(87)

# Adaptive observations: Idealized sampling strategies for improving numerical weather prediction

by

Rebecca Elisabeth Morss

A.B., The University of Chicago
(1993)

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

December 1998
[February 1999]



Signature of Author .....
Department of Earth, Atmospheric, and Planetary Sciences
11 December, 1998

Certified by .....
Kerry A. Emanuel
Professor of Meteorology
Thesis Supervisor

Accepted by .....
Ronald G. Prinn
Chairman, Department of Earth, Atmospheric, and Planetary Sciences

# Adaptive observations: Idealized sampling strategies for improving numerical weather prediction

by

Rebecca Elisabeth Morss

A.B., The University of Chicago
(1993)

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
on 11 December, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

## Abstract

The purpose of adaptive observations is to use information about individual atmospheric situations to identify regions where additional observations are likely to improve weather forecasts of interest. The observation network could be adapted for a wide range of forecasting goals, and it could be adapted either by allocating existing observations differently or by adding observations from programmable platforms to the existing network. In this study, we explore observation strategies in a simulated idealized system with a three-dimensional quasi-geostrophic model and a realistic data assimilation scheme. Several issues are addressed, including whether adapting observations has potential to improve forecasts, how observational resources can be optimally allocated in space and time, how effectively ensemble forecasts can estimate errors in initial conditions, and how much the data assimilation system affects the influence of the observations.

Using simple error norms, we compare idealized non-adaptive observations with adaptive observations for a variety of observation densities. The adaptive strategies implemented incorporate information only about errors in the initial conditions. We test both an idealized adaptive strategy, which selects observation locations based on perfect knowledge of the true atmospheric state, and a more realizable adaptive strategy, which uses an ensemble to estimate errors in the initial conditions.

We find that the influence of the observations, both adaptive and non-adaptive, depends strongly on the observation density. In this simulated system, observations on synoptic scales dominate the average error reduction; above a certain observation density, adding any observations, adaptive or non-adaptive, has a much smaller effect. Results presented show that for non-dense observation networks, the adaptive strategies

tested can, on average in this simulated system, reduce analysis and forecast errors by a given amount using fewer observational resources than the non-adaptive strategies. In contrast, however, our results suggest that it is much more difficult to benefit from modifying the observation network for dense observation networks, for adaptive observations taken infrequently, or for additional observations taken to improve forecasts in individual cases.

The interactions between the observations, the data assimilation system, the errors in the initial conditions, and the forecast model are complex and depend on the specific forecast situation. This leads to a non-negligible risk that forecasts will be degraded when observations are adapted in an individual situation. Further study is needed both to understand these interactions better and to learn to what extent the results from this idealized study apply to more complex, more realistic systems.

Thesis Supervisor: Kerry A. Emanuel
Title: Professor of Meteorology

# Acknowledgments

First, I would like to thank my advisor, Kerry Emanuel, for supporting me while encouraging me to work independently and for providing many opportunities, including the opportunity to work on a project more fascinating than I had ever imagined. I am grateful to Chris Snyder for setting a wonderful example and for his patience during countless hours spent discussing quasi-geostrophic models, adjoints, and data assimilation. I am also indebted to the other members of my thesis committee: Edward Lorenz and Edmund Chang for their advice during the past few years, and John Marshall for his enthusiasm and for joining my committee at nearly the last minute.

I would like to thank those I have met at conferences, at workshops, and during my stays at MMM at NCAR and at EMC at NCEP, for insightful conversations both on science and on life in general. Special thanks to: Zoltan Toth and Istvan Szunyogh for teaching me nearly everything I know about operational forecasting and dodgy plots; David Parrish for helpful discussions on data assimilation; and Rich Rotunno for offering during my practice talks several suggestions that I still use. Thanks also to Michael Morgan, Stan Trier, and Lodovica Illari for their weather knowledge and their encouragement.

I am grateful to the MIT and CMPO/PAOC staff for their administrative and technical assistance, in particular to Joel Sloman for his poetic breaks to otherwise dull days, Jane McNabb for having a friendly answer to nearly every question, and Linda Meinke for solving countless FORTRAN-90 and operating system problems. Thanks also to Stacey Frangos and Beverly Kozol-Tattlebaum for helping schedule and organize events, and Dan Burns for his kind listening ear.

There are many others to whom I am indebted for helping make the good moments of my graduate career enjoyable and the not-so-good moments survivable; unfortunately, I only have room to mention some of them. I am grateful to those who have kept me entertained during my stints on location, including: Dave Rodenbaugh (the e-mail prince), Scott Alan Mittan (see, I still remember!), and friends in Boulder, CO; Joe Moss, Gabe Paal, Chris Gabel, and Marty and Lucy Schlenker in Washington, DC; and the Lear jet FASTEX crew and the Hurricane Hunters in St. John's, NF. In addition, I thank the other friends from afar who kept me going with phone calls, e-mails, meals, visits, and trips, and who provided a much-needed perspective on life outside MIT, including: Angie Mihm, Yoshiki Hakutani, Tania Snyder, Jeff Achter, Chris Velden, Bob Petersen, Ian Renfrew, Rachel Scott, Colleen Webb, and Mark Webb.

I thank my classmates, officemates, and departmentmates at MIT, for providing some useful advice, a crash course in assertiveness training, and even an occasional bit of fun. In particular, thanks to: Albert Fischer and Stephanie Harrington for helping me endure several problem set sessions; Constantine Giannitsis for a reminder that friendships can appear in unexpected places; the lunch clique, including Clint Conrad, Gary Kleiman, Bonnie Souter, Katy Quinn, and Natalie Mahowald, for debates on BLTs with cheese versus cheese sandwiches with bacon; Richard Wardle for keeping me entertained without ever serving me lobster; Christophe Herbaut for occasionally pretending that my French

was as good as his English; and Nili Harnik for her wonderfully original cast art and her infectious smile.

Thanks also to: Mort Webster and Dave Reiner for keeping me up-to-date on climate policy while proving there is such a thing as a free lunch; Mary Agner for her WISEness and for thinking it ridiculous that my ideas were radical; the women's rugby football club for some entertaining song lyrics and the patience learned with two broken limbs; Jody Lyons and Dave Glass for distracting me at a few crucial times; and the Muddy Charles for hours of entertaining procrastination and relaxing thesis editing.

I am grateful to the Tang circuit, the residents of Willow Ave. and W. Newton, and associated parties, for many hours of fun during the past 4+ years. I especially thank: Matthew Dyer for combining taste for bad jokes with a beautiful soul; Karen Willcox for on occasion letting me, no, insisting that I borrow her clothes, and for the upcoming road trip to NYC; and Vanessa Chan for the cheap rates on her carhotel and for not making me lift more than her. I am indebted to Diane Klein Judem, for letting me say what I knew all along and for helping me create another chance. And I thank my roommates during my time at MIT for listening to (and enduring) my occasional griping sessions, particularly: Suzie Wetzel and Nina Shapley for late night conversations about life, love, and the pursuit of happiness; and "Electric Blue" Dave Grundy for his subtle (and not-so-subtle) sense of humor, for cooking me breakfast this morning, and for appreciating the B.J. in me.

There are two people without whom I never would have survived my first two years at MIT. Thanks to: Didier Reiss, for sending a ray of California sunshine my way to carry me through the times when I could not create my own; and Derek Surka, for his good, strong heart and his thoughtfulness about things both large and small.

In addition to those mentioned above, there are several people to whom I am especially grateful for helping make the last few years truly enjoyable. Special thanks to: Marianne Bitler for her friendship, her belief in decompartmentalization, and her encouragement to drink the coffee; Mark "y Mark" van der Helm for being Blotto and a nice guy at the same time (despite his best attempts otherwise) and for cooking me breakfast yesterday morning; and Gerard Roe, for his near-infinite patience, and for having the graciousness to admit when I am right (at least some of the time), the courage to tell me when I am wrong, and most importantly, the wisdom to know the difference.

Finally, thanks to my family for their love, support, and encouragement over the years. I thank my siblings, Sydney, Benjamin, and Alisa (and honorary siblings Norman and Alex), for making me at least half of who I am and for occasionally moving to interesting places for me to visit. And I dedicate this thesis to my parents, Lester and Sue Morss, for buying me more than a few weather books and calendars during my childhood, for listening even when they didn't understand, and for raising me to believe that I could accomplish anything. I love you and thank you, since without that knowledge I would not be writing this.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Motivation and background

For several decades, it has been known that numerical weather forecasts are sensitive to small changes in initial conditions. This means that even if we could model the atmosphere perfectly, errors in the initial conditions would amplify rapidly, leading to forecast errors and forecast failures. The initial conditions for operational numerical weather prediction (NWP) models, called analyses, are based on a combination of short range numerical forecasts with observations. The observations, where available, constrain the atmospheric state in the forecast model to be as close as possible to the true atmospheric state. The current observation network is made up of three types of observation platforms. Fixed platforms, such as rawinsonde stations, take observations at pre-selected times and locations, generally near population centers and thus over land. In the data gaps left by the fixed network, there are observations taken from platforms of opportunity, such as planes and ships, and from remote sensing platforms. At any given time, however, these latter two types of platforms are usually at locations which were selected for reasons other than weather prediction. In many cases, they also have limited vertical coverage or resolution.

Because the observation network is inhomogeneous, in any forecast situation there will be regions where information about the initial conditions is important but where there are insufficient observations. If we could use our knowledge about a specific atmospheric situation to identify these regions, it might be possible to deploy programmable observation platforms to take data in them and improve the initial conditions. These adaptive (also called targeted) observations could both reduce global analysis errors and

help predict important weather phenomena on various temporal and spatial scales.

Forecasts will probably benefit little from observations in regions where the initial conditions are already quite accurate or where errors will have only a small effect on forecasts which interest us. Thus, effective adaptive observation strategies are likely to incorporate information both about probable analysis errors and about rapidly amplifying forecast errors. Various strategies have been proposed to include these criteria, each assuming that somewhat different analysis and forecast errors are the most important to identify and correct first. Subjective methods, such as identifying tropopause or other features which may be pre-cursors of rapidly developing atmospheric systems in data sparse areas, have been suggested (e.g. Snyder 1996). However, because it is not always possible to subjectively identify important features for a certain forecast (for example, when there is downstream energy propagation), several objective techniques have been proposed for adapting observations.

Each individual implementation of each adaptive observing strategy is different, and so we do not discuss the specific proposed strategies in detail. As an overview, however, the objective strategies which are currently being developed and tested include three basic types. First, there are techniques based primarily on estimates of errors in the initial conditions, such as the methods using ensemble spread tested in Lorenz and Emanuel (1998). Second, strategies have been developed directly from adjoint techniques, calculating optimal perturbations (singular vectors) or sensitivities of forecasts to small changes in the initial conditions (e.g. Bergot et al. 1999, Gelaro et al. 1999, Palmer et al. 1998). Adjoint-based targeting strategies emphasize the errors which, if they exist in the initial conditions, will grow most rapidly. Although error statistics can be incorporated, the current implementations of adjoint-based strategies include little or no information about the probability of analysis and forecast errors. Finally, there are strategies which explicitly combine the error probability and error growth criteria. The ensemble transform technique, for example, uses an ensemble of perturbed forecasts to estimate both uncertainty in analyses and forecasts and the time evolution of the uncertainty (e.g. Bishop and Toth 1996, Bishop and Toth 1999).

Intuitively we believe that the best strategies will combine the two criteria in some manner. Unfortunately, adapting observations successfully has turned out to be a much more complex exercise than simply evaluating one or both of the criteria in the theoretically optimal way. How much the different strategies improve forecasts in any atmospheric situation depends on many factors, including the selected "forecast of interest," the limitations of the observing platforms, and the data assimilation procedure

which incorporates the observations into the model state to create the initial conditions for forecasting. This means that the best strategies are likely to vary from situation to situation, and that they are likely to change as forecast models, available observing platforms, and data assimilation schemes evolve. At this time, it is not clear which (if any) of the currently proposed adaptive strategies will be most beneficial in any situation and implementation.

Current work on adaptive observations includes several different approaches, ranging from field experiments in which observations are taken to improve weather forecasts in real time, to idealized studies with simple models. In the real atmosphere, targeted observations are being implemented for tropical cyclone forecasts (Burpee et al. 1996, Aberson 1997; S. Aberson, personal communication), as well as being tested for mid-latitude weather forecasts during the Fronts and Atlantic Storm Track Experiment (FASTEX, in Jan.-Feb. 1997) and the North Pacific Experiment (NORPEX-98, in Jan.-Feb. 1998; NORPEX-99, to occur in Jan.-Feb. 1999). Several types of additional observing platforms have been tested, including dropwindsondes released from aircraft and high-resolution winds derived from satellite water vapor data (Velden et al. 1997), with data potentially gathered from unmanned aircraft in the future (Langford and Emanuel 1993). Unfortunately, because observing platforms and forecast verifications are often highly constrained in the real world, it is difficult to obtain enough good cases to allow one to draw firm conclusions. In addition, operational forecasting systems are complex, data assimilation systems and forecast models vary between numerical weather prediction centers, and the range of possible forecasting goals is large. This combination of factors make the results from real world adaptive observations difficult to interpret. The results are preliminary, and we do not discuss them in detail. In general, however, it is clear that so far the real adaptive observations have had a mixed influence. The assimilated extra observations have improved some forecasts, but in other cases they have had little impact or have even degraded forecasts (see Emanuel and Langland 1998, Gelaro et al. 1999, Langland et al. 1999a, and Szunyogh et al. 1999 for sample FASTEX results; see Langland et al. 1998b and Toth et al. 1999 for initial NORPEX results).

To eliminate many of the practical constraints, observing system simulation experiments (OSSE's) use a forecast model and a data assimilation system to simulate observations and an analysis and forecast cycle. Because simplifying the forecast model both reduces the computational cost dramatically and makes the understanding the results easier, several of these idealized studies have been performed with low-order models. Lorenz and Emanuel (1998), for example, used a one-dimensional model, which predicts

one quantity at 40 locations on a latitude circle, to investigate several adaptive observation strategies. Berliner et al. (1999) proposed a statistical framework and used it with the same one-dimensional model to explore several aspects of observing networks. These and similar experiments have suggested that adapting observations may be beneficial and have raised some important issues to consider when adapting observations. However, because of the one-dimensional dynamics of the model and the associated lack of requirement for a complex data assimilation system, these results may only apply to real atmospheric forecasts in a limited sense.

Both the idealized simple model and the real world approaches have left unanswered several important questions about adaptive observations. Although these questions are much too broad and complex to answer completely within the scope of this study, we begin to address them with the hope that our results will further our theoretical and practical understanding of how we can use observations to improve weather forecasts. This both can help us interpret results from other adaptive observations studies and can suggest which issues are the most important to explore first in future studies. The questions we address are:

- Can adaptive observations improve analyses and forecasts on a statistically significant basis, in a fully three-dimensional dynamical system with a realistic data assimilation system? If adapting observations can improve analyses and forecasts, what kinds of improvements can we expect under different circumstances?

- Which types of strategies are likely to be the most effective in different circumstances? How can the strategies best be implemented, in terms of allocation of observational resources in space and time? How effectively can we estimate errors in initial conditions for use in adaptive observation strategies?

- How important are the limitations of the data assimilation system for adapting observations? How would we like to change the data assimilation procedure to improve the influence of observations on forecasts?

- How important is the risk of degradations when taking extra observations? How can we minimize the risk? With this risk in mind, what is the best way to define our goals when trying to optimally allocate observations?

- Given that we have limited observational resources, as discussed in Emanuel et al. (1997), what is the "optimal" mix of observing systems? In other words, what

role would we like all different types of observations, adaptive and non-adaptive, to play in the future?

The adaptive observing framework for addressing these issues is fairly new. Some of the more basic underlying questions, however, have been addressed previously from somewhat different perspectives. For example, it has recently come to our attention that during the late 1960's and early 1970's, several studies were performed with OSSE's to investigate which types of observations are the most useful for NWP and how these observations can best be assimilated, questions reminiscent of several of those above. More precisely, as satellite data first became available and as optimal interpolation data assimilation schemes became more accepted, several researchers began to explore how asynoptic data (primarily satellite but also including aircraft and other in situ data) might compare to and fill in gaps left by more traditional land-based radiosonde observations. Some of the studies focussed on comparing forecasts from initial conditions generated with observations distributed differently in space and time (Bengtsson and Gustavsson 1971, Bengtsson and Gustavsson 1972, Charney et al. 1969, and Morel et al. 1970). Other studies focussed on prioritizing data types, considering different observing platforms which could observe different meteorological variables with different accuracies (Charney et al. 1969, Jastrow and Halem 1970, Smagorinsky et al. 1970). These early studies were limited by the lack of computational resources; the forecast models are simple and their resolutions coarse by current standards, and only a few experiments could be run. Although the specific issues addressed in these studies are different, some of the general goals and methods are similar to those we use.

## 1.2 Approach

We use an approach which bridges the gap between the full numerical weather prediction system and the idealized simple model experiments described above: an idealized OSSE setup with a fully three-dimensional, yet simplified, quasi-geostrophic forecast model and a realistic data assimilation system. The basic framework for the experiments follows that used by Lorenz and Emanuel (1998) and previous observing system simulation studies (e.g. Jastrow and Halem 1970). We have selected an OSSE setup to avoid many of the logistics which make improving real-time forecasts difficult. With repeatable experiments, simulated observations, and perfect knowledge, we are limited neither by the number of possible cases nor by our ability to sample those cases well.

Later, if we wish, some of the idealizations can be relaxed and the importance of the real-world constraints evaluated.

To address some of the issues raised by both the one-dimensional model and the real world studies, we decided to test different observing networks in a three-dimensional system, but one with as few competing processes as possible. For both our forecast model and for our "real" atmosphere, we chose an idealized multi-level quasi-geostrophic model. Because it has simplified forcing and geometry, the quasi-geostrophic model does not simulate some aspects of real mid-latitude synoptic behavior, such as storm tracks. In other respects, however, the model represents large-scale dynamics realistically, and its relative simplicity both makes a larger number of experiments computationally feasible and simplifies interpreting the results. Since the model dynamics have limited accuracy at sub-synoptic scales, to the extent possible we do not interpret our results for sub-synoptic scale applications.

The forecast model cannot use the data without first incorporating it into a three-dimensional best estimate of the atmospheric state. Consequently, a data assimilation system is a necessary complicating factor. We developed a three-dimensional variational data assimilation system similar to those used in operational NWP so that we can begin to explore how any, necessarily imperfect, data assimilation system interacts with observations and with a forecast model. We describe the data assimilation system in detail, explicitly addressing the strengths and weaknesses of how it incorporates data. This helps us understand both how our choice of data assimilation systems might affect our results and how data assimilation techniques can be improved so that they use observations more effectively.

Originally we designed this study with the goal of determining the "optimal" adaptive strategy. As the study developed, however, it became apparent that this goal was not only impossible to achieve, but was also not the best goal to work towards, for several reasons. First, since the most effective strategy is likely to change from situation to situation and from forecasting system to forecasting system, it is not clear if there is such a thing as an optimal strategy. In addition, the proposed strategies are in early stages of development, both theoretically and practically, and thus did not seem ready for rigorous comparison. Finally, we realized that we understand very little about how any extra observations, adaptive or non-adaptive, interact with data assimilation systems, with forecast errors, and with forecast models. It therefore seemed unwise to compare strategies without understanding more fundamentally what happens to the observational data once they are taken.

As we moved towards understanding more fundamentally how observations are used by a forecast system, we also decided not to attempt explicitly to simulate real observing networks and real observing platforms. Thus, we did not implement a complex "standard" background observation network, varying in time and with different observation densities in different pre-specified regions. Instead, we chose to test only idealized non-adaptive observation strategies. Rather than implementing one or more of the specific proposed adaptive observation strategies described above, we also decided to test only very simple adaptive strategies. The idealized observation strategies allow us to produce results more general than a comparison of how specific strategies added to a given background network of observations can improve specific forecasts in specific forecasting systems.

The primary stated goal of adaptive observations is to improve important forecasts. However, to help us identify the key considerations for adapting observations in general, in this study we also test only idealized error norms. Most of the results presented are aggregated over a large number of cases, in terms of domain- and time-averaged analysis error. These results are still relevant for reducing forecast errors on average because, as we discuss in Section 5.2, in this simulated system average analysis improvement and average forecast improvement are linked. Improving forecasts in individual cases is a more complex issue, one which we only begin to address specifically towards the end of the study.

Adapting observations combines several different areas of numerical weather prediction, including data assimilation techniques, forecast error evolution, ensemble forecasting, and adjoint techniques (including singular vectors). Although these topics are important background material for understanding adaptive observations, each is a broad and complex field on its own. Therefore, in the interest of space, we only discuss them briefly where relevant, directing the reader to some sample appropriate references for further information. To tie our results to real world forecasting, it is also important to have a basic understanding of several other issues, such as the currently proposed adaptive observation techniques and the influence of the adaptive observations tested in field experiments. Because these fields are new and evolving rapidly, and because many of the results are preliminary, again we do not discuss them in detail. When possible, we provide references for the adaptive observations results, although many of them are not yet published (as of late 1998). Often, however, our understanding of the important issues for adapting observations is not based on a single identifiable research study or researcher. Much of our interpretations of how our results might be applied to

real forecasting have developed from our experiences taking data and verifying forecasts during FASTEX (Joly et al. 1997, special issue of *Quart. J. Roy. Meteor. Soc.* to appear) and NORPEX-98 (Langland et al. 1999b) and from presentations and discussions from several workshops on targeted observations. These workshops include those which took place: in Camp Springs, MD on 10-11 Apr. 1997; in Monterey, CA on 8 Dec. 1997 (summarized in Emanuel and Langland 1998); in Toulouse, France on 27-30 Apr. 1998; and in Monterey, CA on 4-5 Nov. 1998 (summary in preparation by R. Langland).

The results from the quasi-geostrophic model, the idealized experimental setup, and the simplified strategies cannot be applied directly to real forecasting. They are important, however, because they can help us understand fundamental aspects of how any atmospheric or related forecasting system might respond to different types of observational information. The basic principles learned from the experiments with the simplified model can then be tested with a more complex, more realistic model and eventually explored with real atmospheric data.

## 1.3 Outline

In Chapters 2 and 3, the quasi-geostrophic model and three-dimensional variational data assimilation system are described. Chapter 4 explains the OSSE setup and how the observation strategies were chosen and implemented. Chapter 5 presents results from networks with different densities of fixed observations, including a summary of the sensitivity of the results to the error norm and to the data assimilation procedure. From these results, we develop a framework for studying further changes in observation networks. We then begin to address the issues discussed above, building on the low-order model results and investigating some unresolved difficulties with adapting observations from the field experiments.

In Chapter 6, we compare global idealized adaptive observations to non-adaptive global observation strategies at different observation densities. We then explore how a limited amount of observational resources might be optimally allocated in space and time. Chapter 7 tests using different-sized ensembles to estimate background error for adaptive observations. We also briefly discuss how the results may be affected if observation locations must be selected well in advance of the observation time. In Chapter 8, observation strategies are tested when the observations are added to a pre-existing network of fixed observations. The influence of the extra observations is illustrated both

with time-averaged results and with results from a time series of individual situations. Finally, in Chapter 9, we present examples and statistics of how individual extra observations affect analyses and forecasts in specific cases. These results demonstrate some of the specific ways in which observations can interact with errors in the initial conditions, with the data assimilation system, and with the forecast model.

Throughout the study, we discuss the extent to which we believe the results from the simplified model and the simplified experimental setup may or may not apply to more complex systems and to more general situations. We also explore the importance of the data assimilation system for the results, and we address what next steps are required before our results can be applied to real world observing networks.

# Chapter 2

# Quasi-geostrophic model

The quasi-geostrophic (QG) model used is a gridpoint channel model on a beta plane, periodic in longitude, developed at the National Center for Atmospheric Research (NCAR) and described in Rotunno and Bao (1996, hereafter referred to as RB96). We have selected this multi-level QG model because it exhibits three-dimensional dynamical behavior that is, to first order, similar to that of the real atmosphere, while it remains simple enough to make a large number of runs computationally feasible. The QG model is forced by relaxation to a specified zonal mean state (described below); it has no orography or seasonal cycle. Dissipation includes fourth-order horizontal diffusion and, at the lower boundary, Ekman pumping. The version for this study has constant stratification and a tropopause with fixed height (but varying temperature) at the upper boundary $z = H$. With the simplified geometry and forcing, this QG model does not simulate some features of real mid-latitude synoptic behavior, such as storm tracks and other stationary wave patterns. Even so, the model exhibits a wide range of chaotic behavior. We have chosen an idealized QG model to help us to gain a basic understanding of how observation networks behave, an understanding which we can later apply to adapting observations in a more complex, more realistic atmosphere.

The QG model variables are potential vorticity ($q$) in the interior and potential temperature ($\theta$) at the upper and lower boundaries. With the geopotential $\psi$ (referred to as streamfunction once non-dimensionalized) and the geostrophic wind $\mathbf{v}_g$ given by

$$\theta = \frac{\overline{\theta}}{g} \frac{\partial \psi}{\partial z}, \tag{2.1}$$

and

$$f\mathbf{v}_g = \mathbf{k} \times \nabla\psi, \tag{2.2}$$

we define the pseudo-potential vorticity as

$$q = \frac{1}{f}\nabla^2\psi + \frac{f}{N^2}\frac{\partial^2\psi}{\partial z^2} + \beta y. \tag{2.3}$$

The equation governing the interior potential vorticity, including the diffusion and relaxation, is then

$$\left(\frac{\partial}{\partial t} + \mathbf{v}_g \cdot \nabla\right) q = -\nu\,\nabla^4 q - \frac{1}{\tau_{\text{relax}}}\left(q - q_{\text{ref}}\right). \tag{2.4}$$

With Ekman pumping at the lower surface, the equation for the potential temperature at the upper and lower boundaries is

$$\left(\frac{\partial}{\partial t} + \mathbf{v}_g \cdot \nabla\right)\theta = -\nu\,\nabla^4\theta - \frac{1}{\tau_{\text{relax}}}\left(\theta - \theta_{\text{ref}}\right) - \begin{cases} \sqrt{\frac{A_v}{2f}}\,\frac{N^2}{g}\,\bar{\theta}\,\zeta_g & \text{at } z = 0 \\ 0 & \text{at } z = H, \end{cases} \tag{2.5}$$

where the vertical component of the geostrophic vorticity is defined as

$$\zeta_g = \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \tag{2.6}$$

and $A_v$ is the vertical eddy diffusion.

The zonal mean reference state ($q_{\text{ref}}$, $\theta_{\text{ref}}$) is a Hoskins-West jet (Hoskins and West 1979), with a zonal wind ($u$) maximum at the tropopause, $u = 0$ at the channel walls, and a corresponding sinusoidally varying meridional temperature gradient. The jet is specified according to Equation (17) in Hoskins and West (1979); for further details on the mathematical formulation, see Hoskins and West (1979) and RB96. In the results below, the maximum zonal wind speed $U$ of the reference state jet is scaled to $60\,\text{ms}^{-1}$, and the model state is relaxed to the reference state over a time scale $\tau_{\text{relax}} = 20$ days. Figure 2.1 depicts meridional cross-sections through the reference state for the potential temperature and zonal wind profiles.

The diffusion coefficient $\nu$ is set to $9.84 \times 10^{15}\,\text{m}^4\text{s}^{-1}$ for the standard resolution runs. Defining the damping time scale for a wavenumber $k$ as

$$\tau_{\text{diff}} = \frac{1}{k^4\,\nu},$$

the damping time scale for the shortest wave at the standard resolution (wavelength $= 2\Delta x = 500\,\text{km}$) is approximately 1 hour.

The Ekman layer is formulated as a well-mixed viscous layer with a bulk aerodynamic

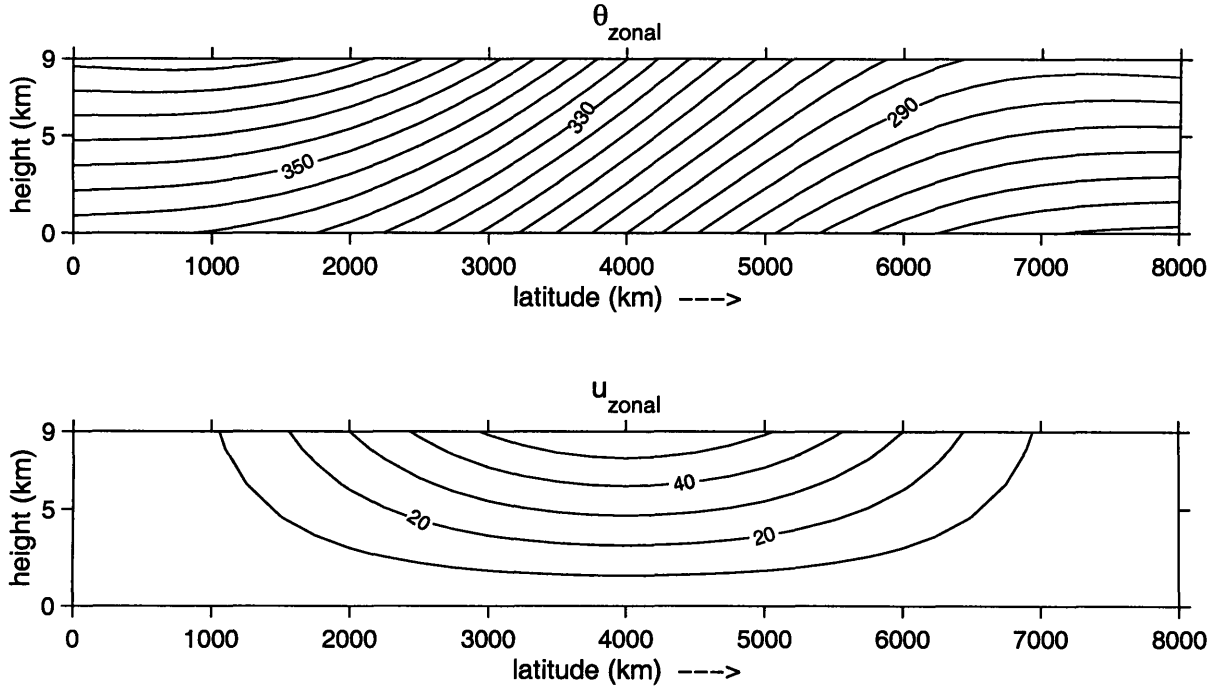



Figure 2.1: Meridional cross-section through the zonal mean reference state for (a) potential temperature (K) and (b) zonal wind ($\mathrm{ms}^{-1}$) in the QG model.

formula for the surface stress, as in RB96. Following Valdes and Hoskins (1988), the viscous boundary layer can be made equivalent to an explicit Ekman layer with a no-slip lower boundary condition. The vertical eddy diffusion coefficient $\mathrm{A_v}$ is related to the surface drag coefficient $\mathrm{c_D}$ by

$$\mathrm{c_D}\,|u_{\mathrm{surf}}| = (\frac{\mathrm{A_v} f}{2})^{\frac{1}{2}},$$

where $|u_{\mathrm{surf}}|$ is the average total horizontal surface velocity. We set $\mathrm{A_v} = 5\,\mathrm{m^2 s^{-1}}$, which is equivalent to $\mathrm{c_D} \approx 1.5 \times 10^{-3}$ and $|u_{\mathrm{surf}}| \approx 10\,\mathrm{ms}^{-1}$. $\mathrm{A_v}$ is also related to the traditional Ekman number $\mathrm{E_v}$ according to

$$\mathrm{E_v} = 2\frac{\mathrm{A_v}}{fH^2}.$$

The equations are non-dimensionalized as in RB96, and at each time step, the boundary potential temperature is incorporated into the interior potential vorticity and Equation (2.3) is inverted to solve for $\psi$. Through Equation (2.1), the boundary potential temperature determines the upper and lower boundary conditions for the inversion. Partitioning the streamfunction into $\bar{\psi}$ (the time-varying zonal mean component) and $\psi'$

(the deviation from $\bar{\psi}$), the boundary conditions at the channel walls are

$$\frac{\partial}{\partial t}\frac{\partial \bar{\psi}}{\partial y} = 0, \tag{2.7}$$

$$\psi' = 0, \tag{2.8}$$

and no diffusion, i.e.

$$\nu = 0. \tag{2.9}$$

After the streamfunction is diagnosed, Equations (2.4) and (2.5) are integrated forward in time using leapfrog time differencing, except just after a new analysis is created, when a forward Euler scheme is used. The time step is approximately 30 minutes for the standard resolution runs, and the integration is time filtered to remove the computational mode (Robert 1966).

The dimensional domain for all results shown has a circumference of 16000 km and a channel width of 8000 km (approximately 70° latitude, large enough so that there is less interaction between the synoptic systems and the channel walls). The depth $H$ is 9 km. The Rossby radius of deformation is defined as

$$\mathrm{R_d} = \frac{NH}{f}. \tag{2.10}$$

With the Brunt-Väisälä frequency $N = .011293$ and $f = 1 \times 10^{-4}$, $\mathrm{R_d} \approx 1000$ km. The Rossby number is defined as

$$\mathrm{R_o} = \frac{U}{fL} = \frac{U}{NH}, \tag{2.11}$$

and since $U = 60\ \mathrm{ms}^{-1}$, $\mathrm{R_o} \approx .59$. The advection time scale

$$t_{\mathrm{adv}} = \frac{L}{U} = \frac{1}{f\mathrm{R_o}} \tag{2.12}$$

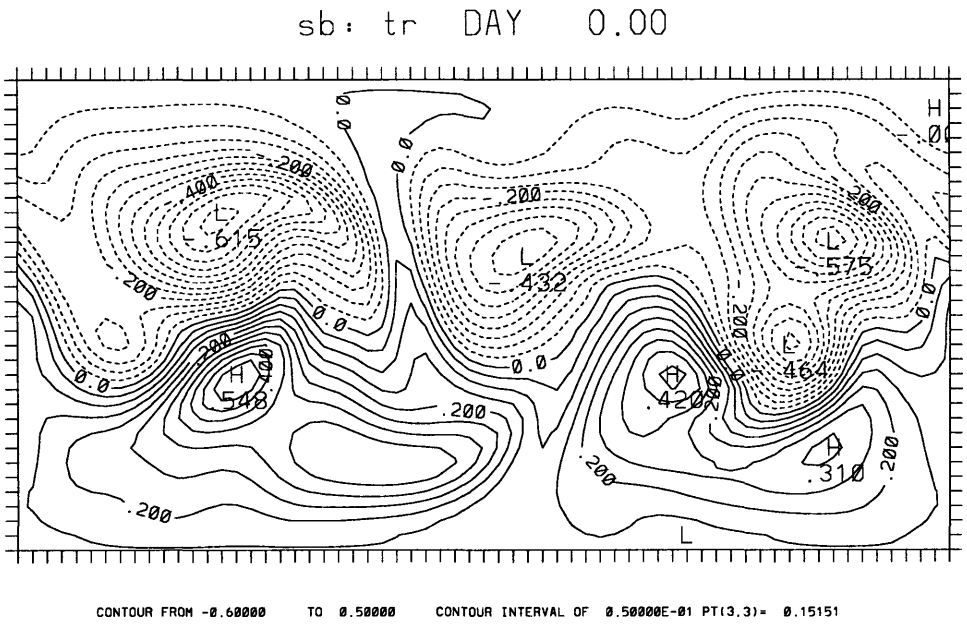
is then approximately 5 hours.

The standard resolution runs have 250 km horizontal resolution and 5 vertical levels. At this resolution, we have estimated from the growth rate of small differences in the QG model state that the error doubling time is approximately 2.5–4 days (for example, compare for dense observations the average magnitude of the initial condition errors in Figure 5.1 with that of the 5 day forecast errors in Figure 5.3). The specific error growth rate depends on the situation, but this value may be on average slow compared to the

real atmosphere. If the resolution is roughly doubled to 125km horizontal resolution and 8 vertical levels, changing to a 7.5 minute time step and a diffusion coefficient $\nu = 2.46 \times 10^{15}\,\mathrm{m^4 s^{-1}}$, the corresponding error doubling time is 1–2 days. To evaluate how sensitive the results are to the model resolution and to the error growth time scale, Section 6.4 presents a few of the results at double resolution. Since it was not computationally feasible to run all of the experiments at higher than the standard resolution, however, we have also compensated for the somewhat slow error growth rate by changing the rate at which observational data are input into the model in Sections 5.4 and 6.3

Figure 2.2 shows the streamfunction and the potential temperature at the upper and lower boundaries for an arbitrarily selected sample QG model state at the standard resolution. Figure 2.3 shows the same fields for an arbitrarily selected state at double resolution. The QG model at higher resolution not only has faster error growth, but comparing Figures 2.2 and 2.3, we see that it also better resolves strong temperature gradients and thus has a more focused jet stream.

a)




b)

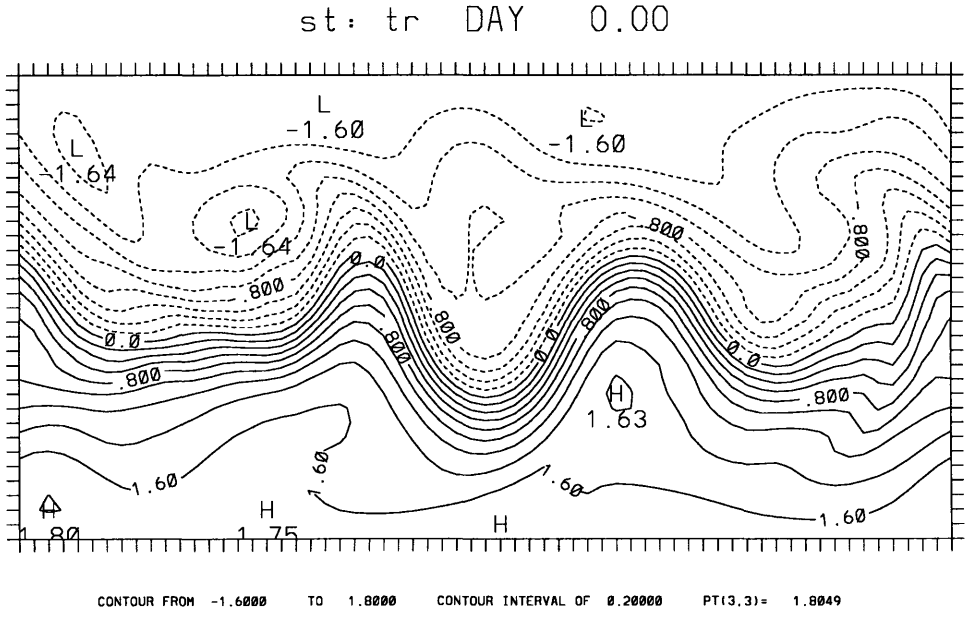
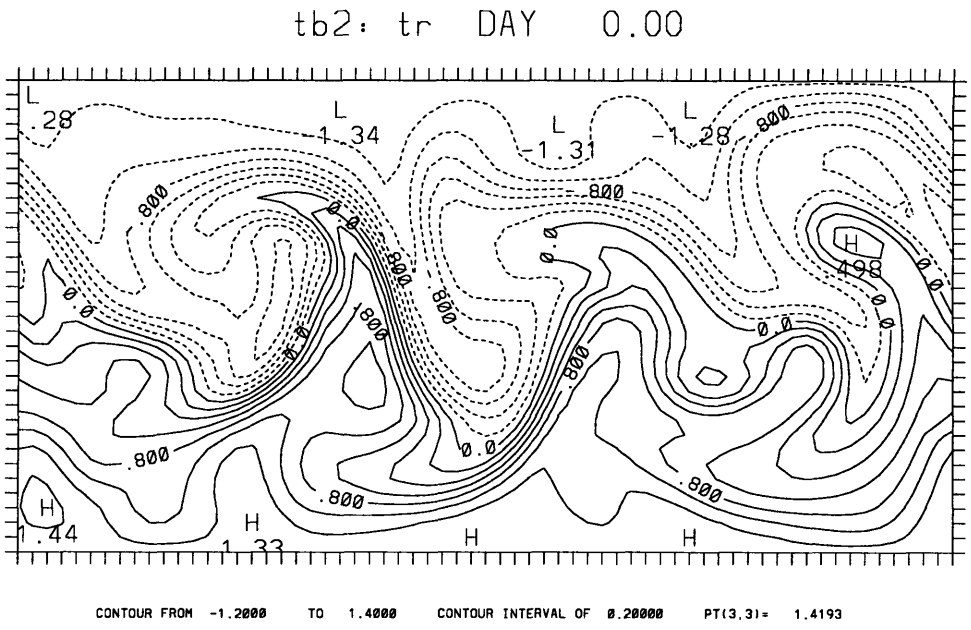



Figure 2.2: Sample vertical cross-sections through the standard resolution QG model state at an arbitrary time, for non-dimensionalized: (a) lower boundary streamfunction, (b) upper boundary streamfunction, (c) lower boundary potential temperature, and (d) upper boundary potential temperature. Longitude (periodic) is plotted on the x-axis and latitude is plotted on the y-axis, each with resolution 250 km.

c)




d)

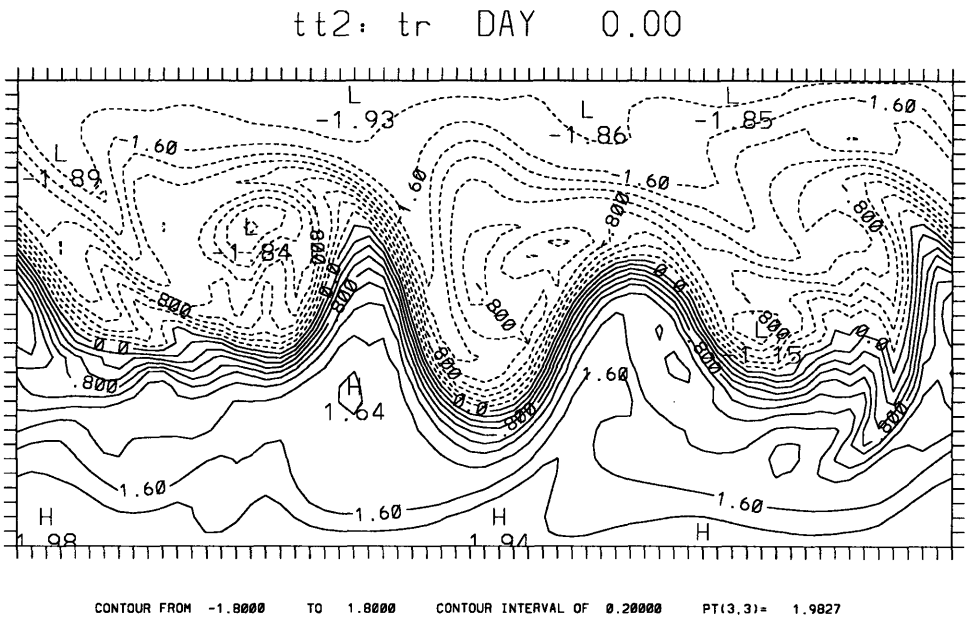
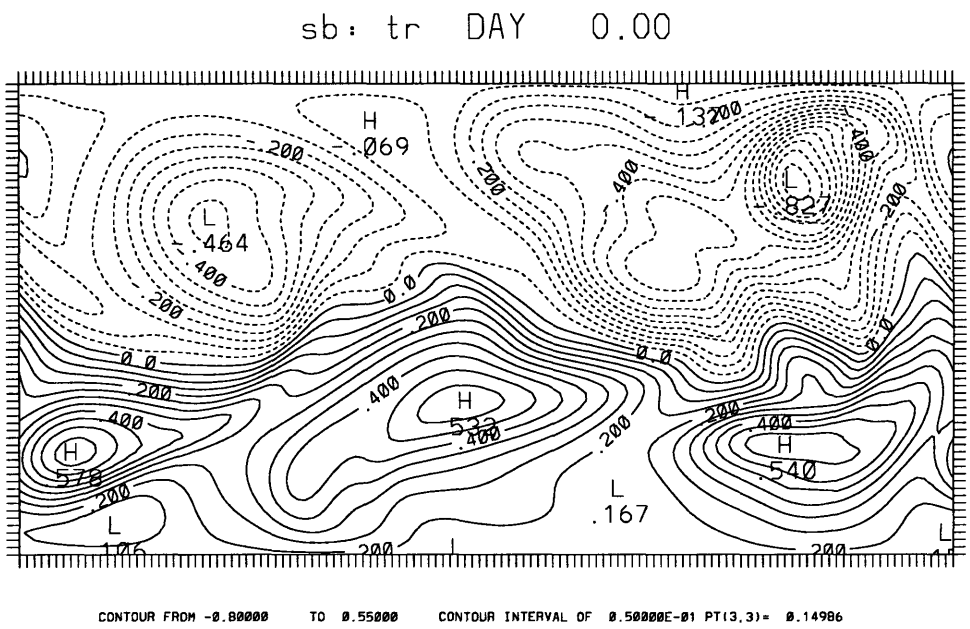



Figure 2.2, continued.

a)




b)

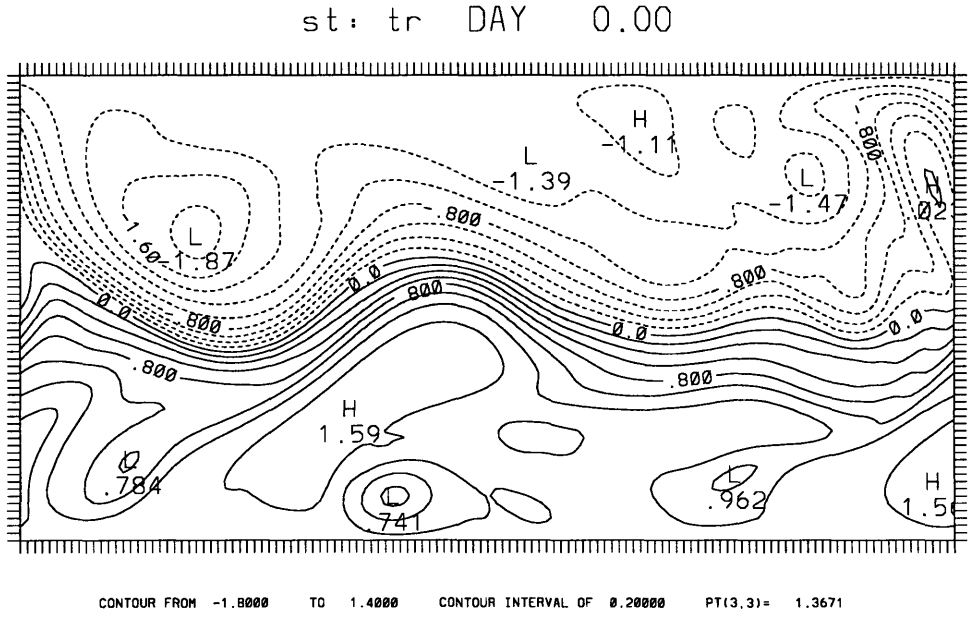
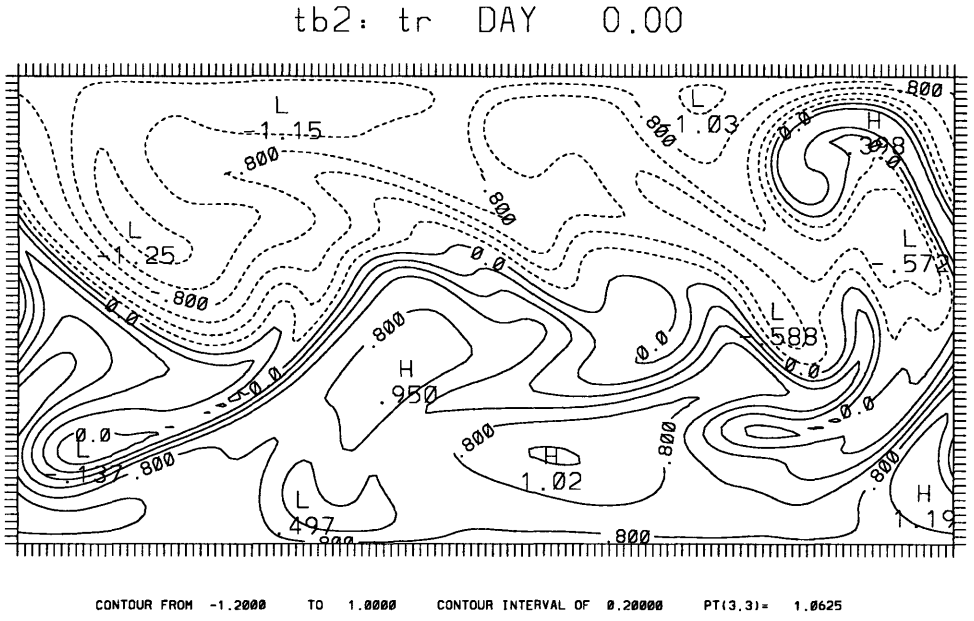



Figure 2.3: Sample vertical cross-sections through the double resolution QG model state at an arbitrary time, for non-dimensionalized: (a) lower boundary streamfunction, (b) upper boundary streamfunction, (c) lower boundary potential temperature, and (d) upper boundary potential temperature. Longitude (periodic) is plotted on the x-axis and latitude is plotted on the y-axis, each with resolution 125 km.

c)




d)

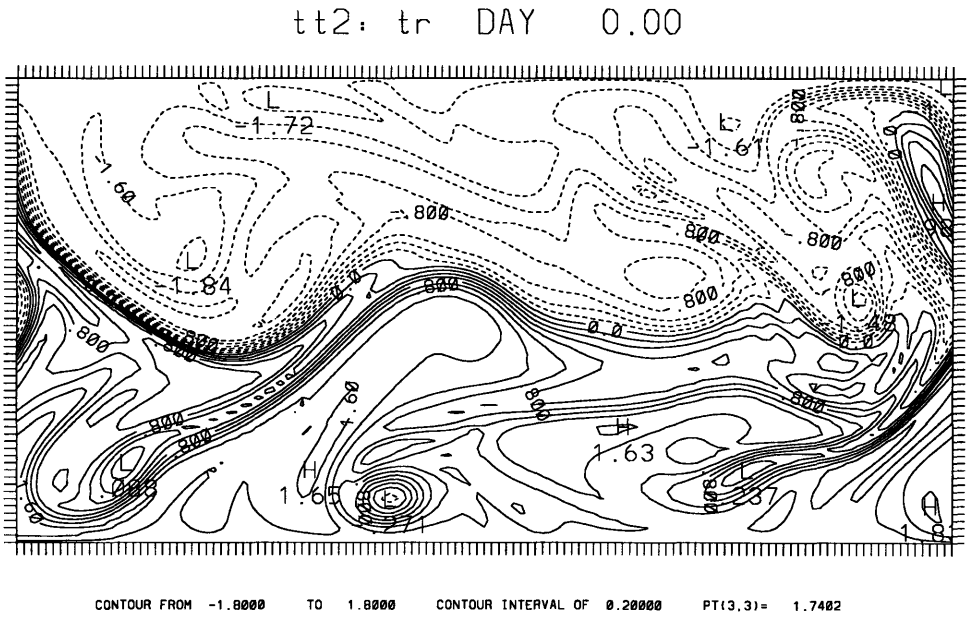



Figure 2.3, continued.

# Chapter 3

# Data assimilation

The data assimilation system controls how observations are incorporated into the forecast model, and therefore it is a key aspect of any data impact study using a relatively complex model. For the initial evaluation of observation strategies, we have implemented a three-dimensional variational (3DVAR) data assimilation scheme, based on the Spectral Statistical-Interpolation (SSI) analysis system currently operational at the National Centers for Environmental Prediction in the U.S. (Parrish and Derber 1992), and similar to those operational at many other weather prediction centers. Although more sophisticated data assimilation systems are currently being developed and implemented, we have selected 3DVAR as the baseline data assimilation system because it has been well-tested, it is similar to operational schemes, it is computationally less intensive than more complex schemes, and it simplifies understanding how the data assimilation affects the observations. Further details on other possible data assimilation schemes and our choice of 3DVAR are discussed in Section 3.1. Sections 3.2, 3.3, and 3.4 present the 3DVAR algorithm and describe its implementation in this study. In Section 3.5, we show several examples of how the 3DVAR assimilates the observations.

## 3.1 Choice of data assimilation

In this section, we briefly outline several different atmospheric data assimilation techniques and explain why we chose to implement 3DVAR. Because data assimilation is a large and rapidly developing field, with each data assimilation implementation slightly different, we have discussed only a few basic atmospheric data assimilation formulations and have avoided going into much detail. For further information on the theory used

to develop variational data assimilation (e.g. Bayesian estimation) and statistical data assimilation in general, the reader is referred to Bengtsson et al. (1981), Daley (1991), and Lorenc (1986), and references therein. For more information on advanced data assimilation topics, the reader is referred to the references listed in Courtier et al. (1993). 3DVAR is described in greater detail in the remaining sections of this chapter, and for details on the other specific data assimilation algorithms discussed, a few sample references are given in the text.

Until several years ago, optimal interpolation (OI) schemes were used to assimilate data at most operational weather prediction centers (e.g. Lorenc 1981). In theory 3DVAR and OI are quite similar. In practice, however, most OI implementations solve several local assimilations, selecting a relevant subset of the observations for each region, while 3DVAR and related variational data assimilation schemes solve for the analysis globally, incorporating all available data at once. The different formulation of OI creates analyses which can be easier to attribute to features in the background field and in the observations. Recently, however, operational centers have been moving away from OI toward variational data assimilation (e.g. Parrish and Derber 1992), and we have chosen 3DVAR for similar reasons. First, OI requires a data selection algorithm, which can be complex and difficult to implement, and which can produce analyses that are inconsistent between regions. In addition, since the 3DVAR analysis is global and is formulated in terms of the analysis variables instead of in terms of pre-defined observation variables, 3DVAR can more easily be adapted to incorporate a variety of data sources (Parrish and Derber 1992). The basic 3DVAR algorithm can also easily be modified to include additional constraints on the analysis, such as a balance condition or other dynamical constraint. Because of this flexibility, we can later adapt the 3DVAR framework to any data assimilation algorithm that can be posed variationally.

Atmospheric variational data assimilation uses predicted error statistics to weight a forecast model state or model trajectory (called the background field) and observations. It then solves for the analysis which has the minimum difference from both the background field and the observations. Although the idealized formulation of variational data assimilation is theoretically appealing and is useful for small-order problems, it cannot currently be used to assimilate atmospheric data in realistic situations for two main reasons. First of all, in atmospheric 3DVAR the full algorithm contains $O(N^2)$ error correlations, where $N$ is equal to the number of analysis variables. In more sophisticated data assimilation formulations the algorithm can be even more complex, with operations on the $O(N^2)$ values. Even in the simplified system studied here, $N$ is $O(10^4)$. Although

we do not have to explicitly invert the equations, it is computationally very expensive to calculate or store all of the values, and it is at best infeasible to solve the full algorithm for a large number of experiments. In addition, because of the large number of error correlations and their detailed time-dependent structure, at each assimilation time very little is known about the error statistics which determine the weights for the inversion. Therefore, to implement a practical variational data assimilation for numerical weather prediction, we must generally simplify the basic algorithm.

Each type of data assimilation and each implementation makes different assumptions, and it is not yet clear which assumptions are the most valid or which work the best. For 3DVAR, we develop a feasible algorithm by simplifying the error statistics, particularly the structures of the background errors; the assumptions used in our version of 3DVAR are described in Section 3.4. Our 3DVAR behaves similarly to data assimilation systems at most operational weather centers. With only these simplified statistics, however, the 3DVAR has little knowledge of the atmospheric dynamics, and thus it is far from optimal. Several more complex data assimilation algorithms have been proposed to incorporate information about time- and space-dependent errors, and by doing so they can greatly improve how the data assimilation recreates the atmospheric state from the observations. To be feasible, these algorithms must be simplified even more than 3DVAR, but as computational resources increase, they will become possible for more implementations.

One possibility, the Kalman filter, has been applied in a variety of fields, and it may be useful for atmospheric data assimilation in a modified form (e.g. Ghil 1997 and references therein). The Kalman filter, instead of specifying the background error statistics externally (as in 3DVAR), uses the forecast model explicitly to update the statistics at each data assimilation time. A second example is four-dimensional variational data assimilation (4DVAR), which is similar to 3DVAR but minimizes in time as well as in space (e.g. Courtier and Talagrand 1987, Talagrand and Courtier 1987, and Thepaut et al. 1993). In 4DVAR the background field is a forecast model trajectory over a specified time period instead of model state at a specific time. Both the Kalman filter and 4DVAR include adjoint models, and simplified forms of 4DVAR using an adjoint or a quasi-inverse model have also been proposed (e.g. Pu et al. 1997). Another possibility is the "ensemble Kalman filter," a 3DVAR-like algorithm which uses an ensemble of perturbed forecasts to estimate the correlations between errors in the background field in real time (Houtekamer and Mitchell 1998). Several other proposed data assimilation strategies, such as those which combine ensemble perturbations directly (K. Emanuel,

D. Parrish, Z. Toth, personal communication; Kalnay and Toth 1994) or try to track coherent atmospheric structures (M. Ghil, personal communication; Ide and Ghil 1998), may also be beneficial for certain atmospheric applications. These techniques, however, are in more preliminary stages of development.

With a different data assimilation system, the same forecasting system uses observations differently. Therefore, our results for different observation strategies are likely to be somewhat sensitive to our choice of data assimilation. We selected 3DVAR among the many schemes currently being developed and implemented for several reasons. First, we required a data assimilation that would allow us to perform a large number of experiments, and 3DVAR is computationally much cheaper than most more sophisticated algorithms. Most of the advanced schemes are also in more preliminary stages of development than 3DVAR, and they must be tested further before they can be well-implemented fairly easily. In addition, preliminary results, from both our simulated system and from the first field tests of adaptive observations, indicate that in order to adapt observation networks effectively, we must understand more about how data assimilation systems and forecast models in general interact with observations limited in space and time. Since the interaction of 3DVAR with observations is not yet fully understood, implementing a more sophisticated scheme early in the study would be premature.

We are most concerned with how the data assimilation incorporates observations and how the forecast model interacts with the resulting analysis increments at times when errors in the initial conditions are likely to be important for a forecast, in other words in precisely the situations when we want to adapt observations. If we do not understand how a data assimilation and forecast model system behaves in a variety of situations, it will be easy to confuse the strengths and weaknesses of how the data assimilation interacts with observations and forecasts with the results from the strategies themselves. Because any realistic data assimilation will always be limited by imperfect information, this problem cannot be avoided. 3DVAR, as a representative data assimilation system, can help teach us how effectively (or ineffectively) any imperfect data assimilation system might use observations in certain situations. It can also help us learn how to think about and how to approach the interactions between initial condition errors, observations, data assimilation, and forecast models in general. This knowledge will help us design not only better observation strategies, but also better data assimilation schemes and well-integrated observing and forecasting systems.

## 3.2 Formulation of three-dimensional variational data assimilation

3DVAR produces an analysis field by combining a weighted short-term model forecast (also called the first guess or the background field) with weighted observations. The weights are determined by statistics both for the errors in the background field and for the errors in the observations. The interpolation and extrapolation of the observations to the analysis locations is determined by the expected magnitudes of the errors at each location (the variances) and the expected correlations between the errors at different locations (the covariances). The observation variables can be of any type at any locations and generally change for each data assimilation time, depending on which observations are available with acceptable quality and timeliness and which we wish to use. In the case of satellite observations, they can even be incorporated over a region rather than at specific locations. The analysis variables are generally related to the variables for the background field (the forecast model variables at model gridpoints). Because in our implementation the analysis variables and the background variables are the same, we refer to them interchangeably.

If at a certain time we have

$M$ = number of observations

and

$N$ = number of analysis variables,

we can define

$\mathbf{L}$ = an operator that transforms from the analysis variables to observation variables

and calculate

$\mathbf{y}$ = an $M$-component vector of observation residuals (the difference between observations and the background field at observation locations).

To assimilate the observations, we then solve for

$\mathbf{x}$ = an $N$-component vector of analysis increments (the difference between the analysis and the background field at analysis locations)

by minimizing the objective function

$$J = \frac{1}{2}[\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{x} + (\mathbf{L}\mathbf{x} - \mathbf{y})^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}(\mathbf{L}\mathbf{x} - \mathbf{y})], \tag{3.1}$$

where

$\mathbf{B}$ = the $N \times N$ matrix of covariances between the background errors (called the background error covariance matrix),

$\mathbf{O}$ = the $M \times M$ observation error covariance matrix, and

$\mathbf{F}$ = the $M \times M$ observation representativeness covariance matrix.

$\mathbf{L}\mathbf{x} - \mathbf{y}$, the difference between the analysis and the observations, is weighted according to the inverse of the observation and representativeness error covariances ($(\mathbf{O} + \mathbf{F})^{-1}$). $\mathbf{x}$, the difference between the analysis and the background field, is weighted according to the inverse of the background error covariances ($\mathbf{B}^{-1}$). Other constraints or conditions can also, if desired, be added to the right hand side of Equation (3.1).

We minimize the weighted difference of the analysis from the observations and from the background field by setting the derivative of $J$ in Equation (3.1) equal to 0:

$$\mathbf{B}^{-1}\mathbf{x} + \mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}(\mathbf{L}\mathbf{x} - \mathbf{y}) = 0. \tag{3.2}$$

Multiplying Equation (3.2) by $\mathbf{B}$ and rearranging, we obtain:

$$[\mathbf{I} + \mathbf{B}\mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}\,\mathbf{L}]\,\mathbf{x} = \mathbf{B}\mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}\,\mathbf{y}. \tag{3.3}$$

At each assimilation time, the 3DVAR solves Equation (3.3) for the analysis increments $\mathbf{x}$ using an iterative conjugate residual solver (described in the Appendix). The analysis increment is then added to the background field to produce the analysis, which becomes the initial conditions for the next model run. In the standard runs shown here, simulated observations are taken and assimilated by inverting Equation (3.3) every 12 hours; this 12 hour period is called the data assimilation interval.

The 3DVAR algorithm can easily be adapted to include many types of observations at many locations. For simplicity, however, all observations in these experiments simulate rawinsondes measuring winds and (potential) temperature at model gridpoints and model levels. $\mathbf{O} + \mathbf{F}$ and $\mathbf{y}$ are defined in observation space, which is then winds and temperature at observation locations. $\mathbf{B}$ and $\mathbf{x}$ are defined in analysis space, which is the

QG model variables (potential vorticity at interior model levels and potential temperature at the upper and lower boundaries) at model gridpoints. The analysis minimization is also performed in analysis space; as discussed in Section 3.1, this is one of the major advantages of the variational formulation. The operator $\mathbf{L}$ converts from analysis space to observation space by inverting potential vorticity to obtain streamfunction and then taking derivatives to calculate winds and temperatures.

## 3.3 Observation error covariances

The observation errors are assumed to be uncorrelated between different rawinsondes and between wind and temperature observations (Dey and Morone, 1985). They are therefore correlated only in the vertical and only for the same variable. The wind and temperature observation error variances at different levels are adapted from the values for rawinsonde observation errors given in Parrish and Derber (1992). Since representativeness of observations is not an issue in this idealized system (there are no sub-grid scale features in the true atmosphere which are unresolved by the model atmosphere), and since representativeness error estimates are generally not separated from rawinsonde instrument error estimates, we do not estimate representativeness errors explicitly. To develop the matrix of observation error covariances $\mathbf{O} + \mathbf{F}$, the variances adapted from Parrish and Derber (1992) are combined with vertical observation error correlations obtained from the simple function given in Equation (3.19) in Bergman (1979) (with vertical correlations a function of the distance between the levels only, not the absolute values of the levels).

Table 3.1 shows the non-dimensionalized covariances for the wind and temperature observation errors in each simulated rawinsonde at the standard resolution. Instead of specifying errors for wind speed and direction, we approximate that the observation errors for zonal wind are .8 $\times$ the total wind observation errors and that errors for meridional wind are .5 $\times$ the total wind errors. The same matrices are shown for double resolution in Table 3.2. We have tested the sensitivity of the results to this formulation and to these specific values of $\mathbf{O}$, and it is minimal.

## 3.4 Background error covariances

The background error statistics ($\mathbf{B}$) specified in the 3DVAR control how the data

$\mathbf{O}_{\text{wind}}$

| *level* | *0* | *1* | *2* | *3* | *4* | *5* | *H* |
|---|---|---|---|---|---|---|---|
| *0* | **1.96** | 2.28 | 1.85 | 1.15 | 0.77 | 0.54 | 0.43 |
| *1* | 2.28 | **3.20** | 3.19 | 1.97 | 1.25 | 0.84 | 0.66 |
| *2* | 1.85 | 3.19 | **6.10** | 5.00 | 3.04 | 1.86 | 1.41 |
| *3* | 1.15 | 1.97 | 5.00 | **7.84** | 6.32 | 3.71 | 2.68 |
| *4* | 0.77 | 1.25 | 3.04 | 6.32 | **9.77** | 7.60 | 5.45 |
| *5* | 0.54 | 0.84 | 1.86 | 3.71 | 7.60 | **11.32** | 9.98 |
| *H* | 0.43 | 0.66 | 1.41 | 2.68 | 5.45 | 9.98 | **10.57** |

$\mathbf{O}_{\text{temperature}}$

| *level* | *0* | *1* | *2* | *3* | *4* | *5* | *H* |
|---|---|---|---|---|---|---|---|
| *0* | **3.24** | 2.53 | 1.02 | 0.51 | 0.38 | 0.29 | 0.25 |
| *1* | 2.53 | **2.58** | 1.32 | 0.63 | 0.44 | 0.32 | 0.28 |
| *2* | 1.02 | 1.32 | **1.69** | 1.07 | 0.66 | 0.43 | 0.35 |
| *3* | 0.51 | 0.63 | 1.07 | **1.69** | 1.39 | 0.80 | 0.62 |
| *4* | 0.38 | 0.44 | 0.66 | 1.39 | **2.82** | 2.18 | 1.60 |
| *5* | 0.29 | 0.32 | 0.43 | 0.80 | 2.18 | **4.18** | 3.90 |
| *H* | 0.25 | 0.28 | 0.35 | 0.62 | 1.60 | 3.90 | **4.76** |

Table 3.1: Observation error covariance matrices used to assimilate simulated rawinsondes into the QG model at the standard resolution (observations at each of the 5 vertical levels and at the upper and lower boundaries), for dimensionalized (a) wind and (b) potential temperature. The diagonal elements are error variances at each level and are bold-faced; the off-diagonal elements are covariances between errors at different levels. The $\mathbf{O}$ used for zonal wind is $(.8)^2 \times \mathbf{O}_{\text{wind}}$, and the $\mathbf{O}$ for meridional wind is $(.5)^2 \times \mathbf{O}_{\text{wind}}$. To recover non-dimensional values, divide $\mathbf{O}_{\text{wind}}$ by $U^2$ $(= (60\,\text{ms}^{-1})^2)$ and divide $\mathbf{O}_{\text{temperature}}$ by $(\theta_o N^2 H R_o / g)^2$ $(= (10.16^\circ\text{K})^2)$, using the constants defined in Chapter 2.

$\mathbf{O}_{\text{wind}}$

| *level* | *0* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *H* |
|---|---|---|---|---|---|---|---|---|---|---|
| *0* | **1.96** | 2.22 | 2.23 | 1.80 | 1.33 | 1.01 | 0.78 | 0.63 | 0.49 | 0.43 |
| *1* | 2.22 | **2.70** | 3.04 | 2.56 | 1.88 | 1.40 | 1.07 | 0.85 | 0.65 | 0.57 |
| *2* | 2.23 | 3.04 | **4.53** | 4.61 | 3.59 | 2.63 | 1.94 | 1.48 | 1.10 | 0.94 |
| *3* | 1.80 | 2.56 | 4.61 | **6.21** | 5.84 | 4.52 | 3.29 | 2.42 | 1.74 | 1.46 |
| *4* | 1.33 | 1.88 | 3.59 | 5.84 | **7.27** | 6.80 | 5.23 | 3.80 | 2.63 | 2.17 |
| *5* | 1.01 | 1.40 | 2.63 | 4.52 | 6.80 | **8.42** | 7.83 | 6.00 | 4.10 | 3.32 |
| *6* | 0.78 | 1.07 | 1.94 | 3.29 | 5.23 | 7.83 | **9.64** | 8.93 | 6.44 | 5.21 |
| *7* | 0.63 | 0.85 | 1.48 | 2.42 | 3.80 | 6.00 | 8.93 | **10.95** | 9.56 | 8.04 |
| *8* | 0.49 | 0.65 | 1.10 | 1.74 | 2.63 | 4.10 | 6.44 | 9.56 | **11.03** | 10.41 |
| *H* | 0.43 | 0.57 | 0.94 | 1.46 | 2.17 | 3.32 | 5.21 | 8.04 | 10.41 | **10.57** |

$\mathbf{O}_{\text{temperature}}$

| *level* | *0* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *H* |
|---|---|---|---|---|---|---|---|---|---|---|
| *0* | **3.24** | 2.86 | 1.72 | 0.97 | 0.62 | 0.46 | 0.38 | 0.32 | 0.28 | 0.25 |
| *1* | 2.86 | **2.82** | 1.97 | 1.15 | 0.72 | 0.52 | 0.42 | 0.35 | 0.29 | 0.27 |
| *2* | 1.72 | 1.97 | **2.06** | 1.52 | 0.98 | 0.67 | 0.52 | 0.41 | 0.33 | 0.30 |
| *3* | 0.97 | 1.15 | 1.52 | **1.69** | 1.38 | 0.97 | 0.71 | 0.53 | 0.41 | 0.36 |
| *4* | 0.62 | 0.72 | 0.98 | 1.38 | **1.69** | 1.51 | 1.13 | 0.81 | 0.59 | 0.51 |
| *5* | 0.46 | 0.52 | 0.67 | 0.97 | 1.51 | **2.01** | 1.92 | 1.41 | 0.98 | 0.82 |
| *6* | 0.38 | 0.42 | 0.52 | 0.71 | 1.13 | 1.92 | **2.74** | 2.56 | 1.83 | 1.50 |
| *7* | 0.32 | 0.35 | 0.41 | 0.53 | 0.81 | 1.41 | 2.56 | **3.59** | 3.24 | 2.74 |
| *8* | 0.28 | 0.29 | 0.33 | 0.41 | 0.59 | 0.98 | 1.83 | 3.24 | **4.39** | 4.33 |
| *H* | 0.25 | 0.27 | 0.30 | 0.36 | 0.51 | 0.82 | 1.50 | 2.74 | 4.33 | **4.76** |

Table 3.2: As in Table 3.1, but for double resolution (observations at each of the 8 vertical levels and at the upper and lower boundaries). The scaling is the same.

assimilation spreads the information from the observations to non-observed locations to increment the analysis, and, as described in Section 3.1, these statistics must be simplified to create a feasible algorithm. To make the solution of Equation (3.3) possible for a variety of observational networks, we assume that:

1. $\mathbf{B}$ is fixed in time;
2. $\mathbf{B}$ is diagonal in horizontal spectral coordinates;
3. $\mathbf{B}$ has separable vertical and horizontal structures and simple vertical correlations; and
4. $\mathbf{B}$ is the same for all observation densities.

