Time-averaged data assimilation for midlatitude climates: towards paleoclimate applications

Angeline G. Pendergrass

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Science

University of Washington

2009

Program Authorized to Offer Degree:
Department of Atmospheric Sciences

University of Washington
Graduate School

This is to certify that I have examined this copy of a master's thesis by

Angeline G. Pendergrass

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

_____
David Battisti

_____
Gregory Hakim

_____
Gerard Roe

Date: _____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Overview

The focus of this study is to explore whether data assimilation techniques commonly employed in numerical weather prediction might be used to improve reconstructions of past climates. We explore this question in the context of midlatitude climate variability using two models. The first is a simple, stochastic, one dimensional representation of the atmosphere and ocean mixed layer (Barsugli and Battisti 1998, hereafter the BB model), and the other is a dynamical, fully non-linear, two-surface quasi-geostrophic model of Hakim (2000) coupled to a slab ocean (hereafter, the QG model).

We establish that the simple BB data assimilation system (which will be referred to as the BB system) shares many behaviors with the QG data assimilation system, and that the simpler model is an effective tool for understanding time-averaged data assimilation and the potential for a dynamical model to assimilate paleoclimate observations, despite lacking many important features like spatial variability.

One key test for data assimilation is whether state-dependent patterns of covariance among climate fields can be taken advantage of, or whether only climatological covariance patterns are useful. An important result is that we show that potential for state-dependent skill is controlled by a key dynamical factor: the accumulation of noise. The accumulation of stochastic noise has a similar effect as more realistic processes, and is likely to prove critical for other, more complicated, systems.

Our experiments with realistic parameters indicate that a slab ocean model could have enough memory to be useful for annual assimilation in the ocean, but not in the

atmosphere. This implies that higher frequency observations are necessary to capture useful atmospheric patterns at these timescales.

This thesis consists of four chapters. The introductory chapter is a very brief review of paleoclimate reconstruction, data assimilation, and how they can fit together. This is followed by a review of literature on stochastic climate modeling relevant to the BB model. The methods chapter contains the details of the setup for the modeling and assimilation systems. It includes a description of the BB model, a description of the QG model with a focus on the development of the slab ocean component, a comparison of the two models, and then a brief introduction to time-averaged data assimilation. The results chapter focuses on original developments presented here, beginning with a framework for dynamical timescales of coupled interactions, followed by the development of and experiments with an idealized time-averaged data assimilation system, and ending with experiments using a data assimilation system built around the QG model. Finally, the discussion will summarize the work presented here and place it in the context of data assimilation for paleoclimate reconstructions.

1.2 Paleoclimate

The goal of applying data assimilation to paleoclimate is to incorporate dynamical constraints into paleoclimate reconstructions. In this section, we will review the available paleoclimate observations at high frequencies. Then, we will review other methods of incorporating dynamics into paleoclimate reconstruction.

Jones et al. (2009) provide a thorough review of currently available high-resolution proxy data. For the most part, the highest frequency observations are annual. Annual dating is possible for records where the seasonal signal leaves a physical or chemical imprint. Sometimes, the record tells us about climate integrated over the entire year; for example, ice sheets contain layers of annual accumulation that are demarcated by isotope signals, or trace element constituents, tell us about precipitation integrated over a calendar year. Some proxies can have sub-annual resolution (e.g., corals);

and these won't be considered here, but they could easily be incorporated with the proposed methods.

People have been creating timeseries of paleoclimate proxy data and interpreting past climate from them for decades (one example is Hays et al. 1976). Mann et al. (1998) and Mann et al. (1999) brought large-scale paleoclimate reconstruction to a new level by using statistical methods to make a quantitative reconstruction of Northern hemisphere mean temperature from a compilation of different types of proxies. Since then, statistical methods (which, here, will mean methods for paleoclimate reconstruction that do not rely upon physics directly) have been used to create reconstructions of specific climate indices and spatial fields (Jones et al. 2009 provides an overview). These reconstructions rely on the covariability of proxies with reanalysis data or instrumental records. However, the instrumental record is limited in time, and it is unclear that these relationships remain fixed over long periods of time.

One of the challenges in paleoclimate reconstruction is that proxy data has larger errors than, for example, meteorological observations. One example of this is coral skeletal oxygen isotope proxies for temperature and salinity, reviewed in Jones et al. (2009). They report errors from vital effects, kinetic or non-equilibrium chemistry, empirical calibration, and noise from non-climatic artifacts. Paleoclimate proxies are indirect records of climate, and they all have errors. Errors in proxy data will not be dealt with explicitly here, but are important to address in future studies.

The next step in climate reconstruction is to incorporate information available from physics-based models. A few studies have begun to attempt this. Some studies have come up with new ways of combining models and observations. Graham et al. (2007) devised the proxy surrogate reconstruction method. In this method, the years in a model integration are reordered to best correspond with proxy data. Goosse et al. (2006) developed the optimal ensemble choice method, where an ensemble of model trajectories is generated and the member that best agrees with proxy data during some time period is selected. This method uses models for reconstruction, but it has

the disadvantage that the model knows nothing about the observations.

Other studies have borrowed methods developed for operational weather prediction. Jones and Widmann (2004) developed a method where a large-scale spatial pattern is reconstructed, and then a model is nudged toward the reconstructed spatial pattern. van der Schrier and Barkmeijer (2005) used another operational weather prediction tool, forcing singular vectors, instead of nudging to coax a model towards a target state reconstructed from proxy data.

Finally, methods from fields outside of earth science have been adapted. Haslett et al. (2006) introduced Bayesian hierarchical modeling in the context of paleoclimate reconstruction. This method finds joint probability distribution functions of proxies and model states.

While these and other developments are ongoing, no method thus far has successfully integrated models and proxy records without the use of instrumental data to generate a climate reconstruction.

1.3 Data assimilation for paleoclimate

We would like to adapt a method used to incorporate observations of the atmosphere into weather models to paleoclimate reconstruction. Here, this method will be called *data assimilation*, and refers specifically to ensemble-based methods derived from the ensemble Kalman filter. There is some confusion over terminology. The term data assimilation is used in the paleoclimate literature to refer to any method of climate reconstruction incorporating models (see, for example, the discussion in Jones et al. 2009). In the context of weather prediction, data assimilation also refers to methods incorporating observations into model integrations; for example, adjoint methods, as well as schemes where no models are used.

We will very briefly review the advances that led to the relevant data assimilation methods for the weather problem. Then, we will describe how data assimilation works. Finally, advances on the path towards adaptation to the paleoclimate problem will

be discussed.

In 1795, Gauss developed the method of least-squares to better predict planetary orbits. Though orbits are deterministic, he could account for errors in the initial conditions by incorporating more observations than degrees of freedom. In 1958 and 1960, two independent papers extended the method of least squares to take into account model error in addition to observational error, generating what is known today as the Kalman filter (Kalman 1960, historical overview in Sorenson 1970). The Kalman filter proliferated in engineering, but it was not directly applicable to geophysical problems due to the enormous number of degrees of freedom in geophysical models. A series of approximations developed. The approximation used here began with Evensen (1994) used a Monte Carlo method to estimate the probability density function, making the method computationally feasible for oceanography and spawning the Ensemble Kalman filter (EnKF). Square-root methods of updating the covariance matrix were then applied to this ensemble method to improve efficiency and accuracy (see review by Whitaker and Hamill 2002). The method used for one of the data assimilation systems here is a version of the square root filter, the Ensemble Adjustment Kalman Filter (Anderson 2001, EAKF). Unlike the EnKF and other square root filters, the EAKF does not add random perturbations to the observations.

1.3.1 Ensemble data assimilation basics

We will now walk through data assimilation for a hypothetical single variable case to introduce the conceptual framework, mathematics, and language that will be used throughout this thesis.

We begin with a model $\mathcal{M}$. The goal will be to find the best estimate of the true system state x^t, of which we have an observation y at time t with some observational error R. We have an ensemble of state estimates $\vec{x}^o$ at time t_o called ensemble members. The ensemble members are arranged into a vector of N model states. This ensemble might have come from a previous forecast, or it might be states drawn

randomly from a long integration of the model.

The data assimilation cycle can be split into two parts: the forecast step and the assimilation step. First, in the forecast step, the ensemble of states is integrated forward, using the model, from time t_o to time t, when the observation is available, which is written mathematically as,

$$\vec{x}(t) = \mathcal{M}(\vec{x}^o).$$

For an atmosphere model, ensemble member trajectories will diverge until saturating at some climatological variance. For an ensemble starting with error less than that of a randomly drawn ensemble, the error will generally increase during the forecast step.

The analysis step takes place at time t. The ensemble at time t is called the background or prior ensemble, denoted here by $\vec{x}^b = \vec{x}(t)$. During the assimilation cycle, this background ensemble will be adjusted towards the true state using the observation. The ensemble is a Monte Carlo estimate of the probability distribution of states. Its mean is $\bar{x}^b$, and its variance, $B = (\vec{x}^b - \bar{x}^b)(\vec{x}^b - \bar{x}^b)^T$, will be used in the adjustment process. The observation operator H is applied to ensemble states to obtain the observation that each state would produce, $y_i^b = Hx_i^b$. The difference between the actual observation and the each model-estimated observation, $y_i^b - y_o$, is the innovation, or new information that is used to update the ensemble members. The innovation is spread to the model state. This spreading is done by a weight K, called the Kalman gain. The updated ensemble, called the analysis, is given by.

$$\vec{x}^a = \vec{x}^b + K(y - H\vec{x}^b).$$

The choice of K is the key to data assimilation. K is chosen to minimize the error of the analysis. We will not show the derivation of K, but it is standard (see, for example, Kalnay 2002). The solution is $K = BH^T[HBH^T + R]^{-1}$.

The error of each analysis ensemble member is $\tilde{x}_i^a = x_i^a - x^t$, and the analysis error

variance A is,

$$A = \vec{\tilde{x}}^a \vec{\tilde{x}}^{aT} = (1 - KH)B.$$

The analysis can now be used as a forecast, a climate reconstruction, or the initial ensemble for a new assimilation cycle.

Figure 1.1 shows a schematic of the data assimilation cycle. This whole system can be visualized by imagining the ensemble members as a cloud of points with the true observed state somewhere near the middle of the cloud. During the forecast step, the cloud spreads out and the true state wanders. The analysis step drags the cloud back toward the true state and shrinking its extent.

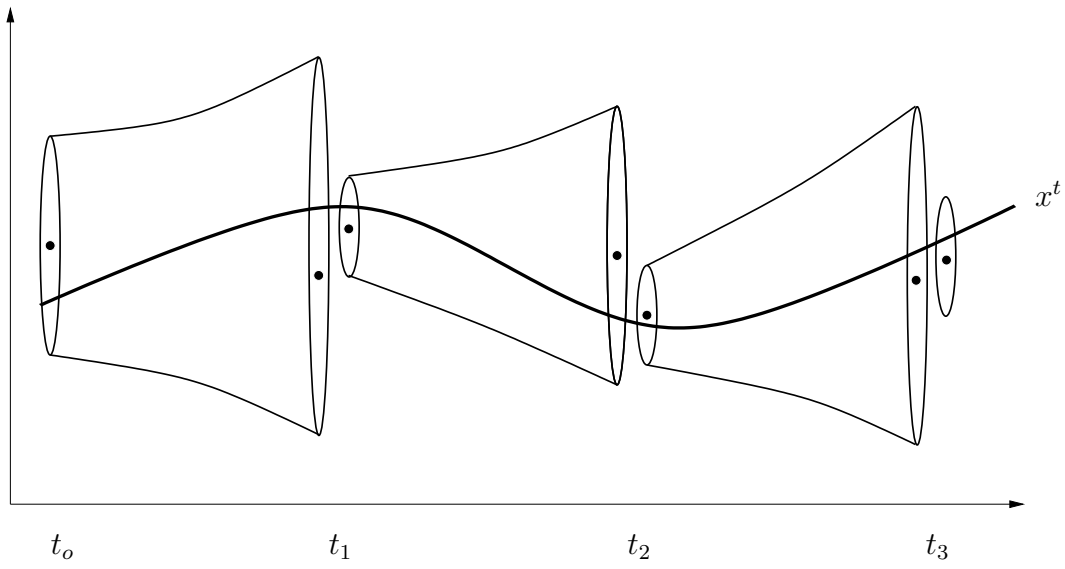

Figure 1.1: **Data assimilation schematic.** The thick line is the true state, and cones are envelopes of ensemble trajectories, and dots are the ensemble mean. See text for accompanying description of the data assimilation cycle.

1.3.2 Adapting the EnKF for paleoclimate

Data assimilation is an apt method for paleoclimate because it can operate with sparse data and takes into account their uncertainty. Additionally, it is extremely flexible: it

can deal with any observation that has a quantifiable relationship to the model state.

One of the first obstacles to assimilating paleoclimate observations is that they are spread over long time periods, often effectively averaged or integrated in time. Dirren and Hakim (2005) developed an algorithm to deal with this by using time-averaged model states in the assimilation scheme and withholding high-frequency residuals. They tested the method in a model designed to have two different dynamical timescales, experimenting with different averaging periods for the observations and the model states, which were not necessarily the same. They found that time-averaged assimilation was skillful for the low frequency states for a range of averaging times, and in some cases there was skill for states time-averaged at frequencies higher than the observation averaging period.

Huntley and Hakim (2009, submitted) tested the time-averaged data assimilation method in the context of the climate system. They used a different version of the QG atmosphere model employed here to experiment with sparse observation networks. They found that small numbers of observations can provide a basis for skillful reconstructions if they are optimally chosen. These results held when error was intentionally introduced into the model. They found that the skill of their reconstructions decreased with averaging time, which was a motivating factor leading to the work in this thesis. Their analysis did not examine assimilation periods longer than a month.

One of the main advantages of the EnKF is that the background error covariances are determined from an evolving ensemble of integrations, so they are dependent on the state of the system at a given time. This is called state-dependence. This would be a very useful property that could make this method far superior to using correlations of proxy data with data from the non-stationary reanalysis period. But state-dependence is only helpful up to a limited horizon that is dependent on the model physics and observations. At observation time, if the model has forgotten its state from the previous observation time, then it is just as well to use the model's climatology instead of integrating the ensemble forward. A key focus of this thesis is

to determine the processes that are responsible for setting this horizon in the context of midlatitude climate dynamics.

1.4 Stochastic climate models in midlatitude climate predictability and variability from air-sea interactions

In this thesis, we will explore state-dependent time-averaged data assimilation by developing an idealized data assimilation system. This system will be built around a linear stochastic model for air-sea interaction developed by Barsugli (1995) and Barsugli and Battisti (1998) (hereafter, the second paper will be referred to as BB98 and the model will be referred to as the BB model). In this section, we review work in stochastic climate modeling for air-sea interaction and predictability that led up to the development of this model and subsequent work employing and extending it.

Stochastic models were introduced to climate by Hasselmann (1976). This paper argued that the timescale separation between the atmosphere and ocean made it possible to approximate the atmosphere as a white noise forcing on the ocean. Then Frankignoul and Hasselmann (1977) brought to bear a very simple stochastic model on the question of how large-scale, persistent sea surface temperature (SST) anomalies could arise in the midlatitude oceans. In their model, SSTs developed large-scale, persistent anomalies forced by only stochastic noise (discussed in Frankignoul 1985). Frankignoul (1985) added a linear damping to the stochastic forcing, which kept the variance from increasing indefinitely at longer times.

By the 1990s, questions of how and what kinds of climate variability arose from air-sea interaction were being addressed with general circulation models (GCMs) with various surface boundary condition configurations. In order to reconcile the effects of different types of boundary conditions, BB98 added a stochastic forcing term to air-sea flux model of Schopf (1985). As discussed in Kushnir et al. (2002), the BB98 model has subsequently been used to interpret and explain results in more complicated and realistic models.

A number of studies have used the BB model and extensions to it to establish predictability based on air-slab ocean coupling, increasingly with the aim of understanding decadal climate variability. Bretherton and Battisti (2000) use the BB98 model for predictability of ensemble means. Saravanan and McWilliams (1998) extended the BB model with an advection term. Scott (2003) subsequently used it to conclude that advection effects can cancel some of the noise. Scott and Qiu (2003) added a deterministic ocean forcing and discussed its predictability. Ferreira et al. (2001) extended the Saravanan and McWilliams (1998) model by coupling it to a two-layer quasi-geostrophic atmosphere model, with a geostrophic ocean and mixed-layer entrainment.

A parallel line of stochastic climate model development continued extending the model based on Frankignoul and Hasselmann (1977). Deser et al. (2003) incorporated entrainment into the model to address mixed-layer reemergence as a mechanism for interannual persistence of SST anomalies. Sura et al. (2006) added a multiplicative noise term to account for the kurtosis in observed SST anomaly distributions.

In summary, while some studies have suggested extensions and modifications, the original concept of Hasselmann (1976) and its implementation in BB98, have both been shown to be successful in capturing important aspects of midlatitude climate variability on interannual timescales. The applicability of BB98 and its analytical tractability render it an ideal tool for this study.

Chapter 2

METHODS

In order to consider predictability as it is relevant to paleoclimate data assimilation, we will use two simplified systems. The first is a stochastic two variable model of atmosphere and slab ocean temperature. Given a set of parameters, solutions at equilibrium of this system are known. In addition, because it is only a two-variable system, the Kalman filter equations can be applied directly. This very simple system allows us to ask questions and find concrete answers. It is, however, not chaotic (all eigenmodes are damped). To perform standard data assimilation experiments, we will use a two-dimensional model with quasi-geostrophic atmospheric dynamics and a slab ocean to carry out experiments that test the ideas that stem from the simpler two-variable model. Both models are designed with midlatitude weather or climate in mind. The quasi-geostrophic model is centered around a jet to emulate midlatitude eddy dynamics, and the two-variable model was developed for air-sea interactions in midlatitude ocean basins.

In this section, both of the models will be described. They will be compared and contrasted. Then a description of the two data assimilation systems will be given.

2.1 Models

2.1.1 Two-variable model

Midlatitude coupled atmosphere-mixed layer interactions are an important driver for interseasonal, interannual, and up to decadal variability (see, for example, Alexander and Deser 1995). For the assimilation of annual paleoclimate observations, a majority of which are in the midlatitudes (see, for example, Mann et al. 1999), these interactions

are candidates for interannual memory.

A very simple model describing the coupled mid-latitude ocean-atmosphere system was developed in Barsugli (1995) and Barsugli and Battisti (1998). A schematic of the model is shown in figure 2.1. This model has a single atmosphere layer and a slab ocean. These layers can exchange heat with fluxes proportional to their temperature difference, and both radiatively damp to presumed equilibrium temperature. Forcing applied to the atmosphere consists of two parts: one proportional to the slab ocean temperature and another due to white noise. The model was developed to describe idealized GCM experiments of Barsugli (1995) and explore the effect of three different SST configurations on climate variability that were, at the time, commonly used in climate modeling: coupled, uncoupled, and specified (time-dependent) SSTs.

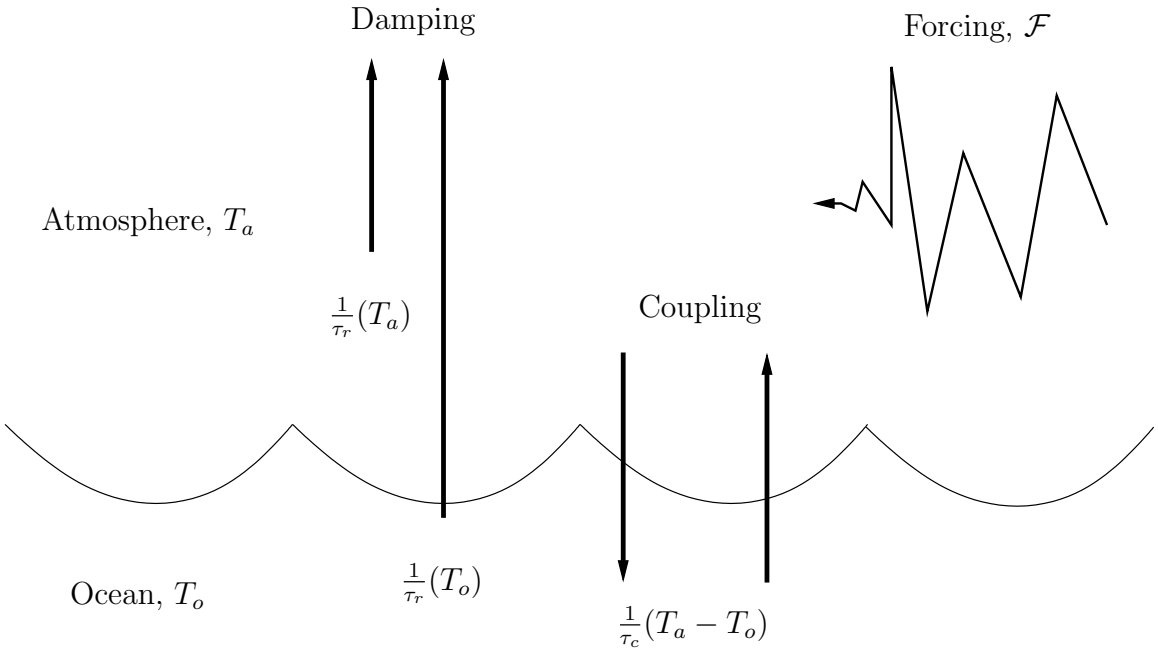

Figure 2.1: **Schematic of BB model.** A schematic of the simple model, similar to figure 2 of BB98.

Derivation from energy balance

The model is derived from a vertically-averaged energy balance model linearized about basic atmosphere and ocean states originally described in Appendix A of BB98. The energy balance model consists of a vertically-averaged atmosphere temperature and slab ocean. Assuming that the surface air temperature is the same as the vertically averaged temperature, the energy balance model is,

$$\gamma_a \frac{\partial T_a}{\partial t} = R_a + \epsilon_a \sigma_b T_o^4 - 2\epsilon_a \sigma_b T_a^4 + \lambda(T_o - T_a) + \mathcal{F}$$
$$\gamma_o \frac{\partial T_o}{\partial t} = R_o - \sigma_b T_o^4 + \epsilon_a \sigma_b T_a^4 - \lambda(T_o - T_a),$$

where γ_a and γ_o are the heat capacities of the atmosphere and slab ocean, T_a and T_o are their temperatures, R_a and R_o are the external radiative forcings, ϵ_a is the emissivity of the atmosphere, σ_b is Boltzmann's constant, λ is a linearized coefficient of air-sea coupling, and $\mathcal{F}$ is the forcing applied to the atmosphere.

Linearizing about the basic state ($T_o = \bar{T}_o + T'_o$, $T_a = \bar{T}_a + T'_a$, $R_o = \bar{R}_o + R'_o$, $R_a = \bar{R}_a + R'_a$) and omitting nonlinear terms gives,

$$\gamma_a \frac{\partial T'_a}{\partial t} = R'_a + 4\epsilon_a \sigma_b \bar{T}_o^3 T'_o - 8\epsilon_a \sigma_b \bar{T}_a^3 T'_a + \lambda(T'_o - T'_a) + \mathcal{F}'$$
$$\gamma_o \frac{\partial T'_o}{\partial t} = R'_o - 4\sigma_b \bar{T}_o^3 T'_o + 4\epsilon_a \sigma_b \bar{T}_a^3 T'_a - \lambda(T'_o - T'_a).$$

Then we neglect the shortwave radiation anomalies R'_a and R'_o and add and subtract a mixed term from each equation to arrive at,

$$\gamma_a \frac{\partial T'_a}{\partial t} = -4\epsilon_a \sigma_b (2\bar{T}_a^3 - \bar{T}_o^3) T'_a + (\lambda + 4\epsilon_a \sigma_b \bar{T}_o^3)(T'_o - T'_a) + \mathcal{F}'$$
$$\gamma_o \frac{\partial T'_o}{\partial t} = -4\sigma_b (\bar{T}_o^3 - \epsilon_a \bar{T}_a^3) T'_o - (\lambda + 4\epsilon_a \sigma_b \bar{T}_a^3)(T'_o - T'_a).$$

At this point, BB98 determine a set of constants from their GCM. These are shown in table 2.1.

