin-up period for Q (**Fig. 3.15c**), and cause the analysis for T and winds to be less accurate than those from assimilating retrieved profiles (**Figs. 3.15a** and **3.15b**). Since our idealized experiment assumes no errors in the retrieval process and error correlation and bias are not considered, the performance of assimilating retrieval profiles is an optimistic expectation comparing to real-data cases.

3.6 Observation impact on CCEWs

The same space-time band-pass filter from section 2 is applied to posterior ensemble mean to evaluate the observation impact for each CCEW mode. The posterior mean fields lack the spatial and temporal coverage to directly filter for large-scale waves, thus the month-long large-domain truth simulation is used to fill in the gaps before applying the band-pass filter. **Figure 3.16** plots the longitude-time diagrams of the

filtered 850-hPa zonal wind from NoDA and CNTL and compares them to the truth, and **Table 3.4** quantifies these errors. Two error sources can lead to a large RMSE, the waves being out of phase with the truth and/or the amplitude of the waves being wrong. The pattern correlation quantifies the contribution from phase errors; if the waves are perfectly in phase with the truth (correlation equals 1), the RMSE is solely due to amplitude errors. Compared to the truth, NoDA does capture most of the large-scale low-frequency wave phases. For the MJO, its correlation is up to 0.98, indicating that using a perfect LBC in the current OSSEs plays a dominant role in modulating the model solution of MJO, which agrees with the sensitivity experiment results from Zhang et al. (2017). The ER, Kelvin, MRG and large-scale WIG waves also have reasonable forecasted phases in NoDA, but errors grow as wave phases propagate away from LBC and further downstream. The ER and MRG westerly signals at day 7 start from middle of the domain, which is not captured by NoDA. The large-scale WIG signals at east side of the domain are much weaker in NoDA than the truth, and for intermediate-scale WIG waves the signals are almost wiped out in the east half of the domain.

By assimilating observations in CNTL, the phase and amplitude of these wave modes are significantly improved. **Figure 3.17** shows the error reduction from NoDA to CNTL as measured by root mean difference total energy (RM-DTE; Melhauser and Zhang 2012). The MJO amplitude error is further reduced in CNTL for all variables. For ER, Kelvin, MRG and large-scale WIG wave, the phase correlation is improved from ~ 0.9 to almost perfect (**Table 3.4**). As for the more challenging intermediate-scale WIG waves, the phase correlation is also much improved for all variables (e.g. from 0.51 to 0.83 for zonal wind). The precipitation has larger phase errors than other variables due to

its more limited predictability. **Table 3.5** shows the relative impacts on WIG waves from assimilating different observations. Since these sensitivity experiments are only run for 9 days, they are not diagnosed for lower-frequency large-scale waves. Assimilating AMV, Met7-Tb, and CYGNSS (NoProfile) significantly improves both large-scale and intermediate-scale WIG waves. The addition of ATOVS profiles further reduces the errors in all variables. The additional merit from higher-resolution profiles appears to be more important for better analysis and prediction of intermediate-scale WIG waves.

3.7 Summary

In this study, a perfect-model OSSE is performed to evaluate the potential impact from assimilating satellite observations on the practical predictability of tropical multiscale weather systems. Following the previous chapter, a regional WRF model is configured to simulate the CCEWs during a 17-day period within the active phase of the October 2011 MJO event. The simulation captures the strong precipitation signals associated with the MJO, Kelvin and IG waves. The ER and MRG waves have relatively weak precipitation but well-defined zonal wind signals. Most wave components are at scales larger than 1000 km, except for IG waves that have a smaller-scale sub-diurnal component. The OSSE assumes that a perfect forecast model is used, and observation errors are also modeled perfectly during data assimilation.

Table 3.3 summarizes the benefits of assimilating each observing network every 3 h using EnKF. According to Zhang et al. (2017), the global circumnavigating signals play an important role in MJO initiation inside their domain. Therefore, the perfect LBC

alone is able to provide the correct phase of most of the large-scale waves (MJO, ER, Kelvin, MRG), and their errors are dominated by amplitude errors. The smaller scale high-frequency waves (WIG), on the other hand, are less well constrained by LBC, and large phase errors occur without data assimilation. The CNTL experiment tests assimilation of ATOVS T and Q profiles and AMV wind observations for the 17-day period, and showed an improvement in winds, T , Q , and hydrometeors for scales larger than 200 km. For large-scale ($l > 1000$ km) CCEWs, their predictability limit can be extended by about 4 days after assimilation. However, the small-scale ($l < 500$ km) waves have more limited predictability with the extra prediction skill lasting for less than 1 day. The EnKF assimilation further improved amplitude for the large-scale waves, and reduced phase errors of intermediate-scale WIG waves as well. Assimilating CYGNSS wind speed retrievals is able to reduce errors in large-scale low-level horizontal wind and T . It provides complementary wind information to the surface and low levels where not many AMVs are available.

The Meteosat-7 infrared T_b has higher horizontal resolution and thus further improves the model variables at smaller scales. The improvement is found mostly in the ice, snow, and graupel mixing ratios near cloud top. However, the 3-h cycling is probably not frequent enough to maintain the extra prediction skill at small scales. Assimilating T_b from only 1 infrared channel does not constrain model dynamic and thermodynamic variables as efficiently as the ATOVS profiles. The later has advantage of combining multiple satellite images that has information at different vertical levels. The uniform horizontal distribution of ATOVS profiles yields a more persistent observation impact than the sporadic sampling of sparser GPSRO profiles. At COSMIC2 resolution, GPSRO

profiles have impacts approaching half of those from ATOVS profiles. Although the number of profiles is likely still 10 times fewer than the reduced-resolution ATOVS, their higher vertical resolution and more accurate retrievals of both T and Q bring more impact per profile. The better horizontal resolution from the ATOVS retrieval profiles and better vertical resolution from the GPSRO profiles can potentially complement each other. With expected much increased future CubeSat resolution, the GPSRO profiles may have an impact comparable or exceeding that from the current ATOVS profiles. The high-resolution profiles can potentially improve smaller-scale weather systems, such as the sub-diurnal WIG waves coupled with transient convective processes.

Table 3.1. Space-time filtering parameters for each tropical weather system (CCEW type).

CCEW type	Spatial scale l (km)	Time scale τ (days)	Propagation direction	Symmetry about the equator
MJO	>2000	>20	eastward	both
ER	>1000	8-20	westward	symmetric
Kelvin	>1000	3-10	eastward	symmetric
MRG	>1000	4-10	westward	anti-symmetric
WIG (L)	>1000	1-3	westward	symmetric
WIG (M)	500-2000	1/4-1	westward	both

Table 3.2. Observed variable, their associated uncertainty (observation error), spatial resolution and approximate data count within a 3-h window of each tested observing network.

Observing network	Observed variable	Observation error	Horizontal spacing	Vertical spacing / coverage	Observation count in 3-h window
ATOVS retrievals	T	~ 2 K	90 km	1 km; surface to 0.1 hPa	~ 35000
	Q	$\sim 30\%$		2 km; surface to 200 hPa	~ 11000
GPSRO retrievals (COSMIC2)	T	~ 1.7 K	irregular (~ 600 km)	~ 100 m in low troposphere 1 km in stratosphere	~ 1660
	Q	$\sim 10\%$			~ 1350
AMV wind	u, v	4.5 m s^{-1}	irregular (~ 100 km)	irregular spacing mostly covers ~ 200 hPa	~ 1700
Met7 IR Tb (channel 3)	T_b	3 K	9 km	column-averaged with weighting function	~ 80000
CYGNSS retrievals	wind speed	2 m s^{-1}	45 km	at ocean surface	~ 3200

Table 3.3. Description of assimilated observing networks and their benefits (improvement in prediction skills in variables, scales and weather systems) found in each experiment.

Experiment	Assimilated observing networks	Benefits in analysis/prediction skill
NoDA	None	<ul style="list-style-type: none"> • Rely on LBC to constrain solution in domain. • Large-scale CCEWs (MJO, Kelvin, MRG, and ER) have some predictability in wave phase but with large amplitude errors.
CNTL	ATOVS profiles + AMV wind	<ul style="list-style-type: none"> • Significantly improve winds, T, Q, and hydrometeors at large to intermediate scales. • Extend the practical predictability limit by ~ 4 days for large scales. • Further reduce the amplitude error for large-scale CCEW and improve intermediate-scale WIG waves.
+Met7	CNTL + Met7-Tb	<ul style="list-style-type: none"> • Further improve ice, snow and graupel mixing ratio near cloud top for smaller scales.
+CYGNSS	CNTL + CYGNSS wind speed	<ul style="list-style-type: none"> • Further improve large-scale low-level u, v, and T.
NoProfile	AMV wind + Met7-Tb + CYGNSS wind speed	<ul style="list-style-type: none"> • AMV+Met7+CYGNSS can still improve u, v, T, and Q, although less effective than ATOVS profiles.
+ATOVS	NoProfile (AMV+Met7+ CYGNSS) + ATOVS profiles	ATOVS profiles are essential in improving u , v , T , and Q , thanks to their fine spatial resolution and coverage.
+ATOVScoarse	NoProfile + ATOVS profiles reduced to 6 hourly and 180-km grid spacing	<ul style="list-style-type: none"> • Although worse than the full-resolution profiles, there are still persistent improvements at large to intermediate scales. • Reduced resolution degrades smaller-scale WIG.
+ATOVSthinV	NoProfile + ATOVS profiles thinned in the vertical	<ul style="list-style-type: none"> • Reduced information content in the vertical does not degrade analysis accuracy as much as reduction in horizontal and temporal resolution.
+ATOVSerror1.5	NoProfile + ATOVS profiles with errors inflated to 150%	
+GPSRO	NoProfile + GPSRO profiles at COSMIC2 resolution	<ul style="list-style-type: none"> • Approaching but less than improvement from low-resolution ATOVS due to sporadic sampling.
+GPSROdense	NoProfile + GPSRO profiles at CubeSat resolution (25 times more data than COSMIC2)	<ul style="list-style-type: none"> • More improvement than ATOVS thanks to finer vertical resolution and more accurate measurements.
+GPSRON +GPSROdenseN	Same as +GPSRO and +GPSROdense but assimilating refractivity (N) profiles	<ul style="list-style-type: none"> • Nonlinear observation operator requires shorter localization distance. • Similar sensitivity to observation density.

Table 3.4. Root-mean-square errors with respect to the truth simulation for u (m s^{-1}), T (K), Q (g kg^{-1}), and precipitation (mm day^{-1}) filtered for CCEW types described in Table 3.1. The errors are averaged over latitude ($0\text{-}5^\circ\text{N}$) and time (17-day period). Pattern correlations with the truth are shown in parentheses. Results are compared between NoDA ensemble mean and CNTL posterior ensemble mean.

CCEW type	Experiment	u	T	Q	Precipitation
MJO	NoDA	0.39 (0.98)	0.063 (0.98)	0.072 (0.99)	1.97 (0.99)
	CNTL	0.17 (1.00)	0.025 (1.00)	0.020 (1.00)	0.53 (1.00)
ER	NoDA	0.50 (0.90)	0.064 (0.90)	0.073 (0.94)	1.18 (0.88)
	CNTL	0.19 (0.99)	0.033 (0.97)	0.027 (0.99)	0.61 (0.98)
Kelvin	NoDA	0.40 (0.92)	0.089 (0.92)	0.078 (0.87)	2.37 (0.83)
	CNTL	0.13 (0.99)	0.025 (0.99)	0.027 (0.98)	0.69 (0.99)
MRG	NoDA	0.33 (0.91)	0.029 (0.85)	0.088 (0.90)	1.00 (0.77)
	CNTL	0.15 (0.98)	0.012 (0.98)	0.021 (0.99)	0.41 (0.97)
WIG (L)	NoDA	0.38 (0.82)	0.067 (0.90)	0.078 (0.80)	2.83 (0.79)
	CNTL	0.14 (0.98)	0.025 (0.99)	0.032 (0.97)	1.20 (0.97)
WIG (M)	NoDA	0.38 (0.51)	0.097 (0.69)	0.060 (0.54)	4.88 (0.35)
	CNTL	0.25 (0.83)	0.061 (0.89)	0.047 (0.77)	2.83 (0.84)

Table 3.5. Similar to Table 3.4, but comparing results from NoDA, NoProfile, +ATOVScoarse and +ATOVS posterior ensemble mean from the first 10 days and for WIG waves.

CCEW type	Experiment	u	T	Q	Precipitation
WIG (L)	NoDA	0.37 (0.74)	0.060 (0.90)	0.072 (0.80)	2.92 (0.62)
	NoProfile	0.19 (0.94)	0.043 (0.95)	0.062 (0.86)	1.61 (0.92)
	+ATOVScoarse	0.14 (0.97)	0.031 (0.97)	0.045 (0.93)	1.06 (0.96)
	+ATOVS	0.13 (0.97)	0.025 (0.98)	0.032 (0.96)	0.90 (0.97)
WIG (M)	NoDA	0.32 (0.59)	0.086 (0.73)	0.056 (0.54)	4.30 (0.32)
	NoProfile	0.26 (0.74)	0.076 (0.80)	0.050 (0.66)	3.04 (0.74)
	+ATOVScoarse	0.24 (0.80)	0.069 (0.84)	0.048 (0.71)	2.56 (0.82)
	+ATOVS	0.23 (0.82)	0.059 (0.88)	0.046 (0.74)	2.10 (0.88)

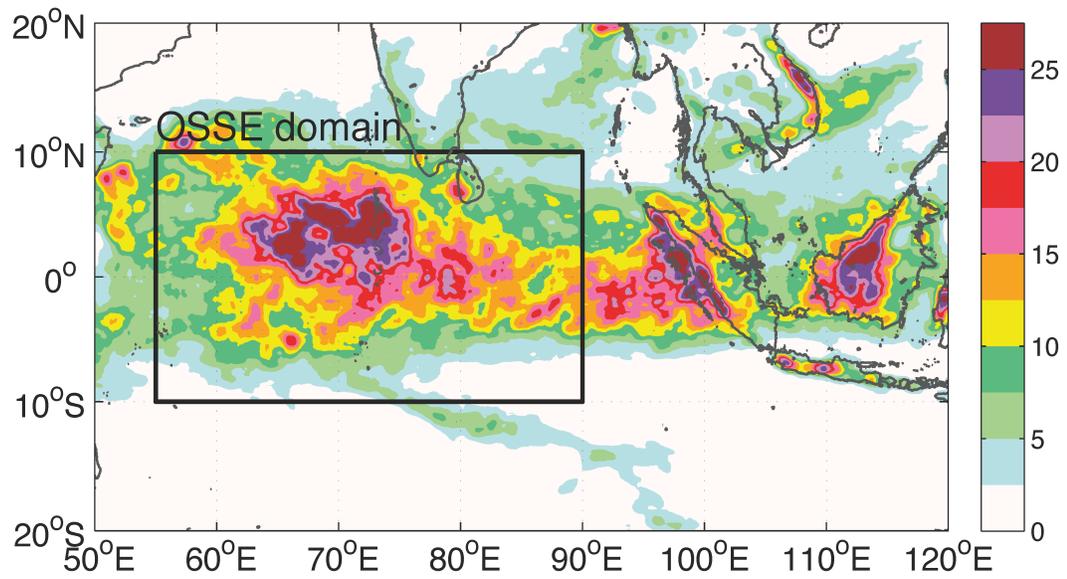


Figure 3.1. Horizontal map of 15-day averaged daily precipitation (mm day^{-1}) shown in the computational domain with ERA-Interim as initial and boundary condition. The OSSE domain is shown in black

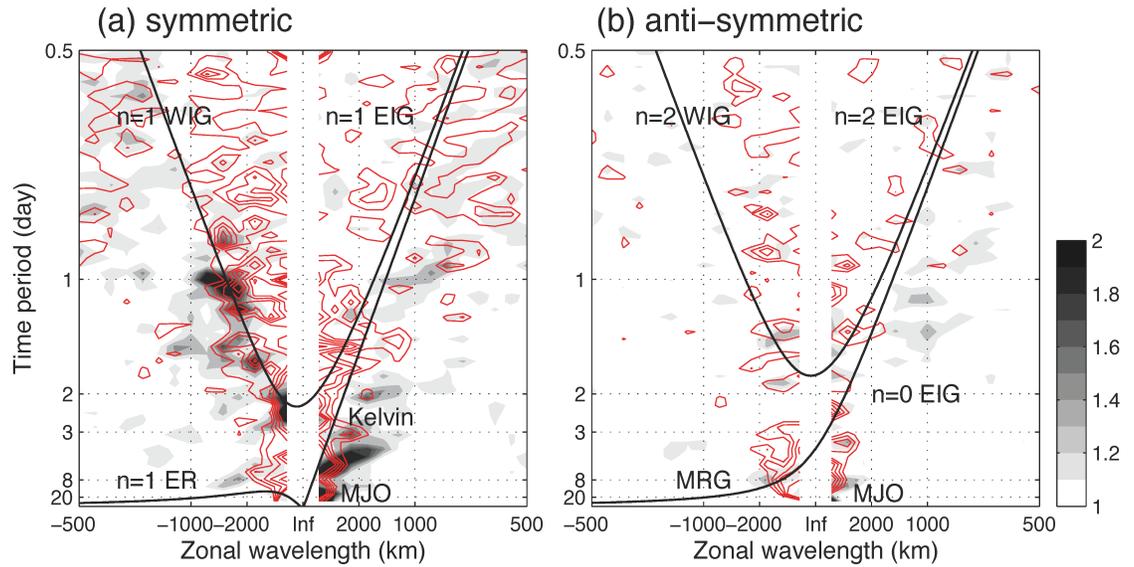


Figure 3.2. Wheeler-Kiladis space-time spectra for the truth simulation: (a) symmetric and (b) asymmetric component about the equator. Signal strength from 1.1 to 2 is shown in gray shadings for precipitation and red contours for 850-hPa zonal wind. Dispersion relations for dry waves are shown with equivalent depth of 25 m. Horizontal (vertical) axis is zonal wavenumber (time frequency) but labeled with zonal wavelength (time period).

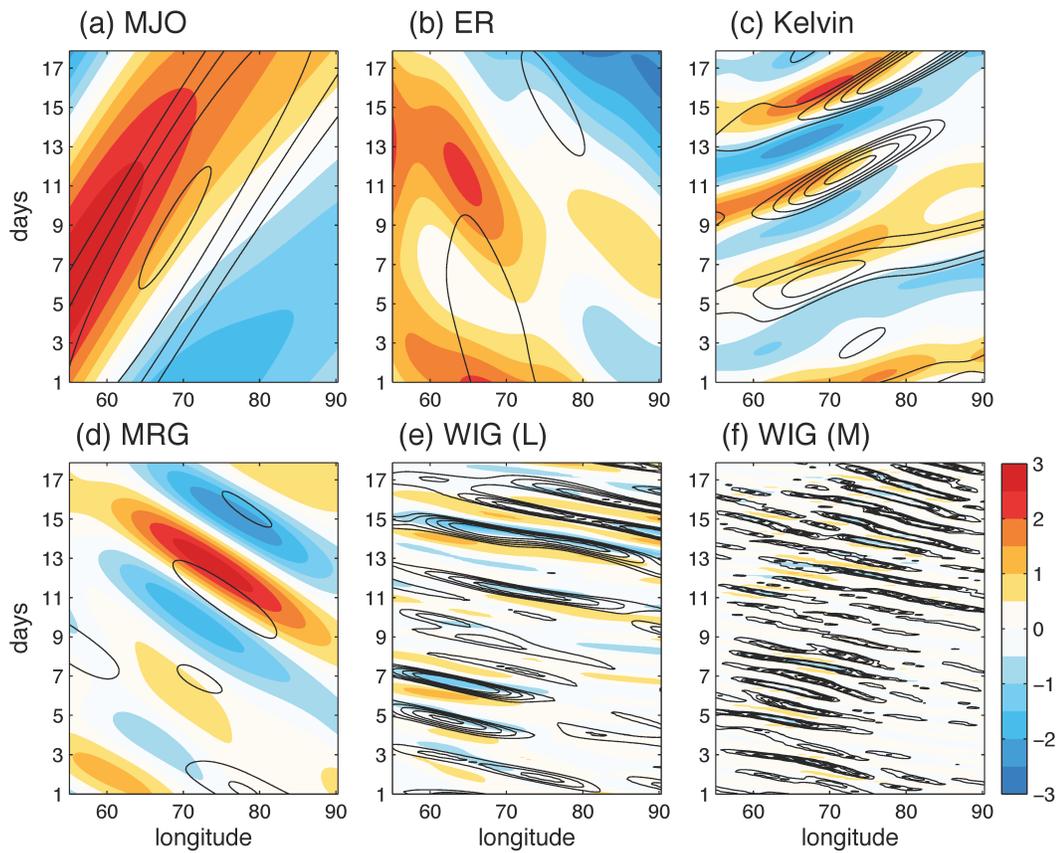


Figure 3.3. Longitude-time Hovmöller diagrams of precipitation (black contours from 2 to 10 mm day^{-1} every 2 mm day^{-1}) and 850-hPa zonal wind (color shadings; m s^{-1}) averaged over $0\text{-}5^\circ\text{N}$ from the truth simulation and filtered for each CCEW: (a) MJO, (b) ER, (c) Kelvin, (d) MRG, (e) large-scale WIG, and (f) intermediate-scale WIG waves.

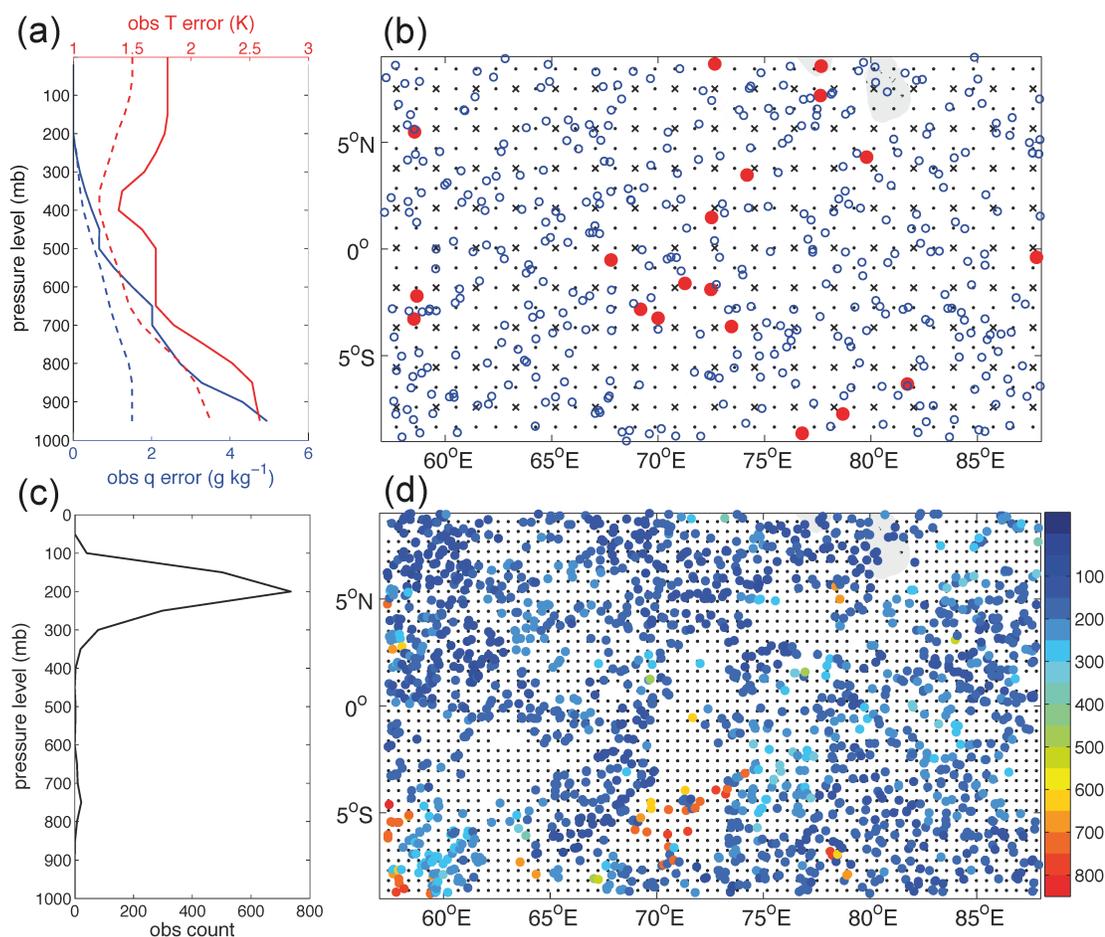


Figure 3.4. (a) Vertical observation error profiles from ATOVS (solid) and GPSRO (dashed) for T (red) and Q (blue); (b) Horizontal map of observation location during a 3-h window for ATOVS (solid black dots), ATOVS with reduced resolution (black cross), GPSRO at COSMIC2 resolution (red dots) and GPSRO at CubeSat resolution (blue circles); (c) AMV observation count (per 50-hPa vertical layer) during a 3-h window; (d) Horizontal map of AMV observation location color-coded with observation pressure height (hPa), black dots show location of CYGNSS wind observations.

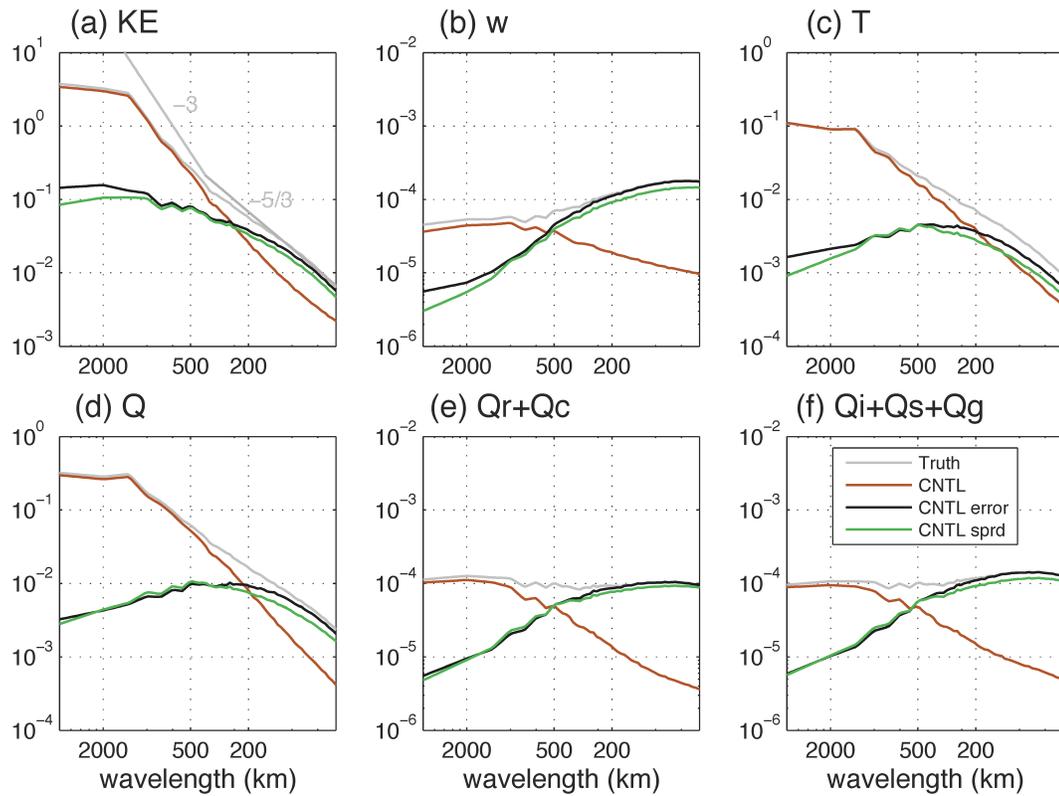


Figure 3.5. Temporally and vertically averaged spectra for (a) kinetic energy ($\text{m}^2 \text{s}^{-2}$), (b) vertical motion ($\text{m}^2 \text{s}^{-2}$), (c) temperature (K^2), (d) specific humidity ($\text{g}^2 \text{kg}^{-2}$), (e) rain+cloud water mixing ratio, and (f) ice+snow+graupel mixing ratios ($\text{g}^2 \text{kg}^{-2}$). Gray lines show spectra of the truth signal; red lines show spectra of the CNTL prior mean; black lines show spectra of CNTL error (difference between CNTL prior mean and the truth); and green lines show spectra of CNTL ensemble variance.

