UNIVERSITY OF READING

School of Mathematical and Physical Sciences

Department of Meteorology

Coupled Atmosphere-Ocean Data Assimilation

Duygu Ular

A dissertation submitted in partial fulfilment of the requirement for the degree of Master of Science in Data Assimilation and Inverse Modelling in Geosciences

August 2014

Abstract

Data assimilation is performed to combine observational data into a numerical model to find an estimate of the present state variables of a dynamical system. In coupled data assimilation more than one subsystem is assimilated into a model at the same time.

A significant difficulty of coupled atmosphere-ocean data assimilation is that the atmosphere and ocean have different time and spatial scales. This dissertation investigates the behaviour of coupled atmosphere-ocean data assimilation when one subsystem is not observed. There are lots of factors which affect the assimilation. We consider only the effects of the number of observations and the observation accuracy.

We study four dimensional data assimilation because it is applicable to very large systems and it reconstructs the states in the unobserved region. The Molteni et al. (1993) model is used to describe the coupled atmosphere-ocean system.

The singular value decomposition is an useful technique to extract the information about the observing system. It is used to understand effects of the observations on the assimilation.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Duygu ULAR

Acknowledgements

I would like to thank both my supervisors Prof. Nancy Nichols and Dr. Amos Lawless for their guidance, help and encouragement during this dissertation. I would also like to thank in special my family and my husband for the support throughout the year.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Data Assimilation

Data assimilation (DA) is performed to combine observational data into a numerical model to find an estimate of the present state variables of a dynamical system (Lawless, 2013). One of the application areas of DA is numerical weather prediction (NWP). It is used to generate an accurate weather forecast which is possible by using accurate initial conditions in a numerical model.

The DA algorithm must be able to assimilate observations which are taken by satellites, weather station etc. These observations have errors so the algorithm must filter the noise and interpolate between observations. For unobserved regions, the algorithm must use a previous weather forecast and the model dynamics to get

accurate data.

One of the significant difficulties of the DA is that the size of the models and the number of observations are very large, of order $\mathbf{O}(10^7 - 10^8)$, $\mathbf{O}(10^5 - 10^6)$ respectively (Nichols, 2010). Therefore to get an accurate estimate we must set up the algorithm carefully.

There are two main categories of DA, which are sequential (e.g. Kalman Filter) and variational DA (e.g. 4D-Var). Sequential DA updates the estimate of the state at each observation time. The main implementing difficulty of the sequential DA is that the covariance matrices have very large dimension so it could be impractical. Therefore most of the sequential DA models try to get rid of this difficulty by using different methods.

The variational DA finds the analysis at the initial time by minimizing a cost function. A significant issue of using variational DA is the need to develop an adjoint model of the linearized state equations (Nichols, 2010). Three dimensional variational DA (3D-Var) uses observations given at only a single point in time. This means there is no model evolution. Four dimensional variational (4D-Var) DA gets a best fit model trajectory through the observations to find the initial state.

In this dissertation we study 4D-Var DA due to it being applicable to very large systems and there are currently not enough studied about 4D-Var in coupled

atmosphere-ocean DA.

1.2 Coupled Data Assimilation

Coupled DA is a new field of DA which aims to combine the observational information covering the atmosphere and ocean into a coupled model (Sugiura et al., 2008).

Han et al.(2013) give a definition of coupled DA as follows.

"When the observed data in one or more media are assimilated into the model,information is exchanged among different media in the coupled system. Such an assimilation procedure is called coupled data assimilation"

In our case the atmosphere and ocean are used as different medias and their information is assimilated into the 4D-Var coupled system. The significant challenge of coupled ocean-atmosphere DA is that the atmosphere and ocean systems have different time and spatial scales.

There are two different types of coupled DA which involve synchronous and asynchronous coupling. In synchronous coupled DA, each subsystem is fed each output of the other one for each time step. In asynchronous coupled DA, one of the subsystems uses averaged output of the other one. Dubois and Yiou (1998) have a study which compares synchronous and asynchronous coupled modelling.

They indicate that asynchronous coupling is often used because of the computing limitations. However synchronous coupling gives a more realistic result. Therefore in this study we use the synchronous coupled atmosphere-ocean DA.

1.3 Key questions addressed

The aim of the dissertation is to understand the effects of the observations on the 4D-Var coupled atmosphere-ocean DA when we observe only one subsystem.

The key questions are:

- When do we get a better analysis result; when we observe only the atmosphere or the ocean?

  We know that 4D-Var is able to reconstruct the state in unobserved regions (Courtier and Talagrand, 1987). We try to apply this reconstruction feature of the 4D-Var to our coupled atmosphere-ocean DA model because the atmosphere and ocean have different spatial and time scales so observing both subsystems could be difficult. We wish to get a good analysis result when we observe only one subsystem. This is the our main key question.

- How is the assimilation affected by the observation accuracy?

  When we observe a system we do not get perfect observations. All of them

have random noise. These noises effect the assimilation significantly. Our aim is to understand these effects.

- How is the assimilation affected by the number of observations?

  As we say in the first key question, it is impossible to have observations from the whole system. Therefore we would like to know the behaviour of the system with different numbers of observations.

1.4 Outline

To address the key questions we focus on nonlinear 4D-Var experiments with the Molteni et al. (1993) model.

Chapter 2 gives background information about nonlinear 4D-Var. We focus on the cost function, adjoint method and error covariance matrices of the 4D-Var separately.

Chapter 3 explains the Molteni Model which describes the atmosphere-ocean coupled system.

Chapter 4 gives the assimilation experiments. We try to find the answers to the key questions

Chapter 5 describes the singular value decomposition (SVD) to understand observation effects. Firstly we give the algorithm of the SVD and then we apply it to the 4D-Var. We show that the 4D-Var analysis can be written as a linear combination of the right singular vectors of the 4D-Var observability matrix.

Chapter 6 gives SVD experiments which provides an understanding how information comes from observations. We try to explain the reasons of the assimilation experiment results.

Chapter 7 makes a conclusion of the thesis with a future work.

Chapter 2

4D-Var

4D-Var was introduced by Le Dimet and Talagrand (1986). It is a method of estimating a set of parameters by running the model and finding the best fit between the model and observations (Bannister, 2001). We need to find initial conditions which is possible by minimizing a cost function. In this chapter we describe the cost function and the adjoint method which is used to find the gradient of the cost function. Then we give a brief description of the error covariance matrices which are an important part of 4D-Var.

2.1 The Cost Function

The cost function ,$\mathcal{J}$, measures the weighted sum of squares of the distance to the background state and to the observations over the time window.

We first consider the nonlinear dynamical system,

$$\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k), \quad k = 0, 1, ..., N-1 \tag{2.1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ represent the state vector at time t_k and $\mathcal{M}$ is the nonlinear model operator. For NWP the vector $\mathbf{x}_k$ would have some variables like temperature, pressure, etc (Lawless, 2013).

We assume we have observations $\mathbf{y}_k \in \mathbb{R}^{p_k}$ at time t_k, $k = 0, 1, ..., N$, which is written as

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \boldsymbol{\epsilon}_k \tag{2.2}$$

where $\mathcal{H}_k$ is an observation operator and $\boldsymbol{\epsilon}_k \in \mathbb{R}^{p_k}$ are observation errors with covariance matrix $\mathbf{R}_k$, which represents uncertainty in the observations.

We use the previous model forecast as a priori information, which is called the background, $\mathbf{x}^b$ and is written as

$$\mathbf{x}^b = \mathbf{x}_0 + \boldsymbol{\varepsilon} \tag{2.3}$$

where $\mathbf{x}_0$ is the initial state and $\boldsymbol{\varepsilon} \in \mathbb{R}$ is the background error with covariance matrix $\mathbf{B}$, which represents uncertainty in the background state.