Now we have linearized equations for the anomalous air and slab ocean temperature, with linear constants. Grouping terms into the new constants λ_{sa}, λ_{so}, λ_a, and

Table 2.1: **BB98 energy balance model constants.**

ϵ_a	0.76
λ	20 W/m^2/K
$\bar{T}_a$	270 K
$\bar{T}_o$	285 K

λ_o gives the following,

$$\gamma_a \frac{\partial T_a}{\partial t} = -\lambda_{sa}(T_a - T_o) - \lambda_a T_a + \mathcal{F} \tag{2.1}$$

$$\gamma_o \frac{\partial T_o}{\partial t} = \lambda_{so}(T_a - T_o) - \lambda_o T_o. \tag{2.2}$$

Defining a time constant $T = \gamma_a/\lambda_{sa}$, we nondimensionalize time by $\hat{t} = t/T$ and write,

$$\frac{\gamma_a}{\lambda_{sa}} \frac{\partial T_a}{\partial t} = \frac{\partial T_a}{\partial \hat{t}} = -(T_a - T_o) - \frac{\lambda_a}{\lambda_{sa}} T_a + \frac{\mathcal{F}}{\lambda_{sa}}$$

$$\frac{\gamma_o}{\lambda_{so}} \frac{\partial T_o}{\partial t} = \left(\frac{\gamma_o}{\lambda_{so}}\right)\left(\frac{\lambda_{sa}}{\gamma_a}\right) \frac{\partial T_o}{\partial \hat{t}} = (T_a - T_o) - \frac{\lambda_o}{\lambda_{so}} T_o.$$

Now, we nondimensionalize the forcing $F = \mathcal{F}/\lambda_{sa}$, and define a scaled heat capacity ratio $\beta = (\gamma_o/\lambda_{so})(\lambda_{sa}/\gamma_a)$. The forcing is assumed have two parts: a purely stochastic white noise with amplitude N, and a deterministic part proportional to the slab ocean temperature controlled by a constant b. Now we have $F = (b-1)T_o + N$ and the system of equations is,

$$\frac{\partial T_a}{\partial \hat{t}} = T_o - \left(\frac{\lambda_a}{\lambda_{sa}} + 1\right) T_a + (b-1)T_o + N \tag{2.3}$$

$$\beta \frac{\partial T_o}{\partial \hat{t}} = T_a - \left(\frac{\lambda_o}{\lambda_{so}} + 1\right) T_o. \tag{2.4}$$

By defining new constants a and d, we now can write a nondimensionalized system

of equations

$$\frac{\partial T_a}{\partial \hat{t}} = -aT_a + bT_o + N \tag{2.5}$$

$$\beta \frac{\partial T_o}{\partial \hat{t}} = T_a - dT_o. \tag{2.6}$$

The most *ad-hoc* part of the model is the assumption for the forcing term. BB98 discussed the rationale for b, but it remains the most conceptually elusive aspect of the model (see Kushnir et al. 2002). In principle, the parameter b varies from 0 to 1: a value of 1 makes the entire forcing stochastic, and a value of 0 means the entire forcing is proportional to SST. The coupled and uncoupled experiments in BB98 differ by values of b (1 and 0). Note that $b = 0$ still leaves a coupled system, but one in which the atmosphere and ocean can be solved independently. Here, we will generally take b to be 0.5 as BB98 did.

The atmosphere actually responds baroclinically to forcing by the ocean, and baroclinic adjustment leads to disturbances that become barotropic on longer timescales, close to a week. This model only represents the longer barotropic timescales.

2.1.2 Quasi-geostrophic model

The second model has a richer set of internal dynamics than the first, though it still represents a dramatic idealization of nature. The model atmosphere was developed by Hakim (2000) to study mid-latitude cyclogenesis. It is a two-surface model, whose prognostic variables are potential temperature at the surface and the tropopause. The potential temperature on these two surfaces is the only variable needed to determine variables available from inverting potential vorticity.

The atmosphere model is derived from the adiabatic primitive equations (shown in chapter 6 of Pedlosky 1987) by making Boussinesq, f-plane, and quasi-geostrophy assumptions. Potential vorticity (PV) is conserved, as in the QG equations. The

model equations as written in Mahajan (2007) are,

$$\begin{aligned}
\nabla \cdot \vec{v} &= 0 \\
\frac{D\vec{u}}{Dt} + f\hat{k} \times \vec{u} &= \nabla_H \Phi \\
\frac{g}{\theta_{OO}}\Theta &= \Phi_z \\
\frac{D\Theta}{Dt} &= 0,
\end{aligned}$$

where $\vec{v}$ is the three-dimensional wind vector, $\vec{u}$ is the horizontal wind vector, Φ is the geopotential, Θ is the potential temperature, g is the rate of gravitational acceleration, f is the Coriolis paramter, θ_{OO} is the basic-state potential temperature, and the total derivative is,

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \vec{u} \cdot \nabla_H + w\frac{\partial}{\partial z}.$$

Defining a streamfunction ψ' and perturbation potential temperature $\theta' = (g/f\Theta_{OO})\Theta$, the equation solved in the atmosphere model is,

$$\theta'_t = -J(\psi', \theta') - \bar{U}\theta'_x - \bar{\theta}_y\psi'_x - \Gamma\nabla^2\psi' - \nu\left(\nabla_H^2\right)^4\theta',$$

where $J(a, b) = a_x b_y - a_y b_x$ is the Jacobian operator, bars indicate the prescribed basic state, Γ is the Ekman parameter and ν is the hyperdiffusion scaling. Additionally there is a damping term, but we'll include this in the equations for the slab ocean instead of here. A version of the atmosphere model that also included a mountain was used by Huntley and Hakim (2009, submitted) to study time-averaged data assimilation; but, that study was restricted to observation times of a month or less. The model was extended for this study by the addition of a very simple slab ocean, analogous to that of the BB model discussed above. In order to maintain zero interior potential vorticity, the flux between the atmosphere and ocean is applied to both the surface and tropopause potential temperature fields. The slab ocean is implemented spectrally, as is the atmosphere.

Heat flux will be proportional to the temperature difference between the atmosphere and ocean. Instead of a vertically-averaged air temperature, we will use the surface potential temperature. The ocean will have no motion, but it will have a damping term so it can come to equilibrium.

We will begin developing the slab ocean from equations 2.1 and 2.2, which describes the evolution of linearized anomalies between the ocean and atmosphere. The QG model atmosphere will be divided in two for the purposes of heat flux exchange: a surface layer temperature T_b and an upper layer temperature T_t. We will assign heat capacities γ to each of these layers by assuming that each represents half of the atmospheric mass, so $\gamma_a = \gamma_b + \gamma_t$, $\gamma_a T_a = \gamma_b T_b + \gamma_t T_t$, and $\gamma_b = \gamma_t = 2c_p p_s/g$.

We need to ensure that we do not severely violate the model assumption of constant internal PV with the heat fluxes from the slab ocean. Potential vorticity is created by differential heating of the surface and tropopause temperatures,

$$\int_{\mathcal{V}} q dV = \int_{\mathcal{V}} \nabla \cdot \nabla \phi dV = \int_{t,s} (\nabla \phi) \cdot \hat{n} dA = \int \theta_t dA - \int \theta_b dA$$

Therefore,

$$\frac{\partial}{\partial t} \int_{\mathcal{V}} q dV = 0 \qquad iff \qquad \int \frac{\partial \theta_t}{\partial t} dA = \int \frac{\partial \theta_b}{\partial t} dA \Rightarrow f_s = f_t,$$

so if we add heat fluxes evenly to the surface and tropopause, no new PV will be created in the interior of the atmosphere. Note that the model does not strictly conserve PV due to hyper-viscosity and Ekman damping (see discussion in Hakim et al. 2002).

Adding heat equally to both layers is required to preserve the model assumptions (conservation of interior PV), and is consistent with the barotropic atmospheric dynamics expected on longer timescales. Although it is not ideal to have to make such an arbitrary assumption, we know of no reason why it would affect the principal results, and the benefits of having such a simple dynamical model are substantial.

While the damping term is usually calculated within the atmosphere, we will

consider it part of these equations. Now, we can rearrange equations 2.1 and 2.2 to

$$\gamma_b \frac{\partial T_b}{\partial t} + \gamma_t \frac{\partial T_t}{\partial t} = -\lambda_{sa}(T_b - T_o) - \lambda_a \frac{\gamma_b T_b + \gamma_t T_t}{\gamma_a} + \mathcal{F}_b + \mathcal{F}_t \tag{2.7}$$

$$\gamma_o \frac{\partial T_o}{\partial t} = \lambda_{so}(T_b - T_o) - \lambda_o T_o. \tag{2.8}$$

$$\begin{aligned}
\gamma_t \frac{\partial T_t}{\partial t} &= -\frac{\lambda_{sa}}{2}(T_b - T_o) - \frac{\lambda_a}{2} T_t + \mathcal{F}_t \\
\gamma_b \frac{\partial T_b}{\partial t} &= -\frac{\lambda_{sa}}{2}(T_b - T_o) - \frac{\lambda_a}{2} T_b + \mathcal{F}_b \\
\gamma_o \frac{\partial T_o}{\partial t} &= \lambda_{so}(T_b - T_o) - \lambda_o T_o.
\end{aligned}$$

Now we can divide by the heat capacity in each equation and use the fact that $\gamma_a = 2\gamma_b = 2\gamma_t$, and define a ratio of heat capacities $d = \gamma_o/\gamma_a$ to find

$$\begin{aligned}
\frac{\partial T_t}{\partial t} &= -\frac{\lambda_{sa}}{\gamma_a}(T_b - T_o) - \frac{\lambda_a}{\gamma_a} T_t + \frac{\mathcal{F}_t}{\gamma_t} \\
\frac{\partial T_b}{\partial t} &= -\frac{\lambda_{sa}}{\gamma_a}(T_b - T_o) - \frac{\lambda_a}{\gamma_a} T_b + \frac{\mathcal{F}_b}{\gamma_b} \\
d\frac{\partial T_o}{\partial t} &= \frac{\lambda_{so}}{\gamma_a}(T_b - T_o) - \frac{\lambda_o}{\gamma_a} T_o.
\end{aligned}$$

d is the nondimensional depth, and is the equivalent to β in equation 2.4, but note that it is not scaled by the slab ocean coupling coefficient ratio $\lambda_{sa}/\lambda_{so}$ because the nondimensionalization for the QG model is different.

The ratio γ/λ has dimensions of time, so we can write each coefficient as a time constant

$$\frac{\partial T_t}{\partial t} = -\frac{1}{\tau_{ca}}(T_b - T_o) - \frac{1}{\tau_{ra}} T_t + \frac{\mathcal{F}_t}{\gamma_t} \tag{2.9}$$

$$\frac{\partial T_b}{\partial t} = -\frac{1}{\tau_{ca}}(T_b - T_o) - \frac{1}{\tau_{ra}} T_b + \frac{\mathcal{F}_b}{\gamma_b} \tag{2.10}$$

$$d\frac{\partial T_o}{\partial t} = \frac{1}{\tau_{co}}(T_b - T_o) - \frac{1}{\tau_{ro}} T_o. \tag{2.11}$$

We set $\lambda_{sa} = \lambda_{so}$ and $\lambda_a = \lambda_o$. This second assumption differs from BB98, where $\lambda_a/\lambda_o \approx 1.5$. Then we define a new forcing with the factor of heat capacity removed

($F = \mathcal{F}/\gamma$), then we can write all of the equations in terms of just two timescales, with the ocean timescales longer than the atmosphere by the product of the heat capacity ratio

$$\frac{\partial T_t}{\partial t} = -\frac{1}{\tau_c}(T_b - T_o) - \frac{1}{\tau_r}T_t + F_t \tag{2.12}$$
$$\frac{\partial T_b}{\partial t} = -\frac{1}{\tau_c}(T_b - T_o) - \frac{1}{\tau_r}T_b + F_b \tag{2.13}$$
$$\frac{\partial T_o}{\partial t} = \frac{1}{d\tau_c}(T_b - T_o) - \frac{1}{d\tau_r}T_o. \tag{2.14}$$

The QG model is formulated in terms of potential temperature and in a different nondimensionalization. The nondimensionalization is easy to apply in this form by multiplying everything by a time constant T_{QG}

$$\frac{\partial T_t}{\partial \hat{t}} = -\frac{1}{\hat{\tau}_c}(T_b - T_o) - \frac{1}{\hat{\tau}_r}T_t + T_{QG}F_t$$
$$\frac{\partial T_b}{\partial \hat{t}} = -\frac{1}{\hat{\tau}_c}(T_b - T_o) - \frac{1}{\hat{\tau}_r}T_b + T_{QG}F_b$$
$$\frac{\partial T_o}{\partial \hat{t}} = \frac{1}{d\hat{\tau}_c}(T_b - T_o) - \frac{1}{d\hat{\tau}_r}T_o.$$

We must also reconcile the difference between the potential temperature in the QG model and the temperature for heat exchange. Let overbars denote the mean state and primes denote perturbations. We can define a mean state potential temperature $\bar{\theta} = \bar{T}\left(p_o/\bar{p}\right)^{R/c_p}$, and if $\bar{p} >> p'$ then working from the definition of potential temperature we find that

$$\theta' = T'\left(\frac{p_o}{\bar{p}}\right)^{R/c_p}\left(1 - \frac{R}{c_p}\frac{p'}{\bar{p}}\right) - \bar{\theta}\frac{R}{c_p}\frac{p'}{\bar{p}}.$$

When $\bar{p} >> p'$, the second scaling term on temperature goes to 1 and the second term on the right-hand side vanishes. At the surface $\bar{p} = p_o$, so $\theta'_b \approx T'_b$. At the tropopause, it is not clear that this is a good approximation, but we will assume it for simplicity anyway.

With these assumptions, and nondimensionalizing by a factor of Θ, we have

$$\frac{\partial \theta_t}{\partial \hat{t}} = -\frac{1}{\hat{\tau}_c}(\theta_b - T_o) - \frac{1}{\hat{\tau}_r}\theta_t + \frac{T}{\Theta}F_t \tag{2.15}$$

$$\frac{\partial \theta_b}{\partial \hat{t}} = -\frac{1}{\hat{\tau}_c}(\theta_b - T_o) - \frac{1}{\hat{\tau}_r}\theta_b + \frac{T}{\Theta}F_b \tag{2.16}$$

$$\frac{\partial T_o}{\partial \hat{t}} = \frac{1}{d\hat{\tau}_c}(\theta_b - T_o) - \frac{1}{d\hat{\tau}_r}T_o, \tag{2.17}$$

where the primes on θ and T have been dropped.

We now present the details of the coupled QG model integrations. The model version used is 2sQG. The time step is 0.01 model days, with daily output. There is no topography. The maximum spectral resolution is 32 meridional waves and 16 zonal waves, with a domain length of 28,000 km zonally, 11,000 km meridionally, and a depth of 10 km. The diffusion parameter is n=8, the diffusion timescale is 10 timesteps, and the Ekman parameter γ=0.075. Anomalies are relaxed to the basic state jet with a damping time of 10 nondimensional model days (equivalent to 11.5 days). The model is initialized with a localized random baroclinic disturbance.

In this study, the effect of changing the degree of coupling and depth, or memory, of the slab ocean will be a focus. Snapshots of integrations with two different values of slab ocean depth are shown in figures 2.2 and 2.3. In both of these figures, the tropopause and surface temperatures show eddies propagating along a baroclinic zone created by the basic state jet covering the center half of the domain. The anomalies have larger amplitude at the tropopause than at the surface. The main visual difference between the $d = 1$ and $d = 20$ integrations is in the slab ocean. The deeper ocean has much more muted anomalies, both in structure and in magnitude. Figure 2.4 shows an example of the evolution of the domain-mean temperature anomaly magnitude for a realistic depth of d=20, as an example of the model's equilibration. As is clear in the figure, the model (particularly the ocean temperature) anomaly magnitude is equilibrated by 500 days of integration. For the analyses presented in the next section, the last 2000 days of 2500-day integrations are used.

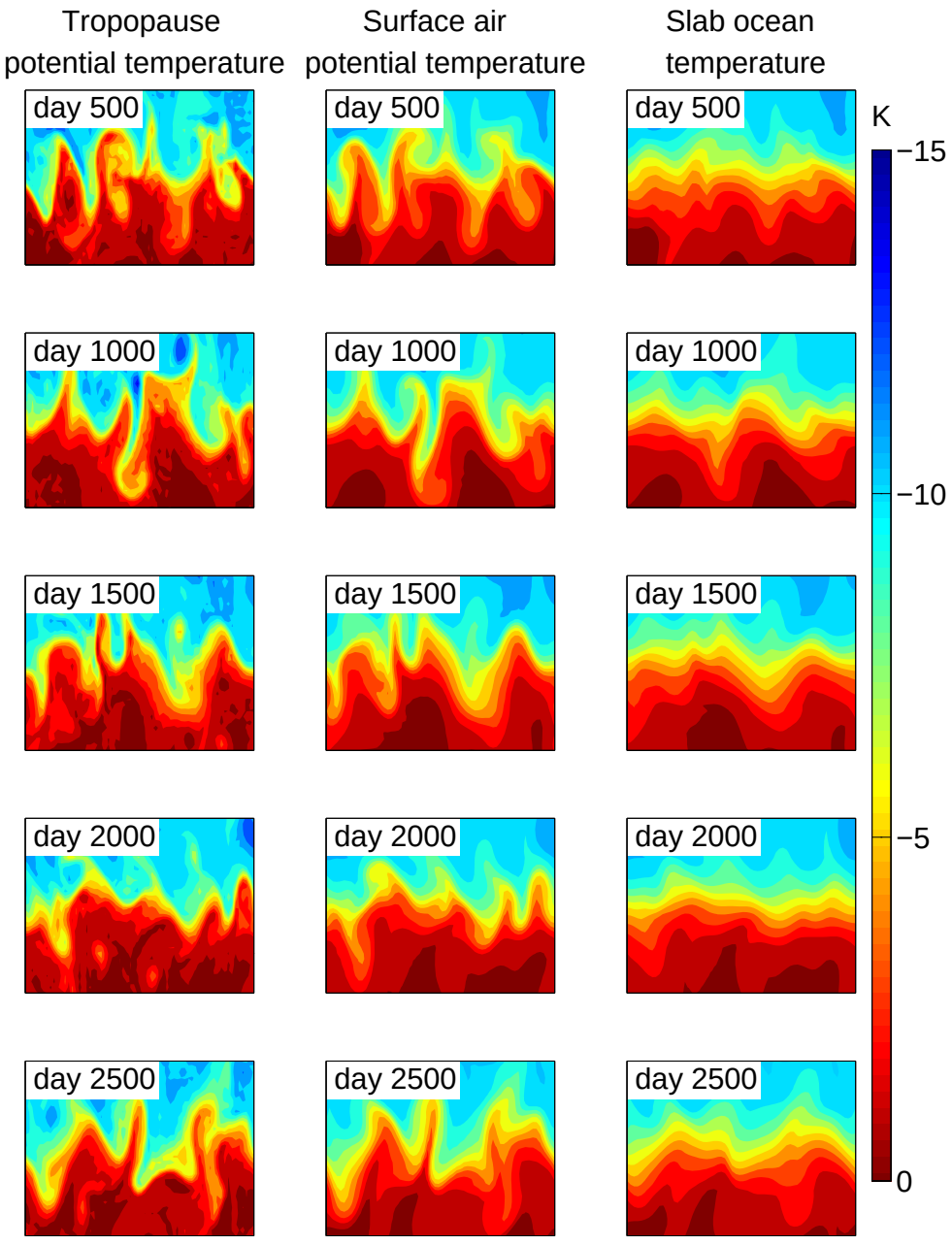

Figure 2.2: **Quasi-geostrophic model integration, d=1.** An example integration of the quasi-geostrophic model integrated for 2500 model days, with $d = 1$ and $\tau_c = 4.7$ days. Instantaneous states are shown once every 500 days, with the tropopause potential temperature on the left, surface air temperature in the center, and slab ocean temperature on the right.

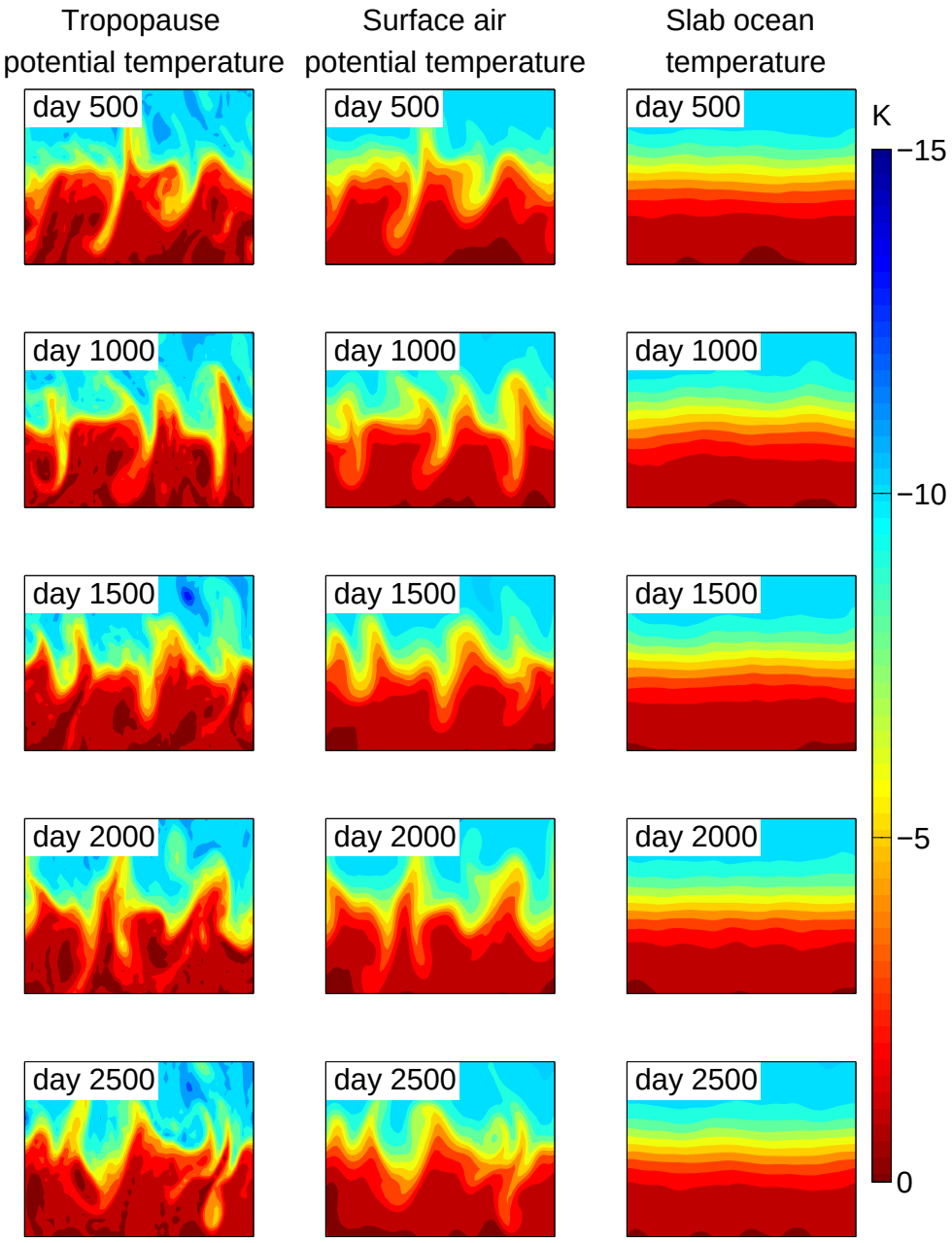

Figure 2.3: **QG model integration, d=20.** As in figure 2.2, but with $d = 20$.