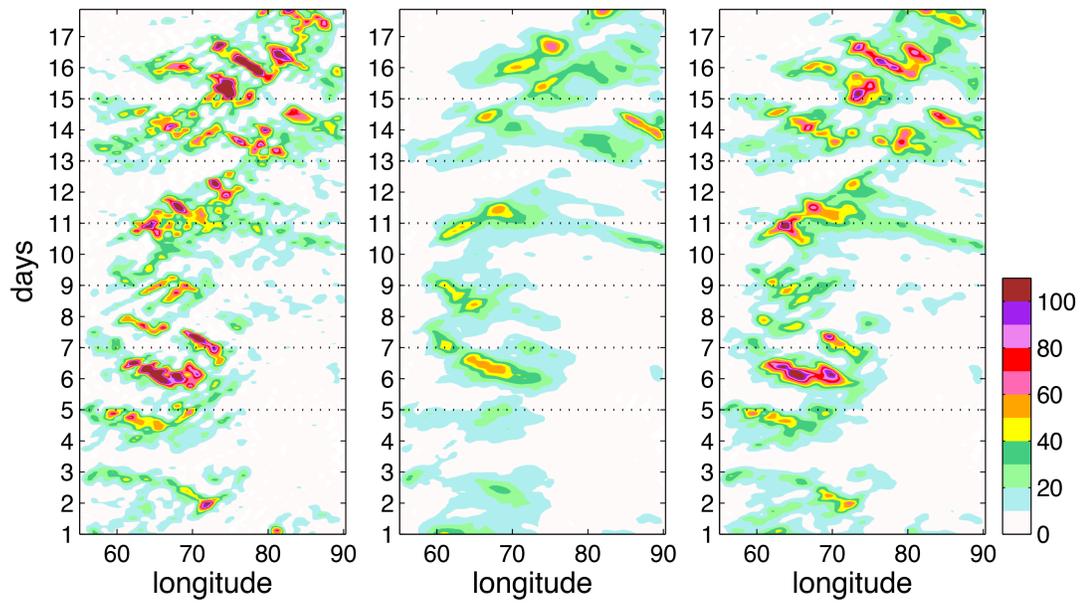


Figure 3.6. Longitude-time Hovmöller diagrams of precipitation (mm day^{-1}) averaged over 0 to 5°N latitudes from (a) truth simulation, (b) NoDA ensemble mean, and (c) CNTL posterior ensemble mean. The precipitation fields are filtered for $l > 200$ km.

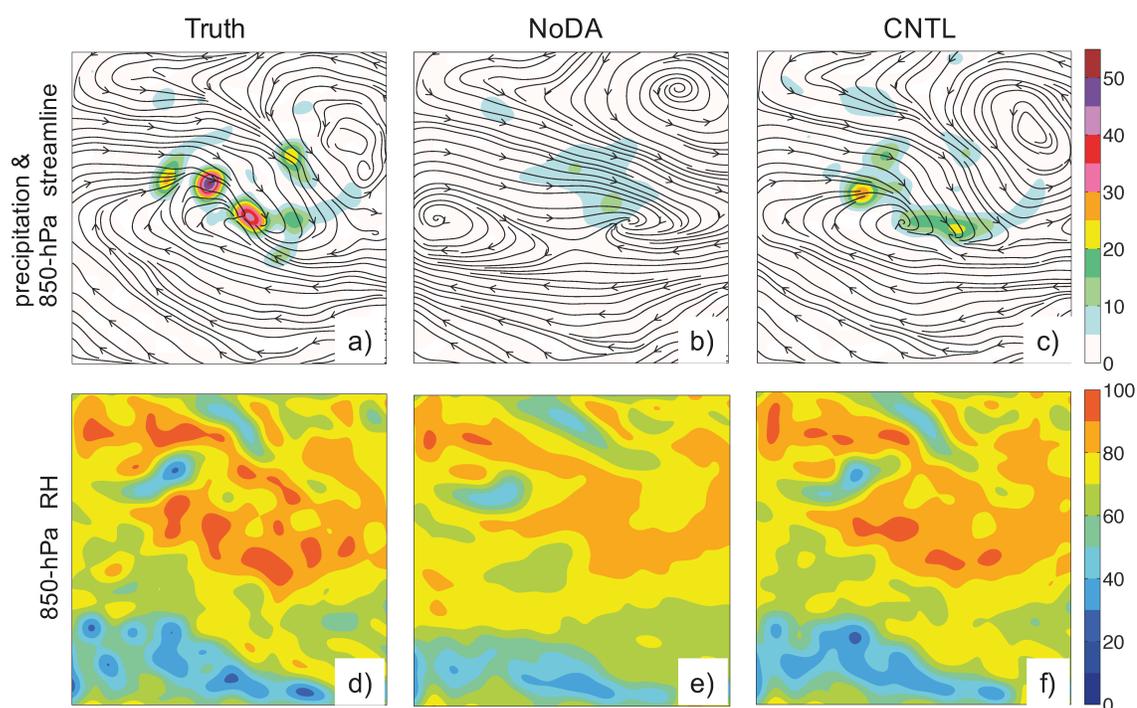


Figure 3.7. (a)-(c) Snapshots of 850-hPa streamlines and 3-h accumulated precipitation (shadings; mm), and (d)-(f) 850-hPa relative humidity (%) after 5.5 days of cycling (valid at 12UTC, 21 Oct). Results are compared for (a, d) truth simulation, (b, e) NoDA ensemble mean, and (c, f) CNTL posterior ensemble mean. All fields are filtered for $l > 200$ km. The left half of the OSSE domain is shown (10°S to 10°N , 55° to 75°E).

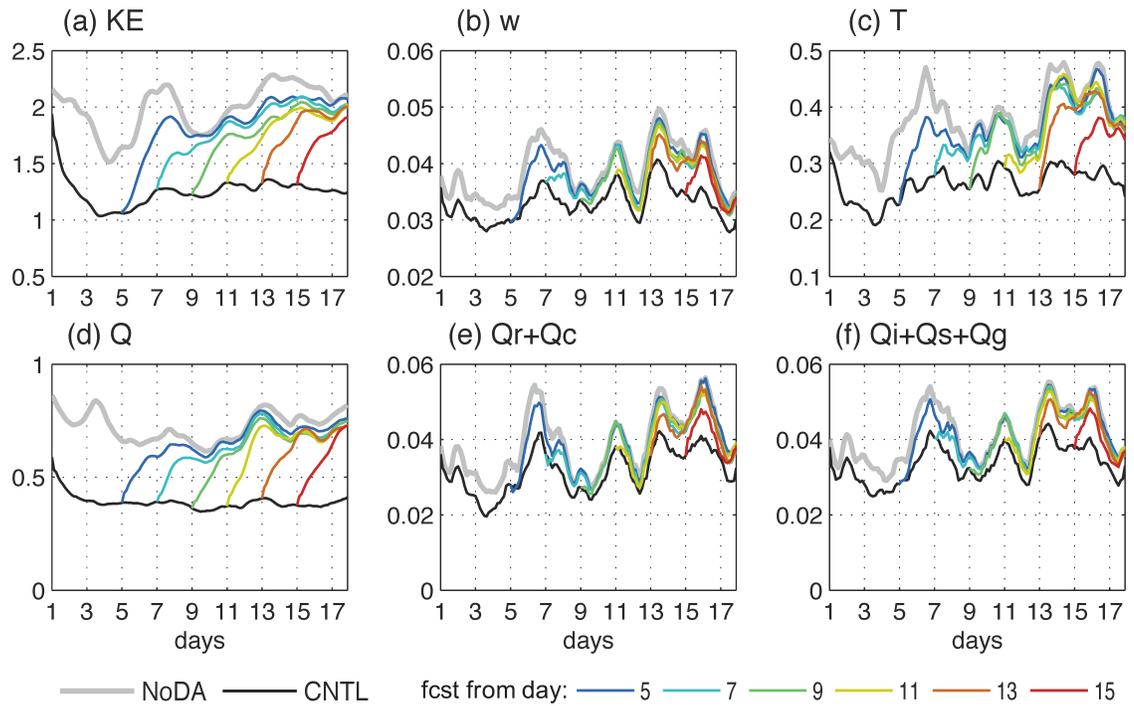


Figure 3.8. Times series of domain-averaged (a) root-mean difference kinetic energy (m s^{-1}), and RMSEs of (b) vertical motion (m s^{-1}), (c) temperature (K), (d) specific humidity (g kg^{-1}), (e) rain+cloud water mixing ratio (g kg^{-1}), and (f) ice+snow+graupel mixing ratio (g kg^{-1}). The errors are also averaged over the vertical column and filtered for $l > 200$ km. Black lines show error from CNTL posterior ensemble mean and gray lines from NoDA ensemble mean. Colored lines show error evolution from ensemble forecasts initiated at day 5, 7, 9, 11, 13, and 15. All ensemble means are calculated from 10 members.

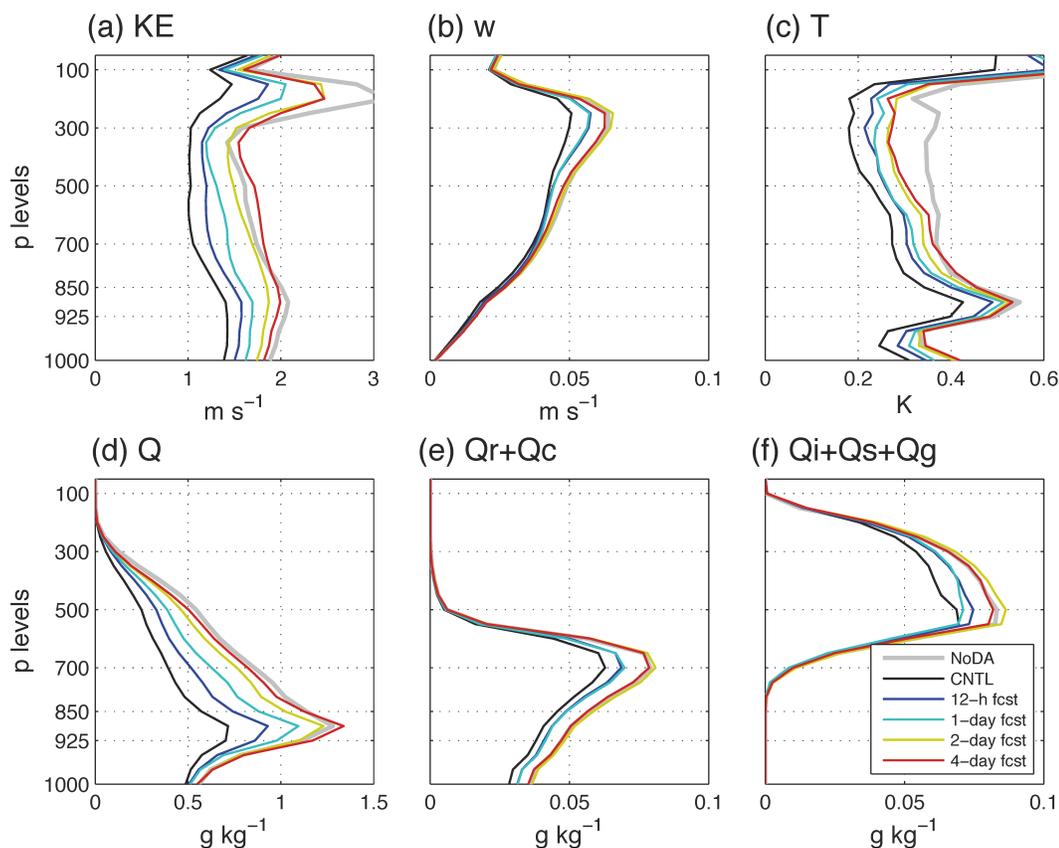


Figure 3.9. Vertical profiles of (a) RM-DKE (m s^{-1}), and RMSEs of (b) vertical motion (m s^{-1}), (c) temperature (K), (d) specific humidity (g kg^{-1}), (e) rain+cloud water mixing ratio (g kg^{-1}), and (f) ice+snow+graupel mixing ratio (g kg^{-1}). The errors are averaged over day 5 to 18 and filtered for $l > 200$ km. Black lines show error from CNTL posterior ensemble mean and gray lines from NoDA ensemble mean. Colored lines show error at forecast lead time of 12 hours, 1, 2, and 4 days. All ensemble means are calculated from 10 members.

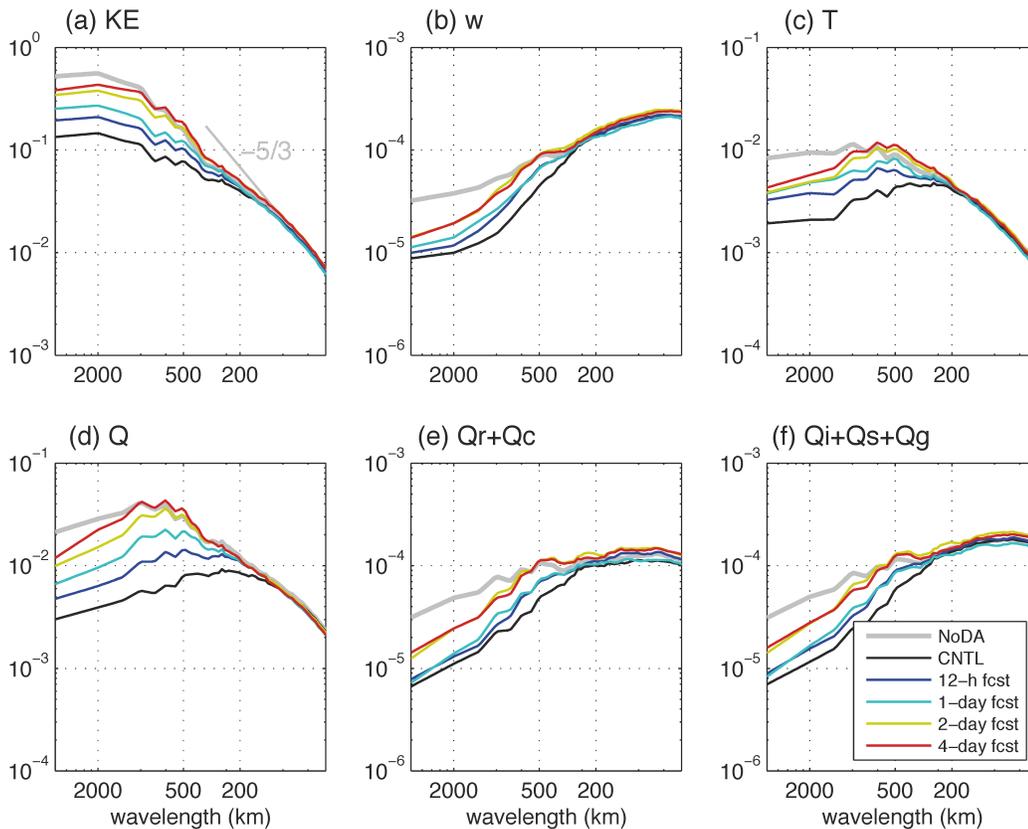


Figure 3.10. Similar to Fig. 3.9 but showing the temporally and vertically averaged error energy spectra. The spectral errors are in variance units, e.g. $\text{m}^2 \text{s}^{-2}$ for DKE and K^2 for temperature. The spectrum are shown with respect to global horizontal wavenumber, $k = \sqrt{k_x^2 + k_y^2}$, but labeled with its corresponding wavelength ($l = k^{-1}$).

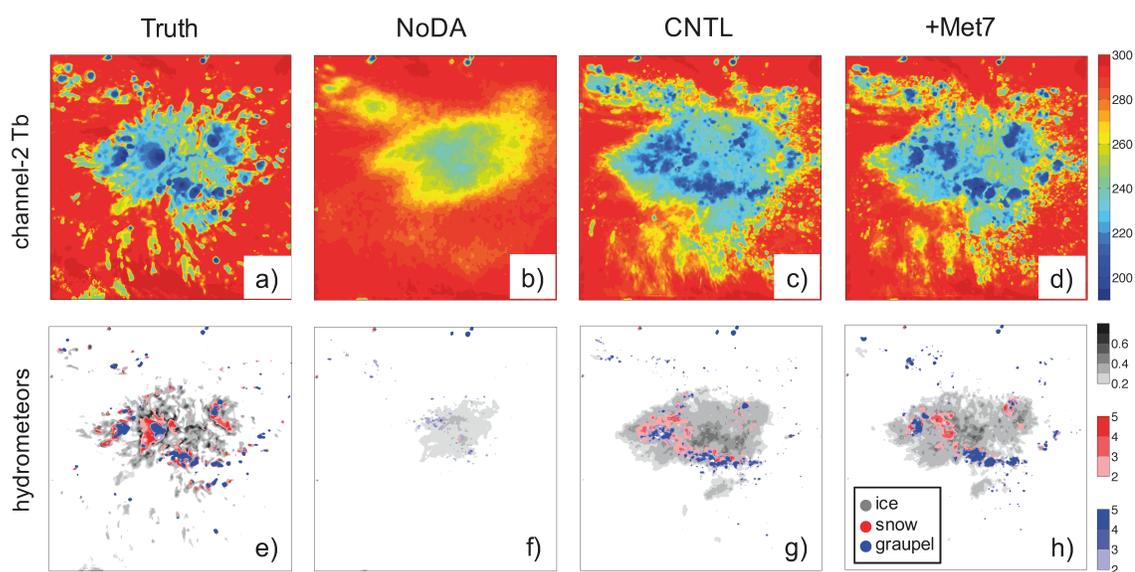


Figure 3.11. (a)-(d) Snapshots of channel-2 Tb (K), and (e)-(h) column integrated ice (gray), snow (red) and graupel (blue) mixing ratios (g kg^{-1}) after 5.5 days of cycling (valid at 12UTC, 21 Oct) from (a, e) truth simulation, (b, f) NoDA ensemble mean, (c, g) CNTL posterior ensemble mean, and (d, h) +Met7 posterior ensemble mean.

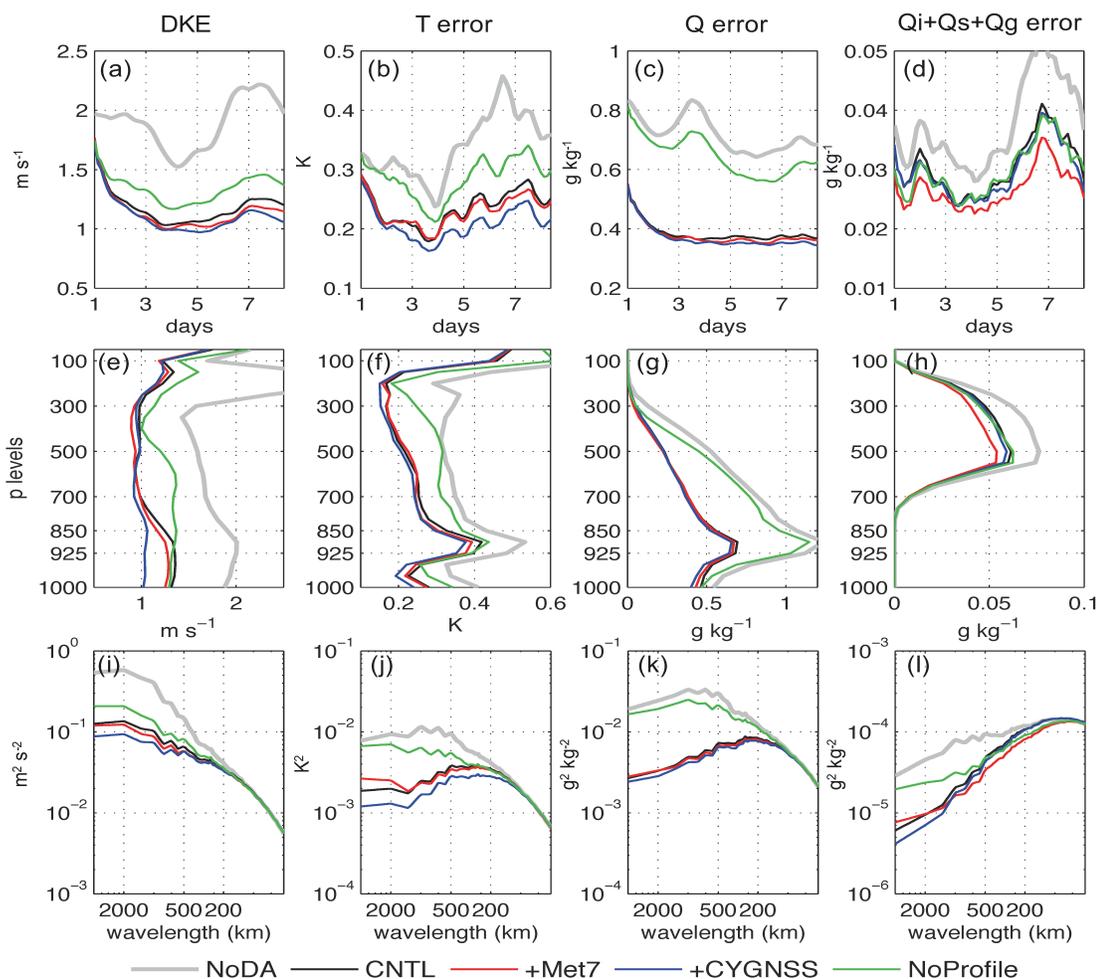


Figure 3.12. (Top row) Time series, (center row) vertical profiles, and (bottom row) spectra of domain-averaged analysis errors in (columns from left to right) kinetic energy, temperature, specific humidity, and summed error from ice, snow, and graupel mixing ratios. Errors are shown as RMSE (RM-DKE for kinetic energy), except that the spectra show error variance. The time series are vertically averaged and filtered for $l > 200$ km; vertical profiles are averaged over time and filtered for $l > 200$ km; spectra are averaged temporally and vertically. Results from NoDA, CNTL, +Met7, +CYGNSS and NoProfile are compared.

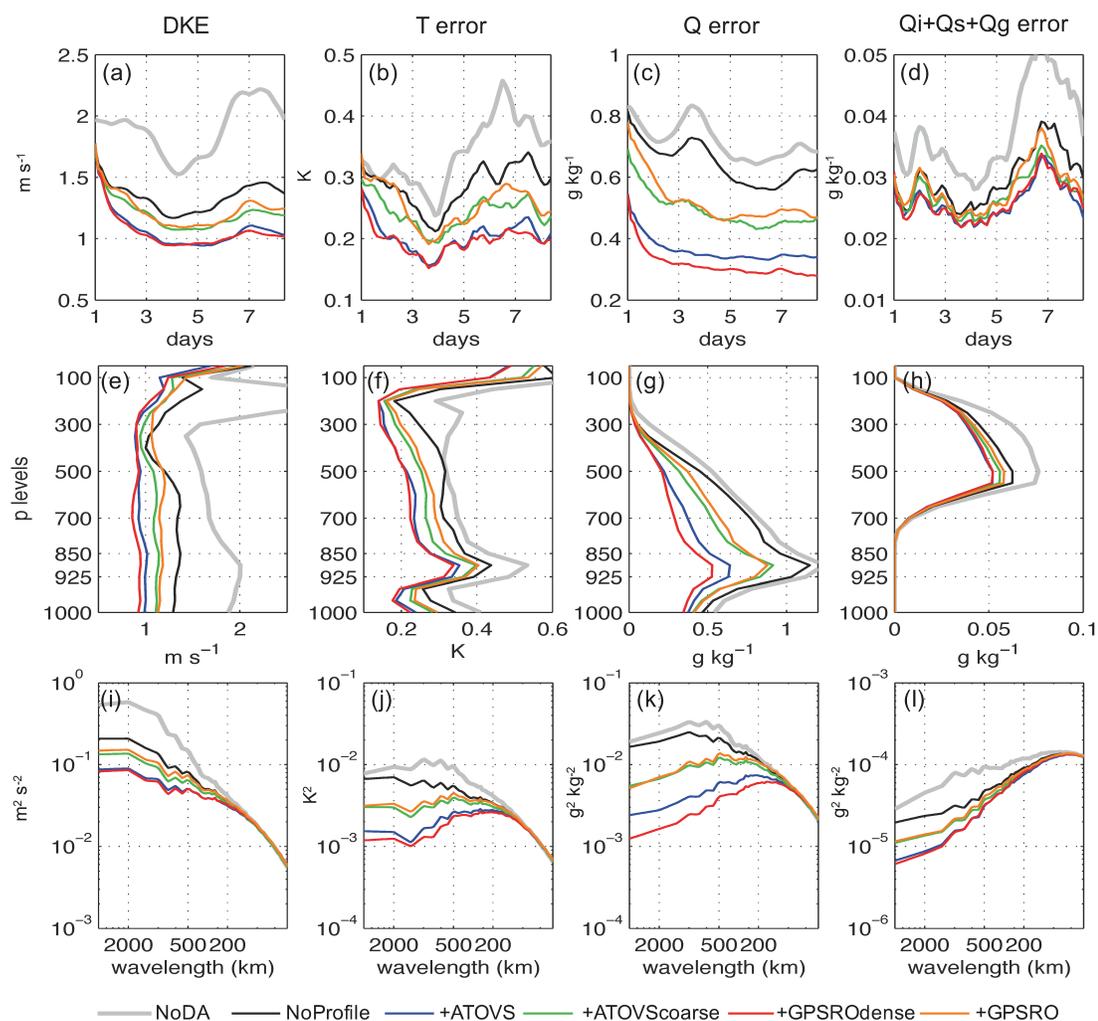


Figure 3.13. Same as Fig. 3.12, but showing results from NoDA, NoProfiles, +ATOVS, +ATOVScoarse, +GPSRO, and +GPSROdense.

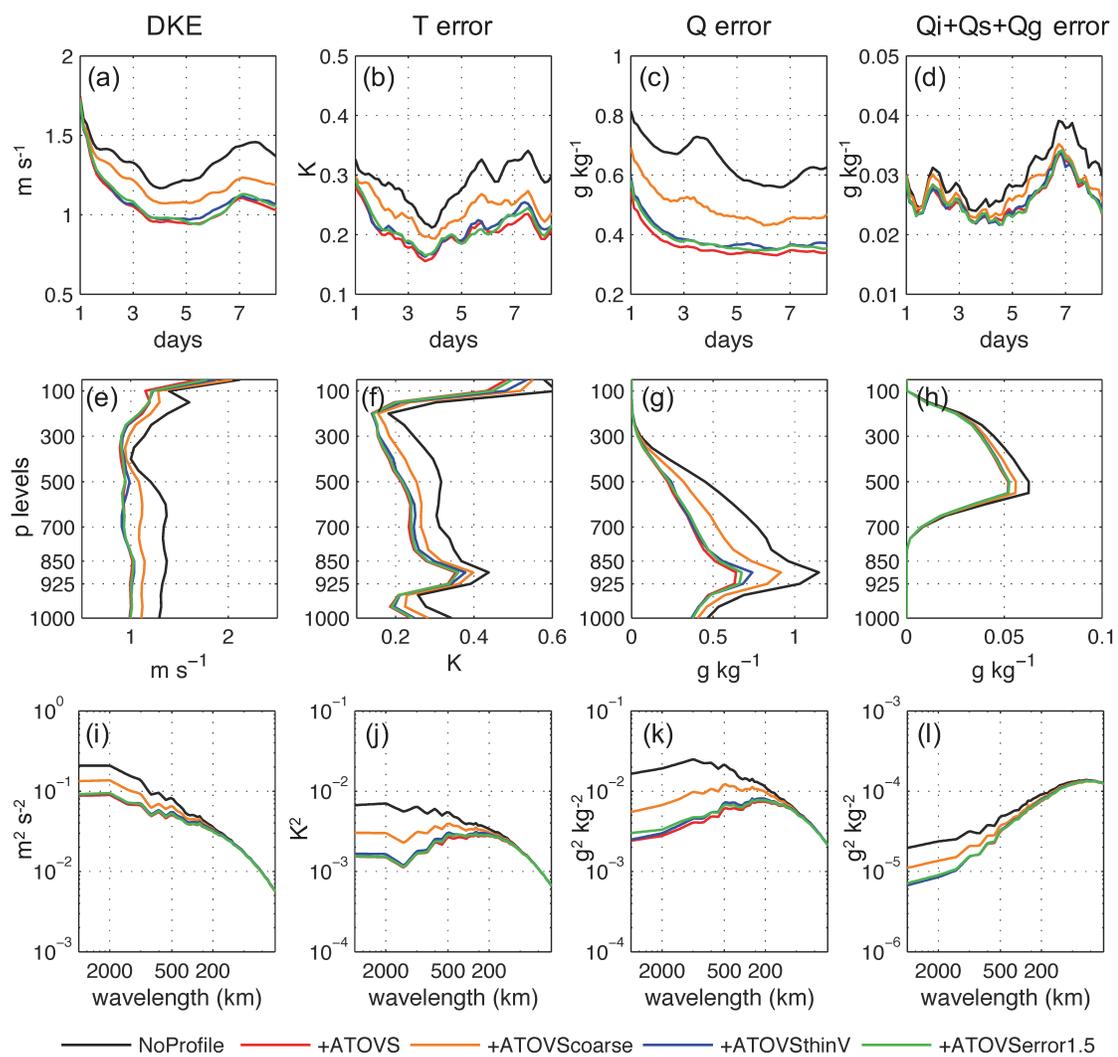


Figure 3.14. Same as Fig. 3.12, but showing results from NoProfiles, +ATOVS, +ATOVScoarse, +ATOVStinV, and +ATOVSerror1.5.