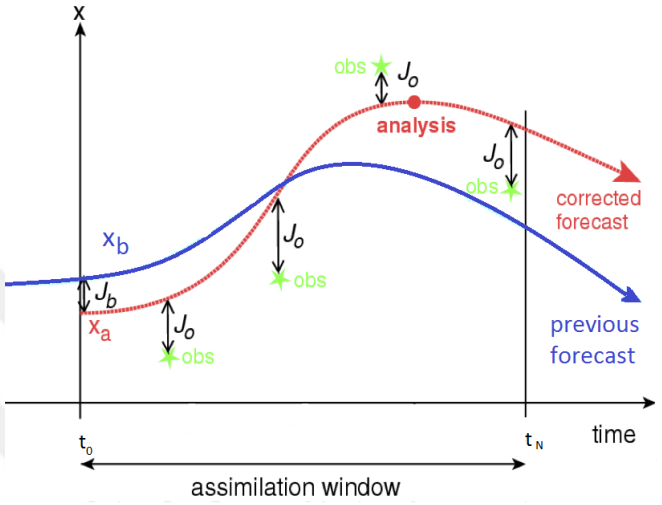

Figure 2.1: Process of the 4D-Var:minimize the squared distance between the analysis and the background state (J_b) at the beginning of the assimilation window and squared distance between the observations and the forecast state (J_o) throughout the assimilation window (Bouttier and Courtier, 1999).

We can now write the usual form of the 4D-Var cost function as

$$\begin{aligned} \mathcal{J}(\mathbf{x}_0) &= \tfrac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \tfrac{1}{2}\sum_{k=0}^{N}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k)^T \mathbf{R}_k^{-1}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k) \\ &= \mathbf{J}_B + \mathbf{J}_o \end{aligned} \tag{2.4}$$

subject to the states $\mathbf{x}_k$ satisfying the nonlinear dynamical system (2.1). The anal-

ysis state $\mathbf{x}_0^a$ at the initial time can be found by minimizing this function.

When we write this cost function, we make two more assumptions. The first one is that the background and observations errors are unbiased and have a Gaussian distributions. The other one is that the model is perfect. Figure 2.1 shows the process of 4D-Var (Bouttier and Courtier, 1999).

To minimize the cost function we consider the method of Lagrange.

2.2 The Adjoint method

To minimize the cost function we use an iterative numerical method such as quasi-Newton or conjugate gradient. These methods use the value and gradient of the cost function at each iteration, so we introduce the Euler-Lagrange equations to calculate the gradient of the 4D-Var.

We define Lagrange multipliers $\boldsymbol{\lambda}_k$ and a Lagrangian $\mathcal{L}$ as follows

$$\mathcal{L} = \mathcal{J}(\mathbf{x}_0) + \sum_{k=0}^{N-1} \boldsymbol{\lambda}_{k+1}^T (\mathbf{x}_{k+1} - \mathcal{M}_k(\mathbf{x}_k)) \tag{2.5}$$

The necessary conditions for a minimum of the cost function are

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}_k} = 0 \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \mathbf{x}_k} = 0$$

The first condition gives us the original model equation and the second one gives us the adjoint equations

$$\boldsymbol{\lambda}_k = \mathbf{M}_k^T \boldsymbol{\lambda}_{k+1} - \mathbf{H}_k^T \mathbf{R}_k^{-1}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k) \quad k = N, N-1, ..., 1 \tag{2.6}$$

with boundary condition $\boldsymbol{\lambda}_{N+1} = 0$ where $\mathbf{M}_k$ is the Jacobian of $\mathcal{M}_k$, which is known as the tangent linear model (TLM) and $\mathbf{H}_k$ is the Jacobian of $\mathcal{H}_k$, which is known as the tangent linear operator with respect to $\mathbf{x_k}$. The adjoint variables $\boldsymbol{\lambda}_k$ give us the measurement of the sensitivity of the cost function to changes in the solutions $\mathbf{x_k}$ of the state equations for each value of k (Nichols, 2010). We can now write the gradient of the cost function with respect to the initial state $\mathbf{x_0}$,

$$\nabla \mathcal{J}(\mathbf{x}_0) = -\boldsymbol{\lambda}_0 + \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) \tag{2.7}$$

This gradient must be equal to zero at the optimum. The numerical optimization methods must run both the model equations for the cost function and the adjoint model for the gradient. Therefore 4D-Var is a very expensive algorithm but it is applicable to very large systems (Lawless, 2013).

2.3 The Error covariance matrices

The error covariance matrices **R** and **B** play a crucial role in the DA process because all models are wrong and all observations are inaccurate. They have essential statistical information about the background field and observations (Bannister, 2001). The diagonal of the matrices **R** and **B** contain the variance of the observation error and the variance of the background error for each state variable, respectively. The off-diagonal terms of **B** are error covariances between different components of the background state vector. Assume an n dimensional state vector with observation errors $(\epsilon_1, \epsilon_2, ..., \epsilon_{p_k})$ and background errors $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ then

$$\mathbf{R}_{ij} = cov(\epsilon_i, \epsilon_j) \quad i = 1, 2, ..., p_k \text{ and } j = 1, 2, ..., p_k \tag{2.8}$$

$$\mathbf{B}_{ij} = cov(\varepsilon_i, \varepsilon_j) \quad i = 1, 2, ..., n \text{ and } j = 1, 2, ..., n \tag{2.9}$$

The background field is used, since it is a forecast from a previous assimilation cycle so it has information about the previously assimilated observations (Lawless, 2013).

Chapter 3

The Model

In this section we introduce the Molteni et al. (1993) model which describes the coupled atmosphere-ocean system.

3.1 A simple atmosphere-ocean model

Molteni et al. (1993) used the Lorenz (1963) system to develop a coupled atmosphere-ocean model. The coupled model is represented by the following ordinary differential equations.

$$
\begin{aligned}
\dot{X} &= -\sigma X + \sigma Y + \alpha V \\
\dot{Y} &= -XZ + rX - Y + \alpha W \\
\dot{Z} &= XY - bZ \\
\dot{W} &= -\Omega V - k(W - W^{\star}) - \alpha Y \\
\dot{V} &= \Omega(W - W^{\star}) - kV - \alpha X
\end{aligned}
\tag{3.1}
$$

where $\sigma = 10$, $r = 28$, $b = 8/3$, $\alpha = 1$, $k = 0.1$, $\Omega = 1.5$ and $W^{\star} = 2$. Here σ, r, b are the Lorenz parameters and the system exhibits chaotic behaviour for these values. The parameter r is called the Rayleigh number, which indicates whether the heat transfer is primarily in the form of conduction or convection. The parameter σ is called the Prandtl number, which is the ratio of momentum diffusivity and thermal diffusivity. The parameter b is a geometric factor. Lorenz (1963) gives an explanation about the variables as follows.

"X is proportional to the intensity of the convective motion while Y is proportional to the temperature difference between the ascending and descending currents. The variable Z is proportional to the distortion of the vertical temperature profile from linearity".

Molteni et al. (1993) explains the other parameters.

The first three nonlinear equations represent the atmosphere, which has a chaotic behaviour and the last two linear equations represent the slow ocean.

3.2 Second order Runge-Kutta Method

For 4D-Var DA we need a numerical method to approximate the solution of the model. In this study we use the second order Runge-Kutta technique to discretize the Molteni et al. (1993) model.

The second order Runge-Kutta method is used to solve an ordinary differential equation of the form

$$\frac{dy}{dt} = f(t, y(t)) \ , \qquad y(0) = y_0 \tag{3.2}$$

where y is an unknown function at time t which we would like to approximate. This method for y_{n+1} is given by

$$\begin{aligned} k_1 &= hf(t_n, y_n) \\ k_2 &= hf(t_n + \tfrac{h}{2}, y_n + \tfrac{k_1}{2}) \\ y_{n+1} &\simeq y_n + k_2 \end{aligned} \tag{3.3}$$

where $h = t_{n+1} - t_n$ is step size and n represents the time step.

The second order Runge-Kutta method is a one step method. The benefit of using a one step method is that it does not store past variables. This means that

the variable at time $n+1$ just depends on the variable at time n.

Chapter 4

Assimilation Experiments

Chapter 2 and 3 give the background information about 4D-Var and the Molteni et al. (1993) model. Now we set up our 4D-Var atmosphere-ocean coupled DA algorithm using this background information to use in our numerical experiments.

For all experiments we fix the time window which has 250 time steps with time length 0.04. The first 150 time steps represent the assimilation window and the other 100 time steps represent the forecast window. We also fix the coupling frequency as 1, which means that we couple the system at each time step.

