# INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600



Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/556

UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

# EFFECTS OF NESTING FREQUENCY AND LATERAL BOUNDARY PERTURBATIONS ON THE DISPERSION OF LIMITED-AREA ENSEMBLE FORECASTS

A Dissertation
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy

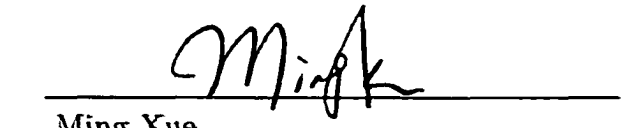
By
PAUL A. NUTTER
Norman, Oklahoma
2003

© by PAUL A. NUTTER 2003
All Rights Reserved.

# EFFECTS OF NESTING FREQUENCY AND LATERAL BOUNDARY PERTURBATIONS ON THE DISPERSION OF LIMITED-AREA ENSEMBLE FORECASTS

A Dissertation APPROVED for the
School of Meteorology

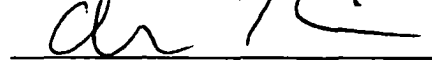
**By**



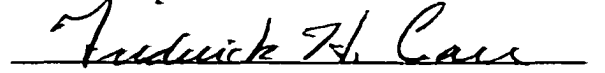
Ming Xue



David J. Stensrud



Alan Shapiro



Frederick H. Carr



Ramkumar N. Parthasarathy

# Acknowledgements

This research was supported by funding from several sources: the University of Oklahoma (OU) School of Meteorology through the Robert E. Lowry endowed chair; the Department of Defense of the U.S.A. under grant N00014-96-1-1112, or Coastal Meteorology Research Program (CMRP); the Williams Energy Marketing and Trading Company through a Williams Graduate Research Fellowship within the OU Center for Analysis and Prediction of Storms (CAPS); the National Science Foundation under grant ATM-0129892, related to the International $H_2O$ Project (IHOP); and the OU Alumni Association Fellowship Program.

The National Severe Storms Laboratory (NSSL) shared desk space and access to a multiprocessor computer. The OU School of Meteorology provided excellent infrastructure and administrative support. Most of the computations were performed on the IBM Regatta p690 symmetric multiprocessor computer operated by the OU Supercompomputing Center for Education and Research (OSCER).

Support, guidance, and encouragement is given by people rather than by financial accounts. Dr. Eugenia Kalnay supported my first year as a research assistant, even though course work prevented me from devoting full attention to research until after her departure. Dr. Brian Fiedler helped guide me through the General Examination and introduced me to several colleagues at the Naval Research Laboratory. He allowed me to continue pursuing my own research interests even though they were not immediately pertinent to the goals of his research programs. Dr. David Stensrud stepped in to help refocus and drive my research program forward after I became sidetracked while curiously exploring numerous related issues. Dr. Ming Xue continued my scientific and financial support after assuming the role of committee chair. My remaining committee members, Drs. Frederick Carr, Alan Shapiro, and Ramkumar Parthasarathy, provided ongoing helpful comments and advice. Although not a committee member, Dr. Kelvin Droegemeier offered me the Williams

Graduate Research Fellowship and helped arrange my transition to CAPS funding. The individuals named above are models of generosity.

Dr. Steven Mullen and David Baumhefner helped inspire my early interest in predictability and ensemble forecasting research while earning my Master of Science degree at the University of Arizona. They also goaded me to return for the Doctoral degree while I was employed with ENSCO, Inc. Dr. Charles Doswell always seemed to appear at the right time to offer additional inspirational advice.

My colleagues at ENSCO, Inc., especially Dr. John Manobianco and Winifred Lambert, helped foster the confidence I needed to return for this degree through daily discussions of unique and challenging problems.

My close friends Dr. Jerry Brotzge, Robert Hale, William Martin, Dr. Ernani Nascimento, Karen Sonntag, and Dr. Steven Weygandt provided continued discussion, inspiration, and occasional outings of much needed humor and relaxation.

Special thanks to the continued encouragement from my family, especially my wife, Elaine, for unselfishly encouraging me to work a few extra hours each evening and accommodating our seemingly perpetual student lifestyle.

# Contents

**Acknowledgements** iv

**List Of Figures** ix

**List of Symbols** xii

**Abstract** xiv

**1 Overview and Background** 1
- 1.1 Introduction . . . 1
- 1.2 Previous Research . . . 3
  - 1.2.1 Predictability Estimates on Global Domains . . . 3
  - 1.2.2 Global Model Ensemble Forecasting . . . 5
  - 1.2.3 Predictability Estimates on Limited-Area Domains . . . 7
  - 1.2.4 Limited-Area Model Ensemble Forecasting . . . 10
- 1.3 Perspectives for Present Research . . . 11

**2 Ensemble Statistics** 13
- 2.1 Definitions . . . 13
  - 2.1.1 Notation . . . 13
  - 2.1.2 Ensemble MSE and Dispersion . . . 14
  - 2.1.3 Total Biased Error Variance . . . 15
  - 2.1.4 Scale Decomposition of Dispersion . . . 16
- 2.2 Statistical Expectation . . . 17
  - 2.2.1 Expected Value of Ensemble MSE . . . 17
  - 2.2.2 Expected Value of Ensemble Dispersion . . . 19
  - 2.2.3 Expected Value of Total Error Variance . . . 20
- 2.3 Normalized Error Variances . . . 20
- 2.4 Variance Budget for Limited-Area Domains . . . 22

**3 Parameterized P.V. Channel Model** 24
- 3.1 Model Dynamics . . . 24
  - 3.1.1 Quasi-geostrophic Potential Vorticity . . . 24

3.1.2 Parameterized Potential Vorticity Equation . . . . . 27
3.1.2.1 Pure Barotropic Model . . . . . 27
3.1.2.2 Parameterized Potential Vorticity Model . . . . . 29
3.1.3 Barotropic and Baroclinic Instability . . . . . 30
3.2 Numerical Solution . . . . . 32
3.2.1 Grid Configurations . . . . . 32
3.2.2 Initialization . . . . . 33
3.2.3 Numerical Methods . . . . . 34
3.2.3.1 Spatial Discretization . . . . . 34
3.2.3.2 Time Integration . . . . . 35
3.2.3.3 Numerical Diffusion . . . . . 36
3.2.4 Lateral Boundary Conditions . . . . . 37
3.3 Example Case . . . . . 39
3.3.1 Impact of Parameterization . . . . . 40
3.3.2 Energetics of Example Case . . . . . 42
3.3.3 Frequency Spectrum Analysis . . . . . 45
3.4 Model Climatology . . . . . 46

**4 Global Ensemble Simulations** **50**
4.1 Ensemble Example Case . . . . . 50
4.1.1 Initial Condition Perturbations . . . . . 50
4.1.2 Example Ensemble Dispersion . . . . . 51
4.2 Normalized Error Variance Spectra . . . . . 55
4.3 Ensemble Bias, Dispersion, and MSE . . . . . 57

**5 Limited-Area Ensemble Simulations** **60**
5.1 Impact of LBC Update Interval . . . . . 60
5.1.1 Example Simulations . . . . . 60
5.1.1.1 LBC Error "Pulsing" . . . . . 61
5.1.1.2 LBC "Sweeping" . . . . . 65
5.1.2 Error Variance Spectra . . . . . 73
5.1.3 Ensemble Error Variance Spectra . . . . . 78
5.1.4 Ensemble Summary Statistics . . . . . 83
5.2 Impact of Spatial Resolution in LBC . . . . . 87
5.2.1 Example of LBC "Sweeping" . . . . . 88
5.2.2 Error Variance Spectra . . . . . 91
5.2.3 Ensemble Error Variance Spectra . . . . . 96
5.2.4 Ensemble Summary Statistics . . . . . 99
5.3 Combined LBC effects . . . . . 102
5.3.1 Ensemble Error Variance Spectra . . . . . 102
5.3.2 Ensemble Summary Statistics . . . . . 105
5.3.3 Example of LAM ensemble dispersion . . . . . 105
5.4 Chapter Summary . . . . . 106

# List Of Figures


3.1 Bickley jet profiles for velocity and vorticity . . . . . 31
3.2 Configuration of grids nested within channel domain . . . . . 33
3.3 Example simulation from the barotropic model . . . . . 40
3.4 Example simulation from the PPV model . . . . . 41
3.5 Example simulation from the PPV model extended to 4 days . . . . . 43
3.6 Time-series of grid-mean kinetic energy, enstrophy, and wavelength for example PPV model simulations . . . . . 44
3.7 Frequency spectrum analysis for $\xi$ obtained from an example PPV model simulation . . . . . 46
3.8 Dimensionless variance spectra for streamfunction, vorticity, $u$ and $v$ wind components, and kinetic energy . . . . . 48

4.1 Example case showing initial vorticity and streamfunction perturbations and subsequent error growth through day six . . . . . 52
4.2 Time series of $\psi$ and $\xi$ drawn from individual ensemble members at a single grid point . . . . . 53
4.3 Example case showing "spaghetti" contours and ensemble dispersion through day six . . . . . 54
4.4 Normalized vorticity error variances for global ensemble simulations . 56
4.5 Global ensemble summary statistics for vorticity . . . . . 58

5.1 Example LAM simulation run on the small, center domain showing effects of 3-hourly updated LBCs in an otherwise perfect simulation . 62
5.2 Minimum resolved spatial scale of aliasing error after temporal filtering due to linear interpolation of LBCs between available updates . . 65
5.3 Example LAM simulation run on the large domain showing effects of 3-hourly updated LBCs in an otherwise perfect simulation . . . . . 67
5.4 Example LAM simulation run on the large domain showing effects of 6-hourly updated LBCs in an otherwise perfect simulation . . . . . 68
5.5 Example LAM simulation run on the medium domain showing effects of 3-hourly updated LBCs in an otherwise perfect simulation . . . . . 69
5.6 Example LAM simulation run on the medium domain showing effects of 6-hourly updated LBCs in an otherwise perfect simulation . . . . . 70

5.27 Normalized error variance for LAM ensemble simulations having low-pass filtered initial conditions and 3-hourly updated, low-pass filtered LBCs . . . . . 104
5.28 Reduction of spread for an example LAM ensemble run on the medium domain having 3-hourly updated, low-pass filtered LBCs . . . . . 107

6.1 Difference in error variance spectra between global ensemble simulations and LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs. . . . . 115
6.2 Example of LBC perturbation fields constructed using the difference in error variance spectra between global ensembles and LAM ensembles having 3-hourly updated, low-pass filtered LBCs. . . . . 117
6.3 Reduction of spread for an example LAM ensemble run on the medium domain having 3-hourly updated, low-pass filtered, and perturbed LBCs. . . . . 119
6.4 Normalized error variance for LAM ensemble simulations having perturbed, hourly updated LBCs. . . . . 121
6.5 Normalized error variance for LAM ensemble simulations having perturbed, 3-hourly updated LBCs. . . . . 122
6.6 Normalized error variance for LAM ensemble simulations having perturbed, 3-hourly updated, low-pass filtered LBCs. . . . . 123
6.7 Summary statistics for LAM ensemble simulations having perturbed, hourly updated LBCs. . . . . 126
6.8 Summary statistics for LAM ensemble simulations having perturbed, 3-hourly updated LBCs. . . . . 127
6.9 Summary statistics for LAM ensemble simulations having perturbed, 3-hourly updated, low-pass filtered LBCs. . . . . 128
6.10 Scale factor applied to difference spectra for LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs. . . . . 131
6.11 Normalized error variance for LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs with scaled perturbations. . . 133
6.12 Summary statistics for LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs with scaled perturbations. . . . . 134

B.1 1-D spectra on a grid . . . . . 156

# List of Symbols

| Symbol | Description | Page Introduced |
|---|---|---|
| $D^2$ | Ensemble dispersion | 15 |
| $E[\cdot]$ | Expected value | 18 |
| $F(\kappa)$ | Complex Fourier coefficient | 15 |
| $J$ | Jacobian operator | 27 |
| $L_x$ | Zonal domain length | 32 |
| $L_y$ | Meridional domain length | 32 |
| $M$ | Number of independent cases | 21 |
| $N$ | Number of ensemble members | 14 |
| $N$ | Brunt-Väisällä frequency | 26 |
| $N_x$ | Number of gridpoints in zonal direction | 32 |
| $N_y$ | Number of gridpoints in meridional direction | 32 |
| $R$ | Ideal gas constant | 25 |
| $S^2$ | Ensemble mean-squared error | 14 |
| $T$ | Temperature | 25 |
| $U_0$ | Mean zonal wind speed | 28 |
| $\vec{V}_g$ | Geostrophic wind vector | 25 |
| $a_k, b_k$ | Real Fourier amplitude coeffients | 35 |
| $c_x$ | Zonal phase speed | 28 |
| $c_{gx}$ | Zonal group velocity | 28 |
| $c_{gy}$ | Meridional group velocity | 28 |
| $f$ | Earth vorticity | 25 |
| $f_0$ | Earth vorticity at 45° north latitude | 25 |
| $g$ | Gravitational acceleration | 26 |
| $i$ | Integer subscript for individual ensemble members | 14 |
| $j$ | Integer subscript for individual case numbers | 21 |
| $i, j$ | Integer subscript for specific gridpoints | 33 |
| $i$ | Imaginary number, $\sqrt{-1}$ | 35 |
| $k$ | Zonal wavenumber | 28 |
| $\hat{k}$ | Unit vector in vertical coordinate direction | 25 |
| $l$ | Meridional wavenumber | 28 |
| $\ell$ | Wavelength | 47 |
| $\ell_x$ | Generalized zonal domain length | 33 |

# Abstract

It is known that "one-way" lateral boundary conditions (LBCs) constrain the growth of initial perturbations in limited-area model (LAM) ensemble forecasts, therefore reducing ensemble dispersion. LBCs provided by independent external models typically lack fine-scale features, and in the case of ensemble forecasts, also lack consistent perturbations. Perturbations growing on the nested domain become displaced by the coarsely resolved LBCs while the domain size itself limits the maximum wavelength attainable by the perturbations. Another aspect of the boundary condition problem that has previously received little direct attention is the impact of boundary condition update frequency. Commonly used linear interpolation between relatively infrequent boundary condition updates acts as a filter that exacerbates the scale deficiency problem.

To examine the above problem in a controlled and efficient manner, a modified barotropic channel model with periodic east-west boundaries is constructed and shown to produce flows bearing resemblance to the real atmospheric circulation. Ten-member ensemble simulations are produced over many cases on the periodic channel domain and each of four smaller nested domains. Lateral boundary effects are specifically isolated since the simulations are perfect except for initial condition perturbations and the use of coarsely resolved and/or temporally interpolated "one-way" LBCs on the nested domains. Statistical results accumulated over 100 independent cases demonstrate clearly the lack of growth in error variance spectra, particularly at wavelengths shorter than 500 km. A new expression is developed that links total error variance to ensemble dispersion while accounting for spatial and ensemble biases. The balances required by this expression are used to show that LBC constraints on small scale error variance growth are directly responsible for underdispersive LAM ensemble simulations.

To help restore error variance and LAM ensemble dispersion, a new method is developed to apply statistically consistent LBC perturbations at each time step that remain spatially and temporally coherent while passing through the boundaries. The amplitude of such perturbations are designed to increase with time while mimicking the error variance lost through the effects of coasely resolved and temporally interpolated LBCs. With a few noted exceptions, the LBC perturbations are shown to capably restore error variance growth and LAM ensemble dispersion without compromising the integrity of the individual solutions.

# Chapter 1

# Overview and Background

## 1.1 Introduction

It is well known that skill in predicting future atmospheric motions is limited due to the natural growth of errors resulting from imperfect observations and analyses (Thompson 1957; Lorenz 1963). More precisely, accepted predictability theories for kinematic fields ascribe an inverse relationship between wave number and time limits of predictive skill (e.g. Lorenz 1969*a*). In apparent contradiction to these theories, several predictability experiments using limited-area models (LAMs) report little or no error growth resulting from perturbed initial conditions (Paegle et al. 1997, references therein). Attempts to explain these seemingly optimistic results have focused either on enhanced local forcing (e.g., topography or surface inhomogeneities) or on the errors introduced by the use of "one-way" lateral boundary conditions (LBCs). The latter effect is favored in the literature and is the subject of the current research.

We begin in the present chapter by providing a limited overview of atmospheric predictability theories that have been tested using global analyses and forecasts. The review specifically highlights the scale dependence of error growth rates and the importance of both analysis uncertainty and model deficiencies as sources of error growth. Lessons learned from predictability studies have direct application to the development and verification of ensemble forecast systems. Hence, we also provide a brief overview of global ensemble forecast systems and summarize how ensemble dispersion may be used as a measure of an ensemble system's efficacy.

The review continues by summarizing previous efforts to refine predictability theories at smaller scales using mesoscale models operating on limited-area domains.

Emphasis in the review is placed on noting differences between error growth behavior in global and mesoscale (limited-domain) models. Results from the LAM studies were somewhat enigmatic and spawned related works to identify mechanisms through which LBCs constrain error growth rates. These known LBC effects are summarized and their subsequent impact on dispersion in short-range ensemble forecasts also is reviewed.

The relationship between LBC constraints on error growth and LAM ensemble dispersion is a crucial theme in this work. As such, a statistical framework is constructed in chapter two that defines ensemble mean square error, total error variance and ensemble dispersion. The individual statistics are then combined in a manner that provides a link between LBC constraints on error variance growth and ensemble dispersion. Chapter two concludes by outlining a hypothetical budget for error variance growth in LAMs that is tested in later chapters.

Many of the previous studies involving LAMs applied fields from discordant models or analyses as LBC forcing for the dependent LAM simulations. Under these configurations, LBCs were identified as the most likely candidate acting to constrain error variance growth. However, the LAM simulations run with these configurations could not completely control or avoid extraneous sources of model error. To examine the impact of LBCs on short-range ensemble dispersion in a more complete and controlled manner, a simplified modeling approach is adopted in chapter three. The simplified configuration employs a single level parameterized potential vorticity (PPV) channel model designed to specifically isolate LBCs as the *only* source of model error. Global and limited-area domains used for this work and the model's climatological characteristics are also described in chapter three.

In chapter four, "classic" predictability experiments are conducted to show that error growth characteristics in the simplified PPV model demonstrate appropriate behaviors. Specifically, statistics defined in chapter two are applied to the output from one-hundred independent ensemble simulations to show that error growth is most rapid at small scales, and reaches a maximum value that is twice the variance of analyses. The results in chapter four also establish benchmark statistics against which the magnitude of the LBC constraint may be measured in later experiments.

Error variance growth and ensemble dispersion characteristics for LAM simulations are considered in chapter five. The first part of the chapter examines the impact

of linear interpolation in time between LBC updates. Impacts seen in selected example simulations are supported by error variance statistics obtained over one-hundred independent cases. Results from LAM ensembles are reported in the second part of chapter five showing that temporal interpolation of LBCs and coarsely resolved LBCs both reduce the saturation value of error variance at small scales. This, in turn, is shown to lead directly to a loss of dispersion in the LAM ensembles.

In chapter six, a new method is proposed to apply small-scale LBC perturbations at every time step of the LAM simulations. With a few noted exceptions, the LBC perturbations are shown to help restore much of the error variance and ensemble dispersion lost due to LBC constraints. The most important and interrelated conclusions obtained from this research are summarized in chapter seven.

## 1.2 Previous Research

### 1.2.1 Predictability Estimates on Global Domains

Before examining the impact of LBC constraints on error growth in LAM simulations, we must first understand error growth on global, or laterally unbounded domains. Predictability studies provide this knowledge by considering the question of how rapidly initially similar atmospheric states diverge with time. The limit of predictability is reached when the difference between atmospheric states grows as large as the difference between random samples from climatology. Differences are usually evaluated using mean square errors, error variances, or anomaly correlations.

The most common approach for obtaining estimates of atmospheric predictability limits has been called the "dynamical" approach (Lorenz 1969*c*). In dynamical, or "classic" predictability experiments, two or more solutions of the equations of motion are obtained using numerical models, each starting from slightly different initial conditions (e.g. Charney et al. 1966; Smagorinsky 1969; Lorenz 1982; Baumhefner 1984; Dalcher and Kalnay 1987; Reynolds et al. 1994, many others). The rate of error growth is then easily evaluated by comparing the different model solutions as they diverge with time. Such experiments suffer from the fact that error growth rates are model dependent, so that different models yield different estimates of atmospheric predictability limits. Baumhefner (1984) pointed out that the dynamical approach is

adequate provided the numerical model used to generate the predictability estimate capably reproduces observed atmospheric variance structures across the spectrum.

Indeed, current general circulation models do reproduce variance structures quite well across observable scales outside the tropics. However, the impact of errors growing up scale from unresolved wavelengths is supported only theoretically using an "empirical-dynamical" approach (Lorenz 1969*c*). The empirical-dynamical method for estimating predictability describes the spectral distribution of errors following a manipulation of simplified equations of atmospheric motion (Lorenz 1969*b*; Lorenz 1969*c*; Leith 1971; Leith and Kraichnan 1972; Boer 1984; Lilly 1984; Thompson 1984, and others). The main conclusion from studies of this type is that any initially small error imposed to any part of the spectrum will grow and spread to other wave numbers via up-scale and down-scale transport of energy. Furthermore, small-scale errors are expected to grow most rapidly so that short wavelengths have short predictability limits.

The theoretically anticipated scale dependence of error growth rates has been verified in computational experiments employing the dynamical approach. For example, Baumhefner (1984) decomposed the total error growth in a global model into a two-dimensional spherical harmonic wavenumber spectrum. His analysis yielded estimated predictability limits ranging from 8.5 days for the zonal vortex (wavenumber 0) to 3.5 days for the smallest scales resolved in that model. Dalcher and Kalnay (1987) obtained similar results, additionally noting that atmospheric motions are more predictable in winter than in summer. Anthes (1986) provides a review of several other related studies, indicating that estimates for overall predictability limits range from about 8 to 16 days. In aggregate, these and many other reports show that predictability decreases at smaller scales and varies with space and time depending on the synoptic setting.

Even if we could provide errorless initial conditions (ICs), atmospheric motions cannot be represented perfectly by numerical models because of incompletely specified or unknown dynamical processes and the use of numerical approximations. Lorenz (1984) showed that these "artificial" model errors reduce predictability limits by about 4 days in a global model. By fitting model data to parametric error growth curves, Dalcher and Kalnay (1987) estimated that model error increases at a rate of about 3% per day at long waves and about 20% per day at short waves

when scaled relative to the maximum error variance. These results show that predictability limits are shortest at small scales regardless of whether errors originated from initial condition uncertainty or from model deficiencies. Boer (1993) derived equations for the evolution of systematic (biased) and random (unbiased) error components and showed that they are not independent. Reynolds et al. (1994) found similar results, but also highlighted the spatial dependence of both natural and artificial error growth. Conclusions from these and other studies indicate that artificial model error accelerates error growth rates variably across different locations, time scales, and spatial scales.

It is regularly asserted that global forecast models have matured enough that IC uncertainty is the dominant source of predictability error growth (e.g. Toth et al. 1997; Harrison et al. 1999). However, a few authors have warned about the need to investigate the importance of model error more carefully using alternative analysis procedures (Tribbia and Baumhefner 1988; Orrell et al. 2002).

### 1.2.2 Global Model Ensemble Forecasting

The recognition that weather prediction is inherently uncertain led to the viewpoint that forecasts are not complete unless accompanied by a probabilistic prediction of forecast skill (Palmer and Tibaldi 1988; Ehrendorfer 1994; Murphy 1998). The Liouville equation forms a basis for such predictions. It governs the temporal evolution of the probability density function (PDF) of the multivariate atmospheric state vector through phase space (Epstein 1969; Gleeson 1970; Thompson 1985; Ehrendorfer 1994). Direct prediction of the PDF contains the statistical moments needed to quantify forecast uncertainty. Epstein (1969) introduced the particular method of stochastic-dynamic prediction wherein the Liouville equation was reduced to equations approximating the time evolution of the PDF mean and covariance. Unfortunately, atmospheric models have so many degrees of freedom that stochastic-dynamic prediction is not feasible.

As an alternative, Leith (1974) proposed a Monte Carlo approach wherein a small ensemble of forecasts is generated from randomly perturbed initial states. The initial perturbations are consistent with analysis uncertainties, and subsequently grow with time as outlined in the predictability studies reviewed above. In the absence of model

error, each individual forecast produced in this manner is designed to represent an equally likely sample from the atmospheric PDF. Leith showed that an adequate estimate of the forecast error covariance, and hence, uncertainty in the forecast, could be obtained from an ensemble of about 8 members. Of course, larger ensembles are needed for more robust statistics. Leith's proposal forms the basis for modern ensemble forecasting techniques.

Hoffman and Kalnay (1983) introduced lagged average forecasting (LAF) as an alternative to the Monte Carlo method. In this approach, the current analysis is grouped with earlier forecasts valid at the current time to form an ensemble. An ensemble mean and covariance could then be estimated using the group of forecasts. Unlike the Monte Carlo method, LAF does not require the expensive production of extra forecasts. Results from LAF experiments indicated an improvement in skill over Monte Carlo methods, especially when scaling was applied to older forecasts (Hoffman and Kalnay 1983; Ebisuzsaki and Kalnay 1991). The improvement was attributed to the idea that the group of forecasts were largely identical except where errors were growing most rapidly.

Operational ensemble forecast systems were started in late 1992 at both the National Meteorological Center and the European Center for Medium-Range Weather Predictions (Tracton and Kalnay 1993; Molteni et al. 1996). Other agencies around the world have since developed global ensemble forecast systems (Sivillo et al. 1997). A variety of different methods are now used for generating IC perturbations. For example, "bred vector" and "singular vector" methods are designed to generate perturbations directed along the most rapidly growing modes to help maximize the rate at which individual ensemble members become different from one another (Houtekamer and Derome 1995; Anderson 1996; Molteni et al. 1996; Buizza 1997; Szunyogh et al. 1997; Toth and Kalnay 1997; Barkmeijer et al. 1998; Errico and Langland 1999*a*; Errico and Langland 1999*b*; Toth et al. 1999; Hamill et al. 2000) Other perturbation methods emphasize uncertainty in the observations or analysis system in a manner that is more consistent with Leith's original proposal (Houtekamer and Derome 1995; Houtekamer et al. 1996; Houtekamer and Mitchell 1998; Hamill and Snyder 2000).

An efficacious ensemble system has the property that the verifying analysis is encompassed within the range of equally likely solutions given by the individual

ensemble members (Leith 1974). Ensemble *spread*, or *dispersion*, is often used to determine if the ensemble supports this desired characteristic. It is a square norm statistic measuring the variability of the ensemble about its mean (see Ch. 2). An overdispersive ensemble will consistently predict scenarios that fall outside the range of weather events observed over many cases. Conversely, an underdispersive ensemble often fails to predict the possibility that a given event will occur (Sivillo et al. 1997; Hamill 2001). Analyses of global ensemble forecast systems have shown that they tend to be somewhat underdispersive (Zhu et al. 1996; Buizza 1997). However this shortcoming is not enough to preclude operational effectiveness of the ensembles (Tracton and Kalnay 1993; Toth et al. 1997) while ongoing research helps to improve the ensemble statistics (Molteni et al. 1996; Toth and Kalnay 1997; Errico and Langland 1999*b*).

### 1.2.3 Predictability Estimates on Limited-Area Domains

The empirical-dynamical predictability theories outlined above lead to a rather pessimistic viewpoint of mesoscale predictability due to the expectation that error growth is more rapid at smaller scales (Tennekes 1978; Anthes et al. 1985). However, Lilly (1984) and Anthes et al. (1985) proffer the alternative hypothesis that enhanced localized forcing and highly organized circulations may help slow error growth rates. From this perspective, Anthes et al. (1985, pg. 164) state that "...there is hope for skillful predictions [of mesoscale motions] using deterministic methods, provided that the synoptic-scale motions are predicted correctly".

An obvious approach for testing these different hypotheses is to use mesoscale, or limited-area models. However, the use of LAMs to conduct predictability studies introduces additional complications over the methods used for classical studies involving global models (Anthes et al. 1985). For example, regional domains may encompass only one energetically active weather system so that error growth rates vary widely from case to case. Furthermore, LAMs often employ more complex physical parameterization schemes that introduce additional sources of model error. Finally, errors introduced by using LBCs can propagate inward at speeds of 20-30° longitude per day (Baumhefner and Perkey 1982).

Anthes et al. (1985) and Anthes (1986) were among the first to conduct predictability experiments using LAMs. In contrast to the global predictability studies reviewed above, they reported *little or no error growth resulting from small variations in the initial conditions*. Furthermore, they showed that their simulations were more sensitive to variations in the LBCs than to the ICs. Errico and Baumhefner (1987), followed on this work, reporting that IC error growth is restricted in a mesoscale model due to dissipation of errors through gravity waves, the "sweeping out" of errors by the LBCs, and numerical dissipation of subgrid-scale errors. Anthes et al. (1989) reported similar results, stating that "...the quality of the LBC is more important than any other factor tested in the temporal evolution of model errors".

Vukicevic and Paegle (1989) used a non-divergent barotropic model to demonstrate that "one-way" LBCs artificially enhanced predictability by limiting the range of scales that could interact with errors. The effect became more pronounced as the size of the regional domain decreased because lateral boundary information crossed the domain more quickly and had less time to interact with initial condition errors growing up scale. Vukicevic and Errico (1990) extended this research to a complex LAM and obtained the similar result that domain size limits predictability error growth by imposing an upper wavelength limit on the horizontal scales to which perturbations may evolve in accordance with model dynamics. In particular, they found that initial perturbations grow only in domains larger than about 4500 $km^2$ and damp in smaller domains. Furthermore, a spectral analysis of the error fields showed that only synoptic scale perturbations having wavelengths longer than about 2000 km contributed to error growth, while those with wavelengths shorter than about 1000 km damped in all experiments. However, while the small-scale errors are themselves damped with time, they still influence synoptic scale error growth through an up-scale transfer of energy (Vukicevic and Errico 1990).

While LBCs are now recognized as an artificial constraint on mesoscale error growth, other studies have emphasized the possibility for enhanced predictability due to stronger dynamical forcing from the lower boundary. For example, sensitivity to small-scale initial condition errors could be small in cases where topographic forcing is dominant (Van Tuyl and Errico 1989; Vukicevic and Errico 1990). Warner et al. (1989) also found that error growth over the interior of the mesoscale domain

decreased most rapidly during the first eight forecast hours, before lateral boundary errors could have propagated into the area tested. They suggested that this decrease in error was due to geostrophic adjustment, surface forcing, and redevelopment of smoothed features. These results indicate that simple turbulence models of error growth which anticipate the most rapid error growth at small scales may not fully explain actual error growth behaviors.

The nature of the initial perturbation also remains an important aspect of consideration. Vukicevic (1991) showed that the initial error evolution is well represented using a mesoscale tangent linear model for periods of about 24 to 36 hours. Ehrendorfer and Errico (1995) applied this result to investigate mesoscale predictability using singular vectors. They found that significant error growth did not occur in earlier mesoscale predictability studies because the random initial perturbations did not project onto the unstable subspace of growing singular vectors. Other constraints on error growth such as boundary "sweeping", numerical dissipation, or geostrophic adjustment were shown to be insignificant compared to the large projection of perturbations onto the stable subspace of diminishing singular vectors. However, Ehrendorfer and Errico (1995) pointed out that their results could not prove that mesoscale processes are inherently more predictable than synoptic scale motions because their linear model did not consider moist or strongly nonlinear physical processes, or error growth in synoptic scale waves. Also, they did not consider time periods beyond 24 hours, a time scale that may be too short to allow a larger subspace of perturbations to grow. However, the fact that the fastest-growing singular vectors yielded growth rates comparable those reported for global models suggests that the mesoscale is not more stable than the scales investigated in global predictability studies.

Results from the studies reviewed above do not indicate with certainty the relative importance of the different effects contributing to artificial constraints on error growth rates in LAMs. The evidence tends to favor LBC effects, but the other mechanisms such as enhanced local forcing, physical parameterizations, or poorly specified initial perturbations have not been eliminated from consideration. To help specifically isolate LBC effects on error growth, Laprise et al. (2000) and De Elía and Laprise (2002) generated "perfect" reference simulations by running a LAM on a large domain with LBCs given by global analyses. These reference fields were

then filtered to emulate the loss of shortwave components in external model fields having lower spatial resolution. The filtered reference fields were then used as LBCs for a second LAM simulation run with the same configuration except on a smaller domain. In this configuration, the only source of error was the use of the spatially filtered one-way LBC. Spectral analysis of the error fields showed that error growth is fastest at small scales, a result that is consistent with global predictability experiments reviewed above. The authors concluded that predictability may be maintained for longer periods in scales that are present in both initial and lateral boundary conditions.

### 1.2.4 Limited-Area Model Ensemble Forecasting

The successful implementation of global ensemble systems led naturally to the development of short-range ensemble forecasting (SREF) systems using mesoscale models on limited-area domains (Mullen and Baumhefner 1989; Brooks et al. 1995; Du et al. 1997; Hamill and Colucci 1997; Leslie and Speer 1998; Stensrud et al. 1999; Stensrud et al. 2000; Hou et al. 2001; Wandishin et al. 2001). Much of the research on global model ensembles has focused on the question of generating appropriate IC perturbations. While this question remains important in SREF (Ehrendorfer and Errico 1995; Hamill and Colucci 1997; Mullen et al. 1999), there are several additional artificial and physical factors that limit mesoscale predictability error growth. Many such factors were reviewed above, including domain size, LBC-induced error growth, scale interaction, enhanced localized forcing, and effects of numerical diffusion and physical parameterizations (Errico and Baumhefner 1987; Vukicevic and Errico 1990; Stensrud et al. 1999).

Recent (SREF) experiments have shown that the ensembles are underdispersive (Hamill and Colucci 1997; Du and Tracton 1999; Hou et al. 2001; Wandishin et al. 2001). That is, the verifying analysis often does not fall within the range of possibilities forecast by the ensemble. Du and Tracton (1999) found that a regional ensemble with a larger domain produces greater spread than does an ensemble with a smaller domain, especially for those variables that were perturbed in the ICs. Furthermore, they found that the contribution to ensemble spread increases with time from LBC perturbations and decreases with time from the IC perturbations.

The effect is rapid for dynamically active variables such as 500-hPa height but is relatively slower for passive (derived) variables such as precipitation. Hou et al. (2001) showed that ensemble spread for mass variables grows quickly for a SREF system configured so that each member forecast has a unique LBC; a SREF system configured so that each member has the same LBC had a slower rate of ensemble dispersion. Other variables such as precipitation were more sensitive to model physics (consistent with Du and Tracton 1999). These and other similar results (Hamill and Colucci 1997; Hou et al. 2001) demonstrate that, with time, the spread of the LAM forecast ensemble becomes increasingly determined by the spread in the global ensemble as high frequency components are "swept" from the LAM domain.

## 1.3 Perspectives for Present Research

In sum, results from the global model predictability experiments reviewed above support the theory that error growth is most rapid at small scales. The studies also showed that natural error growth due to analysis errors dominates model-induced error growth, especially at large scales. In contrast to these results, predictability studies using LAMs consistently reported lack of error growth at small scales while large scale errors amplified as expected. The advection of coarsely resolved or unperturbed LBCs into the regional domain is highlighted as the most probable reason for the constraint on error growth in LAMs. However, other sources of model error such as physical parameterizations, enhanced local forcing, or incorrectly specified initial perturbations may also cause artificial constraints on error growth. There has not yet been enough evidence presented to argue either against or in favor of the hypothesis that mesoscale motions should be more predictable than anticipated by spectral theories.

From these conclusions, the present study begins with the following two assumptions. First, it is assumed that natural error growth at large scales is (or can be) stated accurately using global model forecasts. Second, the assumption is made that error growth rates at small scales in LAMs should behave the same as calculated for global models operating at equivalent resolution.

These assumptions allow a more complete and controlled investigation of the impact of coarsely resolved LBCs on error growth rates in LAMs using a simplified

and idealized "perfect" model configuration. Following an approach similar to that used by Vukicevic and Paegle (1989), Laprise et al. (2000), and De Elía and Laprise (2002), an appropriate high-resolution model is used to construct a known truth on a global, or laterally unbounded domain (see Ch. 3). Then, using the same model, simulations are run on a regional domain with LBCs given by a subset of the control fields. With this approach, the only unavoidable source of model errors are those introduced by the use of the "one-way" LBCs. Additional errors may be introduced by perturbing the ICs, or by filtering the control fields to remove small scale information from the LBCs. If the evolution of the large-scale fields is specified using perfect LBCs, then differences between small scale error growth on the regional and global domains may be attributed directly to LBC effects. In this study, we seek to identify changes in error growth rates in LAMs caused specifically by different domain sizes, inadequate spatial resolution of the LBC, and inadequate temporal resolution of the LBC caused by linear interpolation in time between LBC updates. The latter effect has previously received little direct attention in the literature.

Results from studies reviewed above also show that global ensembles are slightly underdispersive, but not so much as to preclude useful operational implementation. On the other hand, LAM ensembles are more strongly underdispersive, most likely due to the LBC constraints on error growth. From these results, another assumption is made that LAM ensembles should produce the same spread as global ensembles when operated with the same model and initial perturbations. Although other effects may be important, we choose to specifically isolate the influence of LBCs and show that the lack of error growth at small scales is responsible for underdispersive LAM ensembles. Hence, we show in the present study that, in the absence of other effects, SREFs remain underdispersive unless LBCs are given with the same spatial and temporal resolution as the nested model simulation. Because such a configuration is not possible in a realistic operational setting, a new method is proposed to help restore in a statistically consistent manner the variance lost by unresolved scales in LBCs.

# Chapter 2

# Ensemble Statistics

Statistical methods for interpreting results from ensemble predictions on both global and limited-area domains are discussed in this chapter. The chapter begins by introducing basic vector notations, then reviews established statistical measures of performance within the context of this notation. The measures defined are ensemble mean square error, dispersion, and total error variance. The total error variance has equivalent vector and spectral representations, allowing decompositions of error variance into contributions by different scales of motion. A new expression is derived that provides a direct link between ensemble dispersion and the scale decomposition of error variance. This expression allows one to measure how ensemble dispersion is affected by error growth at individual scales of motion.

Expected values of the statistics are then derived to provide a basis for normalizing results and assessing predictive skill relative to the climatological variance. The chapter concludes by proposing a variance *budget* for limited-area models (LAMs) configured with "one-way" lateral boundary conditions. The proposed variance budget is important because it provides a mathematical statement of the hypothesis driving this research.

## 2.1 Definitions

### 2.1.1 Notation

Spatial and ensemble means are described here using notation inspired largely by Stephenson and Doblas-Reyes (2000). Suppose $\mathbf{x}_i$ is a vector field defined on a

$p$-element grid, representing forecasts or analyses obtained from an $N$-member ensemble, where $i = 1, \ldots, N$. The ensemble mean (a $p$-element column vector) is defined

$$\overline{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i. \tag{2.1}$$

The scalar spatial mean for the *ith* ensemble member is given by

$$\langle \mathbf{x}_i \rangle = \frac{1}{p} \sum_{k=1}^{p} \mathbf{x}_{i,k} = \frac{1}{p} \mathbf{1} \cdot \mathbf{x}_i, \tag{2.2}$$

where $\mathbf{1}$ is a $p$-element vector of ones. A useful norm representing the average sum of squares (dot product) over the grid is

$$\|\mathbf{x}_i\|^2 = \frac{1}{p} \mathbf{x}_i \cdot \mathbf{x}_i = \frac{1}{p} \sum_{k=1}^{p} \mathbf{x}_{i,k}^2. \tag{2.3}$$

Note for later derivations that $\|\mathbf{1}\|^2 = (1/p)\mathbf{1} \cdot \mathbf{1} = 1$.

Henceforth, let $\mathbf{f}_i(t)$ denote individual forecast vectors from the ensemble and $\mathbf{a}(t)$ represent the analysis vector corresponding to each forecast. These notations and the derivations that follow are applicable for gridded fields on both global and limited-area domains unless specified otherwise.

### 2.1.2 Ensemble MSE and Dispersion

One of the most basic measures of ensemble forecast accuracy is the *ensemble mean square error* (MSE),

$$S^2 = \frac{1}{N} \sum_{i=1}^{N} \|\mathbf{f}_i - \mathbf{a}\|^2. \tag{2.4}$$

It is useful to manipulate $S^2$ by adding and subtracting the ensemble mean forecast $\overline{\mathbf{f}}$ and analysis $\overline{\mathbf{a}} = \mathbf{a}$ such that

$$\begin{aligned} S^2 &= \frac{1}{N} \sum_{i=1}^{N} \left\| (\overline{\mathbf{f}} - \overline{\mathbf{a}}) + (\mathbf{f}_i - \mathbf{a}) - (\overline{\mathbf{f}} - \overline{\mathbf{a}}) \right\|^2 \\ &= \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{f}_i - \overline{\mathbf{f}} \right\|^2 + \left\| \overline{\mathbf{f}} - \mathbf{a} \right\|^2. \end{aligned} \tag{2.5}$$

Note that cross product terms vanish after taking the ensemble average during expansion of the norm. If $D^2 = \frac{1}{N}\sum_{i=1}^{N} \left\|\mathbf{f}_i - \overline{\mathbf{f}}\right\|^2$ defines the *ensemble dispersion*, or spread, then it follows directly from (2.5) that

$$S^2 = D^2 + \left\|\overline{\mathbf{f}} - \mathbf{a}\right\|^2. \tag{2.6}$$

This result shows that the squared error of the ensemble mean $\left\|\overline{\mathbf{f}} - \mathbf{a}\right\|^2$ is less than the ensemble mean square error $S^2$ because ensemble dispersion $D^2$ allows unpredictable components of the flow to be averaged out in the ensemble mean (Leith 1974; Stephenson and Doblas-Reyes 2000).

### 2.1.3 Total Biased Error Variance

Ensemble MSE and dispersion are commonly used measures of ensemble accuracy, but do not provide for a direct scale decomposition of error growth. The desired scale decomposition is provided by the *total biased error variance*, $\sigma^2$, because it has equivalent vector and spectral definitions. In vector form, $\sigma^2$ is defined as an ensemble average error variance relative to a spatial mean so that

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} \left\|(\mathbf{f}_i - \mathbf{a}_i) - \langle \mathbf{f}_i - \mathbf{a}_i \rangle \, \mathbf{1}\right\|^2. \tag{2.7}$$

The unity vector multiplies the scalar spatial mean to enable vector subtraction and to allow proper application of the norm notation.

Using Eq. (A.10–Appendix A), $\sigma^2$ may be obtained spectrally as (Errico 1985)

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\sum_{\kappa=1}^{K-1} 2\left|F_i(\kappa)\right|^2, \tag{2.8}$$

where $F_i(\kappa)$ is the discrete Fourier transform of $\mathbf{f}_i - \mathbf{a}$ (see Appendix B) and $\kappa = 1, \ldots, K-1$ are the set of Nyquist resolved wavenumbers on the grid. In this form, error variances may be computed individually for specific wavenumbers, or accumulated over a range of scales.

The total error variance $\sigma^2$ can be related to $S^2$ as follows. First, multiply the grid average for each forecast $\langle \mathbf{f}_i \rangle$ and analysis $\langle \mathbf{a} \rangle$ by the unity vector $\mathbf{1}$. Next, add and subtract the resulting vectors within (2.4) and manipulate to obtain

$$S^2 = \frac{1}{N} \sum_{i=1}^{N} \| (\mathbf{f}_i - \mathbf{a}) - \langle \mathbf{f}_i - \mathbf{a} \rangle \mathbf{1} \|^2 + \frac{1}{N} \sum_{i=1}^{N} (\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2. \tag{2.9}$$

The combination of (2.7) and (2.9) reveals that

$$\sigma^2 = S^2 - \frac{1}{N} \sum_{i=1}^{N} (\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2. \tag{2.10}$$

Hence, the total biased error variance is the ensemble MSE reduced by the ensemble mean square spatial bias. The spatial bias term becomes negligible on large or laterally unbounded domains (see Chapter 3) so that the ensemble mean square error can be used as an approximation for the total biased error variance.

### 2.1.4 Scale Decomposition of Dispersion

A key aspect of this work is to study the impact of scale deficient lateral boundary conditions on LAM ensemble dispersion. To study this effect as a function of wavelength, an expression is needed that links ensemble dispersion to the spectral decomposition of total error variance. Such an expression is obtained directly from (2.6) and (2.10) so that

$$\boxed{D^2 = \sigma^2 + \frac{1}{N} \sum_{i=1}^{N} (\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2 - \left\| \overline{\mathbf{f}} - \mathbf{a} \right\|^2.} \tag{2.11}$$

Using the spectral variance (2.8), Eq. 2.11 shows that ensemble dispersion is determined by the accumulated contributions to error variance at all resolved wavelengths, the ensemble mean square spatial bias, and a reduction from the squared error of the ensemble mean. This expression will be applied in later chapters to evaluate the impact of poorly resolved LBCs on ensemble dispersion by comparing the magnitudes of each term obtained for both global and LAM simulations.

forecast expected value, $E[\mathbf{f}]$, and the analysis expected value, $E[\mathbf{a}]$, in (2.4) and expand so that

$$\begin{aligned} S^2 &= \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{f}_i - E[\mathbf{f}]\|^2 + \|\mathbf{a} - E[\mathbf{a}]\|^2 + \|E[\mathbf{f}] - E[\mathbf{a}]\|^2 \\ &\quad + \frac{2}{N}\sum_{i=1}^{N} \frac{1}{p}(\mathbf{f}_i - E[\mathbf{f}]) \cdot (E[\mathbf{f}] - E[\mathbf{a}]) \\ &\quad - \frac{2}{N}\sum_{i=1}^{N} \frac{1}{p}(\mathbf{f}_i - E[\mathbf{f}]) \cdot (\mathbf{a} - E[\mathbf{a}]). \end{aligned} \tag{2.12}$$

Next, take the expected value $E[S^2]$ on a term-by-term basis, noting that the expected value of the ensemble average is the same as that for any individual vector drawn from a common distribution. Specifically, using properties of expectation with $\mathbf{x} = \mathbf{a}$ and $\mathbf{x} = \mathbf{f}_i$ for any $i = 1, \ldots, N$, $E[\overline{\mathbf{x}}] = E[\frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i] = \frac{1}{N}\sum_{i=1}^{N} E[\mathbf{x}_i] = \frac{N}{N}E[\mathbf{x}] = E[\mathbf{x}]$.