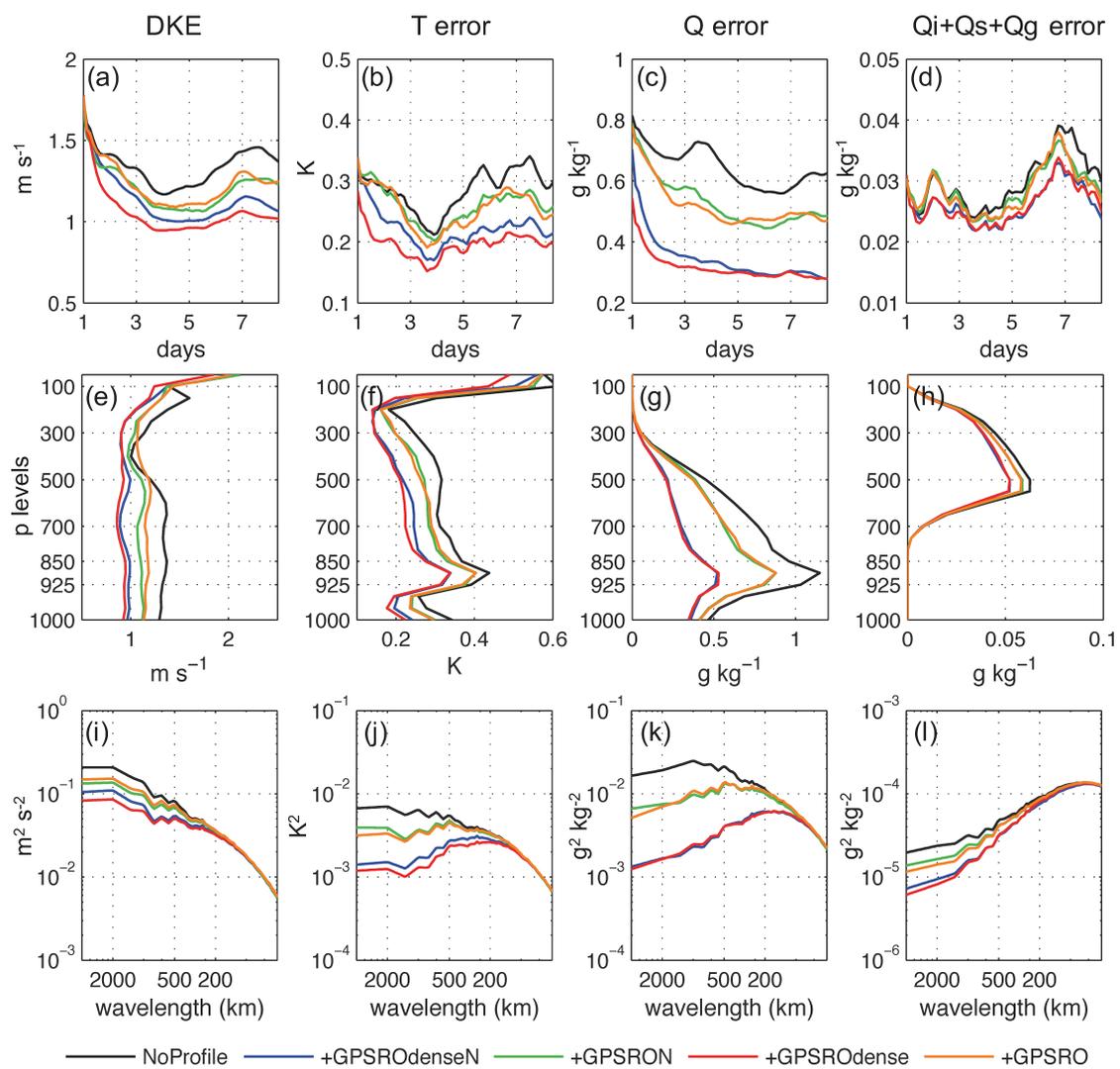


Figure 3.15. Same as Fig. 3.12, but showing results from NoProfiles, +GPSROdenseN, +GPSRON, +GPSROdense, and +GPSRO.

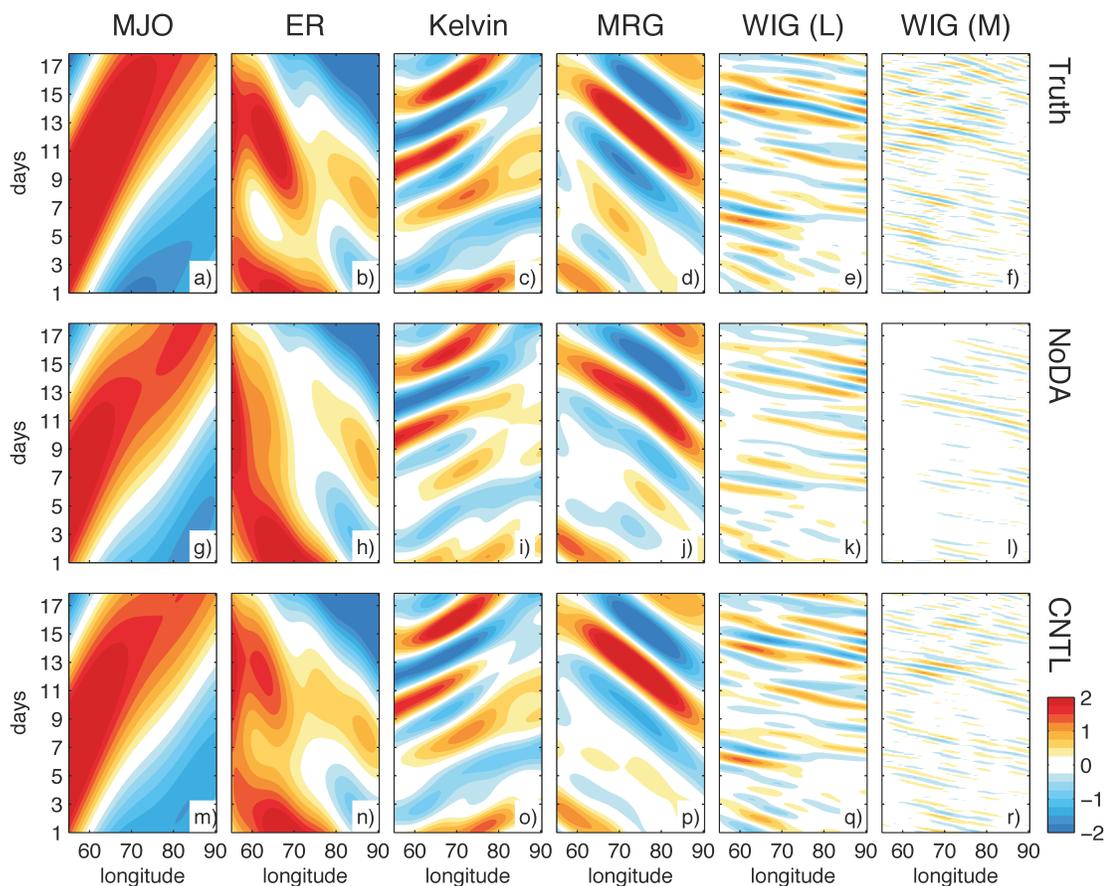


Figure 3.16. Longitude-time Hovmöller diagrams of 850-hPa zonal wind (color shadings; m s^{-1}) averaged over $0\text{-}5^\circ\text{N}$ latitudes from (top row) truth simulation, (center row) NoDA ensemble mean, and (bottom row) CNTL posterior ensemble mean. The fields are filtered for each CCEW shown in columns from left to right.

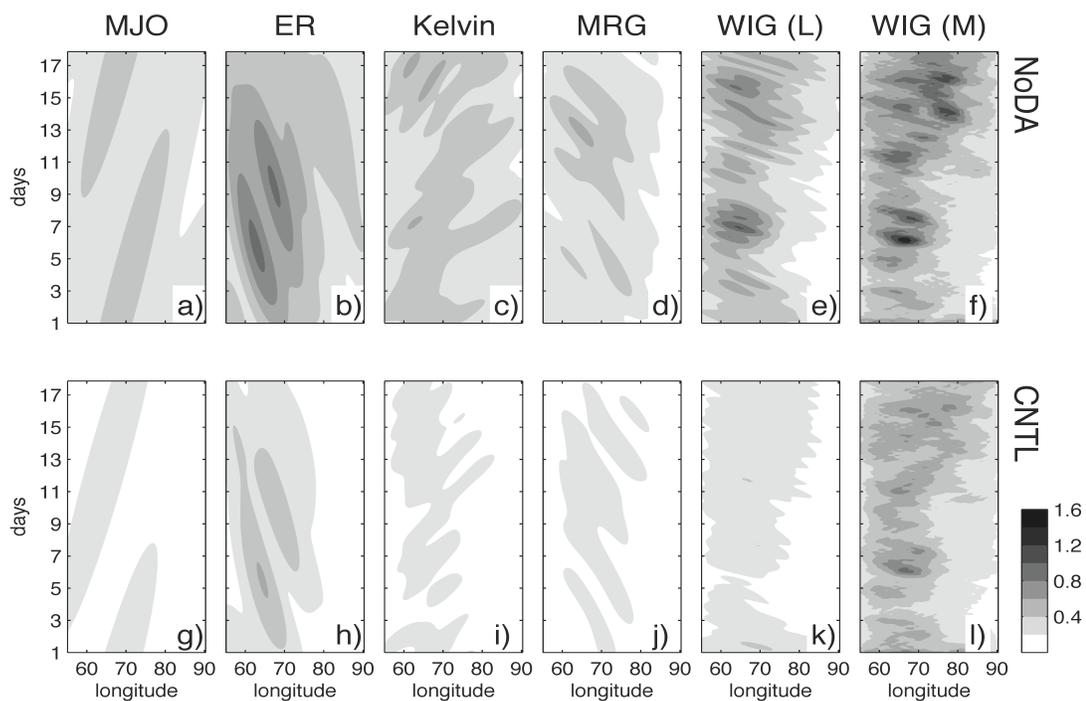


Figure 3.17. Longitude-time Hovmöller diagrams of RM-DTE with respect to the truth (m s^{-1}) averaged over $0\text{--}5^\circ\text{N}$ latitudes and vertical levels for (a)-(f) NoDA ensemble mean and (g)-(l) CNTL posterior ensemble mean. A space-time filter is applied to u , v and T for each CCEW (shown from left to right) before calculating the RM-DTE.

Chapter 4

An adaptive covariance relaxation method for ensemble data assimilation

4.1 Introduction

The ensemble Kalman filter (EnKF) combines information from a prior state estimate and its associated uncertainty (from an ensemble of model realizations) with observations and observation uncertainties to get an improved posterior state estimate and an updated uncertainty (Evensen, 1994). The accuracy of prior error covariance is one of the necessary criteria for optimal filter performance. In the presence of sampling and model error, the ensemble may underestimate the true uncertainty in the prior causing the filter to place too much weight on the prior mean state. Over time, the filter will begin to ignore the observed information, resulting in filter divergence. There are a number of methods for handling unrepresented error sources, among which covariance inflation methods are widely used. Empirical covariance inflation methods include multiplicative inflation (Anderson and Anderson, 1999), which increases the ensemble perturbations by a specified factor, and additive inflation (Mitchell and Houtekamer, 2000), which adds a random perturbation drawn from a specified error distribution to each member. Whitaker and Hamill (2012) hypothesize that multiplicative inflation methods can account for observation network related sampling error, while additive inflation is more suitable for treating model error.

There are also methods specifically designed to handle sampling error due to limited ensemble size. For example, covariance localization (Hamill et al., 2001) treat sampling errors by tapering the Kalman gain far from the observation location, which removes some of the spurious correlations. Houtekamer and Mitchell (1998) proposed a double-ensemble EnKF in which ensemble-estimated covariance is used to update the other ensemble, therefore avoids inbreeding. More recently, a new formulation of the EnKF is proposed to account for the sampling bias due to the use of a limited-size ensemble (Bocquet, 2011; Bocquet and Sakov, 2012). For model error, treatments related to the forecast model can help maintain ensemble spread and ensure filter performance; such treatments include the use of a multi-physics or multi-model ensemble (e.g., Meng and Zhang, 2007), stochastic kinetic energy backscatter (Shutts, 2005; Berner et al., 2009), and stochastically perturbed physics tendencies (Buizza et al., 1999).

Among the aforementioned methods, covariance inflation is favored due to its simplicity. However, given the large dimension of the dynamical systems of interest in atmospheric science, the process of tuning the inflation factor to suit a particular application is often costly. A variety of adaptive covariance inflation (ACI) methods are proposed to estimate inflation factors from innovation (observation-minus-forecast) statistics. The innovation statistics was first used to estimate parameters in forecast and observation error covariance matrices using a maximum-likelihood approach (Dee, 1995; Dee and da Silva, 1999; Dee et al., 1999). For example, Wang and Bishop (2003) formulated an online inflation estimation algorithm within the ensemble transform Kalman filter (ETKF) framework. Li et al. (2009) further extended the algorithm to simultaneously estimate covariance inflation and the observation error, using a set of

observation-minus-forecast diagnostics (Desroziers et al., 2005). They applied temporal smoothing to reduce the sampling bias due to a small innovation sample at each time step. Anderson (2007) provides an alternative algorithm by treating each innovation as a random variable and updates the inflation factor using a hierarchical Bayesian approach. These methods assume observations are uncorrelated. Zheng (2009) and Liang et al. (2011) relaxed this assumption and estimated an inflation factor from the innovation vector at each time step using a maximum-likelihood approach. They found that their algorithm is only efficient when estimating a scalar inflation factor constant in space.

Another important issue for any effective adaptive inflation method is how the method handles a spatially irregular observation network. For areas where no observations are available, applying an inflation factor constant in space may cause the variance of these unobserved variables to keep growing, sometimes even exceeding climatological variance. To alleviate this problem, Anderson (2009) extends the Bayesian approach to estimate a spatially and temporally varying inflation parameter, and assume that the spatial correlation factor of the inflation parameters is the same as the state variables. Miyoshi (2011) also introduces an algorithm that estimates inflation parameters that are spatially varying. He advances Li et al.'s method by including variance of the estimated inflation parameter from Central Limit Theorem, making it a Gaussian approximation of Anderson's method.

In practice, the inflation factor is usually not physically constrained, which will often give rise to imbalance issues for complicated dynamic models. Zhang et al. (2004) developed a method that relaxes the posterior ensemble perturbations to the prior perturbations that effectively inflates the posterior ensemble yet preserves physical

balances of the ensemble perturbations. Following this concept, Whitaker and Hamill (2012) proposed a relaxation-to-prior-spread (RTPS) method that is equivalent to applying a spatially varying multiplicative inflation. As a property of ensemble filters, the posterior ensemble spread should be smaller than prior spread after assimilating observations. Spatially, the reduction in ensemble spread should only be found where observations are available. Therefore, the reduction of spread can serve naturally as a spatial mask for inflation in case of irregular observation network. The scalar relaxation parameter can effectively be an inflation factor given that it is allowed to be larger than 1.

In this chapter, we introduce an adaptive covariance relaxation (ACR) method that estimates the relaxation parameter of RTPS online according to innovation statistics. We will compare this new method to the non-adaptive RTPS as well as Anderson's ACI method. The advantage of RTPS is its simplicity and its ability to provide a spatially varying inflation as in Anderson's method. In the section 4.2, we will provide the mathematical formulation of the proposed ACR method, along with Anderson's method described with the same notations for comparison. Experimental designs are presented in section 4.3 and the results are given in section 4.4. Section 4.5 gives the concluding remarks.

4.2 Methodology

4.2.1. Ensemble Kalman filter

Generally speaking, the inflation methods described here should be applicable to most if not all variations of ensemble Kalman filters. In this study, we will use an ensemble square root filter (EnSRF) introduced by Whitaker and Hamill (2002) for testing. The EnSRF algorithm is described as follows. Let \mathbf{x} be an n -by-1 column vector that holds all state variables. We introduce an ensemble of N members, or state vectors. The ensemble mean and the i th ensemble perturbation are denoted by $\bar{\mathbf{x}}$ and $\mathbf{x}'_i = \mathbf{x}_i - \bar{\mathbf{x}}$, respectively. The background error covariance matrix is calculated from this ensemble using $\mathbf{P} = \frac{1}{N-1} \sum_{i=1}^N \mathbf{x}'_i \mathbf{x}'_i{}^T$. Let \mathbf{y}^o be a p -by-1 column vector that contains all observations at the current time. The observations are drawn from a normal distribution $\mathcal{N}(\mathbf{H}\mathbf{x}^t, \mathbf{R})$, where \mathbf{x}^t is the truth state vector, \mathbf{H} is a p -by- n linear operator that maps the state vector to observation space and \mathbf{R} is the observation error covariance matrix that is assumed to be diagonal. We denote the j th row of \mathbf{H} by \mathbf{H}_j . The EnSRF assimilates observations serially to avoid large matrix inversion. The following equations are applied for $j = 1, 2, \dots, p$ to update $\bar{\mathbf{x}}$ and \mathbf{x}'_i from the prior $[\bar{\mathbf{x}}^b$ and $(\mathbf{x}'_i)^b]'$ to the posterior $[\bar{\mathbf{x}}^a$ and $(\mathbf{x}'_i)^a]'$.

$$\bar{\mathbf{x}}^{\text{new}} = \bar{\mathbf{x}} + \boldsymbol{\rho}_j \circ \mathbf{K}_j (y_j^o - \mathbf{H}_j \bar{\mathbf{x}}), \quad (4.1)$$

$$(\mathbf{x}'_i^{\text{new}})' = \mathbf{x}'_i + \epsilon_j \boldsymbol{\rho}_j \circ \mathbf{K}_j (0 - \mathbf{H}_j \mathbf{x}'_i), \quad \text{for } i = 1, 2, \dots, N, \quad (4.2)$$

where $\mathbf{K}_j = \text{cov}(\mathbf{H}_j \mathbf{x}, \mathbf{x}) / [(\sigma_{y,j}^o)^2 + (\sigma_{y,j}^b)^2]$ is the Kalman gain, ρ_j is a localization

function, the circle denote an element-wise production and $\epsilon_j = \left(1 + \sqrt{\frac{(\sigma_{y,j}^o)^2}{(\sigma_{y,j}^o)^2 + (\sigma_{y,j}^b)^2}} \right)^{-1}$

is a square root modification term to account for the use of unperturbed observations (see Eq. 13 of Whitaker and Hamill 2002, denoted as α in their paper). Here $(\sigma_{y,j}^b)^2$ and $(\sigma_{y,j}^o)^2$ represent the background and observation error variances in observation space; they are the j th diagonal terms of \mathbf{HPH}^T and \mathbf{R} , respectively. The model integrates the posterior estimates forward in time, providing the prior estimates for the next assimilation cycle.

4.2.2. Calculation of inflation factor according to innovation statistics

The first component of an adaptive inflation algorithm is the calculation of inflation factor according to the spread deficiency indicated by innovation statistics. The innovation vector associated with the p observations is $\mathbf{d}^{o-b} = \mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b$. Dee (1995) derived the following expression for the prior innovation statistics:

$$\mathbb{E}[\mathbf{d}^{o-b}(\mathbf{d}^{o-b})^T] = \mathbf{HP}^b \mathbf{H}^T + \mathbf{R}. \quad (4.3)$$

The expected value for $\mathbf{d}^{o-b}(\mathbf{d}^{o-b})^T$ needs multiple realizations to estimate, which typically comes from observations taken at different times. For adaptive inflation, we need an online estimate of the innovation statistics for each time, so a reduced statistic using only the p observations at one time is used. Given that p is large, the following relationship will hold:

$$\text{tr}[\mathbf{d}^{o-b}(\mathbf{d}^{o-b})^T] = (\mathbf{d}^{o-b})^T \mathbf{d}^{o-b} = \text{tr}(\mathbf{HP}^b \mathbf{H}^T) + \text{tr}(\mathbf{R}). \quad (4.4)$$

The deficiency in prior error variance can be expressed in terms of an inflation factor,

$$\lambda_b = \sqrt{[(\mathbf{d}^{o-b})^T \mathbf{d}^{o-b} - \text{tr}(\mathbf{R})] / \text{tr}(\mathbf{H}\mathbf{P}^b \mathbf{H}^T)}. \quad (4.5)$$

Note that, if $(\mathbf{d}^{o-b})^T \mathbf{d}^{o-b} < \text{tr}(\mathbf{R})$, the square root cannot be evaluated. In this case, λ_b is set to 1.

Similarly, covariance inflation can be applied to the posterior ensemble according to the posterior innovation statistics (Desroziers et al., 2005):

$$\mathbb{E}[\mathbf{d}^{a-b}(\mathbf{d}^{o-a})^T] = \mathbf{H}\mathbf{P}^a \mathbf{H}^T, \quad (4.6)$$

where $\mathbf{d}^{o-a} = \mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^a$ and $\mathbf{d}^{a-b} = \mathbf{H}\bar{\mathbf{x}}^a - \mathbf{H}\bar{\mathbf{x}}^b$. And the inflation factor can be calculated as

$$\lambda = \sqrt{(\mathbf{d}^{a-b})^T \mathbf{d}^{o-a} / \text{tr}(\mathbf{H}\mathbf{P}^a \mathbf{H}^T)}. \quad (4.7)$$

Note that the inflation factor in both (4.5) and (4.7) is estimated in observation space, thus additional steps are needed for applying this inflation in state space.

4.2.3. Anderson's adaptive covariance inflation (ACI) method

Anderson (2009) proposes a Bayesian approach for estimating a temporally and spatially varying inflation parameter from the innovation statistics. More specifically, the inflation parameters λ_k are treated as random variables with normal distributions, $\Pr(\lambda_k) = \mathcal{N}(\bar{\lambda}_k, \sigma_{\lambda,k}^2)$. This method uses Bayes' theorem to update the distribution of λ_k by processing each innovation d_j^{o-b} serially. For $j = 1, 2, \dots, p$, the following procedure is performed for each state vector element k .

$$\Pr(\lambda_k | d_j^{o-b}) \propto \Pr(d_j^{o-b} | \lambda_k) \Pr(\lambda_k), \quad (4.8)$$

where d_j^{o-b} is the j th innovation $y_j^o - \mathbf{H}_j \bar{\mathbf{x}}^b$. Each d_j is assumed to be drawn from a zero-mean normal distribution with variance $\theta^2 = (\lambda_k^o \sigma_{y,j}^b)^2 + (\sigma_{y,j}^o)^2$, where λ_k^o is the expected inflation in the observation space given the prior inflation value at k . The method assumes that the inflation parameters and state variables have the same spatial correlation structure,

$$\text{corr}(\mathbf{H}_j \mathbf{x}, x_k) = \text{corr}(\lambda^o, \lambda_k). \quad (4.9)$$

Given the prior value $\bar{\lambda}_k$, the expected inflation λ_k^o can be calculated as

$$\lambda_k^o = 1 + \rho_k \text{corr}(\lambda^o, \lambda_k) (\bar{\lambda}_k - 1), \quad (4.10)$$

where ρ_k is the k th component of the localization function used in (4.1). The observation likelihood in (4.8) can be expressed as

$$\Pr(d_j^{o-b} | \lambda_k) = (\sqrt{2\pi}\theta)^{-1} \exp[-(d_j^{o-b})^2 / 2\theta^2]. \quad (4.11)$$

The updated $\bar{\lambda}_k$ is found when $\Pr(\lambda_k | d_j)$ reaches its maximum, thus $\bar{\lambda}_k$ can be solved by taking the derivative of (4.8) with respect to λ_k and setting it equal to zero. See Appendix A in Anderson (2009) for a detailed numerical method for finding the maximum of $\Pr(\lambda_k | d_j^{o-b})$. Also note that in Anderson's derivation he uses $\sqrt{\lambda}$ as the λ used in this paper.

Anderson's algorithm updates the inflation parameter at k according to the j th innovation and the spatial correlation between the observation and the k th state variable. If the correlation is large, (e.g., for observations near k), λ_k^o is close to $\bar{\lambda}_k$. When the correlation is small, λ_k^o approaches one. The amount of adjustment on $\bar{\lambda}_k$ brought by d_j^{o-b} depends on the selection of $\sigma_{\lambda,k}^2$. A larger $\sigma_{\lambda,k}^2$ indicates that the prior distribution of

$\bar{\lambda}_k$ has larger uncertainty, and should be adjusted more toward the value suggested by the innovation. Due to the nature of Bayesian inference, $\sigma_{\lambda,k}^2$ should decrease after each update, limiting the innovation impact on $\bar{\lambda}_k$ over time. In this study, we implement the method using the posterior $\bar{\lambda}_k$ as the prior value for the next assimilation cycle, and $\sigma_{\lambda,k}^2$ is fixed in time for all k .

4.2.4. Adaptive covariance relaxation (ACR)

Covariance relaxation can be considered as another approach for covariance inflation. Zhang et al. (2004) first proposed to relax the posterior ensemble perturbations to the prior:

$$(\mathbf{x}_i^a)'_{\text{inf}} = (1 - \alpha) (\mathbf{x}_i^a)' + \alpha (\mathbf{x}_i^b)', \quad (4.12)$$

so that the posterior ensemble spread is artificially increased and the posterior retains a certain degree of physical balance from the prior ensemble perturbations. Whitaker and Hamill (2012) suggested to relax the ensemble spread instead of the perturbations, and formulated the RTPS method. For the k th state variable, its ensemble spread is rescaled to a mix of prior and posterior spread values controlled by α :

$$(x_{k,i}^a)'_{\text{inf}} = (x_{k,i}^a)' \left(\alpha \frac{\sigma_k^b - \sigma_k^a}{\sigma_k^a} + 1 \right), \quad (4.13)$$

where $\sigma_k = \sqrt{\frac{1}{N-1} \sum_{i=1}^N x_{k,i}'^2}$ is the ensemble spread of the state variable x_k , and the superscript b and a denote prior and posterior values, respectively. Note that $\alpha = 0$ implies keeping the original posterior spread, while $\alpha = 1$ implies increasing the posterior spread back to the prior spread for each state variable. The rescaling term

$\alpha(\sigma_k^b - \sigma_k^a) / \sigma_k^a + 1$ works effectively as a multiplicative inflation factor for the ensemble perturbations of that state variable. Recall that, after assimilating observations, the posterior ensemble spread σ^a becomes smaller than the prior spread σ^b . This reduction in spread depends on the Kalman gain, which in turn depends on the localization distance and the observation location. When localization is applied, the reduction in spread should be larger in densely observed region and tapers to nearly zero in unobserved region, assuming all else in the Kalman gain (i.e. observation errors, etc.) are equal. Therefore, in RTPS, the relative reduction in ensemble spread, $(\sigma^b - \sigma^a) / \sigma^a$, naturally serves as a spatial mask and a constant α controls the magnitude of inflation.

The ACR method we propose takes advantage of the relative spread reduction term from RTPS, and reduces the problem to estimating a scalar parameter α online. To use innovation statistics, we need to calculate the overall relative reduction in ensemble spread in observation space. The overall relative spread reduction can be expressed as $(\bar{\sigma}_y^b - \bar{\sigma}_y^a) / \bar{\sigma}_y^a$, where $\bar{\sigma}_y = \sqrt{\text{tr}(\mathbf{H}\mathbf{P}\mathbf{H}^T) / p}$ and the overbar indicates an average over the p observations. We match the expected inflation in observation space with the inflation factor suggested by posterior innovation statistics (4.7), and solve for α :

$$\alpha \frac{\bar{\sigma}_y^b - \bar{\sigma}_y^a}{\bar{\sigma}_y^a} + 1 = \lambda. \quad (4.14)$$

Note that the calculated α value does not necessarily fall in the range of (0, 1). If the observation sample size p is small for each cycle, the inflation factor could become very noisy, in this case we apply temporal smoothing to λ in (4.14) to resolve this issue:

$$\lambda_{t,\text{smooth}} = \lambda_{t-1} + (\lambda_t - \lambda_{t-1}) / \tau, \quad (4.15)$$

where λ_t is the λ calculated for the current cycle and τ determines how fast α responds to changes suggested by the innovation statistics. The use of a larger τ is equivalent to using the innovations from a longer period of time to calculate the observed inflation factor. The estimated α are applied in (4.13) in state space to relax the ensemble spread.

4.3 Numerical experiment design

To test the methods described in the previous section, we run trials of cycling data assimilation using the Lorenz 40-variable model (Lorenz, 1996). The forecast model equations are

$$\frac{dx_k}{dt} = a(x_{k+1} - x_{k-2})x_{k-1} - dx_k + F, \quad \text{for } k = 1, 2, \dots, n, \quad (4.16)$$

where a , d and F are model parameters for the nonlinear advection, damping and forcing terms and $n = 40$. The model is defined on a 1D domain with a periodic boundary. The true model sets $a = 1$, $d = 1$ and $F = 8$. We set the model time step to $\Delta t = 0.05$, which corresponds to 6 hours. At each time step, the EnSRF update equations described in section 4.2.1 are applied. The radius of influence is set to 10 grid points for trials when covariance localization is applied.

To create synthetic observations, we add random noise to state variables generated from a simulation designated as the truth run. We set the observation noise σ_y^o to 1.0, commonly used in previous studies (e.g., Anderson 2009; Miyoshi 2011). Two different types of observation networks are tested: fully-observed and half-observed. The fully-observed network has observations located on all state variables (i.e., \mathbf{H} is the identity matrix), while the half-observed network has observations only for the first 20

“land” variables (the rest 20 “ocean” variables are unobserved) following the setup in Lorenz and Emanuel (1998).

For the RTPS, ACI, and ACR inflation methods, we run 10 trials, each trial consisting of 5000 time steps (corresponding to ~ 3.4 years). To evaluate the filter performance of each method, we calculate the analysis root mean square error (RMSE) and the consistency ratio (CR) for the final 1000 time steps of all 10 trials, mitigating issues that may arise during the spin up period of data assimilation. The RMSE is calculated by averaging the squared analysis error $(\bar{x}^a - x^t)^2$ over the n state variables for the final 1000 time steps of all 10 trials, then taking the square root. For the fully-observed cases, all $n = 40$ state variables are used to calculate RMSE, while for the half-observed cases the “land” (“ocean”) variables located at $k = 1, \dots, 20$ ($k = 21, \dots, 40$) are used to calculate RMSE_land (RMSE_ocean). The CR is calculated by averaging the ratio $\sqrt{\text{tr}(\mathbf{HP}^b\mathbf{H}^T + \mathbf{R})/(\mathbf{d}^{\sigma-b})^T\mathbf{d}^{\sigma-b}}$ over these time steps. If $\text{CR} < 1$, the background error variance is underestimated and the ensemble spread is too small, and vice versa. We specify the true \mathbf{R} in our filter so that an erroneous \mathbf{P}^b is the only reason why the statistical relation in (4.4) does not hold and CR deviates from one. We define “filter divergence” as a significant drop in the CR below 1 (ensemble spread is too small) and a growth in the RMSE to a value above the observation noise.

