LAGRANGIAN DATA ASSIMILATION INTO LAYERED OCEAN MODEL

by

Liyan Liu

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics.

Chapel Hill
2007

Approved by

Advisor: Christopher K. R. T. Jones
Reader: Kayo Ide
Reader: Jingfang Huang
Reader: Amit Apte
Reader: Guillaume Vernieres
ABSTRACT

Liyan Liu: Lagrangian Data Assimilation into Layered Ocean Model

(Under the direction of Professor Christopher K. R. T. Jones and Kayo Ide)

Since much surface ocean data is Lagrangian in nature, its assimilation into ocean models is a key element of an ocean forecasting system. We investigate the propagation of information vertically caused by the existing vertical correlations between a stack of layers in the water column, such as the Eulerian velocity field and other dynamical variables, by observing Lagrangian data in the surface. We test the method by using different layered models with the known Lagrangian observations at discrete time intervals in the surface layer and unknown sub-surface layers. We adopt the method for assimilating Lagrangian data in which the model is augmented with drifter advection equations and track the correlations between the flow and the drifters via the Kalman Filter. The experiments show that Lagrangian data assimilation is feasible and effective for layered models.

The technique is first tested on a two layer point vortex flow: a two layer point vortex system of \((N^v_1, N^v_2)\) vortices at each layer with a Gaussian white noise term is modeled by its deterministic counterpart. Positions of \(N_d^{(1)}\) drifter particles in the top layer are observed at regular time intervals and assimilated into the model. Numerical experiments demonstrate successful system tracking for \((N^v_1, N^v_2, N_d^{(1)}, N_d^{(2)}) = (2, 2, 1, 0)\). Our numerical model simulations show that our method is capable of successful tracking of the vortices in both of the layers by observing the Lagrangian data.
from the top layer. It demonstrates that we can capture the Eulerian velocity field of
the point vortex flow in the sub-surface layer by assimilating the Lagrangian data in
the top layer. The method we have developed gives an understanding of the potential
of Lagrangian data assimilation in models with vertical variation.

We further test the method on the two and a half layer reduced gravity shallow
water double gyre unsteady flow configuration. Our numerical simulations show that
the method is capable of correcting both of the active layers even if Lagrangian ob-
servations are only available in the top active layer and the assimilation time interval
is of the order of the Lagrangian auto-correction time scale of the flow. The results
clearly demonstrate that our method is effective when dealing with a more complex
dynamics flow with an unknown sub-surface flow structure. The Lagrangian data
assimilation method that we have developed, therefore, provides an approach that
allows us to fully realize the potential of Lagrangian data for assimilation in more
realistic ocean models.
ACKNOWLEDGEMENTS

I have waited for a long time for this chance to express my deepest gratitude to my advisor, Professor Christopher K. R. T. Jones. He always earns the respect and admiration of his students, and for me, he is the hero in my five years’ journey in the pursuit of the Ph.D. at Chapel Hill. He always supported me both academically and mentally, which was indispensable for me to complete this hard task. His impact on my attitude in research and philosophy about the world will benefit the remainder of my life.

I would also like to thank Dr. Kayo Ide for all these years advisory to me and introducing me into the Lagrangian Data Assimilation. During the process of finishing this dissertation, she gave me priceless suggestions and always kindly encouraged me. Her help was essential in my research all these years.

Thirdly, I truly appreciate Dr. Hayder Salman for kindly supporting me during my research work on two and a half layer reduced gravity shallow water model and the discussion with me about building up the model. I also very appreciate his sharing the code about single layer reduced gravity shallow water model.

Special thanks to Professor Jingfang Huang for valuable suggestions and advices I received from him from the first day I came to Chapel Hill. He always cares about my academical studies at Chapel Hill and encourages me to talk with him for any questions I have.

Many thanks go to Dr. Amit Apte and Dr. Guillaume Vernieres for their useful comments and suggestions and for insightful discussions and kind support. I thank
committee members: Prof. Christopher K. R. T. Jones, Dr. Kayo Ide, Prof. Jingfang Huang, Dr. Amit Apte, and Dr. Guillaume Vernieres for their support.

Special thanks to my friends at our research group from where I have received valuable help, and have gained valuable experience and knowledge from them. I really enjoyed and have learned so much from these past years’ collaborations with them. It was an inseparable part of my student life in North Carolina.

Of course I would also like to thank my parents Wenyu Liu, Shuqin Li and my brother Libin Liu. However, I know that words are redundant here, in comparison to their unconditional love and support, which are far beyond any language in the world.

Without tremendous help and support, I could not have achieved the hard task to complete this thesis in my graduate study. My sincere thanks for all the kindness and supports are truly beyond description.
CONTENTS

LIST OF FIGURES ........................................... xiii
LIST OF SYMBOLS .......................................... xiv

1 Introduction  1
  1.1 Background ........................................ 2
  1.2 Goals .............................................. 5
  1.3 General methodology and Principal results .......... 6
  1.4 Outline ............................................. 7

2 Theory of the Lagrangian Data Assimilation Method 9
  2.1 Sequential Lagrangian Data Assimilation ........... 10
  2.2 The Extended Kalman Filter for Lagrangian data .... 12
  2.3 The Ensemble Kalman Filter for Lagrangian data .... 16

3 Lagrangian Data Assimilation into Two Layer Point Vortex Systems 19
  3.1 Two layer point vortex system ..................... 19
5 Vertical Information Propagation and Parameter Estimation

5.1 Formulation of parameter estimation .............................. 72
5.2 Comparison of vertical correlation and horizontal correlation ... 75

6 Conclusions and Future Research

6.1 Summary and Discussion .............................................. 80
6.2 Future Research ...................................................... 84
   6.2.1 More realistic ocean models ................................. 84
   6.2.2 Covariance localization and filter tuning ................. 85
   6.2.3 Parameter estimation of more realistic ocean models ..... 85

Bibliography .......................... 87
LIST OF FIGURES

3.1 (a): BT mode velocity component and BC mode velocity component at an arbitrary point from each mode as a function of $r$ for $\lambda = 1$, where $r$ is the distance between a grid point and a particular vortex. The upper blue curve shows the velocity magnitude induced by the BT mode, while the red curve shows the velocity magnitude induced by the BC mode. (b): Same layer velocity component and cross layer velocity component at an arbitrary point in each layer as a function of $r$ for $\lambda = 1$, where $r$ is the distance between a grid point and a particular vortex. The upper blue curve is the same layer velocity magnitude, while the lower red curve is the cross layer velocity magnitude. . . . 26

3.2 Streamfunction in the co-rotating frame. Two vortices are located at (-1,0), (1,0) initially in the top layer, and two vortices are located at (0,-1), (0,1) initially in the bottom layer. . . . . . . . . . . . . 28

3.3 Trajectories of the top layer corresponding to the drifters in Figure 3.2 in the Eulerian frame. Two vortices are located at (-1,0), (1,0) initially in the top layer. Red star: initial location of two vortex points; blue star: initial location of drifter 1; black star: initial location of drifter 2; green star: initial location of drifter 3. . . . . . . . . . . . . . . . 29
3.4 The Extended Kalman Filter results in a two layer point vortex system, one drifter in the top layer is observed. \((N_v^{(1)}, N_v^{(2)}, N_d^{(1)}, N_d^{(2)}) = (2, 2, 1, 0), \Delta T = 1.0\), dynamical noise and observation error \((\sigma, \rho) = (0.02, 0.02)\). Actual error in the vortex positions \(|x_{F,i}^J - x_{F,i}^T| (i = 1, 4)\) in both layers, in the model assimilating drifter positions (yellow and green), and in the model without assimilation (red and blue) in both layers. .................................................. 32

3.5 Same assimilation experiment as in Figure 3.4. Predicted rms error in vortex location (yellow and green corresponding to Figure 3.4), vs actual error (blue) and the same for the drifter. ............................ 34

3.6 The Extended Kalman Filter results in a two layer point vortex system, one drifter in the top layer is observed. \((N_v^{(1)}, N_v^{(2)}, N_d^{(1)}, N_d^{(2)}) = (2, 2, 1, 0), \Delta T = 1.0\), dynamical noise and observation error \((\sigma, \rho) = (0.1, 0.1)\). Predicted rms error in vortex location (yellow and green) vs actual error (blue) and the same for the drifter. The system loses track after a certain time. ................................. 36

4.1 Contours of height field in the top active layer corresponding to the true flow. (a): day 0; (b): day 146. ............................... 55

4.2 Contours of height field in the top active layer corresponding to the assimilated flow. (a): day 0; (b): day 146. ................................. 56

4.3 Contours of height field in the top active layer corresponding to the non-assimilated flow. (a): day 0; (b): day 146. ................................. 57

4.4 Contours of height field in the second active layer corresponding to the true flow. (a): day 0; (b): day 146. ................................. 58
4.5 Contours of height field in the second active layer corresponding to the assimilated flow. (a): day 0; (b): day 146. .......................... 59

4.6 Contours of height field in the second active layer corresponding to the non-assimilated flow. (a): day 0; (b): day 146. ......................... 60

4.7 Error in kinetic energy $|KE|$(%). (a): top active layer; (b): second active layer. ......................................................... 62

4.8 Error in height field $|h|$(%). (a): top active layer; (b): second active layer. .......................................................... 63

4.9 Error in drifter positions in the top active layer $|\mathbf{x}_D|$. .................... 64

4.10 Error in drifter positions in time interval $70 < t_k < 90$. Red line: forecast process; Blue line: assimilation process. ......................... 65

4.11 Influence of localization shown in the error of kinetic energy and height field. $l_h^{(2)} = 2000$ km is fixed in each experiment. red: $l_h^{(1)} = 1500$ km; blue: $l_h^{(1)} = 500$ km; green: $l_h^{(1)} = 800$ km; black: No localization. (a): the top active layer; (b): the second active layer. ..................... 69

4.12 Influence of localization shown in the error of kinetic energy and height field. $l_h^{(1)} = 500$ km is fixed in each experiment. red: $l_h^{(2)} = 1000$ km; blue: $l_h^{(2)} = 2000$ km; green: $l_h^{(2)} = 4000$ km; black: No localization. (a): the top active layer; (b): the second active layer. ..................... 70
5.1 Estimation of circulation $\Gamma$ through horizontal propagation (drifter in the top layer) and vertical propagation (drifter in the bottom layer) with the change of drifter positions, under the same noise levels $(\sigma, \rho) = (0.02, 0.02)$ and observation interval $\Delta T = 1$. The parameter $\Gamma$ has an initial 20% uncertainty. The performance of assimilation depends on the horizontal distance between the vortex and the drifter...
LIST OF SYMBOLS

\( N_v^{(1)} \) vortex in the top layer, Abstract
\( N_v^{(2)} \) vortex in the second layer, Abstract
\( N_d^{(1)} \) drifter in the top layer, Abstract
\( N_d^{(2)} \) drifter in the second layer, Abstract
\( x \) state vector, Section 2.1
\( x_F \) flow state vector, Section 2.1
\( x_D \) drifter coordinates, Section 2.1
\( x_F'(t) \) forecast of flow state vector, Section 2.2
\( x_D'(t) \) forecast of drifter coordinates, Section 2.2
\( m_F(x_F', t) \) dynamical operator, Section 2.2
\( m_D(x_D', x_F', t) \) dynamical operator, Section 2.2
\( \sigma(t) \) system noise term, Section 2.2
\( Q \) covariance matrix of system noise, Section 2.2
\( P' \) model error covariance matrix, Section 2.2
\( M \) tangent linear model, Section 2.2
\( y^o(t_j) \) observation vector, Section 2.2
\( h_j \) observation function, Section 2.2
\( \rho^o(t_j) \) observation error, Section 2.2
\( R \) covariance matrix of observation error, Section 2.2
\( H \) linearized observation function, Section 2.2
\( K \) Kalman gain matrix, Section 2.2
\( x^a \) analysis state of state vector, Section 2.2
\( N_E \) ensemble size, Section 2.3
\( x_j^f \) ensemble member, Section 2.3
\( \overline{x_f} \) mean of the ensembles, Section 2.3

\( P_e^f \) ensemble error covariance matrix, Section 2.3

\( d_j(t_k) \) innovation vector, Section 2.3

\( \tilde{\rho}_j^f(t_k) \) generated ensemble noise of drifter, Section 2.3

\( r \) rank, Section 2.3

\( L \) defined in Section 2.3

\( D \) defined in Section 2.3

\( \Gamma \) strength of circulation, Section 3.1.1

\( \lambda \) radius of deformation, Section 3.1.1

\( l \) layer number, Section 3.1.1

\( \psi^{(l)} \) streamfunction, Section 3.1.1

\( \varphi^{[BT]} \) streamfunction of BT mode, Section 3.1.1

\( \varphi^{[BC]} \) streamfunction of BC mode, Section 3.1.1

\( K_0 \) modified Bessel function of the second kind of the zeroth order, Section 3.1.1

\( K_1 \) modified Bessel function of the second kind of the first order, Section 3.1.1

\( \Delta T \) time step, Section 3.2

\( N_F \) number of vortices, Section 3.2

\( N_D \) number of drifters, Section 3.2

\( \Omega \) square domain, Section 4.1

\( f \) Coriolis parameter, Section 4.1

\( \rho \) defined in Section 4.1

\( V_i \) mass flux, Section 4.1

\( U_i, V_i \) defined in Section 4.1

\( u_i, v_i \) defined in Section 4.1
\( h \) layer thickness, Section 4.1
\( \tau^x \) zonal wind forcing term, Section 4.1
\( A \) dissipation term, Section 4.1
\( \tau_0 \) wind stress, Section 4.2
\( R_1, R_2 \) interface friction, Section 4.2
\( \Delta t \) time step, Table 4.1
\((\Delta x, \Delta y)\) horizontal grid scale, Table 4.1
\( g_1, g_2 \) reduced gravity, Table 4.1
\( p_1, p_2, p_3 \) pressure term, Section 4.3.1
\((x_D(t_k), y_D(t_k))\) drifter position, Section 4.4
\( H_1, H_2 \) defined in Section 4.5.1
\( \sigma_h \) variance of Gaussian distribution, Section 4.5.1
\(|KE|^t\) kinetic energy error norms, Section 4.5.1
\(|h|^t\) height field error norms, Section 4.5.1
\(|x_D|^t\) drifter separation distances error norms, Section 4.5.1
\(|KE|^p\) predicted kinetic energy error norms, Section 4.5.1
\(|h|^p\) predicted height field error norms, Section 4.5.1
\(|x_D|^p\) predicted drifter separation distances error norms, Section 4.5.1
\( S_{FD} \) localization matrix, Section 4.7
\( S_{DD} \) localization matrix, Section 4.7
\( C_h \) horizontal localization, Section 4.7
\( r_h \) horizontal distance, Section 4.7
\( l_h^{(1)} \) horizontal localization cut-off radius, Section 4.7.1
\( l_h^{(2)} \) horizontal localization cut-off radius, Section 4.7.1
\( \mu \) parameter defined in Section 5.1.1
Chapter 1

Introduction

In this thesis, I study a Lagrangian data assimilation method in the context of various layered ocean models. Much sub-surface data in the ocean is collected by Lagrangian instruments such as ocean drifters and floats. The primary data are in the form of position information and only reflect the state variables, such as fluid velocity, through the accumulated transport of the instrument under the action of the flow. A novel scheme has been proposed recently for the assimilation of this non-state variable data into the model for the flow field by Ide et al. (2002). In that paper, the analysis can be construed as gleaning information at the same depth as the Lagrangian instruments. Since the models have a single layer, the Lagrangian drifter observations are obtained as the same layer of the flow, and the Lagrangian drifter observations are then assimilated to update the flow in that same layer. The conclusion of that work can then be stated as the fact that the Lagrangian drifter observations are able to carry the information of horizontal motion in the same layer. The question addressed here is whether this information propagates vertically within the framework of this Lagrangian data assimilation scheme. The Lagrangian drifter observations are obtained as the same layer of the flow, and the Lagrangian drifter
observations will be assimilated to update the flow model in that same layer as well as in the bottom layer. The main issue is whether the Lagrangian drifter observations are able to carry the information of horizontal motion to the other depth through the water column.

