ASSIMILATION OF LAGRANGIAN DATA INTO IDEALIZED MODELS OF THE OCEAN MESOSCALE USING ENSEMBLE-BASED METHODS

by
Muhammad-Kassiem Jacobs

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2010

DALHOUSIE UNIVERSITY

DEPARTMENT OF OCEANOGRAPHY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "ASSIMILATION OF LAGRANGIAN DATA INTO IDEALIZED MODELS OF THE OCEAN MESOSCALE USING ENSEMBLE-BASED METHODS" by Muhammad-Kassiem Jacobs in partial fulfillment of the requirements for the degree of Master of Science.

Dated: August 27, 2010

Supervisor: _____

Readers: _____

_____

_____

_____

DALHOUSIE UNIVERSITY

DATE: August 27, 2010

AUTHOR: Muhammad-Kassiem Jacobs

TITLE: Assimilation of Lagrangian Data into Idealized Models of the Ocean Mesoscale Using Ensemble-Based Methods

DEPARTMENT OR SCHOOL: Department of Oceanography

DEGREE: MSc CONVOCATION: October YEAR: 2010

To my wife and my parents for everything!

Table of Contents

List of Tables viii

List of Figures ix

Abstract xi

Acknowledgements xii

Chapter 1 Introduction 1

1.1 Overview of Data Assimilation 3

1.2 Objectives of the Thesis 7

1.3 Structure of the Thesis 7

Chapter 2 Overview of Sequential Data Assimilation 9

2.1 The Linear Stommel Model 9

2.2 Bayesian View of Data Assimilation 11

2.3 Kalman Filters 14

2.3.1 The Linear Kalman Filter 14

2.3.2 The Extended Kalman Filter 18

2.3.3 The Ensemble Kalman Filter 20

2.4 Particle Filters 22

2.4.1 Monte Carlo Integration 23

2.4.2 Sequential Importance Sampling 24

2.4.3 Sequential Importance Resampling 28

2.4.4 Markov Chain Monte Carlo (MCMC) 30

2.5 Discussion 31

Chapter 3 A New Class of Idealized Models of Ocean Eddy and Drifter Position **41**

3.1 The Flow Field 45

3.2 The Drifter Trajectories 48

3.3 The Augmented State-Space Model 50

3.4 Propagation of Uncertainty 51

3.4.1 Experiment I: Uncertain Drifter Release 51

3.4.2 Experiment II: Uncertain Model Dynamics 53

3.5 Discussion 55

Chapter 4 Assimilating Trajectories Into the Idealized Model **70**

4.1 Introduction 70

4.2 An Illustrative Example 72

4.3 Measuring Non-Gaussianity 75

4.3.1 A Simple Measure of Non-Gaussianity 76

4.3.2 Non-Gaussianity and the SAV 77

4.4 Using Δ to Predict EnKF Failure 78

4.5 Sensitivity to Model Parameters 82

4.5.1 Performance Metrics 83

4.5.2 The Sensitivity Maps 84

4.6 Discussion 88

Chapter 5 Discussion **114**

Bibliography **118**

Appendix A The Linear Stommel Model **124**

A.1 Discretization Of The Vorticity Equation 125

Appendix B Pseudo codes for the SIS, generic Particle Filter, SIR Filter and Resample Routine **127**

B.1 Filtering via SIS 127

B.2 Generic Particle Filter 127

B.3 Filtering via SIR 128

B.4 Resample Algorithm 129

Appendix C The Markov Chain Monte Carlo Move Step **130**

List of Tables

Table 2.1 Model parameters used in the Stommel model 17

Table 3.1 List of vortex swirl velocity profiles used in this study 45
Table 3.2 Characterization of the instantaneous stagnation points 46
Table 3.3 Model parameters used in the simulations 52

Table 4.1 Model parameters used in experiments E0 to E7 84

List of Figures

Figure 2.1 Stream functions from the Stommel model 32
Figure 2.2 Stream function estimates from the KF 33
Figure 2.3 Schematic of the EnKF data assimilation procedure 34
Figure 2.4 Stream function estimates from the EnKF 35
Figure 2.5 Example of a poor choice of proposal density prior 36
Figure 2.6 Illustration of Sequential Importance Resampling 37
Figure 2.7 Schematic of the SIR filter 38
Figure 2.8 Stream estimates from the SIR filter 39
Figure 2.9 Stream function estimates from SIR combined with MCMC . . 40

Figure 3.1 Displacement map for the SAVM over one time step 57
Figure 3.2 Trajectory map for the SAVM over 150 time steps 58
Figure 3.3 Stream function for a two point-vortex system 59
Figure 3.4 Probability density of drifter position (I) 60
Figure 3.5 Probability density of drifter position (II) 61
Figure 3.6 Probability density of drifter position (III) 62
Figure 3.7 Probability density of drifter position (IV) 63
Figure 3.8 Probability density of drifter position (V) 64
Figure 3.9 Probability density of drifter position (VI) 65
Figure 3.10 Probability density of drifter position (VII) 66
Figure 3.11 Probability density of drifter position (VIII) 67
Figure 3.12 Sensitivity of drifter dispersion to release location 68
Figure 3.13 Trajectories of 100 drifters for selected released locations . . . 69

Figure 4.1 Basic setup of the illustrative example 91
Figure 4.2 Performance of the EnKF and PF for the illustrative example 92
Figure 4.3 Drifter displacements as a function of release point (I) 93

Figure 4.4 Drifter displacements as a function of release point (II) 94
Figure 4.5 Simple measure of non-Gaussianity 95
Figure 4.6 Map of $\Delta(x_0, y_0|r_0, N)$ (I) 96
Figure 4.7 Map of $\Delta(x_0, y_0|r_0, N)$ (II) 97
Figure 4.8 Performance of the Particle Filter (I) 98
Figure 4.9 Performance of the EnKF (I) 99
Figure 4.10 Performance of the Particle Filter (II) 100
Figure 4.11 Performance of the EnKF (II) 101
Figure 4.12 Performance of the Particle Filter (III) 102
Figure 4.13 Performance of the EnKF (III) 103
Figure 4.14 Error in the EnKF as a function of drifter release point 104
Figure 4.15 Error maps for experiment E0 105
Figure 4.16 Performance metrics for experiment E1 106
Figure 4.17 Performance metrics for experiment E2 107
Figure 4.18 Performance metrics for experiment E3 108
Figure 4.19 Performance metrics for experiment E4 109
Figure 4.20 Performance metrics for experiment E5 110
Figure 4.21 Performance metrics for experiment E6 111
Figure 4.22 Performance metrics for experiment E7 112
Figure 4.23 Final distribution of 200 drifters after 100 time steps 113

Figure C.1 Data assimilation procedure using the MCMC move step . . . 133

Abstract

It is generally accepted that models of the deep ocean must assimilate observations in order to make realistic forecasts in regions dominated by mesoscale variability (i.e., "ocean weather"). The present study is an attempt to quantify the information on ocean weather that is contained in Lagrangian trajectories, and the best way to extract it. Following a review of ocean data assimilation in a Bayesian framework, including the Ensemble Kalman Filter and the Particle Filter, a new class of idealized models of self advecting vortices is introduced. Through a large number of carefully designed Monte Carlo experiments it is shown when, where and why the Ensemble Kalman Filter will fail. The study concludes with a discussion of a hybrid scheme that takes advantage of the lower computational cost of the Ensemble Kalman Filter and the ability of the Particle Filter to handle highly non-Gaussian probability density functions.

Acknowledgements

First, I'd like to thank my supervisor Keith Thompson. His guidance, his patience, his teaching, and his encouragements have made this thesis possible.

I'd also like to thank my external examiner Michael Dowd and my committee members: Allyn Clarke, Barry Ruddick and Bruce Smith, for their constructive criticisms and support.

A special thank you is owed to Dan Wright who was originally part of this committee. Sadly, Dan Wright passed away last month.

Numerous other people have contributed in one way or another to the work presented here. Thank you to all the faculty members that have taught me and to everyone who has helped with technical and administrative issues. In particular, I'd like to thank Jackie, Pamela, Tamara, and Balagopal.

Finally I want to thank professor Huw Davies from the ETH Zurich for helping me to continue my thesis after I moved to Switzerland.

Chapter 1

Introduction

Over the last 20 years there has been a revolution in the way ocean data has been collected. Since the early 1980's satellites (e.g. Nimbus, Geosat, ERS, Topex/Poseidon, Jason-1) have provided near-global monitoring of sea surface height and temperature, ocean color, winds and surface waves (Le Traon [2002]). Satellites generally provide good spatial coverage of surface properties but do not provide direct observations of the interior of the ocean. The repeat time (about 10 days for Jason-1) is comparable to the time-scale of mesoscale ocean and this can cause serious aliasing problems if the eddies move significantly (e.g. Chen and Chen [2004], Gille [2001], Le Traon et al. [1998], Le Traon and Dibarboure [2002]). The altimeter is one of the most important satellite sensor for physical oceanographers (e.g. Ducet and Le Traon [2001], Le Traon et al. [2003], Keppenne et al. [2005], Pascual et al. [2006], Thompson and Demirov [2006]).

Recently a new stream of in situ data has became available from the Argo program, an international effort to deploy profiling floats that drift freely with the ocean currents at about 1000m depth. Before ascending to the surface every ten days, the floats descend to 2000m and then record the vertical profile of temperature and salinity during their ascent. At the surface the float transmits these profiles, and its position, to a data center via a satellite link. The float then sinks back to the 1000m depth and the cycle repeats. The Argo program has presently deployed approximately 3000 floats throughout the world's oceans.

The Argo program is revolutionizing the collection of in situ ocean data. The

temperature and salinity profiles from the Argo floats complement the altimeter observations by providing information on the interior state of the ocean. The combination of these two data streams is providing a wealth of information on large-scale and mesoscale variability. The Argo floats also provide useful information on the motion of the upper ocean in the form of their changing horizontal positions. Guinehut et al. [2004] for example has recently assimilated satellite altimeter and sea surface temperature observations and temperature profiles from Argo floats into the $1/6^{\circ}$ resolution Ocean PArallelise (OPA) ocean model, and Fan et al. [2004] assimilated satellite altimeter observations and float trajectories into the Princeton Ocean Model (POM). In both studies the additional data stream from the floats led to significant improvements in the representation of the ocean state. A major goal of this thesis is to develop ways of extracting useful oceanographic information from the trajectories of freely advected objects like Argo floats.

It is generally accepted that realistic forecast models of the ocean cannot be based solely on the governing equations of fluid flow (or their numerical approximations). As in weather forecasting, errors in the representation of physical processes by the model, and uncertainties in the initial, surface and lateral boundary conditions, make it necessary to assimilate observations. As early as 1963 Lorenz showed that data assimilation is required if unstable systems are to be integrated beyond their deterministic predictability limit (Lorenz [1963a], Lorenz [1963b]). For the synoptic scale of the atmosphere Lorenz calculated that this limit was about 10-14 days (Lorenz [1982], see chapter 6 in Kalnay [2003]). The corresponding predictability limit for the ocean mesoscale is longer, approximately 14-21 days (Powell et al. [2009], Özgökmen et al. [2000]).

Gridded altimeter fields and Argo profiling floats are now assimilated routinely by operational ocean forecast agencies using various methods. The most popular methods are Optimal Interpolation (Cooper and Haines [1996]), Reduced-Order Optimal Interpolation (De Mey and Benkiran [2002]) and the Ensemble Kalman Filter (Evensen [2003]). A detailed description of the ocean model, model resolution and

assimilation methods used by the various ocean forecast agencies can be found at http://strand1.mersea.eu.org/html/strand1/model.html.

There is much interest at the present time in assimilating temperature and salinity profiles from Argo floats (e.g. Guan and Kawamura [2004], Durand et al. [2002], Delcroix et al. [2005]). These data are now used routinely by several forecast agencies, e.g. Mercator and UK Meteorological Office. Relatively little work to date has been completed on the assimilation of Argo float trajectories and this is a major topic for the present study. In particular I will try to quantify the extra information that can be gained from Lagrangian data like observed Argo float trajectories as an additional source of data for assimilation. To this end, data assimilation experiments will be performed with an idealized multiple vortex model. Several assimilation methods will be used in combination with a range of pseudo observations (i.e. gridded and in situ altimeter observations, Lagrangian float trajectories.)

The rest of this chapter provides a brief review of data assimilation, states the objectives of the research and outlines the rest of the thesis.

1.1 Overview of Data Assimilation

Realistic ocean models are often based on coupled partial differential equations that have been discretized in time and space. All models only approximate reality and are subject to errors of various kinds (e.g. incorrect parametrization of physical processes). The effect of these errors is the inevitable drift of the model from reality. Assimilation of data into the model is the only way to counteract such a drift of the model. It has been an essential part of weather forecasting for over 50 years (see Daley [1992]).

Data assimilation methods can be separated into two main categories: variational and sequential. Variational methods (e.g. 4DVar) usually fit dynamical models over a fixed time interval by minimizing a cost function with respect to unknown model parameters or conditions (e.g. the initial state). The model is often assumed to

be error-free "strong constraint") and the minimization is often carried out using a tangent linear and adjoint model, both of which can change with the model state. Any changes in the ocean model requires the appropriate adjustment of the adjoint model and this can make coding difficult. The strong constraint formulation imposes limitations on the length of the time interval for non-linear systems. The limitations are related to the predictability limit of the model (van Leeuwen and Evensen [1996]).

Sequential methods generally allow explicitly for model error ("weak constraint"). They are easier to implement than the variational techniques, and, unlike variational methods, provide error estimates for the estimated ocean state. The time series of covariance is of special interest because it gives insight into how errors in the estimated state evolve in time (Sergers [2002]). The computational cost of sequential methods are however usually higher than variational methods. Two promising sequential methods are the Ensemble Kalman and Particle Filters, both of which are based on Monte Carlo methods. One of the major difficulties in using such methods is that the quality of the models' predictions depend strongly on the assumed statistical properties of the model error; in practice this is usually not well known.

The standard Kalman Filter applies to linear systems and is thus not relevant to complex ocean models (Evensen [1992]). There are however extensions of the standard Kalman Filter that allow for non-linearity: the Extended Kalman Filter (EKF) and the Ensemble Kalman Filter (EnKF). The EKF requires the Tangent Linear Model (TLM) of the underlying model and this makes the EKF unsuitable for strong non-linear models (Evensen [1994]). The EnKF basically uses the same matrix updating equations as the linear Kalman Filter, but it deals with the non-linear dynamics using a Monte Carlo approach. More specifically an ensemble of model states is used to compute the first and second moments of the ensemble which are then used in the analysis step to linearly assimilate the observations. This underlying assumption of Gaussianity limits its applicability (Kivman [2003] and van Leeuwen [2009]). Only in the case of a linear model with Gaussian errors will the prior and posterior distributions be correctly represented and updated as Gaussian distributions.

In Evensen and van Leeuwen [1996] the EnKF is used to assimilate Geosat altimeter data to improve the representation of the Agulhas Current in a quasi geostrophic 5-layer ocean model. It was shown that assimilating the 10 day spaced gridded altimeter observations did limit the rapid meandering growth of the jet, compared to more sophisticated models, and enhanced the eddy shedding which would have been otherwise too small due to the lack of ageostrophic effects.

Keppenne [2000], Keppenne and Rienecker [2002] and Keppenne and Rienecker [2003] describe the implementation and testing of the EnKF in the assimilation of temperature profiles from the Tropical Atmosphere Ocean (TAO) array of moored buoys into the POSEIDON ocean model. Their conclusion is that the EnKF can provide better forecasts than the standard optimal interpolation (OI) method which was widely used at that time.

Molcard et al. [2005] used the EnKF to assimilate float positions into a three-layer primitive equation Miami Isopycnic Coordinate Ocean Model (MICOM) in a simple double-gyre configuration. Experiments were carried out with different sampling time steps (20 minutes, 3 and 6 days) for the observations. In all experiments the ensemble size of the floats was fixed at 24 members. The release positions of the floats on the western boundary included an energetic region which included the recirculation regions and the meandering jet. It was concluded that the model error grows significantly as the sampling time step is increased, in particular when the sampling time step exceeds the Lagrangian time scale of the float trajectories.

Salman et al. [2006] used a 1 1/2-layer quasi geostrophic ocean model to assimilate float trajectories. One focus was to investigate dependencies of the model error with respect to the release domains of the floats and the ensemble size (20, 40 and 80) of the released floats. A further focus was to investigate the influence of the sampling time step (1-20 days) on the model error. They concluded that the performance of the assimilation method degrades with decreasing ensemble size. The assimilation method seems to be insensitive for sampling time steps smaller than 10 days, increasing the sampling time steps beyond 10 days lead to increase in the model error and

finally to a complete failure. A further conclusion was that the assimilation of float observations from eddy-dominated regions surrounding the jet led to smaller model errors compared to assimilating float observations from quiescent parts of the model domain.

The Particle Filter (PF) is another Monte Carlo based method of sequential data assimilation. It imposes virtually no assumptions on either model dynamics or error statistics. The underlying theory for the PF reaches as far back as 1949 (Doucet [2006]) making it one of the oldest data assimilation methods, at least in theory. The only constraint for the application of the PF methods is computational power.

Following the seminal study of Gordon (Gordon et al. [1993]), various particle filter schemes have been developed and applied to low dimensional problems, mainly in robotics and target tracking. The main advantage of the PF is that it can be applied to problems of any complexity and, more importantly, no assumption is made about the probability distribution of the model and observation errors. A practical advantage is that the implementation of the method, at least in its simplest form, is extremely simple, e.g. no matrix inversions are needed. The main reason the PF is not used more widely in fields such as oceanography and atmospheric science is related to the so-called "curse of dimensionality" which refers to the excessive computational cost of dealing with state vectors of high dimension (such as those associated with realistic models of the ocean and atmosphere). However with ever increasing computational power at decreasing costs, and the recognition that it is the dimension of the effective model error rather than the model state that is most important, there is hope that the curse will eventually be overcome.

Particle Filters compute probability distributions from an ensemble of model states that are require to remain "fairly close" to the observations as they become available. To accurately approximate the probability density of the model state the ensemble size usually has to be of order of 10^3 to 10^6 depending on the dynamic system. Given each ensemble member has to be advanced forward in time by the numerical model, the computational burden can be huge if the model state is large. Because of this

limitation, PFs have been mainly used to date in robotics and signal processing using model states of dimension 10 or lower. The state vector of even the simplest ocean model is however of order 10^4.

An early study on the assimilation of gridded altimeter observations using the PF was carried out by van Leeuwen [2003]. Van Leeuwen argued that the Kalman update step in the EnKF is not variance-minimizing. To overcome this problem he introduced a particular form of the PF and showed that it qualitatively outperforms the EnKF when applied to the Korteweg-DeVries equation. He then assimilated gridded altimeter data into a 5-layer QG model using the PF and showed that the model can recover the observed features of the Agulhas Current system. Van Leeuwen also pointed out some potential problems of the PF, namely the tendency for the ensemble to collapse over time and the practical limitation of the ensemble size.

1.2 Objectives of the Thesis

The overall objective is to quantify how much information on basin-scale and mesoscale oceanographic flows can be extracted from Lagrangian float trajectories. An important secondary objective is to compare the performance of two very different assimilation methods, the EnKF and PF, and assess their strengths and weaknesses using a highly idealized model of variability of the ocean mesoscale.

1.3 Structure of the Thesis

The mathematical foundations of the assimilation methods used in this thesis are given in chapter 2 including worked examples based on a highly idealized, linear ocean model (the so-called Stommel model). In chapter 3 an idealized model of the evolution of vortices and drifter position is introduced (the self advecting vortex model). The main characteristics of the SAVM are discussed in detail. Chapter 4 describes the experiments performed with the SAVM. The purpose of the experiments

is to compare the performance of the PF and EnKF and study the sensitivity of estimated quantities to changes in the model's parameters. The results are summarized and discussed in the final chapter.

Chapter 2

Overview of Sequential Data Assimilation

This section provides an overview of some of the more promising assimilation methods presently used in oceanography. The focus is on methods that can deal with non-linearity and non-Gaussianity, namely the Ensemble Kalman Filter and the Particle Filter.

In the first subsection I will introduce briefly the linear Stommel model. The purpose of this model is to demonstrate assimilation methods with a very simple ocean model. In the second subsection the general concept of data assimilation in the Bayesian framework is introduced. The most widely used Kalman filters (linear Kalman filter, extended Kalman filter and Ensemble Kalman filter) are presented in subsection 3. In subsection 4 the general particle filter is introduced including the sequential importance filter and the Markov Chain Monte Carlo method. The last subsection is a brief summary of this chapter.

2.1 The Linear Stommel Model

One of the simplest ocean models is the linear Stommel model. The model includes a simple western boundary current, the steady state Sverdrup balance and is influenced by the propagation of barotropic Rossby waves.

The governing equation can be derived from the equations of motion and the continuity equation and is given by the following vorticity equation

$$\frac{\partial}{\partial t}(v_x - u_y) + f(u_x + v_y) + \beta v = \Phi - \lambda(v_x - u_y) \tag{2.1}$$

where (u, v) denote the velocities in zonal and meridional directions respectively, f

is the Coriolis parameter, Φ is the curl of the wind stress, $\beta = df/dy$, and λ is the friction parameter. Subscripts denote partial derivatives.

After non-dimensionalization followed by discretization (see Appendix A for details) the governing equation takes the form

$$\frac{\partial}{\partial t}[M_{xx} - l^2]\psi_t + M_x\psi_t = \Phi - \lambda[M_{xx} - l^2]\psi_t \tag{2.2}$$

where ψ_t is a column vector describing the stream function at time t, l is the meridional wave number and M is a matrix which describes the spatial discretization of the model. M_x and M_{xx} denote the first and second partial derivatives. The stream function ψ relates directly to the sea surface height of the model. In steady state the second and third term describe the Sverdrup balance of the ocean interior. The first and second term describe the propagation of barotropic Rossby waves. In steady state the second and fourth terms balance in the western boundary current (which has width λ/β in dimensional terms as explained in Appendix A).

The purpose of introducing the Stommel model is to illustrate the data assimilation methods presented in the following sections. The discretized model (2.2) is integrated forward in time and the resulting model fields are taken to represent the 'true' model states. These model fields are perturbed with Gaussian random noise to simulate time series of 'observed' model fields, analogous to observed sea level fields in the real world. From these 'observed' model fields, pseudo observations are subsampled and then assimilated using the EnKF and PF. Figure 2.1 shows the stream function of the true ocean state (panel a) and the stream function of the observed ocean state (panel b). The right panel (b') shows the stream function with friction parameter $\lambda = 0.1$; this value was used in all data assimilation experiments to test if the assimilation methods are capable of recovering the true state even with an incorrect friction parameter.

2.2 Bayesian View of Data Assimilation

Discrete time dynamical systems are typically represented by a state equation which describes the time evolution of the state vector x from time $k-1$ to k:

$$x_k = f_{k-1}(x_{k-1}, v_{k-1}). \tag{2.3}$$

In general f_{k-1} is a known function, often non-linear, that advances the state vector forward one time step. The term v_{k-1} is the model noise which represents errors due to finite resolution, missing physics, imperfect parameterizations, forcing errors, etc.

In the context of the linear Stommel model the state vector x_k represents the stream function ψ_t and the general function f represents a known linear function which includes all processes of the model and describes the integration from time step $k-1$ to k.

The relationship between the state at time k and the observations at the same time is given by the so called measurement equation:

$$z_k = h_k(x_k, w_k) \tag{2.4}$$

where h_k is a known function, often non-linear, that maps the state vector onto the observations. The term w_k is the observation noise that includes the measurement error and the error of representativeness (e.g. Daley [1992]).

The goal of any data assimilation method is to combine equations (2.3) and (2.4)in order to estimate the final (also called *updated* or *analyzed*) state which should recover the true state as closely as possible.

In the Stommel model the vector z_k represents the observations of the stream function (which can be thought of as sea level measured by altimeters).

The state and measurement equation together form the so-called state-space model. Usually the model and measurement errors are assumed to be mutually independent

(with each other and through time) and have known probability density functions, p_v and p_w respectively. This formulation is general and does not assume linearity, Gaussianity or additive errors.

The filtering problem (e.g. Kitagawa, 1998) is to estimate the state at time k given all available observations up to and including time k
($Z_k \stackrel{\text{def}}{=} \{z_i, i = 1, \ldots, k\}$). The solution is given by

$$p(x_k|Z_k) \propto p(z_k|x_k) \int p(x_k|x_{k-1}, Z_{k-1}) p(x_{k-1}|Z_{k-1}) dx_{k-1}. \tag{2.5}$$

If $p(x_{k-1}|Z_{k-1})$ is available at time $k-1$, the prediction density for $p(x_k|Z_{k-1})$ is given by

$$p(x_k|Z_{k-1}) = \int p(x_k|x_{k-1}) p(x_{k-1}|Z_{k-1}) dx_{k-1}. \tag{2.6}$$

The first term in the integral comes from (2.3) and represents the probability density function (pdf) of the model forecast. The second term represents the probability of the analyzed model state at the previous time step, given all measurements up to that time step.

Given (2.3) describes a first order Markov process, $p(x_k|x_{k-1}, Z_{k-1}) = p(x_k|x_{k-1})$, and (2.5) can be rewritten in the following recursive form

$$p(x_k|Z_k) \propto p(z_k|x_k) \int p(x_k|x_{k-1}) p(x_{k-1}|Z_{k-1}) dx_{k-1}. \tag{2.7}$$

The term on the left hand side, $p(x_k|Z_k)$, is the pdf of the analysis at time k. The first factor on the right hand side is the likelihood of the most recent observation (z_k) and is determined by the measurement model, (2.4), and the probability density function p_w of the model error. The transitional density, $p(x_k|x_{k-1})$, controls the evolution of uncertainty in the model state and is determined by the state equation, (2.3), and the probability density function p_v of the measurement error. This term represents the evolution of the model forecast's uncertainty from time $k-1$ to k. The third term on the right hand side, $p(x_{k-1}|Z_{k-1})$, represents the analysis at the previous time step.

The constant of proportionality is fixed by requiring $\int p(x_k|Z_k)dx_k = 1$. Although (2.7) is the conceptual solution to the non-linear filtering problem, it is generally not practical to solve this equation directly to find $p(x_k|Z_k)$.

In practical terms equation (2.7) defines the (model) state x at time k given a (model) forecast for the state and the measurements at time k. The practical estimation of $p(x_k|Z_k)$ will be covered in the following two sections. The estimate depends on the data assimilation method, which itself is determined by the specific form of the functions f and h, and the pdf of the model and observation errors.

The recursive equation for the analysis pdf $p(x_k|Z_k)$ can also be obtained directly as follows (Ristic et al. [2004]):

$$\begin{aligned} p(x_k|Z_k) &= p(x_k|z_k, Z_{k-1}) \\ &= \frac{p(z_k|x_k, Z_{k-1})p(x_k|Z_{k-1})}{p(z_k|Z_{k-1})} \\ &= \frac{p(z_k|x_k)p(x_k|Z_{k-1})}{p(z_k|Z_{k-1})} \end{aligned} \tag{2.8}$$

where the normalizing constant is given by

$$p(z_k|Z_{k-1}) = \int p(z_k|x_k)p(x_k|Z_{k-1})dx_k. \tag{2.9}$$

Equations (2.3), (2.4), (2.6), (2.8) and (2.9) form basis of all sequential filtering methods designed to estimate $p(x_k|Z_k)$.

If f and h are linear, and $x_{k-1}|Z_{k-1}$, v_k , and w_k are all Gaussian it can be shown that $x_k|Z_k$ is also Gaussian. This means that the posterior density at every time step is completely defined by conditional mean and variance of $x_k|Z_k$. This is precisely what is estimated by the linear Kalman Filter as explained below. If f or h, or both, are nonlinear then $p(x_{k-1}|Z_{k-1})$ can have a non-Gaussian form and $p(x_k|Z_k)$ has to be estimated with a nonlinear filter, e.g. Ensemble Kalman filter or particle filter. The same is true if v_k or w_k, or both, have non-Gaussian densities.

2.3 Kalman Filters

The origin of the Kalman Filter (KF) goes back to the fundamental problem in signal processing of separating signal from noise. Kalman solved this problem in state-space for time-varying systems [Kalman, 1960]. The resulting filter technique is called the linear Kalman Filter (KF) and is the optimal solution to a linear-quadratic problem (a linear system with a quadratic cost function, Grewal and Andrews [2001]) of estimating the current state of a linear state-space system.