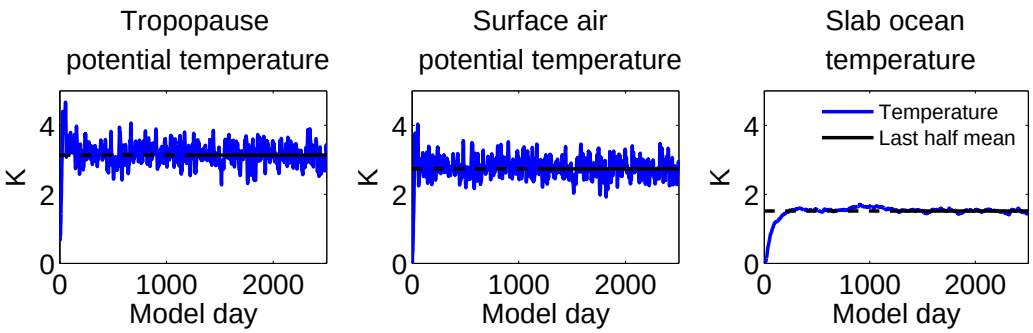

Figure 2.4: **QG model spinup.** Mean anomaly magnitude for entire domain daily for the integration shown in figure 2.3.

2.1.3 Comparing the BB and QG models

These two models are both highly idealized, but the two-variable BB model is much simpler than the QG model. In the BB model, the forcing term is a prescribed stochastic process, whereas the forcing in the QG model is nonlinear quasi-geostrophic atmospheric dynamics. The QG model is two dimensional, while the BB model includes no spatial dimensions. The QG model has an additional pathway for heat to move: lateral advection.

These models also have similarities. In particular, their coupled portions can be written analogously. Equations 2.1 and 2.2 can be rearranged into a form analogous to equations 2.9-11,

$$\frac{\partial T_a}{\partial t} = -\frac{1}{\tau_{ca}}(T_a - T_o) - \frac{1}{\tau_{ra}}T_a + \frac{\mathcal{F}}{\gamma_a} \tag{2.18}$$

$$d\frac{\partial T_o}{\partial t} = \frac{1}{\tau_{co}}(T_a - T_o) - \frac{1}{\tau_{ro}}T_o, \tag{2.19}$$

where $\tau_{ca} = \gamma_a/\lambda_{sa}$, $\tau_{co} = \gamma_a/\lambda_{so}$, $\tau_{ra} = \gamma_a/\lambda_a$, $\tau_{ro} = \gamma_a/\lambda_o$, and $d = \gamma_o/\gamma_a$. For the majority of the work, we will set $\tau_{ca} = \tau_{co} \equiv \tau_c$ and $\tau_{ra} = \tau_{ro} \equiv \tau_r$ for simplicity. Then, in order to compare the two models, we need only to implement consistent nondimensional equivalents of τ_c, τ_r, and d (besides N and b). In this section, we will

compare the spectral properties of the two model for equivalent values of τ_c, τ_r, and d.

Cospectral analyses

Cospectral analysis of the atmosphere and ocean fields in both models will allow us to compare the spectral properties of the variables and their interactions.

For the BB model, the spectra can be determined analytically as shown in BB98 from the Fourier transform of the model equations. The cospectral power and the phase lag of the slab ocean temperature behind the atmosphere temperature can also be determined by extending this algebra, as shown below.

Continuing from the model derivation above, we take the Fourier transform of the nondimensional equations to nondimensional angular frequency $\sigma = T\omega = 2\pi T\omega$ (as opposed to frequency) as BB do. Then the system of equations is

$$\begin{aligned} i\sigma T_a &= -aT_a + bT_o + N \\ i\beta\sigma T_o &= cT_a - dT_o. \end{aligned}$$

This can be rearranged into

$$\begin{aligned} T_a(\sigma) &= \frac{N}{\sigma_a - \frac{bc}{\sigma_o}} \\ T_o(\sigma) &= \frac{cN}{\sigma_o(\sigma_a - \frac{bc}{\sigma_o})}, \end{aligned}$$

by defining

$$\begin{aligned} \sigma_a &= i\sigma + a \\ \sigma_o &= i\beta\sigma + d. \end{aligned}$$

Then the ocean and atmosphere power (equations 14 and 15 in BB98), complex

cospectrum, and phase are,

$$P_a(\sigma) = \frac{\beta^2|N|^2\sigma^2 + d^2|N|^2}{|\sigma_a\sigma_o - \alpha|^2},$$
$$P_o(\sigma) = \frac{c^2|N|^2}{|\sigma_a\sigma_o - \alpha|^2},$$
$$F_{ao}(\sigma) = T_aT_o^* = \frac{c|N|^2(d + i\beta\sigma)}{|\sigma_a\sigma_o - \alpha|^2},$$
$$\phi = \tan^{-1}\frac{\mathrm{Im}F_{ao}}{\mathrm{Re}F_{ao}} = \frac{\beta\sigma}{d},$$

where $\alpha = bc$.

In order to account for our potential uncertainty in b, an envelope of spectra for the extrema in b (0 and 1) are also shown in figure 2.5 by the dotted lines.

The BB model atmosphere and ocean spectra are shown in figure 2.5. For the ocean spectrum, the increase in power going to lower frequencies is characteristic of a red noise process. The plateau at low frequencies is due to the damping term, which also keeps the variance finite as frequency goes to zero. The atmosphere has two plateaus because it is influenced by two damping times, which will be discussed in section 3.2.

The real part of the cospectrum (which is, incidentally, the covariance), has increasing power as frequency decreases, eventually plateauing as the spectra do. The quadrature spectrum first increases in power more rapidly than the cospectrum, but then peaks and approaches zero as frequency goes to zero. The phase lag of the ocean behind the atmosphere describes some of the differences in the co- and quadrature-spectra. At the highest frequencies, the atmosphere and ocean are in quadrature, and as frequency decreases the phase relationship rolls off, until the atmosphere and ocean are exactly in phase at zero frequency. As total power of the system increases from zero at high frequencies, first the quadrature relationship dominates as the atmosphere forces the ocean. As frequency and phase decrease, the power of the quadrature relationship continues to increase, but so does the in-phase part. When the phase is $\pi/4$, the quadrature power peaks, and the forcing of the atmosphere on the ocean decreases

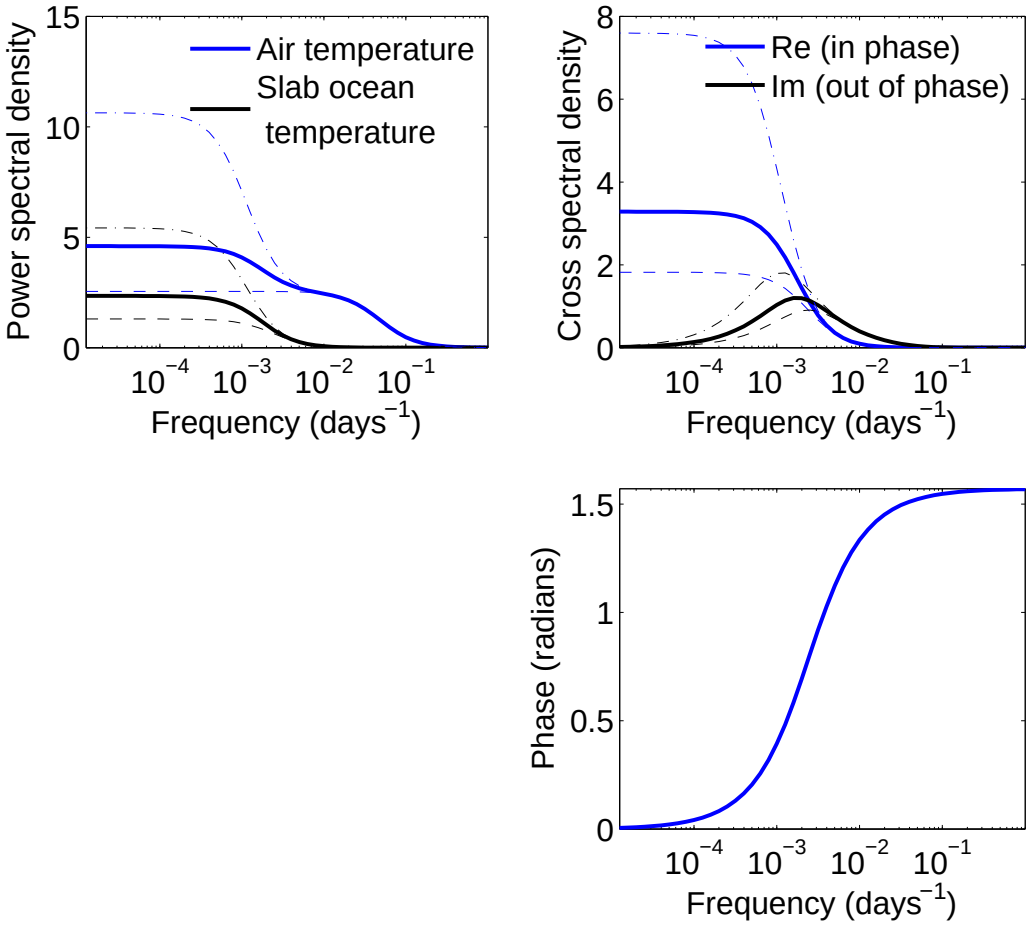

Figure 2.5: **BB model cospectra.** Using $N = 5$, $\tau_c = 4.7$ days, and $d = 20$, the coupled spectral power is calculated according to BB98 equations 14 and 15. Then, the cospectrum is shown in the top right panel, and the phase is in the bottom panel. Thick lines show b=0.5; thin dashed lines show $b = 0$ and $b = 1$.

with further decreases in frequency. Remaining power goes into the in-phase part.

Figure 2.6 shows cospectral (and in the process, spectral) analysis of the ocean and surface air temperature in the quasi-geostrophic model. This analysis employs Matlab's `spectrum` function. For the QG model, histograms of analyses are shown for each gridpoint in the meridional center half of the domain (where the basic state temperature gradient is nonzero). The analyses are done on the last 8000 model-days of an 8500 model-day integration. The window length is 800 model days, and before the spectral analysis 10 model-day chunks are averaged together. The timeseries are not normalized. In figure 2.6, just one line is shown for the simple model, with $b = 0.5$ and $N = 5$, chosen by trial and error to best correspond with the QG model spectra. The rest of the parameters are the same as those for the QG model ($\tau = 10$ model days, $c = 4.7$ days, $d = 20$).

As with the BB model, the spectra of the ocean and atmosphere increase towards lower frequencies because they are red noise processes. The power in the system cannot increase indefinitely because of the damping terms. The cospectral power increases as frequency decreases as in the BB model, and a peak is seen in the quadrature spectrum as well. The atmosphere and ocean show high coherence at frequencies below 100 days, which then degrades with the nonlinear effects at high frequencies. The phase relationship of the QG model follows very closely the BB phase for frequencies where it is well-defined. The correspondence of the characteristics in frequency space of the two models for this set of parameters is not short of remarkable, and lends confidence to using the BB model to represent midlatitude climate dynamics.

2.2 Data assimilation systems

The two models introduced above can be used with a data assimilation algorithm intended for the paleoclimate problem to develop two different types of assimilation systems. See section 1.3.1 for an overview of data assimilation. Because the BB model is simple and linear, it can be combined with the Kalman filter equations

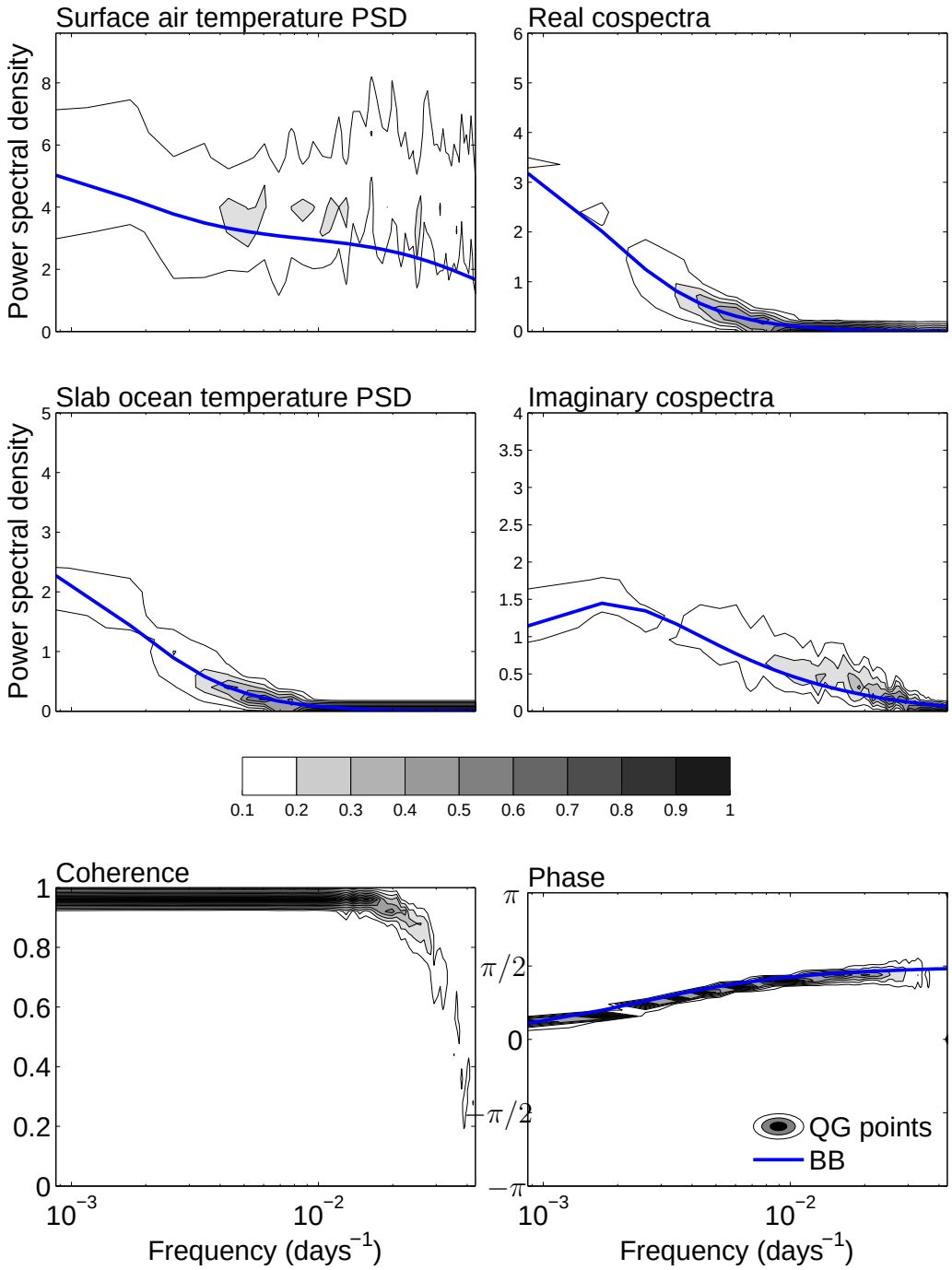

Figure 2.6: **Cospectral analysis intermodel comparison.** Cospectral analysis from 8000 model days of integration from QG model and corresponding cospectrum from BB model (BB coherence is omitted because coherence is always one in linear models). For each frequency, histograms show the relative frequency of spectral lines from each QG model gridpoint. The contour interval is 0.1, or 10%. The vertical integral over each frequency band is 1 in each panel.

to analytically calculate the expected outcomes of data assimilation experiments. This system can be explored and understood in a way that is not possible with more complicated systems, such as the QG system in which we carry out traditional data assimilation experiments. An overview including relevant literature about data assimilation can be found in section 1.3.

The algorithm for data assimilation under consideration is different from those in use for other problems in atmospheric science because it operates with time-averaged states and observations instead of instantaneous ones. This algorithm was first developed in Dirren and Hakim (2005). It is summarized as follows:

1. Integrate model forward in time. Find time mean states.

2. Separate each ensemble member into time-mean and a time-varying residual.

3. Perform update step with time-averaged model states and observations.

4. Recombine updated time-mean state with withheld residual.

5. Integrate from the combined initial condition.

In this section, we will provide an overview of the data assimilation systems. In section 3.2.1, we will develop a system for finding the statistical properties expected from data assimilation. We will not perform individual data assimilation experiments with this model, but instead focus on the expected error from experiments with all possible true states, assuming no model error and an infinite number of ensemble members.

For typical geophysical models, including the QG model, this approach is not possible. For the QG model, we have implemented the algorithm in a typical data assimilation system and performed real experiments. The technical details of this system are described below.

Technical details of quasi-geostrophic data assimilation system

The data assimilation system for the QG model is implemented in the Jamaica release of Data Assimilation Research Testbed (DART, Anderson et al. 2009), a community-maintained data assimilation system. The QG model was interfaced into DART specifically for this project, in addition to the time-averaged assimilation algorithm. The Ensemble Adjustment Kalman Filter (Anderson 2001) was used for the experiments here. Observations, each with an error of 1/10 of the domain-mean climatological variance, were placed at every other gridpoint in the zonal and meridional directions. Ensembles of 48 members were initialized with random draws of instantaneous states from a 2500 day integration of the model. Each experiment consists of one ensemble integration of 100 days with no assimilation (the control case) and another integration with assimilation.

Skill metrics

As a measure of the assimilation skill, we will use the same metrics as Huntley and Hakim (2009, submitted). First, the error variance R is,

$$R = \sqrt{\frac{1}{N}\sum_n [\langle \bar{\theta}(x,t)\rangle - \bar{\theta}_t(x,t)]^2},$$
$$\bar{R} = \frac{1}{M}\sum_m R,$$

where $\langle\cdot\rangle$ denotes an ensemble mean, $\bar{(\cdot)}$ denotes a time average, and $(\cdot)_t$ denotes a true state. The number of points in space used is n and the number of assimilation cycles is m.

The reduction in error variance of an assimilation experiment when compared to a control with no assimilation is one measure of the benefit from assimilating data. This reduction normalized by the control error is called the gain g, given by,

$$g = \frac{\bar{R}_c - \bar{R}}{\bar{R}_c},$$

where $(\cdot)_c$ denotes the control ensemble integration. Gain is 1 for a perfect assimilation ($\bar{R}$=0) and zero for an assimilation showing no improvement over a control integration with no assimilation ($\bar{R}$=$\bar{R}_c$).

For the BB system, the signal variance σ can be defined analogously to the gain, where now R is the error variance of the atmosphere or ocean,

$$\sigma = \frac{R - R_c}{R_c}. \tag{2.20}$$

Chapter 3

RESULTS

So far, we have developed two data assimilation systems - one based on the QG atmosphere model coupled to a slab ocean (the QG system), and another based on a two-variable linear stochastic model for midlatitude atmosphere-ocean interaction (the BB system). In this chapter, we will first identify and discuss three types of timescales at work in the BB model. The fundamental factors controlling the performance of the assimilation system are presented and understood using sensitivity analysis of the BB system. Finally, we will repeat these experiments in the QG system to establish the behaviors common to the two.

The bulk of the discussion will be directed at comparison with the QG model, so most of the experiments will employ a set of parameters equivalent to those used in the QG model. After each step, we will discuss the implications for the assimilation of annual average midlatitude observations for paleoclimate reconstructions. To address this, we will use the same set of parameters that BB used, which were developed to represent midlatitude GCM data. Both sets of parameters can be found in table 3.1.

3.1 Timescales

A key to understanding how data assimilation performs is to establish the important dynamical timescales in the system. We identify three such timescales in the BB system. The first arises naturally from the scaling of the equations. The second can be identified from the eigenvalues of the system. The last, and most important for data assimilation, describes predictability in the face of the accumulation of stochastic noise.

Table 3.1: **BB and QG equivalent parameters.** The parameters from BB98 are reproduced on the left. On the right are the equivalent values for the parameters chosen to correspond with the QG model.

	BB98	QG equivalent	
γ_a	1e+07	1.02e+07	$Jm^{-2}K^{-1}$
γ_o	2e+08	2.05e+08	$Jm^{-2}K^{-1}$
λ_{sa}	23.9	25.6	$Wm^{-2}K^{-1}$
λ_{so}	23.4	25.6	$Wm^{-2}K^{-1}$
λ_a	2.8	10.2	$Wm^{-2}K^{-1}$
λ_o	1.9	10.2	$Wm^{-2}K^{-1}$
a	1.12	1.4	
b	0.5	0.5	
c	1	1	
d	1.08	1.4	
β	20	20	

3.1.1 Processes

The BB model equations can be organized into a form with two components that can be related to physical processes. In equations 2.12-2.14, the atmosphere and ocean temperature tendencies depend on two time constants: a flux or "coupling" term proportional to the difference between surface air and slab ocean temperature (the "coupling timescale" τ_c) and a term proportional to the temperature anomaly itself (the "damping timescale" τ_r). The slab ocean timescales are longer than the atmosphere timescales by a factor of the ratio of their heat capacities, d. The heat capacities of the atmosphere and ocean depend on their masses. In the mixed layer, this is essentially the same thing as the depth so, we will refer to d as the depth.

Equations 2.1 and 2.2 are rearranged to the form of 2.12 and then consolidated as follows:

$$\begin{aligned} \frac{dT_a}{dt} &= \frac{1}{\tau_c}T_o - \frac{\tau_c+\tau_r}{\tau_c\tau_r}T_a + \mathcal{F} && = \frac{1}{\tau_c}T_o - \frac{1}{\tau_{eff}}T_a + \mathcal{F}, \\ \frac{dT_o}{dt} &= \frac{1}{d\tau_c}T_a - \frac{\tau_c+\tau_r}{d\tau_c\tau_r}T_o && = \frac{1}{d\tau_c}T_a - \frac{1}{d\tau_{eff}}T_o. \end{aligned}$$

Recall that T_a is an atmosphere temperature anomaly, T_o is a slab ocean temperature anomaly, and $\mathcal{F}$ is the forcing applied to the atmosphere.

These timescales determine the system dynamics. In this form, several aspects of the dynamics are clear: the ocean forces the atmosphere on the coupling timescale; the atmosphere forces the ocean with a heat-capacity multiple of the coupling timescale; and there are damping timescales of $\tau_{eff} = \tau_c\tau_r/(\tau_c+\tau_r)$ in the atmosphere, and $d\tau_{eff}$ in the ocean.