Assumptions (1)–(2) follow Parrish and Derber (1992). Assumption (1) allows us to use time-averaged statistics to calculate $\mathbf{B}$ once instead of estimating it at each data assimilation time. Assumption (2) produces background errors which are uncorrelated between different horizontal spatial scales and reduces $\mathbf{B}$ from $\mathrm{O}(N^2)$ to $\mathrm{O}(N)$. Since we are using a gridpoint model, to create a $\mathbf{B}$ diagonal in spectral space we must transform to spectral coordinates. As in Parrish and Derber (1992) for their satellite observation error covariance matrix, we write

$$\mathbf{B} = \mathbf{S}\,\mathbf{C}\,\mathbf{S}^{\mathrm{T}}, \tag{3.4}$$

where $\mathbf{S}$ is the transform from spectral to gridpoint coordinates, $\mathbf{S}^{\mathrm{T}}$ is the adjoint operator of $\mathbf{S}$, and $\mathbf{C}$ is the $N \times N$ background error covariance matrix in spectral coordinates. We then set off-diagonal elements of $\mathbf{C}$ to 0.

Forecast errors in the QG model tend to be strongly correlated between potential temperature at the boundary and potential vorticity at the closest interior level. Without vertical correlations in $\mathbf{B}$, the 3DVAR had insufficient information about these correlations from the observations and from $\mathbf{L}$, and it partitioned the analysis increments suboptimally between the upper and lower boundaries and the interior. Since in the inversion to obtain streamfunction, $\mathbf{L}$ treats boundary potential temperature as equivalent to potential vorticity at the closest interior level, this minor problem was not evident in the analyses in terms of streamfunction, wind, or temperature. It was apparent only in the interior potential vorticity analysis increments compared to the interior potential vorticity background errors.

When adding to $\mathbf{B}$ covariances between different levels, we decided to make Assumption (3), constructing the simplest vertical correlations which allow the 3DVAR to perform reasonably well, for several reasons. First of all, since Assumptions (1)–(2) prevent the background error covariances from including vertical correlations which could account for the baroclinic structure in specific atmospheric systems, the vertical structure in $\mathbf{B}$ can make only limited improvements to the data assimilation. In addition, because the vertical levels are strongly coupled by the inversion of QG potential vorticity included in $\mathbf{L}$, the QG model 3DVAR already vertically constrains the analysis increments. Externally specifying vertical structure in the background error covariances could therefore make it perform worse (D. Parrish, personal communication). Assumption (3) is also similar to assumptions made in most atmospheric data assimilation implementations, and it simplifies the specification of $\mathbf{B}$ and the inversion of Equation (3.3) greatly.

With $\mathbf{C}$ defined as in Equation (3.4), we first write

$$\mathbf{C} = \hat{\mathbf{C}}^{\frac{1}{2}} \mathbf{V} \hat{\mathbf{C}}^{\frac{1}{2}}, \tag{3.5}$$

where $\hat{\mathbf{C}}$ is the matrix of horizontal background error covariances at each level and $\mathbf{V}$ is the matrix of background error correlations between different levels for each horizontal spectral component (different horizontal spectral components are assumed to be uncorrelated between different levels according to Assumption (2)). The vertical error correlations are then assumed to be the same for all horizontal scales, reducing $\mathbf{V}$ to block diagonal form with a single $L \times L$ matrix repeated on the diagonal, where $L$ is the number of vertical levels plus the upper and lower boundaries. Instead of developing vertical structure functions (as is often done), we leave $\mathbf{V}$ as correlations between different levels averaged across all spectral components. At the end of this section, we describe in further detail how $\mathbf{V}$ was developed for the QG model and evaluate the validity of the assumption that the vertical correlations are scale-independent.

It was not known a priori how the system would respond when the observation density, the observation strategy, or the data assimilation was altered. Therefore, in order to separate the effects of changes in the observation network from the effects of changes in the data assimilation, we also make Assumption (4), that the background error statistics are the same for all observation densities and all data assimilation intervals. Assumptions (1) and (2) already neglect the space and time variability in background errors associated with any individual observation network. Therefore, Assumption (4) can be

thought of as an extension to neglecting the variability in $\mathbf{B}$ among different observation networks. Although this is clearly not valid given the large changes in the magnitude of the background errors, experiments with the background error statistics show that this assumption does not affect the results qualitatively and that the optimization density selected is reasonable; several examples are shown in Section 5.3.3.

With these assumptions, the background error statistics can be optimized for a fixed density and a standard data assimilation interval and used in the 3DVAR for all observation networks. $\mathbf{B}$ was developed according to the following algorithm:

1. Supply a first guess for $\mathbf{B}$.

2. Run the simulation experiment as described in Section 4.1 for a specified observation density, and, after the spin-up time, calculate the background errors (in this case 12 hour forecast errors) explicitly during a long run.

3. Given the assumptions about the structure of $\mathbf{B}$, accumulate statistics for the background error covariances over many such runs.

4. Return to Step 1, inserting the new statistics into $\mathbf{B}$.

This procedure was iterated several times; after the first iteration, the statistics changed very little. Unlike in operational 3DVAR systems, where the background error is not known and must be estimated, in this idealized system the true background errors can be used to develop the background error covariances directly. Therefore, to the extent that our assumptions about their structure are valid, the weights in our 3DVAR are correct.

For the standard data assimilation, statistics for $\mathbf{B}$ were accumulated with several different distributions of 32 fixed observations (approximately 1.5% on the normalized observation density scale in the figures below) and a 12 hour data assimilation interval. This optimization density was originally chosen because it was the lowest observation density for which the statistics, especially those for the largest horizontal scales, remained reasonably consistent. Figure 3.1 shows the horizontal error variances in $\mathbf{B}$ for lower and upper boundary potential temperature at the standard resolution, as a function of approximate global wavenumber (recall that different spectral components are assumed to be uncorrelated). Separate statistics were also developed for potential vorticity at each interior level, and they look similar to and transition between those shown as one goes from the lower to the upper boundary.

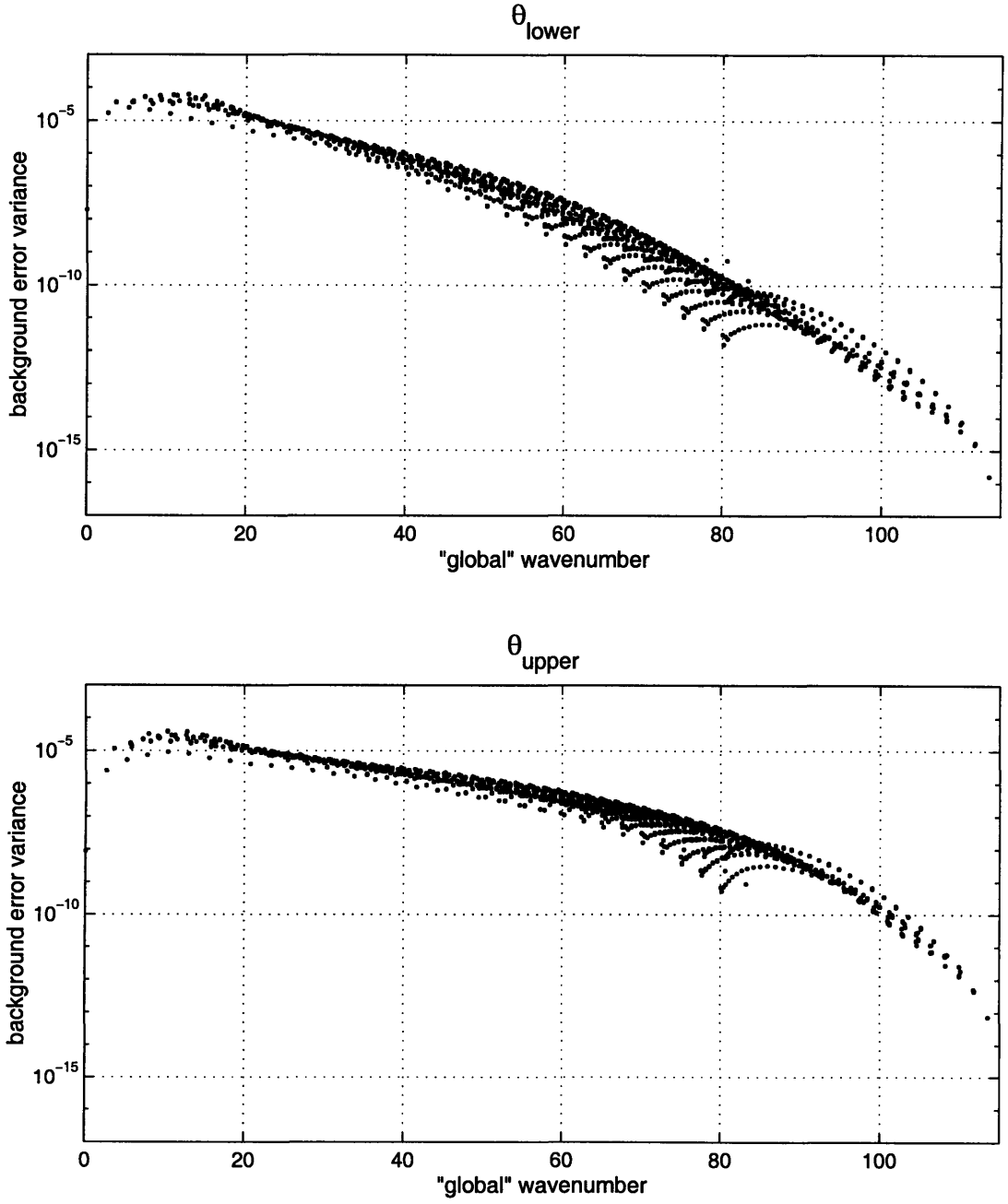



Figure 3.1: Horizontal background error variances (diagonal elements of **B**) in the standard resolution 3DVAR as a function of approximate global wavenumber, for non-dimensionalized (a) lower boundary potential temperature and (b) upper boundary potential temperature. Global wavenumber is not well-defined in the channel model, and it is shown here only to suggest how the error variances compare at different spatial scales. For zonal wavenumber $k$ and meridional half-wavenumber $l$ as represented in the quasi-geostrophic model, we approximate global wavenumber as $( (2.5 \times k)^2 + (5.2 \times .5 \times l)^2 )^{\frac{1}{2}}$. The factors of 2.5 and 5.2 are included to adjust for the zonal and meridional extent of the channel model compared to the "real" globe.

| *level* | *0* | *1* | *2* | *3* | *4* | *5* | *H* |
|---|---|---|---|---|---|---|---|
| *0* | 1.0 | -0.78 | -0.35 | 0.0 | 0.0 | 0.0 | 0.0 |
| *1* | -0.78 | 1.0 | 0.10 | 0.0 | 0.0 | 0.0 | 0.0 |
| *2* | -0.35 | 0.10 | 1.0 | 0.06 | 0.0 | 0.0 | 0.0 |
| *3* | 0.0 | 0.0 | 0.06 | 1.0 | 0.14 | 0.0 | 0.0 |
| *4* | 0.0 | 0.0 | 0.0 | 0.14 | 1.0 | 0.55 | -0.55 |
| *5* | 0.0 | 0.0 | 0.0 | 0.0 | 0.55 | 1.0 | -0.91 |
| *H* | 0.0 | 0.0 | 0.0 | 0.0 | -0.55 | -0.91 | 1.0 |

Table 3.3: Correlations between background errors at different levels averaged over all horizontal scales, at the standard resolution. Correlations which were different signs at different horizontal spatial scales were set to 0. Note the strong negative correlations between errors in the boundary temperature and errors in the interior PV at the closest levels. The correlations also tend to be larger at upper levels.

The forecast error variances in the quasi-geostrophic model decrease rapidly as the horizontal scale becomes small, more so at lower levels where the Ekman pumping has a direct effect. Since the QG model and the experimental setup contain neither realistic small scale dynamics nor model errors (including unresolved, sub-grid-scale processes), it is not clear how the QG model 3DVAR behavior at small spatial scales compares to data assimilation for real atmospheric systems at similar scales. The horizontal structure in the $\mathbf{B}$ used for double resolution is shown for lower and upper boundary potential temperature in Figure 3.2. The background error variances are larger at double resolution, presumably due to the faster error growth discussed in Chapter 2.

Statistics for the vertical background error correlations $\mathbf{V}$ were accumulated in parallel with the horizontal covariances. To develop scale-independent vertical correlations that would not interfere too much with other aspects of the 3DVAR, the correlations between each set of two levels were averaged over all horizontal spectral components. We then set to 0 all correlations which changed sign for different spectral components, as was the case for non-adjacent interior levels. The resulting vertical background error covariances used in the 3DVAR are shown in Table 3.3 for the standard resolution and in Table 3.4 for double resolution.

As discussed in Hollingsworth and Lönnberg (1986), assuming that the vertical and horizontal structures are separable and that the vertical correlations are homogeneous is not automatically valid, but it is a typical assumption and is generally acceptable as long as the levels are reasonably close to each other. An example of how the average

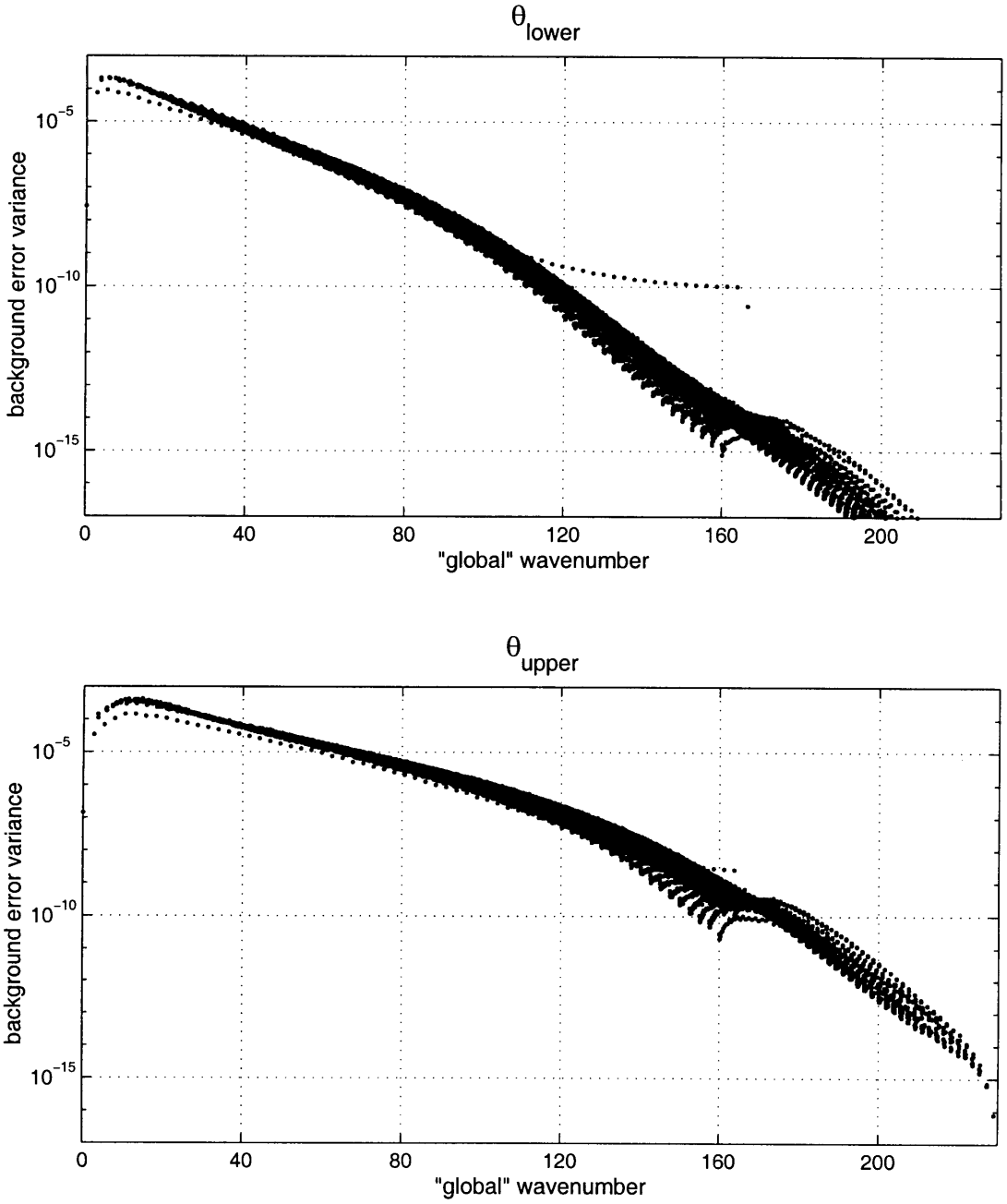



Figure 3.2: As in Figure 3.1, but in the double resolution 3DVAR.

| *level* | *0* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *H* |
|---|---|---|---|---|---|---|---|---|---|---|
| *0* | 1.0 | -0.69 | -0.59 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| *1* | -0.69 | 1.0 | 0.51 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| *2* | -0.59 | 0.51 | 1.0 | 0.30 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| *3* | 0.0 | 0.0 | 0.30 | 1.0 | 0.17 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| *4* | 0.0 | 0.0 | 0.0 | 0.17 | 1.0 | 0.17 | 0.0 | 0.0 | 0.0 | 0.0 |
| *5* | 0.0 | 0.0 | 0.0 | 0.0 | 0.17 | 1.0 | 0.24 | 0.0 | 0.0 | 0.0 |
| *6* | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.24 | 1.0 | 0.41 | 0.0 | 0.0 |
| *7* | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.41 | 1.0 | 0.71 | -0.71 |
| *8* | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.71 | 1.0 | -0.98 |
| *H* | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -0.71 | -0.98 | 1.0 |

Table 3.4: As in Table 3.3, but at double resolution.

vertical correlations vary with the horizontal spectral scale is shown in Figure 3.3 for the correlations between the lower boundary potential temperature and all other levels at the standard resolution. The average correlation between lower boundary potential temperature and lowest interior level potential vorticity is reasonably scale-independent; the other correlations are not. As described earlier, these data were averaged over all spectral components to develop the scale-independent vertical correlations in $\mathbf{B}$. The correlations plotted in Figure 3.3 correspond to the first row or first column of Table 3.3. The entries in Table 3.3 were set to zero for the vertical correlations between lower boundary potential temperature and all but the two lowest interior potential vorticity level because these correlations change sign at different horizontal scales. Correlations between errors at other levels look similar, although correlations between interior levels tend to be more situation-dependent, particularly for lower–middle levels. This produces smaller time-averaged correlations and the smaller scale-averaged correlations in Table 3.3.

The background error statistics given by Assumptions (1)–(4) are nearly isotropic, are zonally invariant, and have limited knowledge of specific atmospheric dynamical situations (see Section 3.5.1 for an example). As described in Section 3.1, more information about the time- and space-dependent error statistics is being gathered and more sophisticated data assimilation schemes are being developed so that these assumptions can be relaxed. Assumptions similar to ours are necessary with the technology and information currently available, however, to develop a data assimilation which is effective and reliable, yet quick enough to run many times. The statistics in our 3DVAR are also similar to those in most operational data assimilation systems. Nevertheless, the relationships

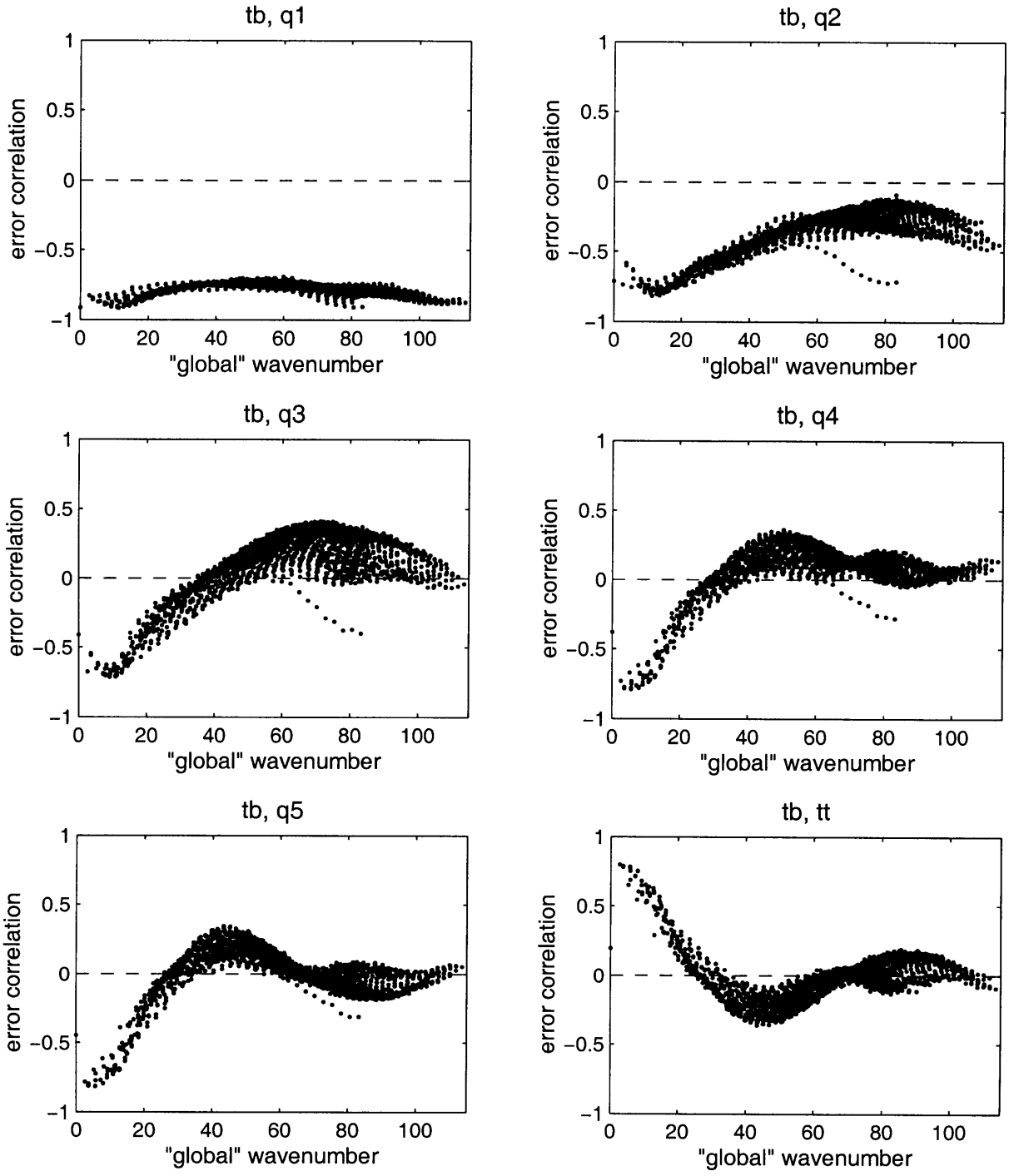



Figure 3.3: Vertical correlations at standard resolution between background (12 hour forecast) errors in lower boundary potential temperature and background errors in: (a) level 1 potential vorticity, (b) level 2 potential vorticity, (c) level 3 potential vorticity, (d) level 4 potential vorticity, (e) level 5 potential vorticity, and (f) upper boundary potential temperature, as a function of "global" wavenumber as defined in Figure 3.1.

between the observations, the data assimilation, and the forecast model are complex, and the data assimilation remains a key aspect of how this simulated system responds to any changes in observations. Therefore, some examples of the sensitivity to the data assimilation are discussed in Sections 5.3 and 5.4, and possible limitations of the data assimilation, and of atmospheric data assimilation in general, are taken into account when evaluating the results.

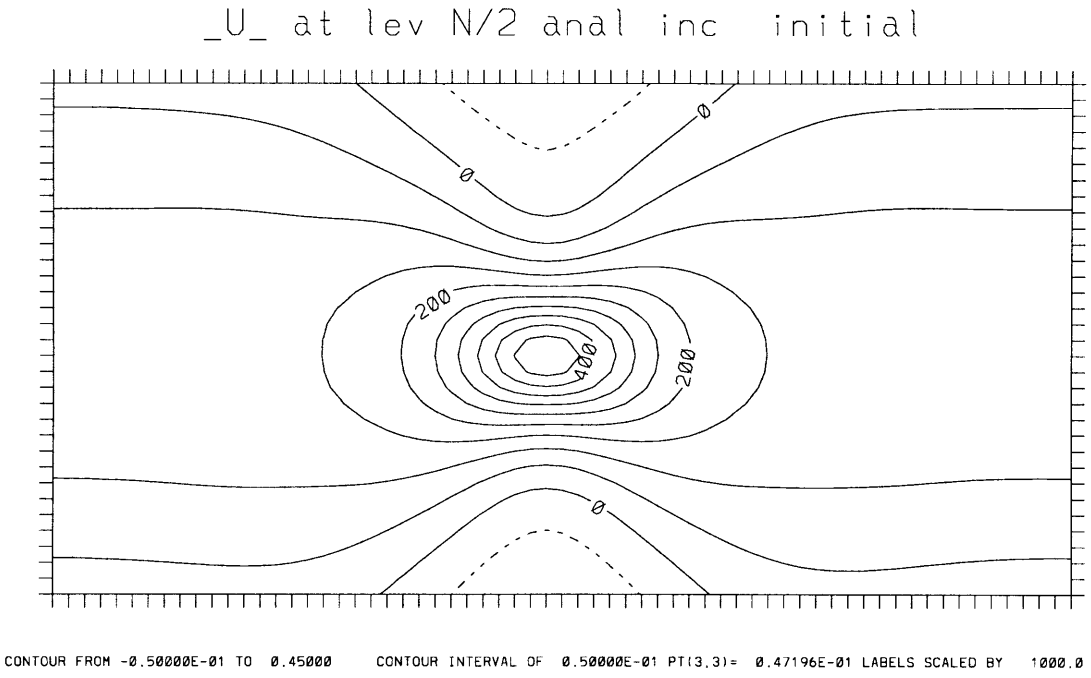
## 3.5 Analysis increments

### 3.5.1 Sample analysis for one observation

With the statistics described in Sections 3.3 and 3.4, the QG model 3DVAR produces analysis increments similar to those shown in Parrish and Derber (1992) for a single observation at mid-latitudes. Figure 3.4 shows an example of the 3DVAR analysis increments at the observation level for one zonal wind observation at the middle model level. Because the analysis variables are interior potential vorticity and boundary potential temperature, the 3DVAR infers (from **B** and **L**) that the zonal wind measurement must have resulted from a dipole in potential vorticity, shown in Figure 3.4d. This potential vorticity dipole in the analysis has wind and temperature structures associated with it, which are recovered using the operator **L**. As in Parrish and Derber (1992), these analysis increments depict the background error correlations in physical space. Since the 3DVAR has no information about the specific atmospheric structure that the observation came from, the single zonal wind observation has been interpolated nearly isotropically.

The 3DVAR interpolates and extrapolates not only horizontally, but also vertically. Figure 3.5 shows, for the same analysis as in Figure 3.4, analysis increments at the lower boundary. In this example, because there is an observation at only one level, the 3DVAR infers the vertical structure of the background error from **L**, from the vertical observation error covariances, and from the vertical structure in **B**. As described in Section 3.4, the analysis increments at different levels are strongly coupled by the inversion of QG potential vorticity included in **L**. Therefore, even though the vertical correlations at middle levels explicitly given to the 3DVAR (shown in Table 3.3) are small, an observation taken at only one level can significantly influence the analysis at other levels. In most situations, however, observations are taken at several levels, so the 3DVAR also determines some of the vertical structure of the analysis increment from the additional

a)




b)

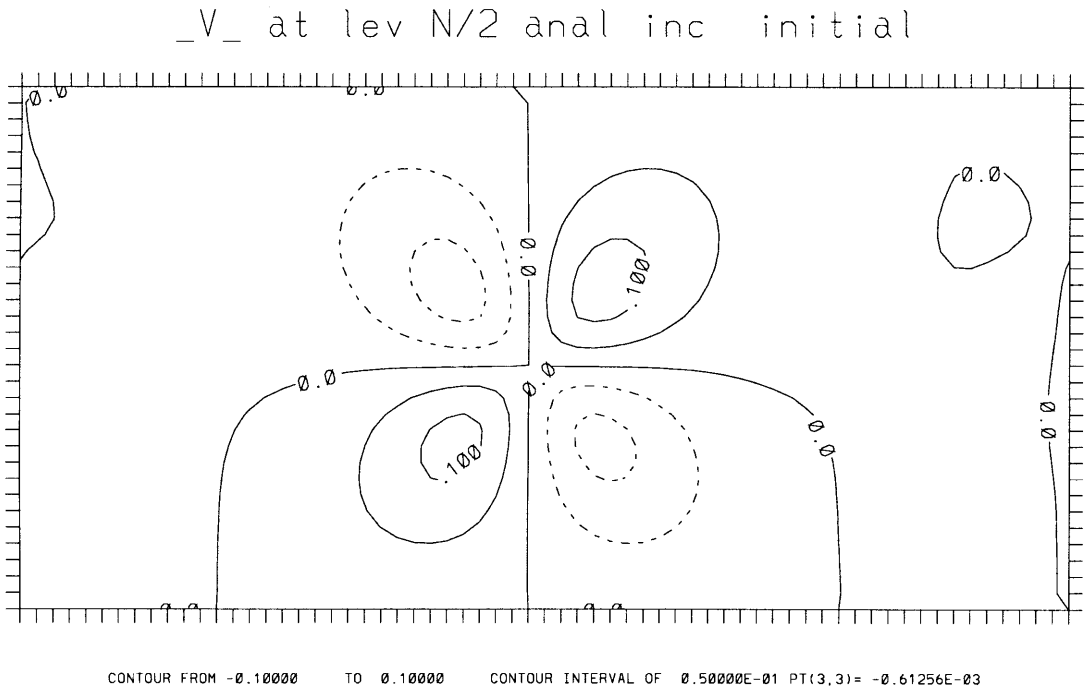
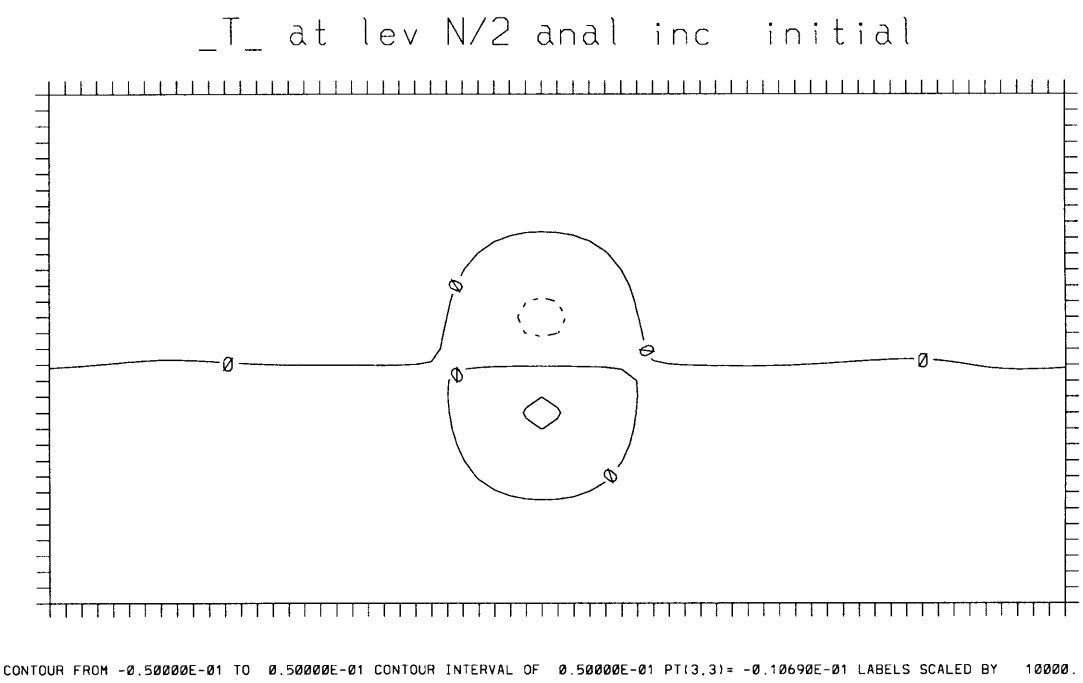



Figure 3.4: 3DVAR analysis increments for a single zonal wind observation $u = 1$ at level 3 (the middle level) in the QG model at standard resolution. The variables plotted are non-dimensionalized: (a) zonal wind, (b) meridional wind, (c) temperature, and (d) potential vorticity, all at level 3. Longitude is plotted on the x-axis and latitude on the y-axis, each with resolution 250 km. Positive contours are solid lines and negative contours are dashed lines; all contour intervals are .05.

c)




d)

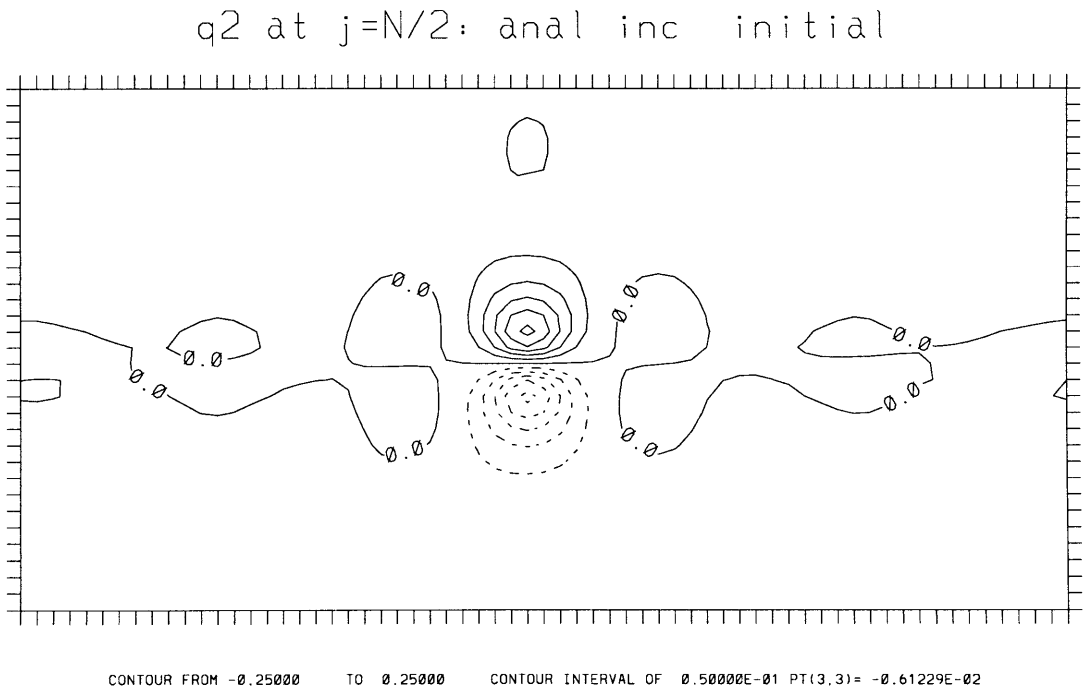
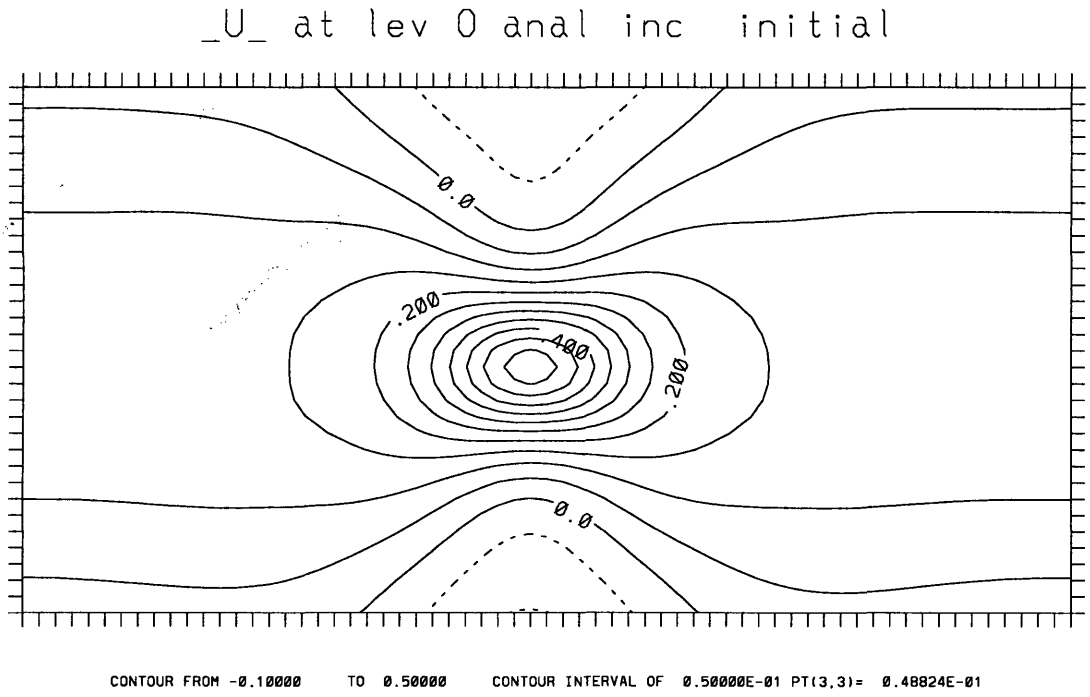



Figure 3.4, continued.

a)




b)

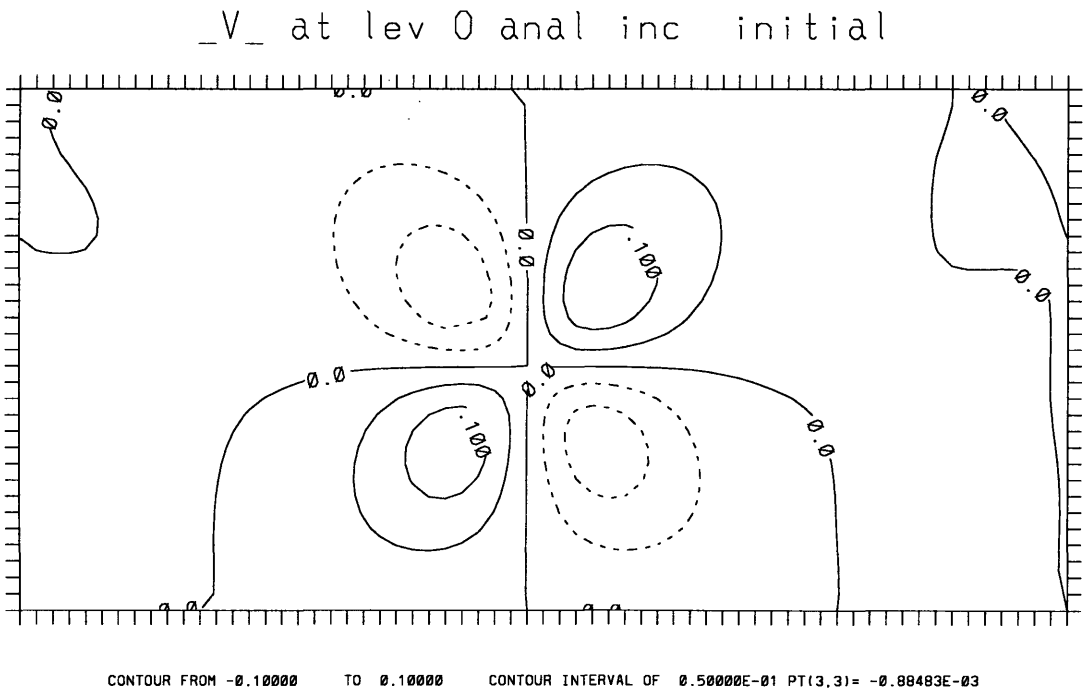
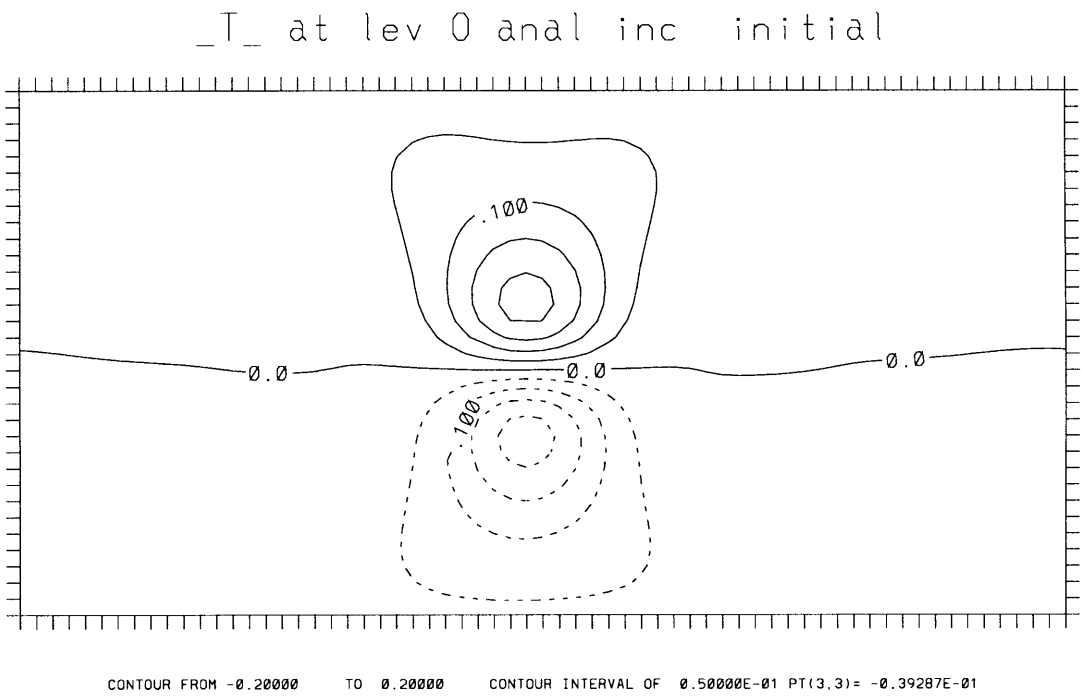



Figure 3.5: 3DVAR analysis increments for the same level 3 zonal wind observation as in Figure 3.4, but plotted at the lower boundary.

c)




d)

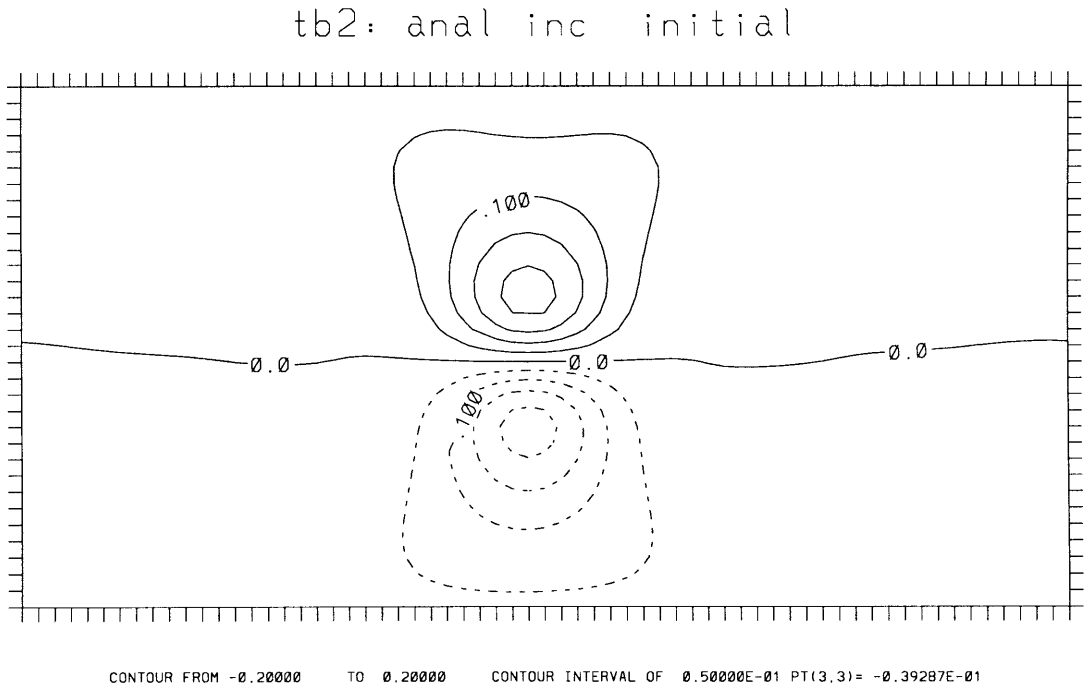



Figure 3.5, continued.

information in the vertical structure in the observations.

### 3.5.2 Sample analysis for many observations

When the 3DVAR has limited observations, it interpolates and extrapolates as described in Section 3.5.1 for one observation, but using all the information from the wind and temperature observations together. When the 3DVAR has many observations, although it needs to interpolate very little between them to reconstruct wind, temperature, and streamfunction fields, it must still infer the associated potential vorticity structure. To help visualize how the 3DVAR reconstructs a model state given wind and temperature observations, in this section we present an analysis for more than one observation taken to correct a realistic background error. The strengths and weaknesses of the 3DVAR can best be demonstrated when it is not constrained by lack of observational information. Therefore, the example we show is for errorless observations taken and assimilated at all (2112) model gridpoints at all model levels. The results shown here are typical, and they are not specific to the errorless observations or to the observation density.

To create a typical background error, we spun up "truth" and "model" states over a period of 60 days, using the experimental setup described in Section 4.1 and a random distribution of 16 fixed observations (approximately .75% on the normalized density curves shown below). The truth state at analysis time is the same as that depicted in Figure 2.2. The background error (the difference between a 12 hour forecast from the previous analysis, and the truth state) at analysis time is shown in Figures 3.6a–3.13a.

The analysis results are presented in Figures 3.6, 3.7, and 3.8 for streamfunction at the lower boundary, at the middle model level, and at the upper boundary, respectively. In these and later figures, panel a depicts the background error, panel b depicts the 3DVAR analysis increments, and panel c depicts the error in the resulting analysis (panel b added to panel a). Analysis results for wind and temperature fields look similar to those shown. Because the 3DVAR assumes some observational error, it is not expected to match the observations exactly. The analysis therefore misses some of the fine-scale features. Overall, however, the data assimilation has sufficient horizontal and vertical structure in the observations to be able to correct streamfunction, wind, and temperature background errors quite well.

The 3DVAR must infer the actual potential vorticity structure in the background error from the observations, the background and observation error covariances supplied to it, and the interpolator $\mathbf{L}$. Figures 3.9–3.13 show the interior potential vorticity and

boundary potential temperature analyses (corresponding to the streamfunction analyses in Figures 3.6–3.8). The potential vorticity background errors tend to occur at smaller spatial scales than the streamfunction errors, both horizontally and vertically, and have more anisotropic structure. Because the background error statistics have a broad, nearly isotropic structure (depicted in Figure 3.4d), and because the observations have more difficulty resolving smaller-scale potential vorticity structures, the 3DVAR has more difficulty correcting errors in terms of potential vorticity than it did in terms of wind, temperature, and streamfunction.

The 3DVAR has direct information about the boundary potential temperature from the temperature observations, so it reproduces the medium-large scale features in the boundary potential temperature reasonably well (Figures 3.9 and 3.13). The 3DVAR is also able to correct some potential vorticity errors at the lower and upper interior levels (Figures 3.10 and 3.12) due to the strongly negative vertical correlations with boundary potential temperature. The vertical correlations in $\mathbf{B}$, which are averaged over all horizontal scales and all situations as described in Section 3.4, are especially important for the potential vorticity analysis at lower levels. For potential vorticity at middle levels, the horizontal and vertical scales of the errors are small and anisotropic, and the 3DVAR is unable to correct most of the background error (Figure 3.11). Although the analysis increment at the middle level is the appropriate sign for correcting the large scale potential vorticity background errors, the potential vorticity analysis increments at middle levels are very small (the contour intervals in panels a–c of Figure 3.11 are the same).

One might expect the analysis to have difficulty resolving fine scale potential vorticity structures given wind and temperature observations. With this many perfect observations, however, we did not expect the data assimilation system to have so much difficulty. We have performed several experiments testing the 3DVAR with different numbers and sets of observations (including observations at one or several levels), with different data assimilation parameters (including changes in the horizontal and vertical background error covariances and the observational error covariances), and with different errors in the initial conditions to be corrected. It is clear that accurate smaller scale horizontal and vertical correlations are needed, but otherwise we have been unable to determine more precisely why this difficulty with potential vorticity analyses occurs. Nor can we say if this is a feature of our specific model, observational mix, and data assimilation system, or if this difficulty occurs more generally.

It has been suggested, based on adjoint calculations, that small scale potential vor-

ticity structures in the lower-middle troposphere can grow sufficiently rapidly that even if the analysis error has only a small projection onto the potential vorticity structure, the rapid error growth may degrade forecasts significantly (for more details on relevant adjoint techniques and singular vector calculations, see for example Farrell 1989, Molteni and Palmer 1993, and Palmer et al. 1998). Several of the adaptive observation strategies currently proposed (discussed in the Introduction) are based on identifying these features and observing them. The assumption behind these strategies is that if we observe in regions containing these features, any rapidly growing analysis errors which might exist can be corrected, and forecast failures can be minimized. However, because our 3DVAR is unable to correct small scale potential vorticity analysis errors, particularly in the lower-middle troposphere, observing in areas with small but potentially extremely rapidly growing small scale potential vorticity errors is likely to have little positive influence, at least in this idealized system. As we will demonstrate in Chapter 9, observing in regions where the data assimilation is not likely to produce the correct structure in the analysis increment can in fact lead to a significant risk of forecast degradation with extra observations.

As mentioned above, it is not clear how important correcting different types of potential vorticity errors is in real atmospheric prediction or how relevant this problem is to real data assimilation systems, even real 3DVAR systems. In addition, more sophisticated data assimilation systems are currently being developed and implemented to incorporate information on the model dynamics or to incorporate time- and space-dependent, anisotropic background error covariances with more knowledge of vertical structure. These data assimilation systems are likely to improve potential vorticity analyses, perhaps significantly. Even more sophisticated data assimilation systems, however, are still likely to have some difficulty correcting potential vorticity errors with small horizontal and vertical scales, both because traditional observations cannot resolve small potential vorticity structures well and because data assimilation systems will always have imperfect estimates of the time- and space-dependent error covariances. When adapting observation networks, then, it may be quite important to consider not only where we would like to improve the initial conditions, but how effectively the data assimilation system being used will (and will not) be able to correct certain types of errors in the initial conditions.

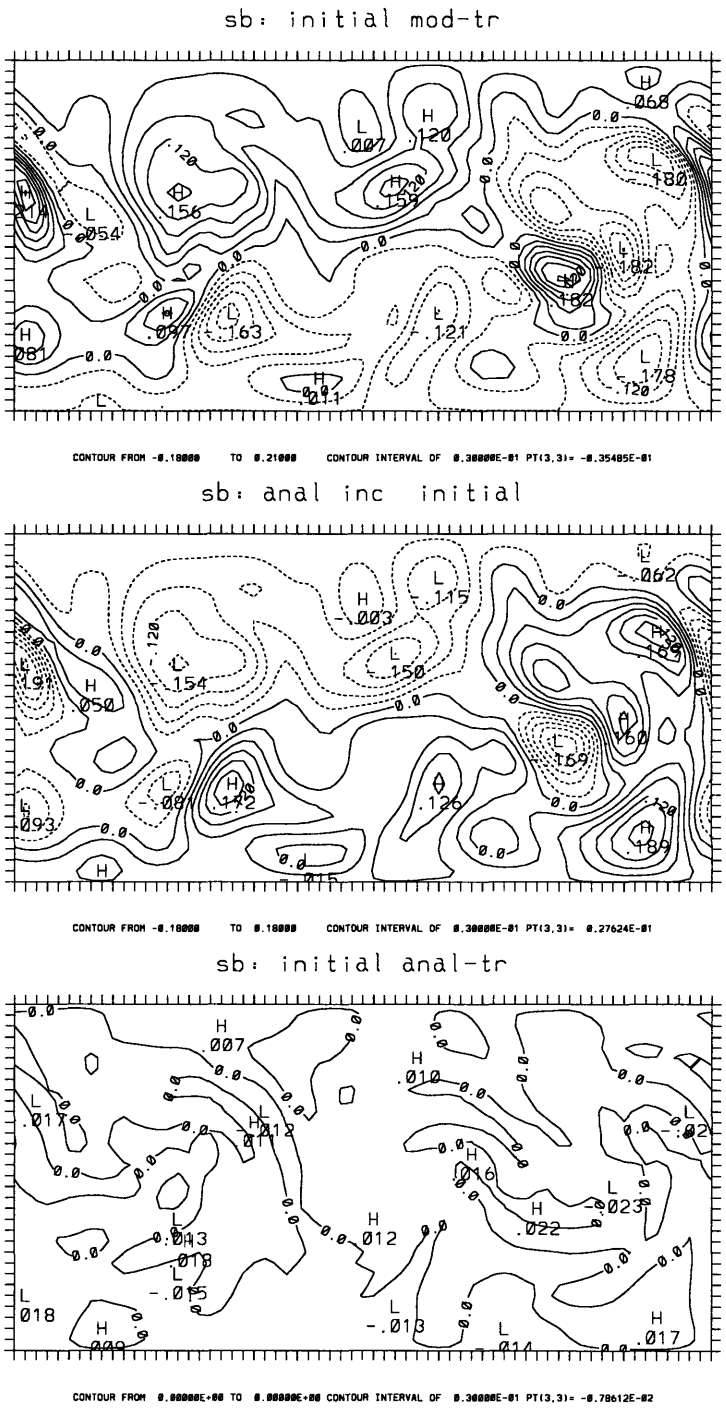



Figure 3.6: Sample 3DVAR analysis for errorless observations taken at all gridpoints at all levels, for a sample forecast error associated with the QG model truth state shown in Figures 2.2 and 2.3. Non-dimensionalized lower boundary streamfunction is shown for: (a) the 12 hour forecast error before data assimilation, (b) the analysis increment produced by the 3DVAR, and (c) the analysis error after data assimilation. The contour interval is .03 for all 3 plots.

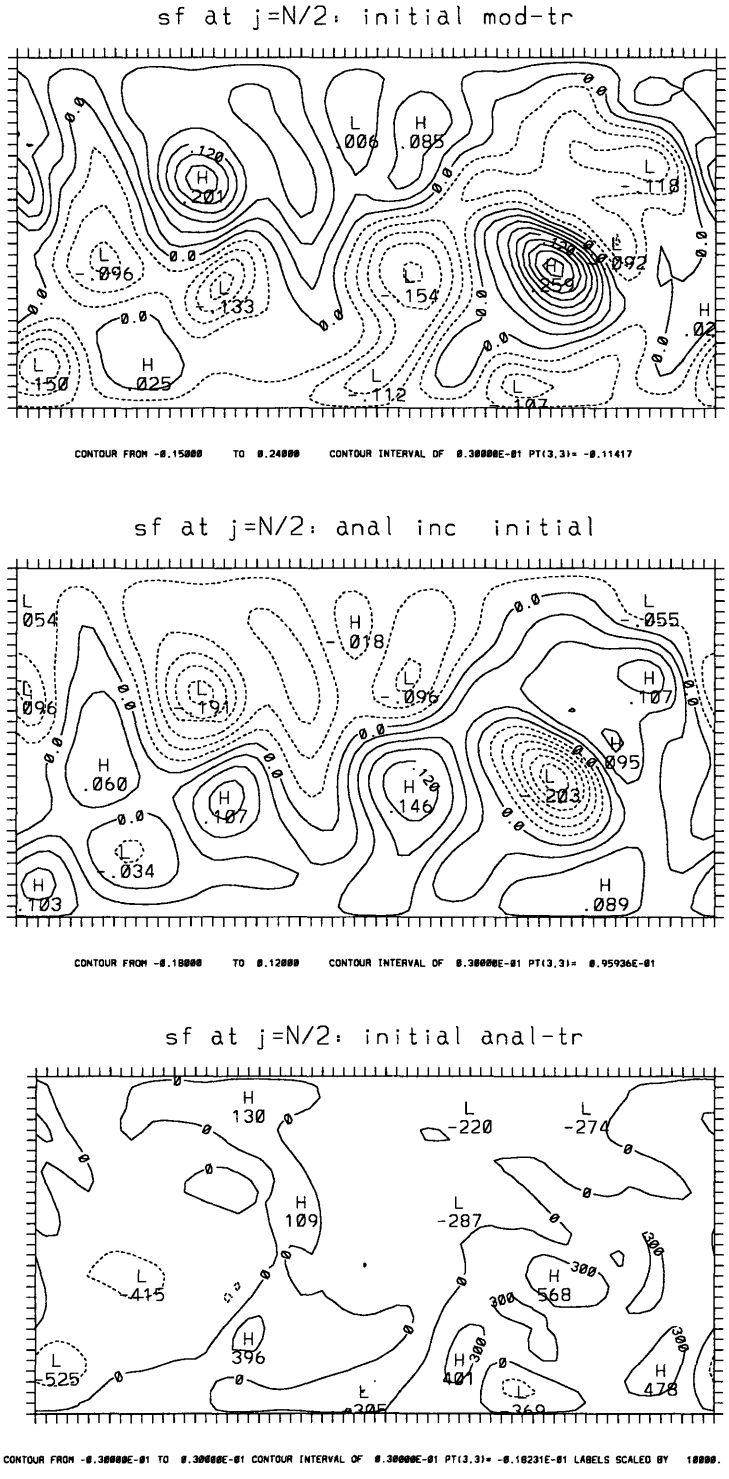



Figure 3.7: As in Figure 3.6, for streamfunction at the middle level (level 3). The contour interval is .03 for all 3 plots, but in (c) the contour labels are scaled by 10000.

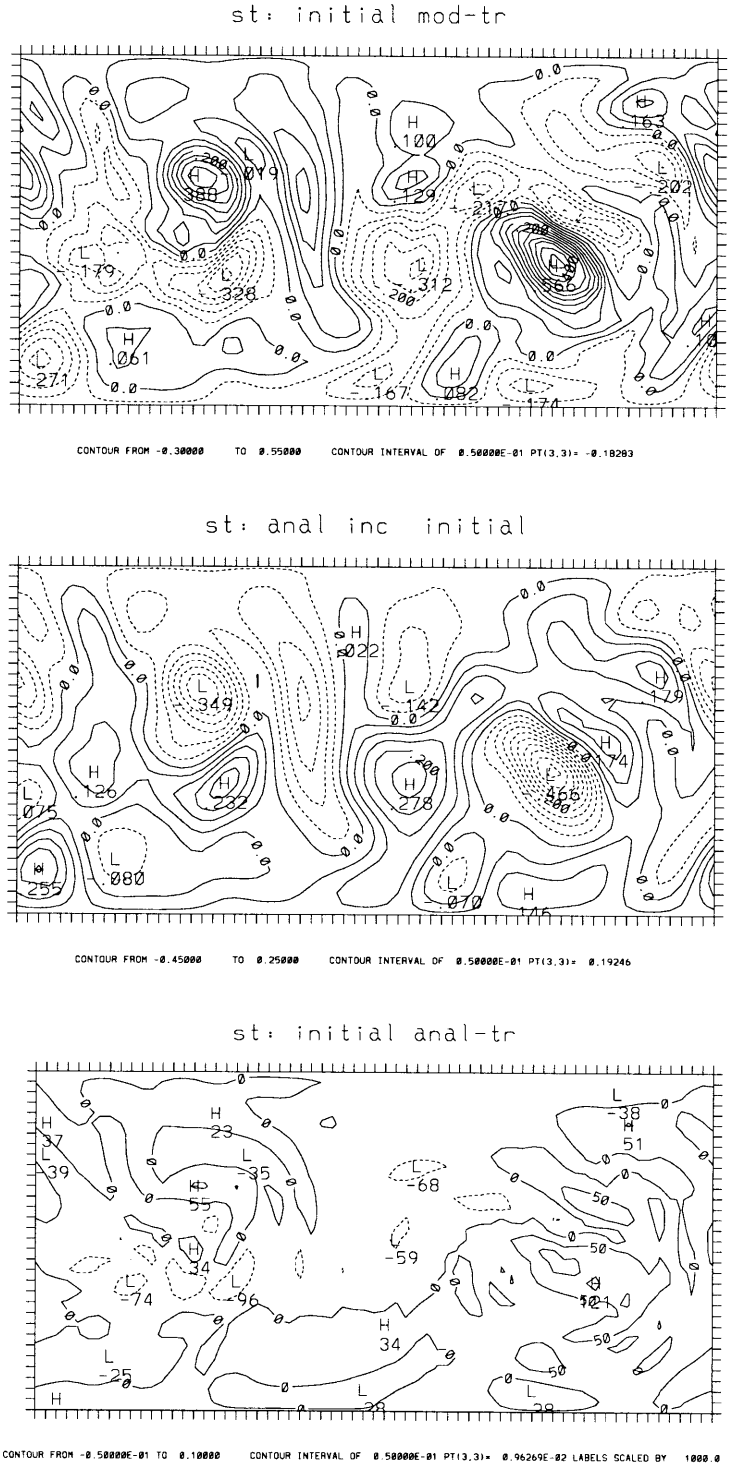



Figure 3.8: As in Figure 3.6, for upper boundary streamfunction. The contour interval is .05 for all 3 plots, but in (c) the contour labels are scaled by 1000.

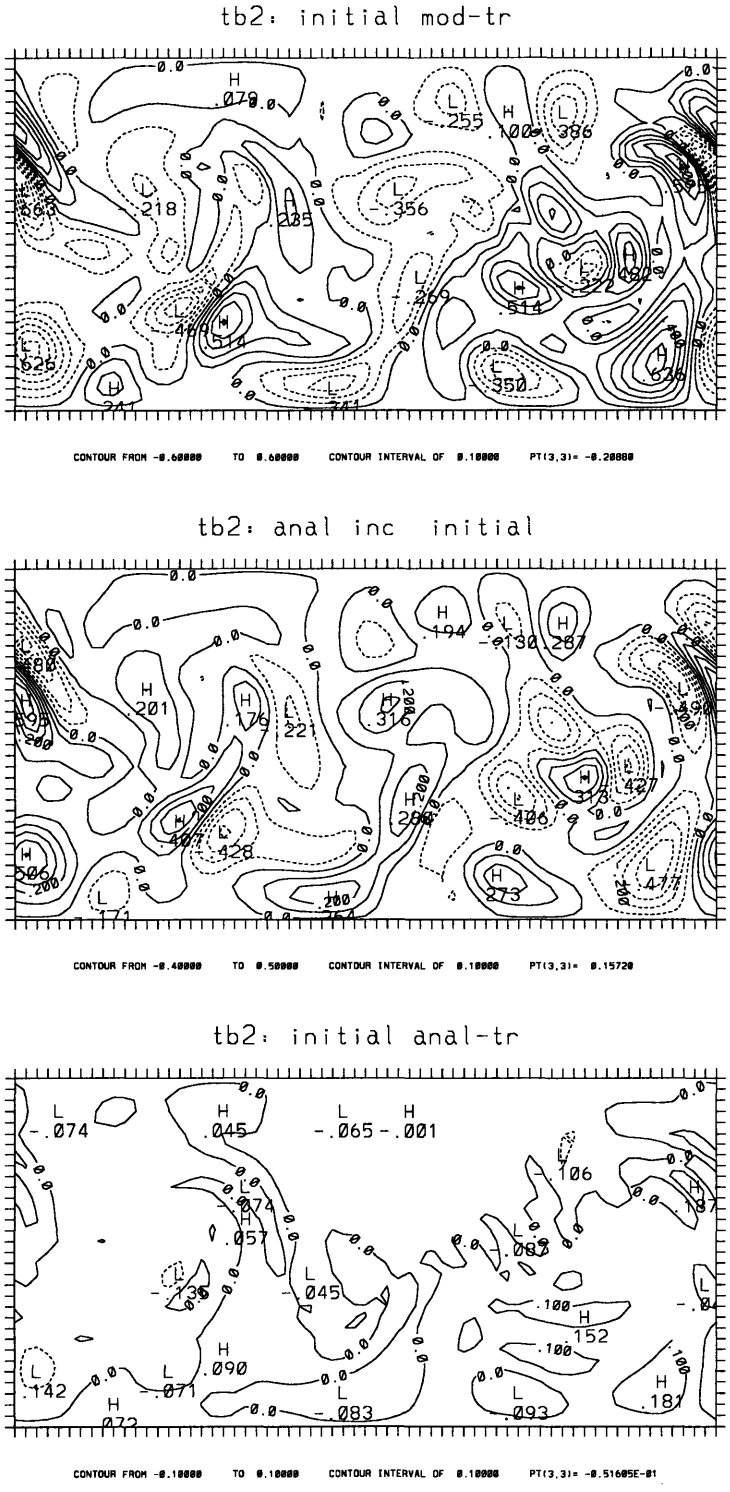



Figure 3.9: As in Figure 3.6, for potential temperature at the lower boundary. The contour interval is .1 for all 3 plots.

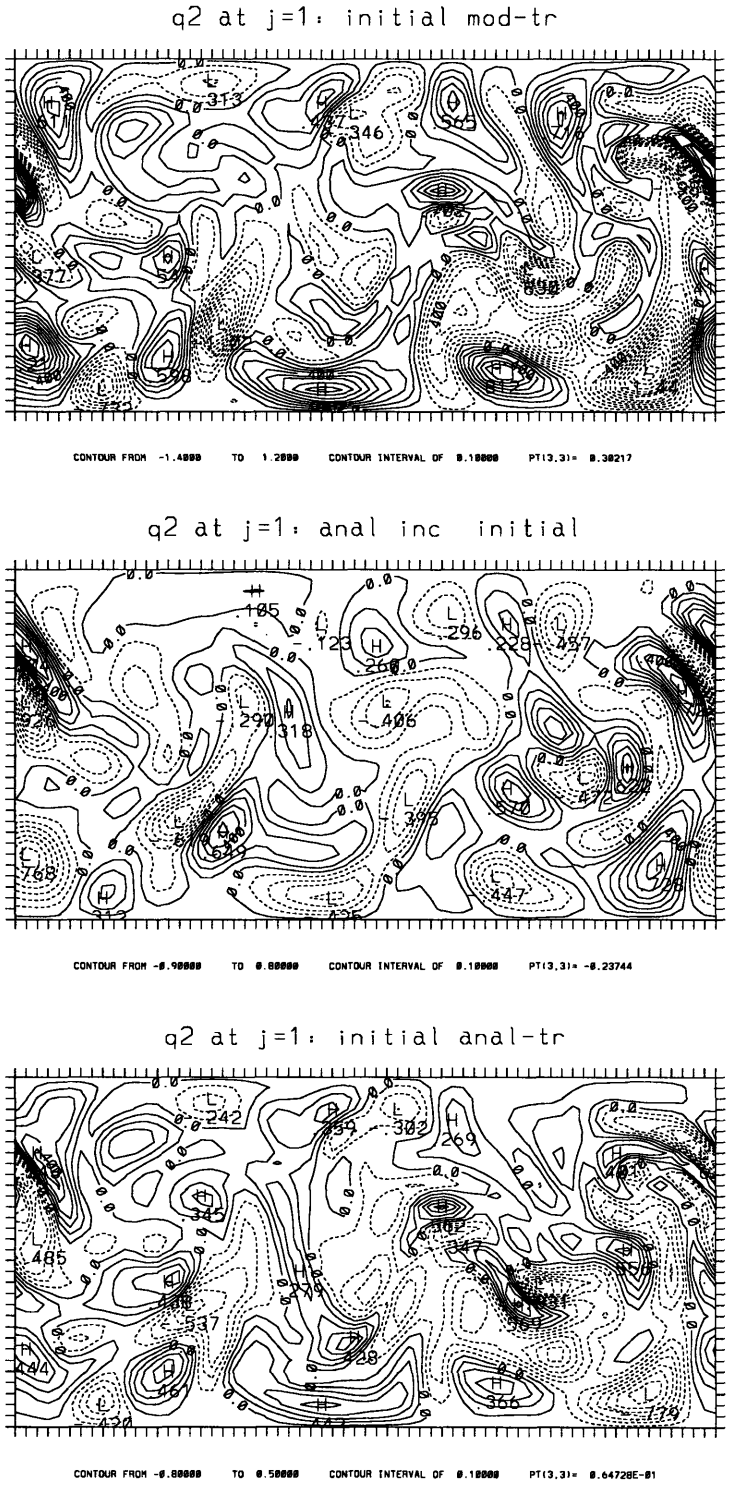



Figure 3.10: As in Figure 3.6, for potential vorticity at the lowest interior level. The contour interval is .1 for all 3 plots.