The expected value for each of the first two terms on the right-hand side of (2.12) is, again using $\mathbf{x} = \mathbf{f}_i$ and $\mathbf{x} = \mathbf{a}$,

$$\begin{aligned} E\,\|\mathbf{x} - E[\mathbf{x}]\|^2 &= E\left[\frac{1}{p}(\mathbf{x} - E[\mathbf{x}]) \cdot (\mathbf{x} - E[\mathbf{x}])\right] \\ &= \frac{1}{p}\sum_{k=1}^{p} E[\mathbf{x}_k^2] - \frac{1}{p}\sum_{k=1}^{p} E[\mathbf{x}_k]^2 \\ &= E[\mathbf{x}^2] - E[\mathbf{x}]^2. \end{aligned} \tag{2.13}$$

Note that individual elements of vectors $\mathbf{x}_k$ in the summations above are random variables drawn from the same distribution, so $E[\mathbf{x}_k^2] = E[\mathbf{x}^2]$ and $E[\mathbf{x}_k] = E[\mathbf{x}]$ for all elements $k = 1, \ldots, p$. Since the right-hand side of (2.13) is the statistical definition of variance, $E\,\|\mathbf{f}_i - E[\mathbf{f}]\|^2 \equiv \mathrm{Var}[\mathbf{f}]$ is the climate variance of forecasts and $E\,\|\mathbf{a} - E[\mathbf{a}]\|^2 \equiv \mathrm{Var}[\mathbf{a}]$ is the climate variance of analyses.

Following a similar approach for the third term in (2.12), it is simple to show that $E\,\|E[\mathbf{f}] - E[\mathbf{a}]\|^2 = (E[\mathbf{f}] - E[\mathbf{a}])^2$. The fourth term vanishes after expanding

the dot product and taking its expected value. Finally, the expectation of the last term in (2.12) yields the covariance between $\mathbf{f}$ and $\mathbf{a}$, Cov[$\mathbf{f}, \mathbf{a}$]. Thus,

$$E[S^2] = \text{Var}[\mathbf{f}] + \text{Var}[\mathbf{a}] + (E[\mathbf{f}] - E[\mathbf{a}])^2 - 2\text{Cov}[\mathbf{f}, \mathbf{a}]. \tag{2.14}$$

Forecasts usually are highly correlated with analyses and have small biases early in the forecast period. Consequently, the covariance term is large, resulting in a small value for $S^2$. As initial forecast errors grow with time, forecasts become uncorrelated with analyses and their covariance tends to zero. If forecasts are unbiased and have the same variance as analyses, then (2.14) shows that *the expected value of $S^2$ converges towards twice the climate variance of analyses*; that is, $E[S^2] = 2\text{Var}[\mathbf{a}]$ at saturation. This is the classic result obtained by Leith (1974) and motivates the common practice of measuring ensemble forecast skill relative to the climate variance of analyses, Var[$\mathbf{a}$].

### 2.2.2 Expected Value of Ensemble Dispersion

To obtain the expected value over many cases for ensemble dispersion, add and subtract the forecast expected value in the definition following (2.5) and use (2.13) so that

$$\begin{aligned} E[D^2] &= \frac{1}{N}\sum_{i=1}^{N} E\left\|(\mathbf{f}_i - E[\mathbf{f}]) - (\overline{\mathbf{f}} - E[\mathbf{f}])\right\|^2 \\ &= \frac{1}{N}\sum_{i=1}^{N} E\left\|\mathbf{f}_i - E[\mathbf{f}]\right\|^2 - E\left\|\overline{\mathbf{f}} - E[\mathbf{f}]\right\|^2 \\ &= \text{Var}[\mathbf{f}] - \text{Var}[\overline{\mathbf{f}}] = (1 - 1/N)\text{Var}[\mathbf{f}]. \end{aligned} \tag{2.15}$$

The additional dependence on $N$ results from the fact that $\text{Var}[\mathbf{f}] = N\text{Var}[\overline{\mathbf{f}}]$. This result shows that for large $N$, the expected value of ensemble dispersion is almost the same as the biased forecast variance averaged over many cases.

Following a similar approach as above, the expected value of the squared error of the ensemble mean (bias) is

$$E\left\|\overline{\mathbf{f}} - \mathbf{a}\right\|^2 = (1/N)\text{Var}[\mathbf{f}] + \text{Var}[\mathbf{a}] + (E[\mathbf{f}] - E[\mathbf{a}])^2 - 2\text{Cov}[\mathbf{f}, \mathbf{a}]. \tag{2.16}$$

When equations (2.14) to (2.16) are used to expand each term after taking the expected value of (2.6), it becomes clear that $E[S^2] = 2E[D^2]$. Thus, *the expected value for ensemble mean square error is twice that of ensemble dispersion.* This result holds at any time during the forecast, *when averaged over many cases*, and does not depend on the forecast error being unbiased or saturated with respect to climatology.

### 2.2.3 Expected Value of Total Error Variance

When averaged over many cases, the expected value for $\sigma^2$ is obtained from (2.10) on a term-by-term basis. Applying the same arguments needed to obtain (2.13),

$$\begin{aligned} E[\sigma^2] &= E[S^2] - \frac{1}{N}\sum_{i=1}^{N} E[\langle \mathbf{f}_i \rangle^2] + \frac{2}{N}\sum_{i=1}^{N} E[\langle \mathbf{f}_i \rangle \langle \mathbf{a} \rangle] - \frac{1}{N}\sum_{i=1}^{N} E[\langle \mathbf{a} \rangle^2] \\ &= E[S^2] - E[\langle \mathbf{f} \rangle^2] + 2E[\langle \mathbf{f} \rangle \langle \mathbf{a} \rangle] - E[\langle \mathbf{a} \rangle^2] \\ &= E[S^2] - E[E[\mathbf{f}]^2] + 2E[E[\mathbf{f}]E[\mathbf{a}]] - E[E[\mathbf{a}]^2] \\ &= E[S^2] - E[\mathbf{f}]^2 + 2E[\mathbf{f}]E[\mathbf{a}] - E[\mathbf{a}]^2 \\ &= E[S^2] - (E[\mathbf{f}] - E[\mathbf{a}])^2. \end{aligned} \tag{2.17}$$

Note that in a perfect (unbiased) model simulation, $E[\mathbf{f}] = E[\mathbf{a}]$ so that $E[\sigma^2] = E[S^2]$. Proceeding with the general case, substitute for $E[S^2]$ using (2.14) so that

$$E[\sigma^2] = \mathrm{Var}[\mathbf{f}] + \mathrm{Var}[\mathbf{a}] - 2\mathrm{Cov}[\mathbf{f}, \mathbf{a}]. \tag{2.18}$$

As with the expectation for ensemble MSE, a large covariance term early in the forecast period reduces the magnitude of total error variance. However, forecast errors grow until the covariance becomes zero so that at saturation, $E[\sigma^2] = E[S^2] = 2\mathrm{Var}[\mathbf{a}]$ for unbiased forecasts having the same variance as analyses.

## 2.3 Normalized Error Variances

The basic statistics defined above are simple measures of accuracy that do not account for skill relative to a standard benchmark. Since the theoretical upper bound

for $S^2$ and $\sigma^2$ at error saturation is twice the climate variance of analyses, the climatological variance provides a natural standard of forecast skill.

The total biased error variance $\sigma^2$ was defined previously (Eq. 2.8) as the ensemble average of the grid mean squared difference between the error and the spatial mean error. Similarly, an estimate for the climate variance of analyses is introduced here as

$$\sigma_a^2 = \frac{1}{M} \sum_{j=1}^{M} \|\mathbf{a}_j - \langle \mathbf{a}_j \rangle \mathbf{1}\|^2 , \tag{2.19}$$

where $M$ is the number of independent analyses used to obtain the variance estimate. As with $\sigma^2$, the climate variance of analyses $\sigma_a^2$ has an equivalent spectral form

$$\sigma_a^2 = \frac{1}{M} \sum_{j=1}^{M} \sum_{\kappa=1}^{K-1} 2 \left| F_j^{\mathbf{a}}(\kappa) \right|^2 , \tag{2.20}$$

where $F_j^{\mathbf{a}}(\kappa)$ is the discrete Fourier transform of $\mathbf{a}_j$ and $\kappa = 1, \ldots, K-1$ are the set of Nyquist resolved wavenumbers on the grid.

Using (2.8) and (2.20), the fraction of total variance contributed by wavenumber(s) $1 \leq k_1 \leq k_2 \leq K-1$ is determined using

$$\eta_{\mathbf{a}}(\kappa) = \frac{\frac{1}{N} \sum_{i=1}^{N} \sum_{\kappa=k_1}^{k_2} 2 \left| F_i(\kappa) \right|^2}{\frac{1}{M} \sum_{j=1}^{M} \sum_{\kappa=k_1}^{k_2} 2 \left| F_j^{\mathbf{a}}(\kappa) \right|^2} . \tag{2.21}$$

The normalization applies equally to all wavenumbers, and has a theoretical upper bound of 2. The notation $\tilde{\eta}_{\mathbf{a}}(\kappa)$ is used when statistics are averaged over all $M$ cases such that

$$\tilde{\eta}_{\mathbf{a}}(\kappa) = \frac{\frac{1}{M} \sum_{j=1}^{M} \frac{1}{N} \sum_{i=1}^{N} \sum_{\kappa=k_1}^{k_2} 2 \left| F_i(\kappa) \right|^2}{\frac{1}{M} \sum_{j=1}^{M} \sum_{\kappa=k_1}^{k_2} 2 \left| F_j^{\mathbf{a}}(\kappa) \right|^2} . \tag{2.22}$$

Both (2.21) and (2.22) may be applied to $N$-member ensembles or to output from a single simulation wherein $N = 1$.

The model used here is a non-forced, dissipative vorticity model (Chapter 3). Therefore, $\sigma_a^2$ is calculated as a function of time to account for the steady reduction of climatological variance. Dissipative effects also influence error variances obtained

from ensemble simulations. Normalization by time-dependent climatological variances help mitigate the time-dependency. Similar effects could be noted in the atmosphere as variances are both time and space dependent.

## 2.4 Variance Budget for Limited-Area Domains

The statistics presented in previous sections may be used to validate results obtained from both global and limited-area forecasts. In this section, the spectral form of total error variance is examined in greater detail with specific application to results from limited-area models (LAMs) configured with "one-way" LBCs. Under this configuration, it is argued that LAM forecasts suffer an artificial reduction of total error variance due to inadequate spatial and temporal resolution in the LBCs.

Consider again Eq. (2.8), but explicitly write out the summation over resolved wavenumbers so that

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} 2\Big[\underbrace{|F_i(1)|^2 + |F_i(2)|^2 + \ldots}_{\text{External (LBC) scales}} + \underbrace{|\not{F}_i^{\epsilon}(K-2)|^2 + |\not{F}_i^{\epsilon}(K-1)|^2}_{\text{LAM scales}}\Big]. \qquad (2.23)$$

In this form, several remarks can be made about the source of error variances with regard to wavelength:

- The variance contribution from longer wavelengths is entirely determined by an external model through the LBCs.
- Nonlinear wave interactions generate variance at all scales resolved on the limited-area domain.
- Variance contributed by short waves declines as coarsely resolved LBCs "sweep" through the domain (indicated by the cross to $\epsilon$).
- Linear interpolation of LBCs between available updates acts as a temporal filter that exacerbates the scale-deficiency problem.

Although nonlinear products generate variance at small scales, experimental results in Ch. 5 show that reduction of total error variance by advection of coarsely resolved or infrequently updated LBCs is the dominant effect.

Equation (2.10) shows that a loss of error variance at small scales will cause an artificial reduction of the ensemble MSE ($S^2$) unless offset by an increase in the square of the grid mean error. The effect also is related to ensemble dispersion through (2.11). Specifically, a loss of error variance at small scales will artificially constrain ensemble dispersion unless offset by changes in the bias terms. The magnitude of these effects will be explored in later chapters.

Equation (2.23) provides a mathematical statement of the hypothesis driving this research. The use of external LBCs having lower spatial or temporal resolution than that produced by the limited area model results in an artificial reduction of ensemble MSE and dispersion. This hypothesis is explored using numerical simulations in a perfect model framework. Finally, a new method is proposed to help restore the error variance lost at these smaller scales.

# Chapter 3

# Parameterized P.V. Channel Model

A simplified model is used to help isolate the impact of lateral boundary condition forcing on the dispersion of limited-area ensemble forecasts. Previous predictability studies using limited-area models often present results based on a small number of cases. The models used in such studies usually contain uncontrollable errors such as those associated with numerics, physical parameterizations, and analysis uncertainties. To avoid these problems, experiments in this work are conducted using a two-dimensional parameterized potential vorticity (PPV) channel model. Although simplified, the model remains nonlinear, dispersive, and sensitive to initial conditions.

The analytic form of the PPV model is introduced in section 3.1. Numerical solution and configuration of the model is discussed in section 3.2. A single example case is shown to help demonstrate the model's behavior. Statistical results obtained from 100 independent simulations representing the model's climatology are shown in section 3.4.

## 3.1 Model Dynamics

### 3.1.1 Quasi-geostrophic Potential Vorticity

The PPV model developed for this work is based on the quasi-geostrophic potential vorticity equation. The quasi-geostrophic potential vorticity equation is developed by conducting scale analyses appropriate for mid-latitude, synoptic-scale motions

and assume $\sigma$ changes slowly with height in the mid-troposphere (Bluestein 1992, Fig. 5.3) so that

$$f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial \psi}{\partial p}\right) = -\frac{f_0 R}{\sigma}\frac{\partial T}{\partial p}. \tag{3.4}$$

An increase in the vertical temperature gradient causes this term to have a larger negative value. Since $q$ is conserved, the relative vorticity must become increasingly positive as the vertical temperature gradient increases provided that changes in planetary vorticity $f$ are small. From this viewpoint, it becomes clear that this term represents the potential vorticity response to temperature changes in a baroclinic atmosphere. It is also called the vortex stretching term because it induces changes in vorticity following the adiabatic effects of vertical motions.

The vortex stretching term (3.4) has an important effect on the horizontal scale at which wave motions develop. To see this effect, first note that static stability is related to the Brunt-Väisälä frequency $N = [(g/\theta)\partial\theta/\partial z]^{1/2}$ in a hydrostatic atmosphere since (Bluestein 1992, pg. 363)

$$\sigma = -\frac{1}{\rho\theta}\frac{\partial \theta}{\partial z}\frac{\partial z}{\partial p} = \frac{1}{\rho^2 g}\frac{1}{\theta}\frac{\partial \theta}{\partial z} = \left(\frac{N}{\rho g}\right)^2. \tag{3.5}$$

The magnitude of the mid-tropospheric static stability parameter typically is

$$|\sigma| \sim \frac{-1}{|\rho||\theta|}\frac{|\Delta\theta|}{|\Delta p|} \sim \frac{-(30\text{ K})}{(0.7\text{ kg m}^{-3})(300\text{ K})(-700\text{ hPa})} \sim 2\times 10^{-6}\text{ m}^2\text{Pa}^{-2}\text{s}^{-2}. \tag{3.6}$$

Given this scale for $\sigma$, let $\lambda^2$ define a scale parameter for the vortex stretching (3.4) so that

$$\lambda^2 \equiv \frac{f_0^2}{\sigma(\Delta p)^2} = \frac{f_0^2}{N^2}\frac{(\rho g)^2}{(\Delta p)^2} = \frac{f_0^2}{N^2}\frac{(-\Delta p/\Delta z)^2}{(\Delta p)^2} = \left(\frac{f_0}{N\Delta z}\right)^2. \tag{3.7}$$

This parameter has magnitude

$$|\lambda^2| \sim \frac{|f_0^2|}{|\sigma||\Delta p|^2} \sim \frac{(10^{-4}\text{ s}^{-1})^2}{(2\times 10^{-6}\text{ m}^2\text{Pa}^{-2}\text{s}^{-2})(10^5\text{ Pa})^2} \sim 5\times 10^{-13}\text{ m}^{-2}. \tag{3.8}$$

The importance of the vortex stretching term becomes clear when noting from (3.7) that the parameter $\lambda$ is inversely related to the Rossby radius of deformation $N\Delta z/f_0$ (Holton 1992; Bluestein 1992). The Rossby radius defines the length scale at which vorticity produced through the vortex stretching associated with vertical

motions becomes comparable in magnitude to the vorticity associated with Earth rotation. Therefore, the relative vorticity response to the baroclinic term in (3.2) is most equally partitioned between changes in vortex stretching and Earth rotation at horizontal length scales approaching $\lambda^{-1} = 1400$ km.

### 3.1.2 Parameterized Potential Vorticity Equation

Numerical solution of (3.1) is reasonably simple since it only requires knowledge of the three-dimensional distribution of streamfunction and the stratification of potential temperature. However, the potential vorticity model could be reduced further by eliminating the dependence on vertical temperature profiles. This simplification leads to a parameterized potential vorticity (PPV) model that still represents many desirable characteristics of mid-tropospheric flow at midlatitudes.

#### 3.1.2.1 Pure Barotropic Model

First consider the most stringent simplification. If we assume there is a level of non-divergence where vertical motions reach their maximum (typically around 500 mb), mass continuity requires that $\nabla \cdot \vec{V}_g = -\partial\omega/\partial p = 0$ at that level. Therefore, the vertical velocity ($\omega$) above and below this level are identical so that contributions to vorticity by stretching vanish locally. If we apply this approximation through the depth of the atmosphere, the baroclinic (vortex stretching) term in (3.2) vanishes to yield the barotropic vorticity equation (Holton 1992; Bluestein 1992)

$$\frac{\partial \zeta}{\partial t} + \vec{V}_g \cdot \nabla(\zeta + f) = 0, \tag{3.9}$$

or

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) + \beta \frac{\partial \psi}{\partial x} = 0, \tag{3.10}$$

where $\beta = df/dy$ evaluated at the same reference latitude as $f_0$ and $J(\psi, \zeta)$ is the Jacobian operator

$$J(\psi, \zeta) = \frac{\partial \psi}{\partial x}\frac{\partial \zeta}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial \zeta}{\partial x}. \tag{3.11}$$

The barotropic vorticity equation states that absolute vorticity is conserved following the horizontal nondivergent (geostrophic) wind. Analytic solutions to (3.10) yield the well-known Rossby wave frequency dispersion relation

$$\omega = kU_0 - \frac{\beta k}{k^2 + l^2} \tag{3.12}$$

where $k$ and $l$ are zonal and meridional wavenumbers and $U_0$ is the constant mean zonal wind speed. Rossby waves are dispersive since the phase speed is a function of wavenumber. The zonal phase velocity is $c_x = \omega/k = U_0 - \beta/(k^2 + l^2)$. Thus, Rossby waves propagate westward relative to the mean flow at a phase speed that increases with wavelength. The group velocity in the zonal and meridional directions are, respectively

$$c_{gx} = \frac{\partial \omega}{\partial k} = U_0 + \beta \frac{k^2 - l^2}{(k^2 + l^2)^2} \tag{3.13}$$

and

$$c_{gy} = \frac{\partial \omega}{\partial l} = \frac{2\beta kl}{(k^2 + l^2)^2}. \tag{3.14}$$

These results indicate that the Rossby wave energy propagates toward the east (west) if $k > l$ ($k < l$), with a northerly (southerly) component if $l$ is positive (negative).

The barotropic model (3.10) is the simplest model capable of producing fully developed wave structures consistent with the character of observed atmospheric motions. Indeed, it was the model of choice for the first successful efforts in numerical weather prediction (Charney 1949; Charney et al. 1950; Charney 1951). Unfortunately, the barotropic model has some rather serious limitations because it merely redistributes relative vorticity through advection by the nondivergent wind. Since motions are purely horizontal, it cannot generate or dissipate vorticity through baroclinic effects or other forcing mechanisms. Barotropic model solutions are dominated by retrograding Rossby wave motions, wherein planetary vorticity advection is balanced locally by changes in relative vorticity. In contrast, the longwave components of real atmospheric motions remain quasi-stationary since the planetary vorticity advection is balanced largely by the vortex stretching associated with vertical motions (Holton 1979).

#### 3.1.2.2 Parameterized Potential Vorticity Model

Minimal treatment of baroclinic effects requires a multi-level modeling system. However, Holton (1979, Sec. 8.4.2) suggests that the effect of vorticity stretching may be included in a single-level barotropic model by parameterizing the baroclinic term of $q$ as

$$f_0^2 \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial \psi}{\partial p}\right) = -\lambda^2 \psi \tag{3.15}$$

where the constant $\lambda^{-1}$ was introduced previously as the the Rossby radius of deformation.

The parameterization (3.15) is a differential equation in Sturm-Liouville form that governs the vertical structure of normal modes (Kundu 1990, pg. 500). The zeroth mode is called the barotropic mode in which velocity and density are independent of height and the eigenvalue is the Rossby wave phase speed. For the barotropic mode, $\lambda^2 = 0$ and again we are left with the pure barotropic vorticity model (3.10). The first baroclinic mode is represented by the approximation (3.15). Solutions to this equation imply vertical variations in velocity and density while the eigenvalue depends on stratification through the buoyancy frequency. Therefore, the parameterization (3.15) represents the *first-order* effects of vertical motions in a baroclinic atmosphere.

Given a simplified, yet realistic approximation of baroclinic effects, the parameterized potential vorticity (PPV) model used in this work is defined by letting $\xi \equiv \zeta - \lambda^2\psi$. Then, with the addition of a numerical diffusion term $D$ (Sec. 3.2.3.3), the quasi-geostrophic potential vorticity equation (3.1) becomes

$$\boxed{\frac{\partial \xi}{\partial t} = -J(\psi, \xi) - \beta\frac{\partial \psi}{\partial x} - D.} \tag{3.16}$$

The PPV model is implemented here in a channel model configuration as discussed in Sec. 3.2. Impermeable, free-slip north and south boundaries enforce zero vorticity and constant streamfunction through specifed Dirichlet boundary conditions. Periodic boundary conditions allow uninterrupted flow along the zonal (east/west) direction so that $\psi(x + L_x, y) = \psi(x, y)$ and $\xi(x + L_x, y) = \xi(x, y)$. The PPV model is also implemented on limited-area subdomains using "one-way" Dirchlet boundary conditions for $\psi$ and $\xi$ (see Sec. 3.2.4).

An example case shown later in Sec. 3.3 demonstrates how the baroclinic parameterization is important to help maintain the intensity of the zonal jet along the center of the channel flow. Simulations using (3.16) are more realistic than the pure barotropic model (3.10) because the vorticity balance includes the first order effects of vortex stretching due to vertical motions.

### 3.1.3 Barotropic and Baroclinic Instability

It is of interest to understand how perturbations imposed onto barotropic flow will grow (or damp) with time. Kuo (1949) states that barotropic instability is possible only when the absolute vorticity changes sign at least once in the region of interest (Haltiner and Williams 1980; Kundu 1990; Holton 1992). A velocity profile satisfying this requirement is the Bickley jet (Bickley 1937):

$$U(y) = u_0 \operatorname{sech}^2\left(\frac{ay}{L_y}\right) \tag{3.17}$$

where $u_0$ is the maximum intensity of the jet and $a$ governs the width of the jet relative to $L_y$. The vorticity of the Bickley jet is

$$\varsigma(y) = -dU/dy = (2u_0 a/L_y) \operatorname{sech}^2\left(\frac{ay}{L_y}\right) \tanh\left(\frac{ay}{L_y}\right). \tag{3.18}$$

The Bickley jet profile shown in Fig. 3.1 is used to initialize the mid-latitude westerly jet stream in later numerical simulations (Sec. 3.2.2). The intensity of the initial velocity profile ($u_0 = 60$ ms$^{-1}$) is somewhat stronger than observed for typical middle tropospheric flow. However, after the simulations are integrated forward for 20 days, local velocity maxima become smaller as barotropic instability extracts kinetic energy from the mean flow field. The example case shown later in section 3.3 illustrates this effect.

Kuo (1973) solved the eigenvalue problem for stability of the Bickley jet (Haltiner and Williams 1980, Sec. 4-4). He showed that barotropically unstable waves in a westerly jet have a positive (eastward) phase velocity with a magnitude that is about half the maximum jet speed. Furthermore, the most unstable wavelength decreases

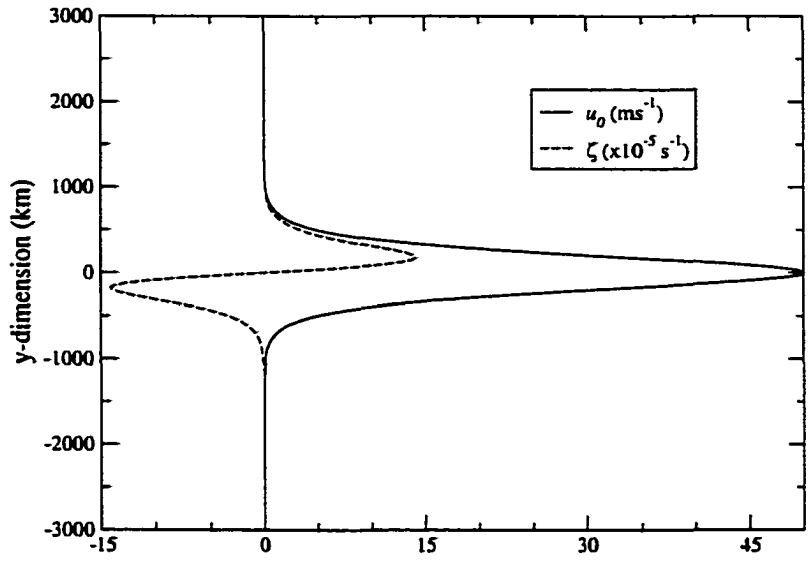



Figure 3.1: Bickley jet profiles for velocity (Eq. 3.17) and vorticity (Eq. 3.18) with $u_0 = 50$ ms$^{-1}$, $L_y = 6000$ km and $a = 11$.

as the jet narrows or becomes more intense. The presence of $\beta$ tends to stabilize mid-latitude westerly jets, but instability must occur for short periods when jets become too sharp (Haltiner and Williams 1980).

Although a simplified single layer model, the baroclinic parameterization in (3.16) acts to modify the barotropic instability mechanism. Holton (1992, pg. 235) describes a special case of the two-layer model for baroclinic instability in which the mean flow is barotropic and there is no vertical shear of the base-state geostrophic wind. For this case, phase speeds of perturbations superimposed on the mean flow are given by

$$c_1 = U_0 - \beta k^{-2} \tag{3.19}$$

and

$$c_2 = U_0 - \beta(k^2 + 2\lambda^2)^{-1}. \tag{3.20}$$

The phase speed $c_1$ represents the dispersion relation for a purely zonal Rossby wave associated with a barotropic perturbation. In contrast, the phase speed $c_2$ represents the dispersion relation for an internal baroclinic Rossby wave. In this case static stability slows the retrogression of the baroclinic mode relative to that

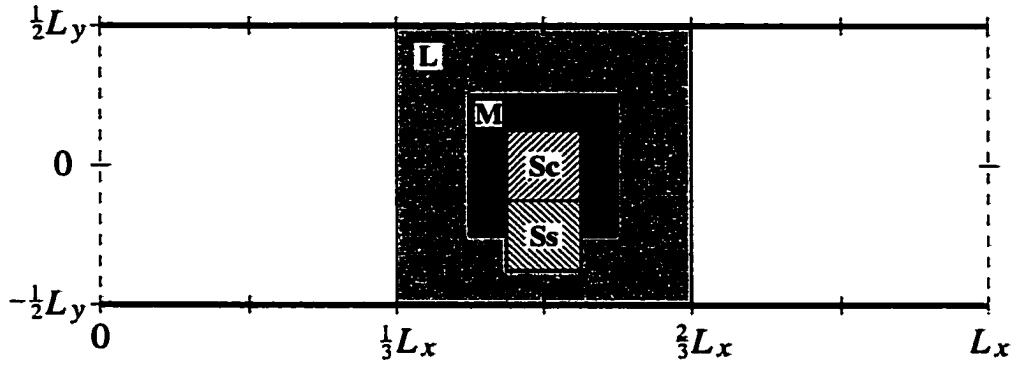



Figure 3.2: Configuration of grids nested within channel domain labeled $L$ (Large), $M$ (Medium), $Sc$ (Small, center), and $Ss$ (Small, south). Each grid is nested only once within the full channel grid; there are no "multiple nests".

| Grid | Dimension | Gridpoints |
|---|---|---|
| Large ($L$) | $(6000 \text{ km})^2$ | $241 \times 241$ |
| Medium ($M$) | $(3000 \text{ km})^2$ | $121 \times 121$ |
| Small, center ($Sc$) | $(1500 \text{ km})^2$ | $61 \times 61$ |
| Small, south ($Ss$) | $(1500 \text{ km})^2$ | $61 \times 61$ |

Table 3.1: Nested grid dimensions for $\Delta x = \Delta y = 25$ km.

periodic east-west boundaries and fourth-order diffusion using two extra "ghost" columns at each end of the channel. Gridpoint locations are determined by $x = i\Delta x$ and $y = j\Delta y - L_y/2$ where $i = 0, \ldots, N_x - 1$ and $j = 0, \ldots, N_y - 1$. Dimensions for the nested domains are listed in Table 3.1.

### 3.2.2 Initialization

All experiments are validated against a high-resolution model "truth" generated on the full periodic channel domain. The model's "climatology" emerges after generating many independent truth simulations. Truth runs begin by specifying a barotropically unstable shear flow as follows:

1. Specify uniformly in $x$ the streamfunction obtained from the Bickley jet profile $U(y)$ shown in Fig. 3.1. Include the base-state zonal wind $U_0$ so that $u = U_0 + U(y) = -\partial\psi/\partial y$. Then, integrate over $y$ and calculate for all $x$

$$\psi_j = \psi_{j-1} - U_0\Delta y - \frac{u_0 L_y}{a}\left[\tanh\left(\frac{ay_j}{L_y}\right) - \tanh\left(\frac{ay_{j-1}}{L_y}\right)\right]$$

   where $u_0$ is the maximum velocity of the jet and $a$ governs the width of the jet relative to $L_y$. At the north and south boundaries, let $\psi(y = \pm L_y/2) = \psi_0 - U_0 y$, where $\psi_0$ is a specified constant for $\psi(y = 0)$.

2. Compute $\xi = \nabla^2\psi - \lambda^2\psi$.

3. Given a set of uniform random numbers $\Pi \in [-1, 1]$, apply perturbations as $\psi = \psi + 0.05\Pi\psi_0$. The perturbations help ensure each case develops uniquely with the onset of barotropic instability.

Restarting and continuing simulations from a saved $\psi$ field simply requires reading $\psi$ from disk and calculating $\xi$ as in step two above.

### 3.2.3 Numerical Methods

After initializing $\xi$, numerical solutions for (3.16) are obtained by computing the tendency of $\xi$, stepping forward in time, and solving the Helmholtz equation $\nabla^2\psi - \lambda^2\psi = \xi$ for $\psi$. Details for each of these steps are provided in the following sections.

#### 3.2.3.1 Spatial Discretization

The potential vorticity tendency is calculated using second order centered finite differences on an un-staggered Cartesian grid. Specifically, the spatial discretization for (3.16) is

$$\frac{\partial\xi}{\partial t} = -J(\psi, \xi) - \frac{\beta}{2\Delta x}(\psi_{i+1,j} - \psi_{i-1,j}) - D. \tag{3.21}$$

Discretization of the numerical diffusion term $D$ is discussed in section 3.2.3.3. The Jacobian operator is discretized using Arakawa's energy and enstrophy conserving scheme (Arakawa 1966; Haltiner and Williams 1980; Durran 1999).

Equation (3.21) shows that $\psi$ must be obtained at each time step to determine the tendency of $\xi$. This is accomplished by solving the Helmholtz equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} - \lambda^2 \psi = \xi(x, y) \tag{3.22}$$

using the Fortran subroutine `hwscrt` (Swarztrauber and Sweet 1975) from the package `FISHPACK`[1] which directly solves a block tridiagonal system of equations using a cyclic reduction algorithm. Dirichlet boundary conditions for $\psi$ are provided for nested-grid simulations. On the full channel domain, Dirichlet conditions are specified on the north and south boundaries as stated in step 1 of Sec. 3.2.2 above, while the east and west boundaries are periodic so that $\psi(x + L_x, y) = \psi(x, y)$.

#### 3.2.3.2 Time Integration

Let $H(\xi) = \partial \xi / \partial t$ represent the tendency of $\xi$, as determined by the right-hand side of (3.21). Time integration begins with a simple Euler forward step

$$\xi^1 = \xi^0 + \Delta t H(\xi^0). \tag{3.23}$$

Subsequent steps are conducted using the second order two step leapfrog-trapezoidal method (Durran 1999)

$$\begin{aligned} \xi^* &= \xi^{n-1} + 2\Delta t H(\xi^n) \\ \xi^{n+1} &= \xi^n + \frac{\Delta t}{2} \left[ H(\xi^*) + H(\xi^n) \right]. \end{aligned} \tag{3.24}$$

The leapfrog-trapezoidal method is slightly dissipative and is computationally stable on a one-dimensional grid if $c\Delta t / \Delta x \leq \sqrt{2}$, where $c$ is the maximum allowable wave speed (Durran 1999). On a two-dimensional grid with $\Delta x = \Delta y$, the stability requirement must be reduced by a factor of $\sqrt{2}$. Thus, for the current configuration, the time step satisfies $c\Delta t / \Delta x \leq 1$. After calculating the maximum allowable $\Delta t$, the step length may be reduced slightly to allow output at even multiples of $\Delta t$.

[1] `FISHPACK` is a shareware source that may be obtained online from `http://www.netlib.org`.

#### 3.2.3.3 Numerical Diffusion

In the real atmosphere, enstrophy accumulates at high wave numbers where it eventually dissipates through viscous effects. During numerical integration of the *inviscid* potential vorticity equation, enstrophy accumulates at the smallest resolved wavelength without dissipation and produces excessive noise. Furthermore, the noise is enhanced by computational errors introduced by finite-difference methods. Therefore, a fourth-order numerical diffusion term $D$ has been added to the simulation model (3.16) where

$$D = -\nu\nabla^4\xi. \tag{3.25}$$

The diffusion term removes shortwave components from the model at a rate determined by the eddy diffusion coefficient $\nu$.

To determine an appropriate value for $\nu$, consider the diffusion equation

$$\frac{\partial\xi}{\partial t} = -\nu\nabla^4\xi. \tag{3.26}$$

Assume solutions of the form $\xi = A(t)\exp[i(kx+ly)]$, where $k$ and $l$ are zonal and meridional wavenumbers. After making the substitution and canceling common exponential terms,

$$\frac{dA}{dt} = -\nu A(k^4+l^4). \tag{3.27}$$

Time integration yields the analytic solution $A = A_0\exp[-\nu(k^4+l^4)t]$. For discrete wavenumbers resolved on a grid, $k = 2\pi/p\Delta x$ and $l = 2\pi/p\Delta y$ where $p = 2,\ldots,(N_x,N_y)$. Then, using discrete time intervals $t = N_t\Delta t$ and $\Delta x = \Delta y$,

$$A = A_0\exp[-2\nu(2\pi/p\Delta x)^4 N_t\Delta t]. \tag{3.28}$$

For the shortest resolved wavelength, $p = 2$ and the diffusion coefficient

$$\nu = \frac{1}{2N_t\Delta t}\left(\frac{\Delta x}{\pi}\right)^4 \tag{3.29}$$

produces an $e$-fold reduction in the amplitude of the $2\Delta x$ wave over $N_t$ time intervals of the analytic solution. For example, with $\Delta x = 25$ km, $\Delta t = \Delta x/c$, where $c = 50$ ms$^{-1}$ and $N_t = 4$ time steps, (3.29) yields $\nu = 1.00\times10^{12}$ m$^4$s$^{-1}$. Note that

the actual $e$-folding time realized during model integration depends on the specific methods of spatial and temporal discretization (Durran 1999; Xue 2000) and is generally shorter than that obtained analytically using (3.28) and (3.29).

### 3.2.4 Lateral Boundary Conditions

Limited-area models employ completely artificial lateral boundaries having no physical counterpart in the real atmosphere. To minimize errors introduced by the use of such boundaries, schemes designed to impose time-dependent lateral boundary conditions must allow information from outside the domain to enter without dissipation. Furthermore, disturbances traveling outward from the domain interior must be allowed to pass through the boundary without reflection. Boundary conditions designed to minimize these spurious wave reflections are called open, wave-permeable, or radiation boundary conditions (Sommerfeld 1949; Orlanski 1976; Givoli 1991; Durran et al. 1993; Staniforth 1997).

Among this class of boundary formulations for limited-area grid point models are the "one-way" and "two-way" approaches. In the two-way approach, solutions are obtained simultaneously and interactively updated on both the external and limited-area grids. In the one-way approach, a previously and independently run external model solution provides the boundary information required to obtain the nested model solution. Since the one-way approach is more commonly used in operational forecast configurations, it is the method used for the PPV model and whose impact is examined in this work.

Most one-way LBC formulations are not well posed mathematically because more conditions are provided by the external model than are actually necessary on the limited-area grid (Williamson and Browning 1974; Perkey and Kreitzberg 1976; Oliger and Sundström 1978; Davies 1983; Staniforth 1997). When the problem is not well posed, variables obtained just inside the domain boundary will generally contradict those values specified *a priori* along the boundary by the external model. Consequently, the solution exhibits a discontinuity at the boundary and admits spurious wave energy. Schemes which over-specify the boundary conditions often apply a "sponge zone" around the edge of the limited domain to damp the spurious waves. (Williamson and Browning 1974; Perkey and

| Parameter | Symbol | Value |
|---|---|---|
| Reference latitude | $\phi_0$ | 45 deg N |
| Earth's mean radius | $a_e$ | 6367650 m |
| Earth's angular velocity | $\Omega$ | $7.292 \times 10^{-5}$ ms$^{-1}$ |
| Coriolis constant | $f_0$ | $2\Omega \sin \phi_0$ |
| Meridional gradient of $f$ | $\beta$ | $2\Omega \cos \phi_0 / a_e$ |
| Streamfunction constant | $\psi_0$ | $5.40 \times 10^8$ m$^2$s$^{-1}$ |
| Inverse of Rossby Radius | $\lambda$ | $7.071 \times 10^{-7}$ m$^{-1}$ |
| Base state zonal wind speed | $U_0$ | 12 ms$^{-1}$ |
| Maximum allowable wave speed | $c$ | 50 ms$^{-1}$ |
| Numerical diffusion period | $N_t$ | 4 time steps |
| Domain length along $x$ | $L_x$ | 18,000 km |
| Domain length along $y$ | $L_y$ | 6,000 km |
| Grid resolution for *truth* | $\Delta x, \Delta y$ | 25 km |

Table 3.2: Constant parameters for the PPV model (3.16). Dimensions and related constants for nested grids were specified previously in Table 3.1.

## 3.3 Example Case

An example simulation highlights characteristics of the fully nonlinear flow produced by the PPV model (3.16). The case is initialized using a perturbed, barotropically unstable shear flow as described in section 3.2.2. The PPV model is then run for 20 days to "spin up" large-scale vortices and ensure the development of a continuous energy spectrum. Following this pre-integration period, time is reset to zero before running the example case for an additional 20 days. Parallel simulations are run from the same initial state using barotropic ($\lambda = 0$) or non-diffusive ($\nu = 0$) models to compare the effects of the vorticity parameterization and numerical diffusion. Model dimensions and constants are fixed according to Table 3.2. Note that $U_0$ and $\psi_0$ are arbitrary constants selected to emulate observed characteristics of the atmosphere along the 500 hPa pressure surface. In particular, $\psi_0$ is selected to provide appropriate values of geopoential height $\Phi$ under the approximation $\Phi = \psi f_0$.

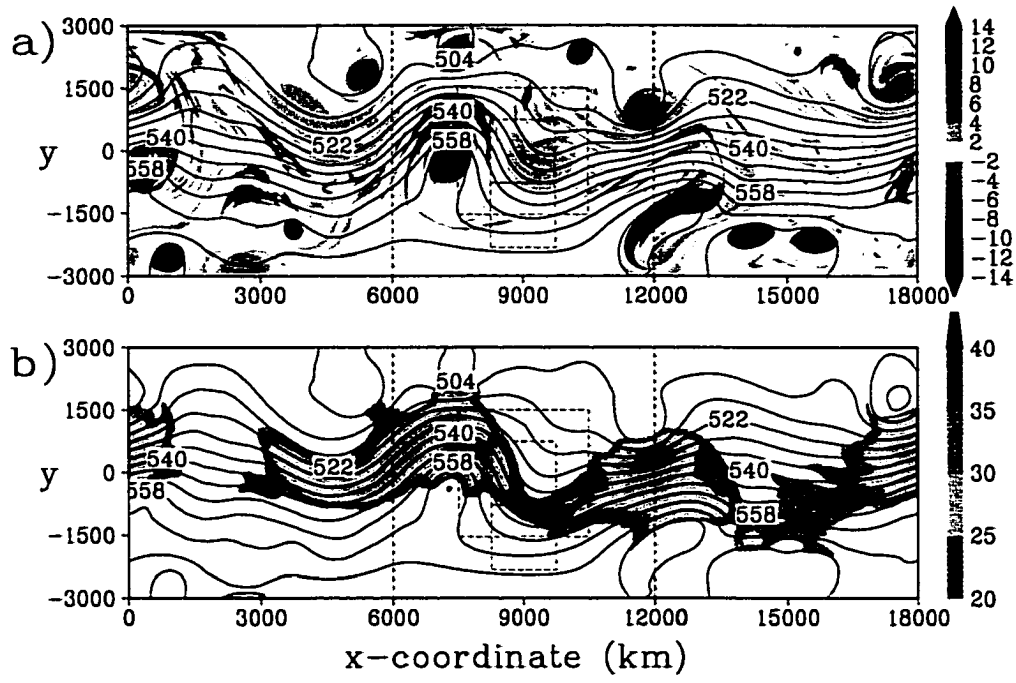



Figure 3.3: Barotropic model simulation ($\lambda = 0$) 20 days after initialization with the Bickley jet profile (Fig. 3.1). Contours show $\psi \times 10^6$ m$^2$s$^{-1}$. Shading in (a) indicates relative vorticity greater than $\pm 2 \times 10^{-5}$ s$^{-1}$ and shading in (b) shows velocity greater than 20 ms$^{-1}$. Nested grid boxes are outlined with dashed gray lines for later reference to limited-area model simulations.

### 3.3.1 Impact of Parameterization

Two cases were initialized with the same perturbed shear flow and integrated for the 20-day spin up period to examine the impact of the vortex stretching parameterization. One case was run using the purely barotropic model ($\lambda = 0$) while the other retains the parameterization term with $\lambda = 7.071 \times 10^{-7}$ m$^{-1}$. Results are shown in Figs. 3.3 and 3.4, respectively.

The relative vorticity fields from each simulation reveal distinctly different patterns. Vorticity in the barotropic case (Fig. 3.3a) generally is more concentrated near the centers of troughs and ridges when compared to the PPV case (Fig. 3.4a). For example, vorticity associated with the ridge located near $x = 7000$ km is much stronger in the barotropic case than its counterpart in the PPV model. As discussed

The time evolution of this case is examined next by continuing in Fig. 3.5 the PPV model simulation from the state shown in Fig. 3.4. The most prominent wave pattern for this case is centered near $x = 6000$ km and has a wavelength of about 4000 km, or slightly longer than the width of the largest sub-domain outlined on the figure. This longwave pattern moves slowly eastward during the 4-day simulation while shortwave features of varying intensity propagate rapidly through the flow. This case reveals that nested-model simulations will require accurate simulation of large-scale changes through lateral boundary conditions while predicting changes in smaller scale features evolving within the confines of the nested domain. Indeed, this is a general requirement for any limited-area model used in weather prediction.

### 3.3.2 Energetics of Example Case

Domain average energy and enstrophy curves for the entire 15-day integration of the example case are shown in Fig. 3.6. An average wavelength

$$\overline{L} = 2\pi\sqrt{\frac{\langle u^2 + v^2 \rangle}{\langle \zeta^2 \rangle}} \tag{3.32}$$

is also shown, where $\langle \cdot \rangle$ denotes the domain average (Durran 1999, pg. 165). Comparison of energetics from three parallel simulations help illustrate the difference between the barotropic and PPV models and the effects of numerical diffusion (Fig. 3.6).

Figure 3.6a shows that domain averaged kinetic energy differences are indistinct for simulations with and without numerical diffusion. This is not surprising since numerical diffusion is effective at the smallest scales while, as shown next in section 3.4, the majority of kinetic energy variance is contained in long wavelengths. Kinetic energy is more strongly affected by the use of the vortex stretching parameterization. Specifically, kinetic energy is nearly conserved throughout the barotropic model simulations but decreases with time when the parameterization is applied. A possible explanation for the decline in energy is that the vortex stretching parameterization helps reduce the amplitude of retrograding Rossby waves, thereby decreasing the contribution to energy from the meridional velocity component.