2.3.1 The Linear Kalman Filter

In the linear case, the discretized form of the state equation (2.3) and observation equation (2.4) can be written as

$$x_k = F_{k-1}x_{k-1} + v_{k-1} \tag{2.10}$$

and

$$z_k = H_k x_k + w_k, \tag{2.11}$$

where F_k comes from the discretized linear model and H_k denotes the linear observation operator which simply selects linear combinations of the elements of the state to approximate observations that are available. The model noise v_k and observation noise w_k are assumed to be independent, normal distributed, and to have zero mean with known covariance Q_k and R_k respectively:

$$v_k \sim \mathcal{N}(0, Q_k) \tag{2.12}$$

and

$$w_k \sim \mathcal{N}(0, R_k). \tag{2.13}$$

(Note the definition of the multivariate normal density is

$$\phi_N(x;\mu,\Sigma) \stackrel{\text{def}}{=} (2\pi)^{-n/2}|\Sigma|^{-1/2}\ \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right] \tag{2.14}$$

where μ, Σ are the mean and variance of the distribution and n is the dimension of x. In the Stommel model n is the spatial discretization and set to $n = 30$). The covariance matrices Q_k and R_k describe the assumed or known spatial structure of the model and observation errors.

It is well known (Ristic et al. [2004]) that if

$$p(x_{k-1}|Z_{k-1}) = \phi_N(x_{k-1}; x_{k-1|k-1}, P_{k-1|k-1}) \tag{2.15}$$

and

$$p(z_k|x_k) = \phi_N(z_k; x_k, R_k) \tag{2.16}$$

then the density of the forecast and the analysis at time k are both multivariate normal:

$$p(x_k|Z_{k-1}) = \phi_N(x_k; x_{k|k-1}, P_{k|k-1}) \tag{2.17}$$

$$p(x_k|Z_k) = \phi_N(x_k; x_{k|k}, P_{k|k}), \tag{2.18}$$

where P is the covariance matrix. The subscripts of the form $k|k-1$ and $k|k$ denote the prediction and analysis (after data assimilation) respectively of the variable. To

update these densities we need to update the conditional means of x_k given the observations up to time $k-1$ and up to time k. Estimating the conditional mean is equivalent to minimizing the functional

$$\begin{aligned} J(x_k) &= \frac{1}{2}\left[(x_k - x_{k|k-1})^T P_{k|k-1}^{-1}(x_k - x_{k|k-1})\right] + \\ &\quad \frac{1}{2}\left[(z_k - H_k x_k)^T R_k^{-1}(z_k - H_k x_k)\right]. \end{aligned} \tag{2.19}$$

The derivative of $J(x_k)$ with respect to x_k gives

$$\frac{\partial J(x_k)}{\partial x_k} = -P_{k|k-1}^{-1}(x_k - x_{k|k-1}) - H_k^T R_k^{-1}(z_k - H_k x_k). \tag{2.20}$$

Solving equation (2.20) for $\frac{\partial J(x_k)}{\partial x_k} = 0$ yields

$$x_{k|k} = x_{k|k-1} + P_{k|k-1}H_k^T(H_k P_{k|k-1} H_k^T + R_k)^{-1}(z_k - H_k x_{k|k-1}) \tag{2.21}$$

which is equivalent to (2.25). Thus, the conditional mean of x_k can be estimated using the following algorithm which is known as the Kalman Filter:

Forecast Step:

$$x_{k|k-1} = F_{k-1} x_{k-1|k-1} \tag{2.22}$$

$$P_{k|k-1} = Q_{k-1} + F_{k-1} P_{k-1|k-1} F_{k-1}^T \tag{2.23}$$

$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \tag{2.24}$$

Analysis Step:

$$x_{k|k} = x_{k|k-1} + K_k(z_k - H_k x_{k|k-1}) \tag{2.25}$$

$$P_{k|k} = P_{k|k-1} + [I - K_k H_k] P_{k|k-1} \tag{2.26}$$

Note that to apply the KF, the initial state ($x_0 \stackrel{\text{def}}{=} x_{0|0}$) and its covariance ($P_0 \stackrel{\text{def}}{=} P_{0|0}$) must be specified.

The terms $P_{k|k-1}$ and $P_{k|k}$ denote the covariance matrices of the model forecast and analyzed state at time k respectively. The Kalman gain K balances the variance of the model estimate against the sum of the variances of the model estimate and observations; it also determines how the information of the measurements is spatially spread over the model state. In the simplest case of a one dimensional state vector ($n = 1$) the Kalman gain is simply a scalar factor given by $K = (\sigma_{k|k-1}^2)/(\sigma_{k|k-1}^2 + \sigma_{obs}^2)$. In the multivariate case, K is a matrix of dimension $n \times N_{obs}$ and the term

$K_k(z_k - H_k x_{k|k-1})$ can be viewed as a nudging term that is added to the model forecast to yield the optimal analyzed model state (given the statistics of v and w).

Table 2.1 lists the model parameters used for the assimilation experiments using the Stommel model. The observation noise covariance matrix R has a diagonal structure with σ^2_{obs} on the diagonal. The model noise covariance matrix Q reflects our assumption that the model errors are autocorrelated. Specifically, Q is a Toeplitz matrix with σ^2_{model} on the diagonal. The off diagonal elements are of the form $\sigma^2_{model}\alpha^{|k|}$ where $\alpha = 0.99$, $k = 0$ on the diagonal, ± 1 on the leading sub-diagonals and so on.

DA method	Panel	λ	N_{obs}	$N_{ensembles}$	$N_{particles}$	σ_{obs}	σ_{model}
True model	2.1a	0.0				0	
KF	2.2c	0.1	2			0.1	0.01
KF	2.2d	0.1	30			0.1	0.01
EnKF	2.4e	0.1	2	100		0.1	0.01
EnKF	2.4f	0.1	30	100		0.1	0.01
PF	2.8g	0.1	2		100	0.1	0.01
PF	2.8h	0.1	30		100	0.1	0.01
PF/MCMC	2.9j	0.1	30		100	0.1	0.01

Table 2.1: Model parameters used in the Stommel model. λ is the non-dimensional friction parameter. The friction parameter used in assimilation experiments was set to $\lambda = 0.1$ and this differs from the value used in the true model. (This is done on purpose to make the task of data assimilation not too trivial for the assimilation methods.) N_{obs} is the number of assimilated observations, $N_{ensembles}$ is the ensemble size used in the ensemble Kalman filter (EnKF), $N_{particles}$ is the number of particles used in the particle filter (PF), σ_{obs} is the standard deviation of the observation noise and σ_{model} is the standard deviation of the model noise. The structure of the observation and model error covariance matrices is described in the text.

Figure (2.2) shows the true model state and estimates recovered by data assimilation using the Kalman filter using only two (panel c) and 30 (panel d) observation locations. The two locations in panel c where chosen to monitor the central and western regions of the ocean basin. The results confirm that the number of available observations does play an important role in recovering the true model state. Indeed

the Kalman Filter can recover the true state reasonably well with only two observation locations (and almost perfectly with 30 observation locations, as expected).

In oceanography most models of interest are highly non-linear and thus the KF is not useful. One possible way to deal with the special case of additive, Gaussian noise in (2.3) and (2.4) is based on using the linearized forms of f_k and h_k to propagate the required error covariance matrices needed by the KF through time. This leads to the Extended Kalman Filter as described below.

2.3.2 The Extended Kalman Filter

In 1960 Stanley F. Schmidt extended the linear Kalman filter to solve the problem of trajectory estimation. The resulting filter was originally named Kalman-Schmidt filter and is today known as the Extended Kalman Filter (Grewal and Andrews [2001]). The Extended Kalman Filter (EKF) is based on the local linearization of the non-linear functions f_{k-1} and h_k in (2.3) and (2.4) respectively. The derivation of the EKF is identical to the KF and can be found in various literature, e.g. Evensen [1994], Brown and Hwang [1997], Grewal and Andrews [2001], Welch and Bishop [2001], Zarchan and Musoff [2002]. The recursive form of the EKF is given by

Forecast Step:

$$x_{k|k-1} = \hat{F}_{k-1}(x_{k-1|k-1}) \tag{2.27}$$

$$P_{k|k-1} = Q_{k-1} + \hat{F}_{k-1} P_{k-1|k-1} \hat{F}_{k-1}^T \tag{2.28}$$

$$K_k = P_{k|k-1} \hat{H}_k^T (\hat{H}_k P_{k|k-1} \hat{H}_k^T + R_k)^{-1} \tag{2.29}$$

Analysis Step:

$$x_{k|k} = x_{k|k-1} + K_k(z_k - h_k(x_{k|k-1})) \tag{2.30}$$

$$P_{k|k} = P_{k|k-1} + \left[I - K_k \hat{H}_k \right] P_{k|k-1} \tag{2.31}$$

where $\hat{F}_{k-1}$ and $\hat{H}_k$ are the discretized, local linearization of the non-linear functions f and h, i.e. the tangent linear models (TLM):

$$\hat{F}_{k-1} = \left[\nabla_{x_{k-1}} f_{k-1}^T(x_{k-1})\right]^T \Big|_{x_{k-1}=x_{k-1|k-1}} \tag{2.32}$$

$$\hat{H}_k = \left[\nabla_{x_k} h_k^T(x_k)\right]^T \Big|_{x_k=x_{k|k-1}} \tag{2.33}$$

where

$$\nabla_{x_k} = \left[\frac{\partial}{\partial x_{k,1}} \cdots \frac{\partial}{\partial x_{k,n}}\right]^T \tag{2.34}$$

with $x_{k,i},\ i = 1 \ldots n$ denoting the ith component of the state vector x_k.

It should be noted that the EKF assumes (2.8) is Gaussian. It has been shown however that for some of the most basic oceanographic applications, such as assimilation of altimeter data into quasi-geostrophic ocean models, the assumption of Gaussianity often does not hold and the EKF can quickly diverge (Evensen [1994]).

A more recent example of filter divergence of the EKF is presented in the paper of Kuznetsov et al. [2003]. In their study the EKF was used to assimilate Lagrangian data, i.e. the trajectories of drifters, in a simple co-rotating vortex model. It was shown that the EKF diverges quickly as the model noise (accounting for non-linear effects neglected by the model) increases.

Related filters based on the EKF are the Singular Evolutive Extended Kalman (SEEK) filter and the Reduced Rank Square Root Kalman Filter (RRSQKF). Both filters make use of matrix methods to reduce the rank of the error covariance matrix P_k and make the filter more stable with respect to strong non-linear system dynamics. The SEEK filter uses a factorization of P_k into its largest eigenvectors/values (Hoteit et al. [2000]). The RRSQKF uses a combination of a square root algorithm as well as a singular value decomposition to reduce the rank of P_k (Madsen and Canizares [1999]).

No data assimilation example using the EKF is given for the linear Stommel

model. Since the Stommel model is *linear* the EKF will give the same results as the linear KF.

2.3.3 The Ensemble Kalman Filter

In an attempt to overcome the shortcomings of the EKF, Evensen introduced the Ensemble Kalman Filter (EnKF) by blending the EKF with Monte Carlo based methods (Evensen [1994]). In the EnKF, an ensemble of forecasts evolves under the system's non-linear dynamics. The ensemble resulting from this forecast step has, in general, a non-Gaussian distribution (2.6), i.e. there is no guarantee that the distribution can be represented accurately by its mean and the variance. At the measurement time, the first two moments are calculated from the ensemble and then used to combine the current ensemble with the observations. A new ensemble is thus constructed which has a Gaussian distribution. The new ensemble is then carried forward one time step using the system's dynamics.

The derivation of the EnKF is, in principle, very similar to the linear Kalman Filter. The main differences are that (1) instead of one state vector, an ensemble of state vectors is carried forward in time, (2) the error covariance matrix of the ensemble is computed directly from the ensemble.

Let x_k^l be one of N ensemble members and let x_0^l denote the ensemble of N initial states where $l = 1, \ldots, N$ indexes the individual ensemble members. The complete recursive formalism for the EnKF is summarized below:

Forecast Step:

$$x_{k|k-1}^l = f_{k-1}(x_{k-1|k-1}^l) + v_{k-1}^l \tag{2.35}$$

$$\bar{x}_{k|k-1} = \frac{1}{N}\sum_{l=1}^{N} x_{k|k-1}^l \tag{2.36}$$

$$P_{k|k-1} = \frac{1}{N-1}\sum_{l=1}^{N}(x_{k|k-1}^l - \bar{x}_{k|k-1})(x_{k|k-1}^l - \bar{x}_{k|k-1})^T \tag{2.37}$$

$$K_k = P_{k|k-1}H_k^T(H_k P_{k|k-1} H_k^T + R_k)^{-1} \tag{2.38}$$

Analysis Step:

$$x^l_{k|k} = x^l_{k|k-1} + K_k(z_k - H_k x^l_{k|k-1}) \tag{2.39}$$

It should be noted that the additive model noise v^l_k represents the spatial structure of the model noise scaled with Gaussian noise, i.e.

$$\begin{aligned} v^l_k &\equiv Q_k^{1/2} u^l_k \\ u^l_k &\sim \mathcal{N}(0,1) \qquad\qquad l = 1, ..., N. \end{aligned} \tag{2.40}$$

The computation of the ensemble mean, $\bar{x}_{k|k}$, and the error covariance matrix $P_{k|k}$ are of no consequence for the EnKF because they are not used. However, they are typically computed for visualization and analysis purposes

$$\bar{x}_{k|k} = \frac{1}{N}\sum_{l=1}^{N} x^l_{k|k} \tag{2.41}$$

$$P_{k|k} = \frac{1}{N-1}\sum_{l=1}^{N}(x^l_{k|k} - \bar{x}_{k|k})(x^l_{k|k} - \bar{x}_{k|k})^T. \tag{2.42}$$

The analysis step of the EnKF produces an ensemble from the posterior conditional distribution (2.9) only for linear dynamics and Gaussian errors. Blending the non-Gaussian predictive density with the Gaussian density from the previous analysis step can cause problems of more than theoretical concern (Kim et al. [2003], Evensen and van Leeuwen [2000]). For cases of extreme non-Gaussianity, the EnKF can produce unrealistic results and sometimes the filter may diverge.

A schematic of the EnKF as implemented in this thesis is given in Figure 2.3 and Figure 2.4 illustrates how well the true state (panel a) can be recovered by this assimilation scheme. Given a sufficient large number of ensemble members, (e.g., $N_{ensemble} = 100$), the EnKF produces almost the same result as the KF. This is to be expected because the Stommel model is linear and thus the KF is the optimal assimilation method.

2.4 Particle Filters

The particle filter (PF) uses sequential Monte Carlo techniques to numerically solve equation (2.8). In contrast to the EnKF, the PF does not form a linear combination of the current ensemble and the observations to estimate the state. In addition, the estimated posterior density (2.8) is not required to be Gaussian.[1]

To capture the posterior density reasonably well, the PF requires large ensemble sizes (e.g., of order 1000 or larger) unless one can impose constraints on the probability density function by limiting the data assimilation to lower dimensional sub spaces of the model or limited (smaller) spatial areas, or both.

With present computer resources, a large ensemble size is not a major limitation for low dimensional problems (e.g., state dimensions less than 100) and an ensemble size of order 10^5 can easily be used. However, the typical ocean model has a state dimension of order 10^4 to 10^7, and given every ensemble member (i.e. model state) has to be integrated forward in time by the ocean model, the ensemble size is limited by the computational load, typically to an order of 10^2. For practical reasons the PF can generally only be applied to small sub domains of interest (using some form of localization). Such sub domains should be characterized by strong non-linear interactions where the assumption of Gaussianity does not hold, otherwise the additional computational cost is not justified. For the rest of the domain less computationally demanding assimilation methods could be used.[2]

Monte Carlo Integration is described in the next subsection. This is followed by a description of Sequential Importance Sampling which is the simplest PF. Two common improvements to Sequential Importance Sampling, namely Sequential Importance Resampling (SIR) and the Markov Chain Monte Carlo (MCMC) step, are then described.

[1] The terms "ensemble member" and "particle" are to be taken synonymous for the rest of this thesis.

[2] An ensemble run (100 members, simulation time of 10 model days) using the Océan PArallélisé (OPA) model, state dimension of $\mathcal{O}(10^6)$, 1/4 degree resolution of the North Atlantic, 40 CPUs) takes roughly 10 real time hours.

2.4.1 Monte Carlo Integration

Monte Carlo (MC) integration is the basis of all sequential MC methods (Doucet et al. [2001b], Ristic et al. [2004]) and all particle filters. To illustrate the approach, assume we want to evaluate the multidimensional integral

$$I = \int f(x)\pi(x)dx \tag{2.43}$$

where I is the expected value of $f(x)$ where x is the realized value of a random variable with probability density function $\pi(x)$. It follows $\pi(x) \geq 0$ and $\int \pi(x)dx = 1$. If it is possible to draw N samples $\{x^i; i = 1, \ldots, N\}$ from $\pi(x)$, where the superscript denotes the ith ensemble member, the MC estimate of I is the sample mean:

$$I_N = \frac{1}{N}\sum_{i=1}^{N} f(x^i). \tag{2.44}$$

As $N \to \infty$, then I_N will converge to I. It should be noted that the integral in (2.43) has the same form as the integral on the right hand side of (2.7) if $\pi(x)$ is taken to be the posterior density, $p(x_{k-1}|Z_{k-1})$.

In many cases $\pi(x)$ is difficult to sample from. Assuming the density function $q(x)$ is known and easy to sample from, and $q(x)$ is similar to $\pi(x)$, then samples can be drawn from $q(x)$ in order to estimate I using

$$I = \int f(x)\frac{\pi(x)}{q(x)}q(x)dx \tag{2.45}$$

provided that $\pi(x)/q(x)$ has an upper bound and the support of $q(x)$ includes the support of $\pi(x)$, i.e., $\pi(x) > 0 \Rightarrow q(x) > 0$. More specifically, the MC estimate of I is given by sampling N independent ensemble members $\{x^i; i = 1, \ldots, N\}$ from $q(x)$ and using

$$I_N = \frac{1}{N}\sum_{i=1}^{N} f(x^i)\tilde{w}(x^i) \tag{2.46}$$

where

$$\tilde{w}(x^i) = \frac{\pi(x^i)}{q(x^i)}. \tag{2.47}$$

The $\tilde{w}(x^i)$ are called importance weights.

Given the normalizing factor of the density $\pi(x)$ is usually unknown, the importance weights need to be normalized. Then I_N is estimated as follows:

$$I_N = \frac{\frac{1}{N}\sum_{i=1}^{N} f(x^i)\tilde{w}(x^i)}{\frac{1}{N}\sum_{j=1}^{N} \tilde{w}(x^j)} = \sum_{i=1}^{N} f(x^i)w(x^i) \tag{2.48}$$

where

$$w(x^i) = \frac{\tilde{w}(x^i)}{\sum_{j=1}^{N} \tilde{w}(x^j)}. \tag{2.49}$$

This technique is termed Importance Sampling. The normalized importance weights $w(x^i)$ can be viewed as weighting factors for the corresponding ensemble member; members with high weights are closer to the true state and thus more "important".

The density $q(x)$ is termed the *importance* or *proposal* density or *transitional prior* and its choice is critical to the success of Monte Carlo Integration.

2.4.2 Sequential Importance Sampling

The idea of evaluating a multidimensional integral using importance weights (e.g., equations (2.46) to (2.49)) is the basis of Sequential Importance Sampling (SIS) which itself is fundamental to particle filtering.

The idea of the SIS filter is to characterize the posterior density by an ensemble of model states and their associated weights. As the ensemble becomes very large the pdf of the ensemble members approaches the true posterior pdf of the model state. The time evolution of the ensemble of random model states is given by (2.35) and (2.40) which are also used by the Ensemble Kalman filter.

The following derivation of SIS follows closely that of Ristic et al. [2004]. Let $X_k = \{x_j, j = 0, ..., k\}$ and $Z_k = \{z_j, j = 0, ...k\}$ denote the sequence of all states

and observations respectively up to time k. The joint posterior density of all the states up to time k, given all available observations, is given by $p(X_k|Z_k)$ from which can be found the marginal density $p(x_k|Z_k)$. The ensemble $\{X_k^i, w_k^i\}_{i=1}^N$ represents the joint posterior $p(X_k|Z_k)$ where $\{X_k^i, i = 1, ...N\}$ is an ensemble (N sequences of states) with associated weights $\{w_k^i, i = 1, ..., N\}$. The weights are normalized such that $\Sigma_i w_k^i = 1$. The joint posterior density at time k is then approximated by

$$p(X_k|Z_k) \approx \sum_{i=1}^{N} w_k^i \delta(X_k - X_k^i) \tag{2.50}$$

which is a discrete approximation of the true posterior $p(X_k|Z_k)$. The weights w_k^i are computed according to (2.49) as described in the previous section.

If the ensemble members X_k^i are drawn from the importance density $q(X_k|Z_k)$ then it follows from (2.47) that

$$w_k^i \propto \frac{p(X_k^i|Z_k)}{q(X_k^i|Z_k)}. \tag{2.51}$$

Assume an ensemble of states is available at time step $k-1$ which approximates $p(X_{k-1}|Z_{k-1})$. Given a new measurement z_k at time k we want to approximate $P(X_k|Z_k)$ by a new ensemble of states. Factorize $q(X_k|Z_k)$ as follows

$$q(X_k|Z_k) \equiv q(x_k|X_{k-1}, Z_k)q(X_{k-1}|Z_{k-1}) \tag{2.52}$$

The ensemble members $X_k^i \sim q(X_k|Z_k)$ can then be drawn by augmenting $X_{k-1}^i \sim q(X_{k-1}^i|Z_{k-1})$ with the new state $x_k^i \sim q(x_k|X_{k-1}, Z_k)$. Applying (2.8) to the posterior

density $p(X_k|Z_k)$ gives

$$\begin{aligned} p(X_k|Z_k) &= \frac{p(z_k|X_k, Z_{k-1})p(X_k|Z_{k-1})}{p(z_k|Z_{k-1})} \\ &= \frac{p(z_k|X_k, Z_{k-1})p(x_k|X_{k-1}, Z_{k-1})p(X_{k-1}|Z_{k-1})}{p(z_k|Z_{k-1})} \\ &= \frac{p(z_k|x_k)p(x_k|x_{k-1})}{p(z_k|Z_{k-1})}p(X_{k-1}|Z_{k-1}) \\ &\propto p(z_k|x_k)p(x_k|x_{k-1})p(X_{k-1}|Z_{k-1}). \end{aligned} \tag{2.53}$$

Substitution of (2.53) and (2.52) into (2.51) gives

$$\begin{aligned} w_k^i &\propto \frac{p(z_k|x_k)p(x_k|x_{k-1})p(X_{k-1}|Z_{k-1})}{q(x_k|X_{k-1}, Z_k)q(X_{k-1}, Z_{k-1})} \\ &= w_{k-1}^i \frac{p(z_k|x_k)p(x_k|x_{k-1})}{q(x_k|X_{k-1}, Z_k)}. \end{aligned} \tag{2.54}$$

Usually only the filtered estimate of the posterior $p(x_k|Z_k)$ is required. Then $q(x_k|X_{k-1}, Z_k) = q(x_k|x_{k-1}, z_k)$, i.e. the importance density is only dependent on x_{k-1} and z_k. Furthermore only x_k^i is needed and X_{k-1}^i and Z_{k-1} are obsolete. The normalized weights then take the form

$$w_k^i \propto w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)} \tag{2.55}$$

and the filtered posterior density $p(x_k|Z_k)$ is approximated by

$$p(x_k|Z_k) \approx \sum_{i=1}^{N} w_k^i \delta(x_k - x_k^i) \tag{2.56}$$

with weights w_k^i given by (2.55). As $N \to \infty$ this estimate approaches the true posterior density, $p(x_k|Z_k)$.

The practical difficulty with SIS is that it is difficult to estimate the reliability of I_N in (2.48), (MacKay [2005]). One reason is that the variance of I_N is unknown

beforehand, and another is the variance of I_N is difficult to estimate because the variances of $f(x^i)\tilde{w}(x^i)$ and $\tilde{w}(x^i)$ may not be good guides to the true variances of $\sum_{i=1}^{N} f(x^i)\tilde{w}(x^i)$ and $\sum_{j=1}^{N} \tilde{w}(x^j)$. If $q(x)$ is too small in a region where $|f(x)\pi(x)|$ is large then even sampling a large number of ensemble members, x^i , from $q(x)$ might not guarantee that one or more members will fall into that region. In that case the estimate of I_N would be wrong.

In most cases $q(x)$ is only an approximation to $p(x_k|Z_k)$. In such cases it can be shown that the variance of the importance weights can only increase over time (Kong et al. [1994]). A complete proof can be found in appendix A of Hermann [2002]. This leads to a problem known as "degeneracy". In practical terms it means that after a certain number of time steps, all but one ensemble member will have negligible weight. The degeneracy problem is inherent to SIS and cannot be avoided. The speed with which the degeneracy progresses can be influenced by a good or bad choice of the proposal distribution (Figure 2.5).

A measure of degeneracy is described by the effective ensemble size N_{eff}:

$$N_{eff} = \frac{1}{\sum_{i=1}^{N}(w_k^i)^2}. \tag{2.57}$$

At one extreme, all of the N ensemble members have the same normalized weight (which must therefore equal $1/N$). In this case $N_{eff} = N$. At the other extreme, all but one ensemble member have zero weight (and thus the remaining ensemble member must have weight of unity). In this case $N_{eff} = 1$. In general, N_{eff} lies between 1 and N. It decreases with the degeneration of the ensemble. If the ensemble degenerates to the point $N_{eff} = 1$, there is no guarantee that the ensemble member with non-zero weight will coincide with the true state and the SIS algorithm will likely diverge with increasing time. The time for the SIS algorithm to diverge depends on model dynamics and the statistical properties of the observation and model errors.

2.4.3 Sequential Importance Resampling

To overcome the degeneracy problem in SIS, a corrective procedure can be used if N_{eff} falls below a prescribed threshold N_{thr}. The basic idea is to eliminate ensemble members with low weights and replace them with copies of ensemble members with high weights. Given ensemble members with high weights are "reused" (i.e. resampled) to form the new ensemble, this procedure is named Sequential Importance Resampling (SIR). The simplest form of resampling is to generate a new ensemble $\{x_k^{i*}\}_{i=1}^N$ by resampling (with replacement) from (2.56) such that $P\{x_k^{i*} = x_k^j\} = w_k^j$. Figure 2.6 illustrates the idea for a particularly simple importance density, $q(x_k|x_{k-1}^i, z_k) = p(x_k|x_{k-1}^i)$.

The combination of the SIS and the resampling of high weighted and elimination of low weighted ensemble members is called the General Particle Filter. All PF are based on SIS and resampling; they differ only by the algorithms used to compute the weights and the method used for resampling (Doucet et al. [2001a], Ristic et al. [2004], Kitagawa [1996]).

One of the most important questions in the design of particle filters is the choice of the importance density $q(x_k|x_{k-1}^i, z_k)$. The optimal choice is

$$\begin{aligned} q(x_k|x_{k-1}^i, z_k)_{optimal} &= p(x_k|x_{k-1}^i, z_k) \\ &= \frac{p(z_k|x_k, x_{k-1}^i)p(x_k|x_{k-1}^i)}{p(z_k|x_{k-1}^i)}. \end{aligned} \tag{2.58}$$

Substitution of (2.58) into (2.55) gives

$$w_k^i \propto w_{k-1}^i p(z_k|x_{k-1}^i). \tag{2.59}$$

In order to use the optimal importance function one has to sample from $p(x_k|x_{k-1}^i, z_k)$ and evaluate

$$p(z_k|x_{k-1}^i) = \int p(z_k|x_k)p(x_k|x_{k-1}^i)dx_k. \tag{2.60}$$

The analytic evaluation of this expression is not possible with the exception of two cases: (i) when the optimal importance function can be used directly, i.e. models for which $p(z_k|x^i_{k-1})$ is Gaussian, and (ii) when the distribution of x_k is discrete. It can be shown that the first case is identical to the linear Kalman Filter (Ristic et al. [2004]). In the second case the integral in (2.60) reduces to a sum and sampling from $q(x_k|x^i_{k-1}, z_k)_{optimal}$ is possible.