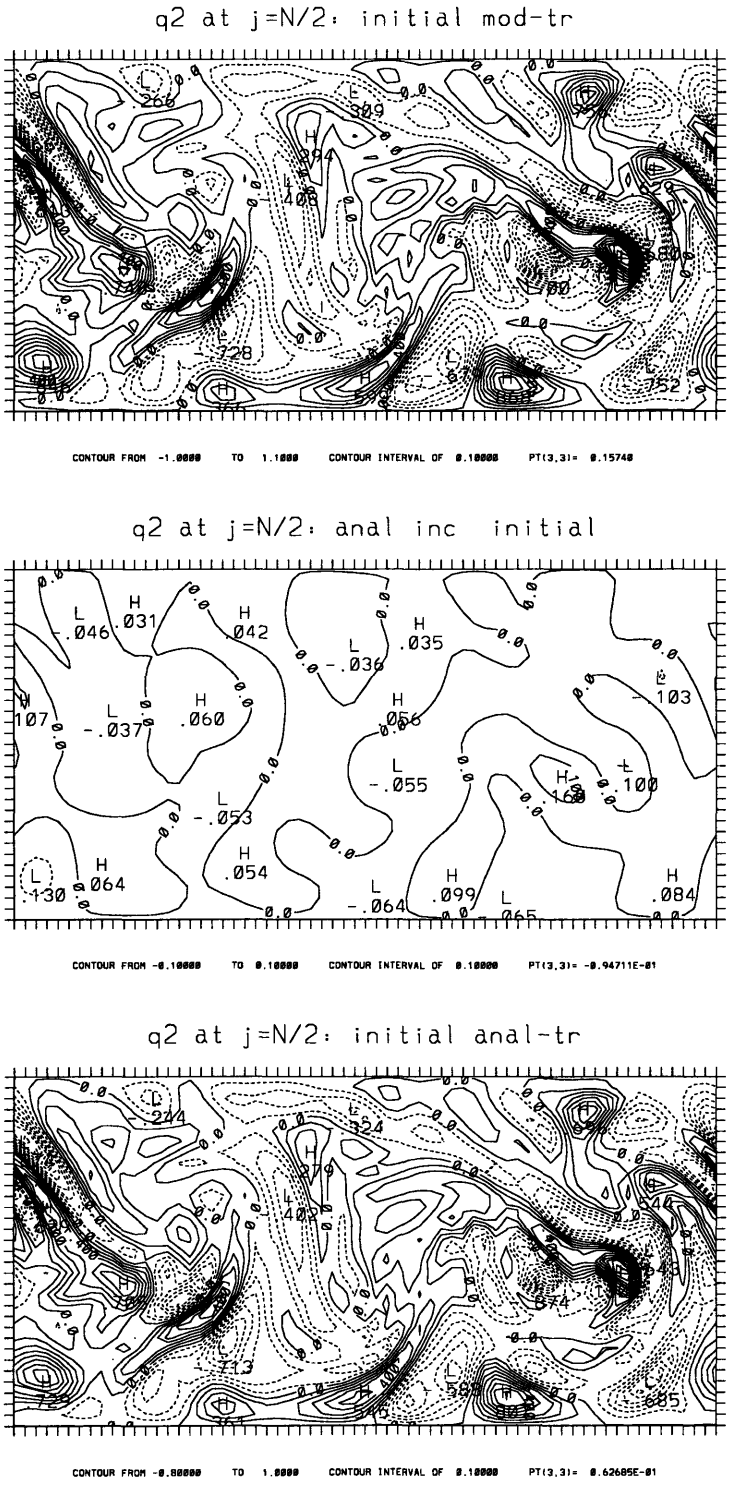



Figure 3.11: As in Figure 3.6, for potential vorticity at the middle level (level 3). The contour interval is .1 for all 3 plots.

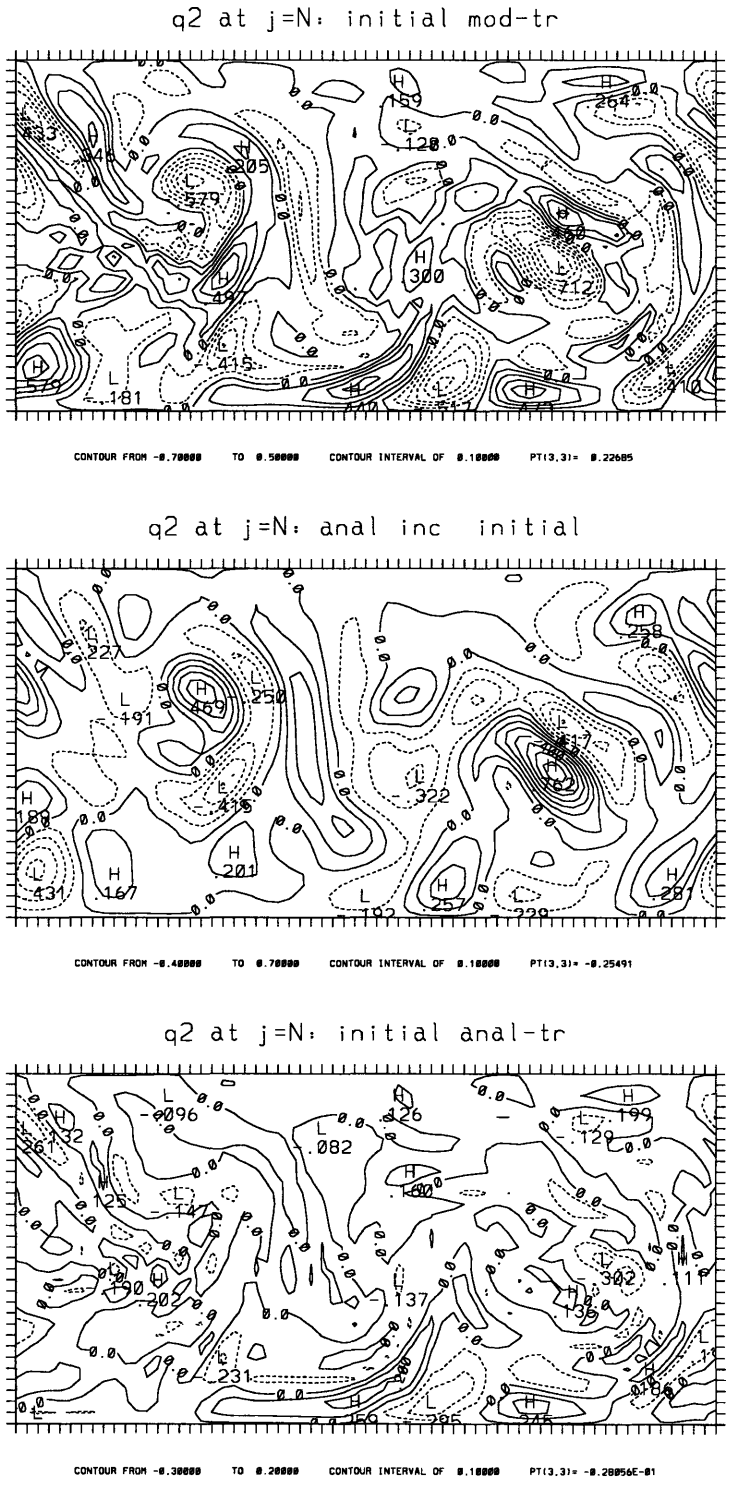



Figure 3.12: As in Figure 3.6, for potential vorticity at the upper interior level. The contour interval is .1 for all 3 plots.

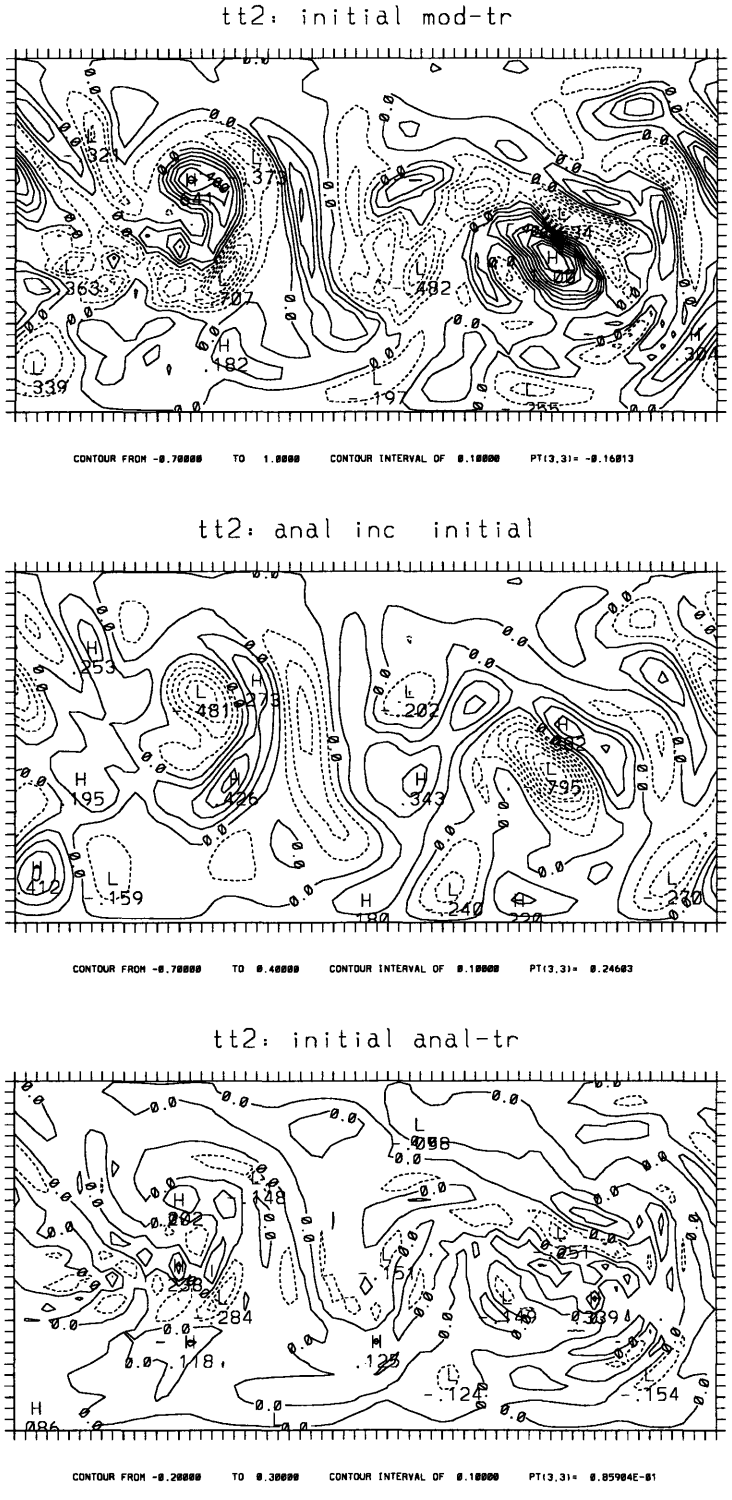



Figure 3.13: As in Figure 3.6, for upper boundary potential temperature. The contour interval is .1 for all 3 plots.

# Chapter 4

# Observing system simulation experiment design

## 4.1 General procedure

We set up the observing system simulation experiments (OSSE's) by first defining a "truth" solution as an arbitrary state of the QG model described in Section 2. An initial "model" (also referred to as "control") solution is produced by adding random noise to the initial truth state. Then:

1. The forecast (QG) model equations are used to integrate both the model and the truth states forward in time until the next data assimilation period.

2. Observation locations are selected (targeted) according to one of the strategies described in Section 4.2.

3. Simulated rawinsonde observations are constructed by sampling winds and temperature from the truth state at all model levels at the selected locations.

4. The observational data are assimilated into the model background field to form an analysis.

5. Using the forecast model, the analysis can be integrated forward in time to create forecasts, and ensembles of perturbed forecasts can be created.

6. The analysis becomes the new model state in Step 1.

This idealized setup mimics the analysis cycle in real atmospheric numerical prediction, with one major advantage: because the actual true state is known at all times, analyses and forecasts can be and are verified using this true state, instead of using limited information from observations or analyses.

Unless otherwise stated, observations are taken at all interior model levels and at the upper and lower boundaries. The observations were added at the boundaries to provide the 3DVAR with more information to partition the boundary potential temperature and the interior potential vorticity, although the boundary observations had only a small effect. Errors are added to the simulated observations both because real observation platforms are imperfect and because the data assimilation procedure cannot fit the observations exactly and thus must assume some observational error. The random observation errors are produced using a Gaussian distribution and the eigenvectors and eigenvalues of the observation error covariance matrix (which was defined in Section 3) as described in Houtekamer (1993), so that the observation error statistics are identical to the error correlations assumed in the data assimilation system. The error structure in the simulated radiosondes, shown in Tables 4.1 and 4.2, is similar to that described in Houtekamer (1993); for example, the first vector (leftmost in the tables) is nearly equivalent barotropic. To minimize confusion between errors in the observational data and the observation locations selected by each strategy, we truncated the Gaussian error distribution at one standard deviation, simulating a quality control algorithm. Although this truncation is not necessary, it prevents the results from being biased by one very "bad" observation in a key location. We have also run the experiments without errors in the observations (but still with observation errors assumed in the data assimilation scheme, since this is required for the data assimilation procedure to converge), and all comparisons are similar to those presented. An example of how the results differ with and without observation error is shown in Section 5.3.1.

In the results shown below, the same equations are used to integrate both the truth and model states, simulating perfect knowledge of the atmospheric dynamic equations. Errors in atmospheric models are likely to be important in real adaptive observations. By limiting forecast errors to result only from errors in the initial conditions, however, assuming a perfect model simplifies understanding the results. Once the basic interactions between observation networks, data assimilation systems, and forecast models are better understood in a perfect model context, adding model error will provide further information about how best to use observational resources in the real atmosphere.

The standard observation and data assimilation intervals are 12 hours; for each of

Wind Observation Error Matrix

| | *eig 1* | *eig 2* | *eig 3* | *eig 4* | *eig 5* | *eig 6* | *eig 7* |
|---|---|---|---|---|---|---|---|
| *lev 0* | 0.39 | -0.71 | -0.82 | -0.66 | -0.37 | -0.09 | 0.25 |
| *lev 1* | 0.61 | -1.10 | -1.08 | -0.61 | -0.13 | 0.01 | -0.25 |
| *lev 2* | 1.26 | -1.83 | -0.75 | 0.50 | 0.59 | 0.11 | 0.08 |
| *lev 3* | 1.96 | -1.68 | 0.70 | 0.63 | -0.51 | -0.20 | -0.03 |
| *lev 4* | 2.72 | -0.44 | 1.23 | -0.75 | 0.15 | 0.28 | 0.01 |
| *lev 5* | 3.08 | 1.21 | -0.21 | -0.15 | 0.29 | -0.47 | -0.01 |
| *lev H* | 2.71 | 1.47 | -0.79 | 0.47 | -0.30 | 0.35 | 0.00 |

Temperature Observation Error Matrix

| | *eig 1* | *eig 2* | *eig 3* | *eig 4* | *eig 5* | *eig 6* | *eig 7* |
|---|---|---|---|---|---|---|---|
| *lev 0* | 0.27 | -0.78 | -0.26 | -0.19 | 0.11 | 0.07 | 0.10 |
| *lev 1* | 0.28 | -0.71 | -0.14 | 0.02 | -0.10 | -0.08 | -0.15 |
| *lev 2* | 0.26 | -0.40 | 0.21 | 0.35 | -0.15 | -0.02 | 0.11 |
| *lev 3* | 0.33 | -0.20 | 0.44 | 0.15 | 0.21 | 0.10 | -0.07 |
| *lev 4* | 0.61 | -0.00 | 0.48 | -0.29 | -0.03 | -0.12 | 0.03 |
| *lev 5* | 0.95 | 0.28 | -0.08 | -0.07 | -0.15 | 0.17 | -0.03 |
| *lev H* | 0.96 | 0.33 | -0.32 | 0.16 | 0.13 | -0.12 | 0.01 |

Table 4.1: Matrices used to generate observation errors for simulated rawinsondes in the QG model at the standard resolution, for dimensionalized (a) wind and (b) temperature measurements. The columns are the eigenvectors of $\mathbf{O}$ multiplied by the square root of the associated eigenvalue, and the rows are the elements of the resulting vector at each level. To generate the errors for each radiosonde, random numbers are generated with a Gaussian distribution, each column of the matrix is multiplied by a different random number, and the columns are summed to produce errors for each level. The errors used for zonal wind are .8 × total wind errors, and the errors for meridional wind are .5 × total wind errors. To recover non-dimensional values, divide the wind observation error matrix by $U$ ( $= 60\ \mathrm{ms}^{-1}$) and divide the temperature observation error matrix by $\theta_o N^2 H R_o / g$ ( $= 10.16\ ^{\circ}\mathrm{K}$).

Wind Observation Error Matrix

| | *eig 1* | *eig 2* | *eig 3* | *eig 4* | *eig 5* | *eig 6* | *eig 7* | *eig 8* | *eig 9* | *eig 10* |
|---|---|---|---|---|---|---|---|---|---|---|
| *lev 0* | 0.42 | -0.67 | -0.76 | 0.66 | 0.45 | 0.27 | 0.15 | 0.09 | 0.04 | 0.07 |
| *lev 1* | 0.57 | -0.91 | -0.95 | 0.71 | 0.36 | 0.11 | -0.01 | -0.05 | -0.04 | -0.09 |
| *lev 2* | 0.96 | -1.43 | -1.11 | 0.36 | -0.21 | -0.33 | -0.22 | -0.08 | -0.02 | 0.05 |
| *lev 3* | 1.42 | -1.78 | -0.69 | -0.43 | -0.54 | -0.10 | 0.16 | 0.14 | 0.05 | -0.03 |
| *lev 4* | 1.88 | -1.70 | 0.21 | -0.80 | 0.02 | 0.36 | 0.05 | -0.13 | -0.07 | 0.02 |
| *lev 5* | 2.35 | -1.17 | 1.04 | -0.26 | 0.54 | -0.06 | -0.21 | 0.07 | 0.09 | -0.01 |
| *lev 6* | 2.77 | -0.22 | 1.18 | 0.62 | 0.04 | -0.32 | 0.20 | 0.00 | -0.09 | 0.01 |
| *lev 7* | 3.02 | 0.89 | 0.39 | 0.68 | -0.55 | 0.27 | -0.05 | -0.07 | 0.09 | -0.01 |
| *lev 8* | 2.82 | 1.59 | -0.68 | -0.21 | -0.03 | 0.13 | -0.15 | 0.15 | -0.11 | 0.00 |
| *lev H* | 2.52 | 1.63 | -0.98 | -0.59 | 0.36 | -0.22 | 0.14 | -0.10 | 0.06 | -0.00 |

Temperature Observation Error Matrix

| | *eig 1* | *eig 2* | *eig 3* | *eig 4* | *eig 5* | *eig 6* | *eig 7* | *eig 8* | *eig 9* | *eig 10* |
|---|---|---|---|---|---|---|---|---|---|---|
| *lev 0* | 0.29 | -0.73 | -0.33 | 0.23 | 0.14 | 0.09 | 0.06 | 0.04 | 0.03 | 0.03 |
| *lev 1* | 0.30 | -0.72 | -0.27 | 0.12 | 0.01 | -0.04 | -0.06 | -0.05 | -0.05 | -0.05 |
| *lev 2* | 0.29 | -0.57 | -0.04 | -0.18 | -0.21 | -0.14 | -0.06 | 0.01 | 0.03 | 0.04 |
| *lev 3* | 0.30 | -0.41 | 0.18 | -0.32 | -0.14 | 0.05 | 0.10 | 0.04 | -0.00 | -0.04 |
| *lev 4* | 0.34 | -0.28 | 0.35 | -0.25 | 0.11 | 0.14 | -0.02 | -0.06 | -0.03 | 0.04 |
| *lev 5* | 0.45 | -0.15 | 0.46 | -0.01 | 0.22 | -0.05 | -0.09 | 0.04 | 0.04 | -0.03 |
| *lev 6* | 0.64 | 0.01 | 0.43 | 0.26 | 0.01 | -0.13 | 0.09 | 0.00 | -0.05 | 0.02 |
| *lev 7* | 0.83 | 0.21 | 0.14 | 0.27 | -0.22 | 0.12 | -0.02 | -0.04 | 0.04 | -0.01 |
| *lev 8* | 0.93 | 0.37 | -0.26 | -0.07 | -0.02 | 0.06 | -0.07 | 0.08 | -0.05 | 0.01 |
| *lev H* | 0.90 | 0.39 | -0.38 | -0.22 | 0.14 | -0.09 | 0.06 | -0.05 | 0.03 | -0.01 |

Table 4.2: As in Table 4.1, but for double resolution. Scaling is the same.

the standard resolution experiments with a 12 hour or shorter interval, the experiment was run for 90 days to equilibrate the model state to the specific observational network, followed by a 90 day run to gather error statistics. For the 1 day and longer data assimilation interval experiments, the spin up and run times were each 180 days. All double resolution runs have spin up and run times of 60 days each. Because the QG model is highly variable, each experiment was performed 3 times, with different initial truth states, different initial model state perturbations, a different set of fixed or random observation locations (for the fixed and random strategies), and a different set of random observational errors. The spread of results from the three runs suggests the possible spread in the errors produced by each observation strategy and density. Since we have only tested a small fraction of possible QG model states and observation configurations, however, and since the results are slightly sensitive to the error norm selected and the data assimilation parameters, experiments with error distributions which cannot clearly be separated are considered indistinguishable within the limitations of this study.

## 4.2 Definitions of observing strategies

In order to compare different types of observation networks, we define three basic strategies for allocating observations: fixed, random, and adaptive. For the non-adaptive networks, instead of attempting to simulate a "standard" real world observation network, we have tested only idealized fixed and random strategies. We chose not to simulate a complex inhomogeneous mix of observation platforms, each with different horizontal and vertical resolutions, for several reasons. First, the numbers and types of observations used vary in time and between different operational weather centers. More importantly, testing less complicated observation networks simplifies attributing the causes of the results. As we will see in later chapters, it is already sufficiently challenging to separate the influence of different observation strategies from the effects of the forecasting system; adding other factors early in the study would only further blur our understanding.

A major purpose of adaptive observations is to obtain data in regions where there is insufficient data. Early experiments with adapting observations in pre-defined data voids, however, convinced us that before testing a mix of non-adaptive and adaptive observations, we needed to understand more generally how the different observation networks interacted with the data assimilation system and the forecast model. Therefore, we decided not to explicitly simulate data dense and data sparse regions. This is also

advantageous because as more satellite, aircraft, and other oceanic observations are assimilated into forecast models, it is no longer clear how "data sparse" non-populated regions are. We have tested a variety of different pre-defined observation spacings, and they produce general results similar to those shown, but otherwise we only peripherally address this issue of explicitly inhomogeneous observation networks.

Although there are many ways to adapt observations, only two simplified sample adaptive strategies — one idealized and the other an approximation to the idealized strategy — are tested in this study. Unless otherwise stated, in the results shown observations are allocated globally, all according to the same strategy, and they are allocated at each data assimilation time just prior to taking the observations. The lack of time lag between selecting adaptive observation locations and observing at them is unrealistic in most scenarios; this issue, however, is addressed only briefly, in Section 7.4.

### 4.2.1 Fixed observations

For fixed observations, observation locations are selected randomly prior to the experiment, then left fixed throughout the experiment. This simulates a fixed, inhomogeneously distributed rawinsonde network. Because the adaptive strategy tends to choose observation locations near the center of the channel, in the jet, the fixed locations are selected randomly with a Gaussian distribution weighting them away from the channel walls. All strategies then have approximately the same overall observation distribution, and only the order in which the locations are selected within that distribution is varied. Except at high observation densities, observation locations are constrained to not be at adjacent gridpoints (i.e. to be at least 350 km apart). The observations are not equally spaced, so there are always some regions with fewer observations than others; there are, however, no systematic data voids. We have tested the sensitivity of the results to the specific choices of spacing and distribution, and it is small. Three sample sets of observation locations for 32 fixed observations (approximately 1.5% of gridpoints observed on the normalized curves below) are shown in Figure 4.1.

### 4.2.2 Random observations

For random observations, observation locations are selected randomly at each targeting time. This strategy is tested in order to differentiate the potential benefits of targeting observations consciously from the potential benefits of simply moving obser-

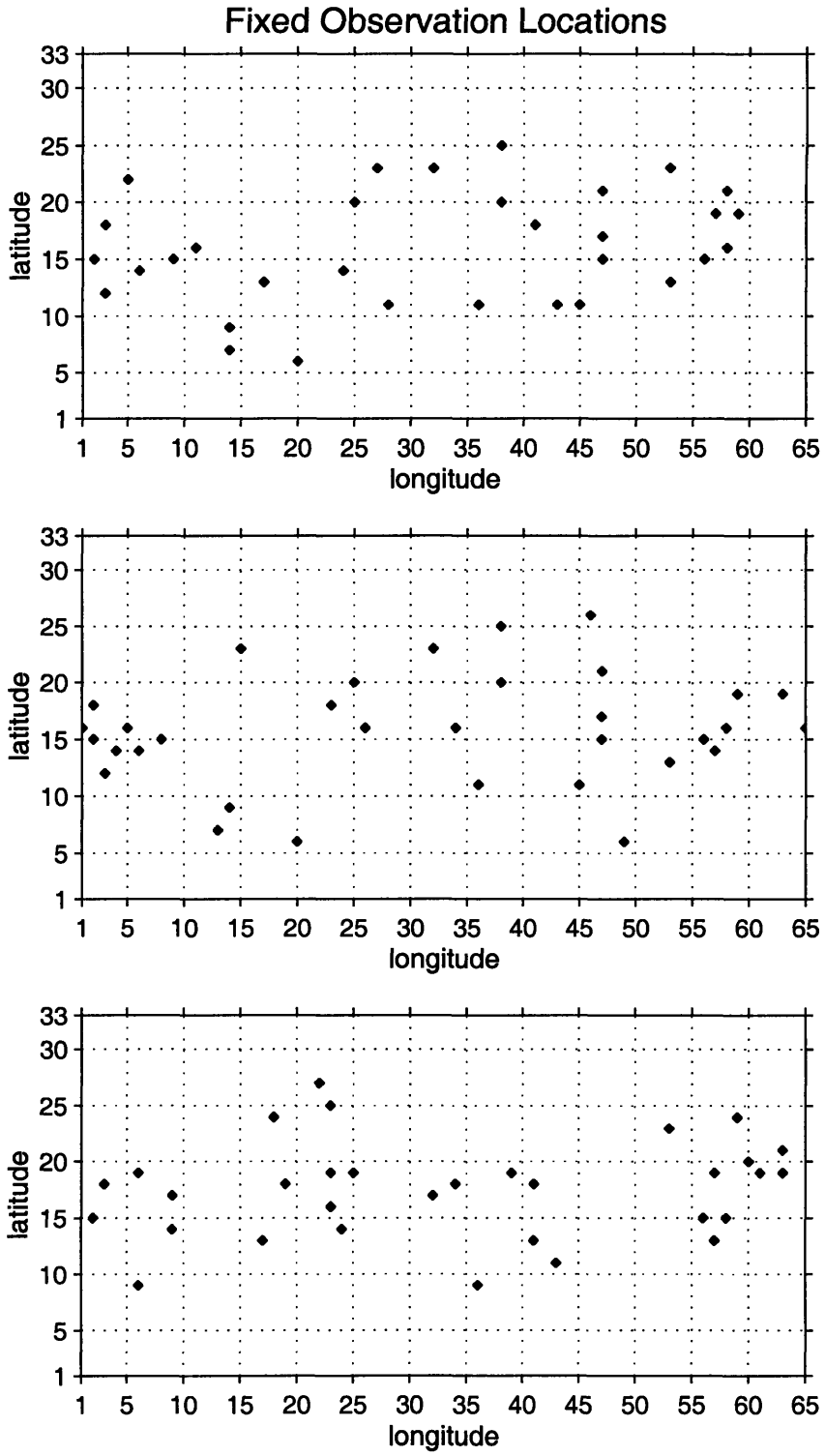



Figure 4.1: Three sample distributions of 32 fixed observation locations. The quasi-geostrophic model is periodic in longitude, so location 1 and location 65 on the x-axis are equivalent. The grid spacing is 250 km.

vation locations. Any effective adaptive strategy should therefore, at minimum, reduce errors more than a random, "null case" targeting strategy. As for the fixed strategy, random observation locations are weighted towards the center of the channel and are constrained to not be at adjacent gridpoints, and sensitivity to the specifics of these choices is small. Figure 4.2 shows the distribution of all observation locations selected during a 180 day period for 32 random observations at each data assimilation time, along with the set of observation locations selected at an individual time.

### 4.2.3 Adaptive observations

Adaptive observations are defined as observation locations selected "intelligently" according to some strategy at each targeting time. The first sample strategy tested is idealized, based on the actual (known) error in the initial conditions. Although this strategy cannot be implemented in the real atmosphere, we can use it to implement adaptive observations in a large number of ways knowing that the strategy is not limited by imperfect knowledge of the true atmospheric state. The second strategy, testing a semi-realistic method of approximating the idealized strategy, estimates errors in the initial conditions from an ensemble of perturbed forecasts.

The only constraint placed on the observation locations selected in the two adaptive strategies is that they not be at adjacent gridpoints. Figure 4.3 shows the distribution of all observation locations selected during a 180 day period for 32 ideal adaptive observations at each data assimilation time, along with the set of observation locations selected at an individual time. As described above, these adaptive strategies often select observation locations near the middle latitudes, in the vicinity of the jet. Note also how, compared to random observations, the adaptive observations are less evenly distributed at each individual time; they tend to cluster to some extent. The sensitivity of the results to constraining the spacing of the adaptive observations has been tested and is small. Observation clusters are important in some circumstances, however, and they are discussed further in Section 6.2 and later chapters.

#### Choice of adaptive strategies

Other criteria besides errors in the initial conditions, such as forecast sensitivity, are important when adapting observations. For several reasons, however, we have chosen to limit the strategies tested to those based on reducing the analysis error. Both of the adaptive strategies tested in this study are similar to the analysis-error based strategies

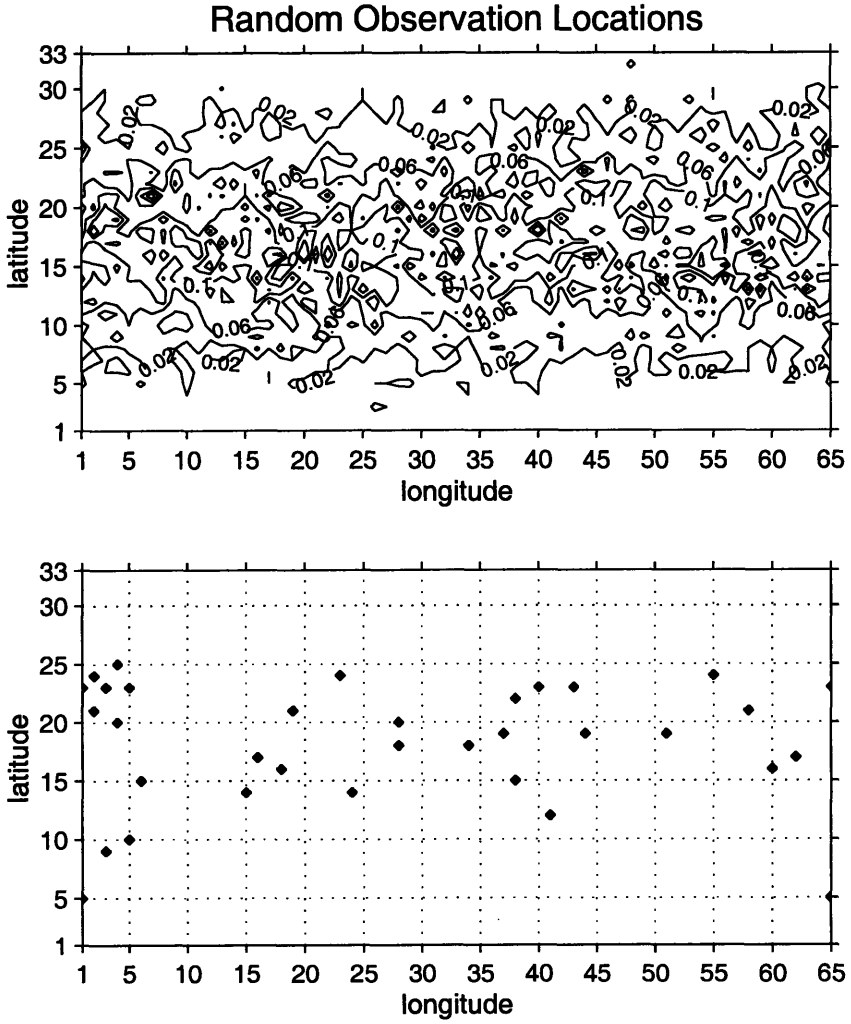



Figure 4.2: (a) Frequency that each location was selected for an observation, as a percentage of the total number of observation locations selected, for 32 random observations selected at each targeting time. (b) The 32 observation locations selected by the random strategy at a sample individual targeting time.

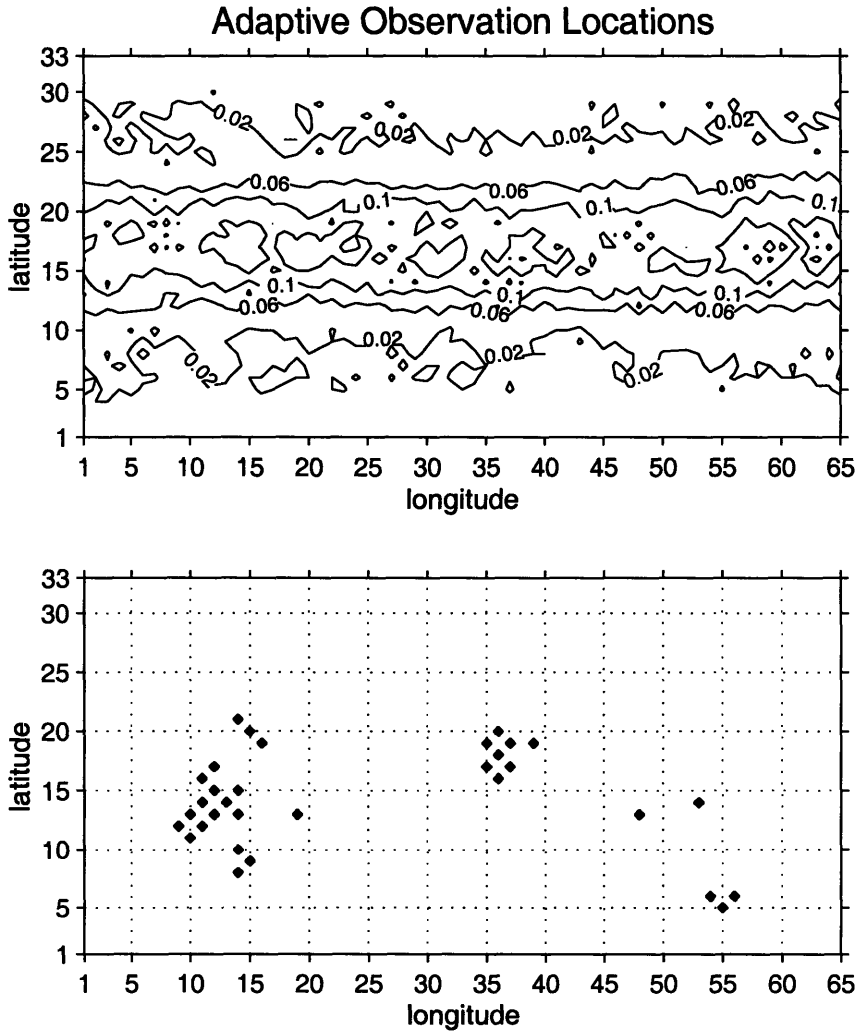



Figure 4.3: (a) Frequency that each location was selected for an observation, as a percentage of the total number of observation locations selected, for 32 ideal analysis error reduction adaptive observations selected at each targeting time. (b) The 32 observation locations selected by the adaptive strategy at a sample individual targeting time.

tested in Lorenz and Emanuel (1998), and much of our reasoning follows theirs. First, the main purpose of this study is to explore how and when analysis and forecast errors can be reduced when *an* intelligent adaptive strategy is used, not to compare different adaptive strategies. Error in the initial conditions is one of the likely important criteria for targeting observations, and since the actual analysis error is known at any time, basing a strategy on it is both very simple and computationally quick.

In addition, when testing strategies which might select observation locations in regions with smaller analysis errors or in one part of a region with large analysis errors, we could easily confuse the limitations of the data assimilation system with the observing network results. Before testing these more complex strategies, it is therefore important both to know that the data assimilation system can use some type of adaptive observations to reduce analysis errors where they are reasonably large and to understand more about how the data assimilation interacts with the observations in the vicinity of large analysis errors. Finally, by estimating the theoretical upper bound of error reduction possible using strategies based on analysis error reduction, we provide results with which to compare more complex and realistic strategies. This estimate is used, for example, to evaluate the effectiveness of different sized ensembles for estimating errors in the initial conditions for use in adapting observations.

### Ideal analysis error reduction

The ideal analysis error reduction (ideal AER) strategy selects observation locations, prior to each observation time, where the background and truth states differ the most. This strategy assumes that in the absence of additional observations, the analysis error will be the largest where the background error is the largest, and that therefore by observing in these regions we will, on average, minimize the analysis error. This reduced analysis error should then result in reduced average forecast error. Since no assessment is made of how important any specific part of the initial conditions is for a future forecast or of how the data assimilation system will incorporate the observations, this strategy is not likely to select the locations at which observations will reduce forecast error the most. The strategy is thus not perfect, but it is "ideal" in the sense of being possible only in idealized studies such as this one.

### Estimated analysis error reduction using an ensemble

The actual background error is not known in a realistic setting. Therefore, we also

test an estimated analysis error reduction (estimated AER) strategy, in which observation locations are selected where the error in the background field is estimated, based on the spread of an ensemble, to be a maximum. To produce the ensemble for the results shown below, we first take the "control" observations (which already include errors) and add sets of positive and negative error perturbations which were generated using the same observation error statistics. We then assimilate a different set of these perturbed observations separately into each ensemble member at each analysis time, at the same observation locations as for the control data assimilation. These perturbed ensemble initial conditions are integrated forward (using the QG model equations) until the next analysis time, when regular observations are assimilated into the control model state and perturbed observations are again assimilated into the individual ensemble members. This simulates the effect of random sampling errors on analyses and forecasts.

The ensemble implemented here is similar to the multiple replication ensemble tested in Lorenz and Emanuel (1998), to the OSSE-MC procedure described in Houtekamer and Derome (1995), and to the system simulation ensemble (Houtekamer et al. 1996) currently operational at the Atmospheric and Environment Service in Canada with an unperturbed forecast model. Assimilating the data into each ensemble member separately is computationally expensive. Nevertheless, we chose this type of ensemble because, based on the ensemble formulation and on the results in Lorenz and Emanuel (1998), we expect it to estimate background error as well as or better then the other ensemble forecasting systems currently operational, such as the breeding cycle ensemble (Toth and Kalnay 1993, Toth and Kalnay 1997) operational at the National Centers for Environmental Prediction in the U.S. (NCEP, formerly NMC) and the singular-vector perturbation ensemble (Mureau et al. 1993, Molteni et al. 1996) operational at the European Centre for Medium-Range Weather Forecasts (ECMWF).

Ensemble forecasting techniques are a subject of active research and debate, and comparing them is outside the scope of this study. In Chapter 7, however, along with testing the estimated AER adaptive strategy, we do evaluate how well the spread of our ensemble estimates background error. To help us provide better ensemble-based estimates of background error in the future, we also discuss what types of background error information might be the most useful for adaptive observations. In addition, we explore the sensitivity of the estimate of background error and of the AER adaptive strategy results to both the ensemble size and the targeting lead time.

# Chapter 5

# Results with global fixed observations

## 5.1 Average errors for different observation densities

It is important that we understand how our simulated observing system adjusts to changes in a typical type of observation network before moving on to testing adaptive observations. Figure 5.1 shows, for a fixed observation network and the standard 12 hour data assimilation interval, how the time- and domain-averaged analysis error decreases as the number of observations increases; note that both axes are logarithmic. In Figure 5.1 and in subsequent figures, the observation density is normalized by the maximum number of observation locations (2112) and the analysis error is normalized by the error in the absence of observations, i.e. the saturation error. Three of the symbols plotted for each observation density are the root-mean-square average analysis errors for each of the three separate runs described in Section 4.1 and indicate of the range of error associated with each observation density; the fourth symbol and the connecting line are the root-mean-square average of the three runs. As described in Section 4.2, the fixed observation networks contain unevenly spaced observations, but they do not explicitly simulate a real world observation network. Three sample fixed observation networks (the three distributions of observation locations used for the experiments with 32 fixed observations, i.e. 1.5% of gridpoints observed) are depicted in Figure 4.1.

This simulated system has three general regimes in observation density. On the left of Figure 5.1, for few observations, the data assimilation system does not have

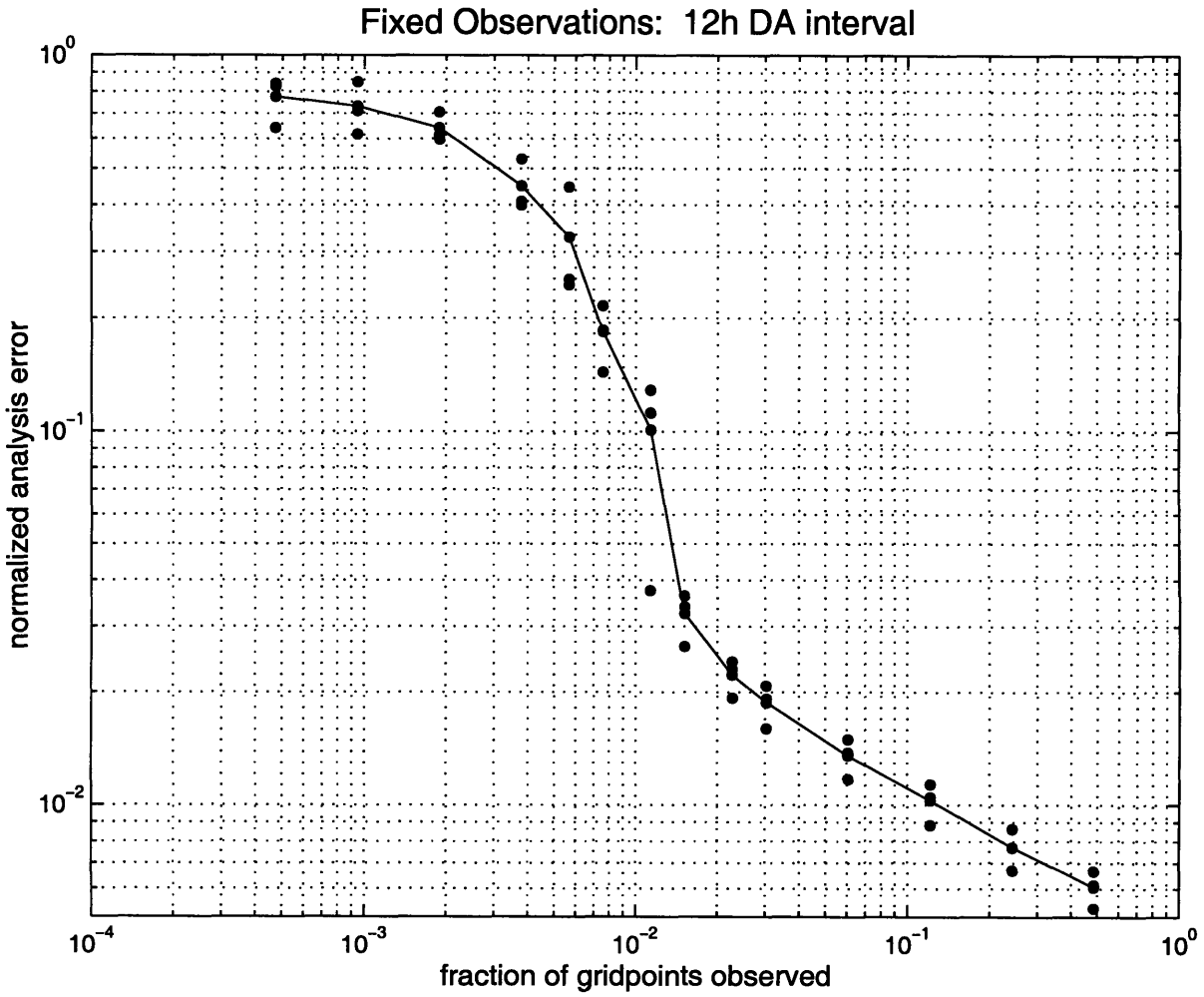



Figure 5.1: Domain- and time- root-mean-square averaged analysis error (streamfunction norm) as a function of the density of fixed observations, for a 12 hour data assimilation interval. The symbols represent the average errors for three different runs and the root-mean-square average of the three, as described in Section 4.1 of the text. For each of the runs, the y-axis is normalized by the average error in the absence of observations, i.e. the saturation error, and the x-axis is normalized by the maximum number of observations (the number of gridpoints in the x-y plane, 2112). 1% of gridpoints observed corresponds to an average of approximately one observation every $(2000\ \mathrm{km})^2$; the Rossby radius of deformation $R_d \approx 1000$ km.

enough information to resolve even the large scale errors in the initial conditions. It can make only small local improvements which are quickly swamped by the error growth, resulting in essentially saturated errors in the model state and only a small benefit accumulated from the observations. At the other extreme, for many observations, the data assimilation system has extracted most of the information that it can from the observations. If we take into account the error in the observations, the assumptions in the data assimilation system, and the relatively small error growth at small scales in the system, the analysis errors are already small. Adding observations to dense observations therefore also produces only a small additional benefit, primarily through error-checking with other observations. This dense observation limit was also observed in observing system simulation experiments performed several decades ago (Bengtsson and Gustavsson 1971, Bengtsson and Gustavsson 1972, Morel et al. 1970). In between our sparse and dense observation limits, there is a middle regime in which, as observations are added, the accumulated error drops off rapidly. Although the specific values change with the experiment, in this example increasing the number of observations from 12 to 32 (from .6% to 1.5% of gridpoints) decreases the analysis errors by an order of magnitude.

One might well ask if the distinctive shape of Figure 5.1 is a consequence of the experimental design. Although we do not discuss the sensitivity of Figure 5.1 to the experimental design in great detail, some sensitivity results are presented in Sections 5.3 and 5.4. For example, we have tested changing details of the experimental setup such as the data assimilation convergence cutoff, the observation error, and the observation spacing, and these changes generally only alter the minimum achievable error and the slope in the data dense regime. Larger changes in the data assimilation system and in other experimental parameters can alter the observation density at which each of the three regimes occurs and the extent and slope of each regime, but generally not the presence of the three regimes. Thus, even though the specific appearance of Figure 5.1 is sensitive to differences in our 3DVAR and the experimental setup, the general shape is not. In this forecast model and data assimilation system, the middle regime can always be differentiated from the sparse and dense observation limits by the large difference in influence from additional observations.

With a 12 hour data assimilation interval, the error decreases rapidly in the range of an average of one observation every $(2000\,\mathrm{km})^2$ in the more dynamically active areas of the domain (this value is approximate because the observations are unevenly spaced). Synoptic scales dominate the error growth in the QG model, and with an average of on the order of one observation per Rossby radius of deformation (1000 km in this

model), the data assimilation procedure receives information which is sufficiently dense in space and time to resolve effectively the dominant features. As we discuss in the following section, the error reduction is dominated by observations at synoptic scales not just for streamfunction errors, but for all other error norms tested. For example, even though atmospheric potential vorticity features and errors tend to have a much smaller spatial scale than streamfunction features and errors, the errors in potential vorticity are, on average, also primarily reduced by observations at synoptic scales. This is likely related both to the difficulty that the 3DVAR data assimilation has in resolving small scale potential vorticity errors (discussed in Section 3.5.2) and to the dynamics in the QG model. Unfortunately, from the experiments performed in this study we cannot be certain why this occurs and to what extent it is due to the simplified dynamics in the model and the idealized experimental setup; a more detailed set of experiments is needed.

In the real atmosphere, this dominance of error reduction at synoptic scales might be limited by errors in the forecast model, such as unresolved sub-grid scale processes. We are also likely, in some situations, to be specifically interested in sub-synoptic scale errors. Further study is required to understand how this result may or may not be extended to more complex models and more sophisticated data assimilation methods. In Section 5.3.3, however, we discuss what might happen with a much more sophisticated data assimilation system, and we hypothesize that the dense observation limit will still occur. Therefore, we believe that it is likely that for a given forecast model, data assimilation method, and forecast situation, there is always an observation density, achievable or not, above which additional observations can reduce errors only a little.

Given Figure 5.1, one might also ask which observation density regime real world synoptic scale weather prediction is in. The specific shape of Figure 5.3.3 and the locations of the three observation density regimes in this simplified system cannot be extended directly to the real atmosphere. However, for real numerical weather prediction data assimilation systems which are similar to our 3DVAR and for forecast models with similar resolution, the estimate of one observation every $(2000\,\mathrm{km})^2$ for the transition regime is certainly well within an order of magnitude. Thus, it seems likely that we have sufficient observations over Northern Hemisphere mid-latitude continents to resolve synoptic scale features, placing us near the dense end of the transition regime or in the dense regime. In less populated Northern Hemisphere regions, we may be in the transition regime or in the dense regime, depending on the number of observations and on the specific atmospheric situation; relevant factors include the error growth rate, the

advection speed, and the type of region from which the errors tend to be propagating. In the Southern Hemisphere, we are probably in some part of the transition regime, perhaps even on the sparse end. If we wish to evaluate the real world observation network more specifically, again further study is needed with a more complex model and with an experimental setup including model error.

We focus on the non-dense observation regime because it is in this regime that additional observations have the greatest benefit, and thus that adaptive observations have the greatest potential. We also focus on the non-dense regime because the specifics of the results in the dense observation regime are sensitive to small changes in the data assimilation system (see Section 5.3). An effective adaptive observation strategy will move the middle regime to a lower observation density (to the left) compared to a fixed or random strategy, producing the same average error reduction with fewer observations. Note that this regime of largest influence from added observations was not selected a priori, but rather was defined by the system. If we had defined a density regime of interest prior to testing the general influence of observations on the analyses and forecasts, we might have expected the regime of largest influence to extend to higher observation densities. Understanding how the system responds to additional observations for a variety of observation densities will be important both when we interpret the results for different global observation strategies in Section 6, and when we test observations added to a pre-existing network of fixed observations in Section 8.

## 5.2 Choice of error norm

The error depicted in Figure 5.1 and nearly all subsequent figures is root-mean-square analysis error averaged in time, throughout the domain, and at all levels, with a streamfunction norm. The general curve shape and the comparisons shown below are similar for all other analysis error norms tested (including energy, potential vorticity, winds, and temperature). As an example, Figure 5.2 shows the same results as in Figure 5.1, but for root-mean-square average analysis error with a potential vorticity norm. Results are also similar for the analysis error at individual levels. We note, however, that for some norms, at some levels (such as lower interior level streamfunction and middle level potential vorticity), dense observations reduce the errors less compared to the saturation value than they do at other levels.

Results for various time- and domain-averaged forecast error norms are also qualita-

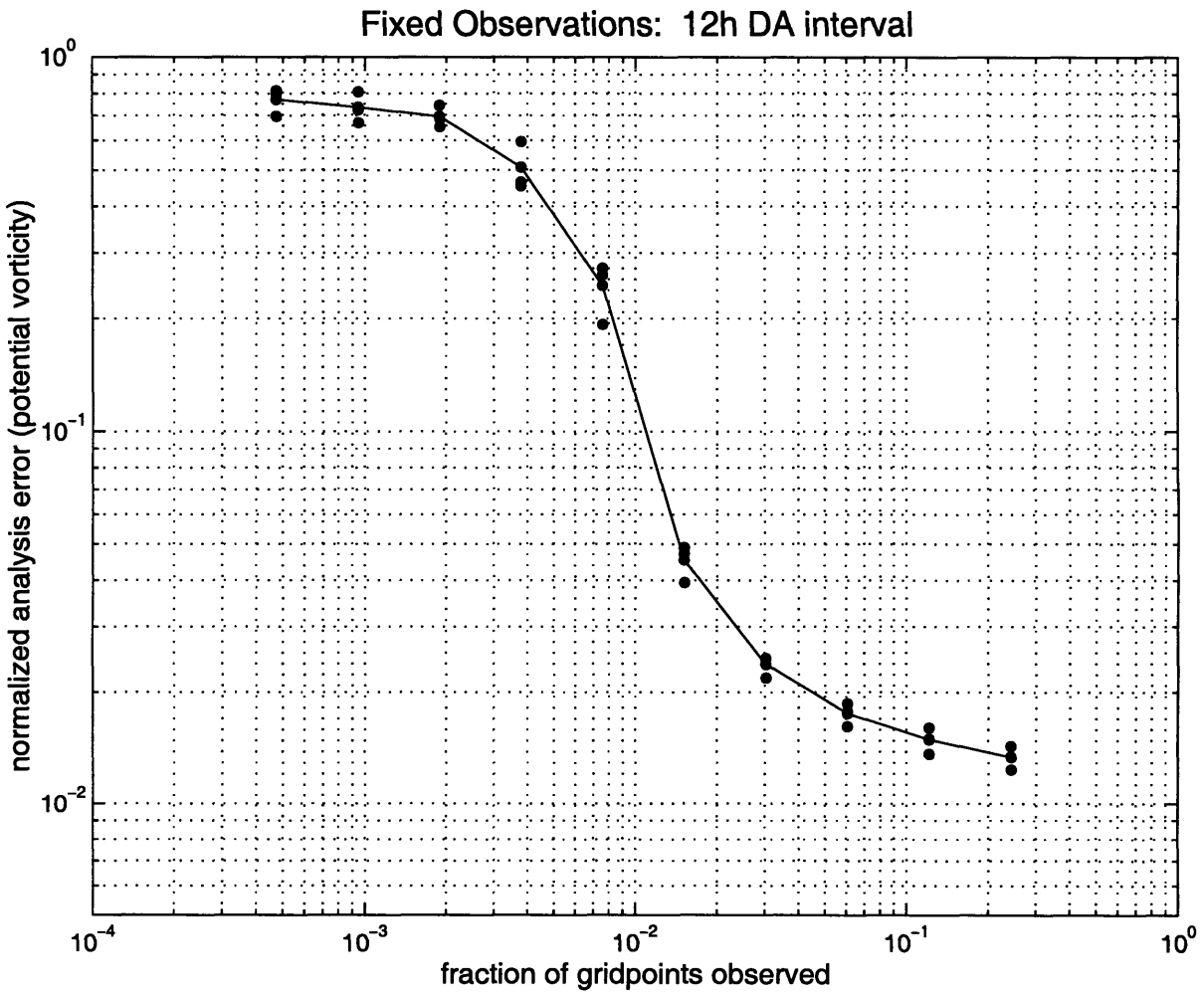



Figure 5.2: As in Figure 5.1, for root-mean-square averaged analysis error with a potential vorticity norm.

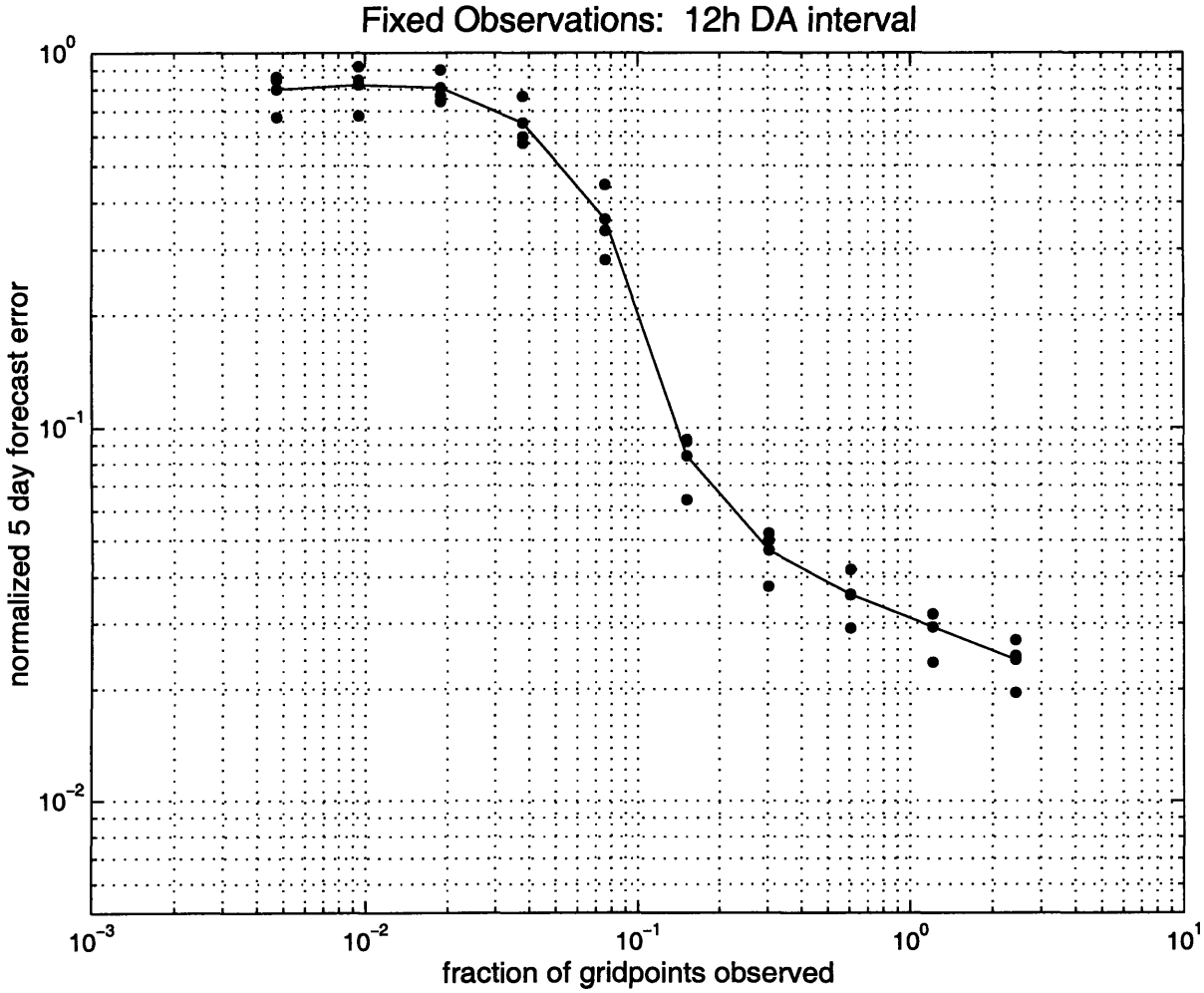



Figure 5.3: As in Figure 5.1, for root-mean-square averaged 5 day forecast error with a streamfunction norm.

tively similar to those presented. As an example, Figure 5.3 shows the same results as in Figure 5.1, but for the 5 day forecast error with a streamfunction norm. The major difference between results for analysis errors and forecast errors, for all of the time-averaged results shown, is that the forecast errors at dense observations are larger than the analysis errors, and thus the overall potential to reduce forecast errors by modifying the observation network is smaller.

Reducing time- and domain-averaged analysis and forecast errors will be beneficial. Adapting observation networks has the most potential, however, when observations are targeted to reduce errors in specific significant forecasts. By defining "significant" as either the deepest or the most rapidly deepening system in the domain according to various norms, selecting one significant system at each error-gathering time, calculating the forecast error locally around that system, and accumulating statistics, we have measured the error associated with the average significant forecast for forecasts up to 10 days. Because the domain tends to be dominated at each time by a few atmospheric systems, each associated with a relatively localized region of analysis and forecast error, the average analysis and forecast error reductions result primarily from reduced errors in one or a few localized regions. Thus, although again the overall decrease in error possible for dense observations is smaller, the time-averaged results for improving local significant forecasts are similar to those for domain-averaged analysis and forecast norms.

The correlation between analysis and forecast errors in these experiments is at least partially caused by our assumption of a perfect forecast model. It is therefore not clear to what extent this result is valid in the real atmosphere, especially on time scales longer than a few days. In addition, the results presented here apply only to the average case, and in some situations we will want to improve one very significant forecast. How to improve forecasts on a case-by-case basis is a complex question, one which we will only begin to address in Chapter 9. On a statistical basis, however, in these experiments, reducing the average analysis error is linked with reducing the domain-averaged forecast error, as well as the forecast error for important individual atmospheric systems. Most results are therefore presented in terms of domain-averaged analysis error with a streamfunction norm.

## 5.3 Sensitivity to data assimilation parameters

In this section, we present several examples of how altering the data assimilation system affects the results for fixed observations. We do so first of all to demonstrate that as described earlier, the shape of the curve is not qualitatively sensitive to the data assimilation parameters. Because the specifics of the results do depend on the data assimilation details, however, particularly when observations are dense, these examples demonstrate one of the main reasons why we decided to focus on non-quantitative results and on comparisons in the low-mid observation density regimes. Changing the data assimilation parameters also allows us to test the importance of some of the choices we made to construct the background error variances in the 3DVAR. Finally, exploring the sensitivity to the data assimilation system helps identify some of the reasons why the system behaves differently in the three observation density regimes. These results will help us interpret the comparisons between different types of observation networks presented in later chapters.

### 5.3.1 Perfect observations

Figure 5.4 shows, along with the results from Figure 5.1, results from the same set of experiments but with perfect observations, i.e. observations taken from the truth state with no errors added. Since the 3DVAR does not converge if the observations are assumed to be errorless, we use the same data assimilation as for imperfect observations. When used with perfect observations, the data assimilation optimized for imperfect observations gives the observations much less weight than it should. Therefore, removing the errors in the observations tests the sensitivity of the system not only to the specific observational errors used, but also to how the data assimilation weights observations relative to the background field when creating the analysis.

The perfect and imperfect observation curves in Figure 5.4 are nearly indistinguishable for low data densities. This suggests that when observational data are limited, improving the accuracy of the observations has little effect — analysis and forecast errors are large because the data assimilation cannot resolve the important atmospheric features with the few observations supplied. As the data density increases, the average analysis errors become much smaller for the perfect observations runs than for the imperfect observations runs. This indicates that observational errors do reduce the ability of the data assimilation to resolve smaller-scale features. Even without errors in the ob-

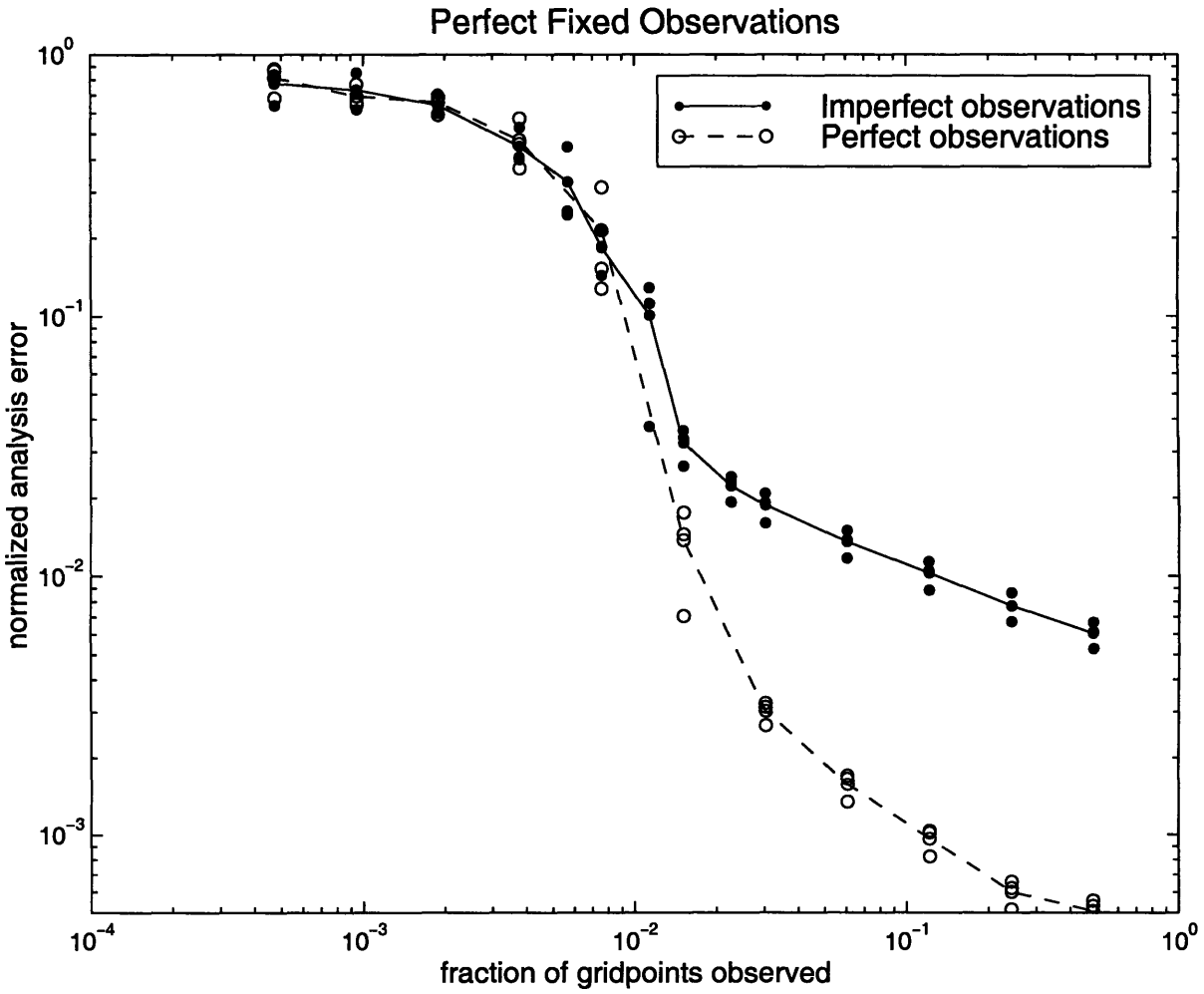



Figure 5.4: As in Figure 5.1, for perfect observations as well as standard (imperfect) observations. The data assimilation remains the same.

servations, however, additional observations still have a smaller influence above a certain observation density, and the general curve shape is similar. Results for other strategies are affected similarly when observation errors are removed, changing primarily for high observation densities. Most of the results are therefore presented only for imperfect observations, and unless otherwise stated, all average comparisons shown look qualitatively the same for perfect observations.

### 5.3.2 Varying convergence cutoff

We consider the data assimilation converged in the standard runs when, in the conjugate residual descent algorithm used to solve Equation (3.1), the maximum residual is less than or equal to 2% of the initial maximum residual. To test if the results are affected by the convergence criterion, Figure 5.5 shows results from the same set of experiments as in Figure 5.1 but with the 3DVAR convergence criterion varied.

When given a stricter convergence criterion (.1% of initial maximum residual), the 3DVAR takes more iterations to converge, but the results change little. This confirms that our original criterion was sufficiently strict. With a less strict criterion (20%), the data assimilation may not extract all of the information from the observations, which we would expect to lead to the analysis errors decreasing less rapidly as the observation density increases. In the low-mid observation density regime, less convergence does in fact produce larger analysis errors. As the observation density becomes large, however, the average analysis errors become smaller for the least strict criterion (20%) than for the other two sets of results in Figure 5.5. For dense observations, then, analysis errors are actually on average *smaller* for fewer data assimilation iterations.

Figure 5.6 tests the same sensitivity to the convergence criterion, but with perfect observations. If there are no errors in the observations, a less strict convergence criterion behaves as expected in all cases, producing larger average analysis errors at all observation densities. This suggests that when the standard data assimilation was allowed to converge more completely for observations with errors, it tried to fit the noise as well as the information in the observations. More iterations therefore created a worse analysis. The 3DVAR tries to extract more information from dense observations than is available at least in part because the background error covariances are too large for this regime, which causes the data assimilation to weight the background too little and the observations too much. The importance of this weighting is discussed further in the next section. Because we focus primarily on results for non-dense observations, however,

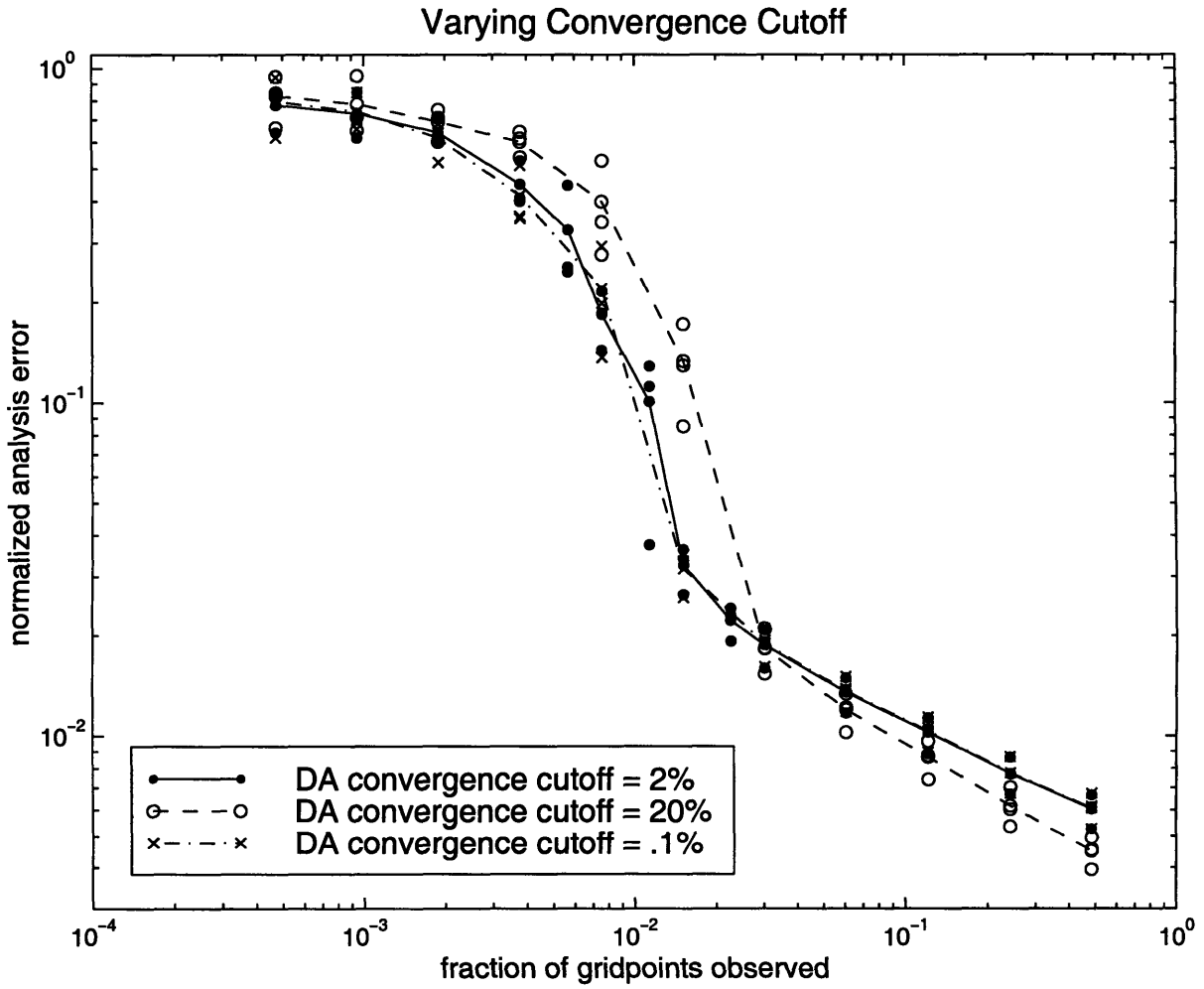



Figure 5.5: As in Figure 5.1, but with different convergence criteria in the conjugate residual solver which inverts the 3DVAR equation. In the standard runs, the data assimilation is considered converged when the maximum residual in the solver is 2% of the initial maximum residual. In the two other experiments shown, the convergence is cut off at 20% and .1% of the initial maximum residual.