This framework shows the timescale separation of the atmosphere and ocean: all ocean timescales are a factor of d longer than atmospheric timescales ($\tau_o = d\tau_a$). If, for example, we know that the predictability timescale in the atmosphere is 2 weeks, then we might expect the ocean timescale to be $d*$(two weeks). However, since the system is coupled, we might also hope that some of this ocean memory could provide

some predictability in the atmosphere.

Throughout the rest of the chapter, we will identify the system parameters by their process timescales τ_c and d. We will not explore values of τ_r, which is fixed at 11.5 days in the QG model, except when considering the BB parameter suite. Table 3.2 shows the different process timescales for the QG equivalent and BB parameters.

Table 3.2: **Process timescales and depth.** The process timescales for the BB parameter choices and for the values equivalent to those used in the QG model.

	BB98	QG equivalent	
d	20	20	Slab ocean depth
τ_{ca}	4.8	4.7 days	Coupling timescale, atmosphere
$\tau_{ra} = \frac{\gamma_a}{\lambda_a}$	41	12 days	Damping timescale, atmosphere
τ_{co}	99	58 days	Coupling timescale, ocean
$\tau_{ro} = \frac{\gamma_o}{\lambda_o}$	1200	230 days	Damping timescale, ocean

3.1.2 Persistence

The process timescales form the physical basis for the time-dependent behavior of the system. A mathematical perspective on timescales comes from an eigenvalue analysis, which identifies basis functions, each with an associated e-folding timescale. Schopf (1985) contains a similar analysis for a slightly different deterministic model.

These timescales of the system can be found by solving the eigenvalue problem

$$\frac{d\vec{T}}{dt} = \mathsf{A}\vec{T} = \begin{bmatrix} -a & b \\ \frac{c}{\beta} & -\frac{d}{\beta} \end{bmatrix} \vec{T} = \begin{bmatrix} -\frac{1}{\tau_{eff}} & \frac{1}{\tau_c} \\ \frac{1}{d\tau_c} & -\frac{1}{d\tau_{eff}} \end{bmatrix} \vec{T} = \lambda\vec{T},$$

the solution of which is,

$$\lambda = \frac{-(d+1)\tau_{eff}^2\tau_c^2 \pm \sqrt{(d+1)^2\tau_{eff}^2\tau_c^2 - 4d\tau_{eff}^2\tau_c^2(\tau_c^2 - \tau_{eff}^2)}}{2d\tau_{eff}^2\tau_c^2}.$$

For the BB model, the eigenvalues are always negative. This means that the system is damped, so in the absence of forcing each eigenmode will return to zero with a timescale that can be found from the eigenvalues by $\tau = 1/\lambda$.

Figure 3.1 shows the time evolution of the initial-value problem that uses the eigenmodes as initial conditions and illustrates how modes of the system evolve. For a set of parameters corresponding to the QG model, there are two modes: fast-decaying ($\tau = 3$ days) and slow-decaying ($\tau = 89$ days). The fast timescale is close the the atmospheric coupling timescale, while the slow timescale is of the same order as the ocean damping timescale. In the fast-decaying mode, the ocean and atmosphere anomalies are in opposition, with a larger atmospheric than oceanic anomaly. The slow-decaying mode has atmospheric and oceanic anomalies of the same sign. The persistence of the slow mode can be regarded as a form of predictability. Once identified, its presence can be anticipated for some time time into the future.

3.1.3 Predictability

In the damped linear deterministic system, any error in an initial condition damps back to climatology. By contrast, in a stochastic system, errors grow because the noise forcing in the atmosphere accumulates, eventually overwhelming any known signal after some finite amount of time. This creates a horizon for the usefulness of state-dependent data assimilation. A key to understanding this is to quantify the timescale on which predictability is lost to the accumulation of stochastic noise. This will allow us to see how the accumulation of noise changes with the system. To accomplish this, we will draw upon the average predictability time (hereafter APT) from Delsole and Tippett (2009). The APT expresses the predictability timescale of an arbitrary model as the equivalent norm-independent exponential decay time of an orthogonal linear stochastic system. APT is based on the Mahalanobis metric,

$$S_\tau = \frac{1}{K} tr\left[(C_\infty - C_\tau)C_\infty^{-1}\right],$$

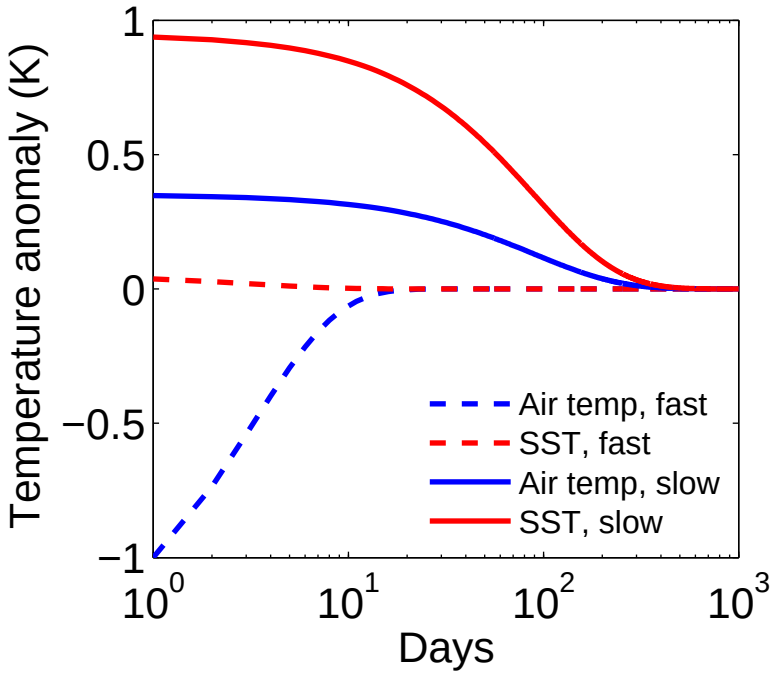

Figure 3.1: **Initial-value problem for eigenmodes of BB model, QG equivalent parameters.** The deterministic initial-value problem, initialized with the two eigenmodes of the BB model (Greg Hakim, personal communication). QG equivalent parameters are used (see table 3.1). Note the logarithmic time axis.

as a measure of predictability (e.g., Delsole and Tippett 2009). In this equation, K is the rank of the system matrix, C_τ is the error covariance for some forecast lead time τ, and C_∞ is the climatological error covariance. Then the APT for the discrete (in time) case is defined by

$$S = 2\sum_{1}^{\infty} S_\tau. \tag{3.1}$$

Perfect skill for one time unit has S_τ=1, so this sum is equal to the number of time units for which there is perfect skill. Delsole and Tippett (2009) also define a continuous version, but we won't use this here.

The BB model contains only two variables, each of which has inherently different thermal inertia, and a primary interest will be to understand what this means for how their predictability timescales differ. To this end, we will additionally consider predictability measures for an atmosphere norm

$$S^a_\tau = \frac{(\epsilon^a_\infty - \epsilon^a_\tau)}{\epsilon^a_\infty},$$

and an ocean norm

$$S^o_\tau = \frac{(\epsilon^o_\infty - \epsilon^o_\tau)}{\epsilon^o_\infty},$$

with corresponding APTs S^a and S^o, defined as for S. Note that the atmosphere and ocean predictability measures are different from the diagonal elements of the Mahalanobis norm, which is normalized by the climatological covariance matrix.

The instantaneous, rather than time-averaged, error covariance will be used to calculate C_τ, ϵ^a_τ and ϵ^o_τ (see equation 3.17 for the definition of instantaneous error covariance). This is for two reasons: for comparison with other timescales, which are internal to the BB model and so defined only for instantaneous error variances, and for simplicity. To find the upper-bounding timescale over which the noise has not yet overwhelmed the signal, we will neglect the residual, or initial, component of the error variance and consider only error due to the accumulation of noise. This means we will only use the error in the second term of equation 3.17.

There is a mathematical relationship between the eigenvalue timescale and the predictability timescale that can be seen in equation 3.17. The persistence timescale is simply the inverse of the eigenvalues of the system matrix A. If D is the matrix with the eigenvalues of A on its diagonal, then,

$$\text{noise accumulation } = \int_0^t \mathsf{V} e^{\mathsf{D}(t-s)} \mathsf{V}^{-1} \mathsf{N}\mathsf{N}^T \left(\mathsf{V} e^{\mathsf{D}(t-s)} \mathsf{V}^{-1}\right)^T ds,$$

as we will see in equation 3.17. So, the integrand of the predictability timescale decays as the square of the persistence timescale.

Figure 3.2 shows an example of the predictability norms and APT calculations. For the QG equivalent parameters, the atmosphere APT is shorter than the atmospheric coupling timescale and the fast-decaying persistence timescale. In the ocean, the APT is longer than the slow-decaying persistence timescale. The APT for the Mahalanobis norm is the timescale for the sum of the total variance (normalized by the climatological covariance matrix) to decay, and it splits the difference between the ocean and atmosphere APT in this case.

3.1.4 Timescales for midlatitude climate

Using the parameters from BB98, we can determine each of the timescales defined above for the midlatitude climate. These parameters are shown in table 3.2. Figure 3.3 shows the initial-value problem for the eigenmodes of the system with BB parameters. The eigenvalue timescales for this set of parameters is 4 days for the atmosphere and 156 days for the ocean.

The predictability timescales for the BB parameters are different than those for the QG equivalent parameters. Figure 3.4 shows calculation of the APT. The predictability timescales for the BB parameters are 4.6 days in the atmosphere and 170 days in the ocean. These timescales are longer because two of the QG process timescales are shorter than their BB98 correspondents: the damping timescale τ_r and the ocean coupling timescale τ_{co}.

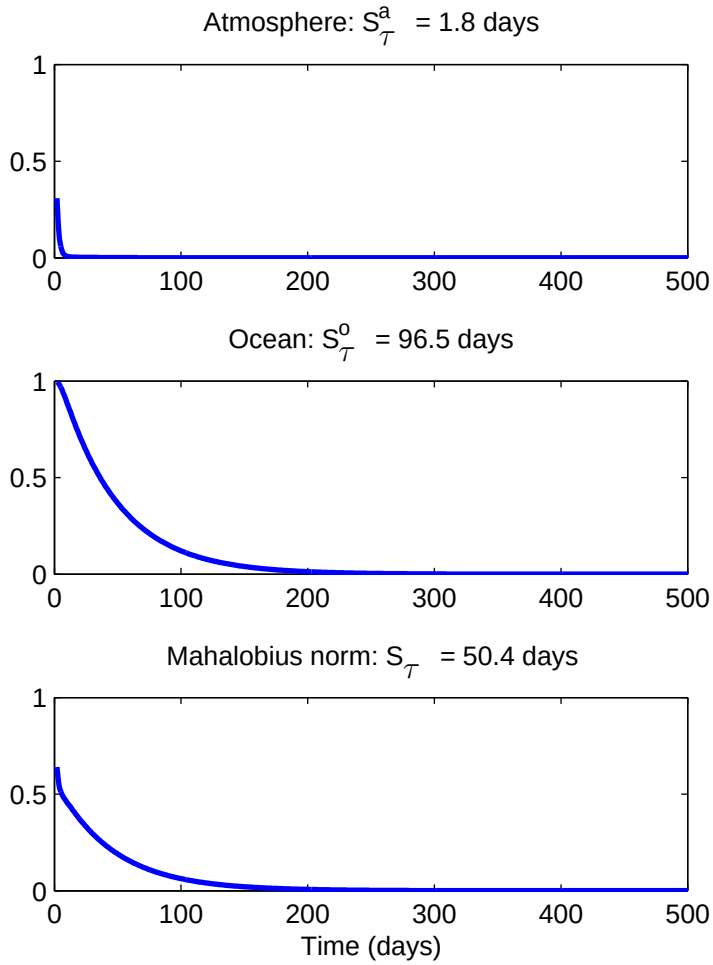

Figure 3.2: **Average predictability time, QG equivalent parameters.** Predictability measures used to calculated average predictability time in the atmosphere, ocean, and with the Mahalanobis norm. Resulting APTs are noted in the title of each panel. The parameters used in this calculation are $\tau_c = 4.7$ days and d=20.

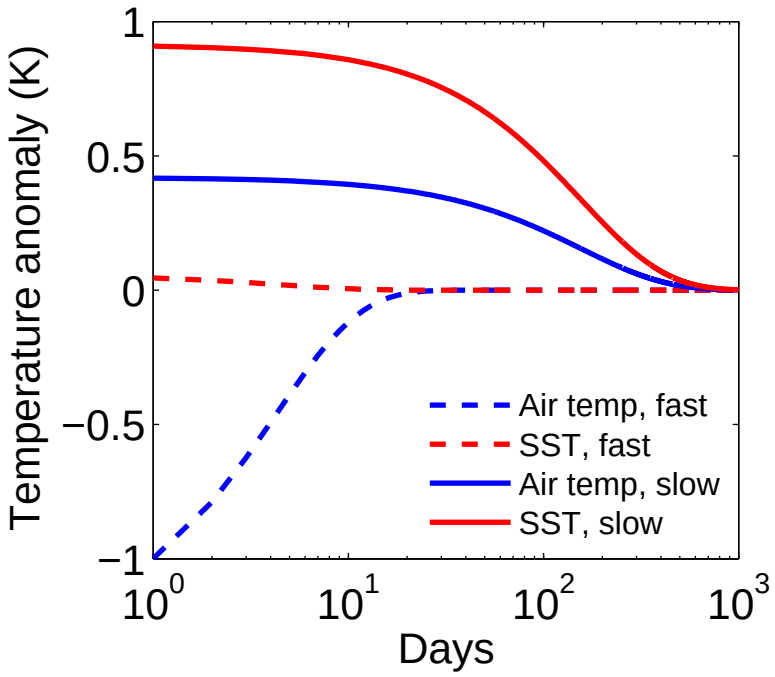

Figure 3.3: **Initial-value problem for eigenmodes of BB model, BB parameters.** As in figure 3.1, but with parameters used in BB98. Note the logarithmic time axis.

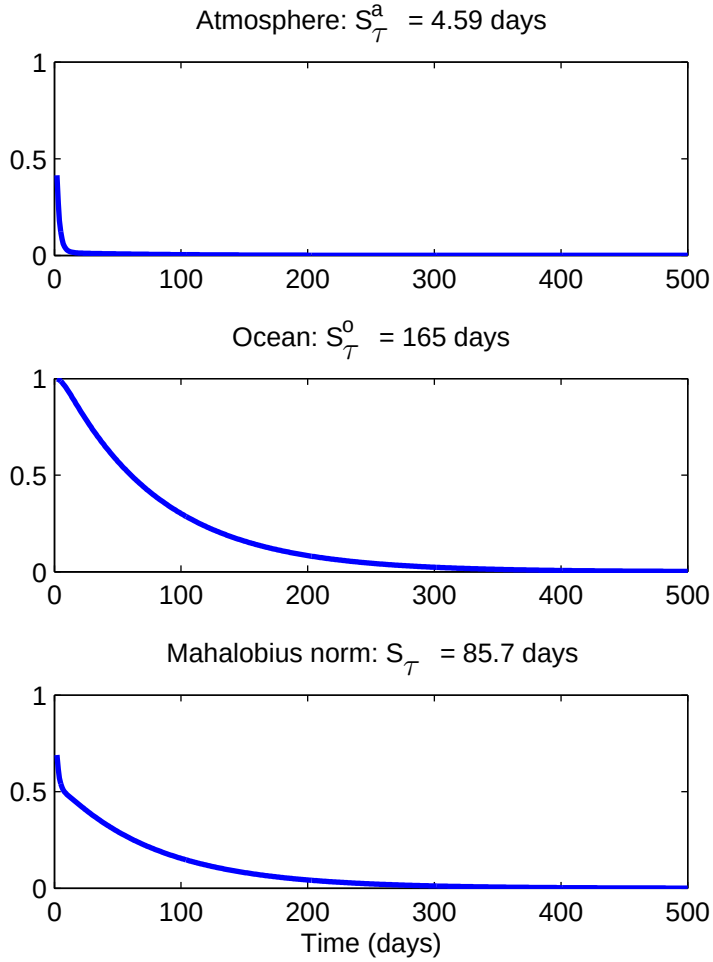

Figure 3.4: **Average predictability time for BB model, BB parameters.** As in figure 3.2, but with BB98 parameters.

3.2 Investigations with the BB data assimilation system

In this section, we will first develop the math behind BB system – the data assimilation system based around the model of BB98. We then explore how the BB system performs as a function of model parameters. The results can be understood in terms of the predictability timescale.

3.2.1 Solving the stochastic evolution of BB error covariances

In this section we will develop the equations for the instantaneous error covariance of BB model states, by using stochastic calculus methods, and then consider how it can be evaluated. We will then go over the development of the covariance of time-averaged BB model states, which is similar. Then we will go over the evaluation of the Kalman filter equation used here. Finally an example data assimilation cycle will be presented and described.

The system is an Ornstein-Uhlenbeck process, which known widely enough in stochastic systems to be discussed in introductory texts. This derivation follows mostly from various examples in Øksendal (2003), but Gardiner (2004) was also used. There are two forms of stochastic calculus: Itô and Stratonovich. Because random processes are not differentiable, stochastic integrals must be defined as approximate integrals. The difference between the two versions of stochastic calculus is in the choice of location of endpoints for the intervals of integration. Itô chose the left end-point of the interval, while Stratonovich chose the midpoints. In general, these choices result in different solutions, which are arrived at from different rules and formulas. For Ornstein-Uhlenbeck processes, the two calculi coincide. This is because the coefficient of the random process $\vec{W}$ does not depend on the independent variable (here, $N \neq f(\vec{T})$). This is stated at the end of Øksendal (2003) chapter 3 (p 37) and discussed in Gardiner (2004) section 4.3 (equations 4.3.39-40, here $b = N$ so $\partial_x b = 0$).

We will begin the derivation from the equations 2.5 and 2.6 for the BB model,

$$\vec{T} = \begin{bmatrix} T_a \\ T_o \end{bmatrix} \qquad \frac{d\vec{T}}{d\hat{t}} = \begin{bmatrix} -a & b \\ \frac{c}{\beta} & -\frac{d}{\beta} \end{bmatrix} \vec{T} + \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \vec{W}$$

$$\frac{d\vec{T}}{d\hat{t}} = \mathsf{A}\vec{T} + \mathsf{N}\vec{W}, \tag{3.2}$$

where $\vec{W}$ is a Gaussian white noise process with mean 0 and variance 1 at all frequencies. Here, the noise term $\vec{W}$ is written as a vector multiplied by a matrix for consistency, but there is only one noise process in this system so it is really just a one-dimensional variable (e.g. $\vec{W} = [W \quad 0]$, for scalar W).

Equation 3.2 is written in the form of Øksendal (2003) equation 3.1.2 (hereafter abbreviated Ø3.1.2). Chapter 3 of Øksendal (2003) examines how to properly define and write the noise term. The end of the chapter shows that the solution to an equation of the form Ø3.1.2 should be in the form of Ø3.3.3. This is the same form as Ø4.1.3, which can be written in shorthand as Ø4.1.6. The shorthand version is

$$dT_t = \mathsf{A}\vec{T}dt + \mathsf{N}dB_t. \tag{3.3}$$

From here, we can closely follow example 5.1.3 to find the solution at time t. We choose $e^{-\mathsf{A}t}$ as an integrating factor and integrate to solve for the evolution of $\vec{T}$,

$$\vec{T} = e^{\mathsf{A}t}\vec{T}(0) + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dB_s. \tag{3.4}$$

We won't take the final step of integrating by parts to get rid of the integral over dB_s, as example 5.1.3 does, so that later we'll be able to exploit Itô's isometry.

Now, we can integrate to find the expected variance over an infinite number of experiments. For comparison with the ensemble Kalman filter (a Monte Carlo algorithm), this expectation can be thought of as being taken over an infinite number of ensemble members. I'll introduce a second, separate noise process U_s, so that we will now have a noise process for each ensemble member and another noise process

for the true states or experiments, where truth is simply one of the members with no data assimilation. We will only calculate expected distributions, and we'll take two expectations: one over "ensemble members" or the p.d.f. of the distribution (denoted by angle brackets $\langle\cdot\rangle$), and another over "assimilation experiments," which are expectations over the noise process for the true state (denoted $E[\cdot]$). The error of an ensemble member is its Euclidean distance from the true state. The error at time t is

$$\vec{\epsilon}(t) = \vec{T}(t) - \vec{T_T}(t) \tag{3.5}$$

$$= e^{\mathsf{A}t}\vec{T}(0) + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dB_s - \left(e^{\mathsf{A}t}\vec{T_T}(0) + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dU_s\right) \tag{3.6}$$

$$= e^{\mathsf{A}t}\vec{\epsilon}(0) + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dB_s - e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dU_s. \tag{3.7}$$

The ensemble mean error is,

$$\langle\vec{\epsilon}(t)\rangle = \left\langle\vec{T}(t)\right\rangle - \vec{T_T}(t) \tag{3.8}$$

$$= e^{\mathsf{A}t}\left\langle\vec{T}(0)\right\rangle + \left\langle e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dB_s\right\rangle - \left(e^{\mathsf{A}t}\vec{T_T}(0) + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dU_s\right) \tag{3.9}$$

$$= e^{\mathsf{A}t}\langle\vec{\epsilon}(0)\rangle - e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dU_s. \tag{3.10}$$

The deviation of the error from the ensemble mean, denoted by $(\cdot)^*$ is,

$$\vec{\epsilon}(t) - \langle\vec{\epsilon}(t)\rangle \equiv \vec{\epsilon}(t)^* = e^{\mathsf{A}t}\vec{\epsilon}(0)^* + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dB_s. \tag{3.11}$$

There are two quantities that will be calculated. First, the relevant data assimilation algorithm uses the error covariance of time-averaged states. But, the evolution of this quantity depends on the total error covariance at the beginning of the forecast period, so we will calculate the instantaneous error covariance as well. The definition of the total error covariance and its relationship to the error covariance of time-averaged

states is,

$$\mathsf{C}_{total} = \left\langle \epsilon^* \epsilon^{*T} \right\rangle \tag{3.12}$$

$$\epsilon = \bar{\epsilon} + \epsilon' \tag{3.13}$$

$$\mathsf{C}_{average} = \left\langle \bar{\epsilon}^* \bar{\epsilon}^{*T} \right\rangle \tag{3.14}$$

$$\mathsf{C}_{residual} = \mathsf{C}_{total} - \mathsf{C}_{average}. \tag{3.15}$$

$$\tag{3.16}$$

where primes $()'$ denote the deviation of $T(t)$ from the time average $\bar{(\cdot)}$ for the forecast period, vector symbols have been dropped for simplicity, and cross terms vanish in the time average in second to last equation. The data assimilation changes only $\mathsf{C}_{average}$ directly for this choice of approximation, as noted in the discussion of the time-averaging algorithm (section 2.2).