We generate the truth state by integrating the coupled system using the second order Runge-Kutta method, which is introduced in section 3.2, with a time step of $h = 0.04$. This method uses the following initial conditions; $x(0) = 1.0$, $y(0) = 1.0$,

$z(0) = 1.0$, $w(0) = 1.1$, $v(0) = 1.1$.

To generate the initial background state firstly we need to specify the background error covariance matrix $\mathbf{B}$, which is introduced in section 2.3. We assume that the background error of each state variable has the same variance which is $\sigma_b^2 = 10^{-2}$ and that there is no correlation between the errors. So our diagonal background error covariance matrix is $\mathbf{B} = \sigma_b^2 \mathbf{I}_5$. Now we can create the initial background state as follows

$$\mathbf{x}_0^b = \mathbf{x}_0 + \sigma_b \boldsymbol{\eta} \tag{4.1}$$

where $\boldsymbol{\eta}$ is a vector of random numbers from normal distribution $N(0, \mathbf{I}_5)$ and σ_b is the standard deviation of the background error. The background trajectory is created from the initial background state by using the second order Runge-Kutta method. To compare the results easily we use the same background field in all experiments.

Similarly, to generate the observations firstly we need to specify the observation error covariance matrix $\mathbf{R}$. We assume that the observation error of each state variable has the same variance and there is no correlation between the errors. So our diagonal observation error covariance matrix is $\mathbf{R} = \sigma_o^2 \mathbf{I}_5$. Now we can generate the observations as

$$\mathbf{y}_k = \mathbf{x}_k + \sigma_o \boldsymbol{\zeta}_k \tag{4.2}$$

where $\boldsymbol{\zeta}$ is a vector of random numbers from normal distribution $N(0, \mathbf{I}_5)$ and σ_o is the standard deviation of the observation error. We generate the observations at certain time steps in the assimilation window. For example, assume we set up the observation frequency= $[50, 50, 50, 50, 50]$. This generates observations at each 50 time steps for each state variable. We define the total time step as 150 in the assimilation window and have no observations at the initial time t_0. This gives us 15 observations in total.

We generate the analysis state at the initial time using the 4D-Var algorithm with this background and observation information. Then again using the second order Runge-Kutta method we produce the analysis trajectory. We can say that we have a good analysis state when the analysis result satisfies the following condition.

$$|| \mathbf{x}_0^a - \mathbf{x}_0 ||_2 < || \mathbf{x}_0^b - \mathbf{x}_0 ||_2$$

where the $\mathbf{x}^a$, $\mathbf{x}$ and $\mathbf{x}^b$ represent the analysis, truth and background states respectively at time t_0. We check the results of all the experiments by considering this condition.

We use the Polack-Ribiere flavour of conjugate gradient methods as an optimization method, which uses only the first derivative of the cost function. Gilbert and Nocedal (1992) give further information about this method. In this method we use

the stopping criteria as

$$\frac{|| \nabla \mathcal{J}^k(\mathbf{x}_0) ||_2}{|| \nabla \mathcal{J}^0(\mathbf{x}_0) ||_2} < 0.001 \tag{4.3}$$

where k represents the number of iterations.

In the following experiments we use perfect and imperfect observation expressions. A perfect observation indicates that the observation does not have error. This means that the observation is generated by the truth state at each time step. The imperfect observation indicates that observational noise is added to the truth state at each time step. When we use the imperfect observations in any experiment, we estimate analysis error by calculating the average of 10 different experiment results due to fact that each experiment has different random noise. Thus we get more realistic results.

In the experiments we compare the analysis errors of the atmosphere, ocean and whole system. We get these errors by calculating the 2-norms of the differences between the analysis and truth states. We also check the results when we normalise the errors by the number of elements to make a fair comparison.

4.1 Experiment 1

In this experiment we try to find an answer to our first key question which is "When do we get we a better analysis result; when we observe only the atmosphere or the ocean? "

Firstly we use perfect observations to get rid of the effect of the random noise. We compare the following two cases which have the same observation covariance matrix $\mathbf{R} = 10^{-4}\mathbf{I}_5$.

- Case 1: There are 30 observations in the atmosphere and there is no observation in the ocean.

$$\text{observation frequency} = [15, 15, 15, 0, 0]$$

- Case 2: There is no observation in the atmosphere and there are 30 observations in the ocean.

$$\text{observation frequency} = [0, 0, 0, 10, 10]$$

It is expected that the observed subsystem gives a better analysis result than the unobserved one.

obs.frequency	error-atm	error-ocean	error-system
(15,15,15,0,0)	0.033906	0.047593	0.058436
(0,0,0,10,10)	0.030208	0.000975	0.030223

Table 4.1: Comparing the error in the analyses using only (perfect) observations of the atmosphere and only (perfect) observations of the ocean when the observation error variance is $\sigma_o^2 = 10^{-4}$ and the background error variance is $\sigma_b^2 = 10^{-2}$.

The error in the analyses are shown in table 4.1 for both cases. As we expect, the error in the atmosphere is smaller than in the ocean in case 1 because we have no observation in the ocean and the error of the ocean is smaller than the atmosphere in case 2. When there is no observation in a subsystem, the analysis state of this subsystem is not close to the truth state so the difference between the analysis and the truth state becomes larger.

The significant points are that the system error in case 1 is higher than in case 2 so we could say that the algorithm gives a better estimate for case 2 than in case 1 and the error in the atmosphere in case 2 is smaller than case 1. According to these experiments, when there are only ocean observations the algorithm reconstructs the states in the unobserved region more successfully than when we have only atmosphere observations.

Up to now in this experiment we have used the perfect observations so we only understand the model behaviour. Now we study with imperfect observations in order to get more realistic results.

obs.frequency	error-atm	error-ocean	error-system
(15,15,15,0,0)	0.032199	0.041788	0.053801
(0,0,0,10,10)	0.030873	0.005984	0.031576

Table 4.2: Comparing the error in the analyses using only (imperfect) observations of the atmosphere and only (imperfect) observations of the ocean when the observation error variance is $\sigma_o^2 = 10^{-4}$ and the background error variance is $\sigma_b^2 = 10^{-2}$.

Table 4.2 shows case 1 and case 2 results with imperfect observations.

From table 4.2 it is clearly seen that the system has a better reconstruction in the unobserved region with the ocean observations than with the atmosphere observations. The results found for the imperfect observations errors are quite similar to the perfect ones in case 1 and case 2 because the observation error variance is 10^{-4} which is quite small, so the observations are very accurate.

As a consequence of experiment 1 we find the answer to the first key question: When we only observe the ocean with more accurate observations, the system gives a better analysis result than when we only observe the atmosphere.

In the next experiment we study with different observation variances to get an idea of the effect of the observation accuracy.

4.2 Experiment 2

In this experiment we try to find an answer to the second key question, which is "how is the assimilation affected by the observation accuracy?". To understand the effect of the observation accuracy on the assimilation we experiment with four different observation error variances which are $10^{-1}, 10^{-2}, 10^{-3}$ and 10^{-4} by considering case 1 and 2 in experiment 1 with imperfect observations.

We expect that we get a better analysis result when the observations are more accurate because the accurate observations have small noise so they are close to the truth states.

Table 4.3 shows the error in the analyses using imperfect observations for both cases. As we expect when we increase the observation error variance, all errors also increase because the system, which has higher observation error variance, has higher random noise. In all cases the unobserved region has larger error than the observed region and when we only observe the ocean we get smaller system error than in the case of only observed atmosphere. Therefore we can still say that when there are only ocean observations, the algorithm reconstructs the states in the unobserved region better than when we have only atmosphere observations.

We can say that the observation accuracy affects the analysis results significantly as an answer of the second key question. To get better estimate we should use

observation frequency:[15,15,15,0,0]				
σ_o^2	10^{-1}	10^{-2}	10^{-3}	10^{-4}
error-atm	0.069829	0.058021	0.040844	0.032199
error-ocean	0.204169	0.188111	0.119447	0.041788
error-system	0.216792	0.197050	0.127194	0.053801
observation frequency:[0,0,0,10,10]				
σ_o^2	10^{-1}	10^{-2}	10^{-3}	10^{-4}
error-atm	0.141991	0.088150	0.047043	0.030873
error-ocean	0.130411	0.082560	0.030017	0.005984
error-system	0.197313	0.121909	0.056673	0.031576

Table 4.3: Comparing the error in the analyses using only (imperfect) observations of the atmosphere and only (imperfect) observations of the ocean for different observation error variances and the background error variance is $\sigma_b^2 = 10^{-2}$.

relatively accurate observations.