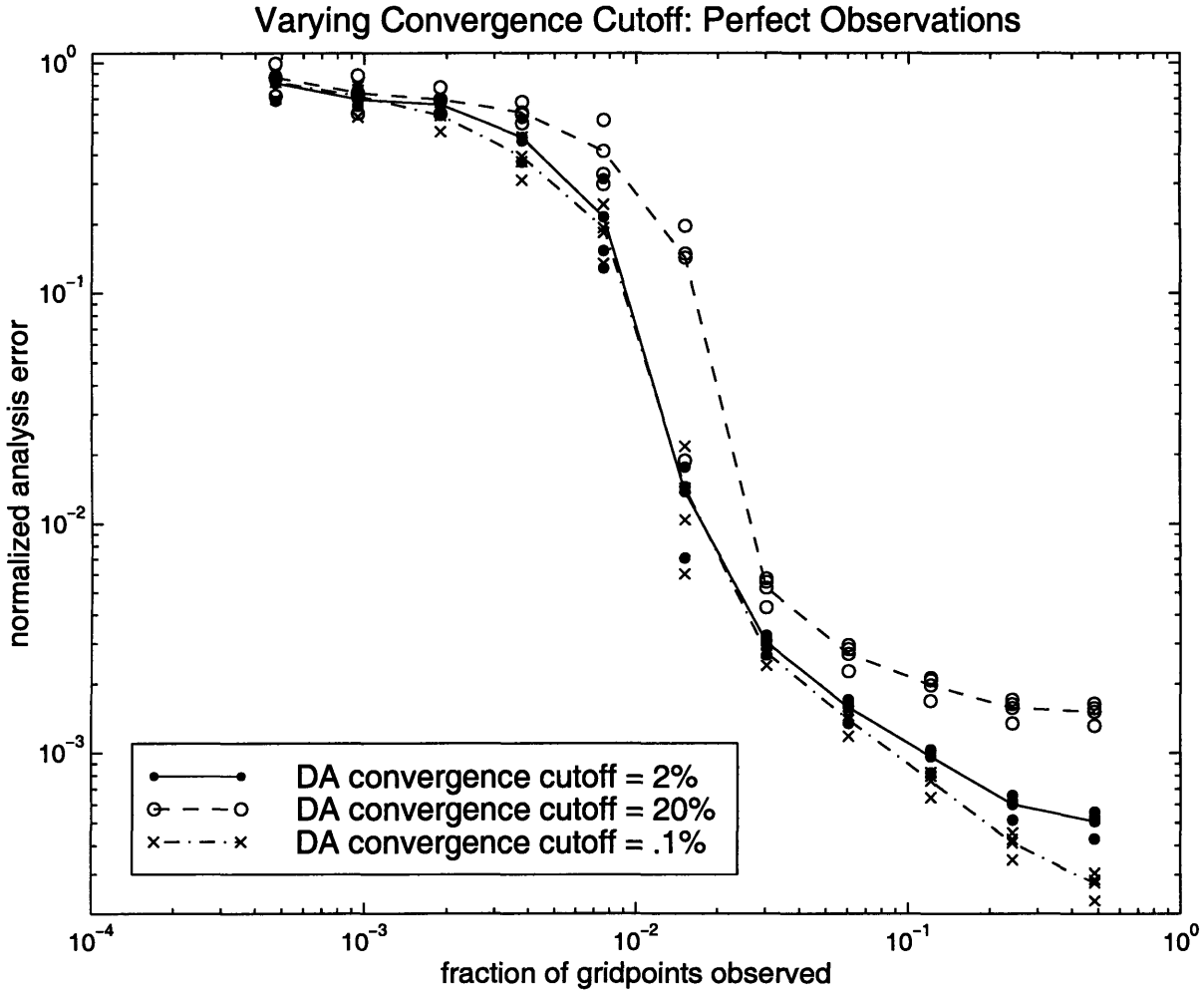



Figure 5.6: As in Figure 5.5, for perfect observations.

this dependence on the convergence criterion in the data dense regime does not affect the comparisons for different observation networks.

### 5.3.3 Varying background error covariances

Next we explore the sensitivity of the fixed observation results to several aspects of the background error covariance matrix ($\mathbf{B}$) used in the 3DVAR. First, we scale $\mathbf{B}$ by different constants to see how our conclusions are affected both by the weighting of the background field compared to the observations and by the assumption that $\mathbf{B}$ is the same for all observation densities. Then, we change the relative weighting of the background errors at different spatial scales to test the importance of specifying $\mathbf{B}$ correctly within our assumptions. Finally, we discuss how the results might be different with a data assimilation which does not assume that the background error statistics are diagonal in spectral space and are time-invariant.

#### Scaling B by a constant

Figure 5.7 shows the average analysis error at different fixed observation densities (the same experiment as in Figure 5.1), but with the background error covariances scaled by different constants. In the data assimilation at each individual time, scaling the background error covariance matrix $\mathbf{B}$ by a constant changes only the magnitude of the analysis increment, leaving its structure the same. Experiments with the 3DVAR have shown that in this simulated system, scaling $\mathbf{B}$ by a constant is roughly equivalent to optimizing $\mathbf{B}$ for observation densities other than the one selected in Section 3.4 (32 fixed observations, approximately 1.5% on the normalized density scale). Therefore, this not only tests how sensitive the results are to the overall weighting of the background data relative to the observation data, but also tests our assumption of constant data assimilation for all observation densities.

Multiplying the standard $\mathbf{B}$ by 100 causes the 3DVAR to assume a larger forecast error, weights the observations more relative to the background field, and optimizes the data assimilation for few observations. Dividing $\mathbf{B}$ by 100 assumes a small forecast error and optimizes the data assimilation for many observations. As we would expect, compared to the other curves in Figure 5.7, each data assimilation (including the standard) performs best near the density for which it was optimized. We can simulate what the curve would look like if $\mathbf{B}$ had been optimized for each individual observation density by connecting the points with the lowest error at each density in Figure 5.7. Without

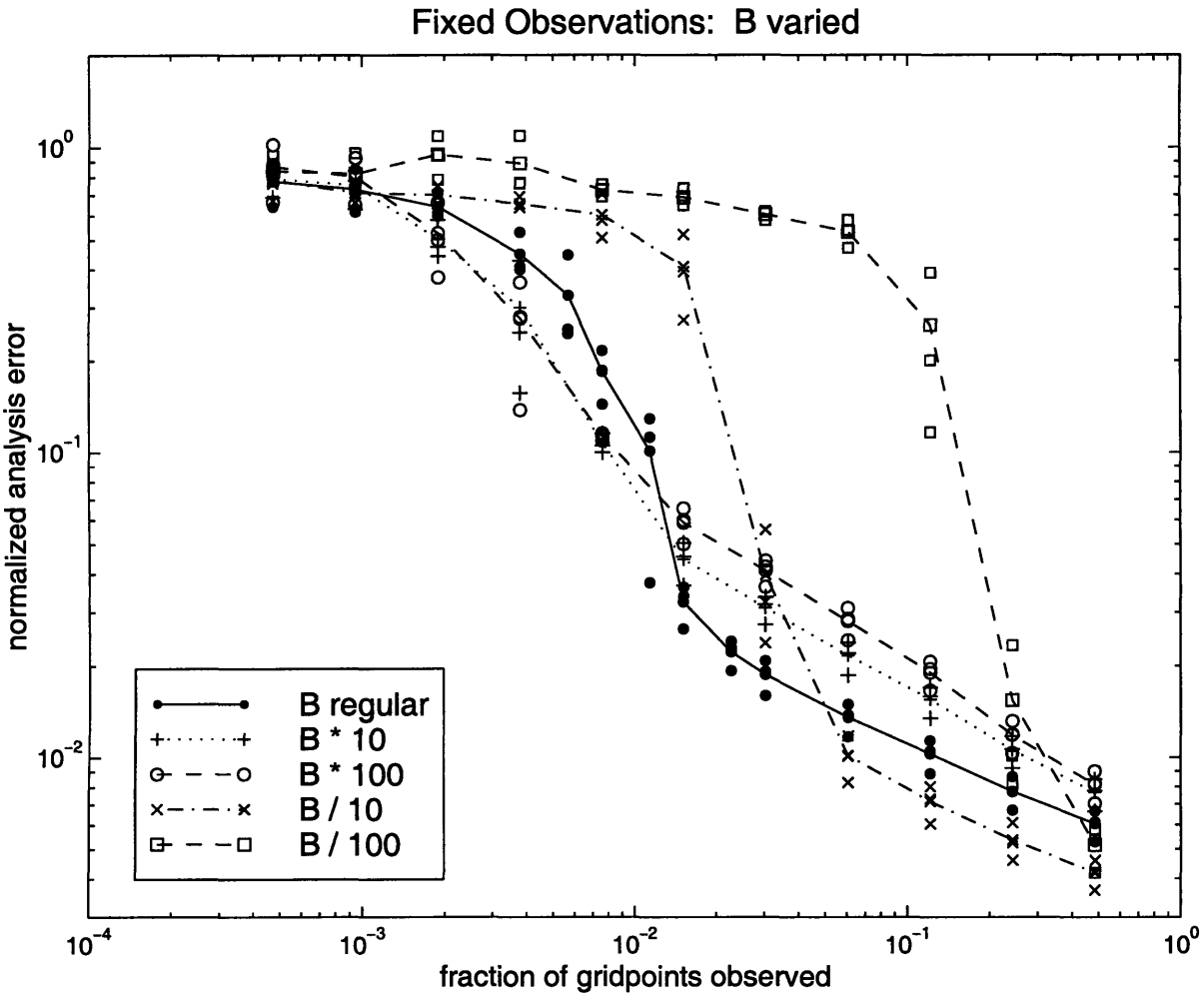



Figure 5.7: As in Figure 5.1, with **B** used in the 3DVAR scaled by different constants. Scaling **B** by a constant changes only the magnitude of each analysis increment and is approximately equivalent to optimizing **B** for different observation densities.

the assumption of constant $\mathbf{B}$ for all observation networks, the fixed observations curve would have smaller errors and steeper slopes in both limits than those in Figure 5.1, a broader and shallower transition regime, and a less clear boundary between the limits and the transition regime, but still a similar overall shape.

Despite the general similarities, the differences between the individual curves in Figure 5.7 show that the background error covariances are a crucial factor in how the system is able to use observations. From the $\mathbf{B}/100$ curve, for example, we can see that if the background error statistics given to the data assimilation are significantly smaller than they are in reality, the data assimilation uses very little information from the observations until they have reached some critical density. Above this critical density, errors are reduced rapidly, but it still takes about 10 times the number of observations as in the standard data assimilation to produce the same error reduction. With $\mathbf{B} \times 100$, on the other hand, the 3DVAR ignores much of the information from more than a few observations. The standard $\mathbf{B}$ is a compromise, in which the 3DVAR does not use data optimally for either sparse or dense observations but is still able to use some of the information from the observations in both limits. This suggests that given our assumption of constant data assimilation, we chose to optimize $\mathbf{B}$ for a reasonable observation density.

Figure 5.7 indicates that our transition between sparse and dense data is fairly abrupt at least in part because the transition regime occurs near the observation density for which the 3DVAR is optimized. To avoid the somewhat exaggerated shape, we could improve the background error weighting at the different observation densities by explicitly optimizing the 3DVAR for different densities. Or, we could improve the weighting by estimating the magnitude of $\mathbf{B}$ during the run (for example, using the ensemble to rescale the standard $\mathbf{B}$ as in Pu et al. 1997, or to estimate the full error covariances as in the ensemble Kalman filter in Houtekamer and Mitchell 1998). We decided to leave the data assimilation fixed, however, optimized for one observation network, because to the extent possible we wanted to isolate the effects of changing observation networks from changes in the data assimilation system. Data assimilation schemes with non-constant $\mathbf{B}$ are also currently still in developmental stages. Since we have optimized $\mathbf{B}$ in the observation density regime in which we focus our study, and since we focus primarily on qualitative comparisons, we do not believe that this is a crucial assumption for our results.

### Scaling B differently at different spatial scales

Figure 5.8 demonstrates how the results for fixed observations can be affected if we change how the data assimilation weights background errors at different spatial scales. To change the relative weighting in these two examples, the error variance in $\mathbf{B}$ for each spectral coefficient is multiplied or divided by the square of the corresponding wavenumber. $\mathbf{B}$ is then multiplied by a constant to keep the maximum background error covariance of the same order of magnitude, in order to avoid drastically changing the overall weighting of the background compared to the observations. Unlike when $\mathbf{B}$ is scaled by a constant, the structure of the analysis increment is changed as well as the magnitude, and the resulting error statistics are incorrect for all observation densities. Multiplying $\mathbf{B}$ by the spectral wavenumber assumes a larger background error at smaller spatial scales than in the standard 3DVAR. The data assimilation then extrapolates more of the information from the observations to smaller scale structures (Figure 5.9 shows a sample analysis increment for one observation, to be compared with Figure 3.4). Dividing $\mathbf{B}$ by the wavenumber assumes a smaller background error at smaller scales and shifts the analysis increments to larger scale structures (Figure 5.10 compared with Figure 3.4).

When there are few observations, the effect of $\mathbf{B}$ is most important since then the horizontal structure in the analysis increments must come almost entirely from $\mathbf{B}$. When there are many observations, $\mathbf{B}$ is used to interpolate between the closely spaced observations but is not needed to extrapolate information, and the three sets of results in Figure 5.8 are therefore similar. For low-mid density observations, increasing the relative magnitude of $\mathbf{B}$ at smaller spatial scales has only a small effect on average. Given Figures 5.9 and 5.10, this indicates that in some cases, misspecifying $\mathbf{B}$ can have a large effect for individual analysis increments, but only a small effect overall. Increasing $\mathbf{B}$ at larger spatial scales, on the other hand, makes the 3DVAR perform significantly worse in the low-mid density regime. In fact, if the analysis increments are shifted to larger scales and there are few observations, the average analysis error is greater than it would be with no observations – the observations actually degrade the analysis. This demonstrates that if $\mathbf{B}$ is misspecified, more data does not always produce better analyses and forecasts, not only in individual cases but also on average.

### Relaxing the assumption that B is time- and space-invariant

The biggest limitation in the data assimilation is not the 3DVAR parameters them-

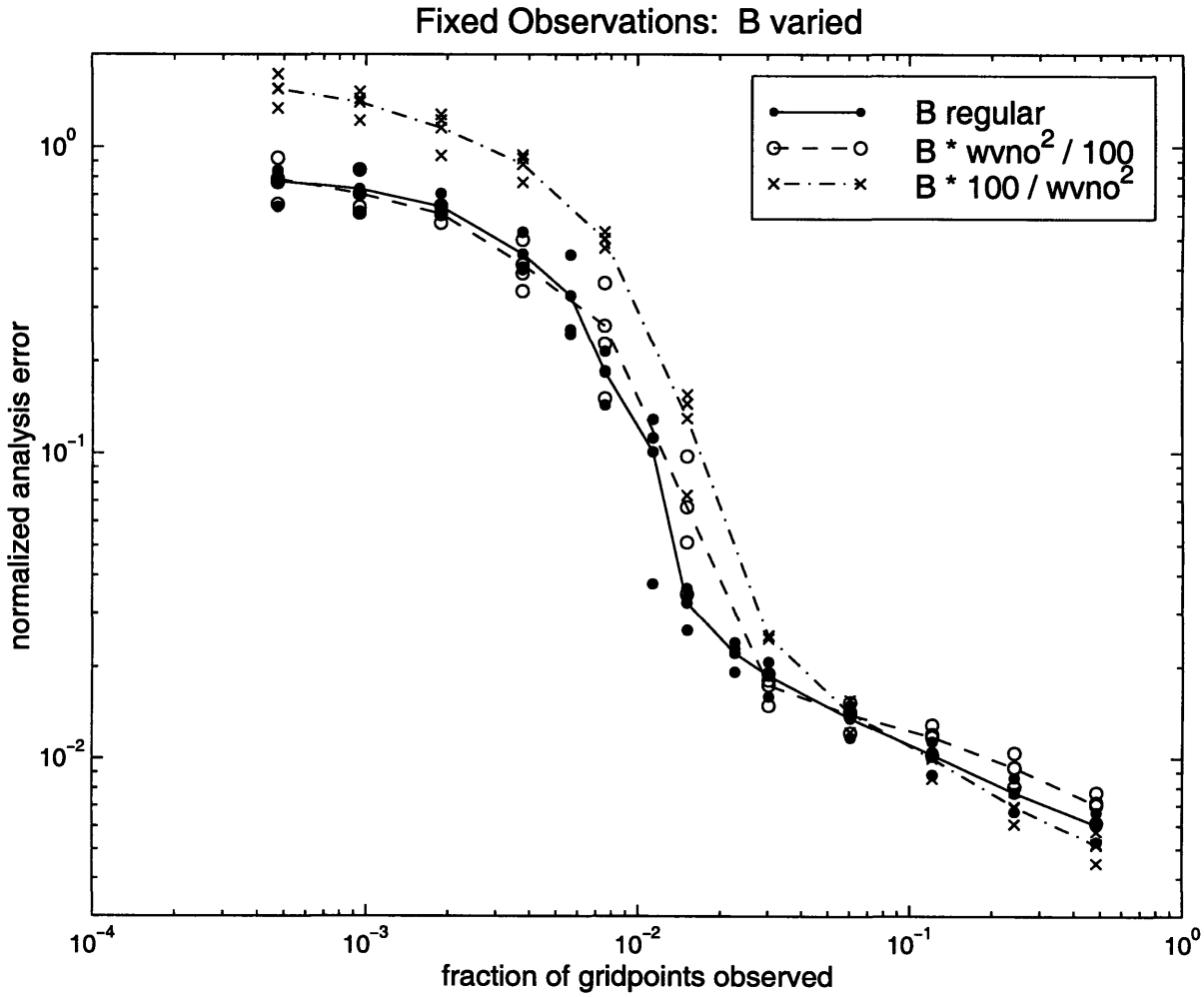
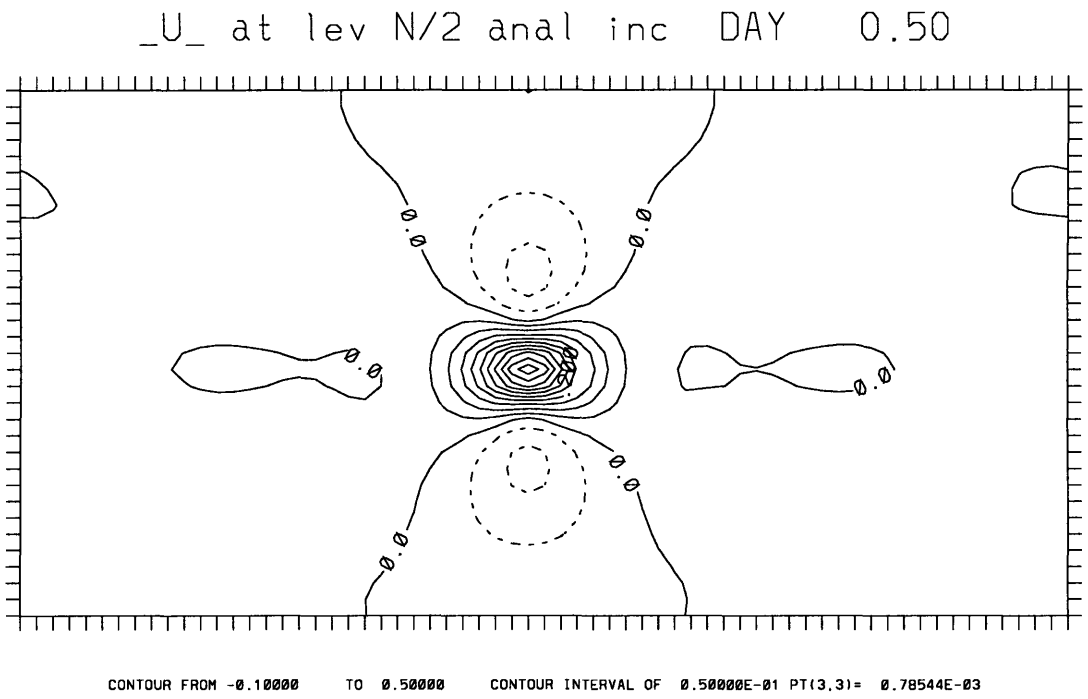



Figure 5.8: As in Figure 5.1, with errors at different horizontal scales weighted differently in $\mathbf{B}$. Using $\mathbf{B} \times$ (wavenumber$^2$/100) in the 3DVAR shifts the analysis increments to smaller scales (see Figure 5.9), and using $\mathbf{B} \times$ (100/wavenumber$^2$) shifts the analysis increments to larger scales (see Figure 5.10).

a)




b)

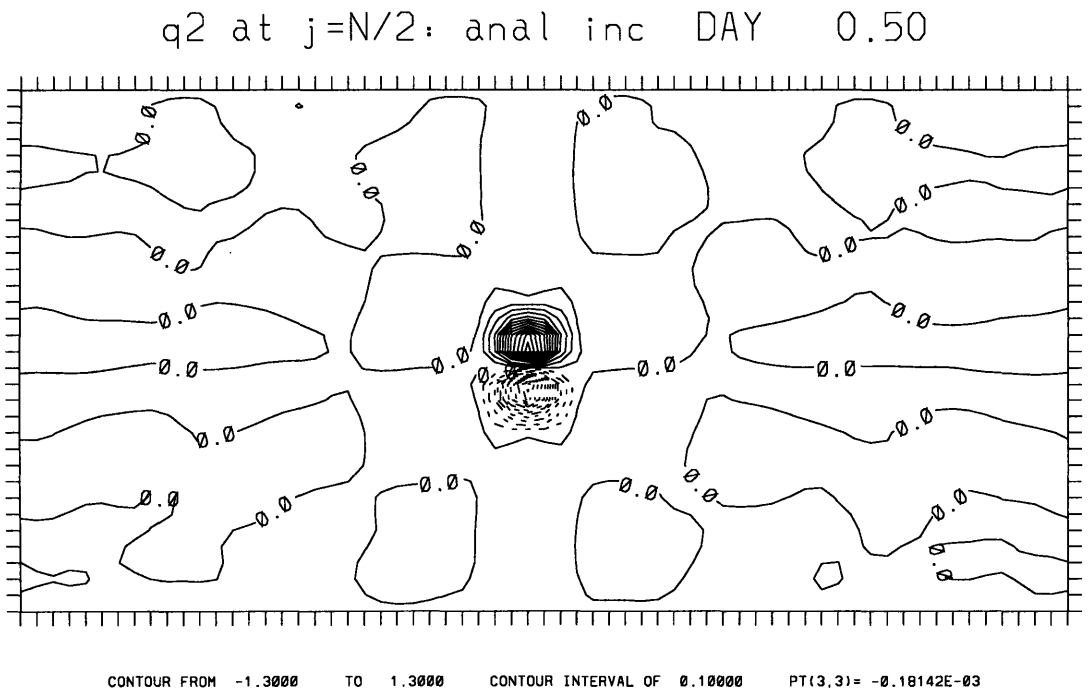
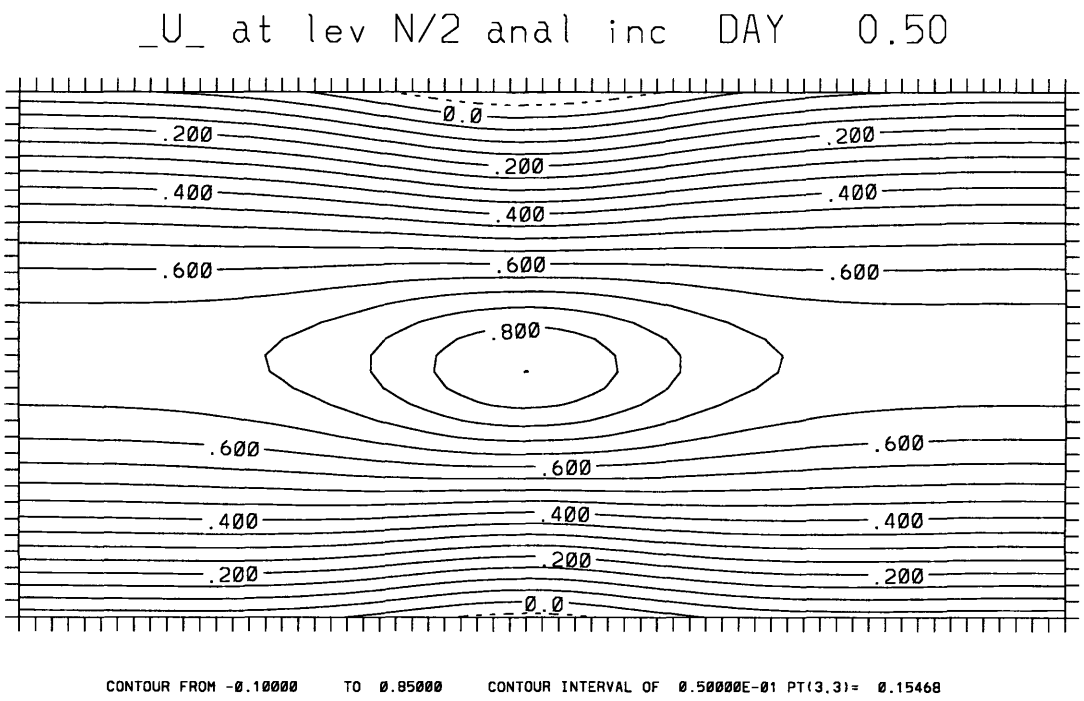



Figure 5.9: Analysis increments produced by the 3DVAR for a single zonal wind observation $u = 1$ at level 3, as in Figure 3.4 but with $\mathbf{B} \times (\text{wavenumber}^2/100)$. The variables plotted are non-dimensionalized (a) zonal wind and (b) potential vorticity, both at level 3. The contour interval is .05 in (a) and .1 in (b).

a)




b)

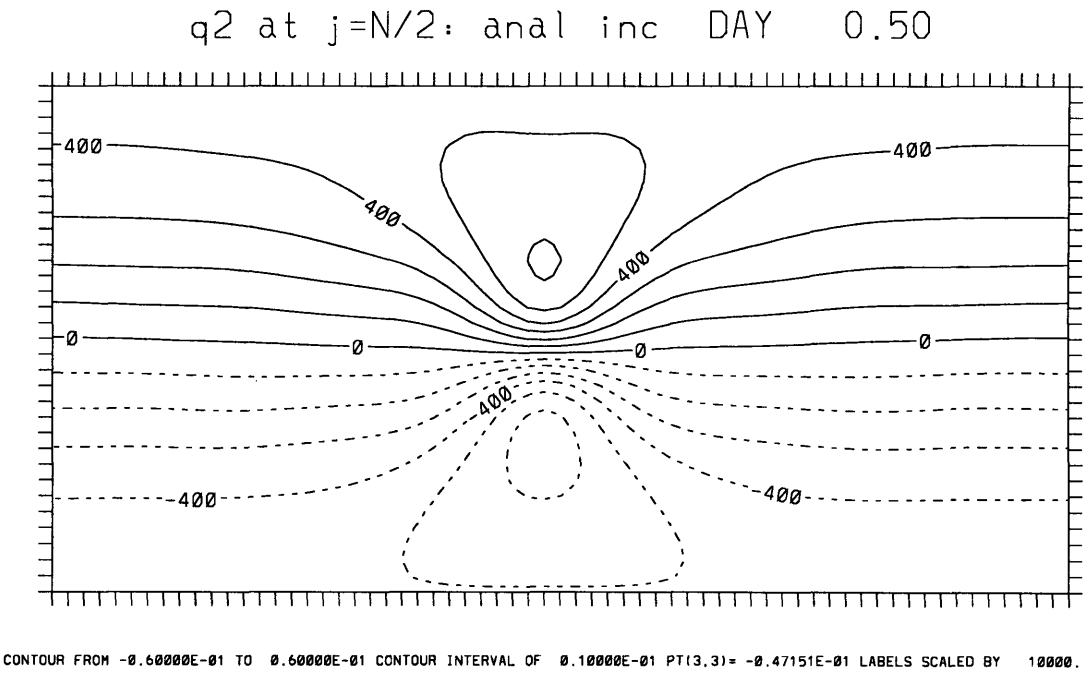



Figure 5.10: Analysis increments produced by the 3DVAR for a single zonal wind observation $u = 1$ at level 3, as in Figure 3.4 but with $\mathbf{B} \times (100/\text{wavenumber}^2)$. The variables plotted are non-dimensionalized (a) zonal wind and (b) potential vorticity, both at level 3. The contour interval is .05 in (a) and .01 in (b).

selves, but rather the assumptions used to develop the 3DVAR statistics. It is difficult to test the sensitivity of the data assimilation to these assumptions, since we have neither the necessary information nor the computational resources to solve the ideal variational algorithm. We have, however, performed preliminary experiments adding more complex time- and space-dependent background error covariances to the 3DVAR, using the full set of time- and space-dependent correlations between the actual errors in the background field at each analysis time. Although we have not studied the results in great detail, they do indicate that, as we would expect, a better data assimilation (with accurate time- and space-dependent error correlations) reduces errors in the data dense regime and shifts the transition regime to a lower observation density.

Obviously, since we can never know the actual background error correlations in the real world, this data assimilation technique is unrealistic; we would instead have to estimate the correlations using an ensemble Kalman filter, a modified Kalman filter, a variant of 4DVAR, or another technique. It is not clear how much having to estimate the time- and space-dependent background error covariances will limit the data assimilation's effectiveness. In addition, the results from these experiments suggest that without good information about long-distance background error correlations, the data assimilation cannot extract sufficient information from only a few fixed observations to reduce analysis errors in separate synoptic systems. Because accurate long-distance error correlations are likely to be difficult to obtain in the real atmosphere (Houtekamer and Mitchell 1998), even with a more sophisticated data assimilation, the sparse observation limit seems likely to extend to more complex systems.

Similarly, in any atmospheric observing and forecast system there is likely a limit in which observations are sufficiently dense that adding observations produces little benefit. This could be due to the errors in or the representativeness of the observations, the model resolution or accuracy, the error growth rates at different scales, or the limitations of the data assimilation. Even in extreme situations, one can imagine a hypothetical, although likely completely irrelevant for real world forecasting, dense observation limit. For example, if we had perfect observations with a nearly perfect data assimilation system (at all spatial scales), analyses for many observations would be nearly the same or as good as those for very few observations. For a realistic data assimilation system and a forecast model with limited resolution, taking more observations than the number of relevant degrees of freedom (such as taking several observations for each gridpoint) would benefit the analysis at best only a small amount. The dense observation limit therefore also seems likely to exist in real numerical weather prediction. To shift this

limit, we will need to improve the forecasting system, for example by improving the data assimilation system, the observation network or accuracy, or the forecast model or model resolution.

Naturally, for any given model and data assimilation configuration in the real atmosphere, we may always be in an observation density regime in which we do not see the sparse or the dense observation limit. Given the inhomogeneous observation network, however, and the variety of atmospheric systems with errors growing on many different time and space scales, these limits are likely to be more than just theoretically important in some situations. Understanding the observation density regime that one is in for a certain atmospheric situation, observation network, forecast model, and data assimilation system may therefore be the most important part of optimizing how any type of observational resources are used.

## 5.4 Sensitivity to data assimilation interval

As more complex data assimilation schemes, new data sources, and faster computers are developed, there are likely to be changes in the rate at which observational data are input into numerical forecast models. To explore the possible effects of more frequent data input, and to test how sensitive our results are to the ratio between the data input interval and the time scales for advection and error growth, we have experimented with changing the interval between data assimilations. Figure 5.11 shows, along with the same results as in Figure 5.1, the analysis error for different spatial densities of fixed observations when data are taken and assimilated at intervals other than the standard 12 hours. Consistent with the assumption of fixed data assimilation for all observation densities, the 3DVAR remains fixed, optimized for a 12 hour interval.

The x-axis in Figure 5.11 is the spatial density of observations at each data assimilation time. For each density along the x-axis, as the data assimilation interval is halved (say from 12 hours to 6 hours), the same number of observations is taken twice as often. Intuitively, one might think that twice as much data would produce a better analysis, but it is clear from Figure 5.11 that this is not always the case. In the dense observation limit, for example, all of the curves nearly converge. This indicates that if the observations are sufficiently dense, just as adding observations at new locations at the same data assimilation interval benefits the analysis only a little, adding more frequent observations at the same locations also has a small influence. Note, however, that in the

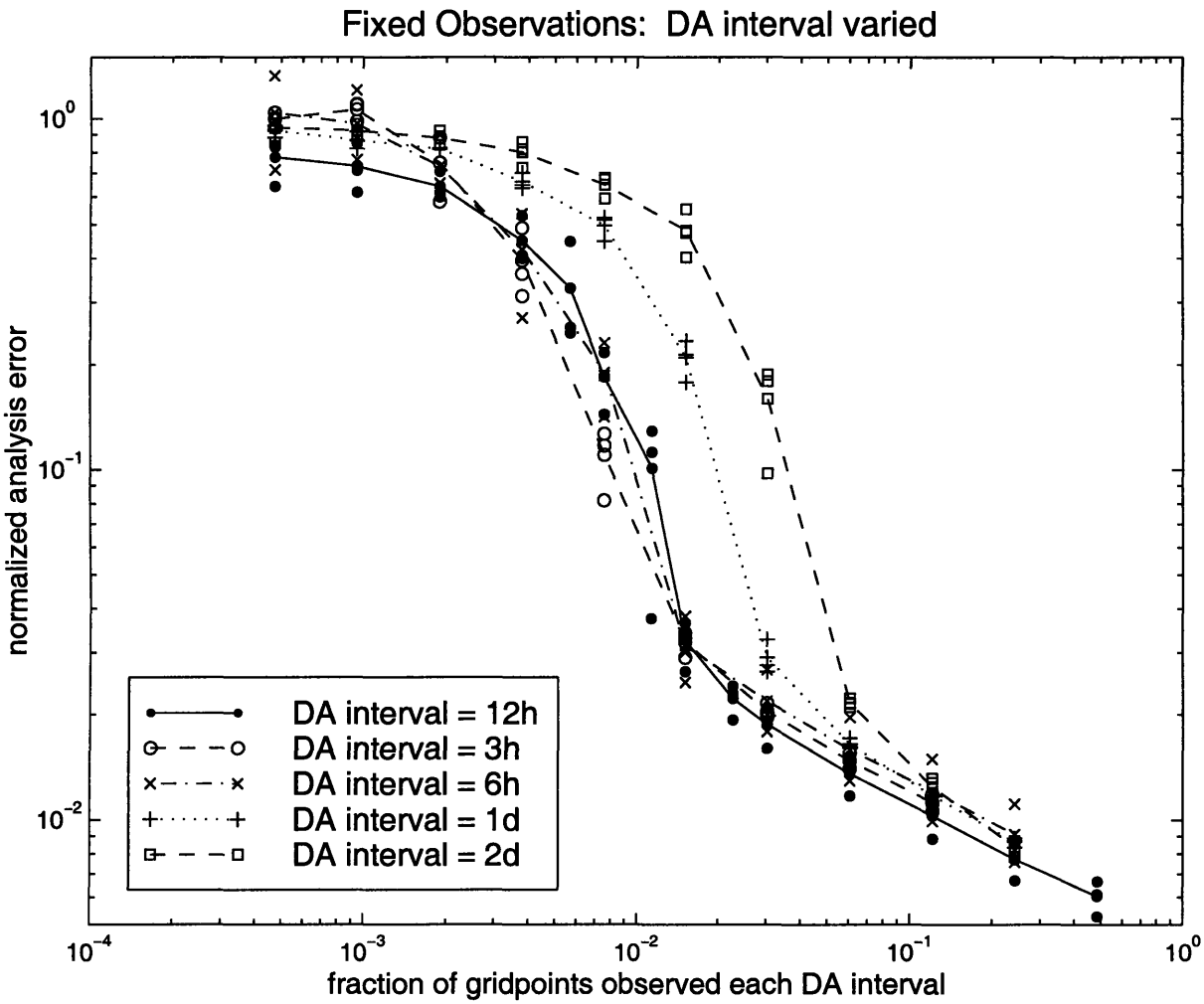



Figure 5.11: As in Figure 5.1, for fixed observations, but with 3 hour, 6 hour, 1 day, and 2 day data assimilation intervals, in addition to the standard 12 hour interval. The x-axis is normalized by the number of gridpoints and represents the spatial observation density at each data assimilation time.

data dense regime, the standard data assimilation system uses additional observational data suboptimally, and so the specific results may be due to how we have chosen to optimize the data assimilation system.

Even for non-dense observations, adding more frequent data may have only a small influence. The 12, 6, and 3 hour data assimilation results in Figure 5.11 nearly overlap, suggesting that atmospheric features move and change sufficiently slowly that a fixed location sees redundant information every 6 hours. In this model setup, the error doubling time is a few days and the advective time scale is 5 hours. At all spatial observation densities, then, when observations are taken at fixed locations more frequently than every 12 hours, the additional data helps the data assimilation system very little with identifying the types of features which can be resolved by the scales in the model. Only in the middle regime in data density, and only for data assimilation intervals of 12 hours or longer, does changing the rate at which data are taken significantly affect the results.

Figure 5.12 presents the same results as in Figure 5.11, but with the x-axis of each curve normalized by the number of data assimilation times per day. The x-axis is now observation density in space *and* time, and might be thought of as the rate at which observation resources are expended. Each observation density represents a constant amount of data input per day but at different data assimilation intervals; for example, 2 observations every 6 hours for a 6 hour data assimilation interval is equivalent to 8 observations per day for a 1 day interval. We can now begin to answer questions such as: given a certain number of observations each day, how frequently would we like to observe, at the expense of observing at fewer locations, to maximize error reduction? These results for fixed observation networks will be compared with those for random and for adaptive networks in Section 6.3.

If we are restricted to fixed observation locations, Figure 5.12 shows that, given a constant total number of observations taken more frequently than every 12 hours, the analysis error is reduced the most when we sample at more locations less frequently instead of at fewer locations more frequently. This is as might be expected from the redundancy described earlier for frequent data gathering at fixed locations. Allocating many observations at few times might also be advantageous because it provides the data assimilation system with more information on the structure of analysis errors at each time. However, since it is difficult to distinguish among data assimilation intervals less frequent than 12 hours, the system likely does not prefer a certain data assimilation interval for fixed observations as long as the observations are not redundant.

We do not believe that this result indicates that rawinsondes more frequent than

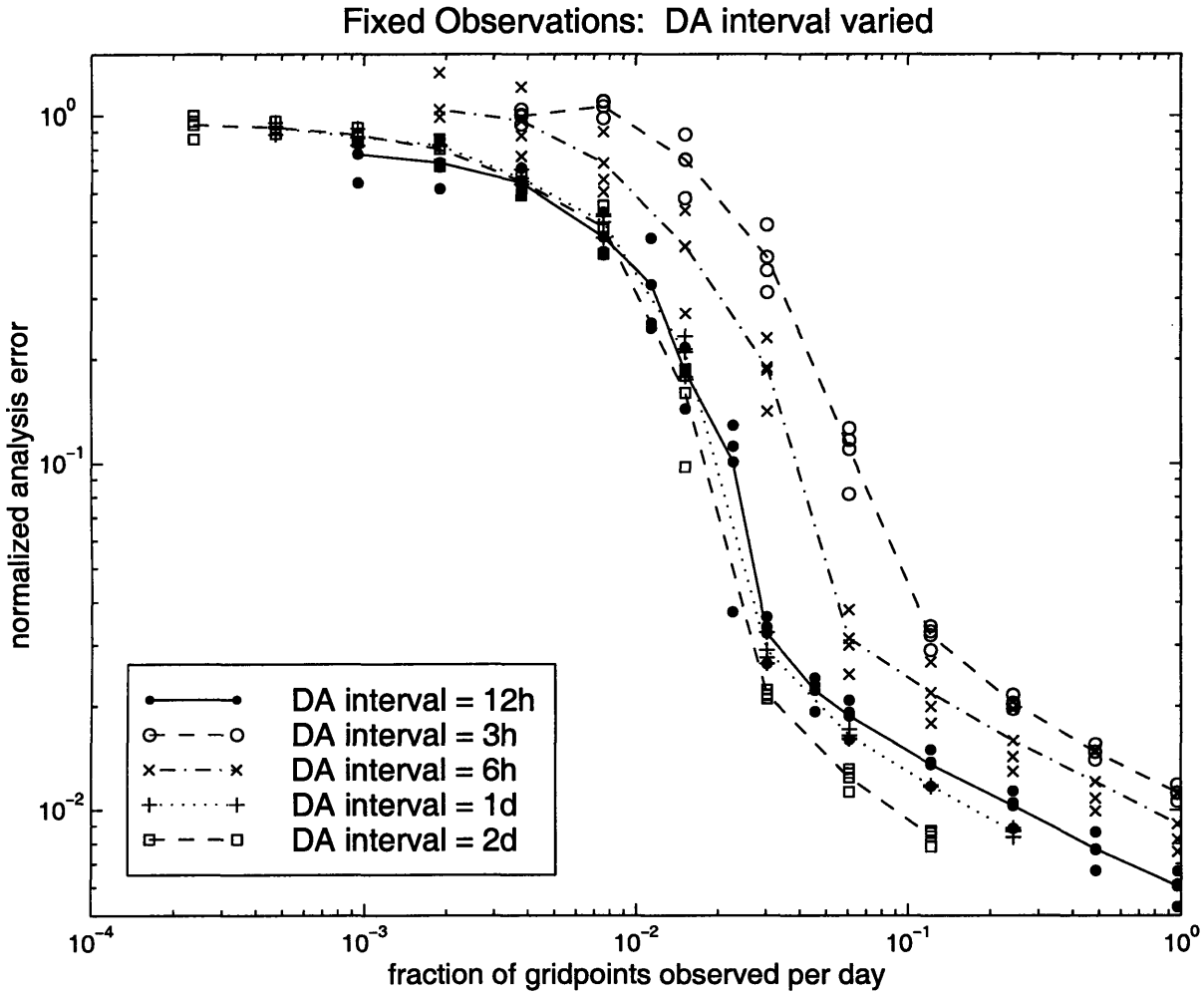



Figure 5.12: As in Figure 5.11, but with the x-axis normalized both by the number of gridpoints and by the number of data assimilation intervals per day. The x-axis represents the observation density in space and time, i.e. allocation of a fixed amount of observational resources.

12 hours are not useful in the real world. First, because the error growth in the standard resolution QG model is relatively slow and because the QG model dynamics are unrealistic at sub-synoptic scales, observations may not be as clearly redundant in real numerical weather prediction, and they are not likely to be redundant at similarly low frequencies. The influence of the additional observations is also strongly limited by the model resolution and by the lack of time- and space-dependent error covariances in the data assimilation. In addition, more frequent data from fixed platforms could be quite beneficial for several other weather prediction applications not addressed by this study, such as subjective forecasting or numerical modeling and forecasting of mesoscale and smaller features. Finally, even when the data are sufficiently dense that the data assimilation system and forecast model can no longer make efficient use of new information on average, more frequent observations could improve an analysis and forecast in individual situations with significantly faster than average advection or error growth. These potential benefits are especially important considering how little fixed platforms cost compared to other observation platforms.

On a more fundamental level, Figure 5.12 demonstrates how — given an atmosphere with different error growth rates at various spatial scales, a model with a certain spatial resolution, and a data assimilation system — the effect of adding observations in space and time can depend strongly on the observation density and error regime. Before these results can be applied to real atmospheric prediction, further study is needed with a more realistic forecasting system, particularly one with a more complex dynamical model and with an experimental setup including forecast model error. Nevertheless, for this simulated system, the results for different observation densities will be important when interpreting the results for adaptive observations presented in the following chapters.

# Chapter 6

# Results with global non-fixed observations

Using the framework developed from the fixed observation experiments in Chapter 5, we now compare results when observation locations, instead of being fixed, are moved randomly or adaptively. For this set of experiments, there is no standard observation network to which observations are added; all observations are allocated throughout the domain according to only one strategy. As in the fixed observation experiments, the data assimilation procedure remains fixed for all observation networks.

The three strategies tested in this chapter are described in detail in Section 4.2; here we provide only a brief review. Fixed observation locations are selected arbitrarily prior to the experiment and remain fixed. Random observation locations are selected arbitrarily at each targeting time. Ideal AER adaptive locations are selected at each targeting time where the actual background error (the most recent forecast error) is a maximum before the data assimilation is performed, in other words where the actual analysis error would be the largest if there were no observations. Because the ideal AER adaptive strategy requires perfect information, it can only be implemented in an idealized system such as this one; a more realizable implementation of the ideal AER adaptive strategy is tested in Chapter 7.

## 6.1 Comparison of fixed, random, and ideal adaptive strategies

Figure 6.1 compares the average analysis error as a function of observation density

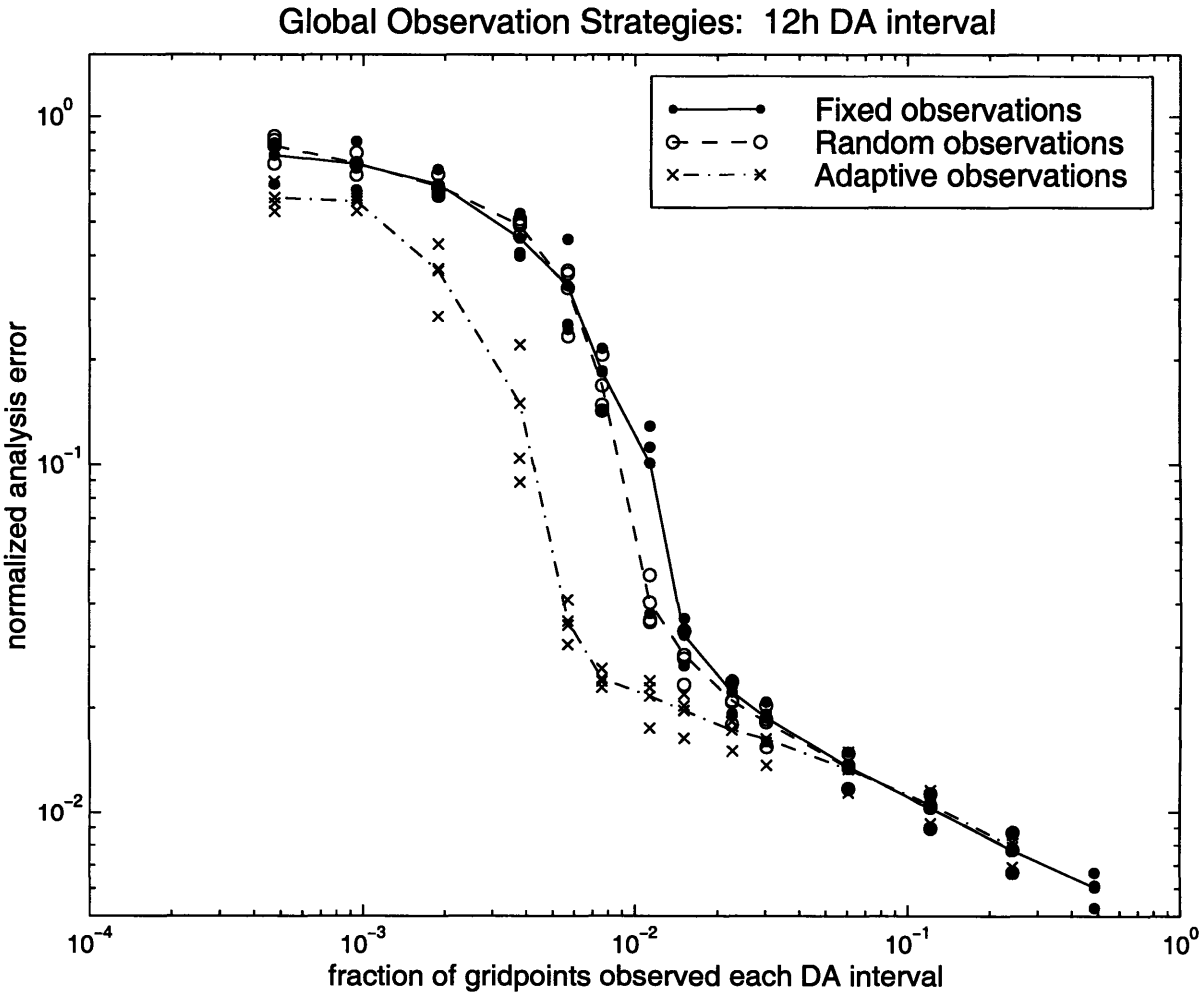



Figure 6.1: As in Figure 5.1, but with error as a function of observation density for random and ideal AER adaptive observations in addition to fixed observations, all for a 12 hour data assimilation interval. The observation strategies are defined in Section 4.2.

for three types of global observation networks: fixed (the same results as in Figure 5.1), random, and ideal AER adaptive. The data assimilation and targeting intervals are 12 hours. The influence of the observations is again measured in terms of streamfunction analysis error root-mean-square averaged throughout the domain and over many model states. Recall from Chapter 5 that in the data dense regime, additional fixed observations provide only a small additional benefit. In Figure 6.1, we see that if observations are spatially dense, additional randomly or adaptively moving observations also provide only a small additional benefit. In the data dense regime, the results for the three strategies are also similar. This indicates that for spatially dense observations, the data assimilation and forecast model system has little preference among the observation strategies tested.

As discussed in Section 5.2, results for other time-averaged analysis and forecast error norms are qualitatively similar to those shown for a streamfunction analysis error norm, with differences primarily in the data dense regime. As an example, Figure 6.2 shows results from the same set of experiments as in Figure 6.1, but for reduction in the domain- and time-averaged analysis error with a potential vorticity norm. Figure 6.3 shows the same results, but for the 5 day forecast streamfunction error. We do not discuss the results for different error norms in detail, but rather present them only for a general comparison, to supplement the discussion in Section 5.2. For most of the other results in this study, since comparisons between different error norms are similar to those shown here, we only present and discuss results for analysis errors in terms of streamfunction.

In the examples shown, the strategies cannot be clearly distinguished if there are more than an average of approximately one observation every $(1000\,\mathrm{km})^2$ at a 12 hour data assimilation interval. Because the error reduction for dense observations is sensitive to changes in the experimental setup, data assimilation system (as discussed in Section 5.3), and the error norm, this specific observation density may vary. Figure 6.1 suggests, however, that for data above a certain density, it is difficult to improve analyses and forecasts significantly by adding *any* type of observations, adaptive or non-adaptive. This result supports some of the FASTEX and NORPEX results (currently unpublished), that it is difficult for additional observations to have a significant positive influence when the background error without the observations is relatively small (such as when the observations are taken just downstream of a data dense region). To benefit significantly on average from adaptive observations at higher observation densities, we will likely require some combination of a more effective data assimilation system and a more sophisticated

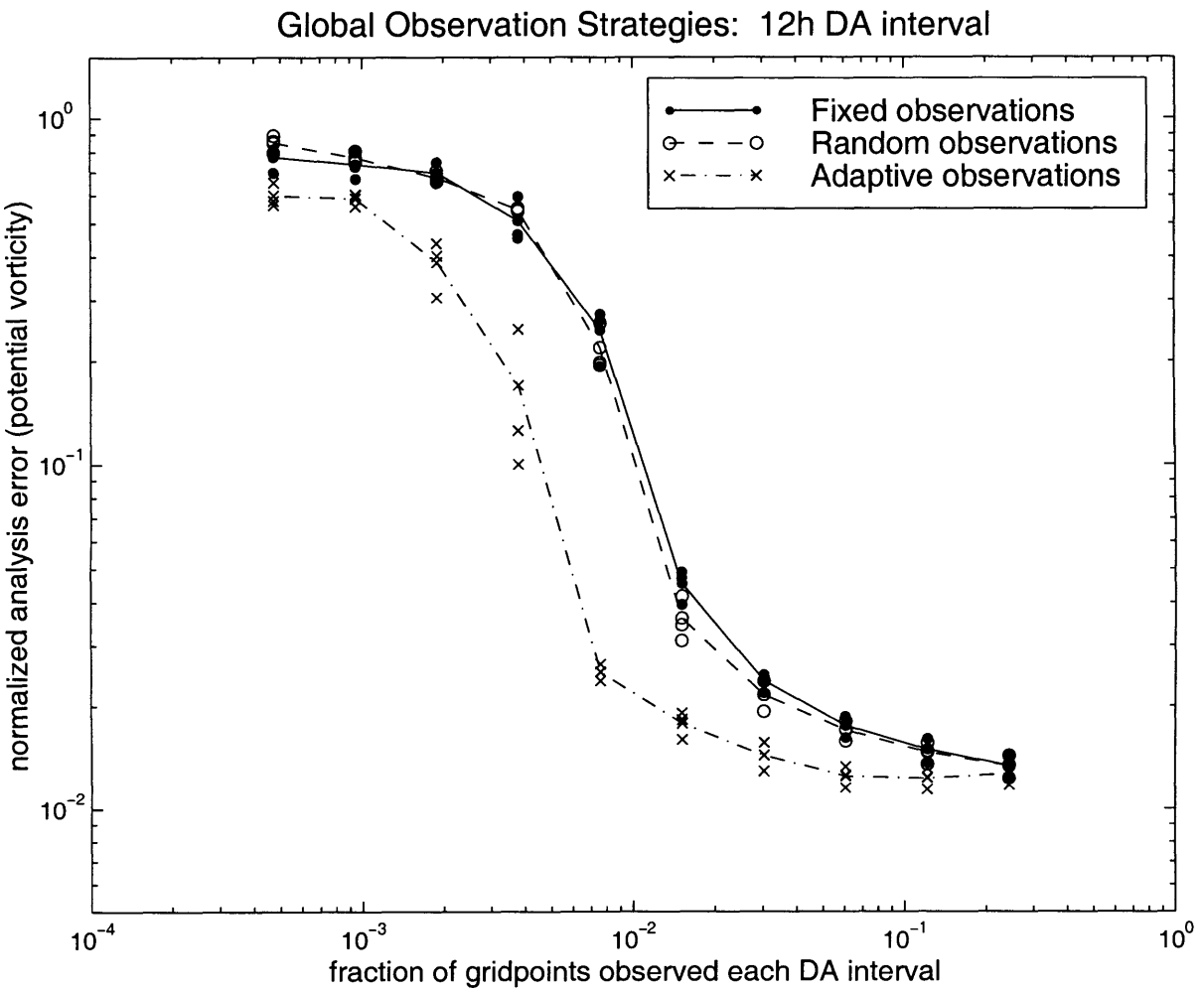



Figure 6.2: As in Figure 6.1, for root-mean-square averaged analysis error with a potential vorticity norm.

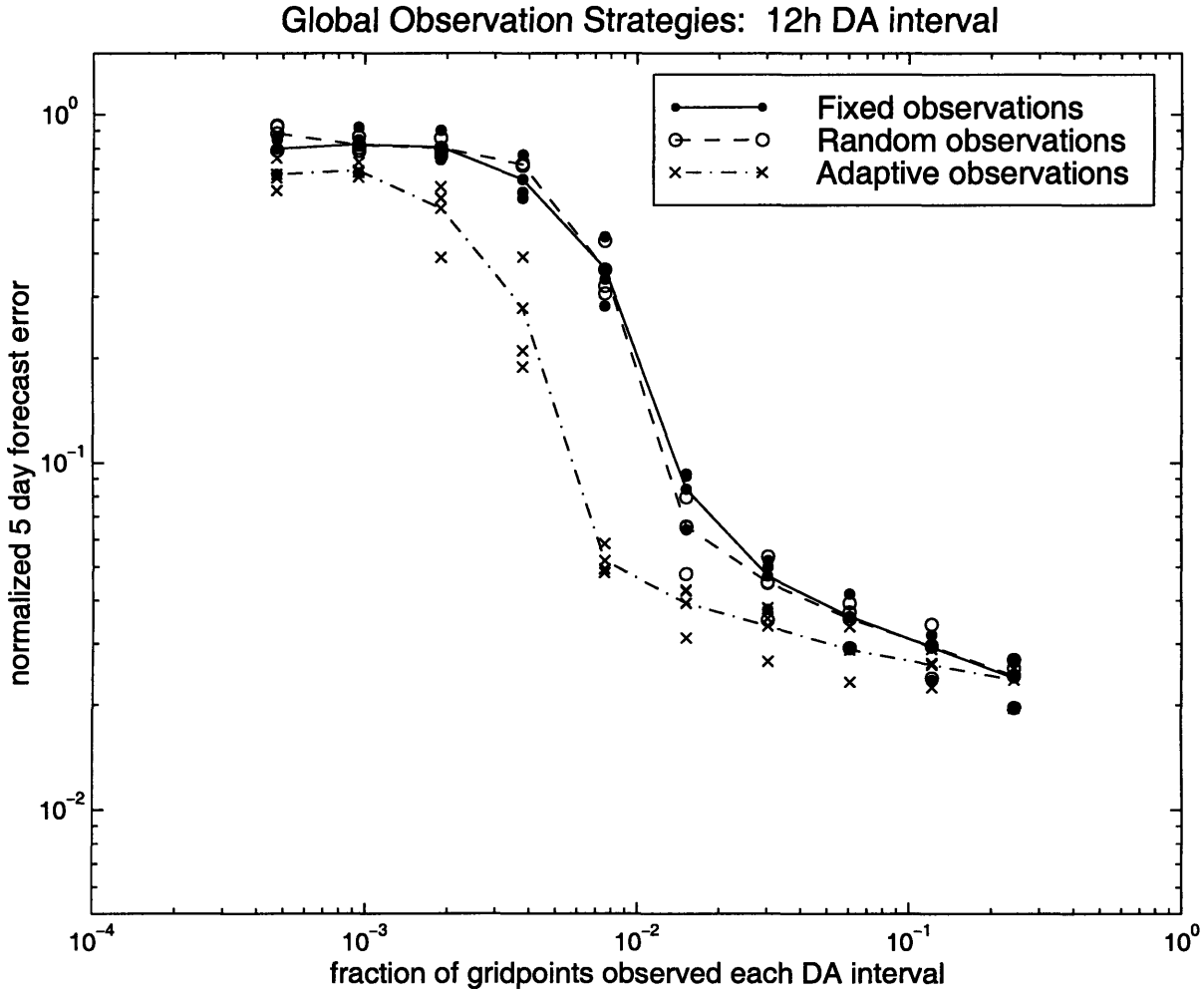



Figure 6.3: As in Figure 6.1, for root-mean-square averaged 5 day forecast error with a streamfunction norm.

strategy, and we may have to measure the benefits with a more specific error norm. With the current data assimilation system, we focus in the spatially non-dense data regime, where adding and adapting observations has the greatest potential to reduce errors.

For non-dense observations taken at a 12 hour data assimilation interval, in some situations a random observation network performs slightly better than a fixed network. Although this is not apparent from Figure 6.1, if the experimental setup or the data assimilation parameters are changed slightly, random observations do in some cases clearly perform better than fixed observations at a 12 hour data assimilation interval. The possible benefit from randomly moving observations is not a surprise given the redundancy of fixed observations at shorter data assimilation intervals (described in Section 5.4). Because arbitrarily moving the observations can reduce average error, the random strategy will be our targeting "null case." Adaptive strategies will have to improve upon both the fixed and random strategies to be considered effective.

Non-dense adaptively allocated observations improve the analysis, on average, significantly more than the same number of randomly allocated or fixed observations. This holds true even taking into account the approximated spread of the results for each network, as indicated by the symbols on Figure 6.1. The adaptive strategy tested here is the ideal AER strategy using energy averaged at all levels as the norm for targeting. Most other targeting norms tested, including background error measured in terms of root-mean-square averaged potential vorticity (averaged enstrophy), winds, and temperature at all levels, and each of these norms at individual levels, produced similar average improvements. A streamfunction norm was less effective for analysis error-based targeting, presumably because it smooths smaller scale structures.

Although the specific values depend on the details of the experiment and the adaptive strategy tested, Figure 6.1 illustrates the range of error reduction we can expect from the ideal AER adaptive strategy for non-dense observations. In the example shown, in order to reduce the average analysis error by a factor of approximately 30 compared to saturation, we only need to take about one half as many adaptive observations as we would fixed or random observations. Essentially, by using the adaptive strategy, the rapid dropoff of error followed by the near saturation of error reduction has been shifted to lower observation densities. Because of the substantial savings in observational resources, we consider the ideal adaptive strategy effective for non-dense observations. As discussed in Section 5.2 and demonstrated in Figure 6.3, average comparisons for forecast errors are qualitatively similar to those shown (although quantitatively somewhat less favorable). Therefore, Figure 6.1 shows that, at least in the context of a QG model,

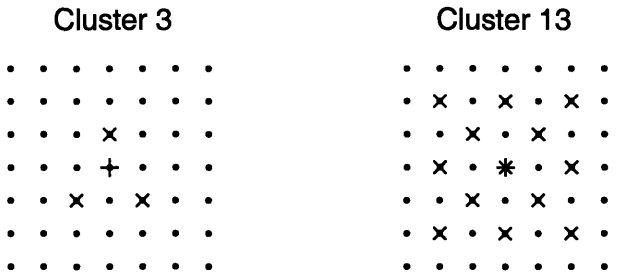



Figure 6.4: Cluster 3 and cluster 13 patterns used for clustered observations. Dots represent gridpoints, a "+" is plotted at the targeted location (selected by the appropriate strategy), and "x"s are plotted at observed locations. Cluster 13 has an observation at the targeted location; cluster 3 does not.

with good information about the background errors and in the appropriate regime of data density in space and time, it is possible to objectively adapt observation networks to reduce average analysis and forecast errors.

## 6.2 Clustered vs. single observations

The forecast errors tend to focus in a few regions at any given time, so the adaptive strategy tested tends to cluster the observations in one or several groups. To ensure that the difference between the random and ideal adaptive strategies is not due to their different observation spacing tendencies, we have tested taking several observations simultaneously in a pre-specified pattern around each target, which we will refer to as "clustered" observations. The clusters simulate, for example, one or more aircraft sent with infinite speed to drop rawinsondes at several locations within a specific region. As the targets are selected, their spacing is constrained to be such that the clusters do not overlap; the total number of non-redundant observations then remains consistent between experiments. Most of the clustered observation results shown use two sample clusters, depicted in Figure 6.4. The first, which we will call "cluster 3," is a triangle of three observations surrounding the targeted location. The second, "cluster 13," observes at the 13 alternating gridpoints in a 1000 km $\times$ 1000 km grid around the target. Cluster 3 was selected because it is one of the more effective small clusters, and cluster 13 was selected because it is one of the more effective clusters that provides a near-complete survey of a targeted region. Unless otherwise stated, all average comparisons shown are qualitatively similar for single observations and for various sized and shaped observation clusters.

In most cases, the random and the ideal adaptive strategies reduce average analysis error as much or more by allocating observations singly than by allocating the same number of observations among one or more clusters of pre-specified size and shape. This result is somewhat counter-intuitive, since with the limited information about the background error in the data assimilation procedure, in most individual cases we would expect more error reduction with clusters of observations large enough to sample more of the atmospheric structure (discussed further in Section 9.5). At least the most effective analysis error-based strategies, however, seem to prefer on average to be allowed to decide how many observations to allocate to each target based on the specific circumstances. The strategy can then select its own observation pattern, according to the shapes, structures, spatial extents, and relative magnitudes of the background errors.

Only at longer data assimilation intervals or for less effective adaptive strategies (such as the estimated analysis error reduction strategy tested in Chapter 7 or strategies using a streamfunction targeting norm) do observations in pre-specified patterns reduce errors more than the same number of single observations. For some strategies, global adaptive observations are effective *only* when observations are taken in clusters. Because it is not possible to measure the error reduction produced when observation locations are selected in a certain way without the context of how the observational data from these locations are used, it is difficult to separate the effects of the strategies and the observation clusters from the effects of the data assimilation system and the forecast model. These less effective AER adaptive strategies, however, appear on average to be able to identify general regions, but not specific locations, where an observation will improve the analysis and forecasts.

As discussed above and in Section 9.5, clusters of observations may be more effective than single observations if we are trying to improve forecasts not on average, but in certain individual situations. It is also likely that real world targeted observing platforms (such as planes and satellites) will take multiple observations in a region, since the incremental cost of taking additional observations is small once a platform has been targeted. To try to identify the "optimal" patterns for adaptive observations, we tested many different cluster patterns, with various shapes and numbers of observations ranging from 250 km (the grid spacing) to about 3000 km in diameter. Although certain cluster patterns do seem to be preferred, we do not present any results comparing specific patterns for two reasons. First, the effectiveness of each cluster is very dependent on the data assimilation system. Second, the optimal patterns tend to vary from situation to situation, again illustrating the preference of the adaptive strategy to choose its own

observation pattern based on the circumstances. In general, however, we found that adding more observations per cluster tends to reduce errors for a constant number of targeted clusters, but that adding observations at distances greater than about 1250 km from the targeted location (i.e. using clusters greater than 2500 km in diameter) or placing observations at adjacent model gridpoints is usually not beneficial.

## 6.3 Sensitivity to data assimilation interval

Given a certain number of observations, we would like to know not only how to distribute them in space, among strategies and targeted locations, but also how to distribute them in time. For example, if there are eight observation platforms available, each of which can be sent anywhere to observe but only once per day, should they be used all at once or at different times? We will address this issue by comparing how the same three observation strategies (fixed, random, and ideal AER adaptive) behave as the time interval between observations is changed. As in Section 5.4, this comparison also tests how sensitive the results are to the ratio between the data input interval and the error growth time scale.

Figure 6.5 shows the average analysis error as a function of observation density for the same three strategies shown in Figure 6.1, but with observation locations selected and observations assimilated every 3 hours, four times as often as the standard. Just as for the 12 hour data assimilation interval, the three strategies cannot be distinguished if the observations are more dense than an average of approximately one observation every $(1000\,\mathrm{km})^2$. For non-dense observations, since observations taken more frequently than every 12 hours are redundant at fixed locations, randomly moving the observations produces a significantly better analysis than leaving the observations fixed. As in Figure 6.1, moving observations intelligently, using the adaptive strategy, improves the analysis even more. In fact, comparing Figures 6.1 and 6.5, we see that as the time interval between data assimilations decreases, the benefit from adapting the observation network increases. Recall that to reduce the average analysis error by a factor of about 30 compared to saturation, for a 12 hour data assimilation interval we require about 2 times as many fixed or random observations as we do adaptive observations. For a 3 hour interval, to reduce the errors by the same factor of 30 we require about 4 times as many random observations or about 8 times as many fixed observations as we do adaptive observations.

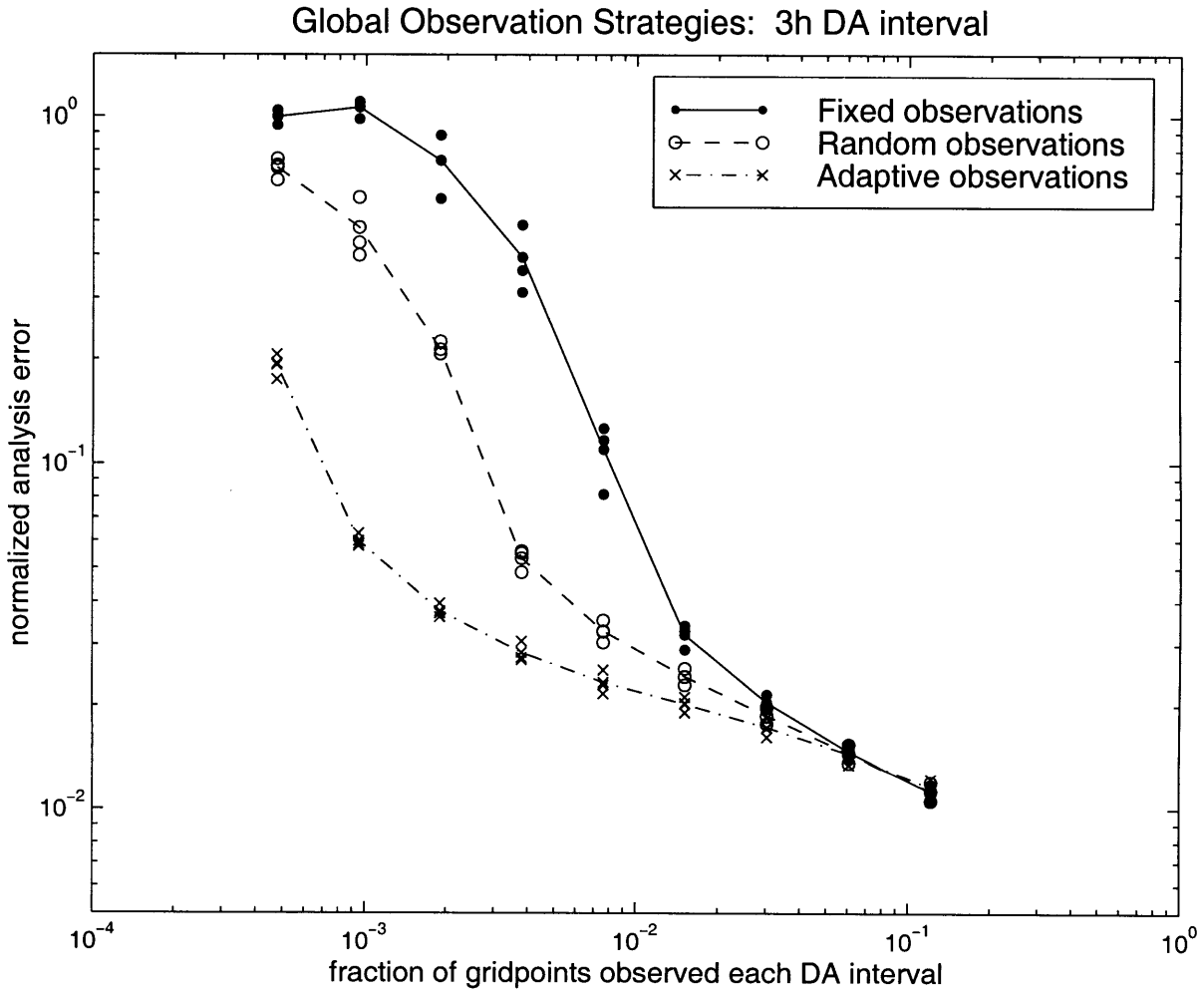



Figure 6.5: As in Figure 6.1, for a 3 hour data assimilation interval.