1.1 Background

Data assimilation is a novel, versatile methodology that can be used for estimating and forecasting ocean variables. Data assimilation involves combining observational data with the underlying physical principles governing the system. It incorporates observations into the forecast models. The observations could be measurements from surface stations, radars and satellites. In our research, the observations come from Lagrangian meters such as ocean drifters and floats. In this case, the problem is to combine the state of the ocean model with the Lagrangian drifter observations to produce an updated estimate of the ocean.

The physical state variables of an ocean model are usually the velocity components, height field, pressure, density, temperature and salinity. Lagrangian data assimilation is the procedure by which through assimilating Lagrangian drifters we can estimate the ocean state and can improve its estimation. Lagrangian drifters can thought of as providing an understanding of the motion of water parcels in the ocean. The Lagrangian data assimilation method has been tested on single layer models, such as a one layer point vortex system by Kuznetsov et al. (2003) as well as a reduced gravity, shallow water, double gyre flow by Salmon et al. (2006). I am interested in whether Lagrangian drifters can also provide an understanding of horizontal motion at other depths. Since the open ocean is well stratified and the layers are dynamically coupled, I might ask whether Lagrangian drifter information can exploit this capacity
to improve forecasts at depths other than those where the Lagrangian information is collected. In this thesis, I focus on the assimilation of Lagrangian drifter observations in the top layer, and ask whether this process can improve the prediction and estimate of the sub-surface ocean state. An augmented state space will be used that includes the drifter coordinates as extra variables added into the original flow state. This will allow us to assimilate the Lagrangian drifter observations directly into the system.

Data assimilation techniques, such as 3D-Var (Daley 1991; Kalnay 2003) and 4D-Var are currently widely used at national weather and ocean forecasting centers. These are variational methods that minimize a cost function which is a weighted measure of the forecast and observations. In this thesis, sequential Kalman Filters are implemented on nonlinear ocean dynamics. One of our approaches is based on applying the Extended Kalman Filter (EKF; Jazwinski 1970) to the dynamics. The Extended Kalman Filter is a generalization of the (linear) Kalman Filter to the nonlinear context. It uses a tangent linear model to approximate the nonlinear system by a linear one. It uses an error covariance matrix that acts as a measurement of the uncertainty in the estimates. This error covariance matrix is evolved along with the state estimates up to time, where observations are made. When employing a Kalman Filter method to assimilate data into an ocean model, the question arises is to how to efficiently compute the evolution of the error covariance matrix. The Extended Kalman Filter scheme uses the tangent linear model (TLM) to compute the error covariance matrix, which can be viewed as an application of Taylor expansion, which gives a linear approximation of the evolution of the error covariance matrix in nonlinear systems. Therefore, for highly nonlinear and complex ocean systems, the TLM can significantly misrepresent the true error. It will lead to the divergence of the filter. The resulting divergence of the filter might lead to the failure of the assimilation process. Another difficulty associated with TLM is that the computational cost of computing the error covariance matrix is extremely high. It is therefore not a
practical way of dealing with realistic ocean models since the state space dimension may be as large as $10^7$.

An effective alternative method was proposed by Evensen (1994, 2003) in the form of the Ensemble Kalman Filter (EnKF), which I also implement in this thesis. The key idea of the Ensemble Kalman Filter is to calculate the error covariance matrix from a set of generated ensembles, instead of computing the evolution of the error covariance matrix through the linearization. Each ensemble member is evolved under the fully nonlinear flow model. By integrating an ensemble of model states forward in time, it is possible to calculate statistical moments such as the mean and error covariance.

In the Ensemble Kalman Filter, the ensemble error covariance matrix is computed by mean of the ensembles and each of the ensemble member. The analysis state after the assimilation process is calculated as the mean of the updated ensembles. I can, therefore, expect a better performance using the Ensemble Kalman Filter than using the Extend Kalman Filter especially when dealing with highly nonlinear systems. A great advantage of the Ensemble Kalman Filter is that the computational cost is significantly less than for the Extended Kalman Filter. Suppose I have a system with $N = 10^4$ unknowns with $N$ equations, the corresponding error covariance matrix is of size $N^2 = 10^8$. TLM is not feasible to compute such a large matrix. However, by using an ensemble size of $O(80)$, I can significantly reduce the computational cost and, meanwhile, reflect the nonlinearity of the system. Different variations of the Ensemble Kalman Filter include the Ensemble Transform Kalman Filter (ETKF; Bishop et al., 2001), and the Ensemble Adjustment Kalman Filter (EAKF; Anderson, 1999). The Ensemble Kalman Filter with covariance localization is also implemented in this thesis. The localization function will retain the correlations estimated by the ensembles within a local neighborhood but eliminate all correlations beyond a
specified cut-off radius. It can be used to reduce the errors caused by the long distance weak correlations and noise.

1.2 Goals

By performing experiments with the implemented filters, such as the Extended Kalman Filter, on a layered point vortex system, we will show that by assimilating the Lagrangian drifter observations in the top layer of the two layer point vortex system, I am able to correct the vortex positions in both of the layers. We shall also perform experiments with the implemented filters, such as the Ensemble Kalman Filter, to a layered reduced gravity shallow water double gyre flow system, and show that by, assimilating the Lagrangian drifter observations in the top active layer of the two and a half shallow water model, I am able to correct the flow fields in both of the active layers. Moreover, we shall perform experiments to estimate parameters in the layered point vortex system to compare the horizontal correlation and vertical correlation.

The ultimate goal that motivates this work is the forecasting and estimation of coherent structures of the three dimensional ocean and the providing of regular descriptions of the pressure, temperature, salinity and velocity. The idea is to develop and test Lagrangian data assimilation methods for the systematic development of ocean circulation models and the assimilation of Lagrangian data in ocean models in support of the ocean modeling and prediction. Since oceans are strongly stratified away from the coast, the three-dimensional ocean flow can be modeled as stacks of two-dimensional layers. This thesis is focusing on Lagrangian data assimilation of layered ocean models. I implemented the scheme on layer models to investigate whether the proposed Lagrangian data assimilation scheme is able to propagate information vertically. When Lagrangian data are assimilated at some fixed depth, information is
gathered at one level, it can then be used to improve the prediction of the flow field at other levels.

1.3 General methodology and Principal results

To investigate the questions described in the previous section about ocean modeling and Lagrangian data assimilation, Lagrangian data assimilation experiments are performed on a two layer point vortex system and also on a two and a half layer reduced gravity shallow water double gyre system. The Extended Kalman Filter is implemented into the two layer point vortex system since it is a lower dimensional model. The Ensemble Kalman Filter is implemented in the two and a half layer reduced gravity shallow water double gyre system. This is appropriate and effective since it is a highly nonlinear system and a large size of model. The Lagrangian drifter observations are coming from the drifter trajectories at discrete time intervals in the top layer of the model. The Lagrangian drifter observations are used to assimilate and to update the flow states in the top layer and also in the bottom layer.

The Lagrangian data assimilation experiments in this thesis are tested through the identical twin experiments that I set up. A true run state is created and Lagrangian drifter observations are simulated by adding Gaussian noise that represent observational error. An assimilation run is performed by assimilating the observations at each assimilation time step. The identical twin experiments show that Lagrangian data will be used to correct the flow field in the top layer where Lagrangian drifters are deployed and the bottom layers where no Lagrangian observations are directly available. The most important result of the dissertation is that Lagrangian drifter data can be used to not only improve the predictions of the ocean flow states in the same layer with the drifters but also in the sub-surface layers where no drifters are
Another goal of ocean data assimilation is the estimation of parameters. The fundamental properties of the ocean system appear in the field equations as parameters. In principle the parameters of the ocean system can be estimated directly from measurements. However, in practice, directly measuring the parameters of an ocean system is difficult because of the difficulties surrounding data acquisition in the ocean. Data assimilation provides a powerful methodology for parameter estimation. By including the key parameters into the flow state vectors, they can be estimated along with the state vectors whenever observations are available. For example, the key parameters in the point vortex system are vortex strength of circulation. In practice, this parameter is hard to measure directly. Data assimilation provides a feasible approach by including such parameter into the state space and improving the estimation of parameter in the process of assimilation.

1.4 Outline

Chapter 2 describes the theory of Lagrangian data assimilation using an augmented system. It gives the formulation of the Extended Kalman Filter and the Ensemble Kalman Filter for Lagrangian data assimilation framework.

Chapter 3 firstly describes the formulation of the two layer point vortex system in terms of the decomposition into a barotropic mode and a baroclinic mode. It also presents the results of experiments using the Extended Kalman Filter scheme to assimilate the Lagrangian data from the top layer of the two layer point vortex system to improve the prediction of the vortex motions in both of the layers.

Chapter 4 firstly describes the formulation of the two and a half layer reduced
gravity shallow water double gyre flow configuration. It also presents the results of experiments using the Ensemble Kalman Filter scheme to assimilate the Lagrangian data from the top layer of the two and a half layer reduced gravity shallow water system to correct the velocity field and height field in both of the active layers.

Chapter 5 firstly describes the formulation of parameter estimation using Lagrangian data assimilation. It also presents the results of experiments using the Extended Kalman Filter to estimate the key parameters, i.e. the strength of vortex circulation $\Gamma$ in the two layer point vortex system to compare the horizontal correlation and vertical correlation.

The dissertation concludes in Chapter 6 with a review of preceding chapters and some suggestions for further work.
Chapter 2

Theory of Lagrangian Data

Assimilation Method

Lagrangian meters, such as ocean drifters and floats, have been used for some time to provide information about the ocean (Mariano, 2000). They revealed key phenomena in the ocean, for instance the discovery of meddies, etc (Dutkiewicz et al. 2001; Garfield et al. 2001; Carr and Rossby 2001; Bower and Hunt 2000). Such meters also provide a clear picture of the horizontal motion of water parcels. It is argued here that the trajectories of Lagrangian drifters also contain detailed quantitative information about the dynamics of the underlying flow. The suggestion is that Lagrangian data can be exploited as a predictive tool in ocean forecasting. Several other attempts have been made to assimilate Lagrangian data to correct the evolution of dynamical models and for estimating Eulerian velocity field by Molcard et al. (2003) and Özgökmen et al. (2003). The problem of using Lagrangian data is that most assimilation schemes in oceanography and meteorology use model variables computed on a fixed grid in space (Ghil and Malanotte Rizzoli 1991), whereas the Lagrangian observations are distributed non-uniformly over the space and do not give the data
Recently, in the study of Ide et al. (2002), a new method was presented for the assimilation of Lagrangian data. The essential idea behind their approach was to augment the state space of the model by including drifter coordinates as additional variables. In doing so, the problem for Lagrangian data that they do give data directly in terms of model variables is resolved. By introducing the drifters into the dynamical system and tracking the correlation between the state vector and drifters, they were able to extract information about the flow from drifter observations. This chapter presents a unified exposition of various formulations of the Kalman Filters for Lagrangian data assimilation, including the Extended Kalman Filter and the Ensemble Kalman Filter.

2.1 Sequential Lagrangian Data Assimilation

This section reviews the theory of Lagrangian data assimilation methods that have been defined in recent research. For nonlinear systems, the classical Kalman Filter has been extended using approximations of the nonlinear system, such as the Extended Kalman Filter (EKF). The Extended Kalman Filter is based on the linearization of the model evolution. The Extended Kalman Filter handles small nonlinearities, but is rather inefficient in case of very nonlinear systems and large scale computations, such as mesoscale shallow water double gyre model. The other difficulty is to write a tangent linear model for complex models or for highly nonlinear models. For complex models and large dimensional systems, difficulties arise as how to reduce the computational cost and how to represent the error covariance matrix. Sub-optimal schemes are designed to deal with high dimensional systems and highly nonlinear models. For example, the Ensemble Kalman Filter method (Anderson 2001; Evensen 1994;
Heemink et al. 2001; Pham 2001), where the covariance matrix is approximated by a set of realizations (the ensemble), is based on a Monte Carlo simulation. The ensembles of model forecasts are performed from which the error covariance matrix can be derived by construction. Each ensemble member is evolved by the fully nonlinear flow model, so no linearization is made. Thus, I can expect a better performance for the Ensemble Kalman Filter than with the Extended Kalman Filter.

For Lagrangian data assimilation, I consider an augmented state space that includes the drifter coordinates as extra variables (Ide et al. 2002). This augmented state vector $\mathbf{x} = (\mathbf{x}_F, \mathbf{x}_D)$ combines the Eulerian part $\mathbf{x}_F$, which describes the flow state vector, and the Lagrangian part $\mathbf{x}_D$, which describes the drifter coordinates. The equations of model evolution of the flow state variables are correspondingly augmented by the advection equations of drifters $\mathbf{x}_D$. Thus, the error covariance matrix is augmented by including the correlation between the errors in flow variables and the errors in drifter positions. Since the augmented error covariance matrix includes the correlation between the state vector $\mathbf{x}_F$ and the drifter position $\mathbf{x}_D$, the assimilation of the drifter positions is able to improve not only the forecast of the drifter positions $\mathbf{x}_D$, but also able to correct the forecast of the flow variables $\mathbf{x}_F$. So, by including the drifters into the dynamical model and tracking them and their correlations with the flow state, I can correct the flow field every time that Lagrangian drifter information are available.

When the Lagrangian drifter observations are taken only from the surface layer, or the flow at a certain depth, the suggestion is that they might also contain information on the flow field at sub-surface layer or other depths because of the dynamical coupling of the layers. When Lagrangian data are assimilated at this fixed depth, information gathered at one level can hopefully be used to improve the prediction of the flow field at other levels. I apply the proposed scheme here to layered ocean models, which
shows the successful tracking of the flow field in both of the layers.

2.2 The Extended Kalman Filter for Lagrangian data

I will mainly adopt the Extended Kalman Filter scheme as discussed in the paper by Kuznetsov et al. (2003). I consider a numerical layered ocean model that is represented by state vector $x^f_F(t)$ in each layer, which is the Eulerian flow variables, $f$ denotes the forecast state. In general, $x^f_F(t)$ could contain information about the velocity field, height field, salinity, pressure, etc. for each grid point in each layer in the flow domain. The evolution of the flow state vector can be represented as:

$$\frac{dx^f_F}{dt} = m_F(x^f_F, t) \quad (2.1)$$

where $m_F$ is the corresponding dynamics operator.