Given the optimal importance function is usually not available, a known but suboptimal importance density function $q(x_k|x^i_{k-1}, z_k)$ is chosen. A very popular suboptimal choice is the so-called prior distribution:

$$q(x_k|x^i_{k-1}, z_k) = p(x_k|x^i_{k-1}). \tag{2.61}$$

Given an additive noise model (2.3), sampling from the prior, $p(x_k|x^i_{k-1})$, amounts to sampling from the distribution of the model error, i.e. $p(v_{k-1})$ or

$$p(x_k|x^i_{k-1}) = \mathcal{N}(x_k, Q_{k-1}). \tag{2.62}$$

Substitution of (2.61) into (2.55) gives

$$w^i_k \propto w^i_{k-1} p(z_k|x^i_k). \tag{2.63}$$

The weights w^i_k are determined by the likelihood $p(z_k|x^i_k)$, i.e. the proximity of the ensemble members to the observations (see equation (2.4)). The same does not apply to equation (2.59). Note that the right hand sides of (2.61) and (2.63) can be calculated from the state equation (2.3) and measurement equation (2.4) respectively.

One important reason the prior distribution is a poor choice as a proposal distribution is that the current observation z_k is not taken into account. Figure 2.5 illustrates how the choice of $p(x_k|x^i_{k-1})$ can lead to poor performance. In this example the likelihood $p(z_k|x_k)$ is much more peaked than $p(x_k|x^i_{k-1})$ resulting in negligible weights (see (2.63)) for most ensemble members.

A schematic of the SIR filter (i.e. the SIS algorithm with resampling) as implemented in this thesis is shown in Figure 2.7. Pseudo codes for the SIS, General Particle Filter, SIR and resampling are given in Appendix B.

The performance of the SIR filter on the linear Stommel problem is illustrated in Figure 2.8. This figure compares the true state (panel(a)) with the model states recovered by the SIR filter given 2 and 30 observations (panels (g) and (h) respectively). Note the SIR filter recovers well the true state, although not quite as well as the KF or EnKF near the eastern boundary. However, if the SIRF is combined with an MCMC step (see below) the recovered model state is improved and comparable with the results from the KF and EnKF (see Figure 2.9). Tuning of the MCMC step can increase the accuracy further.

2.4.4 Markov Chain Monte Carlo (MCMC)

The previous section described how resampling can reduce the degeneracy problem in the SIS algorithm. However, it can be shown that resampling can, in turn, induce another problem: the loss of diversity among the ensemble members. The problem arises due to the fact that repeated resampling of a finite number of ensemble members through time can lead to the collapse of the ensemble to a single member, i.e. all N members occupy the same point in the state space.

This "particle collapse" is due to the design of the SIS algorithm. The fact that the forecast of each ensemble member includes a random part (via the model error v_k^l, see (2.40)) does not always prevent particle collapse. If the SIS algorithm has discarded all but one particle, the new ensemble will be entirely made up of identical clones. After the forecast for the next time step (see (2.35)), these clones will differ only by their model error v_k^l. Since the statistics are the same for each v_k^l the ensemble does not carry more information than one single ensemble member.

Particle collapse can be avoided by a Markov Chain Monte Carlo (MCMC) step following the SIR step. The MCMC is designed to improve sample diversity without changing the probability distribution of the ensemble, i.e. the density of the ensemble

after the SIR step, and after the subsequent MCMC step, are identical. (A more detailed description of MCMC is given in Appendix C.) Figure 2.9 shows how the accuracy of the recovered model state in the Stommel model example can be improved by MCMC.

2.5 Discussion

The Bayesian framework provides a natural and elegant way of describing and developing data assimilation schemes. Using this framework, a unified description has been given in this chapter of Kalman Filters (linear, extended and ensemble) and particle filters (sequential importance resampling, with and without a Markov Chain Monte Carlo step).

To illustrate the implementation and effectiveness of the various assimilation schemes, a very simple model of the ocean has been introduced: the Stommel model. Although the model is linear, it does support the intensification of western boundary currents, westward propagating Rossby waves, and also an interior ocean in Sverdrup balance.

The assimilation experiments with the Stommel model show that all of the assimilation methods work well and produce comparable results. The assimilation methods all used a "wrong" ocean model (i.e. the friction parameter differed from the true model) and noisy observations. For all assimilation schemes the true model state was recovered with high accuracy. While the EnKF represents a model state only by its first two statistical moments, particle filters are able to present the true pdf of a model state. For many practical applications (e.g. prediction of extreme events) knowledge of the pdf of the model state, particularly the shape of the tails, can be as important as the knowledge of the mean and variance. It is encouraging to note that the SIR filter with an MCMC step performs almost as well as the EnKF. This motivates the investigation described in subsequent chapters of the performance of the particle filter when assimilating data into more complex nonlinear systems.

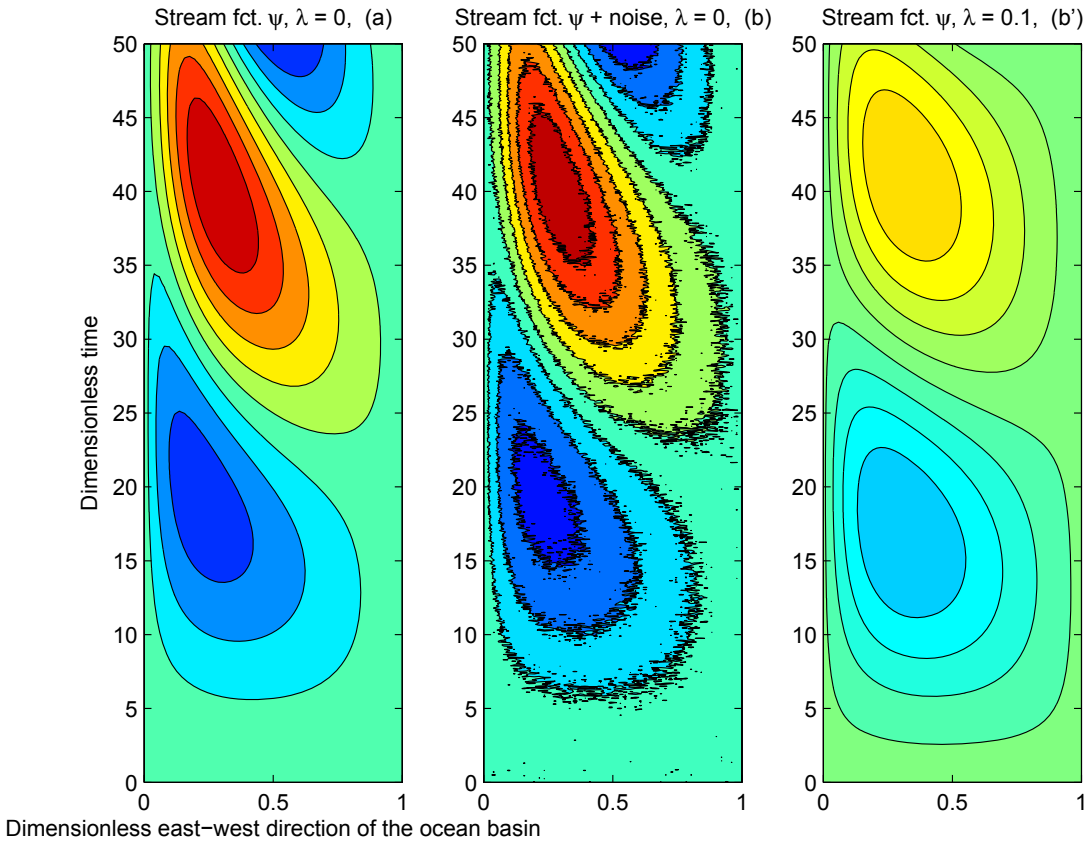

Figure 2.1: Stream function for the Stommel model. Panel (a) shows the true stream function calculated with the friction parameter $\lambda = 0$. Panel (b) shows the "observed" stream function obtained by adding observation error to the true stream function shown in Panel (a). Panel (b') shows the stream function calculated with the friction parameter $\lambda = 0.1$. By comparing panels (a) and (b') it is clear that increasing friction decreases the magnitude of the stream function and increases the width of the western boundary current as expected.

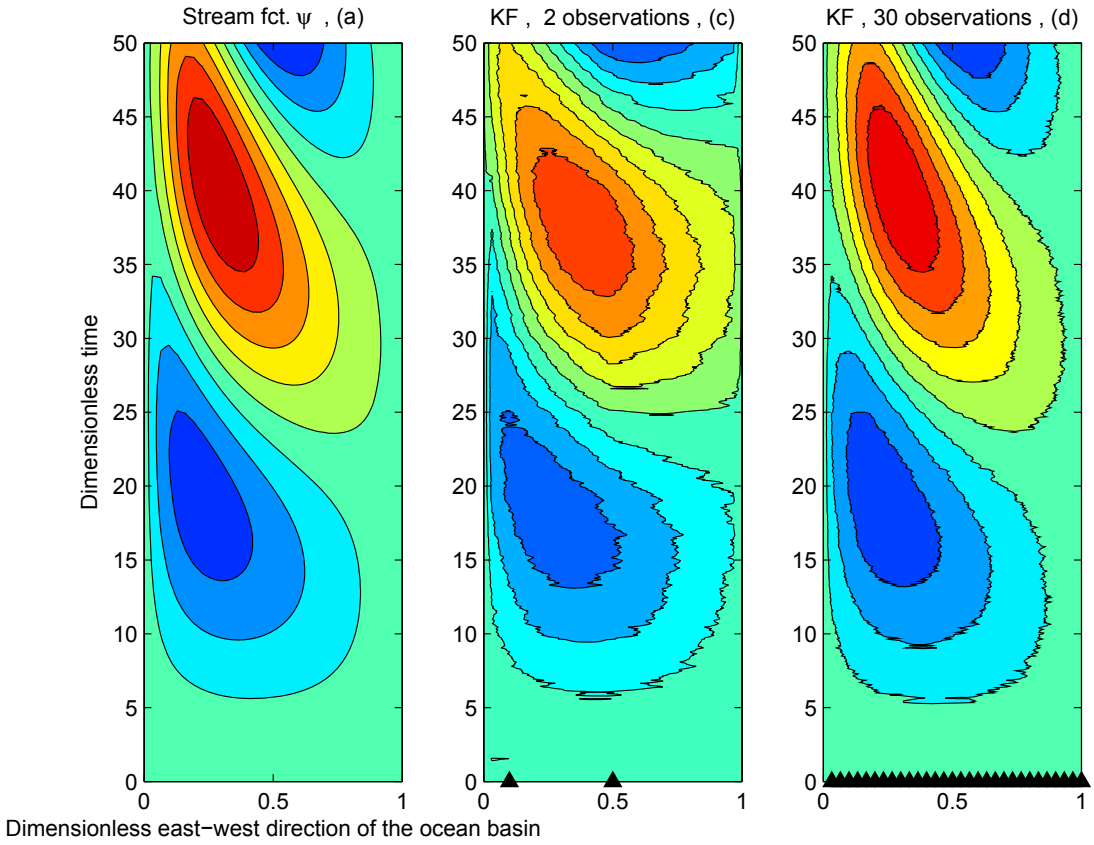

Figure 2.2: True stream function from the Stommel model and estimates from the KF. Panel (a) shows the true stream function ψ calculated from the Stommel model assuming no model error. The remaining two panels show the estimates of ψ from the Kalman Filter. For panel (c) the two observation locations were grid cells 3 and 15; for panel (d) observations were assumed available from all 30 model grid points. The main model parameters used in the twin experiments are listed in Table 2.1.

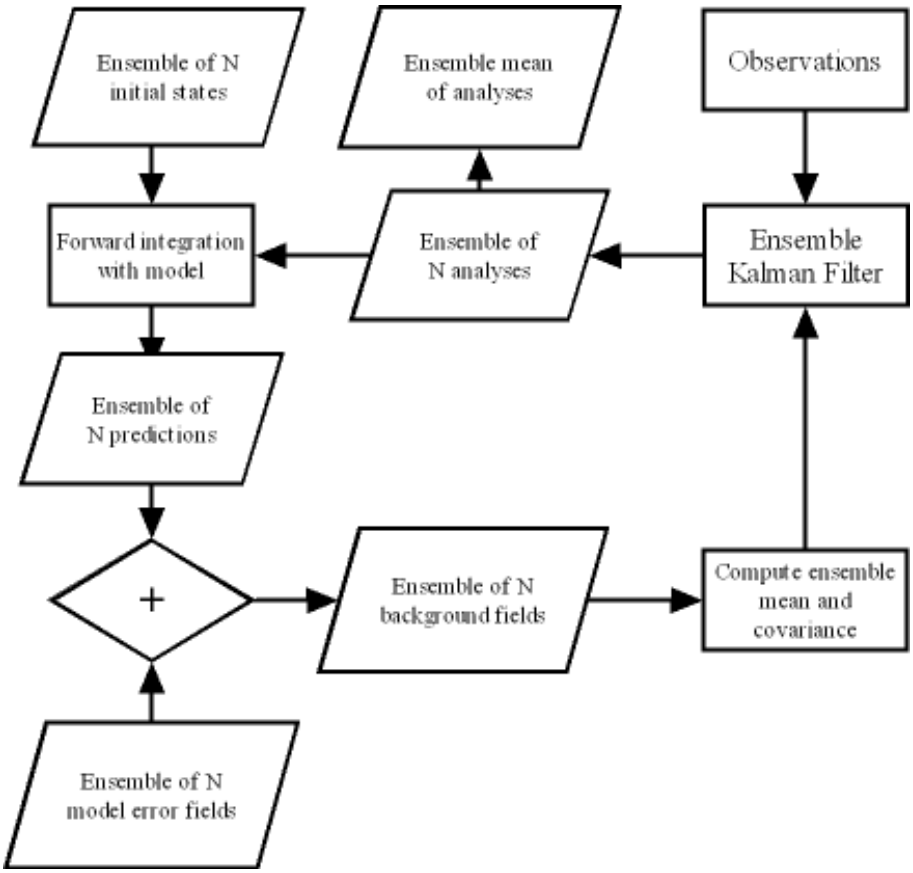

Figure 2.3: Schematic of the EnKF data assimilation procedure as implemented in the present study.

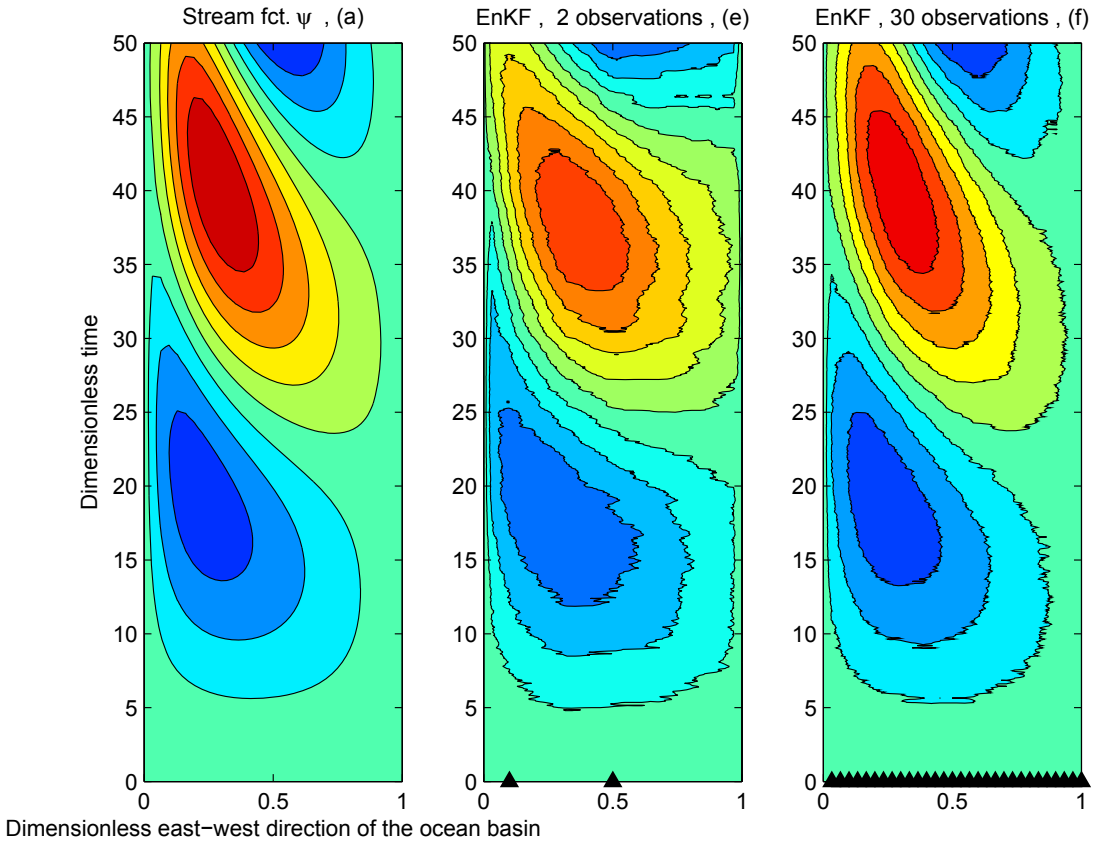

Figure 2.4: True stream function from the Stommel model and estimates from the EnKF. Identical format to Figure 2.2.

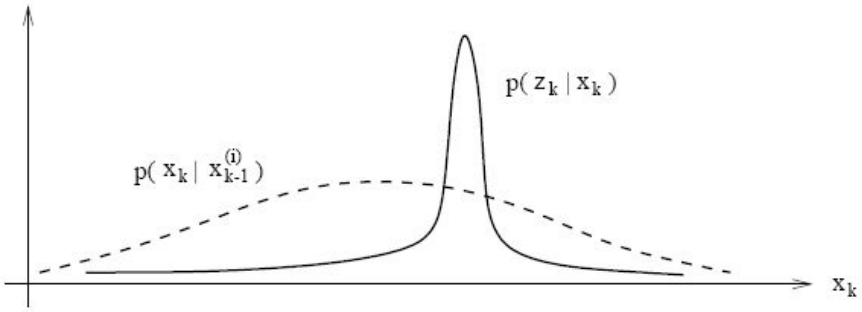

Figure 2.5: Example of how a poor choice of proposal density prior can lead rapidly to degeneracy of the importance weights, the w^i. (Figure from Hermann [2002].) In cases were the model error is much larger than the observation error the prior $p(x_k|x^i_{k-1})$ is a poor choice of proposal density (shown by the pdf with dashed lines and large "spread" compared to the likelihood of the observation, z_k).

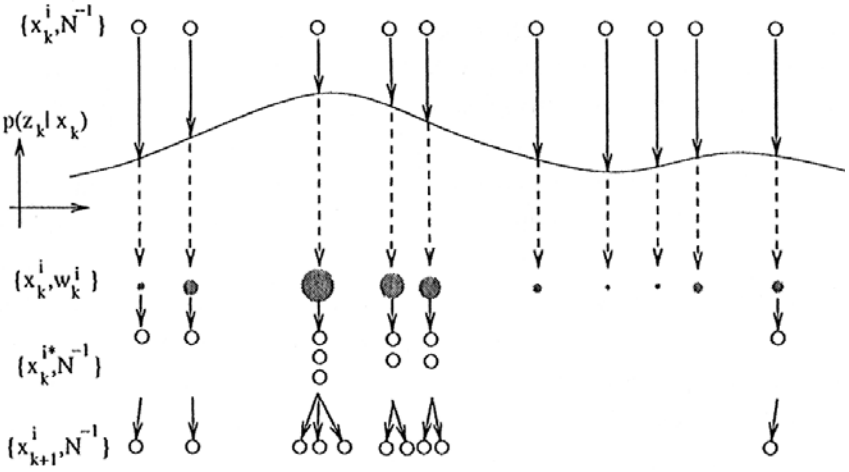

Figure 2.6: Illustration of Sequential Importance Resampling. Initially each ensemble member is assigned the same weight $w_k^i = N^{-1}$ which approximates (2.6). The measurement z_k is then used to compute the importance weight for each ensemble member using (2.55) resulting in the N pairs $\{x_k^i, w_k^i\}$. In the resampling step, the "important" ensemble members are selected more frequently to obtain the approximation of (2.8). The new ensemble members are equally likely and so the state-weight pairs are now $\{x_k^{i*}, N^{-1}\}$. To move to the next time step the model equation is used to give the $\{x_{k+1}^i, N^{-1}\}$ pairs (which can be used to approximate (2.6) at time $k = k+1$. (Figure is from Ristic et al. [2004].)

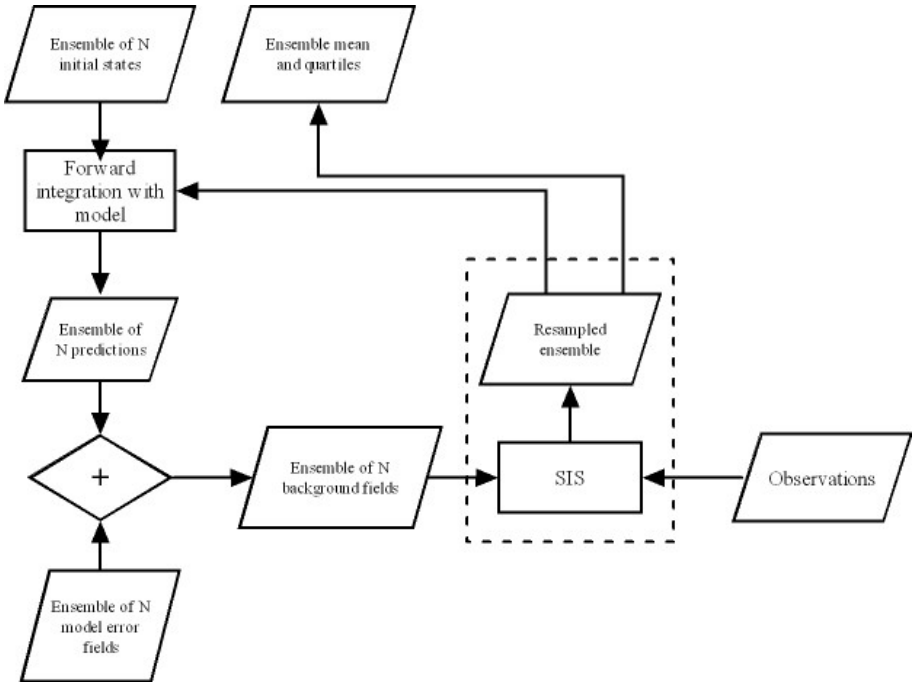

Figure 2.7: Schematic of the SIR filter. This approach is based on the SIS algorithm with resampling. The SIR step is shown by the box enclosed by the dashed line.

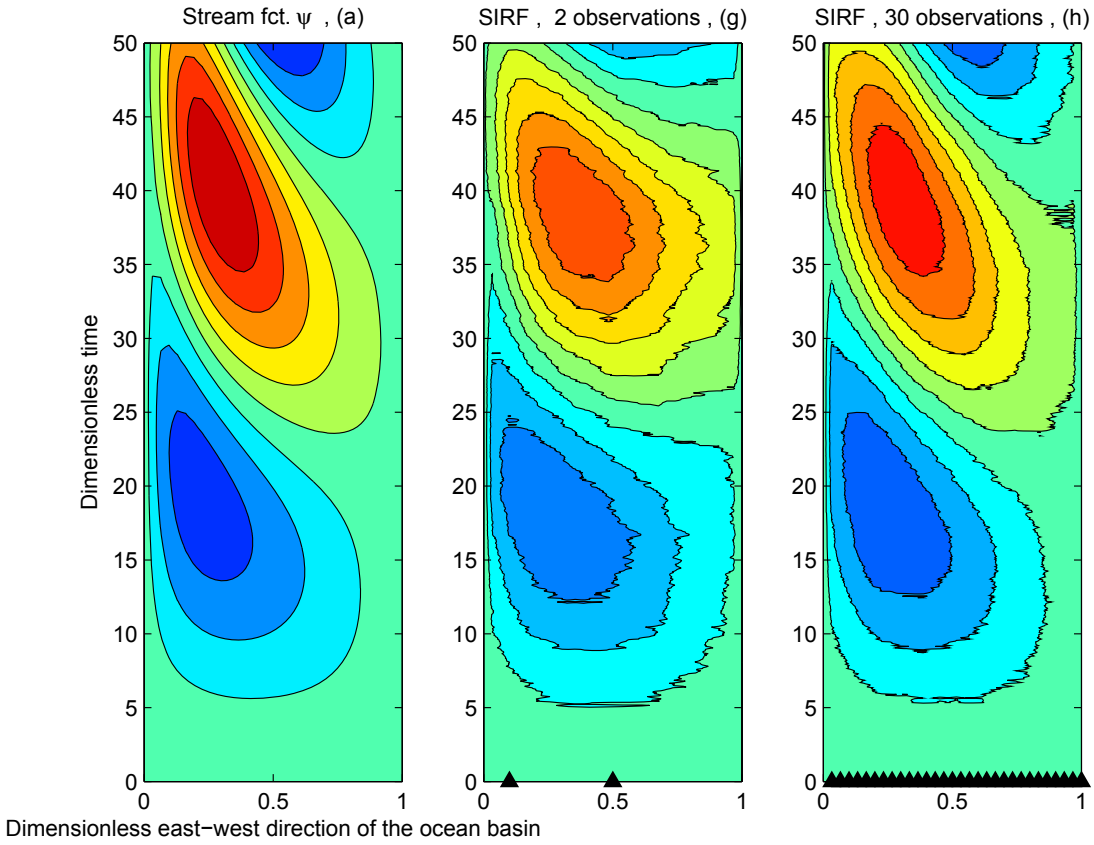

Figure 2.8: True stream function from the Stommel model and estimates from the SIR filter using $N_{particles} = 100$ particles (ensemble members). Identical format to Figure 2.2.

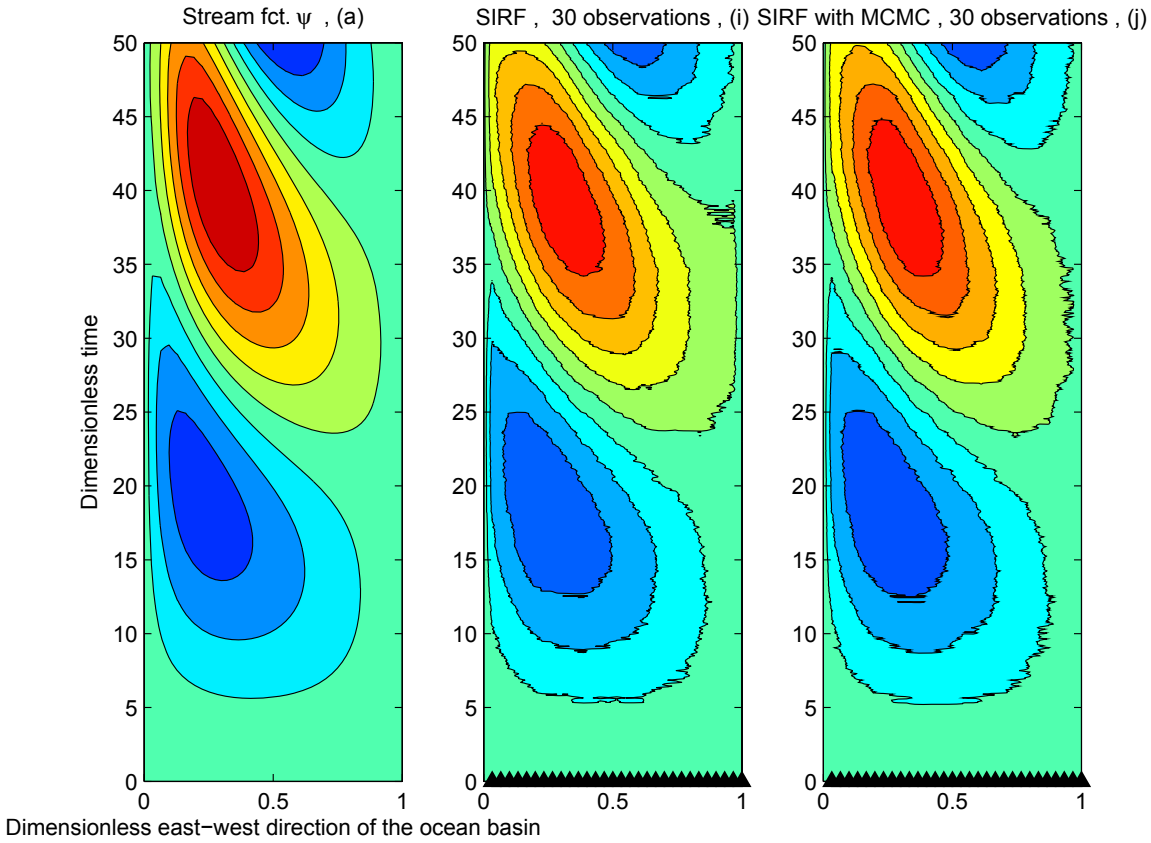

Figure 2.9: True stream function from the Stommel model and estimates from the SIR filter combined with MCMC. Identical format to Figure 2.2. This figure illustrates how the quality of the data assimilation using the SIR filter can be improved without increasing the number of particles, $N_{particles}$. Panel (i) is the same as panel (h) in Figure 2.8: $N_{particles} = 100$, no MCMC step. Panel (j) shows the recovered model state when the SIR filter is used in combination with an MCMC step (all other model parameters remain unchanged).