Up to now in all the experiments we have 30 observations in total. The next experiment investigates the effect of the number of observations on the analysis by taking fewer observations.

4.3 Experiment 3

In this experiment we try to find an answer to the third key question, which is "how is the assimilation affected by the number of observations?".

We consider the second experiments with fewer observations to see how the number of observation affects the assimilation results.

We study with the following two cases.

- Case 3: There are 6 observations in the atmosphere and there is no observation in the ocean.

$$\text{observation frequency} = [75, 75, 75, 0, 0]$$

- Case 4: There is no observation in the atmosphere and there are 6 observations in the ocean.

$$\text{observation frequency} = [0, 0, 0, 50, 50]$$

The error in the analyses for the imperfect observations are shown in table 4.4 with a comparison of different observation variances. As in the result of experiment 2, when we increase the observation error variance from 10^{-4} to 10^{-1} all errors in the analysis also increase in both cases. Therefore we have better estimate with more accurate observations. In addition it is clearly seen that when we have less observation, the atmosphere could not reconstruct the unobserved ocean successfully for all variances. The significant point is that when we only observe the ocean the analysis error is not smaller than when we only observe the atmosphere in the case of $\sigma_o^2 = 10^{-1}$ so in this case we get better estimate with the atmosphere observations.

We can say that if we only observe the atmosphere with a smaller number of imperfect observations and higher observation error variances, the system gives a better estimate than in the case of we only observe the ocean. To explain the reasons

observation frequency:[75,75,75,0,0]				
σ_o^2	10^{-1}	10^{-2}	10^{-3}	10^{-4}
error-atm	0.062000	0.059554	0.058214	0.055346
error-ocean	0.212299	0.209983	0.209382	0.209092
error-system	0.221665	0.218837	0.216381	0.216298
observation frequency:[0,0,0,50,50]				
σ_o^2	10^{-1}	10^{-2}	10^{-3}	10^{-4}
error-atm	0.175125	0.122109	0.073270	0.043484
error-ocean	0.167505	0.117718	0.072131	0.034511
error-system	0.244595	0.172329	0.104259	0.056375

Table 4.4: Comparing the error in the analyses using only (imperfect) observations of the atmosphere and only (imperfect) observations of the ocean for different observation error variances and the background error variance is $\sigma_b^2 = 10^{-2}$.

for these three experiment results, in the next chapter we introduce the SVD which helps us to understand how the system uses the observation information. Then we try to understand the reasons by considering the same cases with the SVD.

Chapter 5

Singular Value Decomposition(SVD)

To get a better idea of the effect of the observations on the coupled atmosphere-ocean DA, we should consider how we get information from the observations. One of the techniques is singular value decomposition (SVD) which is useful in understanding the effects of the observations. This technique is introduced by Mateer (1965).

In this chapter we start by giving some information about the SVD. Then we apply the technique to 4D-Var.

5.1 SVD

The following description is given by Golub and Van Loan (1996). Assume $\mathbf{H}$ is an $m \times n$ ($m \geq n$) matrix with rank r. $\mathbf{H}$ can be written as the product of the three matrices

$$\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{5.1}$$

where

- $\mathbf{U}$ is an $m \times m$ orthogonal matrix, the m columns $\mathbf{u}_i$ of $\mathbf{U}$ are called the left singular vectors (LSVs), which form orthonormal bases so $\mathbf{U}^T\mathbf{U} = \mathbf{I}_m$ and they are eigenvectors of $\mathbf{H}\mathbf{H}^T$
- $\mathbf{V}$ is an $n \times n$ orthogonal matrix, the n columns $\mathbf{v}_i$ of $\mathbf{V}$ are called the right singular vectors (RSVs), which form orthonormal bases so $\mathbf{V}^T\mathbf{V} = \mathbf{I}_n$ and they are eigenvectors of $\mathbf{H}^T\mathbf{H}$
- $\mathbf{S}$ is an $m \times n$ diagonal matrix which has r positive singular values $\mathbf{s}_i$ which are ordered $s_1 \geq s_2 \geq ... \geq s_r > 0$

Golub and Van Loan (1996) prove that $\mathbf{H}$ can be written as

$$\mathbf{H} = \sum_{i=1}^{r} s_i \mathbf{u}_i \mathbf{v}_i^T \tag{5.2}$$

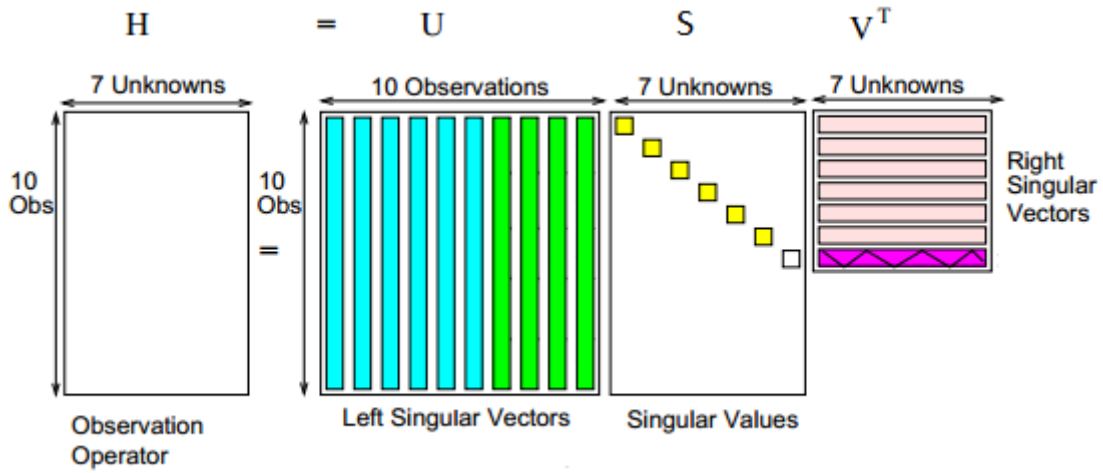

Figure 5.1: Graphic illustrating the SVD of the observation operator **H**. There are m=10 observations so the length of the LSVs is 10 and there are n=7 unknown variables so the length of the RSVs is 7 (Johnson, 2003).

Figure 5.1 gives a graphical description of the SVD of a matrix (Johnson,2003).

The SVD allows us to determine the important part of the observations. We now apply this useful technique to 4D-Var to get observation information from our coupled atmosphere-ocean DA algorithm.

5.2 Applying to 4D-Var

We introduced 4D-Var in chapter 2. This method is hard to use with the SVD technique because of the nonlinear $\mathcal{M}$ but it could be adapted and approximated using the tangent linear hypothesis. This provides a quadratic cost function. $\mathcal{M}$

can be linearized;

$$\mathcal{M}(t_k, \mathbf{x}_k + \delta\mathbf{x}_k) = \mathcal{M}(t_k, \mathbf{x}_k) + \mathbf{M}(t_k, \mathbf{x}_k)\delta\mathbf{x}_k + O(\delta\mathbf{x}_k^2) \tag{5.3}$$

The hypothesis neglects terms of order two

$$\mathcal{M}(t_k, \mathbf{x}_k + \delta\mathbf{x}_k) \simeq \mathcal{M}(t_k, \mathbf{x}_k) + \mathbf{M}(t_k, \mathbf{x}_k)\delta\mathbf{x}_k \tag{5.4}$$

where $\mathbf{M}$ is the tangent linear model.

Now we write the background error covariance matrix as $\mathbf{B} = \sigma_b^2\mathbf{I}$ and the observation error covariance matrix as $\mathbf{R} = \sigma_o^2\mathbf{I}$ into the cost function, where σ_b^2 defines the background error variance and σ_o^2 defines the observation error variance.