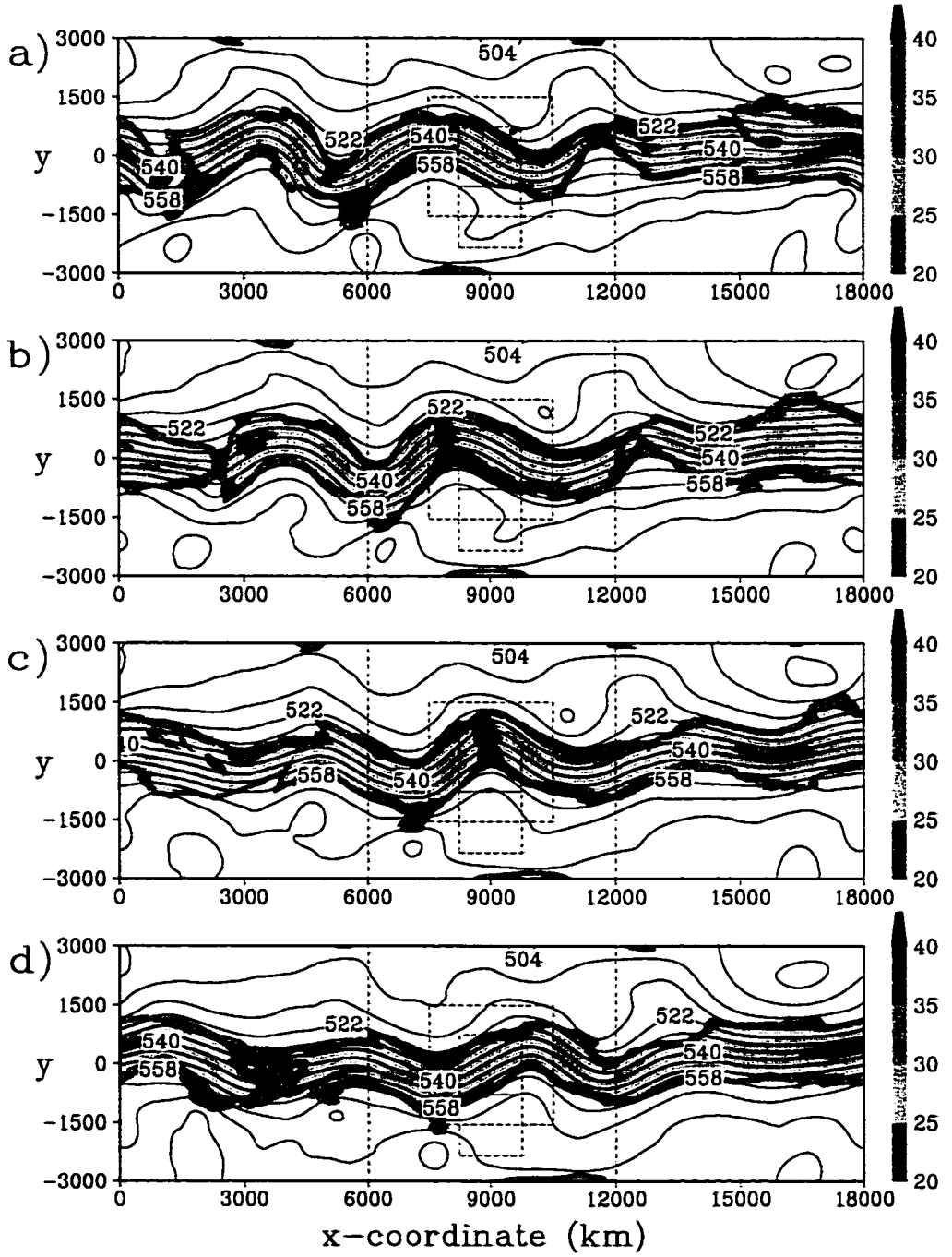



Figure 3.5: Continuation of the case example from Fig. 3.4b for (a) 1 day, (b) 2 days (c) 3 days, and (d) 4 days. Contours and shading are as in Fig. 3.4.

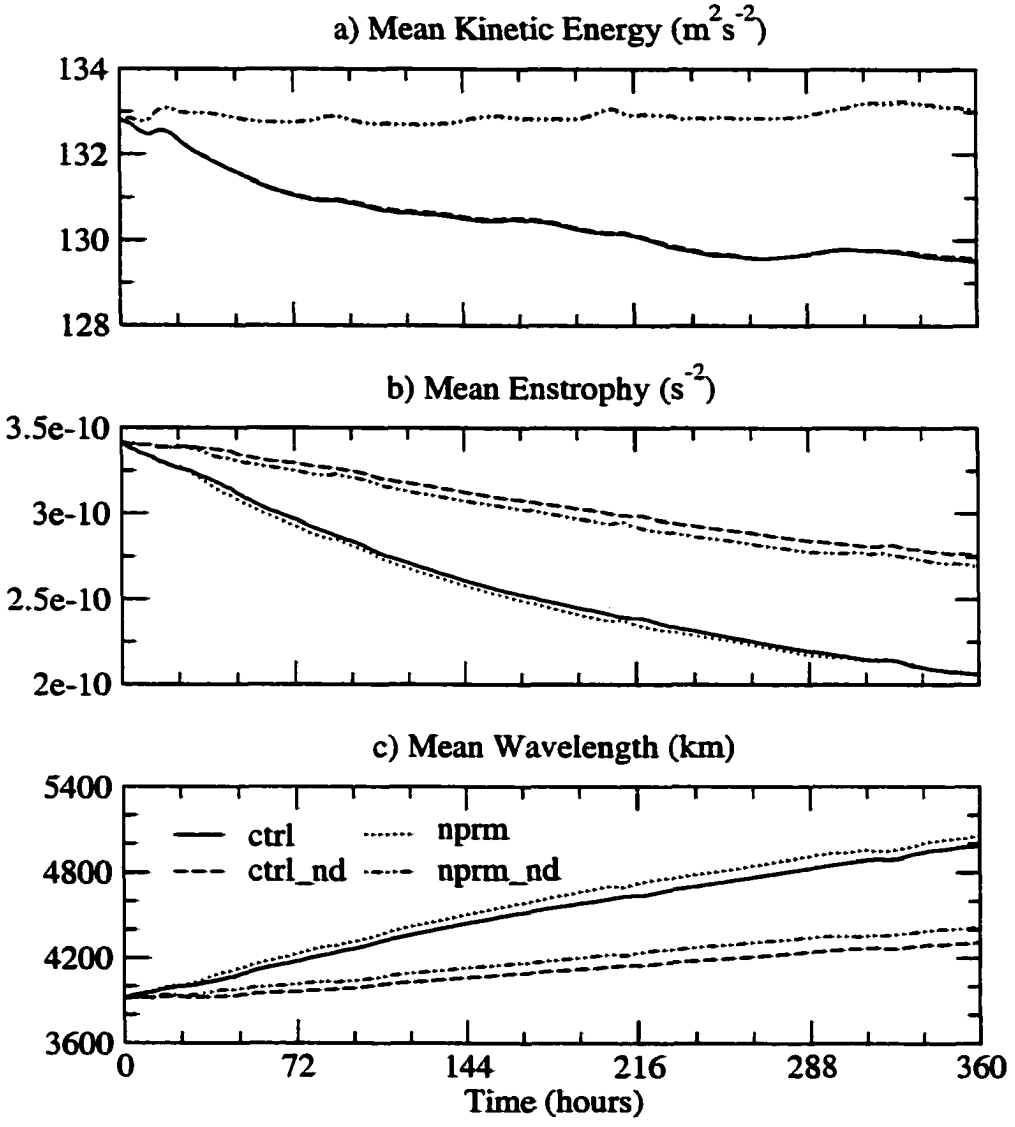



Figure 3.6: Grid-mean (a) kinetic energy [$m^2s^{-2}$], (b) enstrophy [$s^{-2}$] and (c) wavelength [km] for 15-day simulations starting from the state shown in Fig. 3.4. Simulations labeled "ctrl" have $\lambda = 7.071 \times 10^{-7}$ $m^{-1}$ while those labeled "nprm" have $\lambda = 0$. Numerical diffusion is excluded ($\nu = 0$) for simulations labeled "nd".

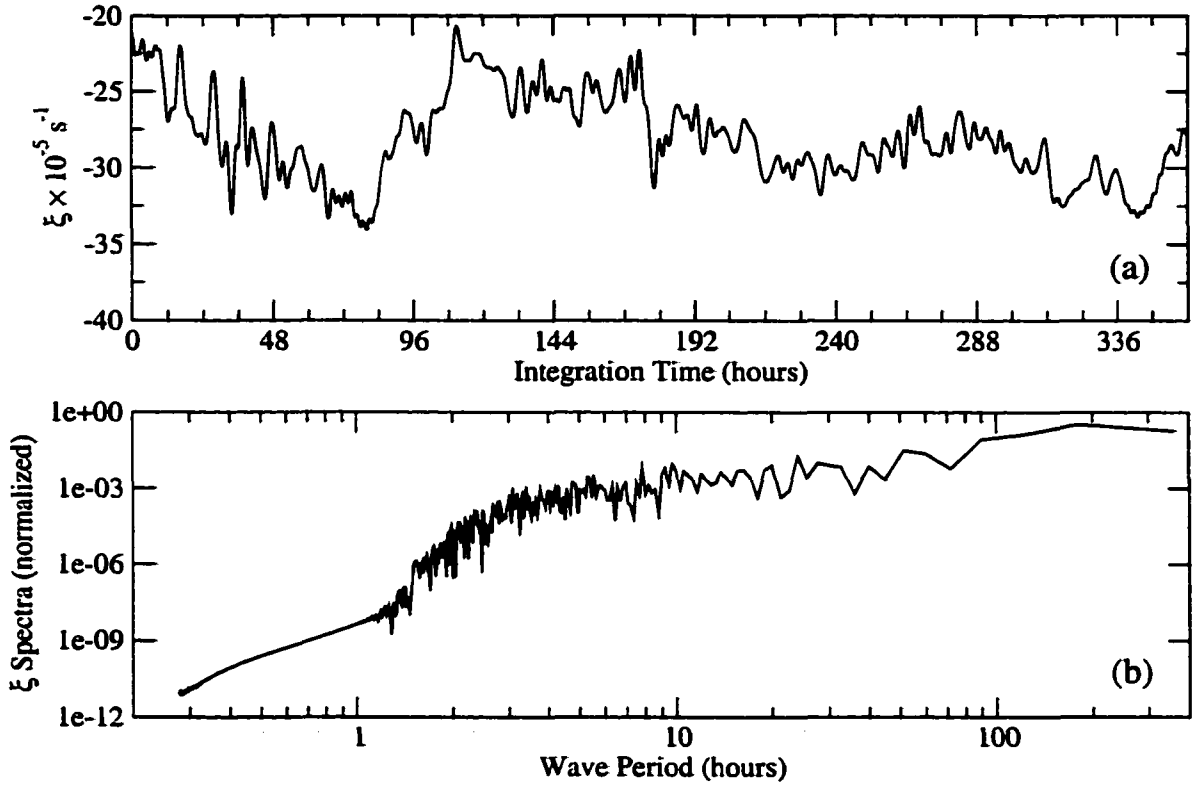



Figure 3.7: a) Time series of $\xi$ obtained at the center point of the example case. b) Frequency spectrum of the $\xi$ time series, normalized by the sum over all wave frequencies.

the total variance. These results suggest that LBC nesting intervals of 3-6 hours are needed to sample wave structures in this part of the flow. If the spectrum were analyzed away from the center of the channel where the flow is less variable, dominant wave periods would be longer. Example cases in Ch. 5 help illustrate the nature of error growth associated with various nested domain sizes and LBC nesting intervals.

## 3.4 Model Climatology

Climatology is the standard benchmark used to measure limits of skill in numerical weather prediction. Since most simulations conducted here are compared to a model

generated truth, it is important to establish the climatological characteristics of the PPV model. For this work, "climatology" refers collectively to the statistical moments obtained from one hundred independent cases produced in the same way as the example case discussed above. Specifically, model-generated "truth" states are initialized with uniquely perturbed shear flows (Sec. 3.2.2) and run for 20 days on the full channel domain. After this spin-up period, time is reset to zero and the truth runs are continued up to 15 days further.

The variance spectrum for the PPV model is shown in Fig. 3.8, normalized by the total climate variance for each variable. Specifically, the fraction of total variance contributed by each wavenumber $\kappa$ is determined using

$$\chi_{\mathbf{a}}(\kappa) = \frac{\frac{1}{M}\sum_{j=1}^{M} 2\left|F_j^{\mathbf{a}}(\kappa)\right|^2}{\sigma_a^2}, \tag{3.33}$$

where $M = 100$ independent cases and $\sigma_a^2$ is given by (2.20). Kinetic energy spectra are easily obtained from the variance spectra of velocity components by calculating $\frac{1}{2}[\chi_{\mathbf{u}}(\kappa) + \chi_{\mathbf{v}}(\kappa)]$.

The spectrum is continuous across all wavelengths, a feature indicative of any nonlinear chaotic system. The kinetic energy spectrum has an $\ell^3$ slope[2] at wavelengths $\ell$ between 500 and 3000 km. The $\ell^3$ ($\kappa^{-3}$) slope for this part of the spectrum is consistent with the downscale enstrophy cascade expected for geostrophic turbulence. The spectrum shows no evidence of a $\ell^{5/3}$ ($\kappa^{-5/3}$) slope accompanying an upscale energy cascade. However, the energy cascade is not expected since there are no sources for injecting energy into this model.

The spectra shown in Fig. 3.8 for the PPV model reveal that more than 99% (59%) of the climate variance of $\psi$ ($\zeta$) is contained in waves longer than about 1500 km. About 90% of the climate variance for $\zeta$ is represented by waves longer than 300 km. The smallest nested domain used in this work is 1500 km$^2$, thereby emphasizing the importance of providing accurate contributions to variance at large scales through the LBCs. Vukicevic and Errico (1990) made a similar statement.

[2]Since wavenumber $\kappa = 2\pi/\ell$, a $\kappa^{-3}$ slope in $(\log\chi, \log\kappa)$ coordinates has an $\ell^3$ slope in $(\log\chi, \log\ell)$ coordinates because $\log\chi = \log\kappa^{-3} = -3\log 2\pi + \log\ell^3$. Note that the abscissa in Fig. 3.8 is inverted to enable direct comparison with traditional $(\log\chi, \log\kappa)$ coordinates.

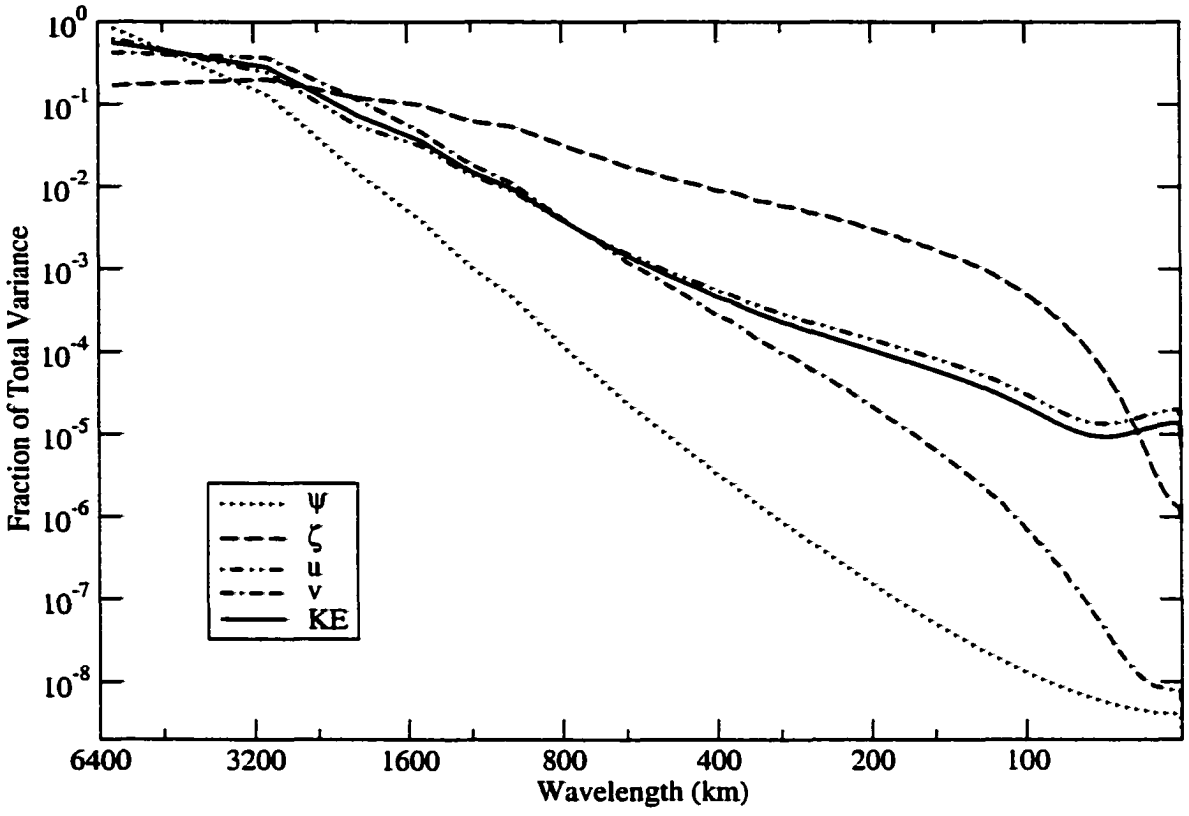



Figure 3.8: Dimensionless fraction of total variance contributed by different wavelengths (Eq. 3.33) for streamfunction $\psi$, vorticity $\zeta$, $u$ and $v$ wind components, and kinetic energy averaged over 100 independent samples.

Note that specific values for the portion of accumlated variance above specified wavelengths are model dependent.

When considering error growth on limited-area domains, it must be noted that variance calculations depend on the domain size *and* location within the flow. To illustrate this point, climate variances are computed on four subgrids (Fig. 3.2) extracted from the whole domain climatology simulations. These variances are listed in Table 3.3. Values in parentheses show the amount of variance contributed by waves resolved within the nested domains relative to the total variance on the full channel domain. Variances for velocity components are not shown in the table since they are represented by kinetic energy. Results shown in Fig. 3.8 and Table 3.3 reveal that vorticity is more evenly distributed across the spectrum than streamfunction

| Domain | $\sigma_a^2(\psi)$ [$m^4s^{-2}$] | $\sigma_a^2(\zeta)$ [$s^{-2}$] | $\sigma_a^2$(k.e.) [$m^2s^{-2}$] |
|---|---|---|---|
| Global | $6.455\times10^{13}$ | $6.847\times10^{-10}$ | 59.88 |
| Large | $5.491\times10^{13}$ (.851) | $7.148\times10^{-10}$ (1.04) | 59.94 (1.00) |
| Medium | $1.660\times10^{13}$ (.257) | $6.528\times10^{-10}$ (.953) | 34.27 (.572) |
| Small, center | $1.844\times10^{12}$ (.029) | $3.828\times10^{-10}$ (.559) | 7.50 (.125) |
| Small, south | $1.445\times10^{12}$ (.022) | $3.466\times10^{-10}$ (.506) | 5.16 (.086) |

Table 3.3: Total variance on each nested domain for streamfunction $\psi$, vorticity $\zeta$, and kinetic energy averaged over 100 cases from climatology. Values in parentheses are the ratio of variances on each nested domain to the variance on the full channel (global) domain.

or kinetic energy. In fact, even the medium size domain (3000 $km^2$) accounts for a greater proportion of total variance in $\zeta$ than does the largest subgrid (6000 $km^2$). The variance in streamfunction is almost negligible at scales resolved within the smallest domains (1500 $km^2$). Variances for the zonal wind component $u$ are inflated at small scales because of the rapid increase in flow speed near the frictionless north and south boundaries of the channel. Vorticity carries the greatest power at small scales, so it will be used as the variable of choice for examining spectral error variance growth rates in later chapters.

The PPV model has the advantage of simplicity but transfers energy more slowly across the spectrum compared to true atmospheric flows. The PPV model lacks forcing due to factors such as moist physics, radiation, thermal gradients, or interaction with the lower boundary. Therefore, the climate variance for the PPV model is not stationary, but tends lower with time. For these reasons, climate variances are computed as a function of time and used to normalize all error variances presented in later chapters. In spite of these limitations, normalized error growth behaviors presented in the following chapters are at least qualitatively consistent with results obtained from more realistic models.

# Chapter 4

# Global Ensemble Simulations

Ensemble simulations using the "global" periodic channel configuration are needed to establish benchmark error growth characteristics for the PPV model. An example case is presented to help illustrate the method used for generating initial condition perturbations and the subsequent behavior of perturbation error growth. The statistics defined in Ch. 2 are then used to analyze results from one-hundred independent 10-member ensemble simulations.

## 4.1 Ensemble Example Case

### 4.1.1 Initial Condition Perturbations

Recall from the previous chapter that a model-generated climatology has been produced consisting of 100 independent cases. This climatology is used to obtain initial conditions and perturbations following the method used by Schubert and Suarez (1989). Specifically, two unique states are randomly selected from the model's climatology. Perturbations are then formed by scaling the difference between the two samples by a factor of 0.10. Finally, the perturbation field is added to an initial condition field represented by any one of the climatological states. This perturbation procedure is repeated ten times for a given initial field to create the starting condition for 10-member ensemble simulations. One-hundred independent 10-member ensemble simulations are constructed by assigning perturbations to initial conditions given by each of the available climatological cases.

An example initial perturbation for one ensemble member, and its subsequent evolution, is shown in Fig. 4.1. The unperturbed control simulation for this case is the same as that shown previously in Fig. 3.5. As designed, the streamfunction perturbation field generally features smooth, small amplitude, synoptic-scale structures with wavelengths of about 3000 to 6000 km (Fig. 4.1a). The streamfunction perturbations evolve slowly in accordance with barotropic dynamics. However, the vorticity field does evolve more rapidly, especially at small scales.

The perturbation method used here is rather simplistic, but it effectively creates error growth and dispersion among the ensemble members. The error growth rates are slow because the model does not contain a fast mode of instability and also because the initial perturbations have rather small amplitudes. Test simulations run using perturbations imposed after scaling difference fields by 0.20 yielded little change in the qualitative behavior of the ensemble. Indeed, for a given scale of motion, error growth rates should be the same regardless of the amplitude of the initial perturbation applied to that scale. Hence, it is not expected that the choice of perturbation method will prevent conclusions from being applied more generally, even if only in a qualitative sense.

### 4.1.2 Example Ensemble Dispersion

Time series of $\psi$ and $\xi$ for individual ensemble members are shown in Fig. 4.2 to help visualize the ensemble dispersion for this example case. Wiggles in the time series for $\psi$ around 18 and 36 hours reveal successive short waves passing by this location that are highly correlated among the individual ensemble members. As the large-scale ridge moves into position over this grid location around 72 hours, the smaller scale details have begun to decorrelate. The change in the large-scale pattern from ridge to trough is handled well by all the ensemble members. This indicates a slow evolution of $\psi$ perturbations at large scales.

The time series for $\xi$ shown in Fig. 4.2b reveals more rapid evolution of perturbation error growth. The $\xi$ time series begin to decorrelate after about 12 hours, and appear completely uncorrelated by around 36 hours. Notice the time-scale for Fig. 4.2b has been truncated to 72 hours because of the rapid dispersion of $\xi$. Beyond this time, the $\xi$ series all continue to follow the long-period pattern associated

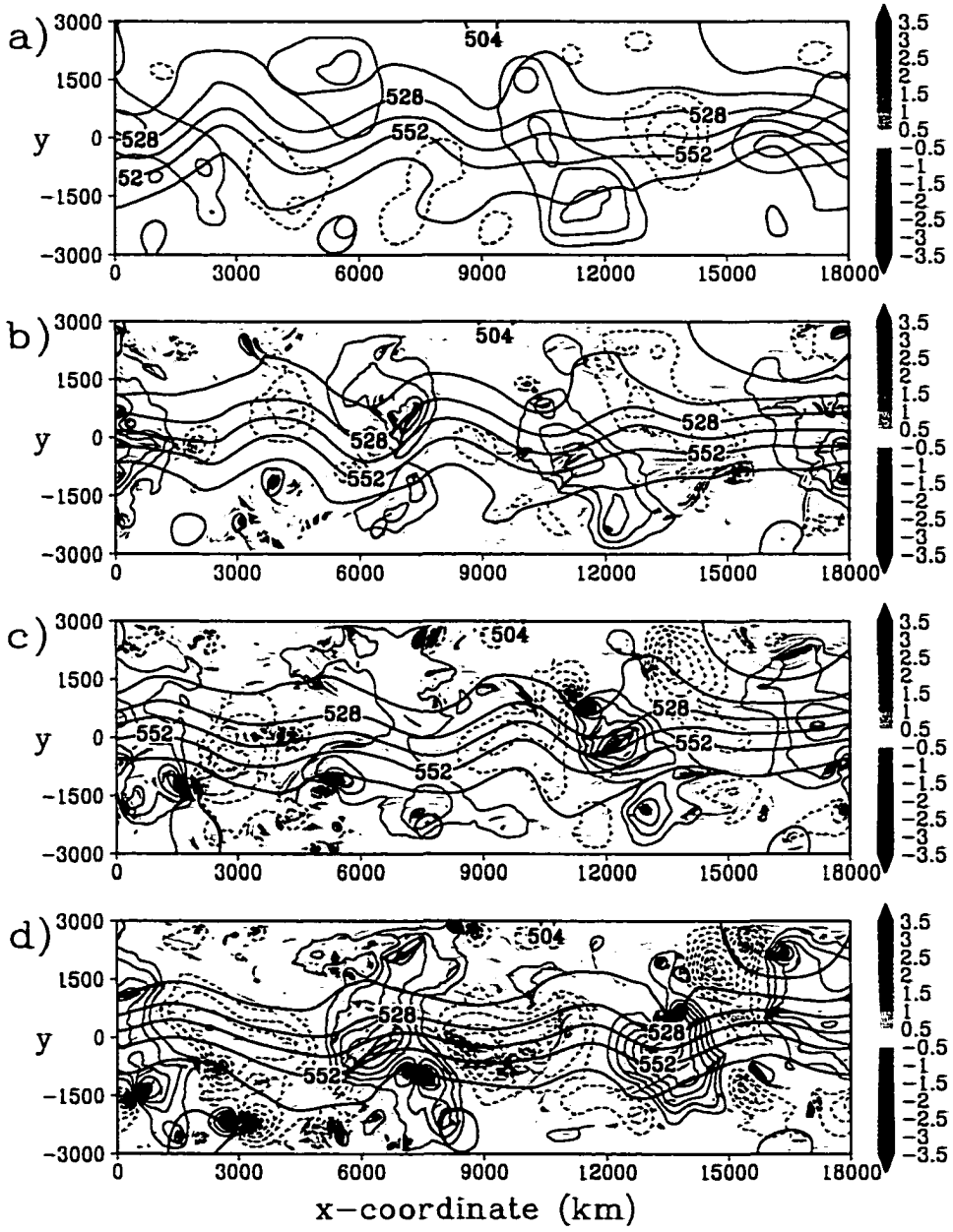



Figure 4.1: Example case showing initial perturbations and subsequent error growth for (a) day 0, (b) day 2, (c) day 4, and (d) day 6. Smooth, labeled curves are streamlines from the control simulation drawn at $12 \times 10^6$ $m^2s^{-1}$ intervals. Vorticity perturbations ($\zeta \times 10^{-5}$ $s^{-1}$) are shaded while (positive/negative) streamfunction perturbations are shown with (solid/dashed) contours at $\pm 5 \times 10^5$ $m^2s^{-1}$ intervals, or ~0.1% (zero line omitted).

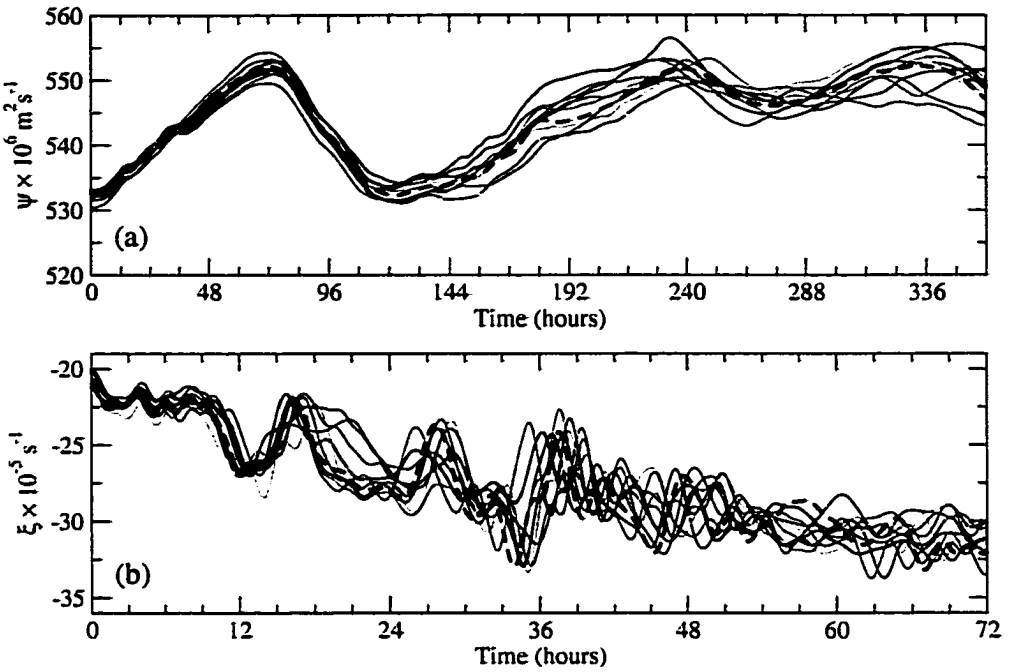



Figure 4.2: Time series of (a) $\psi$ and (b) $\xi$ drawn from individual ensemble members at a single grid point at the center of the channel domain. The unperturbed control simulation is indicated with the heavy dashed line.

with the movement of the large-scale $\psi$ field. Thus, although $\xi$ perturbations grow rapidly at small scales, they also evolve more slowly at large-scales.

Ensemble dispersion was defined previously as a scalar statistic following Eq. 2.5. A slight alteration to this definition yields a vector form of dispersion, denoted $\mathbf{d}^2$, that may be viewed as a two-dimensional contour plot. Specifically, at each grid point $p$ in the domain, field dispersion for an $N$-member ensemble is given by

$$\mathbf{d}_p^2 = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{f}_{p,i} - \overline{\mathbf{f}_p})^2. \tag{4.1}$$

Figure 4.3 shows the results of this calculation for the example ensemble case. Dispersion for the streamfunction is again notably slow. However, vorticity dispersion grows very quickly in the first two days, and gradually organizes with time toward larger scales. Dispersion varies widely in space and time for both variables, a characteristic that is common to all ensemble systems.

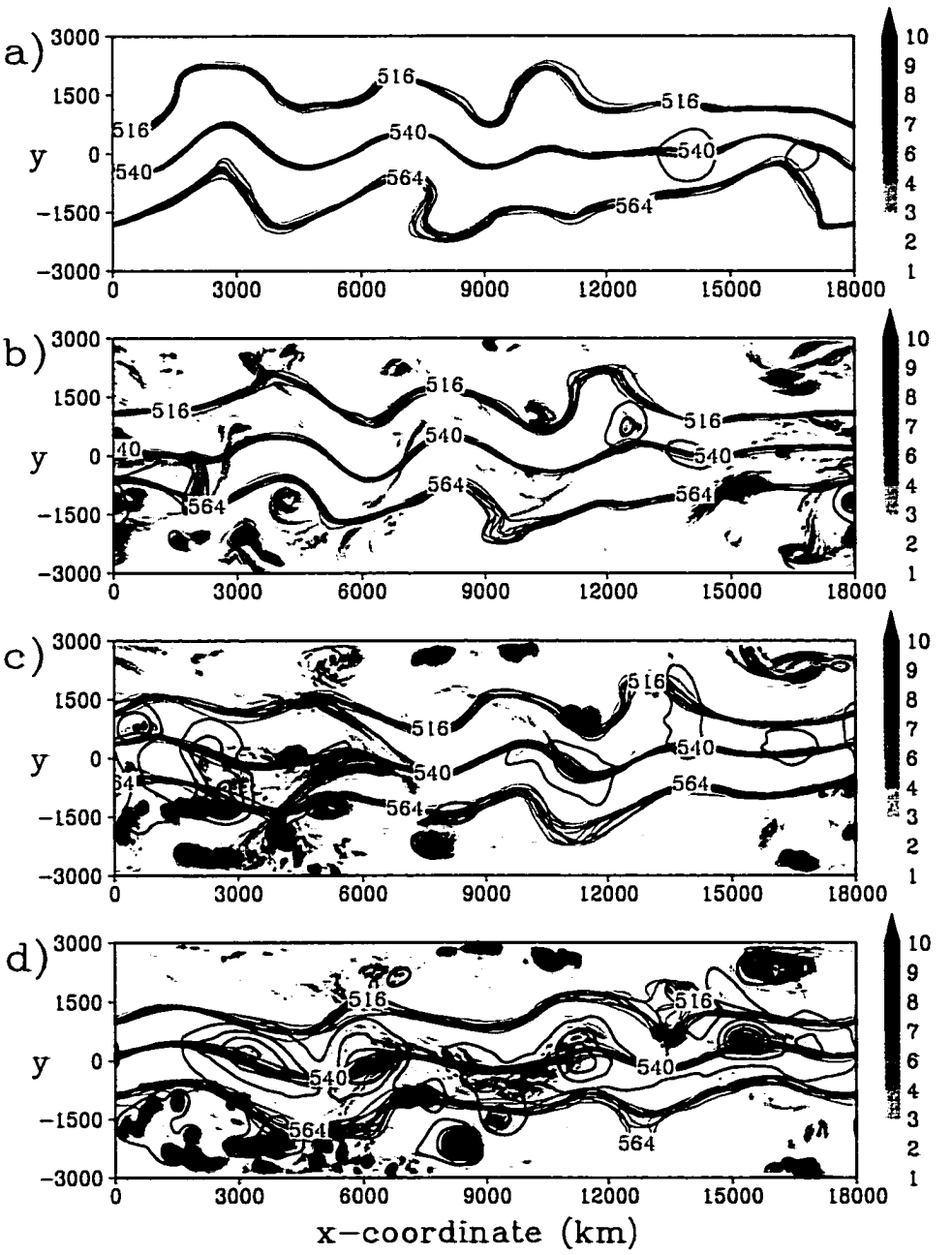



Figure 4.3: Example case showing ensemble dispersion (Eq. 4.1) for (a) day 0, (b) day 2, (c) day 4, and (d) day 6. A "spaghetti" plot drawn with solid black lines shows the (516,540,564)$\times 10^6$ $m^2 s^{-1}$ streamlines from each of the 10 ensemble members. Vorticity dispersion is shaded ($\times 10^{-10}$ $s^{-2}$) while streamfunction dispersion is shown with solid blue contours at $2 \times 10^{12}$ $m^4 s^{-2}$ intervals.

## 4.2 Normalized Error Variance Spectra

Results from one-hundred independent 10-member global ensembles are examined as a function of wavelength using the spectral form of normalized error variance (2.21). The statistics are calculated hourly for each ensemble, then averaged over all 100 cases. Although simulations are conducted on the full periodic channel domain, variance spectra are calculated separately for each of the four subdomains outlined previously in Fig. 3.2. These calculations enable direct comparisons to results from LAM simulations conducted in chapters 5 and 6. Statistics from the full channel domain are not shown since they appear almost identical to those obtained on the large (6000 km$^2$) domain. All statistics are presented in terms of vorticity since error growth for $\zeta$ is more rapid than for $\psi$. Furthermore, the vorticity spectrum contains more power at small scales than other variables (Fig. 3.8), which enables greater resolution of small-scale LBC effects as discussed in Sec. 2.4.

Results shown in Fig. 4.4 reveal, as expected, that error growth due to initial perturbations is most rapid at short wavelengths. Error variance contributions at wavelengths $50 \leq \lambda \leq 250$ km approach the expected maximum value of 2 within about 48 hours on all subdomains. In general, error growth rates become progressively slower at longer wavelengths. However, an interesting behavior is seen in the 1500 km wavelength band when compared on the different subdomains. On the large and medium grids (Fig. 4.4a,b), the normalized error variance for this band reaches 1.5 by the end of the simulations. Corresponding values are about 1.2 on the small, centered grid and 1.8 on the small, southern grid (Fig. 4.4c,d). This result suggests that synoptic-scale waves are less predictable when displaced outside the central shear zone. To see this, note the enhanced spread among the $516\times10^6$ m$^2$s$^{-1}$ and $564\times10^6$ m$^2$s$^{-1}$ "spaghetti" streamlines near troughs and ridges in Fig. 4.3 compared to the central $540\times10^6$ m$^2$s$^{-1}$ streamlines. The large and medium domains encompass broader regions that include areas outside the central shear zone, and therefore have error variance values at 1500 km wavelengths that are between the two extremes.

Sample size becomes important as the number of available gridpoints decreases on smaller domains. With fewer gridpoints, less coefficients are obtained in the spectral variance calculations. Consequently, error growth curves are quite erratic

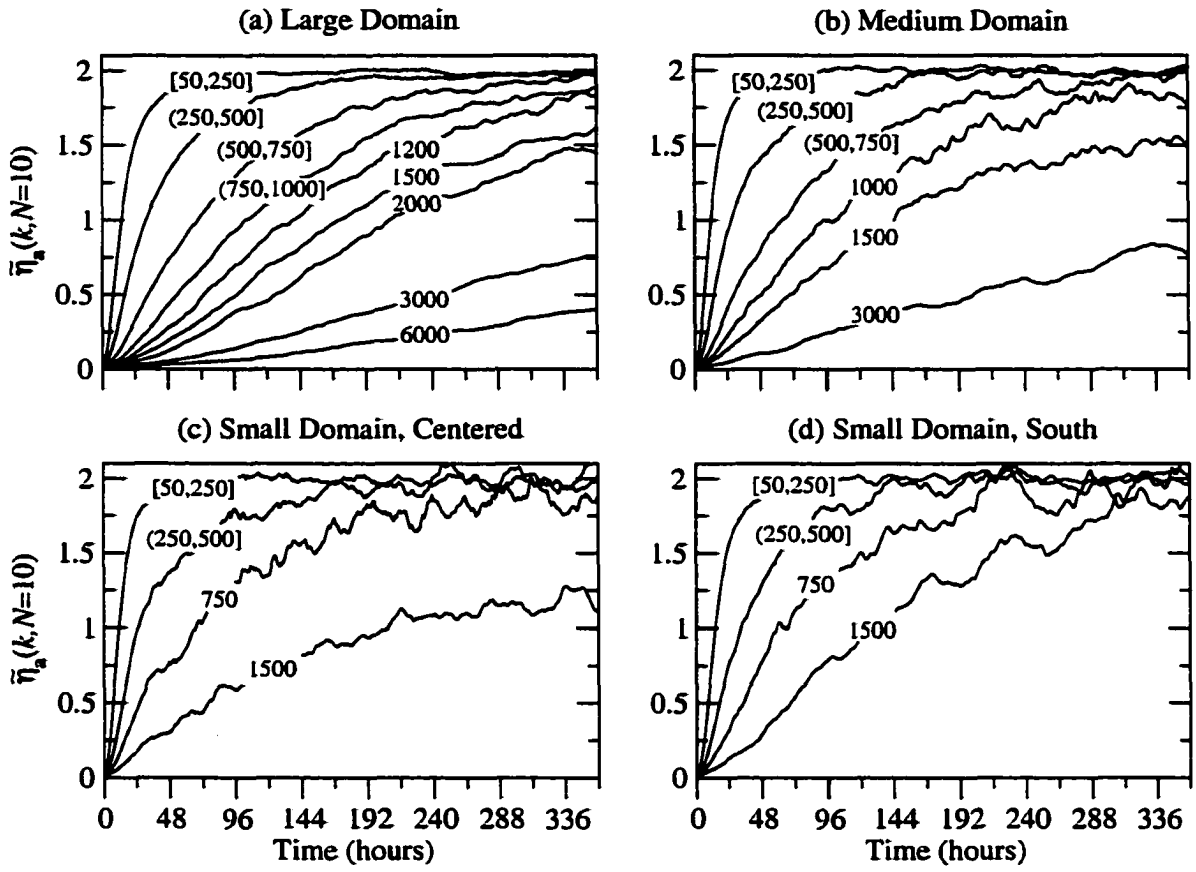



Figure 4.4: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member global ensemble simulations. Error variances are shown for specific wavelengths, or accumulated over 250 km wavelength bands as indicated by labels on each curve.

for individual ensembles and necessitates the use of averages over many cases to help smooth the curves for useful interpretation.

Several of the curves in Fig. 4.4 slightly exceed at times the expected maximum value of 2. This result is explained by noting that initial condition perturbations introduced additional vorticity into the system. Consequently, the vorticity variance of forecasts averaged over all cases is up to 10% greater than the climate vorticity variance of analyses. If error variances are normalized by the variance of forecasts rather than the variance of analyses, all the curves remain less than or equal to the maximum expected value of two (not shown). Since the difference is less than 10%,

error variances continue to be normalized by analysis variances to remain consistent with the statistics as defined in chapter 2.

## 4.3 Ensemble Bias, Dispersion, and MSE

Equation 2.11 was derived to help link error variance spectra to ensemble dispersion and is repeated here for convenience:

$$D^2 = \sigma^2 + \frac{1}{N}\sum_{i=1}^{N}(\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2 - \left\|\overline{\mathbf{f}} - \mathbf{a}\right\|^2.$$

To compare the relative magnitudes of each term in this expression, the following ensemble summary statistics are shown in Fig. 4.5:

- Ensemble dispersion, $D^2 = \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{f}_i - \overline{\mathbf{f}}\right\|^2$ (precedes Eq. 2.6)
- Total biased error variance, $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\sum_{\kappa=1}^{K-1} 2\left|F_i(\kappa)\right|^2$ (Eq. 2.8)
- Ensemble mean squared spatial error, $(\text{sme})^2 = \frac{1}{N}\sum_{i=1}^{N}(\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2$
- Grid mean square of the ensemble mean error, $(\text{eme})^2 = \left\|\overline{\mathbf{f}} - \mathbf{a}\right\|^2$
- Ensemble mean square error, $S^2 = \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{f}_i - \mathbf{a}\right\|^2$. (Eq. 2.4)

The ensemble mean-square error is included to highlight the balance of terms in equations 2.6 and 2.10, respectively:

$$S^2 = D^2 + \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{f}_i - \overline{\mathbf{f}}\right\|^2 \quad \text{and} \quad \sigma^2 = S^2 - \frac{1}{N}\sum_{i=1}^{N}(\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2.$$

The latter equation (2.10) suggests that $\sigma^2$ is a good approximation for the ensemble MSE, provided that the ensemble average spatial-mean error is small. Since the spatial mean error term is positive, $\sigma^2$ must be less than or equal to $S^2$. Figure 4.5a,b reveals the contradictory result that $\sigma^2$ is up to 15% greater than $S^2$ on the large and medium domains. This contradiction is explained by the fact that removal of linear trends from the data fields prior to calculating Fourier transforms introduces additional variance in higher wavenumbers (Errico 1985). Indeed, the

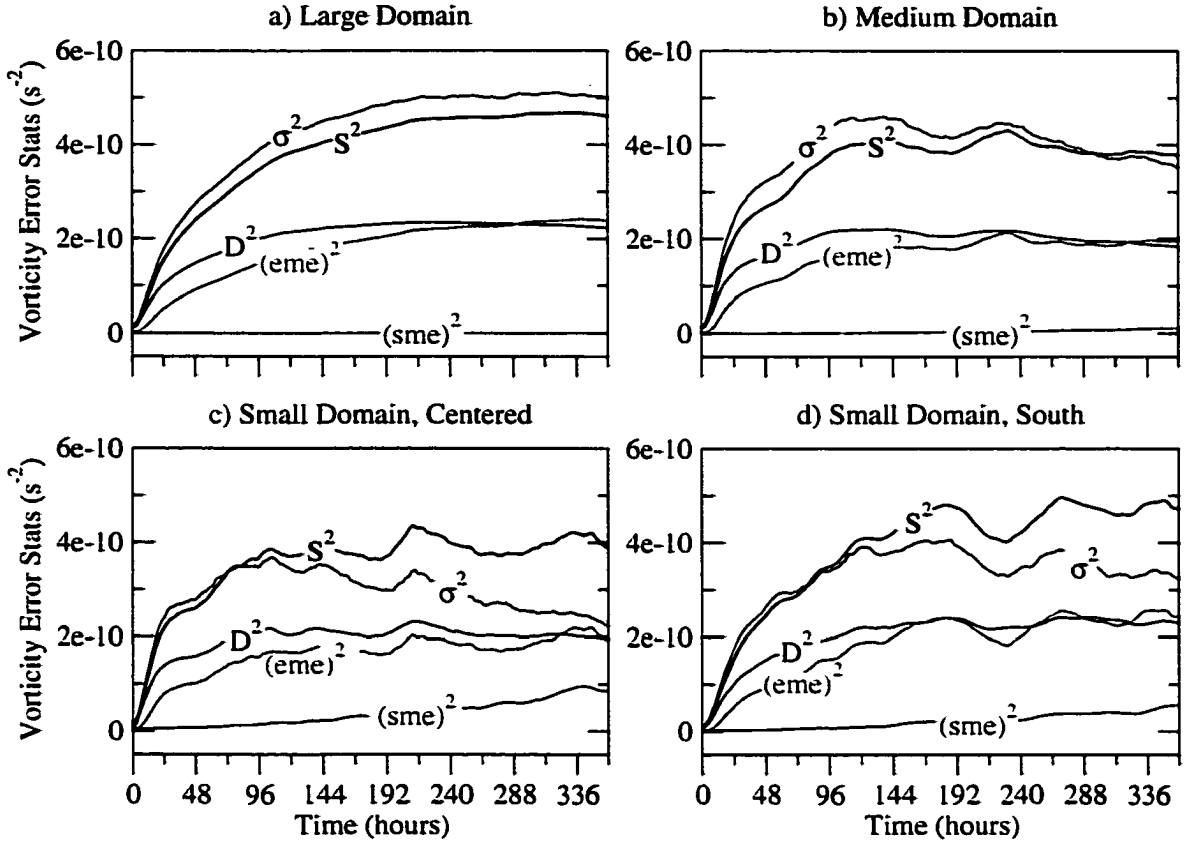



Figure 4.5: Ensemble summary statistics for vorticity, averaged over 100 independent 10-member global ensembles. See text (page 57) for a description of variables shown.

error variance curve obtained from the periodic global domain —where detrending is not required— is almost identical to the $S^2$ curve in Fig. 4.5a (not shown). On both the small domains (Figs. 4.5c,d), the spatial bias term becomes nontrivial about 5 days into the simulations. At this point, the relationship between $\sigma^2$ and $S^2$ predicted by Eq. 2.10 is confirmed. Thus, as errors grow toward longer wavelengths exceeding the size of the smaller subdomains (1500 km$^2$), the spatial mean error grows larger and contributes to a reduction of $\sigma^2$ relative to $S^2$.