Chapter 3

A New Class of Idealized Models of Ocean Eddy and Drifter Position

The Stommel model of ocean circulation was introduced in the previous chapter to illustrate the main sequential assimilation methods used in this thesis. The Stommel model is linear in the sense that the change of the ocean state at a given time step is a linear function of the state and forcing from the previous time step. It is therefore not surprising that all the assimilation methods performed well when applied to this simple model.

In this chapter a new class of idealized models of circulation and drifter motion is introduced. One important feature of this class of models is that the dimension of the state vector is low (typically 4) and so many sensitivity studies can be performed. Another feature is that the models are nonlinear in that the change in drifter position (dx/dt) is a function of position (because the velocity is also a function of x). This introduces significant complexity and richness to the model. In particular the quality of the nowcasts and forecasts of the eddies and drifter position will be shown to vary amongst the assimilation methods. It will become clear that the results obtained in this chapter will provide insight into the performance of the assimilation schemes when using more complex ocean models. In particular the results will help determine the conditions (e.g. intensity of eddies, drifter release point, ensemble size, priors) under which the various data assimilation schemes break down and the model forecasts become unreliable.

The new model has essentially two components: one component describes the evolution of a field of vortices, and the other describes the changing positions of an arbitrary number of drifters advected by the vortices. Overall the goal is to accurately

nowcast and forecast the changes in the positions of the vortices from the observed positions of the drifters. It is important to note that the observations are assumed to be noisy (the true position of the drifters is not known exactly) and they are intermittent (to simulate Argo drifters that only report their positions on a 10 day schedule). In general the position of the vortices are fixed with respect to a reference point that will, for reasons to become clear later, referred to as the saddle point. In general the individual vortex centers follow the movement of the saddle point and so if the movement of saddle point can be inferred from the drifter position observations, so can the evolution of the vortices and thus the complete flow field. This is at the heart of the low dimensional representation of the instantaneous flow field: all that is required in the position of the saddle.

The basic idea behind the eddy flow component of the flow field comes from Aref (1984) who was interested in "chaotic advection" of particles by simple flow fields. The eddies in Aref's system are modeled as point-vortices which create a potential flow. Aref showed that if the position of the vortices is kept fixed, the motion of the particle is integrable and the system does not stir very efficiently. If the vortices are moved (or "blink") such that the potential flow becomes unsteady, the motion of the particles can become chaotic and this can lead to efficient stirring.

Aref constructed a simple and illuminating model of particle motion in terms of a sequence of analytically-specified displacements about a small number of simple flow structures (vortices). The alternating switching on and off of the vortices lead to the "blinking vortex" terminology. One attraction of Aref's approach is that there is no error in the integration of $dx/dt = u(x,t)$ because the displacements over each "blink" are specified analytically.

In this thesis we introduce a new class of blinking vortex models to simulate the complicated flow patterns, and drifter motion, that are observed in the real ocean and atmosphere as a result of interaction of vortices. The simplicity and computational efficiency of this class means that it provides an excellent way of testing and comparing various assimilation schemes (e.g. EnKF and PF). The success or failure of a

particular assimilation scheme will shed light on its performance in a more realistic ocean model.

The blinking vortex class can accommodate different configurations of an arbitrary number of vortices, drifters and release points. It assumes the vortices are fixed with respect to a reference point that has coordinates subsequently denoted by c^s (where c denotes center, and superscript s denotes saddle). One special case is the Self Advecting Vortex Model (SAVM). This model has two point-vortices with centers at $(p, 0)$ $(-p, 0)$ and fixed radii R. The polarity of the two vortices is the same. As each vortex blinks on, it advects the position of the other vortex (hence the terminology "self advecting"). For simplicity the vortices are viewed in a frame of reference that aligns the x-axis with the two vortex centers (see below for details).

To increase the relevance of the blinking vortex class of models, the evolution of the saddle point and drifters are allowed to evolve stochastically. The saddle point is assumed to follow a random walk; this implies that each vortex center follows the same random walk because their locations are fixed with respect to the saddle. This random walk will be shown to have a significant effect on the trajectories of the drifters. The position of each drifter, following advection by the vortices, has also been assumed to be changed by the addition of a random forcing which represents unknown and unresolved dynamical processes of the model. The physical motivation for adding stochasticity is to allow for model error in the motion of the vortices and the particles. Note that by following an ensemble of saddle and drifter positions we can propagate uncertainty in their positions and estimate their evolving probability density functions.

A summary of a single update of the position of the saddle and the drifters for the blinking vortex class of models is given below:

1. Wind each drifter about each vortex center in turn. The length of the arc is determined by the distance of each drifter from each vortex center (r), the polarity and swirl velocity profile ($U(r)$) of the vortex (e.g. Rankine, exponential),

and the model time step (Δt). The drifter is wound separately around each vortex. After each wind, the drifter reaches an intermediate position. The drifter reaches its final position after it is wound around the last vortex. The sequence in which the rotations occur is fixed but not critical provided the time step is small.

2. For the SAVM, the reference frame is rotated so that the position of the two vortex centers are always aligned with the x-axis.

3. Update the position of the saddle (and thus the position of each vortex) by adding a random displacement. The standard deviation in x and y is denoted by σ_c. The displacement is assumed Gaussian with zero mean and zero correlation between the x and y displacements. The variance of the displacement controls the drift of the saddle.

4. Update the position of the drifters by adding random displacements to each drifter. The standard deviation in x and y is denoted by σ_d. The displacement is assumed Gaussian with zero mean and zero correlation between the x and y displacements, and between drifters. The variance of the displacement controls the "diffusion" of uncertainty in drifter position.

It is shown below that the blinking vortex class of models can be written in standard state space form. The state vector (x_k) defines the position of the saddle, and each of the drifters at time step k. The saddle and drifter positions depend only on conditions at the previous time step, $x_k = \mathcal{F}(x_{k-1})$ and so the dynamical model is first-order Markov. The function $\mathcal{F}$ is nonlinear if the swirl profile is a nonlinear function of distance from the center.

In the next chapter it will be assumed that drifter positions have been observed and these data are then assimilated into the model to estimate the trajectory of the saddle (and thus the complete flow field) and the true drifter positions. The rest of this chapter provides details on the SAVM and some typical model output. In

section 3.1 the vortex models and the flow fields of the SAVM are introduced and used to explain the terms *saddle*, *elliptic point*, *hyperbolic point*, *separatrix* and their properties. In section 3.2 the drifter model is introduced. In section 3.3 the vortex and drifter model are combined to form, together with the observation equation, the state-space form of the model. The results of various sensitivity experiments with the SAVM are presented in section 3.4.

3.1 The Flow Field

The equations for the flow field of the SAVM are defined in terms of the swirl velocity profile associated with each individual vortex. Let r denote the distance between the center of a vortex and a given drifter, R denote the "radius" of the vortex, ω provides a scale for the swirl speed (positive for an anticyclonic vortex), and $\theta(r/R)$ defines the swirl profile. The swirl velocity is given by

$$U(r/R) = \omega r \theta(r/R). \tag{3.1}$$

The following table lists the vortex profiles used in this study:

Name	$\theta(r)$
Solid Body	1
Gaussian	$\exp(-r^2/R^2)$
Point Vortex	$(\epsilon + r/R)^{-1}$
Exponential	$\exp(-r/R)$

Table 3.1: List of vortex swirl velocity profiles used in this study. Note that $0 < \epsilon \ll 1$ is introduced in the point vortex profile to avoid a singularity at $r = 0$.

I will assume, without loss of generality, that the horizontal position of the vortices is fixed at $|p| = 1$. This means that the two vortices of the SAVM are centered at $(-1, 0)$ and $(1, 0)$. The vortex radius R is thus effectively measured in term of half the distance between the two vortices. In principle the parameters which define the swirl velocity U can be different for each vortex. However, for simplicity I will assume

$|\omega|$ and R to be the same for each vortex.

Figure 3.1 shows a typical displacement map for a fixed time, t_{frozen}. The black vectors in this figure show how far a particle moves with respect to its position at the earlier time step. The displacement maps are generated by evaluating the instantaneous displacement of drifters deployed on a regular field, thereby providing a view of the overall flow. It is important to note that although the displacement maps do not allow one to readily deduce the Lagrangian trajectories or their properties, they do provide useful information in the form of "instantaneous stagnation points" (ISPs). The location of an ISP, x_{isp} say, is defined implicitly by $v(x_{isp}, t_{frozen}) = 0$ where $v(x, t)$ denotes the instantaneous velocity field (e.g. Mancho et al. [2006]).

For two dimensional incompressible flow, the eigenvalues (λ_1 and λ_2) of the Jacobian $\partial v/\partial x$ evaluated at $x = x_{isp}$, $t = t_{frozen}$ determine the type of ISP. The most important cases are given in the following table.

Name of ISP	Condition on the λ_i
Hyperbolic with stable node	Both $\lvert\lambda_i\rvert < 1$
Hyperbolic with unstable node	Both $\lvert\lambda_i\rvert > 1$
Saddle	$\lvert\lambda_1\rvert > 1$ and $\lvert\lambda_2\rvert < 1$ or vice versa
Elliptic	Both $Re(\lambda_i) = 0$ and $Im(\lambda_i) \neq 0$

Table 3.2: Characterization of the instantaneous stagnation points of a frozen flow field ($v(x, t_{frozen})$) in terms of the eigenvalues of $\partial v/\partial x$.

The movement over time of a saddle-like hyperbolic point is called a hyperbolic trajectory (which too can be a stationary point). As noted by Wiggins [2005], curves that intersect with the saddle and partition the flow in different regions are termed *separatrices*. Separatrices are distinguished into *stable* and *unstable manifolds*. Particles moving on the stable manifolds are attracted towards the saddle and particles moving on the unstable manifold are repelled from the saddle. A separatrix acts as a barrier to transport, i.e. particles can not cross the separatrix and move from one region into another region (e.g. the boundary of an eddy in a flow field is a separatrix).

In complex dynamical systems the analytic form of the flow field is generally not known; only flows at spatial and temporal increments are known. (See for example Figure 3.1.) However, it is still possible to compute the hyperbolic points and separatrices of these discrete systems. Knowing their separatrices and hyperbolic points can allow the flow to be partitioned into distinct flow regions and this can lead to a low dimensional representation of the original dynamical system.

An instantaneous flow field from the SAVM is shown in Figure 3.1. The red circles mark the positions of the two, self advecting vortices. The blue squares mark the centers (elliptic points) of the so-called ghost vortices resulting from the superposition of individual vortices. The green triangles mark saddle-like hyperbolic points which too are a result of superposition. Although the flow fields of the SAVM look fairly simple, the trajectories of particles advected by these flow fields can be very complex. This is especially true if the vortex centers are subject to random displacements. This is illustrated by the associated drifter trajectory map shown in Figure 3.2. Four flow regions around the elliptic points of the real and ghost vortex centers are clearly evident, one "figure-eight" flow region around the two real vortices, and an almost circular flow in the far field (computational details below). Similar structures are found in the more complex, self advecting model of Kuznetsov et al. [2003] (see Figure 3.3) which is defined in continuous time and requires numerical integration to estimate drifter trajectories (unlike the simpler SAVM).

The partitioning of the flow into distinct regions raises the possibility of applying different data assimilations methods in different regions. They could correspond, for example, to regions that respond in a strongly nonlinear fashion to small perturbations (e.g. along unstable manifolds, or in the vicinity of saddles). If this can be shown to be true for idealized models (e.g. the SAVM) then this approach (different assimilation schemes in different regions) may possibly be applied to more realistic ocean models.

The positions of the two vortices of the SAVM are defined with respect to the

saddle c_k^S which may undergo a random walk:

$$\begin{aligned} c_k^S &= c_{k-1}^S + u_{k-1} \\ u_k &\sim \mathcal{N}(0, \sigma_c^2 I) \end{aligned} \tag{3.2}$$

where $c_k^S \in \mathbb{R}^2$ is the location of the saddle at time step k, $u_k \in \mathbb{R}^2$ is the additive random displacement, and I is the identity matrix. The movement of the vortex centers is controlled by the movement of the saddle because all vortices are referenced with respect to this point. This means that separate equations are not needed to describe the evolution of the vortex centers. More specifically let $L^v = (L_x^v, L_y^v)^T$ denote the *fixed* difference in position of the saddle and the vth vortex. The position of the center of this vortex, c_k^v, is then simply given by

$$c_k^v = c_k^S + L^v. \tag{3.3}$$

3.2 The Drifter Trajectories

The rotation of a single drifter about the first vortex is denoted by

$$x_k' = c_{k-1}^{(1)} + T(x_{k-1} - c_{k-1}^{(1)}) \tag{3.4}$$

where T is the rotation matrix denoted by

$$T = \begin{bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta) \end{bmatrix}. \tag{3.5}$$

The superscript of $c^{(.)}$ denotes the vortex around which the drifter rotates. Intermediate drifter positions are denoted by a dash, i.e. $x^{'}$. The angle $\Delta\theta$ denotes the drifter rotation around the vortex center over one time step and is given by

$$\Delta\theta = \frac{U}{r}\Delta t. \tag{3.6}$$

Combining (3.1) and (3.6) gives

$$\Delta\theta = \omega\Delta t\theta(r/R) \tag{3.7}$$

If θ is constant we obtain a solid body rotation and the drifter sub model becomes *linear* because θ no longer depends on drifter position.

Winding about two or more vortices is straightforward. For example, to wind about the second vortex we have

$$x_k'' = c_{k-1}^{(2)} + T(x_k' - c_{k-1}^{(2)}).$$

One subtlety of the SAVM is that the reference frame is adjusted, by a solid body rotation, to align the x-axis with the line connecting the two vortex centers after the second wind has been completed.

After the drifter has been wound around the second vortex, a random displacement is added to the drifter position

$$x_k = x_k'' + w_{k-1}. \tag{3.8}$$

Note that this equation can be written in the form

$$\begin{aligned} x_k &= \mathcal{F}_d(x_{k-1}, c_{k-1}^S) + w_{k-1} \\ w_k &\sim \mathcal{N}(0, \sigma_d^2 I) \end{aligned} \tag{3.9}$$

where $\mathcal{F}_d$ is, in general, a nonlinear function. $x_k \in \mathbb{R}^2$ is the location of the drifter at time step k, $w_k \in \mathbb{R}^2$ is the drifter noise (with standard deviation σ_d).

Figure 3.2 shows the trajectories over 150 time steps of a large number of particles in the SAVM for the case of no model error, i.e. $\sigma_d = 0$ and $\sigma_c = 0$. The black dots denote the release points of the drifters. Note that the SAVM produces a trajectory map that is qualitatively very similar to the published stream function

map of Kuznetsov et al. [2003], although the SAVM is much simpler than the two point-vortex model used in Kuznetsov's study.

3.3 The Augmented State-Space Model

Following Ide et al. [2002] and Kuznetsov et al. [2003] we augment the state equation for the evolution of the center (3.2) with the equations for the evolution of each drifter position (3.9). This is done by stacking the equations for the center and drifter evolution. The state equation for the augmented system now takes the form

$$\begin{pmatrix} c_k^S \\ x_k \end{pmatrix} = \mathcal{F}(c_{k-1}^S, x_{k-1}) + \begin{pmatrix} u_{k-1} \\ w_{k-1} \end{pmatrix}. \tag{3.10}$$

The corresponding observation equation for the drifters is of the form

$$y_k = H \begin{pmatrix} c_k^S \\ x_k \end{pmatrix} + v_k \tag{3.11}$$

where y is the observation vector, v is the observation error and

$$H = [0_{n\times 2} \quad I_{n\times n}]$$

where n is twice the number of drifters and 0 and I denote zero and identity matrices of the indicated dimension.

The important point to note is that equations (3.10) and (3.11) are in state-space form which means that any of the data assimilation methods (e.g. EnKF or PF) can now be readily tested.

As an aside, note that for a single drifter, and a single vortex, we can write

$$\mathcal{F}(c^S_{k-1}, x_{k-1}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & \cos(\theta) & -\sin(\theta) \\ 0 & 1 & \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} c^S_{k-1} \\ x_{k-1} - c^S_{k-1} \end{pmatrix}$$

and

$$H = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

3.4 Propagation of Uncertainty

To illustrate the complexity of drifter dispersion in the SAVM, and also assess predictability, we now describe a set of simulations with the SAVM using point-vortices with radius $R = 1$ and swirl speed scale of $\omega = \pi/30$. In each experiment 1000 drifters were released around a specified position and their positions computed for subsequent time steps. This allowed the density of drifter position to be estimated and plotted as a function of space and time. The two main sets of experiments (I and II) are described in Table 3.3.

The first set of experiments explores sensitivity of the pdf of drifter position to its initial condition. The second set explores sensitivity to model error (in both the evolution of the saddle of the SAVM and also drifter position. Within each set of experiments the influence of the release locations of the drifters was explored. Drifters were released at the origin, near elliptic and hyperbolic points, and elsewhere.

3.4.1 Experiment I: Uncertain Drifter Release

Figures 3.4 to 3.6 show the spreading of the uncertainty in drifter location for different values of σ_{prior_d} and different release locations (x_0, y_0). In all six experiments the

Figure	σ_d	σ_c	σ_{prior_d}	σ_{prior_c}	$x_0,\ y_0$	Time steps
EXPERIMENT I: UNCERTAIN DRIFTER RELEASE						
3.4 left	0	0	0.01	0	0, 0	51
3.4 right	0	0	0.1	0	0, 0	51
3.5 left	0	0	0.1	0	-1, 0	51
3.5 right	0	0	0.1	0	0, 1.6	51
3.6 left	0	0	0.1	0	-2, 0	51
3.6 right	0	0	0.1	0	-1, -1.6	51
EXPERIMENT II: UNCERTAIN MODEL DYNAMICS						
3.7 left	0.05	0	0	0	0, 0	51
3.7 right	0.05	0	0	0	-1, 0	51
3.8 left	0.05	0	0	0	-2, -2	51
3.8 right	0.05	0	0	0	-1, -1	51
3.9 left	0	0.05	0	0	0, 0	51
3.9 right	0	0.05	0	0	-1, 0	51
3.10 left	0	0.05	0	0	-2, -2	51
3.10 right	0	0.05	0	0	-1, -1	51
3.11 left	0.05	0.05	0	0	0, 0	51
3.11 right	0.05	0	0	0	0, 0	301

Table 3.3: Model parameters used in the simulations. EXPERIMENT I: The initial drifter positions are not known exactly, rather they have Gaussian pdfs centered on (x_0, y_0) with standard deviation σ_{prior_d}. The saddle of the SAVM is the origin (i.e., the standard deviation, $\sigma_{prior\_c} = 0$), and the model is error free ($\sigma_c = \sigma_d = 0$). EXPERIMENT II: The initial drifter positions coincide exactly with (x_0, y_0) ($\sigma_{prior\_d} = 0$). The origin of the SAVM is the origin (i.e., the standard deviation, $\sigma_{prior\_c} = 0$). The model has non zeros errors for the drifters and/or center (σ_d and σ_c are nonzero).

model is assumed error free, i.e. no model error is added to the drifters or center. The underlying vector map shows the instantaneous displacement field. This field is constant since the vortex parameters were chosen to be time independent.

All figures show that the drifter distribution is Gaussian initially (ellipses) as expected. Whether the drifter distribution stays Gaussian or becomes non-Gaussian depends strongly on the release location of the drifters. For example, the distribution of drifters released around an elliptic point (e.g. centers of real or ghost vortices) tends to stay Gaussian (Figure 3.5). Furthermore the distribution tends to be stationary

in space and in time. Drifters trapped near elliptic points will generally move around this point for a long time. By way of contrast, the distribution becomes strongly non-Gaussian (e.g. "banana" shaped) if the drifters are released close to hyperbolic points, e.g. the origin of the system (Figures 3.4 and 3.6). This is especially true if the drifters pass through or near hyperbolic points or pass regions with strong velocity gradients. The drifter distributions released around locations away from hyperbolic and elliptic points tend to maintain their Gaussian form (right panel of Figure 3.6).

3.4.2 Experiment II: Uncertain Model Dynamics

To investigate the effect of model error, the previous experiments were repeated but with σ_c and σ_d non-zero. However in the new experiments all the drifters are released at the same known locations, i.e. the prior standard deviation of the initial drifter distribution is zero.

Figures 3.7 to 3.8 show that the distribution of the drifters spreads as time increases. The drifter distribution becomes more non-Gaussian, even if the drifters are released near an elliptic point. The nearer the drifters are released to hyperbolic points, or regions with a strong velocity gradients, the faster the distribution becomes non-Gaussian.

The introduction of model error for the saddle c^S transforms the initial Gaussian distribution of the drifters into a non-Gaussian distribution with time. Figures 3.9 to 3.10 show that the speed of the transition depends mainly on the release location of the drifters. The spread of the distribution of the centers increases too over time but note that it maintains its Gaussian (elliptic/circular) shape. This is to be expected since the centers are only perturbed with Gaussian noise and not subject to any other linear or nonlinear transformations. Since drifters depend in a nonlinear way on the position of the centers, it is to be expected that the drifter distribution evolves in a much more complex way.

Figure 3.11 (left panel) shows how the combined influence of model error for both drifter and center leads to an increased non-Gaussian spread in the drifter density

(compare to Figure 3.7). Figure 3.11 (right panel) shows the same experiment as Figure 3.7 but for a longer model integration of 300 time steps. This experiment confirm the result that Aref noted in his paper (Aref [1984]), namely that small model error in drifter position will increase the total mixing of the system ensuring that the drifters will eventually become evenly distributed over the model domain. The speed of the mixing depends mainly on the magnitude of the drifter model errors. It is noticeable that the distribution is not completely symmetric. This is due to the asymmetry in the flow field that results from the order in which the vortices in the SAVM blink.

To summarize the sensitivity of drifter dispersion to release location, a reference drifter and an ensemble of 1000 additional drifters were released at a fixed point from a grid that covered the model domain. The differences is position between the reference and each ensemble drifter were calculated, accumulated and stored. This was repeated 50 times and the mean and median of the 50 accumulated errors were calculated, resulting in one total error associated for the chosen release location. This procedure was repeated for each grid point of a grid covering the model domain.

Figure 3.12 shows the maps of dispersion assuming an integration time of 51 steps, $\sigma_{prior_c} = 0.05$, $\sigma_{prior_d} = 0.05$, $\sigma_c = 0.05$ and $\sigma_d = 0.05$. The left panel shows the mean of the squared total error and the right panel shows the median of the square of total error. These "predictability maps" indicate release locations directly with drifter dispersion. The color bars show the total squared error. High values are associated with high dispersion and thus also associated with high uncertainty of the drifter location and low predictability. Both panels show high drifter dispersion if drifters are released in the incoming manifold of the saddle point at the origin. The white contour lines show sample drifter trajectories for a complete error free model. It should be noted that the method of deriving the total error does play a role in the outer regions of the domain. Using the median to compute the total error seems to decrease the magnitude of the total error. This fact should be kept in mind when ensemble based data assimilation methods are used.

Figure 3.13 emphasizes the validity of the predictability map. The figure shows the trajectories of an ensemble of 100 drifters released on the incoming manifold at $(-0.25, 0.25)$ (left panel) and the outgoing manifold at $(-0.25, -025)$ (right panel). The left panel indicates high dispersion of the drifters: approximately 50% of the drifters follow the left part of the outgoing manifold and 50% the right part. The right panel shows that drifter dispersion is low until the drifters reach the vicinity of the incoming manifold of the left saddle at $(-2, 0)$.

3.5 Discussion

A new class of idealized models of for the evolution of ocean eddies and drifter evolution is introduced. The model is very economical both in terms of the dimension of the state vector and also the computational cost of integrating it forward in time. Following the pioneering study of Aref, the model is based on a sequence of analytically defined displacements of the drifters and so there is no numerical error beyond roundoff associated with the calculation of the drifter trajectories. A particular form of model, the so-called Self Advecting Vortex Model (SAVM), is shown to give very similar results to the more complex, continuous-time model of Kuznetsov et al. (2003) This latter model is more expensive to run and subject to the accumulation of errors arising from the numerical integration scheme. The simplicity and accuracy of the SAVM makes it ideal for carrying out a large number of sensitivity studies involving parameters such as model noise and drifter release points.

It has been shown that the drifter dispersion can be non-Gaussian even for a simple model such as the SAVM for large areas in the model domain. It follows that any method that assimilates infrequent drifter data must be able to account for the non-Gaussian probability density functions of drifter position. The results also show that drifters are trapped near elliptic points and their density functions tend to conserve their Gaussian shape for long periods.

By partitioning the flow domain in regions using separatrices, and identifying the elliptic and hyperbolic points, a hybrid assimilation method may be developed.

In regions where the density functions generally conserve their Gaussian form, the EnKF may be applied for data assimilation; for the smaller regions where the density functions become strongly non-Gaussian, the computationally demanding PF may be more appropriate.

In the next chapter both the EnKF and PF will be applied separately to the SAVM. The setup of the experiments will be similar to the ones performed for this section. The goal of the experiments will be to find the strength and weakness of each method and, if possible, find a blend of the two methods that takes advantage of relative efficiencies and accuracies in an effective manner.