We assume the observation operator $\mathbf{H}$ is linear i.e $\mathcal{H}_k(\mathbf{x}) = \mathbf{H}\mathbf{x}$ so the cost function can be written as

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T\sigma_b^{-2}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{k=0}^{N}(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k)^T\sigma_o^{-2}(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k) \tag{5.5}$$

subject to the $\mathbf{x}_{k+1} = \mathbf{M}_k(\mathbf{x}_k)$

Writing it in unconstrained form we get,

$$
\begin{aligned}
\mathcal{J}(\mathbf{x}_0) &= \tfrac{1}{2}\sigma_b^{-2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T(\mathbf{x}_0 - \mathbf{x}_0^b) \\
&+ \tfrac{1}{2}\sigma_o^{-2}(\mathbf{H}_0\mathbf{x}_0 - \mathbf{y}_0)^T(\mathbf{H}_0\mathbf{x}_0 - \mathbf{y}_0) \\
&+ \tfrac{1}{2}\sigma_o^{-2}(\mathbf{H}_1\mathbf{M}_1\mathbf{x}_0 - \mathbf{y}_1)^T(\mathbf{H}_1\mathbf{M}_1\mathbf{x}_0 - \mathbf{y}_1) \\
&+ \ ... \\
&+ \tfrac{1}{2}\sigma_o^{-2}(\mathbf{H}_N\mathbf{M}_N\mathbf{x}_0 - \mathbf{y}_N)^T(\mathbf{H}_N\mathbf{M}_N\mathbf{x}_0 - \mathbf{y}_N)
\end{aligned} \tag{5.6}
$$

After rearranging we have,

$$
\begin{aligned}
\mathcal{J}(\mathbf{x}_0) &= \tfrac{1}{2}\sigma_b^{-2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T(\mathbf{x}_0 - \mathbf{x}_0^b) \\
&+ \tfrac{1}{2}\sigma_o^{-2}\begin{pmatrix} (\mathbf{H}_0\mathbf{x}_0 - \mathbf{y}_0)^T & ... & (\mathbf{H}_N\mathbf{M}_N\mathbf{x}_0 - \mathbf{y}_N)^T \end{pmatrix}\begin{pmatrix} (\mathbf{H}_0\mathbf{x}_0 - \mathbf{y}_0) \\ \vdots \\ (\mathbf{H}_N\mathbf{M}_N\mathbf{x}_0 - \mathbf{y}_N) \end{pmatrix}
\end{aligned} \tag{5.7}
$$

Now we multiply both side of the cost function equation by σ_o^2, then cost function becomes

$$
\mathcal{J}_2(\mathbf{x}_0) = \mu^2 \mid\mid \mathbf{x}_0 - \mathbf{x}_0^b \mid\mid_2^2 + \mid\mid \hat{\mathbf{H}}\mathbf{x}_0 - \hat{\mathbf{y}} \mid\mid_2^2 \tag{5.8}
$$

where $\mathcal{J}_2 = \sigma_o^2\mathcal{J}$, $\mu^2 = \frac{\sigma_o^2}{\sigma_b^2}$, $\hat{\mathbf{y}}$ is the time sequence of observations, $\hat{\mathbf{H}}$ is the observability matrix and $\hat{\mathbf{R}}$ is generalized observation error covariance matrix which is

block diagonal:

$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{pmatrix}, \quad \hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_0 \\ \mathbf{H}_1\mathbf{M}_1 \\ \vdots \\ \mathbf{H}_N\mathbf{M}_N \end{pmatrix}, \quad \hat{\mathbf{R}} = \begin{pmatrix} \mathbf{R}_0 & & & \\ & \mathbf{R}_1 & & \\ & & \ddots & \\ & & & \mathbf{R}_N \end{pmatrix}. \tag{5.9}$$

To find the analysis at time t_0, we set the gradient of the equation (5.8) equal to zero,

$$\nabla \mathcal{J}_2(\mathbf{x}_0) = \mu^2(\mathbf{x}_0 - \mathbf{x}_0^b) + \hat{\mathbf{H}}^T(\hat{\mathbf{H}}\mathbf{x}_0 - \hat{\mathbf{y}}) \tag{5.10}$$

After the following rearrangements

$$\mu^2(\mathbf{x}_0 - \mathbf{x}_0^b) + \hat{\mathbf{H}}^T(\hat{\mathbf{H}}\mathbf{x}_0 - \hat{\mathbf{H}}\mathbf{x}_0^b + \hat{\mathbf{H}}\mathbf{x}_0^b - \hat{\mathbf{y}}) = 0$$

$$(\mu^2\mathbf{I} + \hat{\mathbf{H}}^T\hat{\mathbf{H}})(\mathbf{x}_0 - \mathbf{x}_0^b) = \hat{\mathbf{H}}^T(\hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0^b)$$

$$\mathbf{x}_0 - \mathbf{x}_0^b = (\mu^2\mathbf{I} + \hat{\mathbf{H}}^T\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^T\hat{\mathbf{d}}$$

where $\hat{\mathbf{d}} = \hat{\mathbf{y}} - \hat{\mathbf{H}}\mathbf{x}_0^b$ is called the innovation vector. We can apply the SVD to the

observability matrix $\hat{\mathbf{H}}$,

$$\mathbf{x}_0 - \mathbf{x}_0^b = (\mu^2\mathbf{I} + (\mathbf{USV}^T)^T(\mathbf{USV}^T))^{-1}(\mathbf{USV}^T)^T\hat{\mathbf{d}} \tag{5.11}$$

Using the features of the RSVs and LSVs, $\mathbf{UU}^T = \mathbf{I}$, $\mathbf{VV}^T = \mathbf{I}$ we get

$$\mathbf{x}_0 - \mathbf{x}_0^b = \mathbf{V}(\mu^2\mathbf{I} + \mathbf{S}^2)^{-1}\mathbf{SU}^T\hat{\mathbf{d}} \tag{5.12}$$

Finally the 4D-Var analysis can be written as a linear combination of RSVs

$$\mathbf{x}_0^a - \mathbf{x}_0^b = \sum_{i=1}^{r} f_i c_i \mathbf{v}_i \tag{5.13}$$

where

$$f_i = \frac{s_i^2}{\mu^2 + s_i^2} \tag{5.14}$$

$$\text{and} \quad c_i = \frac{\mathbf{u}_i^T\hat{\mathbf{d}}}{s_i}. \tag{5.15}$$

Before the assimilation, the error covariance matrices, the model and the observations specify the structure of the analysis. This structure is given by RSVs which are determined by the two factors, f_i, c_i (Johnson, 2005).

The terms f_i, which are called Tikhonov filter factors, decrease with small singular values. This is illustrated in figure 5.2. These terms plays a key role in

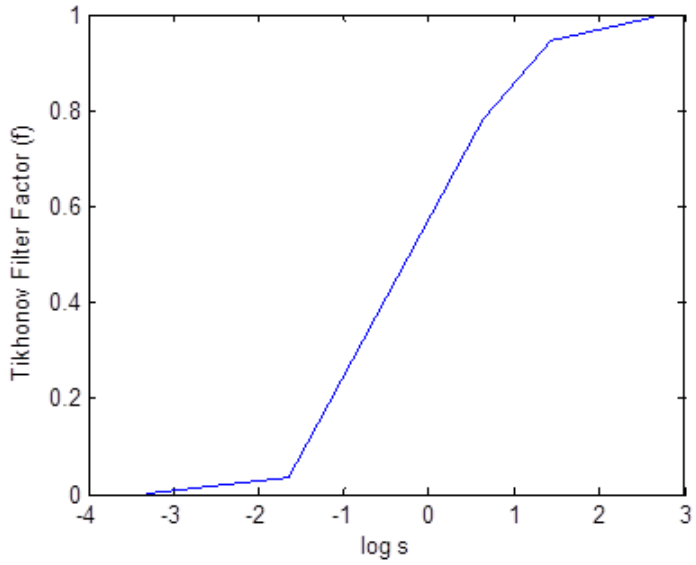

Figure 5.2: The values of Tikhonov filter factor as a function of the logarithm of the singular values

the assimilation to extract the information from the observations. Particularly, the value of μ^2 gives us the information about the interpolating and filtering aspects of 4D-Var.