With less frequent data assimilation (not shown), random observations are no longer an improvement over fixed observations, confirming that observations at fixed locations are not redundant in this forecasting system at intervals longer than 12 hours. Adapting observations is also less beneficial at longer data assimilation intervals, and the adaptive observation results depend more on the background error norm used to target and the prespecified clustering and spacing of the observations.

Why are adaptive observations better at shorter data assimilation intervals, more so than random and fixed observations? In Figures 6.6 and 6.7, we allocate a constant number of observations at different data assimilation intervals, as in Figure 5.12 for fixed observations but using the random and the ideal AER adaptive strategies, respectively. Although we focus on the results for non-dense observations, we note that for dense observations, less frequent data assimilation is the most beneficial for all three observation strategies. This occurs because for spatially dense observations, the random and adaptive results are nearly the same no matter how often the same number of observations are taken and assimilated. In other words, if we renormalize Figures 6.6 and 6.7 so that the results are a function of only the spatial observation density instead of the total amount of observational resources, then all of the data assimilation intervals produce the same results in the spatially dense regime, as in Figure 5.11 for fixed observations.

For non-dense observations, if observation locations are randomly selected at each time (Figure 6.6), this forecast model and data assimilation system has no preference for how the observations are allocated in time. This means that arbitrary distribution of observations in space and in time can be interchanged. Recall from Figure 5.12 that for data assimilation intervals shorter than 12 hours, fixed observations perform best when more observations are taken less often. In Figure 6.7, we see that adaptive observations behave in the opposite sense — as long as the observations are below a certain density, the adaptive strategy performs best when it takes observations more frequently.

In order to understand this behavior, we have experimented with selecting the adaptive locations at every other data assimilation time. If, for example, observation locations are targeted every 24 hours but observations are taken and assimilated every 12 hours, at the same locations at two consecutive data assimilation intervals, Figure 6.8 shows that the average analysis error falls between that for the 12 hour and 24 hour interval curves. This suggests that adaptive observations in this simulated system are more beneficial when they are allocated at more times rather than simply at more locations for two reasons: first, because observations can be taken frequently in potentially important regions (at the same locations or at slightly different locations); and second, because the

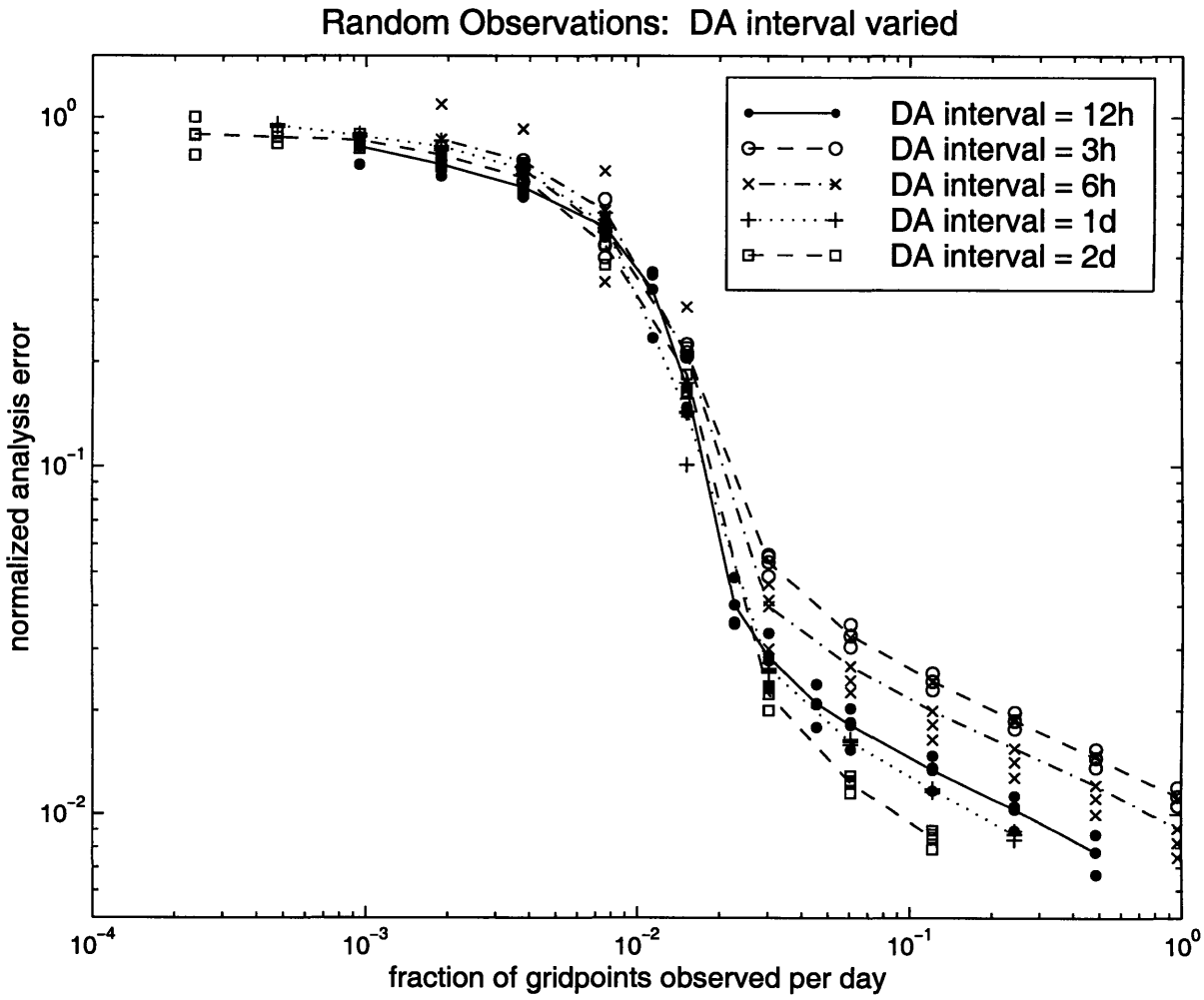



Figure 6.6: As in Figure 5.12, for random observations. The x-axis represents observation density in space and time, i.e. allocation of a fixed amount of observational resources.

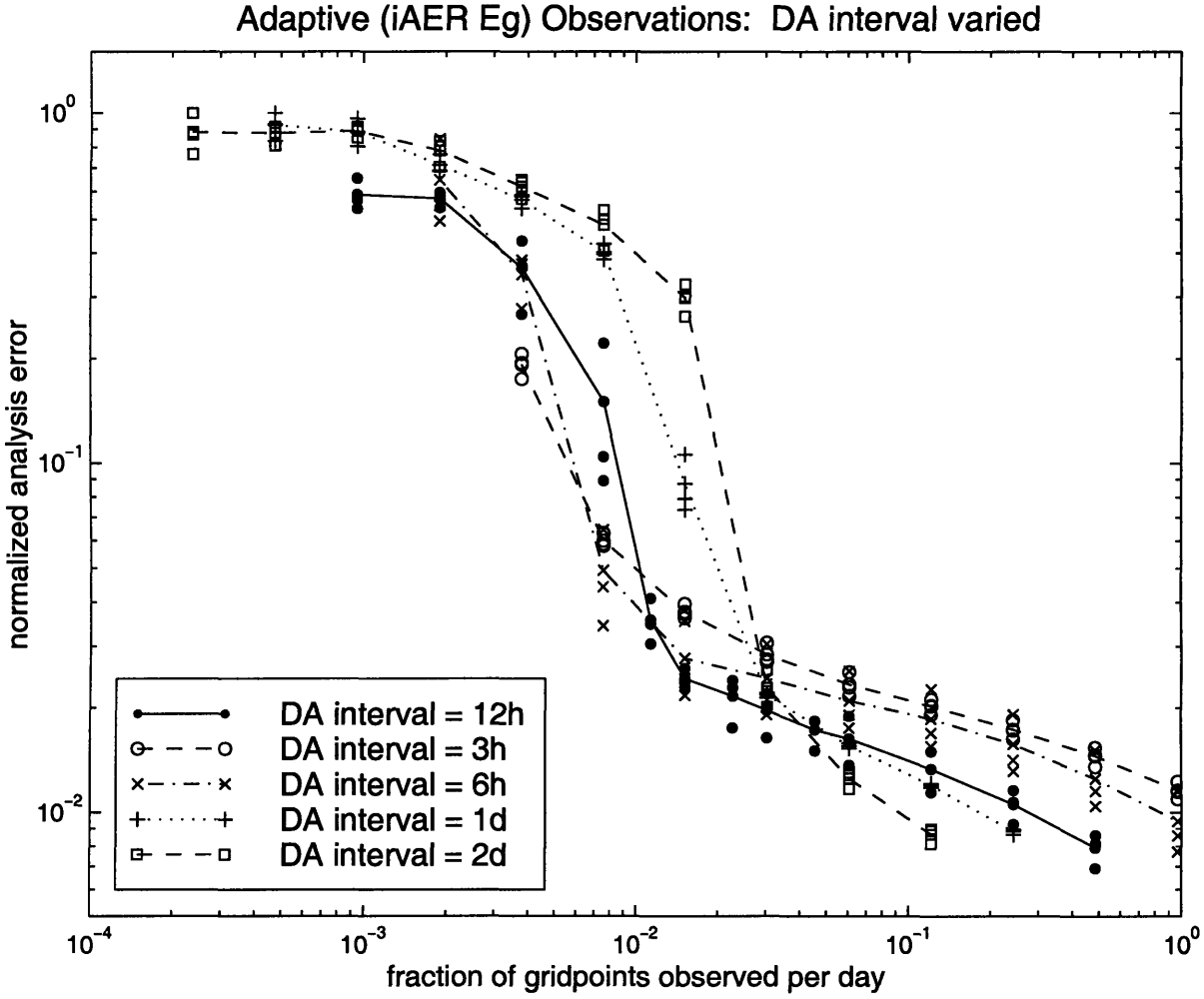



Figure 6.7: As in Figure 5.12, for ideal AER adaptive observations. The x-axis represents observation density in space and time, i.e. allocation of a fixed amount of observational resources.

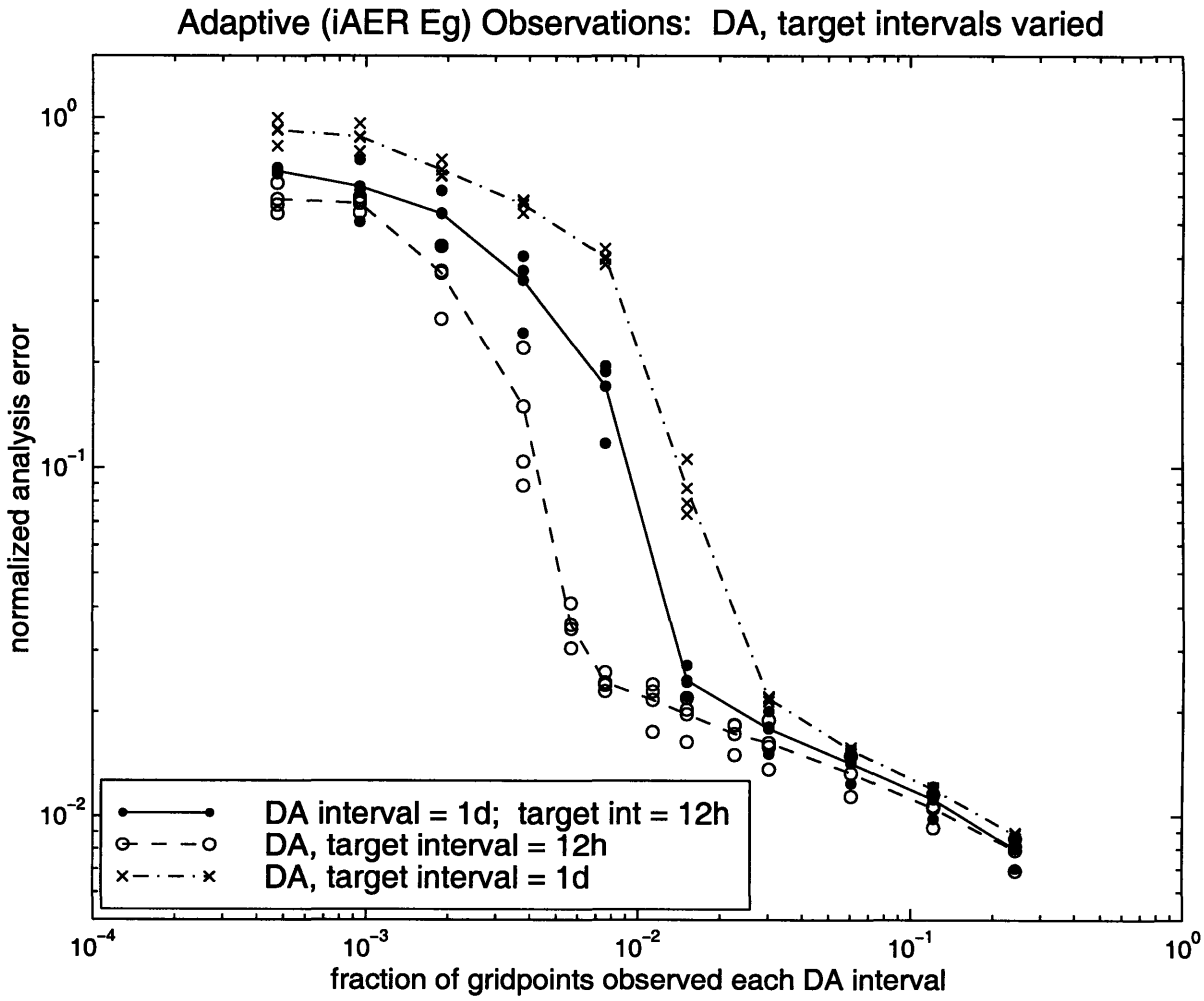



Figure 6.8: Average analysis error vs. *spatial* observation density at each data assimilation time, for ideal AER adaptive observations targeted and taken every 12 hours, targeted and taken every 24 hours, and targeted every 24 hours but taken every 12 hours (at the same locations two times in a row).

potentially important regions can be selected more frequently. Taking observations at several different times in regions with large background errors gives the data assimilation system more information to help correct the initial conditions, similar to the clustering of observations in space described in Section 6.2. As we learned from fixed observations, however, because atmospheric systems move and because errors evolve and propagate, repeating locations is not enough. More frequent targeting is also beneficial, since it gives the adaptive strategy the opportunity to select the most important locations, those not just with analysis errors but with non-decaying analysis errors, more often.

This adaptive strategy selects observation locations in regions with large background errors and lets the data assimilation system correct the errors as well as it can. After the forecast model integrates the new analysis to the next data assimilation time, the strategy evaluates if large background errors have remained in the same locations or reappeared nearby, or if larger errors have developed in other regions. It can then choose whether it should observe in the same locations or in different ones. Without knowing specifically what the data taken will be, how the data assimilation system will incorporate data into the model, and how the forecast model will integrate the resulting analysis increment forward in time, the adaptive strategy tested in this study, and perhaps any feasible realistic strategy, cannot know whether analysis or forecast errors will actually be reduced once an observation is taken. With more frequent targeting, then, this strategy can better identify the analysis errors that are reasonably large and have been growing or have persisted in the recent past, whether it is because there have been no observations in the region or because the observations have not reduced the important errors.

Figure 6.9 shows the observation locations selected during a 3 day sequence of an experiment with 2 single ideal AER adaptive observations targeted and assimilated every 3 hours. Even without including the synoptic situation or the evolution of the analysis and forecast errors, Figure 6.9 depicts how the adaptive strategy continually changes how it allocates observation locations. In some situations the adaptive strategy clusters the observations in space, selecting two locations near each other (in any pattern), and in some situations it observes in two distinct regions. In some situations the adaptive strategy clusters the observations in time, observing at or very close to the same location at two or more consecutive data assimilation times, and in some situations it selects different regions at consecutive times. Combining clustering in space and time, there are many possibilities just within the few examples shown. For example, the adaptive strategy may repeat an observation in one region while moving the other observation

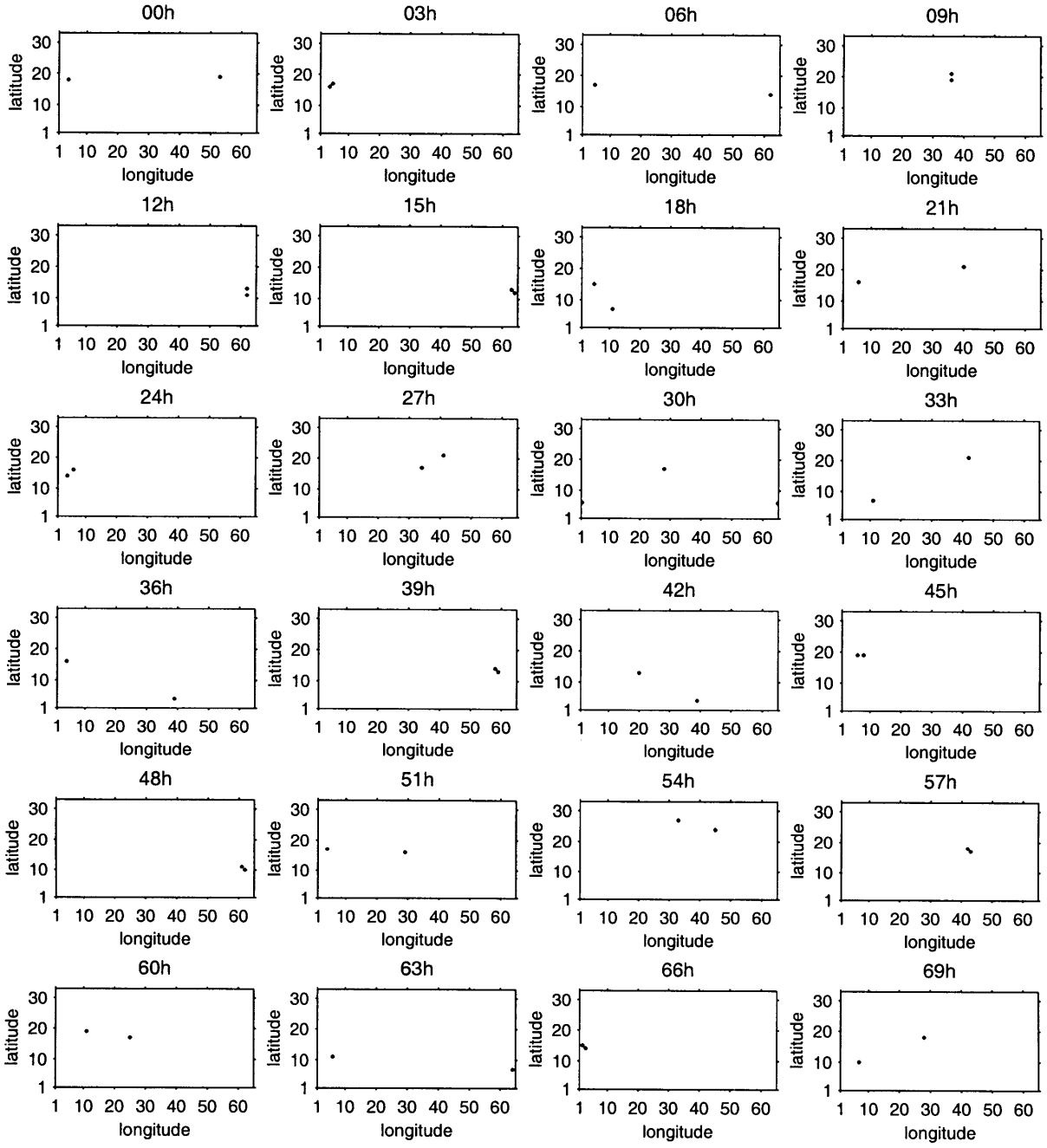



Figure 6.9: Sample sequence of adaptive observation locations selected at 3 hour intervals over the course of 3 days, from an experiment with 2 single ideal AER adaptive observations assimilated every 3 hours.

around, as in hours 0-6 and hours 18-24. Or, the adaptive strategy may observe in one region, switch to another region, then return to near the first location, as in hours 36-42 and hours 63-69. How the locations are selected in each situation depends on the relative magnitudes of the background errors in different locations, how well the data assimilation system corrected the previous errors, and how the resulting analysis increments have since evolved. Even in a system with perfect knowledge such as this one, it would be difficult to predict the sequence of locations selected by the adaptive strategy.

Along with the results for spatial clustering of observations discussed in Section 6.2, the results in this section suggest that the adaptive strategy usually reduces errors the most if it is allowed to determine how to allocate the observations in both space and time according to the specific situation. Certainly, less frequent adaptive strategies might be more effective if, incorporating more information from adjoints or one of the other proposed methods that deem to determine which errors are growing rapidly, they knew more about which analysis errors were likely to be important in the near future. Adding information about the data assimilation procedure to the strategy and improving the data assimilation system may also help adaptive observations reduce errors when taken less often. However, since both data assimilation and the criteria for selecting observations locations are statistical procedures in the real atmosphere, neither a perfect data assimilation nor a perfect adaptive strategy exists. The risk that observations will degrade or not influence a forecast will therefore always have to be considered when assimilating observations and when developing and choosing adaptive strategies.

There are many complications not included in this study, such as model errors, more complex dynamics, and practical limitations of targeting with imperfect information and imperfect observation platforms, which limit the effectiveness of adaptive observations and of data assimilation in the real world. If these are taken into account, the chance that observations will not improve a forecast grows, and an adaptive strategy is even more likely to perform best when taking observations more frequently. More frequent observations might also be especially beneficial with a data assimilation scheme which attempts to fit the initial conditions to the observations over a specified time period, as the model state evolves dynamically, such as a four-dimensional variational data assimilation system.

Unfortunately, very frequent observations are impractical in the real world both because many observation platforms require advance notice for targeting and because data assimilation can be computationally very expensive. Therefore, when adapting

observations in the real world, there will have to be some tradeoff between the costs and benefits of more frequent targeting. It is also important to remember that, as for clustered observations, these results are only valid on an average basis and are not necessarily the best way to improve any individual forecast. Analysis errors and specific forecast errors are certainly linked on average (as discussed in Section 5.2), but they are not interchangeable in any realistic system (as we will demonstrate in Section 8.2). In specific cases, the best way to allocate adaptive observations in space and time is likely to depend on both the forecast situation and the observational resources available.

## 6.4 Double model resolution results

To ensure that the results presented in this chapter are not due to the somewhat slow error growth or the limited resolution of the QG model, we have also performed some OSSE's with the forecast model resolution doubled (125 km horizontal resolution with 8 vertical levels). As discussed in Chapter 2, doubling the QG model resolution approximately halves the error doubling time, to 1–2 days. The advection time scale remains the same, approximately 5 hours. The data are assimilated with the same 3DVAR, but with the statistics optimized at double resolution (shown in Table 3.2 and Figure 3.2) and with the convergence criterion set to 5% of the initial maximum residual instead of 2%. The parameters for the experimental setup are the same as for the standard resolution, except the statistics are accumulated over a shorter time period as described in Section 4.1.

The observation strategies are the same as those discussed in Section 4.2, but with slightly different spacing: for 100 or fewer observations, the observations are constrained to be more than 3 gridpoints (375km) apart; for 101–500 observations, more than 2 gridpoints (250km) apart; for 501–2000 observations, more than 1 gridpoint (125km) apart; and for more than 2000 observations, the spacing is not constrained. Again, this is to prevent the adaptive strategy from clustering the observations too close together. Even with the weaker data assimilation convergence criterion and the shorter run time, the double resolution runs are still much more computationally expensive than the standard runs, for both the forecast model integration and the data assimilation. Consequently, at double resolution we show results for only 12 hour and 6 hour data assimilation intervals.

Figure 6.10 corresponds to Figure 6.1, comparing the same three strategies (fixed, random, and ideal AER adaptive) for a 12 hour data assimilation interval but for double

model resolution. Figure 6.11 compares the strategies at double resolution for a 6 hour data assimilation interval. Compared to the standard resolution results, the double resolution examples shown have a less distinct transition regime, a larger separation between the fixed and random strategy results for a 12 hour data assimilation interval, and a small but possibly significant benefit from adapting observations even in the dense observation regime. However, the double resolution experiments have not only doubled error growth, but also different observation spacing, different data assimilation parameters, and a weaker convergence criteria for the data assimilation. As discussed in Section 5.3, these types of changes can affect the specific shape of the error reduction vs. observation density curve. Therefore, the standard and double resolution results cannot be strictly compared, and rather than addressing these specific features we focus on the more general results.

Although the transition and data dense regimes are less distinct at double resolution than at single resolution, above a certain observation density, taking observations at more locations still provides a smaller benefit. Comparing Figures 6.10 and 6.11, we see that above a certain observation density there is also little or no benefit from taking more frequent observations. Thus, we again focus on the non-dense observation regimes, where adding and adapting observations is most likely to have an influence.

Comparisons between the random and fixed strategies are similar at double resolution to comparisons at the standard resolution. At both resolutions, the random strategy is more of an improvement over the fixed strategy for a 6 hour data assimilation interval than for a 12 hour data assimilation interval. If the fixed and random strategy results from Figures 6.10 and 6.11 are compared for a constant amount of observational resources (as in Figures 5.12 and 6.6 at the standard resolution), we again find that random observations have no preference for allocation in space and time, while observations at fixed locations are redundant when taken more frequently than every 12 hours. The time scale in the comparison between fixed and random observations is also similar at both resolutions, in other words the comparison is similar for the two resolutions at the same data assimilation interval. Therefore, we infer that the advection time scale (which remains the same) is more important than the error growth time scale (which is halved at double resolution) for determining the redundancy of fixed observations, i.e. for determining the benefit from simply moving observations. This makes sense intuitively, since the advection time scale is shorter than the error growth time scale and thus is likely to dominate how quickly the flow changes at a fixed location.

The behavior of the adaptive strategy at double resolution is also qualitatively similar

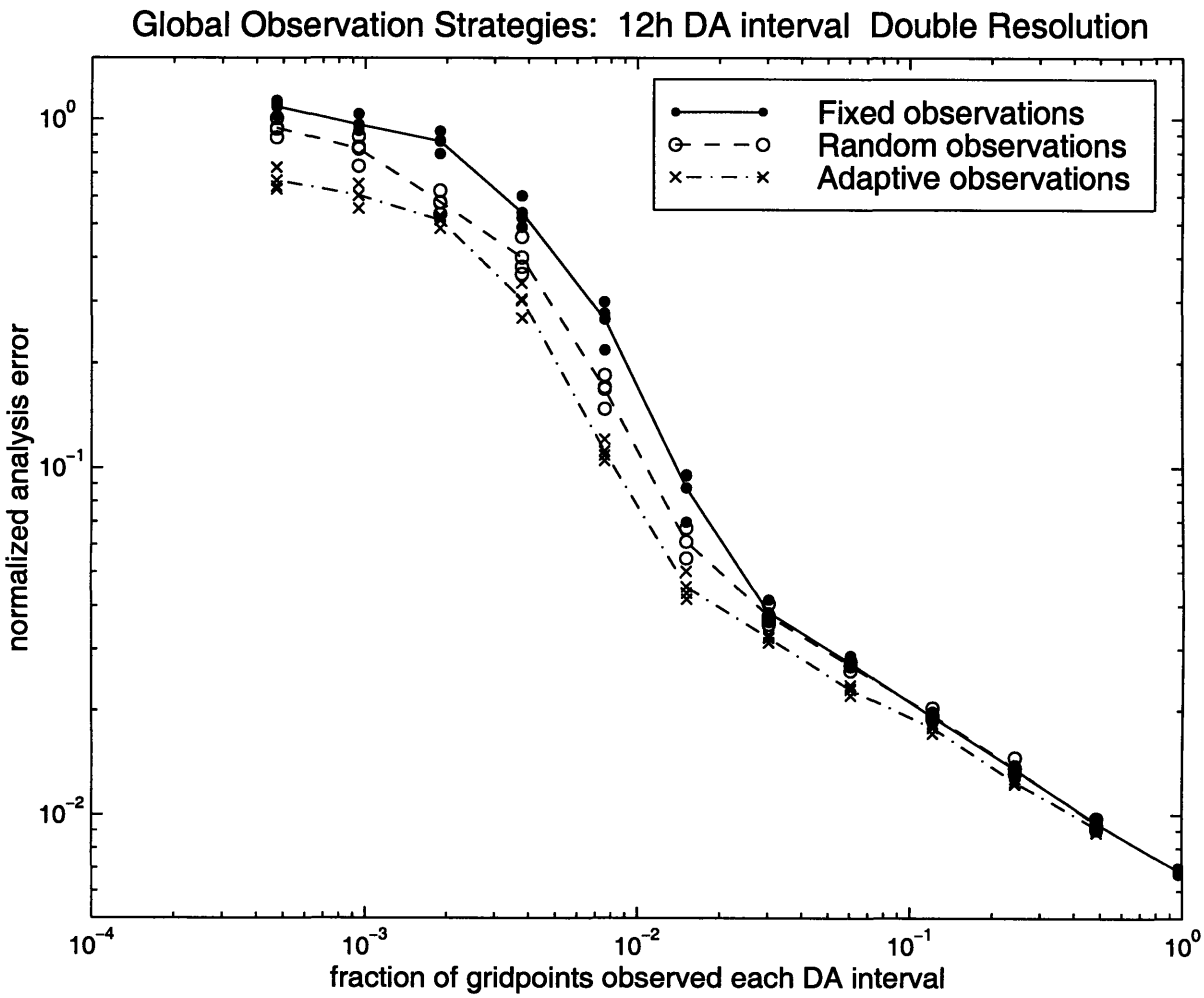



Figure 6.10: As in Figure 6.1, for the QG model at double resolution and a 12 hour data assimilation interval.

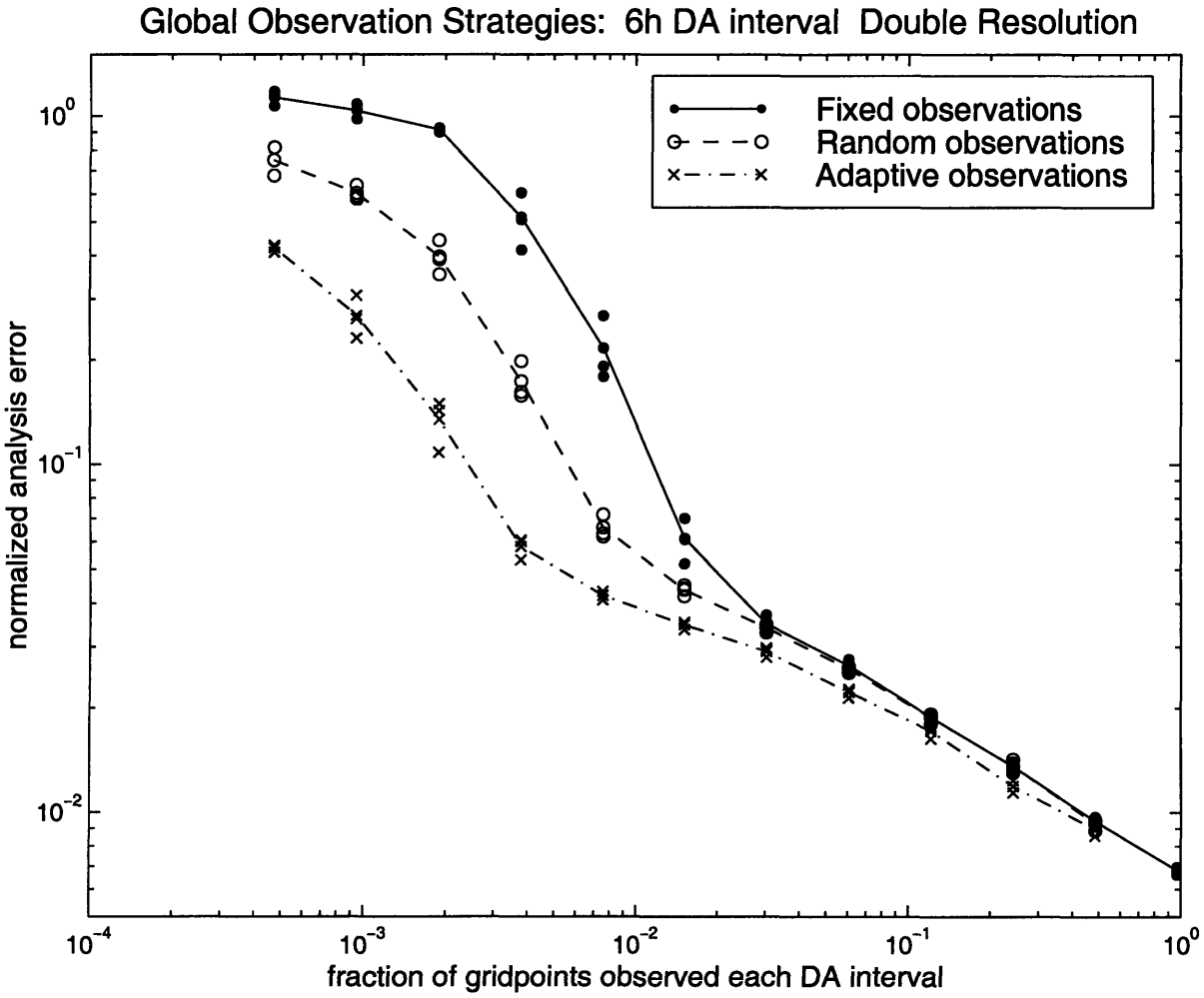



Figure 6.11: As in Figure 6.10, for a 6 hour data assimilation interval.

to its behavior at single resolution. At both resolutions, the adaptive strategy is more of an improvement over the random strategy when the data assimilation interval is shorter. Comparing the double resolution adaptive results for a constant amount of observational resources, as at standard resolution in Figure 6.7, the adaptive strategy performs best when a few observations are taken more frequently. However, at double resolution, twice as frequent targeting and data assimilation is necessary to obtain approximately the same benefit from adaptive observations. In other words, the improvement of adaptive observations over random observations at double resolution for a given data assimilation interval (e.g. 6 hours) is most similar to the improvement at single resolution for twice the data assimilation interval (e.g. 12 hours). This suggests that the error growth time scale, more than the advection time scale, controls the effectiveness of ideal AER adaptive observations compared to random observations. When evaluating how effective adaptive observations may be compared to fixed observations in a certain forecast system or in a certain situation, then, both the advection and error growth time scales are important, but for different reasons.

Most importantly, in both Figures 6.10 and 6.11, for non-dense observations the adaptive strategy performs better on average than the random strategy. This indicates that even with higher than the standard quasi-geostrophic model resolution (and the associated faster error growth), it is possible, in the appropriate circumstances in this simplified system, to reduce average analysis errors by objectively adapting observation networks.

# Chapter 7

# Adaptive sampling using ensemble spread to estimate background error

The ideal analysis error reduction adaptive strategy is a useful tool for exploring a large number of adaptive observation configurations. Since the actual error in the background field is not known in the real atmosphere, however, this idealized strategy is impossible to implement in any real atmospheric situation. Therefore, we have also tested a more realizable approximation to the ideal AER adaptive strategy, which we call the estimated analysis error reduction (AER) adaptive strategy. The estimated AER adaptive strategy estimates the background error from the spread of an ensemble of perturbed forecasts, then targets observations at the locations where the ensemble spread is the largest. It is similar to the ensemble-based strategies tested in Lorenz and Emanuel (1998) and is described in further detail in Section 4.2.3. Because the ensemble is an imperfect estimate of background error, we do not expect the estimated AER adaptive strategy to identify the locations with large background errors as well as the idealized strategy identifies them. The estimated strategy will therefore most likely, but not necessarily, be less effective than the idealized strategy.

As in Chapter 6, in this chapter we compare results for global observation strategies. Results for observations added to a pre-existing fixed observation network are discussed in Chapter 8. Observations are assimilated every 12 hours, again using the standard 3DVAR data assimilation system described in Chapter 3. First, we test selecting the observation locations only 12 hours before they are taken, based on the spread of an ensemble of 12 hour forecasts generated at the data assimilation time prior to the observation time; we call this targeting observations with a 12 hour lead time. Results with

observation locations selected more than 12 hours in advance are discussed in Section 7.4.

The ensemble used here is generated by assimilating perturbed observations into each of the ensemble members at each data assimilation time, simulating errors propagating through an analysis and forecast cycle. This is called a multiple replication ensemble by Lorenz and Emanuel (1998) and an OSSE-Monte Carlo ensemble by Houtekamer and Derome (1995); further details on the ensemble are given in Section 4.2.3. For the results shown here, the ensemble consists of the control run and several perturbed runs. The ensemble spread is defined as the average difference between the ensemble members and the mean of the ensemble, with an energy norm (the same norm as for the background error in the ideal AER strategy results shown). We chose to define the ensemble spread as the average of the ensemble perturbations instead of as the root-mean-square average of the perturbations because the perturbation energy is already a squared, semi-positive definite, quantity. Several other definitions of ensemble spread have been tested and are discussed. In the first two sections, the ensemble has 13 members (12 perturbed trajectories in addition to the control); results for different ensemble sizes are shown in Section 7.3.

## 7.1 Comparison with global fixed, random, and ideal adaptive strategies

For a single observation at each targeted location, the estimated AER adaptive strategy produces the same results as the random strategy results shown in Figure 6.1. For clustered observations (described in Section 6.2), on the other hand, the estimated AER adaptive strategy is more effective than the random strategy. Figure 7.1 shows the average error in the control analysis as a function of observation density for random, estimated AER adaptive, and ideal AER adaptive clustered observations; in this example, we use cluster 3 (depicted in Figure 6.4), taking a triangle of 3 observations around each targeted location. Therefore, when implemented globally, the estimated AER adaptive strategy is on average an improvement over the random strategy only if a cluster of observations is taken at each targeted location. Even for clustered observations, however, the estimated AER adaptive strategy is less effective than the ideal AER adaptive strategy. These results suggest that, as discussed in Section 6.2, on average the estimated strategy identifies locations near (but not exactly at) locations at which observations will improve the analyses and forecasts.

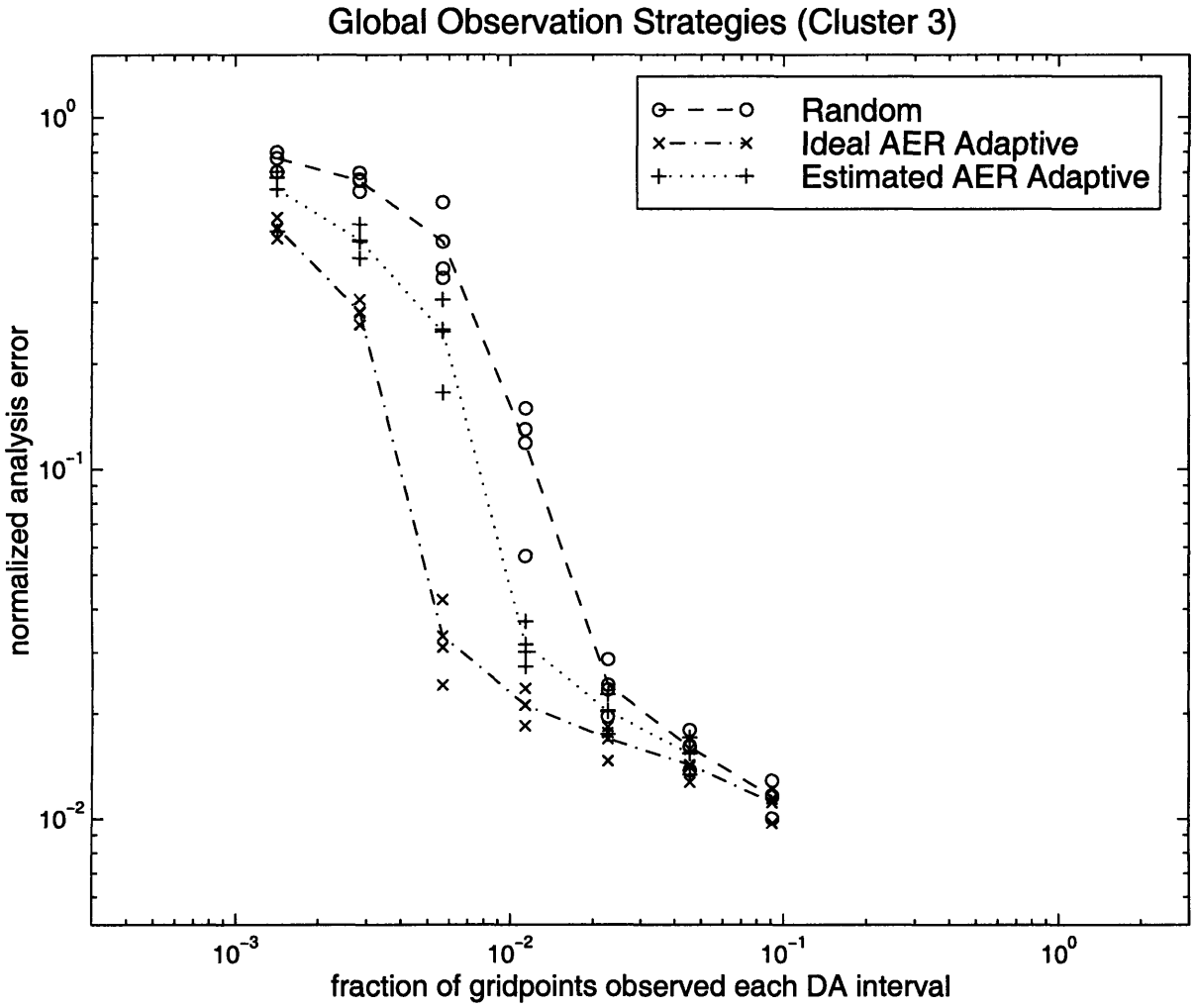



Figure 7.1: Analysis error reduction as a function of observation density for random, ideal AER adaptive, and estimated AER adaptive observations, with cluster 3 observations taken for each targeted location for all of the strategies. The observation and data assimilation intervals are 12 hours. Cluster 3 is shown in Figure 6.4. For the estimated strategy, the ensemble has 13 members (12 perturbed forecasts in addition to the control). The x-axis is normalized both by the number of gridpoints (2112) and by the number of observations taken for each targeted location (3); however, the x-axis is also shifted so that the maximum observation density is 100% $\times$ the number of observations taken for each targeted location.

In this study, the control forecast is included in the ensemble mean and the ensemble spread. We have experimented both with calculating the ensemble perturbations with respect to the control forecast and with not including the control forecast in the ensemble. These different definitions of the ensemble produce estimated AER adaptive results similar to those shown, although in some cases slightly less positive.

For the results shown, the ensemble spread is measured in terms of energy averaged at all levels over all the ensemble perturbations. We have also tested ensemble-based targeting with potential vorticity and streamfunction norms, defining the ensemble spread as the root-mean-square average of the ensemble perturbations (instead of the average). As with the different background error norms tested for the ideal AER strategy (discussed in Chapter 6), using a potential vorticity norm for estimated AER adaptive observations produces similar results to an energy norm, while a streamfunction norm is less effective. In addition, based on the results in Buizza and Palmer (1998), we tested an extremum ($L^{\infty}$) energy norm. The extremum norm measures the ensemble spread by taking the maximum value among the ensemble perturbations at each location (instead of the average of the ensemble perturbations at each location). Results with the extremum norm are similar to those shown and are discussed further in Sections 7.2 and 7.3.

In the estimated AER adaptive strategy tested here, we use the ensemble only to estimate background error at the observation time. Ensembles of perturbed forecasts contain information besides an estimate of background error, and the background error is not the only criterion important for adapting observations. Therefore, we could almost certainly use ensembles to develop a more sophisticated, more effective adaptive observation strategy. For example, future error growth can be estimated using the time evolution of the ensemble, as in the ensemble transform technique proposed by Bishop and Toth (1999). In this study, we do not explore how one might optimally use ensembles; we only test if ensembles might be used to approximate the ideal AER adaptive strategy in a more realistic setting. Even though the estimated AER strategy incorporates limited information from the ensemble, Figure 7.1 demonstrates that at least in this idealized system, and again with the appropriate observation configuration, it is possible to use an ensemble to estimate background error effectively for adapting observations. We have also only tested verifying the control analyses and forecasts; if we were to measure the influence of the observations by verifying probabilistic information from the ensemble, such as the ensemble mean or the ensemble spread, we might see even more of an effect from adapting observations. As a first attempt at ensemble-based

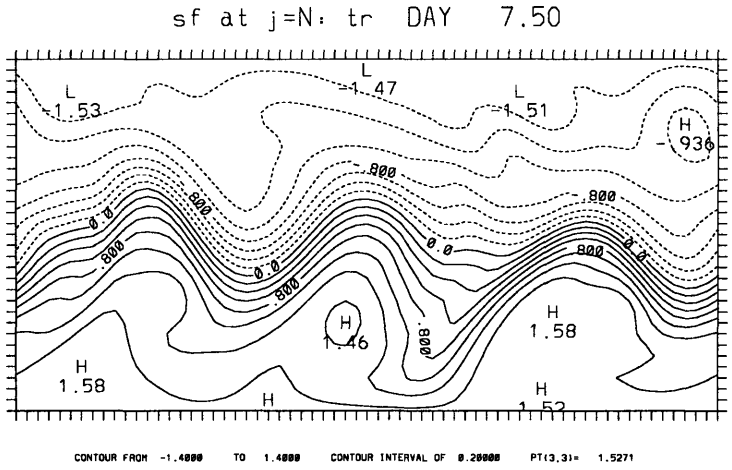



Figure 7.2: QG model truth state at the time at which Figures 7.4, 7.5, 7.9, and 7.10 are valid. Streamfunction is shown at the upper interior level.

targeting in the system, therefore, the results shown in Figure 7.1 are encouraging.

## 7.2 Ensemble spread as an estimate of background error

To understand how the estimated AER adaptive strategy approximates the idealized strategy, next we compare the ensemble spread and the background error on a case-by-case basis. As discussed in the previous section, the ensemble spread is only evaluated according to how well it estimates background error and how useful this background error estimate is for adapting observations. In order to avoid biasing the comparisons by using one of the two adaptive strategies, we spun up a control forecast and an ensemble with observations at 16 fixed locations (1.2% of gridpoints), without taking any adaptive observations. Figure 7.2 depicts the synoptic situation at a sample observation time, and Figure 7.3 depicts the 16 fixed locations where observations have been taken every 12 hours during the spin up period. Figure 7.4 shows a sample background error (equivalent to the error in the 12 hour control forecast) and a sample ensemble spread (for a 7 member ensemble), both valid at the same observation time as Figure 7.2 and generated with a 12 hour lead time. The error in the ensemble mean 12 hour forecast valid at the same time (shown in Figure 7.5 for a 7 member ensemble, verified versus the truth state) is similar to the control forecast error. The fields are shown for energy at the upper interior level; at lower levels, the background errors tend to be less localized, and comparisons

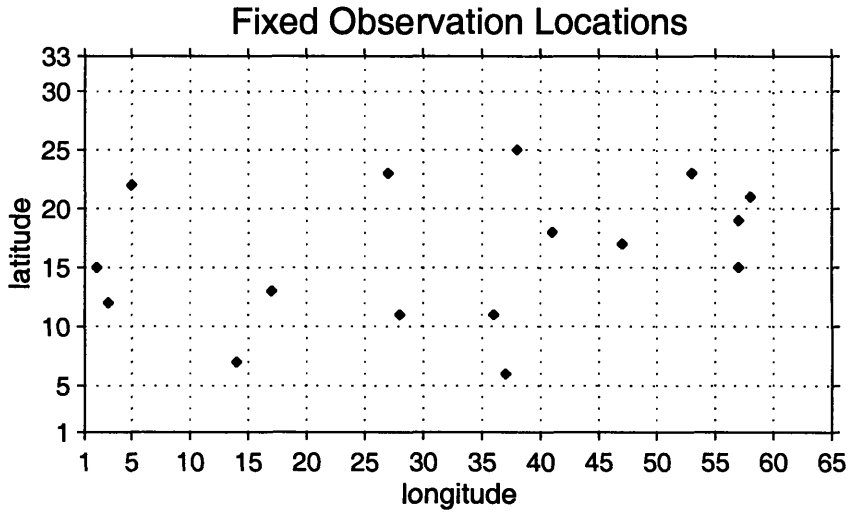



Figure 7.3: The 16 fixed observation locations for the results shown in Figures 7.4, 7.5, 7.9, and 7.10.

look similar or slightly less favorable. Although the ensemble behaves differently in each situation, Figure 7.4 demonstrates the most common basic features of how ensemble spread estimates background error in this simulated system.

Large forecast errors tend to be associated with the few synoptic systems in the domain, and in these dynamically active regions the ensemble members diverge, producing large spread. Therefore, the ensemble spread in Figure 7.4 is large in many of the regions in which the actual background errors are large. Despite identifying some of the same general features, the ensemble spread estimates background error imperfectly in several ways. First, because the control forecast has errors related to its specific initial conditions, the actual errors may be large in only one part of a region identified by the ensemble spread as likely to have errors. The ensemble spread in Figure 7.4 thus tends to have less small scale structure than the background error field and to be less focussed on specific subregions. Based on the ensemble spread, the estimated AER strategy then at times identifies locations close to but not exactly at the locations with large background errors. In extreme cases, the ensemble dispersion can be large but the control forecast can, by chance, still be close to the truth state in the whole region (Buizza and Palmer 1998; Z. Toth, personal communication). Consequently, the ensemble spread also sometimes identifies regions where there is only a very small background error, as in the left half of the domain in Figure 7.4.

A related issue is that although the ensemble spread and the background error often identify several similar main regions, they often prioritize the regions differently. In

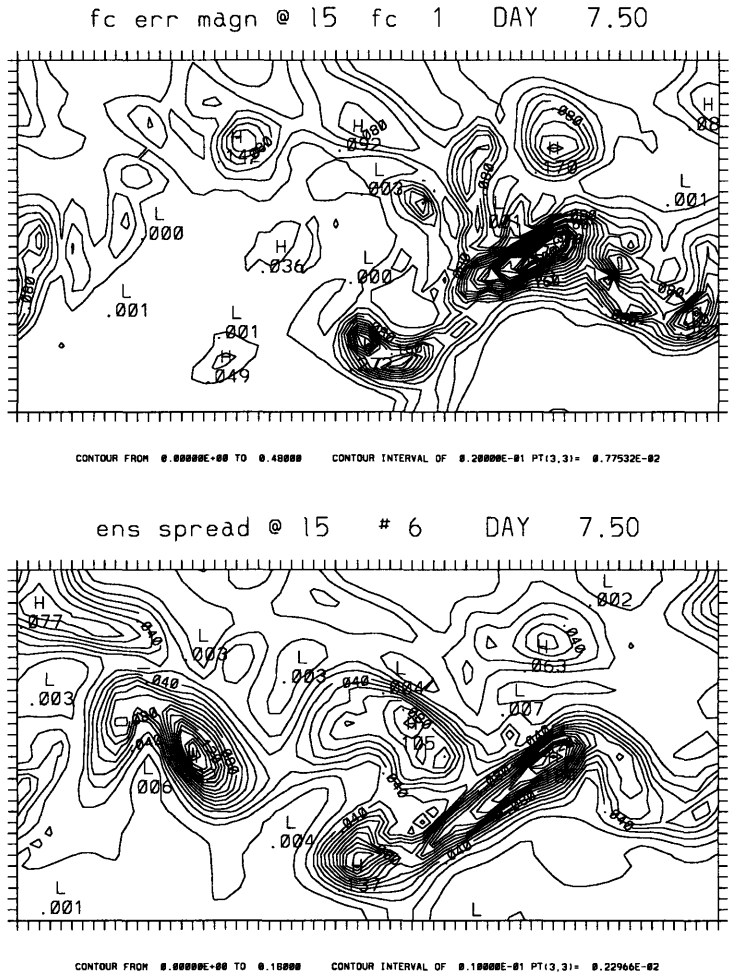



Figure 7.4: (a) Background (12 hour control forecast) error and (b) spread of an ensemble of 12 hour forecasts, both valid for the same sample observation time. Both were generated with a 12 hour lead time, and both are shown for an energy norm at the upper interior level. The QG model truth state at the observation time is shown in Figure 7.2. The control and ensemble trajectories were spun up with the 16 fixed observation locations shown in Figure 7.3. The ensemble has 7 members (6 perturbed members in addition to the control). The x-axis is longitude (periodic) and the y-axis is latitude, and the contour interval is .02 in (a) and .01 in (b).

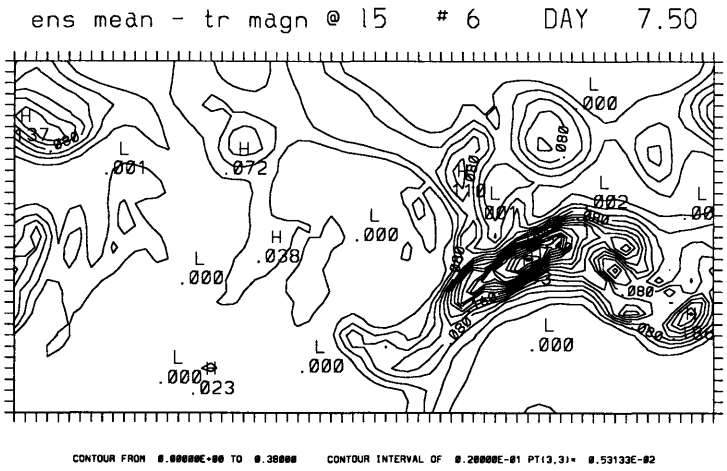



Figure 7.5: Ensemble mean 12 hour forecast error, valid at the same sample observation time for the same control forecast error as in Figure 7.4. The ensemble has 7 members. The norm is energy at the upper interior level, and the contour interval is .02.

the case shown, for example, from the ensemble spread we would be able to identify the region with the largest background error, but we might find it difficult to select the most important region for extra observations among the next few regions identified by the ensemble. This problem is common and is usually not fixed by increasing the ensemble size (as we will see in Section 7.3). At times, there are also regions with small ensemble spread but non-negligible background errors, in other words regions where the ensemble spread underestimates background error. Finally, in some cases a region of large ensemble spread is somewhat offset from a region with large background errors. This occurs in the upper left hand corner of the domain in Figure 7.4.

Because ensemble forecasting is a statistical procedure, the first few ways in which the ensemble spread is an imperfect predictor of background error are, in some sense, inherent to and expected in ensemble prediction. Some of the misidentification, however, may be at least partially due to our ensemble configuration or our ensemble spread norm. Although which ensemble techniques are the best for different purposes is an area of active debate (see for example Anderson 1996, Anderson 1997, Houtekamer and Derome 1995, Mureau et al. 1993, and Toth and Kalnay 1997), results in Lorenz and Emanuel (1998) and Hamill et al. (1999) suggest that the multiple replication ensemble is a good choice for estimating forecast error statistics. Nevertheless, because we have not tested the statistics of our ensemble extensively to ensure that they match the statistics of the analysis and forecast errors, even with the appropriate generation technique our ensemble is still probably not ideal.

For any ensemble technique, how we choose to measure the ensemble spread can also be important. Buizza and Palmer (1998), for example, demonstrate that measuring the ensemble spread with an extremum ($L^\infty$) norm instead of with an average norm minimizes the number of times when the ensemble spread underestimates forecast error, especially for a large ensemble. Although we have tested several different norms and found on average only small differences (further discussed in Section 7.3, we have not explored how to measure the ensemble spread in great detail. In the process of evaluating the ensemble, however, we realized that to determine how to estimate background errors most effectively for adapting observation networks, we must first understand more generally which aspects of background error we really wish to estimate for adapting observations.

The data assimilation procedure handles the observations and the errors in the initial conditions differently in each situation, and the forecast model will use the resulting analysis increments with varying success. Because of the complex interactions between the background errors, the observations, the data assimilation system, and the forecast model, it is not clear which errors in the initial conditions are the most important to estimate well when adapting observations. Alternately, we might say that it is not clear which errors are the most important *not* to estimate *in*correctly. For example, would we prefer to overestimate background error, risking observing in a region with small errors? Or would we prefer to underestimate background error and risk not observing in a region with large errors? Or perhaps the most risky regions are those where the ensemble spread is offset from a region with background errors, since then the data assimilation system may be most likely to extrapolate into the region with large errors and degrade the analysis and forecast?

How to best estimate background error for adaptive observations depends fundamentally on how well the data assimilation scheme uses observations in different situations. For example, because the 3DVAR background error statistics and the 3DVAR algorithm have little knowledge of the dynamical situation, and because observation errors limit how well the data assimilation can resolve small errors, our data assimilation system performs poorly in regions with small background errors. We have performed preliminary experiments taking observations at locations with different sized background errors, and the results suggest that in this idealized system, the analysis is on average most improved when we observe in some subset of the regions with medium to large background errors. Therefore, given the limitations of the data assimilation system, we may not want our ensemble spread to significantly overestimate background error, since then we

will observe at some locations with small errors.

A more sophisticated data assimilation method will better be able to improve analyses and forecasts from observations in regions with smaller background errors. With a different data assimilation system, then, we may want slightly different information from the ensemble. However, no data assimilation system is perfect. In realistic atmospheric prediction, there are many more difficulties than in this idealized system: ensembles and forecast models are imperfect, observations have errors, and actual background errors are not known. Therefore, when selecting adaptive observation locations, we will always have to consider how the limitations of the observations and of the data assimilation system may interact with the limitations of the background error estimate.

## 7.3 Sensitivity to ensemble size

Next, we investigate how large an ensemble is required to estimate background error well enough to adapt observations effectively. To make testing different ensemble sizes computationally feasible, we have selected, for single observations and for each of the observation cluster patterns, the observation density at which the random strategy results differ the most from the ideal AER adaptive strategy results, i.e. the observation density at which adapting observations based on background error produces the most improvement. Then, we have run the experiments for the estimated AER adaptive strategy at that one observation density but varying the ensemble size. The average analysis error reduction produced by the estimated AER adaptive strategy is shown as a function of ensemble size for 12 single observations in Figure 7.6. Figure 7.7 shows the same set of experiments for 4 targeted locations and cluster 3 (3 observations in a triangle around the targeted location), and Figure 7.8 shows the same set of experiments for 2 targeted locations and cluster 13 (13 observations at alternating gridpoints in a 1000 km $\times$ 1000 km area around the targeted location); both clusters are depicted in Figure 6.4. For comparison, the random and ideal AER adaptive strategy results are also shown in each figure for the same observation density and the same cluster pattern. We have performed the same experiments for average error reduction in forecasts up to 3 days, and the comparisons between different ensemble sizes are similar to those shown for analysis error. The ensemble spread norm is again average energy at all levels; results for the other targeting norms tested are similar.

In Figure 7.6 we see that, as described earlier in this chapter, for single observations

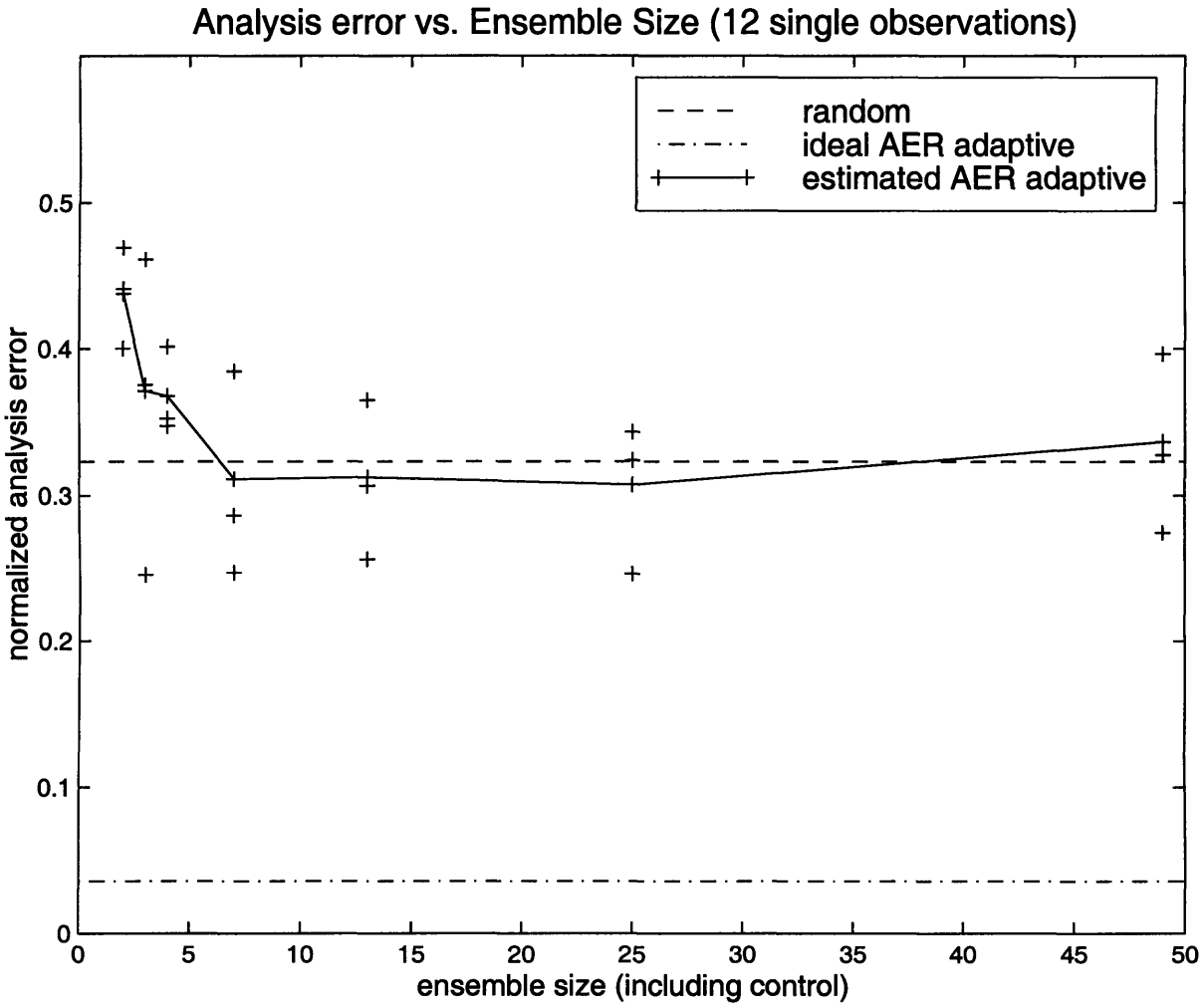



Figure 7.6: Analysis error reduction as a function of ensemble size (including the control forecast) for the estimated AER adaptive strategy with single observations at 12 targeted locations (.57% of gridpoints observed). For comparison, the random strategy and ideal AER adaptive strategy results are also shown for the same observation density.

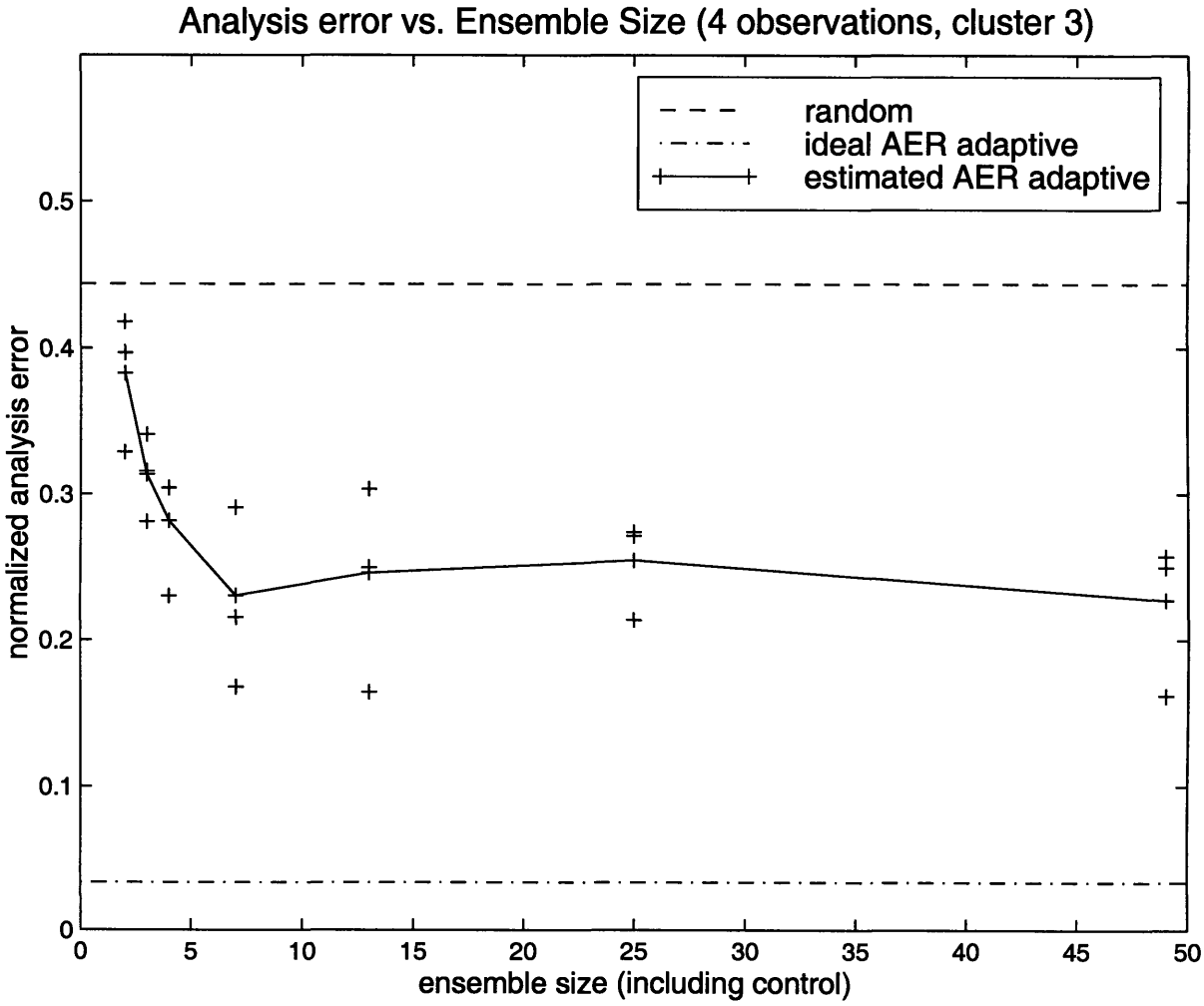



Figure 7.7: As in Figure 7.6, but with cluster 3 observations at 4 targeted locations (.57% of gridpoints observed). The random and ideal AER adaptive strategy results are shown for the same observation density and the same cluster pattern. The 13 member ensemble results are the same as those for the appropriate observation density in Figure 7.1.