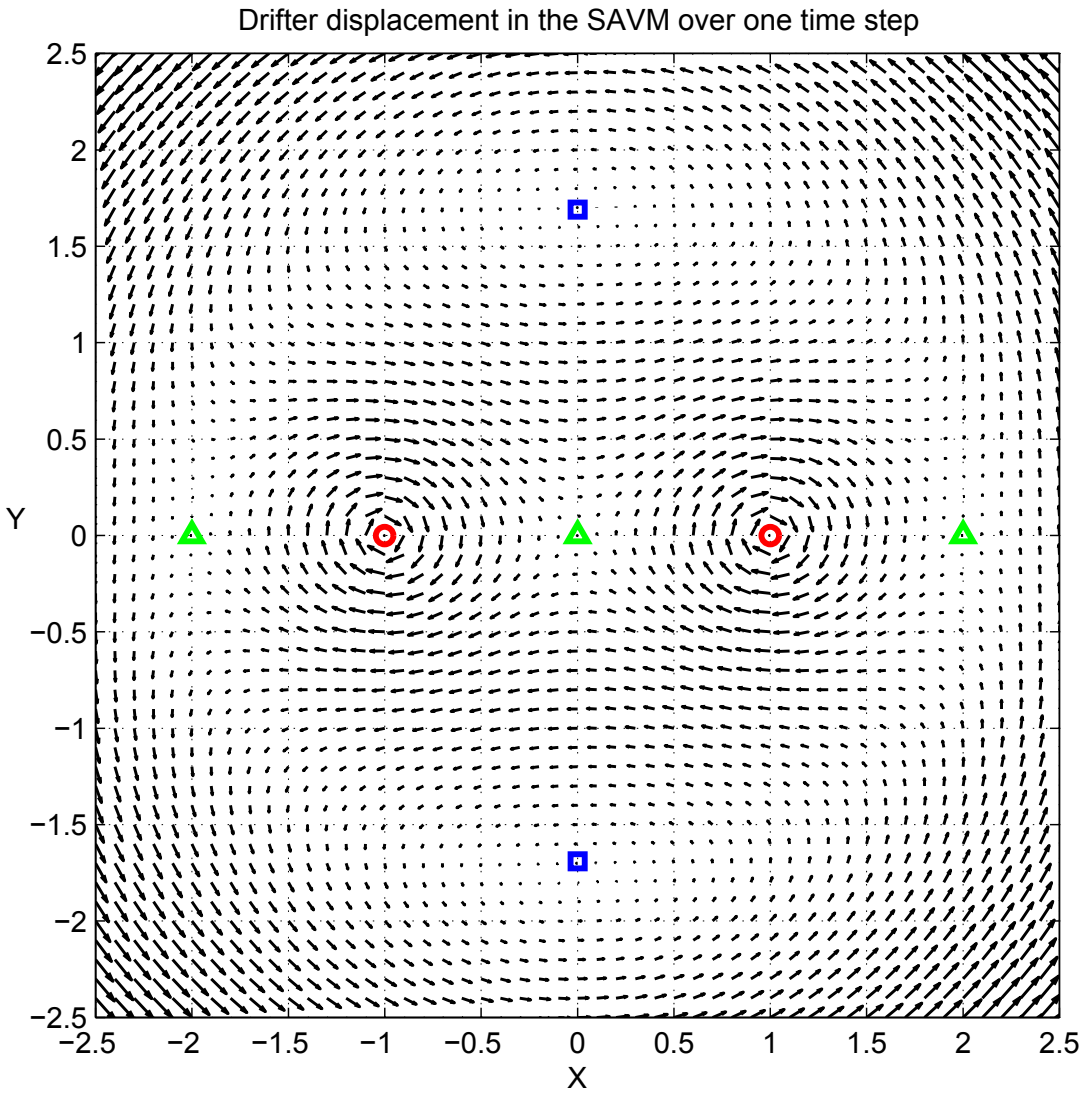

Figure 3.1: Displacement map for the SAVM over one time step. The red circles mark the centers of the physical vortices. They are point-vortices with radius $R = 1$, strength $\omega = \pi/30$, and $\epsilon = 10^{-4}$. The green triangles mark the saddle (hyperbolic) points, and the blue squares mark the centers of ghost vortices.

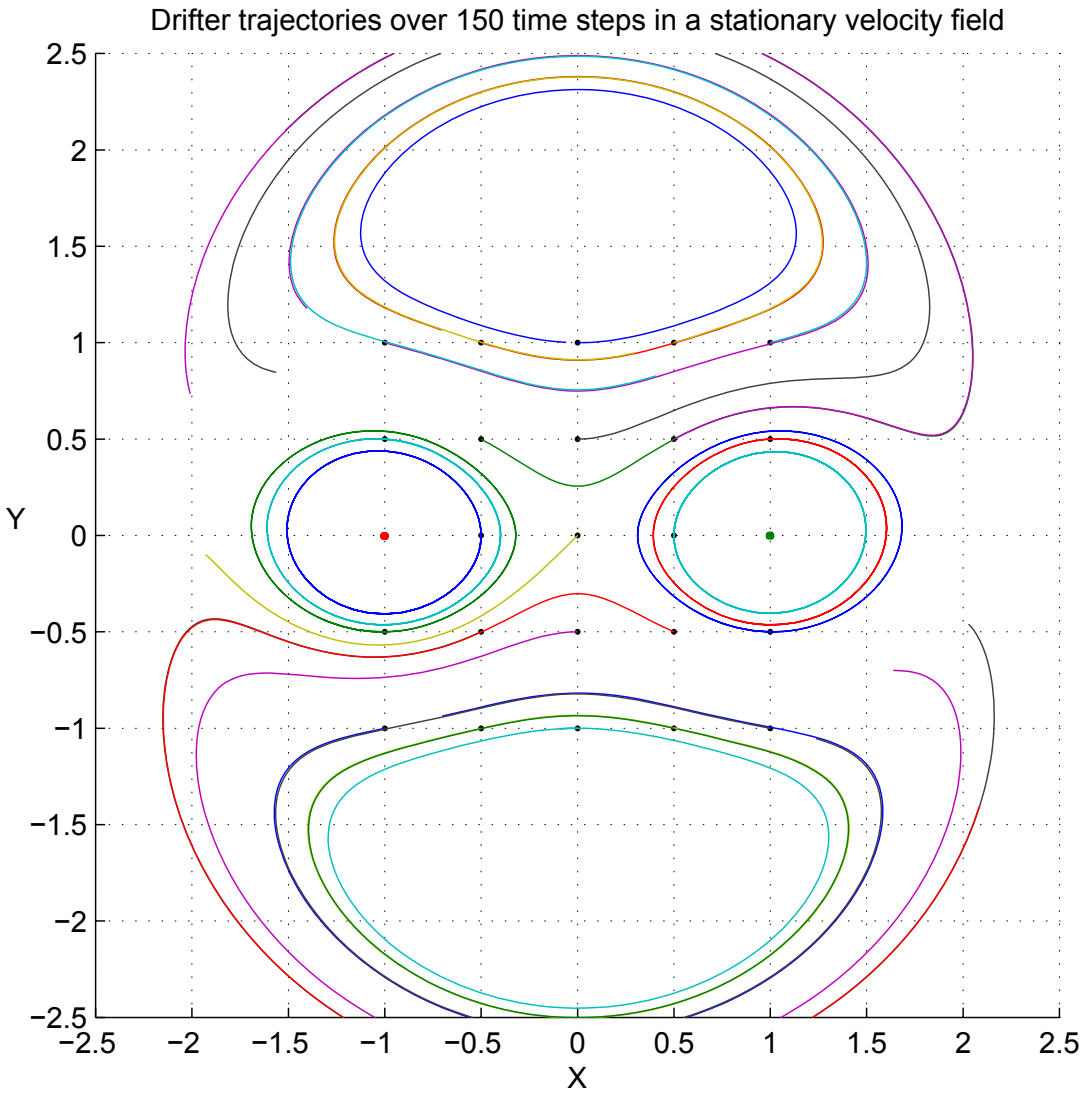

Figure 3.2: Trajectory map for the SAVM over 150 time steps. The black dots mark the release points of the drifters. Point-vortices were used with $R = 1$, $\omega = \pi/30$ and $\epsilon = 10^{-4}$. Note the trajectories match qualitatively the stream function map shown in Figure 3.3.

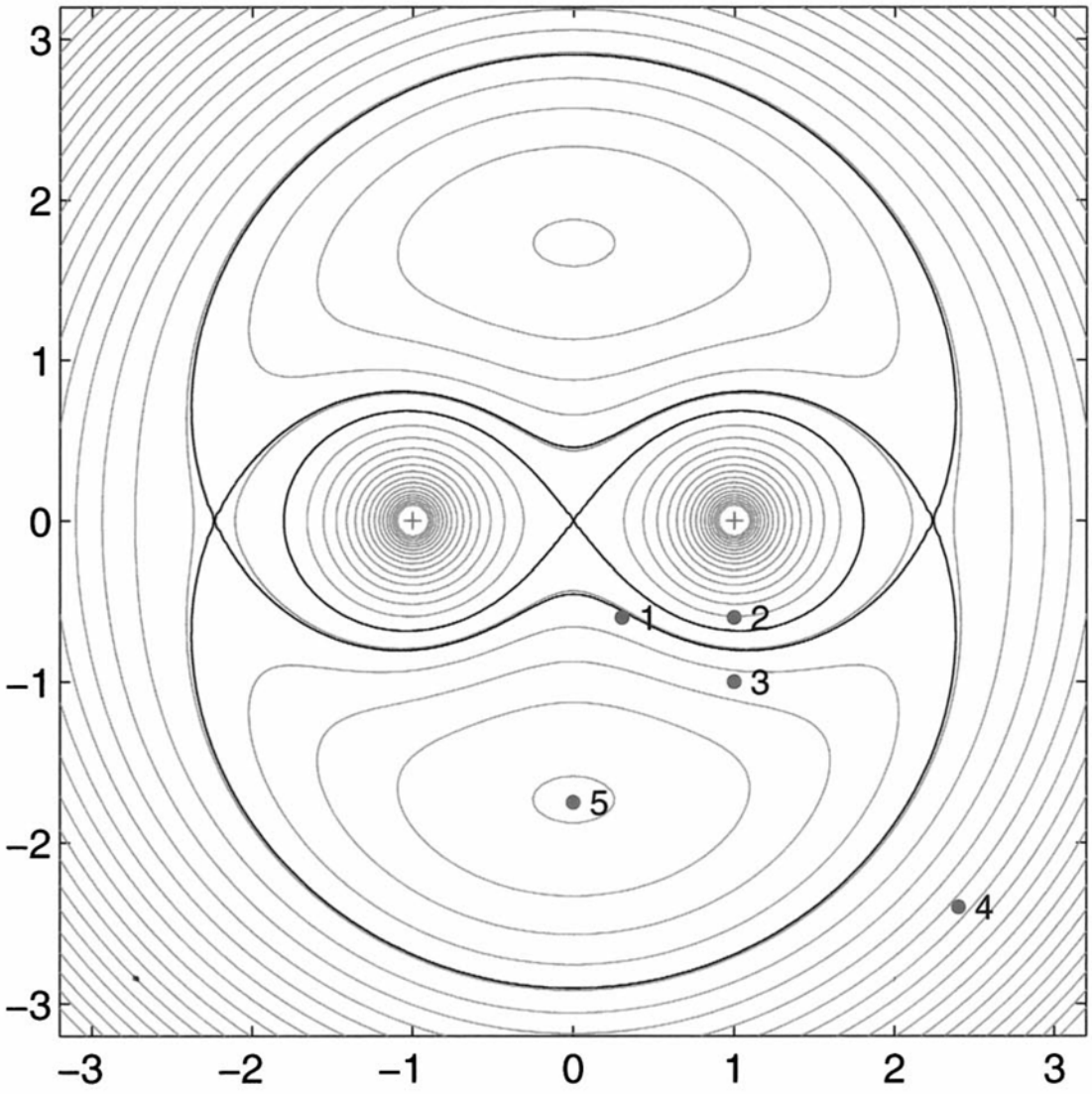

Figure 3.3: Stream function for a two point-vortex system and the associated 5 separatrices (thick black lines). From Kuznetsov et al. [2003].

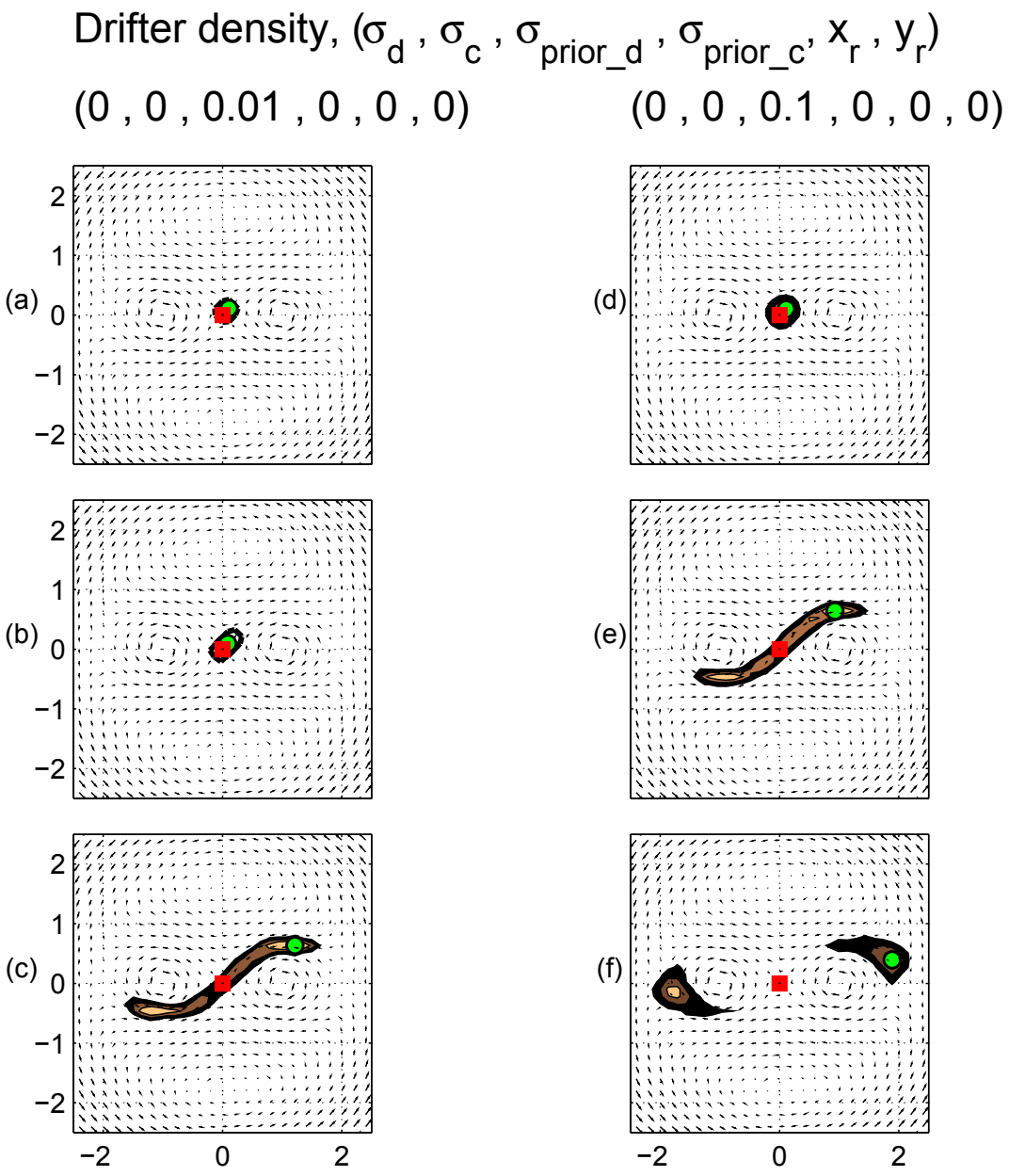

Figure 3.4: Probability density of drifter position as a function of space and time steps $t = 1$, (a,d), $t = 26$, (b,e) and $t = 51$, (c,f). The standard deviation of the drifter prioris $\sigma_{prior\_d} = 0.01$ (left panels) and $\sigma_{prior\_d} = 0.1$ (right panels). The mean release location of the drifter is the origin (red square). The green dot marks the peak of the drifter distribution. The shaded regions encompass 30% (light brown), 60% (dark brown) and 90% (black) of the drifter ensemble.

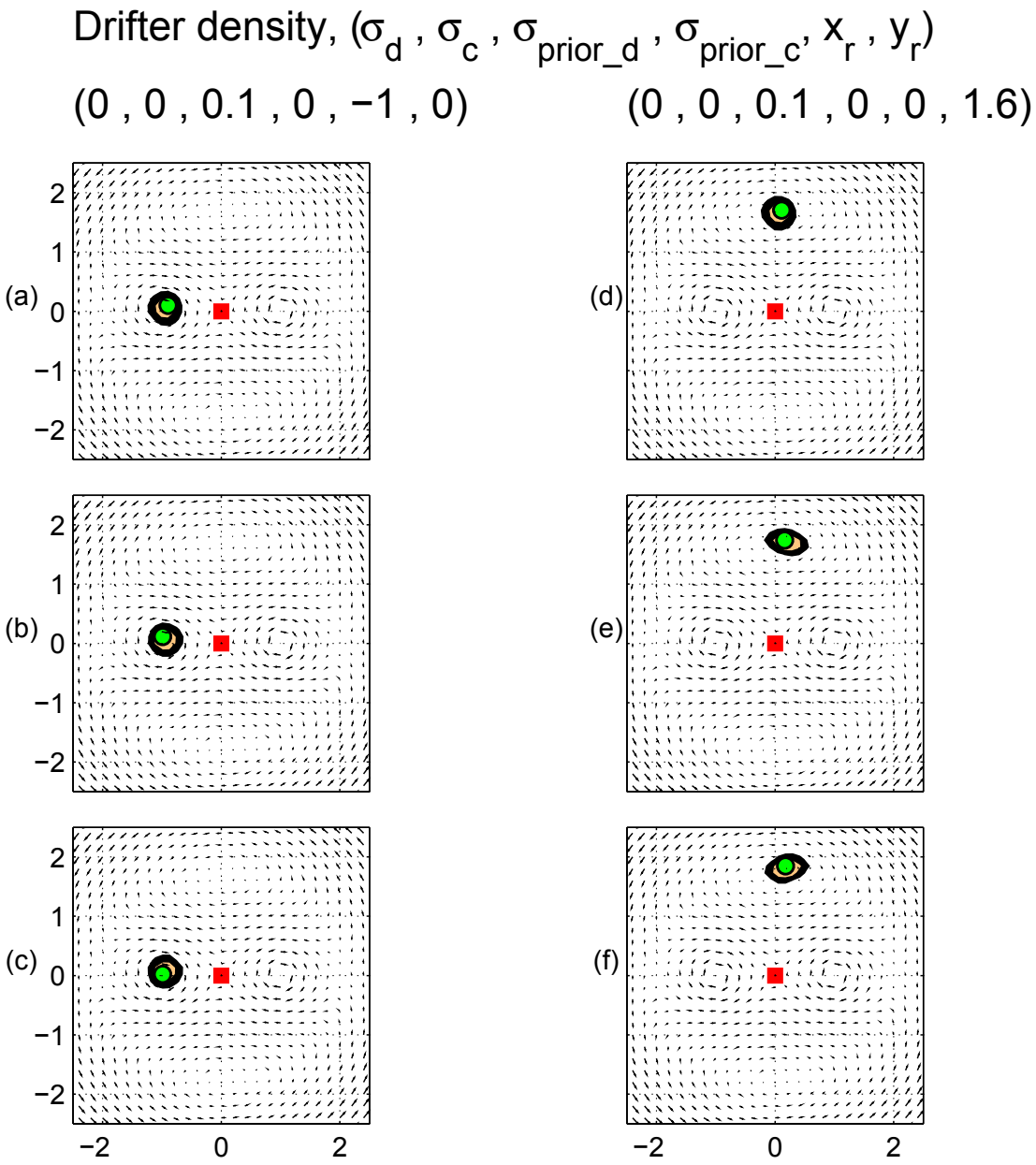

Figure 3.5: Same format as Figure 3.4 but for different drifter release locations. Drifters shown in the left panel side were released at $(x_0, y_0) = (-1, 0)$, the center of the left vortex, and for the right panel at $(x_0, y_0) = (0, 1.6)$, the center of the top ghost vortex. The standard deviation of the prior for the drifters is $\sigma_{prior\_d} = 0.1$.

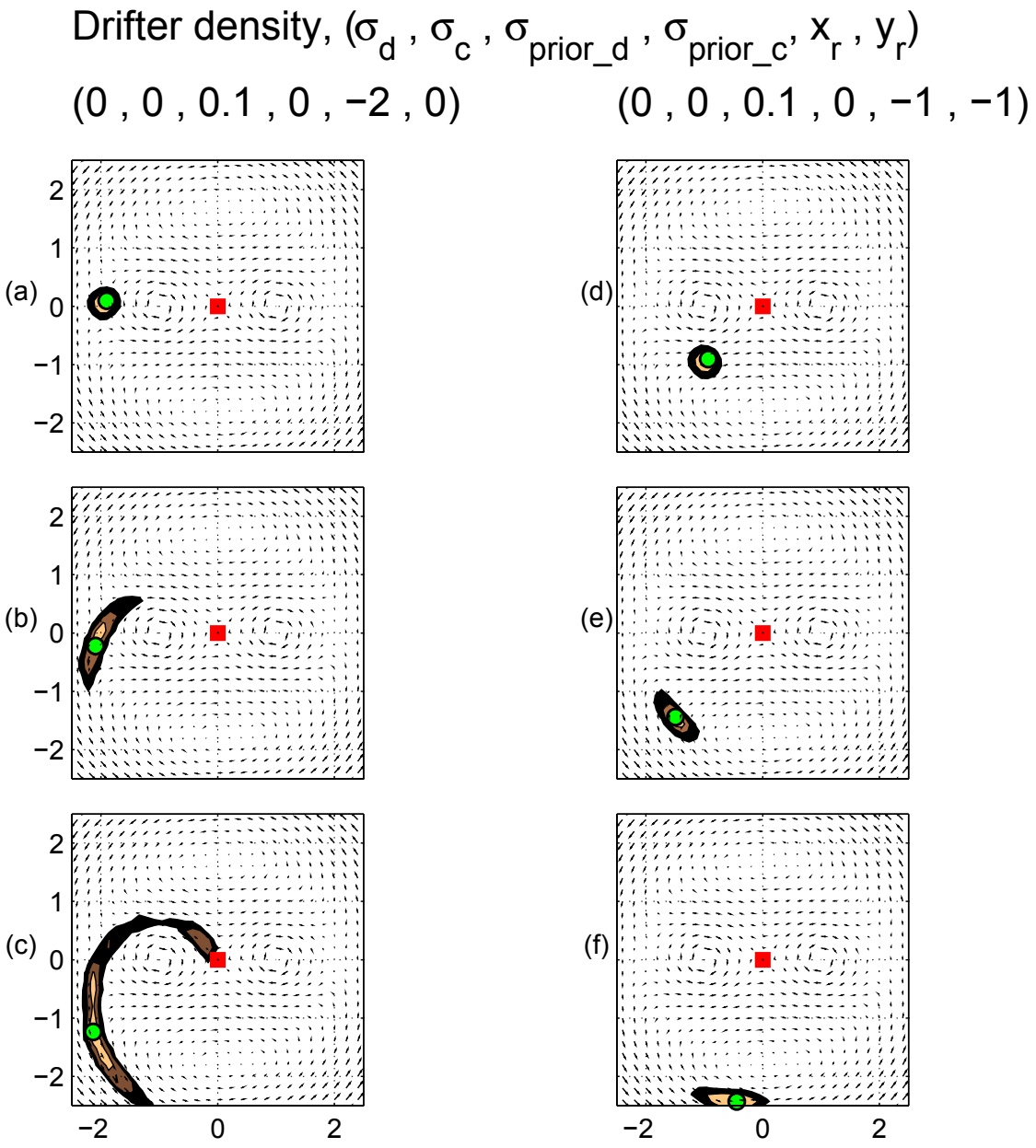

Figure 3.6: Same format as Figure 3.4 but different drifter release locations. Drifters shown in the left panel side were released at $(x_0, y_0) = (-2, 0)$, the left saddle, and for the right panel at $(x_0, y_0) = (-1, -1)$, the center of the top ghost vortex. The standard deviation of the prior for the drifters is $\sigma_{prior\_d} = 0.1$.

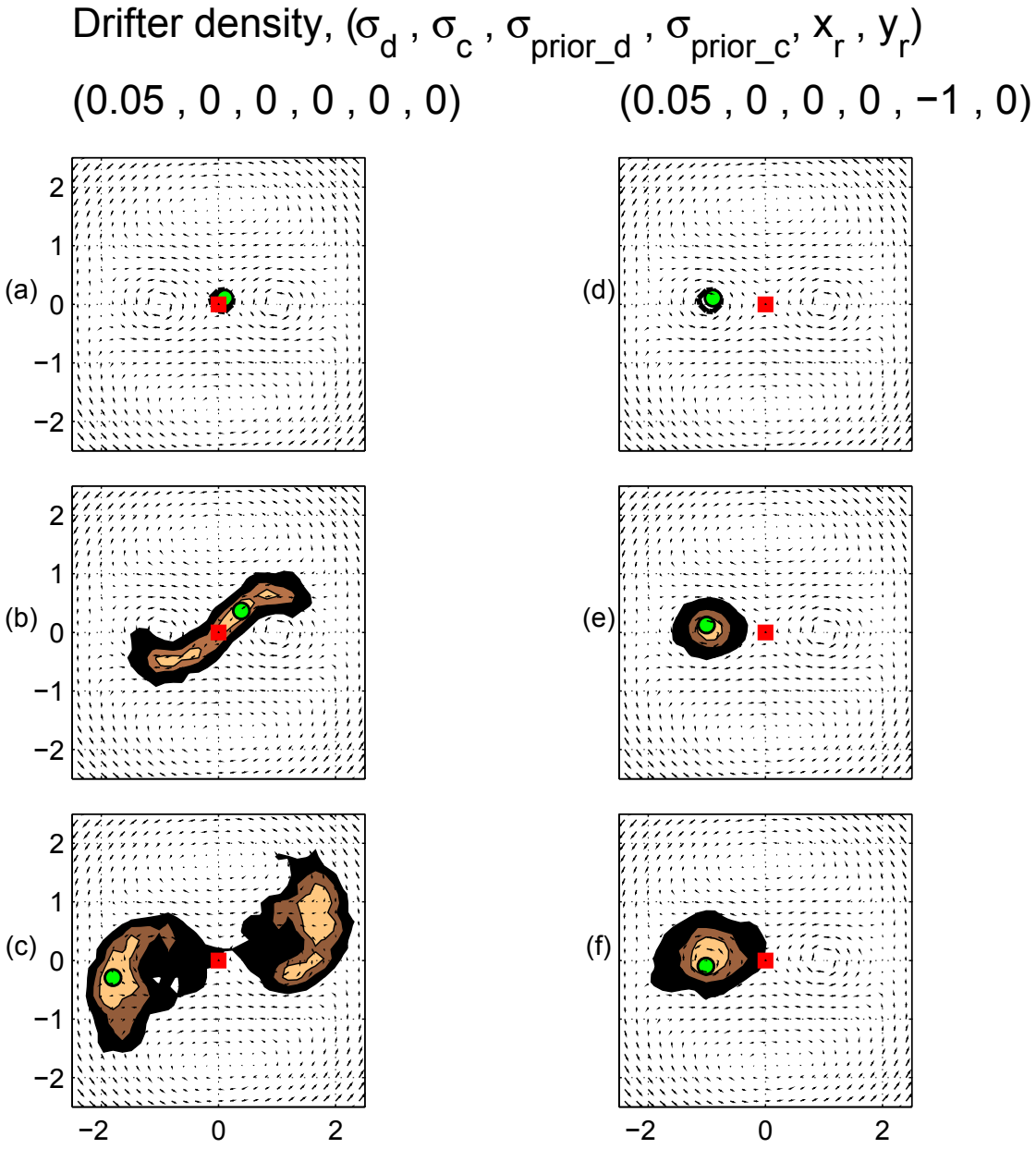

Figure 3.7: Same format as Figure 3.4 but nonzero drifter model error ($\sigma_d = 0.05$). Drifters were released exactly at the origin (left panels) and $(x_0, y_0) = (-1, 0)$ (right panels).