In this work, I consider the observations as coming from the positions of Lagrangian drifters. At time $t_k$, the horizontal positions of the drifters $x^f_D(t_k)$ are observed. The evolution of the drifter state vector can be represented as:

$$\frac{dx^f_D}{dt} = m_D(x^f_D, x^f_F, t) \quad (2.2)$$

To track the information about the flow from the drifter observations I augment the model state space:

$$x = \begin{pmatrix} x_F \\ x_D \end{pmatrix} \quad (2.3)$$
combines the state vector of the flow $x_F$ and the vector of the drifter coordinate $x_D$. Tracer advection equations are then added to the model, so our new augmented system is:

$$\frac{dx_F}{dt} = m_F(x_F, t)$$
$$\frac{dx_D}{dt} = m_D(x_D, x_F, t) \quad (2.4)$$

The first part of the system equation is the unchanged original model, the second part represents the drifter advection equations. The discretized dynamics of $x^t$, (the superscript $t$ denotes the true state vector), is assumed to be a stochastic system with an unresolved small scale noise term $\sigma(t)dt$, and I assume below that it is a zero-mean Gaussian white noise: $E[\sigma(t)\sigma(t)^T] = Q$.

$$dx_F^t = m_F(x_F^t, t)dt + \sigma(t)dt$$
$$dx_D^t = m_D(x_F^t, x_D^t, t)dt + \sigma(t)dt \quad (2.5)$$

The key element of the EKF is tracking the evolution of the model error covariance matrix:

$$P^f = E[(x^f - x^t)(x^f - x^t)^T] = E[\begin{pmatrix} (x_F^f - x_F^t)(x_F^f - x_F^t)^T & (x_F^f - x_F^t)(x_D^f - x_D^t)^T \\ (x_D^f - x_D^t)(x_F^f - x_F^t)^T & (x_D^f - x_D^t)(x_D^f - x_D^t)^T \end{pmatrix}] \quad (2.6)$$

using the tangent linear model (TLM), which is approximated by a closed equation for the predicted error covariance matrix $P^{f'}$:

$$\frac{dP^f}{dt} = (MP^f) + (MP^f)^T + Q \quad (2.7)$$
where

\[
P_f = \begin{pmatrix} P_{ff} & P_{fd} \\ P_{df} & P_{dd} \end{pmatrix}
\]  
(2.8)

\[
M = \left. \frac{\partial m(x, t)}{\partial x} \right|_{x(t)}
\]  
(2.9)

\[
M = \begin{pmatrix} M_{ff} & 0 \\ M_{df} & M_{dd} \end{pmatrix}
\]  
(2.10)

\(M\) is the linearized dynamics operator, evaluated at \(x_f(t)\). Sequential data assimilation updates the model every time an observation of the true state becomes available. The observation at time \(t_j\) can be written in terms of:

\[
y^o(t_j) = h_j[x(t_j)] + \rho(t_j)
\]  
(2.11)

where \(h_j\) is an observation function and \(\rho(t_j)\) are random variables representing errors of the observations, which are typically assumed to be uncorrelated zero-mean Gaussian distribution with a covariance \(R\):

\[
E[\rho(t_j)] = 0, \quad E\{\rho(t_j)[\rho(t_j)]^T\} = R
\]  
(2.12)

where \(R\) is the covariance matrix of observation error.

The observation function \(h\) corresponding to drifter positions is linear in our method because of the nature of Lagrangian data:

\[
h[x] = Hx, \quad H = (0 \ I)
\]  
(2.13)
This results in a special form of the Kalman gain matrix:

\[
K = P' H^T (HP' H^T + R)^{-1}
\]

\[
= \begin{pmatrix}
P_{FD}' \\
\quad \\
\quad \\
\end{pmatrix}
(\begin{pmatrix}
P_{FD}' \\
\quad \\
\quad
\end{pmatrix} + \begin{pmatrix}
R \\
\quad
\end{pmatrix})^{-1}
\]

(2.14)

The weights with which drifter observations correct the flow variables \(x_F\) are given by the first \(N\) rows of \(K\), they are proportional to \(P_{FD}'\), the correlations between the flow state and the drifter positions. These correlations always appear because drifter paths depend on the flow, \(M_{DF} \neq 0\).

The updated error covariance matrix is given by:

\[
P^a = (I - KH) P'
\]

(2.15)

The update \(x^a\) can be written as, \((a\) denotes the analysis state):

\[
x^a = \begin{pmatrix}
x^a_F \\
x^a_D
\end{pmatrix}
= \begin{pmatrix}
x_F' \\
\quad \\
x_D'
\end{pmatrix} + K (y^o - Hx_f')
\]

(2.16)

Sequential Kalman Filter updates the model every time an observation of true state becomes available. The update \(x^a\) will be used as the new forecast state for the model evolution at next time step.
2.3 The Ensemble Kalman Filter for Lagrangian data

I will mainly adopt the Ensemble Kalman Filter scheme as discussed in the paper by Salman et al (2006). The evolution of the model error covariance matrix $P^f$ estimated by the Ensemble Kalman Filter is still defined as:

$$
P^f = E[(x^f - x^i)(x^f - x^i)^T]$$

$$= \begin{bmatrix}
(x_F^f - x_F^i)(x_F^f - x_F^i)^T (x_F^D - x_D^f)(x_F^D - x_D^f)^T
(x_D^f - x_D^i)(x_F^f - x_F^i)^T (x_D^D - x_D^i)(x_D^D - x_D^i)^T
\end{bmatrix}$$

(2.17)

The background error covariance matrix of the Ensemble Kalman Filter is calculated directly from the ensembles instead of solving an evolution equation of $P^f$. In the Extended Kalman Filter employed by Kuznetsov et al., the evolution of $P^f$ is generated by the TLM, which is a linear approximation for the evolution of the covariance matrix. In the Ensemble Kalman Filter, there is no need to compute the evolution equation explicitly. Instead, the forecast covariance matrix $P^f$ is constructed by all the ensemble members and the ensemble mean.

$$P^f \approx E[(x^f - \overline{x^f})(x^f - \overline{x^f})^T]$$

(2.18)

where the ensemble mean is defined as:

$$\overline{x^f} = \frac{1}{N_E} \sum_{j=1}^{N_E} x^f_j$$

(2.19)
So the ensemble covariance matrix can be defined as reference here:

$$P^f_e = \frac{1}{N_E - 1} \sum_{j=1}^{N_E} (x^f_j - \overline{x})(x^f_j - \overline{x})^T$$

where $N_E$ is the ensemble size. And $x^f_j$, $(j = 1, ..., N_E)$, are the individual ensemble members. Nonlinear effects for the evolution of the covariance matrix are included in the Ensemble Kalman Filter. And each ensemble member is evolved through the fully nonlinear system of dynamical equations. So the Ensemble Kalman Filter is more appropriate for highly nonlinear systems than the Extended Kalman Filter, because the error covariance matrix $P^f$ of the Extended Kalman Filter is computed by the linearization of the evolution. However, the computational costs prevent us from using a very large ensemble size. I will test the scheme using an ensemble size about $O(100)$ members or less to assess the feasibility of our Lagrangian data assimilation algorithm with the Ensemble Kalman Filter.

The first order approximation of the analysis state can be constructed through

$$x^a_j(t_k) = x^f_j(t_k) + K(t_k)d_j(t_k)$$

The update for each ensemble member can be interpreted as a combination of the model forecast and observation of system with observation error in it. The expression $d_j(t_k)$ is the innovation vector, the vector giving the difference between the observations and the prediction of the observed quantities by the model.

$$d_j(t_k) = y^o(t_k) - h_j[x^f_j(t_k)] + \tilde{\rho}_j(t_k)$$
The Kalman gain matrix is defined as:

$$K = P_e^f H^T (HP_e^f H^T + R_e)^{-1}$$

$$= \begin{pmatrix} (P_{FD})_e \\ (P_{DD})_e \end{pmatrix} [(P_{DD})_e + R_e]^{-1}$$

(2.23)

The additional error $\tilde{\rho}_j(t_k)$, introduced to each ensemble member of drifter, is required to circumvent the problem of generating an updated ensemble which has a variance that is too low. If the variance keeps decreasing with each update, the filter will assume greater confidence in the model and less confidence in the observations. So eventually it will lead to a model trajectory that is not related to the observations. This additional error is added in a way that ensures it has zero mean and an uncorrelated Gaussian with variance given by

$$\frac{1}{N_E - 1} \sum_{j=1}^{N_E} [\tilde{\rho}_j(t_k) [\tilde{\rho}_j(t_l)]^T = \delta_{kl} R_e$$

(2.24)

where $R_e$ converges to $R$ (error covariance matrix of observations noise). In all, the perturbed observation ensembles are generated from the same distributions as the observation errors.

So, although the error covariance matrix of the augmented system $P_e$ is computationally expensive in construction, I note that the above results reduce the computational cost substantially. Instead of computing the whole error covariance matrix $P_e$, only the elements of $(P_{FD})_e$ and $(P_{DD})_e$ need to be computed. The analysis state will be computed as the mean of the updated ensemble members:

$$\bar{x}^a = \frac{1}{N_E} \sum_{j=1}^{N_E} x_j^e$$

(2.25)
Chapter 3

Lagrangian Data Assimilation into Two Layer Point Vortex System

3.1 Two layer point vortex system

This section is mainly a brief description of the two layer point vortex model as given in the work of Hogg and Stommel (1985), Legg and Marshall (1993, 1996, 1998). The dynamics of the two layer model is given by a weighted sum of both modes (barotropic and baroclinic). All the vortices will interact with each other through the modes. The vortices induce a velocity field around them with specifics determined by the model parameters such as the strength of circulation associated with each vortex $\Gamma$ and the radius of deformation of each layer, called $\lambda$. 
3.1.1 Point vortex system in the two layer model

Point vortex flow is a singular solution of the Euler equation that consists of a flow with zero vorticity everywhere except at a finite number of points. Suppose there are $N_v^{(l)}$ vortices in the $l$-th layer $l (l = 1, 2)$, and the $k$-th vortex in layer $l$, circulation $\Gamma_k^{(l)}$, is located at $x_k^{(l)} \ (k=1,...,N_v^{(l)})$, where $x = (x, y)^T$ is the two dimensional coordinate, unless otherwise noted. The quantity $\lambda$ denotes the strength of the coupling for the two layers. The equations for determining the streamfunctions $\psi^{(l)} \ (l = 1, 2)$ in each of the two layers can be written as:

$$\nabla^2 \psi^{(1)}(x) - \frac{1}{2} \lambda^{-2} (\psi^{(1)}(x) - \psi^{(2)}(x)) = \sum_{k=1}^{N_v^{(1)}} \Gamma_k^{(1)} \delta(x - x_k^{(1)})$$

$$\nabla^2 \psi^{(2)}(x) - \frac{1}{2} \lambda^{-2} (\psi^{(2)}(x) - \psi^{(1)}(x)) = \sum_{k=1}^{N_v^{(2)}} \Gamma_k^{(2)} \delta(x - x_k^{(2)}) \quad (3.1)$$

where the superscript $(l)$ denotes the layer number $(l = 1, 2)$, and $\delta$ is the delta function.

Each layer is dynamically coupled with the adjacent layer and the strength of that coupling is given by the radius of deformation $\lambda$. The two layer point vortex system can be decomposed into a combination of barotropic and baroclinic mode. So the dynamics in layer $l \ (l = 1, 2)$ is given by weighted sum of both modes, where the vortices in both layers contribute to vortex interaction in each individual mode. So the two layer point vortex system can be solved by decomposing the system into a barotropic mode and a baroclinic mode, and then sum the two modes up with appropriate weights.
I can convert the layer model to a mode model by multiplying by $\mathbf{V}$.

\[
\nabla^2 \phi(x) - \Lambda^2 \phi(x) = \mathbf{q}(x)
\]

(3.2)

where

\[
\Lambda^2 = \begin{pmatrix}
0 & 0 \\
0 & \lambda^{-2}
\end{pmatrix}
\]

\[
\mathbf{V} = \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} \\
1 & -1
\end{pmatrix}
\]

\[
\phi(x) = \begin{pmatrix}
\phi^{[\text{BT}]}(x) \\
\phi^{[\text{BC}]}(x)
\end{pmatrix} = \mathbf{V} \psi(x)
\]

\[
\mathbf{q}(x) = \begin{pmatrix}
\mathbf{q}^{[\text{BT}]}(x) \\
\mathbf{q}^{[\text{BC}]}(x)
\end{pmatrix} = \mathbf{V} \begin{pmatrix}
\omega^{(1)}(x) \\
\omega^{(2)}(x)
\end{pmatrix}
\]

(3.3)

Where

\[
\omega^{(1)}(x) = \sum_{k=1}^{N^{(1)}} \Gamma_k^{(1)} \delta(x - x_k^{(1)})
\]

\[
\omega^{(2)}(x) = \sum_{k=1}^{N^{(2)}} \Gamma_k^{(2)} \delta(x - x_k^{(2)})
\]

(3.4)

Or:

\[
\nabla^2 \phi^{[\text{BT}]}(x) = \frac{1}{2} \omega^{(1)}(x) + \frac{1}{2} \omega^{(2)}(x)
\]

\[
\nabla^2 \phi^{[\text{BC}]}(x) - \lambda^{-2} \phi^{[\text{BC}]}(x) = \omega^{(1)}(x) - \omega^{(2)}(x)
\]

(3.5)
where the right-hand side terms show the contributions of the vortices in the individual layer to the barotropic mode and baroclinic mode.

Here I decomposed the two layer point vortex system into a combination of barotropic mode and baroclinic mode. So the dynamics in layer $l$, ($l = 1, 2$), are given by a weighted sum of all modes, where the vortices in both layers contribute to vortex interaction in each individual mode.

The barotropic (BT) mode model for the point vortex system has the form:

$$\nabla^2 \phi^{[\text{BT}]}(x) = \sum_{k=1}^{N_v} \Gamma_k \delta(x - x_k)$$  \hspace{1cm} (3.6)

where $\phi^{[\text{BT}]}$ represents the streamfunction for a mode model, the superscript in $[\text{BT}]$ denotes the type of mode. A point vortex system is described by a number $N_v$, and the $k$-th vortex of the circulation $\Gamma_k$ is located at $x_k$ ($k=1,\ldots,N_v$) as in the right hand side of the equation.

The streamfunction of the BT model is:

$$\phi^{[\text{BT}]}(x) = \sum_{k=1}^{N_v} \frac{\Gamma_k}{2\pi} \log |x - x_k|$$  \hspace{1cm} (3.7)

Where $|x - x_k|^2 = (x - x_k)^2 + (y - y_k)^2$. A barotropic point vortex system induces an angular velocity which decays like $1/r$, as shown in Figure 3.1(a), where $r$ is the distance between the vortex point $x_k$ and an arbitrary point $x$.