We introduce sampling error by reducing ensemble size (N) from 80 to 5. To introduce model error, we change the default parameters a , d and F in the forecast model. The forcing parameter F is gradually reduced from the true value 8 to 5, incrementally producing a less chaotic model. We change parameter a from 1 to 0.8 (d from 1 to 1.2) to

introduce error in the advection (damping) process in the forecast model. The inflation methods are also tested with combinations of errors in a , d , and/or F to examine if they can handle model errors from a mixture of processes.

For each error regime, we test the performance of Anderson’s method and the proposed ACR method. For simplicity, we will denote Anderson’s method as “ACI σ_λ^2 ”, where σ_λ^2 is the selected value for the variance of inflation parameter. We have tested ACI with $\sigma_\lambda^2 = 0.1, 1, \text{ and } 10$, and show that results are sensitive to this parameter. For ACR, we have tested with the temporal smoothing parameter $\tau = 1, 10, \text{ and } 100$, and only show results for $\tau = 100$. Although the estimated inflation factor depends on the choice of τ , we found that the analysis RMSEs, on average, are not sensitive to this choice for the Lorenz 40-variable toy model. We also show results using the RTPS method with a range of α values (0-1) to serve as a benchmark. Note that $\alpha = 0$ corresponds to the case when no inflation is applied. For both sampling and model errors, we test the inflation methods either with or without applying covariance localization.

4.4 Results

4.4.1. Performance with sampling error due to limited ensemble size

Sampling error may occur when using a limited-size ensemble to estimate the background error covariance. **Figure 4.1** shows the results from the perfect-model fully-observed cases where we gradually increase sampling error by reducing sample size N . **Figures 4.1a** and **4.1c** show the analysis RMSE as a function of sample size for the case without and with localization, respectively, and **Figs. 4.1b** and **4.1d** show the

corresponding CR. RMSE values in **Fig. 4.1a** are also documented in **Table 4.1**. Generally, as N decreases, the filter diverges (RMSE jumps above observation noise while CR drops below 1) at some point. We can define three ranges of N values: 1) stable range in which the filter is stable even without inflation; 2) unstable range in which the filter diverges even for the best-tuned inflation; and 3) the transition range in between unstable and stable range. We are interested in the transition range, because within this range a properly tuned inflation method can help restore filter stability. For the case without localization (**Figs. 4.1a** and **4.1b**), the transition range is (15, 40). Results from non-adaptive RTPS show that, within this range, a larger amount of inflation (larger α for RTPS) is needed to restore filter stability for larger sampling error (smaller N). At $N = 30$, $\alpha = 0.1$ is enough, while at $N = 17$, $\alpha = 0.3$ is needed. In the stable range, no inflation is needed and $\alpha = 0$ results in the lowest RMSE; increasing inflation (larger α) will increase RMSE. In the unstable range, we see that using a larger α can help reduce RMSE slightly, but cannot prevent filter divergence. The ACR method is able to restore filter stability through the transition region, although the resulting RMSE is not as small as the best-tuned RTPS. For the ACI method, we found that relatively large $\sigma_\lambda^2 = 1$ is needed to bring similar improvement as ACR. Recall that the value of σ_λ^2 tunes how much an innovation impacts the final inflation field, thus a large σ_λ^2 will lead to a final inflation that is close to the value suggested by innovation statistics. Note that Anderson (2009) sets $\sigma_\lambda^2 = 0.01$ in his study. We found that this value only works well in the stable range. With localization (**Figs. 4.1c** and **4.1d**), the filter's tolerance to sampling error is greatly

increased. The transition range shifts to (4, 10). The ACR method can also restore filter stability in this case, and the ACI still needs at least $\sigma_\lambda^2 = 1$ to bring such improvement.

For the half-observed case, we only show results when localization is applied (**Fig. 4.2**), because the model becomes unstable when the unobserved variables are also inflated; localization helps limit the inflation within the observed region in RTPS. In this case, we calculated analysis RMSE separately for the observed variables (RMSE_land, **Fig. 4.2a**) and unobserved variables (RMSE_ocean, **Fig. 4.2b**), and **Fig. 4.2c** shows CR in observation space. Comparing RMSE_land to RMSE in **Fig. 4.1c**, we can see that the analysis RMSE is smaller for the land variables in the half-observed case compared with the fully-observed case. The ocean variables maintain a relatively large ensemble spread that propagates down-stream, effectively inflating the ensemble covariance over land. We found that the filter does not diverge for the land variables until N reaches 5, while the ocean variables maintain a noise level higher than the observation noise that remain within the climatological bound. Similar to the fully-observed case, the ACR method also restores filter stability for $N < 6$. The ACI still needs a larger σ_λ^2 value to exhibit the same improvement. The RMSEs for ocean variables are also improved for $N < 6$ when adaptive inflation is applied to the land variables, which may result from the more accurate state estimates over land that propagates down-stream to the ocean.

4.4.2. Performance with an imperfect forecast model

In this subsection, we test the inflation methods under the condition where an imperfect forecast model is used. **Figure 4.3** shows results from the fully-observed cases with $N = 40$ and no localization. RMSE values in **Fig. 4.3a** are also documented in

Table 4.1. In **Figs. 4.3a** and **4.3b**, the F parameter is the only model error source. In this case, we see that only a tiny model error ($F = 7.9$) will lead to filter divergence without inflation. The RTPS method shows that, for larger model error, larger α values are needed to restore filter stability. For example, $\alpha = 0.2$ is enough for $F = 7.9$, while $\alpha = 0.9$ is needed for $F = 5$. Similar to the sampling error cases, we found the ACR method able to estimate an α close to the optimal value and reduce the RMSEs significantly. As for ACI, $\sigma_\lambda^2 = 0.1$ only brings limited improvement, while $\sigma_\lambda^2 = 1$ brings improvement that is comparable to ACR method. To show the performance under a combination of model error sources, we also increase damping by 20% ($d = 1.2$, **Figs. 4.3c** and **4.3d**) and decrease nonlinear advection by 20% ($a = 0.8$, **Figs. 4.3e** and **4.3f**). In both cases, the optimal value for α is much larger (~ 0.8) with a smaller reduction in RMSE, with both cases RMSE still under the observation noise. We found that the ACR method handles all types of model error well; a similar improvement can be achieved by ACI only when using a relatively large σ_λ^2 .

Figure 4.4 shows results from half-observed cases for $N = 10$ with localization. In this case, similar to sampling error cases in **Fig. 4.2**, the filter tolerates model error more when localization is applied. The filter does not diverge without inflation until F is as small as 6 (**Fig. 4.4a**). We also show a combination of error sources from F , a and d as in **Fig. 4.3**, and found that the ACR method also handles all types of model error well, while ACI needs large σ_λ^2 to achieve this improvement.

4.4.3. Spatial and temporal behavior of inflation parameters

In this subsection, we examine the spatial and temporal behavior of the inflation parameters estimated by the ACR method, namely the $\alpha(\sigma_k^b - \sigma_k^a)/\sigma_k^a + 1$ term in (13). It is desirable that the adaptive methods automatically adjust the inflation parameters both in space and time to account for the inhomogeneity. **Figure 4.5** compares the spatial pattern of inflation parameters averaged over time between different inflation methods for two half-observed cases with large error sources. The inflation from RTPS can be thought of as a scalar α value ranging from 0 to 1 that tunes a spatial pattern (the relative ensemble spread reduction term). We set the localization distance to 10, so that observations over land ($k = 1, \dots, 20$) do not impact the center of the ocean ($k = 30$). The average inflation is larger over land than over the ocean, with a maximum near $k = 1$ and tapering to zero from $k = 20$ to $k = 30$. Since the domain is cyclic (periodic boundary condition), the maximum is at the downstream side of the ocean, where large uncertainties propagate toward the land. For the ACR method, the resulting inflation pattern has the same properties as the RTPS method. On the other hand, the ACI method uses the σ_λ^2 value to tune the overall amount of inflation. In the presence of large model error (**Fig. 4.5b**), $\sigma_\lambda^2 = 0.1$ leads to overall inflation of only ~ 1.01 over land. Increasing σ_λ^2 to 1 increases the inflation to ~ 1.05 , but the maximum near $k = 1$ is not well captured. Increasing σ_λ^2 to 10 produces the inflation pattern similar to ACR, but it appears to be noisier.

Figure 4.6 shows the time series of λ , the inflation factor estimated in observation space from ACR, for the cases with no error sources (**Fig. 4.6a**) and with large model

error (**Fig. 4.6b**). The temporal smoothing τ value is set to 1, 10, and 100. As expected, larger τ value leads to a more smoothed λ in time. For the case with no error sources, the λ value is close to 1 throughout the time period, while for the case with large model error, λ is elevated to ~ 1.2 . In our experiments, although λ has different temporal variability depending on the τ value used, the resulting temporally averaged λ and analysis RMSE are not particularly sensitive to the choice of τ (not shown). One caveat arises if τ is set too large, e.g. as large as the model integration time period, so the algorithm adjusts the inflation too slowly causing the filter to diverge before inflation reaches the required magnitude. Therefore, the optimal τ should not be too large so that the adjustment of the inflation magnitude is faster than the accumulation rate of unrepresented errors, but not too small so that inflation becomes too noisy in time.

The ACI method treats each innovation as one observation and uses it to update the inflation field. The updated inflation will be a weighted average of its prior value and the “observed” value suggested by the innovation, with the weight dependent on the selection of σ_λ^2 values. The ACR method will be equivalent to ACI if we use one innovation at a time in (7) to get an “observed” inflation factor and also update the prior inflation factor by the weighted average. We can reformulate the temporal smoothing equation (15) as an update equation that weights the observed and prior inflation factors, similar to the reformulation of the temporal smoothing equation in Li et al. (2009) by Miyoshi (2011) as a Gaussian approximation to the ACI method. In doing so, we should expect both methods to have similar sensitivity to the choice of σ_λ^2 or τ values. However, in our experiments, we used all p innovations at each time step to derive the “observed”

inflation factor deterministically in (7), rather than treating each innovation as a random variable. Even if $\tau = 1$, the inflation factor is still derived from innovation statistics using p observations, which is equivalent to using a large σ_λ^2 in the ACI method. In consequence, the analysis RMSE is insensitive to the choice of τ from 1 to 100. However, increasing τ reduces the noise and variability in the estimated inflation factor. From **Fig. 4.6** we see that the inflation factor has larger variance and frequently exceeds 1.5 when τ is set to 1. Though the noisy inflation factor works for the Lorenz 40-variable model, further research is needed to show whether it still works for other more complex models and/or observation network. It is possible that an inflation factor that has too large a variance in time may be a potential issue for less autonomous models.

4.5 Summary

Covariance inflation can improve filter performance by inflating the ensemble to account for the unrepresented sampling/model errors. Tuning the inflation parameter can be costly, therefore adaptive algorithms that estimate the inflation online according to innovation statistics have been introduced in previous studies (Wang and Bishop, 2003; Anderson, 2007; 2009; Li et al., 2009; Miyoshi, 2011). The online estimated inflation parameter adjusts to the value suggested by innovation statistics. When a spatially irregular observation network is used leaving some state variables unobserved, a spatially varying inflation parameter is likely needed. Inflating the unobserved variables may cause their ensemble spread to keep increasing, and therefore likely leads to filter/model instability. To introduce spatially varying covariance inflation, the ACI method

(Anderson, 2009) treats the inflation parameters as individual random variables and updates them using “observed” inflation from innovations. In this study, an alternative ACR method is introduced. We start with the RTPS method (Whitaker and Hamill, 2012), in which a spatially varying relative ensemble spread reduction term is used to mask the inflation to only the observed variables. We show that this reduction term is a direct result from assimilating observations whose spatial pattern depends on observation location and localization distance. The ACR method estimates the relaxation parameter in RTPS by matching innovation statistics in observation space, therefore it is equivalent to a spatially and temporally varying inflation. Numerical experiments with the Lorenz 40-variable model show that the ACR method can restore filter stability with the presence of sampling/model error for a range of severity. It is found that the spatial inflation pattern estimated by ACI is noisier compared to ACR and does not capture the location where ensemble spread deficiency is most severe. Also, ACI results are sensitive to the tunable variance of the inflation parameter, while ACR results are less sensitive to the tunable temporal smoothing parameter.

Both the RTPS and the algorithm for online estimation of relaxation parameter are easy to implement for the proposed ACR method. Therefore, it has the potential for real-data atmospheric model applications. The results in this study only apply to the Lorenz 40-variable model, thus there are several potential issues regarding the use of real observations and more complicated models that we want to stress. First, the real observations usually contain variables with different units. In this case, the innovation statistics (4.5) and (4.7) should use normalized observations so that each type of observation has comparable contribution (see Eq. 18 in Wang and Bishop, 2003). Second,

for complicated models, the dynamics can be also inhomogeneous in space. In this case, one needs to verify that the system of interest (e.g. a hurricane) is well covered by observations and the observations outside of the system are excluded so the innovation statistics truly reflect the inflation needed for the system of interest. Finally, for real observations, the observation operator may be nonlinear. Since we estimate the relaxation parameter in observations space and apply it in state space, the nonlinear relation between observation and state space may cause the relaxation parameter to be either under- or overestimated. In this case, one may need a tangent linear observation operator to map state space to observation space.

Table 4.1. Analysis RMSEs using no inflation, ACR, best-tuned RTPS and ACI methods (columns) for different error sources (rows) including sampling error and model error. The fully-observed case without localization is shown (corresponding to Figs. 4.1a and 4.3a).

	No inflation	ACR	RTPS (optimal α)	ACI 0.1	ACI 1
$N = 80$	0.1920	0.2163	0.1851 (0.1)	0.1925	0.1955
40	0.2181	0.2275	0.1821 (0.1)	0.2106	0.1977
20	4.0032	0.2766	0.1926 (0.2)	2.4608	0.3541
17	4.1459	0.4561	0.2198 (0.3)	2.7773	0.5846
15	4.2028	1.6785	1.5101 (0.9)	3.0480	0.8755
10	4.4331	3.3941	2.9290 (0.9)	3.8840	3.3389
5	4.7771	4.5219	3.7310 (1.0)	4.6173	4.4831
$F = 8$	0.2154	0.2378	0.1880 (0.1)	0.1979	0.2012
7.9	3.9566	0.2918	0.2221 (0.3)	1.9667	0.3285
7.5	4.0564	0.4435	0.3424 (0.6)	2.0047	0.6668
7	4.0107	0.5835	0.4231 (0.7)	2.2781	0.8577
6	4.0423	0.7783	0.5234 (0.8)	2.7079	1.0630
5	4.1770	0.9044	0.5939 (0.9)	2.9864	1.1891

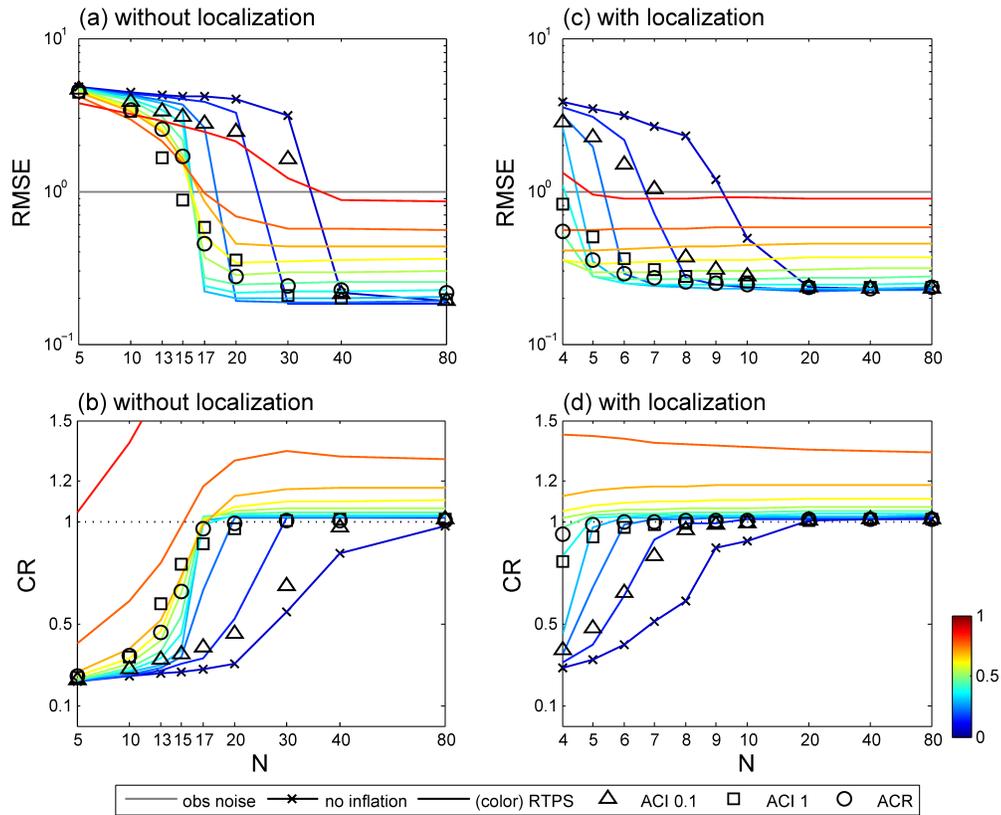


Figure 4.1. Analysis RMSE (top row) and consistency ratio (bottom row) for the fully-observed cases with different degrees of sampling errors introduced by changing the ensemble size (N). Left (right) column shows result without (with) localization ($ROI = 10$). The gray line shows the observation noise level. The colored lines show inflation parameters for RTPS with α ranging from 0 to 1, the cross markers (no inflation) is equivalent to $\alpha = 0$. Triangle and square markers correspond to the ACI methods with $\sigma_\lambda^2 = 0.1$ and 1, respectively. Circle markers correspond to the ACR method. Note that we show more data points in the transition region ranging from the filter being stable to diverging. The transition region is different for the cases without localization ($N = 15-40$) and with localization ($N = 4-10$).

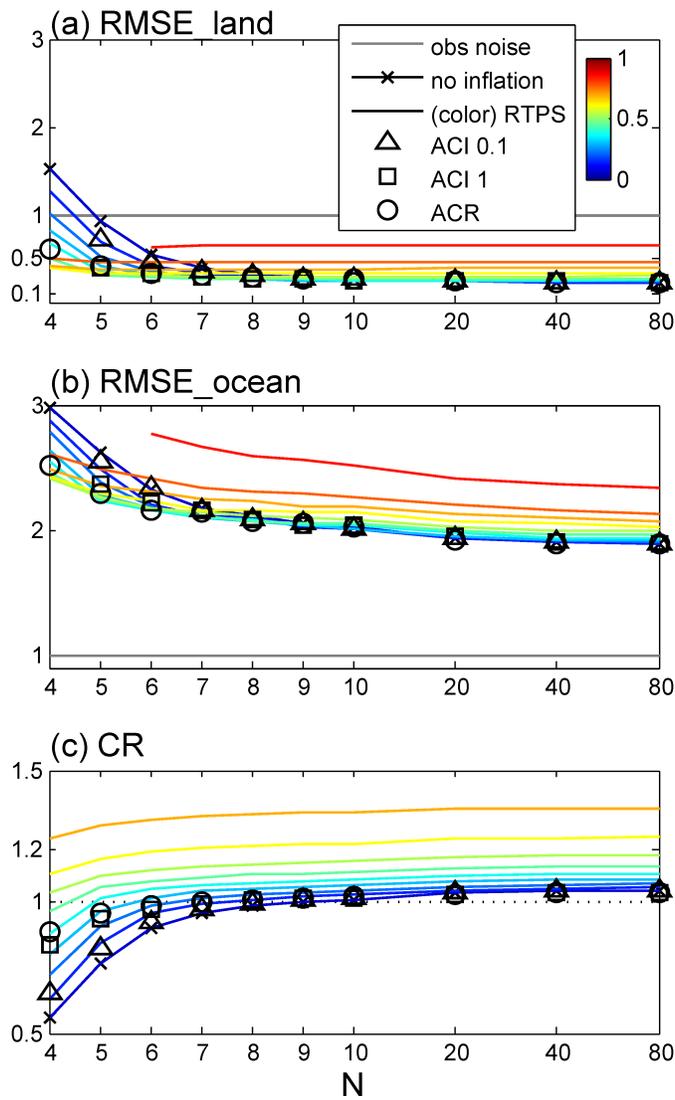


Figure 4.2. Analysis RMSE and consistency ratio for the half-observed cases ($N=10$ with localization). RMSE_land (RMSE_ocean) is calculated using the observed (unobserved) variables along with the truth, while the consistency ratio is calculated in observation space (observed variables only). The legend is the same as Fig. 4.1.

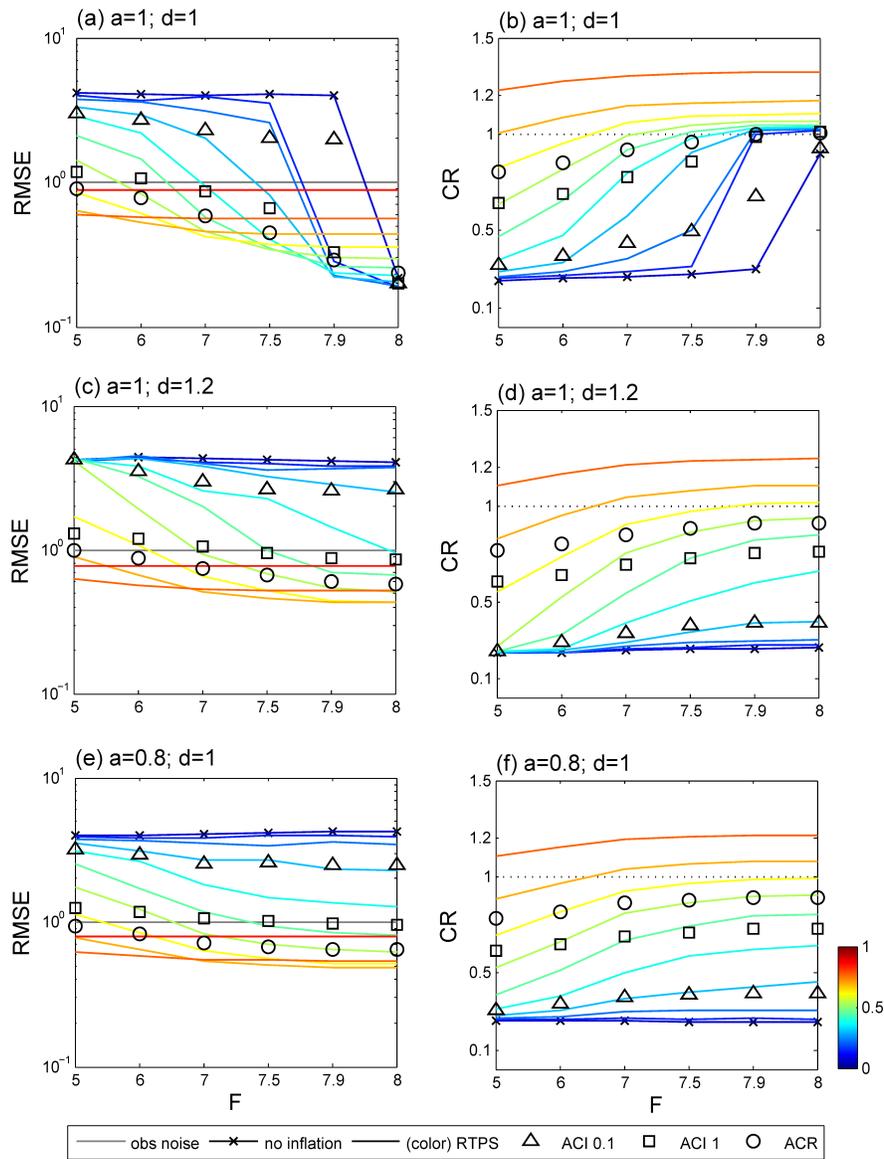


Figure 4.3. Analysis RMSE (left column) and consistency ratio (right column) for the fully-observed cases ($N=40$ without localization) with different degree of model errors introduced by varying parameters F , a and d . Within each panel, F varies from 5 to 8. For each row, a different combination of a and d values are used. The legend is the same as Fig. 4.1.

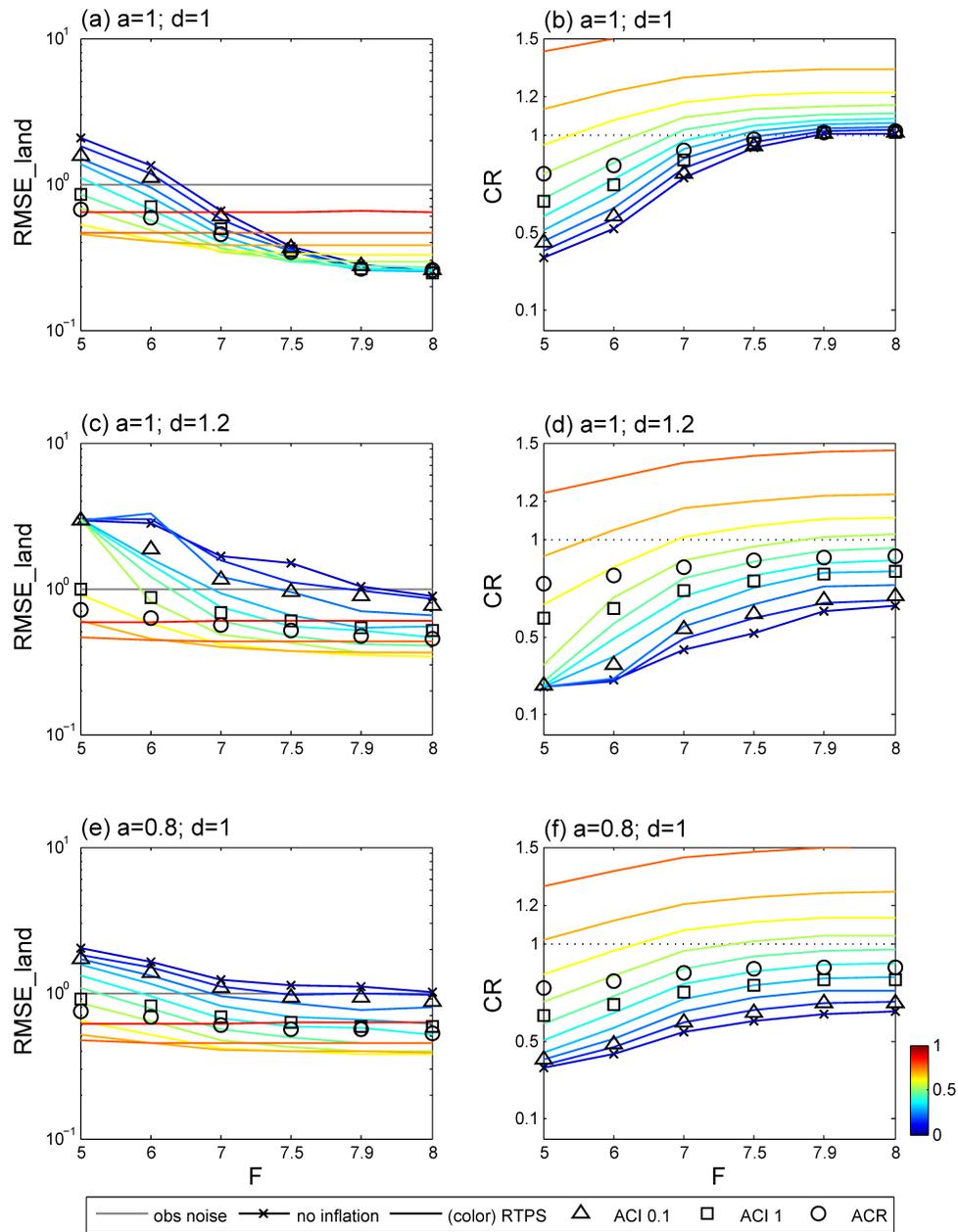


Figure 4.4. Same as Fig. 4.3, but showing half-observed cases ($N = 10$ with localization).