Instantaneous error covariance

The error covariance for a particular experiment at a particular time is defined as $\left\langle \epsilon^* \epsilon^{*T} \right\rangle$. Here we will find a solution for the expected covariance over an infinite number of experiments, $E[\left\langle \epsilon^* \epsilon^{*T} \right\rangle]$. Employing equation 2.4,

$$\begin{aligned} \bar{\epsilon}^* \bar{\epsilon}^{*T} = e^{\mathsf{A}t} \vec{\epsilon}(0)^* \vec{\epsilon}(0)^{*T} \left(e^{\mathsf{A}t}\right)^T + \left(e^{\mathsf{A}t} \int_0^t e^{-\mathsf{A}s} \mathsf{N} dB_s\right) \left(e^{\mathsf{A}t} \int_0^t e^{-\mathsf{A}s} \mathsf{N} dB_s\right)^T \\ + \left(e^{\mathsf{A}t} \int_0^t e^{-\mathsf{A}s} \mathsf{N} dB_s\right) \left(e^{\mathsf{A}t} \vec{\epsilon}(0)\right)^T. \end{aligned}$$

When expectation operators over ensembles and experiments are applied, all cross terms (those with only one term integrated over dB_s) will vanish because,

$$E\left[\int_S^T f dB_t\right] = 0 \qquad \text{(Øksendal 2003 theorem 3.2.1.iii)}.$$

Taking the expectation over the ensemble, the terms with one integral over dB_s will vanish because the expectation of the stochastic process is 0. Noting that $\left\langle e^{\mathsf{A}t} \right\rangle = e^{\mathsf{A}t}$,

the error covariance will be,

$$\left\langle \vec{\epsilon}^{*}\vec{\epsilon}^{*T} \right\rangle = e^{\mathsf{A}t} \left\langle \vec{\epsilon}(0)^{*}\vec{\epsilon}(0)^{*T} \right\rangle \left(e^{\mathsf{A}t}\right)^{T} + \left\langle \left(e^{\mathsf{A}t}\int_{0}^{t} e^{-\mathsf{A}s}\mathsf{N}dB_{s}\right)\left(e^{\mathsf{A}t}\int_{0}^{t} e^{-\mathsf{A}s}\mathsf{N}dB_{s}\right)^{T} \right\rangle.$$

Now we take the expectation over infinite experiments. After distributing two transposes, we have,

$$\begin{aligned} E\left[\left\langle \epsilon\epsilon^{T} \right\rangle\right] = {} & e^{\mathsf{A}t} E\left[\left\langle \vec{\epsilon}(0)\vec{\epsilon}(0)^{T} \right\rangle\right] \left(e^{\mathsf{A}t}\right)^{T} \\ & + e^{\mathsf{A}t} \left\langle \left(\int_{0}^{t} e^{-\mathsf{A}s}\mathsf{N}dB_{s}\right)\left(\int_{0}^{t} e^{-\mathsf{A}s}\mathsf{N}dB_{s}\right)^{T} \right\rangle \left(e^{\mathsf{A}t}\right)^{T}. \end{aligned}$$

Itô's isometry transforms the expectation of the product of two integrals into time integrals of the square of the integrands. The relevant formulation of Itô's isometry is Gardiner (2004) 4.2.42 (which is more general than Øksendal 2003 3.1.11) is,

$$E\left[\int_{S}^{T} f(t)dB_{t} \int_{S}^{T} g(t)dB_{t}\right] = E\left[\int_{S}^{T} f(t)g(t)dt\right].$$

Noting that the transpose commutes with the integral, the expected error covariance is,

$$\begin{aligned} E\left[\left\langle \vec{\epsilon}^{*}\vec{\epsilon}^{*T} \right\rangle\right] = {} & e^{\mathsf{A}t} E\left[\left\langle \vec{\epsilon}(0)^{*}\vec{\epsilon}(0)^{*T} \right\rangle\right] \left(e^{\mathsf{A}t}\right)^{T} \\ & + e^{\mathsf{A}t} \int_{0}^{t} \left(e^{-\mathsf{A}s}\mathsf{N}\right)\left(e^{-\mathsf{A}s}\mathsf{N}\right)^{T} ds \left(e^{\mathsf{A}t}\right)^{T}, \end{aligned}$$

$$\boxed{E\left[\left\langle \epsilon^{*}\epsilon^{*T} \right\rangle\right] = e^{\mathsf{A}t} E\left[\left\langle \vec{\epsilon}(0)^{*}\vec{\epsilon}(0)^{*T} \right\rangle\right] \left(e^{\mathsf{A}t}\right)^{T} + \int_{0}^{t} e^{\mathsf{A}(t-s)}\mathsf{N}\mathsf{N}^{T} \left(e^{\mathsf{A}(t-s)}\right)^{T} ds} \quad (3.17)$$

The two terms equation 3.17 govern the two ways the error covariance evolves. The first is growth or decay of the initial error covariance, depending on the eigenvalues of A. The second term is the accumulation of noise, which increases error covariance.

Evaluating the solution

Since A is linear and constant, we can solve this equation for any initial instantaneous error covariance as long as A has an inverse. For the BB model, A is readily invertible.

The matrix with columns that are eigenvectors of A diagonalizes A ($\mathsf{D} = \mathsf{V}^{-1}\mathsf{AV}$), and the entire time dependence can be condensed into the terms of the diagonal matrix D ($e^{\mathsf{A}t} = \mathsf{V}e^{\mathsf{D}t}\mathsf{V}^{-1}$). An advantage of the simplicity of the BB model is that we can use a brute force element-by-element method to solve the integral. The time-independent coefficient calculation will be shown only once. All subsequent noise-integral terms can be found in a similar way by evaluating the appropriate integral over the exponential terms. We will work with the elements of three matrices: $\mathsf{M}(t,s)$, which contains all of the time-dependent parts; C_o, which is the initial error covariance matrix; and N, which is the matrix of coefficients for the noise term. To find the elements of M, we'll have to work with the elements of V and D.

$$\mathsf{M}(t,s) = e^{\mathsf{A}(t-s)} = \begin{bmatrix} m_{11}(t,s) & m_{12}(t,s) \\ m_{21}(t,s) & m_{22}(t,s) \end{bmatrix} \qquad \mathsf{C}_o = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \qquad \mathsf{N} = \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix}.$$

The first term in the exact equation can be evaluated directly. But the time integral in the second term requires more work.

$$\begin{aligned}
&\int_0^t e^{\mathsf{A}(t-s)}\mathsf{N}\mathsf{N}^T\left(e^{\mathsf{A}(t-s)}\right)^T ds \\
&\qquad = \int_0^t \mathsf{M}(t,s)\mathsf{N}\mathsf{N}^T\mathsf{M}(t,s)^T ds \\
&\qquad = \int_0^t \begin{bmatrix} m_{11}(t,s) & m_{12}(t,s) \\ m_{21}(t,s) & m_{22}(t,s) \end{bmatrix} \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} m_{11}(t,s) & m_{21}(t,s) \\ m_{12}(t,s) & m_{22}(t,s) \end{bmatrix} ds \\
&\qquad = \int_0^t \begin{bmatrix} m_{11}(t,s) & m_{12}(t,s) \\ m_{21}(t,s) & m_{22}(t,s) \end{bmatrix} \begin{bmatrix} N^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} m_{11}(t,s) & m_{21}(t,s) \\ m_{12}(t,s) & m_{22}(t,s) \end{bmatrix} ds \\
&\qquad = N^2 \int_0^t \begin{bmatrix} m_{11}^2 & m_{21}m_{11} \\ m_{21}m_{11} & m_{21}^2 \end{bmatrix} ds \\
&\qquad = N^2 \begin{bmatrix} \int_0^t m_{11}^2 ds & \int_0^t m_{21}m_{11} ds \\ \int_0^t m_{21}m_{11} ds & \int_0^t m_{21}^2 ds \end{bmatrix}.
\end{aligned}$$

The next step is to find the elements of M. Though we only need two elements (m_{11} and m_{21}) right now, we'll need the rest later. The rank-two system makes it

possible to directly use the inverse of V, which would be impractical for large matrices. The integral evaluates to

$$\begin{aligned}\mathsf{M}(t) = e^{\mathsf{A}t} = \mathsf{V}e^{\mathsf{D}t}\mathsf{V}^{-1} &= \frac{1}{\det(\mathsf{V})}\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22}\end{bmatrix}\begin{bmatrix} e^{d_1 t} & 0 \\ 0 & e^{d_2 t}\end{bmatrix}\begin{bmatrix} v_{22} & -v_{12} \\ -v_{21} & v_{11}\end{bmatrix} \\ &= \frac{1}{\det(\mathsf{V})}\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22}\end{bmatrix}\begin{bmatrix} v_{22}e^{d_1 t} & -v_{12}e^{d_1 t} \\ -v_{21}e^{d_2 t} & v_{11}e^{d_2 t}\end{bmatrix} \\ &= \frac{1}{\det(\mathsf{V})}\begin{bmatrix} v_{11}v_{22}e^{d_1 t} - v_{12}v_{21}e^{d_2 t} & -v_{11}v_{12}e^{d_1 t} + v_{12}v_{11}e^{d_2 t} \\ v_{21}v_{22}e^{d_1 t} - v_{21}v_{22}e^{d_2 t} & -v_{12}v_{21}e^{d_1 t} + v_{22}v_{11}e^{d_2 t}\end{bmatrix}.\end{aligned}$$

Now it's clear that each element of the integral will have three exponential terms: $e^{2d_1 t}$, $e^{(d_1+d_2)t}$, and $e^{2d_2 t}$. I won't carry out the element multiplication any farther, because it can be implemented more cleanly (with fewer errors) this way. The integral can be carried out separately for each element, where everything but the exponential parts is constant. So now we have the sum of many integrals of the form $\int_0^t e^{as}ds$ for $a = 2d_1, d_1 + d_2$, and $2d_2$

$$\int_0^t e^{as}ds = \frac{1}{a}\left(e^{at} - 1\right).$$

Error covariance of time-averaged states

Recognizing that the error of the time-averaged states is the same as the time-average of the error, and then employing equation 3.5 above, we can formulate the error covariance of the time-averaged states,

$$\begin{aligned}\bar{\epsilon} &= \frac{1}{\tau}\int_0^\tau \epsilon(t)dt = \frac{1}{\tau}\int_0^\tau \left(e^{\mathsf{A}t}\bar{\epsilon}(0) + e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dB_s - e^{\mathsf{A}t}\int_0^t e^{-\mathsf{A}s}\mathsf{N}dU_s\right)dt \\ \bar{\epsilon}^*\bar{\epsilon}^{*T} &= \left(\frac{1}{\tau}\int_0^\tau \epsilon(t)^*dt\right)\left(\frac{1}{\tau}\int_0^\tau \epsilon(t)^*dt\right)^T.\end{aligned}$$

The cross-terms will cancel as they did for instantaneous error covariance. Then we are left with,

$$E[\langle \bar{\epsilon}^* \bar{\epsilon}^{*T} \rangle] = \frac{1}{\tau^2} \int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \left(\int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \right)^T + \frac{1}{\tau^2} \left(\int_0^\tau \left(\int_0^t e^{\mathsf{A}(t-s)} \mathsf{N} dB_s \right) dt \right) \left(\int_0^\tau \left(\int_0^t e^{\mathsf{A}(t-s)} \mathsf{N} dB_s \right) dt \right)^T.$$

In order to use Itô's isometry again, we can bring the dt integral inside the dB_s, dU_s integrals:

$$\int_0^\tau \int_0^t e^{\mathsf{A}(t-s)} ds dt = \int_0^\tau \int_s^\tau e^{\mathsf{A}(t-s)} dt ds \text{ so}$$

$$E[\langle \bar{\epsilon}^* \bar{\epsilon}^{*T} \rangle] = \frac{1}{\tau^2} \int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \left(\int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \right)^T + \frac{1}{\tau^2} \left(\int_0^\tau \int_s^\tau e^{\mathsf{A}(t-s)} \mathsf{N} dt dB_s \right) \left(\int_0^\tau \int_s^\tau e^{\mathsf{A}(t-s)} \mathsf{N} dt dB_s \right)^T.$$

Now we can use Itô's isometry to arrive at

$$E[\langle \bar{\epsilon}^* \bar{\epsilon}^{*T} \rangle] = \frac{1}{\tau^2} \int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \left(\int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \right)^T + \frac{1}{\tau^2} \int_0^\tau \left(\int_s^\tau e^{\mathsf{A}(t-s)} \mathsf{N} dt \right) \left(\int_s^\tau e^{\mathsf{A}(t-s)} \mathsf{N} dt \right)^T ds. \tag{3.18}$$

Manipulating the first term into something more manageable is accomplished by distributing transposes and integrating (note that I is the identity matrix and $(\cdot)^{-T}$ indicates an inverse and a transpose),

$$\begin{aligned} \frac{1}{\tau^2} \int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \left(\int_0^\tau e^{\mathsf{A}t} \epsilon(0)^* dt \right)^T &= \frac{1}{\tau^2} \left(\int_0^\tau e^{\mathsf{A}t} dt \right) \epsilon(0)^* \epsilon(0)^{*T} \left(\int_0^\tau \left(e^{\mathsf{A}t} \right)^T dt \right) \\ &= \frac{1}{\tau^2} \mathsf{A}^{-1} \left(e^{\mathsf{A}\tau} - I \right) \epsilon(0)^* \epsilon(0)^{*T} \left(e^{\mathsf{A}\tau} - I \right)^T \mathsf{A}^{-T}. \end{aligned}$$

For the second term, we return to M and its elements,

$$\begin{aligned}
&\frac{1}{\tau^2}\int_0^\tau \left(\int_s^\tau e^{\mathsf{A}(t-s)}\mathsf{N}dt\right)\left(\int_s^\tau e^{\mathsf{A}(t-s)}\mathsf{N}dt\right)^T ds \\
&\quad= \frac{1}{\tau^2}\int_0^\tau \left(\int_s^\tau \mathsf{M}(t,s)\mathsf{N}dt\right)\left(\int_s^\tau \mathsf{M}(t,s)\mathsf{N}dt\right)^T ds \\
&\quad= \frac{1N^2}{\tau^2}\int_0^\tau \left(\begin{bmatrix}\int_s^\tau m_{11}(t,s)dt & 0\\ \int_s^\tau m_{21}(t,s)dt & 0\end{bmatrix}\right)\left(\begin{bmatrix}\int_s^\tau m_{11}(t,s)dt & 0\\ \int_s^\tau m_{21}(t,s)dt & 0\end{bmatrix}\right)^T ds \\
&\quad= \frac{1N^2}{\tau^2}\int_0^\tau \left(\begin{bmatrix}\int_s^\tau m_{11}(t,s)dt & 0\\ \int_s^\tau m_{21}(t,s)dt & 0\end{bmatrix}\right)\left(\begin{bmatrix}\int_s^\tau m_{11}(t,s)dt & \int_s^\tau m_{21}(t,s)dt\\ 0 & 0\end{bmatrix}\right) ds \\
&\quad= \frac{1N^2}{\tau^2}\int_0^\tau \begin{bmatrix}\left(\int_s^\tau m_{11}(t,s)dt\right)^2 & \int_s^\tau m_{11}(t,s)dt\int_s^\tau m_{21}(t,s)dt\\ \int_s^\tau m_{21}(t,s)dt\int_s^\tau m_{11}(t,s)dt & \left(\int_s^\tau m_{21}(t,s)dt\right)^2\end{bmatrix} ds.
\end{aligned}$$

Similarly to the instantaneous case, each of the elements of this matrix will be the product of a sum of exponential terms. Each of the exponential integral parts can be broken down as,

$$\begin{aligned}
&\int_0^\tau \left(\int_s^\tau e^{a(t-s)}dt\int_s^\tau e^{b(t-s)}dt\right) ds \\
&\quad= \int_0^\tau \left(\frac{1}{a}\left(e^{a(\tau-s)}-1\right)\frac{1}{b}\left(e^{b(\tau-s)}-1\right)\right) ds \\
&\quad= \frac{1}{ab}\int_0^\tau \left(e^{(a+b)(\tau-s)}-e^{a(\tau-s)}-e^{b(\tau-s)}+1\right) ds \\
&\quad= \frac{1}{ab}\left(-\frac{1}{a+b}\left(1-e^{(a+b)\tau}\right)+\frac{1}{a}\left(1-e^{a\tau}\right)+\frac{1}{b}\left(1-e^{b\tau}\right)+\tau\right).
\end{aligned}$$

Careful Matlab implementation allows calculation of the error covariance of quantities averaged over time τ from a total initial error covariance.

Expected data assimilation: Kalman Filter error covariance update for a two variable system

A tremendous advantage of the two-variable BB system is that the simplest form of the Kalman filter equations can be used, and practical issues of implementation can

be avoided. Below the equations are evaluated for the case where both the atmosphere and slab ocean are observed. The case where only one variable is observed is simpler because the denominator consists only of scalars; it is the standard sequential problem.

First, we define the background error variances and covariance,

$$
\begin{aligned}
|A| &= \left\langle \left(\bar{T}_a - \langle \bar{T}_a \rangle\right)^2 \right\rangle, \\
|B| &= \left\langle \left(\bar{T}_o - \langle \bar{T}_o \rangle\right)^2 \right\rangle, \\
C &= \left\langle \left(\bar{T}_a - \langle \bar{T}_a \rangle\right) \left(\bar{T}_o - \langle \bar{T}_o \rangle\right) \right\rangle.
\end{aligned}
$$

We will evaluate the analysis error covariance A, which requires the Kalman gain K,

$$
\begin{aligned}
\mathsf{A} &= (\mathsf{I} - \mathsf{KH})\mathsf{B}, \\
\mathsf{K} &= \mathsf{BH}^T[\mathsf{HBH}^T + \mathsf{R}]^{-1},
\end{aligned}
$$

where B is the background error covariance, H is the observation operator, R is the observation error variance, and I is the identity matrix.

First, we define B, H and R as ,

$$
\mathsf{B} = \begin{bmatrix} |A| & C \\ C & |B| \end{bmatrix}, \qquad \mathsf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \qquad \mathsf{R} = \begin{bmatrix} r_a & 0 \\ 0 & r_b \end{bmatrix}.
$$

Then we can build up the Kalman gain matrix,

$$
\begin{aligned}
\mathsf{HBH}^T &= \mathsf{IBI}^T = \mathsf{B}, \\
\mathsf{HBH}^T + \mathsf{R} = \mathsf{B} + \mathsf{R} &= \begin{bmatrix} |A| + r_a & C \\ C & |B| + r_b \end{bmatrix}.
\end{aligned}
$$

The inverse of $[\mathsf{HBH}^T + \mathsf{R}]$ can be carried out exactly for a two-variable system, enabling an exact Kalman gain matrix calculation,

$$
[\mathsf{HBH}^T + \mathsf{R}]^{-1} = [\mathsf{B} + \mathsf{R}]^{-1} = \frac{1}{(|A| + r_a)(|B| + r_b) - C^2} \begin{bmatrix} |B| + r_b & -C \\ -C & |A| + r_a \end{bmatrix},
$$

$$
\mathsf{BH}^T = \mathsf{B}.
$$

$$\mathsf{K} = \mathsf{B}[\mathsf{B}+\mathsf{R}]^{-1} = \frac{1}{(|A|+r_a)(|B|+r_b)-C^2}\begin{bmatrix}|A| & C\\ C & |B|\end{bmatrix}\begin{bmatrix}|B|+r_b & -C\\ -C & |A|+r_a\end{bmatrix}$$

$$\boxed{\mathsf{K} = \frac{1}{(|A|+r_a)(|B|+r_b)-C^2}\begin{bmatrix}A(|B|+r_b)-C^2 & -Cr_a\\ -Cr_b & |B|(|A|+r_a)-C^2\end{bmatrix}}.$$

Then we can build the analysis error covariance

$$\mathsf{I}-\mathsf{KH} = \mathsf{I}-\mathsf{K} = \frac{1}{(|A|+r_a)(|B|+r_b)-C^2}\begin{bmatrix}r_a(|B|+r_b) & -Cr_a\\ -Cr_b & r_b(|A|+r_a)\end{bmatrix},$$

$$\begin{aligned}\mathsf{A} &= (\mathsf{I}-\mathsf{KH})\mathsf{B}\\ &= \frac{1}{(|A|+r_a)(|B|+r_b)-C^2}\begin{bmatrix}r_a(|B|+r_b) & -Cr_a\\ -Cr_b & r_b(|A|+r_a)\end{bmatrix}\begin{bmatrix}|A| & C\\ C & |B|\end{bmatrix}\end{aligned}$$

$$\boxed{\mathsf{A} = \frac{1}{(|A|+r_a)(|B|+r_b)-C^2}\begin{bmatrix}|A|r_a(|B|+r_b)-C^2r_a & Cr_ar_b\\ Cr_br_a & -C^2r_b+|B|r_b(|A|+r_a)\end{bmatrix}}.$$

The pieces are now in place to implement data assimilation with the BB model.

Schematic illustration of BB system

Each data assimilation cycle, the Kalman filter operates on the error covariance from the ensemble of time-averaged BB model states. To recall what this is, each member of an infinite ensemble of BB model states is averaged over the forecast period. The true time-averaged state is taken to be the average over one particular model trajectory; the error is then the difference between each ensemble member and the true state. The expected error variance of time-averaged states is taken over an infinite number of data assimilation experiments. The instantaneous error is the corresponding error variance of ensemble states at any given moment in time. In order to show how these calculations are made, the top panel of figure 3.5 shows the internal workings of one component of the BB system during an assimilation cycle of length τ'. The

thin lines show instantaneous error variance; the thick lines show the error variance of time-averaged states.

There will always be some initial error at $t = 0$. This component of the error will decay with time because the system is damped. However, there is another component, the accumulation of noise from the stochastic forcing, which quickly becomes the dominant error source. Thus, the total error variance (calculated in equation 3.17), which is the sum of these two components, increases over time. As it asymptotically approaches the control error variance, the signal decays towards zero (see section 2.2 for the definition of signal). Note that the error does not grow indefinitely as it would for a random walk, and again this is because the system is damped.

At time τ', the average state of each ensemble member is taken over the assimilation cycle; its error variance is calculated. This error variance is related to the average of the instantaneous errors by equations 3.12. Each component and the total of this time-average are shown as a horizontal line in Figure 1b. This is the background error covariance used in the assimilation.

The system is now able to proceed with the next step of the assimilation cycle. With the final instantaneous error variance and error of the time-averaged states, we can calculate the residual error; this is the difference between the final instantaneous error and the error of time-averaged states. Now, hypothetical time-averaged observations can be assimilated to decrease the error, resulting in an analysis. This updated error is added back to the residual error to calculate the new initial instantaneous error. The cycle begins again.

We can explore how the total error and its components vary for other averaging times, and this is shown in the bottom panel of Figure 3.5. The particular averaging time used in the top panel is indicated by a thin vertical line. As averaging time increases, the error of time-averaged states becomes larger, and signal becomes smaller. The accumulation of noise dominates over residual errors for long averaging times.

An example integration of the BB system to convergence is shown in figure 3.6.