The relationship between ensemble dispersion and ensemble MSE was defined by Eq. 2.6, which shows the importance of ensemble mean error. For about the first 12 hours of the simulations, the ensemble mean error (bias) remains small, so the dispersion and MSE have similar magnitudes. However, the mean error grows quickly

as individual ensemble members become increasingly uncorrelated and unpredictable features are averaged out. These results are seen on all four subdomains by the end of the forecast period.

Although we cannot decompose the contributions made by various wavelengths to ensemble dispersion, it has been possible to decompose the error variance, which is then related back to dispersion and ensemble MSE. The only limitation to this argument is the effect of spatial biases. The spatial biases in Fig. 4.5 are small, thus enabling the link between ensemble dispersion and error variance contributions at different wavelengths.

# Chapter 5

# Limited-Area Ensemble Simulations

When LBCs are provided from an external model simulation, they typically are given at, for example, 3-, 6-, or even 12-hour intervals. Since LAM solutions require LBCs at every time step, the LBCs are usually interpolated linearly in time between the available updates. This interpolation causes aliasing that acts as a temporal filter, removing the high frequency wave components whose wave period is shorter than twice the LBC update interval. Consequently, we expect to see a constraint on error variance growth and ensemble dispersion as discussed in section 2.4.

In addition to the temporal deficiency, LBCs are almost always provided on a grid having lower spatial resolution than the interior LAM domain. Consequently, small-scale features are progressively swept from the domain as the coarsely resolved fields propagate through the upstream boundary (Errico and Baumhefner 1987; Vukicevic and Errico 1990; Warner et al. 1997). This effect was also highlighted in section 2.4 as an additional constraint that limits error variance growth and ensemble dispersion.

The impact of these two effects on error variance growth and ensemble dispersion are considered in the following sections. The effects are first considered independently, then combined to evaluate the collective impact.

## 5.1 Impact of LBC Update Interval

### 5.1.1 Example Simulations

The following example simulations help visualize the impact of infrequent LBC updates in the context of model configurations used for this work. The first effect

considered below has received little direct and/or systematic attention in the published literature. The effect is dubbed LBC "pulsing", and is a direct consequence of using infrequently updated, but otherwise perfect LBCs. As the LBC errors introduced by pulsing begin propagating through the LAM domain, error growth within the LAM domain becomes controlled by the more widely recognized "sweeping" effect.

#### 5.1.1.1 LBC Error "Pulsing"

The example case shown in Fig. 5.1 is perfect except for errors created using 3-hourly interpolated LBCs. The initial and lateral boundary conditions are subsets of output from a full-channel simulation running at the same resolution as the LAM model. The initial condition is not shown because it does not contain error.

During the first 1.5 hours of the simulation (Fig. 5.1-a,b,c), errors develop within the peripheral 7-point wave absorbing zone (Sec. 3.2.4). Then, from 2 to 3 hours (panels d,e,f), errors within the buffer zone decline and vanish. This error growth "pulse" repeats between subsequent 3-hourly LBC updates, with the peripheral errors vanishing at 3, 6, and 9 hours (Fig. 5.1-f,l,r). Errors in the buffer zone are greatest near the midpoint of the LBC update cycle when the respective linearly and nonlinearly evolving external and internal solutions are most inconsistent. Once introduced, the LBC pulse errors continue to propagate inward and modify the LAM solution. As the LAM solution becomes more infected with each successive error "pulse", the LBC inconsistency becomes stronger and generates larger errors, which then propagate farther inward.

The LBC update interval determines the minimum spatial scale of errors that can be introduced by aliasing of fields passing through the lateral boundary, provided that advection or wave propagation is the main cause of local changes in the solution. Suppose a wave passes through the boundary with speed $|c_x|$. If the LBC is updated with frequency $f_{LBC}$, then the minimum wavelength that can be fully sampled is $\lambda_{LBC} = 2|c_x|/f_{LBC}$. Thus, the minimum spatial scale of aliasing errors introduced while waves pass through the boundary increases with less frequent LBC updates. If this minimum scale is longer than the smallest resolved wavelength on the LAM domain, then LBC pulsing errors will contribute to constraints on error variance

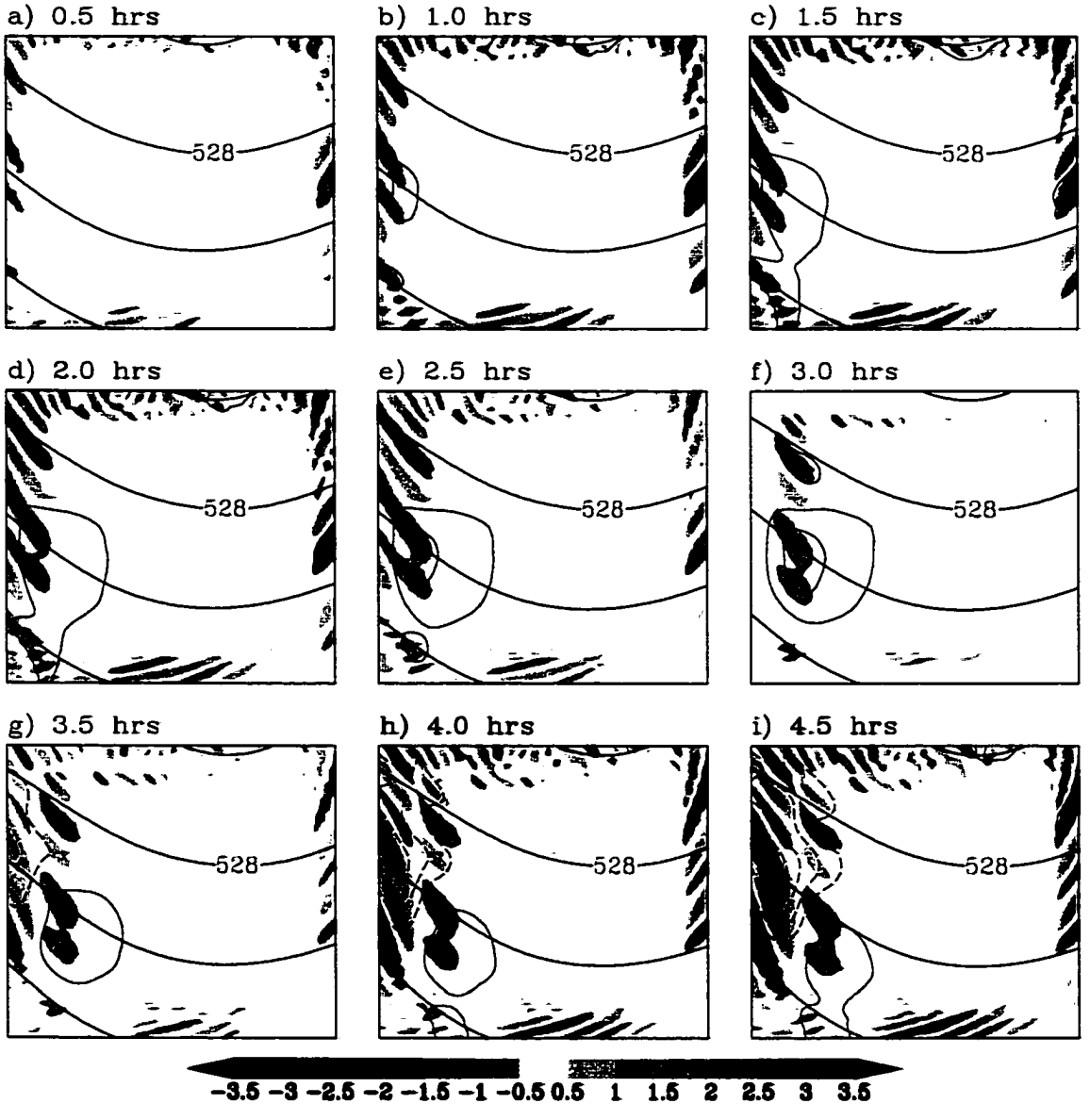



Figure 5.1: Example LAM simulation run on the small, center domain (1500 km$^2$, Fig. 3.2) showing effects of LBC update interval in an otherwise perfect simulation. LBCs are given every three hours, and linearly interpolated at each step between the updates. Gentle curves are streamlines from the control simulation that provides the LBCs ($12 \times 10^6$ m$^2$s$^{-1}$ intervals). Vorticity errors ($\times 10^{-5}$ s$^{-1}$) are shaded while (positive/negative) streamfunction errors are shown with (solid/dashed) contours at $\pm 5 \times 10^4$ m$^2$s$^{-1}$ intervals, or $\sim$0.01% (zero line omitted).

growth (recall Sec. 2.4) while "sweeping" through the domain. LBC pulsing errors are maximized at small scales due to the filtering effect, but may also form at larger scales by slightly reducing the amplitude of all non-stationary waves passing through the lateral boundary.

The Rossby wave phase relation (Eq. 3.12) can help quantify the temporal filtering effect for the PPV model used in this study. Zonal phase speeds for waves passing through the boundary are given by $c_x = U_0 - \beta/(k^2+l^2)$. Assuming zonal flow ($l = 0$), $k = 2\pi/L_x$, and $|c_x|$ defined above, we find that $\lambda_{LBC} = 2|U_0 - \beta L_x^2/4\pi^2|/f_{LBC}$. Results obtained from this expression for selected nesting frequencies are shown in Fig. 5.2. Constants $U_0$ and $\beta$ are specified in Table 3.2. There is a cusp in the curves for external wavelengths of 5400 km where the Rossby waves are stationary and do not suffer any filtering due to temporal interpolation of LBCs. All values of $\lambda_{LBC}$ are positive because the absolute value of $c_x$ was used.

The results in Fig. 5.2 show that for 1/2-hourly updated LBCs, all waves are sampled well enough to have adequate spatial resolution on the LAM domain. When using hourly-updated LBCs, external wavelengths less than 2500 km move fast enough to produce aliasing errors at scales up to about 100 km. Since less than 1% of the total vorticity variance is represented by these scales (Fig. 3.8), LBC pulsing errors remain negligible for this model if LBCs are updated at least once per hour. When using 3-hourly LBC updates, external wavelengths that are shorter than 3000 km generate aliasing errors at scales no smaller than 200 km while passing through the lateral boundary. By applying the argument of section 2.4, we see that this loss of spatial resolution in error growth caused by temporal filtering of LBCs contributes to constraints on error variance growth and ensemble dispersion.

The vorticity errors in the example case (Fig. 5.1) produce approximately 5 to 10 anomaly couplets along each side of the domain. Since each side has a length of 1500 km, this result suggests that the vorticity errors in this example form at wavelengths of about 150 to 300 km. This estimate is consistent with those indicated in Fig. 5.2 for 3-hourly LBC updates.

The LBC pulse error is seen on each of the four domains tested (not shown in examples), but takes longer to propagate across the medium (3000 km$^2$) and large (6000 km$^2$) domains. As expected, the pulse effect is less pronounced when using hourly LBC updates, and more pronounced when using 6-hourly LBC updates. Note

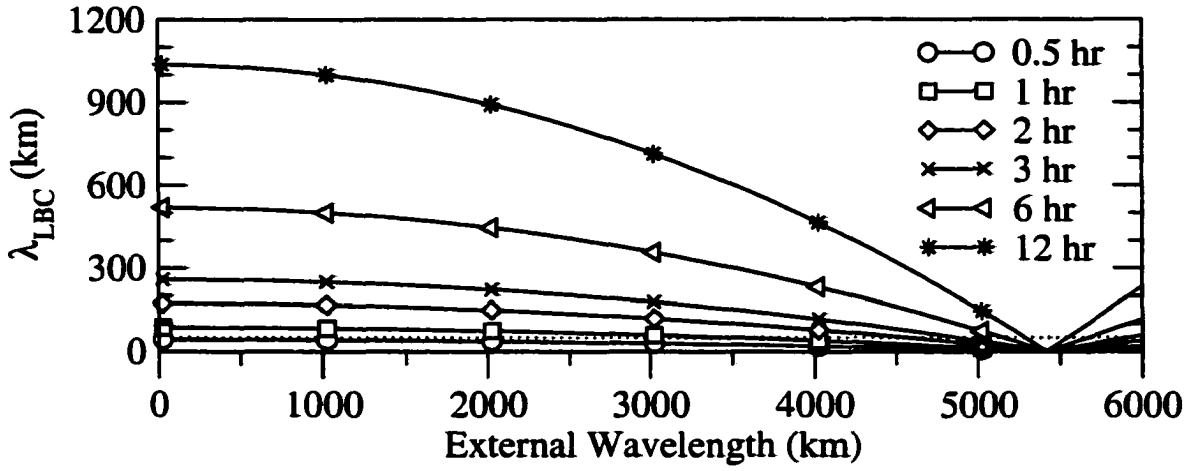



Figure 5.2: Minimum resolved spatial scale of aliasing error after temporal filtering due to linear interpolation of LBCs between available updates. The abscissa marks wavelengths on the external domain passing through the LBC with Rossby phase speed $|c_x|$. The ordinate shows minimum resolved wavelength $\lambda_{LBC} = 2|c_x|/f_{LBC}$ for different LBC nesting frequencies $f_{LBC}$. The dotted horizontal line marks the smallest $2\Delta x = 50$ km wavelength resolved on the LAM domain.

that in practical applications the LBC pulse effect could be swamped by other errors when the LBCs are given by discordant external analyses or forecasts. Furthermore, note that the distribution of phase speeds varies with the complexity of the synoptic setting. Therefore, the intensity at which LBC pulse errors are produced also varies with time and location.

#### 5.1.1.2 LBC "Sweeping"

In a widely recognized paper, Baumhefner and Perkey (1982) showed that propagation of LBC errors toward the interior of the LAM domain can account for up to 50% of the total forecast error. As features originating along the inflow boundary pass through the LAM domain, they limit error variance growth and ensemble dispersion by "sweeping-out" other features evolving within the LAM (Errico and Baumhefner 1987; Vukicevic and Errico 1990). Whereas these previous studies emphasized error removal by "sweeping" of unperturbed LBCs, the emphasis here is

on the propagation of LBC error into an otherwise perfect LAM simulation. Specifically, the example case in this section focuses only on the sweeping, or throughput, of errors generated by LBC "pulsing" from infrequent LBC updates. Boundary sweeping by coarsely resolved LBCs is considered later in section 5.2.

The LBC sweeping effect is illustrated here by continuing the example LAM simulation from the previous section to 96 hours on the large, medium, and small-centered domains (Figs. 5.3 to 5.8). As before, the only source of error results from temporal interpolation between LBC updates given every 3 or 6 hours. The LBC error pulsing effect is not directly evident in these figures because results are shown at 12-hour intervals, a time when the LBCs are perfect and errors vanish within the boundary zone. The following discussion emphasizes the manner in which LBC sweeping constrains error growth for each of the different domains and update intervals.

The first example (Fig. 5.3) is run on the large (6000 $km^2$) domain with 3-hourly LBC updates. From 0 to 12 hours, LBC errors resulting from temporal interpolation begin to appear along the upstream boundary. After 60 hours (panel f), the LBC errors have propagated almost completely through the domain and have also appeared to the north and south outside the main gradient flow. Streamfunction errors develop as a trough passes through the upstream boundary between 36 and 48 hours. This error subsequently broadens to larger scales and moves further into the domain.

The example simulation was also run on the large domain, but with 6-hourly LBC updates (Fig. 5.4). As anticipated from the previous section, more intense errors are generated for this case along the upstream boundary. The propagation speed of the vorticity error has not changed, as it still takes about 60 hours for the errors to sweep through the domain. However, streamfunction errors appear after just 12 hours and evolve upscale while propagating through the domain.

LBC interpolation errors sweep through the medium domain (Fig. 5.5) in about half the time needed to cross the large domain. The propagation speed of the errors remains unchanged since the medium domain is half the size of the large domain. However, once inside the LAM domain, the errors have little time to evolve and/or amplify before exiting the downstream boundary. Hence, errors do not grow as large as they could have on an unbounded domain. The same effect is noted for

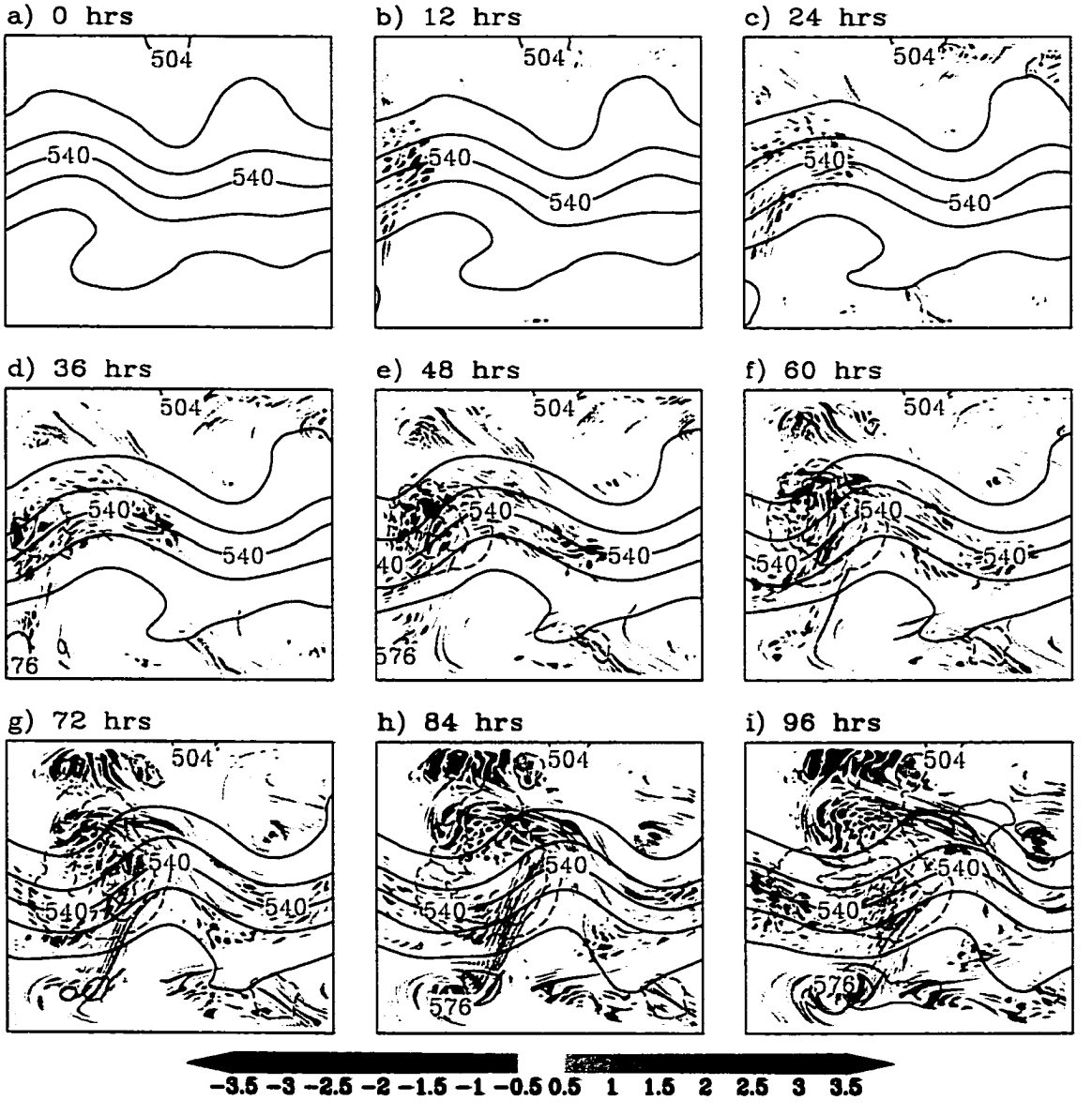



Figure 5.3: Example LAM simulation run on the large domain (6000 km$^2$, see Fig. 3.2) showing effects of LBC update interval in an otherwise perfect simulation. LBCs are given every three hours, and linearly interpolated at each step between the updates. Gentle curves are streamlines from the control simulation that provides the LBCs ($12 \times 10^6$ m$^2$s$^{-1}$ intervals). Vorticity errors ($\times 10^{-5}$ s$^{-1}$) are shaded while (positive/negative) streamfunction errors are shown with (solid/dashed) contours at $\pm 2 \times 10^5$ m$^2$s$^{-1}$ intervals, or $\sim$0.04% (zero line omitted).

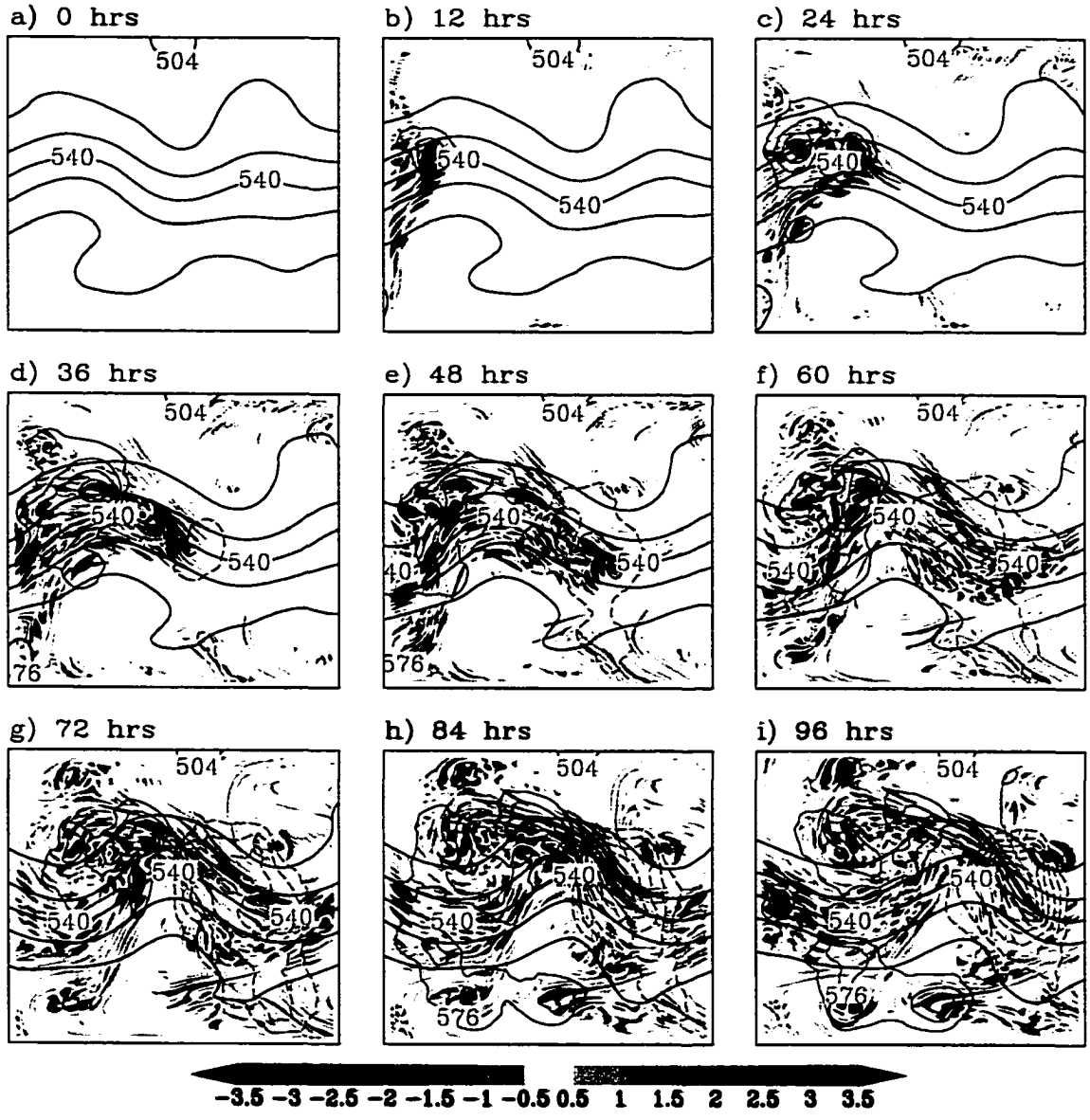



Figure 5.4: As in Fig. 5.3, except lateral boundary conditions are updated every six hours.

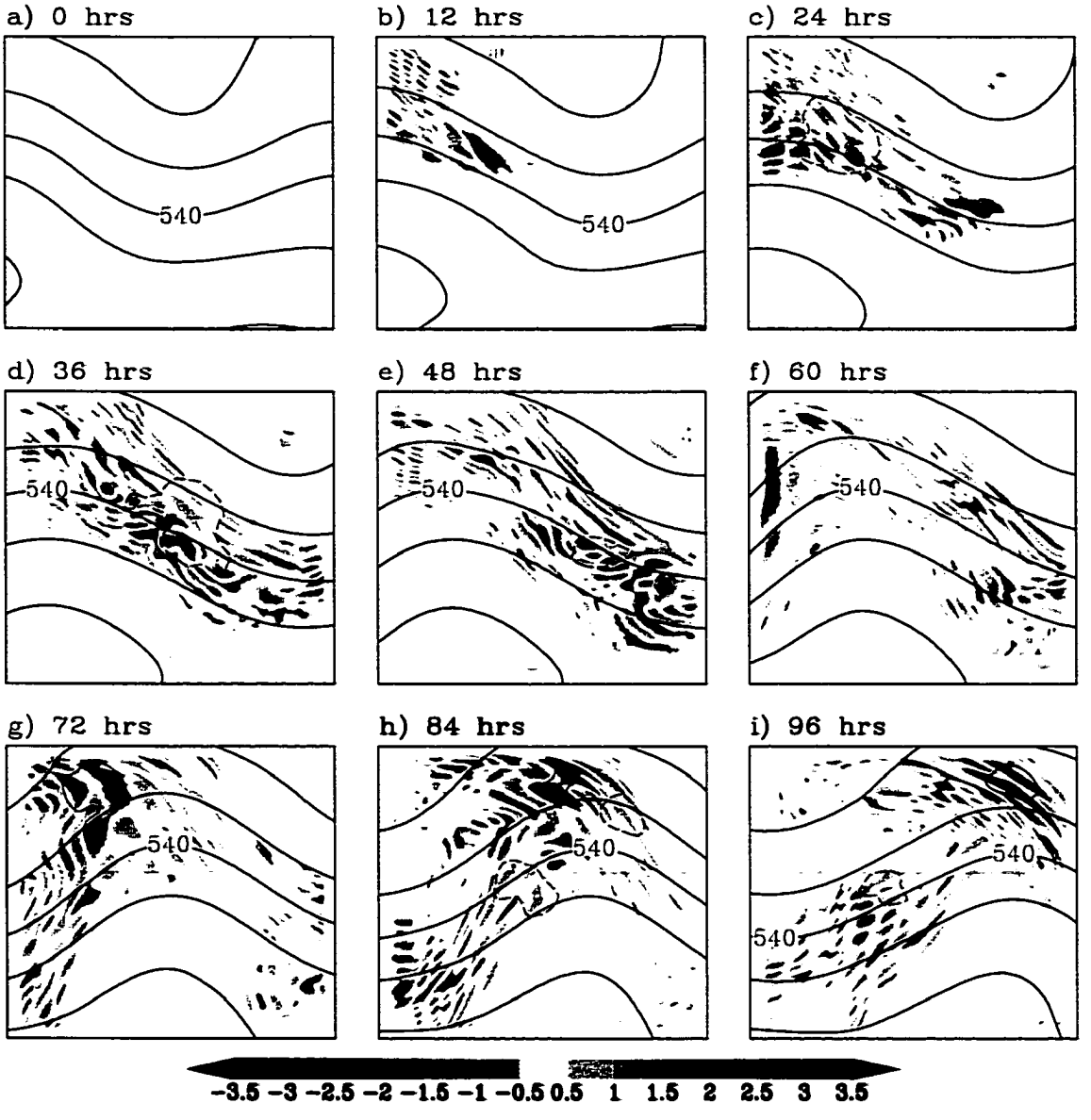



Figure 5.5: As in Fig. 5.3, except LAM simulation is run on the medium domain (3000 km$^2$).

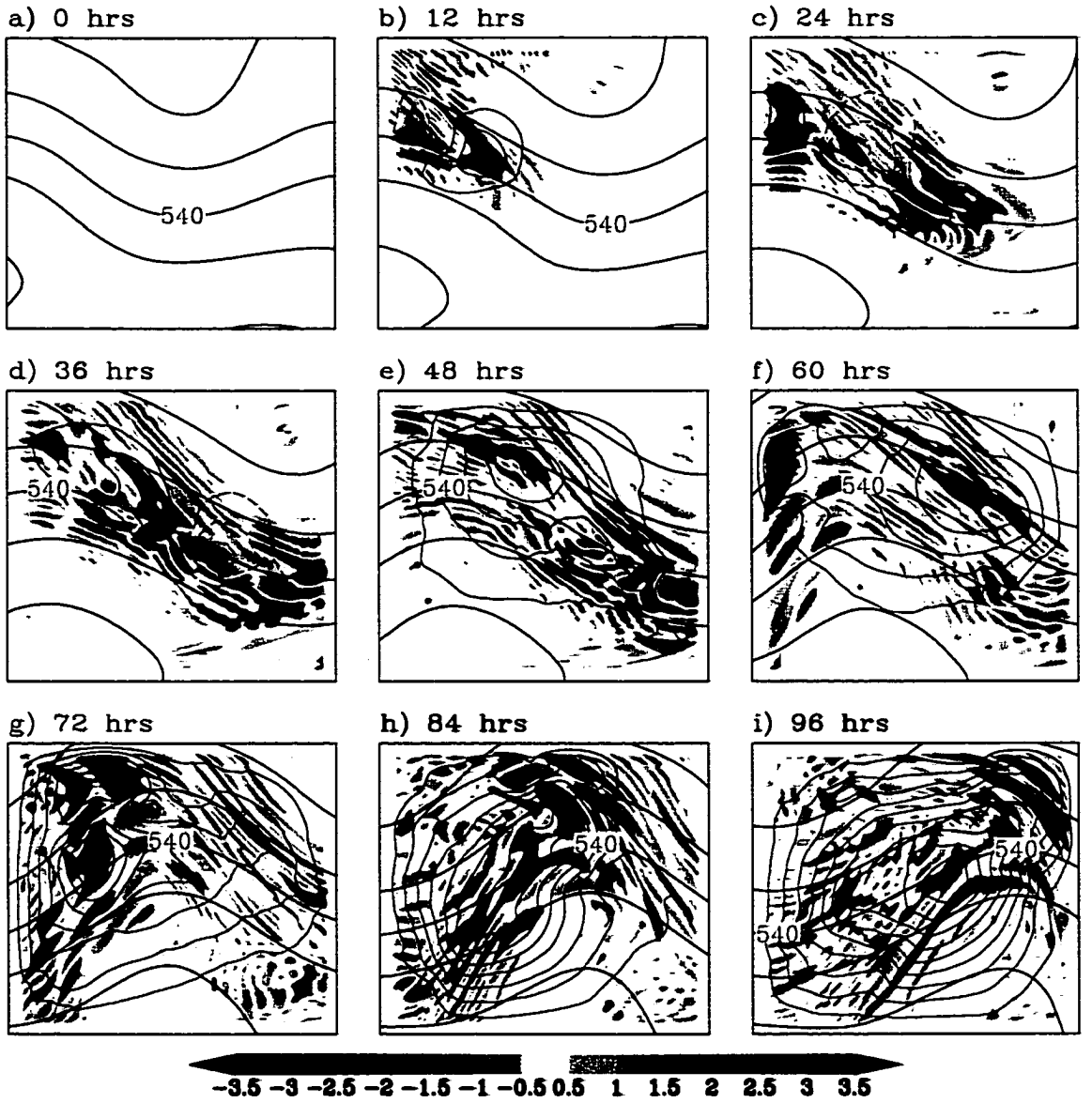



Figure 5.6: As in Fig. 5.3, except LAM simulation is run on the medium domain (3000 km$^2$) and LBCs are updated every six hours.

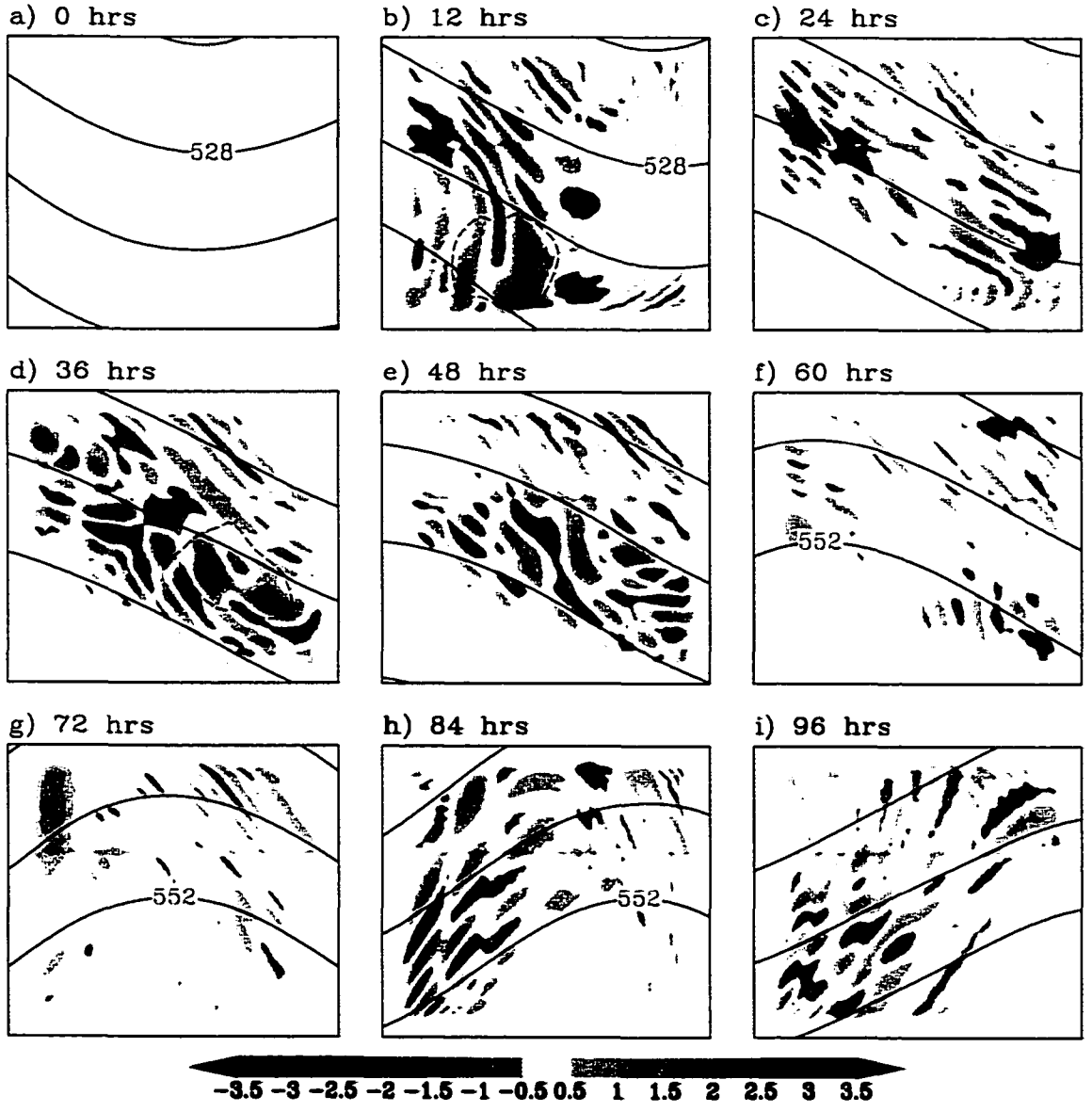



Figure 5.7: As in Fig. 5.3, except LAM simulation is run on the small, centered domain (1500 km$^2$).

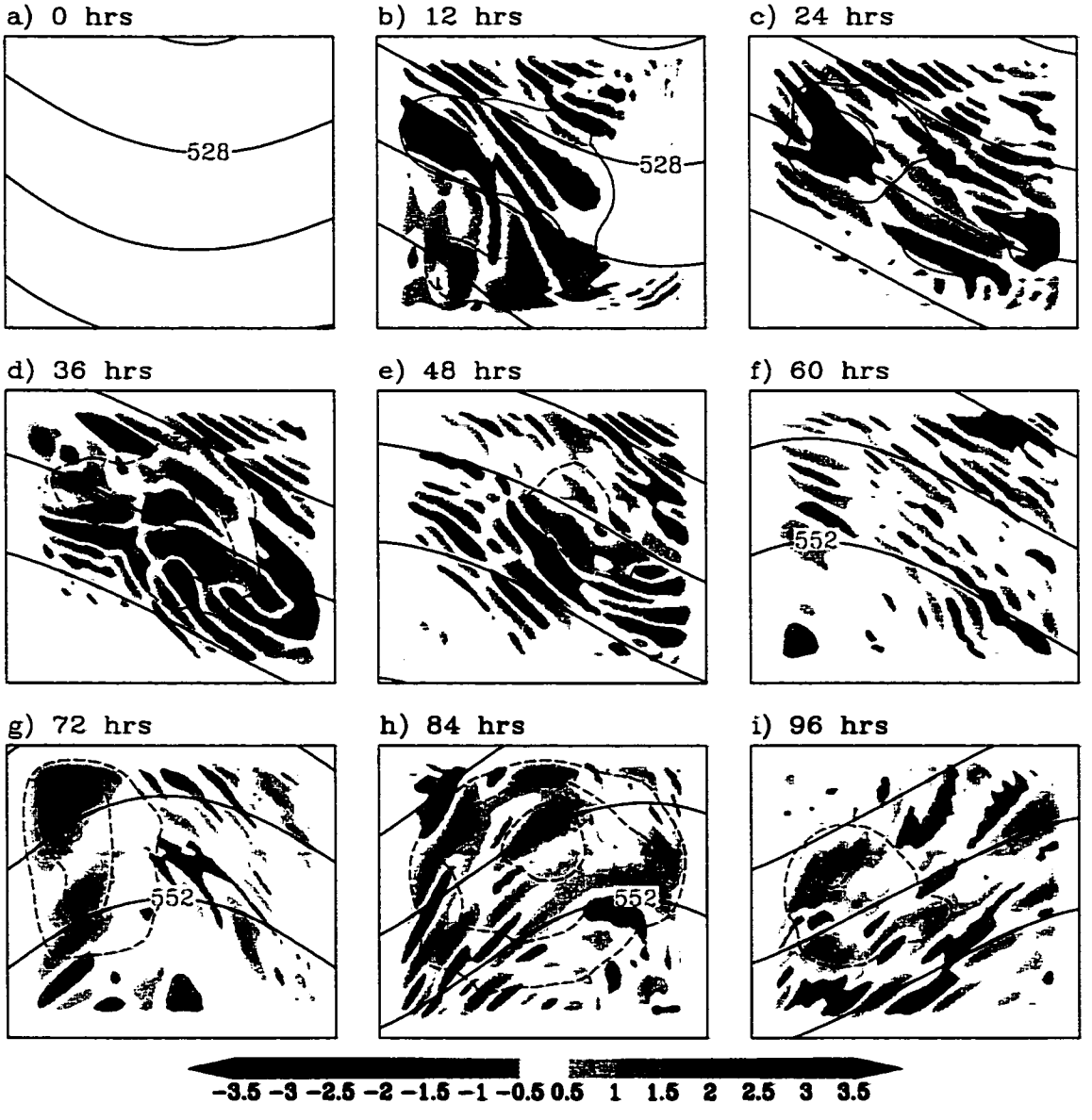



Figure 5.8: As in Fig. 5.3, except LAM simulation is run on the small, center domain (1500 km$^2$) and LBCs are updated every six hours.

streamfunction errors, which are small compared to those on the large domain. The position of the large-scale wave pattern also appears to modulate the amplitude of the errors through variations in the LBC error pulse as discussed above. The most intense errors seem to form between the passage of ridges and troughs (compare panels f and h).

Error growth is more vigorous on the medium domain when LBCs are given at 6-hourly intervals (Fig. 5.6). Vorticity errors completely sweep through the domain by 36 hours, but a streamfunction error couplet has also formed whose spatial scale nearly fills the extent of the domain. As the errors continue to propagate upscale, they eventually reach a spatial scale limit determined physically by the extent of the LAM domain itself. This is a secondary type of error growth constraint discussed previously by Vukicevic and Errico (1990).

Similar effects are noted on the small, center domain for both 3-hourly (Fig. 5.7) and 6-hourly (Fig. 5.8) LBC update intervals. LBC errors pass through the domain within about 12 hours after they first appear, hence placing strong limits on error growth. Furthermore, the lack of continuity between images emphasizes both the rapid LBC error sweeping and the dependence of errors on the larger scale pattern.

### 5.1.2 Error Variance Spectra

Results from the previous example simulations showed that temporal interpolation of LBCs initiates error growth by filtering small scale waves from fields passing through the lateral boundaries (Sec. 5.1.1.1). Once LBC errors propagate into the LAM domain, sweeping of new errors from the upstream boundary constrains the amplitude of error growth. From this perspective, we expect that error variance growth on LAM domains is less than it would be in the absence of LBC effects, especially at small scales.

To document the loss of error variance at smaller scales due only to LBC interpolation, 100 independent LAM simulations are conducted using configurations shown in the previous examples. Initial and lateral boundary conditions are provided by direct subsets of the global simulations running with the same numerical configuration as the LAM. The simulations are perfect except for the use of one-way LBCs interpolated linearly in time between updates. Error variances are computed

spectrally using Eq. (2.22). LBCs are the only source of model error for this configuration. The only available benchmark for this LAM configuration is the expected value at saturation of 2.0 which applies equivalently to all wavelengths.

Results from LAM simulations with hourly updated LBCs are shown in Fig. 5.9. Note that the error variances are all much smaller than the saturation value of 2.0. Error growth at large scales greater than about 750 km remains negligible throughout the simulations. Results at smaller scales vary by domain size. On the large domain (Fig. 5.9a), shortwave error variances continue to increase with time without reaching a limiting value. However, on the medium and small-center domains (Fig. 5.9b-c), error growth in the 50-250 km wavelength interval ceases after about 48 and 24 hours, respectively. This effect is a manifestation of LBC sweeping; errors propagate completely through the domain before they have time to grow larger. In general, error variance growth behaves as expected for this case. The hourly LBC update interval is sufficient to completely sample most waves passing through the lateral boundary (see Figs. 3.8 and 5.2). Hence, nearly negligible errors are introduced through LBC "pulsing". Errors that do form are constrained well below maximal growth by LBC "sweeping" on all but the large domain.

When LBCs are updated every 3 hours (Fig. 5.10), shortwave error variances on the medium and small domains are approximately 10 to 30 times larger than the previous hourly updated case. Error variances on the large grid have increased only slightly. The same basic error growth characteristics are observed as in Fig. 5.9. Specifically, small scale errors grow freely on the large domain but reach a limiting value on the smaller domains due to LBC sweeping.

A new attribute appearing with 3-hourly updated LBCs is the dominant impact of LBC error pulsing (Sec. 5.1.1.1). The oscillations in Fig. 5.10 are clearly attributed to LBC error pulsing because they have a period of three hours. On the small domains (panels c,d) oscillations from LBC pulsing account for about half of the error variance magnitudes. Most verification studies exclude the boundary zone from calculations to avoid these oscillations. We retain these points to enhance the number of data points in our sample, but also realize the effect is real and should be included in the statistics. For readers wishing to ignore the statistical oscillation, an error growth curve may be traced through the local minima as a good approximation of scores that would be obtained on a grid exclusive of the boundary zone.