The baroclinic (BC) mode model for the point vortex system has the form:

$$\nabla^2 \phi^{[\text{BC}]}(x) - \lambda^2 \phi^{[\text{BC}]}(x) = \sum_{k=1}^{N_v} \Gamma_k \delta(x - x_k)$$  \hspace{1cm} (3.8)

The BC mode is controlled by $\lambda$, as I mentioned.
The streamfunction of the BC model is:

$$\phi^{[\text{BC}]}(x) = -\sum_{k=1}^{N_v} \frac{\Gamma_k}{2\pi} K_0\left(\frac{|x - x_k|}{\lambda}\right)$$  \hspace{1cm} (3.9)$$

where $K_0$ is the modified Bessel function of the second kind of the zeroth order. A BC point vortex system induces an angular velocity field that decays like $(r/\lambda)K_1(r/\lambda)/r$, as shown in Figure 3.1(a), where $K_1$ is the modified Bessel function of the second kind of the first order. Because $\hat{r}K_1(\hat{r})$ decreases monotonically from $\lim_{\hat{r} \to 0} \hat{r}K_1(\hat{r}) = 0$ to $\lim_{\hat{r} \to \infty} \hat{r}K_1(\hat{r}) = 1$, the BC velocity decreases like the BT velocity for small $\hat{r} = r/\lambda$ but much faster than the BT velocity for large $\hat{r} = r/\lambda$, as shown in Figure 3.1(a).

3.1.2 Velocity field induced by the vortices

The solution of the two layer point vortex system is based on the BT mode and BC mode solutions. So the streamfunction of the two layer point vortex system in each layer $l$ ($l = 1, 2$) is:

$$\psi^{(l)}(x) = \frac{1}{2} \sum_{k=1}^{k=2} \sum_{j=1}^{j=N_v^{(l)}} \frac{\Gamma_j^{(k)}}{2\pi} \left(\log\left(\left| x - x_j^{(k)} \right| \right) - \theta^{(l,k)} K_0\left(\frac{|x - x_j^{(k)}|}{\lambda}\right)\right)$$  \hspace{1cm} (3.10)$$

where the first term in the parenthesis corresponds to the BT mode and the second term corresponds to the BC mode. $l$ denotes the layer number of $(x^{(l)}, y^{(l)})$, and $\theta^{(l,k)}$ is a quantity with the value 1 if $l = k$, indicating that $(x^{(l)}, y^{(l)})$ is in the same layer with the vortex point $(x_j^{(k)}, y_j^{(k)})$, or -1 if $l \neq k$.

Suppose $x^{(l)} = (x^{(l)}, y^{(l)})$ is an arbitrary point in the two layer model, the Eulerian
The velocity induced at this particular point is:

\[
\begin{align*}
\frac{dx^{(l)}}{dt} &= -\sum_{k=1}^{K} \sum_{j=1}^{N(k)} \frac{1}{2} \frac{1}{2\pi} \left( \frac{y^{(l)} - y_j^{(k)}}{|x^{(l)} - x_j^{(k)}|^2} \right) \left[ 1 + \theta^{(l,k)} \left( \frac{|x^{(l)} - x_j^{(k)}|}{\lambda} \right) \right] K_1 \left( \frac{|x^{(l)} - x_j^{(k)}|}{\lambda} \right) \\
\frac{dy^{(l)}}{dt} &= \sum_{k=1}^{K} \sum_{j=1}^{N(k)} \frac{1}{2} \frac{1}{2\pi} \left( \frac{x^{(l)} - x_j^{(k)}}{|x^{(l)} - x_j^{(k)}|^2} \right) \left[ 1 + \theta^{(l,k)} \left( \frac{|x^{(l)} - x_j^{(k)}|}{\lambda} \right) \right] K_1 \left( \frac{|x^{(l)} - x_j^{(k)}|}{\lambda} \right)
\end{align*}
\]

The vortices also move according to the above expression except that there is no self-induced velocities.

From the expression of the velocity field in the two layer as shown in the above expression, I first can see that there are two components: BT mode which has a magnitude proportional to \(1/r\), and BC mode which has a magnitude proportional to \((r/\lambda)K_1(r/\lambda)/r\) as I explained in the previous Section, which influence the velocity at an arbitrary point \(x^{(l)} = (x^{(l)}, y^{(l)})\). And the velocity field is a weighted sum of both of the BT mode and BC mode. The similarity for small ratio \(r/\lambda\) of both mode velocities and the dissimilarity for large ratio \(r/\lambda\) of both mode velocities determine the drastically different behaviors for the velocity field.

The velocity induced from a vortex which is located at the same layer with \(x^{(l)} = (x^{(l)}, y^{(l)})\), is the average of the BT mode and BC mode velocities, thus has a magnitude proportional to \((1 + (r/\lambda)K_1(r/\lambda))/2r\), which can be viewed as the same-layer velocity component. It decays like \(1/r\) for small \(r/\lambda\) where the BT mode and BC mode velocities are nearly the same, but like \(1/2r\) for large \(r/\lambda\) where the BC mode velocity becomes negligible faster than the BT mode velocity. The velocity induced from a vortex which is located in the other layer from \(x^{(l)} = (x^{(l)}, y^{(l)})\), has the magnitude proportional to \((1 - (r/\lambda)K_1(r/\lambda))/2r\), i.e., it is half of the difference between the BT mode and BC mode velocities because the BC mode acts against the
BT mode in the cross-layer interaction, which can be viewed as the cross-layer velocity component. It is 0 at $r/\lambda = 0$ where the BC mode velocity completely cancels out the BT mode velocity, increases until it reaches a maximum at the point where the gradient vanishes, i.e. $\frac{d(1-(r/\lambda)K_1(r/\lambda))/2r}{dr} = 0$, which might lead to the break down of the TLM, then decreases to zero with the rate of the decay proportional to $1/2r$ for large $r/\lambda$ as shown in Figure 3.1(b).

The velocity field is also dependent on radius of layer deformation $\lambda$. This is not hard to understand since the small radius of layer deformation indicates that the layers are strongly dynamically coupled. As the radius of deformation increases, the contribution from the vortices in the same layer dominates the Eulerian velocity field since a large radius of layer deformation indicates that the layers are weakly dynamically coupled. The contribution from the vortices in the different layers becomes weaker. As the radius of deformation becomes larger and larger, vertical interaction is inhibited.
Figure 3.1: (a): BT mode velocity component and BC mode velocity component at an arbitrary point from each mode as a function of $r$ for $\lambda = 1$, where $r$ is the distance between a grid point and a particular vortex. The upper blue curve shows the velocity magnitude induced by the BT mode, while the red curve shows the velocity magnitude induced by the BC mode. (b): Same layer velocity component and cross layer velocity component at an arbitrary point in each layer as a function of $r$ for $\lambda = 1$, where $r$ is the distance between a grid point and a particular vortex. The upper blue curve is the same layer velocity magnitude, while the lower red curve is the cross layer velocity magnitude.
3.1.3 Experiment model setup

I start our experimental model from the case of two identical vortices in each layer with circulations $\Gamma^{(1)}_1 = \Gamma^{(1)}_2 = \Gamma^{(2)}_1 = \Gamma^{(2)}_2 = \Gamma = 2\pi$. The initial positions of vortices $x^{(1)}_k (k = 1, 2)$ are located at (-1,0) and (1,0) in the top layer and $x^{(2)}_k (k = 1, 2)$ are located at (0,-1) and (0,1) in the bottom layer. The radius of deformation between two layers $\lambda = 1$. In this case, the original equation is simplified with $N^{(1)}_v = 2$ and $N^{(2)}_v = 2$. This is a very special case which avoids the chaotic behavior of vortices. The streamfunction of equation and velocity field of equation can also be simplified in the same way. The velocity field at an arbitrary point is the sum of the velocity components induced by each of the four vortices as shown in equation. Since the vortices are holding the same strength of circulations and are located symmetrically initially, the vortices will rotate around a common center, i.e. the origin point in this particular setup, with a constant angular velocity in each layer, and the period for each individual vortex is 10.5112 computed from the velocity fields.
Figure 3.2: Streamfunction in the co-rotating frame. Two vortices are located at (-1,0), (1,0) initially in the top layer, and two vortices are located at (0,-1), (0,1) initially in the bottom layer.

The streamfunction of the flow in the corotating frame where the vortices are fixed at the initial positions is shown in Figure 3.2. In the corotating frame, the vortex points are fixed at the initial positions in both of the layers and the frame is
rotating with the same constant angular velocity as vortex does in the Eulerian frame. If I release a drifter at an initial position in the flow field, the drifter would follow the streamlines and its motion would be periodic. Three drifter launch locations in the top layer are chosen and the motion will follow the streamlines in the corotating frame. The trajectory of each drifter in the Eulerian frame is also shown in Figure 3.3. The figure shows that the trajectories of drifters are fairly complex and nonlinear in the Eulerian frame. Such a deterministic two layer system was used in the following sections to demonstrate assimilation process.

Figure 3.3: Trajectories of the top layer corresponding to the drifters in Figure 3.2 in the Eulerian frame. Two vortices are located at (-1,0), (1,0) initially in the top layer. Red star: initial location of two vortex points; blue star: initial location of drifter 1; black star: initial location of drifter 2; green star: initial location of drifter 3.
3.2 Assimilation of one drifter in the top layer with the Extended Kalman Filter

In this section, I will implement the Extended Kalman Filter scheme to the two layer point vortex system by assimilating the Lagrangian data in the top layer. The deterministic model I am using is the one setup in the last section. The dynamics of $\mathbf{x}^t$ can be written as a stochastic system as shown in equation (2.5). Without assimilation, the deterministic model loses track of the full system in both layers after some time $t$. The error in vortex positions grows in both of the layers.

I will employ the identical twin experiment that are used to provide both the model forecast $\mathbf{x}^f$ and the true state $\mathbf{x}^t$. To produce two different sets of the flows from our two layer point vortex system, our true state is a stochastic system with an unresolved small scale noise term as I mentioned before. While the forecast state is a deterministic system propagated with time. I will test whether our method can recover the true value of the vortex positions in both layers, by simply propagating the information from the drifter measurements in the top layer of point vortex system through the error covariance matrix to improve the prediction of the vortex positions in both layers.

To track the vortex positions for a long period in both layers, new information has to be injected into the model. Here the new information is Lagrangian data of the drifter I collected from the top layer. In all the experiments of this section, the initial condition of the drifter will be fixed at (-0.2, 1.2) in the top layer as numbered 1 in Figure 3.2. Since in the true system, an unresolved small scale noise is present, the trajectories of drifters are even more complex and hard to predict. The issue is that by tracking the drifter trajectories, I intend to estimate the vortex positions in both of the layers.
As shown in the paper by Kuznetsov et al. (2003), the drifter observations collected in the top layer can be used to improve the prediction of the state vectors in the top layer. In this multi-layer system, I assimilate the drifter observations in the top layer to demonstrate that it can also improve the prediction of the velocity field in the bottom layer which does not contain Lagrangian data. More specifically, for the two layer point vortex system, the observation is the drifter’s position I released into the top layer. It is the dynamical interaction of vortices between the layers that enables the drifter in the top layer to provide enough information to improve the prediction of the vortex positions in the other layer.

In this case, I found that one drifter in the top layer is enough to provide the necessary information to estimate vortices in two layers, given dynamical noise $\sigma = 0.02$ and observation error $\rho = 0.02$, under the assimilation time interval $\Delta T = 1$. Here I use the model with $(N^{(1)}_v, N^{(2)}_v, N^{(1)}_d, N^{(2)}_d) = (2, 2, 1, 0)$, i.e. four vortices in total and one drifter in the top layer. So by only observing one Lagrangian drifter in the top layer, I can track the model information in both of the layers. The vertical correlation of layer information can be tracked which means that the Lagrangian drifter in the top layer not only captures the model dynamics in the top layer but also captures the model dynamics in the bottom layer. As I mentioned in Section 3.1, the vortices in both layers induce a velocity flow field in this layered model. The velocity field of the drifter is a weighted combination of each vortex from each layer. The information of the vortex positions in the bottom layer propagated to the top layer can be captured by Lagrangian data. Or I can say that information propagated vertically in the water column can be captured by Lagrangian data. Moreover, by collecting the Lagrangian data in the top layer, the Eulerian velocity field in the bottom layer can be consistently assimilated through the vertical correlation of the layers.
Figure 3.4: The Extended Kalman Filter results in a two layer point vortex system, one drifter in the top layer is observed. \((N_v^{(1)}, N_v^{(2)}, N_d^{(1)}, N_d^{(2)}) = (2, 2, 1, 0), \Delta T = 1.0,\) dynamical noise and observation error \((\sigma, \rho) = (0.02, 0.02).\) Actual error in the vortex positions \(|x_{F,i} - x_{F,i}^*| (i = 1, 4)\) in both layers, in the model assimilating drifter positions (yellow and green), and in the model without assimilation (red and blue) in both layers.

Figure 3.4 shows the actual vortex position error norms of both layers in the deterministic model without assimilation and in the model that assimilates Lagrangian data in the top layer. Without assimilation, the errors grow in both of the layers. Assimilation of the drifter positions in the top layer affords simultaneously successful tracking of the vortices in both layers; the errors do not grow in both of the layers beyond tolerance 0.2 for the full length of the assimilation. The information collected by Lagrangian drifter can be propagated both horizontally in the same level and vertically through the water column. Moreover, the performance of vertical correlation is as good as the performance of horizontal correlation, since I can see that, the vortex position errors can be reduced to the same level in both layers simultaneously. It is an
essential feature for the Lagrangian data assimilation. The Lagrangian drifters can be used to improve the predictions of surface layer and sub-surface layer simultaneously.

A comparison of the predicted root-mean-square (rms) error from $P'$ with actual errors for a given noise realization in figure 3.5 shows a good agreement between the two in both of the layers.
Figure 3.5: Same assimilation experiment as in Figure 3.4. Predicted rms error in vortex location (yellow and green corresponding to Figure 3.4), vs actual error (blue) and the same for the drifter.

If I increase the noise realization to $\sigma=0.1$ and $\rho=0.1$, the assimilating model tracks the full system for awhile but fails after some time as shown in Figure 3.6, which means that the assimilating model is affected by the increase of noise. The
expected rms errors predicted by the variance equation do not correspond to actual
errors and remain misleadingly moderate. The failure of the filter was also observed
in the study of Kuznetsov et al. It is caused by the breakdown of the TLM. The
expected rms errors predicted by the error covariance matrix do not correspond to
actual errors and remain misleadingly moderate. The malfunction of the filter is
cau by the breakdown of the TLM. Our scheme would work properly only when
the noise level $\sigma$ and $\rho$ are moderate so that $\Delta x$ remains small enough and the TLM
remains a good approximation.
Figure 3.6: The Extended Kalman Filter results in a two layer point vortex system, one drifter in the top layer is observed. \((N_v^{(1)}, N_v^{(2)}, N_d^{(1)}, N_d^{(2)}) = (2, 2, 1, 0), \Delta T = 1.0,\) dynamical noise and observation error \((\sigma, \rho) = (0.1, 0.1).\) Predicted rms error in vortex location (yellow and green) vs actual error (blue) and the same for the drifter. The system loses track after a certain time.