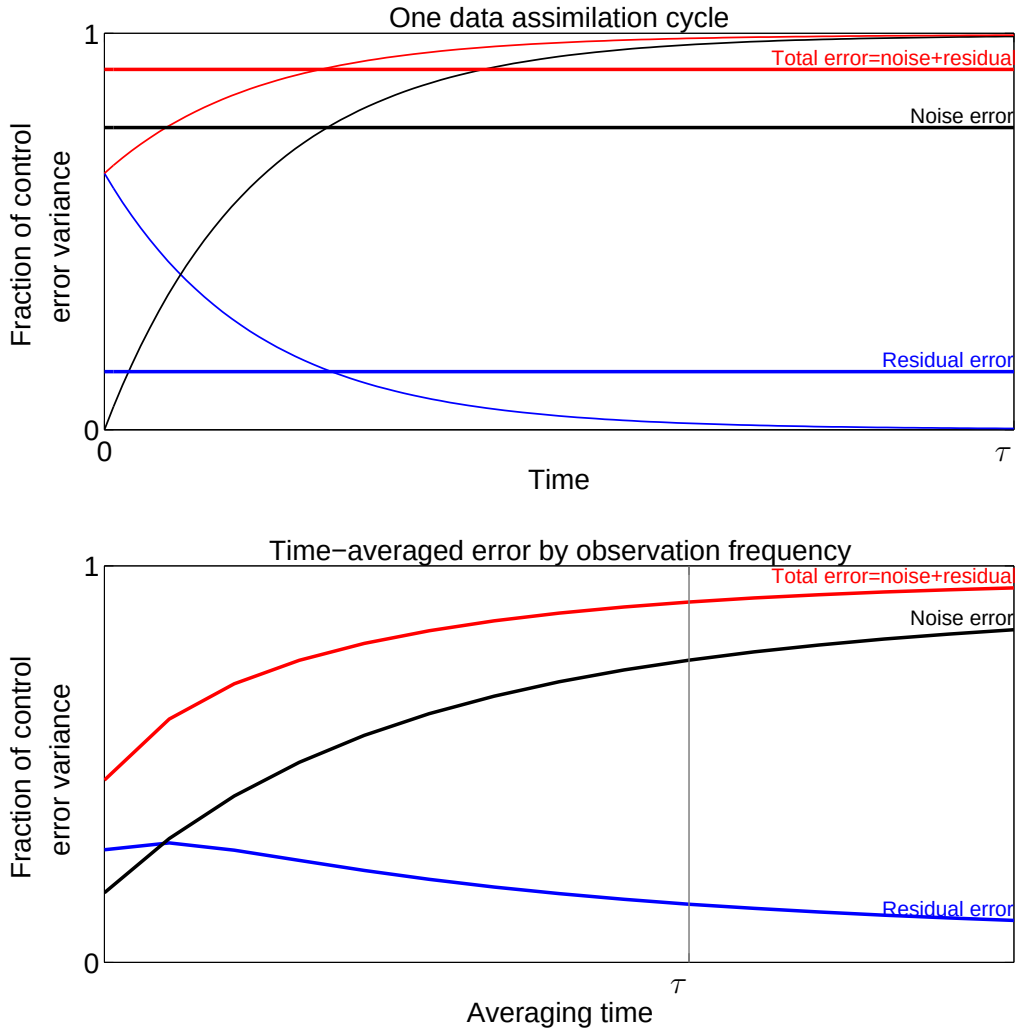

Figure 3.5: **The internal workings of the BB system.** The top panel illustrates the errors instantaneous and time-averaged states for one assimilation cycle of length τ'. The bottom panel shows how the error and its components varies for different averaging times. Thin lines are instantaneous error variance normalized by the control; and thick lines are time-averaged normalized error variance.

The instantaneous and time-averaged background and analysis errors are found by integrating the system for enough data assimilation cycles to converge on an equilibrium solution for a given set of system parameters, averaging period, and observational error. Because the variance of the model varies greatly for different parameters and averaging times, observational error is always specified as a fraction of the control error variance. Initially, the error starts from zero and grows, until asymptoting to an equilibrium value. When it has stopped changing (judged by visual inspection), which happens by 700 days for all integrations considered here, then the error variances for the control case (with no data assimilation) are declared found. The final values are used for analysis. To find the equilibrium error variance for assimilation experiments, the assimilation system starts from the background errors and iterates for the same number of assimilation cycles as the spinup, though in practice it converges much faster.

3.2.2 BB data assimilation system investigations

We will now perform experiments with the completed BB system to establish the behavior of time-averaged data assimilation. More specifically, we characterize how state-dependent data assimilation skill varies with different model parameters and averaging times, and compare the skill in the ocean and atmosphere.

In all experiments shown, both the atmosphere and the ocean are observed. The error variance of the ocean and atmosphere in these experiments is 10% of the climatological variance for the averaging time of the experiment. We will examine the skill of the background, or prior states, which are the time-averaged states at the end of a forecast period before observations are assimilated. We will compare the background states to a control case with no observations. In cases where background states still have skill, state-dependent data assimilation is an improvement over using climatological covariance as background. We will not consider the analysis skill here, though an actual reconstruction would be based on the analysis rather than the background.

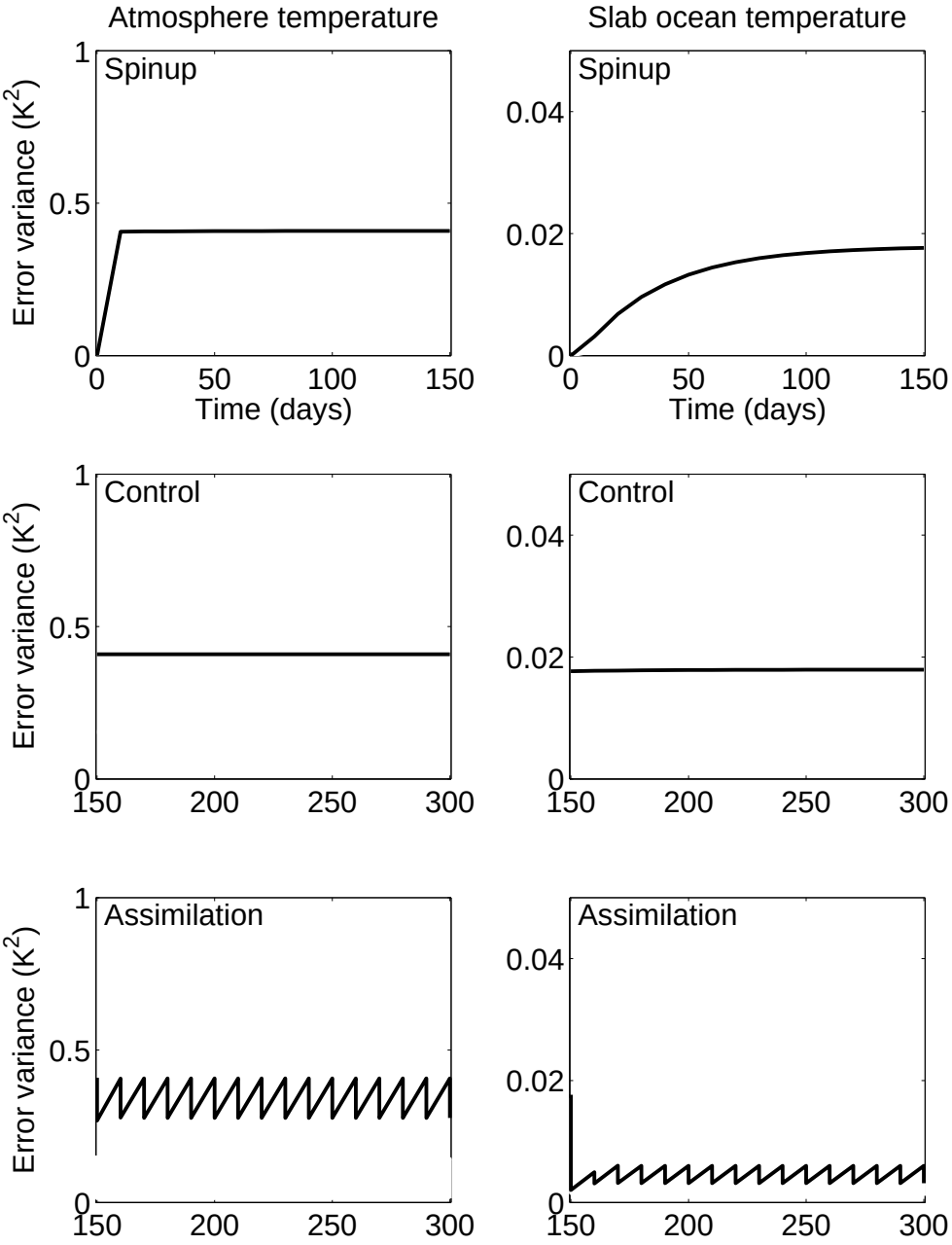

Figure 3.6: **BB system example.** The evolution of the error variance is shown for instantaneous and time-averaged states (atmosphere left and ocean right) for the BB data assimilation system. The top panels show the "Spinup" phase (30 assimilation cycles). The middle panels show the "Control," case with no assimilation. The bottom panels show sawtooth plots for the "Assimilation" (observational error is 1/10 of final spinup variance. Model parameters are $\tau_c = 2.9$ days, $d = 20$, and an averaging time of 10 days.

Instead, we focus on when the model has state-dependent skill, which is most directly evaluated by examining the background error covariance.

The metric for skill in the BB system will be the fractional reduction in total error variance of the background state compared to the control, which is called the signal variance, defined in equation 2.20. It is closely related to the gain, the metric used in the QG system experiments, discussed in section 2.2.

We will proceed with the experiments by first establishing some expectations for parameter sensitivity based on physical reasoning from the processes. Then we will consider an overview of how background skill varies with averaging time, slab ocean depth and coupling timescale. From this overview, we will identify some behaviors that confirm our expectations and some that don't. Finally, we will use the insights gained to characterize the data assimilation skill for annual observations in the mid-latitudes and consider how it would change for plausible parameter adjustments.

Parameter sensitivity overview

How might we expect skill to vary with averaging time? If we were concerned with instantaneous data assimilation, it would be obvious that skill should decrease with observation frequency. Higher frequency observations mean more frequent observations, which have more skill. For time-averaged data assimilation, the climatological variance of the states also decreases with averaging time. In these experiments, we scale observation error with climatological variance to maintain constant error in a fractional sense. So indeed we expect that skill will decrease with averaging time.

How might skill vary with depth? In response to an increase in slab ocean depth, we certainly expect the skill in the ocean to increase because of a deeper ocean has more thermal inertia and moves more slowly. The system is coupled, so we would also expect the increased skill in the ocean to influence the atmosphere.

For an increase in coupling (a decrease in coupling timescale), we could consider the influence of the atmosphere on the ocean and the influence of the ocean on the

atmosphere. Both have the same result: increasing skill in the atmosphere and decreasing skill in the ocean. First, the ocean would "feel" the atmosphere, including its stochastic forcing more, which would decrease skill in the ocean, possibly with a compensating increase it in the atmosphere. Second, the atmosphere would "feel" the longer timescale in the ocean which (without considering point 1) would increase the skill in the atmosphere, possibly with a compensating decrease in the ocean. In reality, both of these influences are at play.

We will now perform the experiments, taking care to note where our expectations are borne out and where they are not. Note that the overview experiments in figures 3.7 and 3.8 are designed to be analogous to the experiments with the QG system shown later in figures 3.21 and 3.22.

Figure 3.7 shows how skill varies with averaging time and ocean depth for τ_c=4.7 days. As expected in both the atmosphere and ocean, skill declines as averaging time increases. It does so more rapidly for the atmosphere than for the ocean. A more notable result is that the atmosphere and ocean respond very differently to increasing slab ocean depth. In the ocean, the skill for any given averaging time increases with depth. In contrast, the skill in the atmosphere decreases slightly as the ocean depth increases. This counter-intuitive result will be further explored in section 3.2.2.

Figure 3.8 shows how skill varies with coupling and averaging time. A similar dependence of skill on averaging time is again apparent. In response to increased coupling (decreased coupling timescale), the skill in both the atmosphere and ocean decreases.

In the next four subsections, we will explore these results in more detail. First we will examine the differences between the ocean and atmosphere skill. Then we will look at the averaging time, which confirms our expectations. Then, we will look at why skill in the atmosphere decreases with increasing depth, followed by understanding the results as a function of coupling strength.

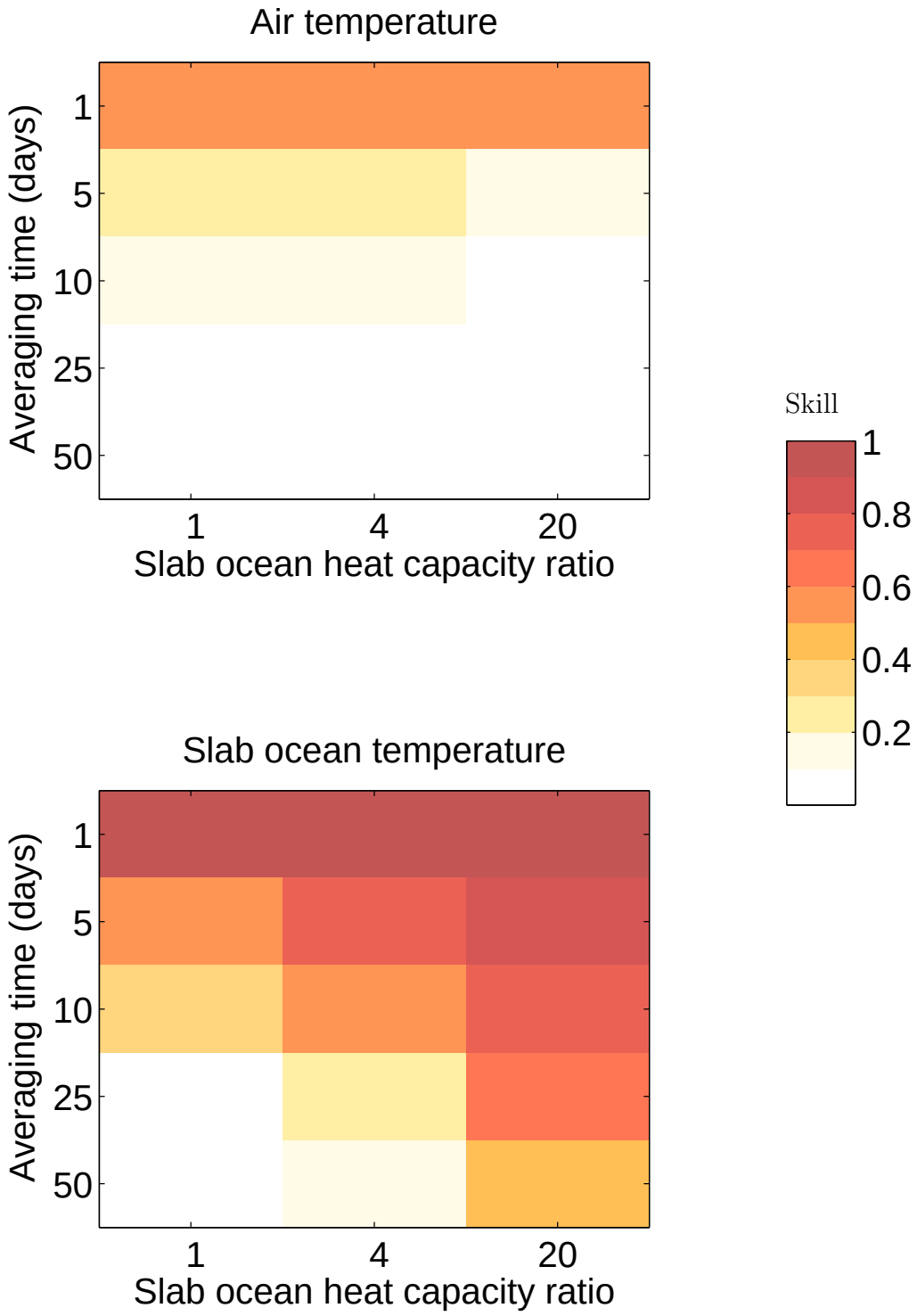

Figure 3.7: **BB system – dependence of background skill on slab ocean depth and averaging time.** Parameter sensitivity for the BB data assimilation system, with $\tau_c = 4.7$ days. Compare with figure 3.21.

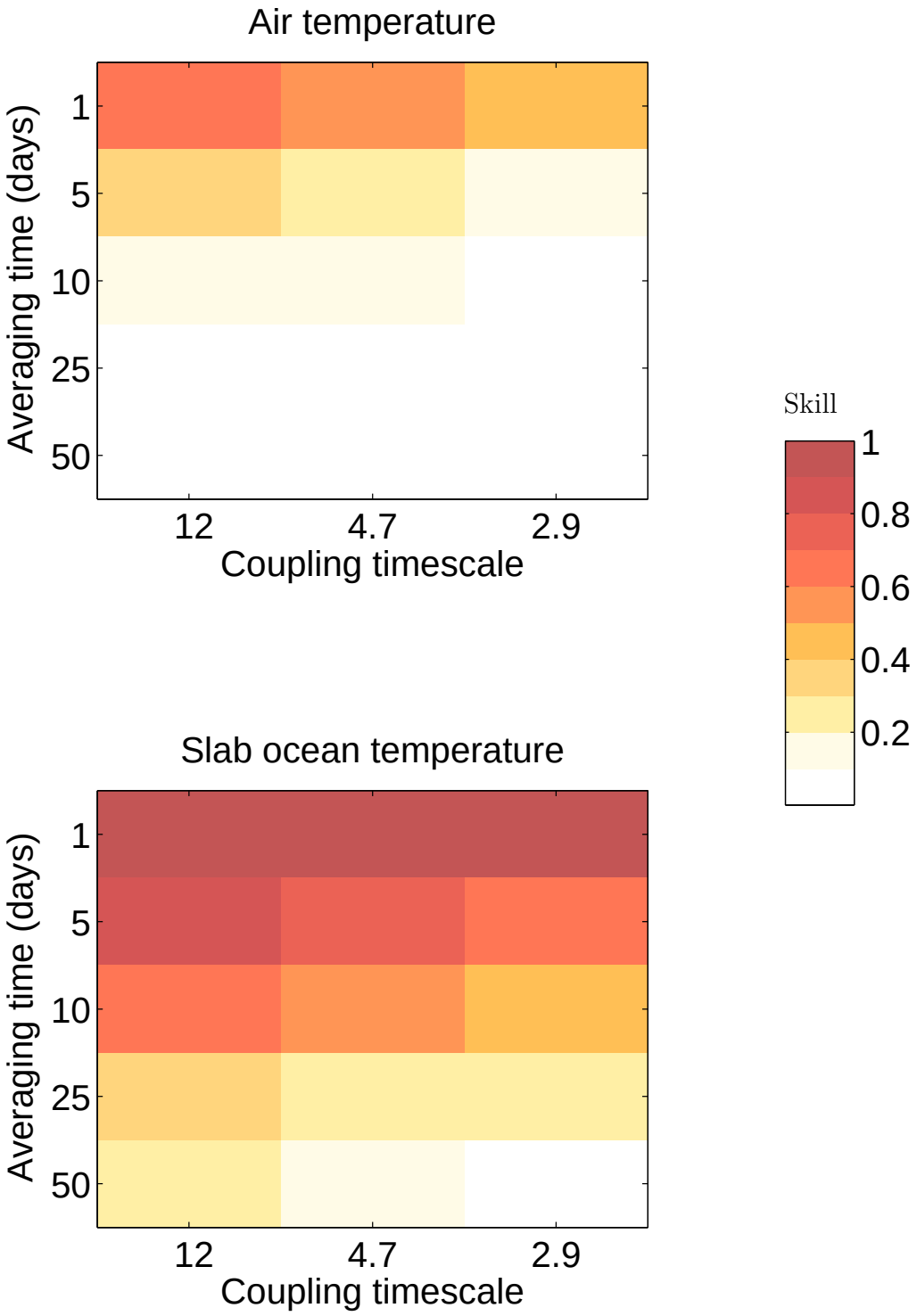

Figure 3.8: **BB system – dependence of background skill on coupling and averaging time.** Parameter sensitivity for the BB data assimilation system, with $d = 4$. Compare with figure 3.22.

Ocean and atmosphere differences

In figures 3.7 and 3.8 there is a striking difference between skill in the ocean and atmosphere. This is due to two factors. The first is simply the longer thermal inertia of the ocean. The process timescales in the ocean are each a factor of d longer than in the atmosphere, so persistence and predictability in the ocean should be longer. But even for d=1, when the ocean and atmosphere have the same thermal inertia, the ocean is has longer predictability than the atmosphere. Figures 3.10 and 3.11 show two examples of ocean and atmosphere differences when d=1. For τ_c=4.7 days, the APT is 2.5 days for the atmosphere and 9.2 days for the ocean. From this we can conclude that there is another difference. The clearest explanation is simply that the noise forcing is applied to the atmosphere and takes time to be communicated to the ocean.

Sensitivity to averaging time

We have reasoned that skill should decrease with increasing averaging time, and the overview experiments agree with our expectation. Figure 3.9 quantifies how skill varies with τ for d=20, τ_c=4.7 days. For the shortest averaging times, we see that both ocean and atmosphere have skillful background states. The skill in the atmosphere (the blue line) decreases very rapidly, but the skill in the ocean (the black line) decreases much more slowly toward zero. Recalling that the APTs for the atmosphere was 1.8 days and for the ocean was 97 days, this is roughly the rate that the skill falls off with averaging time. This is a direct result of the ocean's greater memory.

Figures 3.10 and 3.11 show the total error, its two components (discussed in section 3.2.1), and the signal, when d=1 and τ_c=4.7 days, and for two different averaging times, τ; τ=5 (Figure 3.10), and τ=25 (Figure 3.11). d=1 means the ocean heat capacity equals the atmosphere's. This is chosen in order to deliberately focus on the atmosphere ocean differences when their process timescales are equal. Because the

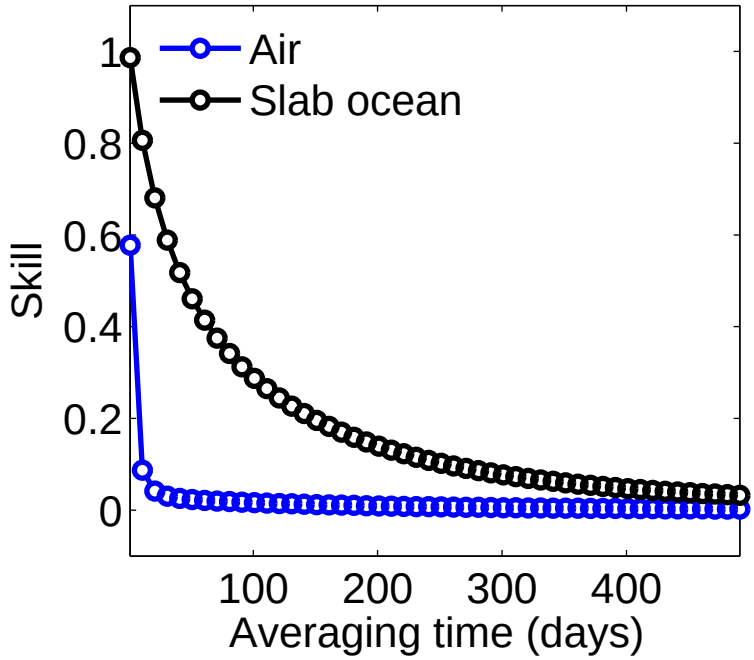

Figure 3.9: **BB system background skill – averaging time.** Skill varies with averaging time. The parameters are τ_c=4.7 d and d=20.