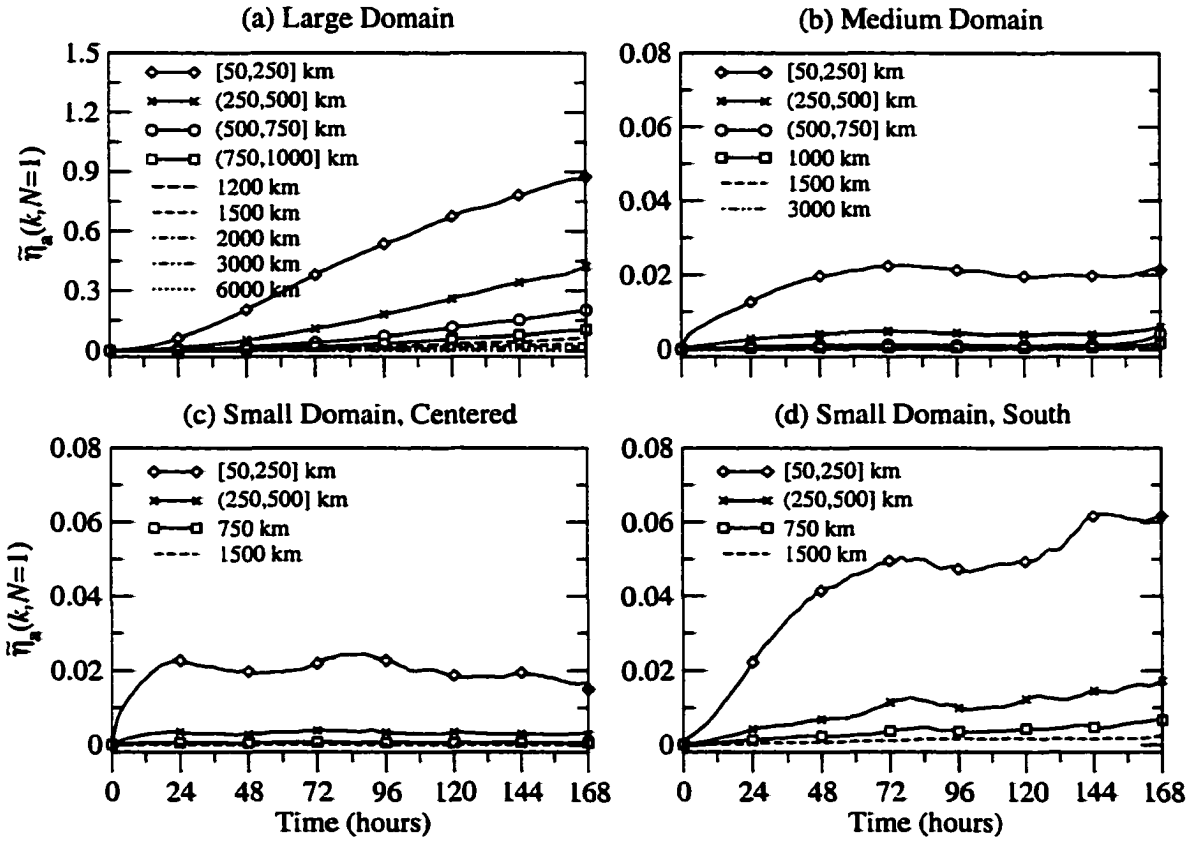



Figure 5.9: Normalized vorticity error variance (Eq. 2.22) for specific wavelengths, or accumulated over 250 km wavelength bands. Results are averaged over results from 100 independent LAM simulations that are perfect except for the use of one-way, hourly updated LBCs.

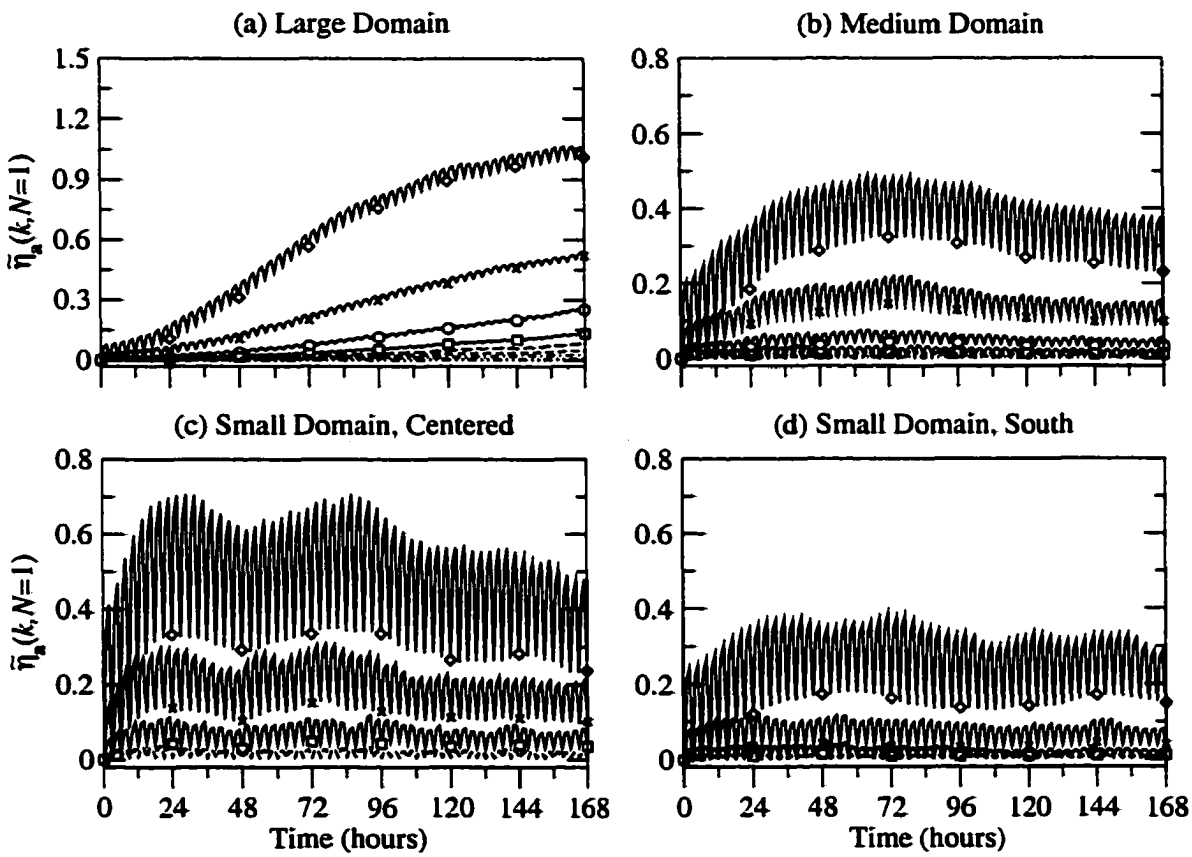



Figure 5.10: As in Fig. 5.9, including line legends, except LBCs are updated once every three hours.

When the LBC updates are provided every 6 hours (Fig. 5.11), error variances increase even more in response to stronger LBC pulsing. In particular, note that error variances have grown in all wavelengths up to about 1000 km. As larger LBC errors propagate into the domain interior, they modify the LAM solution, and subsequently lead to even greater LBC error pulses. All other attributes of error growth for this case are qualitatively the same as described above for case having 3-hourly LBC updates.

Statistical results presented in this section have demonstrated that linear interpolation of one-way LBCs between available updates generates error growth in an otherwise perfect simulation. The amplitude and spatial scale of the error growth increases with longer update intervals. Errors grow freely on the large domain, but are restricted by LBC sweeping on the medium and small domains.

### 5.1.3 Ensemble Error Variance Spectra

Results from the previous section considered error growth caused solely by interpolation in time between LBC updates. Now we consider the impact of LBC interpolation on error growth in LAM ensemble simulations having perturbed initial conditions. Statistics are averaged over 100 independent 10-member ensemble simulations. Initial conditions and LBCs for individual LAM ensemble members are obtained as direct subsets of the corresponding global ensemble members (Sec. 4.1.1). In the current configuration, both global and LAM ensemble simulations are run with the identical numerical configurations. The only sources of error growth are initial condition perturbations and the use of temporally interpolated one-way LBCs. The impact of LBCs on the LAM ensemble simulations is determined by comparing error variance spectra directly to those obtained from the laterally unbounded global simulations.

When LBCs are updated hourly (Fig. 5.12), error variance spectra for the LAM ensembles are nearly the same as those from global ensembles at wavelengths greater than about 500km. At smaller scales, error variances saturate at values less than those from the global ensembles. On the medium and small-centered domains in particular (panels b,c), error variances in the 50-250 km wavelength band reach a saturation level of just over 1.75 after 24 hours, or about 10% smaller than those

in the global ensembles. We saw previously in Fig. 5.9 that error sweeping limits the error growth at these scales; the effect is also seen here in the LAM ensembles. Specifically, the spatial filtering effect caused by temporal interpolation has removed small scale perturbations from the LBCs that would otherwise have continued to amplify in the global simulations. Aside from these relatively minor impacts at small scales, the use of hourly updated LBCs does not induce major changes in the error variance characteristics of the LAM ensemble simulations.

Given 3-hourly LBC updates (Fig. 5.13), further losses of error variance occur at small scales due to ongoing LBC sweeping of spatially filtered inflow. Saturation values in the shortest wavelength band range from about 1.5 on the small-centered grid to about 1.75 on the large domain. Thus, spatial filtering from 3-hourly LBC interpolation reduces maximum shortwave error variance growth by about 10 to 25%. Smaller reductions are seen in spatial scales up to around 750 km. The LBC error pulse is also evident, although the magnitude of oscillations accounts for at most about 5% of the error variance values in the 50-250 km wavelength band at saturation. Hence, initial condition perturbations are the dominant source of error growth for these ensemble simulations. A new attribute also appears in this configuration wherein the LAM ensemble error variances exceed those of the global ensembles during the first 12 hours of the simulations. A probable explanation for this effect is the stronger impact of the LBC error pulse.

When LBCs are updated every 6 hours (Fig. 5.14), the error variance characteristics appear almost the same as for the 3-hourly case. The saturation values in the 50-250 km wavelength band are slightly lower than in the previous 3-hourly updated case due to LBC sweeping. The initial LBC error pulse generates strong error variances during the first 12-18 hours, more than doubling values from the global ensembles.

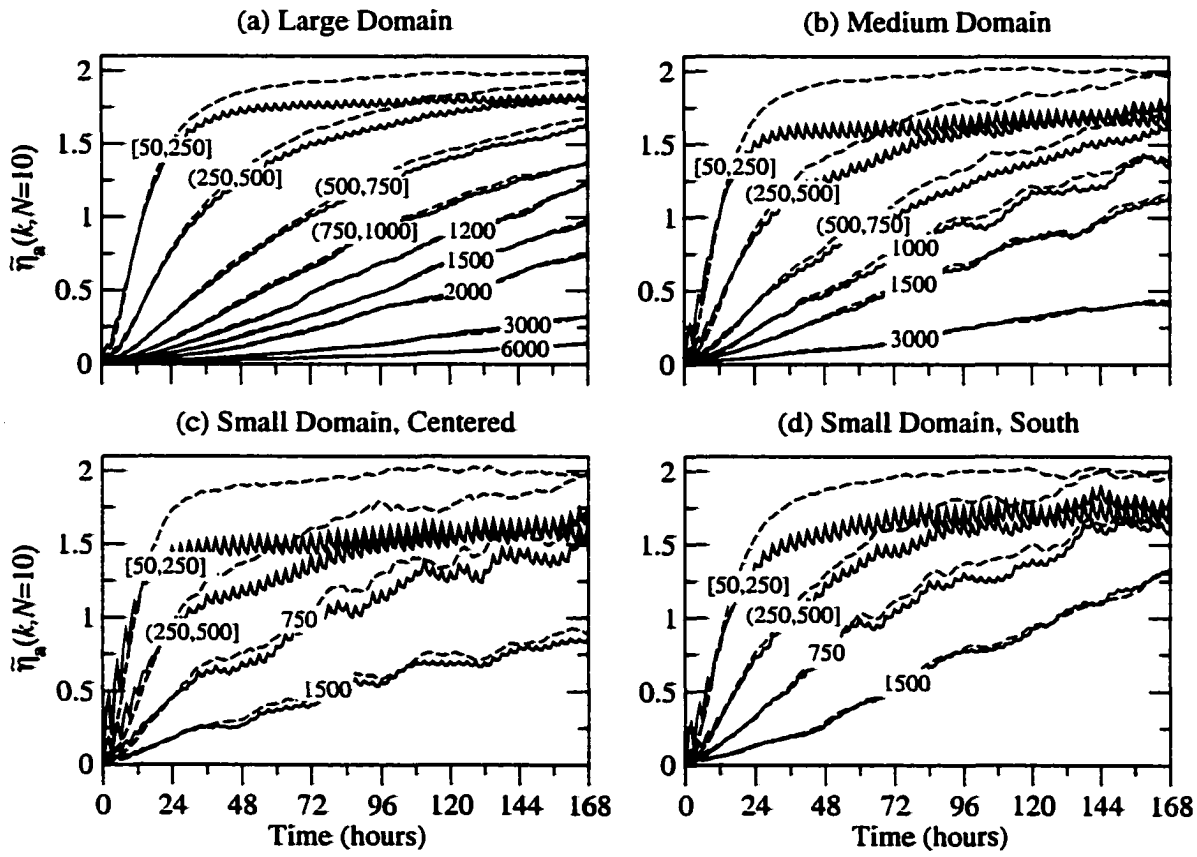



Figure 5.13: As in Fig. 5.12, except LBCs are updated once every three hours.

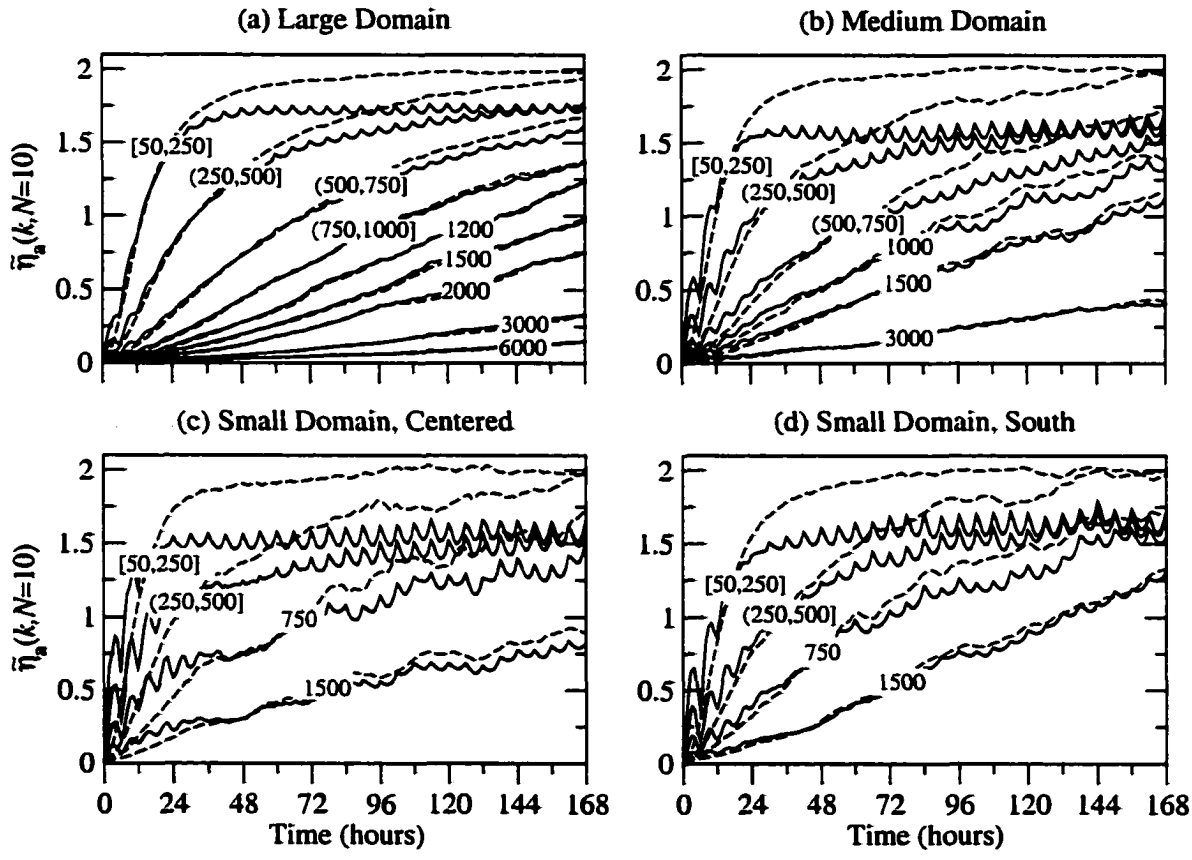



Figure 5.14: As in Fig. 5.12, except LBCs are updated once every six hours.

### 5.1.4 Ensemble Summary Statistics

Equation (2.11) provides a relation between ensemble dispersion and the error variance spectra presented above. Repeating again for convenience,

$$D^2 = \sigma^2 + \frac{1}{N}\sum_{i=1}^{N}(\langle \mathbf{f}_i \rangle - \langle \mathbf{a} \rangle)^2 - \left\| \bar{\mathbf{f}} - \mathbf{a} \right\|^2 .$$

The last two terms in this expression are spatial and ensemble bias terms that are not easily decomposed to show their contributions to $D^2$ at different scales. To attribute changes in error variance spectra ($\sigma^2$) to ensemble dispersion at specific wavelengths, we must account for differences in the bias terms obtained from global and LAM ensemble simulations. If these terms remain nearly unchanged under different model configurations, then we can attribute losses in dispersion directly to constraints on error variance growth at particular scales.

As discussed previously (page 57), removal of linear trends on the LAM domain introduces additional variance at small scales. Thus, as suggested by Eq. 2.10, $\sigma^2$ is not a fair approximation for $S^2$ even when the spatial bias is small on limited-area domains. However, the variance added through detrending does not affect the spatial bias or ensemble mean error terms. Therefore, the relationship between changes in ensemble dispersion and in total error variance that are attributable to LBC impacts is not altered by the detrending procedures.

For the LAM ensemble case having hourly updated LBCs (Fig. 5.15), the loss of error variance at small scales yields only minimal reductions in the ensemble summary statistics when compared to those from the global simulations. Indeed, the reductions are trivial except for ensemble mean-square error ($S^2$), total error variance ($\sigma^2$), and ensemble dispersion ($D^2$) on the medium and small-centered domains (panels b,c). The spatial bias and ensemble mean error terms are nearly identical in both LAM and global ensemble simulations. These results show that LAM ensemble dispersion is minimally impacted by hourly updated LBCs since waves passing through the lateral boundary are sampled reasonably well. This conclusion is consistent with results shown above in Fig. 5.12 where the small-scale error variances decreased only by about 10% for LAM ensemble simulations.

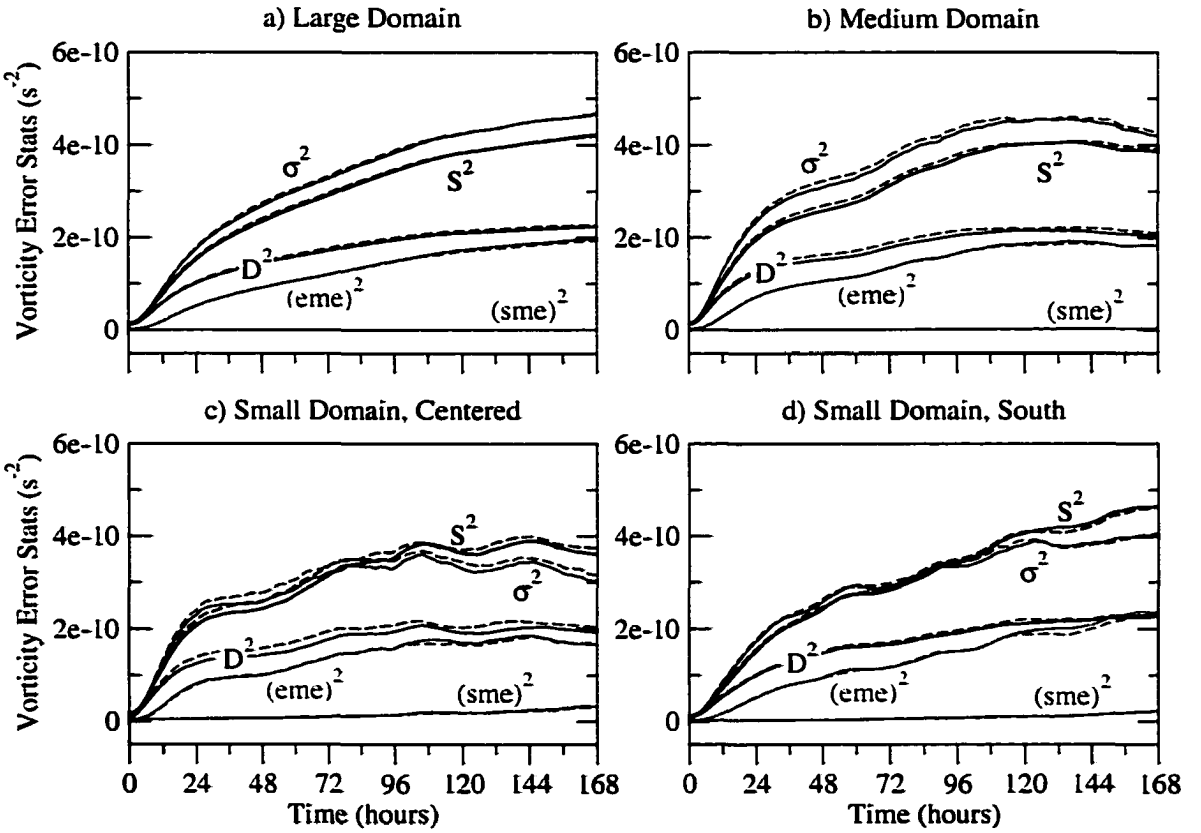



Figure 5.15: Ensemble summary statistics for vorticity, averaged over 100 independent 10-member LAM simulations having hourly updated LBCs. See text (page 57) for a description of variables shown. Dashed reference lines are reproduced from Fig. 4.4, showing statistics from the global ensemble simulations.

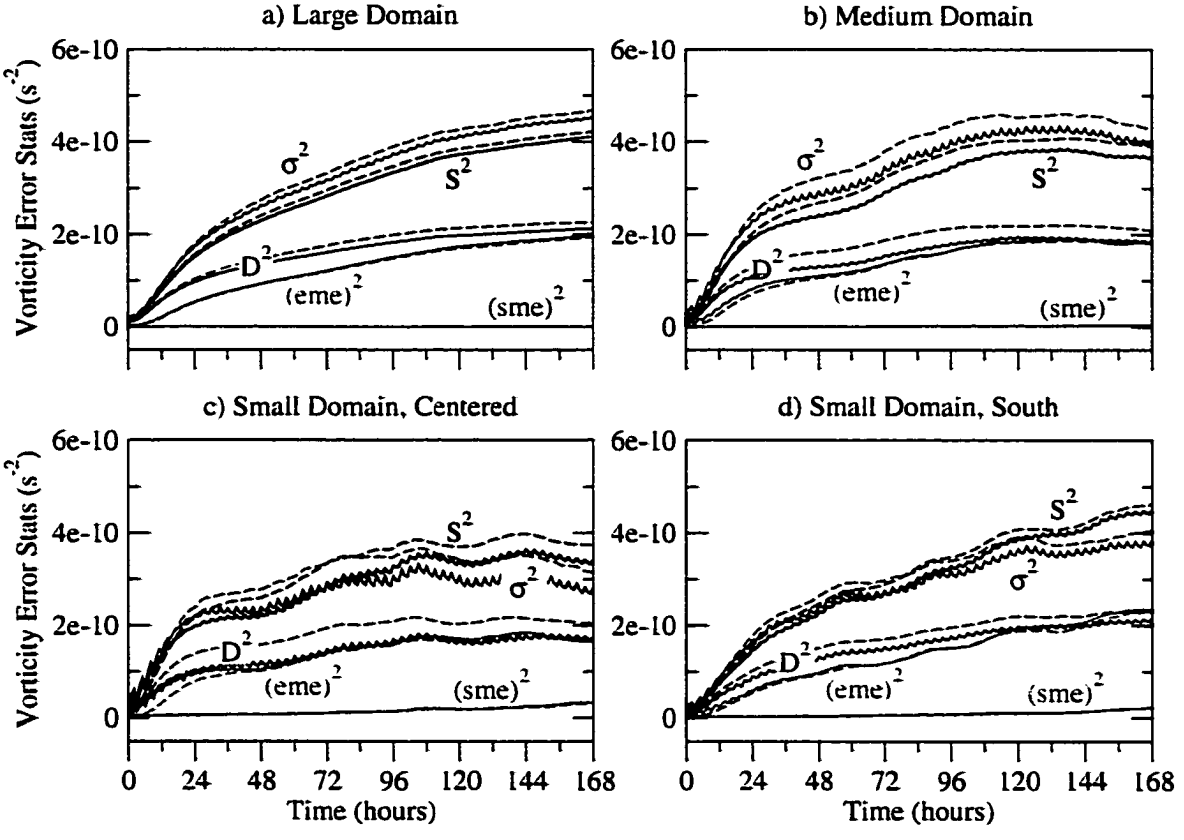



Figure 5.16: As in Fig. 5.15, except LBCs are updated once every three hours.

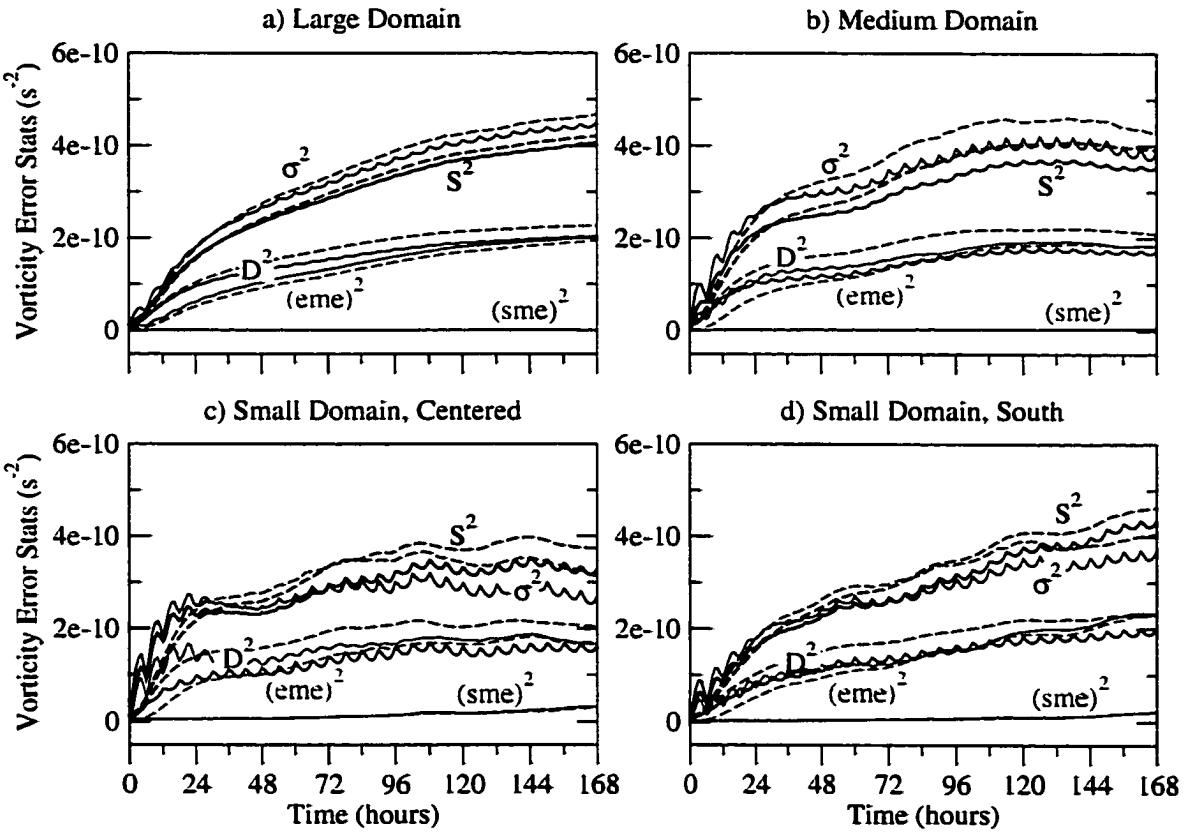



Figure 5.17: As in Fig. 5.15, except LBCs are updated once every six hours.

Figure 5.16 shows that LAM ensemble dispersion is more strongly affected when LBCs are updated every 3 hours. The spatial bias again remains nearly identical for both LAM and global ensembles. The ensemble mean error increases slightly for the first 36 hours of simulations on the medium and small-centered domains (panels b,c), probably in response to the LBC error pulse between the first few LBC updates. This increase in ensemble mean error also yields an increase in $\sigma^2$ and $S^2$. However, LAM ensemble dispersion remains less than global ensemble dispersion because the increase in $\sigma^2$ is negated by the increase in ensemble mean error (see Eq. 2.11). LBC error sweeping of filtered scales begins to reduce ensemble dispersion almost immediately by constraining error variance growth at small scales. The LBC error pulse seen in $\sigma^2$ also affects ensemble dispersion, although the magnitude of the oscillation is small compared to the overall loss of dispersion.

Given 6-hourly LBC updates (Fig. 5.17), the effect is about the same as above. One conclusion that might be drawn from this statement is that ensemble dispersion at saturation will be about the same if LAM ensembles are given LBCs every 3 hours or every 6 hours. However, the use of less frequent LBC updates does increase the ensemble mean errors. But again, these errors do not affect dispersion until LBC sweeping acts to constrain error variance growth at small scales.

Results presented in this section show that loss of error variance at small scales due to LBC interpolation is responsible for a loss of LAM ensemble dispersion when using LBC update intervals of 3-hours or longer. One might have expected that the loss of error variance at small scales identified above is not important because most of the variance in atmospheric flows is contained at large scales. Contrary to this expectation, the summary statistics presented here show that the loss of of error variance at small scales is enough to produce a decline in the total error variance ($\sigma^2$), and hence, the ensemble mean square error ($S^2$) and ensemble dispersion ($D^2$).

## 5.2 Impact of Spatial Resolution in LBC

The previous section highlighted the loss of error variance and LAM ensemble dispersion caused exclusively by linear interpolation between LBC updates. The analysis is repeated presently, but this time with LAM configurations designed to isolate the impact of coarsely resolved LBCs.

Coarsely resolved LBC fields are generated by applying a low-pass spatial filter (see Appendix B) to fields from global model simulations running at the same resolution as the LAM. The low-pass filter removes completely all wavelengths shorter than 150 km while perfectly retaining the amplitudes of wavelengths longer than 450 km. The filtering process preserves the accuracy of large scale motions while removing those which would not be present on a grid having three times less spatial resolution as the LAM grid. All simulations in this section are run with hourly LBC updates to minimize the interpolation error documented above. Test runs using more frequently updated LBCs did not change the results presented below. This analysis procedure is similar to that used previously by Laprise et al. (2000) and De Elía and Laprise (2002).

### 5.2.1 Example of LBC "Sweeping"

The example in section 5.1.1.1 (Fig. 5.1) showed how interpolation of LBCs between 3-hourly updates introduced LBC error through a pulsing effect. That same case is run again, but this time using the configuration described above with low-pass filtered, hourly-updated LBCs.

Error growth during the first 1.5 hours of the new example simulation (Fig. 5.18) is much like that of the previous example (Fig. 5.1). Errors are introduced within the peripheral boundary zone because of the lack of shortwave features in the LBC. The errors continue propagating inward without interruption, rather than impulsively as before. Indeed, the LBC error pulse seen previously is not defined for the present configuration because the low-pass filtered LBC always contains error relative to the fully resolved control simulation.

The rate at which errors propagate into the LAM domain is about the same as the previous case (compare panels f, j, and r in both figures). However, vorticity errors are less intense for this example and streamfunction errors are barely discernible. Whereas temporal interpolation of LBCs affects the amplitudes of all nonstationary wavelengths, the low-pass filtered LBCs are perfect at wavelengths greater than 450 km. This explains the difference in intensity of the errors since fewer wavelengths are affected by spatial filtering than by temporal interpolation.

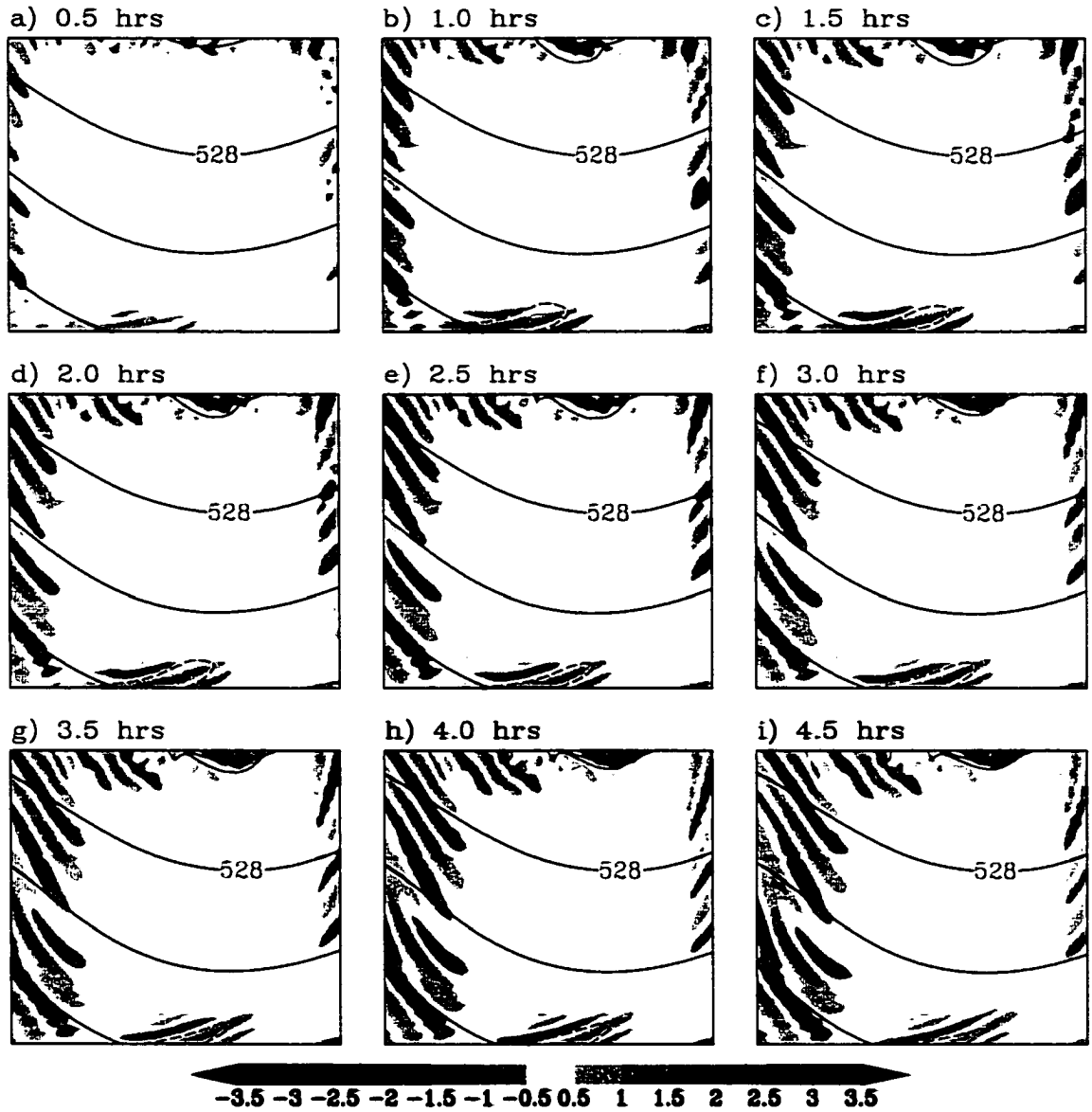



Figure 5.18: Example LAM simulation run on the small, center domain (1500 km$^2$, Fig. 3.2) configured with hourly updated LBCs that are low-pass filtered to remove spatial scales smaller than 150 km. Gentle curves are streamlines from the control simulation run on the full global domain ($12 \times 10^6$ m$^2$s$^{-1}$ intervals). Vorticity errors ($\times 10^{-5}$ s$^{-1}$) are shaded while (positive/negative) streamfunction errors are shown with (solid/dashed) contours at $\pm 5 \times 10^4$ m$^2$s$^{-1}$ intervals, or $\sim$0.01% (zero line omitted).

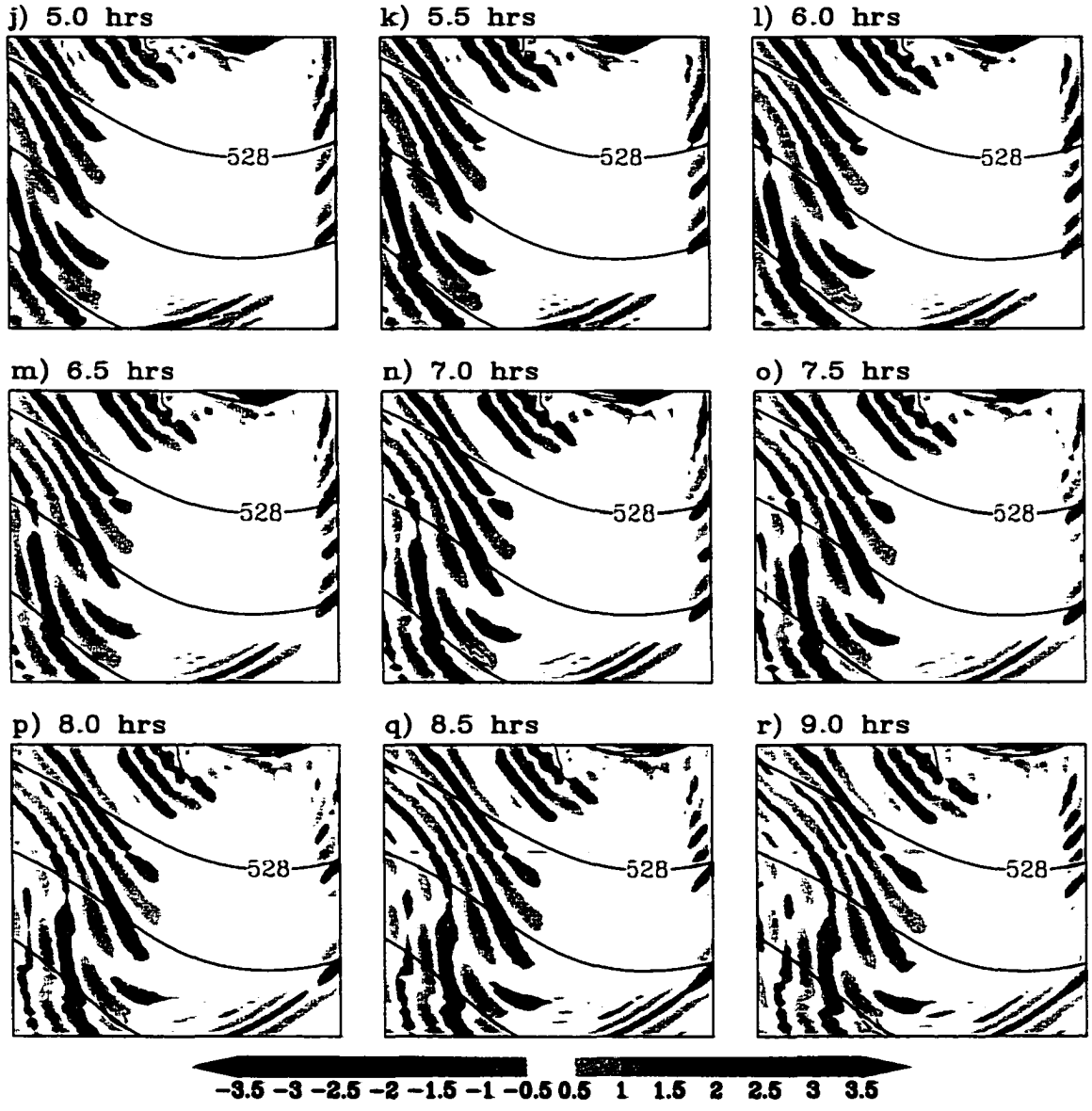



Figure 5.18 (continued).

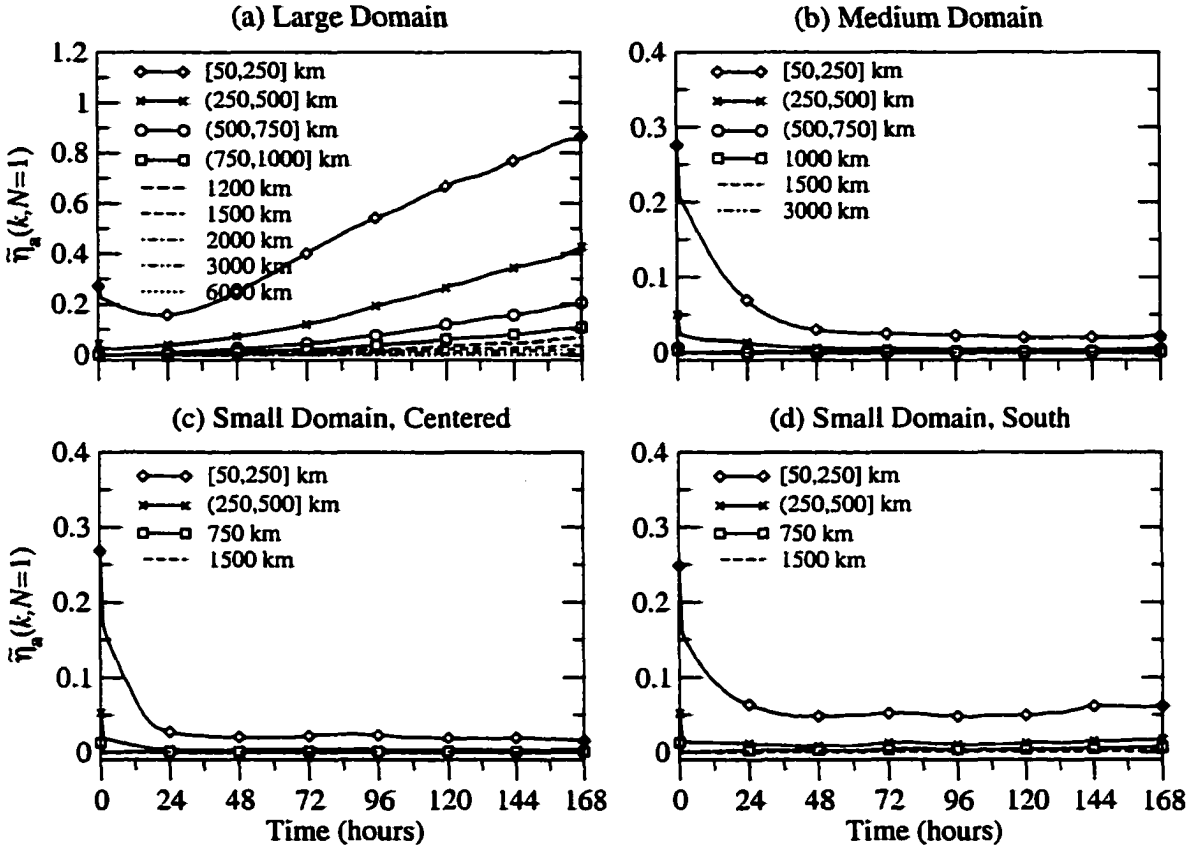



Figure 5.19: Normalized vorticity error variance (Eq. 2.22) obtained as an average over 100 independent LAM simulations. Hourly-updated LBCs are provided by a direct subset of the global fields. Initial conditions are also given by the global simulation, but low-pass filtered to remove wavelengths smaller than 150 km.

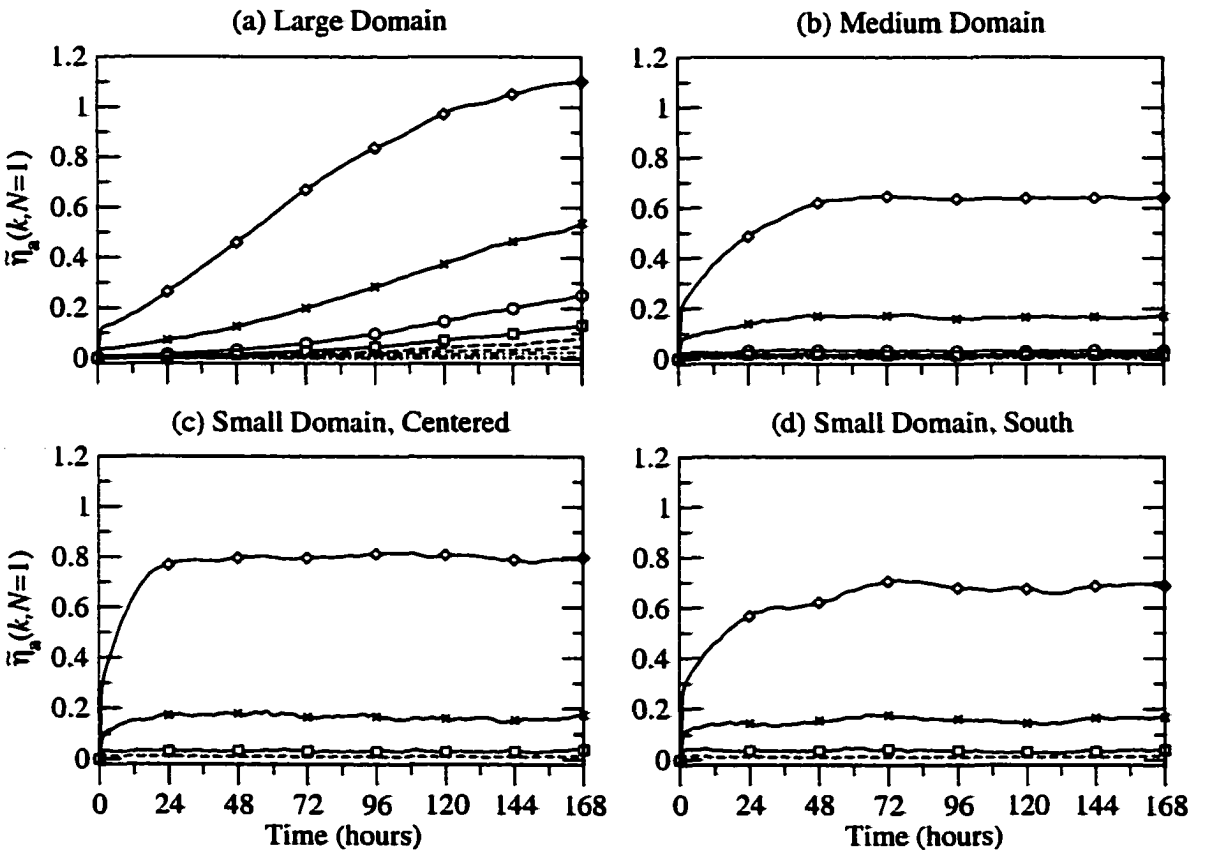



Figure 5.20: As in Fig. 5.19, including line labels, except the initial condition is perfect and the LBCs are low-pass filtered to remove wavelengths smaller than 150 km.