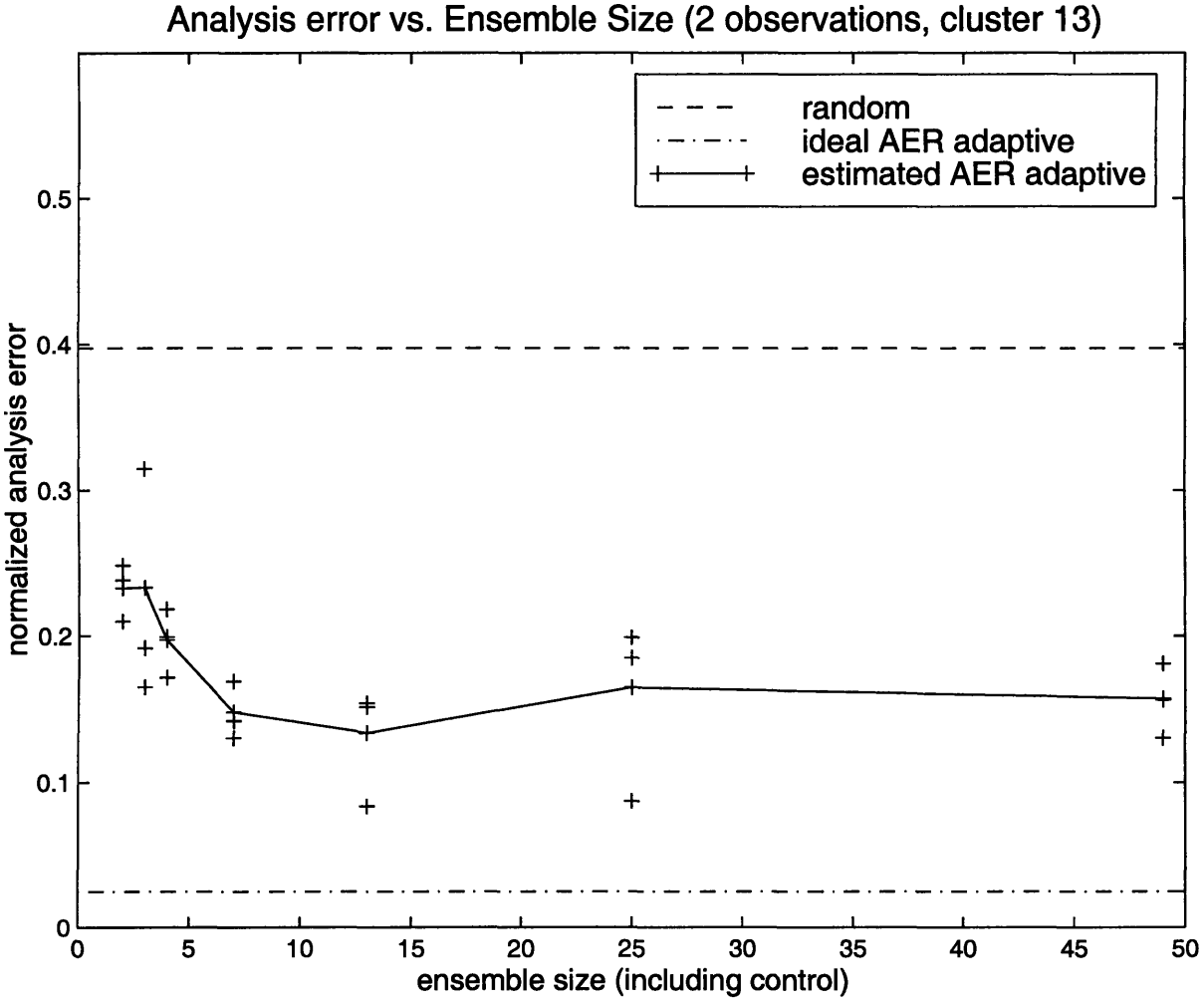



Figure 7.8: As in Figure 7.6, but with cluster 13 observations at 2 targeted locations (1.2% of gridpoints observed). The random and ideal AER adaptive strategy results are shown for the same observation density and the same cluster pattern.

the estimated AER adaptive strategy tested here is no more effective than the random strategy. This is true even in the limit of large ensemble size. For cluster 3 observations, Figure 7.7 indicates that to produce most of the average analysis error improvement possible for the estimated AER adaptive strategy, we require only 6 perturbed forecasts (7 ensemble members). For cluster 13 observations, Figure 7.8 shows that large clusters of estimated AER adaptive observations, in addition to benefiting little from more than 6 perturbed forecasts, are also on average able to reduce the analysis error by about 40% with only one or two perturbed forecasts. The differences between 7 member and larger ensembles in Figures 7.6–7.8 are not statistically significant, given the large spread among the different runs and the limited number of cases sampled. Therefore, for all of the ensemble and observation configurations tested, in this simulated system adding more than 6 perturbed forecasts benefits the estimated AER adaptive strategy little or not at all.

Buizza and Palmer (1998) find that large ensembles are only beneficial if an extremum norm is used for the ensemble spread rather than an average norm. Our ensemble also sometimes underestimates background error (see Section 7.2), which an extremum spread norm can help correct. With an extremum norm for the ensemble spread, however, the estimated AER adaptive observations are no more effective, and ensembles with more than 7 members still produce the same results. As discussed in Section 7.2, this suggests that before attempting to extract additional information from significantly larger ensembles, it is important to know what type of background error estimate is optimal for adaptive observations in the relevant prediction system.

Large ensembles do not benefit the estimated AER adaptive strategy because with only a few perturbed members, an ensemble in this idealized system is on average able to identify the few general regions with large background errors. To demonstrate this, Figure 7.9 shows, for the same situation and the same background error as in Figure 7.4, the spread of ensembles of different sizes. Recall that the ensemble includes the control forecast and that the spread is calculated with respect to the ensemble mean with an energy norm (which is a squared quantity). With only 2 members (Figure 7.9a), the ensemble spread is able to identify at least some portion of the regions with large background errors. With 4 members (Figure 7.9b), the ensemble spread captures most of the regions that the 7 member ensemble (Figure 7.4b) does, although it weights the regions somewhat differently. With 25 members (Figure 7.9c), the ensemble spread looks nearly identical to the 7 member ensemble spread (in fact, most of the cases studied look more identical than this one).

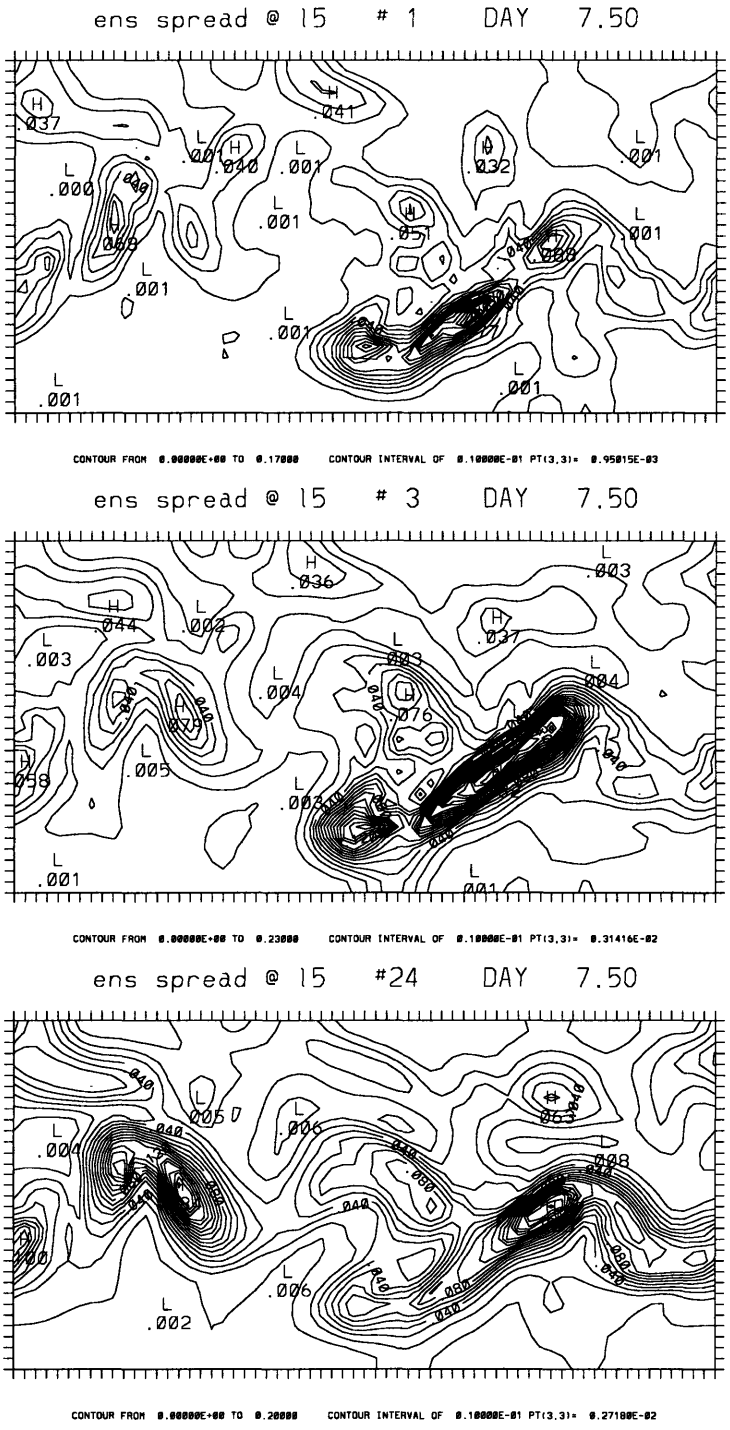



Figure 7.9: Ensemble spread valid at the same time for the same control forecast error as in Figure 7.4, for an ensemble with: (a) 2 members (1 perturbed member in addition to the control), (b) 4 members, and (c) 25 members. The corresponding spread for a 7 member ensemble is shown in Figure 7.4b. The targeting lead time is again 12 hours. The contour interval is .01 in all 3 plots.

The comparisons in Figure 7.9 help explain the lack of benefit from larger ensembles in the time-averaged estimated AER adaptive results, and they suggest that the number of dominant degrees of freedom that the ensemble can identify in this simulated system is, in an individual situation, small. The specific number of ensemble members we require, however, is likely dependent on the simplified dynamics and the simplified geometry of the QG forecast model. With a larger domain and more complex dynamics, we might expect to need a larger ensemble. These factors may be at least partially offset in more realistic systems, however, by features such as storm tracks which tend to focus the largest error growth in certain regions. We have only tested very simple uses of the ensemble; larger ensembles may also still be useful for calculating higher order moments than the ensemble spread and for identifying more specific unpredictable features.

## 7.4 Sensitivity to increased targeting lead time

In the real world, many adaptive observation platforms (particularly moving in situ platforms) require that observation locations be selected well in advance of the observation time. In recent field experiments, for example, because of aircraft constraints the adaptive observation locations generally had to be selected 36-48 hours in advance, with preliminary flight planning beginning even earlier. Advance lead time may be problematic when adapting observations for two reasons. First, adaptive observations are by definition most important in uncertain forecast situations, and it is in precisely these situations that errors in forecast models and in initial conditions will make it very difficult to plan in advance. In addition, current ensemble-based estimates of initial condition errors incorporate no information about where observations will be taken in the future. If observations are taken in important regions between the targeting time and the observation time, an ensemble will overestimate background error. There has been some concern that the current techniques for estimating probable errors are therefore not very useful for longer lead times.

All the results shown so far in this chapter are for observation locations selected with a 12 hour (or shorter) lead time. We have not evaluated in great depth how much longer lead times limit adaptive observation strategies; however, to provide a basic idea of how extendable our results are to advance targeting scenarios, we have tested both the ideal and estimated AER adaptive strategies at lead times greater than 12 hours. First, we examine the same case as in Figure 7.4, valid at the same observation time and spun up

with the same 16 fixed observation locations (and no adaptive observations), but with the background error estimated further in advance. For a 24 hour lead time (not shown), there is only one set of observations taken between targeting and taking observations, and the ensemble spread and forecast errors look qualitatively similar to those for a 12 hour lead time.

For longer lead times, even with a perfect forecast model we can begin to see the limitations of predicting background error further in advance. Figure 7.10a shows the 48 hour forecast error, and Figure 7.10b shows the 48 hour ensemble spread valid at the same time. These fields would be used to select ideal and estimated AER adaptive observations for a 48 hour lead time, compared to Figure 7.4 for a 12 hour lead time. Both the forecast error and the ensemble spread identify the same major areas for a 48 hour lead time as they did for 12 hours, but for longer lead times these few areas are much more dominant and the fields have significantly less small scale structure. This is likely both because the growth of a few errors dominates over longer time periods and because the 48 hour lead time fields do not know about the 16 fixed observations which have been taken in the interim. Despite the difficulties for the longer lead time, however, predictions of background error made 48 hours in advance still have some skill.

To see how useful these background error predictions are for adapting observations, next we examine how well the two adaptive strategies perform on average for different lead times. To simulate targeting observations more than 12 hours in advance, we have tested selecting estimated AER adaptive observations based on the spread of an ensemble of forecasts generated more than 12 hours prior to the observation time, and selecting ideal AER adaptive observations based on the actual error in the control forecast generated more than 12 hours prior. Figure 7.11 shows results from the ideal and estimated AER adaptive strategies as a function of targeting lead time, for 4 targeted locations with cluster 3 observations (the same observation density as in the ensemble size experiments in Figure 7.7). Because for longer lead times both the intervening error growth time and the intervening data inputs are important, we have plotted the lead time in terms of number of data assimilation intervals; each data assimilation interval is equivalent to 12 hours. Again, for comparison, the random strategy results are shown for the same observation density.

Although both the estimated and the ideal AER adaptive strategies are less beneficial for longer lead times, on average they still improve analyses more than random observations as long as the observations are targeted 4–8 or fewer data assimilation intervals in advance. This estimate may be somewhat optimistic, since our idealized system exhibits

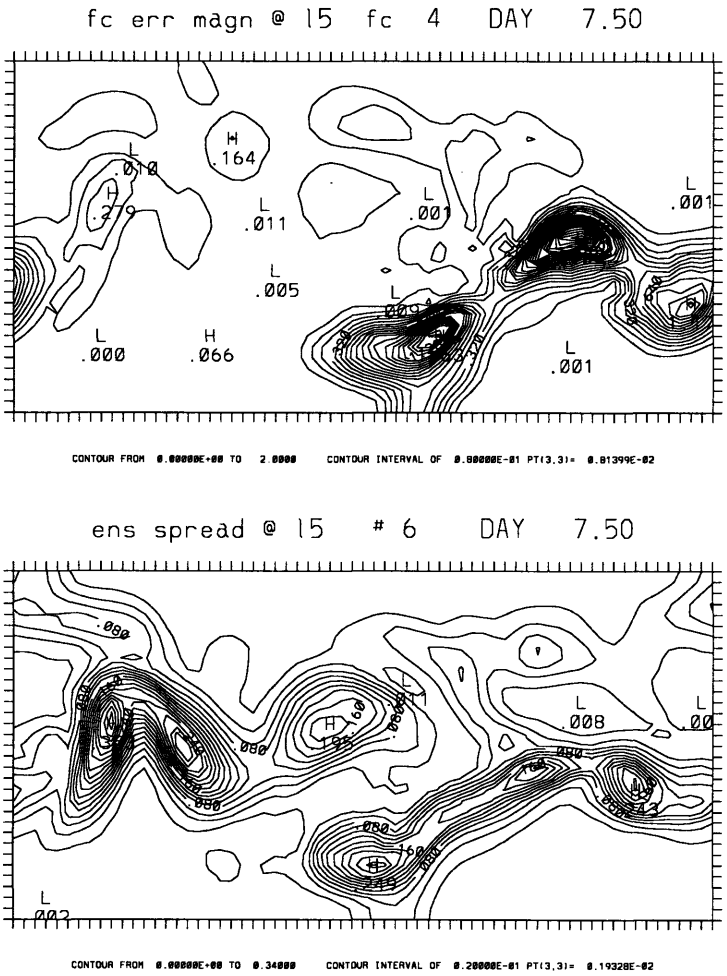



Figure 7.10: Control forecast error and 7 member ensemble spread valid at the same observation time as in Figure 7.4, but with the forecast and the ensemble spread generated 48 hours prior, i.e. with a 4 data assimilation interval lead time for targeting observations. The contour interval is .08 in (a) and .02 in (b).

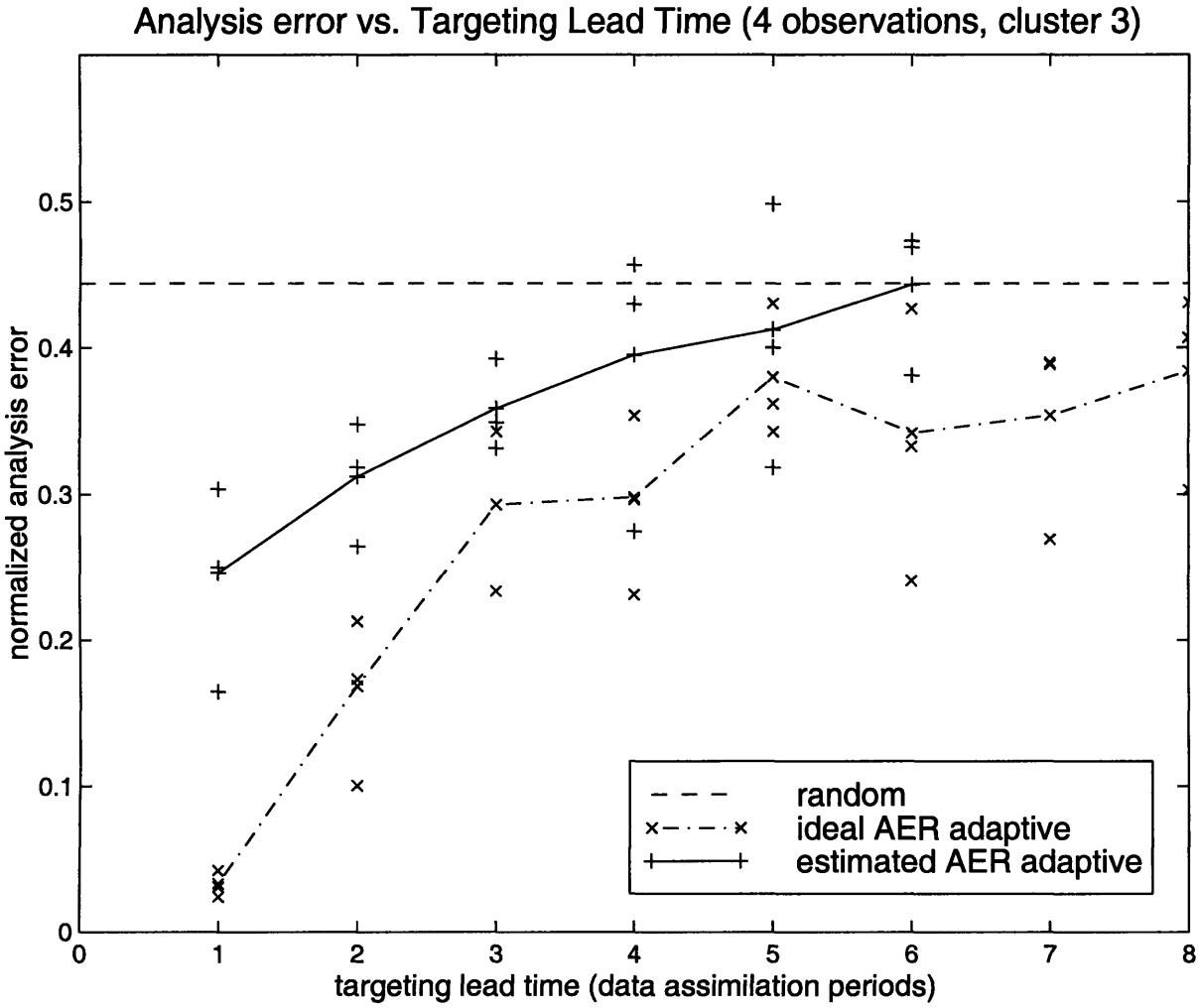



Figure 7.11: Analysis error reduction for the estimated AER adaptive and ideal AER adaptive strategies as a function of the targeting lead time, i.e. how many data assimilation intervals the observation locations are selected in advance of the observation time. The data assimilation interval is 12 hours, so a 1 data assimilation interval lead time is equivalent to a 12 hour lead time. As in Figure 7.7, cluster 3 observations are taken at 4 targeted locations. The ensemble has 13 members. For comparison, the random strategy results are also shown for the same clustered observation density. The 1 data assimilation interval lead time results are the same as those for the appropriate observation density in Figure 7.1. The kinks in the ideal AER adaptive results are likely due to insufficient statistics and can be ignored.

somewhat slow error growth and has no forecast model error. However, there are also factors that tend to make this estimate somewhat pessimistic. For example, in each case we are taking all adaptive observations, all at locations which have been estimated to have large background errors. By continually adapting observations (instead of leaving them fixed as in Figure 7.10), we are likely to decrease the background errors at the targeted locations between when the observation locations are selected and when the observations are taken, even further limiting the accuracy of our advance estimate. These results suggest that it may be possible to effectively select adaptive observation locations (with a strategy based on errors in the initial conditions) several data assimilation periods in advance of when they are taken. Since lead time is an important limitation of adapting observations with aircraft, further study is needed in a more realistic system before the adaptive strategy results presented here can be applied to the real world.

# Chapter 8

# Observations added to a pre-existing network of fixed observations

It is unlikely that all real atmospheric observations will be taken adaptively in the near future. We initially allocated observations according to only one strategy at a time, however, to avoid potentially confusing interactions between fixed and moving observations. Building on our basic understanding, we now compare observation strategies in a more realistic scenario: observations added to a pre-existing network of fixed observations. In contrast to the "global" observation experiments presented earlier, we call these experiments "added" observations. To prevent any overlapping observation locations, added observations are spaced from each other and from the pre-existing observations according to the spacing constraints described in Section 4.2. All results shown in this chapter allocate observations and take and assimilate data every 12 hours.

## 8.1 Results for time-averaged errors

In this section, the influence of the added observations is measured as in Chapters 5–7, with a time- and domain-averaged streamfunction analysis error norm. As discussed in Section 5.2 and as demonstrated for global observation strategies in Figures 6.2 and 6.3, results are similar in this simulated system for other time-averaged analysis and forecast error norms. Results for error reduction in a time series of forecasts are shown in Section 8.2.

Figure 8.1 compares average results from the same four strategies discussed in previous chapters (fixed, random, ideal AER adaptive, and estimated AER adaptive) for dif-

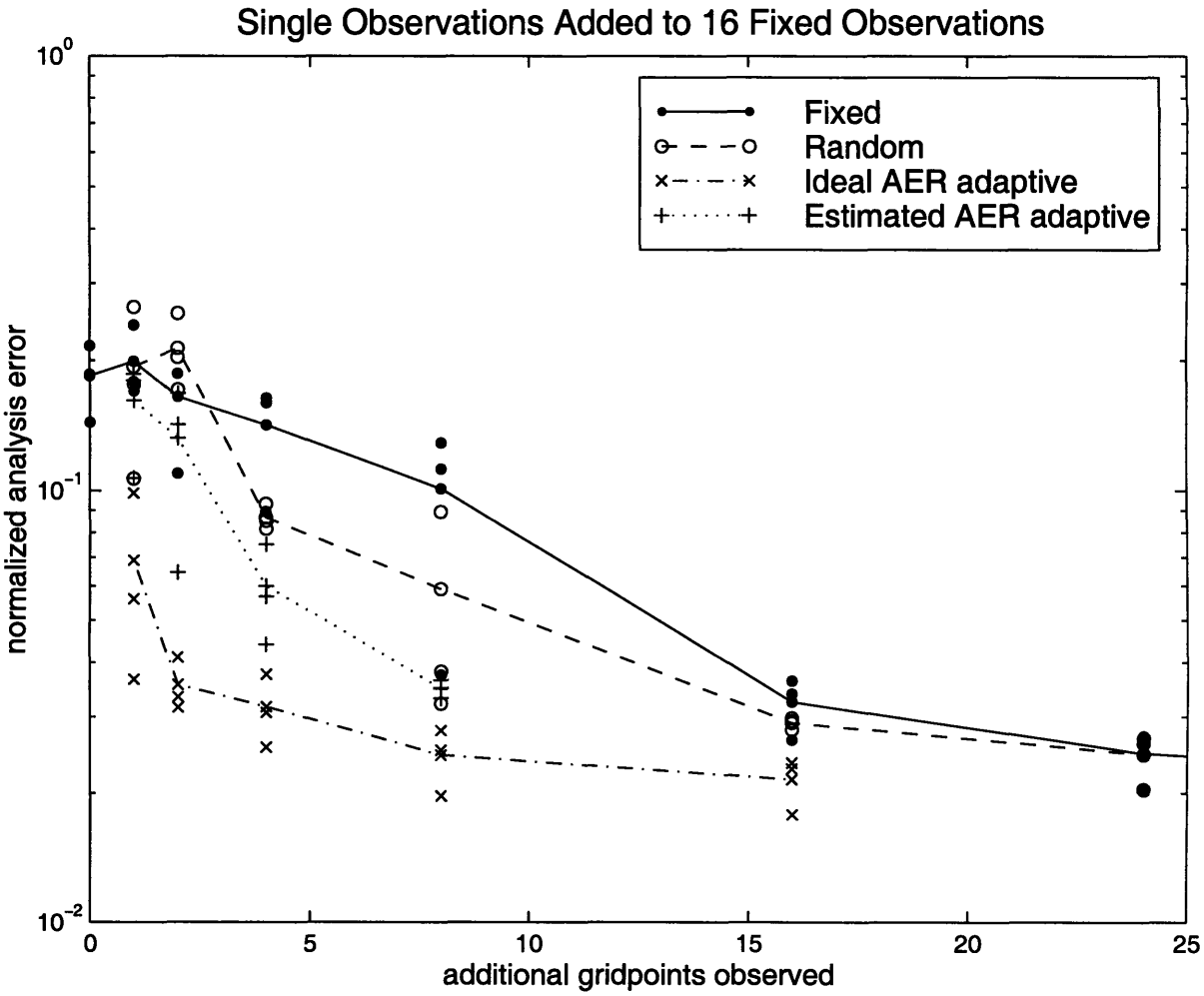



Figure 8.1: Analysis error reduction as a function of the number of additional observations for single fixed, random, ideal AER adaptive, and estimated AER adaptive observations (with a 13 member ensemble). The observations are added to a pre-existing network of 16 fixed observations (approximately .76% on the normalized observation density x-axis). All observations are taken and assimilated every 12 hours. Note that the x-axis in the added observations figures is linear.

ferent numbers of observations added to a pre-existing network of 16 fixed observations. A single observation is taken for each selected location. When observations are added to a reasonably sparse pre-existing observation network instead of all allocated adaptively, the ideal AER adaptive strategy is still an improvement over the random strategy. In fact, in terms of additional observational resources required, the ideal adaptive strategy is actually *more* efficient when the observations are added to a fixed network. In Figure 8.1, for example, only 2 added adaptive observations reduce the average analysis error as much as 16 added fixed or random observations.

Figure 8.2 compares the same results as in Figure 8.1 for added cluster 3 observations (except fixed observations, which are still taken unclustered); cluster 3 is depicted in Figure 6.4. Results are similar for other cluster patterns. The ideal AER adaptive strategy is very effective for clustered observations, so much so that only one or two adaptive observation clusters are needed to reduce the error to nearly the dense observation limit. The results in Figures 8.1 and 8.2 suggest that the 16 fixed observations reduce the analysis error sufficiently that in an average situation, only a few regions contain fairly large analysis errors. Given the limitations of the data assimilation system and the limitations of the adaptive strategy, the analysis error can then be reduced nearly as much as is possible with observations at only a few well-selected extra locations. This is also demonstrated in Figure 8.3, which compares the global observation strategy results from Figure 6.1 with the added ideal adaptive observation results from Figure 8.1 as a function of the density of all (fixed and moving) observations. With only a few adaptive observations added to the 16 fixed observations, the results from the added adaptive observation network asymptote to the results from the all adaptive observation network.

Recall from Chapter 7 that for global observations, the estimated AER adaptive strategy is an improvement over the random strategy only for clustered observations. For observations added to a reasonably sparse pre-existing network, on the other hand, Figures 8.1 and 8.2 show that the estimated adaptive strategy (with a 13 member ensemble) is effective for single as well as for clustered observations. The effectiveness for added but not global observations also occurs for another less effective strategy tested, the ideal AER strategy with a streamfunction norm for targeting (discussed in Section 6.1). As discussed in Section 6.2, it is difficult to separate the effects of the strategies and the observation clusters from the effects of the data assimilation system and the forecast model. That the estimated AER adaptive strategy is more effective for added than for global observations, however, suggests that the fixed observations "anchor" the ensemble in some sense. This allows the estimated strategy to better identify not just general

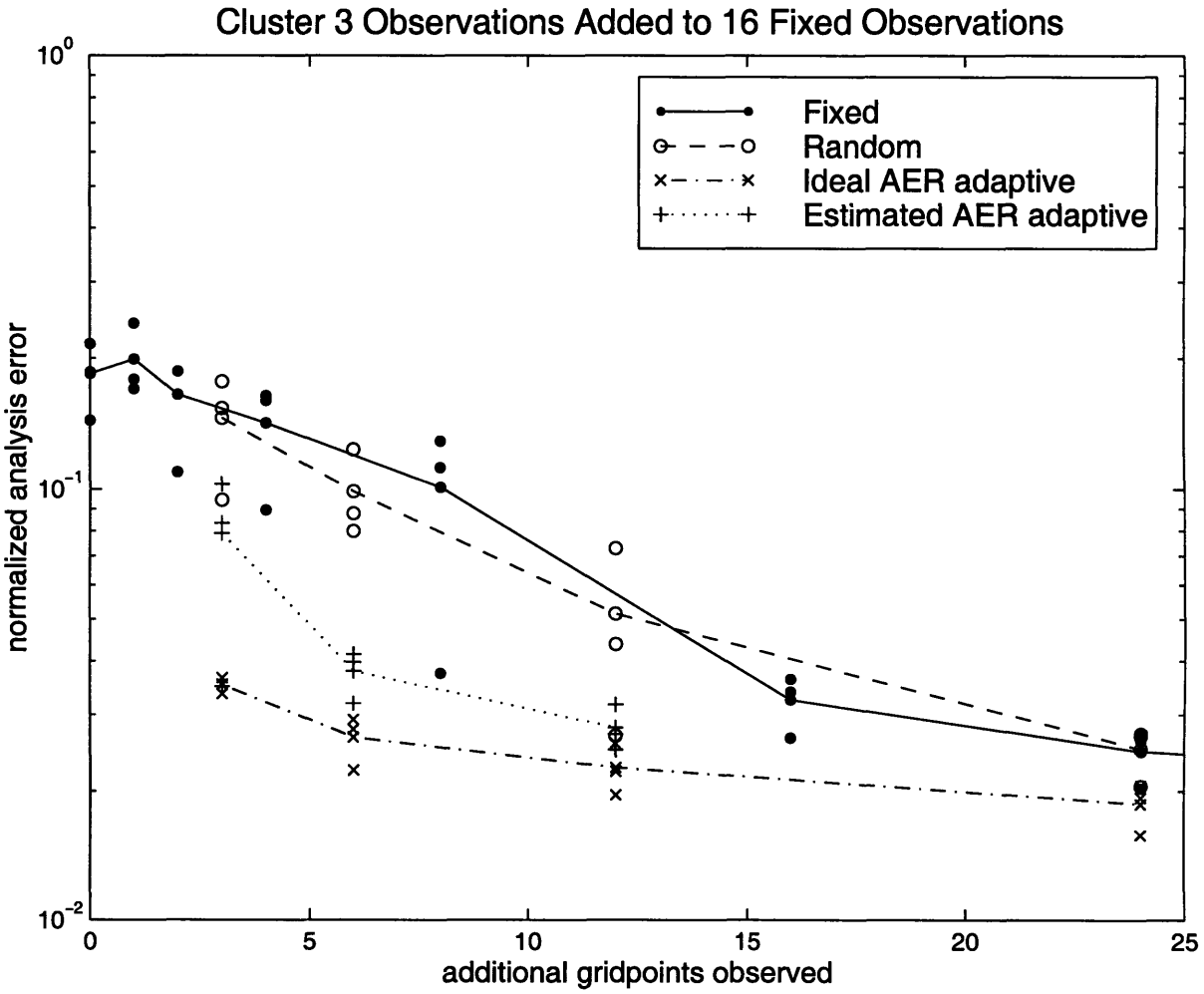



Figure 8.2: As in Figure 8.1, for cluster 3 observations (described in Section 6.2) taken for the random and adaptive strategies (single observations are still taken for the fixed strategy). The x-axis is the number of additional observations, which is equal to the number of targeted locations $\times$ 3.

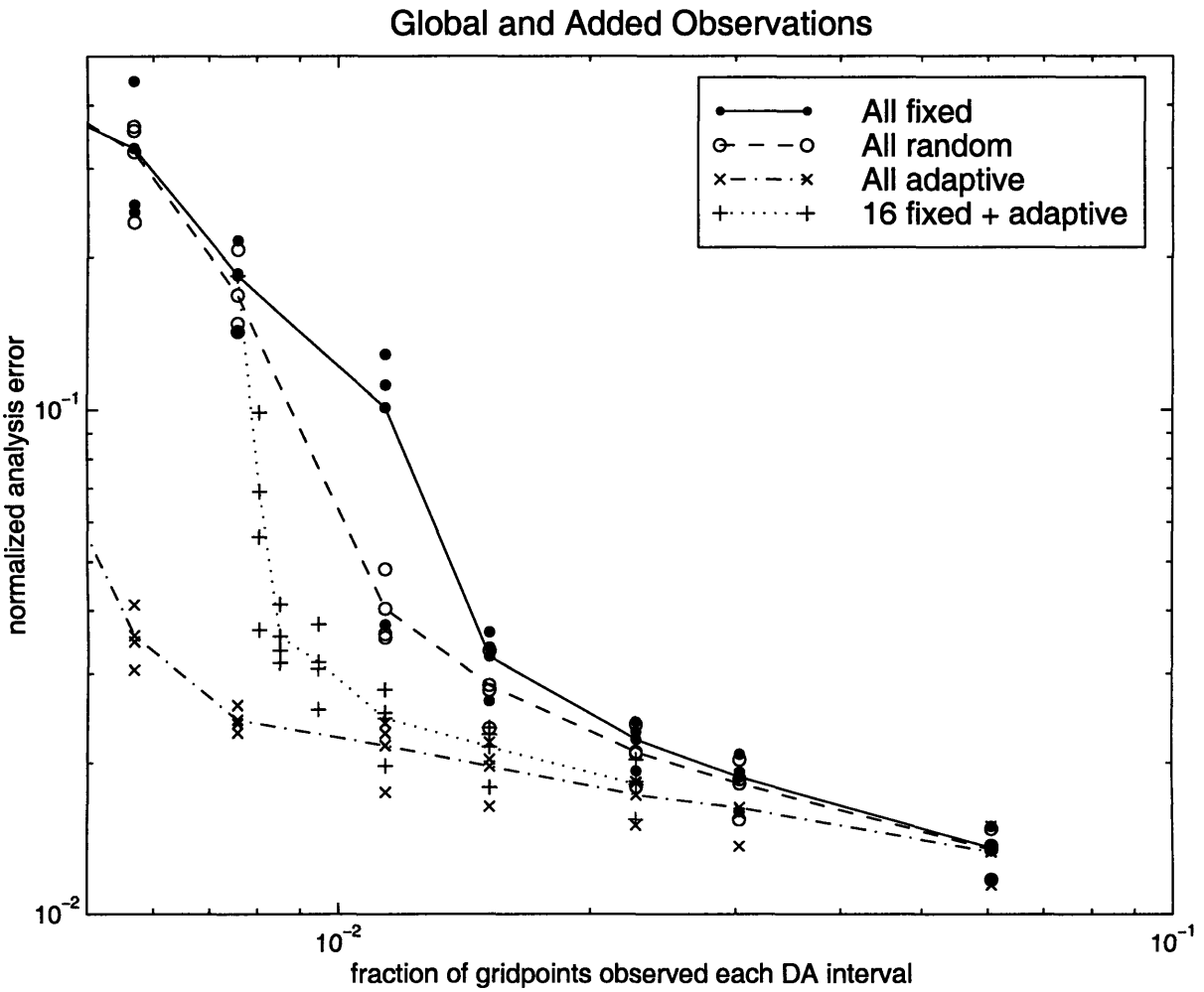



Figure 8.3: Analysis error reduction as a function of the total number of observations for single global fixed, random, and ideal AER adaptive observations, and for single adaptive observations added to a pre-existing network of 16 fixed observations. The global observation strategy results are the same as those in Figure 6.1 but are shown only for a subset of the observation densities. The added adaptive observation results are the same as those in Figure 8.1 but are plotted as a function of the total observation density, i.e. the density of the fixed and adaptive observations together. The axes are magnified to show the regime of interest.

regions, but also more specific locations with large background errors.

We have not explored in detail how or why this anchoring happens. However, from our examination of background errors and ensemble spread for different observation networks, we suspect that it occurs in large part because the fixed observations consistently reduce both large analysis errors and large ensemble spreads in certain regions, giving the adaptive strategies (but particularly the estimated strategy) fewer locations among which to choose. The fixed observations may also be beneficial because the ensemble sometimes selects locations with small background errors at a few consecutive targeting times, and in these situations the fixed observations prevent the ensemble spread from diverging too rapidly from the true background error while keeping the control trajectory reasonably close to the true trajectory. In a more realistic atmosphere, storm tracks and other stationary waves may play a similar role to the fixed observations in this idealized system, physically "anchoring" the ensemble and producing an ensemble spread which better estimates background error with fewer or no fixed observations assimilated into the ensemble.

Comparing Figures 8.1 and 8.2, we see that the results from the estimated strategy are closer to the results from the ideal strategy for added clustered observations than they are for added single observations. For larger observation clusters than the results shown, the estimated strategy is an even better approximation to the ideal strategy. Thus, for added observations as for global observations, the estimated AER strategy is more effective when it takes more than one observation around the targeted location. As discussed in Chapter 7, on average the estimated AER strategy is able to identify locations near but not exactly at locations where the background error is large. The benefit of clustering observations for the estimated AER strategy is added to the benefit from the pre-existing fixed observations.

Figures 8.1 and 8.2 are for a fairly sparse pre-existing network, with an observation density of .76% of gridpoints observed (on the normalized observation density scale). For a pre-existing network of 1.5% of gridpoints (32 observations), the adaptive strategies produce only a small improvement over the random or fixed observation strategies. For higher density pre-existing networks, the strategies are nearly indistinguishable. As discussed in Chapters 5 and 6, above a certain spatial observation density, adding observations has a much smaller effect. In this dense observation regime, the possible error reduction is limited not only by the amount of observational data, but also by the errors in the observations, by the quasi-geostrophic model resolution, and by the data assimilation system. This is true for any type of added observations, whether the

other observation locations in the domain have been selected adaptively or otherwise. Therefore, adaptive observations added to a pre-existing network are only effective when the observation density is reasonably low. However, as long as one is an observation density regime in which additional observations can have a reasonable influence, the comparisons for global observation strategies shown in previous sections are qualitatively valid for the same strategies when observations are added to a pre-existing observation network.

## 8.2 Results for time series of errors

We would like to adapt observations not only to reduce errors on average, but also to reduce the number of situations with very large forecast errors, i.e. the number of forecast "busts." To understand how we might be able to reduce errors in different circumstances, in this section we compare the day-to-day variability in the errors resulting from different observation networks. We do so by, instead of showing time-averaged errors, showing a time series of errors from an individual sequence of forecasts. In all of the error time series shown, the experiment is run for the same arbitrarily selected 360 day sequence of QG model truth states. The data is assimilated and the analyses and forecasts are verified (versus the true state) every 12 hours. For each of the observation networks, we start with the same initial conditions at some past time, then assimilate data every 12 hours to equilibrate the errors in the model state to the observation network prior to $t = 0$.

In previous chapters and in Section 8.1, we have presented most results in terms of reduced analysis error, explaining (and in Figure 6.3, demonstrating) that the results for average forecast error reduction are qualitatively similar. For the time series results, however, it is in some cases easier to subjectively differentiate among the observation networks for forecast errors than for analysis errors. Therefore, in this section most of the figures present time series of domain-averaged 3 day forecast errors. Recall from Chapter 2 that the error growth rate in the standard resolution QG model is somewhat slower than that in the real atmosphere, so a 3 day time scale in these results is equivalent to a shorter time scale in the real atmosphere. As we show for one example (Figure 8.6), time series of analysis errors look similar to those for forecast errors, although the differences are in some cases slightly less pronounced. The tables summarize the results from the figures, comparing the means and standard deviations of the analysis errors and of the

| | Analysis error | | 3 day forecast error | |
|---|---|---|---|---|
| observation network | mean | std deviation | mean | std deviation |
| none | 29.84 | 7.06 | 29.79 | 7.05 |
| 8 fixed | 19.48 | 5.61 | 23.98 | 6.89 |
| 16 fixed | 8.17 | 3.61 | 12.81 | 5.94 |
| 32 fixed | 0.96 | 0.28 | 1.63 | 0.62 |
| 64 fixed | 0.61 | 0.11 | 1.06 | 0.34 |

Table 8.1: Means and standard deviations of domain-averaged analysis and 3 day forecast errors for different fixed observation networks, during the 360 day time series shown in Figure 8.4. The error norm is streamfunction in non-dimensional units $\times 100$. Because we have only tested one observation distribution for each observation density, small changes between results for different observation networks are not statistically significant.

3 day forecast errors during the 360 day period for each of the observation networks.

Figure 8.4 shows time series of 3 day forecast errors for different densities of fixed observations. Table 8.1 shows the corresponding error means and standard deviations during the 360 day period. In Figure 8.4 and Table 8.1, we see that even in the QG model (with simplified dynamics, forcing, and geometry), the error variability can be quite large. There are two main reasons why the forecast errors vary from day to day in this idealized system, both of which are also important in the real atmosphere. First, as indicated by the error variability when there are no observations, the error growth rate in the QG model (and thus in the forecasts) varies from day to day. Second, the fixed observation locations sample the important features in the atmospheric state much better in some situations than in others. This error variability is responsible for the variability in the time-averaged results for each observation network shown earlier, as indicated by the spread between the symbols in previous figures (such as Figure 5.1). Non-dense observation networks, particularly those in the transition regime (e.g. 16 fixed observations in Figure 8.4), also tend to produce analysis and forecast errors with significant variability on time scales of one or several months. Consequently, the spread in the time-averaged results (which are averaged over a 90 day period) is particularly large for these observation networks.

Figure 8.4 also demonstrates how, as discussed in previous chapters, the time-averaged results are not necessarily valid in any specific case. Because of the imperfect data assimilation system, having more observations generally improves the analysis and

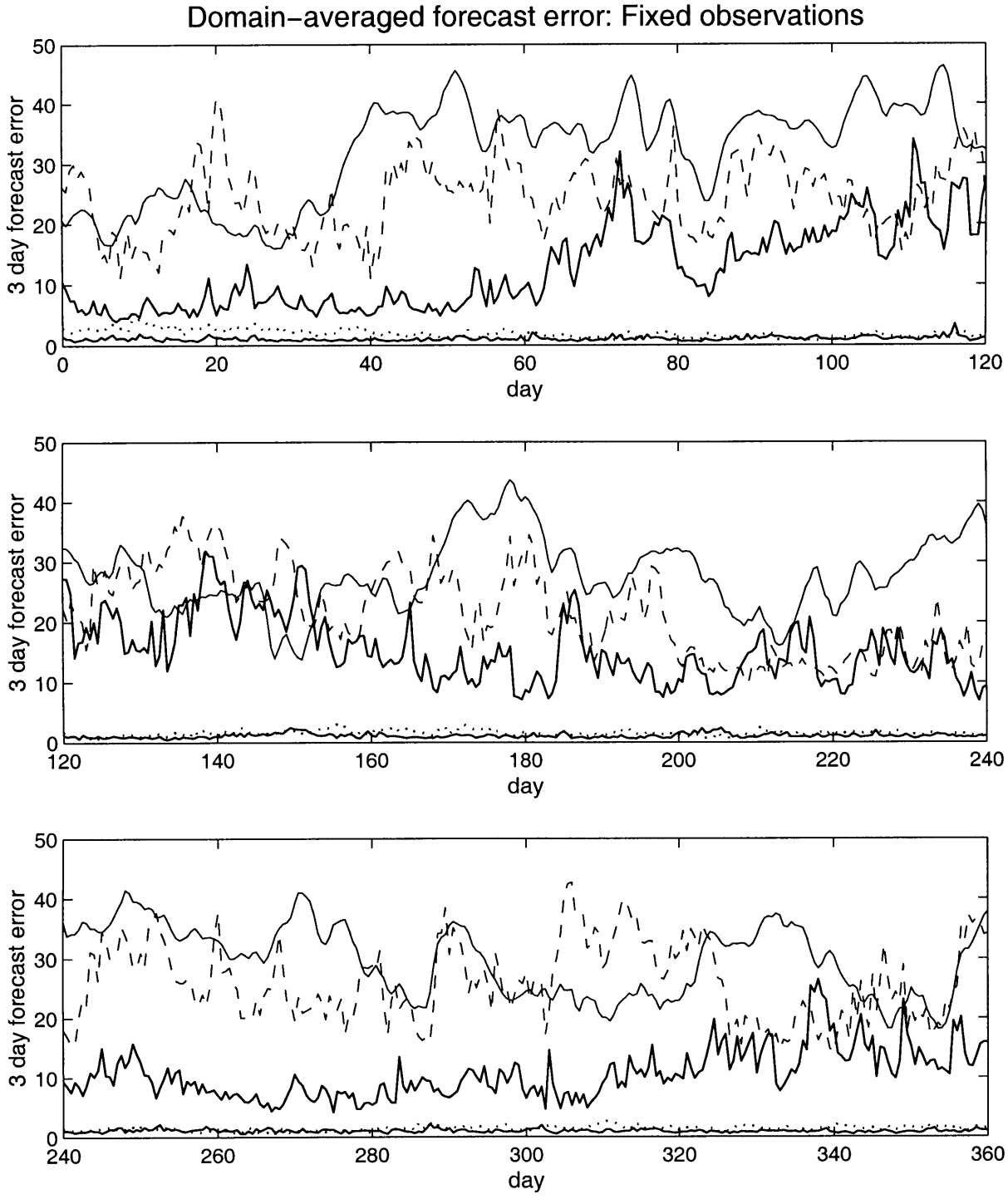



Figure 8.4: Domain-averaged 3 day forecast error plotted every 12 hours during one 360 day period for: no observations (medium solid line); 8 fixed observations (dashed line, 0.38% of gridpoints observed); 16 fixed observations (upper thick solid line, 0.76%); 32 fixed observations (dotted line, 1.5%); and 64 fixed observations (lower thick solid line, 3.0%). Only one distribution of fixed observations and one set of initial conditions is shown for each observation density. The model state was spun up for 60 days prior to $t = 0$ for each observation network. The error scale is streamfunction in non-dimensional units $\times 100$.

| observation network | Analysis error | | 3 day forecast error | |
|---|---|---|---|---|
| | mean | std deviation | mean | std deviation |
| 16 fixed | 8.17 | 3.61 | 12.81 | 5.94 |
| 17 fixed | 7.03 | 3.08 | 11.17 | 5.09 |
| 16 fixed + 1 random | 7.75 | 3.23 | 12.07 | 5.21 |
| 16 fixed + 1 estimated adaptive | 4.91 | 1.72 | 7.88 | 3.20 |
| 16 fixed + 1 ideal adaptive | 3.06 | 1.21 | 5.13 | 2.14 |
| 18 fixed | 5.86 | 2.62 | 9.25 | 4.22 |
| 16 fixed + 2 random | 5.28 | 2.23 | 8.34 | 3.61 |
| 16 fixed + 2 estimated adaptive | 3.85 | 1.36 | 6.66 | 2.59 |
| 16 fixed + 2 ideal adaptive | 1.32 | 0.35 | 2.23 | 0.82 |

Table 8.2: As in Table 8.1, for 1 or 2 single observations added to 16 fixed observations (the results shown in Figures 8.5, 8.6, and 8.7). The ensemble for the estimated strategy has 7 members. Again, small changes are not statistically significant.

forecast, but not always. Finally, Figure 8.4 and Table 8.1 clearly show the transition between the sparse and dense observation regimes, which we first discussed in Chapter 5. The rapid error dropoff in the time-averaged results in Figure 5.1 and later figures is due to a dramatic decrease in both the mean error and the variability in the error between 16 and 32 fixed observations.

In the previous section (for example Figure 8.1), we saw that by adding only a few adaptive observations to a pre-existing network of 16 fixed observations, we can on average significantly improve analyses. To see how adding adaptive observations to a fixed network changes analyses and forecasts on a day-to-day basis, Figure 8.5 shows the error time series for domain-averaged 3 day forecast errors during the same 360 day sequence as in Figure 8.4, for the same network of 16 fixed observations and for a single fixed, random, estimated AER adaptive, or ideal AER adaptive observation added to the 16 fixed observations every 12 hours. For comparison, Figure 8.6 shows the same results for the domain-averaged analysis error (without the estimated strategy, for clarity). Table 8.2 summarizes the analysis and 3 day forecast error time series results during the 360 day period.

Figure 8.5 demonstrates that adding an observation of any type, adaptive or non-adaptive, to a pre-existing network does not improve the forecast in each situation.

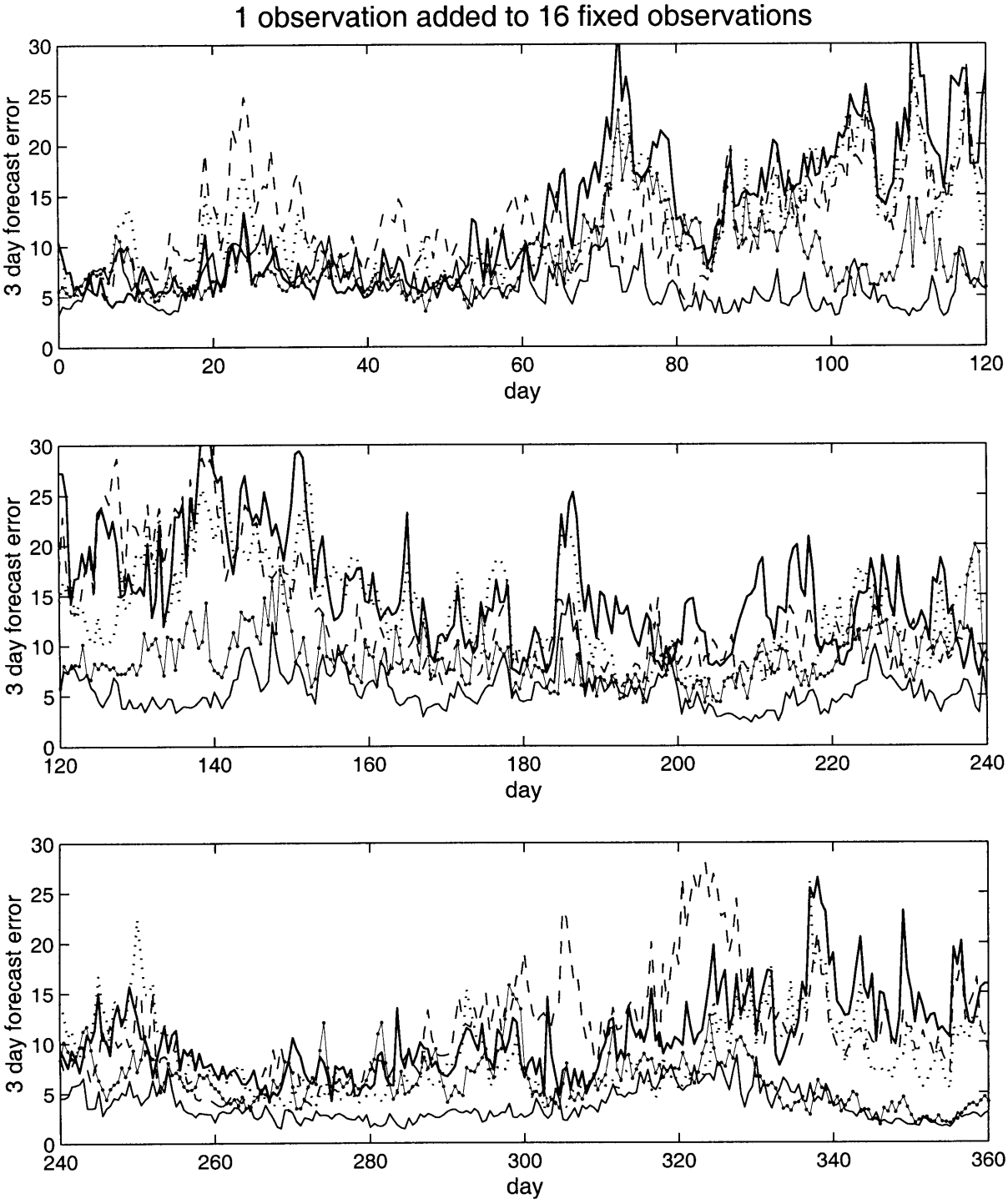



Figure 8.5: As in Figure 8.4, for domain-averaged 3 day forecast error during the same 360 day period, for: 16 fixed observations (thick solid line); 17 fixed observations (dotted); 16 fixed observations + 1 random observation (dashed); 16 fixed observations + 1 estimated AER adaptive observation (thin solid with circles); and 16 fixed observations + 1 ideal AER adaptive observation (medium solid). The ensemble for the estimated strategy in this and later time series results has 7 members.

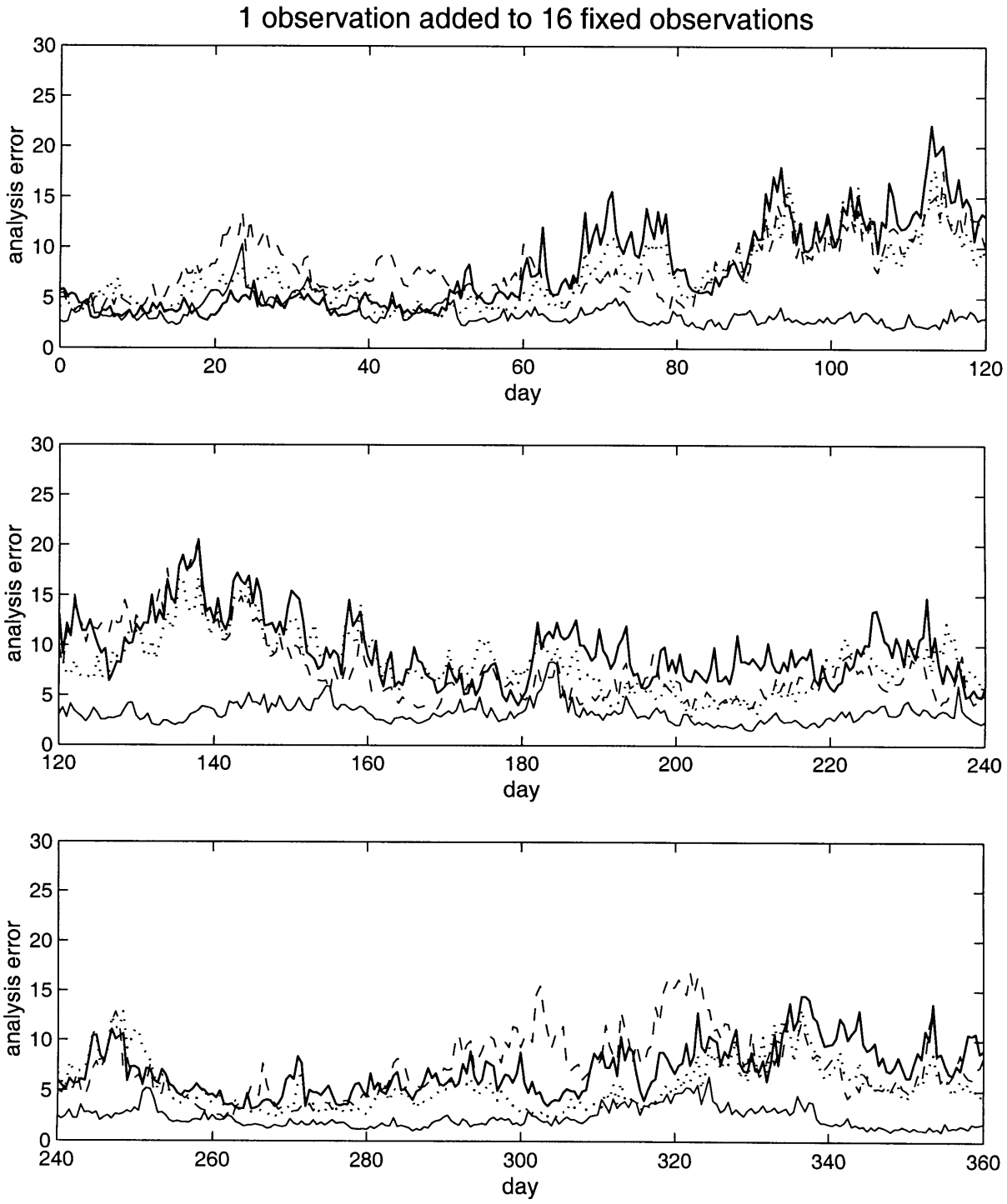



Figure 8.6: As in Figure 8.5, during the same 360 day period, but for domain-averaged analysis error. Results are depicted for all of the same observation networks, except that the added estimated AER adaptive observation is not shown for clarity.

Figure 8.5 and Table 8.2 also show that an extra fixed or random observation every 12 hours improves the forecast in some situations but degrades it in others, nor does the extra non-adaptive observation significantly decrease the error variability. An extra ideal AER adaptive observation every 12 hours, on the other hand, reduces the error variability by a large amount. It does so by dramatically reducing the error in the situations when the error would otherwise be very large. Therefore, adding adaptive observations to a reasonably sparse observation network can not only reduce errors on average, but also reduce the number of forecast busts.

Figure 8.7 and the lower half of Table 8.2 show the same results, but for 2 single observations added to the 16 fixed observations every 12 hours. The fixed and random strategies reduce both the average error and the error variability more with 2 added observations than with 1 added observation, but they still leave medium–large forecast errors in many situations. The added fixed or random observations also still noticeably degrade some forecasts. In contrast, two added ideal adaptive observations again dramatically reduce not only the average error, but also the error variability and the error in situations when the error would otherwise have been large. As a further benefit, in the 720 cases shown the 2 added ideal adaptive observations never degrade the forecast.

In Figures 8.5 and 8.7 and Table 8.2, the estimated AER adaptive strategy, which is closer to a strategy which could implemented in the real world, is also tested. For single observations added to a pre-existing network, the estimated adaptive strategy is not as effective as the ideal AER adaptive strategy. It is, however, still more effective than the random strategy, both on average and in most individual situations. Although added estimated adaptive observations do not always improve the forecast, and although they sometimes perform worse than added random observations, they are also advantageous because they are less likely than the added fixed and random observations to produce severe forecast degradations.

Figure 8.8 and Table 8.3 show results for the same 360 day time series, for the same network of 16 fixed observations and for 1 cluster 3 (as depicted in Figure 6.4) of random, estimated adaptive, or ideal adaptive observations added to the fixed network every 12 hours. For an added cluster of observations, the random strategy performs better on average than it did for single observations, and it is less likely to significantly degrade the forecast. The ideal adaptive strategy, however, remains significantly more effective. The clustered estimated adaptive strategy not only usually performs better than the random strategy, but in many cases it also improves the forecasts nearly as much or more than the ideal adaptive strategy.

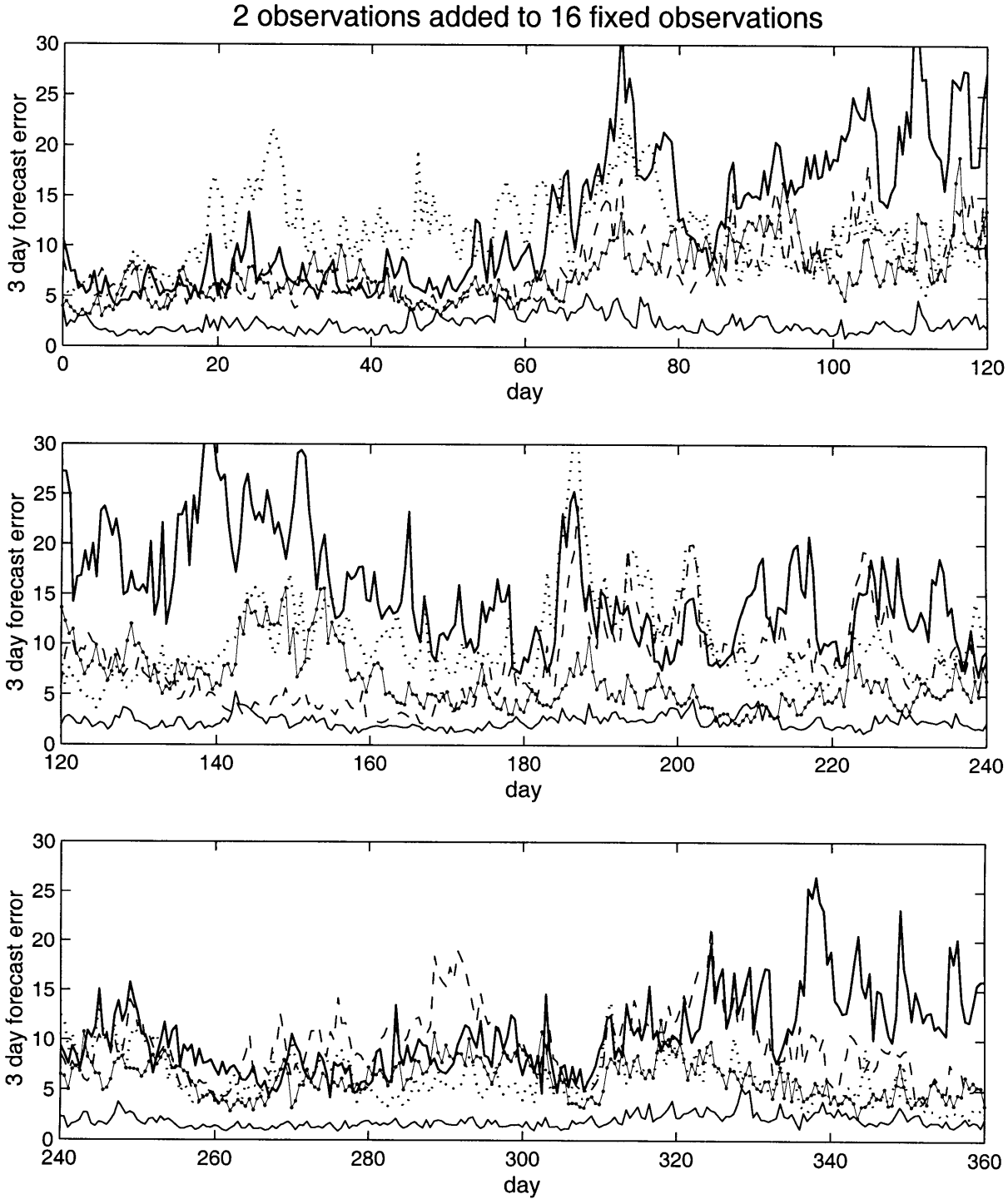



Figure 8.7: As in Figure 8.5, for 3 day forecast error during the same 360 day period for: 16 fixed observations (thick solid line); 18 fixed observations (dotted); 16 fixed observations + 2 random observations (dashed); 16 fixed observations + 2 estimated AER adaptive observations (thin solid with circles); and 16 fixed observations + 2 ideal AER adaptive observation (medium solid).

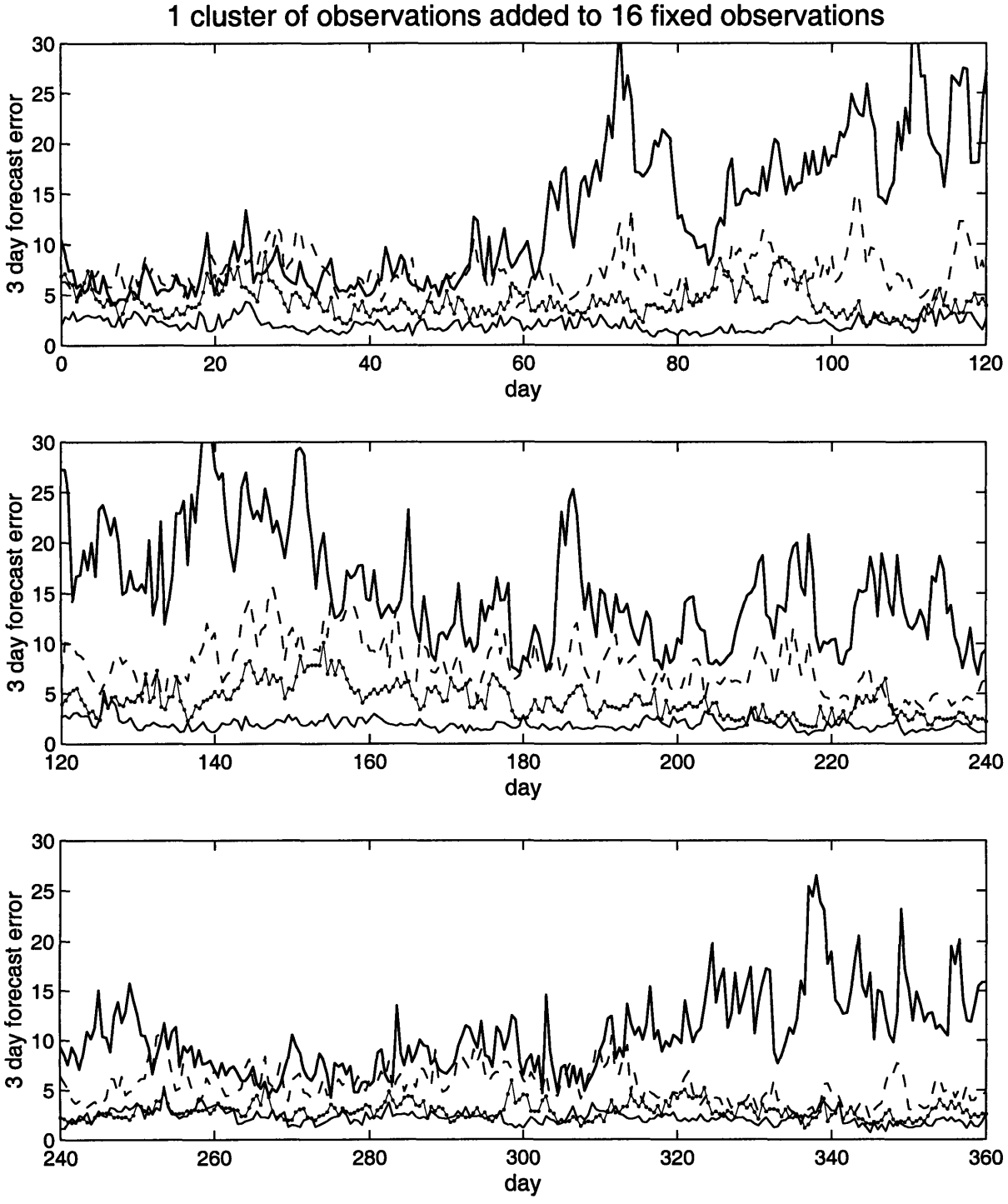



Figure 8.8: As in Figure 8.4, for 3 day forecast error during the same 360 day period for: 16 fixed observations (thick solid line); 16 fixed observations + 1 cluster 3 of random observations (dashed); 16 fixed observations + 1 cluster 3 of estimated AER adaptive observations (thin solid with circles); and 16 fixed observations + 1 cluster 3 of ideal AER adaptive observations (medium solid).

| observation network | Analysis error | | 3 day forecast error | |
|---|---|---|---|---|
| | mean | std deviation | mean | std deviation |
| 16 fixed | 8.17 | 3.61 | 12.81 | 5.94 |
| + 1 cluster 3 random | 4.15 | 1.39 | 6.79 | 2.36 |
| + 1 cluster 3 estimated adaptive | 2.21 | 0.72 | 3.83 | 1.51 |
| + 1 cluster 3 ideal adaptive | 1.25 | 0.34 | 2.07 | 0.66 |

Table 8.3: As in Table 8.2, for 1 cluster 3 of observations added to 16 fixed observations (the results shown in Figure 8.8). Again, small changes are not statistically significant.

Figure 8.9 depicts the same 360 day time series as in Figure 8.7, but for 2 single observations added to 64 fixed observations every 12 hours (the estimated AER strategy results are not plotted for clarity). The mean and standard deviations of the errors are summarized in Table 8.4. With observations taken at 64 fixed locations (3.0% of gridpoints) every 12 hours, we are in or near the data dense regime. Both the mean error and the day-to-day variability in the error (in absolute terms) are much smaller than for 16 fixed observations (as illustrated in Figure 8.4 and Table 8.1). As we discussed in the previous section, adaptive observations do not on average reduce errors by a large amount when added to a pre-existing network of 64 fixed observations. It is possible, however, that in relative terms the variability is still sufficiently large that adaptive observations can reduce errors in the situations with the largest errors.

For a reasonably dense pre-existing fixed observation network, Figure 8.9 and Table 8.4 show that added adaptive observations are at best only a small improvement over added fixed or random observations, either on a time-averaged or a case-by-case basis. In some situations, when the error with no extra observations is larger than average for several days, the 2 additional adaptive observations do improve the forecast. In these cases, however, adding fixed or random observations generally reduces the error about as much. It is usually simply adding observations, therefore, not adapting them, that is beneficial. In addition, adding any type of observations (adaptive or non-adaptive) sometimes degrades the forecast.