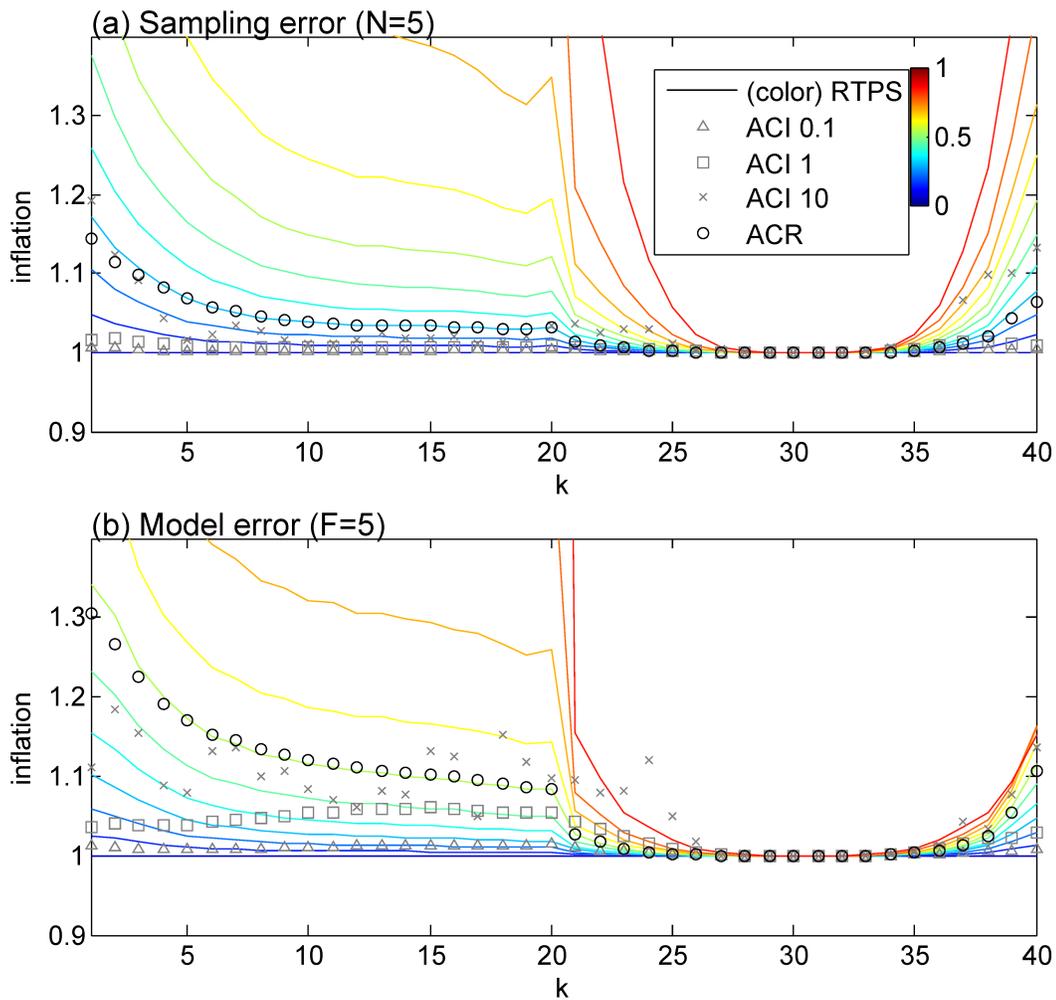


Figure 4.5. The spatial structure of inflation parameters for the half-observed cases ($N = 10$ with localization) with two types of severe error sources: (a) sampling error with $N = 5$ and (b) model error with $F = 5$. The colored lines show inflation parameters for RTPS with α ranging from 0 to 1. Gray triangle, square and cross markers correspond to ACI methods with $\sigma_\lambda^2 = 0.1, 1$ and 10 , respectively.

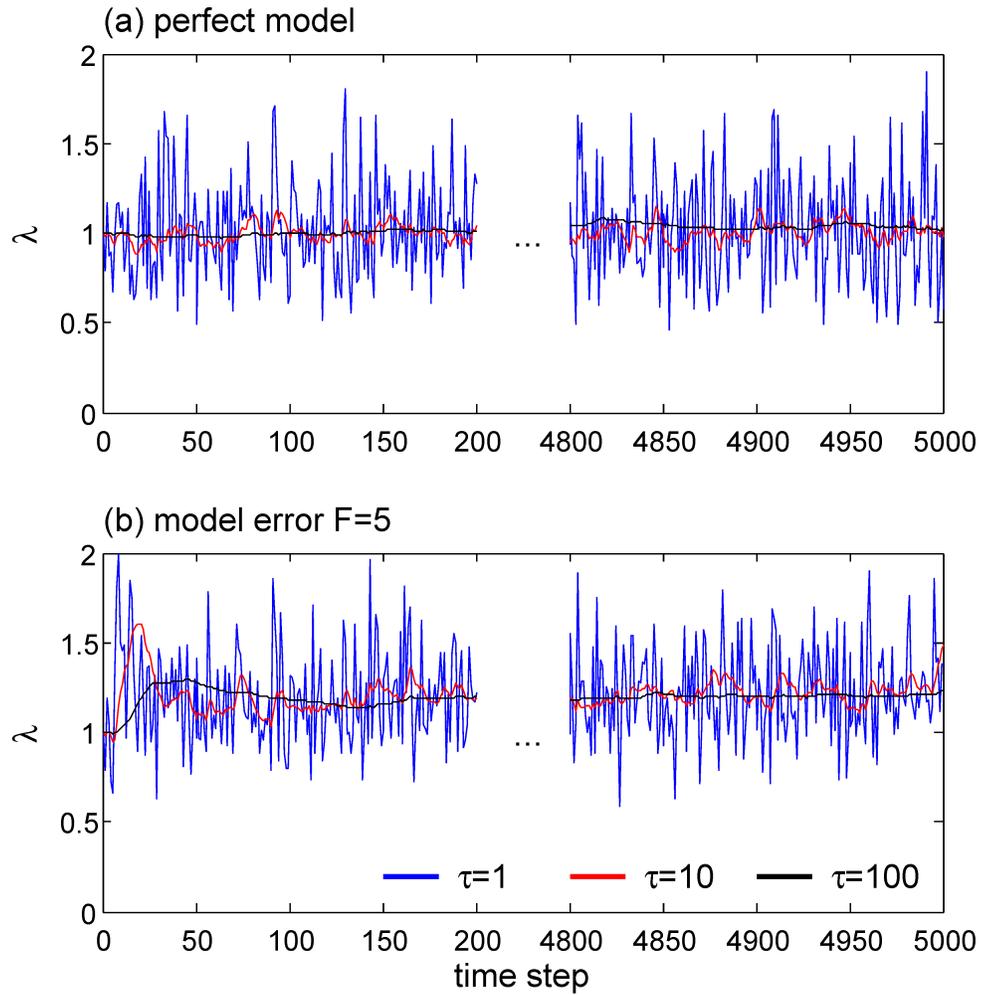


Figure 4.6. Time series of the inflation factor (λ) from ACR method for half-observed cases ($N=10$ with localization) with (a) perfect model and (b) model error $F=5$. Only the first and last 200 steps are shown. Blue, red and black lines correspond to $\tau = 1, 10$ and 100 , respectively.

Chapter 5

On the selection of localization radius in ensemble filtering for multiscale quasi-geostrophic dynamics

5.1 Introduction

Covariance localization (Hamill et al. 2001) is a pragmatic procedure for ensemble filters (Houtekamer and Zhang 2016) to remedy sampling errors due to limited ensemble size. In practice, the ensemble size is usually much smaller than the number of state variables and observations, and is not large enough to span all possible directions in phase space for the update, which is known as the rank problem (Lorenz 2003). If no (or insufficient) localization is applied, the spurious long-distance sample-estimated correlation is not removed and contaminates the filter performance. On the other hand, if too much localization is applied, it not only discards useful observations, but also introduces noise at small scales due to the artificial tapering of the analysis impact. The noise introduced by localization also causes physical imbalance of the analysis (Kepert 2009; Greybush et al. 2011; Lange and Craig 2014). A simple localization function is the Gaspari and Cohn (1999) fifth-order polynomial with a specified cutoff radius (hereafter referred to as the GC function). This tapering function can be either applied directly to the background error covariance in model space (Houtekamer and Mitchell 1998), or to the analysis increment limiting the impact of observations to nearby state variables (Houtekamer and Mitchell 2001; Hamill et al. 2001). The result of either approach is

similar but the effective localization radius is smaller for the latter (Sakov and Bertino 2011; Nerger et al. 2012). The optimal localization radius finds a balance between the removal of sampling noise and the preservation of useful observation information. Operationally, the tuning of localization radius considers multiple factors including fit to observations, balance, computational cost, model resolution, and observation density. Manual tuning of the localization radius by trial and error for a particular model and ensemble filtering system can be very costly. One may desire a more flexible adaptive scheme to determine the localization a priori. However, to design such a scheme requires knowledge of the complex codependence of localization on the underlying correlation scale determined by the model dynamics, the ensemble size, and the observing network.

Anderson (2007) demonstrated that the optimal localization function could be quite different from a smoothed Gaussian. This motivates the exploration of localization functions that adapt to the flow-dependent error covariance (Anderson 2007, 2012; Bishop and Hodyss 2007, 2009). Anderson and Lei (2013) derived an Empirical Localization Function (ELF) based on the information from an Observing System Simulation Experiment (OSSE). The physical intuition behind these adaptive methods is that localization should only reduce the observation impact when the signal-to-noise ratio is low. In practice, localization distance is often tuned to scale with the overall correlation length. The tuned localization distance is $O(1000)$ km for global modeling and data assimilation systems, but a much shorter localization distance of $O(10)$ km is found more suitable for convective weather systems using high-resolution models and observations (Zhang et al. 2009; Sobash and Stensrud 2013). Since the horizontal correlation length scale increases with height, several studies have found that localization radius should also

increase with height (Zhu et al. 2013; Houtekamer et al. 2014; Kleist and Ide 2015). Lei et al. (2015) demonstrated that a narrower horizontal but wider vertical localization scale is preferred for precipitating regions. For atmospheric flows with multiple spatial scales (e.g. high-resolution model domains capturing both synoptic- and convective-scale flows), the data assimilation scheme should handle multiple correlation lengths simultaneously. Zhang et al. (2009) first proposed a Successive Covariance Localization (SCL) approach that localizes the observation impact with a hierarchy of radii to account for different physical length scales. Multiscale localization methods are further explored and found more advantageous than normal single-scale localization in several other studies (Miyoshi and Kondo 2013; Li et al. 2015; Buehner and Shlyayeva 2015).

Despite all these efforts to develop a better localization scheme, a comprehensive theory of localization is still lacking. It is still not clear what makes a smaller localization radius more suitable for convective-scale weather systems, whether it is due to the higher model and observation resolution or the shorter overall correlation lengths. Zhen and Zhang (2014) systematically explored the codependence of localization on the underlying physical scales and the ensemble size for a single observation. As a result, an optimal localization scheme was derived and tested in the Lorenz (1996) model framework. Flowerdew (2015) proposed a similar method but allowed the localization function to have a shape other than Gaussian. These methods are yet to be tested in more complex models that contain multiple physical scales. A theory for optimal localization with a dense observing network may be more complicated than the theory for a single observation (Flowerdew 2015). Previous studies have documented the sensitivity of localization to observation density. From a perfect-model OSSE, Anderson (2007)

showed that the optimal localization function is broader for regions with dense observations than for an isolated observation. However, in real-data experiments, Dong et al. (2011) and other studies found that a smaller localization radius is necessary to achieve better analysis accuracy for denser observing networks. Kirchgessner et al. (2014) suggested that optimal localization radius is obtained when the effective observation dimension is about equal to the ensemble size for dense observations. Perrianez et al. (2014) derived an optimal localization radius by high-level heuristic arguments assuming a uniform observing network, and they also suggest using a smaller localization radius for denser observations. These studies suggest there may be a more complicated relation between observing network and localization.

In this chapter, we seek to systematically explore the relative importance of physical correlation lengths, model resolution, ensemble size, and observing networks to the selection of localization radius, which will provide insights on the development of a better localization scheme for multiscale weather systems. A series of sensitivity experiments are conducted using the two-layer quasi-geostrophic (QG) model. The QG model, although simple, can capture the essence of multiscale atmospheric dynamics. GC functions with fixed cutoff radii are investigated in a serial ensemble Kalman filter (EnKF), and the best localization radius for a given scenario is determined by trial and error. Section 5.2 describes the QG model and EnKF configuration and the design of sensitivity experiments. Section 5.3 demonstrates the scale dependency of the best localization radius, and is followed by sensitivity experiment results in section 5.4 that show how the best localization radius varies in response to changes in model resolution, ensemble size, and observing network. Our findings are summarized in section 5.5.

5.2 Experimental design

5.2.1 Two-layer quasi-geostrophic model

The two-layer QG model described in Smith et al. (2002) is adopted to perform numerical experiments in this study. Previous literature has comprehensively documented its dynamical processes (Larichev and Held 1995; Held and Larichev 1996; Salmon 1998; Smith et al. 2002). Harlim and Majda (2010) used this model to investigate the assimilation of a sparse observing network for the atmosphere. The model is defined on a doubly periodic square domain and simulates the large-scale atmospheric flow with baroclinic instability induced from an imposed vertical wind shear. Background streamfunction is defined as $\Psi_1 = -Uy$ for the top layer and $\Psi_2 = Uy$ for the bottom layer, where U is the mean flow. The prognostic equations for perturbations around this background state can be written as

$$\frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U \frac{\partial q_1}{\partial x} + (\beta + k_d^2 U) \frac{\partial \psi_1}{\partial x} = 0, \quad (5.1)$$

$$\frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U \frac{\partial q_2}{\partial x} + (\beta - k_d^2 U) \frac{\partial \psi_2}{\partial x} + r \nabla^2 \psi_2 = 0, \quad (5.2)$$

$$q_1 = \nabla^2 \psi_1 + \frac{k_d^2}{2} (\psi_2 - \psi_1), \quad (5.3)$$

$$q_2 = \nabla^2 \psi_2 - \frac{k_d^2}{2} (\psi_2 - \psi_1), \quad (5.4)$$

where subscript 1 denotes the top layer and 2 the bottom layer; ψ is the perturbation streamfunction and q is the perturbation QG potential vorticity; $J(\psi, q) = \partial_x \psi \partial_y q - \partial_y \psi \partial_x q$ is the Jacobian term representing the nonlinear advection; β is the meridional gradient of the Coriolis parameter; k_d is the Rossby deformation wavenumber; r is the

strength of linear Ekman drag that removes large-scale energy buildup from the bottom layer. Two characteristic wavenumbers corresponding to the Rossby deformation scale and the Rhines scale are defined as

$$k_d = (L/2\pi)f/\sqrt{g'H}, \quad (5.5)$$

and

$$k_\beta = (L/2\pi)\sqrt{\beta/U_0}, \quad (5.6)$$

respectively, where f is the Coriolis parameter, g' is the reduced gravity, H is the vertical scale height, U_0 is the horizontal velocity scale, and $L/2\pi$ is a scaling factor. The kinetic energy spectrum of the QG model features an energy injection due to baroclinic instability near the deformation scale k_d^{-1} . For scales larger than the deformation scale ($k < k_d$), there is an inverse cascade of kinetic energy with a $-5/3$ power law. The cascade halts at a scale $k_{\beta,r}^{-1}$ that is determined by both the β effect and the bottom drag r . At scales smaller than the deformation scale ($k > k_d$), the enstrophy cascades forward and dissipates at the smallest scales, resulting in a -3 power law for kinetic energy. The model mimics this dissipation with an exponential cutoff filter that removes energy buildup at the smallest scales (see Appendix B of Smith et al. 2002).

A baseline configuration used in this study sets the model parameters as $k_d = 20$, $k_\beta = 4$, $U = 0.2 U_0$ and $r = 0.5$. Let $L = n dx$ be the domain size, where n is the number of grid points in both zonal and meridional directions and dx is the grid spacing. The model resolution is set to $n = 128$, which resolves $k_{\max} = 63$ modes in each direction. The exponential filter cutoff wavenumber is $k = 40$. These parameters are chosen similarly to the atmospheric case from Harlim and Majda (2010) except that the deformation

wavenumber here is larger to produce a wider range of scales with baroclinic instability. The typical deformation length scale for the large-scale atmospheric flow is 1000 km. With this scaling, the model grid spacing dx corresponds to ~ 80 km in real atmospheric models. The average eddy turnover time is ~ 0.15 non-dimensional time units, which corresponds to ~ 2 days.

5.2.2 Ensemble filter

The data assimilation method used in this study is the ensemble square root filter (Whitaker and Hamill, 2002), a serial EnKF, which is described as follows. Let x be the state variable vector, and let y^o be the observation vector. An ensemble of N members is introduced to estimate the flow-dependent background error covariance. For each observation indexed with subscript j , the following equations are applied to update the ensemble serially to reach the final analysis.

$$(\bar{x})_{\text{update}} = \bar{x} + \rho_j \circ K_j (y_j^o - H_j \bar{x}), \quad (5.7)$$

$$(x'_i)_{\text{update}} = x'_i + \gamma_j \rho_j \circ K_j (0 - H_j x'_i), \quad \text{for } i = 1, 2, \dots, N, \quad (5.8)$$

where \bar{x} is the ensemble mean, x'_i is the ensemble perturbation for the i -th member, ρ_j is a localization function, K_j is the Kalman gain defined as

$$K_j = \frac{\text{cov}(H_j x, x)}{\text{var}(y_j^o) + \text{var}(H_j x)}, \quad (5.9)$$

γ_j is a square-root modification term defined as

$$\gamma_j = \left(1 + \sqrt{\frac{\text{var}(y_j^o)}{\text{var}(y_j^o) + \text{var}(H_j x)}} \right)^{-1} \quad (5.10)$$

Here $\text{cov}(H_j x, x)$ is the background error covariance between observations and state variables, $\text{var}(y_j^o)$ is the observation error variance, and $\text{var}(H_j x)$ is the associated background error variance. The linearized observation operator H_j is not used in this study. Instead, the $H_j \bar{x}$ and $H_j x_i'$ terms are approximated by $\overline{h_j(x)}$ and $h_j(x_i) - \overline{h_j(x)}$, respectively, where h_j is the nonlinear observation operator and overbars denote ensemble averages.

GC functions with fixed localization radii are adopted as the localization function ρ . Localization radius (radius of influence, or ROI) is defined as the physical distance at which the analysis increments are tapered to zero. The adaptive covariance relaxation method introduced in the chapter 4 is also applied to maintain ensemble spread and prevent catastrophic filter divergence.

5.2.3 Observing network and assimilation experiments

A control (CNTL) experiment is first performed using the baseline configuration in section 5.2.1. The truth (nature run) is generated by initializing the QG model with white noise and running the model for 50 time units until it reaches a quasi-steady state. A 15-time-units model run (~ 200 days) during the quasi-steady state is taken as the truth. Cycling data assimilation is performed using an ensemble of $N = 64$ members and assimilating synthetic observations simulated from the truth every $\Delta t = 0.05$ time units,

which yields 300 cycles in total. Under a perfect-model assumption, the truth model is used as the forecast model during cycling data assimilation. Using the trial-and-error method, the best-performing ROI that yields the lowest analysis error variance is determined. Analysis errors are also decomposed into spectral components to evaluate the scale dependency in filter performance. Sensitivity experiments are designed to test how the best-performing ROI changes in response to changes in model resolution, ensemble size and observing network. **Table 5.1** summarizes the model and filter parameters used in each experiment. ROI is defined as a physical distance and expressed in terms of the number of grid points in CNTL (i.e. ROI = 8 means a cutoff distance of 8 dx). Cases with ROI = 8, 16, 32, 64, as well as no localization (ROI = ∞), are tested for each experiment. Some additional values of ROI are tested in the neighborhood of the minimum analysis error to more accurately estimate the best-performing ROI.

Instead of assimilating the state variable ψ directly, a more challenging scenario is considered where temperature (θ) or horizontal winds (u and v) are observed from the top model layer. Let k_x and k_y be the zonal and meridional wavenumbers, respectively. The nonlinear observation operators that convert ψ to u , v , and θ can be described in spectral space.

$$\hat{u} = -ik_y \hat{\psi}, \quad (5.11)$$

$$\hat{v} = ik_x \hat{\psi}, \quad (5.12)$$

$$\hat{\theta} = -\sqrt{k_x^2 + k_y^2} \hat{\psi}, \quad (5.13)$$

where hats denote the two-dimensional Fourier transform of a variable. **Figure 5.1** shows snapshots of the top-layer θ and u in comparison to the corresponding ψ from the truth

simulation. The θ field is overall in phase with ψ but has more small-scale details. On the other hand, u is not in phase with ψ and their peaks are not collocated. Synthetic observations are generated by first converting ψ from the truth simulation to observations according to (5.11)-(5.13), and then adding simulated observation errors randomly drawn from a zero-mean normal distribution with error variance of σ^2 . The observation errors are considered uncorrelated both in space and in time. For CNTL, a uniform observing network of top-layer θ with $\sigma = 3$ is used. Although the bottom layer is not observed, the top-layer observations provide information for both layers because the flow simulated in CNTL is mostly barotropic, especially for the large scales. However, not all model grid points on the top layer are observed. Therefore, the horizontal propagation of information is important for good filter performance. Let Δx be the spacing between observations in both directions, the number of observations in each direction is $n_o = n dx/\Delta x$. For CNTL, the observation density is set to $\Delta x = 3 dx$. For a uniform observing network with independent random instrument errors, the spectral variance of its observation error can be exactly calculated. The observation error variance associated with wavenumber k , $R(k)$, is inversely proportional to the number of observations n_o^2 ,

$$R(k) = 2\pi k \sigma^2/n_o^2. \quad (5.14)$$

With fixed instrument error σ , a denser observing network yields lower observation error spectral variance thanks to the larger number of observations. Sensitivity to changes in observing network, including its density, accuracy and spatial homogeneity, will be tested in section 5.4.3.

To diagnose filter performance, the analysis error (difference between posterior ensemble mean and the truth) variance is averaged over the cycling data assimilation period. When θ is assimilated, the state variables will be converted to θ to calculate the analysis error variance. When u and v are assimilated, the analysis error variance for u and v will be calculated and then combined as an error kinetic energy, which is defined as $(u^2 + v^2)/2$. The errors are further decomposed into spectral components to facilitate the comparison across scales. The analysis error variance associated with the wavenumber k component is defined as

$$E(k) = \sum_{k_x^2 + k_y^2 = k^2} [\hat{\varepsilon}(k_x, k_y)]^2, \quad (5.15)$$

where $\hat{\varepsilon}$ is the two-dimensional Fourier transform of the analysis error and $k = \sqrt{k_x^2 + k_y^2}$ is the global wavenumber. Summing $E(k)$ over a range of wavenumbers will yield the mean square error (MSE), or root mean square error (RMSE) if the square root is taken, associated with this scale range. In this study, the spectrum is divided into three bands: large (L; $k = 1-4$), intermediate (M; $k = 5-20$), and small (S; $k = 21-63$) scales.

To generate the prior ensemble, random white noise is added to the truth initial condition for each member and an ensemble forecast is run for a spin-up period of 1.5 time units until the ensemble spread reaches climatological level. **Figure 5.2a** shows the kinetic energy spectrum from the truth (black line) and time evolution of error kinetic energy from the ensemble mean (colored lines) during this spin-up period. Error saturates when its kinetic energy reaches the level of the reference kinetic energy. The M scale follows a $-5/3$ power law associated with the inverse energy cascade, where the small-

scale errors saturate faster than the large-scale errors and an overall upscale error growth is present. As errors grow upscale, they have more large-scale components, which results in an increasing overall correlation length scale. An estimation of this correlation length is shown in **Fig. 5.2b** for θ observations and **Fig. 5.2c** for u observations. The mean absolute correlation (MAC) estimated from the ensemble is averaged over the observing network and plotted as a function of horizontal distance. Both MACs for θ and u are broadening over time. The MAC for u has a nonlocal peak due to the fact that u and ψ maxima are not collocated, and the correlation between u and ψ is overall much lower than the correlation between θ and ψ . The ensemble after spin-up is used as the prior for the ensuing data assimilation cycles.

5.3 Scale-dependent best localization radius

Figure 5.3 shows the time evolution of θ MSE at L, M, and S scales from CNTL using ROI = 8, 16, 32 and 64. The first several cycles undergo an adjustment (filter spin up) period of approximately 20 cycles before the filter performance reaches steady state. The cases using smaller ROIs experience longer filter spin up periods than the cases using larger ROIs. Earlier in the spin up period, the larger ROI = 32 and 64 cases perform better than smaller ROIs. However, for the steady state, the ROI = 16 case appears to improve, while the ROI = 64 case is clearly suboptimal. At L scale, ROI = 32 is more favorable. ROI = 16 occasionally performs as well as ROI = 32, but on average is worse. On the other hand, at M scale, the ROI = 16 and 32 cases are competitive at steady state. For the S scale, the ROI = 16 case appears to have better performance. These results

indicate a scale dependency for favorable ROIs. The following calculation of time-averaged analysis errors will exclude the first 50 cycles to make sure that results reflect steady-state filter behavior, and the long record (250 cycles) ensures robust statistics not influenced by transient behavior. **Figure 5.4a** shows the time-averaged θ spectral error variance $E(k)$ from CNTL. Along with the analysis error variance, the observation error variance $R(k)$ and error variance from a free ensemble without data assimilation (NoDA) are also plotted for reference. The observation error variance follows a +1 power law associated with white noise, and the NoDA error is fully saturated at all scales and thus resembles the reference spectrum from the truth. Note that the θ spectrum follows a similar power law to the kinetic energy. From L to S scales, the best ROI that minimizes analysis error variance at that scale appears to shift toward smaller values. The S scale is not observed due to coarser observation resolution than the model grid, thus the analysis errors remain mostly saturated at this scale.

Two experiments, M_Scale and S_Scale, similar to CNTL but with changed reference kinetic energy spectra are conducted to further demonstrate the scale dependency of the best ROI. **Figures 5.4b** and **5.4c** show their resulting error spectra. The CNTL reference spectrum has a peak at the L scale ($k = 3$), and this spectral peak is shifted toward M and S scales by changing model parameters r , k_d , and U . As bottom drag r increases, more energy is removed from the large-scale end of the spectrum, and the halting scale $k_{\beta,r}^{-1}$ becomes smaller. The zonal wind shear ($\pm U$) is increased to inject more baroclinic instability to ensure that eddies have similar amplitudes as the CNTL. For S_Scale, the deformation scale (k_d^{-1}) is also shifted to a smaller scale ($k_d = 35$) to

allow baroclinic instability to develop. The resulting M_Scale and S_Scale experiments simulate eddies with different sizes from CNTL. M_Scale is dominated by eddies at M scale and S_Scale by eddies at S scale, which is also reflected in their averaged eddy turnover time (~ 0.1 for M_Scale and ~ 0.07 for S_Scale).

Results confirm that the best-performing ROI decreases as the dominant scale of the system becomes smaller. For a given system, if there is one dominant scale, our results suggest that this scale will determine the best localization radius. A fixed ROI works well when there is only one dominant scale. **Figure 5.5a** plots the analysis errors for all scales with respect to ROIs. The overall best ROI that minimizes domain-averaged analysis RMSE is near 24 for CNTL, and it shifts to 16 for M_Scale and 12 for S_Scale. When a wider range of scales is present, a fixed ROI may become insufficient to minimize errors at all scales, and a different ROI should be specified for each scale, as suggested by previous studies (Zhang et al. 2009; Miyoshi and Kondo 2013; Li et al. 2015; Buehner and Shlyayeva 2015). **Figure 5.5b** illustrates this by plotting the CNTL analysis RMSE filtered for the L, M, and S scales with respect to ROIs. The L-scale component favors ROI = 32, the M scale favors ROI = 24, and the S scale favors ROI = 16.

The relationship between correlation length scale and localization distance is not necessarily one-to-one (Anderson and Lei 2013; Anderson 2016). In this study, we hypothesize that the best localization radius scales with the overall correlation length. One can consider a multiscale data assimilation problem as successively constraining from the large to small scales that have decreasing correlation lengths. **Figure 5.6** plots the MAC functions filtered for L and S scales from CNTL. The shape of correlation

functions on average is very different for L and S scales. The correlation function is broader at L scale than at S scale. The magnitude of overall correlation at zero distance is lower at L scale than at S scale. When a larger ROI is used, the averaged correlation remains lower than when a smaller ROI is used, which indicates larger ROIs constrain the analysis ensemble more than smaller ROIs. The analysis ensemble also tends to be more dispersive for cases using smaller ROIs, which can be inferred from the amount of inflation determined by the adaptive covariance relaxation algorithm. The relaxation coefficient averaged over time for the ROI = 8, 16, 24, 32, and 64 cases are $\alpha = -0.23, 0.17, 0.27, 0.35, \text{ and } 0.57$, respectively. A larger α means more inflation is applied to the analysis ensemble.

5.4 Sensitivity experiments

5.4.1 Model resolution

Numerical models usually cannot resolve all the scales of an atmospheric flow. In this section, we investigate cases using models that have lower resolutions than the truth model that generates the synthetic observations. For the QG model used in CNTL, the S scale features a forward enstrophy cascade, and the M scale features an inverse energy cascade. A model with resolution reduced to $k_{\max} = 31$ is first investigated. Comparing to the CNTL model, it has most of its S scale truncated and cannot accurately resolve the forward enstrophy cascade. Model parameters are selected to ensure that the low-resolution model has a large-scale energy spectrum matching with the high-resolution model in CNTL. Without the forward enstrophy cascade, a stronger enstrophy filter

(exponential cutoff at $k = 20$) is applied, and the baroclinic instability is slightly increased. Such model tuning is typically done in real atmospheric models, too, although more sophisticated methods such as parameterization are used to account for processes that these models cannot resolve. The low-resolution model is first tested in a perfect-model scenario in the LowRes experiment, where the observations are generated from a truth run using the same low-resolution model. **Figure 5.7a** shows the analysis error spectra from the cases using different ROIs. Other than the slightly elevated smallest-scale energy in LowRes, the results are very similar to the CNTL experiment (**Fig. 5.4a**). In LowRes_Model experiment, the low-resolution model is only used in the forecast step of the data assimilation and observations sampled from the high-resolution truth in CNTL are assimilated. The resulting analysis error spectra are shown in **Fig. 5.7b**. **Figure 5.8** plots the analysis RMSE filtered for L and M scales with respect to ROIs. Comparing LowRes (red), and LowRes_Model (blue) to CNTL (black), the reduced model resolution does not appear to influence the best ROI at the well-resolved L scale, while the best ROI at M scale slightly decreases due to the representation errors. The time-averaged adaptive relaxation coefficients are $\alpha = 0.25$ for LowRes (similar to CNTL where $\alpha = 0.27$) and $\alpha = 0.32$ for LowRes_Model, indicating more inflation is applied to the ensemble spread when representation errors are present.