To analyze the influence of the system parameters I have performed 100 experi-
ments with different noise realizations for each parameter set. The mean of the actual error in the vortex positions $|x_{F,i}^f - x_{F,i}^t| (i = 1, 4)$ of two layers are listed in Table 3.1. It summarizes how much the noise level $\sigma$ and $\rho$, and observation interval $\Delta T$ can influence the results of the Extended Kalman Filter scheme. From the table, the dependence on dynamical noise $\sigma$ is considerably stronger than the dependence on the observation error $\rho$. This is reasonable because the observations are only introduced every time interval $\Delta T=1$ while the dynamical noise is evolved through the system all the time. The scheme can also be affected by the observation time interval $\Delta T$. With the increase of $\Delta T$, which means the drifter positions are observed less often, the scheme results in a relatively larger error in vortices tracking.

Table 3.1: The average error of vortex positions $|x_{F,i}^f - x_{F,i}^t| (i = 1, 4)$ at time $t=60$, (results of 100 experiments with different noise realizations), under the different noise level and observation intervals for each parameter set.

<table>
<thead>
<tr>
<th>$\Delta x$</th>
<th>$\Delta T=0.5$</th>
<th>$\Delta T=1$</th>
<th>$\Delta T=1.5$</th>
<th>$\Delta T=2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\sigma, \rho)=(0.02,0.02)$</td>
<td>0.01</td>
<td>0.03</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>$(\sigma, \rho)=(0.02,0.05)$</td>
<td>0.01</td>
<td>0.04</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>$(\sigma, \rho)=(0.05,0.02)$</td>
<td>0.02</td>
<td>0.15</td>
<td>0.25</td>
<td>0.57</td>
</tr>
<tr>
<td>$(\sigma, \rho)=(0.05,0.05)$</td>
<td>0.18</td>
<td>0.53</td>
<td>1.03</td>
<td>1.59</td>
</tr>
</tbody>
</table>
Chapter 4

Lagrangian Data Assimilation into Two and a Half Layer Shallow Water Double Gyre System

In this chapter, I adapt the Ensemble Kalman Filter method for the assimilation of Lagrangian data from the top layer into the layered ocean model. The method is tested by using a set of identical twin experiments on the two and half layer reduced gravity shallow water model with double gyre flow configuration. The use of a layered shallow water model is more challenging than a layered point vortex system as the dimension of the underlying system becomes extremely large. The Extended Kalman Filter can be used for nonlinear systems, but the computational demand resulting from the error covariance matrix limits the size of the problem. I employ the Ensemble Kalman Filter to avoid the difficulty of dealing with such a large number of unknowns that would be needed to evolve the error covariance matrix were a TLM used. The Ensemble Kalman Filter takes advantage of using an ensemble size of $O(100)$ members to represent the ensemble error covariance matrix. The Ensemble
Kalman Filter circumvents the expensive integration of the error covariance matrix by propagating an ensemble of state vector from which the required error covariance matrix information is only obtained at each assimilation time step.

4.1 Two and a half layer reduced gravity shallow water double gyre flow model

The dynamical behavior of the midlatitude upper ocean is often handled by a reduced gravity system, where one or several active layers lie above a motionless layer of infinite depth. Thus, an \( n + \frac{1}{2} \) model describes an upper ocean with \( n \) distinct active layers; the half layer refers to the infinitely deep, quiescent layer. I consider an idealized ocean model with two and a half layers in a square domain configuration. The reduced gravity shallow water model is a comprehensive mesoscale ocean model, which has two active layers and a motionless layer in the bottom. The shallow water model we chose as our basic flow description is widely used in oceanographic modeling and its main assumption is that the ocean behaves as having different layers where density is constant in each layer. Here a reduced gravity version of this layered model is preferred. The flow model is driven by a zonal wind forcing term in the surface layer. Friction is present at each of the layer interfaces.

Neglecting thermodynamics, I consider the reduced gravity shallow water layer model was driven by a steady wind stress in a closed rectangular basin \( \Omega = \{0 \leq x \leq L, 0 \leq y \leq D\} \). The lengths of the domain in the zonal direction and meridional direction are denoted by \( L \) and \( D \). The layer model consists of two active layers of constant density \( \rho_1 \) and \( \rho_2 \) in each layer overlying an infinitely deep and motionless layer of constant density \( \rho_3 \). The \( \beta \) plane approximation (Gill 1982; Pedlosky 1987), was used to describe the midlatitude oceans in Cartesian coordinates, where the
Coriolis parameter $f$ is given by $f = f_0 + \beta_0 y$.

Vertically integrating the Navier-Stokes equations over the three layers and using the hydrostatic assumption for each layer, I derive the following governing partial differential equations in the reduced gravity configuration as shown in Simonnet (2002):

\[
\begin{align*}
\frac{\partial U_1}{\partial t} + \nabla \cdot (U_1 V_1) &= -g_1 h_1 \frac{\partial h_1}{\partial x} - g_2 h_1 \frac{\partial h_2}{\partial x} + (f_0 + \beta_0 y)V_1 \\
&+ A \triangle U_1 - R_1(U_1 - U_2) + \frac{\tau^x}{\rho_1} \\
\frac{\partial V_1}{\partial t} + \nabla \cdot (V_1 U_1) &= -g_1 h_1 \frac{\partial h_1}{\partial y} - g_2 h_1 \frac{\partial h_2}{\partial y} - (f_0 + \beta_0 y)U_1 \\
&+ A \triangle V_1 - R_1(V_1 - V_2) \\
\frac{\partial U_2}{\partial t} + \nabla \cdot (U_2 V_2) &= -g_2 \rho_1 h_2 \frac{\partial h_1}{\partial x} - g_2 h_2 \frac{\partial h_2}{\partial x} + (f_0 + \beta_0 y)V_2 \\
&+ A \triangle U_2 - R_2 U_2 + R_1(U_1 - U_2) \\
\frac{\partial V_2}{\partial t} + \nabla \cdot (U_2 V_2) &= -g_2 \rho_2 h_2 \frac{\partial h_1}{\partial y} - g_2 h_2 \frac{\partial h_2}{\partial y} - (f_0 + \beta_0 y)U_2 \\
&+ A \triangle V_2 - R_2 V_2 + R_1(V_1 - V_2) \\
\frac{\partial h_1}{\partial t} &= -\frac{\partial U_1}{\partial x} - \frac{\partial V_1}{\partial y} \\
\frac{\partial h_2}{\partial t} &= -\frac{\partial U_2}{\partial x} - \frac{\partial V_2}{\partial y}
\end{align*}
\]

Here:

\[
\begin{align*}
V_l &= (u_l i + v_l j) \\
U_l &= h_l u_l \\
V_l &= h_l v_l
\end{align*}
\]

\((U_l, V_l)\) is the mass-flux vector in each active layer, where \((u_l, v_l), l = 1, 2,\) represent the eastward and northward components respectively, and \(l\) is the layer number.
While $h_1$, $h_2$ are the top and second layer thicknesses. The two reduced gravities are:

$$g_1 = g(\rho_3 - \rho_1)/\rho_3$$

$$g_2 = g(\rho_3 - \rho_2)/\rho_3$$

(4.2)

The quantity $\rho_l$ is the layer density, $g$ is the gravitational acceleration. $\tau^x$ is the zonal wind forcing term.

The right hand side of the momentum equation consists of the pressure gradient terms contributed from both of the active layers; the Coriolis force term in the $\beta$ plane approximation; the diffusion term; and the friction term between each layer interface. A zonal wind forcing term is also shown in the momentum equation of the top layer. The interface friction between the top active layer and the second active layer is represented as $R_1$, and the interface friction between the second active layer and the bottom friction is represented as $R_2$. The interface friction is proportional to the differences of the velocities. There are three versions of the dissipation terms frequently used in ocean models (Hurlburt and Thompson 1980; Bleck and Boudra 1981; Killworth et al. 1991; Oberhuber 1993):

$$Ah\Delta u, Ah\Delta v$$

$$\Delta(Ah\Delta u), \Delta(Ah\Delta v)$$

$$A\Delta U, A\Delta V$$

(4.3)

where $\Delta$ is the two-dimensional Laplace operator of a two-dimensional vector. Here I will adopt the dissipation term used in Simonnet (2002); $A\Delta U, A\Delta V$ to represent the diffusion term.

The above momentum equations are written in conservative form, but I can also transform them into primitive form by subtracting the continuity equations and then
dividing both sides of the equations by $h_i$. 

$$\frac{\partial u_1}{\partial t} + u_1 \cdot \nabla V_1 = -g_1 \frac{\partial h_1}{\partial x} - g_2 \frac{\partial h_2}{\partial x} + \left(f_0 + \beta_0 y\right)u_1$$

$$+ \frac{A}{h_1} \left(\frac{\partial^2 (h_1 u_1)}{\partial x^2} + \frac{\partial^2 (h_1 u_1)}{\partial y^2}\right) - R_1 \left(u_1 - \frac{h_2 u_2}{h_1}\right) + \frac{\tau_x}{\rho_1 h_1}$$

$$\frac{\partial v_1}{\partial t} + v_1 \cdot \nabla V_1 = -g_1 \frac{\partial h_1}{\partial y} - g_2 \frac{\partial h_2}{\partial y} - \left(f_0 + \beta_0 y\right)u_1$$

$$+ \frac{A}{h_1} \left(\frac{\partial^2 (h_1 v_1)}{\partial x^2} + \frac{\partial^2 (h_1 v_1)}{\partial y^2}\right) - R_1 \left(v_1 - \frac{h_2 v_2}{h_1}\right)$$

$$\frac{\partial u_2}{\partial t} + u_2 \cdot \nabla V_2 = -g_2 \frac{\rho_1 \partial h_1}{\rho_2 \partial x} - g_2 \frac{\partial h_2}{\partial x} + \left(f_0 + \beta_0 y\right)v_2$$

$$+ \frac{A}{h_2} \left(\frac{\partial^2 (h_2 u_2)}{\partial x^2} + \frac{\partial^2 (h_2 u_2)}{\partial y^2}\right) - R_2 u_2 + R_1 \left(\frac{h_1 u_1}{h_2} - u_2\right)$$

$$\frac{\partial v_2}{\partial t} + v_2 \cdot \nabla V_2 = -g_2 \frac{\rho_1 \partial h_1}{\rho_2 \partial y} - g_2 \frac{\partial h_2}{\partial y} - \left(f_0 + \beta_0 y\right)v_2$$

$$+ \frac{A}{h_2} \left(\frac{\partial^2 (h_2 u_2)}{\partial x^2} + \frac{\partial^2 (h_2 u_2)}{\partial y^2}\right) - R_2 v_2 + R_1 \left(\frac{h_1 v_1}{h_2} - v_2\right)$$

$$\frac{\partial h_1}{\partial t} = -\frac{\partial (u_1 h_1)}{\partial x} - \frac{\partial (v_1 h_1)}{\partial y}$$

$$\frac{\partial h_2}{\partial t} = -\frac{\partial (u_2 h_2)}{\partial x} - \frac{\partial (v_2 h_2)}{\partial y}$$

No normal flow and no slip conditions are applied along all the boundaries. The boundary conditions and initial conditions of each layer are given by

$$u_l(x, y, t)|_{\partial \Omega} = 0$$

$$v_l(x, y, t)|_{\partial \Omega} = 0$$

$$\frac{\partial u_l(x, y, t)}{\partial n}|_{\partial \Omega} = 0$$

$$\frac{\partial v_l(x, y, t)}{\partial n}|_{\partial \Omega} = 0$$

$$u_l(x, y, 0) = 0$$

$$v_l(x, y, 0) = 0$$

$$h_l(x, y, 0) = H_0$$

(4.4)
4.2 Configuration of the mathematical model

The parameter values used here are taken largely from Simonnet (2002). The ocean model is set up in an idealized, mesoscale configuration. The model is configured in a square domain of 2000 km × 2000 km in each layer. The horizontal eddy diffusion is represented by the coefficient $A = 400 \text{ m}^2 \text{ s}^{-1}$. The interface friction between the top active layer and the second active layer $R_1$, and the interface friction between the second active layer and the bottom friction $R_2$, are both of Rayleigh type $R_1 = R_2 = 5 \times 10^{-8} \text{ s}^{-1}$.

Standard midlatitude values for the Coriolis parameter are used: $f_0 = 5 \times 10^{-5} \text{ s}^{-1}$, $\beta_0 = 2 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$. The top layer is driven by a steady, zonal wind forcing term with a sinusoidal profile:

$$\tau^x(y) = -\frac{\tau_0}{H_0}\cos(2\pi y/D) \quad (4.5)$$

where $\tau_0$ is the wind stress, $\rho_0$ is the density of water. This leads to the formation of a double gyre circulation in both of the active layers, western boundary currents, mid-latitude jet, and mesoscale eddies (Verron and Leprovost 1978; Miller et al. 1987; Jiang et al. 1995). The assigned parameters are given in Table 4.1.
Table 4.1: Reference values for the model’s parameters.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basin size ((L, D))</td>
<td>(2000 km, 2000 km)</td>
</tr>
<tr>
<td>Layer thickness ((H_1, H_2))</td>
<td>(500 m, 500 m)</td>
</tr>
<tr>
<td>Coriolis parameter (f_0)</td>
<td>(5 \times 10^{-5} \text{ s}^{-1})</td>
</tr>
<tr>
<td>(\beta_0)</td>
<td>(2 \times 10^{-11} \text{ m}^{-1} \text{s}^{-1})</td>
</tr>
<tr>
<td>Eddy viscosity (A)</td>
<td>400 m² s⁻¹</td>
</tr>
<tr>
<td>Layer frictions (R_1)</td>
<td>(5 \times 10^{-8} \text{ s}^{-1})</td>
</tr>
<tr>
<td>Layer frictions (R_2)</td>
<td>(5 \times 10^{-8} \text{ s}^{-1})</td>
</tr>
<tr>
<td>Top layer density (\rho_1)</td>
<td>1025 kg m⁻³</td>
</tr>
<tr>
<td>Second layer density (\rho_2)</td>
<td>1028 kg m⁻³</td>
</tr>
<tr>
<td>Bottom layer density (\rho_3)</td>
<td>1029 kg m⁻³</td>
</tr>
<tr>
<td>Reduced gravity (g_1)</td>
<td>(4 \times 10^{-2} \text{ m} \text{s}^{-2})</td>
</tr>
<tr>
<td>Reduced gravity (g_2)</td>
<td>(8 \times 10^{-3} \text{ m} \text{s}^{-2})</td>
</tr>
<tr>
<td>Horizontal grid scale ((\Delta x, \Delta y))</td>
<td>(20 km, 20 km)</td>
</tr>
<tr>
<td>Time step (\Delta t)</td>
<td>12 minutes</td>
</tr>
</tbody>
</table>

The flow equations are discretized on an \((n_x \times n_y) = (100 \times 100)\) grid in each active layer, so that each cell size is \((\Delta x, \Delta y) = (20 \text{ km, 20 km})\). The flow state vector then has the size of \(N = 2(n_x - 1)(n_y - 1) + n_x n_y\) in each layer considering the velocity field and height field. Note that the velocity vectors are stored at cell nodes whereas the height field is stored at cell centers. Thus the total size of the state vector \(x_F\) is \(2N = 4(n_x - 1)(n_y - 1) + 2n_x n_y\). It can be seen as an extension of the C-grid staggering by Arakawa and Lamb (1979). Note that the velocity field at domain boundaries are not included in the state vector. The flow state vector \(x_F\) is given by \(x_F = (u_{1,1}^{(1)}(t), ..., u_{nx-1,ny-1}^{(1)}(t), v_{1,1}^{(1)}(t), ..., v_{nx-1,ny-1}^{(1)}(t), h_{1,1}^{(1)}(t), ..., h_{nx,ny}^{(1)}(t), u_{1,1}^{(2)}(t), ..., u_{nx-1,ny-1}^{(2)}(t), v_{1,1}^{(2)}(t), ..., v_{nx-1,ny-1}^{(2)}(t), h_{1,1}^{(2)}(t), ..., h_{nx,ny}^{(2)}(t))\).
4.3 Geostrophic balance

4.3.1 Pressure gradients in multi-layer models

In this section I derive the pressure gradient terms in each active layer shown in the governing equation of the layered shallow water double gyre flow model. For most of the derivation of the governing equations, each layer can be treated separately, in the manner used to deal with the single layer case. It is the horizontal pressure gradient term that couples the layers together. Since the layered model is dynamically coupled, it has the feasibility of the assimilation of Lagrangian data observation from the top active layer into the sub-surface of the layered flow model.