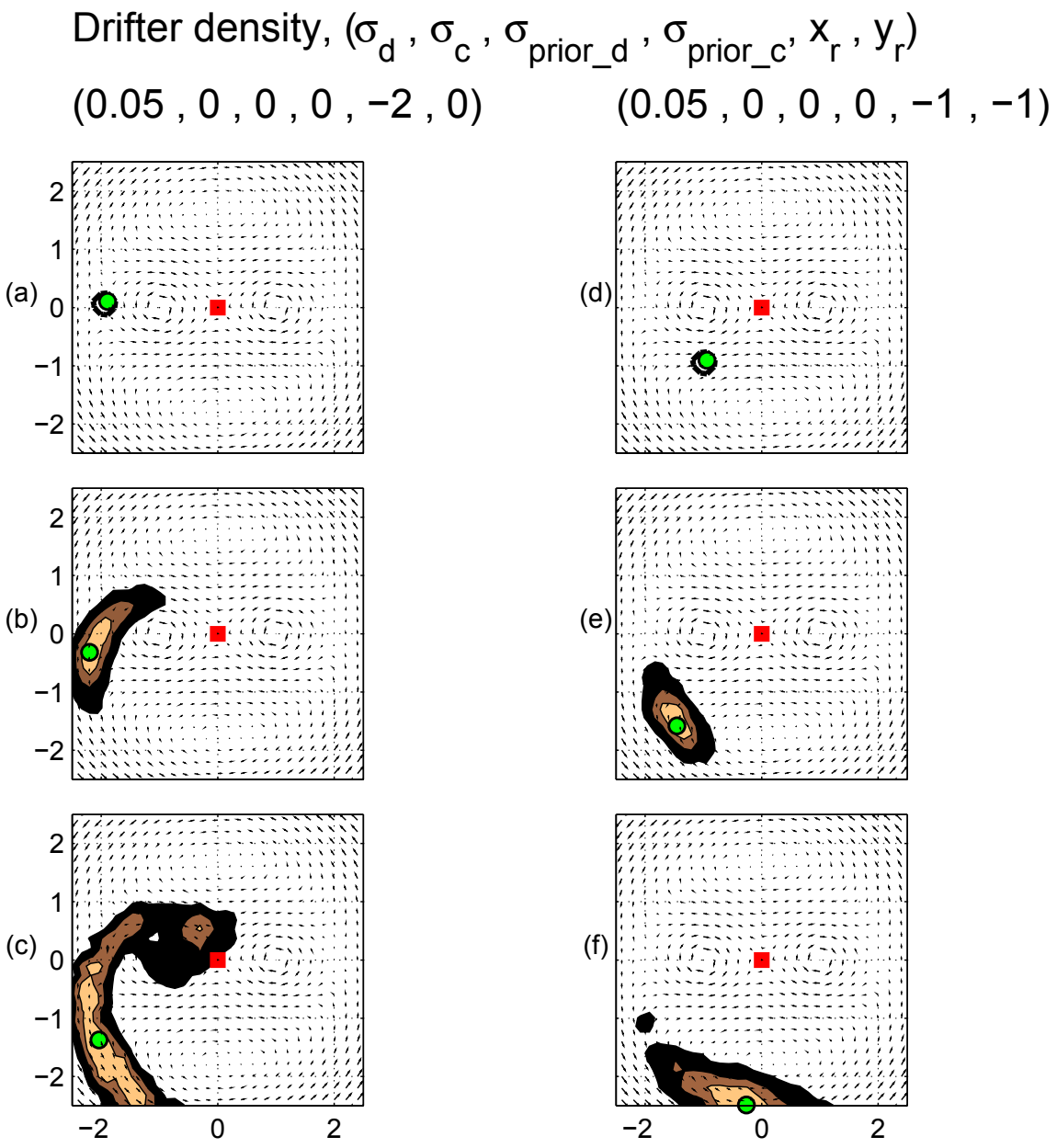

Figure 3.8: Same format as Figure 3.4 but nonzero drifter model error ($\sigma_d = 0.05$). Drifters on the left panel side were released at $(x_0, y_0) = (-2, 0)$ (left saddle point) and $(x, y) = (-1, -1)$ (right panel). The standard deviation for the model error of the drifters is $\sigma_d = 0.05$ for both cases.

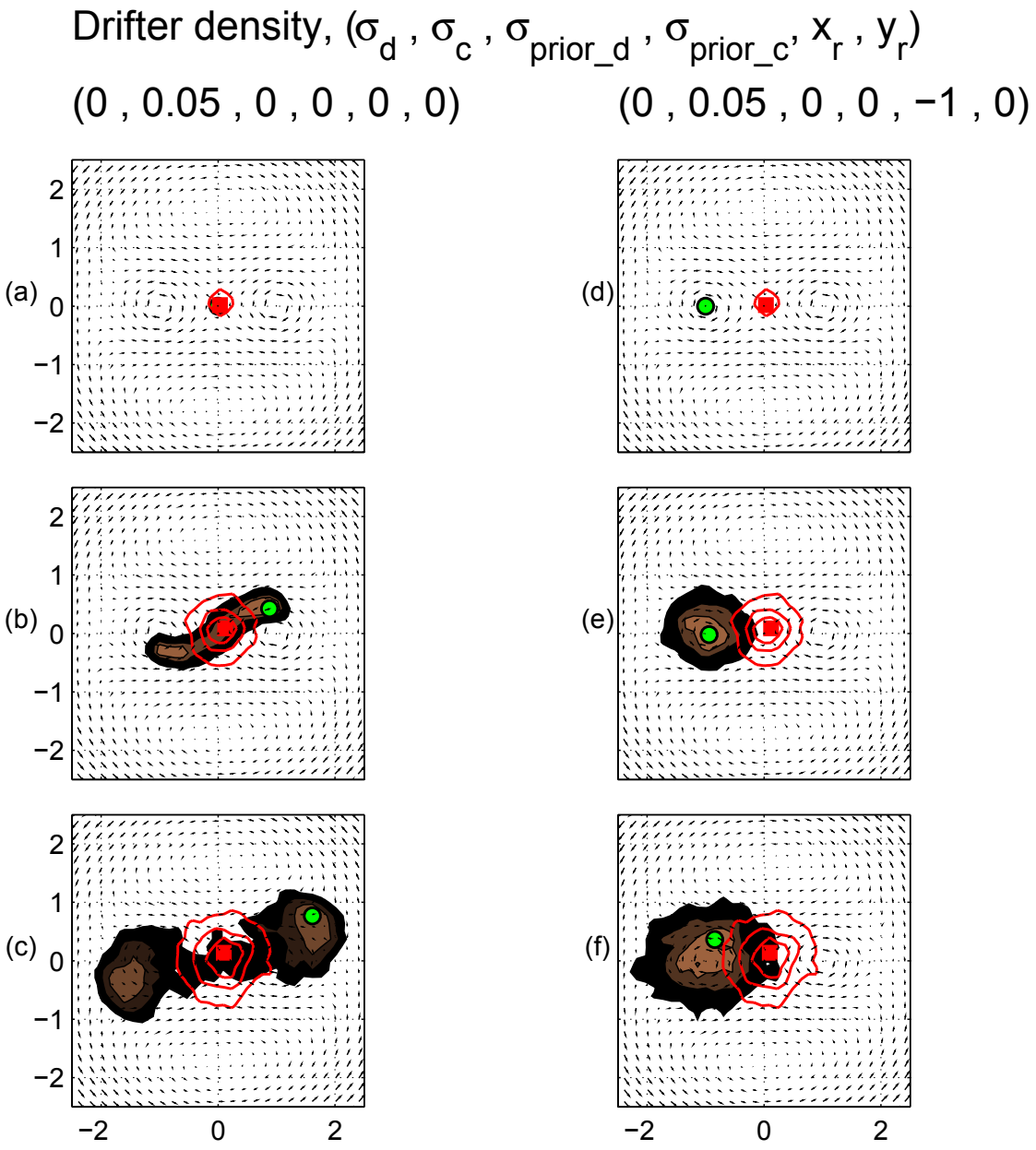

Figure 3.9: Same format as Figure 3.4 but nonzero center model error ($\sigma_c = 0.05$). The release location of the drifters is the origin (left panel) and $(x_0, y_0) = (-1, 0)$ (right panel). The red contours show the 30, 60 and 90 percent probability for the pdf of center location.

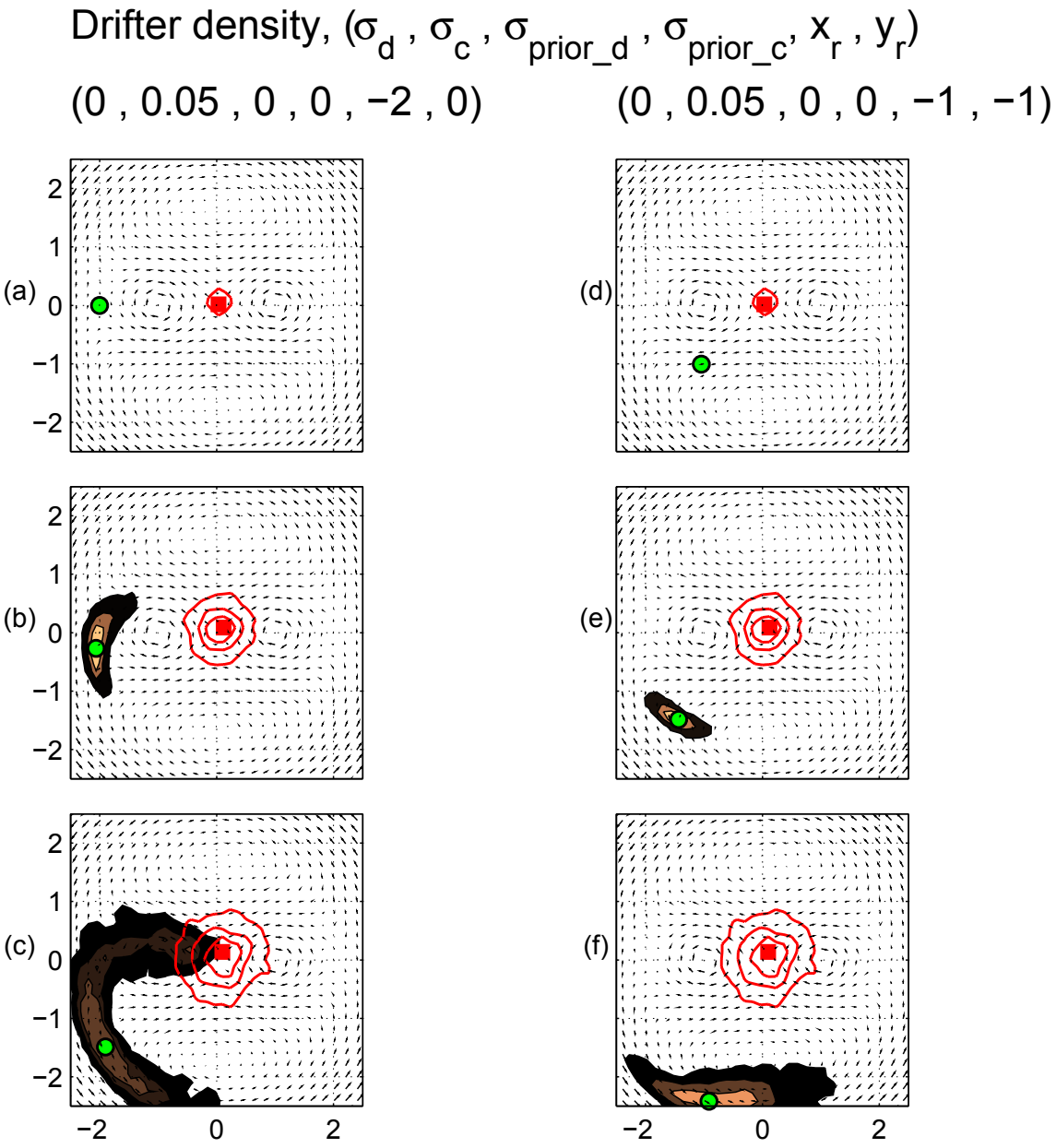

Figure 3.10: Same format as Figure 3.9 but different drifter release locations. Drifters for the left panels were released at $(x, y) = (-2, 0)$ (left saddle point) and for the right panels $(x_0, y_0) = (-1, -1)$.

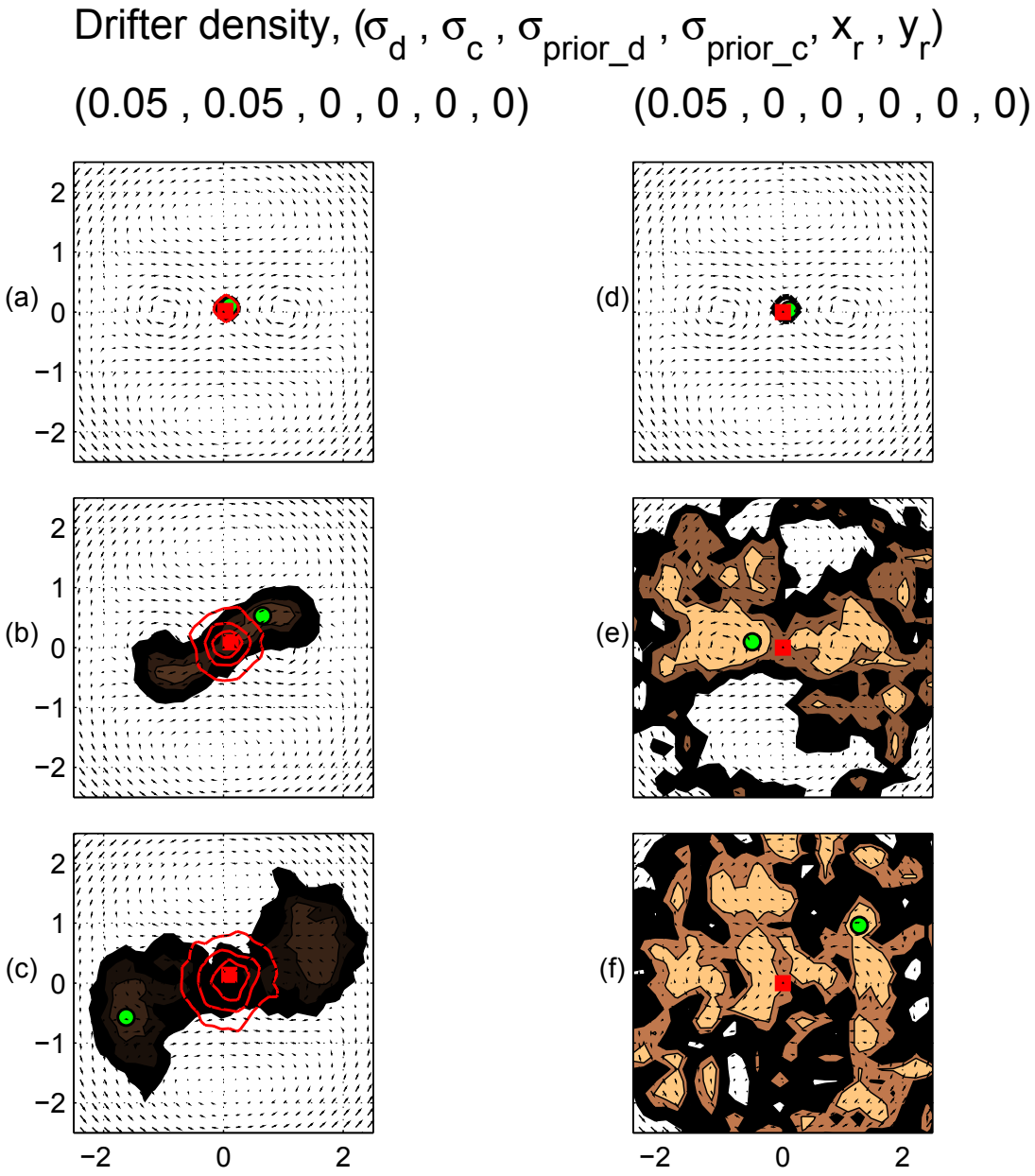

Figure 3.11: Left side: Same format as left panels of Figure 3.7 but with uncertain model dynamics with $\sigma_d = 0.05$ and $\sigma_c = 0.05$ respectively. Right side: Same format as left panels of Figure 3.7 but at $t = 1$, (d), $t = 151$, (e) and $t = 301$, (f)). The standard deviation for the model error of the drifters and center is $\sigma_d = 0.05$ and $\sigma_c = 0.0$. Compared to Figure 3.7 the combined effect of noise in the drifters and centers leads to an increased non-Gaussian spread of the drifter distribution (left panel). The experiment (right panel) confirms the result from Aref [1984] that a small model error for the drifters will lead to complete mixing.

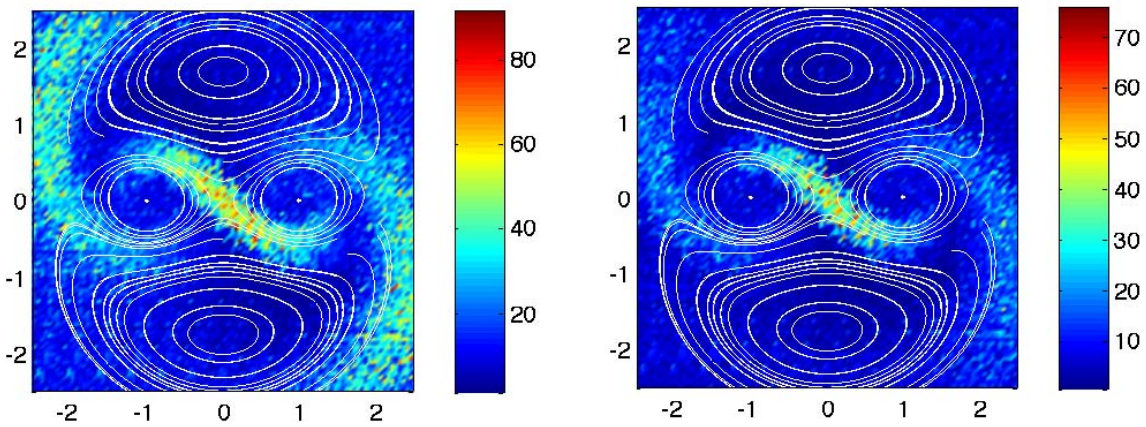

Figure 3.12: Sensitivity of drifter dispersion to release location in the form of predictability maps. Left panel: Mean of the square of the total error (i.e., difference between position of reference drifter and the ensemble). Right panel: Median of the square of the total error. The white contour lines show selected drifter trajectories from a completely error free model. $\sigma_{prior_c} = 0.05$, $\sigma_{prior_d} = 0.05$, $\sigma_c = 0.05$ and $\sigma_d = 0.05$.

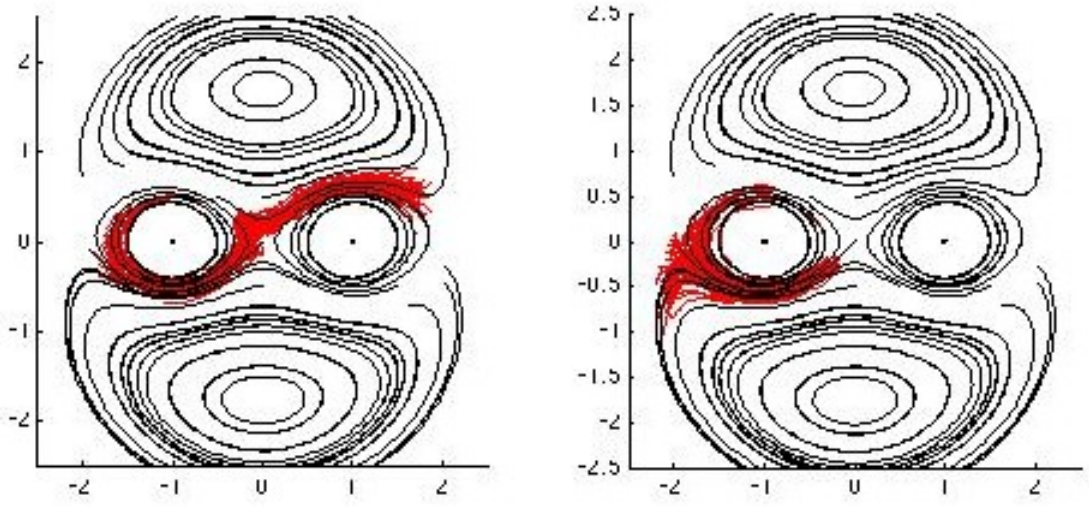

Figure 3.13: Trajectories of 100 drifters for selected released locations. The left hand panels are for an incoming manifold, $(x_0, y_0) = (-0.25, 0.25)$ and the right hand panels are for an outgoing manifold, $(x_0, y_0) = (-0.25, -0.25)$. The drifters shown in the left panels disperse immediately after release. The drifters shown in the right panels only disperse when they are in the vicinity of the left saddle point $(-2, 0)$ and its incoming manifold. $\sigma_{prior_c} = 0.05$, $\sigma_{prior_d} = 0.05$, $\sigma_c = 0.05$ and $\sigma_d = 0.05$.

Chapter 4

Assimilating Trajectories Into the Idealized Model

4.1 Introduction

It was shown in the previous chapter that the probability density function (pdf) of drifter position depends strongly on the release point of the drifters, the initial spread of the pdf around this release point (i.e., the prior), and the model error. It was also shown that the pdf of drifter position can quickly evolve into non-Gaussian shapes. Given that the instantaneous drifter position is a function of its flow history, it is relevant to ask if observed drifter positions, and their pdfs, can be inverted to estimate the flow field. This is the main motivation for this chapter.

One of the attractions of the self advecting vortex model (SAVM) is that the Eulerian flow field is defined by a low dimensional state space model. In fact, as discussed in the previous chapter, the instantaneous flow field is completely determined by the horizontal position of a "center" about which the various vortices rotate and "blink" on and off. If the true center is fixed (as is the case here), then the separatrices, hyperbolic and elliptic points of the system (and thus the Lagrangian character of the flow) are also determined by the position of the center. For the SAVM the dimension of the state vector is 2 (one dimension for each of the horizontal coordinates of the center) and the goal of the assimilation is to use observed drifter trajectories to infer the true positions of the center and also the drifters.

It will be assumed that only intermittent, noisy observations of drifter position are available, and that the position of the drifters and the center (and thus the instantaneous flow field) evolve according to a model which is subject to error. Optimal

blending of the noisy observations with the imperfect model requires careful attention to propagation of uncertainty in the position of the center and the drifters. The more accurately the position of the center can be estimated by the assimilation of the intermittent drifter observations, the better the complete instantaneous flow field can be reconstructed.

The assimilation procedure also provides the pdf of drifter position for past, present and future times. It is important to note that this type of information is potentially useful for a number of applications including marine search and rescue and pollution control, where a person lost at sea or an oil slick, would be represented by passively advected particles in the model. The pdfs of future particle position will map regions where, for example, the probability of a successful rescue operation is high.

Two data assimilation methods will be compared and contrasted: the ensemble Kalman filter (EnKF), which assumes all of the pdfs are Gaussian, and the more flexible particle filter (PF). One goal of this chapter is to identify the flow conditions under which the EnKF works well and when it fails. To quantify the performance of the assimilation schemes, two sets of metrics are introduced: one measures the evolution of non-Gaussianity in the advected pdfs, and the other measures the overall performances of the assimilation schemes.

Molcard [2003, 2005] has carried out related work using an idealized QG model. Molcard et al. [2003] showed how the quality of the assimilation of Lagrangian data depends on the observation interval, number of drifters and release domains of the drifters. The data assimilation method used by Molcard is a form of optimal interpolation (OI) and thus differs strongly from the methods used in this thesis. Salman et al. [2006] performed a similar study using a QG model for the background flow and the EnKF for data assimilation. Salman et al. [2006] suggested that the assimilation scheme fails in the vicinity of hyperbolic points, especially when the observation interval is long. This result was confirmed in the study of Apte et al. [2008] in which the performances of MCMC based methods and the EnKF for assimilating Lagrangian

data in an idealized shallow water model with similar characteristics as the self advecting vortex model introduced in Chapter 3.

The Lagrangian assimilation scheme of Molcard [2003] has subsequently been implemented into the primitive equation model MICOM (Molcard et al. [2005]). In the studies of Molcard et al. [2003], Molcard et al. [2005] and also of Salman et al. [2006] the drifters were released in regularly-spaced grids. In Molcard et al. [2006] this approach was dropped in favor of defining a launch strategy that improved the performance of the assimilation scheme. It was found that the assimilation scheme significantly improved if the drifters were initially launched along the outflowing manifolds of hyperbolic points. We will return to this point later.

The outline of the chapter is as follows. In Section 2 an illustrative example is presented that shows the type of information generated by the EnKF and PF and the different results that they can generate. In Section 3 a new metric is introduced for measuring the evolution of non-Gaussianity. In Section 4 this metric is used with to the PF and EnKF in some simple data assimilation experiments. The results give an initial indication about the conditions under which the EnKF will break down. In Section 5 various sensitivity experiments are performed to determine which parameters have the largest impact on the accuracy of the data assimilation methods. The chapter concludes with a brief summary and discussion of the results.

4.2 An Illustrative Example

A simple example of the assimilation of drifter data is given in this section to motivate the subsequent discussion of the performance of the PF and EnKF.

The black lines in Figure 4.1 show the trajectories of selected particles that have been advected by the SAVM model. The center of the SAVM was at the origin (shown by the black circle) and the Rankine vortices were centered at $y = 0$ and $x = \pm 1$ (shown by the black crosses). The swirl velocity of the vortices was $\omega = \pi/30$ and the threshold for the vortex was $\epsilon = 0.05$. (See Chapter 3 for definition of these

parameters.)

The trajectory of a drifter deployed at $x = -2.04$ and $y = 1$ over 100 time steps is shown by the red line in Figure 4.1. (The release point is marked by the red cross.) The red circles show the position of the drifter at observation times of $N_{obs} = 31$, 61 and 91.

In the following assimilation experiment it is assumed that the prior for the center is a Gaussian centered on $x = 0.4$ and $y = 0$ with an isotropic standard deviation of 0.4. In other words, before the data are assimilated it is believed that the most likely position for the center is $x = 0.4$, $y = 0$ but there is considerable uncertainty concerning this value (reflected by the standard deviation of the prior of 0.4). As the drifter observations are assimilated the pdf of center position will change and should move toward the true value of $x = 0$, $y = 0$ used to generate the observations. The more quickly the assimilation scheme predicts the center is at the origin, the more effective the scheme.

The model error for the center was taken to be $\sigma_c = 0$. This means that the center does not move (but the assimilation scheme still updates the pdf of the center as more data are assimilated.) The model error for the drifter was assumed to be $\sigma_d = 0.05$ and is applied only at observation times. The drifter observations to be assimilated by the EnKF and PF are given by the circles in Figure 4.1 i.e. the observations are perfect. However the assimilation scheme does not assume that the observations are perfect; the observation error is assumed to be $\sigma_{obs,d} = 0.05$. No useful observations of the center are available and so it is assumed that $\sigma_{obs,c}$ is a very large number. The size of the ensemble for the PF was taken to be 10000. The EnKF was run with ensemble sizes of 100, but averaged over 100 ensembles for an effective ensemble size that equalled that of the PF.

The time variation of the center and drifter positions estimated by the EnKF and PF are shown in Figure 4.2. The upper panels show results from the EnKF, and the lower panels show results from the PF. The first two columns of panels are for the x and y positions of the drifter; the third and fourth columns are for the x and

y positions of the center. For each panel, the red lines show the true position of the drifter or center, with drifter observations shown by the red circles. The black lines show the mean and standard deviation of the positions estimated by the two assimilation schemes (calculated from the ensembles).

The effect of the assimilation of the drifter observations is clearly evident in Figure 4.2; the drifter ensemble is pulled to the observation and the pdf of the center is updated. The shaded area shows that, as expected, the uncertainty decreases overall with the assimilation of data. Note however that after the assimilation step the uncertainty in drifter position grows again with the advection of the ensemble by the complex SAVM flow field. An important conclusion that can be drawn from Figure 4.2 is that the PF outperforms the EnKF in terms of estimating the present and future position of the drifter and center. Note that after the first observation is assimilated, the PF has accurately inferred the true position of the center and this results in accurate forecasts of drifter position. The EnKF on the other hand has difficulty in pinning down the correct position of the center (the black and red lines do not agree, and the shaded areas are wide) leading to problems in forecasting the true position of the drifter.

The above experiment (and Figures 4.1 and 4.2) are strongly dependent on the release point of the drifter. There are regions where the performance of PF and EnKF are almost the same. In the rest of this section an attempt is made to map out when, where and why the performance of the EnKF is lower than that of the PF. This will require highly idealized numerical experiments in order to understand what is happening e.g., the assimilation of a single drifter observation, with no model error for the center. It is important to note however that the numerical codes that have been developed can handle an arbitrary number of drifter and observation times, and non zero error for the evolution equations for the drifters and center.

4.3 Measuring Non-Gaussianity

In this section a new metric is defined to quantify the evolution of non-Gaussianity in the pdf of particle position. The new metric, denoted by Δ, is then used to predict the conditions under which the EnKF will fail. To motivate the definition of Δ some illustrative examples of the evolution of the pdf of drifter position by the SAVMare presented. For these examples (and all subsequent applications of the SAVM in this chapter) the swirl intensity is taken to be $\omega = \pi/30$ and the Rankine vortex threshold is taken to be $\epsilon = 0.05$. (See Chapter 3 for definition of these parameters.) The Rankine vortices are located at $y = 0$ and $x = \pm 1$; the separation of the vortices provides a horizontal scale for all the subsequent maps.