When μ tends to zero, the filter factor is one so only the term c_i determines the weight of the RSVs. When $\mu^2 > 0$, there are three different cases,

$$f_i \cong \begin{cases} 1 & s_i \gg \mu \\ \frac{1}{2} & s_i = \mu \\ \frac{s_i^2}{\mu^2} & s_i \ll \mu \end{cases} \tag{5.16}$$

For a fixed value of s_i, if the observations variance is large relative to the background variance ($\sigma_o^2 > \sigma_b^2$), μ^2 becomes large so the value of f_i is close to zero. Hence the analyses reject the observations and take values close to the background. In contrast if the observation variance is small relative to the background variance ($\sigma_o^2 < \sigma_b^2$), μ^2 becomes small so the value of f_i is close to 1 and the analyses are close to the observations. Therefore the value of μ^2, which is determined by the background and observation error variances, is significant to get important information from the observations and get rid of the observational noise (Johnson, 2005).

Chapter 6

SVD Experiments

In this chapter we try to get answers to the following questions. First is why do we get a better analysis when we only observe the ocean rather than when we only observe the atmosphere? Second is why do we get a better analysis when we only observe the atmosphere with fewer and relatively inaccurate observations rather than when we only observe the ocean. To find answers to these questions we use the SVD technique which is useful to extract information from the observations.

In the following experiments there are some specific notations which are $\mathbf{H}_A$ and $\mathbf{H}_O$

$$\mathbf{H}_A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{H}_O = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.1}$$

In place of $\mathbf{H}$ we use $\mathbf{H}_A$ when we have just the atmosphere observations and $\mathbf{H}_O$ when we have just the ocean observations.

We have no observations at the initial time t_0 so the observability matrix $\hat{\mathbf{H}}$ is

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{HM}_1 \\ \mathbf{HM}_2 \\ \vdots \\ \mathbf{HM}_N \end{pmatrix}. \tag{6.2}$$

To make a comparison easily we study with the same cases as in chapter 4. The observations and the background are generated as in previous experiments.

The matlab SVD command gives the LSVs, RSVs and the singular values. In our algorithm there are 5 unknown variables (X, Y, Z, W, V) so the number of non-zero singular values is 5 and the RSVs are represented column by column in the 5x5 matrix $\mathbf{V}$. The first column of the matrix is given weight by the first singular value, second column of the matrix is given weight by the second singular value and so on. The first three elements of the each RSV represent the atmosphere variables (X, Y, Z) and the last two elements represent the ocean variables (W, V). The absolute values of these elements give us an idea of the amount of information which comes from the corresponding observations. For example, if the element $v_{5,1}$

of the matrix $\mathbf{V}$ has the highest value among the elements of first RSV, this means that the most contribution to the analysis comes from observation of the ocean variable V.

As we explained in chapter 5 the RSVs are weighted by the terms f_i and c_i. In the experiment, we use

$$\log(c_i) = \log(\frac{|\mathbf{u}_i^T \hat{\mathbf{d}}|}{s_i}) \tag{6.3}$$

which is known as the Picard ratio. In the all experiments there is not significant change in value of this term so we do not focus this term.

In the following experiments we illustrate the $\mathbf{V}$ matrix as a colour figure. We consider values above $+0.2$ and below -0.2 to be high values. The higher absolute values are indicated blue and red colours (see figures 6.2, 6.4, 6.6 and 6.8).

6.1 Experiment 1

In this experiment we start with considering case 1 in chapter 4 which is the observation frequency= $[15, 15, 15, 0, 0]$ and observation error variance is 10^{-4}. There are only atmosphere observations so in this case the observability matrix is calculated

in the following way

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_A\mathbf{M}_{15} \\ \mathbf{H}_A\mathbf{M}_{30} \\ \vdots \\ \mathbf{H}_A\mathbf{M}_{150} \end{pmatrix}. \tag{6.4}$$

We apply the SVD to this matrix and the results are shown in figure 6.1 and figure 6.2. It is clearly seen that although the first two singular values have high values, the others are extremely small. The value of filter factor is 1 for the first four RSVs and 0 for the last one so the last RSV is filtered. This means that the RSVs that correspond to the first two large singular values have large amplitudes in the system. Figure 6.2 shows the RSVs. For the first RSV which is the first column of the $\mathbf{V}$, the third and fourth elements have relatively higher values than the others so the vector mostly contributes to the analysis of the variables Z and W. On the other hand, for the second RSV, the first and second elements have relatively higher values so the most contribution come from the observations of the variables X and Y. Therefore although the system can reach the information about the variables X, Y, Z and W, it could not have enough information of variable V so the algorithm does not give a better estimate for the ocean. This is also seen in the table 4.1.

Now we consider case 2 in chapter 4 which is observation frequency=[0,0,0,10,10].

In this case the observability matrix is calculated in the following way

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_O\mathbf{M}_{10} \\ \mathbf{H}_O\mathbf{M}_{20} \\ \vdots \\ \mathbf{H}_O\mathbf{M}_{150} \end{pmatrix} \tag{6.5}$$

Figures 6.3 and 6.4 show that the singular value, filter factor, picard ratio and the RSVs of this matrix. We have again extremely small singular values except the first two of them and the filter factor filters just the last RSV. Therefore the first two RSVs that correspond to the large two singular values have large amplitudes in the system. When we analyse the first two RSVs, which are the first two columns of the matrix $\mathbf{V}$, the third and fourth elements are relatively higher than the others in the first RSV and the first, second, fourth and fifth elements are relatively higher than the third one in the second RSV. This means that the system gets the information of the variables X, Y, Z, W and V so in this case the algorithm gives a better analysis for both subsystems.

When we compare the results of these two cases, case 2 has information about all the variables although the case 1 does not have information about the variable V so we expect that the case 2 gives better reconstruction for the unobserved regions

than the case 1. This explains why we get a better analysis when we only observe the ocean than when we only observe the atmosphere.

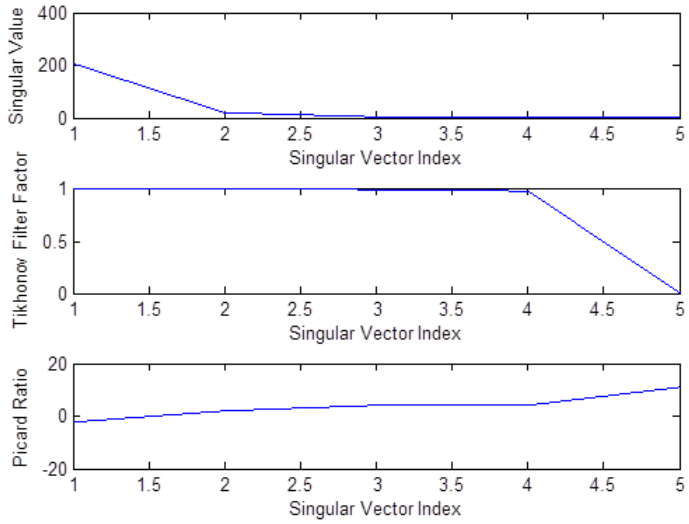

Figure 6.1: Singular values, filter factor and picard ratio of the observability matrix $\hat{\mathbf{H}}$ as a function of the singular vector index, when the observation frequency is $[15, 15, 15, 0, 0]$, the observation error variance is $\sigma_o^2 = 10^{-4}$ and the background error variance is $\sigma_b^2 = 10^{-2}$.

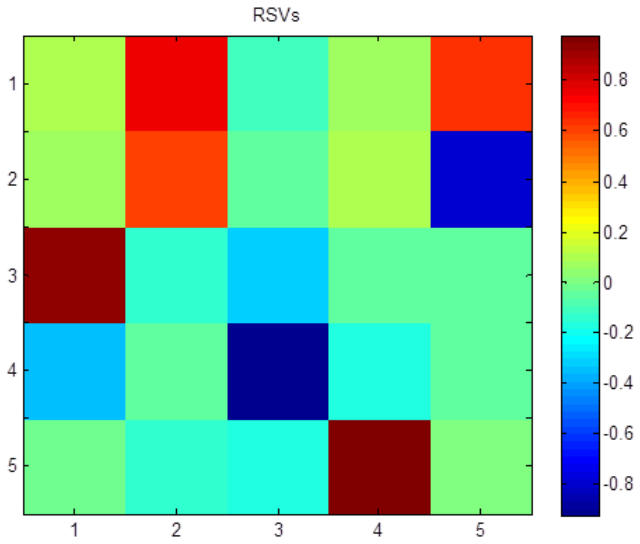

Figure 6.2: The matrix $\mathbf{V}$ which includes the RSVs, when the observation frequency is $[15, 15, 15, 0, 0]$, the observation error variance is $\sigma_o^2 = 10^{-4}$ and the background error variance is $\sigma_b^2 = 10^{-2}$. The matrix $\mathbf{V}$ is shown as an image so that each element corresponds to a colour as shown by the colour bar.