Another model with even lower resolution $k_{\max} = 21$ is tested in LowRes2 (**Fig. 5.7c**) and LowRes2_Model (**Fig. 5.7d**). In this case, the lack of resolution starts to influence the representation of baroclinic instability near the deformation wavenumber. To fully resolve the baroclinic instability, the deformation wavenumber is changed to $k_d = 14$. Although some tuning is made to match the model climatology with CNTL, the

error growth rate is slightly higher at larger scales for this low-resolution model. This is a common situation for real atmospheric models where key dynamical processes cannot be well-represented at the small scale and therefore model forcing is biased. In the perfect-model scenario, LowRes2 (**Fig. 5.7c**) shows that the best ROI becomes larger due to the increased error growth rate at larger scales. **Figure 5.8** also shows this shift in best ROI from CNTL to LowRes2 for both L and M scales. When observations from CNTL are assimilated instead, LowRes2_Model (**Fig. 5.7d** and green line in **Fig. 5.8**) shows that although the steady-state analysis RMSE is larger due to the presence of larger model errors, the best ROI stays relatively unchanged from CNTL for the L scale. The averaged adaptive relaxation coefficient is larger for LowRes2_Model ($\alpha = 0.58$) than for LowRes2 ($\alpha = 0.23$) due to the presence of a large representation error.

Previous studies (e.g. Aksoy et al. 2012) have demonstrated the negative impact from representation errors when assimilating observations from a high-resolution nature run while the model is at a coarser resolution. Various methods are proposed to account for these representation errors in data assimilation algorithms to reduce their negative impact (Janjić and Cohn 2006; Bocquet et al. 2011; Hodyss and Nichols 2015; van Leeuwen 2015). Our results indicate that localization is not sensitive to model resolution as long as a dynamical process is well resolved and/or model representation errors are well-accounted for in the data assimilation scheme. Note that we define the localization ROI in terms of the physical length which does not change as model resolution reduces. If the ROI is defined as a number of grid points, it will change as model resolution changes (i.e fewer grid points for lower resolution). Also note that our conclusions are

drawn from QG model experiment results. Whether this can be generalized to other systems is yet to be confirmed with further experimentation.

5.4.2 Sampling error due to limited-size ensemble

Given the same observing network and a perfect model, a larger ensemble size provides more accurate sample-estimated error covariances and therefore improves the accuracy of the analysis. The negative impact of sampling error on filter performance can be demonstrated by decreasing the ensemble size N . To test the sensitivity of localization to N , experiments are conducted with $N = 16, 32, 256, 1024$, and compared to $N = 64$ as in CNTL. **Figures 5.9a-c** plot the resulting L-, M-, and S-scale analysis RMSEs as a function of ROI. The trend that the ROI that minimizes analysis RMSE at a certain scale becomes larger as N increases is seen for all three bands. These results are consistent with previous studies (Houtekamer and Mitchell 2001; Lorenc 2003; Zhang et al. 2006; Anderson 2007, 2012). When a larger N is used, there is also a wider range of ROIs around the best ROI that produce similar filter performance, indicating less sensitivity to localization scale in this case (Lei and Whitaker 2017). Not surprisingly, the minimum analysis RMSE is achieved using the largest $N = 1024$. This is consistent with the recent study of Kondo and Miyoshi (2016) who demonstrated with their 10240-member ensemble that with large enough ensemble size one can achieve the best filter performance without localization. **Figure 5.10a** shows the steady-state MACs estimated from analysis ensembles of different sizes. Correlation decreases as horizontal distance increases. Beyond the decorrelation length scale, the true correlation should be zero on average. Due to sampling error, the limited-sized ensembles have MACs that asymptote

to non-zero correlation values, and the asymptotic value increases as N decreases. Since data assimilation constrains the observed scales and reduces error, the analysis error has less large-scale contribution and a shorter overall correlation length compared to errors from a free ensemble without data assimilation (Yoon et al. 2010). Ensemble forecasts from the analyses will have correlation lengths that increase as errors grow upscale and eventually saturate again as shown in **Fig. 5.10c**.

Assimilating u and v observations is more challenging than θ observations since their correlation functions are nonlocal. The sensitivity experiments mentioned above are repeated using u and v as observations, and the resulting analysis RMSEs are plotted with respect to ROIs in **Figs. 5.9d-f**. **Figure 5.10b** shows the steady-state MACs for each experiment, and the time evolution of a MAC during ensemble forecasts is shown in **Fig. 5.10d**. Similar to the θ observation cases, the asymptotic correlation values are unchanged when u and v observations are assimilated. The overall correlation is much lower compared to θ observations. As forecast errors grow and larger scales contribute more, the nonlocal correlation peak shifts toward longer distances. For larger ROIs, the trend that the best ROI increases as N increases is unchanged. However, an extra penalty is present for smaller ROIs that are too short and exclude the nonlocal correlation peaks that are important sources of information for this case. As a result, as N decreases, the best ROI does not decrease beyond the distance at which these nonlocal correlations peak.

5.4.3 Accuracy and density of the observing network

The sensitivity of localization to changes in the observing network is tested in this section. In CNTL, we considered a uniform observation grid coarser than the model grid ($\Delta x = 3 dx$). When observation density increases (n_o increases), the observation error spectral variance $R(k)$ will decrease according to (5.14). In the ObsSparse experiment, the horizontal observation interval is set to $\Delta x = 9 dx$, which yields an $R(k)$ 9 times larger than that from CNTL. Changes in network density have two different effects on filter performance. First, the observation error variance $R(k)$ is changed (“accuracy effect”). Second, the number of independent pieces of information available within the localization scale to constrain each state variable is also changed (“number effect”). To separate these two effects, another experiment, ObsErrorX3, is conducted in which the observing network is of the same density as CNTL but its observation error standard deviation σ is increased by a factor of 3 resulting in the same $R(k)$ as ObsSparse (i.e. only the accuracy effect is present in ObsErrorX3). In both ObsSparse and ObsErrorX3, the information provided by the observation is less accurate. The other effect of increasing observation accuracy is tested in ObsDense where the horizontal observation interval is set to $\Delta x = 1 dx$, and ObsError/3 where observation density is unchanged but observation error standard deviation σ is decreased by a factor of 3. **Figure 5.11** shows error spectra from these experiments, and **Figs. 5.12a-c** plot their analysis RMSEs filtered for L, M, and S scales with respect to ROIs. For ObsSparse (**Fig. 5.11a**), the ROI = 8 case is excluded because this ROI is smaller than the observation interval which results in only one observation available to constrain each state variable.

With both accuracy and number effects, a sparser observing network favors a larger ROI to achieve best filter performance, and a denser network favors a smaller ROI (compare CNTL, ObsSparse, and ObsDense in **Figs. 5.12a-c**), which is consistent with findings from several previous studies (Dong et al. 2011; Zhu et al. 2013; Perri  nez et al. 2014; Kirchgessner et al. 2014; Snook et al. 2015). It is evident that the observation error variance $R(k)$ influences the lowest analysis error variance $E(k)$ achievable, and the best-performing ROI is larger when a less accurate observing network is used. The accuracy effect alone is responsible for some of the sensitivity in localization at L and M scales as $R(k)$ is reduced. A more accurate observing network with smaller $R(k)$ yields a smaller $E(k)$ that has reduced contribution from large scales, which causes its correlation length to be shorter so that a smaller ROI is favorable. However, for S scale, the accuracy effect does not change the best ROI because prior error is still saturated at this unobserved scale despite a lower $R(k)$. The number effect can be seen by comparing ObsSparse to ObsErrorX3 (also ObsError/3 to ObsDense) in **Figs. 5.12a-c**. The former has lower observation density than the latter while they share the same $R(k)$. As fewer pieces of independent information are available within the localization scale (ROI), the range of ROIs with good performance around the best ROI gets narrower. This implies that the sensitivity to localization increases as fewer independent observations are available. Both accuracy and number effects contribute to the sensitivity of localization to changes in the observing network. The accuracy effect appears to be more important when observations are dense. As observations become sparse, the number effect becomes more important.

Spatially inhomogeneous observing networks are more common than uniform observation grids. The irregularity in the horizontal location of observations precludes the

calculation of observation error spectral variance. We repeated the experiments in this section using observations randomly located in the domain but keeping the number of observations and instrument error unchanged. **Figures 5.12d-f** show results from using these irregular observing networks compared to their regular network counterparts (**Figs. 5.12a-c**). Overall, the behavior of an irregular network is very similar to a regular one with the same accuracy and density. The only difference is that an irregular network yields larger analysis error, especially for sparse networks where there are occasionally fewer observations within the localization scale. For a sparse irregular observing network, the best ROI is larger than the best ROI for a regular network with the same number of observations. These results suggest that an irregular observing network is less efficient in reducing analysis error than a uniform network with the same number of observations.

5.5 Summary

In this study, numerical experiments are conducted to test the scale dependency of localization and its sensitivity to several aspects of an ensemble modeling and data assimilation system. The two-layer quasi-geostrophic (QG) model is employed as the forecast model. It is a simple model but captures the essence of large-scale atmospheric dynamics with a realistic spectral energy distribution. Compared to standard test models such as the Lorenz (1996) system, the QG model has better representation of multiscale dynamics and therefore is better suited for testing when scale is the key concern. The data assimilation method considered here is the square root filter (Whitaker and Hamill, 2002)

with an adaptive covariance relaxation (chapter 4) that ensures filter stability. The Gaspari and Cohn (1999) function with fixed localization radius is applied in the filter, and the localization radius that minimizes analysis error is determined by trial and error. The findings are summarized as follows.

- 1) The best localization radius is scale dependent. For a weather system, the overall best localization radius scales with its dominant correlation length scale. When multiple scales of motion are present simultaneously, the localization that minimizes analysis error at a certain scale also depends on the correlation length for that scale. For example, larger scale favors larger localization radius. A multiscale localization can achieve better filter performance by specifying a scale-dependent localization radius.
- 2) A lower model resolution does not change the best localization radius (defined as a physical distance) for the resolved scales as long as the model representation of the dynamical processes is correct. If the low-resolution model cannot adequately resolve some key dynamical processes, the incorrect model dynamics will give rise to representation errors when assimilating observations. However, if adaptive inflation methods are applied to account for these representation errors, the localization radius is not sensitive to model resolution changes.
- 3) Consistent with previous studies, a decrease in ensemble size is found to cause the best localization radius to shift to smaller values, and the range of localization radii with good performance also becomes narrower. However, this behavior changes when assimilating observations whose correlations with state variables are nonlocal. For correlation functions with peaks at a certain distance (i.e. correlation between u

and ψ), further reducing the localization radius will not remedy sampling error as ensemble size decreases.

- 4) Increasing the density of observing networks without changing the instrument error causes the best localization radius to shift to smaller values. Two effects contribute to this behavior. (i) A denser network yields lower spectral observation error variance for the large scales, which lowers the large-scale prior error. As a result, the overall correlation length scale decreases for the prior ensemble due to less contribution from large-scale errors. (ii) A denser network provides more pieces of independent observation information within a localization radius, and a smaller radius is enough to allow the same number of observations to constrain each state variable. Irregular networks behave similarly to regular ones with the same density in terms of best localization radius, except that a larger radius is favored for irregular networks when observations are too sparse to sample a certain scale.

Table 5.1. Description of model and filter parameters used in control and sensitivity experiments. In each experiment, a range of ROIs is tested and the ROI that minimizes analysis RMSE is determined.

Experiment	Model and filter parameters (changes relative to CNTL)
CNTL	$k_{\max} = 63$, $k_d = 20$, $k_\beta = 4$, $U = 0.2U_0$, $r = 0.5$ (spectral peak at large scale) ensemble size $N = 64$ assimilate θ observations every $3dx$ with $\sigma = 3$ every $t = 0.05$
Scale dependence:	
M_Scale	$U = 0.5U_0$, $r = 6$ (spectral peak at intermediate scale)
S_Scale	$k_d = 35$, $U = 0.6U_0$, $r = 20$ (spectral peak at small scale)
Sensitivity to model resolution:^a	
LowRes	$k_{\max} = 31$ for both truth and forecast model
LowRes_Model	$k_{\max} = 31$ forecast model but truth is from CNTL
LowRes2	$k_{\max} = 21$, $k_d = 14$ for both truth and forecast model
LowRes2_Model	$k_{\max} = 21$, $k_d = 14$ forecast model but truth is from CNTL
Sensitivity to sampling error:^b	
N16, N32, N256, N1024	ensemble size N changed to 16, 32, 256, and 1024
Sensitivity to observing network:^c	
ObsSparse, ObsDense	horizontal observation interval changed to $9dx$ and $1dx$, respectively
ObsErrorX3, ObsError/3	observation error std changed to $\sigma = 9$ and $\sigma = 1$, respectively ^d

^a The ROI is in physical distance units (number of grid points from CNTL model), e.g. ROI = 8 means 8 grid points in CNTL, which corresponds to 4 grid points in LowRes, and only 2.67 grid points in LowRes2.

^b The ensemble size experiments are also repeated for the case where u and v are assimilated instead of θ .

^c The observing network experiments are first performed with regular (horizontally uniform) networks, then repeated with irregular networks (randomly located) that match the density and accuracy of the regular ones.

^d The observation error level is the same for ObsErrorX3 and ObsSparse, they both have 3 times larger RMSE than CNTL.

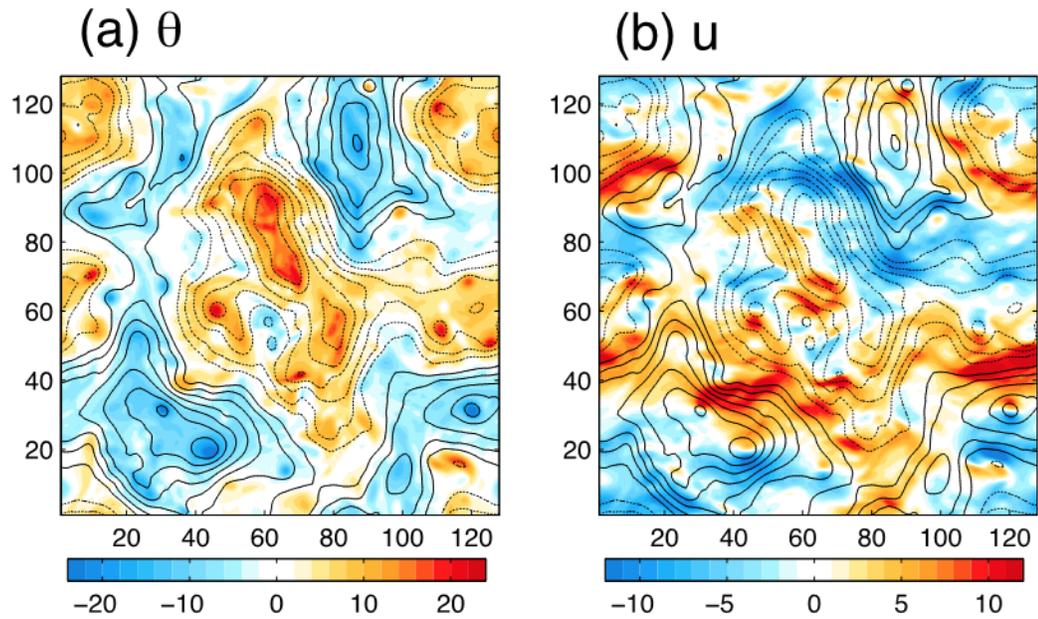


Figure 5.1. Snapshots of (a) θ and (b) u in color shadings and state variable ψ in black contours (solid lines for positive and dotted lines for negative values). The ψ , θ , and u variables are nondimensionalized. The scaling of the QG model is described in section 5.2.1.

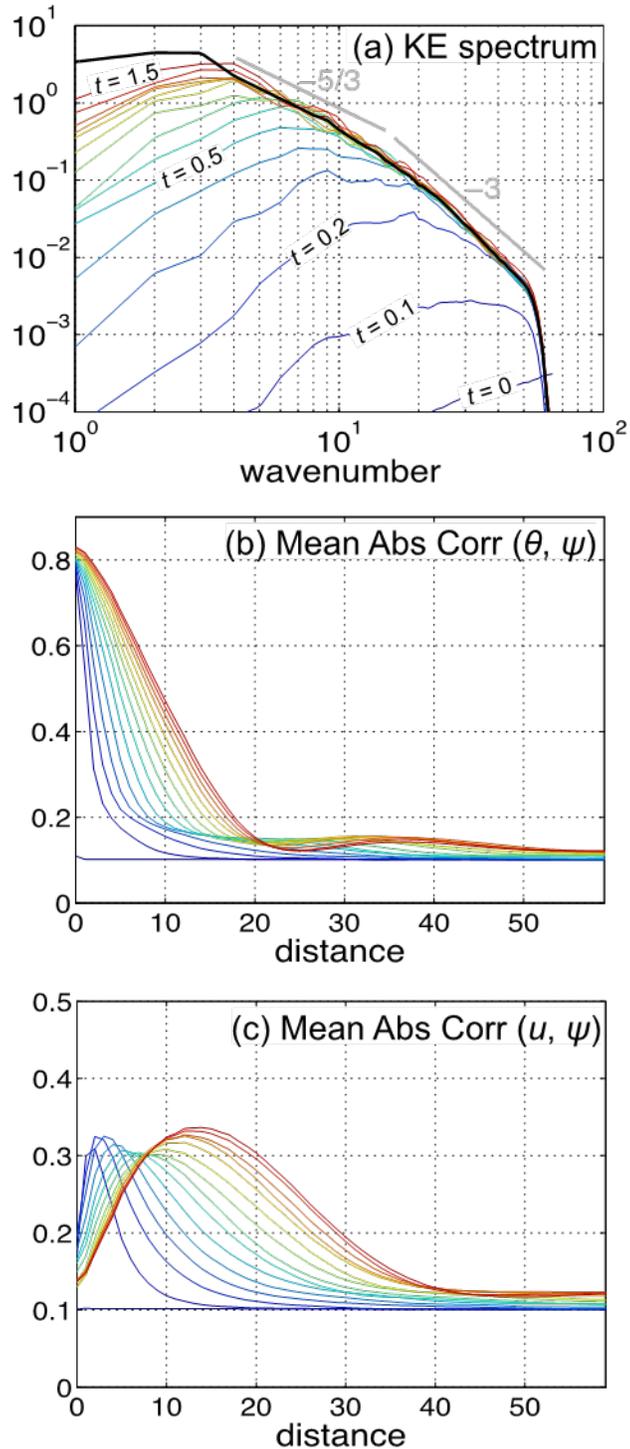


Figure 5.2. (a) Spectra of error kinetic energy from ensemble mean (color-coded with forecast time) and the time-averaged reference kinetic energy from the truth (black) during an ensemble spin-up period of 1.5 time units. (b) Spatially averaged mean absolute correlation (MAC) between θ and ψ plotted as a function of horizontal distance (number of grid points) also color-coded with forecast time as in (a). (c) Same as (b) but showing correlation between u and ψ .

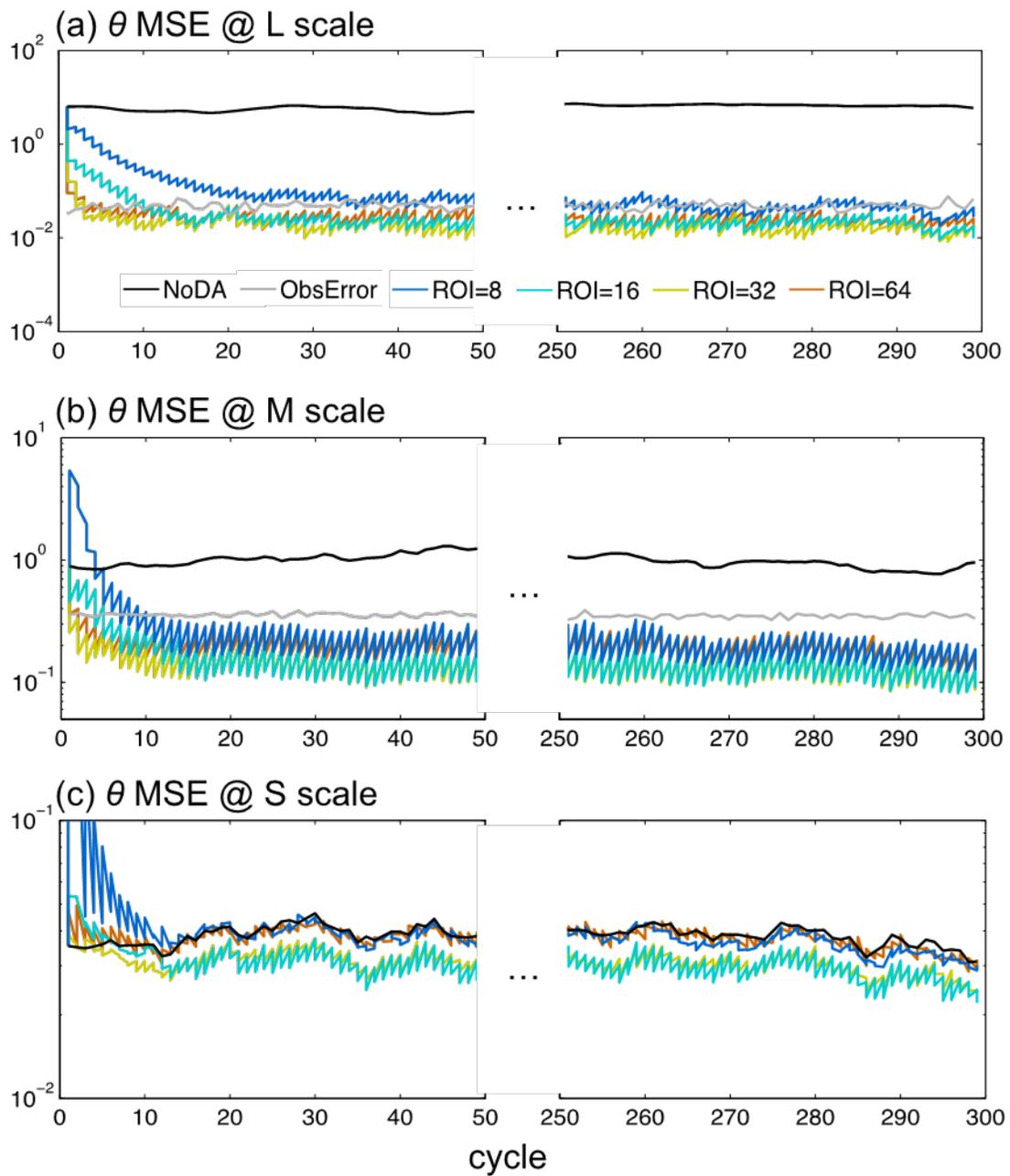


Figure 5.3. Time series of θ MSE from CNTL filtered for the (a) L, (b) M and (c) S scales. Results are shown for errors from free ensemble (NoDA; black), observation (gray), and from analysis mean using ROI = 8, 16, 32, and 64 (colored).

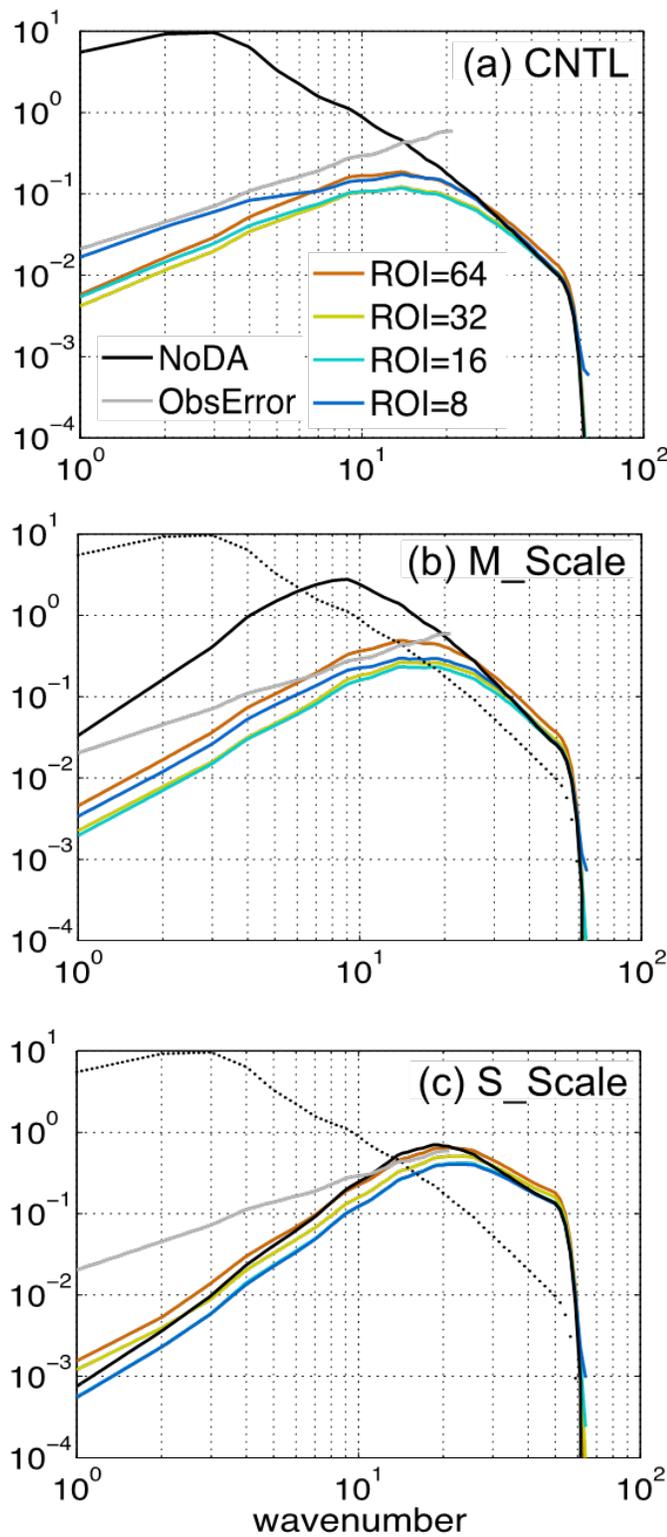


Figure 5.4. θ error spectra from (a) CNTL, (b) M_Scale, and (c) S_Scale. Results are shown for errors from a free ensemble (NoDA; black), observation (gray), and from analysis mean using ROI = 8, 16, 32, and 64 (colored). The CNTL NoDA error is shown as dotted lines in (b) and (c) for reference.

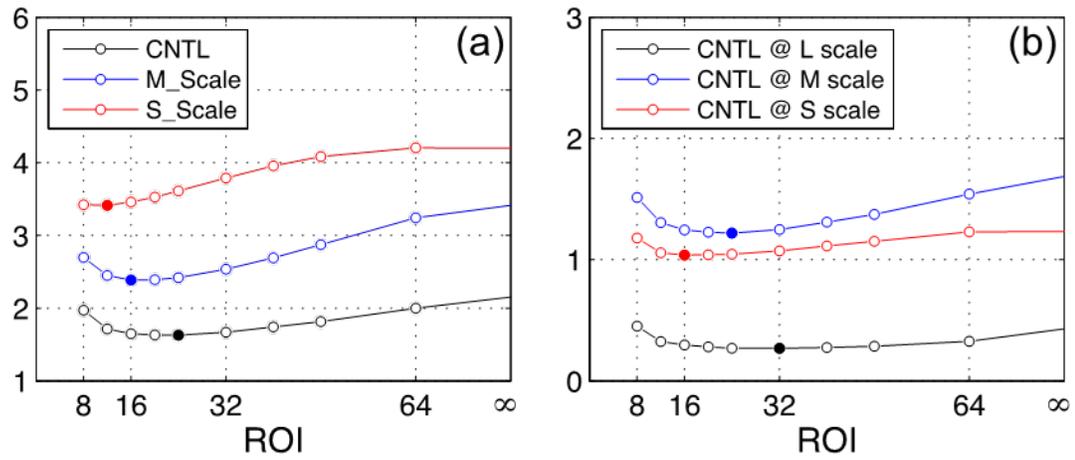


Figure 5.5. (a) θ RMSE plotted with respect to ROI from CNTL (black), M_Scale (blue) and S_Scale (red); (b) Same as (a) but showing θ RMSE from CNTL filtered for the L (black), M (blue) and S (red) scales. The filled circles indicate the ROI with lowest analysis error.

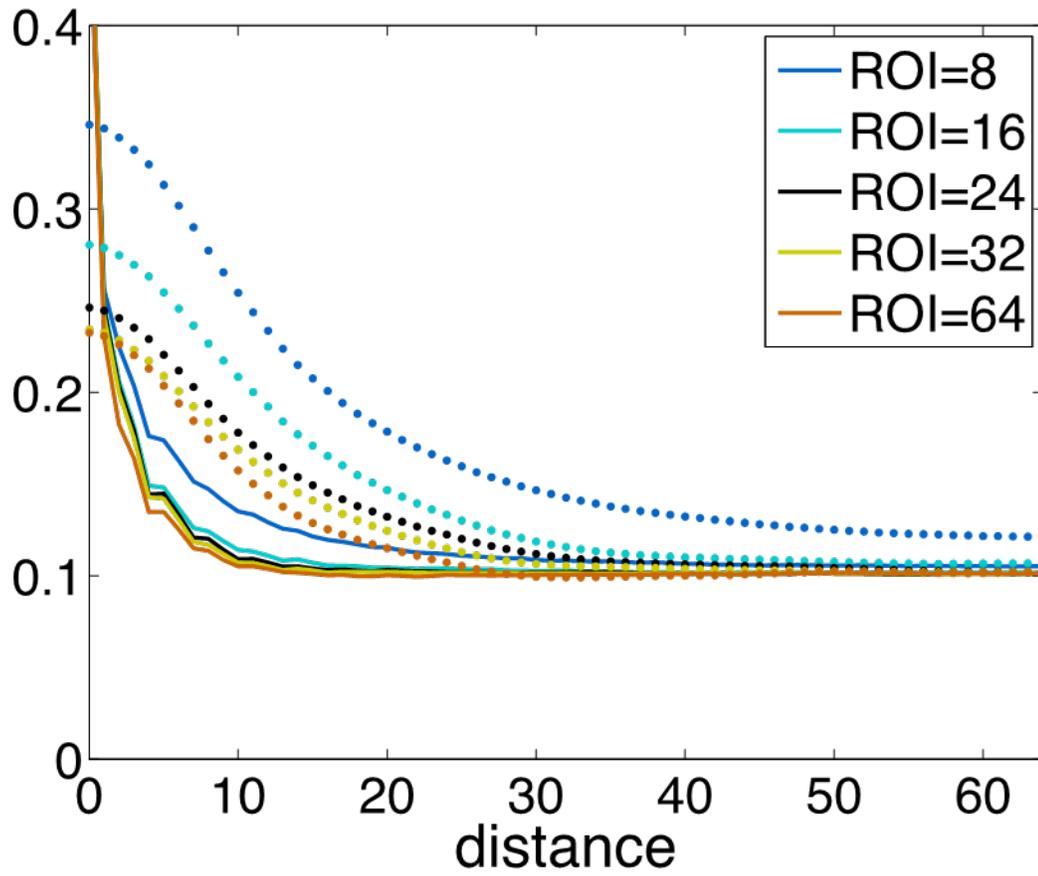


Figure 5.6. MAC between θ and ψ filtered for L (dotted) and S (solid) scales estimated by the prior CNTL ensemble using ROI = 8, 16, 24, 32, and 64. The MACs are averaged over space and time and plotted as a function of horizontal distance (number of grid points).