Consider first the upper left panel of Figure 4.3. In this case 100 drifters were released in a circular pattern coinciding with the red circle of radius $r_0 = 0.1$ centered on $x_0 = -2$, $y_0 = 0.4$. The drifter positions are subsequently integrated for $N = 80$ time steps. No model error is added to the position of the drifter ($\sigma_d = 0$) or the center ($\sigma_c = 0$). In effect this is a deterministic integration with the center remaining fixed and the drifters simply advected by the steady flow field. (The stochastic versions of the equations will be considered later.) The positions of the drifters after $N = 80$ time steps are shown by the small blue crosses. These release points for the drifter ensemble were centered on the inflowing manifold of the leftmost hyperbolic point and, as expected, approximately half of the drifters are quickly wrapped around the leftmost Rankine vortex and the other half are slowly swept around the lower ghost vortex. The non-linear evolution of drifter position is clearly evident.

A complex distribution of drifter positions is also obtained if the drifters are released near the inflowing manifold near the center ($x_0 = -0.1$, $y_0 = 0.1$, see lower left panel of Figure 4.3). The drifters in this case are essentially split into two equal groups that are carried around the ghost vortices.

The effect of releasing drifters in the vicinity of outflowing manifolds is shown by the right panels of Figure 4.3. In general the final distribution of the drifter positions

is simpler for these two sets of runs although some deviation from an elliptical shape is evident, suggesting some nonlinearity in the transformation from initial to final drifter position.

The effect of releasing the drifters in the vicinity of the Rankine and ghost vortices is shown in Figure 4.4. Drifters released in the ghost vortices (right panels) are hardly dispersed and the elliptical shape of the distribution suggest that the transformation from initial to final drifter position is nearly linear. Drifters released in the Rankine vortices (left panels) are strongly stirred due to the high shears in the vortex (with exception of the center where the velocity gradients are, by design, small). Although the overall final shape of the distribution is approximately elliptical the transformation from the initial to final drifter position is clearly not linear (e.g. a slight 'banana' is evident); the initial circular shape of drifter release locations has been transformed into a spiral.

4.3.1 A Simple Measure of Non-Gaussianity

Figure 4.5 shows the effect of various flow fields on the shape of an initially circular distribution of particles (denoted by the boundary of the blue patch). Note that for the upper left pair (A), an almost linear mapping has transformed the circle into an ellipse and so the mean of the final positions of the particles (denoted by the red cross) coincides almost exactly with the final position of the particle placed initially at the center of the circle (position is marked by the red circle). The second transformation (B) illustrated in Figure 4.5 shows the transformation to a "banana" like distribution after advection, and a non-zero distance between the mean of the particles initially around the circle and the center point after advection i.e. the red cross and red circle do not coincide. We will subsequently use this distance as a simple measure of the strength of non-gaussian transformations and denote it by Δ. In general Δ will depend on the position of the center of the initial circle (x_0, y_0), the radius of the circle (r_0), and the number of advecting time steps (N) so we write $\Delta(x_0, y_0|r_0, N)$.

Figure 4.5 also shows that there are situations where $\Delta \approx 0$ even though the

transformation is not linear e.g., the spiral and "S-shaped" transformations illustrated by C and D respectively. Thus $\Delta = 0$ is a necessary but not sufficient condition for a linear transformation. This is not surprising because it is impossible to find a single scalar measure that will cover all possible types of transformation. Two practical advantages of Δ are its ease of calculation and interpretation.

4.3.2 Non-Gaussianity and the SAV

The drifter distributions shown in Figure 4.3 and 4.4 suggest that the EnKF will likely perform well in the ghost vortices where the Gaussian assumption is valid and $\Delta \approx 0$. To obtain a more precise delineation of regions where the EnKF is expected to fail, and perform well, Figure 4.6 maps $\Delta(x_0, y_0|r_0, N)$ over a 241×241 grid of release points (x_0, y_0) centered on a circle of radius $r_0 = 0.2$, for 4 integration times ($N = 20,\ 40,\ 60$ and 80).

The regions of large $\Delta(x_0, y_0|r_0, N)$ in Figure 4.6 generally correspond with regions of strong curvature and high speed. Overall the largest values of Δ are found either around the Rankine vortices or along the inflowing manifolds that approach the 3 hyperbolic points. This becomes increasingly evident as the length of the integration (i.e., N) increases. These are also the regions where the EnKF may break down. This will be checked in the following sections. Figure 4.7 shows the effect on Δ of increasing r_0 from 0.2 to 0.3. Overall the map retains the same features although the magnitude of Δ and extent the of regions of large Δ generally increases with r_0.

One feature evident in Figures 4.6 and 4.7 is that Δ is zero exactly along the inflowing manifold approaching the hyperbolic point at the origin. This is due to the generation of the "S-shaped" distributions discussed above which, while non elliptical, still have $\Delta = 0$. It is also clear from Figure 4.6 that as N increases, the values of Δ approach zero around the Rankine vortices. This is the spiral effect evident in case C in Figure 4.5.

The drifter distributions shown in Figures 4.3 and 4.4 are consistent with the lower right panel of Figure 4.6 corresponding to $N = 80$. (The magenta circles in this panel

mark the drifter release points in Figures 4.3 and 4.4.) Drifters released at locations where $\Delta(x_0, y_0|r_0, N)$ is near zero (e.g., the magenta circles in the dark blue regions in Figure 4.6) tend to retain an elliptical shape. On the other hand, drifters released at locations where $\Delta(x_0, y_0|r_0, N)$ is high, (e.g., the magenta circles in the light blue to red regions) evolve into "banana" shaped contours.

4.4 Using Δ to Predict EnKF Failure

Some illustrative examples of data assimilation are now presented for both the EnKF and the PF. It will be shown that the map of Δ is useful in interpreting the performance of the EnKF.

Some relevant data assimilation experiments have been performed by Kuznetsov et al. [2003] and Ide et al. [2002]. In these works a continuous time analogue of the SAVM was used in conjunction with the extended Kalman filter. The results with respect to the release locations found by Kuznetsov et al. [2003] generally agree with the results for the EnKF presented in this section. The main differences with the present work is the use of the EnKF rather than the simpler EKF, the use of the SAVM which is much simpler and more computationally efficient than its continuous time analogue, and finally the use of Δ to predict when the EnKF will perform poorly.

The experiments performed in this subsection are similar to the runs illustrated in Figures 4.3 and 4.4. The main difference is that a data assimilation step is included halfway through the model run. Each experiment consists of a set of two assimilation runs, one using the PF and the other using the EnKF. The only parameter that varies amongst the sets of experiments is the release point of the drifters.

For each set of runs, all drifters in the ensemble are released at the same point and so the standard deviation of prior drifter position, $\sigma_{pr,d}$, is 0 (i.e., the release point of the drifter is assumed known). The prior for the center was Gaussian with a mean at the origin and an isotropic standard deviation of 0.2, i.e. the exact position of the center, and thus the flow field, is not known exactly and is to be inferred from the

drifter observations. The standard deviation of the model error for the drifter and center are taken to be $\sigma_d = 0.025$ and $\sigma_c = 0$ respectively. The model is integrated for $N = 80$ time steps and an observation of drifter position data is assimilated half way through the integration (at time step $N_{obs} = 41$).

To generate the observed drifter position at time step 41, a reference run was made with (i) the same drifter release location, (ii) the center at the origin, and (iii) $\sigma_d = \sigma_c = 0$. The form of the EnKF implemented in this thesis does not generate pseudo observations because this leads to an inflation of the observation error variance (the observations are already subject to error) but it was assumed that the observation noise was not zero i.e., $\sigma_{obs,d} = 0.05$. Thus the assimilation scheme is fed perfect observations but told that the observations are imperfect. No observations of the center are used; this corresponds to taking a very large value for $\sigma_{obs,c}$. The initial assumed position of the center is centered on the origin with an uncertainty given by $\sigma_{obs,c}$. The goal of the experiments is to determine how quickly the two data assimilation methods can determine the center is at the origin given observations associated with different drifter release points. The position of the center defines the complete flow of the SAVM and so the faster the assimilation method finds the true location of the center (i.e. the origin) the better the assimilation scheme can predict the flow and drifter position.

The PF was run with $N_{PF} = 10000$ ensemble members. The EnKF was run by taking 100 ensembles each of size 100 and averaging the forecast error covariance matrices. In this way the (statistical) number of ensemble members for the PF and the EnKF are effectively the same. Applying the EnKF with one ensemble of size 10000 instead would 'blow' up the model due to the large matrices and the subsequent memory usage involved.

Figure 4.8 shows the performance of the PF for a release point in the vicinity of inflowing manifolds (see corresponding magenta circle in the upper right panel of Figure 4.6). The Δ map for $N = 40$ and $r_0 = 0.2$ in Figure 4.6 suggests that this release point will be dispersed in a non-gaussian shape and the EnKF may encounter

difficulties. The top left panel shows the prior pdf for the center (shown by the circle contours centered on the origin), and also the release point of the drifters. The top right panel shows the marginal pdf of drifter position just before the data are assimilated (blue lines). Given $\sigma_c = 0$ the marginal pdf of the position of the center has not changed. The bottom left panel shows the posterior pdf of drifter position after the drifter observation (shown by the green dot) has been assimilated. Note that the pdf has collapsed onto the observation point with a standard deviation that is effectively set by $\sigma_{obs,d}$. The posterior pdf of the center would be shown by the red contours, which in this case have collapsed on the red cross, i. e., the PF has effectively pinned down the correct position of the center (shown by the posterior pdf concentrated at the origin). The red and blue crosses show the mean of the pdf for the center and drifter positions respectively. The bottom right panel shows the marginal pdf after 40 additional time steps following the assimilation step. Note the pdf is again starting to develop non-gaussian contours.

Figure 4.9 shows the corresponding results for the EnKF. In this case the pdf at time steps 40, 41 and 80 are shown as a Gaussian pdf with the same mean and standard deviation as the ensemble. It is clear (compare the upper right panels of Figure 4.8 and 4.9) that the Gaussian assumption in this case is highly suspect with the result that there are significant errors in the marginal pdf of drifter position and the center at time step 80 (lower right panel). It should be noted that the ensemble of the EnKF is resampled from the posterior pdf (i.e. only the mean and covariance are used to regenerate the ensemble) and this avoids the collapse of the ensemble; the resampling also destroys any information beyond the second moment, i. e. any information on the non-Gaussian structure of the pdf. It is important to note that the PF will always give the correct pdf for the drifters and center if the model (including error parameters are correct and a sufficiently large number of particles are used. Hence any discrepancies between the results from the PF and EnKF are due to errors in the EnKF.

Figures 4.10 and 4.11 show the performance of the PF and EnKF respectively for

drifters released near the inflowing manifold at $(-0.25, 0.25)$, close to the center (see corresponding magenta circle in the upper right panel of Figure 4.6). As the drifters approach the center they can follow, with approximately equal probability, one of the two outflowing manifolds. The right top panel of Figure 4.10 shows the marginal pdf (blue contour lines) of the drifter locations just before the data assimilation step. As expected the marginal pdf follows approximately the contour lines of the outflowing manifold. The corresponding Gaussian pdf in the upper right panel of Figure 4.11 covers a larger region, indicating greater uncertainty in the true drifter location when expressed as a Gaussian distribution.

The assimilation at time step 41 resets the positions of the drifters in the PF and EnKF at approximately at the same location (blue cross). The PF estimates the center (red cross) of the SAVM almost exactly at the true location $(0, 0)$ with a small uncertainty (Figure 4.10, lower left panel). The EnKF estimates the center to be slightly south of the true center $(0, 0)$ but with a significant uncertainty remaining. As a result, the drifter location at time step 80 is estimated with a high uncertainty too, compared to the PF (compare lower right panels of Figure 4.10 and 4.11). The results of the data assimilation methods confirm, similar to the previous experiment, the predictions made by the map of Δ.

Although the EnKF can fail in regions of strong nonlinearity, it can perform well elsewhere. One region where the EnKF performs as well as the PF is in the ghost vortices. Figures 4.12 and 4.13 show the marginal pdf plots for the same four time steps as above. The pdf of drifter position computed by the EnKF match the results of the PF almost perfectly. Given that the EnKF uses an order of magnitude fewer ensemble members than the PF it is the preferred assimilation method in such regions.

To provide an overall picture of where the EnKF fails, and where it performs well, Figure 4.14 shows a map of the distance between the mean position of the center calculated by the EnKF and the PF just after the drifter observation is assimilated ($|x_{KF} - x_{PF}|$). To produce this map the experiments illustrated in Figures 4.8 to 4.12 were repeated for a grid of release points. At each release point, i.e. grid point

of the SAV, an ensemble of drifters were released. At a specified time step one drifter observation was assimilated by the PF and EnKF respectively. This procedure was applied in sequence to each release point individually. For each release point the distance between the mean position of the center calculated by the EnKF and PF was calculated and contoured over the SAVM domain. Four different maps were produced for a single drifter observation assimilated at time steps $N_{obs} = 21$, 41, 61 and 81, respectively.

There is good qualitative agreement between the maps of Δ and $|x_{KF} - x_{PF}|$ for the center (compare Figure 4.6 and Figure 4.14). In particular the map of Δ correctly predicts that the EnKF will perform poorly in the vicinity of the Rankine vortices and along the inflowing manifolds. It also predicts correctly that the region of poor performance of the EnKF (compared to the PF) extends with increasing N_{obs}. Hence the map in Figure 4.6 can be used to anticipate which assimilation method will be most effective for particular regions and observation intervals. This point is reconsidered in the final section of this chapter.

4.5 Sensitivity to Model Parameters

The sensitivity of the performance of the EnKF to changes in model parameters, and the time to the first observation, N_{obs}, is explored. The methodology is similar to that used above: for a given set of parameters the EnKF and PF are used to assimilate a single drifter observation generated by running the SAVM without model error and with the center exactly at the origin. The release point of the drifter is assumed known ($\sigma_{pr,d} = 0$) and the position of the center is assumed uncertain with a Gaussian prior with a mean at the origin and a standard deviation of $\sigma_{pr,c}$. Although an error free observation of drifter position is fed to the assimilation scheme, the standard deviation of the observation error, $\sigma_{obs,d}$, is not assumed zero. As above, the standard deviation of the model error for the center is taken to be $\sigma_c = 0$.

4.5.1 Performance Metrics

The metrics used to assess the performance of the assimilation schemes are based on distances between positions of the center or drifter, based on truth or estimates by the assimilation schemes. The following four metrics are used:

1. $|\overline{x}_{PF} - x_T|_d$ for the drifter: This metric is the distance between the mean position of the drifter calculated using the PF and the true drifter position (x_T).

2. $|\overline{x}_{PF} - x_T|_c$ for the center: Similar to the above metric but for mean position of the center.

3. $|\overline{x}_{KF} - \overline{x}_{PF}|_d$ for the drifter: This metric is the distance between the mean position of the drifter calculated using the PF and the EnKF. This is taken as the error in drifter position due to using the EnKF.

4. $|\overline{x}_{KF} - \overline{x}_{PF}|_c$ for the center: Similar to the above metric but for mean position of the center.

All of the metrics were evaluated at time step $2N_{obs}$. The reason for not using time N_{obs} is that the assimilation pulls the drifter ensemble back to the observation and the spread of the ensemble is effectively controlled by $\sigma_{obs,d}$. (This is evident in the lower left panels of Figures 4.8 to 4.13) By evaluating the metrics at $2N_{obs}$ allows the errors in the center to affect the drifter position after the assimilation time.

For a given set of parameters the above calculation was performed for a set of release points on a 64×64 grid that covered the model domain. The above metrics for each release point were then contoured across the SAVM domain. The PF used an ensemble size of 10000. The PF also uses 30 MCMC steps and the prior as the proposal distribution. The EnKF used an ensemble size of 100 particles but the experiments were repeated 100 times and the results averaged to give an effective sample size of 10000.

A typical set of plots for the four metrics is shown in Figure 4.15. The upper panels show the distance between the mean drifter position and truth (left panel)

and center and truth (right panel). The lower panels show the distances between the drifter position calculated with the EnKF and PF (left panel) and center calculated with the EnKF and PF (right panel). By varying the parameters it has been possible to map out the sensitivity of the four metrics to changes in the prior for the center ($\sigma_{pr,c}$), the time to the first observation (N_{obs}), the size of the model error for the drifter (σ_d), and the observation error ($\sigma_{obs,d}$). The exact parameter settings for the various experiments, and the numbers of the figures that map the metrics, are given in Table 4.1.

Exp.	σ_d	$\sigma_{pr,c}$	$\sigma_{obs,d}$	N_{obs}	Figure
E0	0.025	0.2	0.05	20	4.15
E1	0.025	**0.1**	0.05	20	4.16
E2	0.025	**0.3**	0.05	20	4.17
E3	0.025	0.2	0.05	**10**	4.18
E4	0.025	0.2	0.05	**40**	4.19
E5	**0.050**	0.2	0.05	20	4.20
E6	**0.075**	0.2	0.05	20	4.21
E7	0.025	0.2	**0.10**	20	4.22

Table 4.1: Model parameters used in experiments E0 to E7. σ_d is the drifter model error, $\sigma_{pr,c}$ is the standard deviation of the prior pdf for the center, $\sigma_{obs,d}$ is the drifter observation error, N_{obs} is the time to the first observation. The total length of each run is $2N_{obs}$. The threshold and swirl speed for the Rankine vortices were 0.05 and $\pi/30$ respectively. The model noise for the center is zero. The position of the center is not observed and so $\sigma_{obs,c}$ is effectively infinite.

4.5.2 The Sensitivity Maps

Maps of $|\overline{x}_{PF}-x_T|$ and $|\overline{x}_{KF}-\overline{x}_{PF}|$ for the drifter and center are plotted in Figures 4.15 to 4.22. The format is the same for each of the figures. The upper panels show errors in PF estimates of position relative to their true values. The lower panels show the distance between the estimates of drifter and center position calculated by the EnKF and PF. Given the PF will always give the correct answer given a sufficiently large

ensemble size, the lower panels can be interpreted as errors associated with the use of the KF.

Overall similar features can be seen in all of the plots. The upper right panels show $|\overline{x}_{PF} - x_T|$ for the center. The largest errors made by the PF when predicting the position of the center are associated with drifters that are released in the far field (i.e., $|x|$ greater than about 2). The reason is that the flow in the far field is not sensitive to changes of the position of the center and so drifters moving in this region do not effectively constrain the center. Closer to the center, and the positions of the Rankine vortices, the values of $|\overline{x}_{PF} - x_T|$ for the center show a complex structure. There are clearly some release points in this region for which the PF has difficulty pinning down the center. The PF generally does well if the drifters are released in the ghost vortices (centered on $x = 0$). This can be explained by the small velocity gradients in this region and hence the slow, and easily predictable, movement of the drifter trajectories. Another region of drifter release for which the PF can accurately predict the drifter position is close to the center of the Rankine vortices. For computational reasons the Rankine vortices were designed to have near solid body rotation around their centers (i.e., through the use of the threshold value). One consequence is that within the threshold region the drifter trajectories are simple and the PF has little difficulty inferring the true position of the center.

The upper left panels of Figures 4.15 to 4.22 show $|\overline{x}_{PF} - x_T|$ for the drifter. As for the center, the PF has difficulty pinning down the true position of the drifter in the far field. This is due to the difficulty in pinning down the center, and also the strong velocities in the far field that tend to amplify errors. The PF also has difficulty pinning down the position of the drifter for specific release regions close to the Rankine vortices. More detailed calculations (not shown) were unsuccessful in providing a simple explanation of these features in the error map. Close to the centers of the Rankine vortices the stronger velocity gradients may result in a so called chaotic advection (see Aref [1984]).

The upper left panels also show for almost all experiments two (symmetric) locations at about $(x, y) = (1.5, 045)$ and $(x, y) = (1.5, -0.45)$ with local maxima of high errors $|\overline{x}_{PF} - x_T|_d$. The path of drifters released in the near vicinity of these two locations tend to be split up into two or more directions: One path spirals the drifters clockwise towards the center of the Rankine vortex, the other paths evolve along the outflowing manifolds of the hyperbolic points which the drifters pass by. The magnitude of these two local maxima are proportional with the initial spread of the center locations, $\sigma_{pr,c}$, and the length of the observation interval, N_{obs}. Figure 4.23 shows the final distribution (blue dots) of 200 drifters after 100 time steps. The figure shows that the drifters follow three distinct paths. The drifters were released on a circle (red), centered on $(x, y) = (-1.5, 0.45)$, of radius=0.2 (corresponding to $\sigma_{pr,c} = 0.2$ as used in most experiments).

The lower panels of Figures 4.15 to 4.22 show the error in position of drifter and center made by the EnKF relative to the PF. The most important point to note is that, for all of the figures, the error maps for the center and drifter (left and right panels) are qualitatively similar to each other and to the maps of Δ shown in Figures 4.6 and 4.7 . This means that the complex spatial structure of the error in the KF can be explained in terms of the simple measure of non-Gaussianity, Δ. In general the EnKF performs poorly in the vicinity of the Rankine vortices and on the inflowing manifolds as mentioned earlier. Figures 4.15 to 4.22 show that this conclusion is robust with respect to changes in the parameters listed in Table 4.1.

The rest of this subsection describes some of the differences between the various figures. To simplify the discussion experiment E0 (Figure 4.15) is taken as the reference case against which the other sensitivity runs are compared.

Experiments E1 and E2: These experiments explore the importance of prior information about the position of the center (see Figures 4.16 and 4.17). In experiment E1 the value of $\sigma_{pr,c}$ is smaller than the reference, thus reflecting greater initial certainty about the position of the center. In E2 the prior standard deviation reflects less certainty about the initial position of the center.

As expected, reducing $\sigma_{pr,c}$ benefits both the EnKF and PF; all 4 performance metrics generally decrease with lower $\sigma_{pr,c}$. It is noticeable that the performance of the EnKF relative to the PF (lower panels) retains the same qualitative shape when $\sigma_{pr,c}$ is changed. It is also apparent from experiment E1 (see Figure 4.16) that the EnKF can perform well if the center is initially well known.

Comparison of the lower panels of Figures 4.15 to 4.17 with the Δ maps (upper left panels of Figures 4.6 and 4.7) shows that nonlinearity around the Rankine vortices has the largest impact on the estimated position of the center and drifter. Comparing Figures 4.6 and 4.7 shows the effect on Δ of changing the uncertainty in the release point of the drifter with respect to the center. Similar behavior is also evident in the maps of $|\overline{x}_{KF} - \overline{x}_{PF}|$; as $\sigma_{pr,c}$ increases so does the magnitude and extent of $|\overline{x}_{KF} - \overline{x}_{PF}|$ for the drifter and center.

Experiments E3 and E4: These experiments explore the importance of the observation interval N_{obs} (see Figures 4.18 and 4.19). In experiments E3 and E4 the observation intervals are half and double of the reference observation time of 20 time steps respectively. Generally the accuracy of the assimilation improves with decreasing observation interval as expected. To quantify the dependence of the error on N_{obs} the total error in the domain region of interest (i.e. the far flow field is excluded) was calculated for E3, E0 and E4, respectively. The evolution of the total error with respect to N_{obs} shows a linear dependence.

In E3 (with relatively frequent observation) the accuracy of the predicted drifter positions is almost identical for the EnKF and PF as expected. This is because there is insufficient time for nonlinearity in the flow to generate a non-Gaussian pdf. This behavior is also evident in the Δ maps; as the time interval decreases so does the magnitude and extent of the high Δ regions.

Experiments E5 and E6: These experiments explore the importance of the model error for the drifter, σ_d (see Figures 4.20 and 4.21). In experiments E5 and E6 the drifter model error is twice and three times the reference value, respectively. The

main result is that the results for E5 and E6 are nearly identical to E0. This implies that the drifter model error has little impact on the data assimilation as long as the error is within certain bounds. The reason is that the drifter is brought back to the observed value when the data are assimilated. To quantify the dependence of the error on σ_d the total error in the domain region of interest (.i e. the far flow field is excluded) was calculated for experiments with $\sigma_d = 0.025$ $(E0)$, $0.05(E5)$, 0.075 $(E6)$, 0.15, 0.3 and 0.5, respectively. The evolution of the total error with respect to $\sigma_{obs,d}$ slightly oscillates. From $\sigma_d = 0.025$ to 0.05 the error increases, decreases for 0.075 and increases thereafter until $\sigma_d = 0.3$. For $\sigma_d = 0.5$ the error decreases again but the assimilation nevertheless breaks down and the center position can not longer be determined.

EXPERIMENT E7: In this experiment the impact of drifter observation error $\sigma_{obs,d}$ is examined. The observation error of the drifter, $\sigma_{obs,d}$, is doubled compared to E0. Comparing Figures 4.22 and 4.15 shows only minimal differences in the results. The reason is as above: the drifter is brought back to the observed value when the data are assimilated. To quantify the dependence of the error on $\sigma_{obs,d}$ the total error in the domain region of interest (i.e., the far flow field is excluded) was calculated for experiments with $\sigma_{obs,d} = 0.1$ (E8), 0.15, 0.2 and 0.3, respectively. The evolution of the total error with respect to $\sigma_{obs,d}$ shows a weak quadratic dependence. For a value of $\sigma_{obs,d} > 0.3$ the assimilation breaks down and the center position can no longer be determined.

4.6 Discussion

Although the dimension of the state vector describing the SAVM is only 2, the SAV's flow fields are quite complex and the pdf of passively advected particles can quickly become non-Gaussian for particles released in some regions. One way to describe the steady flow field of the SAVM is using concepts taken from dynamical systems theory. By tracing the trajectories of passively advected particles it is possible to show that the flow field has 4 elliptic points and 3 hyperbolic points. The hyperbolic

points each have inflowing and outflowing manifolds. It is also possible to define separatrices that partition the flow domain into separate regions that do not allow exchange of advected particles. While these concepts are fundamental, they are of limited usefulness when the flow field, and the drifter trajectories, evolve with time according to a stochastic updating equation, i.e., when the model for drifter and center motion includes a random error term.

In order to quantify the evolution of non-Gaussianity in the pdfs of drifter and center position, the metric Δ was introduced. The metric has two attractive features: It is easy to calculate and simple to interpret. It was shown that Δ is useful in terms of understanding why the EnKF will perform poorly compared to the PF. In general the regions of poor EnKF performance for the flow fields considered here coincide with inflowing manifolds and the fast, strongly curved flows around the Rankine vortices. If the drifters are released in these regions the flow field should be estimated using the computationally more expensive PF. It is important to note that inflowing manifolds will not always be associated with poor EnKF performance. This is clear when one considers a simple flow field with velocity field given by $u = x$, $v = -y$. This flow field has a hyperbolic point at the origin but the linearity of the transformation means that the performance of the EnKF will match that of the PF everywhere. This is consistent with the map of Δ which is zero everywhere for this flow field.

The four performance metrics gave some insight into the best regions to deploy drifters in order to estimate the flow field of the SAV. It was shown that deploying the drifters in the ghost zones generally gives good estimates of the positions of the center (and thus the positions of the self advecting vortices). For the parameter settings considered in this study, if the self advecting vortices are centered on $x = \pm 1$, $y = 0$ then the drifter should ideally be released in the vicinity of $x = 0$, $y = \pm 1$.

Sensitivity experiments were performed to understand how the performance of the data assimilation schemes depend on the model parameters and the way the observations are collected. Overall the parameters that have the strongest impact on the performance of the data assimilation are (i) the release point of the drifter, (ii)

the time between the drifter observations, and (iii) the prior pdf for the flow field. For the studies undertaken here, the observation error for drifter position, and also model error for drifter evolution did not have much impact on the performance of the schemes.