The results in Figure 8.9 do suggest that if we can predict which situations will have larger forecast errors, adding observations might be beneficial even for dense observations. With an adaptive strategy focussed on improving the specific predicted significant

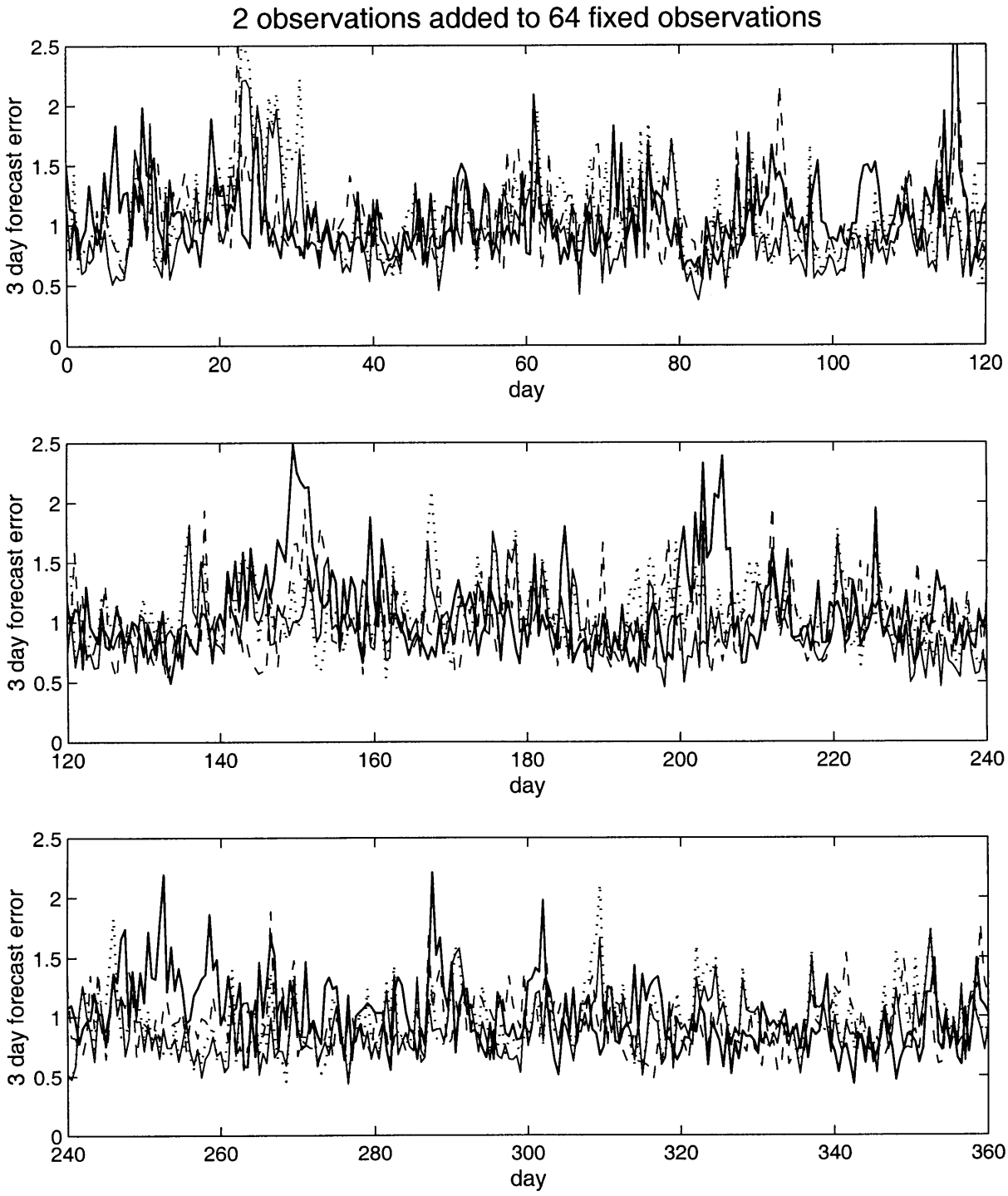



Figure 8.9: As in Figure 8.7 for 3 day forecast error during the same 360 day period for: 64 fixed observations (thick solid line); 66 fixed observations (dotted); 64 fixed observations + 2 single random observations (dashed); and 64 fixed observations + 2 single ideal AER adaptive observations (medium solid). The estimated AER strategy results are not plotted for clarity. Note the large change in the y-axis scaling.

| observation network | Analysis error | | 3 day forecast error | |
|---|---|---|---|---|
| | mean | std deviation | mean | std deviation |
| 64 fixed | 0.61 | 0.11 | 1.06 | 0.34 |
| 66 fixed | 0.61 | 0.11 | 1.04 | 0.31 |
| 64 fixed + 2 random | 0.60 | 0.11 | 1.00 | 0.27 |
| 64 fixed + 2 estimated adaptive | 0.59 | 0.10 | 1.00 | 0.29 |
| 64 fixed + 2 ideal adaptive | 0.57 | 0.10 | 0.94 | 0.28 |

Table 8.4: As in Table 8.2, for 2 single observations added to 64 fixed observations (the results shown in Figure 8.9) and the estimated AER adaptive strategy results. Again, small changes are not statistically significant

error, additional observations in these situations might be particularly effective. As we discussed in previous chapters, however, using extra observations to improve analyses and forecasts at higher observation densities is difficult. In order to effectively adapt observations at these densities, we will likely need a better strategy, a better data assimilation system, and a more specific norm. We will also need to study many more cases to see if the adaptive observations are effective on a statistically significant basis. In contrast, for reasonably sparse observation networks, analysis and forecast errors can be significantly reduced using a simple strategy and a relatively simple data assimilation scheme.

The 3DVAR data assimilation system performs best when observations are taken at locations with large background errors, and the ideal adaptive strategy tested here is particularly effective because it observes at precisely these locations. In realistic situations, we must estimate the background error. the estimated adaptive strategy results demonstrate, with an imperfect estimate of background error the forecasts will benefit less from adding adaptive observations to a pre-existing network. Another possible major limitation of these results is that we assume a perfect forecast model. In a realistic system with forecast model errors, adding adaptive observations is not likely to improve forecasts as significantly on similar time scales. Our strategy, however, is dependent only on our estimate of background error, not on the accuracy of our forecast model trajectories. Therefore even with an imperfect forecast model, for the appropriate observation density, adding adaptive observations can still improve analyses and shorter lead time forecasts.

The time series results demonstrate that adding adaptive observations to a pre-existing fixed observation network not only reduces average analysis and forecast errors, but also reduces the day-to-day variability of the errors — as long as the pre-existing observation density is low enough that the extra observations can have an influence. In the examples shown here, for a sparse observation network, we reduced the largest domain-averaged analysis and forecast errors simply by observing at every data assimilation time at the few locations with the largest background errors, without explicitly taking future error growth into account.

# Chapter 9

# Influence of extra observations on analysis and forecast errors in individual situations

The results presented in previous chapters are nearly all for error reduction averaged over a large number of cases. The adaptive strategies we have tested are also based only on errors in the initial conditions, without explicitly taking into account future error growth. The results shown so far have therefore left unanswered two fundamental questions, both of which we believe need to be addressed before observations can be adapted effectively. First, how do the extra observations interact with the data assimilation system and with the forecast model to change analyses and forecasts in individual situations? Second, what criteria should we include in the "optimal" adaptive observation strategy, where the optimal strategy is likely to vary from situation to situation?

To begin addressing these issues, in this chapter we explore some of the different ways in which any extra observations can change analyses and forecasts. Because the influence of observations depends strongly on the specific forecasting system, the specific forecast situation, and the specific errors in the initial conditions, any detailed results we find in individual cases may not be applicable for adapting observations in general. Therefore, we do not focus on the details of how individual observations affect specific forecasts. Instead, we test taking extra observations at each location in the domain, then summarize the results by calculating how much extra observations at each location change the analysis and the forecasts. To avoid biasing our results by the specific errors in the initial conditions at and near each extra observation location, we test extra

observations both in different synoptic situations and for the same synoptic situation but with different errors in the initial conditions.

Even for an individual synoptic situation with one set of errors in the initial conditions, extra observations can affect analyses and forecasts in a variety of complex ways. Although we would like to understand how and why observations improve and degrade forecasts, we have found many of the effects difficult to trace to a particular cause. To draw firm conclusions and to provide detailed answers to the questions above, a much more in depth dynamical study is needed. Consequently, from these results we only try to suggest general ways in which extra observations can affect analyses and forecasts in any forecasting system which, similar to our system, is sensitive to errors in the initial conditions and assimilates data imperfectly.

We would like to incorporate the results from this chapter and from previous chapters into a single framework. If we are to do so, however, we must emphasize that while the results in this chapter can help us interpret the results from Chapters 5–8, they are not strictly comparable. In the results presented in previous chapters, additional observations were taken at every data assimilation interval during a long run; in this chapter, extra observations are tested at only one time. As discussed in Section 6.3, the adaptive strategy is most effective when it takes observations more frequently, reducing large errors in the initial conditions and forecasts as they grow. In this chapter we will see that without the opportunity to adapt observations frequently to correct degraded analyses or forecasts, the additional observations are much less likely to have a positive influence. In particular, the results shown in this chapter are much more pessimistic than those in Section 8.2, even for observations added to a reasonably sparse network. Thus, we should keep in mind that at least for the type of adaptive strategy tested in this study, observations added regularly and observations added at one time can produce quite different results even if they are taken according to the same strategy.

## 9.1 Experimental design

All results presented in this chapter are from case studies during the same 8 day period of the QG model "truth" state. Prior to the 8 day period, we generate 6 different distributions of 16 fixed observation locations and 6 different sets of perturbed initial conditions for the "model" state. We then spin up the OSSE (as described in Section 4.1) for approximately 60 days, assimilating observations every 12 hours. This

produces, at day 0, 6 different sets of initial conditions, one set equilibrated to each of the distributions of fixed observations. During each 8 day case study, although we test assimilating observations at different locations, only the same 16 fixed observations are actually assimilated into the control analysis. As in previous chapters, all observations are simulated rawinsondes, with observations taken at one or more levels at each horizontal gridpoint selected for observation.

Every 12 hours during each 8 day case study, we test taking, in addition to the 16 fixed observations, extra observations arranged in a pattern centered at each location in the domain (except at the channel walls), one by one. For each set of extra observations, we calculate the global average analysis error (with an energy norm) and compare it to the global analysis error without the extra observations. The choice of global average energy for the error norm is discussed in Section 9.2. We define the change in analysis error with the extra observations as

$$\Delta \text{error} = \frac{\text{global average error with extra observations}}{\text{global average error without extra observations}} - 1. \tag{9.1}$$

Next, we run forecasts from each of the analyses with extra observations, calculate the global average error in the resulting forecast, and compare it to the global error in the same forecast without extra observations. For each forecast interval, the change in forecast error with each set of extra observations is again defined according to Equation (9.1). Results are only presented for changes in forecasts up to 2 days since they are nearly identical for longer range forecasts.

With the definition in Equation (9.1), negative values of error change indicate that the extra observations decrease the error, i.e. improve the analysis or forecast. Positive values indicate that the extra observations increase the error, i.e. degrade the analysis or forecast. At the 16 fixed observation locations, extra observations are not tested and the analysis and forecast error changes are set to 0.

After running several of these experiments, we realized that observations affected analysis and forecast errors in very different ways in different situations. Unfortunately, since it is computationally very expensive to assimilate data and run forecasts for extra observations explicitly at each location in the domain, we can only run a limited number of experiments. To evaluate most efficiently how important different criteria are for selecting extra observation locations, we decided to test not only different synoptic situations, but also the same synoptic situation with different errors in the initial conditions. The terminology we use to differentiate among the situations is as follows. We

refer to each 12 hour time during the 8 day QG truth evolution as a "synoptic situation"; there are 17 synoptic situations tested. Because the synoptic situations are separated by only 12 hours, they are not truly independent. Each synoptic situation usually has 2–3 identifiable synoptic systems in the domain. We refer to the different distributions of fixed observations and the corresponding evolution of errors in the initial conditions through the 8 days as "fixed observation distributions 1–6." For each intersection of a synoptic situation and a fixed observation distribution, we have 1 "case." We therefore present results for extra observations at each location in the domain in $6 \times 17 = 102$ cases.

We show two types of results. First, we look at how changes in analysis and forecast errors with extra observations (with the changes averaged globally as defined in Equation (9.1)) depend on the location where the extra observations are taken. To do this, we examine plots of error change as a function of the extra observation location in different cases. Because the results vary significantly from case to case, we also present some statistical results accumulated over all 8 days and all 6 observation distributions. Although we have tested only a limited number of cases, the statistical results suggest how often in general in this simulated system we can expect randomly selected observations added to a reasonably sparse pre-existing observation network to improve or degrade analyses and forecast by a certain percentage.

With the limited dynamical information in the data assimilation procedure, we know that 1 extra observation location is insufficient for the data assimilation to resolve errors in the initial conditions effectively. To give the data assimilation system a better opportunity to correct errors, therefore, we test not only adding single observations for each location, but also adding sets of observations with different cluster patterns (as defined in Section 6.2) centered around each location. Unlike in previous experiments, we do not space the observation clusters from the fixed observations, nor do we remove any observations which overlap with the fixed locations. When the cluster extends outside the domain (for observation locations near the channel walls), the part of the cluster outside the walls is ignored. Most of the results shown (the individual examples in Section 9.3 and the statistical results in Section 9.4) are for cluster 13 extra observations (13 observations in a 1000 km $\times$ 1000 km grid). In Section 9.5, we compare both types of cluster 13 results with results for single extra observations and for two other cluster patterns.

To avoid confusing the influence of extra observations with the influence of observational errors, in the results shown here we also take all perfect (errorless) observations.

As described in Chapter 3 and Section 5.3.1, even with perfect observations the data assimilation system still assumes some observational error and therefore does not match the observations exactly. We have also run same experiments with imperfect observations, and the results are very similar but in some cases slightly less optimistic.

## 9.2 Choice of error norm

With the large number of forecasts tested, it would be difficult to track the propagation of changes in analyses and forecasts and to verify the appropriate synoptic system for each set of extra observations. Therefore, we chose to measure the influence of the observations on domain-averaged errors instead of on errors in localized forecasts. As described in Section 5.2, large analysis and forecast errors tend to be relatively localized, and the global error is dominated by errors in these few localized regions. Thus, by using a global average error norm we are primarily measuring the influence of extra observations in the few regions with large errors. The results shown are for an average energy error norm; results are similar for root-mean-square average potential vorticity and streamfunction error norms.

When interpreting the results, we generally assume that if observations in a certain region significantly change analysis and forecast errors, the extra information in that region had a large effect. With a global error norm, however, in some situations this interpretation may not be precisely correct. Because the data assimilation system solves for the analysis globally, extra observations in one region tend to produce small changes in the analysis in other areas of the domain. Although the effect on the global average analysis error is small, the changes outside the observation region could grow rapidly, changing the forecast significantly outside the region of the extra observations even before the direct signal from the observational data propagates there. Thus, a large change in the forecast error means that the assimilated data has a large influence, not necessarily that information from the observation region itself has influence. We have evaluated this subjectively in a few cases, and it does not seem to be a major factor. Because most modern analyses are global, and because our results themselves do not depend on our interpretation of the source region of the influence being strictly true, we do not believe that this is a major limitation of our results. It should be kept in mind, however, that in some situations the global data assimilation procedure may be responsible for some of the changes which we attribute to observations in a certain region.

## 9.3 Results in individual cases

In this section we illustrate how, in general, extra observations at different locations can affect analysis and forecast errors. As discussed in the introduction to this chapter, the results are complex, and understanding them in detail is not possible within the scope of this study. Our first goal, therefore, is simply to present several examples to help the reader appreciate the range of possibilities both for observations at different times and for observations at different locations at the same time. Every example is unique, and even within the few examples shown here, we are unable to describe the variety of interactions between observations and forecasts. We do, however, describe some general features of all of the cases studied, using the cases shown as examples. While doing so, we discuss what the results might mean for adapting observations in general. Although we have been unable to identify any strict predictors for where and when extra observations can most improve (or degrade) forecasts, we also briefly discuss how each of the examples shown suggests a potentially important issue to consider when adapting observations.

The experiments are set up as described in Section 9.1, with extra observations tested centered around each location in the domain during an 8 day evolution of the truth state (17 synoptic situations). The 8 day case study was run for six different distributions of 16 fixed observations, producing 102 cases. The results shown in this section are for extra cluster 13 observations (13 observations at alternating gridpoints in a $1000\,\mathrm{km} \times 1000\,\mathrm{km}$ grid, depicted in Figure 6.4). Recall that the observations are errorless and the forecast model is perfect.

For each case, we examine contour plots of fractional changes in analysis and forecast errors (as defined in Equation (9.1)) as a function of the center location for a set of extra observations. The changes can be compared between different plots, to see how the background errors and the influence of extra observations varies between cases. The changes can also be compared within each plot, to see at which locations extra observations actually have the maximum and minimum influence on the same analysis and the same forecast. In each case, we can determine which locations different adaptive strategies would select, and we can compare the changes produced by observations in these locations with the changes produced by observations at any other location in the domain. For example, from the background error we can identify which regions the ideal AER adaptive strategy would observe in. Although these experiments could be used to compare different specific observing strategies, in this study we mainly speculate about

which criteria might be important for any effective observation strategy.

The examples shown here are all for the same synoptic situation, but for different distributions of fixed observations. In other words, extra test observations are taken at the same time in the truth state evolution, to try to predict the same atmospheric state, but with different analysis and forecast errors to correct. The initial time is 2.5 days into the 8 day run, and forecasts are run from each of the analyses at day 2.5 to predict the atmospheric state up to day 4.5. The synoptic evolution in the truth state for the relevant time period, days 2.5–4.5, is shown in Figures 9.1 and 9.2. As in previous chapters, all analyses and forecasts are verified versus the true state.

Figures 9.3–9.6 show the changes in analysis and forecast errors produced by extra cluster 13 observations at day 2.5 for 4 of the fixed observation distributions. In each of the figures, panels a and b show the background error, i.e. the error in the initial conditions without any observations (fixed or extra). The fixed observation locations are also depicted. Panel c shows the fractional change in the global analysis error (as defined in Equation (9.1)) as a function of the center location for the extra test cluster of observations. Panel d shows the change in the 12 hour forecast error (the error in the forecast valid at day 3) as a function of the location of the extra cluster of observations at day 2.5, and panel e shows the corresponding change in the 2 day forecast error (the error in the forecast valid at day 4.5). Recall from Section 9.1 that negative changes represent analysis and forecast error reductions, so the dashed contours in panels c–e indicate regions where an extra cluster 13 of observations improves the analysis or forecast. The solid contours are error increases, indicating degradations. Table 9.1 summarizes the maximum improvements and degradations produced by any extra cluster 13 of observations at day 2.5 for the each of the 6 fixed observation distributions.

Before examining the results, we briefly discuss how the plots in panels c–e compare to adjoint sensitivity plots. For a given forecast norm, adjoint sensitivities indicate how sensitive the forecast is to small changes in the initial conditions at each location, assuming linear evolution (see for example Errico and Vukicevic 1992). The plots we present indicate how sensitive the global average forecast error is to assimilated observations taken around each location, for the full non-linear model. Because our plots are explicitly dependent not only on the model dynamics but also on the data assimilation system used, and because our plots identify forecast error reduction or increase instead of the (unsigned) change in a forecast norm, the two calculations have different interpretations. This is true even if we assume that the linear assumption is valid. As an example of how the results are different, if we calculate our results for the sensitivity of

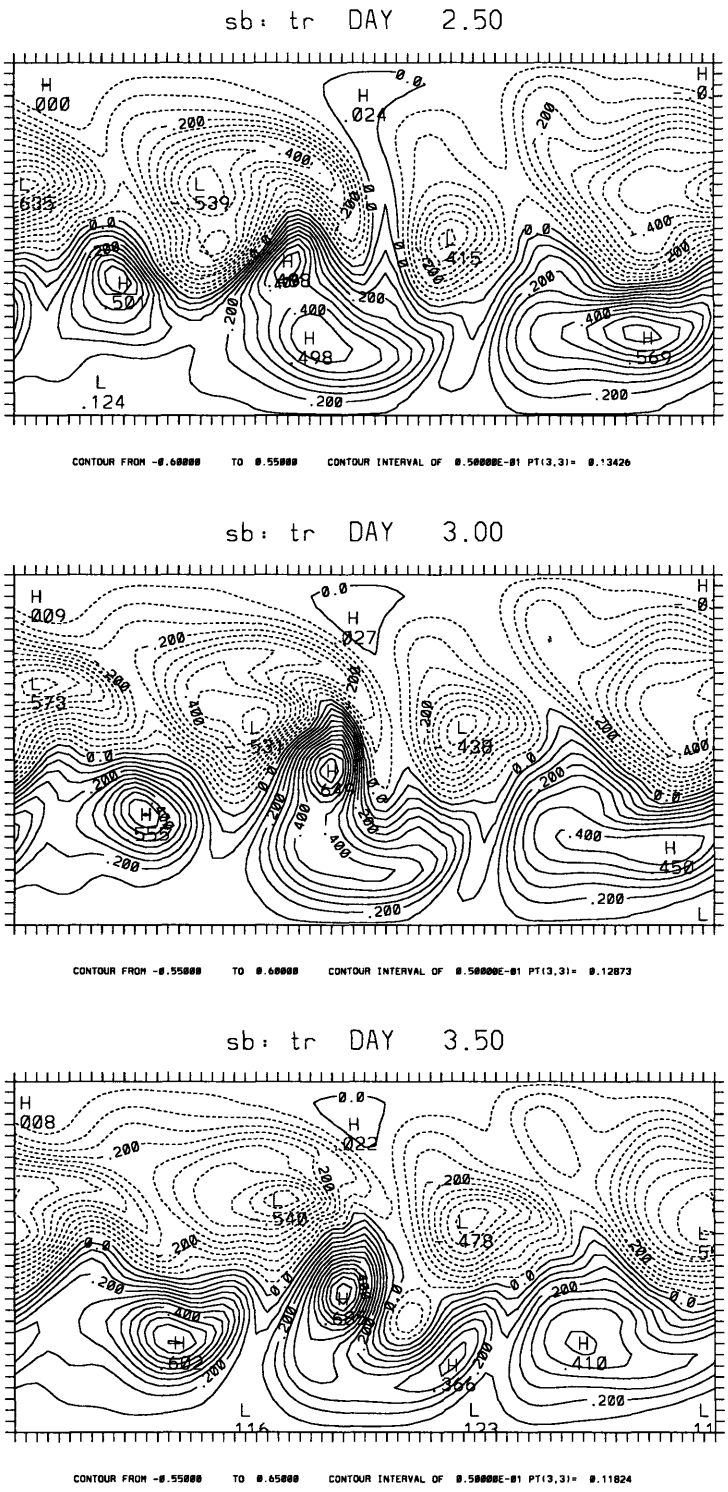



Figure 9.1: Evolution of the QG model truth state, beginning at the extra observation time for the examples shown in Figures 9.3–9.6 and 9.9–9.11. The truth state is plotted every 12 hours for 2 days. Streamfunction is shown at the lower boundary.

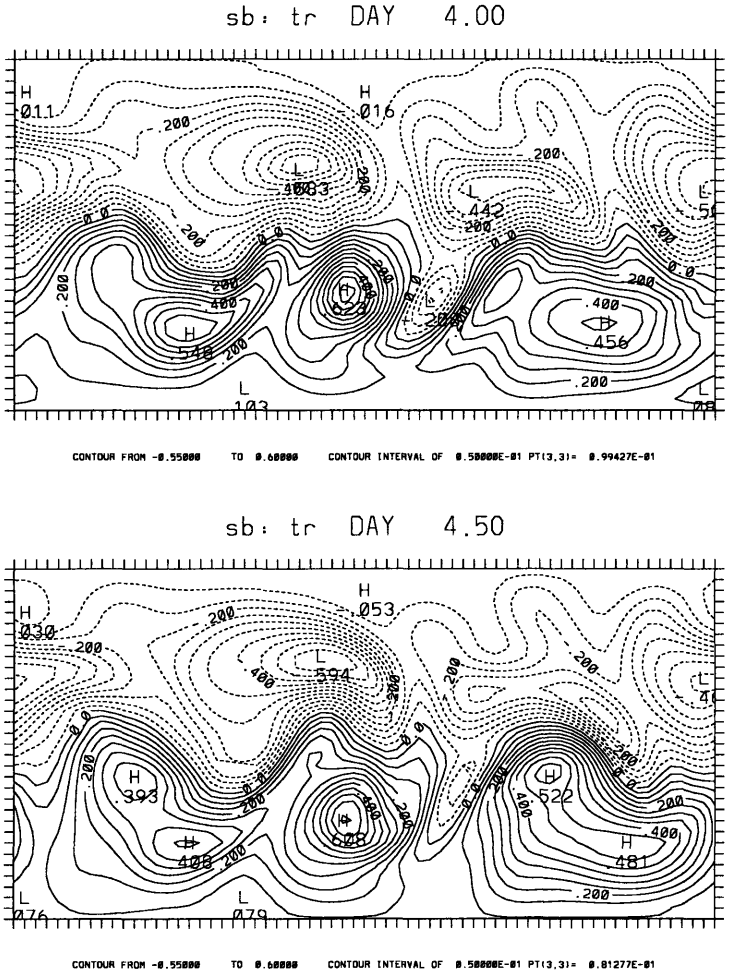



Figure 9.1, continued.

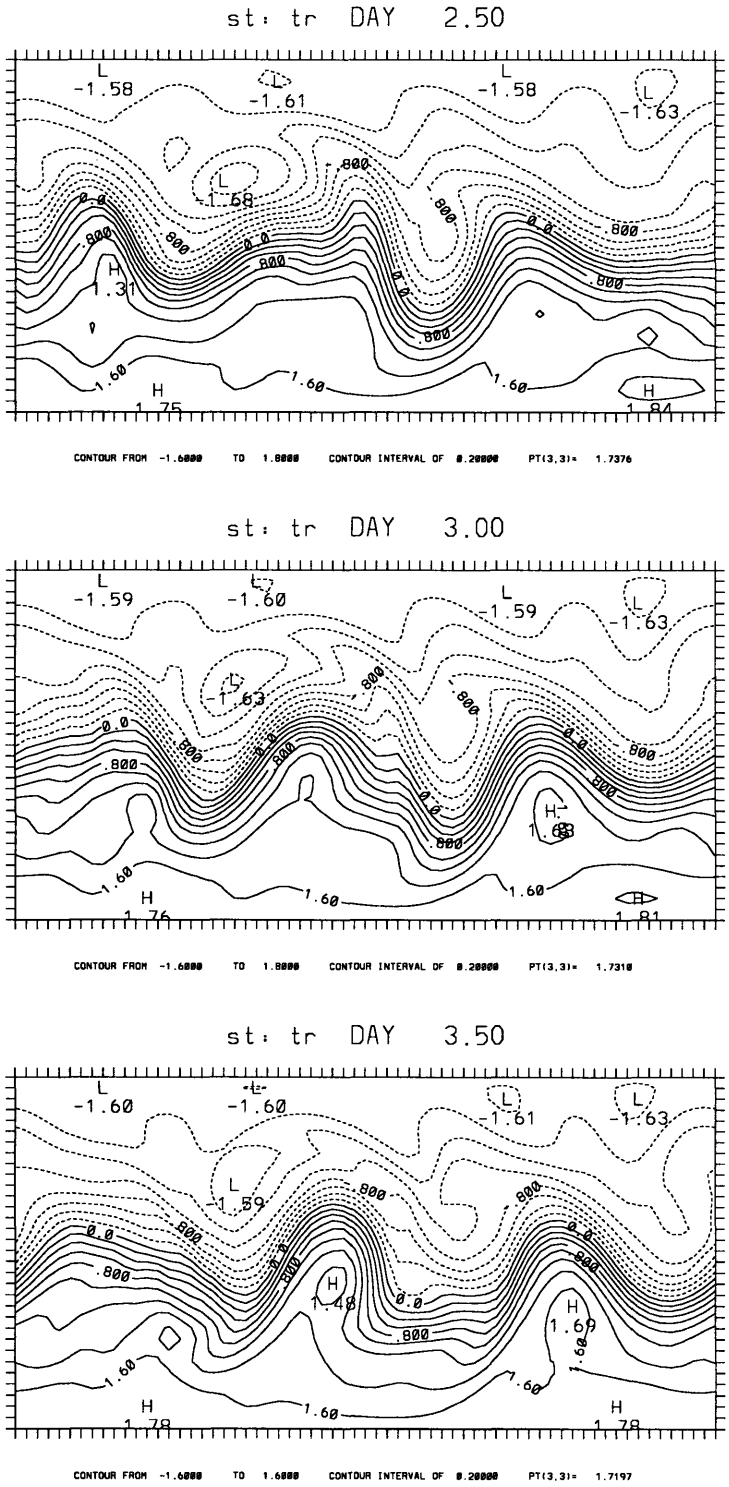



Figure 9.2: As in Figure 9.1, for streamfunction at the upper boundary.

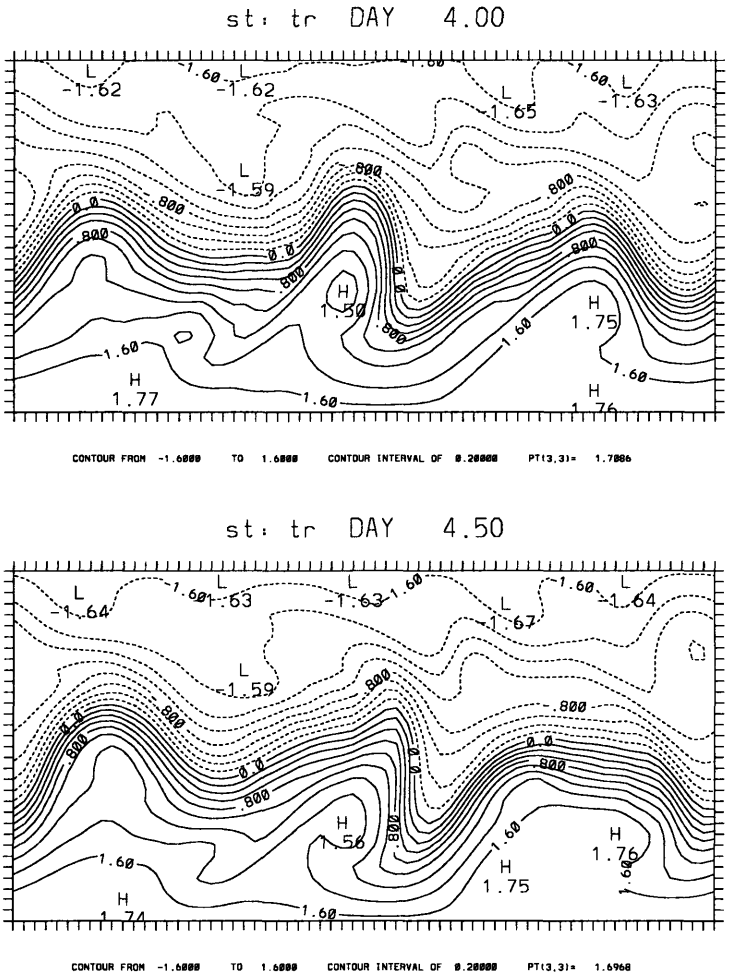



Figure 9.2, continued.

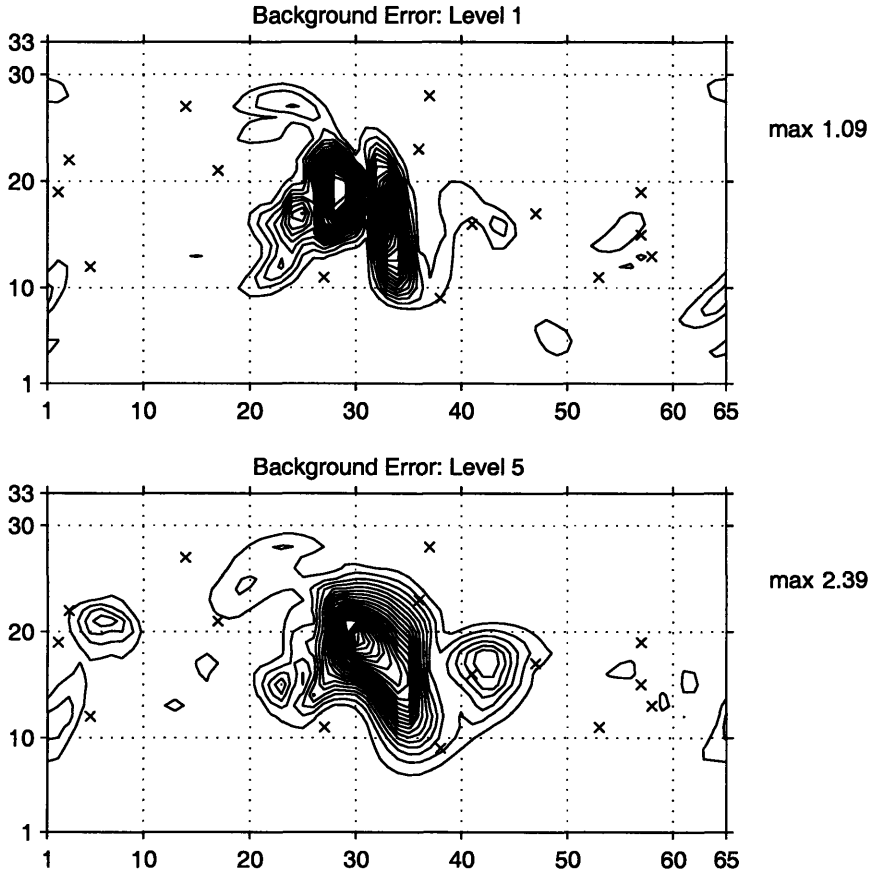



Figure 9.3: Background error (before any observations are taken) for fixed observation distribution 1 at day 2.5 at (a) the lower interior level and (b) the upper interior level, both with an energy norm. The locations of the 16 fixed observations are marked with an $\times$. The x-axis is longitude (periodic) and the y-axis is latitude. The contour interval is .05 in (a) and .1 in (b), and the maximum values are given to the right of the plot.

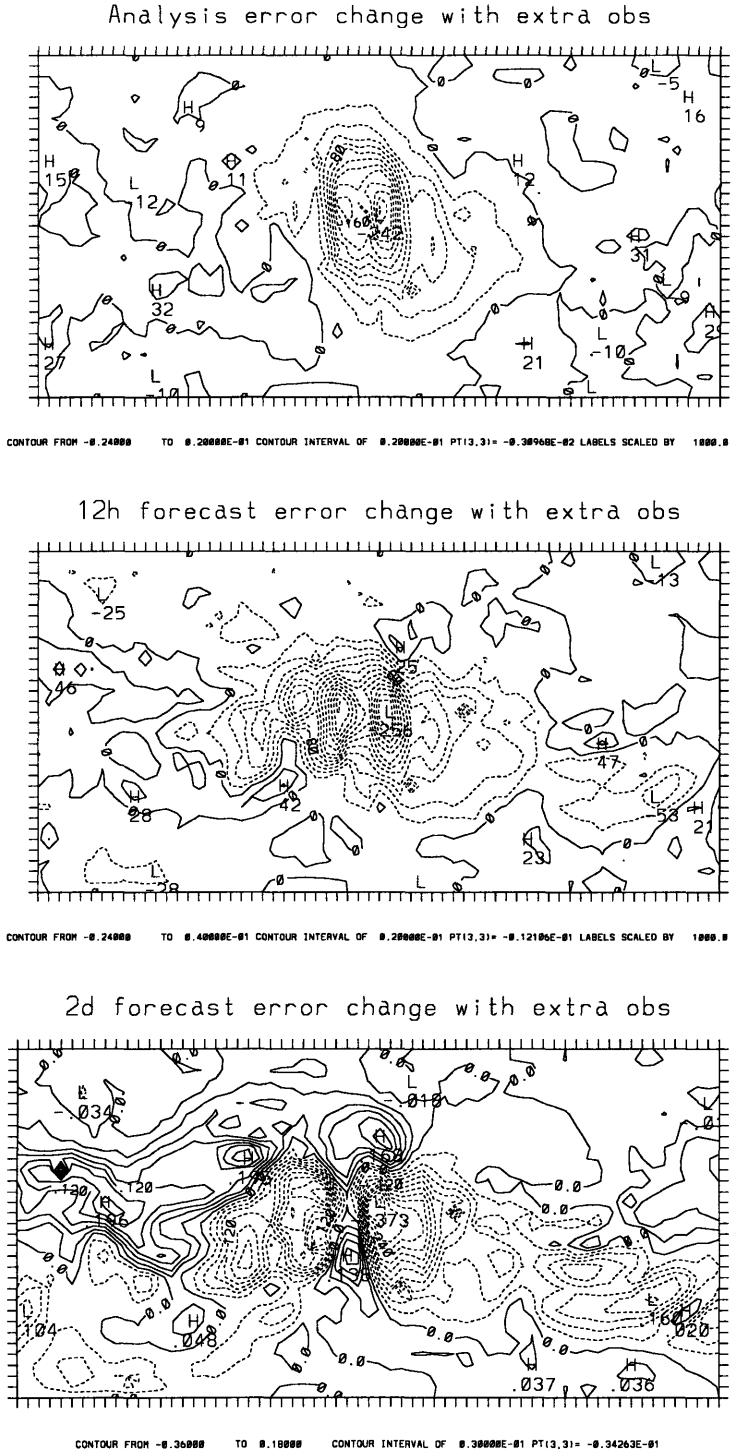



Figure 9.3, continued: Change in the global average error produced by an extra cluster 13 of observations at day 2.5, plotted at the center location for the extra observations, for the change in the (c) analysis error, (d) 12 hour forecast error, and (e) 2 day forecast error. Error change is defined as in Equation (9.1). Negative contour lines are dashed and indicate that the extra observations improve the analysis or forecast; positive contour lines are solid and indicate degradations. The contour interval is .2 in (c) and (d) and .3 in (e). The maximum and minimum values are summarized in Table 9.1.

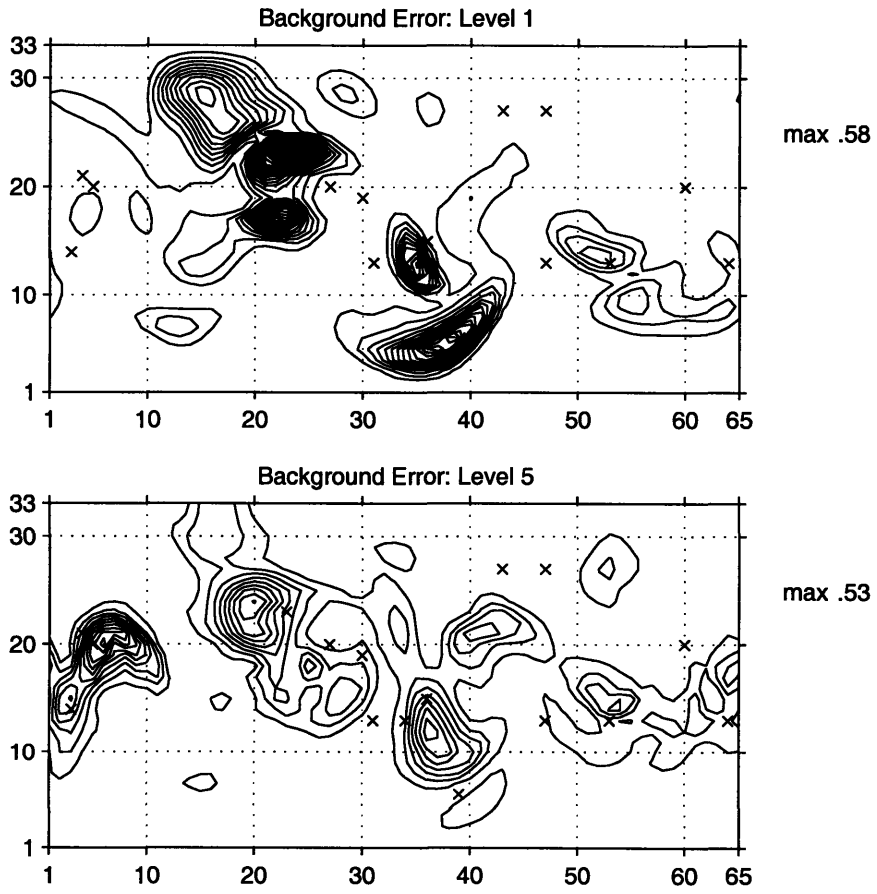



Figure 9.4: As in Figure 9.3a–b, for fixed observation distribution 5. The contour interval is .025 in (a) and .05 in (b).

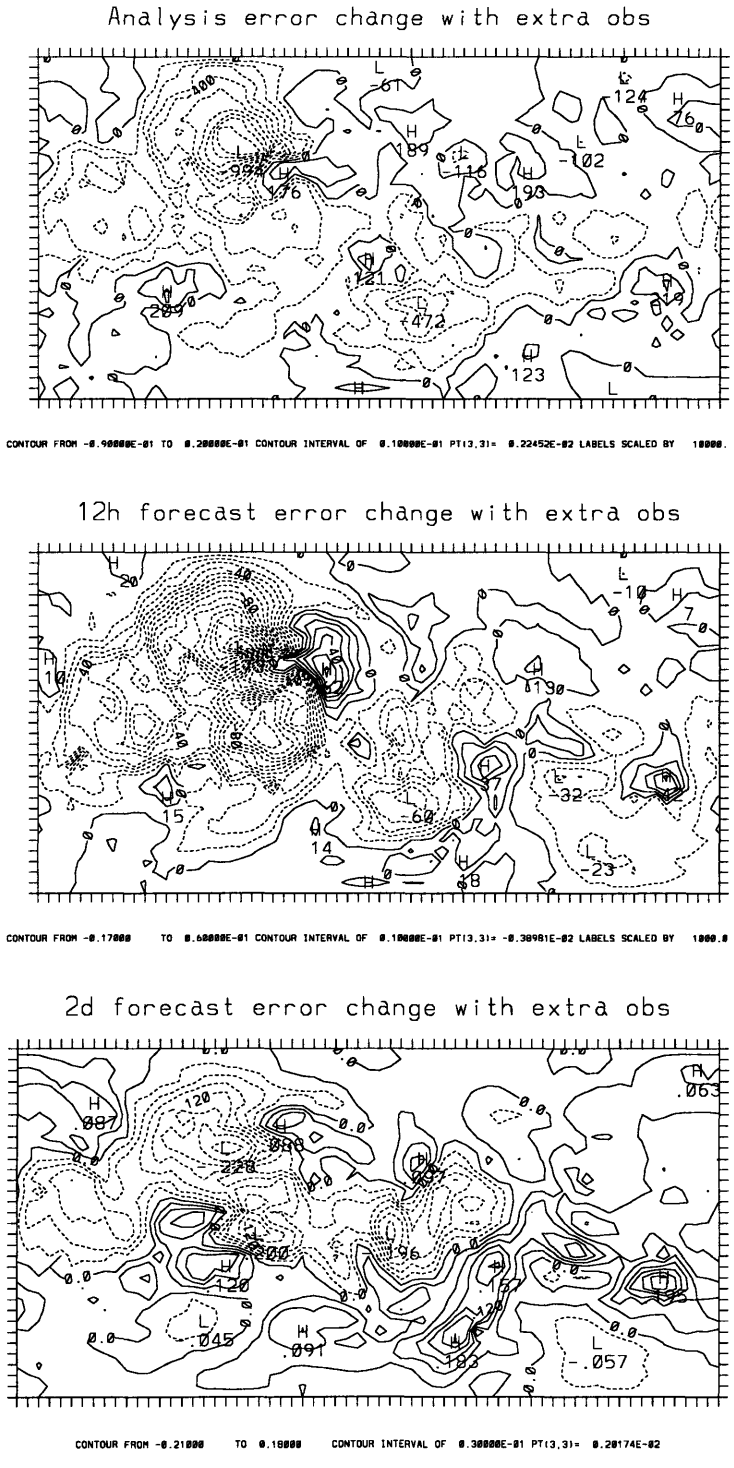



Figure 9.4, continued: As in Figure 9.3c–e, for fixed observation distribution 5. The contour interval is .1 in (c) and (d) and .3 in (e). The maximum and minimum values are summarized in Table 9.1.

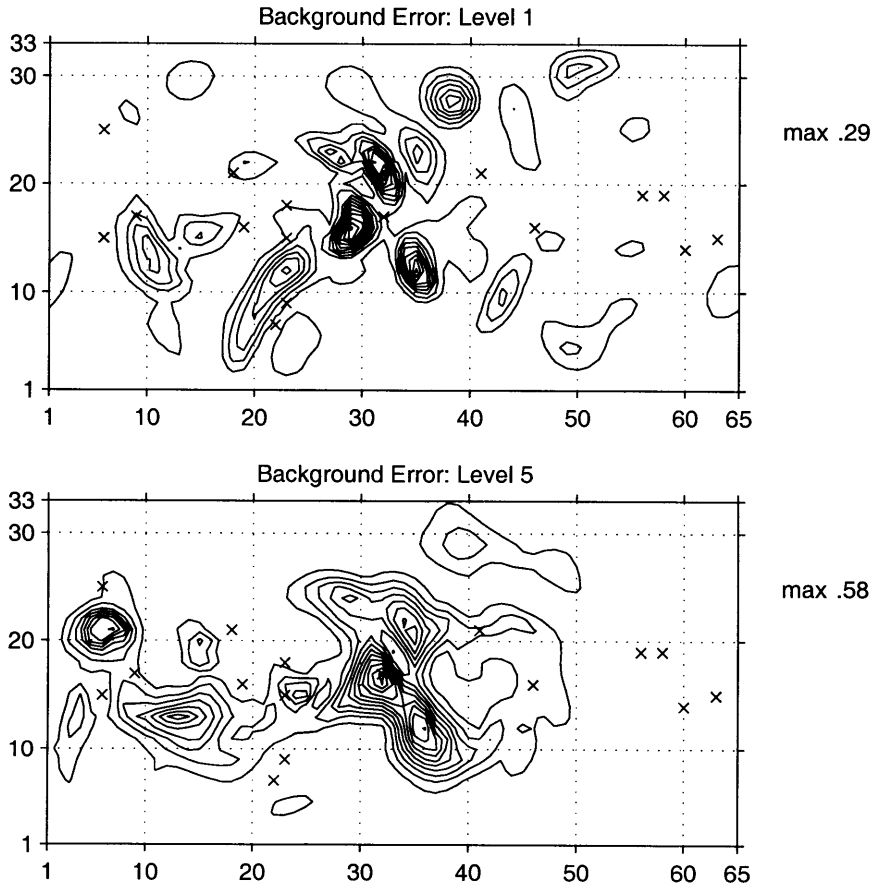
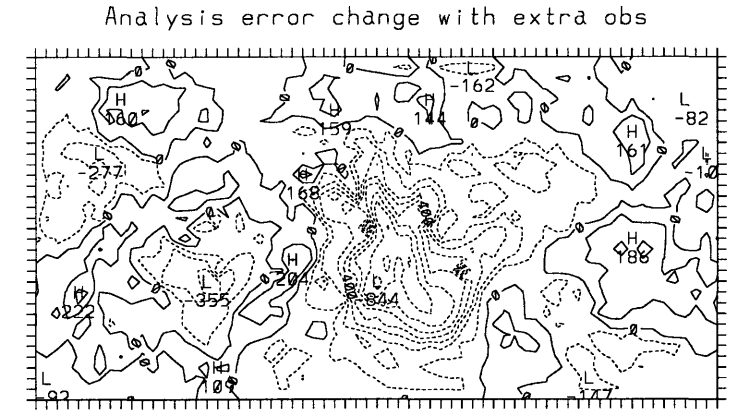
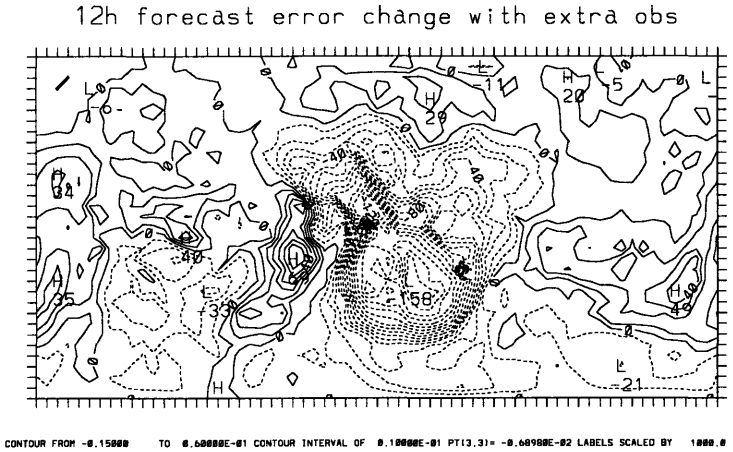



Figure 9.5: As in Figure 9.3a–b, for fixed observation distribution 2. The contour interval is .025 in (a) and .05 in (b).



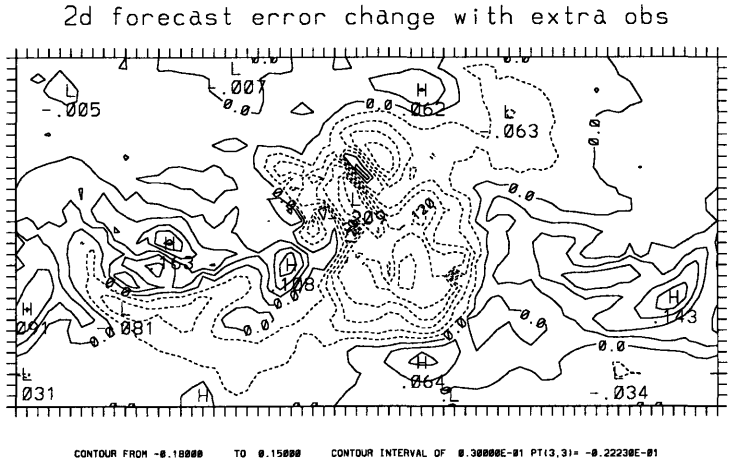




Figure 9.5, continued: As in Figure 9.3c–e, for fixed observation distribution 2. The contour interval is .1 in (c) and (d) and .3 in (e). The maximum and minimum values are summarized in Table 9.1.

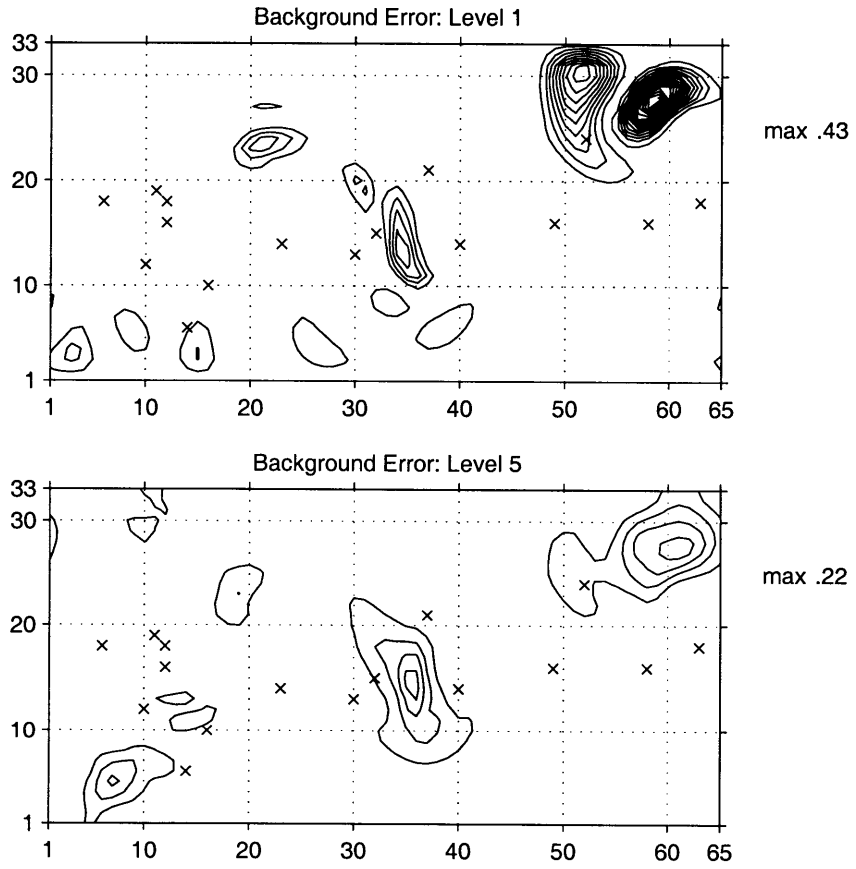



Figure 9.6: As in Figure 9.3a–b, for fixed observation distribution 6. The contour interval is .025 in (a) and .05 in (b).

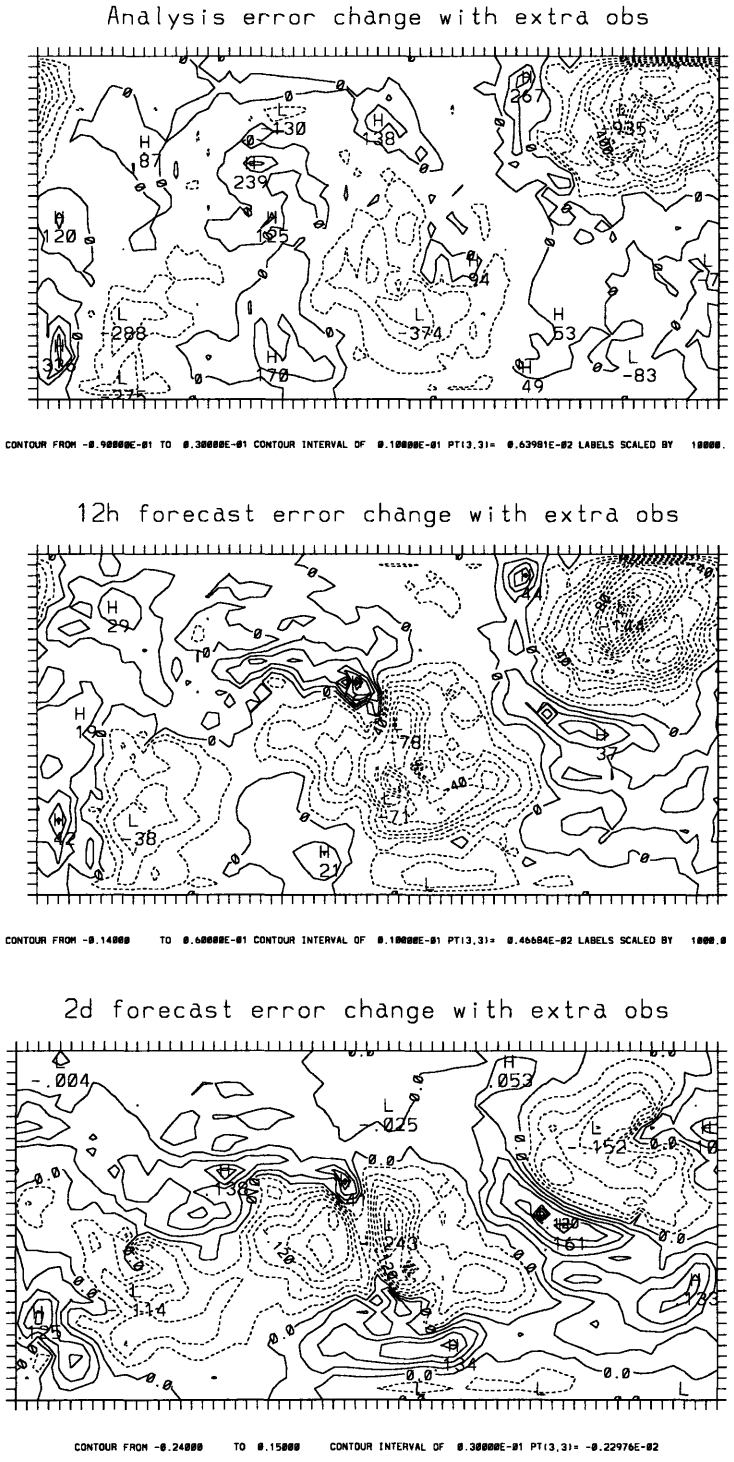



Figure 9.6, continued: As in Figure 9.3c–e, for fixed observation distribution 6. The contour interval is .1 in (c) and (d), and .3 in (e). The maximum and minimum values are summarized in Table 9.1.

Maximum improvements and degradations from extra cluster 13 observations
(fractional change in global error)

| Fixed observation | Analysis | | 12 hr forecast | | 2 day forecast | |
|---|---|---|---|---|---|---|
| distribution number | improv. | degrad. | improv. | degrad. | improv. | degrad. |
| 1 | -.24 | +.03 | -.26 | +.05 | -.37 | +.20 |
| 2 | -.08 | +.02 | -.16 | +.07 | -.21 | +.16 |
| 3 | -.12 | +.01 | -.16 | +.05 | -.20 | +.21 |
| 4 | -.08 | +.04 | -.11 | +.14 | -.14 | +.37 |
| 5 | -.10 | +.02 | -.17 | +.06 | -.23 | +.19 |
| 6 | -.09 | +.03 | -.14 | +.06 | -.24 | +.16 |

Table 9.1: Maximum improvements and degradations (global error reductions and increases) in analyses, 12 hour forecasts, and 2 day forecasts produced by adding a cluster 13 of observations to 16 fixed observations at day 2.5. Results for all 6 fixed observation distributions tested are included; more detailed results for distributions 1, 2, 5, and 6 are shown in Figures 9.3–9.6.

forecasts to observations at individual levels, because the data assimilation procedure extrapolates both horizontally and vertically, we do not see the same type of horizontal or vertical structures that adjoint sensitivity calculations often have. Therefore, even though our results calculate sensitivities of forecast errors to observations, we do not refer to them as "sensitivities" to prevent confusion. We note that although including the data assimilation produces results which are in some ways more practically useful, it also produces results which are more complex to interpret theoretically and which are specific to the data assimilation system used. Including the data assimilation also increases the computational burden significantly.

In Figures 9.3–9.6 panels a–b, we see that the background errors are very different in different cases; the background errors depend strongly on where the fixed observations have been taken in the past. In addition to the large variability in the horizontal structure of the background errors, note that the background errors tend to be more localized at lower levels than at upper levels, and that the errors may either be nearly barotropic or have significant vertical structure. Despite the differences, some of the fixed observation distributions do have significant background errors in similar regions. For example, all 4 cases shown have errors in the center of the domain, and several of the cases have relatively large background errors in similar regions in the left portion of the

domain. In regions where several different fixed observation distributions have significant background errors for the same synoptic situation, there is generally an identifiable major synoptic feature, as we see in Figure 9.2.

The similarities in Figures 9.3–9.6 a–b suggest that in a more realistic setup without knowledge of the true state, we may be able to identify some of the general regions with large background errors from information only about the synoptic situation. This requires that our estimate of the atmospheric state be reasonably accurate, so that we know what the synoptic situation is. We could try to identify these regions by subjectively evaluating the rapidly evolving synoptic systems, for example, or by evaluating a set of evolved singular vectors. However, the large differences between the figures suggest that even if we can identify one or two important regions, we will have difficulty estimating the magnitudes or structures of the background error if we do not also incorporate information about the probability of errors in the initial conditions. As demonstrated in Section 7.2, even an ensemble spread designed to estimate background errors often has difficulty identifying the particular structure of the errors in an individual case. In some situations (discussed below), it may be important to know more about the specific structure of the background error before adapting observations. Currently, it is not clear how well the background error structure can be estimated.

We would like to take more observations because intuitively we believe that more data will improve our best guess of the current atmospheric state. Unfortunately, data assimilation procedures must reconstruct a complex and rapidly-changing atmospheric state from limited observational information. This means that the benefit from extra observations will always be limited to some extent by our ability to assimilate the information into the forecast model. Data assimilation is also a statistical procedure. Therefore, even if we could develop and solve an ideal data assimilation algorithm with ideal statistics, there is always a possibility that extra observations will degrade an analysis instead of improving it. Panel c in Figures 9.3–9.6 clearly shows that in many situations, adding observations in certain locations in fact degrades the analysis.

It is sometimes stated that more data "should" improve the analysis, and that if extra observations do not improve analyses most of the time, there must be something wrong with the data assimilation system. Certainly with the simplified statistics, our data assimilation does not perform as well as an more sophisticated data assimilation system, or even a better 3DVAR data assimilation, would. Looking at panel c and Table 9.1 more closely, however, we see that our data assimilation system, despite its statistical nature and its simplified statistics, performs quite well in two ways. First of all, at the

majority of locations, adding observations does improve the analysis. In addition, the maximum analysis degradations are significantly smaller than the maximum analysis improvements. The distributions and magnitudes of the changes vary quite a bit from case to case, but in nearly all cases the data assimilation system performs well in both of these ways. In Section 9.4, we present statistical results which indicate that this occurs in general.

Comparing panels a and b with panel c, we also see that extra observations are much more likely to reduce the analysis error significantly if they are taken in regions with a large background error. The comparisons are even closer for background errors and analysis changes at individual levels. This is so prevalent that in most cases, we could even determine which regions have large background errors (although not the structure of the errors) simply by looking at whether extra observations in that region actually significantly improve the analysis. Unfortunately, this is not possible in more realistic situations since we cannot verify the analysis versus the true state.

The analysis benefits more from observations in regions with large background errors for two reasons. First, the data assimilation system tends to correct large background errors better than it corrects small background errors. Second, we measure the influence of the observations with a global average error norm, which emphasizes changes in regions with large errors. These results suggest that the data assimilation system actually does fairly well at reducing the analysis error if we take extra observations in regions where the data assimilation procedure is best able to use the information.

Comparing the analysis error changes in panel c with the forecast error changes in panels d and e, we see that even small changes in initial conditions can produce significant changes in forecasts. Often, changes amplify rapidly during the first 12 hours, but error magnitudes also increase significantly in longer range forecasts. In regions where the influence of extra observations grows rapidly, we would expect singular vectors or sensitivity calculations with a domain-averaged forecast error norm to have significant amplitude at the initial time, indicating that changing the initial conditions in this region will have a large (positive or negative) influence on the forecast. Unfortunately, both small improvements and small degradations in the analysis grow rapidly. Consequently, while in most cases the magnitude of the improvements far outweighs the magnitude of the degradations at the initial time and in the 12 hour forecast, by the 2 day forecast the improvements and the degradations are nearly as frequent and are of the same order of magnitude. Again, we discuss this on a statistical basis in Section 9.4.

With a perfect model, forecast errors can be caused only by errors in the initial

conditions. One might infer therefore that improving analyses will improve forecasts. Figures 9.3–9.6 c–e show, however, that even in this idealized system, extra observations can improve the analysis but degrade forecasts. This can occur if the extra observations reduce the global or local average analysis error but change the structure of the analysis error in some way that amplifies into a forecast error. For example, the extra observations might improve the global analysis, but at the same time they could increase the projection of the analysis error onto one of the leading singular vectors. Even if this increased projection is very small, it will grow rapidly and may significantly degrade the forecast.

Unfortunately, adjoint-based calculations can only tell us which changes will grow rapidly, not which will lead to rapidly growing improvements or degradations. It is left to the data assimilation system to determine which sign of a rapidly growing structure we should add to produce a positive change. The 3DVAR data assimilation system has precisely the most difficulty in resolving structures similar to those produced by adjoints, which often have small scale horizontal and vertical structure and small amplitude at initial time. A 4DVAR or other more sophisticated data assimilation scheme may reduce the likelihood that assimilated data will improve the analysis but degrade the forecast. However, in order to adapt observations effectively to improve individual forecasts, more work is needed on designing data assimilation systems and adaptive strategies which minimize the risk of forecast degradations.

As discussed earlier, the results are complex and are different in each situation. Thus, even if we know the true dynamical situation and the actual errors in the initial conditions, we have found it difficult to predict subjectively how a specific set of extra observations will influence an analysis and forecast. Each of the examples shown, however, demonstrates a particular feature of how extra observations can interact with analyses and forecasts in this system, features which we have also seen in other cases. For the first two examples, we discuss the two ways in which we have been able to predict, in some cases, that extra observations are likely to degrade a forecast. These predictors only suggest a large risk of forecast degradation; they do not always lead to degradations when observations are actually taken in the region. The second two examples are briefly described to illustrate some difficulties with trying to attribute analysis and forecast improvements and degradations. As in all of the cases studied, in each of these four examples there are a variety of interesting ways in which extra observations change forecast errors, many of which we do not discuss.

Fixed observation distribution 1 (Figure 9.3) is an example with one very large,

dominant background error (note the change in contour interval between panels a–b in Figure 9.3 and in Figures 9.4–9.6). In the region with large background errors, extra observations improve the analysis and the 12 hour forecast up to about 25%. In much of the region, extra observations also improve the 2 day forecast, up to about 37%. In areas with a local minimum in background error within the larger region, however, even though extra observations improve the analysis, they significantly degrade the 2 day forecast. In a few areas, only a small horizontal displacement (several hundred kilometers) of the center location for the extra observation cluster can change a much better forecast into a much worse forecast. The degradations most likely occur because the observation cluster only gathers data near the edge of the region with a large background error, and then the data assimilation system extrapolates incorrectly (recall that these results are for cluster 13; these degradations are discussed further for different cluster patterns in Section 9.5). This extrapolation effect commonly (but not universally) occurs when background errors have a lobed structure similar to the one in Figure 9.3. When other fixed observation distributions have large background errors in this region for the same synoptic situation, the lobed structure of the background errors is different, and the degradations occur at different locations. Therefore, when adapting observations in some situations, it may be important to know not just a general region with large background errors, but also the more specific locations within that region where they are larger and smaller. Alternately, we may need to develop a more sophisticated data assimilation system or a more sophisticated adaptive strategy to circumvent this problem.

For the fixed observation distribution 5 example (Figure 9.4), we focus on one specific geographic region. Between longitudes 30–40 and latitudes 1–10, there is a significant background error at lower levels, but only a very small background error at upper levels. By the 2 day forecast, we can see that observations in this region have significantly degraded the forecast. This also occurs in other cases; in regions where the background error is large at lower levels but negligible at upper levels, extra observations often (but again not always) degrade the forecast. Degradations occasionally, but not as often, occur in regions where the background error is significant only at upper levels. These degradations occur at least in large part because, as demonstrated in Section 3.5, the data assimilation system used in this study is least effective at correcting background errors at lower levels. Therefore, to prevent these degradations we need either to avoid observing in regions with background errors primarily at lower levels, or to improve our data assimilation system so that it is able to correct these errors without degrading the analysis structure or degrading the analysis at other levels.

In fixed observation distribution 2 (Figure 9.5), the largest background errors at day 2.5 are in the middle of the domain and in the left portion of the domain. In both regions, extra observations reduce the analysis error by approximately the same amount. By the 2 day forecast, however, observations taken at the initial time in the middle part of the domain significantly reduce the forecast error, while many test observations taken in the left portion of the domain significantly increase the forecast error. Based on the results from the previous two examples, both of the regions have potentially problematic structure in the background error. Thus, even if we were able to assess the actual background errors prior to testing the observations, we would be as likely to expect assimilated extra observations to perform poorly in the center of the domain as in the left half of the domain. Without examining the specific analysis increments and forecast changes produced by observations in each region, it is difficult to understand why some sets of observations improve forecasts and others degrade forecasts. This difficulty with attributing degradations frequently occurs in the cases studied.

Fixed observation distribution 6 (Figure 9.6) has several fixed observations in the middle and on the left side of the domain, in both of the important regions identified by the other examples. Because in these regions the observations continuously reduce the errors, the background errors are small at day 2.5. The largest background errors are in the upper right hand corner of the domain, in a region which does not have significant background errors in any of the other fixed observation distributions tested. Not only are the background errors in this region significant, but observing in the region has a significant effect on the forecasts. This suggests that correcting the background error in this region has a large influence, even though observing in this region had a small or no effect in any of the other examples shown. Even in this case, however, it is not only correcting the most significant background errors that has a significant effect. Correcting much smaller background errors in other regions of the domain also changes the forecast error by a similar amount.

## 9.4 Statistical results over several examples

As the examples in Section 9.3 show, the influence of extra observations varies significantly from case to case. Therefore, we also present some statistics gathered over all 102 cases, including the cases shown in Figures 9.3–9.6. With observations tested at $1984 - 16 = 1968$ locations in each case, we have tested analyses and forecasts for

$102 \times 1968 \approx 2 \times 10^5$ extra sets of cluster 13 observations. Although the statistics are accumulated over only a limited number of synoptic situations, we believe that our sample is large enough to be fairly representative for this simulated system.

The histogram in Figure 9.7 illustrates how often we can expect any extra cluster 13 observations added to 16 fixed observations to degrade analyses and forecasts. The error change is again defined according to Equation (9.1), with positive error changes indicating degradations. The histogram also illustrates how often we can expect the extra observations to improve or degrade the analyses and forecasts by more than a certain amount. In Figure 9.7, we see first of all that extra observations are much more likely to improve than degrade the analysis, and that they are much more likely to do so by a significant amount. As discussed in the previous section, this indicates that the data assimilation system used in this study, despite its faults, is reasonably successful at reducing analysis errors.

When forecasts are made from these analyses, as we saw in the individual examples, both the improvements and the degradations tend to increase. In the 12 hour forecasts, degradations occur slightly more frequently than in the analysis, but the degradations are usually not very severe. By the 2 day forecasts, on the other hand, the distribution of error changes with extra observations is nearly symmetric; extra observations are only slightly more likely to improve the 2 day forecast than to degrade it. This is true both for any improvements or degradations, and for improvements or degradations larger than a certain amount. Thus, as discussed in the previous section, although the extra observations do generally improve analyses and short term forecasts, this does not mean that extra observations generally improve longer term forecasts. This is true even with a perfect forecast model.