I can start the derivation from the (sea) surface $z = \eta$, and integrate the hydrostatic relation downward $\frac{\partial P}{\partial z} = \rho g$, which assumes an exact equilibrium in the vertical between the pressure gradient force and the gravitational force. The pressure in the top, second and bottom layers, respectively are:

\[
\begin{align*}
p_1 & = p_a - \rho_1 g(z - \eta) \\
p_2 & = p_a + \rho_1 g(\eta + h_1) - \rho_2 g(z + h_1) \\
p_3 & = p_a + \rho_1 g(\eta + h_1) + \rho_2 g h_2 - \rho_3 g(z + h_1 + h_2)
\end{align*}
\]

where $p_a = \text{const}$ is the atmospheric pressure.

The pressure gradient in each layer can be written as:

\[
\begin{align*}
\nabla p_1 & = \rho_1 g \nabla \eta \\
\nabla p_2 & = \rho_1 g \nabla \eta - (\rho_2 - \rho_1) g \nabla h_1 \\
\nabla p_3 & = \rho_1 g \nabla \eta - (\rho_3 - \rho_1) g \nabla h_1 - (\rho_3 - \rho_2) g \nabla h_2
\end{align*}
\]
Assuming the lowest layer is infinitely deep, so that the pressure gradient in the lowest layer is negligible. If we assume \( \nabla p_3 = 0 \), \( \nabla \eta \) can be written in term of \( \nabla h_1 \) and \( \nabla h_2 \).

\[
\begin{align*}
\frac{1}{\rho_1} \nabla p_1 &= g_1 \nabla h_1 + g_2 \nabla h_2 \\
\frac{1}{\rho_2} \nabla p_2 &= g_2 \frac{\rho_1}{\rho_2} \nabla h_1 + g_2 \nabla h_2
\end{align*}
\]

(4.8)

Where \( g_1 \) and \( g_2 \) are reduced gravities in the top layer and the second layer:

\[
\begin{align*}
g_1 &= \frac{g(\rho_3 - \rho_1)}{\rho_3} \\
g_2 &= \frac{g(\rho_3 - \rho_2)}{\rho_3}
\end{align*}
\]

(4.9)

So these are exactly the pressure gradient terms shown in the governing equations. The presence of \( \nabla h_1 \) and \( \nabla h_2 \) in each layer express the strong coupling of the layers. It makes the assimilation of Lagrangian data from the top active layer into both of the active layers feasible.

### 4.3.2 Geostrophic balance relation

If the flow is such that the Rossby number \( R_0 \) is small, where \( R_0 = \frac{U}{fL} \) (here \( U \) is a typical horizontal current speed, \( f \) is the Coriolis parameter and \( L \) is a typical horizontal scale over which \( U \) varies), then the Coriolis force is balanced by the pressure gradient force in the horizontal component of the momentum equation. One interesting phenomenon about the two and half layer model is that geostrophic balance has a different form in the second active layers because of the dynamical coupling.

In the pressure gradient terms shown up in the second active layer, the ratio of the reduced densities of two active layers can be approximated as 1, i.e. \( \frac{\rho_1}{\rho_2} = 1 \). So the top layer pressure gradient term \( g_2 \frac{\rho_1}{\rho_2} \nabla h_1 \) and the second layer pressure gradient term
\( g_2 \nabla h_2 \), are basically of the same order. The gradient of the height field in the second layer \( g_2 \nabla h_2 \) no longer balances the Coriolis force by itself. Instead, it is a weighted combination of the top layer pressure force and the second layer pressure force that balance the Coriolis force term \( (f_0 + \beta_0 y) V_2 \). In this sense, the streamlines in the second layer do not reflect the height field of second layer uniquely, it is actually a reflection of both layers’ influences.

The geostrophic balance relation is thus generalized to a relation describing different behavior at layered models. The generalization to the geostrophic balance relation used in the second active layer is different from the usual geostrophic balance. The dynamics of the surface layer are essentially the dynamics of a shallow water flow with depth equal to the equivalent depth calculated above. The generalization used in the second active layer is, however dependent on both active layers. The dynamics of the second layer are essentially the dynamics of a combination of the surface layer and the second layer. Under these circumstances, the surface layer has a significant effect on the results of the sub-surface layers as long as the thickness of the sub-surface layer is at the same order as the thickness of the surface layer. This co-balance happening in the second active layer provides us the feasibility of assimilating Lagrangian drifter in the top active layer and improving the prediction of the flow field in the second active layer.

### 4.4 Computational Lagrangian trajectory

Drifters in the two and a half layer model were released in a specific model layer (i.e. in the top layer) and always remained within the layer. The track of drifters are approximated during the model run by using interpolation in space and integration in time schemes. Interpolation in space is performed using a sixteen-point grid box
surrounding the drifter. Bicubic interpolation in space as a higher order smoothness method is used if a sufficient number of good grid points are available. The order of the interpolation in space is reduced to bilinear when the float is surrounded by four grid points. Linear interpolation in time is performed to construct the intermediate time stage velocity field and the integration with a fourth order Runge-Kutta method is used.

Bicubic interpolation (Press et al., 1992) estimates the value at a position in the destination drifter trajectory by an average of sixteen grid points surrounding the closest corresponding grid. To compute the velocity of the drifter \( (u_D(t_k), v_D(t_k)) \) at position \( (x_D(t_k), y_D(t_k)) \) at time \( t_k \) in a certain grid cell such as:

\[
\frac{dx_D(t_k)}{dt} = u_D(t_k) \\
\frac{dy_D(t_k)}{dt} = v_D(t_k)
\]

(4.10)

Four velocities in the nearest rectangular grids are needed. The grid points are \( (x(i), y(i)) \), \( (x(i + 1), y(i)) \), \( (x(i), y(i + 1)) \) and \( (x(i + 1), y(i + 1)) \) respectively.

\[
x(i) \leq x_D(t_k) \leq x(i + 1) \\
y(j) \leq y_D(t_k) \leq y(j + 1)
\]

(4.11)

The bicubic formulas have the form:

\[
u_D(t_k) = \sum_{i=1}^{4} \sum_{j=1}^{4} c_u(i, j) p^{i-1} q^{j-1}
\]

\[
v_D(t_k) = \sum_{i=1}^{4} \sum_{j=1}^{4} c_v(i, j) p^{i-1} q^{j-1}
\]

(4.12)
where

\[
p^i = \frac{x_D(t_k) - x(i)}{x(i + 1) - x(i)}
\]

\[
q^j = \frac{y_D(t_k) - y(i)}{y(i + 1) - y(i)}
\]  \hspace{1cm} (4.13)

c_u(1 : 4, 1 : 4) and c_v(1 : 4, 1 : 4) are computed by the gradients and the cross derivatives at all of these four grid points. So the velocity of drifter is computed as a weighted average of the nearest sixteen grid points in a rectangular grid (a $4 \times 4$ array). Here, two cubic interpolation polynomials, one for each horizontal direction, are used.

The drifter position is advected using a fourth-order Runge-Kutta method. Linear interpolation in time is needed when computing the intermediate stages in the Runge-Kutta method. At each intermediate stage $t = t_{k+\frac{1}{2}}$, the velocity field is computed as the mean of the velocity field at time $t = t_k$ and time $t = t_{k+1}$ in each direction at each grid point.

\[
\begin{align*}
    u(i, j; t_{k+\frac{1}{2}}) &= \frac{1}{2}u(i, j; t_k) + \frac{1}{2}u(i, j; t_{k+1}) \\
    v(i, j; t_{k+\frac{1}{2}}) &= \frac{1}{2}v(i, j; t_k) + \frac{1}{2}v(i, j; t_{k+1})
\end{align*}
\]  \hspace{1cm} (4.14)

At each intermediate stage, the bicubic interpolation will also be implemented to compute the velocity at drifter position.
4.5 Assimilation of Lagrangian drifter in the top active layer with the Ensemble Kalman Filter

4.5.1 Identical twin experiment

In this setup, I will employ the identical twin experiment in which the same flow and drifter equations are used to provide both the model forecast $x^f$ and the true state $x^t$. To produce two different sets of the flows from our layered reduced gravity shallow water model, I employ a different initial condition for the averaged water depth $H_1$ and $H_2$ into the two systems, the velocity fields will also be perturbed along with the perturbation of the height field. I will test whether our method can recover the true value of the height fields, $H_1$ and $H_2$, and the true value of the velocity field, by simply propagating the information from the drifter measurements in the top layer through the error covariance matrix to improve the prediction of the height fields and velocity fields in both of the active layers.

I have set the initial heights $H_1 = 500$ m and $H_2 = 500$ m for the true system in both of the layers. I initially integrated the system for a period of 10 years to establish a fully developed flow. The initial run produces one true flow and a set of ensembles of our model forecast. The ensembles are considered a 10% error in the mean value of the model in both layers. The ensembles are generated by the perturbation of the initial height field. I used a Gaussian distribution with a variance of $\sigma_h = 50$ m for both layers to generate $H_1$ and $H_2$ for the members of ensembles.

After 10 years of spinning up, drifters are released into the true system in the top layer. Each drifter is initialized at a specified location. The observations are taken to be distributed by an independent Gaussian white noise with the same statistics with a variance of 200 m. The drifters are integrated with the true flow to generate
a set of true trajectories to be used in the assimilation. The drifter observations are coming from drifter positions at specified discrete time intervals.

After obtaining the drifter observations, I then integrate each ensemble over the same time period. I update each ensemble member when drifter observations come into the system at each specified time interval. The drifters are also perturbed with the same error statistics as the ensemble distributions. The analysis state of flow field is the mean of the updated ensembles.

The errors for the kinetic energy, height field, and drifter separation distances are computed to quantify the performance of the method. The number of drifters, their locations, their releasing layers, assimilation time intervals, as well as the number of ensembles can all affect the performance of our scheme. For simplicity, I will refer to the region \( \{(x, y)|0 \text{ km} < x < 1000 \text{ km}, 500 \text{ km} < y < 1500 \text{ km}\} \) as the energetic region where I will mainly deploy the drifters. And I will generate \( O(80) \) ensemble members for the assimilation. In our experiment, 36 drifters will be distributed uniformly in \((6 \times 6)\) arrays in this energetic region.

The measurements of errors in velocity field, height field, and drifter locations are defined as the same in the paper of Salman et al. (2006) in both of the active layers. The true errors for the kinetic energy, height field, and drifter separation distances in the active layers are given by \((l = 1, 2 \text{ is the layer number.})\)

\[
|KE|^l_t = \sqrt{\frac{\sum_{i=1}^{n_x-1} \sum_{j=1}^{n_y-1} (u_{ij}^l - u_{i,j}^t)^2 + (v_{ij}^l - v_{i,j}^t)^2}{\sum_{i=1}^{n_x-1} \sum_{j=1}^{n_y-1} (u_{ij}^l)^2 + (v_{ij}^l)^2}} 
\]

\[
|h|^l_t = \sqrt{\frac{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} (h_{ij}^l - h_{i,j}^t)^2}{\sum_{i=1}^{n_x} \sum_{j=1}^{n_y} (h_{ij}^l)^2}} 
\]

\[
|x_D|^l = \sqrt{\frac{\sum_{i=1}^{N_D} (x_{D,i}^l - x_{D,i}^t)^2 + (y_{D,i}^l - y_{D,i}^t)^2}{N_D}} 
\] (4.15)
The predicted errors for the kinetic energy, height field, and drifter separation distances in the active layers are given by

\[
|KE|_l^p = \sqrt{ \sum_{k=1}^{N_E} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left[ (u_{l,ij}^f - (u_{l,k,ij}^f))^2 + (v_{l,ij}^f - (v_{l,k,ij}^f))^2 \right]/(N_E - 1) } \\
|h|_l^p = \sqrt{ \sum_{k=1}^{N_E} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left[ (h_{l,ij}^f - (h_{l,k,ij}^f))^2 \right]/(N_E - 1) } \\
|\mathbf{x}_D|^p = \sqrt{ \sum_{k=1}^{N_E} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left[ (x_{D,i}^f - (x_{D,k,i}^f))^2 + (y_{D,i}^f - (y_{D,k,i}^f))^2 \right]/N_D(N_E - 1) } \quad (4.16)
\]

where

\[
\begin{align*}
\overline{u_{l,ij}^f} & = \frac{1}{N_E} \sum_{k=1}^{N_E} (u_{l,k,ij}^f) \\
\overline{v_{l,ij}^f} & = \frac{1}{N_E} \sum_{k=1}^{N_E} (v_{l,k,ij}^f) \\
\overline{h_{l,ij}^f} & = \frac{1}{N_E} \sum_{k=1}^{N_E} (h_{l,k,ij}^f) \\
\overline{x_{D,i}^f} & = \frac{1}{N_E} \sum_{k=1}^{N_E} (x_{D,k,i}^f) \\
\overline{y_{D,i}^f} & = \frac{1}{N_E} \sum_{k=1}^{N_E} (y_{D,k,i}^f)
\end{align*}
\quad (4.17)
\]

\((\overline{u_{l,ij}^f}, \overline{v_{l,ij}^f})\) are the mean of the ensembles for the velocity field; and \(\overline{h_{l,ij}^f}\) is the mean of the ensembles for the height field; and \((\overline{x_{D,i}^f}, \overline{y_{D,i}^f})\) is the mean of the ensembles for the drifter position.
4.6 Results

I begin with a detailed analysis of our result by setting the assimilation time step to 1 day, and by releasing a set of $6 \times 6$ drifters in the energetic region of the flow in the top layer I defined before as $\{(x, y)|0 \text{ km} < x < 1000 \text{ km}, 500 \text{ km} < y < 1500 \text{ km}\}$. So I will assimilate 36 flow drifters in the top active layer every day to update the velocity field and height field in both of the active layers. Please note that our Lagrangian data are completely coming from the top active layer, while no Lagrangian observations are obtained from the second layer. The reason I release the drifters into this particular region was motivated by the findings of Kuznetsov et al. (2002) and Molcard et al. (2003) who identified the importance of sampling the complex spatial-temporal structure of the flow for the assimilation to be successful.