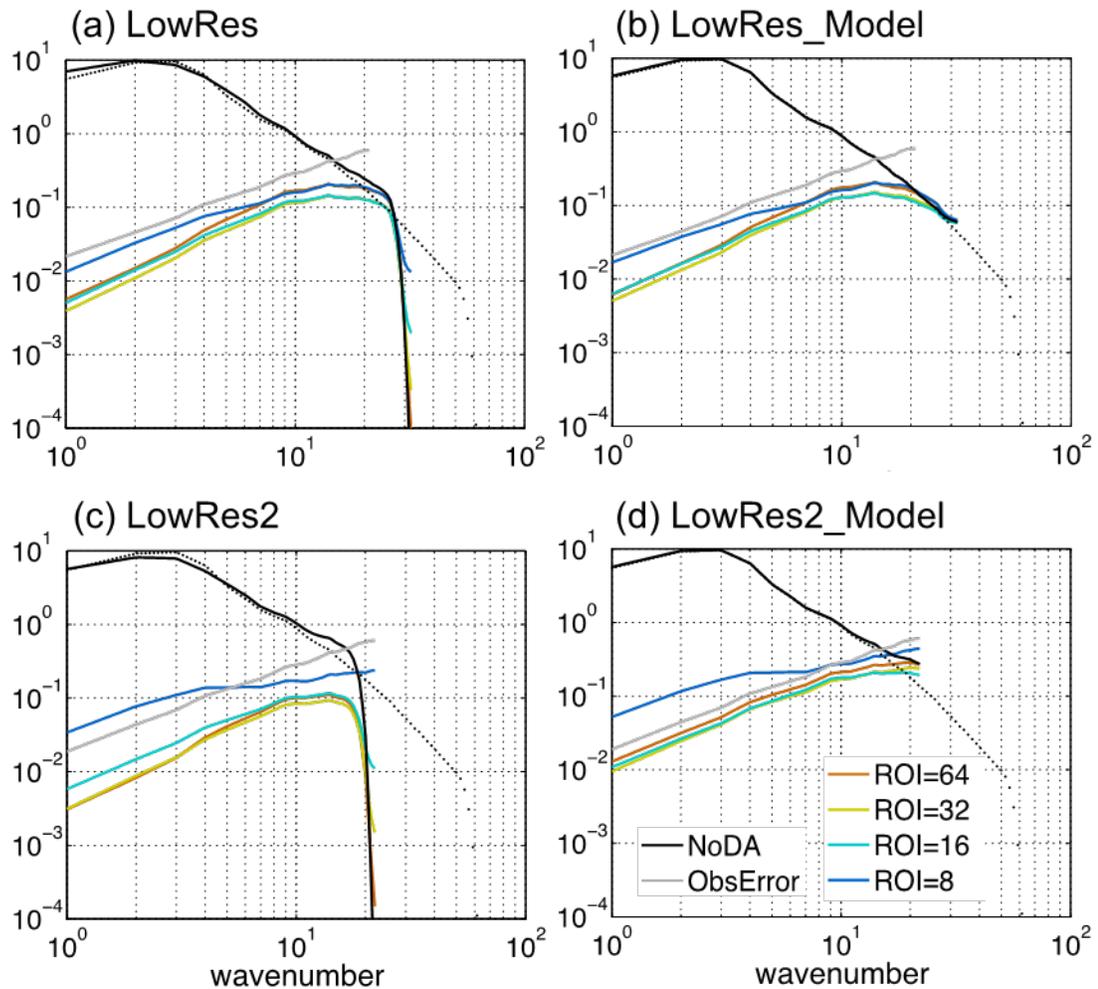


Figure 5.7. θ error spectrum of NoDA (black), observation (gray), and analysis ensemble mean using ROI = 8, 16, 32, and 64 (colored) from (a) LowRes, (b) LowRes_Model, (c) LowRes2, and (d) LowRes2_Model experiments. The CNTL NoDA error spectrum is shown as dotted lines for reference. The ROI is in physical distance units (number of grid points from CNTL model), e.g. ROI = 8 means 8 grid points in CNTL, which corresponds to 4 grid points in LowRes, and 2.67 grid points in LowRes2.

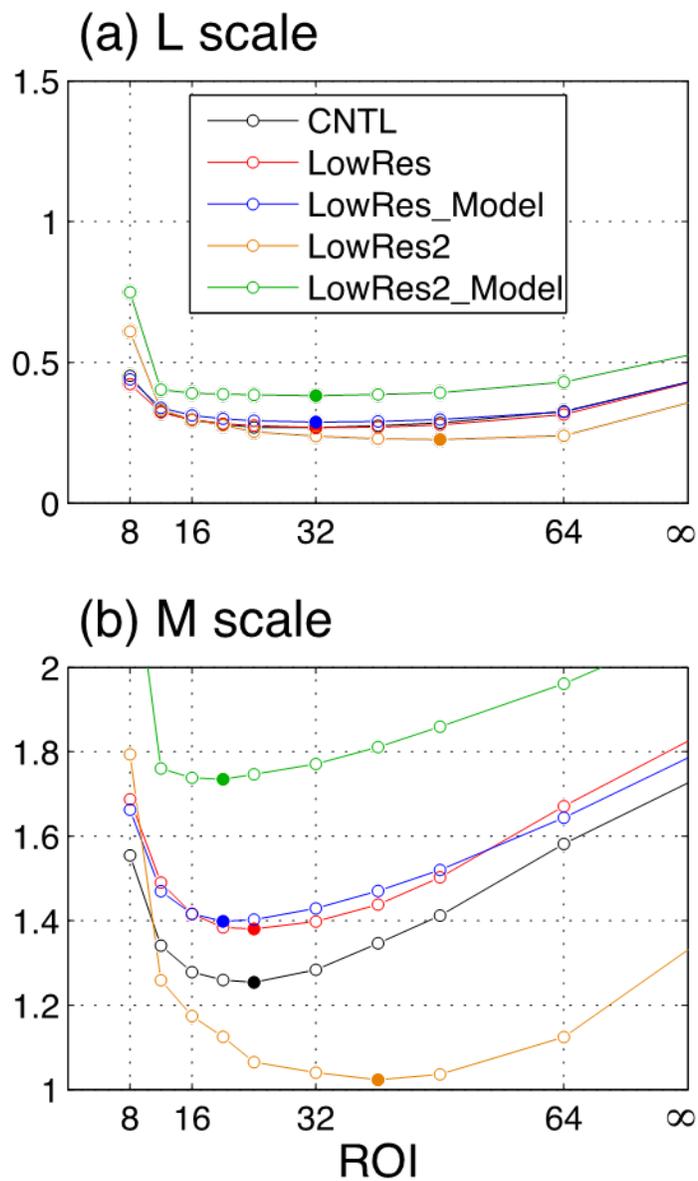


Figure 5.8. θ RMSE filtered for (a) L and (b) M scales plotted with respect to ROI.

Results are compared for CNTL, LowRes, LowRes_Model, LowRes2 and LowRes2_Model.

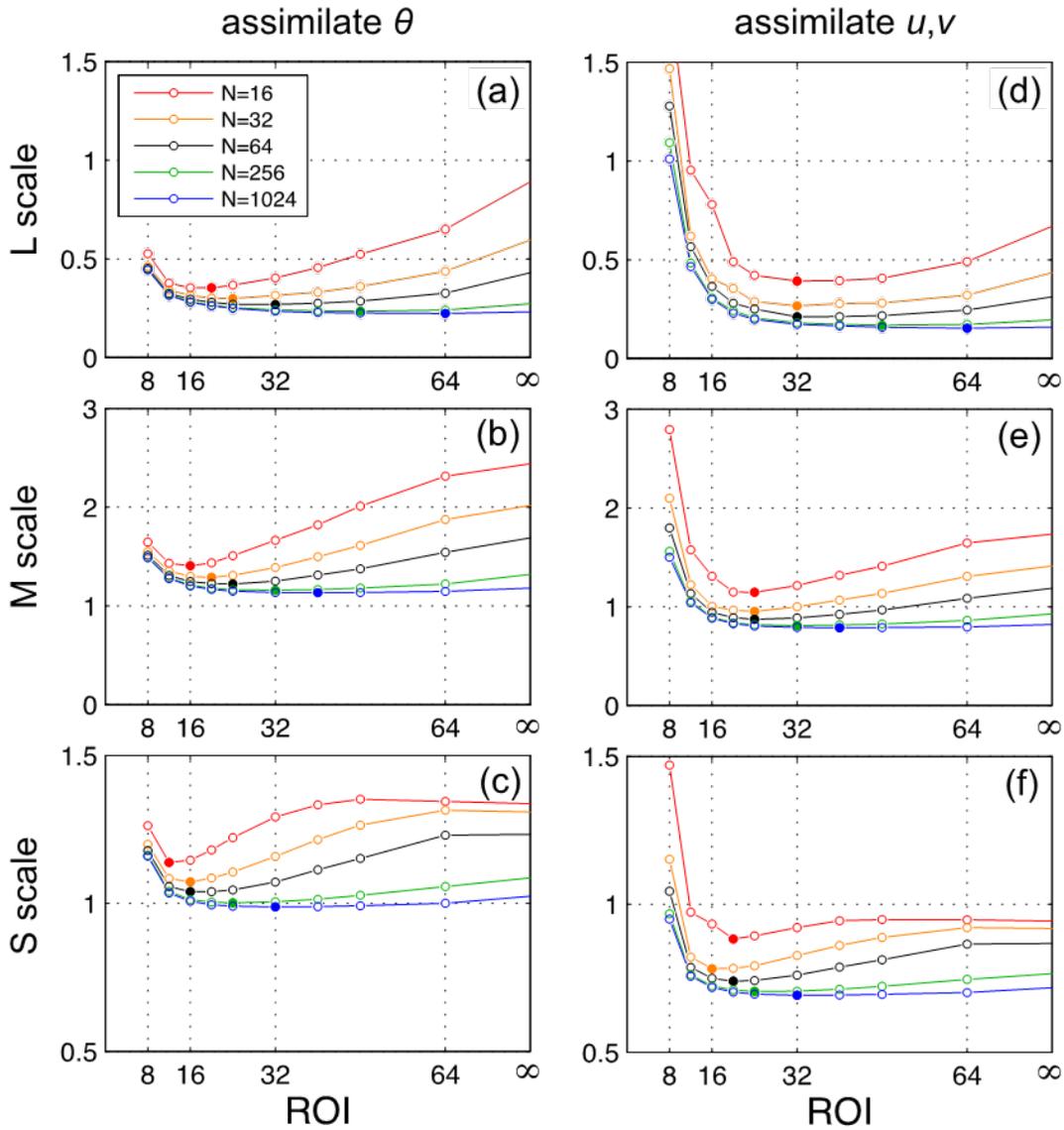


Figure 5.9. θ RMSE filtered for (a) L, (b) M and (c) S scales plotted with respect to ROI compared for the $N=16$, $N=32$, $N=64$ (CNTL), $N=256$ and $N=1024$ cases. Filled circles indicate the minimum error corresponding to the best ROI. (d)-(f) Same as (a)-(c) but showing root mean error kinetic energy from the cases assimilating u and v .

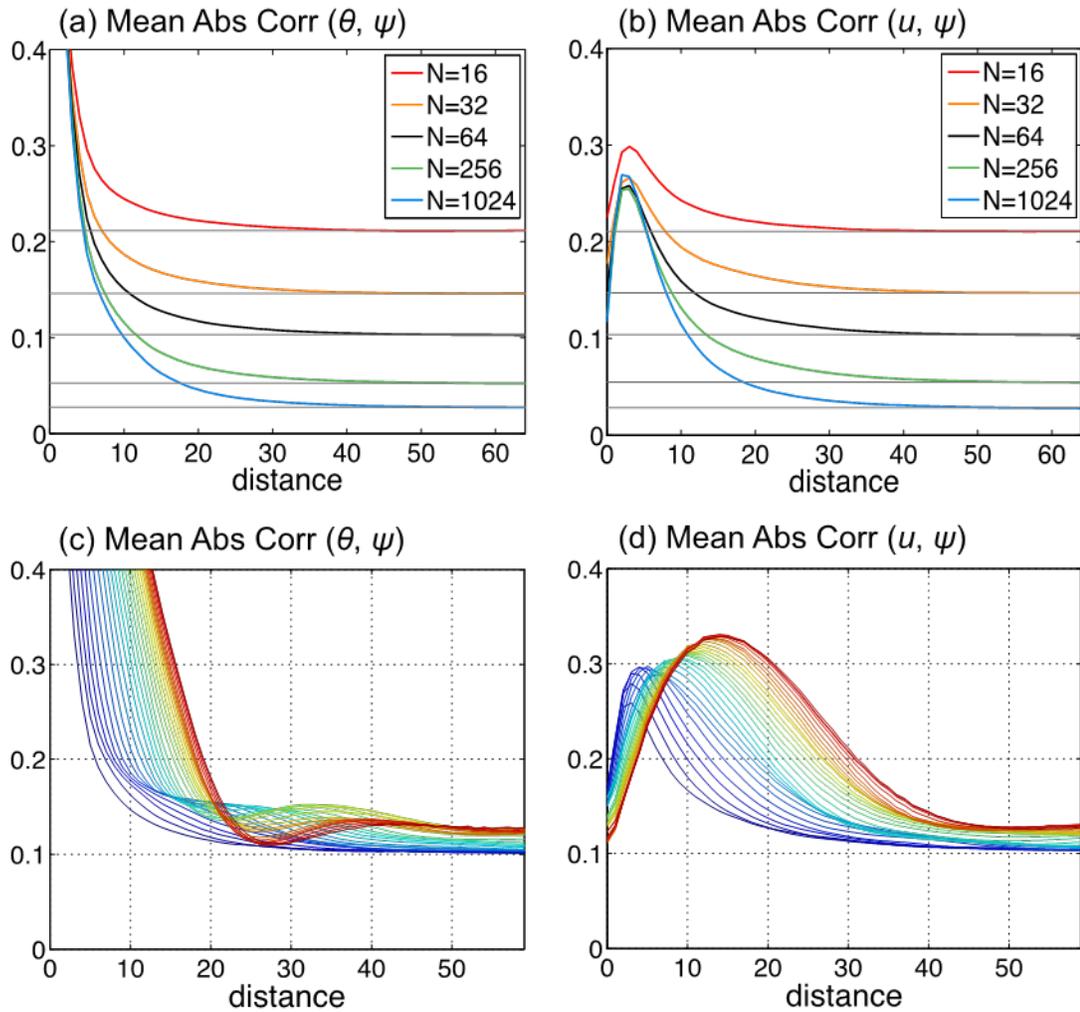


Figure 5.10. (a) MAC between θ and ψ from the posterior ensemble with changing ensemble size $N = 16, 32, 64, 256, 1024$ using the best ROI. (b) Same as (a) but showing MAC between u and ψ from the cases assimilating u and v . (c) Time evolution of MAC between θ and ψ from ensemble forecasts initialized with posterior from the $N = 64$ case, the lines are color-coded with forecast time $t = 0$ to $t = 1.5$ every 0.05 time units. (d) same as (c) but showing MAC between u and ψ from the cases assimilating u and v .

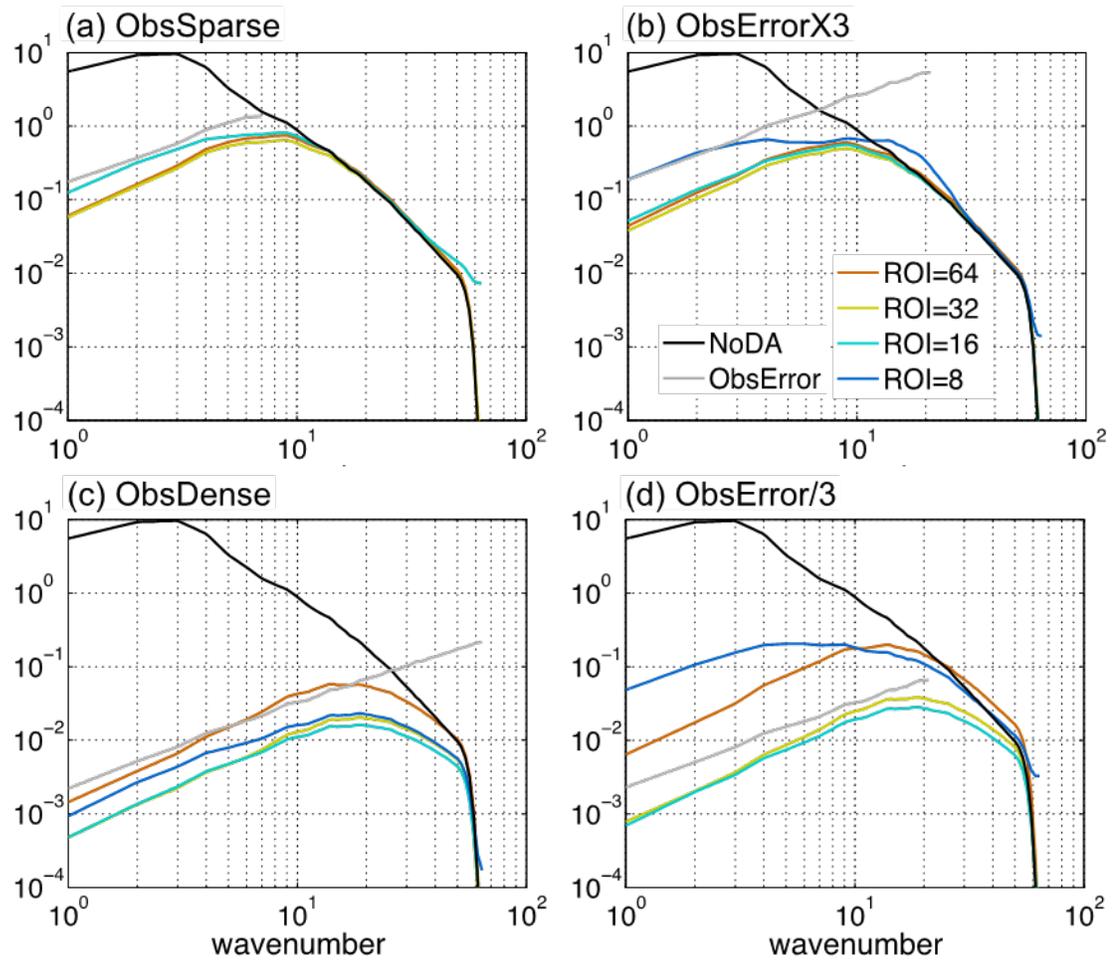


Figure 5.11. Error spectra of free ensemble mean (NoDA; black), observation (gray), and analysis mean using ROI = 8, 16, 32, and 64 (colored) from (a) ObsSparse, (b) ObsErrorX3, (c) ObsDense, and (d) ObsError/3 cases.

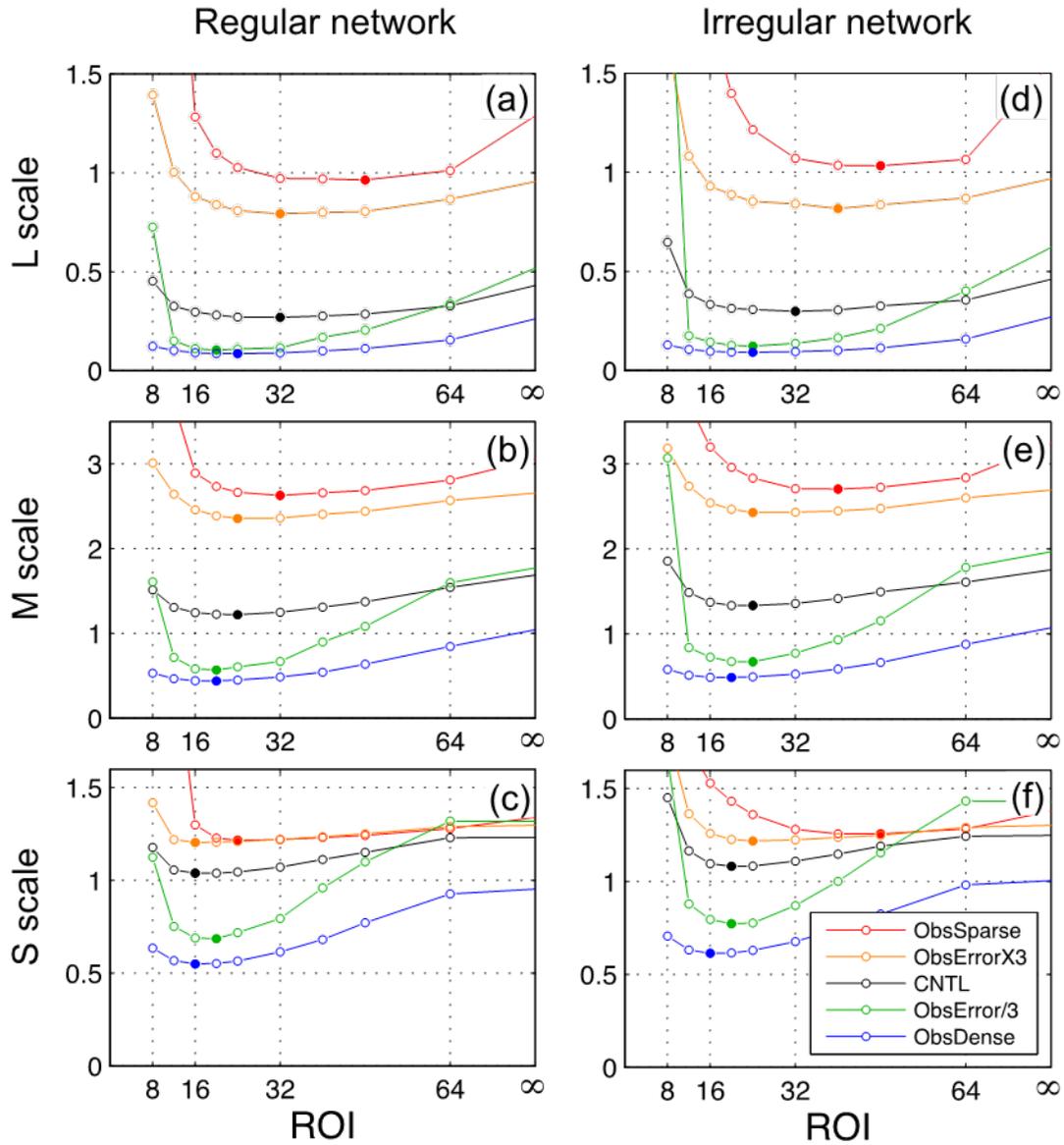


Figure 5.12. θ RMSE filtered for (a) L, (b) M and (c) S scales plotted with respect to ROI for the cases with changing observing networks: ObsSparse, ObsDense, ObsErrorX3 and ObsError/3. Filled circle markers indicate the minimum error corresponding to the best ROI. (d)-(f) Same as (a)-(c) but showing results from using irregular observing networks.

Chapter 6

Concluding Remarks

6.1 Predictability of multiscale tropical weather systems

This dissertation seeks to identify the practical and intrinsic predictability limits for multiscale tropical atmosphere, and evaluate the potentials in assimilating current and future satellite observations to improve the practical predictability of tropical weather at different scales. Using a regional convection-permitting model, Weather Research and Forecasting (WRF), predictability limits are estimated for each component of the convectively coupled equatorial waves (CCEW) through systematic ensemble sensitivity experiments with perturbed initial and boundary conditions. At each spatial scale, the divergence rate of ensemble members tells us how fast errors grow; the faster error growth the more limited predictability. Predictability limit is reached when the ensemble spread is indistinguishable from random climatological draws. With realistic errors sampled from the operational global model forecasts, we estimated that the current practical predictability limit is about 8 days for waves at large scales (>2000 km). For smaller-scale (<200 km) waves, the predictability limit decreases to less than 1 day. The intrinsic limit of predictability, an upper bound of the expected prediction skills, was also estimated by running the ensemble experiments with perturbation variances reduced to only 1% the realistic level. The intrinsic predictability limit is found achievable beyond 2 weeks for planetary-scale weather systems, while the smaller scale phenomena still have

an intrinsic limit of less than 3 days. There is a sharp transition from slow to fast error growth at the intermediate scales (~500 km), separating the more predictable large-scale wave components from the less predictable small-scale weather systems.

The current findings encourage the future development in data assimilation and modeling system to further improve the predictability of CCEWs. However, further experiments are required to evaluate the impact to tropical weather predictability from several error sources not accounted for in the current idealized experiment. The convection-permitting WRF model, although providing a reasonable representation of CCEWs, is still not perfect. In this study, only the uncertainties from model IC and LBC are accounted for when estimating practical predictability. While the IC and LBC uncertainties are sampled from realistic forecast errors, the estimated practical predictability may potentially differ from those estimated from other modeling systems that account for additional error sources (e.g., model dynamics and physics parameterizations, low boundary condition forcings, etc.). The predictability estimates from the MJO active phase event in this study may also differ from those estimated for other events.

The current study only simulates a 15-day period within an MJO active phase, which is not long enough to estimate the predictability of MJO itself. In previous MJO predictability studies using global model simulations, the practical predictability estimates range from 15 to 45 days depending on the models and diagnostics used (Gottschalck et al. 2010; Vitart and Molteni 2010; Neena et al. 2014b; Hamill and Kiladis 2014). The MJO predictability metric is usually based on the Real-time Multivariate MJO (RMM) index (Wheeler and Hendon 2004), which captures more the large-scale features

of the MJO signal. Ling et al. (2014) suggested that, when finer-scale local features are included, the estimated predictability becomes more limited than using a global measure. The predictability of MJO is also found to be dependent on its phase (Waliser et al. 2003; Nasuno et al. 2013; Neena et al. 2014a, b). While active phase of the MJO has better predictability, the models have difficulty in accurately predicting the timing of convection onset during its suppressed phase. As an extension of these global model MJO predictability studies, the current study provides a comprehensive predictability estimate for the multiscale CCEWs during an MJO active phase. How the predictability estimates change for a suppressed MJO phase may be a future research topic.

The selection of MJO diagnostics is also nontrivial in predictability studies. Waliser et al. (2009) proposed several candidate MJO diagnostics that can capture more small-scale details of MJO than traditional ones. Results in this paper show that predictability is variable-dependent. Therefore, process-oriented diagnostics (Kim et al. 2014) that targets the tropospheric moisture or even precipitation may result in very different predictability estimates compared to just using dynamic variables. A longer regional simulation with convection-permitting resolution may be the next step to facilitate a comprehensive predictability study that compares different diagnostics.

6.2 Potentials in assimilating satellite observations

The potentials in improving the analysis and prediction of multiscale tropical weather systems are further evaluated in this dissertation through assimilating synthetic satellite-based observations using the Penn State WRF ensemble Kalman filter (EnKF)

data assimilation system. The observing networks investigated include the retrieved temperature and humidity profiles from the Advanced TIROS Operational Vertical Sounder (ATOVS) and Global Positioning System Radio Occultation (GPSRO), the atmospheric motion vectors (AMVs), infrared brightness temperature from Meteosat-7 (Met7-Tb), and retrieved surface wind speed from the Cyclone Global Navigation Satellite System (CYGNSS). It is found that assimilating synthetic ATOVS thermodynamic profiles and AMV winds improves the accuracy of wind, temperature, humidity and hydrometeors for the scales larger than 200 km. The skillful forecast lead times can be extended by as much as 4 days for scales larger than 1000 km. Assimilation of synthetic Met7-Tb further improves the analysis of cloud hydrometeors even at scales smaller than 200 km. Assimilating synthetic CYGNSS surface winds further improves the low-level wind and temperature. Meanwhile, the impact from assimilating the current-generation synthetic GPSRO data with better vertical resolution and accuracy is comparable to assimilating half of the current ATOVS profiles, while a hypothetical 25-time increase in the number of GPSRO profiles can potentially exceed the impact from assimilating the current network of retrieved ATOVS profiles. These results not only show great promises in further improving practical predictability of multiscale tropical weather systems but also shed lights on the need, design and cost-benefit analysis of current and future space-borne observing systems for better tropical weather prediction.