Figure 9.8 shows the same histogram as in Figure 9.7, but with the y-axis on a logarithmic scale to allow us to look at the tails of the distribution. From this view, we again see that extra cluster 13 observations are much more likely to improve than to degrade analyses and 12 hour forecasts by a significant amount. The same sets of extra observations are about equally likely to improve or to degrade 1 day forecasts very significantly. Despite the approximately symmetric distribution shown in Figure 9.7 for the most frequent error changes in the 2 day forecasts, however, Figure 9.8 suggests that extra observations are actually more likely to degrade forecasts very severely (by more than 30%) than to improve them by a similar percentage. Due to the small number of cases with these error changes, we are not certain that the results are statistically significant. However, they do suggest that when adapting observations to improve forecasts

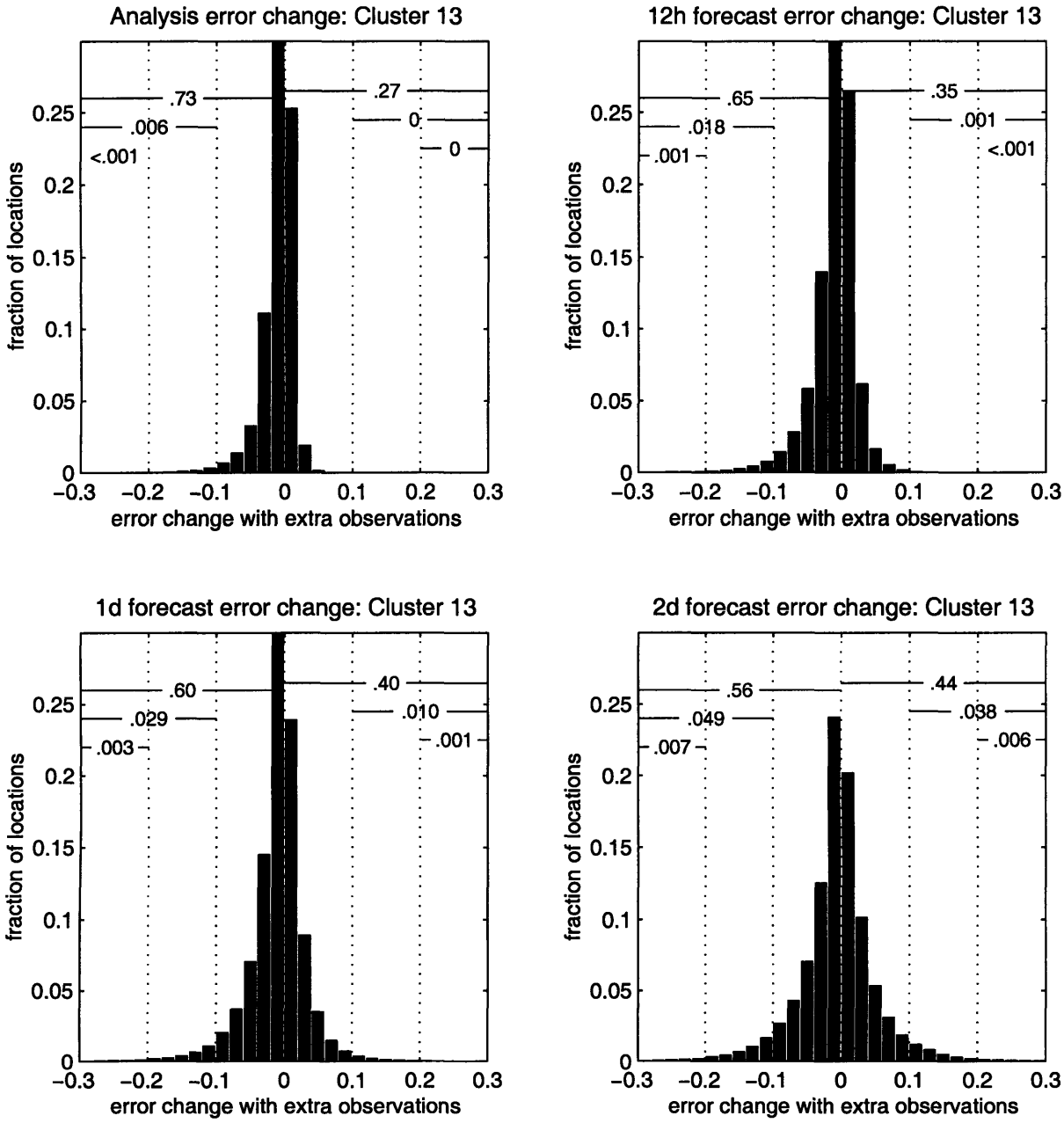



Figure 9.7: Histogram of the likelihood that an extra cluster 13 of observations, added anywhere in the domain to 16 fixed observations, changes the global error in (a) the analysis, (b) the 12 hour forecast, (c) the 1 day forecast, or (d) the 2 day forecast, by the indicated fraction (as defined in Equation (9.1). Recall that negative error changes indicate improvements and positive error changes indicate degradations. The statistics are accumulated over 8 days (17 synoptic situations) and 6 distributions of fixed observations. The numbers near the top of the histogram indicate the fraction of the locations at which extra observations changed the error by the amount in the indicated band. For example, in panel a, at 73% of the locations extra observations reduced the analysis error by any amount, while at .6% of the locations extra observations reduced the analysis error by more than 10%. A 0 means that there were no locations in that section of the histogram; <.001 means that there were some locations, but less than .1% of the total number tested.

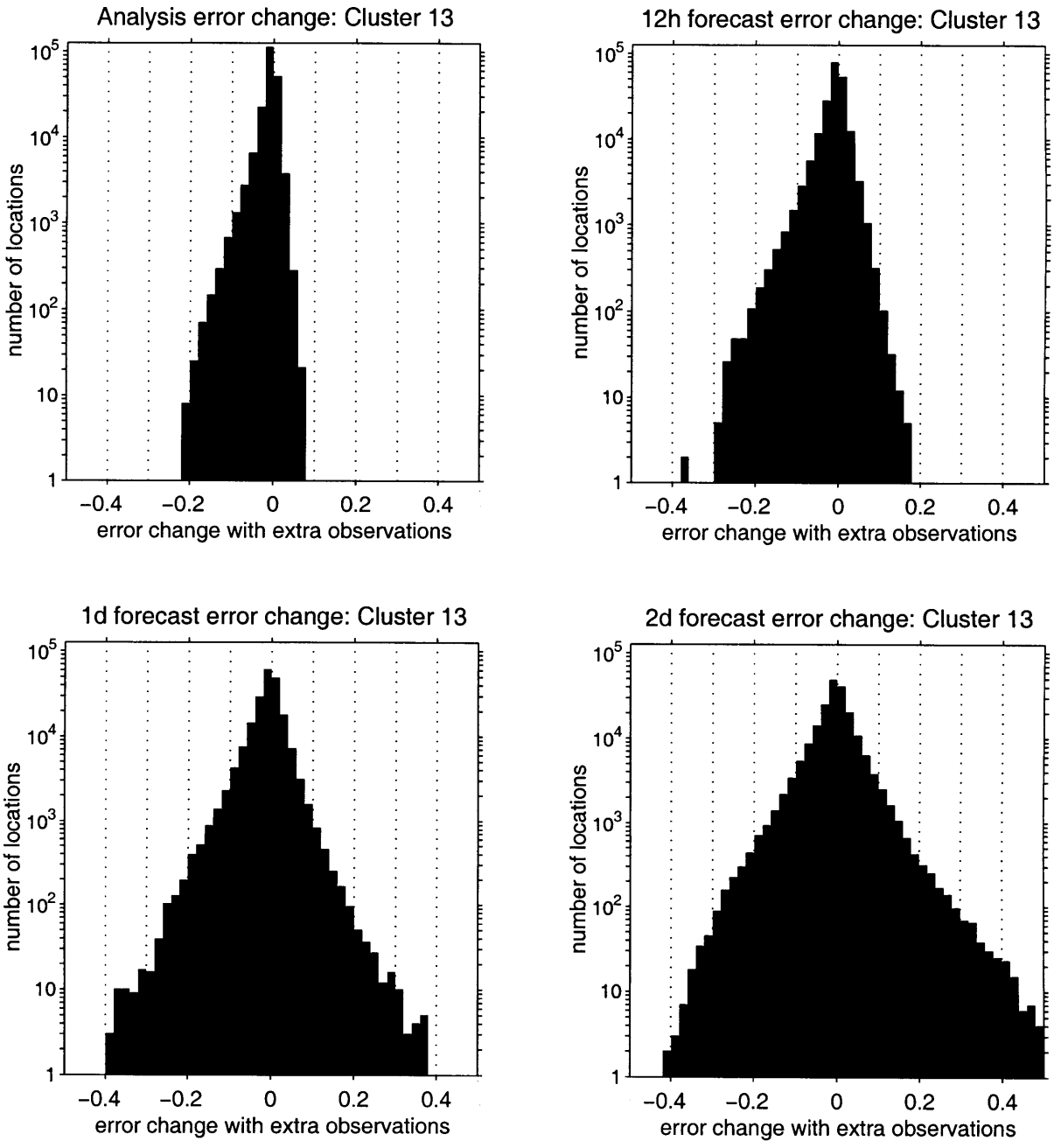



Figure 9.8: As in Figure 9.7, but with the y-axis the number of locations on a logarithmic scale. Note the change in the x-axis scale.

at longer lead times, there is not only a significant risk of degrading the forecast, but there may also be a small risk of very severely degrading the forecast.

## 9.5 Sensitivity to observation cluster pattern

In Section 9.3, we discuss how taking extra cluster 13 observations at a location with a local minimum in background error (within a larger region with large background errors) can degrade forecasts. We attribute this to the observation cluster observing at the edge of a significant feature, and the data assimilation system then incorrectly extrapolating the information incorrectly into the rest of the feature. This extrapolation effect can also occur when an extra cluster of observations is taken in the vicinity of a large background error. Since the analysis increment from a single observation is spread over a smaller area, these degradations are likely to occur less often for single observations

In general, we believe that an extra cluster of observations gives the data assimilation system more information about the structure of analysis errors, and thus that an extra cluster of observations is more likely to improve analyses and forecasts. Therefore, we chose to show the majority of the examples for extra clustered observations instead of for single observations. In this section, to explore how the size of the observation cluster affects the potential for forecast improvement and forecast degradation, we present results similar to those in Sections 9.3 and 9.4 for different cluster patterns. We do not discuss the results in great detail, but rather present them to provide a general idea of how results from adapting observations in individual situations may or may not depend on the size of the observation cluster.

The results presented in previous sections tested a cluster 13 of observations, 13 observations at alternating gridpoints in a 1000 km $\times$ 1000 km grid, centered at each location. Here we test single observations and 2 other cluster patterns. Cluster 1 refers to single observations tested at each location. Cluster 3, depicted along with cluster 13 in Figure 6.4, is 3 observations in a triangle around the observation location. Cluster 81 is 81 observations in a 9 observation $\times$ 9 observation, i.e. a 2000 km $\times$ 2000 km, square around the observation location. As described in Section 6.2, all observations in each cluster pattern are taken simultaneously.

Figures 9.9, 9.10, and 9.11 show the influence of observations centered at different locations in the domain for the same time and the same fixed observation distribution

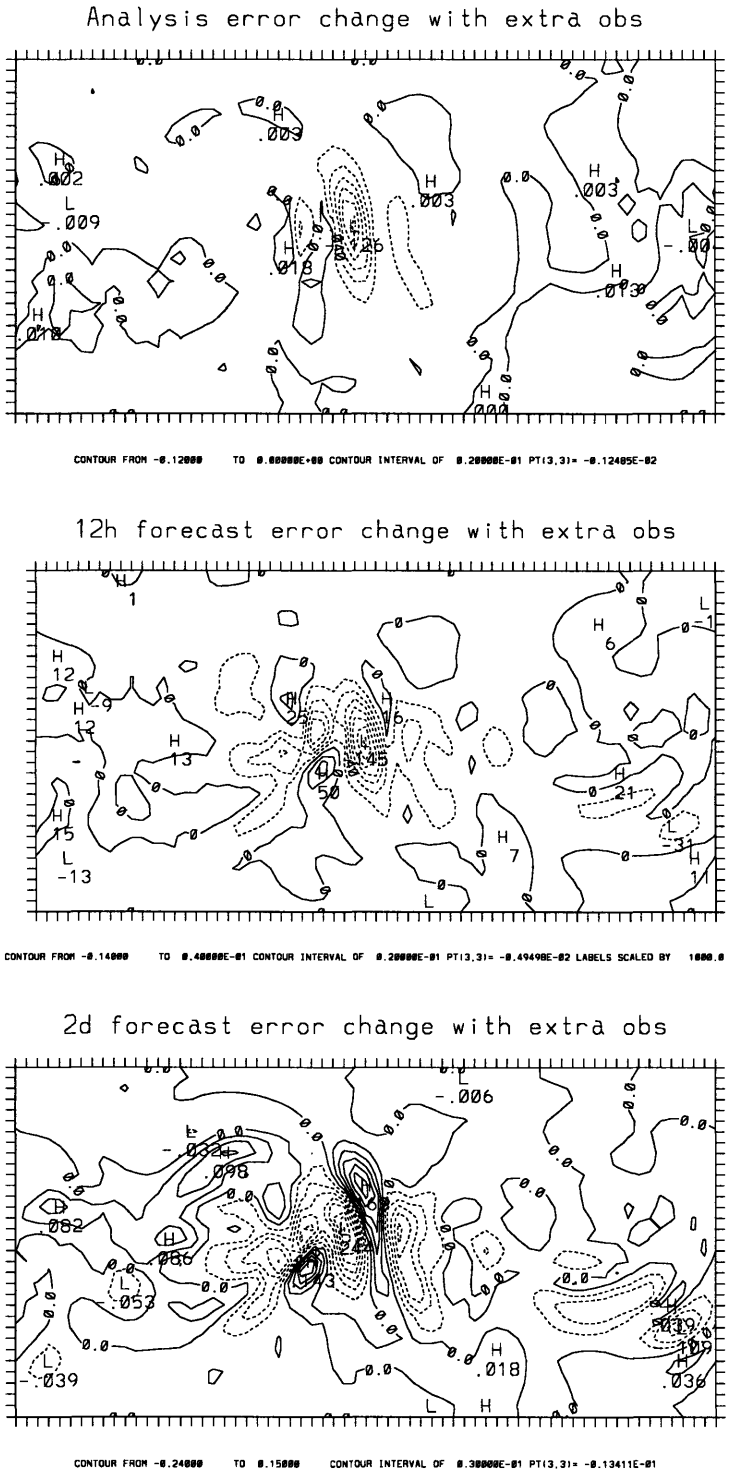



Figure 9.9: As in Figure 9.3c–e, for fixed observation distribution 1 and extra single (cluster 1) observations taken at day 2.5. The background error before any observations are taken is the same as in Figure 9.3a–b. The contour interval is .2 in (c) and (d), and .3 in (e). The maximum and minimum values are summarized in Table 9.2.

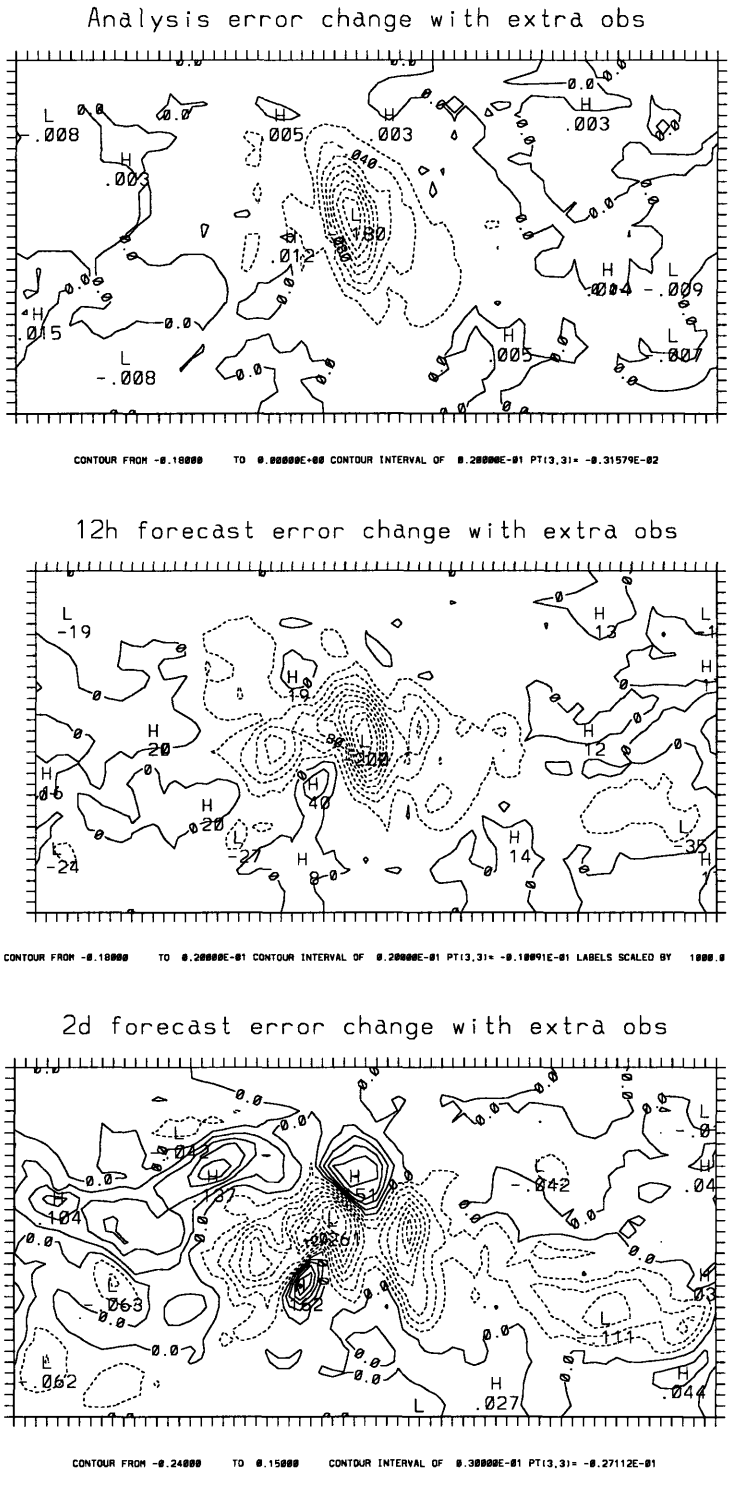



Figure 9.10: As in Figure 9.3c–e, for fixed observation distribution 1 and extra cluster 3 observations taken at day 2.5. The contour interval is .2 in (c) and (d), and .3 in (e). The maximum and minimum values are summarized in Table 9.2.

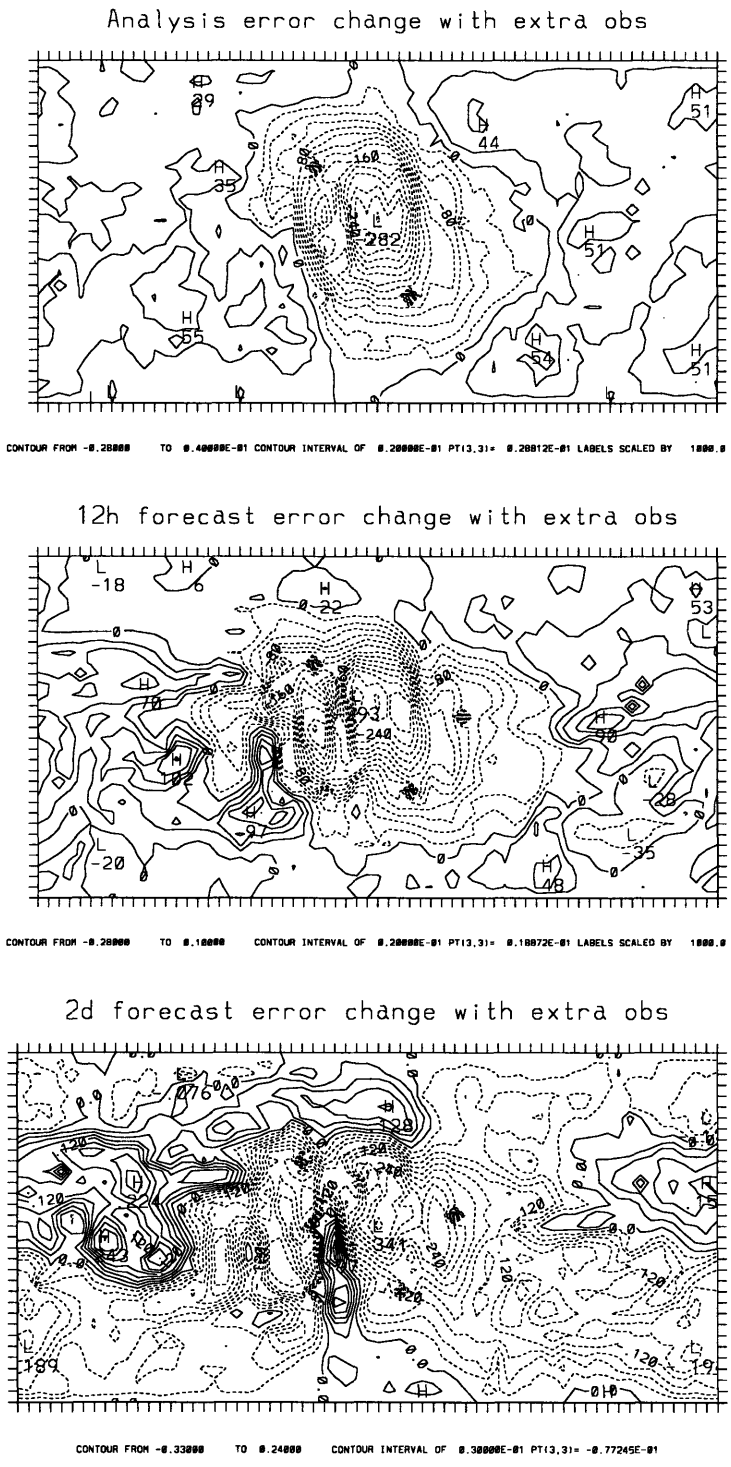



Figure 9.11: As in Figure 9.3c–e, for fixed observation distribution 1 and extra cluster 81 observations taken at day 2.5. The contour interval is .2 in (c) and (d), and .3 in (e). The maximum and minimum values are summarized in Table 9.2.

Maximum improvements and degradations from observations added to distribution 1 (fractional change in global error)

| | Analysis | | 12 hr forecast | | 2 day forecast | |
|---|---|---|---|---|---|---|
| Cluster pattern | improv. | degrad. | improv. | degrad. | improv. | degrad. |
| 1 | -.13 | +.02 | -.14 | +.05 | -.24 | +.16 |
| 3 | -.18 | +.02 | -.20 | +.04 | -.26 | +.16 |
| 13 | -.24 | +.03 | -.26 | +.05 | -.37 | +.20 |
| 81 | -.28 | +.06 | -.29 | +.10 | -.34 | +.24 |

Table 9.2: Maximum improvements and degradations (global error reductions and increases) in analyses, 12 hour forecasts, and 2 day forecasts produced by observations added to fixed observation distribution 1 at day 2.5. The cluster 13 results are the same as in Table 9.1; the other results are for the extra observations taken either singly (cluster 1), with cluster 3, or with cluster 81. More detailed results are shown in Figures 9.3 and 9.9–9.11.

(the same background error) as in Figure 9.3, but for single, cluster 3, and cluster 81 observations, respectively. Table 9.2 presents the maximum improvements and degradations in Figures 9.3 and 9.9–9.11, similar to Table 9.1. Although the general features described in Section 9.3 for cluster 13 observations are similar for all of the cluster patterns, there are some differences in the results. First of all, as one would expect, larger observation clusters tend both to produce larger analysis and forecast error changes and to produce large analysis and forecast error changes from observations at more locations. The locations at which extra observations improve or degrade forecasts are often similar among the cluster patterns tested, particularly for the 2 smaller clusters. In some situations, however, one cluster pattern improves the forecast with observations at a certain location, and another cluster pattern degrades the forecast with observations at the same location in the same situation. In some situations, there is also a tendency for a region where observations degrade forecasts to occur further away from a region of large background error for a larger pattern than it does for a smaller pattern. This occurs, for example, in the northern half of the middle longitudes of the domain. This agrees with our interpretation that some of the degradations occur because observations are taken at the edge of a significant region.

Figures 9.12, 9.13, and 9.14 show the same histogram of how often we can expect extra observations to change analyses and forecasts by a certain amount as in Figure 9.7, but

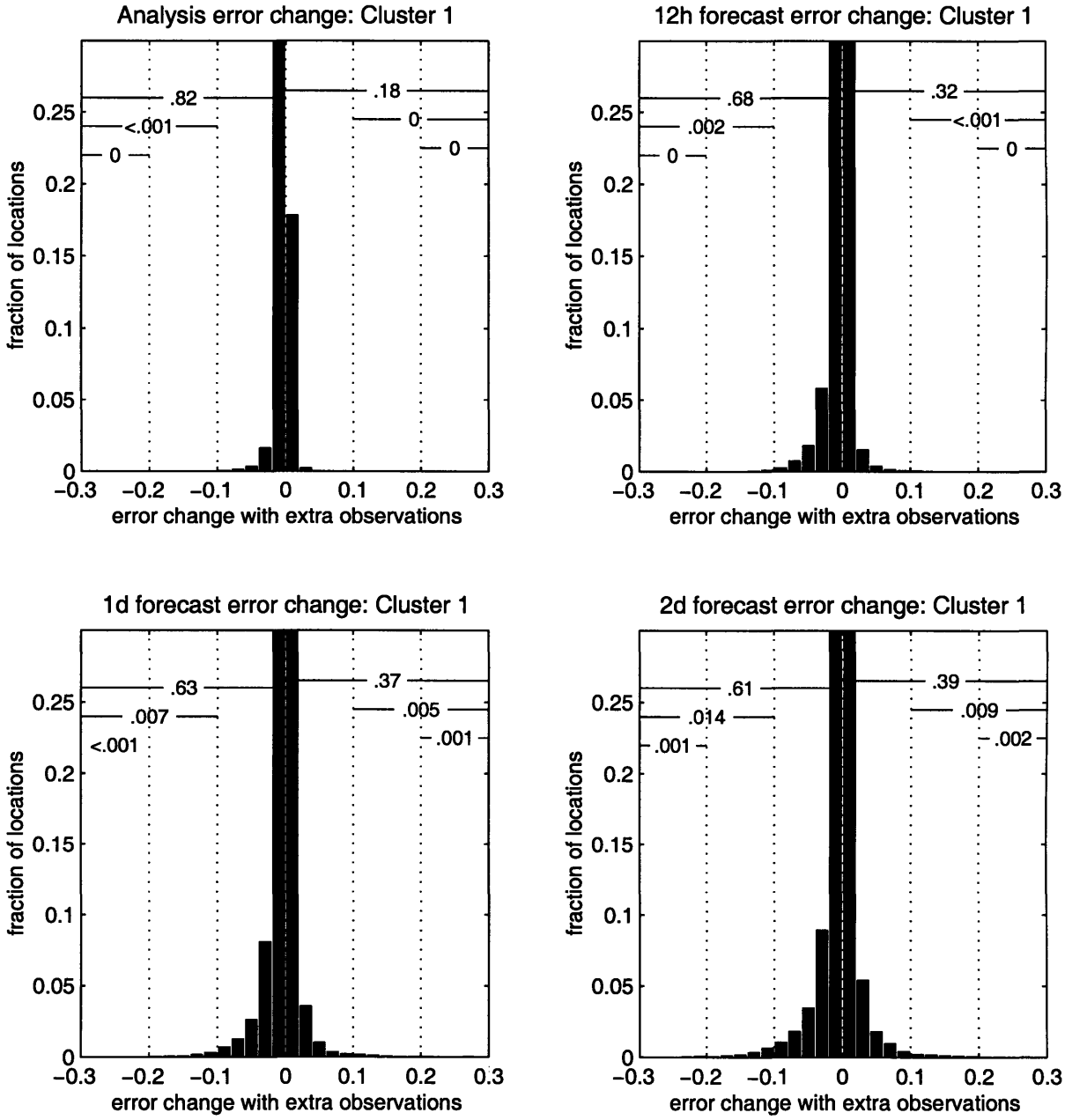



Figure 9.12: As in Figure 9.7, for an extra cluster 1 (single) observation tested at each location. The statistics are accumulated over only one distribution of the fixed observation locations (distribution 1).

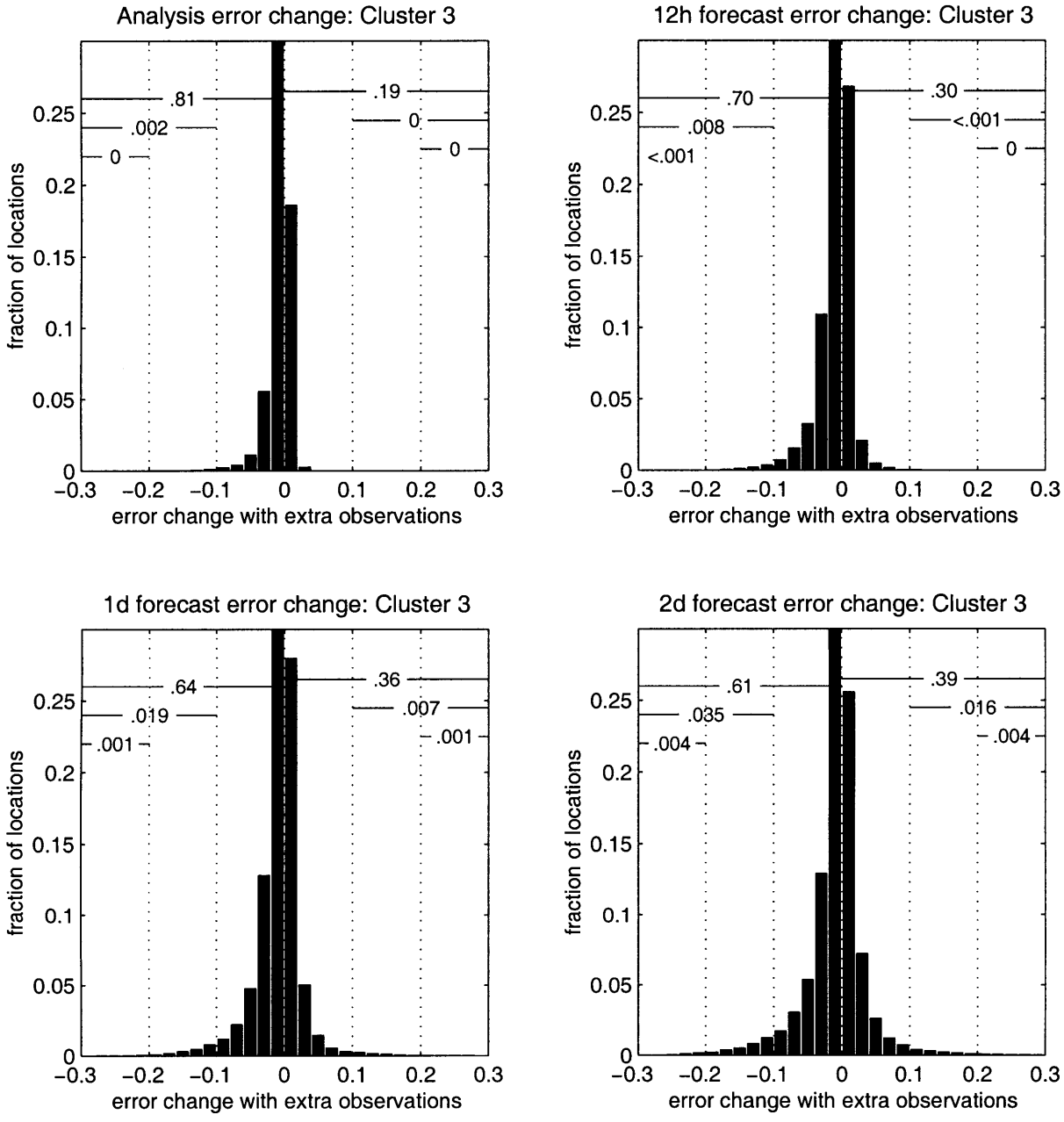



Figure 9.13: As in Figure 9.12, for an extra cluster 3 of observations tested around each location.

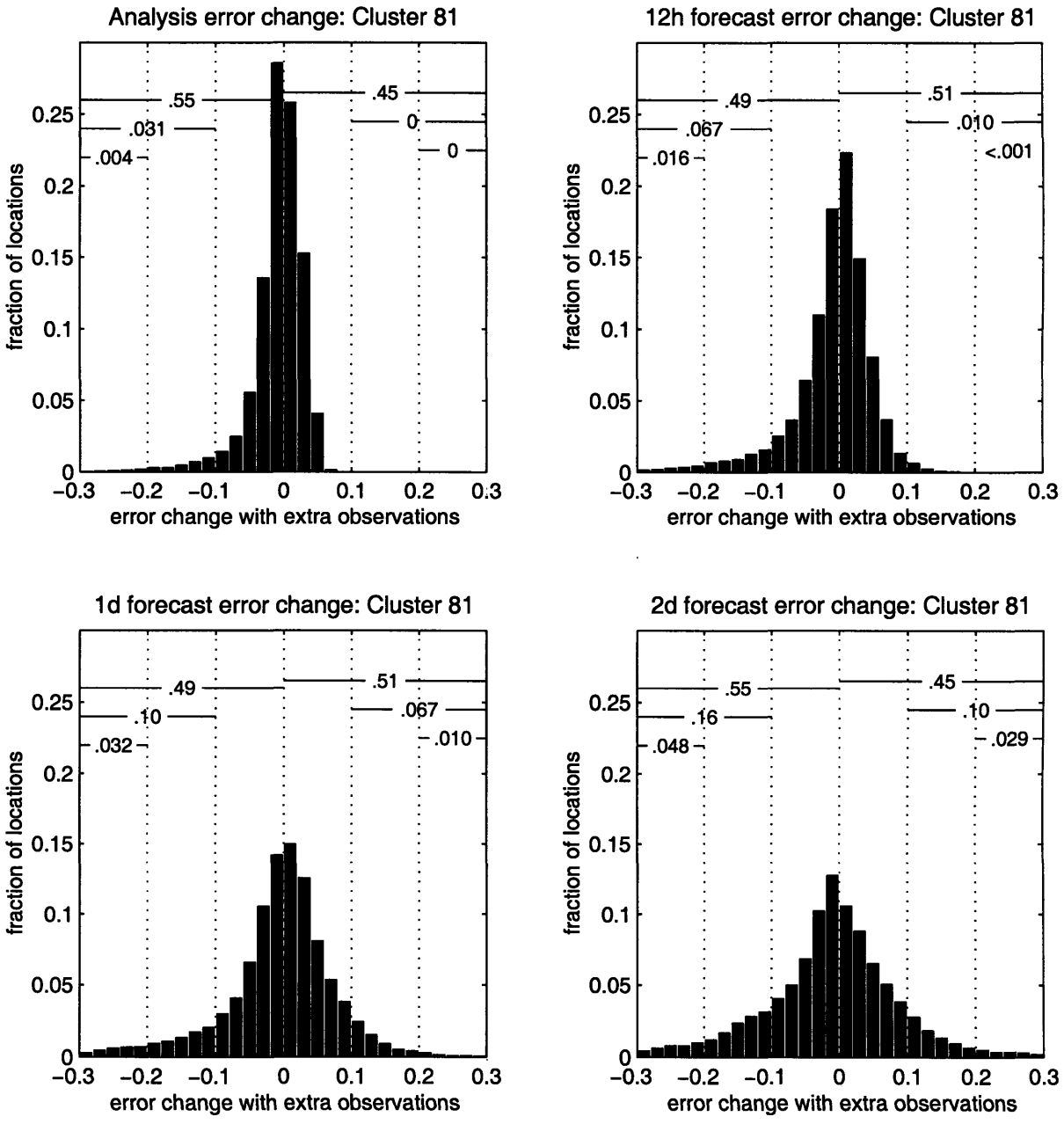



Figure 9.14: As in Figure 9.12, for an extra cluster 81 of observations tested at each location.

for single, cluster 3, and cluster 81 observations, respectively. The statistics in Figures 9.12–9.14 are gathered over only 1 distribution of fixed observations, or $17 \times 1 \times 1968 \approx 3 \times 10^4$ extra sets of observations. The cluster 13 results for only one distribution of fixed observations (i.e. the same $\approx 3 \times 10^4$ extra sets of observations) look qualitatively similar to Figure 9.7; therefore, the cluster 13 and other cluster results can be compared. Because of the smaller number of statistics, however, we focus only on the general appearance of the histograms.

The general statistical results are the same as those discussed in Section 9.4; for all cluster patterns tested, extra observations are much more likely to improve rather than to degrade analyses significantly, but they are not more likely to improve rather than to degrade 1 or 2 day forecasts significantly. Larger cluster patterns are beneficial in some situations, since they are more likely than smaller cluster patterns to reduce analysis and forecast errors significantly. However, larger cluster patterns can also be detrimental, since they are also more likely to increase analysis and forecast errors significantly. Thus, with larger cluster patterns there may be a trade-off between potential benefits and potential risks when adapting observations. Also, we note that as discussed for the choice of observation clusters in Section 6.2, clusters 3 and 13 tend to be more effective (with a smaller risk of degradation relative to the frequency of improvements) than cluster 81.

## 9.6 Applicability of results to other forecasting systems

How extra observations influence analyses and forecasts is very dependent on the forecasting system, particularly on the data assimilation system. Therefore, our results should be interpreted as a general indication of when and where observations may improve or degrade analyses, not as an absolute guide. One could argue that since we use a relatively unsophisticated data assimilation scheme, our results are pessimistic, overestimating the risk that extra observations will degrade forecasts. Certainly, with a better data assimilation system one would hope that extra observations would be more likely to improve both analyses and forecasts. However, because the influence of extra observations will always be limited to some extent by how the data are assimilated into the forecast model, the risk of degradations is inherent in adding observations to improve weather prediction.

In addition, the results shown are for perfect observations, perfect observing platforms, and a perfect forecast model. These are the circumstances in which the observations are most likely to have a positive influence, circumstances which are not realizable in real numerical weather prediction. The error growth rate in the standard resolution quasi-geostrophic model is also somewhat slower than in the real atmosphere, so real world forecasts are likely to degrade on shorter time scales than those depicted here. Therefore, for a data assimilation system which behaves similarly to our 3DVAR, the results we present are quite optimistic — in real forecasting systems, with more complex forecast models and significant practical constraints, there is probably an even greater chance that extra observations will degrade analyses and forecasts.

As discussed in the Introduction, real world adaptive observations taken to improve forecasts have also produced a mix of improvements and degradations. There are two possible ways to reduce the risk of degradation. First, we can develop better data assimilation systems, data assimilation systems which are less likely to degrade analyses and are less likely to improve analyses but degrade forecasts. No data assimilation system is perfect, however, and any errors which exist in the initial conditions will always eventually grow into forecast errors. Consequently, improving data assimilation systems can decrease the risk of degradation, but it cannot prevent degradations. In order to improve forecasts in individual situations while avoiding forecast degradations to the extent possible, we will need to develop observation strategies which recognize and incorporate the risk of degradation. These strategies could do so by simply taking observations more frequently, as we did in previous chapters, and by explicitly including information about how the data assimilation system and the forecast model will use the observational information after the data are taken.

# Chapter 10

# Summary

Instead of reviewing the major results of the thesis in the order presented, we summarize by discussing how our results begin to answer each of the questions raised in the Introduction. We emphasize, however, that all of the results discussed are for our idealized simulated system; they are not necessarily valid in any more complex forecasting system. The general results can provide a framework for understanding how different types of observation networks might behave in the real atmosphere, but the specific results, particularly those relating to specific time and space scales, are likely not directly applicable.

In particular, we believe that two major aspects of our simulated system limit the applicability of these results to the real atmosphere: 1) the perfect model assumption; and 2) the limited quasi-geostrophic dynamics in the forecast model. Real world observing platforms also have many limitations, such as requirements that observations in a region be taken non-simultaneously and that observation locations be selected in advance of the targeting time; we have tested only idealized platforms. These and other issues not addressed in this study will likely affect the potential for adapting observations. Therefore, further study is needed in a more realistic system before our results can be extended to real numerical weather forecasting and real observation networks.

How any forecast model uses any observations is also highly dependent on the data assimilation system used. Therefore, as more sophisticated data assimilation schemes are developed for and implemented in numerical weather prediction, the most important considerations for adapting observations and the best observing strategies are likely to change. Where important, we mention which results we believe are likely to be system-dependent and which are likely to be more general.

- Can adaptive observations improve analyses and forecasts on a statistically significant basis, in a fully three-dimensional dynamical system with a realistic data assimilation system? If adapting observations can improve analyses and forecasts, what kinds of improvements can we expect under different circumstances?

Our results show mixed potential for adaptive observations. In Chapters 6, 7, and 8, we demonstrate that for observation networks which are sufficiently sparse in space in this simulated system, adaptive observations can reduce analysis and forecast errors, averaged over a large number of cases, by using fewer observational resources than fixed or random observations. This is true for both globally adapted observations and for adaptive observations added to a pre-existing network of fixed observations. The adaptive strategies tested are based on analysis error reduction only, with no attempt to estimate future error growth. Although both the strategies and the experimental setup are idealized, the success of a simple strategy with a relatively simple data assimilation system makes us optimistic that it may be possible, given the appropriate considerations, to adapt observation networks to improve weather forecasts. The time series results in Section 8.2, which show how a few adaptive observations added regularly to a non-dense pre-existing network can significantly reduce the number of forecasts with large errors, are especially encouraging.

For fairly dense observation networks, on the other hand, adapting observations in this simulated system is not very effective. In fact, if there are enough observations to resolve synoptic scale features reasonably well, adding or reallocating observations according to *any* strategy, adaptive or non-adaptive, improves analyses and forecasts only a small amount. Again, this is true for both global observations and for observations added to a pre-existing network. The results in Figures 6.2, 6.3, and 6.10–6.11 suggest that analyses and forecasts may still benefit from adapting observations at higher observation densities, but that the benefit is likely to be much smaller than it is for non-dense observations. In the real world there are many constraints, such as imperfect observing platforms and imperfect knowledge, on adaptive observations. Consequently, these smaller possible benefits are difficult to achieve in realistic situations. In order to improve analyses and forecasts from adapting observations at sub-synoptic scales, we are likely to require some combination of better model resolution and a better data assimilation system. In addition, we may need a better observing strategy and a more specific error norm.

We can extend to more complex forecasting systems the general principle that the

effectiveness of adaptive observations, in fact of any observations, depends on the density of all of the observations in the system. The influence of observations at any specific observation density, however, will differ among forecast models, data assimilation systems, and experimental setups. For dense observations, when our forecast model is unrealistic (both because of the dynamics and because of the lack of model error) and our data assimilation does not use observations well, the results are particularly sensitive to the details of our simulated system. The type of observing platform to be used will also be a major factor in determining the potential of different observation strategies in different circumstances.

- Which types of strategies are likely to be the most effective in different circumstances? How can the strategies best be implemented, in terms of allocation of observational resources in space and time? How effectively can we estimate errors in initial conditions for use in adaptive observation strategies?

Our data assimilation system performs best when there is a relatively large background error to correct (discussed further below). Thus, in our simulated system the most effective strategies observe at locations with medium–large background errors. Adding information about future error growth to the strategy could certainly help us choose among the locations with reasonably sized background errors. However, with a data assimilation system which behaves similarly to ours, targeting strategies based on error growth alone are unlikely to be effective. As data assimilation systems evolve and are able to better correct small background errors (without a severe risk of forecast degradation), incorporating information about error growth into a strategy may be more beneficial.

An ideal observing strategy, in any circumstances, would incorporate information about how both the data assimilation system and the forecast model will use the observational data once they are gathered. Unfortunately, it is very difficult to include these criteria in a realistic system with many degrees of freedom. In situations with small background errors or where the data assimilation system is likely for some other reason to have difficulty incorporating the observations well, however, we believe that it may be especially important to integrate the observation network, the data assimilation system, and the forecast model in some manner.

Results in Section 6.3, Section 8.2, and Chapter 9 suggest that the type of adaptive

strategy we test prefers to allocate the observational resources in both space and time according to the specific situation. By doing so, it can attempt to reduce errors in the analysis and forecast cycle wherever and whenever they begin to grow, before they become too large. Any adaptive strategy, unless it has good knowledge about how the data assimilation and the forecast model will use the observational data, is also likely to perform best when taking fewer observations frequently. One can imagine, however, that an unintelligent adaptive strategy could behave similarly to the random strategy we test, with no preference for how the observations are allocated in space and time. An even less effective adaptive strategy might behave similarly to the fixed strategy for frequent observations, preferring to take more observations at fewer times; this can occur if the observations tend to be redundant, as in the adjoint-based strategy for improving the forecast at a specific location which was tested (for a one-dimensional system) in Lorenz and Emanuel (1998). Thus, different types of adaptive strategies may work best for differently allocated observations.

In Chapters 7 and 8, we test using a multiple replication ensemble (described in Section 4.2.3) to estimate errors in initial conditions for adapting observations. The results suggest that an ensemble-based estimate of background error is reasonably useful for adapting observations, but by no means ideal. The ensemble estimate is more useful when a cluster of observations is taken at each targeted location instead of a single observation, and when fixed observations are taken regularly to "anchor" the ensemble. Only a few ensemble members are needed to obtain a reasonable estimate of background error, at least with the simple norm we test. Results in Section 7.4 suggest that the estimate of background error becomes less accurate as the lead time increases, but that even with a requirement for advance lead time, we may still be able to estimate background error with some skill.

The ensemble is successful at identifying a few regions where background errors are likely to be large. The ensemble is less successful at identifying regions with smaller background errors and at estimating smaller scale structures within regions with large background errors. Although in some situations we may only want to identify a few general regions with probable errors, in other situations we are likely to want more detailed information about the background errors, information that the ensemble may estimate inaccurately in several different ways. An ensemble could almost certainly be configured to estimate background error better than our first attempt does. If we are to develop a significantly better technique for estimating background error for adaptive observations, however, we must first understand which types of background errors are

the most important to estimate well (and not to estimate poorly). In other words, further work is needed to learn more about which errors the data assimilation system and the forecast model are most likely to be able to use observations to correct.

- How important are the limitations of the data assimilation system for adapting observations? How would we like to change the data assimilation procedure to improve the influence of observations on forecasts?

Even for perfect observations, the data assimilation does not always improve analyses with extra observations for two main reasons. First, data assimilation is a statistical procedure, with the inherent risk of degradations. Second, we have poor knowledge of the true forecast error statistics which are needed to assimilate the observations well. Because more observations do not always produce better analyses and forecasts once they are assimilated, the data assimilation system is one of the primary factors which limits the effectiveness of adaptive observations. If the background errors to be corrected are smaller, the signal-to-noise ratio in the observations is smaller, and so the data assimilation is even more of a limiting factor when adapting observations. Thus our 3DVAR data assimilation system, and perhaps any reasonable data assimilation system, is more likely to correct large than small background errors.

As discussed in Chapter 9, in many cases it is difficult to determine why or how an assimilated observation degraded an analysis or a forecast. We have only been able to identify two specific types of situations when our data assimilation system is likely to perform poorly: when there is significant small scale horizontal structure in the background error, and when there is significant vertical structure in the background error. This is likely related to the difficulty our data assimilation has in correcting interior potential vorticity structures (discussed in Section 3.5.2). A more sophisticated data assimilation scheme, one which incorporates reasonably accurate time- and space-dependent estimates of either the background error structure or the dynamical situation, can help fix these problems. Several schemes that include this type of information (described in Section 3.1) are currently being developed and implemented and are likely to be necessary if we wish to use observations to correct small magnitude or small scale errors. However, although these better data assimilation schemes can improve the influence of all observations, they cannot prevent degradations.

Data assimilation is an unavoidable complication when incorporating observations

into a realistic atmospheric prediction system. Therefore, we will always have consider the limitations of the data assimilation system taking observations to improve numerical forecasts. An ideal data assimilation for adaptive observations would recognize these limitations, somehow minimizing the likelihood that an assimilated observation will degrade an important forecast. Unfortunately, with our limited knowledge about time- and space-dependent analysis and forecast errors and with the large number of degrees of freedom in an atmospheric system, this is a difficult problem to solve, both theoretically and practically. Developing data assimilation systems which better combine observations with model forecasts but are still feasible to implement is, therefore, currently one of the major challenges in numerical weather prediction.

- How important is the risk of degradations when taking extra observations? How can we minimize the risk? With this risk in mind, what is the best way to define our goals when trying to optimally allocate observations?

As discussed above, there is always a risk that an atmospheric data assimilation system will degrade rather than improve the initial conditions when it assimilates data. Despite this, we found in Chapter 9 that our relatively unsophisticated 3DVAR data assimilation system performs fairly well on average, in the sense that when it assimilates observations it generally improves the average analysis. Our data assimilation is particularly likely to improve the analysis if the observations are taken in regions where the background error is large. Thus, as long as the background errors are not small (which generally means that the past observation network is not dense), the risk that the data assimilation will degrade the analysis outright is not large.

Even if the observations improve the average analysis, when the analysis increment is added to the initial conditions and the resulting analysis is integrated forward in time by the forecast model, any part of the analysis increment which projects onto a growing forecast error can lead to a forecast degradation. Consequently, as demonstrated in Chapter 9, the risk that extra observations will degrade a forecast is significant. Further study is needed on how observations interact with data assimilation systems, background errors, and forecast models, so that we can better understand how and why some of these degradations occur.

It may be possible, for a given data assimilation system and forecast model, to identify situations in which observations are likely to risk degrading the analysis or

forecast. One could then minimize the risk of degradation by not observing in these regions, essentially by observing at locations where the data assimilation system and forecast model are most likely to use the additional information well. In Chapter 9, we identify a few of these risky types of regions for this simulated system: regions with small background errors, and regions with background errors with significant horizontal or vertical structure. However, the most risky regions are likely to vary from forecasting system to forecasting system, and without perfect information about the background error they are also difficult to identify.

Even if we could identify the risky regions, in any individual situation it is very difficult, if not impossible, to know in advance whether certain observations will definitely degrade a forecast. Degradations can also appear on any time scale, from the analysis to a long range forecast. Therefore, another way to minimize the risk of degradation is to adapt observations regularly and frequently, as we found in this idealized system. By taking observations regularly and frequently the adaptive strategy can try to decrease these degradations as they appear, minimizing the risk that a degradation will propagate through a forecast uncorrected. This strategy can work similarly in any forecasting system.

Since it is always possible that extra observations will degrade a forecast, it is not clear how we should define the "best" locations to take observations, in other words what our goal should be when selecting adaptive observation locations. As an example, without considering the limitations of assimilating the data, one might try to adapt observations by identifying the locations where additional information will theoretically have the most influence on a forecast. Because this goal increases the risk of severe degradation along with increasing the chance of significant improvement, however, it seems unwise. To avoid maximizing the risk of degradation, in most of this study we have worked towards minimizing the average analysis error. This goal is easy to define, and in this simulated system it also on average reduces forecast error. While beneficial, this goal is probably not ideal for improving "significant" forecasts.

There are a wide range of alternate possible goals, each of which may lead to a different observation strategy. For example, would we like to reduce errors on average for certain types of forecasts? Reduce forecast errors in individual important situations? Reduce the risk of large (or of any size) forecast busts? Or would we like to optimize some more probabilistic forecast, such as a forecast of the range of possible weather situations or of the likelihood that a certain weather event will occur? Many of the possible goals are nebulous and difficult to work towards, certainly more so than those

which have been attempted to date. If we are to maximize the forecast improvement produced by a limited amount of observational resources, however, we believe that it is first important to define a goal which to the extent possible accounts for the limitations of different aspects of the forecasting system.

- Given that we have limited observational resources, as discussed in Emanuel et al. (1997), what is the "optimal" mix of observing systems? In other words, what role would we like all different types of observations, adaptive and non-adaptive, to play in the future?

This question is very complex, and as discussed earlier, we do not believe that the results from our idealized simplified system can be directly applied to answer it. We can, however, use our general results to suggest some possible effective configurations for current observation networks, which can then be explored in a more realistic system.

A fixed observation network is both very inexpensive and useful for a variety of purposes in addition to numerical weather prediction. Therefore, we begin by assuming that we have a regular rawinsonde observation network over populated regions. Along with the fixed network, we are also likely to have non-targeted satellite, aircraft, and other observations at different locations at different times. Together these consist of our non-adaptive observation network.

Given this pre-existing network, as discussed above we may be able to improve analyses and forecasts while minimizing the risk of degradation by adding a few adaptive observations on a regular basis. For our forecast model and data assimilation system, this only works well when the observation network is not dense, where dense is defined as enough observations to observe synoptic scale features reasonably well. For a more complex model, however, and a more sophisticated data assimilation system, adding a few adaptive observations regularly may be beneficial for other observation densities.

Adaptive observations could be taken from several types of observing platforms. We could, for example, deploy manned or unmanned aircraft to take data in the identified important regions. Or, if an important region happened to be near land, we could deploy one or more additional rawinsondes (manned or automated). It may also be possible at some times to gather higher resolution (in space or time) satellite or aircraft data than are standard, data which for some reason cannot always be taken, processed, or assimilated in all regions. If so, we could preferentially assimilate extra data from

these sources in the identified regions, at the best resolution for the data assimilation system and forecast model to use effectively. In the future, targeted observations may also be available from other platforms, such as dropsondes from commercial aircraft or satellites which can be focussed on a specific region. Each observing platform has different strengths and weaknesses, and thus in each situation a different type of platform will be the most useful. Ideally, we would have some combination of different observing platforms available and could choose which mix to deploy based on the situation and the platform constraints.

Further work is needed to determine the most effective strategy for selecting these additional observation locations. The results from this study suggest that, with current data assimilation systems, we should first observe at locations with medium–large background errors. The best strategies would likely also include criteria about which forecast errors are likely to be growing rapidly and affect a forecast of interest and about at which locations the data assimilation system is likely to be able to incorporate additional observational data effectively. Whatever the specific strategy, our results suggest that the adaptive observations will be most effective if they are taken regularly, so that errors do not have an opportunity to grow unchecked. To correct smaller background errors, a more sophisticated strategy (if we have a more sophisticated data assimilation system) may be necessary and could be implemented in addition to the basic strategy.

This proposed observation network can, first of all, on average reduce the error in the initial conditions. Reducing the average analysis error has several advantages. First, it is likely on average to reduce forecast error. Second, it will provide an improved estimate of the basic state trajectory, which can help us decide when and where we would like to adapt observations. It can do this both by providing a better forecast of whether the significant forecast we are interested in is likely to occur, and by providing better information about where we should place the observations (for example, by improving the accuracy of adjoint calculations indicating which errors are growing fastest). In addition, with a better analysis we will better be able to estimate the time- and space-dependent forecast error covariances which are necessary for an improved data assimilation scheme. Not only can adding a few adaptive observations regularly to a pre-existing network reduce average analysis error, but the time series results in Section 8.2 suggest that it may also reduce the number of severe forecast busts.

Results in Chapter 9 suggest that at least with current data assimilation systems, it is difficult to adapt observations to improve forecasts in individual situations without a significant risk of degradation. If additional observations cannot be or have not been

taken regularly, however, it may still be possible to effectively adapt observations for individual significant forecasts by beginning to take extra observations several days in advance of when we would like to have an improved forecast. Of course, the significant event for which we would like to improve the forecast is, by definition, difficult to predict well in advance. By simply on average improving the analysis in advance of when we are likely to want an improved specific forecast, however, this strategy may work well for similar reasons to the regular adaptive observing strategy described above. Even if some of the extra observations degrade the forecast, it is unlikely that they will all do so. At the very least, we would have a series of forecasts made from initial conditions with different sets of data in potentially important regions; these forecasts can help us evaluate the range of possible weather events or the probability that a certain significant weather event will occur.

The effectiveness of adaptive observations, particularly for networks which cannot add observations regularly, is likely to depend strongly on how large the risk of degradation is. The risk of degradation is strongly affected not only by the strategy, but also by the limitations of the observing platform used, by how the observations taken interact with the data assimilation procedure, and by the accuracy of the forecast model. Therefore, as observing platforms, data assimilation systems, and forecast models evolve, the optimal observation network is likely to evolve as well. At this time, we can only suggest observation networks for forecasting systems similar to those currently available; as aspects of the forecasting system change, the observing strategy should be reevaluated. We do believe, however, that as long as there are limited observations, an ideal observation network will know as much as possible about the data assimilation system, the forecast model, and the forecast goals. The more that observation networks, data assimilation systems, and forecast models are integrated, the better each is likely to perform.

# Appendix A

# Conjugate residual solver

The matrices in the data assimilation are too large to store and invert explicitly. Therefore, to invert Equation (3.3) for the 3DVAR, we have implemented an iterative conjugate residual solver. Conjugate residual methods are a type of conjugate gradient method; they are designed to converge for the inversion of matrices which are either definite or symmetric, but not necessarily both. Equation (3.3) can be written in the form

$$\mathbf{A}(\mathbf{x}) = \mathbf{f}, \tag{A.1}$$

where

$$\mathbf{A} = \mathbf{I} + \mathbf{B}\mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}\mathbf{L}$$

and

$$\mathbf{f} = \mathbf{B}\mathbf{L}^{\mathrm{T}}(\mathbf{O} + \mathbf{F})^{-1}\mathbf{y}.$$

To assimilate the data we calculate $\mathbf{f}$, then invert $\mathbf{A}$ to solve for the analysis increment $\mathbf{x}$.

Our conjugate residual solver is based on the CR2 scheme in Smolarkiewicz and Margolin (1994). For each

$i$ = iteration number,

we define

$\mathbf{x}_i$ = a trial solution to Equation (A.1) for the $i$th iteration; and

$\mathbf{r}_i$ = a residual vector for the $i$th iteration ($= \mathbf{A}(\mathbf{x}_i) - \mathbf{f}$).

The residual $\mathbf{r}_i$ measures how well the trial solution $\mathbf{x}_i$ satisfies Equation (A.1). The solver is considered converged when the residual satisfies some criterion; $\mathbf{x}_i$ then becomes the solution to Equation (A.1), i.e. the analysis increment.

If we wish to accelerate the convergence, we can also define

$\mathbf{P}$ = a preconditioner for $\mathbf{A}$.

$\mathbf{P}$ is an operator which is based on some knowledge of $\mathbf{L}$, but which is easier to invert than $\mathbf{L}$ itself.

To invert Equation (A.1) using the conjugate residual solver, we first select an initial guess for the trial solution $\mathbf{x}_0$ (we set $\mathbf{x}_0 = 0$). After calculating the initial residual $\mathbf{r}_0 = \mathbf{A}(\mathbf{x}_0) - \mathbf{f}$, preconditioning the initial residual $\mathbf{p}_0 = \mathbf{P}^{-1}(\mathbf{r}_0)$, and evaluating $\mathbf{A}(\mathbf{p}_0)$, we iterate through the following loop, beginning with $i = 1$:

1. Evaluate the optimization constant $\beta$ and use it to iterate $\mathbf{x}$ and $\mathbf{r}$:

$$\beta_i = -\frac{(\mathbf{r}_{i-1}\,,\,\mathbf{A}(\mathbf{p}_{i-1}))}{(\mathbf{A}(\mathbf{p}_{i-1})\,,\,\mathbf{A}(\mathbf{p}_{i-1}))}$$

$$\mathbf{x}_i = \mathbf{x}_{i-1} + \beta_i \mathbf{p}_{i-1}$$

$$\mathbf{r}_i = \mathbf{r}_{i-1} + \beta_i \mathbf{A}(\mathbf{p}_{i-1})$$

2. Test for convergence. If the algorithm has converged (if $\|\mathbf{r}_i\| \leq \varepsilon$), set $\mathbf{x} = \mathbf{x}_i$ and exit the loop.

3. If the residual does not yet satisfy the convergence criterion, precondition the residual and evaluate $\mathbf{A}$:

$$\mathbf{q}_i = \mathbf{P}^{-1}(\mathbf{r}_i)$$

$$\text{evaluate } \mathbf{A}(\mathbf{q}_i)$$

4. Evaluate the second optimizing coefficient $\alpha$ and use it to iterate $\mathbf{p}$ and $\mathbf{A}(\mathbf{p})$:

$$\alpha_i = -\frac{(\mathbf{A}(\mathbf{q}_i)\,,\,\mathbf{A}(\mathbf{p}_{i-1}))}{(\mathbf{A}(\mathbf{p}_{i-1})\,,\,\mathbf{A}(\mathbf{p}_{i-1}))}$$

$$\mathbf{p}_i = \mathbf{q}_i + \alpha_i \mathbf{p}_{i-1}$$

$$\mathbf{A}(\mathbf{p}_i) = \mathbf{A}(\mathbf{q}_i) + \alpha_i \mathbf{A}(\mathbf{p}_{i-1})$$

5. Return to Step 1.

In this implementation, the inner product ( , ) is defined as the dot product.

In our standard 3DVAR, we define convergence of the solver as when the maximum residual is less than 2% of the initial maximum residual. Because convergence is reasonably quick, we did not implement a preconditioner, i.e. we set $\mathbf{P} \equiv \mathbf{I}$ (and therefore $\mathbf{q}_i \equiv \mathbf{r}_i$). The rate of convergence depends on the specific situation, but assimilating more observations usually requires more iterations. For the standard resolution, the algorithm generally converges in 2–20 iterations. For the double resolution 3DVAR, convergence was defined as 5% of the initial maximum residual and only a few more iterations were required. Other convergence criteria (such as the root-mean-square average of all residuals) were tested and produced similar results; Section 5.3.2 shows several examples of the sensitivity of the results to the magnitude of the convergence criterion.

The major cost of the conjugate residual solver is in evaluating the operator $\mathbf{A}$. Because $\mathbf{A}$ is stored implicitly and is never inverted, however, the exact form of $\mathbf{A}$ need never be known. Evaluating $\mathbf{A}$ requires computational resources of the same order of magnitude as advancing the quasi-geostrophic model one time step, since the major computational burden in both is in the inversion of potential vorticity to calculate streamfunction (in the operator $\mathbf{L}$ and in the QG model advection). The algorithm requires only one application of $\mathbf{A}$ during each iteration, in Step 3; the second application of $\mathbf{A}$, in Step 4, is not evaluated explicitly, but iteratively. The time step in the standard resolution QG model is 30 minutes. Therefore, at most it takes the same order of magnitude of time to solve the 3DVAR as it takes to integrate one 12 hour QG model forecast. This is comparable to operational 3DVAR data assimilation systems.

As described in Smolarkiewicz and Margolin (1994), the algorithm converges as long as $\beta_i \leq 0$, i.e. as long as the matrix $\mathbf{A}$ is definite (our sign convention for $\beta$ is reversed because we have defined our residual as the negative of theirs). In all but a few situations, $\beta$ remains negative and the solver converges. In most non-convergent cases, the background error covariance matrix $\mathbf{B}$ was specified incorrectly, either because the sensitivity to changes in $\mathbf{B}$ was being tested or because $\mathbf{B}$ was being developed. Non-convergence is also more likely when there are very few or very many observations. Even in these situations, however, positive $\beta$ and non-convergence are rare. If $\beta$ becomes positive, it does so when the residual is already fairly small. The resulting analysis increment is therefore affected only a small amount.

As discussed in Section 3.1, one of the major advantages of the variational data assimilation formulation is its flexibility. The same holds true for the conjugate residual solver. To implement a different variational data assimilation, one must only redefine $\mathbf{A}$ and $\mathbf{f}$ in Equation (A.1), then use the same algorithm — in fact the same FORTRAN

code — to assimilate the data.

# References


Aberson, S. D., 1997: Adaptive observations in a hurricane environment. Preprints, *22nd Conf. on Hurricanes and Tropical Meteorology*, 19-23 May 1997, Ft. Collins, CO, Amer. Meteor. Soc., 308-309.

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

Anderson, J. L., 1997: The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results. *Mon. Wea. Rev.*, **125**, 2969-2983.

Bengtsson, L., and N. Gustavsson, 1971: Assimilation of data in dynamical analyses. *Tellus*, **23**, 328-336.

Bengtsson, L., and N. Gustavsson, 1972: Assimilation of non-synoptic observations. *Tellus*, **24**, 383-399.

Bengtsson, L., M. Ghil, and E. Källén (Eds.), 1981: *Dynamic Meteorology: Data Assimilation Methods*, Springer-Verlag, New York, 330 pp.

Bergman, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.*, **107**, 1423-1444.

Bergot, T., G. Hello, A. Joly, and S. Marlardel, 1999: Adaptive observations: A feasibility study. *Mon. Wea. Rev.*, under review.

Berliner, L. M., Z.-Q. Lu, and C. Snyder, 1999: Statistical design for adaptive weather observations. *J. Atmos. Sci.*, accepted.

Bishop, C. H., and Z. Toth, 1996: Using ensembles to identify observations likely to improve forecasts. Preprints, *11th Conf. on Numerical Weather Prediction*, 19-23 August 1996, Norfolk, VA, Amer. Meteor. Soc., 72-74.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, under review.

Buizza, R., and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**, 2503-2518.

Burpee, R. W. J. L. Franklin, S. J. Lord, R. E. Tuleya, and S. D. Aberson, 1996: The impact of Omega dropwindsondes on operational hurricane track forecast models. *Bull. Amer. Meteor. Soc.*, **77**, 925-933.

Charney, J., M. Halem, and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160-1163.

Courtier, P., J. Derber, R. Errico, J.-F. Louis, and T. Vukicevic, 1993: Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. *Tellus*, **45A**, 342-357.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113**, 1329-1347.

Daley, R., 1991: *Atmospheric Data Analysis*, Cambridge Univ. Press, Cambridge, UK, 457 pp.

Dey, C. H., and L. L. Morone, 1985: Evolution of the National Meteorological Center global data assimilation system: January 1982 - December 1983. *Mon. Wea. Rev.*, **113**, 304-318.

Emanuel, K. A., and coauthors, 1997: Observations in aid of weather prediction for North America: Report of Prospectus Development Team Seven. *Bull. Amer. Meteor. Soc.*, **78**, 2859-2868.

Emanuel, K. A., and R. Langland, 1998: FASTEX adaptive observations workshop meeting summary, Dec. 8, 1997, Naval Research Laboratory, Monterey, Calif. *Bull. Amer. Meteor. Soc.*, **79**, 1915-1919.

Errico, R. M., and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU-NCAR Mesoscale Model. *Mon. Wea. Rev.*, **120**, 1644-1660.

Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46**, 1193-1206.

Gelaro, R., R. Langland, G. D. Rohaly, and T. E. Rosmond, 1999: An assessment of the singular vector approach to targeted observing using the FASTEX data set. *Quart. J. Roy. Meteor. Soc.*, under review.

Ghil, M., 1997: Advances in sequential estimation for atmospheric and oceanic flows. *J. Meteor. Soc. Jpn.*, **75**, 289-304.

Hamill, T. M., C. Snyder, and R. E. Morss, 1999: Error characteristics of bred, singular vector, and perturbed observation ensemble forecasts. *Mon. Wea. Rev.*, to be submitted.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 11-136.

Hoskins, B. J., and N. V. West, 1979: Baroclinic waves and frontogenesis, part 2, Uniform potential vorticity jet flows: cold and warm fronts. *J. Atmos. Sci.*, **36**, 1663-1680.

Houtekamer, P. L., 1993: Global and local skill forecasts. *Mon. Wea. Rev.*, **121**, 1834-1846.

Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.*, **123**, 2181-2196.

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

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

Ide, K., and M. Ghil, 1998: Extended Kalman filtering for vortex systems. Part I: Methodology and point vortices. *Dyn. Atmos. Oceans*, **27**, 301-332.

Jastrow, R., and M. Halem, 1970: Simulation studies related to GARP. *Bull. Amer. Meteor. Soc.*, **51**, 490-513.

Joly, A., and coauthors, 1997: The Fronts and Atlantic Storm-Track Experiment (FASTEX): scientific objectives and experimental design. *Bull. Amer. Meteor. Soc.*, **78**, 1917-1940.

Kalnay, E., and Z. Toth, 1994: Removing growing errors in the analysis cycle. Preprints, *10th Conf. on Numerical Weather Prediction*, 18-22 July 1994, Portland, OR, Amer. Meteor. Soc., 212-215.

Langford, J. S., and K. A. Emanuel, 1993: An unmanned aircraft for dropwindsonde deployment and hurricane reconnaissance. *Bull. Amer. Meteor. Soc.*, **74**, 367-375.

Langland, R. H., R. Gelaro, G. D. Rohaly, and M. A. Shapiro, 1999a: Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOPs 17 and 18. *Quart. J. Roy. Meteor. Soc.*, under review.

Langland, R. H., Z. Toth, R. Gelaro, I. Szunyogh, M. A. Shapiro, S. Majumdar, R. Morss, G. D. Rohaly, C. Velden, N. Bond, and C. Bishop, 1999b: The North Pacific Experiment (NORPEX-98): Targeted observations for improved North American weather forecasts. *Bull. Amer. Meteor. Soc.*, to be submitted.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701-721.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177-1194.

Lorenz, E. N., and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.*, **55**, 399-414.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73-119.

Molteni, F., and T. N. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, **119**, 269-298.

Morel, P., G. Lefevre, and G. Rabreau, 1970: On initialization and non-synoptic data assimilation. *Tellus*, **23**, 197-206.

Mureau, R., F. Molteni, and T. N. Palmer, 1993: Ensemble prediction using dynamically conditioned perturbations. *Quart. J. Roy. Meteor. Soc.*, **119**, 299-323.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633-653.

Parrish, D. F., and J. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Pu, Z.-X., E. Kalnay, J. C. Derber, and J. G. Sela, 1997: Using forecast sensitivity patterns to improve future forecast skill. *Quart. J. Roy. Meteor. Soc.*, **123**, 1035-1053.

Pu, Z.-X., E. Kalnay, D. Parrish, W. Wu, and Z. Toth, 1997: The use of bred vectors in the NCEP global 3D variational analysis system. *Wea. and Forecasting*, **12**, 689-695.

Robert, A. J., 1966: The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Jpn.*, **44**, 237–245.

Rotunno, R., and J.-W. Bao, 1996: A case study of cyclogenesis using a model hierarchy. *Mon. Wea. Rev.*, **124**, 1051-1066.

Smolarkiewicz, P. K., and L. G. Margolin, 1994: Variational solver for elliptic problems in atmospheric flows. *Appl. Math. and Comp. Sci.*, **4**, 527-551.

Smagorinsky, J., K. Miyakoda, and R. F. Strickler, 1970: Variables in initial conditions of dynamical weather prediction. *Tellus*, **22**, 141-157.

Snyder, C., 1996: Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Amer. Meteor. Soc.*, **77**, 953-961.

Szunyogh, I., Z. Toth, K. A. Emanuel, C. Bishop, J. Woolen, T. Marchok, R. Morss, and C. Snyder, 1999: Ensemble-based targeting experiments during FASTEX: The impact of dropsonde data from the LEAR jet. *Quart. J. Roy. Meteor. Soc.*, in revision.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311-1328.

Thepaut, J.-N., R. N. Hoffman, and P. Courtier, 1993: Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.*, **121**, 3393-3414.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297-3319.

Toth, Z., I. Szunyogh, S. Majumdar, R. Morss, C. Bishop, and S. Lord, 1999: Ensemble-based targeted observations during NORPEX. Preprints, *3rd Symposium on Integrated Observing Systems*, 10-15 January 1999, Dallas, TX, Amer. Meteor. Soc., to appear.

Valdes, P. J., and B. J. Hoskins, 1988: Baroclinic instability of the zonally averaged flow with boundary layer damping. *J. Atmos. Sci.*, **45**, 1584-1593.

Velden, C. S., C. M. Hayden, S. J. Nieman, W. P. Menzel, S. Wanzong, and J. S. Goerss, 1997: Upper-tropospheric winds derived from geostationary satellite water vapor observations. *Bull. Amer. Meteor. Soc.*, **78**, 173-195.