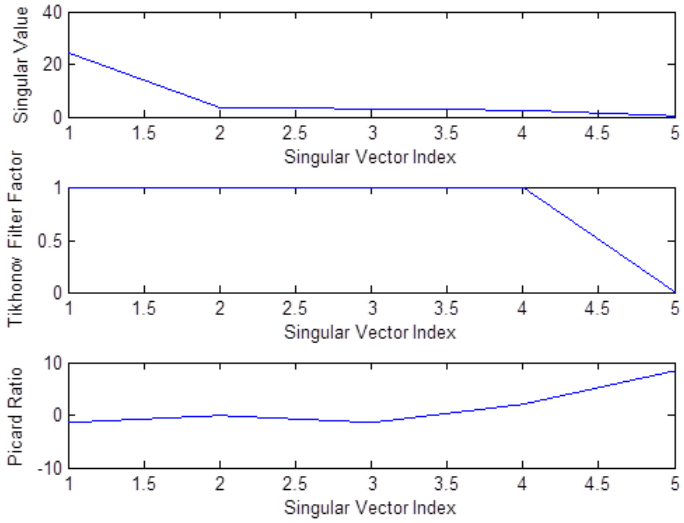

Figure 6.3: Singular values, filter factor and picard ratio of the observability matrix $\hat{\mathbf{H}}$ as a function of the singular vector index, when the observation frequency is $[0, 0, 0, 10, 10]$, the observation error variance is $\sigma_o^2 = 10^{-4}$ and the background error variance is $\sigma_b^2 = 10^{-2}$.

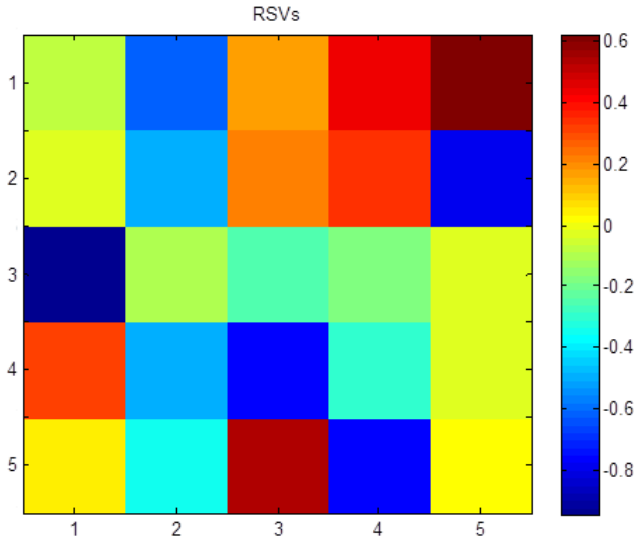

Figure 6.4: The matrix $\mathbf{V}$ which includes the RSVs, when the observation frequency is $[0, 0, 0, 10, 10]$, the observation error variance is $\sigma_o^2 = 10^{-4}$ and the background error variance is $\sigma_b^2 = 10^{-2}$. The matrix $\mathbf{V}$ is shown as an image so that each element corresponds to a colour as shown by the colour bar.

6.2 Experiment 2

In this experiment we investigate why the algorithm gives a better analysis result when we only observe the atmosphere with less and inaccurate observations than when we observe only the ocean.

Now we consider case 3 in chapter 4 which has observation frequency= $[75, 75, 75, 0, 0]$ and the observation error variance, 10^{-1}. In this case our observability matrix is

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_A\mathbf{M}_{75} \\ \\ \mathbf{H}_A\mathbf{M}_{150} \end{pmatrix} \tag{6.6}$$

We apply the SVD technique to this matrix. The results are shown in figures 6.5 and 6.6. In this experiment, we have only one high singular value and the third, fourth and fifth RSVs are filtered so the contribution comes from only the first two RSVs. The third and fourth elements of the first RSV have relatively higher values than the others and the first and second elements of the second RSV have relatively higher values than the others.

On the other hand when we consider case 4 in chapter 4, which has observation frequency= $[0, 0, 0, 50, 50]$ and observation error variance as 10^{-1}, the observability

matrix is

$$\hat{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_O\mathbf{M}_{50} \\ \mathbf{H}_O\mathbf{M}_{100} \\ \mathbf{H}_O\mathbf{M}_{150} \end{pmatrix} \tag{6.7}$$

The SVD results in this case are shown in figures 6.7 and 6.8. In this case the significant point is that the filter factor filters all of the RSVs except for the first one so only first RSV contributes to the system and we lose a lot of observation information. The system extracts the information using the first RSV. This explains why when we have less accurate and fewer observations, we can not get a better analysis in case 4 than in case 3.

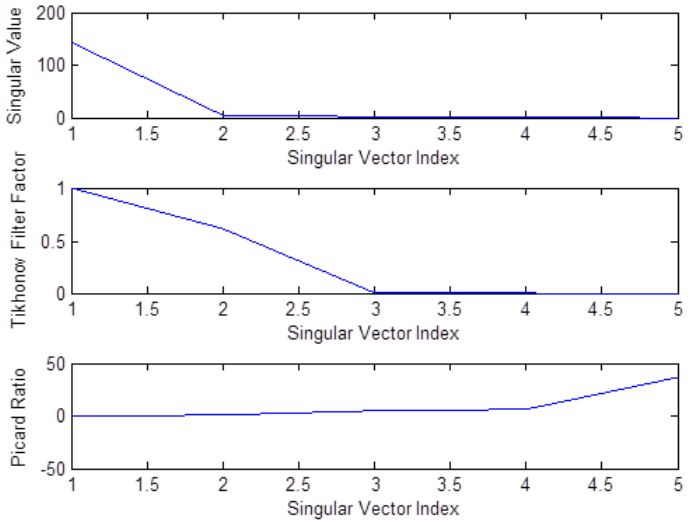

Figure 6.5: Singular values, filter factor and picard ratio of the observability matrix $\hat{\mathbf{H}}$ as a function of the singular vector index, when the observation frequency is $[75, 75, 75, 0, 0]$, the observation error variance is $\sigma_o^2 = 10^{-1}$ and the background error variance is $\sigma_b^2 = 10^{-2}$.

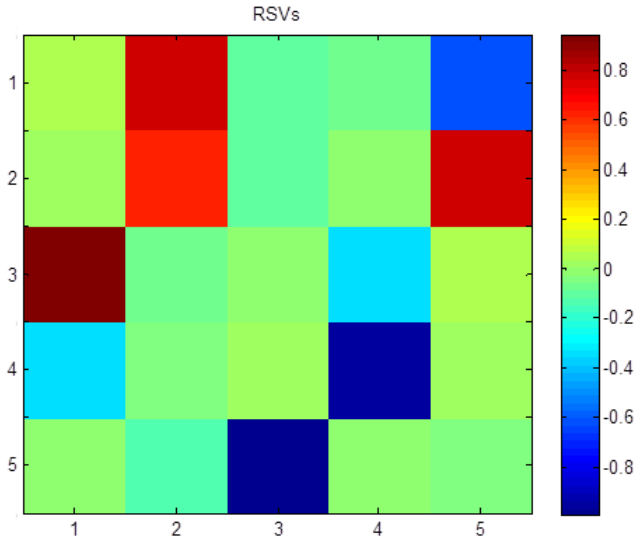

Figure 6.6: The matrix $\mathbf{V}$ which includes the RSVs, when the observation frequency is $[75, 75, 75, 0, 0]$, the observation error variance is $\sigma_o^2 = 10^{-1}$ and the background error variance is $\sigma_b^2 = 10^{-2}$. The matrix $\mathbf{V}$ is shown as an image so that each element corresponds to a colour as shown by the colour bar.