The idealized OSSE provided a proof of concept of using satellite observations to improve tropical weather predictability. However, future real-data application will face more challenges. There may be non-trivial uncertainties in model dynamics, physics schemes, and forward operator not accounted by the current study. The actual retrieval

algorithms are not used in this study to generate ATOVS and GPSRO profiles, thus the simulated errors for these synthetic profiles do not account for complicated error sources from the retrieval processes. Thus, the idealized experiment results should be interpreted with these caveats in mind. The assimilation of these satellite observations in real-data cases needs to be tested in future studies. In this study, a limited-area domain is used and a perfect LBC is specified, and uncertainties from LBC and forecast model itself are not accounted for. Therefore, the NoDA experiment is not intended to account for all realistic sources of forecast errors that can impact practical predictability. Since we only focused on the impact of assimilating data in a limited-area domain for a selected period, one should be cautious in generalizing our findings. Although the 9-km horizontal grid spacing used in is comparable to the current operational global models, it will be interesting to further test the sensitivity of EnKF performance in satellite data assimilation to the horizontal resolution, especially for smaller-scale tropical weather phenomena and CCEWs. In this dissertation, we highlighted the different performance from ATOVS and GPSRO profiles due to their spatial coverage and resolution. A similar OSSE can be useful for assessing the cost-effectiveness of other future space-based sensors. We also believe that a more systematic direct assimilation of satellite Tb is needed for substituting the assimilation of retrieved profiles. The synthetic observations that are available evenly in space and time allow better quantification of multiscale analysis errors. One caveat of this assumption is that results are rather optimistic in terms of the information content and availability of the current satellite observing networks. On the other hand, although methods like the AOEI make the direct assimilation of Tb

possible, we believe the filter performance is still suboptimal and a better methodology in dealing with nonlinearity in forward operator is still much desired.

6.3 Adaptive multiscale data assimilation methods

Ensemble filters require covariance inflation to account for unrepresented errors, and covariance localization to remove sampling errors due to limited ensemble size. In practice, data assimilation methods are tuned to achieve best performance for the application to a certain weather system and using a certain type of observations. These tuning efforts are laborious and thus more adaptive methods are desired. The satellite observation assimilation experiment exposed the disadvantage of using a data assimilation method that cannot handle spatially inhomogeneous observations and the simultaneous presence of weather systems at multiple scales. For example, the tuned inflation factor may work best for one type of observations, but not the others; and the tuned localization radius may work best for a weather component at certain scale, but not for other scales. This motivates the development of data assimilation methods that can adaptive to variations in observing networks and to weather systems at different scales (i.e. multiscale data assimilation), which will perform better in the tropical weather prediction scenario.

An Adaptive Covariance Relaxation (ACR) method was introduced and tested in the simplified Lorenz (1996) 40-variable system. The new method is compared to the Adaptive Covariance Inflation (ACI) method (Anderson 2009) that estimates a spatially and temporally varying inflation factor from innovation statistics using a hierarchical

Bayesian approach. The ACR method estimates a covariance relaxation parameter from the same innovation statistics. Instead of treating inflation parameters as spatially varying random variables as in ACI, the relaxation-to-prior-spread method provides an ensemble spread reduction term that serves as a spatial mask to account for an irregular observation network. We demonstrated that the ACR method is able to improve filter performance with the presence of sampling/model errors over a range of severity. Its reliability and ease of implementation suggest potential for future applications beyond the tropical weather systems.

For covariance localization, previous studies suggest that the optimal localization radius depends on ensemble size, observation density and accuracy, as well as the correlation length scale determined by model dynamics. A comprehensive localization theory for multiscale dynamical systems with varying observation density remains an active area of research. Using a two-layer quasi-geostrophic (QG) model, this study systematically evaluates the sensitivity of the best Gaspari-Cohn localization radius to changes in model resolution, ensemble size and observing networks. Numerical experiment results show that the best localization radius is smaller for smaller-scale components of a QG flow, indicating its scale dependency. The best localization radius is rather insensitive to changes in model resolution, as long as the key dynamical processes are reasonably well represented by the low-resolution model with inflation methods that account for representation errors. As ensemble size decreases, the best localization radius shifts to smaller values. However, for nonlocal correlations between an observation and state variables that peak at a certain distance, decreasing localization radii further within this distance does not reduce analysis errors. Increasing the density of an observing

network has two effects that both reduce the best localization radius. First, the reduced observation error spectral variance further constrains prior ensembles at large scales. Less large-scale contribution results in a shorter overall correlation length, which favors a smaller localization radius. Second, a denser network provides more independent pieces of information, thus a smaller localization radius still allows the same number of observations to constrain each state variable.

The particular type of localization used in this study is observation-space localization (Houtekamer and Mitchell 2001; Hamill et al. 2001), which tapers the sample covariance between an observation and state variables and between an observation and other observation priors. Model-space localization (Houtekamer and Mitchell 1998) directly tapers the background error covariance before the observation operator is applied. Although the similarity of the two approaches is proven (Sakov and Bertino 2011; Nerger et al. 2012), these approaches can have different behavior under certain conditions such as small ensemble size and short localization distance (Lei and Whitaker 2015). For example, when assimilating satellite observation impact in the vertical, Campbell et al. (2010) suggested that model-space localization is superior. However, Lei and Whitaker (2015) showed that the opposite can be true for some cases. In order to develop a more robust localization theory, both types of localization need to be further investigated, especially for cases where nonlocal observation operators involve spatial averaging. Correlated observation errors may also change the behavior of localization. The impact from observation operators and error model can also be a future research topic.

Although the flow simulated by the QG model possesses many scales, its dynamical process (baroclinic instability) is relatively simple. In real atmospheric models, there could be multiple sources of instability at different scales. For example, the addition of moist convective instability may change the error growth rate at small scales, and the scale interaction may be more complicated. The current study motivates further exploration of the behavior of localization in different contexts in pursuit of a more generalized theory and development of adaptive multiscale localization methods.

Bibliography

- Aksoy, A., S. Lorusso, T. Vukicevic, K.J. Sellwood, S.D. Aberson, and F. Zhang, 2012: The HWRF Hurricane Ensemble Data Assimilation System (HEDAS) for high-resolution data: The impact of airborne Doppler radar observations in an OSSE. *Mon. Wea. Rev.*, **140**, 1843–1862.
- Anderson, J. L., 2007. An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus* **59A**: 210-224.
- , 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99-111.
- , 2009. Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus* **61A**: 72-83.
- , 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **140**, 2359-2371.
- , 2016: Reducing correlation sampling error in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **144**, 913-925.
- , Anderson SL. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Weather Rev.* **127**: 2741-2758.
- , and L. Lei, 2013: Empirical localization of observation impact in ensemble Kalman filters. *Mon. Wea. Rev.*, **141**, 4140-4153.

- Bei, N., and F. Zhang, 2007: Impacts of initial condition errors on mesoscale predictability of heavy precipitation along the Mei-Yu front of China. *Q.J.R. Meteorol. Soc.*, **133**, 83–99.
- , and ———, 2014: Mesoscale predictability of moist baroclinic waves: Variable and scale-dependent error growth. *Adv. Atmos. Sci.* **31**, 995-1008.
- Berner J, Shutts GJ, Leutbecher M, Palmer TN. 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.* **66**: 603-626.
- 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.
- , and ———, 2009: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus A*, **61**, 84-96.
- Bocquet M. 2011. Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlin. Processes Geophys.* **18**: 735-750.
- , and Sakov P. 2012. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems. *Nonlin. Processes Geophys.* **19**: 383-399.
- , L. Wu, and F. Chevallier, 2011: Bayesian design of control space for optimal assimilation of observations. I: Consistent multiscale formalism. *Quart. J. Roy. Meteor. Soc.*, **137**, 1340-1356.
- Buehner, M., and A. Shlyayeva, 2015: Scale-dependent background-error covariance localization. *Tellus A*, **67**, 28027.

- Buizza R, Miller M, Palmer TN. 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.* **125**: 2887-2908.
- 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.
- Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585.
- Chou, M. D., and M. J. Suarez, 1999: A solar radiation parameterization for atmospheric studies. *Technical Report Series on Global Modeling and Data Assimilation*, NASA/TM-19990104606, **15**.
- Collins, W. D., and Coauthors, 2004: Description of the NCAR Community Atmosphere Model (CAM 3.0), *NCAR Technical Note*, NCAR/TN-464+STR, 226pp.
- Cook, K. L. B., P. Wilczynski, C. J. Fong, N. L. Yen and G. S. Chang, 2011: The Constellation Observing System for Meteorology Ionosphere and Climate follow-on mission. *Aerospace Conference, 2011 IEEE*, 1-7.
- Dee, D.P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Weather Rev.* **123**: 1128-1145.
- , and A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Weather Rev.* **127**:1822-1834.
- , G. Gaspari, C. Redder, L. Rukhovets, A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part II: Applications. *Mon. Weather Rev.* **127**:1835-1849.

- , and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteor. Soc.*, **137**, 553–597.
- Desroziers, G., L. Berre, B. Chapnick, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.* **131**: 3385-3396.
- Dickinson, M., and J. Molinari, 2002: Mixed Rossby-gravity waves and western Pacific tropical cyclogenesis. Part I: Synoptic evolution. *J. Atmos. Sci.*, **59**, 2183-2196.
- Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteorol. Atmos. Phys.*, **112**, 41-61.
- Dunkerton, T. J., and F. X. Crum, 1995: Eastward propagating ~2- to 15-day equatorial convection and its relation to the tropical intraseasonal oscillation, *J. Geophys. Res.*, **100**, 25781-25790.
- English, S. J., Renshaw, R. J., Dibben, P. C., Smith, A. J., Rayer, P. J., Poulsen, C., Saunders, F. W. and Eyre, J. R., 2000: A comparison of the impact of TOVS arid ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Q.J.R. Meteorol. Soc.*, **126**, 2911–2931.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162.
- Flowerdew, J., 2015: Towards a theory of optimal localization. *Tellus A*, **67**, 25257.
- Gaspari, G., and S. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723-757.

- Geer, A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037.
- Gottschalck, J., and Coauthors, 2010: A framework for assessing operational Madden–Julian oscillation forecasts: A CLIVAR MJO Working Group Project. *Bull. Amer. Meteor. Soc.*, **91**, 1247–1258.
- Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522.
- Haertel, P. T., and G. N. Kiladis, 2004: Dynamics of two day equatorial waves. *J. Atmos. Sci.*, **61**, 2707–2721.
- Hamill, T. M., and G. N. Kiladis, 2014: Skill of the MJO and northern hemisphere blocking in GEFS medium-range reforecasts. *Mon. Wea. Rev.*, **142**, 868–885.
- Held, I. M., and V. D. Larichev, 1996: A scaling theory for horizontally homogeneous, baroclinically unstable flow on a beta plane. *J. Atmos. Sci.*, **53**(7), 946–952.
- Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341.
- Hou, A. Y., R.K. Kakar, S. Neeck, A.A. Azarbarzin, C.D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi, 2014: The global precipitation measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.*, **113**, D13103.

- Johnson, R. H., and P. E. Ciesielski, 2013: Structure and properties of Madden-Julian oscillations deduced from DYNAMO sounding arrays. *J. Atmos. Sci.*, **70**, 3157-3179.
- Judt, F., S. S. Chen, and J. Berner, 2016: Predictability of tropical cyclone intensity: scale-dependent forecast error growth in high-resolution stochastic kinetic-energy backscatter ensembles. *Quart. J. Roy. Meteor. Soc.*, **142**, 43-57.
- Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003.
- Kim, D., P. Xavier, E. Maloney, M. Wheeler, D. Waliser, K. Sperber, H. Hendon, C. Zhang, R. Neale, Y. Hwang, and H. Liu, 2014: Process-Oriented MJO Simulation Diagnostic: Moisture Sensitivity of Simulated Convection. *J. Climate*, **27**, 5379–5395
- Kirchgessner, P., L. Nerger, and A. Bunse-Gerstner, 2014: On the choice of an optimal localization radius in ensemble Kalman filter methods. *Mon. Wea. Rev.*, **142**, 2165-2175.
- Kursinski, E. R., Hajj, G., Schofield, J., Linfield, R., and Hardy, K. 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, *J. Geophys. Res.*, **102**, 23429–23465.
- Hamil, T. M., J. S. Whitaker and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776-2790.
- Harlim, J., and A. J. Majda, 2010: Filtering turbulent sparsely observed geophysical flows. *Mon. Wea. Rev.*, **138**, 1050-1083.

- Hodyss, D., and N. Nichols, 2015: The error of representation: basic understanding. *Tellus A*, **67**, 24822.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.
- , and ——, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.
- , X. Deng, H. L. Mitchell, S.-J. Baek, and N. Gagnon, 2014: Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 1143–1162.
- , and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489-4532.
- Janjić, T., S. E. Cohn, 2006: Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134**, 2900-2915.
- Keper, J. D., 2009: Covariance localisation and balance in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 1157–1176.
- Kirchgeßner, P., L. Nerger, and A. Bunse-Gerstner, 2014: On the choice of an optimal localization radius in ensemble Kalman filter methods. *Mon. Wea. Rev.*, **142**, 2165-2175.
- Kleist, D. T., and K. Ide, 2015a: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Mon. Wea. Rev.*, **143**, 433–451.
- Kondo, K., and T. Miyoshi, 2016: Impact of removing covariance localization in an ensemble Kalman filter: Experiments with 10,240 members using an intermediate AGCM. *Mon. Wea. Rev.*, **144**, 4849-4865.

- Lange, H., and G. C. Craig, 2014: The impact of data assimilation length scales on analysis and prediction of convective storms. *Mon. Wea. Rev.*, **142**, 3781–3808.
- Larichev, V. D., and I. M. Held, 1995: Eddy amplitudes and fluxes in a homogeneous model of fully developed baroclinic instability. *J. Phys. Ocean.*, **25**, 2285–2297.
- van Leeuwen, P. J., 2015: Representation errors and retrievals in linear and nonlinear data assimilation. *Quart. J. Roy. Meteor. Soc.*, **141**, 1612–1623.
- Lei, L. and J.S. Whitaker, 2015: Model Space Localization Is Not Always Better Than Observation Space Localization for Assimilation of Satellite Radiances. *Mon. Wea. Rev.*, **143**, 3948–3955.
- , and ——, 2017: Evaluating the trade-offs between ensemble size and ensemble resolution in an ensemble-variational data assimilation system, *J. Adv. Model. Earth Syst.*, **9**, 781–789.
- , J. L. Anderson and G. S. Romine, 2015: Empirical localization functions for ensemble Kalman filter data assimilation in regions with and without precipitation. *Mon. Wea. Rev.*, **143**, 3664–3679.
- Li, Z., J. C. McWilliams, K. Ide, and J. D. Farrara, 2015: A multiscale variational data assimilation scheme: Formulation and illustration. *Mon. Wea. Rev.*, **143**, 3804–3822.
- Li, J., W.W. Wolf, W.P. Menzel, W. Zhang, H. Huang, and T.H. Achtor, 2000: Global soundings of the atmosphere from ATOVS measurements: The algorithm and validation. *J. Appl. Meteor.*, **39**, 1248–1268.
- Li H, Kalnay E, Miyoshi T. 2009. Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Q. J. R. Meteorol. Soc.* **135**: 523–533.

- Liang X, Zheng X, Zhang S, Wu G, Dai Y, Li Y. 2011. Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble Kalman filter assimilation. *Q. J. R. Meteorol. Soc.* **138**: 263-273.
- Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models. *Mon. Wea. Rev.*, **138**, 1587–1612.
- Ling, J., P. Bauer, P. Bechtold, A. Beljaars, R. Forbes, F. Vitart, and C. Zhang, 2014: Global versus local MJO forecast skill of the ECMWF model during DYNAMO. *Mon. Wea. Rev.*, **142**, 2228–2247.
- 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.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- , 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646.
- , 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505-513.
- , 1996: Predictability—A problem partly solved. *Proc. Seminar on Predictability*, Reading, United Kingdom, ECMWF, 1–18.
- , and K. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.*, **55**, 399-414.
- Love, B. S., A. J. Matthews, and G. M. S. Lister, 2011: The diurnal cycle of precipitation over the Maritime Continent in a high-resolution atmospheric model. *Quart. J. Roy. Meteorol. Soc.*, **137**, 934-947.

- Madden, R. A., and P. R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702-708.
- , and ——, 1972: Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci.*, **29**, 1109-1123.
- Mapes, B. E., T. T. Warner, and M. Xu, 2003: Diurnal patterns of rainfall in northwestern South America. Part III: Diurnal gravity waves and nocturnal convection offshore. *Mon. Wea. Rev.*, **131**, 830-844.
- , S. Tulich, T. Nasuno, and M. Satoh, 2008: Predictability aspects of global aquaplanet simulations with explicit convection. *J. Meteorol. Soc. Japan.*, **88A**, 175-185.
- Mannucci, A. J., J. Dickson, C. Duncan, and K. Hurst, 2010: GNSS Geospace Constellation (GGC): A Cubesat Space Weather Mission Concept. Jet Propulsion Laboratory, California Institute of Technology, Pasadena.
- Melhauser, C., F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69**, 3350-3371.
- Meng, Z., and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Weather Rev.* **135**: 1403-1423.
- , and ——, 2008: Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part III: Comparison with 3DVAR in a Real-Data Case Study. *Mon. Wea. Rev.*, **136**, 522–540.
- Mitchell, H. L., and P. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Mon. Weather Rev.* **128**: 416-433.

- Minamide, M. and F. Zhang, 2017: Adaptive Observation Error Inflation for Assimilating All-Sky Satellite Radiance. *Mon. Wea. Rev.*, **145**, 1063–1081.
- Miyakawa, T., M. Satoh, H. Miura, H. Tomita, H. Yashiro, A. T. Noda, and K. Yoneyama, 2014: Madden-Julian oscillation prediction skill of a new-generation global model demonstrated using a supercomputer. *Nature Communications*, **5**, 3769.
- Miyoshi T. 2011. The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Weather Rev.* **139**: 1519-1535.
- , E. Kalnay, and H. Li, 2013: Estimating and including observation-error correlations in data assimilation. *Inverse Problems in Science and Engineering*, **21**, 387-398.
- , and K. Kondo, 2013: A multi-scale localization approach to an ensemble Kalan filter. *SOLA*, **9**, 170-173.
- Nakazawa, T., 1988: Tropical super clusters within intraseasonal variations over the western Pacific. *J. Meteor. Soc. Japan*, **66**, 823-836.
- Nasuno, T., 2013: Forecast skill of Madden-Julian oscillation events in a global nonhydrostatic model during the CINDY2011/DYNAMO observation period. *Sola*, **9**, 69-73.
- Neena, J. M., J. Y. Lee, D. Waliser, B. Wang, and X. Jiang, 2014a: Predictability of the Madden-Julian oscillation in the Intraseasonal Variability Hindcast Experiment (ISVHE). *J. Climate*, **27**, 4531–4543.

- , X. Jiang, D. Waliser, J. Y. Lee, and B. Wang, 2014b: Eastern Pacific intraseasonal variability: A predictability perspective. *J. Climate*, **27**, 8869–8883.
- Nerger, L., T. Janjic, J. Schroeter, and W. Hiller, 2012: A regulated localization scheme for ensemble-based Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **138**, 802–812.
- Nieman, S. J., W. P. Menzel, C. M. Hayden, D. Gray, S. Wanzong, C. S. Velden, and J. Daniels, 1997: Fully automated cloud-drift winds in NESDIS operations. *Bull. Amer. Meteor. Soc.*, **78**, 1121–1134.
- Parrish, D.F. and J.C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Perianez, A., H. Reich, and R. Potthast, 2014: Optimal localization for ensemble Kalman filter systems. *J. Meteor. Soc. Japan*, **92**, 585–597.
- Reynolds, C. A., and P. J. Webster, 1994: Random error growth in NMC's global forecasts. *Mon. Wea. Rev.*, **122**, 1281–1305.
- Reale, A. L., 2001: NOAA operational sounding products from Advanced-TOVS polar orbiting environmental satellites. *NOAA Technical Report NESDIS 102*.
- Reynolds, C. A., and P. J. Webster, 1994: Random error growth in NMC's global forecasts. *Mon. Wea. Rev.*, **122**, 1281–1305.
- Rotunno, R., and C. Snyder, 2008: A generalization of Lorenz's model for the predictability of flows with many scales of motion. *J. Atmos. Sci.*, **65**, 1063–1076.
- Ruf, C.S., R. Atlas, P.S. Chang, M.P. Clarizia, J.L. Garrison, S. Gleason, S.J. Katzberg, Z. Jelenak, J.T. Johnson, S.J. Majumdar, A. O'brien, D.J. Posselt, A.J. Ridley, R.J. Rose, and V.U. Zavorotny, 2016: New Ocean Winds Satellite Mission to Probe Hurricanes and Tropical Convection. *Bull. Amer. Meteor. Soc.*, **97**, 385–395.

- Sakov, P., and L. Bertino, 2011: Relation between two common localization methods for the EnKF. *Comput. Geosci.*, **15**, 225-237.
- Salmon, R., 1998: *Lectures on geophysical fluid dynamics*, Oxford University Press.
- Selz, T., and G. C. Craig, 2015: Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Mon. Wea. Rev.*, **143**, 813-827.
- Schwartz, C.S., G.S. Romine, K.R. Smith, and M.L. Weisman, 2014: Characterizing and Optimizing Precipitation Forecasts from a Convection-Permitting Ensemble Initialized by a Mesoscale Ensemble Kalman Filter. *Wea. Forecasting*, **29**, 1295–1318.
- Shutts G. 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Q. J. R. Meteorol. Soc.* **131**: 3079-3102.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. *NCAR Tech. Note*, NCAR/TN-475+STR, 113 pp.
- Smith, E. K., and S. Weintraub, 1953: The constants in the equation for atmospheric refractivity index at radio frequencies. *J. Res. Natl. Bur. Stand.*, **50**, 39–41.
- Smith, K. S., G. Boccaletti, C. C. Henning, I. Marinov, C.Y. Tam, I. M. Held, and G. K. Vallis, 2002: Turbulent diffusion in the geostrophic inverse cascade. *J. Fluid Mech.*, **469**, 13-48.
- Snook, N., X. Ming, and Y. Jung, 2015: Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Mon. Wea. Rev.*, **143**, 1035-1057.

- Sobash, R. A., and D. J. Stensrud, 2013: The impact of covariance localization for radar data on EnKF analyses of a developing MCS: Observing system simulation experiments. *Mon. Wea. Rev.*, **141**, 3691-3709.
- Sun, Y. Q., and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *J. Atmos. Sci.*, **73**, 1419-1438.
- Takayabu, Y. N., and T. S. Nitta, 1993: 3-5 day-period disturbances coupled with convection over the tropical Pacific Ocean. *J. Meteorol. Soc. Japan*, **71**, 221-246.
- Tao, D., and F. Zhang, 2015: Effects of vertical wind shear on the predictability of tropical cyclones: Practical versus intrinsic limit. *J. Adv. Model. Earth Syst.*, **7**, 1534-1553.
- Tavolato, C., and L. Isaksen, 2015: On the use of a Huber norm for observation quality control in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **141**, 1514–1527.
- Velden, C. S., and K. M. Bedka, 2009: Identifying uncertainty in determining satellite-derived atmospheric motion vector height attribution. *J. Applied Meteor. Climat.*, **48**, 450-463.
- , C. Hayden, S. Nieman, W. Menzel, S. Wanzong, and J. Goerss, 1997: Upper-tropospheric winds derived from geostationary satellite water vapor observations. *Bull. Amer. Meteor. Soc.*, **78**, 173-195.
- Vitart, F., and F. Molteni, 2010: Simulation of the Madden–Julian Oscillation and its teleconnections in the ECMWF forecast system. *Q. J. R. Meteorol. Soc.*, **136**, 842–855.

- Waliser, D. E., K. M. Lau, W. Stern, and C. Jones, 2003: Potential predictability of the Madden-Julian oscillation. *Bull. Amer. Meteor. Soc.*, **84**, 33–50.
- , and Coauthors (CLIVAR MJO Working Group) 2009: MJO simulation diagnostics. *J. Climate*, **22**, 3006–3030.
- Wang, B. R., X. Y. Liu, and J. K. Wang, 2013: Assessment of COSMIC radio occultation retrieval product using global radiosonde data. *Atmos. Meas. Tech.*, **6**, 1073–1083.
- Wang, S., A. H. Sobel, F. Zhang, Y. Q. Sun, Y. Yue, and L. Zhou, 2015: Regional simulation of the October and November MJO events observed during the CINDY/DYNAMO field campaign at gray zone resolution. *J. Climate*, **28**, 2097–2119.
- Wang X, Bishop CH. 2003. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60**: 1140-1158.
- Wheeler, M. C., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.*, **56**, 374-399.
- , and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917-1932.
- Whitaker, J. S., and T. M. Hamil, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913-1924.
- , and ———, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Weather Rev.* **140**: 3078-3089.

- Yoneyama, K., C. Zhang, and C. N. Long, 2013: Tracking pulses of the Madden-Julian oscillation. *Bulletin of the American Meteorological Society*, **94**, 1871-1891.
- Yoon, Y., E. Ott, and I. Szunyogh, 2010: On the propagation of information and the use of localization in ensemble Kalman filtering. *J. Atmos. Sci.*, **67**, 3823-3834.
- Yunck, T. P., E. J. Fetzer, A. M. Mannucci, C. O. Ao, W. Irion, B. D. Wilson, and G. J. Manion, 2009: Use of radio occultation to evaluate atmospheric temperature data from spaceborne infrared sensors, *Terr. Atmos. Ocean. Sci.*, **20**, 71–85.
- Zeng, X., and A. Beljaars, 2005: A prognostic scheme of sea surface skin temperature for modeling and data assimilation. *Geophys. Res. Lett.*, **32**, L14605.
- Zhang, C., 2005: The Madden-Julian oscillation. *Rev. Geophys.*, **43**, RG2003.
- , J. Gottshalck, E. D. Maloney, M. W. Moncrieff, F. Vitart, D. E. Waliser, B. Wang, and M. C. Wheeler, 2013: Cracking the MJO nut. *Geophysical Research Letters*, **40**, 1223-1230.
- Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the “surprise” 24–25 January 2000 snowstorm. *Mon. Wea. Rev.*, **130**, 1617–1632.
- , C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173-1185.
- , C. Snyder, and J. Sun, 2004: Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **132**, 1238–1253.
- , A. M. Odins, and J. W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Weather and Forecasting*, **21**, 149-166.

- , Z. Meng, and A. Aksoy, 2006: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722-736.
- , N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.*, **64**, 3579-3594.
- , and J. A. Sippel, 2009: Effects of moist convection on hurricane predictability. *J. Atmos. Sci.*, **66**, 1944-1961.
- , Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105-2125.
- , M. Minamide, and E. E. Clothiaux, 2016: Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954-2963.
- , S. Talaphdar, and S. Wang, 2017: The role of global circumnavigating mode in the MJO initiation and propagation. *J. Geophysical Research - Atmosphere*, accepted.
- Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2016: Intrinsic predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma at storm scales. *Mon. Wea. Rev.*, **144**, 127-1298.
- Zhen, Y., and F. Zhang, 2014: A probabilistic approach to adaptive covariance localization for serial ensemble square-root filters. *Mon. Wea. Rev.*, **142**, 4499-4518.
- Zheng, X. G., 2009: An adaptive estimation of forecast error covariance parameters for Kalman filtering data assimilation. *Adv. Atmos. Sci.* **26**: 154-160.

Zhu, K., Y. Pan, M. Xue, X. Wang, J. S. Whitaker, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A regional GSI-based ensemble Kalman filter data assimilation system for the Rapid Refresh configuration: Testing at reduced resolution. *Mon. Wea. Rev.*, **141**, 4118–4139.

Vita

Yue Ying

Department of Meteorology and Atmospheric Science
Pennsylvania State University, University Park, PA, 16802
Email: xyy159@psu.edu Tel: 814-753-0104

Research Interests

- Numerical modeling of atmospheric phenomena across scales
- Dynamics and predictability of coupled multiscale systems
- Data assimilation techniques that adapt to dynamical systems and observations

Education

2018: Ph.D. Meteorology, Computational Science (minor), Pennsylvania State University

2012: M.S. Meteorology, Peking University

2009: B.S. Atmospheric Sciences, Peking University

Journal Publications

Ying, Y., and F. Zhang, 2018: Potentials in improving predictability of multiscale tropical weather systems evaluated through ensemble assimilation of simulated satellite-based observations. *J. Atmos. Sci.*, accepted.

Ying, Y., F. Zhang, and J. L. Anderson, 2018: On the selection of localization radius in ensemble filtering for multiscale quasi-geostrophic dynamics. *Mon. Wea. Rev.*, accepted.

Ying, Y., and F. Zhang, 2017: Practical and intrinsic predictability of multi-scale weather and convectively-coupled equatorial waves during the active phase of an MJO. *J. Atmos. Sci.*, 74, 3771-3785.

Ying, Y., and F. Zhang, 2015: An adaptive covariance relaxation method for ensemble data assimilation. *Quart. J. Roy. Meteor. Soc.*, 141, 2898-2906.

Ying, Y., and Q. Zhang, 2012: A modeling study on tropical cyclone structural changes in response to ambient moisture variations. *J. Meteorol. Soc. Japan*, 90, 755-770.

Teaching Experience

2016-2017: Guest Lecturer of *Data Assimilation*, Pennsylvania State University.

2011: Teaching Assistant for *Computer Algorithms and Data Structure*, Peking University.

Professional Memberships

Since 2017: Chi Epsilon Pi National Meteorology Honors Society.

Since 2012: American Meteorological Society.

Honors and Awards

2018: Best Student Presentation, 22nd AMS Conference on IOAS-AOLS.