Contour results for the height in the top active layer $H_1$ are given in Figure 4.1, Figure 4.2 and Figure 4.3 at two different times. Figure 4.1 is the true flow of the top layer which I generated from the identical twin experiment, with an initial height field of $H_1 = 500$ m. Figure 4.3 is the free run flow of the top layer, with a Gaussian distribution variance of $\sigma_h = 50$ m of height field for the different members of ensemble generated initially to perturb the flow field in both of the active layers, which is called free run without assimilation. Figure 4.2 is the assimilation run of the top layer, which has the same initial condition with the free run top layer flow for each different member of ensembles. So that the flow is perturbed initially to differ from the true top layer flow. While it will be corrected by assimilating Lagrangian observations in the top layer every time step, here as 1 day. In each figure, there are two different sets of plots included: (a) is the height field at day 0 measured from the beginning of the assimilation cycle, which is 10 years after spinning up; (b) is the height field measured at 146 days after the beginning of the assimilation cycle. The results are based on the mean of the updated ensembles.
At the beginning of the assimilation process, which is day 0, I note that the height field of the true flow in Figure 4.1(a) differs from the other two flows in both of the layers, i.e. non-assimilated flow in Figure 4.3(a) and the assimilated flow in Figure 4.2(a). The difference is caused by the perturbation I generated initially in both of the layers for the non-assimilated flow. The non-assimilated flow and the assimilated flow are identical since the assimilation process has not yet started. And the appearance of the non-assimilated flow and the assimilated flow are computed as the mean of the ensembles.

After 146 days, there are 146 assimilation steps that performed since the assimilation step time is 1 day. The assimilation case now resembles the true flow more closely than the non-assimilated case in the top active layer. The contours of height field between the true flow and the assimilated flow are essentially identical especially in the energetic region. While the non-assimilated flow are significantly different from the true flow. Initially, both Figure 4.2(a) and Figure 4.3(a) are identical and distinct from Figure 4.1(a). However, as the assimilation of Lagrangian data into the top active layer is performed every day, the assimilated flow in Figure 4.2(b) gradually evolves to produce flow structures close to the true flow system in Figure 4.1(b). The non-assimilated flow in Figure 4.3(b), on the other hand, is very different from the true flow.

Figure 4.4 to Figure 4.6 depicted the flow structures of the second active layer where no Lagrangian drifters are released at all. By only assimilating the Lagrangian drifters from the upper layer, the contours of height field between the true flow in Figure 4.4(b) and the assimilated flow in Figure 4.5(b) are also essentially identical especially in the energetic region in this second layer. While the non-assimilated flow in Figure 4.6(b) is significantly different from the true flow. It shows that Lagrangian drifter can be used to correct the sub-surface flow structures where more complex
flow dynamics is present.

Figure 4.1: Contours of height field in the top active layer corresponding to the true flow. (a): day 0; (b): day 146.
Figure 4.2: Contours of height field in the top active layer corresponding to the assimilated flow. (a): day 0; (b): day 146.
Figure 4.3: Contours of height field in the top active layer corresponding to the non-assimilated flow. (a): day 0; (b): day 146.
Figure 4.4: Contours of height field in the second active layer corresponding to the true flow. (a): day 0; (b): day 146.
Figure 4.5: Contours of height field in the second active layer corresponding to the assimilated flow. (a): day 0; (b): day 146.
Figure 4.6: Contours of height field in the second active layer corresponding to the non-assimilated flow. (a): day 0; (b): day 146.

To provide a more quantitative measure of the convergence of our assimilation experiments, I have computed the error norms defined in terms of the kinetic energy, height field, and drifter separation distances as defined before, for both of the active
layers. I also include the case without assimilation to reflect the relative improvement in assimilation process. Figure 4.7 is the measurement in terms of the kinetic energy in both of the active layers and Figure 4.8 is the measurement in terms of the height field in both of the active layers. Please note that the true errors, $|KE|_t$ and $|h|_t$, ($l = 1, 2$), are computed from the difference between the true flow state and the mean of the updated ensembles in the assimilated flow. The predicted errors, $|KE|_p$ and $|h|_p$ are computed from the difference between the mean of the ensembles and each ensemble member I generated in the assimilated flow. They are the errors that the Ensemble Kalman Filter reduces.

I note that within a time-scale of around 200 days, both the true errors and the predicted errors in kinetic energy and height field have almost been reduced drastically in both of the active layers with assimilation. Both the true errors and the predicted errors have essentially the same time-history. It is also shown as that the errors are reduced simultaneously in the two active layers. The Lagrangian drifters can be used to improve the predictions of surface layer and sub-surface layer simultaneously. If I compare Figure 4.8(a) and Figure 4.8(b), which are the errors of height field in the top layer and second layer respectively, I note that the errors in the second active layer are undergoing a larger oscillation than the errors in the top active layer during the assimilation process since the Lagrangian observations are coming from the top layer. However, despite the oscillations of the errors in the second layer, the assimilation process still results in a decay of an error around 4% remains in the height field errors at the end of 1 year.
Figure 4.7: Error in kinetic energy $|KE|$(%). (a): top active layer; (b): second active layer.
Figure 4.8: Error in height field $|h|$(%). (a): top active layer; (b): second active layer.

Figure 4.9 is the measurement in terms of the drifter separation distances in the top active layer where the Lagrangian drifters are released. The predicted errors and the true errors in the assimilated run are at the order of the error variance used in our Ensemble Kalman Filter indicating successful tracking of the drifter trajectories.
along the entire assimilation process.

To illustrate how our method successfully tracks the true Lagrangian drifter trajectories, and reduces the error caused by the free run forecast, Figure 4.10 shows the assimilation process at time interval [70 days, 90 days] for four drifters I chose randomly in the top active layer. Each blue line then represents the assimilation update process and the each red line represents the forecast evolution process. For example, an assimilation update step is performed at day $t_k$, $(70 < t_k < 90)$, reducing the error between the forecast and the true system. The error does reduce from $x_D^f(t_k) - x_D^t(t_k)$ to $x_D^a(t_k) - x_D^t(t_k)$ by assimilating the Lagrangian drifter observations in the top layer at day $t_k$. Immediately at the following time step $t_{k+1}$, the model forecast leads to an increase in the error of $x_D^f(t_{k+1}) - x_D^t(t_{k+1})$. Consequently, the assimilation update process at $(t_{k+1})$ successfully reduces the error to $x_D^a(t_{k+1}) - x_D^t(t_{k+1})$ by assimilating the Lagrangian drifter observations in the top layer at day $t_{k+1}$.
Figure 4.10: Error in drifter positions in time interval $70 < t_k < 90$. Red line: forecast process; Blue line: assimilation process.
4.7 Covariance localization

The background error covariance matrix of the Ensemble Kalman Filter is estimated fully from the ensembles. Under certain circumstances, this finite ensemble size can degrade the performance of the filter at the assimilation step. The low-rank assumption that enables the Ensemble Kalman Filter to be computed with relatively small ensemble size is a key source of sampling errors of error covariance matrix estimation, especially with the error correlation between distant points, which can be addressed by localization. Cohn and Parrish (1991), Mitchell et al. (2002) have analyzed the problem and have noted that a small ensemble size produces noisy correlations between remote points within the flow. A localization function with local support that retains the correlations within a local neighborhood but suppresses all correlations beyond a specified cut-off radius has been used in the work of Salman (2006). The localization matrix is computed at each assimilation step and is a function of the drifter positions.

The localization was done by defining a cut-off radius and by forcing zero correlation beyond twice the radius. To prevent the sampling errors at large distances, covariance localization is a straightforward and simple solution. Houtekamer and Mitchell (1998) introduced a cut-off radius that defines the distance that an observation affects. Houtekamer and Mitchell (2001) applies a smooth function known as a fifth-order piecewise rational function (Gaspari and Cohn (1999)) to the terms $P_e H^T$ and $H P_e H^T$. This localization method is also used in Anderson (2001).

It was found that the optimal value of the cut-off radius increased as the number of available ensemble members increased. Arguing heuristically, the Ensemble Kalman Filter is computationally more efficient than the usual Kalman Filter at the expense of not precisely estimating correlations, especially the small correlations at
large distances.

The compactly supported correlation function is actually a matrix of horizontal correlation function $C_h(r_h)$ as defined in Gaspari and Cohn (1999), and Houtekamer and Mitchell (2000). $r_h$ is the horizontal distance. $C_h(r_h)$ is implemented to filter out the long range horizontal correlations. As noted in Houtekamer and Mitchell (2001), Hamill, Whitaker and Snyder (2001), the function $C_h(r_h)$ depends on the observation location of $x_D$; it has a maximum of 1.0 at the observation location and typically decreases monotonically to zero at some finite distance from the observation. I then modify the Kalman gain matrix by introducing a localization matrix $S$ such that

$$
K = \left( \begin{array}{c}
(S_{FD} \circ \mathbf{P}_{FD})_e \\
(S_{DD} \circ \mathbf{P}_{DD})_e
\end{array} \right) \left[ ((S_{DD} \circ \mathbf{P}_{DD})_e + \mathbf{R}_e)^{-1}
\right]
$$

(4.18)

$S_{FD}$ and $S_{DD}$ are corresponding to a distance dependent cut off matrix respectively, and computed at each assimilation step.

The horizontal correlation function is implemented to filter out spurious long range correlations between grid points $(x(i, j), y(i, j))$ and the observation location $(x_D(t_k), y_D(t_k))$ in the horizontal direction. The correlation function $C_h(r_h)$ is defined as:

$$
r_h(i, j, t_k) = \left[ (x(i, j) - x_D(t_k))^2 + (y(i, j) - y_D(t_k))^2 \right]^{\frac{1}{2}}
$$

(4.19)

where $r_h$ is the Euclidean distance between the grid point and the location of the Lagrangian drifter observation.

The localization function $C_h(r_h)$ is defined as a fifth order piecewise function as
follows:

\[
C_h(r_h) = \begin{cases} 
1 - \frac{1}{4}(\frac{r_h}{l_h})^5 + \frac{1}{2}(\frac{r_h}{l_h})^4 + \frac{5}{8}(\frac{r_h}{l_h})^3 - \frac{5}{3}(\frac{r_h}{l_h})^2 & 0 \leq r_h \leq l_h \\
\frac{1}{12}(\frac{r_h}{l_h})^5 - \frac{1}{2}(\frac{r_h}{l_h})^4 + \frac{5}{8}(\frac{r_h}{l_h})^3 + \frac{5}{3}(\frac{r_h}{l_h})^2 & l_h \leq r_h \leq 2l_h \\
-5(\frac{r_h}{l_h}) + 4 - \frac{2}{3}(\frac{r_h}{l_h})^{-1} & r_h > 2l_h
\end{cases}
\]  

(4.20)

Then the horizontal correlation matrix \(C_h(r_h)\) is defined for every grid point \((i, j)\) in the domain as defined in Hamill, Whitaker and Snyder (2001). The quantity \(l_h\) is the horizontal correlation scale for the assimilation. In our layered model system, \(l_h\) can also be different when computing the horizontal localization between the Lagrangian drifters and the top layer grid points which is the cut-off radius \(l^{(1)}_h\), and the horizontal localization between the Lagrangian drifters and the second layer grid points which is the cut-off radius \(l^{(2)}_h\). Figure 4.11 shows the results of keeping \(l^{(2)}_h\) fixed and the errors of kinetic energy and height field changing along with the variation of \(l^{(1)}_h\). The results show that there is no big difference between the cases where localization and no localization are implemented. However, by implementing localization, the errors can be reduced more smoothly and monotonically decreasing than in the no localization case since we could filter out the small background error correlations associated with the remote observations. The no localization case shows bigger fluctuation and it is more noisy. So, by providing the local support, the small and noisy correlations associated with remote observations are filtered out. Figure 4.12 shows the results of keeping \(l^{(1)}_h\) fixed and the errors of kinetic energy and height field changing along with the variation of \(l^{(2)}_h\). The experiments show that a larger cut-off radius is needed, i.e. \(l^{(2)}_h > l^{(1)}_h\) for the cross layer localization since the correlation between the drifters and the second layer is relatively weaker than the correlation between the drifters and the top layer. This localization strategy, like the cutoff radius we used, greatly improves the conditioning of the matrices \(P_{FD}\) and \(P_{DD}\).
Figure 4.11: Influence of localization shown in the error of kinetic energy and height field. \( l_h^{(2)} = 2000 \) km is fixed in each experiment. red: \( l_h^{(1)} = 1500 \) km; blue: \( l_h^{(1)} = 500 \) km; green: \( l_h^{(1)} = 800 \) km; black: No localization. (a): the top active layer; (b): the second active layer.
Figure 4.12: Influence of localization shown in the error of kinetic energy and height field. $l_h^{(1)} = 500$ km is fixed in each experiment. red: $l_h^{(2)} = 1000$ km; blue: $l_h^{(2)} = 2000$ km; green: $l_h^{(2)} = 4000$ km; black: No localization. (a): the top active layer; (b): the second active layer.

Since our model is a reduced gravity shallow water model, the vertical variation of each layer can be neglected. So the implementation of the vertical correlation
function has small influence of the errors of kinetic energy and height field. However, I would expect an improvement would happen for the non-shallow water flow models where the vertical variation has to be considered and plays an important role.
Chapter 5

Vertical Information Propagation and Parameter Estimation

5.1 Formulation of parameter estimation

In this chapter, I will mainly show how to implement the Lagrangian data assimilation method to estimate parameters in the layered model. Basically, I will still use the two layer point vortex system and assimilate the Lagrangian drifter observations in the top layer of the two layer point vortex system by using the Extended Kalman Filter scheme. To track the information about the parameters, I augment the model state space by treating the parameters as part of the state vector $x$. So that the new augmented model state vector is:

$$
x = 
\begin{pmatrix}
  x_F \\
  x_D \\
  \mu
\end{pmatrix} \quad (5.1)
$$
which combines the state vector of the flow $x_F$, the vector of the drifter $x_D$, and the parameter vector $\mu$. The parameter evolution equation and drifter advection equation are added to the model:

$$
\begin{align*}
\frac{dx^f_F}{dt} &= m_F(x^f_F, t; \mu^f) \\
\frac{dx^f_D}{dt} &= m_D(x^f_D, x^f_F, t; \mu^f) \\
\frac{d\mu^f}{dt} &= 0
\end{align*}
$$

The first part of the augmented equation is the unchanged original flow model, the second part of the augmented equation represents drifter advection equation, and the third part of the augmented equation represents that the parameters are constant.