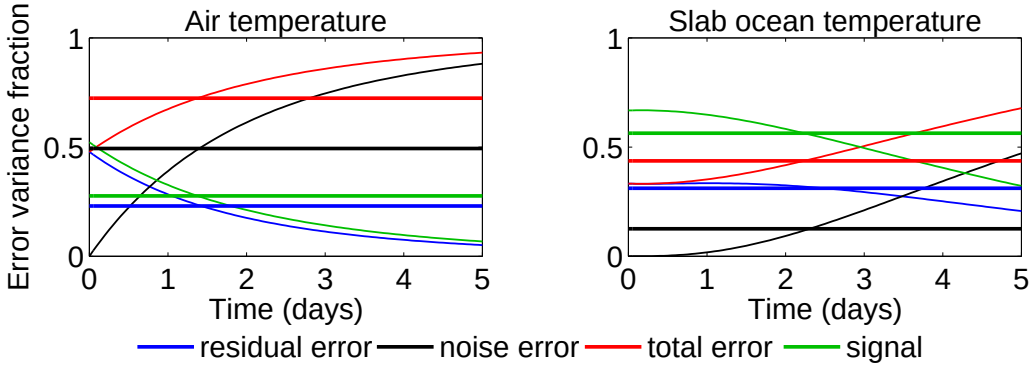

Figure 3.10: **Skill and error – τ=5.** Skill, error and its components for τ_c=4.7 days and d=1. Thin lines are instantaneous error variance normalized by the control; and thick lines are time-averaged normalized error variance.

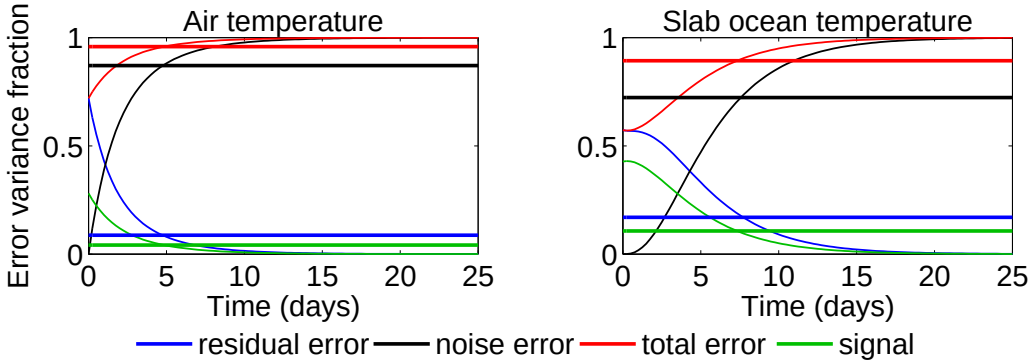

Figure 3.11: **Skill and error – τ=25.** As in figure 3.10 but with $\tau = 25$.

stochastic forcing is applied directly to the atmosphere, there is a lag between the error accumulation in the atmosphere and its communication to the ocean. This is seen comparing the thin black lines in both panels of both Figure 3.10 and Figure 3.11. For short averaging time, this ocean lag causes a big difference in the time-averaged errors of the atmosphere and ocean, seen by comparing the thick horizontal lines in Figure 3.11. As a result the skill is considerably greater in the ocean than the atmosphere. For longer averaging times the ocean lag matters less, and so the skill in atmosphere and ocean is comparable.

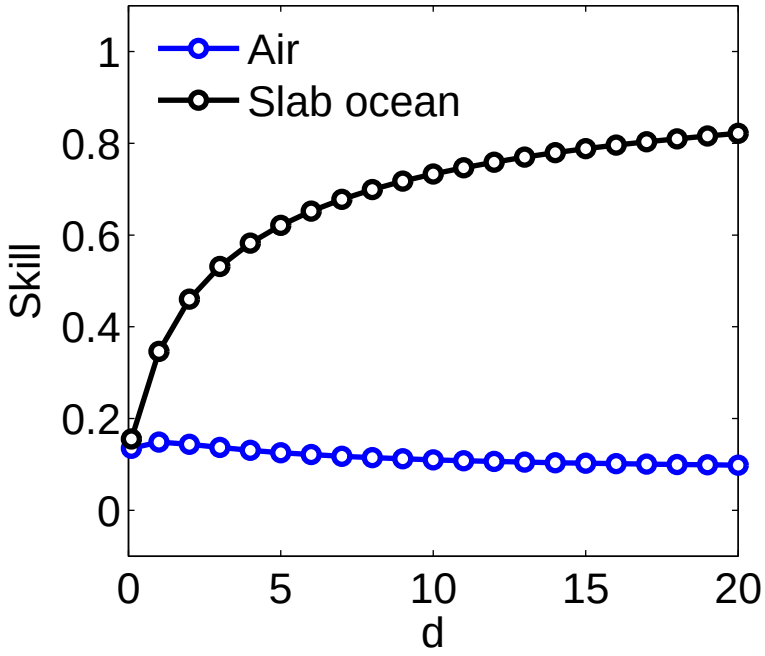

Figure 3.12: **BB system background skill – slab ocean depth.** Skill varies with d, for τ_c=4.7 days and τ = 10 days.

Sensitivity to depth

We now turn to the first of two results that did not meet our expectations. Why does skill in the atmosphere decrease as the depth of the slab ocean increases? Figure 3.12 is a cross section though Figure 3.7 and shows that as depth varies, there is a strong increase in ocean skill, and a slight decline in atmosphere skill.

Figures 3.13 and 3.14 show the components of error for two different slab ocean depths, d=1 and d=20. For the deeper mixed layer, there is a small but identifiable decrease in the time it takes for error due to noise forcing to accumulate in the atmosphere. The APT (defined in section 3.1.3) quantifies this accumulation of noise. In the atmosphere, the APT is 2.5 days for d=1 and 1.8 days for d=20. The eigenvalue timescales are longer for both modes with higher d: the fast-decaying mode increases

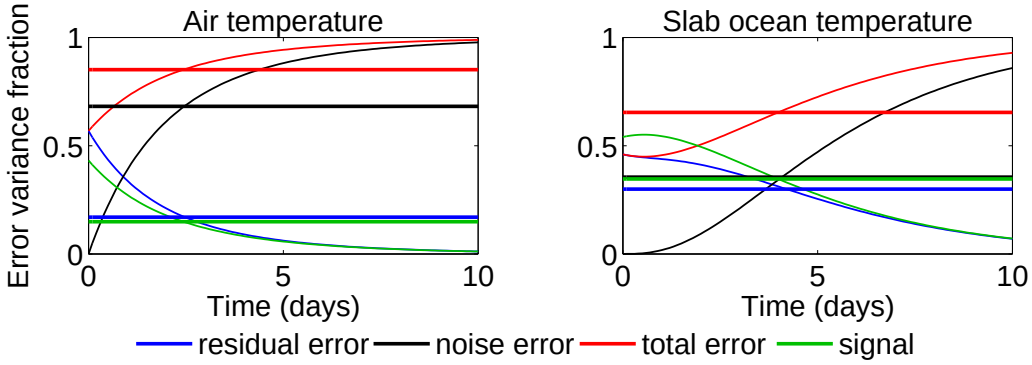

Figure 3.13: **Skill and error – d=1.** Skill, error and its components for τ_c=4.7 days, τ=10 days, and d=1. Thin lines are instantaneous error variance normalized by the control; and thick lines are time-averaged normalized error variance.

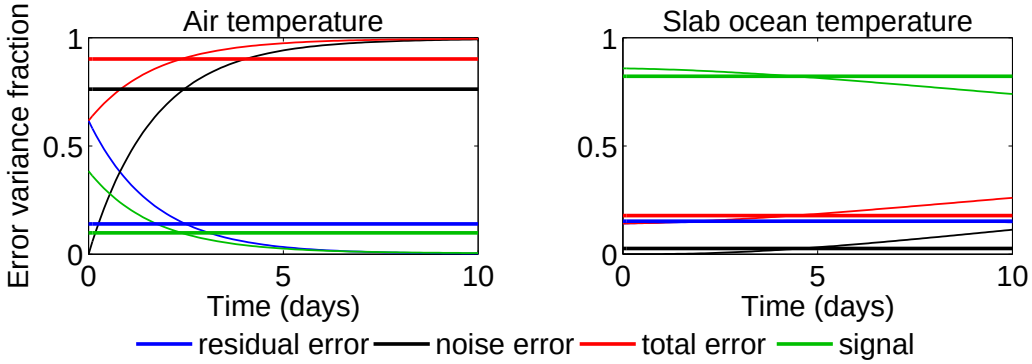

Figure 3.14: **Skill and error – d=20.** As in figure 3.13 but for d=20.

from 2.2 to 3.3 days, and the slow-decaying mode increases from 6.7 to 90 days. The eigenmode initial-value problem is shown in figure 3.15. As the thermal inertia of the ocean increases, the persistence timescales of the atmosphere increase, but the predictability timescales decrease, indicating that the atmosphere loses predictability because of noise accumulates more quickly with increasing depth.

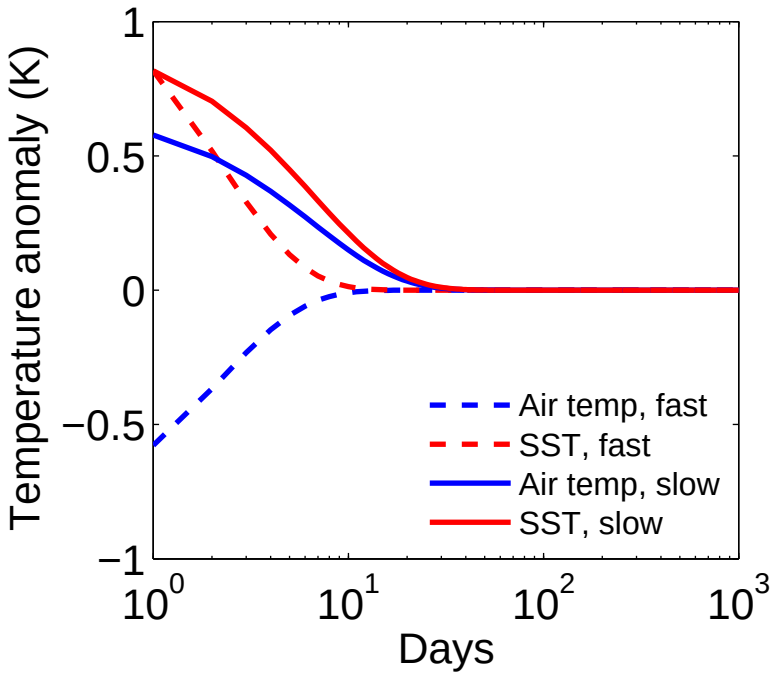

Figure 3.15: **Initial-value problem for eigenmode of BB model, QG equivalent parameters, d=1.** The deterministic initial-value problem, initialized with the two eigenmodes of the BB model, with d=1. Note the logarithmic time axis.

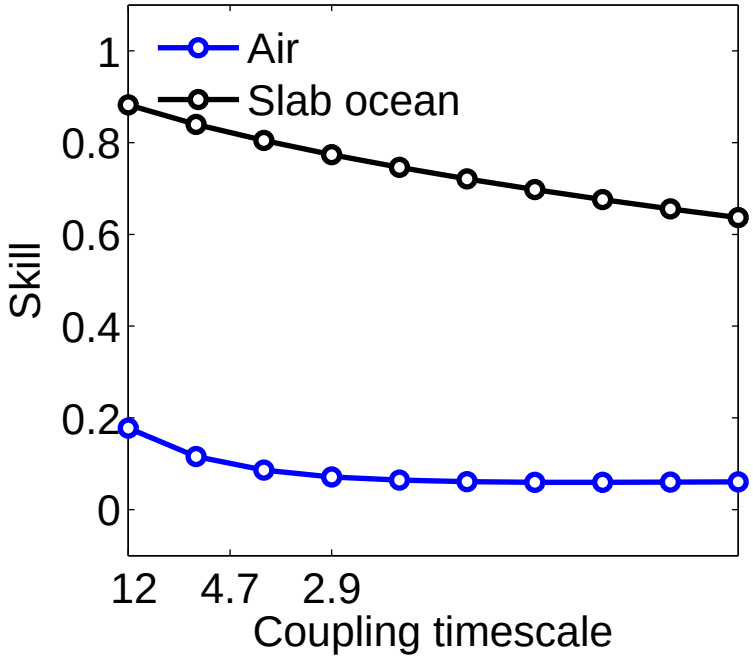

Figure 3.16: **BB system background skill – coupling.** Skill varies with coupling τ_c for τ=10 days and d=20.

Sensitivity to coupling

Next, we turn to the second of our counter-intuitive results. Figure 3.16 shows how skill changes with coupling. Why does skill decrease in the atmosphere with increased coupling?

As with the first conundrum, the explanation is found by considering the noise. The errors due to noise accumulate faster in the case where there is more coupling. In contrast to variations with depth, noise accumulates faster in both the atmosphere and in the ocean. Figures 3.17 and 3.18 show the components of error for coupling timescales of 12 and 2.9 days. We can quantify the changes with the persistence timescale andthe APT. As the coupling timescales decreases from 12 to 2.9 days, the APT in the ocean decreases from 150 days to 73, and in the atmosphere it decreases

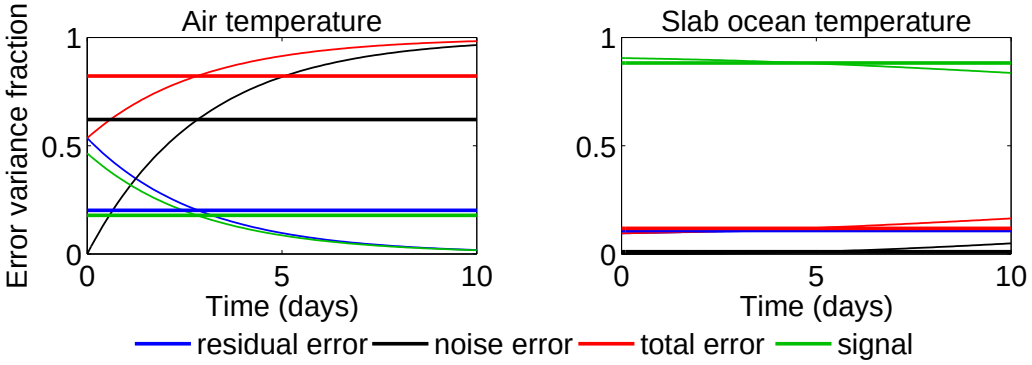

Figure 3.17: **Skill and error – τ_c=12 days.** Signal, noise and its components for d=20, τ=10 days, and τ_c=12 days. Thin lines are instantaneous error variance normalized by the control; and thick lines are time-averaged normalized error variance.

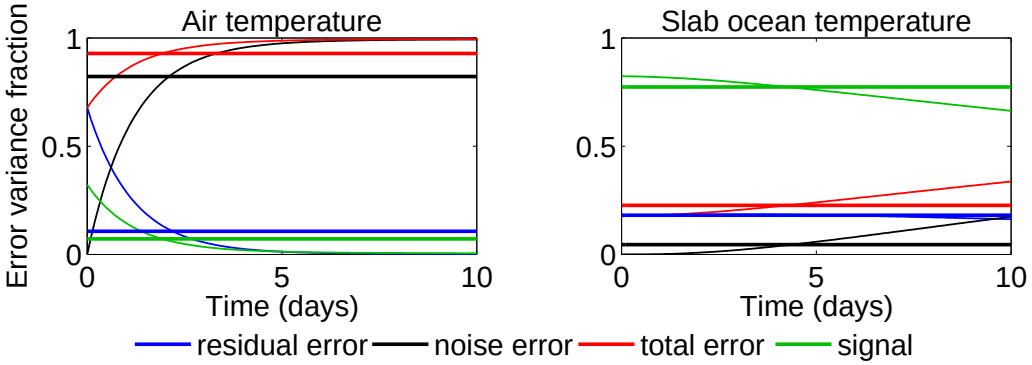

Figure 3.18: **Skill and error – τ_c=2.9 days.** As in figure 3.17, but for τ_c=2.9 days.

from 3.7 days to 1.2. The eigenvalue timescales also decrease from 5.7 to 2.3 days for the fast-decaying mode and 130 to 69 days for the slow-decaying mode.

Annual mean data assimilation in the midlatitudes

We have explored how data assimilation skill varies across a wide range of parameters, and we are now in a position to evaluate the potential skill for parameters appropriate to the reconstruction of midlatitude, interannual paleoclimate variability.

Figure 3.19 shows the signal and noise for the standard BB model parameters and

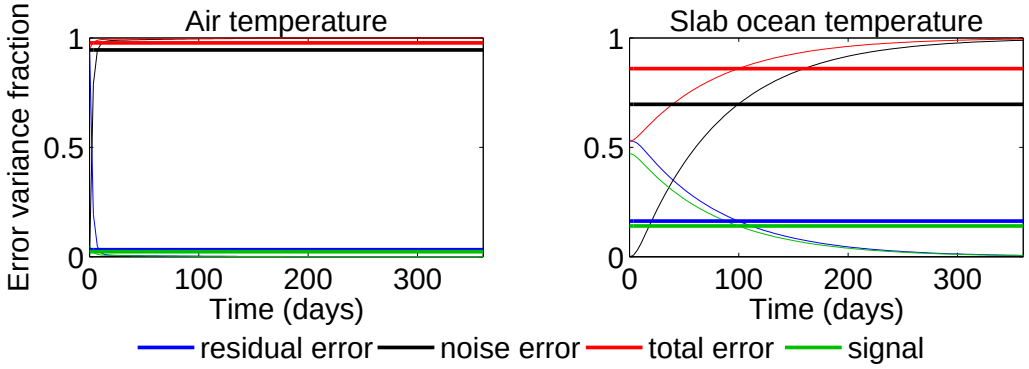

Figure 3.19: **Skill and error, BB parameters, perfect observations.** Signal and noise in the BB system for parameters used in BB98, with perfect observations. Thin lines are instantaneous error variance normalized by the control; and thick lines are time-averaged normalized error variance.

annual assimilation with perfect observations. This is an upper bound on the skill of the background forecast for annual observations. The signal variance is almost zero in the atmosphere but 14% in the ocean. Figure 3.20 shows the slightly more realistic case where observational error variance is the same as the climatological error variance. This is more characteristic of paleoclimate observations, which often have high uncertainty. In this case the skill is still substantial; the signal variance of time-averaged states in the atmosphere is still almost zero but still 9% in the ocean. Skill has not yet been completely lost for the ocean.

Based on the results for the parameter sensitivity investigations in section 3.2.2, how might the potential for assimilation change with different parameters? A mixed layer depth of $d = 20$, corresponding to 40 m, is on the lower end of realistic values. In another example from the literature, Bretherton and Battisti (2000), use $d = 40$ (100 m). Above, we saw that increasing the thermal inertia of the mixed layer increases the predictability time of the ocean. But, we also saw that it does not improve the predictability time in the atmosphere, and even shortens it. For a balance of skill between the atmosphere and ocean, data assimilation will be more skillful for

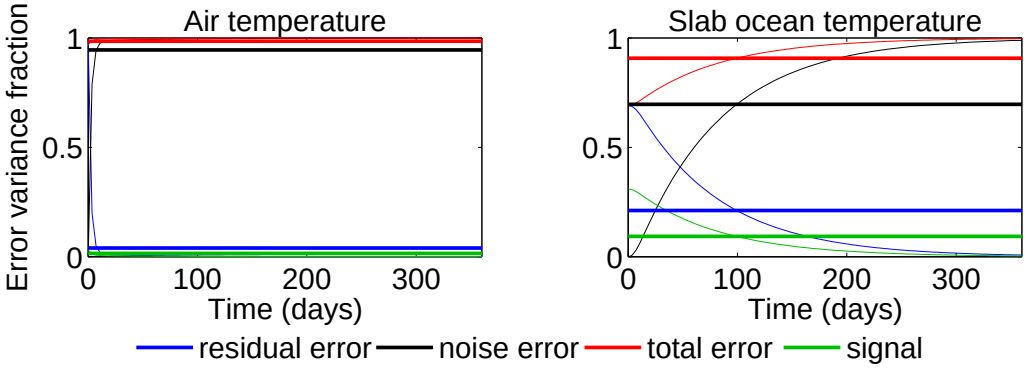

Figure 3.20: **Skill and error, BB parameters, observations with error.** Signal and noise in the BB system with observational error variance equal to climatological error variance.

shallower oceans (subject to the energy balance approximation used here), though of course we can't choose the depth of the real ocean.

What about variations in coupling? The coupling coefficient from BB98 is a function of both the wind speed from the air-sea flux exchange and the basic state temperature about which the anomalies are linearized. The linearized wind speed would certainly change with location and model. BB98 estimated a value by aggregating all locations in their model. It is definitely important to bear in mind how its variation would change the results.

As discussed above, increases in coupling correspond to decreases in skill in both the atmosphere and the ocean. Higher values of coupling coefficient (lower values of coupling timescale) will decrease the skill, while decreasing coupling will increase the skill.

3.3 Generalizability: QG model

Now that we have developed an understanding of the BB system, we turn to the QG system to apply the lessons just learned to a model with different forcing in a practical, rather than idealized, data assimilation scheme.

3.3.1 Comparison of experiments

The stochastic forcing in the BB model was designed to understand climate variability in a GCM, which, like the QG model, has nonlinear deterministic forcing. Additionally, heat in the QG model and in GCMs can be laterally advected by fully nonlinear dynamics, whereas in the BB model it can only go up, down, or out. These are substantial increases in complexity, so while it is an impressive result that the BB emulates aspects of midlatitude variability, it should not be taken for granted that all results should generalize.

In addition to simplified forcing, the BB system allows us to make powerful simplifications in the data assimilation. The data assimilation scheme is able to provide the expected error over an infinite number of experiments with an infinite number of ensemble members, and suffers none of the practical implementation issues that ensemble data assimilation was designed to address. In contrast, data assimilation in the QG system is a standard implementation of one flavor of the EnKF. The QG system performs real data assimilation experiments for particular true states. The QG model is a more complex, and arguably a better physical representation of midlatitude dynamics, and so understanding the commonalities and differences between these two systems is important.

Recall the details of the QG data assimilation experiment, which were discussed in section 2.2. We integrate one model instance for 100 assimilation cycles (all averaging times start from the same initial state) and take consecutive time-averages of the model states to be our series of observations. These observations are taken at every other gridpoint in the zonal and meridional directions, and are assigned an error of 10% of the climatological error variance for the averaging time. Then, an ensemble of 48 members is integrated without any data assimilation for 100 assimilation cycles. The error of the ensemble mean is determined by taking its difference from the "true" perfect integration, which is then squared and averaged over the center of the domain

to find the mean-squared error. Finally, the initial ensemble is integrated and the observations are assimilated for 100 assimilation cycles. The metric for skill is gain, discussed in section 2.2.

Figures 3.21 and 3.22 show the skill results for the QG system varying model parameters, and are the equivalent to Figure 3.7 and 3.8 where the BB system was used.

Figure 3.21 shows how skill varies with averaging time and slab ocean depth. Figure 3.21 looks a lot like figure 3.7; skill is generally greater for the ocean than for the atmosphere, and skill increases with averaging time. For increasing depth skill increases in the ocean, and decreases in the atmosphere.

From these similarities we conclude several things. First, in agreement with BB98, and with an extension to time-averaged data assimilation, fast weather variations in the QG system can be emulated well by stochastic forcing. Second, like the BB system, noise accumulated faster in the atmosphere than in the ocean, and the larger the ocean inertia the better the ocean skill relative to the atmosphere. As for BB system, skill is generally greater for the ocean than for the atmosphere, and skill increases with averaging time. Applying the lesson from the BB model here, it seems that the weather variations are not communicated to the ocean as effectively and instead accumulate in the atmosphere.

In contrast, skill as a function of coupling strength in the QG system is different from the BB system. Figure 3.22 shows how skill varies with averaging time and coupling. Note that we use d=4 for this figure because a deeper slab ocean would not respond on the timescales used in this data assimilation experiment.