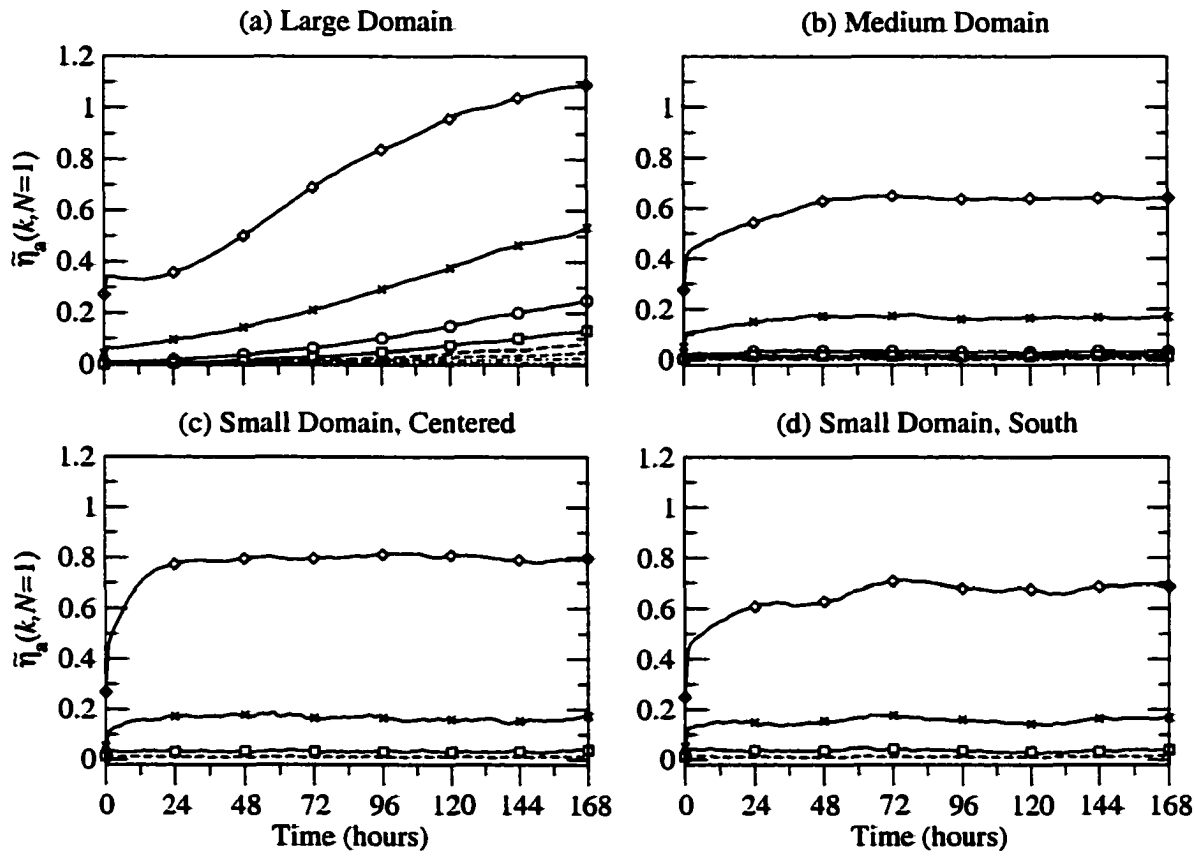



Figure 5.21: As in Fig. 5.19, including line labels, except both initial and lateral boundary conditions are low-pass filtered to remove wavelengths smaller than 150 km.

time. Although the low-pass filter does not affect wavelengths longer than 450 km, variance spectra amplify at larger scales because the LBC errors have adequate time to grow upscale while passing through the large domain. On the medium and small domains, errors appear immediately and grow toward saturation at wavelengths less than 500 km. Error variances in the 50-250 km wavelength band saturate at values from about 0.65 to 0.8. These values are more than 60% smaller than the expected value of 2. This behavior is similar with that seen previously in Figs. 5.10 and 5.11 — although without the pulsing effect — for error growth due to 3- or 6-hour temporal interpolation of LBCs. In drawing this comparison, note that the saturation values on the small and medium domains are less than those for the cases having temporally interpolated LBCs. The smaller errors were seen in the example case above (Fig. 5.18) and may be attributed to the sweeping through of LBCs that are perfect at scales greater than 450 km.

The third LAM configuration (Fig. 5.21) combines the previous two, wherein both the initial and lateral boundary conditions are low-pass filtered to remove small-scale features. This configuration is analogous to situations where initial and lateral boundary conditions for a LAM forecast are obtained by interpolating fields from a lower resolution global model forecast. The results exhibit characteristics featured in both the previous configurations. Specifically, error variances first jump sharply, then grow toward saturation at values less than 1 on the smaller domains. On the large grid, error variance spectra in the 50-250 km wavelength band decline slightly in the first 24 hours as the flow organizes to smaller scales, but then grows without limitation along with other scales.

Error growth behaviors presented for the experiments in this section are consistent with those from previous studies considering the impact of LBCs on error growth in LAMs (e.g. Vukicevic and Errico 1990; Laprise et al. 2000; De Elía and Laprise 2002). Specifically, coarsely resolved LBCs emulated through low-pass filtering constrain the saturation value of error growth by removing variance at small scales while sweeping through the LAM domain. The error growth constraints imposed by coarsely resolved LBCs are strongest on smaller domains and overwhelm those due to coarsely resolved initial conditions within about 24- to 48-hours.

### 5.2.3 Ensemble Error Variance Spectra

One-hundred independent 10-member ensemble simulations were run using two of the configurations above. One configuration has low-pass filtered initial conditions while the other does not. Both have hourly updated, low-pass filtered LBCs. Initial conditions and LBCs for individual LAM ensemble members are obtained as direct subsets of the corresponding global ensemble members (Sec. 4.1.1), with filtering applied where specified. Error growth in the ensemble occurs due to initial condition perturbations, use of "one-way" LBCs, and lack of small scales in filtered initial or lateral boundary conditions. Comparison of error variance spectra to those from the global ensemble simulations helps to measure the impact of the coarsely resolved LBCs as shown in section 5.1.3.

When simulations are configured with unfiltered initial conditions and low-pass filtered LBCs (Fig. 5.22), error variances in the 50-250 km wavelength band grow rapidly and reach saturation after about 24 to 36 hours. The shortwave error variances saturate at values that are about 20% smaller than those from the global ensembles on the large grid and about 40-50% smaller on the medium and small grids. Furthermore, the error variances are constrained at larger scales up to about 750 km. It is clear from these results that the absence of small scales in the LBCs places severe constraints on the maximum values of variance spectra at small scales. Indeed, the constraint is stronger than that caused by the filtering effect of interpolating between LBC updates (Fig. 5.13).

It is interesting that the spatial filtering of LBCs imposes stronger constraints on error growth even though fewer wavelengths are removed compared to those lost by LBC interpolation. Both LBC effects limit error growth in the ensembles by sweeping the initial perturbations from the LAM domain at small scales. However, the LBC errors caused by temporal interpolation appear strong enough to partially restore error variance growth due to the *artificial* effects of aliasing errors in place of the natural growth of initial condition uncertainties. From this perspective, temporal interpolation between LBC updates has the greater impact on ensemble error variance, if only because natural perturbation error growth is supplanted by artificial

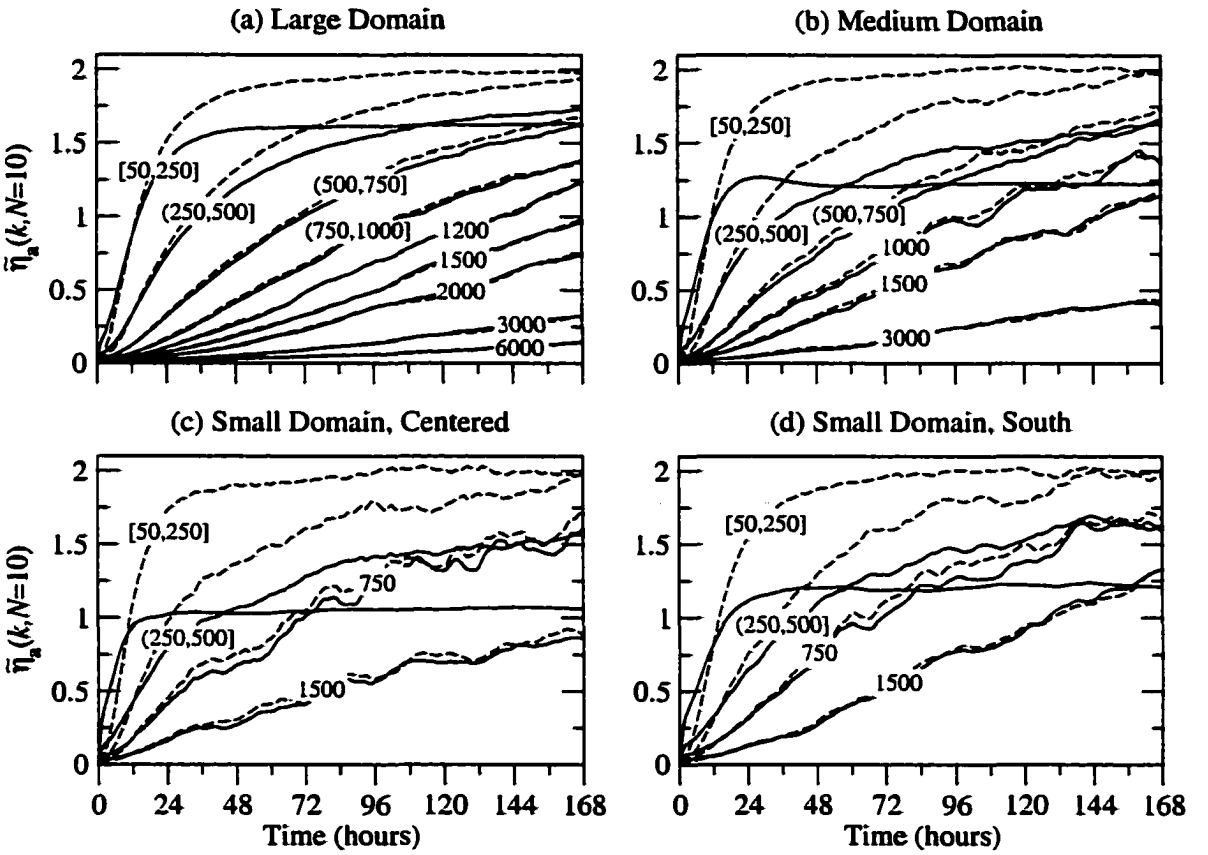



Figure 5.22: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member LAM ensemble simulations having hourly-updated, low-pass filtered LBCs. Line labels (km) indicate wavelengths(s) contributing to error variances. Dashed reference lines are reproduced from Fig. 4.4, showing error variances from subsets of the global ensemble simulations.

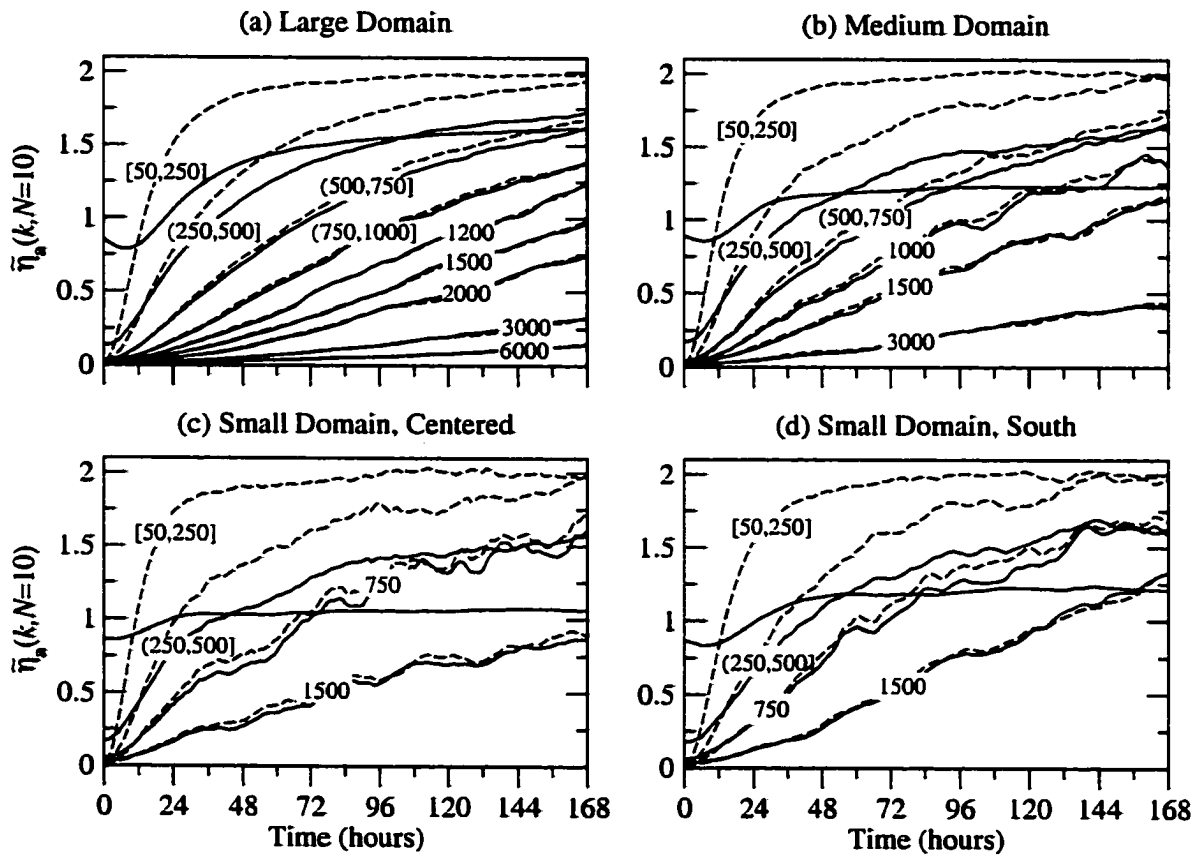



Figure 5.23: As in Fig. 5.22, except both initial and lateral boundary conditions are low-pass filtered.

LBC error growth. Hence, any attempt to construct dynamically conditioned initial perturbations for LAM ensembles could be mooted by the sweeping of artificial errors associated with temporal interpolation of LBCs.

Similar results are obtained when both initial conditions and LBCs are low-pass filtered (Fig. 5.23). The primary difference for this configuration is that the shortwave error variance spectra are already near saturation at the start of the simulations. There is some reduction of variance in the 50-250 km wavelength band during the first 12 hours as wave structures redevelop downscale. However, the perturbation error growth and LBC sweeping effects take over quickly and bring the error variances to the same saturation limits seen above with unfiltered initial conditions.

### 5.2.4 Ensemble Summary Statistics

The same set of ensemble summary statistics are presented as in section 5.1.4 to quantify the impact of error variance constraints on LAM ensemble dispersion. Recall that if the spatial bias and ensemble mean error terms remain unchanged between the global and LAM ensemble simulations, then changes in dispersion are directly linked to the changes in the variance spectra.

Consider summary statistics from the configuration having fully resolved initial conditions and low-pass filtered LBCs from the global ensemble (Fig. 5.24). The LAM ensemble dispersion is less than that of the global ensembles on all domains. The loss of dispersion is most notable on the medium and small domains, where it is reduced by about 10 to 40%. Since the bias terms remain nearly unchanged, the loss of dispersion is attributed directly to the loss of error variance at small scales. Hence, the use of coarsely resolved LBCs imposes a similar, but slightly stronger constraint on LAM ensemble dispersion as compared to constraints imposed by the use of 3- or 6-hourly updated LBCs (compare to Fig. 5.16).

Similar results are obtained for the configuration having low-pass filtered initial conditions and LBCs from the global ensemble (Fig. 5.25). In this configuration, the coarsely resolved initial fields on all LAM domains produce an increase in $S^2$, $\sigma^2$, and the ensemble mean error for the first 24 hours of the simulations. As discussed previously (Sec. 5.1.4), LAM ensemble dispersion remains less than global ensemble

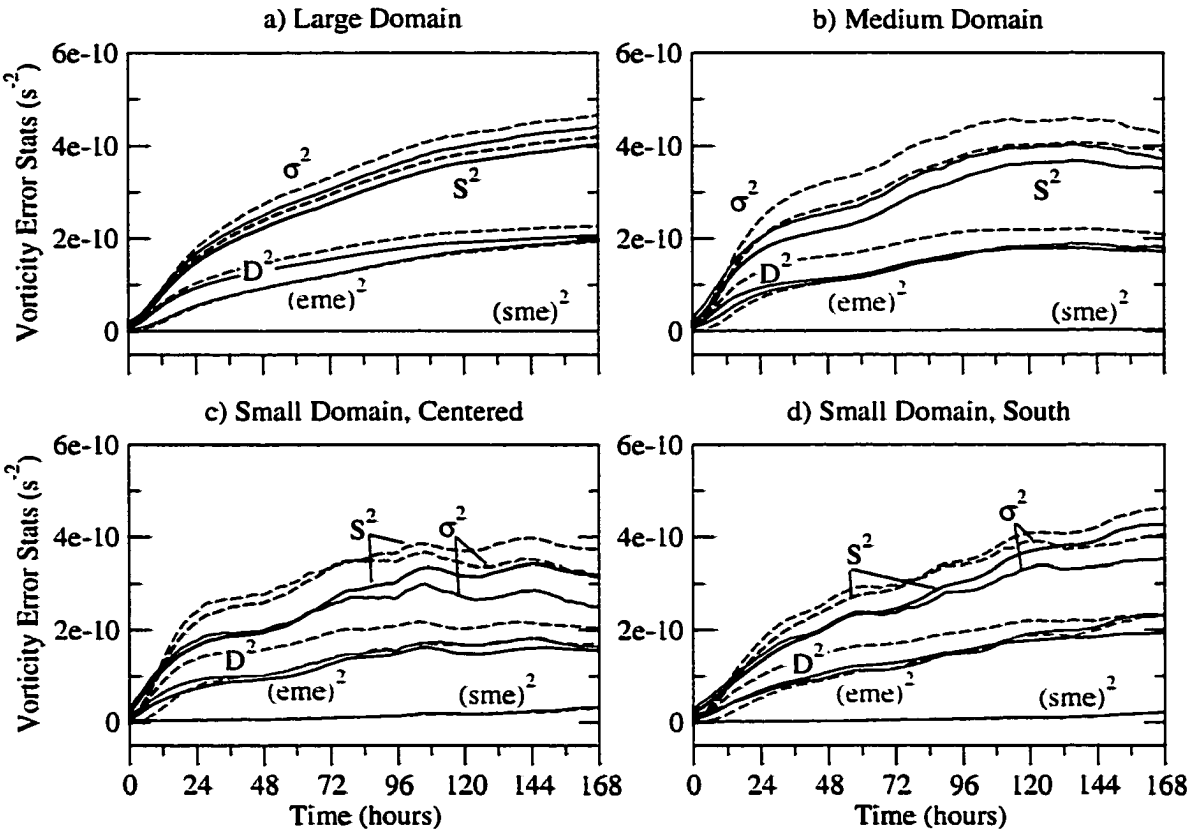



Figure 5.24: Summary statistics for vorticity, averaged over 100 independent 10-member LAM ensemble simulations. LBCs are updated hourly and low-pass filtered to remove wavelengths less than 150 km. See text (page 57) for a description of variables shown. Dashed reference lines are reproduced from Fig. 4.4, showing statistics from the global ensemble simulations.

dispersion over this period because the increase in $\sigma^2$ is negated by the increase in ensemble mean error (see Eq. 2.11). Note that after about 24-36 hours, the ensemble statistics are exactly the same as in the previous case (Fig. 5.24). This result has an interesting implication; initializing LAM ensembles with high resolution initial conditions is not enough to reduce the constraints imposed by LBCs on error variance growth or ensemble dispersion.

## 5.3 Combined LBC effects

LAM configurations run in the previous two sections were designed specifically to isolate the impacts of temporal interpolation between LBC updates and coarsely resolved LBCs. The combined impact of these two individual effects is now evaluated using model configurations having low-pass filtered initial and lateral boundary conditions updated at 3-hour intervals. Simulations were also run using low-pass filtered, 6-hourly updated LBCs. Statistics from these simulations are not shown because they are nearly the same as those from simulations having 3-hourly updated LBCs.

### 5.3.1 Ensemble Error Variance Spectra

Results shown in Figs. 5.26 and 5.27 having low-pass filtered, 3-hourly updated LBCs are nearly the same as in Figs. 5.22 and 5.23. Specifically, error variance spectra on the medium and small domains are reduced by up to 50% at small scales due to constraints imposed by sweeping of coarsely resolved exterior fields through the domain. This similarity of results supports the argument that inadequate spatial resolution in LBCs is the dominant effect contributing to restraints on error growth in LAMs.

There are two unique attributes in the error variance spectra resulting from the combined spatial and temporal filtering of LBCs. First, the LBC error pulse errors associated with temporal interpolation are absent from the 50-250 km wavelength band. This is not surprising since wavelengths less than 150 km have been removed from the external fields by the low-pass spatial filter. The second attribute is that the LBC filtering has constrained error variance growth into longer wavelengths up

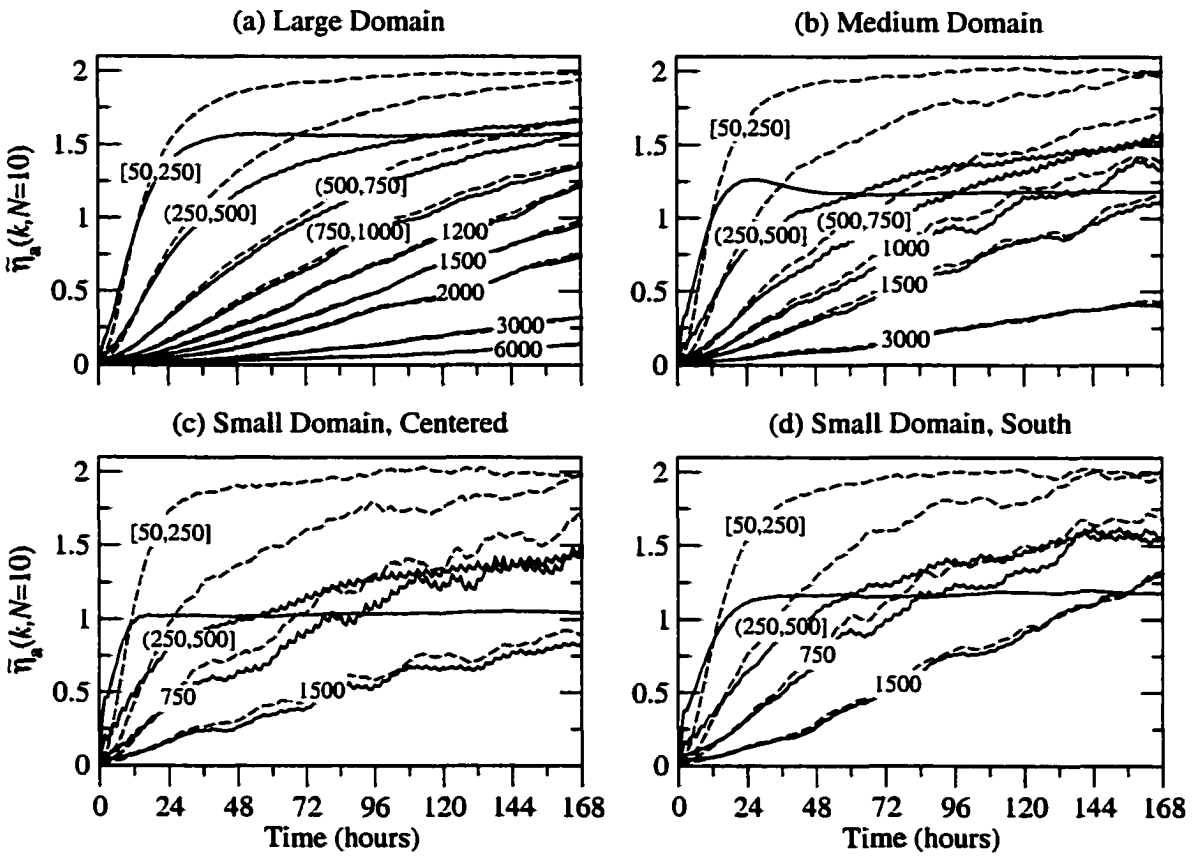



Figure 5.26: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs. Line labels (km) indicate wavelengths(s) contributing to error variances. Dashed reference lines are reproduced from Fig. 4.4, showing error variances from subsets of the global ensemble simulations.

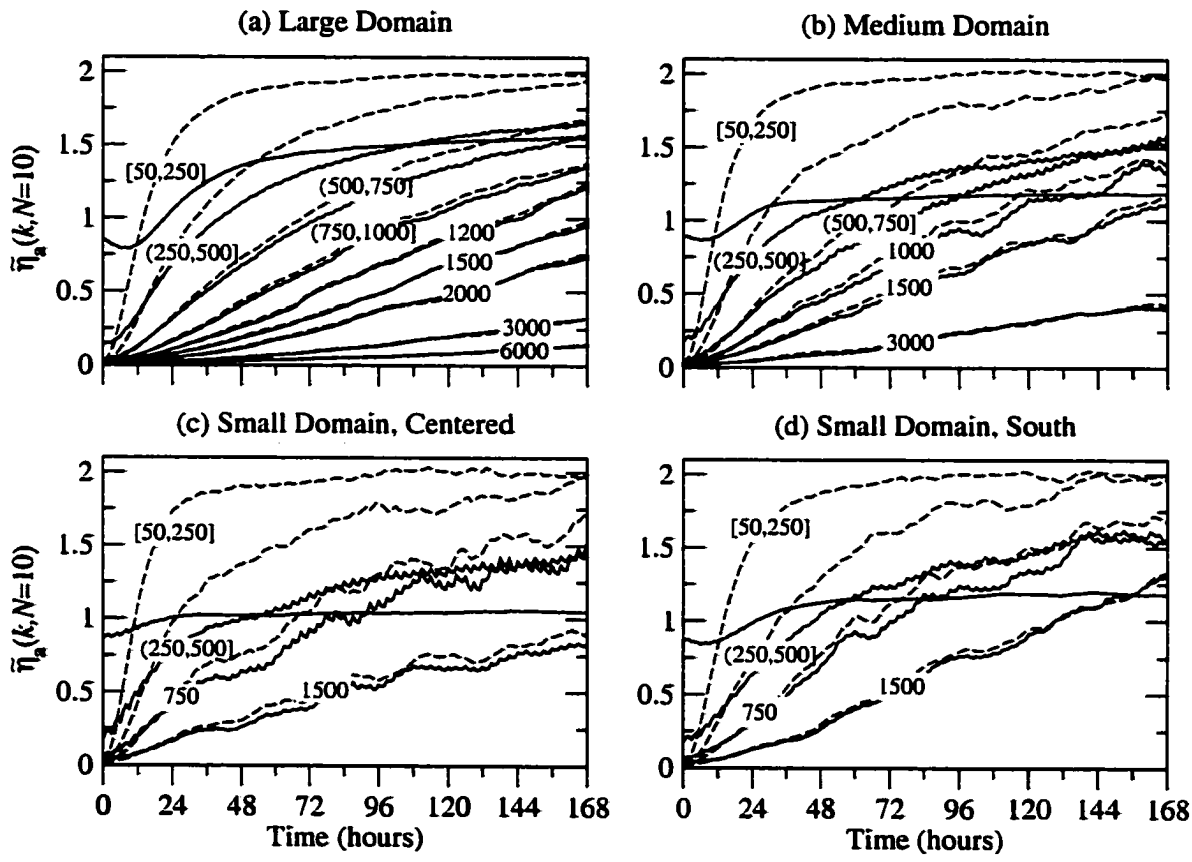



Figure 5.27: As in Fig. 5.26, except both initial and lateral boundary conditions are low-pass filtered.

time, the dispersion loss ratio remains evenly distributed and near zero within the downstream portion of the domain. The reductions near the upstream boundary is obviously attributed to the spatial and temporal filtering effects associated with the LBCs. The area impacted by loss of dispersion grow with time as LBC errors sweep through the domain from west to east. Locally, the LAM ensemble dispersion loss ratio shows reductions of a factor of eight or larger. When averaged across the domain, the reduction of LAM ensemble dispersion shown by this example is consistent with the ensemble summary statistics shown above (e.g., Fig. 5.24).

## 5.4 Chapter Summary

The LAM simulations presented in this chapter were designed specifically to isolate the effects of infrequently updated or coarsely resolved LBCs. A simplified model configuration was used so that the accuracy of LAM simulations could be measured relative to a model-generated truth in the absence of dynamical or numerical errors. Errors were evaluated using a variety of statistics averaged over 100 independent cases. Specifically, the scale dependence of LBC effects was measured in terms of error variance spectra. A new expression was developed and applied to show for the first time that constraints on LAM ensemble dispersion are controlled by constraints on error variance growth at small scales.

Example simulations presented in Sec. 5.1.1 showed that linear interpolation in time between available LBC updates generates error growth through a mechanism dubbed LBC error "pulsing". Once LBC errors form within the peripheral boundary zone, they propagate, or "sweep" completely through the LAM domain over a time period determined by domain size and advection speed. Temporal interpolation acts as a spatial filter, removing completely all wavelengths passing through the boundary between LBC updates. Rapidly moving waves that are not adequately sampled between LBC updates are aliased to larger scales, therein placing a lower bound on the minimum scale at which LBC errors are introduced. Interpolation may also reduce the amplitude of any nonstationary wave moving through the lateral boundary.

Statistical results showed that hourly-updated LBCs produced minimal error growth on the medium and small domains. In contrast, short wavelength (50-250 km)

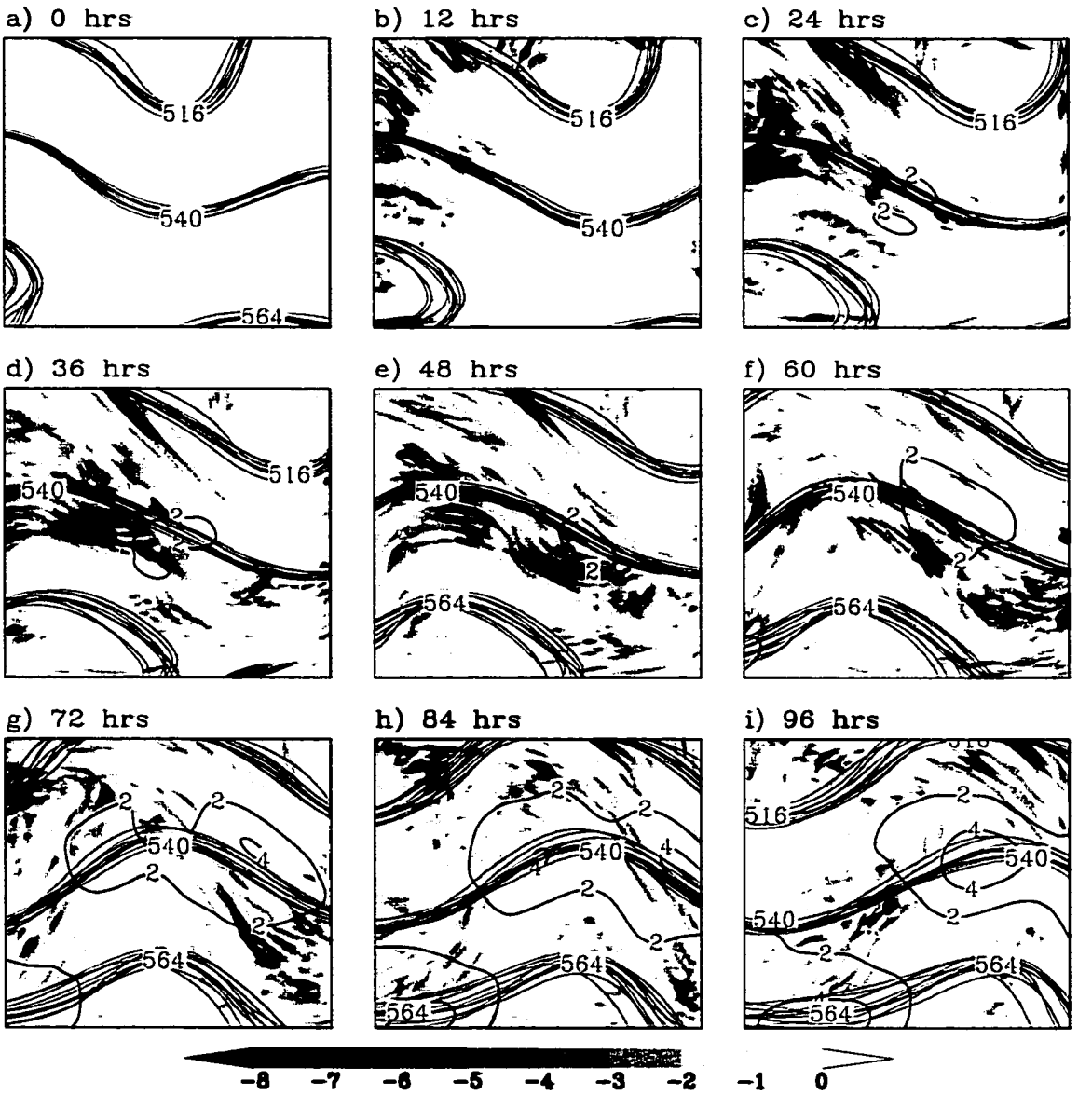



Figure 5.28: Example case run on the medium (3000 km$^2$) domain showing reduction of LAM ensemble dispersion (Eq. 5.1). A "spaghetti" plot drawn with solid black lines shows the (516,540,564)$\times 10^6$ m$^2$s$^{-1}$ streamlines from each of the 10 LAM ensemble members. Reduction of vorticity dispersion is shaded, while streamfunction dispersion is shown with solid blue contours at $2 \times 10^{12}$ m$^4$s$^{-2}$ intervals.

ensemble simulations are nearly the same as those from the global ensemble simulations. Therefore, the loss of error variance at small scales is directly responsible for the constraints on LAM ensemble dispersion.

The impact of coarsely resolved LBCs or temporal interpolation of LBCs on error growth in LAM simulations is quite similar. Both act to remove small scale features from the external fields passing through the lateral boundary. In the case of ensembles, the LBC effects filter out perturbations growing at small scales and thus acting to constrain ensemble dispersion. These results confirm the hypothesis based on the error variance budget outlined in Sec. 2.4. Of the two effects considered here, the impact of coarsely resolved LBCs was shown to have a stronger impact than temporal interpolation. The simplified model approach used here removes other factors from consideration that could contribute toward total error. Nonetheless, the error growth behaviors shown here are certainly present to some extent in current atmospheric models. The unresolved question is how large is the impact relative to other modeling and LBC data deficiencies.

# Chapter 6

# Lateral Boundary Perturbations

Statistical results from the previous chapter showed that spatial and temporal filtering effects of "one-way" LBCs imposed constraints on error variance growth and LAM ensemble dispersion. In this chapter, a method is proposed to apply small-scale LBC perturbations at every time step of the LAM simulation. As designed, the perturbations are shown to help restore the small-scale error variance and ensemble dispersion lost due to the LBC constraints.

## 6.1 LBC Perturbations

### 6.1.1 Overview of Method

The basic motivation for developing LBC perturbations is the loss of small scale error growth due to the LBC filtering and sweeping effects documented in the previous chapter. It is impossible to construct an efficacious LAM ensemble system unless these constraints are relaxed. To this end, LBC perturbations are designed to counter these effects by creating a statistically consistent source of error growth along the lateral boundaries. The LBC perturbations continually propagate into the domain and replace small-scale errors that are otherwise swept away through the downstream boundary.

Application of LBC perturbations has not previously been attempted, and is different from using an ensemble of unique LBCs provided by individual members of an external model ensemble. Indeed, all the LAM ensemble configurations used in the previous chapter used LBCs given by the corresponding global ensemble members.

where $a_k$ and $b_k$ are real amplitude coefficients. The phase-angle form of the Fourier series is obtained by letting $a_k = c_k \cos(-\theta_k)$ and $b_k = c_k \sin(-\theta_k)$, where $c_k = \sqrt{a_k^2 + b_k^2}$ and $\theta_k$ is the phase angle for wavenumber $k$. Apply these definitions for $a_k$ and $b_k$ in (6.1) and manipulate so that

$$\begin{aligned} f(x) &= a_0 + \sum_{k=1}^{\infty} [c_k \cos(\theta_k) \cos(2\pi kx/L) - c_k \sin(\theta_k) \sin(2\pi kx/L)] \\ &= a_0 + \sum_{k=1}^{\infty} [c_k \cos(2\pi kx/L + \theta_k)] \\ &= a_0 + \frac{1}{2} \sum_{k=1}^{\infty} \left[ c_k e^{i(2\pi kx/L + \theta_k)} + c_k e^{-i(2\pi kx/L + \theta_k)} \right] \\ &= a_0 + \frac{1}{2} \sum_{k=1}^{\infty} \left[ c_k e^{i\theta_k} e^{i(2\pi kx/L)} + c_k e^{-i\theta_k} e^{-i(2\pi kx/L)} \right]. \end{aligned} \tag{6.2}$$

Equation (6.2) shows that a periodic function can be synthesized simply by specifying a real amplitude coefficient ($c_k$) and phase angle ($\theta_k$) for each wavenumber $k$. This form is useful because the one-dimensional variance spectra computed previously have retained only the magnitudes of complex Fourier coefficients. The phase angles remain unknown but may be specified randomly.

Fast Fourier transform (FFT) algorithms use the complex form of the Fourier series. To convert (6.2) to the more useful complex form, introduce complex Fourier coefficients

$$F(0) = a_0, \quad F(k) = \frac{1}{2} c_k e^{i\theta_k} = \frac{1}{2}(a_k - ib_k), \quad F(-k) = \frac{1}{2} c_k e^{-i\theta_k} = \frac{1}{2}(a_k + ib_k). \tag{6.3}$$

Apply these coefficients in (6.2) so that

$$\begin{aligned} f(x) &= F(0) + \sum_{k=1}^{\infty} \left[ F(k) e^{2\pi ikx/L} + F(-k) e^{-2\pi ikx/L} \right] \\ &= \sum_{k=0}^{\infty} F(k) e^{2\pi ikx/L} + \sum_{k=-1}^{-\infty} F(k) e^{2\pi ikx/L} \\ &= \sum_{k=-\infty}^{\infty} F(k) e^{2\pi ikx/L}. \end{aligned} \tag{6.4}$$

The extension of Eqs. (6.1) to (6.4) into their two-dimensional forms is not difficult but involves many additional terms. While details of the derivation are omitted (see Walker 1988), the complex form of the Fourier series for a two-dimensional periodic field $f(x,y) = f(x + L_x, y + L_y)$ is

$$f(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F(k,l) \exp[2\pi i(kx/L_x + ly/L_y)]. \tag{6.5}$$

The discrete Fourier series used for the calculations is

$$f(x,y) \simeq \sum_{\substack{k= \\ -N_x/2+1}}^{N_x/2} \sum_{\substack{l= \\ -N_y/2+1}}^{N_y/2} F(k,l) \exp[2\pi i(kx/L_x + ly/L_y)], \tag{6.6}$$

where $x = (k + N_x/2 + 1)\Delta x$ and $y = (l + N_y/2 + 1)\Delta y$. Even integers $N_x$ and $N_y$ denote the number of grid points along each dimension of the domain. In practice, Fourier series approximation of real fields makes use of complex conjugate symmetries so that the negative $l$ wavenumbers are omitted (Press et al. 1996).

Equation (6.6) can be used to synthesize a field having pre-determined variance spectra $|F'(k,l)|^2$ and random phase angles $\theta_{k,l}$ by specifying

$$F(k,l) = \sqrt{\frac{|F'(k,l)|^2}{2}}(\cos\theta_{k,l} + i\sin\theta_{k,l}), \tag{6.7}$$

except the factor of one-half is omitted for $k = 0$ and $k = N_x/2$. This factor is required since error variance spectra obtained previously using FFT algorithms were multiplied by two because of the complex conjugate symmetry in transforms of real data.

#### 6.1.2.2 Amplitude of Perturbations

Amplitudes of the LBC perturbations are determined by the loss of error variance at specific wavelengths due to LBC effects. Thus, if $\sigma_\kappa^2$ denotes the one-dimensional

error variance spectra obtained from previous global and LAM simulations (see Appendix B and Sec. 2.1.3), amplitudes of the perturbation spectra are determined using (6.7) with

$$|F'(k,l)|^2 = \sigma_\kappa^2(\text{global}) - \sigma_\kappa^2(\text{LAM}), \tag{6.8}$$

where $\kappa = \sqrt{k^2 + l^2}$. The perturbation spectra are distributed equally among all the wavenumber pairs $(k,l)$ contained within each annular wavenumber ring $\aleph(\kappa) \pm (1/2)\delta\kappa$, where $\aleph$ denotes nearest integer. Variances are set to zero for wavenumber pairs where $\kappa$ exceeds that of the smallest resolved wavelength since these were not accumulated in the one-dimensional spectra. Negative values of $|F'(k,l)|^2$ are set to zero because the error variance in the LAM simulations already exceeds that of the global simulations.

Results in Fig. 6.1 show $|F'(k,l)|^2$ obtained for the LAM ensemble configuration having 3-hourly updated, low-pass filtered LBCs. These data reflect differences between the solid and dashed curves shown in Fig. 5.26, although now presented as a function of wavelength at specific times. Given these difference spectra, amplitudes of the LBC perturbation field are set to zero for about the first 12 hours, depending on wavelength. The amplitude of the perturbation is greatest at wavelengths between about 100 and 1000 km. Indeed, these are the scales that were most strongly effected by the filtering effects associated with spatial and temporal filtering of LBCs. Difference spectra are not shown beyond 72 hours because there is minimal additional growth beyond this time. The LBC perturbation field constructed using the difference spectra in Fig. 6.1 begin with zero amplitude, then begin to grow after about 12 hours until reaching a nearly constant value around 72 hours.

The difference spectra for these simulations were computed each hour. These spectra were interpolated linearly in time before generating LBC perturbation fields at every 7.5-minute time step. Temporal interpolation of the spectra before generating the perturbation field does not reduce small-scale variance as does interpolation between external LBC fields. Furthermore, temporal changes in the difference spectra are small since statistics were obtained as averages over 100 cases. In practical applications, data will not be available hourly, perhaps only every 3, or 6 hours. A possible approach for these scenarios is to fit analytic curves to the set of difference spectra (e.g. Lorenz 1982; Dalcher and Kalnay 1987; Schubert and Suarez 1989; Stroe

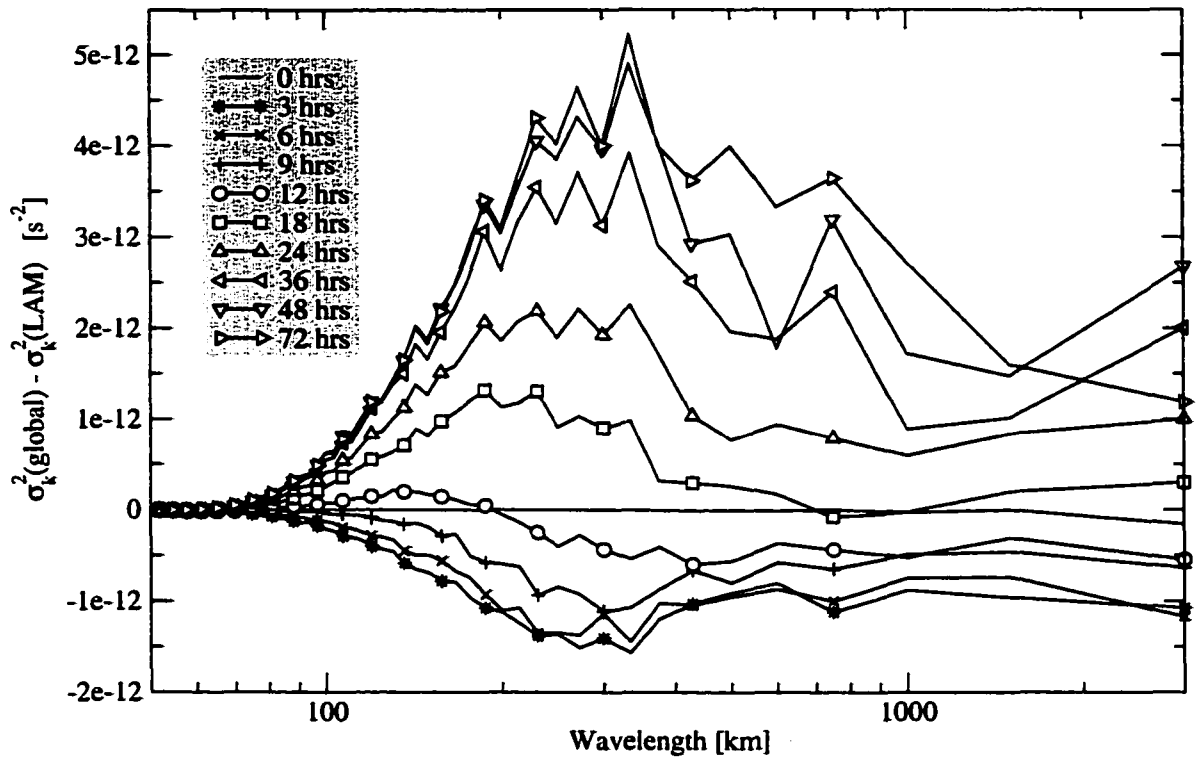



Figure 6.1: Difference in error variance spectra between global ensemble simulations and LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs. Compare to Fig. 5.26.

and Royer 1993; Reynolds et al. 1994; De Elía and Laprise 2002). Attempts were made to fit such curves here, but it was difficult to obtain parameters that produced accurate fits across all scales of motion. This is an issue that should be addressed further with application to more complex atmospheric LAMs.