Tracking the evolution of the augmented error covariance matrix by using the TLM, which gives a closed equation for $P^f$:

$$
\begin{align*}
\frac{dP^f}{dt} &= MP^f + (MP^f)^T + Q
\end{align*}
$$

The linearized dynamics operator and the error covariance matrix can be written in
The observations are still coming from the Lagrangian drifter observations. So the observation function $H$ is linear:

$$y^o(t_j) = h_j[x^t(t_j)] + \rho^t(t_j)$$

$$= \begin{pmatrix} 0 & H_{DD} & 0 \end{pmatrix} \begin{pmatrix} x_F^t \\ x_D^t \\ \mu^t \end{pmatrix} + \rho^t(t_j)$$

$$(5.5)$$

The Kalman gain matrix is augmented in the same way by including the correlation
between the Lagrangian drifter and the parameters $\mathbf{P}_{\mu D}$:

$$
\mathbf{K} = \mathbf{P}^{f} \mathbf{H}^T (\mathbf{H} \mathbf{P}^{f} \mathbf{H}^T + \mathbf{R})^{-1}
$$

$$
= \begin{pmatrix}
\mathbf{P}^{f}_D \mathbf{H}_{DD}^T \\
\mathbf{P}^{f}_D \mathbf{H}_{DD}^T \\
\mathbf{P}_{\mu D} \mathbf{H}_{DD}^T
\end{pmatrix}
\left[
\mathbf{H}_{DD} \mathbf{P}^{f}_D \mathbf{H}_{DD}^T + \mathbf{R}
\right]^{-1}
$$

(5.6)

In particular, it is $\mathbf{P}_{FD}$ and $\mathbf{P}_{\mu D}$ that enables $\mathbf{x}_F^a$ and $\mu^a$ to be estimated.

The updated error covariance matrix is given by:

$$
\mathbf{P}^a = (\mathbf{I} - \mathbf{KH}) \mathbf{P}^f
$$

(5.7)

The update state $\mathbf{x}^a$ can be written as:

$$
\mathbf{x}^a = \begin{pmatrix}
\mathbf{x}_F^a \\
\mathbf{x}_D^a \\
\mu^a
\end{pmatrix}
$$

$$
= \begin{pmatrix}
\mathbf{x}_F^f \\
\mathbf{x}_D^f \\
\mu^f
\end{pmatrix}
+ \mathbf{K}(\mathbf{y}^a - \mathbf{H}_{DD} \mathbf{x}_D^f)
$$

(5.8)

### 5.2 Comparison of vertical correlation and horizontal correlation

In this section, I focus on the performance of using information provided by Lagrangian drifters in the surface layer to improve the prediction of the flow state variables in the other layers. Not only vortex positions, but also physical parameters
can be estimated. From another point of view, I examine the performance of vertical information propagation by estimating key physical parameters.

Firstly, a particular model is set up to illustrate how information propagates both horizontally and vertically. Unlike the layer model I used in Chapter 3, here no vortex is located in the bottom layer and a single vortex is located in the top layer. Although the bottom layer is free of vortex circulation, the dynamic coupling of the two layers will induce a velocity flow field in the bottom layer from the vortex in the top layer.

To analyze the ability of the system to propagate information vertically and horizontally, the strength of circulation $\Gamma$ of the single vortex is to be estimated given an initial guess. When a Lagrangian drifter is released in the top layer, by observing the drifter’s position, the horizontal correlation between the vortex and the drifter enables us to estimate the parameter. In the same way, when the Lagrangian drifter is released into the bottom layer, by observing the drifter’s position in the bottom layer, the vertical correlation between the top layer vortex and the bottom layer drifter can also lead to a successful estimation of the parameter.

(1) Lagrangian drifter in the top layer: mainly horizontal correlation effect. Lagrangian data is used to track the vortex positions in the top layer, and to estimate the vortex circulation $\Gamma$ in the top layer.

(2) Lagrangian drifter in the bottom layer: vertical correlation effect. Lagrangian data from the bottom layer is used to track the vortex positions in the top layer, and to estimate the vortex circulation $\Gamma$ in the top layer.

By augmenting the model state vector with the parameter: circulation of the
vortices $\Gamma$, and the drifter coordinates, the augmented state vector has the form of:

$$
\mathbf{x} = \begin{pmatrix}
\mathbf{x}_F \\
\mathbf{x}_D \\
\Gamma
\end{pmatrix}
$$

(5.9)

In this case, the corresponding observation operator is:

$$
\mathbf{H} = \begin{pmatrix} 0 & I & 0 \end{pmatrix}
$$

(5.10)

As Kuznetsov et al. (2002) have mentioned, understanding the correlation between the information value of a drifter and its initial position is a crucial step in devising optimal strategies. The efficiency of the scheme depends critically on the launch locations of the drifter. The performance of assimilation error of circulation estimation using the Extended Kalman Filter shown in Figure 5.1 also depends on the drifter launch location. Suppose the initial guess of the parameter $\Gamma$ has a 20% uncertainty. The target of the assimilation is to reduce the uncertainty of the parameter. The figure shows that the performance depends on the launch location of the drifter whether it is released in the top layer or in the bottom layer. If the drifter is taken too close to a vortex in the top layer, for instance the drifter numbered 2 in Figure 3.3, it rotates around the vortex. Its fast rotation (compared with the vortex rotation) leads to a rapid growth in the drifter error. Unless the observation frequency is substantially increased the error covariance predicted by the TLM would become unreliable. Especially for the point vortex system, since it is strongly nonlinear and singular, the large nonlinearity has a great impact near the vortex. So the way to treat such a drifter is to use a shorter observation time interval to capture enough information. When the drifter is released into the bottom layer, there is no strong rotation for the drifter even the horizontal distance is small, so the error of parameter
estimation only depends on the horizontal distance between the top vortex and the bottom drifter. If I take the drifter launch location far away from the vortex in the top layer, for instance the drifter numbered 3 in Figure 3.3, or in the bottom layer, the weak correlation between the vortex and the drifter would not provide enough constraints on the individual vortex induced by system noise. The drifter is not able to reduce the error of the forecast of vortex. Suppose that there is an arbitrary point in the two layer model, the magnitude of velocity is a weighted sum of same layer velocity component \((1 + (r/\lambda)K_1(r/\lambda))/2r\) and cross layer velocity component \((1 - (r/\lambda)K_1(r/\lambda))/2r\), induced from same layer or different layer respectively as I mentioned in Chapter 3. Whatever it is, if there is a large distance between drifter and vortex, the velocity of vortex at this point is trivial as shown in Figure 3.1, so that it cannot carry adequate information of the vortex. This is somehow similar to the case without assimilation, so that the assimilated errors go up with the non-assimilated errors. Because of the linear behavior of circulation in the vortex flow, estimation of the circulation converges linearly.
Figure 5.1: Estimation of circulation $\Gamma$ through horizontal propagation (drifter in the top layer) and vertical propagation (drifter in the bottom layer) with the change of drifter positions, under the same noise levels $(\sigma, \rho) = (0.02, 0.02)$ and observation interval $\Delta T = 1$. The parameter $\Gamma$ has an initial 20% uncertainty. The performance of assimilation depends on the horizontal distance between the vortex and the drifter.

For this particular setup, I can claim that information can be propagated vertically, and is effectively equivalent to the propagation of horizontal information. Recall that, in Chapter 3, I investigated the magnitude of the velocity in each layer, from Figure 3.1, I can see that, as the distance between the vortices grows, the magnitudes of the velocity are almost of the same order whether from the same layer or different layers. As the distance $r$ increases, the two magnitudes of the velocity components: $(1 + (r/\lambda)K_1(r/\lambda))/2r$ and $(1 - (r/\lambda)K_1(r/\lambda))/2r$, are both tending to $1/r$. The same strength of velocity magnitudes leads to the same strength of correlations.
Chapter 6

Conclusions and Future Research

6.1 Summary and Discussion

Our goal was to develop an application of Lagrangian data assimilation, which is expected to produce improved ocean structure forecasting system. Several approaches to Lagrangian data assimilation for ocean models have been suggested, for instance, the Extended Kalman Filter with the augmented system and the Ensemble Kalman Filter with the augmented system. The present research is aimed at exploring one of the main questions in Lagrangian data assimilation: given the Lagrangian drifter observations in the top layer, are they able to correct the flow field in the same layer and also correct the flow field in the sub-surface layer through the water column.

Chapter 2 introduced several formulations of the Lagrangian data assimilation methods. It introduced the Extended Kalman Filter and the Ensemble Kalman Filter for Lagrangian data assimilation with an augmented system. Difficulties in the assimilation of Lagrangian data arise because the model is generally described in terms of Eulerian variables computed on a fixed grid in space, as a result there is
no direct connection between the model variables and Lagrangian observations that carry time-integrated information. In order to capture the correlation between the Lagrangian drifter observations and model variables, the vector space is augmented by including the Lagrangian drifter coordinates into the system. The evolution model is also augmented with drifter advection equations to track the correlations between the flow and the drifters. The error covariance matrix carries the correlation of flow state vectors and Lagrangian drifter observations to update the flow state vectors. The formulation and implementation of the Extended Kalman Filter and the Ensemble Kalman Filter for Lagrangian data assimilation are described in Chapter 2.

The major difference between the Extended Kalman Filter and the Ensemble Kalman Filter is that the approximation of the nonlinearities of the flow model and the measurement process, for instance, the Extended Kalman Filter uses a linearized equation for the error covariance matrix and the Ensemble Kalman Filter propagates a finite ensemble size of error covariance matrix nonlinearly. The implementation of the Extended Kalman Filter requires derivatives of the flow model which is evaluated numerically from a TLM. The implementation of the Ensemble Kalman Filter requires propagation of the nonlinear system for each ensemble.

Chapter 3 describes how Lagrangian data assimilation was implemented into the two layer point vortex systems by using the Extended Kalman Filter. The first main step is the formulation of a two layer point vortex system by introducing the barotropic mode and baroclinic mode. The dynamical coupling between the layers of two layer point vortex system provides us the possibility of assimilating the Lagrangian drifter observations in the top layer to correct the flow motions, here mainly the position of vortices, in both of the layers. The error covariance matrix is calculated by using tangent linear model of the evolution dynamics. It is a first order approximation of Taylor expansion to the evolution system. The experiment shows
that by only assimilating the Lagrangian drifters in the top layer of the model, we could not only correct the vortex positions in the top layer, but also correct the vortex positions in the bottom layer where no Lagrangian drifter observations directly come from. The results provide us an understanding of how the information propagates vertically through the water column.

Chapter 4 introduced the two and a half layer reduced gravity shallow water double gyre flow configuration. It is derived from the Navier-Stokes equations over the three layers and using the hydrostatic assumption for each layer. The flow is driven by a wind forcing term in the top active layer. And the layer frictions are present at each layer interface. The system consists of two active layers where velocity fields and height fields are changing with time, and a motionless bottom layer which is infinitely deep. Again, the flow together with the drifter equations comprise the augmented system and the drifter coordinates are also augmented to the state space. The flow equations are discretized on an $n_x \times n_y = (100 \times 100)$ grid. This spatial discretization produces a set of $10^4$ ordinary differential equations. Obviously it is impossible to implement the Extended Kalman Filter and compute the tangent linear model. The computational cost is extremely expensive with the use of the TLM to evolve the error covariance matrix for such a large size system. So the Ensemble Kalman Filter was implemented to the two and a half layer reduced gravity shallow water model. Each ensemble is evolved using the fully nonlinear flow model, no linearizations are made in the Ensemble Kalman Filter. The experiment shows that an ensemble of $O(80)$ members are capable of producing an accurate error covariance matrix. It greatly reduces the computational cost compared with the Extended Kalman Filter. The observations are still coming from the Lagrangian drifter information in the top active layer of the layered shallow water. The drifters are released into the most energetic western region in the top layer since it is more effective to collect information in dynamically active regions of the flow. To quantify the performance of Lagrangian
data assimilation using the Ensemble Kalman Filter, the kinetic energy error norms and height error norms are calculated. The experiments show that by assimilating the Lagrangian drifters in the eddy dominated region surrounding the jet in the top layer, both the kinetic energy error norms and the height field error norms can be reduced greatly from the forecast in both of the active layers. It demonstrated that given the Lagrangian drifter observations at discretized time intervals, it can be used to correct the flow field where the Lagrangian drifters are coming from, and also can be used to correct the sub-surface flow field where Lagrangian drifters come from.

The idea of covariance localization for Eulerian data assimilation has also been implemented in this work to reduce the noisy correlations between the remote sites within our flow domain. The noisy correlations between the Lagrangian drifters in the top layer with the remote points in the top layer, and the noisy correlations between the Lagrangian drifters in the top layer with the remote points in the second active layer are both considered in this work. The element of the localization matrix corresponds to a distance dependent on cut-off function. The localization function is computed by the drifter positions at each assimilation step. The experiments show that with the localization, I can reduce the kinetic energy error norms and height field error norms further and smoothly.

Chapter 5 describes the application of Lagrangian data assimilation to estimate the key parameter of two layer point vortex model such as the strength of the vortex circulation to compare the horizontal correlation and vertical correlation. The vertical correlation is as strong as the horizontal correlation so that we could track it and use it to estimate the state variables and parameters. The existence of dynamical coupling is of fundamental importance for efficient and accurate field and parameter estimation. The experiments show that provided the layered system is dynamically coupled between the layers, Lagrangian data assimilation is able to re-
duce the uncertainty of the key parameters \( \Gamma \) and estimate the parameter \( \Gamma \) correctly. The experiments also state that Lagrangian drifters can be used to capture the vertical information propagations through the water column if the layers are dynamically coupled. The performance of the estimation depends on the correlation between the drifter position and the vortex position. So the launch location of drifter has to be considered. The stronger correlation between the drifter and the vortex provides us a better performance when estimating the parameter \( \Gamma \).

6.2 Future Research

Lagrangian data assimilation is only a few years old and has to develop a path towards an operational application. A Lagrangian data assimilation system consists of three components: a set of observations which comes from the Lagrangian drifters, an augmented dynamical model including the Lagrangian drifter coordinates into the system, and a Lagrangian data assimilation scheme. The specific uses of Lagrangian data assimilation depend upon the quality of Lagrangian drifter data sets and ocean models, and the desired purposes of the filtering schemes.

Possible further investigations relevant to Lagrangian data assimilation are as follows:

6.2.1 More realistic ocean models

The Lagrangian data assimilation method has been developed and applied to oceanographic models of increasing complexity and realism, from single layer models to multi-layer models, for instance, the point vortex system and the reduced gravity shallow water double gyre flow configurations. The main goal of Lagrangian data
assimilation is to develop a simple and portable method that can be applied to realistic models and configurations. I intend to construct more complicated and more realistic models which describes the motion of ocean and apply Lagrangian data assimilation into the system.

6.2.2 Covariance localization and filter tuning

Covariance localization is a practical aspect of the Ensemble Kalman Filter. It restricts the use of ensemble-based error covariance matrix to relatively small subsets of grid points in the whole flow region. So the background error covariance matrix is actually calculated on a small size of ensembles. For the single layer models or shallow water models, horizontal localization are considered only since vertical localization would not have a big positive impact on the results. The cut-off radius eliminates the need to have to estimate the small correlations associated with remote observations. However, vertical localizations should also be considered in the open, stratified ocean where vertical variations can not be ignored and play an important role. The vertical localization should decrease the dimension of the local regions, allowing for a better local representation of the background structures by the relatively small ensembles.

6.2.3 Parameter estimation of more realistic ocean models

Parameter estimation via data assimilation is making an increasingly significant impact on ocean science. Lagrangian data assimilation-based parameter estimation in the ocean sciences has benefit for model tuning. Moreover, it can also be used to help account for model inadequacy during ensemble forecasting. Data acquisition in the ocean is sufficiently difficult and costly so as to make field estimates by direct measurements sparse. How to assimilate Lagrangian data obtained from Lagrangian
drifters to estimate the key parameters of the flow ocean is then a challenging issue in ocean data assimilation.
Bibliography