Some results are similar: skill is still generally larger in the ocean than in the atmosphere, skill decreases with averaging time, and skill decreases with coupling coefficient in the ocean. However, in contrast with the BB system, skill in the atmosphere actually increases with coupling coefficient. In the BB model, we diagnosed that the reason atmospheric skill decreased as coupling increased was because noise

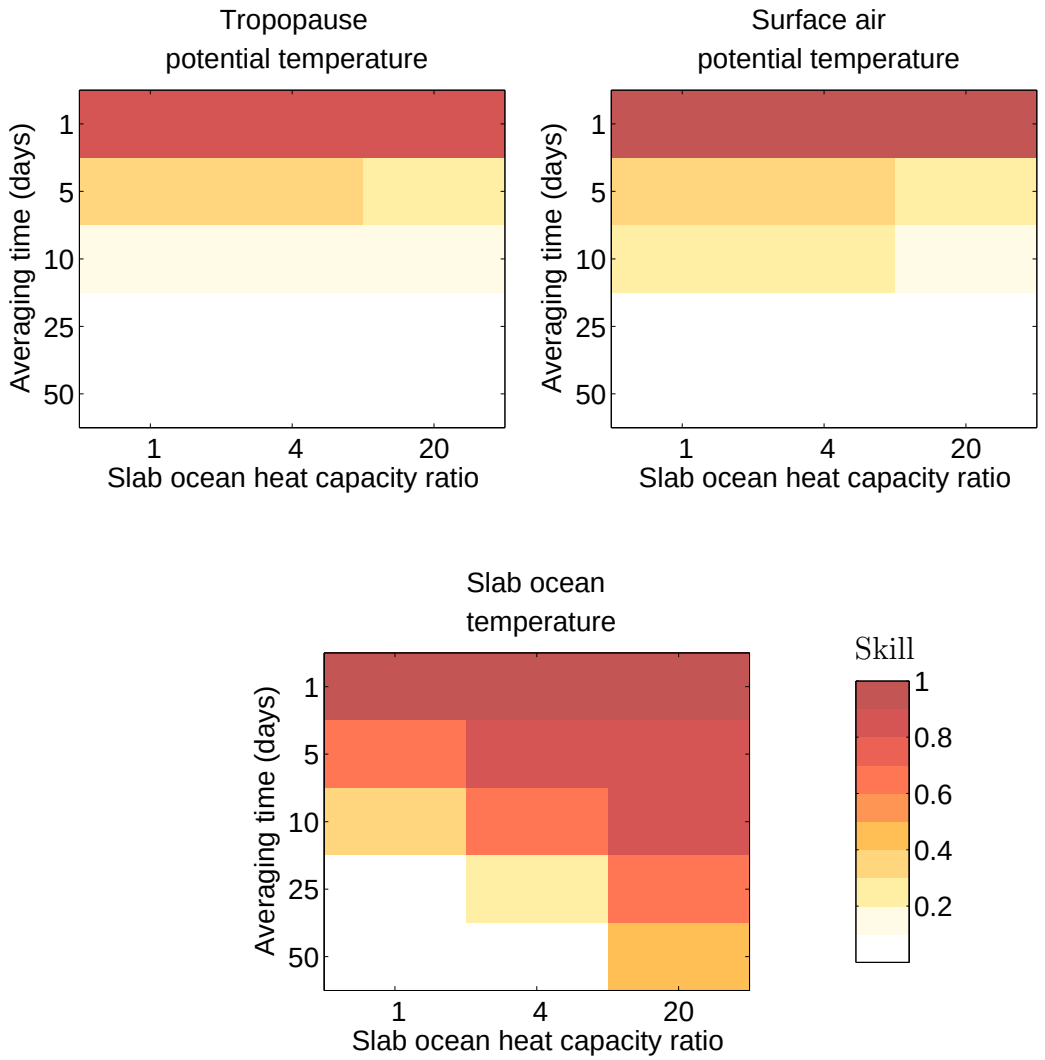

Figure 3.21: **QG system background skill – slab ocean depth and averaging time.** Each pixel is the skill of the background time-averaged state for a data assimilation experiment of 100 assimilation cycles, area-averaged over the meridional center of the domain (the region of basic state meridional temperature gradient). A score of 0 indicates that the error is the same as that of a control run with no assimilation, and score of 1 means there is no error. Coupling timescale τ_c=4.7 days.

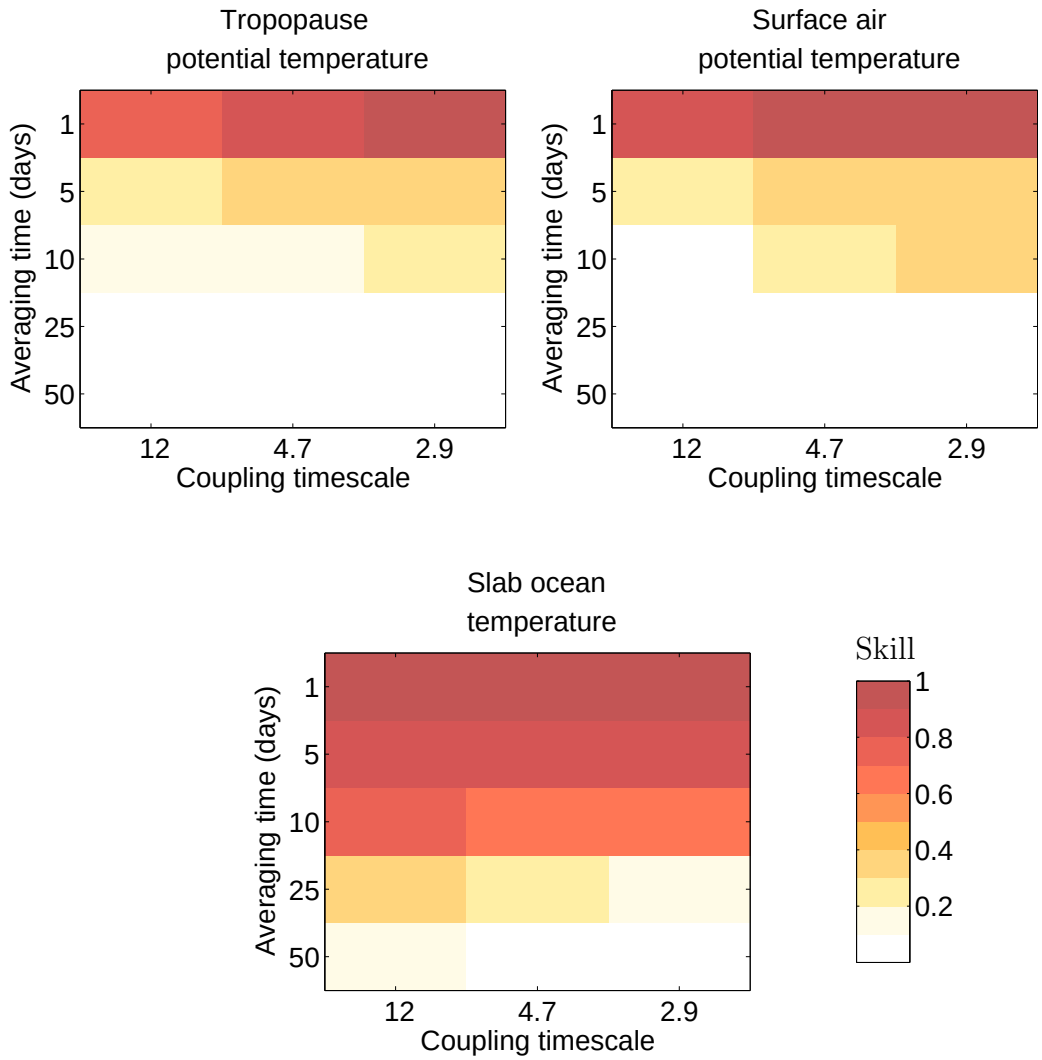

Figure 3.22: **QG system background skill – coupling and averaging time.** As in figure 3.21, except coupling coefficient varies. Slab ocean depth is d=4.

accumulated faster. Here the opposite appears to be true. The reasons for this difference are unclear at the time of writing. It may be that the atmosphere responds to changes in coupling in a way that is not captured here, for example, through changes in advection. We could calculate the advection of heat in the QG model and see how it changes with coupling, but this has not been done here.

3.3.2 Lessons from comparing the QG model and the BB model

We have seen that the BB system was useful for understanding variations in the QG system for variations in time-averaging and depth, but not for coupling. Based on this experience, what lessons can we expect to generalize from the linear stochastic BB system to nonlinear deterministic systems, and which do not seem to be applicable? How might we expect assimilation with an atmospheric GCM coupled to a slab ocean to differ from the QG model?

First, skill for the ocean exceeds skill for the atmosphere. This is fundamentally due to the atmosphere being subject to high-frequency forcing (either stochastic, or arising out of nonlinear chaotic synoptic dynamics). This forcing is only secondarily felt by the ocean as a result of air-sea fluxes. The result is that noise accumulates faster in the atmosphere than in the ocean. As this and other studies have shown, this conceptual picture appears to be very effective in characterizing the real midlatitude climate dynamics, and so we can have some confidence in its applicability to the real system.

Second, we found that both systems predicted a decrease of atmospheric skill, and an increase of ocean skill, with increasing mixed layer depth. This is due to increased accumulation of noise. However, it is possible that this would be different for models where the atmosphere adjusts baroclinically to SST anomalies.

Third, the two systems did not agree about how coupling affected skill. We clearly can no longer have confidence in the assertion that skill in the atmosphere should decrease with coupling in the real system. This does not happen in the QG system.

Additionally, the formulation of the slab ocean part of the QG model is only applicable to timescales longer than baroclinic adjustment of the atmosphere to an SST anomaly, which takes about a month in GCMs. On the basis of the QG and BB systems, it is not possible to say how skill would vary with coupling for a GCM coupled to a slab ocean.

Finally, other aspects of ocean dynamics may come into play in the real ocean or on linger timescales. The absence of dynamical ocean memory due to say, gyre circulations, mixed layer re-entrainment (e.g., Deser et al. 2003), or meridional overturning, and the absence of internally-generated ocean noise might make a difference in the real system.

Chapter 4
DISCUSSION

4.1 Summary

In this thesis, we developed an idealized system to understand expected time-averaged data assimilation. This system used a linear stochastic model for air-sea interactions, the BB model. We performed experiments to explore when time-dependent data assimilation is useful, and how this varies with model parameters. We compared the BB system experiments to experiments with a time-averaged extension of a standard data assimilation system using a model with a two-surface quasi-geostrophic atmosphere coupled to a slab ocean, the QG system.

We defined a predictability timescale for the atmosphere and ocean based on the average predictability time. This predictability timescale was useful for understanding the behavior of the BB system.

We used the BB system to develop an understanding of time-averaged data assimilation for a system with two variables of differing timescales. The atmosphere loses state-dependent skill faster than the ocean because stochastic errors accumulate more quickly there. When averaging time and slab ocean depth vary, the skill of background states in the BB system vary consistently with those in the QG system. Skill decreases with averaging time in both the atmosphere and ocean. With increasing slab ocean depth, skill in the ocean increases, but skill in the atmosphere decreases. For changes in coupling, responses are not consistent between the models, and this remains to be understood.

We performed calculations with the BB system using the set of parameters describing midlatitude climate variability developed by Barsugli and Battisti (1998) ex-

periments to characterize the prospects for useful state-dependent data assimilation with annual frequency observations. The calculations indicate that, for the assumptions and parameter choices made here, the signal variance of time-averaged states in the atmosphere is less than 2% and in the ocean is 14% for perfect observations.

4.2 Further investigations with the BB system

The BB and QG system are powerful frameworks for exploring the potential of data assimilation. We have focused on understanding one critical factor–skill in background error variance. But this is only one small part of the overall challenge. So, how in principle, might data assimilation help paleoclimate reconstruction? And what are natural next steps to take?

Data assimilation allows model states to be blended with paleoclimate proxy records to develop reconstructions. Ideally, data assimilation could use information about the forcing an internal state of the climate system to reconstruct climate. Incorporating information due to forcing means that relationships between observed and unobserved variables change both due to prescribed forcings, for example greenhouse gases and boundary conditions. Incorporating information based on internal variability means allowing relationships between observed and unobserved variables to change depending on the estimate of the internal state of the system. In this sense, data assimilation is only useful for constraining internal variability when there is state-dependent skill.

This thesis has focused on the background error covariances to test when it is possible to constrain internal variability using data assimilation, for the case with observations of the entire model state and annual frequency observations.

Even when there is no state-dependent skill, climatological covariances or time-dependent forcing changes obtained from models could still be used in data assimilation. The analysis, rather than the background, would be the final reconstruction product. The BB system could be used to investigate analysis skill for the case of

climatological covariances.

Below are a few possible paths of investigation into the intersection of climatological and state-dependent skill, and how observation information is spread between fast and slow components of the BB system.

If we can observe one part of the state, how do observations influence the unobserved part? In reality, there are few observations in the ocean, and they are spatially limited (see Jones et al. 2009), in contrast to the experiments here where all state variables are observed. Most of the land-based paleoclimate proxies (e.g. trees and ice) primarily respond to atmospheric variables. For proxy records which average over longer than the atmospheric predictability time, these observations may not provide any state-dependent information. With very few observations in the ocean, it might be a formidable challenge to maintain state-dependent skill for paleoreconstructions. Because the covariance of the variables depends on their phase relationship (see section 2.1.3), this should be a key to understanding and characterizing how information (both state-dependent and otherwise) from observed variables can spread to others.

We have just seen that for annual frequency observations it is possible that there is background skill in the ocean but not in the atmosphere. In this case, ocean reconstructions should be better than if climatological covariances had been used. But do atmospheric reconstructions improve in this case? This important question will be investigated by comparing analysis errors for the two types of background states.

4.3 Challenges for paleoclimate data assimilation

The work presented here exists within the broader context of moving towards an attempt at paleoclimate data assimilation with sophisticated models and actual paleoclimate observations. To this end, what has been done so far, and what challenges still need to be overcome?

- The first step to adapting weather data assimilation to the paleoclimate problem was enabling the use of observations that occur over a period of time. Dirren and Hakim (2005) developed and tested the method for assimilating time-averaged observations.

- Huntley and Hakim (2009, submitted) used the time-averaged algorithm in an atmosphere model and found it to have state-dependent skill. They showed that just a few observations, when taken in the right places, can be used to reconstruct time-averaged model states.

- This thesis addressed the conditions for which state-dependent time-averaged data assimilation is an improvement over climatological background states. We found that there is state-dependent skill in the ocean, but not in the atmosphere, for annual observations.

- A clear relationship between the proxy observations and atmospheric state variables, in other words a good forward model, is critical. Our understanding of these relationships is constantly improving, though there is also much room for progress. Hughes and Ammann (2009) emphasizes their importance in combining paleo proxies with models.

- Bias is often present in both climate models and paleoclimate proxies, and it can be difficult to quantify. Data assimilation theory is based on unbiased observations. Methods have been developed for dealing with bias (see, for example Evensen 2007), but with so few degrees of freedom from observations, this could impede the potential skill of reconstructions from data assimilation.

- Paleoclimate records almost always contain dating uncertainty. Of course this is an issue for all paleoclimate reconstruction techniques, but as of yet, there aren't any studies into the effects of this uncertainty in the context of data assimilation.

The data assimilation framework presented here readily affords the opportunity to do such a study. Two possibilities include increasing the observational error covariance going back in time and perturbing the dates of different ensemble members. Addressing this will be an essential part of establishing with what skill paleoclimate reconstructions can be performed using state-dependent data assimilation.

BIBLIOGRAPHY

Alexander, M. and C. Deser, 1995: A mechanism for the recurrence of wintertime midlatitude SST anomalies. *J. Phys. Oceanogr.*, **25 (1)**, 122–137.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The data assimilation research testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90 (9)**, 1283–1296.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129 (12)**, 2884–2903.

Barsugli, J. J., 1995: Idealized models of intrinsic midlatitude atmosphere–ocean interaction. Ph.D. thesis, University Of Washington.

Barsugli, J. J. and D. S. Battisti, 1998: The basic effects of atmosphere–ocean thermal coupling on midlatitude variability. *J. Atmos. Sci.*, **55 (4)**, 477 – 493.

Bretherton, C. S. and D. S. Battisti, 2000: Interpretation of the results from atmospheric general circulation models forced by the time history of the observed sea surface temperature distribution. *Geophys. Res. Lett.*, **27 (6)**, 767 – 770.

Delsole, T. and M. K. Tippett, 2009: Average predictability time. Part I: Theory. *J. Atmos. Sci.*, **66 (5)**, 1172–1187.

Deser, C., M. A. Alexander, and M. S. Timlin, 2003: Understanding the persistence of sea surface temperature anomalies in midlatitudes. *J. Climate*, **16 (1)**, 57–72.

Dirren, S. and G. J. Hakim, 2005: Toward the assimilation of time-averaged observations. *Geophys. Res. Lett.*, **32 (4)**.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99 (C5)**, 10 143–10 162, doi:10.1029/94Jc00572.

Evensen, G., 2007: *Data Assimilation: The Ensemble Kalman Filter*. Springer Verlag.

Ferreira, D., C. Frankignoul, and J. Marshall, 2001: Coupled ocean–atmosphere dynamics in a simple midlatitude climate model. *J. Climate*, **14 (17)**, 3704–3723.

Frankignoul, C., 1985: Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes. *Rev. Geophys.*, **23 (4)**, Pages 357 – 390.

Frankignoul, C. and K. Hasselmann, 1977: Stochastic climate models. Part II: Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, **29 (4)**, 289 – 305.

Gardiner, C. W., 2004: *Handbook Of Stochastic Methods: For Physics, Chemistry And The Natural Sciences*. 3d ed., Springer.

Goosse, H., H. Renssen, A. Timmermann, R. Bradley, and M. Mann, 2006: Using paleoclimate proxy-data to select optimal realisations in an ensemble of simulations of the climate of the past millennium. *Climate Dyn.*, **27 (2)**, 165–184, doi:10.1007/S00382-006-0128-6.

Graham, N., et al., 2007: Tropical Pacific – mid-latitude teleconnections in medieval times. *Climatic Change*, **83 (1)**, 241–285, doi:10.1007/S10584-007-9239-2.

Hakim, G. J., 2000: Role of nonmodal growth and nonlinearity in cyclogenesis initial-value problems. *J. Atmos. Sci.*, **57 (17)**, 2951 – 2967.

Hakim, G. J., C. Snyder, and D. J. Muraki, 2002: A new surface model for cyclone – anticyclone asymmetry. *J. Atmos. Sci.*, **59 (16)**, 2405–2420.

Haslett, J., S. Bhattacharva, M. Salter-Townshend, S. P. Wilson, J. Allen, B. Huntley, and F. Mitchell, 2006: Bayesian palaeoclimate reconstruction. *Journal Of The Royal Statistical Society. Series A. Statistics In Society*, **169**, 395 – 438.

Hasselmann, K., 1976: Stochastic climate models. Part I: Theory. *Tellus*, **28**, 473 – 485.

Hays, J. D., J. Imbrie, and N. J. Shackleton, 1976: Variations in the earth's orbit: Pacemaker of the ice ages. *Science*, **194 (4270)**, 1121 – 1132.

Hughes, M. and C. Ammann, 2009: The future of the past — an earth system framework for high resolution paleoclimatology: Editorial essay. *Climatic Change*, **94 (3)**, 247–259, doi:10.1007/S10584-009-9588-0.

Huntley, H. S. and G. J. Hakim, 2009: Assimilation of time-averaged observations in a quasi-geostrophic atmospheric jet model. *Climate Dyn.*, Submitted.

Jones, J. and M. Widmann, 2004: Reconstructing large-scale variability from palaeoclimatic evidence by means of data assimilation through upscaling and nudging (DATUN). *The KIHZ Project: Towards A Synthesis Of Holocene Proxy Data And Climate Models*, H. Fischer, T. Kumke, G. Lohmann, G. Flösser, H. Miller, H. von Storch, and J. Negandank, Eds., Springer, 171–193.

Jones, P. D., et al., 2009: High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects. *Holocene*, **19 (1)**, 3.

Kalman, R. E., 1960: A new approach to linear filtering and prediction theory. *Trans. ASME J. Basic Eng.*, **82**, 35 – 45.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation And Predictability.* Cambridge University Press.

Kushnir, Y., W. A. Robinson, I. B. E, N. M. J. Hall, S. Peng, and R. Sutton, 2002: Atmospheric GCM response to extratropical SST anomalies: Synthesis and evaluation. *J. Climate*, **15 (16)**, 2233–2256.

Lemke, P., 1977: Stochastic climate models, part 3. application to zonally averaged energy models. *Tellus*, **29 (5)**, 385–392.

Mahajan, R., 2007: Spatial growth of perturbations in turbulent baroclinic flows. Ms, University Of Washington, 32 pp.

Mann, M. E., R. S. Bradley, and M. K. Hughes, 1998: Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, **392 (6678)**, 779 – 787.

Mann, M. E., R. S. Bradley, and M. K. Hughes, 1999: Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.*, **26 (6)**, 759 – 762.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multi-decadal experiments. *Climate Dyn.*, **20 (2)**, 175 – 191.

North, G. R. and R. F. Cahalan, 1981: Predictability in a solvable stochastic climate model. *J. Atmos. Sci.*, **38 (3)**, 504–513.

Øksendal, B. K., 2003: *Stochastic Differential Equations: An Introduction With Applications*. Sixth ed., Springer.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics*. 2d ed., Springer-Verlag.

Penland, C. E. C., 2003: A stochastic approach to nonlinear dynamics: A review (extended version of the article - "Noise out of chaos and why it won't go away"). *Bull. Amer. Meteor. Soc.*, **84 (7)**, 925–925.

Saravanan, R. and J. C. McWilliams, 1998: Advective ocean – atmosphere interaction: An analytical stochastic model with implications for decadal variability. *J. Climate*, **11 (2)**, 165–188.

Schopf, P. S., 1985: Modeling tropical sea-surface temperature: Implications of various atmospheric responses. *Coupled ocean–atmosphere Models*, J. C. J. Nihoul, Ed., Elsevier Oceanography Series, Vol. 40, 727 – 734.

Scott, R. B., 2003: Predictability of SST in an idealized, one-dimensional, coupled atmosphere – ocean climate model with stochastic forcing and advection. *J. Climate*, **16 (2)**, 323–335.

Scott, R. B. and B. Qiu, 2003: Predictability of SST in a stochastic climate model and its application to the Kuroshio extension region. *J. Climate*, **16 (2)**, 312–322.

Sorenson, H. W., 1970: Least-squares estimation: From Gauss to Kalman. *Ieee Spectrum*, **7 (7)**, 63 – 68.

Sura, P., M. Newman, and M. A. Alexander, 2006: Daily to decadal sea surface temperature variability driven by state-dependent stochastic heat fluxes. *J. Phys. Oceanogr.*, **36 (10)**, 1940–1958.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131 (7)**, 1485–1490.

van der Schrier, G. and J. Barkmeijer, 2005: Bjerknes 'hypothesis on the coldness during ad 1790 – 1820 revisited. *Climate Dyn.*, **25 (5)**, 537–553, doi: 10.1007/S00382-005-0053-0.

von Storch, H., J. von Storch, and P. Müller, 2000: Combining paleoclimatic evidence and GCMs by means of data assimilation through upscaling and nudging (DATUN). *Proc 11Th Symposium On Global Change Studies*, Long Beach, Ca.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130 (7)**, 1913–1924.