#### 6.1.2.3 Translating the Perturbation Field

The perturbation field is initialized by specifying uniform random phase angles $0 \leq \theta_{k,l} \leq 2\pi$ in (6.7). Once initialized, the phase angles are stored and incremented at each time step to cause a translation of the perturbation field when it is synthesized using (6.6). This translation is important for providing temporally and spatially coherent wave structures as they pass through the lateral boundary.

The field is translated at some characteristic speed $c_\kappa$ by incrementing the phase angles such that

$$\theta_{k,l}(t + \Delta t) = \theta_{k,l}(t) + \kappa c_\kappa \Delta t, \tag{6.9}$$

where again, $\kappa = \sqrt{k^2 + l^2}$.

An appropriate choice of translation speed for this work is the Rossby phase speed. Since the available error variance spectra are one-dimensional, we use the isotropic, or uni-directional phase speed

$$c_\kappa = (U_0 k - \beta k/\kappa^2)/\kappa \tag{6.10}$$

Rossby phase speeds calculated using the parameters in Table 3.2 are less than 12 ms$^{-1}$. The $x$-component of group velocity remains near 12 ms$^{-1}$ while the $y$-component of group velocity is generally less than 0.001 ms$^{-1}$. Thus, the entire perturbation field translates from west to east at about 12 ms$^{-1}$. Other choices for the translation speed could be more appropriate in applications using full primitive equation models. This question remains beyond the scope of the present study.

### 6.1.3 Example Simulation with Perturbed LBCs

An example perturbation vorticity field ($\zeta'$) was constructed at multiple times using Eq. 6.6 with the difference spectra in Fig. 6.1. Results are shown in Fig. 6.2. Streamfunction perturbations are obtained by solving the Poisson equation $\nabla^2\psi' = \zeta'$. The solution to the Poisson equation is unique to within a constant value when using periodic boundary conditions. Therefore, the spatial mean $\langle\psi'\rangle$ was subtracted from each solution to ensure that the perturbation streamfunction remains unbiased.

As explained in Sec. 6.1.2.2, error variances from the LAM simulation are greater than those of the global simulations for about the first 12 hours. Therefore, the amplitude of vorticity perturbations are set to zero since the LAM simulation already has excessive error variance during this time. The impact of this choice is seen in Fig. 6.2a,b as the perturbation field does not begin to amplify until about 15 hours have passed. Careful examination of the vorticity perturbation field reveals about 10 to 20 wave couplets across the breadth of the 3000 km$^2$ domain. This result is

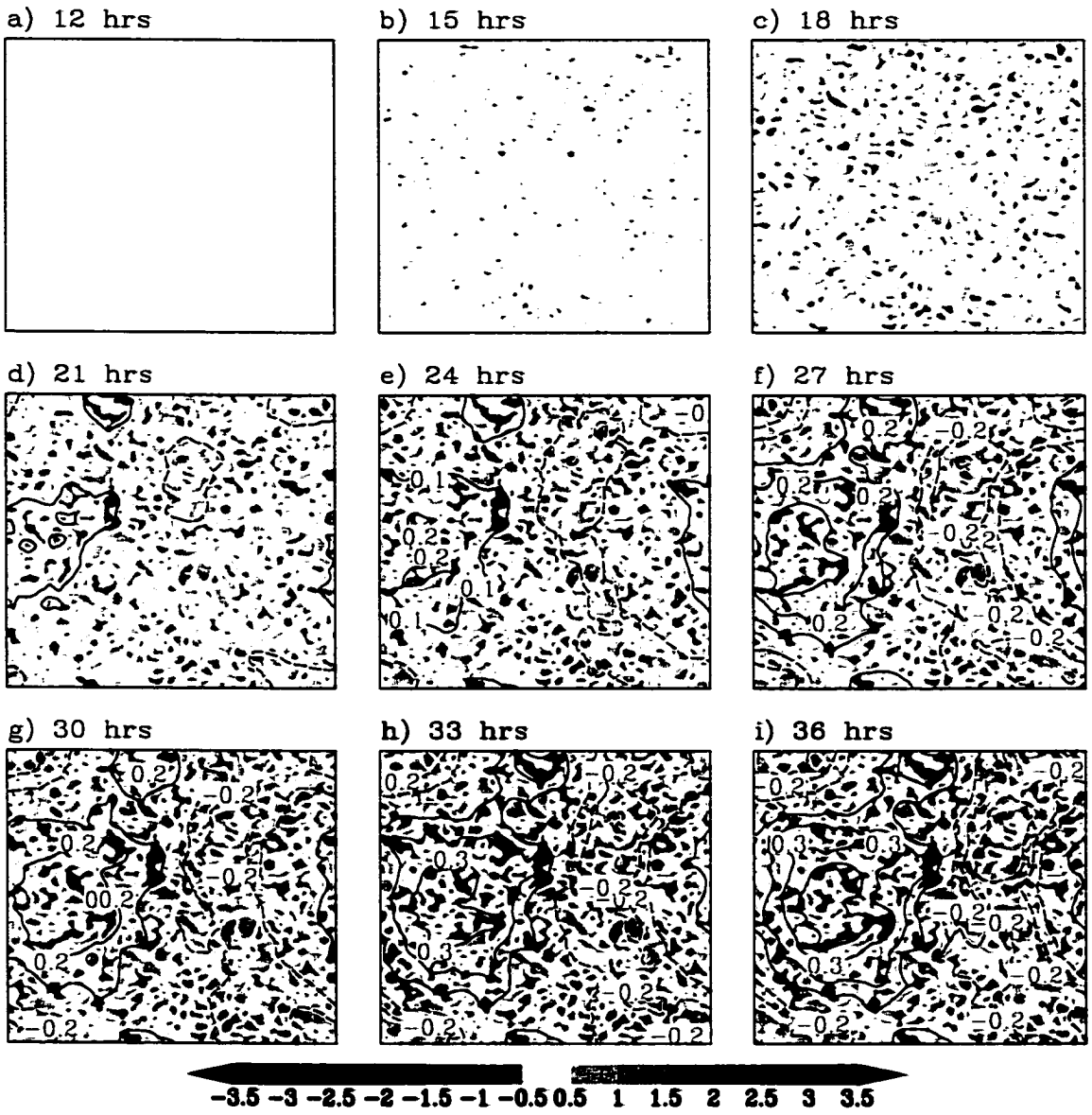



Figure 6.2: Example of LBC perturbation fields constructed for the medium domain using the difference in error variance spectra between global ensembles and LAM ensembles having 3-hourly updated, low-pass filtered LBCs (Fig. 6.1). Vorticity perturbations are shaded $\zeta' \times 10^{-5}$ s$^{-1}$ while streamfunction perturbations are contoured at $0.1 \times 10^6$ m$^2$s$^{-1}$ intervals (zero streamline omitted).

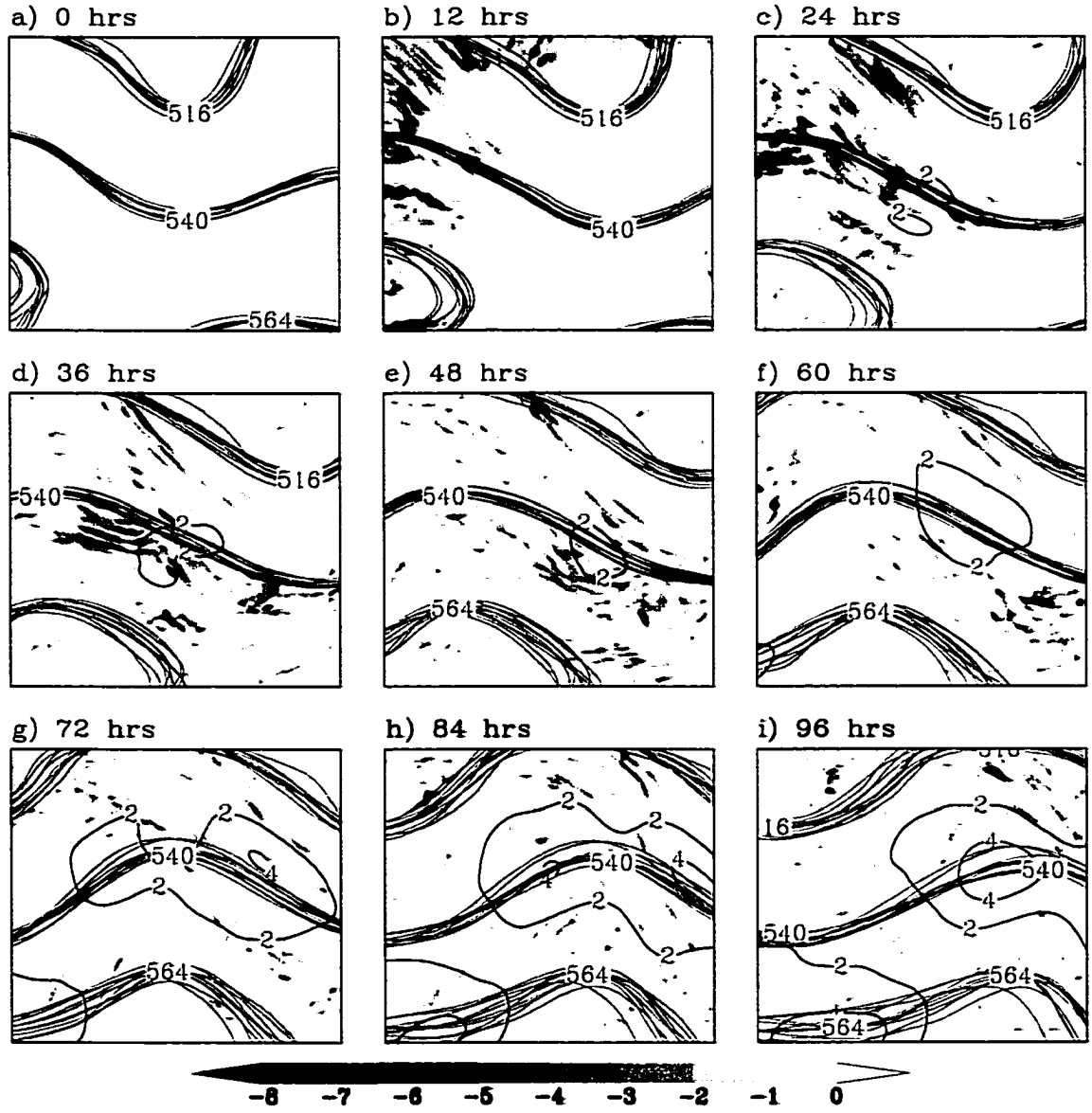



Figure 6.3: Example case run on the medium domain with 3-hourly updated, low-pass filtered, and perturbed LBCs showing reduction of LAM ensemble dispersion (Eq. 5.1). A "spaghetti" plot drawn with solid black lines shows the (516,540,564)$\times 10^6$ $m^2 s^{-1}$ streamlines from each of the 10 LAM ensemble members. Reduction of vorticity dispersion is shaded, while streamfunction dispersion is shown with solid blue contours at $2 \times 10^{12}$ $m^4 s^{-2}$ intervals. Compare to Fig. 5.28.

appear evenly distributed throughout the domain. The LBC perturbations do not apply instantaneously across the breadth of the domain, but instead propagate inward to restore those scales that have been filtered out by LBC filtering and sweeping effects.

## 6.2 Statistical Results

The LAM ensemble simulations run in chapter 5 are repeated here, except LBC perturbations are created and applied at each time step during the simulations as discussed above. The initial condition perturbations and the global ensemble members providing LBCs are exactly the same as in the previous simulations. As in chapter 5, statistical results are obtained as averages over 100 independent 10-member LAM ensemble simulations.

### 6.2.1 Ensemble Error Variance Spectra

Normalized error variances obtained from simulations having perturbed, hourly updated LBCs are shown in Fig. 6.4. Remarkably, error variances from the perturbed LAM simulations are nearly superimposed with those from the laterally unbounded global ensemble simulations. The LBC perturbations propagate into the LAM domain and restore almost all the error variance lost due to the filtering effects of temporal interpolation between LBC updates.

When LBCs are updated every three hours (Fig. 6.5), application of LBC perturbations completely restores error variances at wavelengths longer than about 500 km. The LBC perturbations are less effective for smaller scales, where the proportion of error variance restored depends on domain size. For example, on the large domain (panel a), the LBC perturbations restore about 1/3 of the error variance lost at saturation in the smallest scales due to temporal interpolation of LBCs. Compare this to the small, centered domain (panel c), where the LBC perturbations restore more than 3/4 of the error variance lost due to LBC effects in the unperturbed LAM simulations.

To explain these results, note that difference spectra used to determine the amplitude of LBC perturbations (Eq. 6.8) are based on error variance calculations obtained

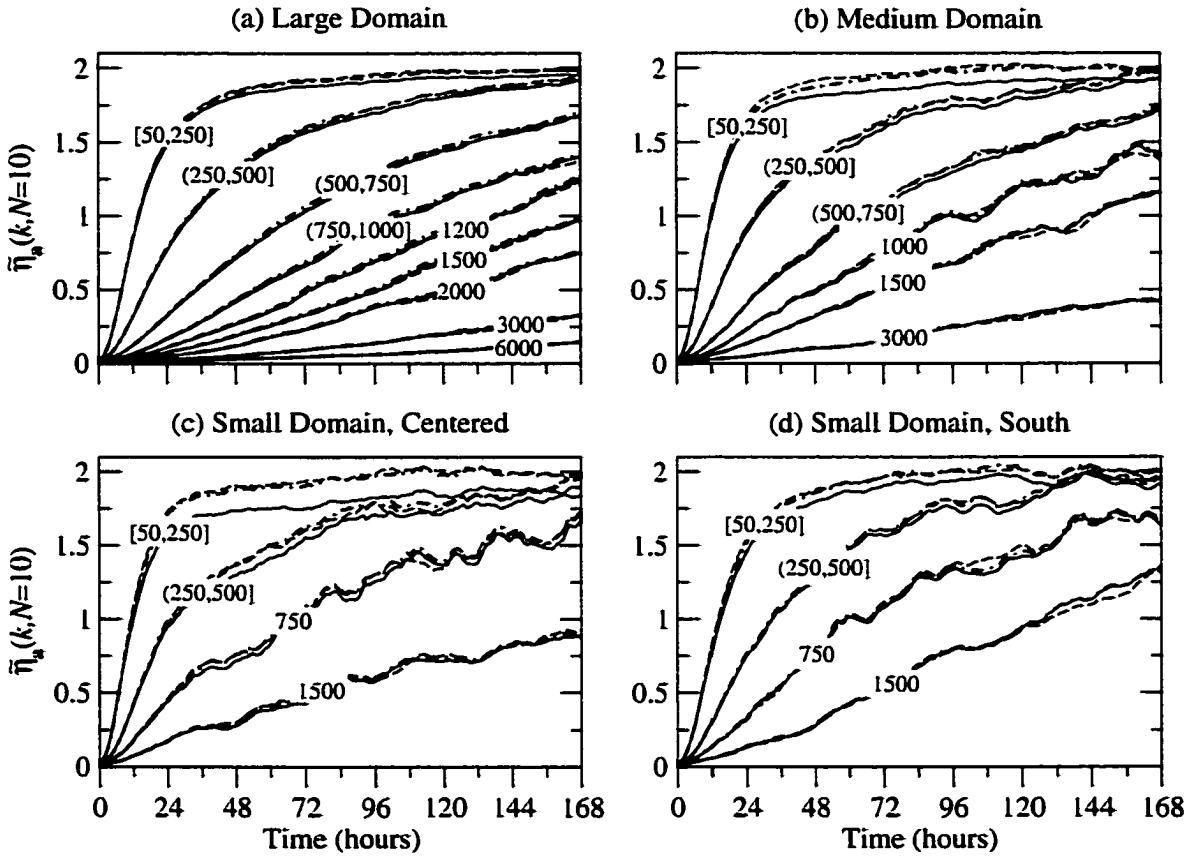



Figure 6.4: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member LAM ensemble simulations having perturbed, hourly updated LBCs (red, dash-dot lines). Line labels (km) indicate wavelength(s) contributing to error variances. Dashed reference lines are reproduced from Fig. 4.4, showing error variances from subsets of the global ensemble simulations. Solid lines are reproduced from Fig. 5.12, showing error variances from the corresponding LAM ensemble simulations run without LBC perturbations.

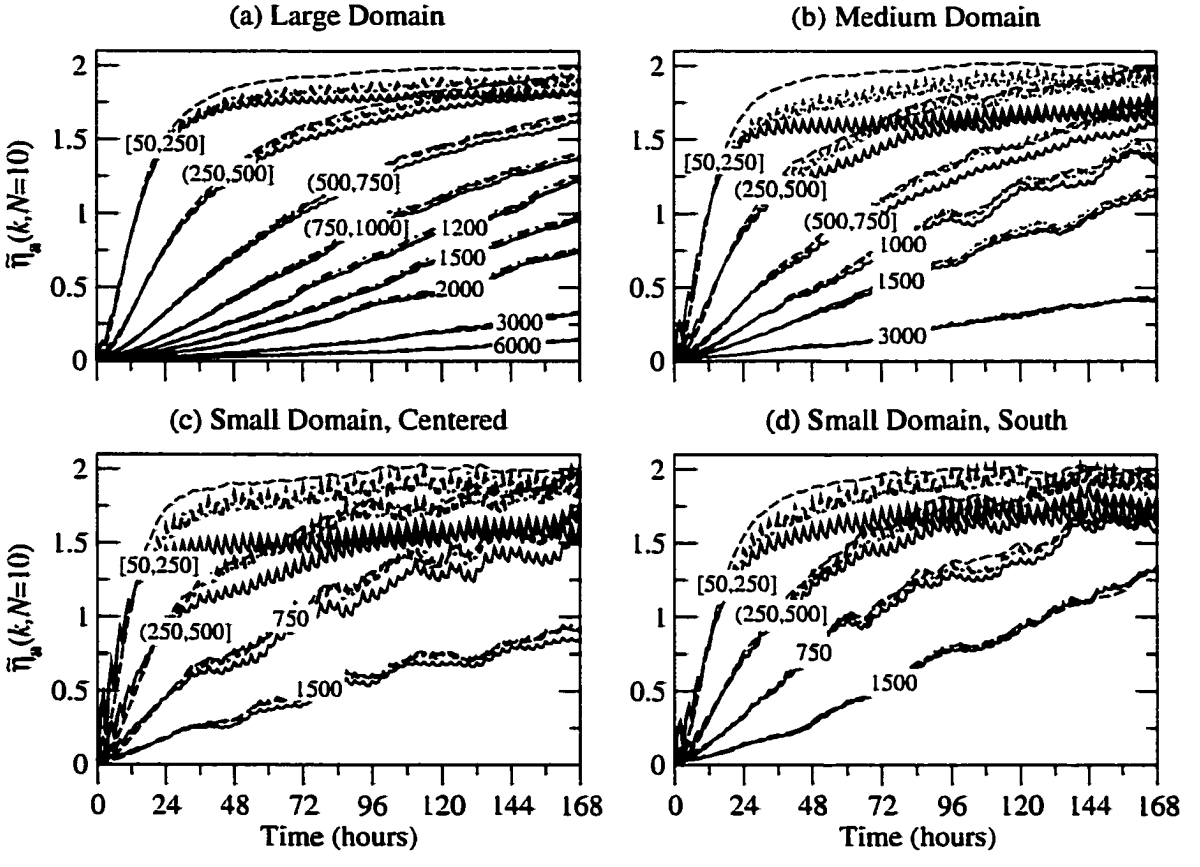



Figure 6.5: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member LAM ensemble simulations having perturbed, 3-hourly updated LBCs (red, dash-dot lines). Line labels (km) indicate wavelength(s) contributing to error variances. Dashed reference lines are reproduced from Fig. 4.4, showing error variances from subsets of the global ensemble simulations. Solid lines are reproduced from Fig. 5.13, showing error variances from corresponding LAM ensemble simulations run without LBC perturbations.

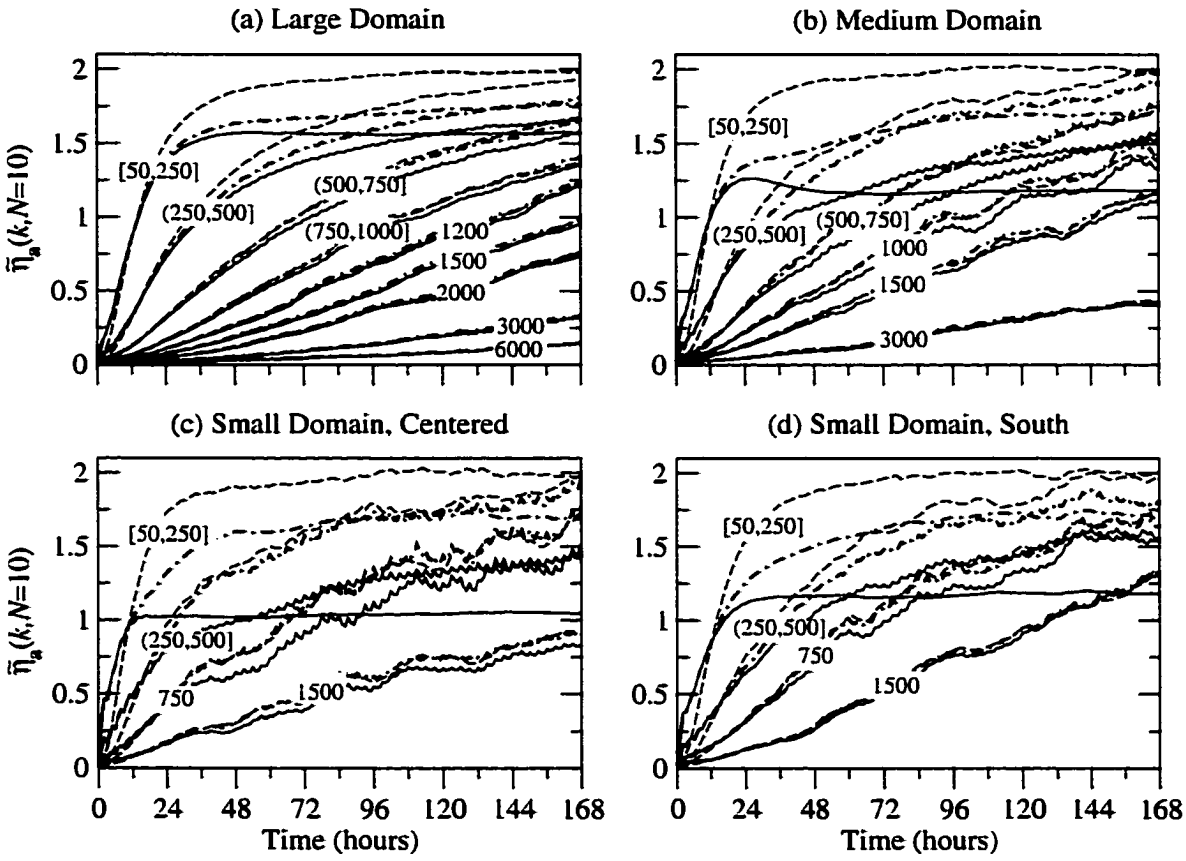



Figure 6.6: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member LAM ensemble simulations having perturbed, 3-hourly updated, low-pass filtered LBCs (red, dash-dot lines). Line labels (km) indicate wavelength(s) contributing to error variances. Dashed reference lines are reproduced from Fig. 4.4, showing error variances from subsets of the global ensemble simulations. Solid lines are reproduced from Fig. 5.26, showing error variances from corresponding LAM ensemble simulations run without LBC perturbations.

from data over the full extent of the LAM domain. However, LBC perturbations are applied only within the peripheral 7-point wave-absorbing zone. The perturbations subsequently disperse and/or dissipate while propagating through the LAM domain. Therefore, the difference spectra likely underestimate the amplitude of LBC perturbations needed to fully restore LAM error variances to those obtained from global simulations. The LBC perturbations are more effective on smaller domains because there is less time for dispersion and dissipation to reduce their impact while passing through the LAM domain. The perturbations also restore variance more effectively for larger scale waves since these have slower dispersion and less dissipation. Furthermore, the small-scale perturbations help maintain error variance at larger scales through nonlinear wave interactions and the upscale and downscale transfers of energy.

There are other interesting features seen in Fig. 6.5. First, note that variance spectra in the perturbed simulations are identical to those for unperturbed simulations over the first 12 to 24 hours. As discussed above, the amplitude of the perturbation field is set to zero during this time because the LAM variance spectra exceed those of the global simulations. Note also that the error variance curves continue to oscillate because of the LBC error "pulse" caused by temporal interpolation between otherwise perfect LBCs (see Sec. 5.1.1.1). Results are not shown from the LAM configuration having 6-hourly updated LBCs because they reveal similar features as those in Fig. 6.5.

Normalized error variances are shown in Fig. 6.6 for the LAM configuration having perturbed, 3-hourly updated, low-pass filtered LBCs. The effects of the LBC perturbations are similar to those noted previously in Fig. 6.5. Error variances are fully restored at wavelengths longer than 500 km, but the perturbations are less effective at smaller scales, especially on the large domain. As before, the LBC perturbation amplitudes needed to restore the lost error variance are likely underestimated by the difference spectra since the perturbations are applied only at the boundary and disperse and dissipate while passing through the LAM domain.

Comparison of error variances in the 50-250 km wavelength band in Figs. 6.5 and 6.6 shows that the LBC perturbations are least effective in combination with low-pass filtered external LBC fields. Similar results were seen in Sec. 5.2.3, wherein error variance constraints at small scales caused by coarsely resolved LBCs were

stronger than those due to temporal interpolation of LBCs. To explain these results, recall that fully resolved (unfiltered) external LBC fields are perfect each time new updates are given. The small-scale features present in the LBCs at these update times have greater amplitudes that those of the LBC perturbations constructed based on error variance difference spectra. Since they have greater amplitudes, shortwaves passing through the boundary at the LBC update times are less affected by dispersion and dissipation. Hence, error variances are almost completely restored as shown by the peaks in the 50-250 km error variance curve in Fig. 6.5c,d. In contrast, small-scale features are never retained in the low-pass filtered external LBC fields. Small scale error variance growth under this configuration is produced entirely by small amplitude LBC perturbations that are damped and/or dispersed quickly while passing through the LAM domain. Thus, error variances are restored less effectively compared to the LAM simulations configured with fully resolved external LBC fields (see Fig. 6.6c,d).

### 6.2.2 Ensemble Summary Statistics

The use of LBC perturbations effectively restored much of the error variance lost by coarsely resolved and temporally interpolated external LBC fields, especially on smaller domains and at wavelengths greater than 250 km. A primary goal in applying the perturbations is to restore LAM ensemble dispersion without adversely impacting the individual ensemble members. To see if this goal is achieved, ensemble summary statistics are presented for direct comparison to those seen previously in Sec. 4.3.

Results in Fig. 6.7 show that the total error variance is fully restored when LBC perturbations are applied in LAM configurations having hourly updated LBCs. This result is expected since error variance spectra were restored at all wavelengths for this configuration (Fig. 6.4). The use of LBC perturbations does not introduce additional spatial bias or ensemble mean error. Most importantly, ensemble dispersion is fully restored to values obtained from corresponding subsets of global ensemble simulations. These results fulfill the balances required by Eq. 2.11 and the error variance budget for LAMs outlined in Sec. 2.4.

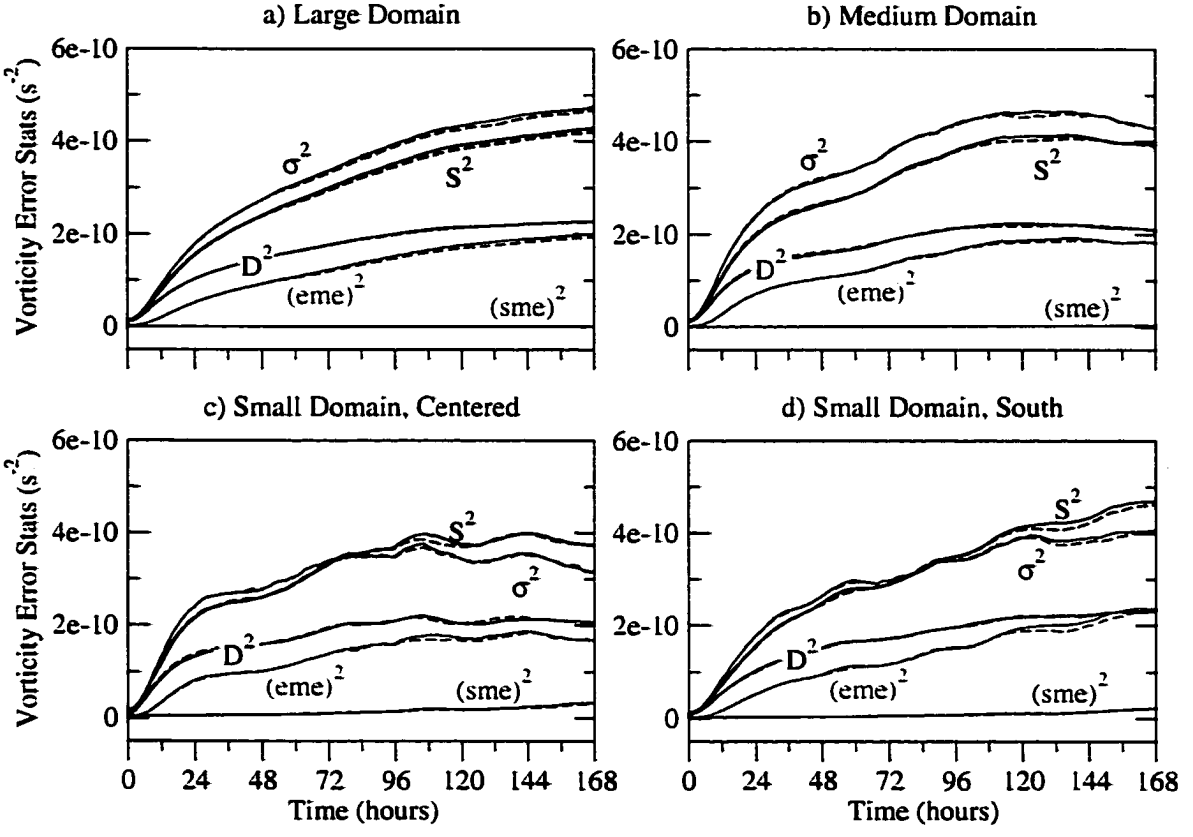



Figure 6.7: Ensemble summary statistics for vorticity, averaged over 100 independent 10-member LAM simulations having perturbed, hourly updated LBCs. See text (page 57) for a description of variables shown. Dashed reference lines are reproduced from Fig. 4.4, showing statistics from the global ensemble simulations. Compare to Fig. 5.15.

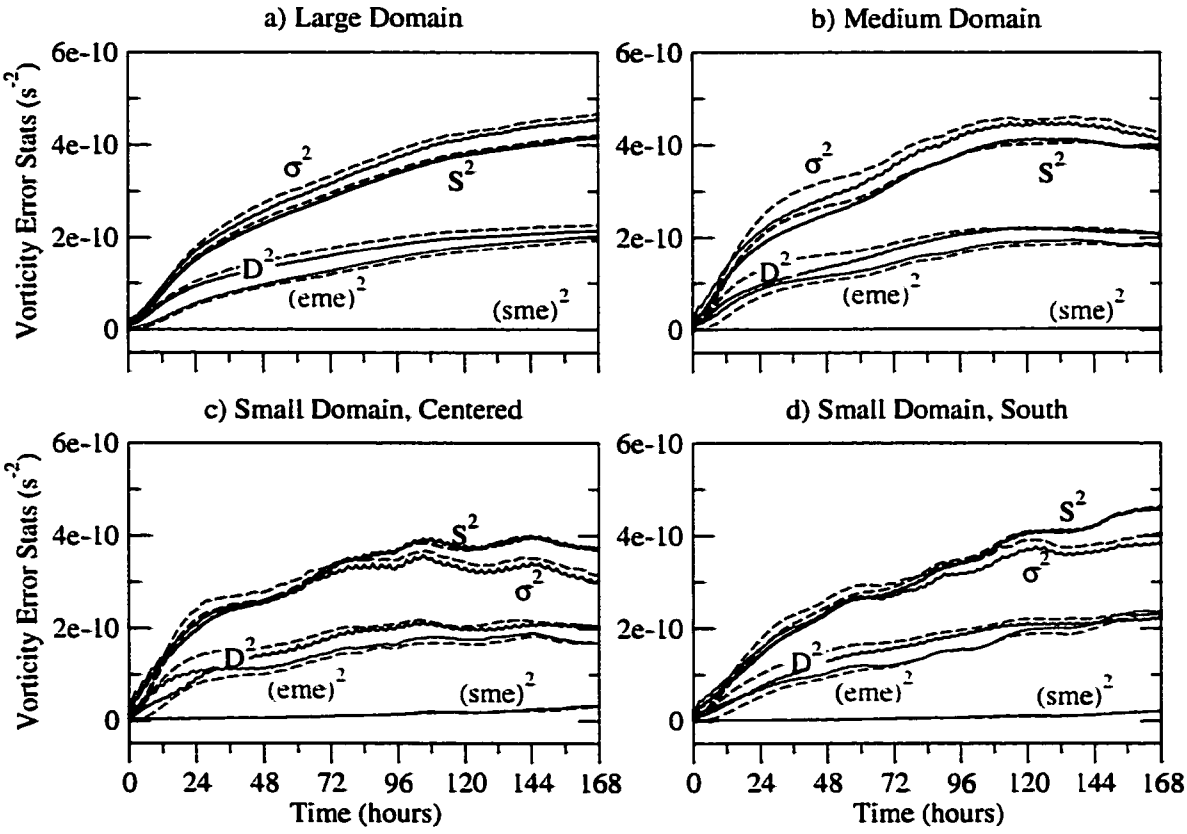



Figure 6.9: Ensemble summary statistics for vorticity, averaged over 100 independent 10-member LAM simulations having perturbed, 3-hourly updated, low-pass filtered LBCs. See text (page 57) for a description of variables shown. Dashed reference lines are reproduced from Fig. 4.4, showing statistics from the global ensemble simulations. Compare to Fig. 5.24.

When LBCs are updated every three hours (Fig. 6.8), the total error variance from LAM ensembles having perturbed LBCs closely matches that obtained from corresponding subsets of global ensemble simulations. The statistics continue to oscillate with time due to the LBC error pulse caused by temporal interpolation. Indeed, $\sigma^2$ most closely matches that of the global ensembles when it peaks at 3-hourly LBC update times for reasons discussed above (page 124). Additional spatial biases or ensemble mean errors introduced by the LBC perturbations are not discernible when comparing Figs. 6.8 and 5.16. Hence, after about 48 hours, the recovery of total error variance using LBC perturbations directly restores much of the LAM ensemble dispersion on all domains.

One of the more interesting features of Fig. 6.8 is the fact that $D^2$ remains less than that obtained from global ensembles between about 12-48 hours of the simulation. To explain this result, note that the LAM ensemble mean error exceeds that of the global simulations during this same period. This increase is caused by temporal interpolation of external fields between LBC updates, which removes small scale features from the forecasts compared to the fully-resolved analysis fields. The large domain (panel a) is less affected by interpolation errors because it takes longer than 48 hours for LBC sweeping to remove small scale features throughout the domain. The LBC perturbations contribute very little to changes in the ensemble mean error because they have such small magnitude compared to interpolation errors. Recall from Eq. 2.6 that

$$D^2 = S^2 - \left\|\bar{\mathbf{f}} - \mathbf{a}\right\|^2. \tag{6.11}$$

Hence, an increase in ensemble mean error causes a decrease in ensemble dispersion which cannot be recovered by the use of LBC perturbations as applied herein.

The same arguments apply to results from the LAM ensemble configuration having perturbed, 3-hourly updated, low-pass filtered LBCs (Fig. 6.9). Most of the total error variance has been restored through LBC perturbations. In turn, ensemble dispersion is largely recovered, except when reduced by the ensemble mean error created by coarsely resolved and temporally interpolated external LBC fields.

## 6.3 Scaled LBC Perturbations

Results shown above in section 6.2.1 reveal that the amplitude of LBC perturbations needed to fully restore error variance growth is underestimated under certain LAM configurations. Specifically, the perturbation amplitudes are underestimated at wavelengths shorter than 500 km, especially when applied to large domain simulations having coarsely resolved external LBCs updated at intervals of 3 hours or longer. In an attempt to remedy this deficiency, the perturbation amplitudes are increased by applying a scale factor to the difference spectra $|F'(k,l)|^2$ (Eq. 6.8). The scale factor is defined as ratio of error variances obtained from perturbed LAM ensembles to those obtained from global ensemble simulations. Thus, the difference spectra are redefined by introducing the scaling factor $\Lambda$ so that

$$\Lambda = \sigma_\kappa^2(\text{global})/\sigma_\kappa^2(\text{perturbed LAM}) \tag{6.12}$$

and

$$|F'(k,l)|^2 = \Lambda[\sigma_\kappa^2(\text{global}) - \sigma_\kappa^2(\text{unperturbed LAM})]. \tag{6.13}$$

Examples of the scale factor ($\Lambda$) that are applied to the error variance difference spectra seen previously in Fig. 6.1 are shown in Fig. 6.10. The scale factor is less than unity for about the first 12 hours because error variances from the LAM ensemble simulations are greater than those from the global simulations. Although less than unity, the scaling factor does not change the sign of $|F'(k,l)|^2$. Hence, the perturbation fields still are assigned zero amplitude while $|F'(k,l)|^2$ remains negative. The scale factors also are less than one for wavelengths shorter than about 60 km. This is of no significance since the difference spectra at these scales are so small (see Fig. 6.1). The scale factor exceeds unity after about 12 hours and grows toward values near 1.4 for wavelengths between about 70 and 500 km. These are the scales at which the amplitude of LBC perturbations was underestimated and error variances in LAM ensembles with LBC perturbations was not fully restored. After about 72 hours, error growth rates stabilize, so the scaling factor remains generally constant.

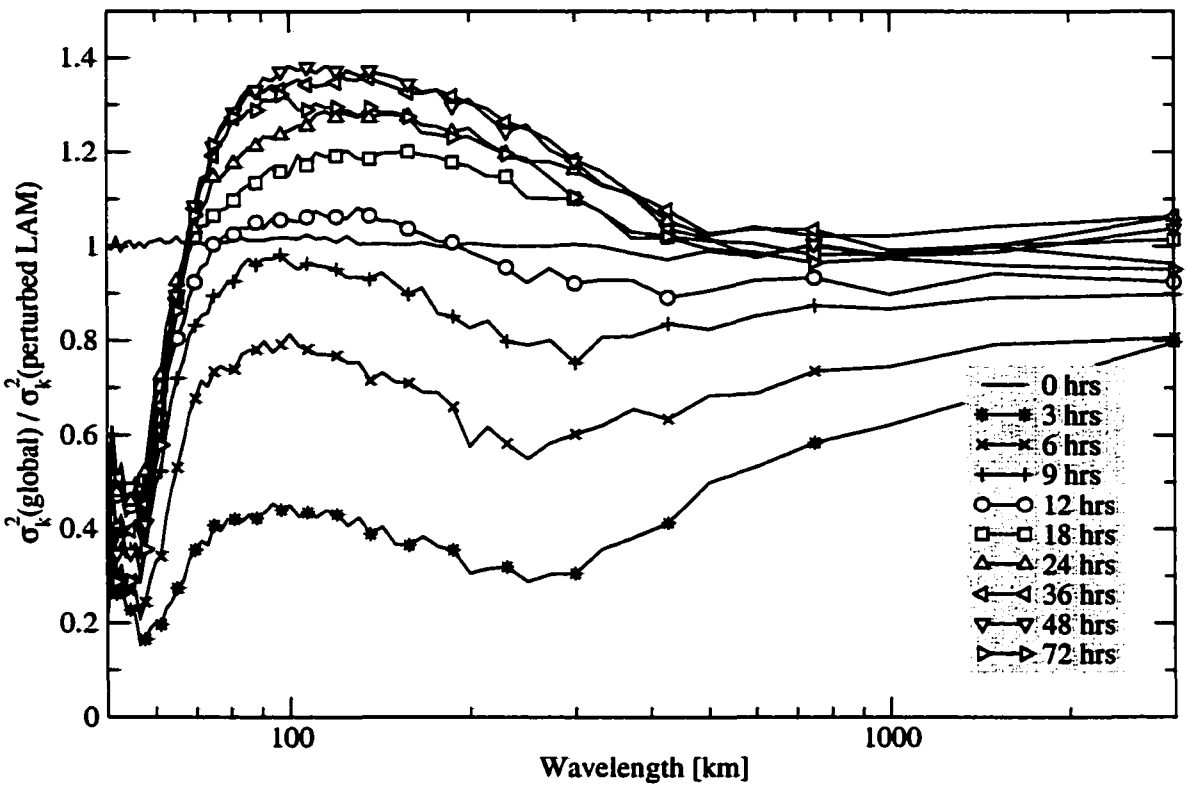



Figure 6.10: Scale factor $\Lambda$ (Eq. 6.12) applied to difference spectra for LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs.

The same set of LAM ensemble simulations were run as above in Sec. 6.2, except scaled LBC perturbation fields were constructed using Eq. 6.6 with amplitudes determined by Eq. 6.13. Results from the case having hourly updated external LBC fields are not shown because error variances lost due to LBC constraints were effectively restored simply using unscaled LBC perturbations (see Figs. 6.4 and 6.7). Indeed, scale factors for this configuration are near unity when scaled LBC perturbations are applied and the statistical results are nearly unchanged.

Results from the case having 3-hourly updated external LBC fields with scaled perturbations also are not shown because they appear nearly identical to those in Figs. 6.5 and 6.8. In Sec. 6.2.1, these simulations were shown to underestimate the amplitude of perturbations needed to fully restore error variances, especially on the large domain. However, the use of scaled LBC perturbations did not yield significant

improvements. As before, the perturbation amplitudes remain so small that when applied along the lateral boundaries, the perturbations dissipate and/or disperse while subsequently passing through the LAM domains.

The use of scaled LBC perturbations did yield a small improvement for the LAM configuration having low-pass filtered LBCs that are updated every 3-hours (Fig. 6.11). Given an increase in perturbation amplitudes of up to 40% (Fig. 6.10), the increase in error variance at scales less than 500 km is about 0 to 6% (compare Figs. 6.6 and 6.11). Although tiny improvements are seen in the error variance spectra, the ensemble summary statistics in Fig. 6.12 show that ensemble dispersion is almost unchanged compared to the configuration having unscaled LBC perturbations. Hence, the significance of these improvements using scaled LBC perturbations is questionable.

Scaled LBC perturbations did not yield significant improvements in the amount of error variance restored at small scales and are not recommended for use in practical applications. The only way to completely restore the error variance in the smallest scales — especially those that have been removed by LBC filtering effects — is to provide LBC perturbations having the same amplitudes as the true fields. However, this would contribute more to the ensemble mean error which, in turn, would cause a loss of ensemble dispersion. If the LBC perturbations are too large, they begin to act in opposition to the original intent of restoring ensemble dispersion through increases in error variance.

## 6.4 Chapter Summary

A new method was developed in this chapter to apply LBC perturbations at every time step of LAM simulations. The perturbations are intended to restore the small-scale error variances and ensemble dispersion lost due to coarsely resolved and temporally interpolated external LBC fields. The LBC perturbations introduced herein are different from using an ensemble of unique LBCs given by corresponding global ensemble members. Indeed, results from Ch. 5 show that even when using an ensemble of LBCs, error growth rates are constrained because the external LBC fields are still coarsely resolved and temporally interpolated in time between updates.