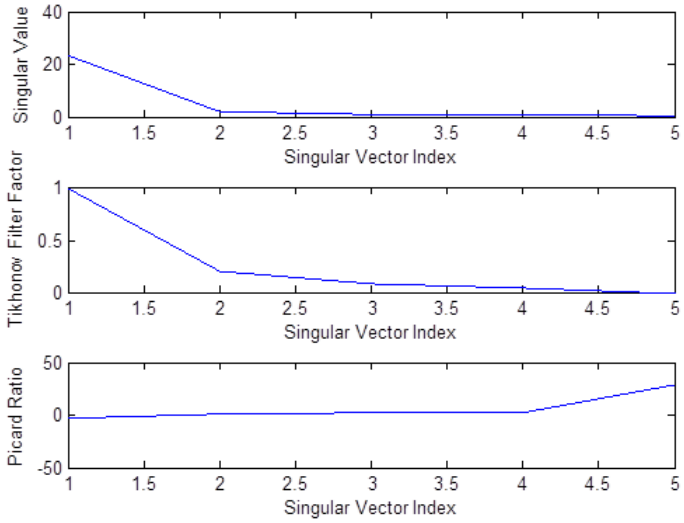

Figure 6.7: Singular values, filter factor and picard ratio of the observability matrix $\hat{\mathbf{H}}$ as a function of the singular vector index, when the observation frequency is $[0, 0, 0, 50, 50]$, the observation error variance is $\sigma_o^2 = 10^{-1}$ and the background error variance is $\sigma_b^2 = 10^{-2}$.

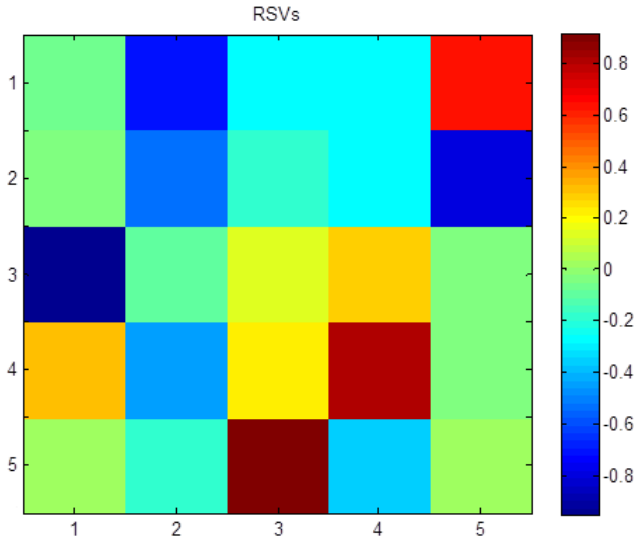

Figure 6.8: The matrix $\mathbf{V}$ which includes the RSVs, when the observation frequency is $[0, 0, 0, 50, 50]$, the observation error variance is $\sigma_o^2 = 10^{-1}$ and and the background error variance is $\sigma_b^2 = 10^{-2}$. The matrix $\mathbf{V}$ is shown as an image so that each element corresponds to a colour as shown by the colour bar.

Chapter 7

Discussion

7.1 Summary and Conclusion

This dissertation examines the behaviour of a simplified coupled atmosphere-ocean DA system. We particularly consider why we get better analysis results in the case where there is no observation in the ocean or no observation in the atmosphere. To understand this behaviour we set up a 4D-Var algorithm with the Molteni et al.(1993) model which describes the coupled atmosphere-ocean system.

From the first experiment of chapter 4 we have two important results which are that the atmosphere and ocean reconstruct the states in the unobserved region and when there are only the ocean observations, the algorithm gives better results than when we only observe the atmosphere. To understand the reason for these results we

use the SVD, which is a useful technique to extract the observational information. Therefore we applied the SVD to our situation in experiment 1 of chapter 6. The results clearly show that when we only observe the atmosphere, the algorithm could not have enough ocean information so it does not reconstruct the unobserved ocean very well. However when we only observe the ocean, the algorithm has information about all state variables, so it reconstructs the unobserved atmosphere successfully.

Then we seek the effects of the observation accuracy on the assimilation. Experiment 2 in chapter 4 gives the result that when the system uses the relatively accurate observations, the algorithm gives a better result than in the case of using inaccurate observations. Another significant result is that even when we increase the value of the observation error variance, there is still a better reconstruction in the unobserved region by the ocean observations. We reach all these results using 30 observations in total.

To understand the effect of the number of observations we considered fewer observations in experiment 3 of chapter 4. We have an important result from this experiment that is when we observe only the ocean with fewer and relatively inaccurate observations, we can not get a better analysis result than when we observe only the atmosphere. This shows that the number of observations affects the assimilation result significantly. Again to understand the reason for this result we used the SVD

technique. Experiment 2 of chapter 6 illustrates that in this case the algorithm could not extract most of the information from the observations. However when we only have atmosphere observations, the algorithm can reach almost all state variables.

In this dissertation there is a limitation which is caused by random numbers. In the experiments with imperfect observations the results can show some differences because when we do any experiment more than once, the observations are generated by different random numbers for each of these experiments. Although we use the average of the 10 different experiment results, it is impossible to reach an exact conclusion.

7.2 Future work

There are lots of different factors which affect the analysis result in 4D-Var coupled atmosphere-ocean DA algorithm. In this dissertation we have investigated the observations effects when we only observe one subsystem. We have also done a large number of experiments where observations from both subsystems were available to reconstruct all of the states. No clear conclusions could be derived in these cases, but further investigation would be valuable. In particular it would be a good idea to consider each coordinate variable separately because in this case we can get more information about the model behaviour.

The background error covariance matrix is an important part of the 4D-Var. In our study we assume there is no correlation between the background errors. For further study we can suggest using different background error covariance matrix.

Bibliography

Bannister, R. N. (2001). Elementary 4d-var. Reading: University of Reading. http://www.met.rdg.ac.uk/∼ross/Documents/Var4d.pdf

Bouttier, F., Courtier, P. (2002). Data assimilation concepts and methods March 1999. *Meteorological training course lecture series*. ECMWF.

Courtier, P., Talagrand, O. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quarterly Journal of the Royal Meteorological Society*, 113(478), 1329-1347.

Dubois, M. A., Yiou, P. (1999). Testing asynchronous coupling on simple "ocean-atmosphere" dynamic systems. *Climate dynamics*, 15(1), 1-7.

Gilbert, J. C., and Nocedal, J. (1992). Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on optimization*, 2(1), 21-42.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix computations*. The John Hopkins University Press, fourth edition, 76-80

Han, G., Wu, X., Zhang, S., Liu, Z., and Li, W. (2013). Error Covariance Estimation for Coupled Data Assimilation Using a Lorenz Atmosphere and a Simple Pycnocline Ocean Model. *Journal of Climate*, 26(24), 10218-10231.

Johnson, C. (2003). Information Content of Observations in Variational Data Assimilation. PhD thesis, University of Reading.

Johnson, C., Hoskins, B. C. and Nichols, N. K. (2005). A singular vector perspective of 4DVar: Filtering and interpolation. *Quarterly Journal of the Royal Meteorological Society* 131.605 : 1-19.

Lawless, A. S. (2013). *Variational data assimilation for very large envi-*

ronmental problems. In *Large Scale Inverse Problems: Computational Methods and Applications in the Earth Sciences* (2013), Eds. Cullen, M.J.P., Freitag, M. A., Kindermann, S., Scheichl, R., Radon Series on Computational and Applied Mathematics 13. De Gruyter, pp. 55-90.

Le Dimet, F. X., and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2), 97-110.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2), 130-141.

Mateer, C. L. (1965). On the information content of Umkehr observations. *Journal of the Atmospheric Sciences*, 22(4), 370-381.

Molteni, F., Ferranti, L., Palmer, T. N., and Viterbo, P. (1993). A dynamical interpretation of the global response to equatorial Pacific SST anomalies. *Journal of climate*, 6(5), 777-795.

Nichols, N. K. (2010). Mathematical concepts of data assimilation. *In Data Assimilation* (pp. 13-39). Springer Berlin Heidelberg.

Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., Igarashi, H. and Ishikawa, Y. (2008). Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *Journal of Geophysical Research: Oceans* (1978–2012), 113(C10).