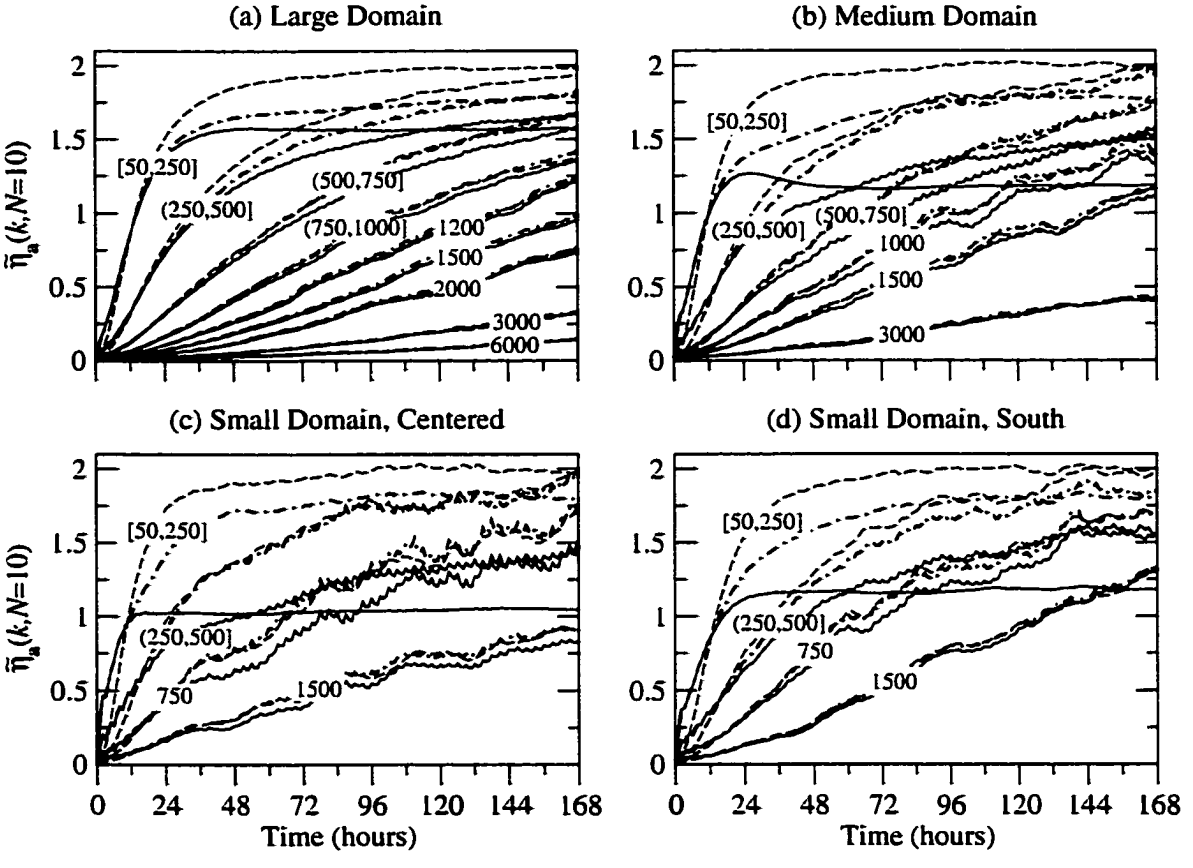



Figure 6.11: Normalized vorticity error variance (Eq. 2.22), averaged over 100 independent 10-member LAM ensemble simulations having 3-hourly updated, low-pass filtered LBCs with scaled perturbations (red, dash-dot lines). Line labels (km) indicate wavelength(s) contributing to error variances. Dashed reference lines are reproduced from Fig. 4.4, showing error variances from subsets of the global ensemble simulations. Solid lines are reproduced from Fig. 5.26, showing error variances from corresponding LAM ensemble simulations run without LBC perturbations. Compare to Fig. 6.6.

Two-dimensional LBC perturbation fields are generated on the LAM grid using inverse Fourier transforms. Amplitudes of the perturbations are determined by the loss of error variance at specific wavelengths due to LBC effects. The fields are initialized with random phase angles, then translated in time using the Rossby phase relation. This method ensures that the perturbation fields remain coherent in both space and time while passing through the lateral boundaries. At each time step, the perturbation field is added to the interpolated LBC fields obtained from subsets of global model simulations. The freshly perturbed LBC field is then blended with the LAM solution across the peripheral 7-point wave absorbing zone.

Results showed that the application of LBC perturbations in LAM ensemble simulations is highly effective at restoring error variances and ensemble dispersion to the values obtained from subsets of global ensemble simulations. Two exceptions were noted. First, error variances were not fully restored at wavelengths shorter than 500 km, especially for large domain simulations having coarsely resolved external LBC fields updated at intervals of 3 hours or longer. This deficiency was attributed to the fact that short-wavelength LBC perturbations have small amplitudes and subsequently disperse and/or dissipate while propagating through the LAM domain. The second notable exception is that, in spite of the gain in error variance, ensemble dispersion for the small domains could not be fully recovered between about 12 and 48 hours of the simulations. This deficiency was explained by the increase in ensemble mean error caused by coarsely resolved and temporally interpolated external LBC fields. The increase in ensemble mean error causes a decrease in ensemble dispersion that cannot be recovered by the use of LBC perturbations as applied herein.

In an attempt to increase the fraction of error variance restored at small scales, a scaling factor was applied to increase the amplitudes of the LBC perturbations. The scaled LBC perturbations produced incremental improvements without degrading the individual LAM simulations. However, the improvements were insignificant, and not enough to justify using scaled perturbations for practical applications.

The LBC perturbations are based solely on differences in error variance spectra. Their only relationship to the dynamical evolution of flow inside the LAM domain is through the nonlinear advection term that transfers energy and enstrophy upscale and downscale throughout the spectrum. To ensure that the LBC perturbations do

not overwhelm the quality of the LAM solution, it is necessary that their amplitudes remain small. Furthermore, the ensemble mean error can become inflated if the perturbations are too strong. An increase in ensemble mean error contributes to a loss of ensemble dispersion, which opposes the effort to restore dispersion through increases in error variance. Although not perfect in every aspect, the LBC perturbations developed in this chapter appear small enough to satisfy these concerns while restoring most of the ensemble dispersion and error variance lost through LBC constraints.

# Chapter 7

# Summary

This study examined the constraints on limited-area model (LAM) error growth and ensemble dispersion caused by the use of "one-way" lateral boundary conditions (LBCs) obtained from coarsely resolved, periodically updated external model fields. Analytic calculations and statistical results from simplified model simulations showed that temporal interpolation of coarsely resolved external fields acts to filter short wavelength features from the LBCs. The spectral definition of error variance was used to show that the loss of shortwave features is directly responsible for the constraints on error growth in LAMs. A new expression was developed that links error variance to ensemble dispersion while accounting for spatial and ensemble biases. The balances required by this expression were used to show that LBC constraints on small scale error variance growth are directly responsible for underdispersive LAM ensemble simulations. To help restore the error variance and ensemble dispersion lost through LBC constraints, a new method was developed to apply statistically consistent LBC perturbations that remain spatially and temporally coherent while passing through the boundaries. With a few noted exceptions, the LBC perturbations were shown to be highly effective at restoring error variance growth and LAM ensemble dispersion without compromising the integrity of the individual solutions.

The study began in Ch. 1 with a review of existing work in the subject. The review highlighted an apparent contradiction between "classic" models of upscale error growth (Lorenz 1982) and the lack of error growth and ensemble dispersion in LAMs despite initial condition perturbations (Anthes et al. 1985; Hamill and Colucci 1997). Attempts to explain these results generally favor the artificial errors introduced by the use of "one-way" LBCs (Vukicevic and Errico 1990). Recent

of linear interpolation in time between LBC updates. This is an issue that has previously received little direct and/or systematic attention in the published literature. Results showed that temporal interpolation between available LBC updates causes an error "pulsing" effect due to inconsistencies between the exterior LBC field and the interior LAM solution. Once the LBC pulsing errors form within the peripheral wave-absorbing zone, they propagate or "sweep" completely through the LAM domain over a time period determined by the domain size and propagation speed. Temporal interpolation of external LBC fields was shown to act as a spatial filter by aliasing waves passing completely through the boundary between updates to larger scales. It also reduces the amplitude of any nonstationary wave moving through the lateral boundary. As smaller scales are progressively removed from the LAM domain through LBC interpolation, the error variances at these scales reach a limiting value that is less than that attained in global simulations.

The second part of Ch. 5 examined the impact of having coarsely resolved external LBC fields by applying a low-pass spatial filter that removed all wavelengths less than 150 km. The approach was not new, and follows that used by Laprise et al. (2000) and De Elía and Laprise (2002). As in the case with temporally interpolated LBCs, the loss of variance at small scales imposes a constraint on the maximum error growth. These results are consistent with those reported using more complex models by (Vukicevic and Errico 1990; Laprise et al. 2000; De Elía and Laprise 2002). Therefore, error growth behaviors established for other aspects of this work should generalize at least qualitatively to more complex model systems.

Statistical results in Ch. 5 also showed that error growth constraints caused by coarsely resolved LBCs are stronger than those caused by the filtering effect of temporal interpolation. However, the temporal interpolation introduces a greater component of artificial model error that may sweep out dynamically consistent initial perturbations. For the first time, the deficiency in error variance caused by both of these LBC filtering effects is identified as a direct cause for limiting dispersion in limited-area ensemble forecasts. The error growth behaviors identified herein are certainly present to some extent in current atmospheric models operating with "one-way" LBCs. The unresolved question is how large is the impact relative to other modeling and data deficiencies.

To help restore the error variance lost at small scales, a new method was developed in Ch. 6 to apply spatially and temporally coherent LBC perturbations at scales affected by interpolation of external LBC fields. Statistical results showed that the application of LBC perturbations in LAM ensemble simulations were highly effective at restoring error variances and ensemble dispersion to the values obtained from subsets of global ensemble simulations. Two exceptions were noted. First, LBC perturbations were applied only within the peripheral wave absorbing zone. Since the short wavelength perturbations have small amplitudes compared to those contained in the original fields, they disperse and/or dissipate while passing through the domain. Therefore, the LBC perturbations were unable to fully restore error variances for scales less than 500 km, especially for large domain simulations. The second notable exception was that ensemble dispersion could not be fully recovered between about 12 and 48 hours of the simulations. The deficiency was explained by the increase in ensemble mean error caused by coarsely resolved and temporally interpolated of external LBC fields. Although not perfect in every aspect, the LBC perturbations applied in Ch. 6 appear small enough to maintain the integrity of the LAM simulations while restoring most of the error variance and ensemble dispersion lost through LBC constraints discussed above. The LBC perturbations are most effective on small domains where, indeed, the LBC constraints on error growth are most severe.

We conclude with comments on how this work may apply to more realistic modeling systems. Two fundamental assumptions were made at the start of this work. First, it was assumed that natural error growth at large scales is (or can be) stated accurately using global model forecasts. This assumption is needed to ensure that error variances and ensemble dispersion are correct at wavelengths exceeding the breadth of the LAM domain. LBC perturbations are only effective at scales up to the size of the LAM domain and cannot correct deficiencies caused by improper error growth rates at larger scales. Second, the assumption was made that error growth rates at small scales in LAMs should behave the same as those in global models operating at equivalent resolution. The amplitude of the LBC perturbations was determined by the difference in error variance spectra between global ensembles and LAM ensembles having unperturbed LBCs. Hence, the LBC perturbations may be less effective if there are differences in the variance spectra caused by dynamical or

artificial discrepancies between external and LAM models. A related assumption is that the model simulations are unbiased, since ensemble dispersion is linked to the total error variance, the ensemble mean error and the spatial bias. This secondary assumption is less important because corrections can be applied for systematic model errors.

The PPV model used for this work is a single-level model and does not contain fast modes for error growth. Therefore, the error growth rates obtained in the PPV model are slower than those in more complex atmospheric models. However, statistical results in Ch. 5 suggest that errors grow in a manner that is at least qualitatively consistent with those obtained in several previous studies. Therefore, additional conclusions drawn from this work should apply equally well to more complex modeling systems. To support this statement, note that Vukicevic and Paegle (1989) tested LBC sensitivities in a barotropic model. The work was later expanded to a full primitive-equation model (Vukicevic and Errico 1990). Results from the later study supported conclusions obtained using the former barotropic model. When considering an extension of the current study to other modeling systems, the magnitude of the LBC constraint will depend on the speed at which waves pass through the lateral boundaries and on the amplitudes of initial condition perturbations. Furthermore, the timescale needed to reach saturation is likely to decrease, but the shape of the normalized error growth curves should remain unchanged.

The greatest challenge faced in applying the LBC perturbation technique to other modeling systems is the determination of appropriate amplitude coefficients needed for constructing the perturbation fields. Global ensemble systems have been available for more than a decade. It should not be difficult to obtain error variance spectra over many cases for these ensembles, especially since most are integrated using spectral methods. The greater challenge is to obtain error variance spectra from LAM ensemble systems. LAM ensemble systems have existed for several years, but most do not include the statistical verification packages needed to calculate one-dimensional error variance spectra. Such packages would need to be developed, then results accumulated over many cases. Once an appropriate set of verification data has been accumulated, corrections for systematic errors must be applied before obtaining difference spectra. Finally, an issue that requires additional research is how to determine the vertical structure of LBC perturbations.

Temporal interpolation of coarsely resolved external LBC fields has been shown to remove small-scale features from LAM solutions and quickly sweep out any set of initial condition perturbations. LBC perturbations applied at every time step are essential to the design of an efficacious LAM ensemble system. The effort will be most rewarding on smaller domains where LBC sweeping effects act most quickly to constrain error growth rates. The additional expense of applying LBC perturbations may be offset by the ability to integrate LAM ensembles over smaller domains.

# Reference List


Anderson, J. L. (1996), 'Selection of initial conditions for ensemble forecasts in a simple perfect model framework', *J. Atmos. Sci.* **53**, 22–36.

Anthes, R. A. (1986), The general question of predictability, *in* 'Mesoscale Meteorology and Forecasting', American Meteorological Society, Boston, chapter 27.

Anthes, R. A. and Baumhefner, D. P. (1984), 'A diagram depicting forecast skill and predictability', *Bull. Amer. Meteor. Soc.* **65**, 701–703.

Anthes, R. A., Kuo, Y.-H., Baumhefner, D. P., Errico, R. M. and Betgge, T. W. (1985), 'Predictability of mesoscale atmospheric motions', *Advances in Geophysics* **28**, 159–202.

Anthes, R. A., Kuo, Y.-H., Hsie, E.-Y., Low-Nam, S. and Bettge, T. W. (1989), 'Estimation of skill and uncertainty in regional numerical models', *Q. J. Roy. Met. Soc.* **115**, 763–806.

Arakawa, A. (1966), 'Computational design for long-term numerical integrations of the equations of atmospheric motion', *J. Comput. Phys.* **1**, 119–143.

Barkmeijer, J., van Gijzen, M. and Boutier, F. (1998), 'Singular vectors and the analysis error covariance metric', *Q. J. Roy. Met. Soc.* **124**, 1695–1713.

Barnes, S. L. (1986), 'On the accuracy of omega diagnostic computations', *Mon. Wea. Rev.* **114**, 1664–1680.

Baumhefner, D. P. (1984), The relationship between present large-scale forecast skill and new estimates of predictability error growth, *in* G. Holloway and B. J. West, eds, 'Predictability of Fluid Motions', number 106 *in* 'AIP conference proceedings', American Institute of Physics, New York, pp. 169–180.

Baumhefner, D. P. and Perkey, D. J. (1982), 'Evaluation of lateral boundary errors in a limited-domain model', *Tellus* **34**, 409–428.

Bickley, W. G. (1937), 'The plane jet', *Phil. Mag. (7)* **23**, 727–731.

Bluestein, H. B. (1992), *Synoptic-Dynamic Meteorology in Midlatitudes, Volume I: Principles of Kinematics and Dynamics*, Oxford University Press.

Boer, G. J. (1984), 'A spectral analysis of predictability and error in an operational forecast system', *Mon. Wea. Rev.* **112**, 1183–1197.

Boer, G. J. (1993), 'Systematic and random error in an extended-range forecasting experiment', *Mon. Wea. Rev.* **121**, 173–188.

Brooks, H. E., Tracton, M. S., Stensrud, D. J., DiMego, G. and Toth, Z. (1995), 'Short-range ensemble forecasting: Report from a workshop, 25-27 july 1994', *Bull. Amer. Meteor. Soc.* **76**, 1617–1624.

Buizza, R. (1997), 'Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system', *Mon. Wea. Rev.* **125**, 99–119.

Carr, F. H. (1977), Numerical simulation of a mid-tropospheric cyclone, Ph.D. dissertation, Florida State University. Available from: University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106.

Charney, J. G. (1949), 'On a physical basis for numerical prediction of large-scale motions in the atmosphere', *J. Meteor.* **6**, 371–385.

Charney, J. G. (1951), Dynamical forecasting by numerical process, *in* 'Compendium of Meteorology', Amer. Meteor. Soc., Boston, MA.

Charney, J. G., Fjørtoft, R. and von Neumann, J. (1950), 'Numerical integration of the barotropic vorticity equation', *Tellus* **6**, 309–318.

Charney, J. G., Fleagle, R. G., Lally, V. E., Riehl, H. and Wark, D. Q. (1966), 'The feasibility of a global observation and analysis experiment', *Bull. Amer. Meteor. Soc.* **47**, 200–220.

Dalcher, A. and Kalnay, E. (1987), 'Error growth and predictability in operational ecmwf forecasts', *Tellus* **39A**, 474–491.

Davies, H. C. (1976), 'A lateral boundary formulation for multilevel prediction models', *Quart. J. Roy. Meteor. Soc.* **102**, 405–418.

Davies, H. C. (1983), 'Limitations of some common lateral boundary schemes used in regional NWP models', *Mon. Wea. Rev.* **111**, 1002–1012.

De Elía, R. and Laprise, R. (2002), 'Forecasting skill limits of nested, limited-area models: A perfect-model approach', *Mon. Wea. Rev.* **130**, 2006–2023.

Du, J., Mullen, S. L. and Sanders, F. (1997), 'Short-range ensemble forecasting of quantitative precipitation', *Mon. Wea. Rev.* **125**, 2427–2459.

Du, J. and Tracton, M. S. (1999), Impact of lateral boundary conditions on regional-model ensemble prediction, *in* H. Ritchie, ed., 'Research activities in atmospheric and oceanic modeling, Report 28', TD-942, WMO, pp. 6.7–6.8.

Durran, D. (1999), *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*, Vol. 32 of *Texts in Applied Mathematics*, Springer-Verlag.

Durran, D. R., Yang, M.-J., Slinn, D. N. and Brown, R. G. (1993), 'Toward more accurate wave-permeable boundary conditions', *Mon. Wea. Rev.* **121**, 604–620.

Ebisuzsaki, W. and Kalnay, E. (1991), Ensemble experiments with a new lagged analysis forecasting scheme, *in* 'Research activities in atmospheric and oceanic modeling, report 15', WMO, WMO, Geneva, Switzerland, pp. 6.31–6.32.

Ehrendorfer, M. (1994), 'The Liouville equation and its potential usefulness for the prediction of forecast skill. part I: Theory', *Mon. Wea. Rev.* **122**, 703–713.

Ehrendorfer, M. and Errico, R. M. (1995), 'Mesoscale predictability and the spectrum of optimal perturbations', *J. Atmos. Sci.* **52**, 3475–3500.

Epstein, E. S. (1969), 'Stochastic dynamic prediction', *Tellus* **21**, 739–759.

Errico, R. M. (1985), 'Spectra computed from a limited area grid', *Mon. Wea. Rev.* **113**, 1554–1562.

Errico, R. M. (1987), 'A comparison between two limited-area spectral analysis schemes', *Mon. Wea. Rev.* **115**, 2856–2861.

Errico, R. M. and Baumhefner, D. (1987), 'Predictability experiments using a high-resolution limited-area model', *Mon. Wea. Rev.* **115**, 488–504.

Errico, R. M. and Langland, R. (1999*a*), 'Notes on the appropriateness of "bred modes" for generating initial perturbations used in ensemble prediction', *Tellus* **51A**, 431–441.

Errico, R. M. and Langland, R. (1999*b*), 'Reply to: Comments on "Notes on the appropriateness of 'bred modes' for generating initial perturbations used in ensemble prediction"', *Tellus* **51A**, 450–451.

Givoli, D. (1991), 'Non-reflecting boundary conditions', *J. Comput. Phys.* **94**, 1–29.

Gleeson, T. A. (1970), 'Statistical-dynamical prediction', *J. Appl. Meteorology* **9**, 333–344.

Haltiner, G. J. and Williams, R. T. (1980), *Numerical Prediction and Dynamic Meteorology*, 2nd edn, John Wiley & Sons.

Hamill, T. M. (2001), 'Interpretation of rank histograms for verifying ensemble forecasts', *Mon. Wea. Rev.* **129**, 550–560.

Hamill, T. M. and Colucci, S. J. (1997), 'Verification of Eta-RSM short-range ensemble forecasts', *Mon. Wea. Rev.* **125**, 1312–1327.

Hamill, T. M. and Snyder, C. (2000), 'A hybrid ensemble Kalman filter-3d variational analysis scheme', *Mon. Wea. Rev.* **128**, 2905–2919.

Hamill, T. M., Snyder, C. and Morss, R. E. (2000), 'Comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles', *Mon. Wea. Rev.* **128**, 1835–1851.

Harrison, M., Palmer, T., Richardson, D. and Buizza, R. (1999), 'Analysis and model dependencies in medium-range ensembles: two transplant case studies', *Quart. J. Roy. Meteor. Soc.* **125**, 2487–2515.

Hoffman, R. N. and Kalnay, E. (1983), 'Lagged average forecasting, an alternative to Monte Carlo forecasting', *Tellus* **35A**, 100–118.

Holton, J. R. (1979), *An Introduction to Dynamic Meteorology*, Vol. 23 of *International Geophysics Series*, 2nd edn, Academic Press.

Holton, J. R. (1992), *An Introduction to Dynamic Meteorology*, Vol. 48 of *International Geophysics Series*, 3rd edn, Academic Press.

Hou, D., Kalnay, E. and Drogemeier, K. K. (2001), 'Objective verification of the SAMEX '98 ensemble forecasts', *Mon. Wea. Rev.* **129**, 73–91.

Houtekamer, P. L. and Derome, J. (1995), 'Methods for ensemble prediction', *Mon. Wea. Rev.* **123**, 2181–2196.

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H. L. (1996), 'A system simulation approach to ensemble prediction', *Mon. Wea. Rev.* **124**, 1225–1242.

Houtekamer, P. L. and Mitchell, H. L. (1998), 'Data assimilation using an ensemble Kalman filter technique', *Mon. Wea. Rev.* **126**, 796–811.

Kundu, P. K. (1990), *Fluid Mechanics*, Academic Press.

Kuo, H. L. (1949), 'Dynamical instability of two-dimensional nondivergent flow in a barotropic atmosphere', *J. Meteor.* **6**, 105–122.

Kuo, H. L. (1973), 'Dynamics of quasi-geostrophic flows and instability theory', *Adv. in Appl. Mech.* **13**, 247–330.

Laprise, R., Varma, M. R., Denis, B., Caya, D. and Zawadzki, I. (2000), 'Predictability of a nested limited-area model', *Mon. Wea. Rev.* **128**, 4149–4154.

Leith, C. E. (1971), 'Atmospheric predictability and two-dimensional turbulence', *J. Atmos. Sci.* **28**, 145–161.

Leith, C. E. (1974), 'Theoretical skill of Monte Carlo forecasts', *Mon. Wea. Rev.* **102**, 409–418.

Leith, C. E. and Kraichnan, R. H. (1972), 'Predictability of turbulent flows', *J. Atmos. Sci.* **29**, 1041–1058.

Leslie, L. M. and Speer, M. S. (1998), 'Short-range ensemble forecasting of explosive australian east coast cyclogenesis', *Wea. Forecasting* **13**, 822–832.

Lilly, D. K. (1984), Some facets of the predictability problem for atmospheric mesoscales, *in* G. Holloway and B. J. West, eds, 'Predictability of Fluid Motions', number 106 *in* 'AIP conference proceedings', American Institute of Physics, New York, pp. 287–294.

Lorenz, E. N. (1963), 'Deterministic nonperiodic flow', *J. Atmos. Sci.* **20**, 131–140.

Lorenz, E. N. (1969*a*), 'Atmospheric predictability as revealed by naturally occurring analogues', *J. Atmos. Sci.* **26**, 636–646.

Lorenz, E. N. (1969*b*), 'The predictability of a flow which possesses many scales of motion', *Tellus* **21**, 289–307.

Lorenz, E. N. (1969*c*), 'Three approaches to atmospheric predictability', *Bull. Amer. Meteor. Soc.* **50**, 345–349.

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

Lorenz, E. N. (1984), Estimates of atmospheric predictability at medium range, *in* G. Holloway and B. J. West, eds, 'Predictability of Fluid Motions', number 106 *in* 'AIP conference proceedings', American Institute of Physics, New York, pp. 133–140.

Molteni, R., Buizza, R., Palmer, T. N. and Petroliagis, T. (1996), 'The ECMWF ensemble prediction system: Methodology and validation', *Quart. J. Roy. Meteor. Soc.* **122**, 73–119.

Mullen, S. L. and Baumhefner, D. P. (1989), 'The impact of initial condition uncertainty on numerical simulations of large-scale explosive cyclogenesis', *Mon. Wea. Rev.* **122**, 1548–1567.

Mullen, S. L., Du, J. and Sanders, F. (1999), 'The dependence of ensemble dispersion on analysis-forecast systems: Implications to short-range ensemble forecasting of precipitation', *Mon. Wea. Rev.* **127**, 1674–1686.

Murphy, A. H. (1998), 'The early history of probability forecasts: Some extensions and clarifications', *Mon. Wea. Rev.* **13**, 5–15.

Nobile, A. and Roberto, V. (1986), 'MFFT: A package for two- and three-dimensional vectorized discrete Fourier transforms', *Comp. Phys. Comm.* **42**, 233.

Oliger, J. and Sundström, A. (1978), 'Theoretical and practical aspects of some initial boundary value problems in fluid dynamics', *S.I.A.M. J. Appl. Math.* **35**, 419–446.

Orlanski, I. (1976), 'A simple boundary condition for unbounded hyperbolic flows', *J. Comp. Phys.* **21**, 251–269.

Orrell, D., Smith, L., Barkmeijer, J. and Palmer, T. (2002), 'Model error in weather forecasting', *Nonlinear Proc. Geophys.* .

Paegle, J., Yang, Q. and Wang, M. (1997), 'Predictability in limited area and global models', *Meteorol. Atmos. Phys.* **63**, 53–69.

Palmer, T. N. and Tibaldi, S. (1988), 'On the prediction of forecast skill', *Mon. Wea. Rev.* **116**, 2453–2480.

Perkey, D. J. and Kreitzberg, C. W. (1976), 'A time-dependent lateral boundary scheme for limited area primitive equation models', *Mon. Wea. Rev.* **104**, 744–755.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1996), *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd edn, Cambridge University Press.

Reynolds, C. A., Webster, P. J. and Kalnay, E. (1994), 'Random error growth in NMC's global forecasts', *Mon. Wea. Rev.* **122**, 1281–1305.

Schubert, S. D. and Suarez, M. (1989), 'Dynamical predictability in simple general circulation model: Average error growth', *J. Atmos. Sci.* **46**, 353–370.

Sivillo, J. K., Ahlquist, J. E. and Toth, Z. (1997), 'An ensemble forecasting primer', *Wea. Forecasting* **12**, 809–818.

Smagorinsky, J. (1969), 'Problems and promises of deterministic extended range forecasting', *Bull. Amer. Meteor. Soc.* **50**, 286–311.

Sommerfeld, A. (1949), *Partial Differential Equations in Physics*, Academic Press.

Staniforth, A. (1997), 'Regional modeling: A theoretical discussion', *Meteorol. Atmos. Phys.* **63**, 15–29.

Stensrud, D. J., Bao, J.-W. and Warner, T. T. (2000), 'Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems', *Mon. Wea. Rev.* **128**, 2077–2107.

Stensrud, D. J., Brooks, H. E., Du, J., Tracton, M. S. and Rogers, E. (1999), 'Using ensembles for short-range forecasting', *Mon. Wea. Rev.* **127**, 433–446.

Stephenson, D. B. and Doblas-Reyes, F. J. (2000), 'Statistical methods for interpreting Monte Carlo ensemble forecasts', *Tellus* **52A**, 300–322.

Stroe, R. and Royer, J. F. (1993), 'Comparison of different error growth formulas and predictability estimation in numerical extended-range forecasts', *Ann. Geophysicae* **11**, 296–316.

Swarztrauber, P. and Sweet, R. (1975), Efficient FORTRAN subroutines for the solution of elliptic equations, Technical Report NCAR/TN-109+IA, National Center for Atmospheric Research.

Szunyogh, I., Kalnay, E. and Toth, Z. (1997), 'A comparison of Lyapunov and optimal vectors in a low-resolution GCM', *Tellus* **49A**, 200–227.

Tennekes, H. (1978), 'Turbulent flow in two- and three-dimensions', *Bull. Amer. Meteor. Soc.* **59**, 22–28.

Thompson, P. D. (1957), 'Uncertainty of initial state as a factor in the predictability of large-scale atmospheric flow patterns', *Tellus* **9**, 275–295.

Thompson, P. D. (1984), A review of the predictability problem, *in* G. Holloway and B. J. West, eds, 'Predictability of Fluid Motions', number 106 *in* 'AIP conference proceedings', American Institute of Physics, New York, pp. 1–10.

Thompson, P. D. (1985), 'A statistical-hydrodynamical approach to problems of climate and its evolution', *Tellus* **37A**, 1–13.

Toth, Z. and Kalnay, E. (1997), 'Ensemble forecasting at NCEP and the breeding method', *Mon. Wea. Rev.* **125**, 3297–3319.

Toth, Z., Kalnay, E., Tracton, S. M., Wobus, R. and Irwin, J. (1997), 'A synoptic evaluation of the NCEP ensemble', *Wea. Forecasting* **12**, 140–153.

Toth, Z., Szunyogh, I., Kalnay, E. and Iyengar, G. (1999), 'Comments on: "Notes on the appropriateness of 'bred modes' for generating initial perturbations"', *Tellus* **51A**, 442–449.

Tracton, M. S. and Kalnay, E. (1993), 'Operational ensemble prediction at the national meteorological center: Practical aspects', *Wea. Forecasting* **8**, 379–398.

Tribbia, J. J. and Baumhefner, D. P. (1988), 'The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modeling deficiencies', *Mon. Wea. Rev.* **116**, 2276–2288.

Van Tuyl, A. H. and Errico, R. M. (1989), 'Scale interaction and predictability in a mesoscale model', *Mon. Wea. Rev.* **117**, 495–517.

Vukicevic, T. (1991), 'Nonlinear and linear evolution of initial forecast errors', *Mon. Wea. Rev.* **119**, 1602–1611.

Vukicevic, T. and Errico, R. M. (1990), 'The influence of artificial and physical factors upon predictability estimates using a complex limited-area model', *Mon. Wea. Rev.* **118**, 1460–11482.

Vukicevic, T. and Paegle, J. (1989), 'The influence of one-way interacting lateral boundary conditions on predictability of flow in bounded numerical models', *Mon. Wea. Rev.* **117**, 340–350.

Walker, J. S. (1988), *Fourier Analysis*, Oxford University Press.

Wandishin, M. L., Mullen, S. L., Stensrud, D. J. and Brooks, H. E. (2001), 'Evaluation of a short-range multimodel ensemble system', *Mon. Wea. Rev.* **129**, 729–747.

Warner, T. T., Key, L. E. and Lario, A. M. (1989), 'Sensitivity of mesoscale-model forecast skill to some initial-data characteristics, data density, data position, analysis procedure and measurement error', *Mon. Wea. Rev.* **117**, 1281–1310.

Warner, T. T., Peterson, R. A. and Treadon, R. E. (1997), 'A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction', *Bull. Amer. Meteor. Soc.* **78**, 2599–2617.

Williamson, D. L. and Browning, G. L. (1974), 'Formulation of the lateral boundary conditions for the ncar limited area model', *J. Appl. Meteor.* **13**, 8–16.

Xue, M. (2000), 'High-order monotonic numerical diffusion and smoothing', *Mon. Wea. Rev.* **128**, 2853–2864.

Zhu, Y., Iyengar, G., Toth, Z., Tracton, M. S. and Marchok, T. (1996), Objective evaluation of the NCEP global ensemble, *in* 'Preprints, 11th Conf. on Numerical Weather Prediction', Amer. Meteor. Soc., Norfolk, VA, pp. 29J–30J.

# Appendix A

# The Variance Spectrum and Total Variance

A key element of this work is the partitioning of variance into contributions made by motions of specific wavelength bands. A mathematical basis for the *variance spectrum* is provided here, followed by a link to the more common statistical definition of *total variance*.

Given two points in a field $f(\mathbf{x})$ and $f(\mathbf{x}+\mathbf{x}')$ separated by a distance $\mathbf{x}'$, the autocorrelation function is (Kundu 1990, pg. 425)

$$a(\mathbf{x}') = \langle f(\mathbf{x}) f(\mathbf{x}+\mathbf{x}') \rangle \tag{A.1}$$

where the angle brackets denote an average over all points in the field. The autocorrelation is related to variance through the *Schwartz inequality* (Walker 1988; Kundu 1990)

$$a(\mathbf{x}') \leq \left\langle f^2(\mathbf{x}) \right\rangle^{\frac{1}{2}} \left\langle f^2(\mathbf{x}+\mathbf{x}') \right\rangle^{\frac{1}{2}} . \tag{A.2}$$

Specifically, if $\mathbf{x}' = 0$, (A.2) becomes an equality and defines the variance

$$a(0) = \left\langle f^2(\mathbf{x}) \right\rangle . \tag{A.3}$$

The *variance spectrum* is obtained by taking the Fourier transform of the autocorrelation function. To demonstrate this fact, introduce a two-dimensional wavenumber vector $\boldsymbol{\kappa} = (k, l)$ so that the Fourier transform of $a(\mathbf{x}')$ is

$$\mathcal{A}(\boldsymbol{\kappa}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} e^{-i\boldsymbol{\kappa}\cdot\mathbf{x}'} a(\mathbf{x}') d\mathbf{x}, \tag{A.4}$$

where $d\mathbf{x}$ is a generalized notation for $dxdy$. For (A.4) to be true, the inverse Fourier transform must exist so that

$$a(\mathbf{x}') = \int_{-\infty}^{\infty} e^{i\boldsymbol{\kappa}\cdot\mathbf{x}'} \mathcal{A}(\boldsymbol{\kappa}) d\boldsymbol{\kappa}, \tag{A.5}$$

where $d\boldsymbol{\kappa}$ is a generalized notation for $dldk$. When $\mathbf{x}' = 0$, the variance spectrum is obtained from (A.3) and (A.5) since

$$\langle f^2(\mathbf{x}) \rangle = \int_{-\infty}^{\infty} \mathcal{A}(\boldsymbol{\kappa}) d\boldsymbol{\kappa}, \tag{A.6}$$

Thus, (A.6) shows that the variance spectrum is obtained from the Fourier transform of the autocorrelation function.

The *Wiener-Khinchin theorem* (Walker 1988; Press et al. 1996) states that $\mathcal{A}(\boldsymbol{\kappa})$ is obtained for real functions by multiplying the Fourier transform of $f(\mathbf{x})$ by its complex conjugate. That is,

$$\mathcal{A}(\boldsymbol{\kappa}) = \mathcal{F}(\boldsymbol{\kappa})\mathcal{F}^*(\boldsymbol{\kappa}) = |\mathcal{F}(\boldsymbol{\kappa})|^2 \tag{A.7}$$

where $\mathcal{F}(\boldsymbol{\kappa})$ is the Fourier transform of $f(\mathbf{x})$. Using (A.7) in (A.6), the variance spectrum for real $f(\mathbf{x})$ becomes

$$\langle f^2(\mathbf{x}) \rangle = \int_{-\infty}^{\infty} |\mathcal{F}(\boldsymbol{\kappa})|^2 \, d\boldsymbol{\kappa}. \tag{A.8}$$

Thus, the variance contained within wavenumbers $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa} + d\boldsymbol{\kappa}$ is the magnitude of the Fourier coefficient $|\mathcal{F}(\boldsymbol{\kappa})|^2$.

The notation used above generalizes to multi-dimensional space. Henceforth, consideration is restricted to one-dimensional spectra obtained from an appropriately detrended two-dimensional field (see Appendix B). Let $F(k, l)$ be the complex Fourier coefficients obtained from the two-dimensional discrete Fourier transform of $f(\mathbf{x})$. Then, using $\kappa = \sqrt{k^2 + l^2}$, $|F(\kappa)|^2$ is the sum of all $|F(k, l)|^2$ such that

$\aleph(\kappa) - (1/2)\delta\kappa \leq \kappa \leq \aleph(\kappa) + (1/2)\delta\kappa$, where $\aleph$ denotes the nearest integer and $\delta\kappa$ is defined in Appendix B. It follows from (A.8) that

$$\left\langle f^2(\mathbf{x}) \right\rangle \approx |F(0)|^2 + \sum_{\kappa=1}^{K-1} 2\,|F(\kappa)|^2 . \tag{A.9}$$

The factor of two results from the fact that $|F(\kappa)| = |F(-\kappa)|$ for real $f(\mathbf{x})$. To obtain the *total biased variance*, introduce the variable $\sigma^2$ such that

$$\sigma^2 \equiv \left\langle f^2(\mathbf{x}) \right\rangle - |F(0)|^2 = \sum_{\kappa=1}^{K-1} 2\,|F(\kappa)|^2 . \tag{A.10}$$

Note that $|F(0)|^2$ represents the square of the grid mean of $f(\mathbf{x})$. Thus, $\sigma^2$ quantifies how much the field $f(\mathbf{x})$ fluctuates around the field average.

The total biased variance is more commonly presented and understood within a statistical framework. Specifically, given a symmetrically distributed random variable $\varphi$ and its expected value $E(\varphi)$, variance is defined

$$\sigma^2 \equiv E([\varphi - E(\varphi)]^2) = E(\varphi^2) - E(\varphi)^2. \tag{A.11}$$

Note the similarity to (A.10). Now, following (A.11), the *total biased variance* for any two-dimensional field $f(i,j)$ defined on the grid $0 \leq i \leq N_x - 1,\ \ 0 \leq j \leq N_y - 1$ is computed as

$$\sigma^2 = \frac{1}{N_x N_y} \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} (f_{i,j} - \langle f \rangle)^2 \tag{A.12}$$

where $N_x$ and $N_y$ are the number of grid points along each dimension and $\langle f \rangle$ is the grid mean. The variance as defined here is a biased estimate of the population variance because the sample mean is used rather than the expected value. The *unbiased* variance may be obtained by simply multiplying (A.12) by $N_x N_y/(N_x N_y - 1)$. Diagnostic checks during calculations of vorticity error variance spectra shown in previous chapters indicate that Eq. A.10 retains more than 99.99% of the vorticity error variance obtained by Eq. A.12.

# Appendix B

# Spectra From a Limited Area Grid

Much of the analysis in this work makes use of one dimensional spectra produced using data from a two dimensional gridded field. Procedures for constructing these one dimensional spectra are provided in this appendix.

The first step in conducting any spectral analysis using discrete Fourier transforms is to ensure the data are periodic over the interval of interest. Consider a two dimensional grid $f(i,j)$, $0 \leq i \leq N_x - 1$, $0 \leq j \leq N_y - 1$ with grid spacing $\Delta x = \Delta y$. Now, the periodic field is obtained by subtracting linear trends along each row and column (Errico 1987) so that

$$\begin{aligned} f_p(i,j) &= f(i,j) - (f_{I,j} - f_{0,j}) \left(\frac{i}{I} - \frac{1}{2}\right) - (f_{i,J} - f_{j,0}) \left(\frac{j}{J} - \frac{1}{2}\right) \\ &+ (f_{I,J} - f_{0,J} - f_{I,0} + f_{0,0}) \left(\frac{i}{I} - \frac{1}{2}\right) \left(\frac{j}{J} - \frac{1}{2}\right), \end{aligned} \tag{B.1}$$

where $I = N_x$ and $J = N_y$. The full channel domain is periodic along the $x$-direction, so linear trends are removed only across the $y$-direction so that

$$f_p(i,j) = f(i,j) - (f_{i,J} - f_{j,0}) \left(\frac{j}{J} - \frac{1}{2}\right). \tag{B.2}$$

After removing the linear trends across the field, $f_p(i,j)$ is periodic with $f_p(0,j) = f_p(I,j)$ and $f_p(i,0) = f_p(i,J)$. Errico (1985, 1987) shows that detrending eliminates the aliasing onto smaller scales of power associated with wavelengths that extend beyond the limited-area domain. However, the detrending does introduce additional

power at smaller scales due to differences in patterns along each boundary. The additional short wavelength power is minimized in the absence of rapid field variations along the boundaries. Barnes (1986) proposed a detrending scheme that resulted in less modification of the field, but Errico (1987) showed that it still suffered from serious aliasing effects. Although there is no perfect approach for detrending data on a limited area grid, results are presented here in a normalized manner (Eq. 2.21) so that conclusions are minimally altered when comparing different model configurations.

Having produced a detrended field $f_p(i,j)$, the two-dimensional discrete Fourier transform is computed using available Fast Fourier Transform routines[1]. The output from such routines is subsequently stored in a complex valued array $F(m,n)$ representing wavenumber indices $0 \leq m \leq I/2$ and $-J/2 \leq n \leq J/2$ (Fig. B.1). Actual wavenumbers $k$ and $l$ are obtained as $k = 2\pi m/(I-1)\Delta x$ and $l = 2\pi n/(J-1)\Delta y$. Note that computations are needed for just half the total Nyquist limited wavenumbers due to the symmetry of Fourier coefficients obtained from real data (Errico 1985; Press et al. 1996).

Two dimensional power spectra are obtained by computing $2|F(m,n)|^2$. As an exception, $|F(0,n)|^2$ and $|F(I/2,n)|^2$ are not multiplied by two. To obtain the total biased variance (Eq. A.10), add the power obtained at each point in $F$ except the grid mean $|F(0,0)|$.

One-dimensional spectra, like those shown throughout chapters 4, 5, and 6 are computed by summing power from points on $F$ that lie within annular rings of width $\delta\kappa = 2\pi/(J-1)\Delta y$. The center of each ring represents wavenumber $\kappa = \sqrt{k^2+l^2}$. Unlike Carr (1977) or Errico (1985) who determine $\delta\kappa$ relative to $I-1$, $\delta\kappa$ is determined here along the shorter axis, resulting in wider rings (see Fig. B.1). On the square nested domains where the number of gridpoints is the same along each direction, this choice is inconsequential. On the rectangular channel grid, the different way of computing $\delta\kappa$ excludes the longest waves in the zonal direction. In all situations, the mean $F(0,0)$ is excluded from the summations since it does not contribute to the total variance.

[1]Computational details common to most FFT routines are provided by (Press et al. 1996). The Fortran routine `MFFT` (Nobile and Roberto 1986) was applied for this work, obtained online from `http://www.fftw.org/benchfft/doc/ffts.html`.

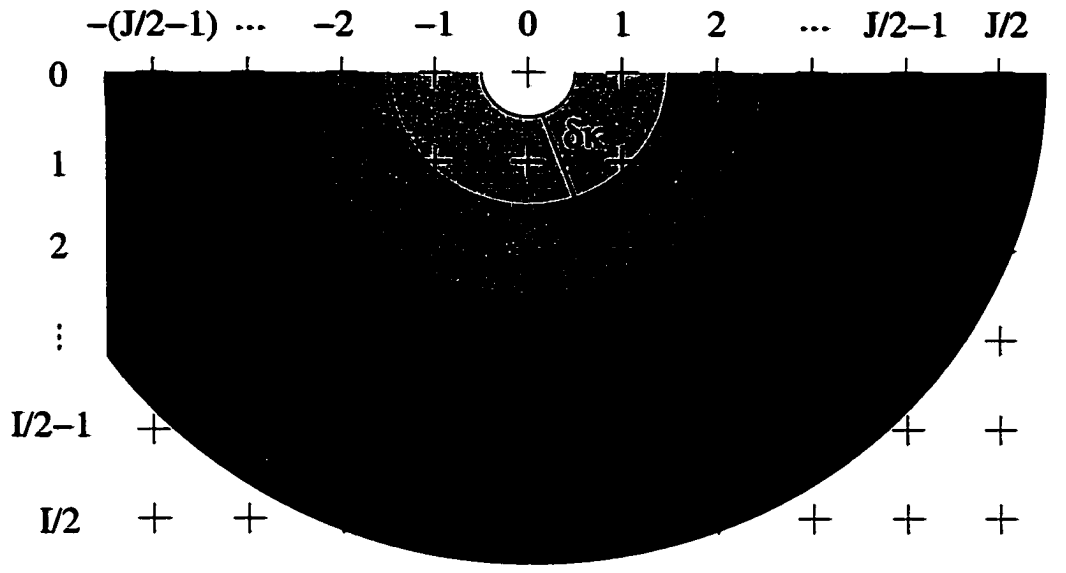



Figure B.1: Schematic diagram showing the arrangement of complex coefficients obtained from a discrete two dimensional Fast Fourier Transform. Shaded regions of width $\delta\kappa$ show the annular rings used to construct one dimensional spectra (see text for explanation). Crosses located outside any shaded regions reveal those points excluded from power spectra summations.

Having completed the Fourier transform, a low-pass filter may be applied to the field when desired. Specifically, the raised cosine filter is applied to $F$ with weights determined by

$$w(m,n) = \begin{cases} 1 & \text{if } \kappa < \kappa' - \delta\kappa'/2 \\ 0 & \text{if } \kappa > \kappa' + \delta\kappa'/2 \\ \frac{1}{2}[1 - \sin(\pi(\kappa - \kappa')/\delta\kappa')] & \text{otherwise} \end{cases} \tag{B.3}$$

where $\delta\kappa'$ is the width of the filter and $\kappa'$ is the cutoff wavenumber at which wave amplitudes are reduced to half their original value. In practice, the filter width and cutoff wavenumbers are specified in terms of wavelenghts before being converted to wavenumbers in computation. After the filter is applied, an inverse discrete Fourier Transform is computed and the linear trends restored to construct the new filtered field.