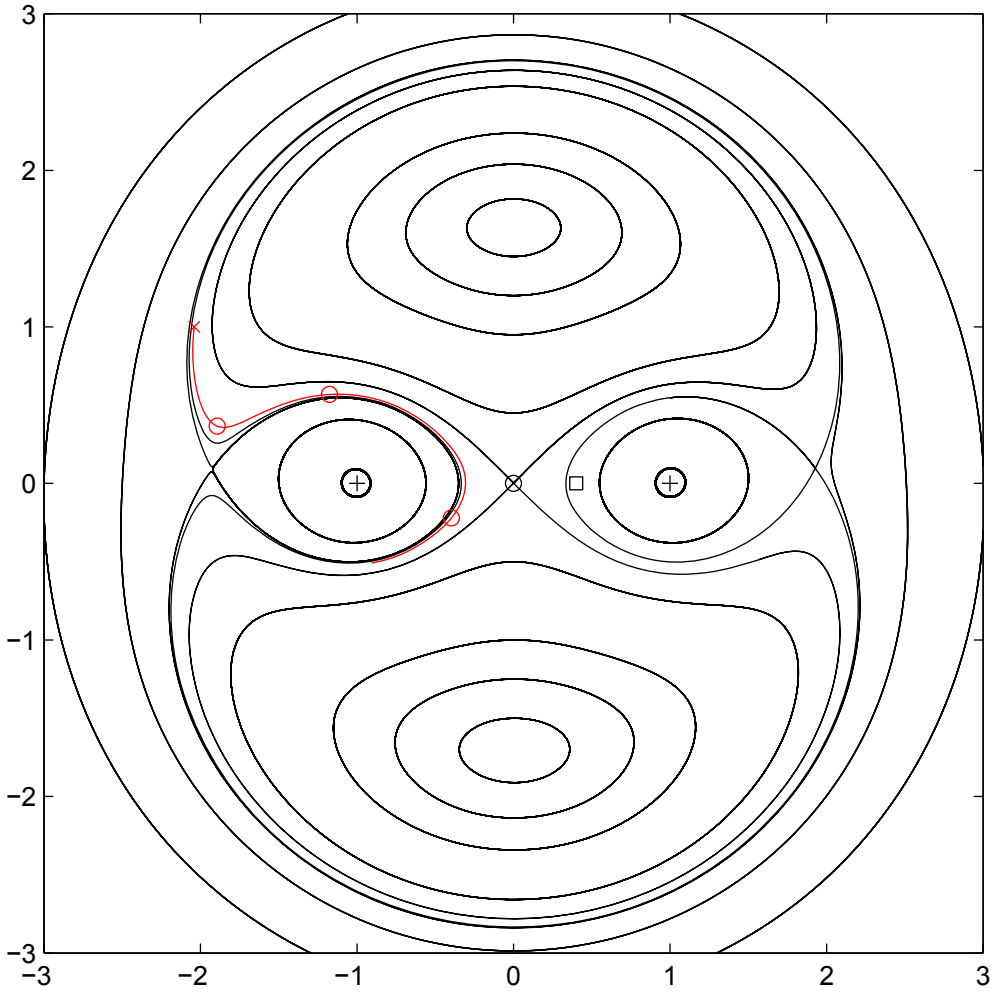

Figure 4.1: Basic setup of the illustrative example. The black lines show the trajectories of selected particles advected by the SAVM and are used to visualize the flow field. The red line is the trajectory of a drifter over a time span of 100 time steps after being released at the red cross. The red circles show the positions of the particle at times $N_{obs} = 31, 61$ and 91. The black circle shows the center of the SAVM and the black crosses show the center of the Rankine vortices used to generate the trajectories. The black square marks the center of the Gaussian prior for the center used in the example.

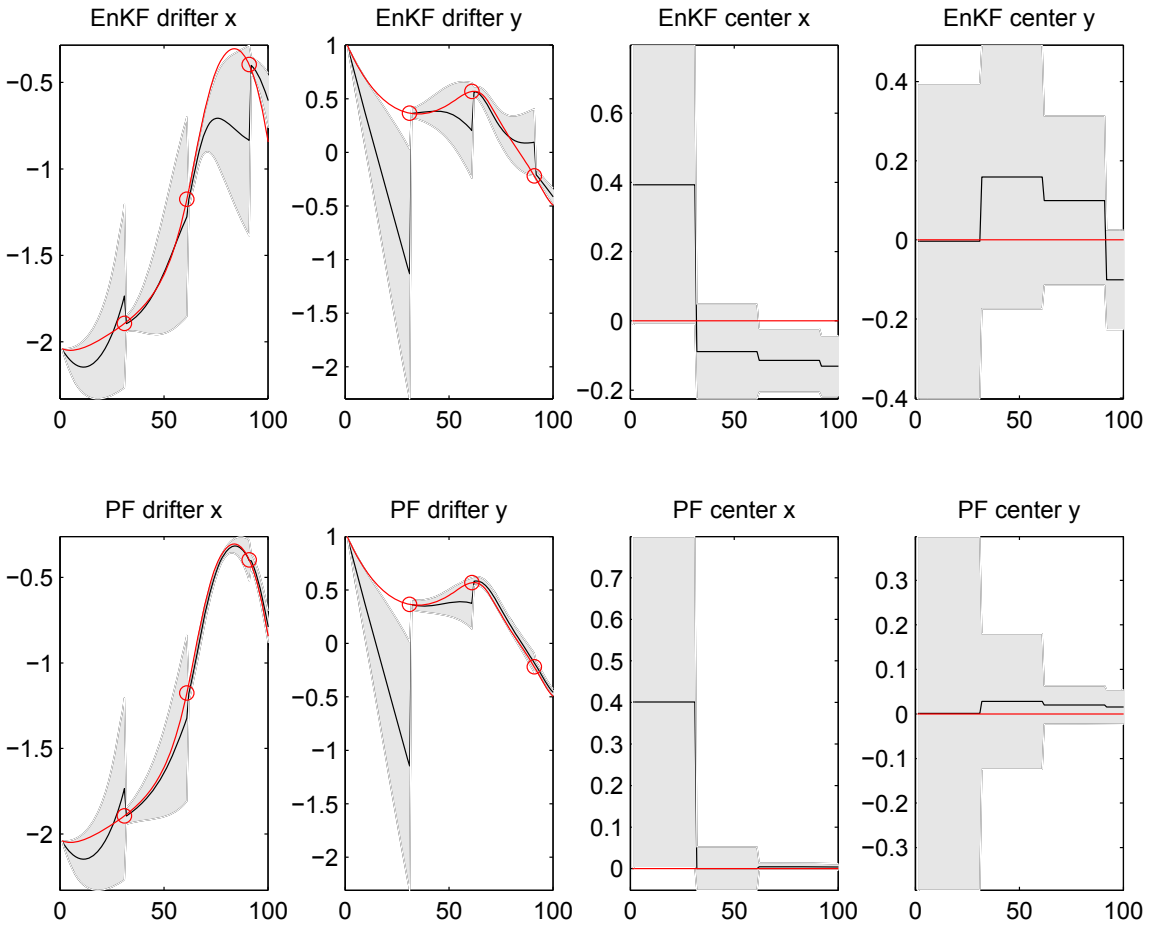

Figure 4.2: Performance of the EnKF and PF for the illustrative example. The red lines show the true trajectory of the drifter and the center. (Note that the true center does not move in this experiment.) The red circles correspond to the circles in Figure 4.1 and mark the observed drifter locations that were assimilated. The black lines show the trajectory of the mean drifter position estimated by the EnKF and PF; the shaded areas show the standard deviation of the marginal pdf of drifter and center position calculated from the ensembles.

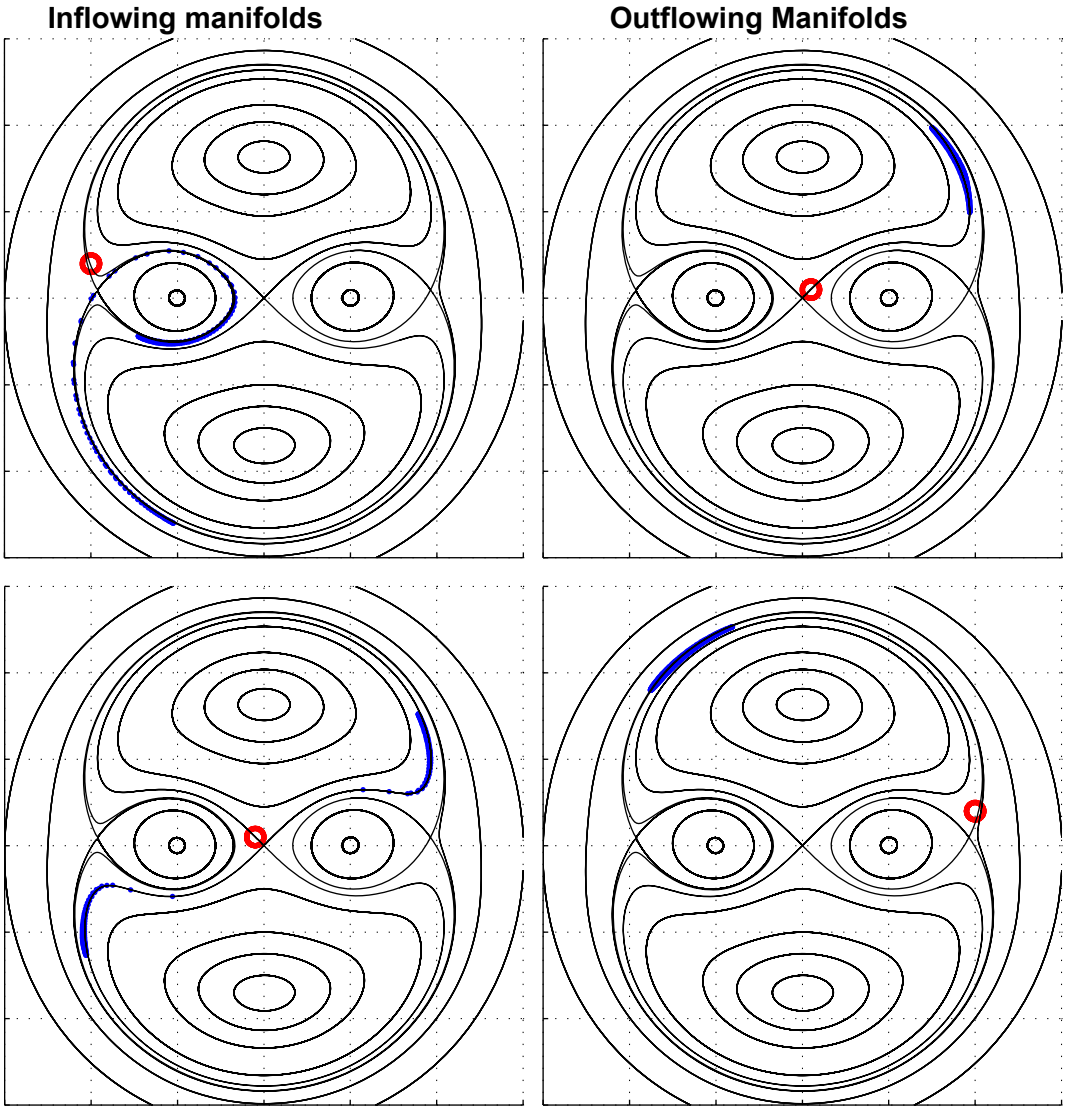

Figure 4.3: Typical drifter displacements as a function of release point. For each panel the red circle marks the release points of 100 drifters evenly distributed around a circle of radius $r_0 = 0.1$. The blue crosses show the positions of the drifters after $N = 80$ time steps. The left panels show typical distributions of drifters released in the vicinity of inflowing manifolds (circles centered on $(-2.0, 0.4)$ and $(-0.1, 0.1)$). The right panels correspond to outflowing manifolds (circles centered on $(2.0, 0.4)$ and $(0.1, 0.1)$). The black lines show the trajectories of drifters advected with no model error (σ_d, σ_c both zero) and a center exactly at the origin.

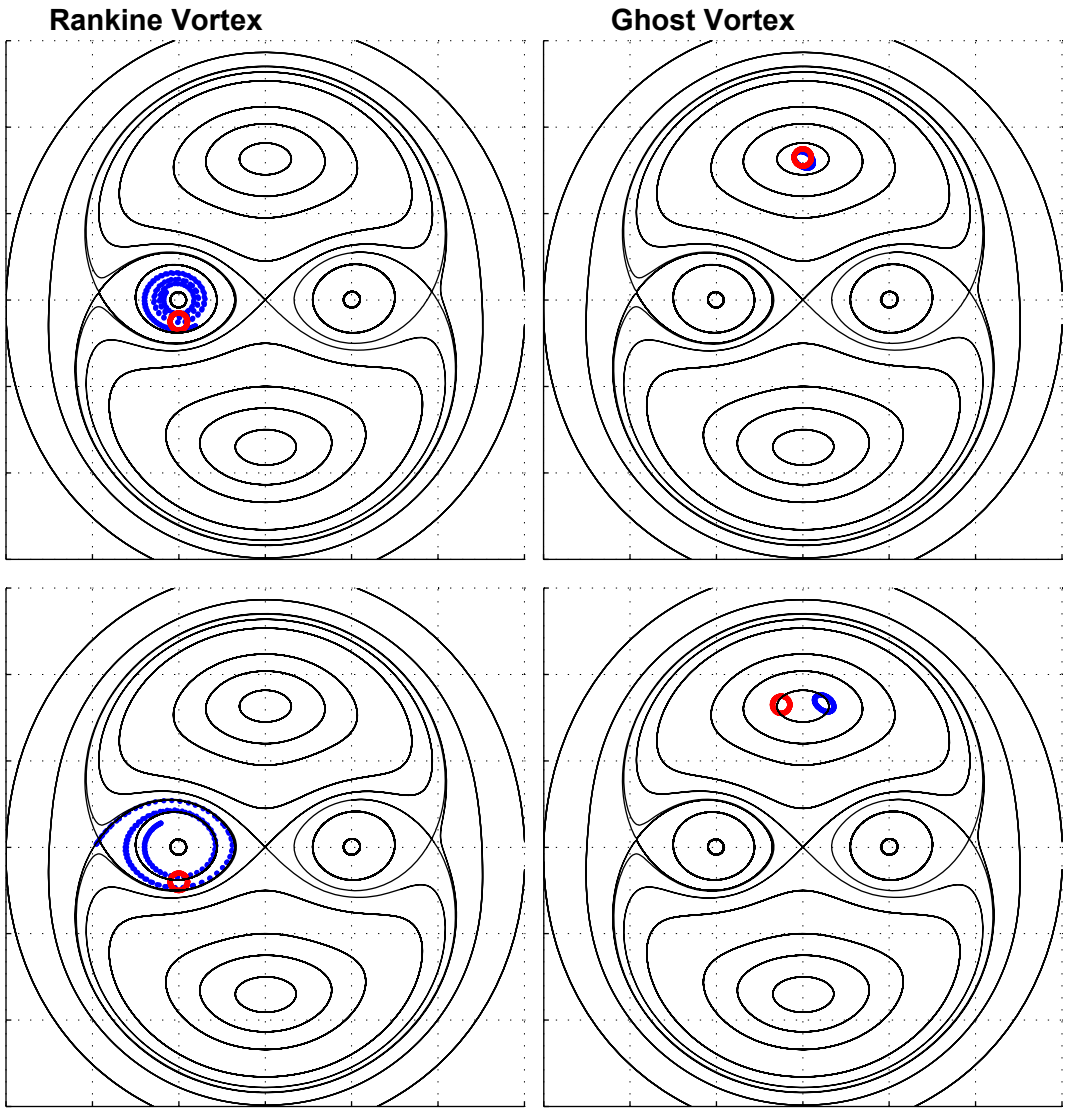

Figure 4.4: Typical drifter displacements as a function of release point. Same format as Figure 4.3. The left panels show flow pattern of drifters released in the vicinity of a Rankine vortex, $(-1.0, -0.25)$ and $(-1.0, -0.4)$ respectively; the right panels show the flow patterns of drifters released in the ghost vortices, $(0.0, 1.65)$ and $(-0.25, 1.65)$ respectively.

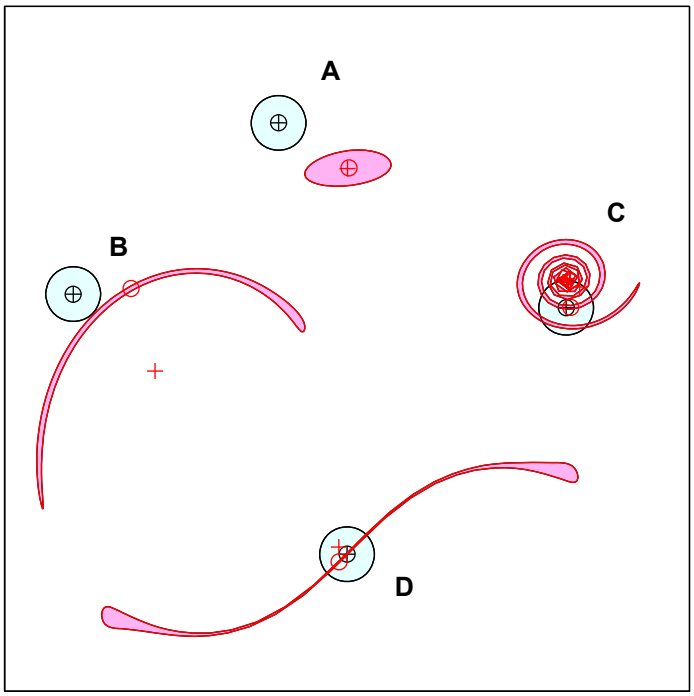

Figure 4.5: A simple measure of non-Gaussianity. Four pairs of typical pdf are shown before and after advection. The circular boundary of the blue patch shows the initial distribution of a large number particles centered on the black cross. The mean position of the particles is marked by the inner circle which, for the blue patch, coincides with the cross. The boundary of the red patch shows the position of the particles after advection. The red circle marks the position of the particle that was initially at the center of the circle; the red cross is the mean position of the drifters that initially formed the boundary of the blue patch. The nonlinearity measure Δ is the distance between the red cross and the red circle. Note that for these examples the area of the red and blue patches are identical because the mapping is area preserving (and generated by the SAV).

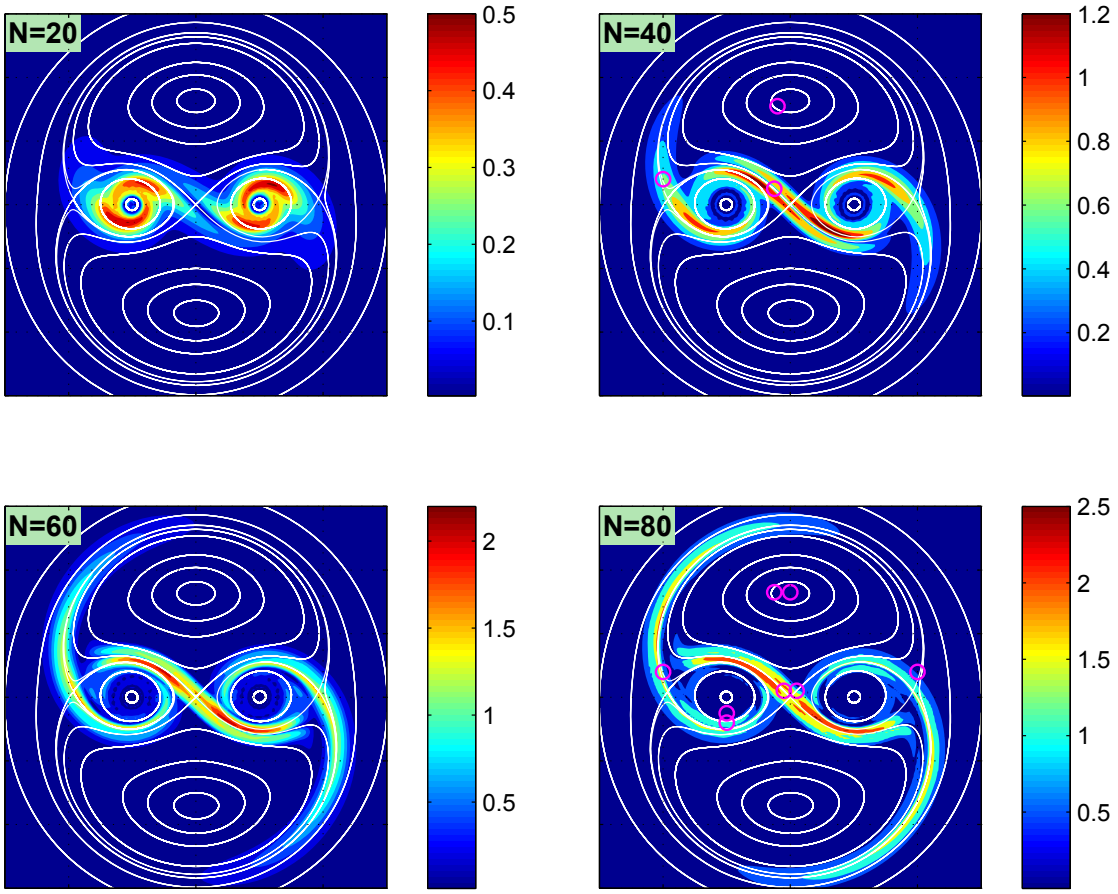

Figure 4.6: Map of $\Delta(x_0, y_0|r_0, N)$ for $N = 20,\ 40,\ 60$ and 80 as a function of release point (x_0, y_0). The radius r_0 of the release circle is $r_0 = 0.2$. The 8 magenta circles in the lower right panel ($N = 80$) show the drifter release locations in Figures 4.3 and 4.4. The 3 magenta circles in the upper right panel ($N = 40$) show the release locations as used for figures 4.8 4.9, 4.10, 4.11, 4.12 and 4.13

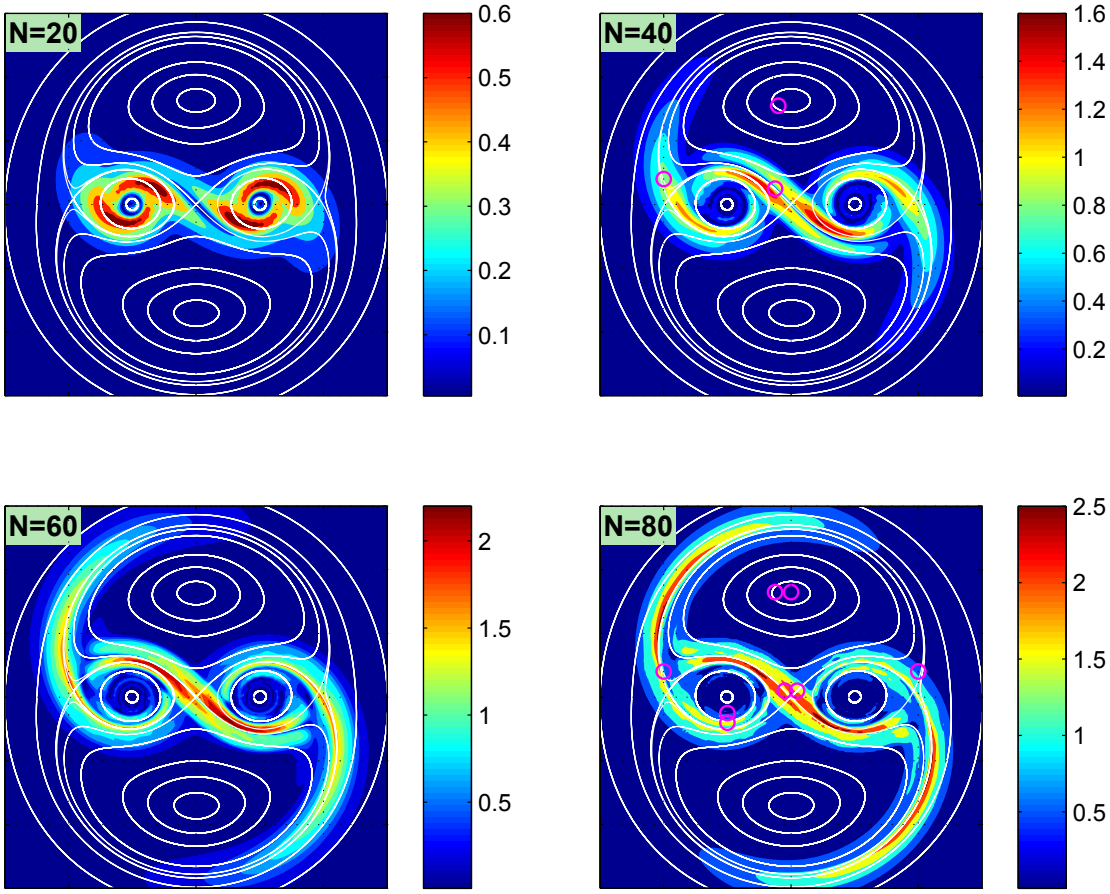

Figure 4.7: Map of $\Delta(x_0, y_0|r_0, N)$, same format as for Figure 4.6 but with an increased $r_0 = 0.3$.

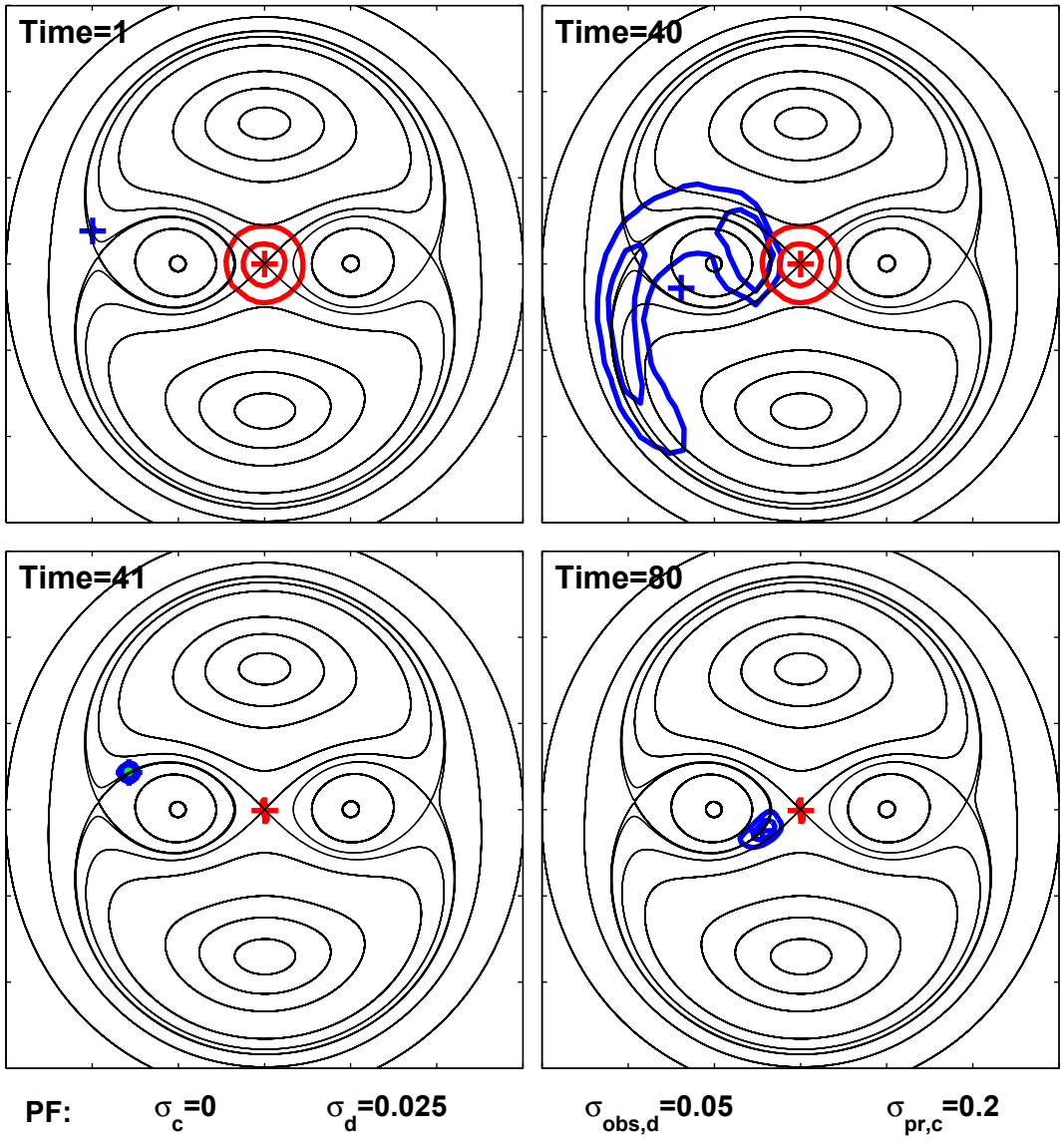

Figure 4.8: Typical results from the Particle Filter. The upper left panel shows the release point of the drifters $(-2.0, 0.4)$ and the prior pdf of the center (red lines). For all marginal pdfs the isolines mark the 50 and 90 percentiles. In the remaining panels the red and blue contours show the pdf of the position of the center and drifter respectively. One single drifter observation (green dot) is assimilated at time step 41 (lower left panel). The red and blue crosses show the mean position of the ensemble of centers and drifter positions respectively.

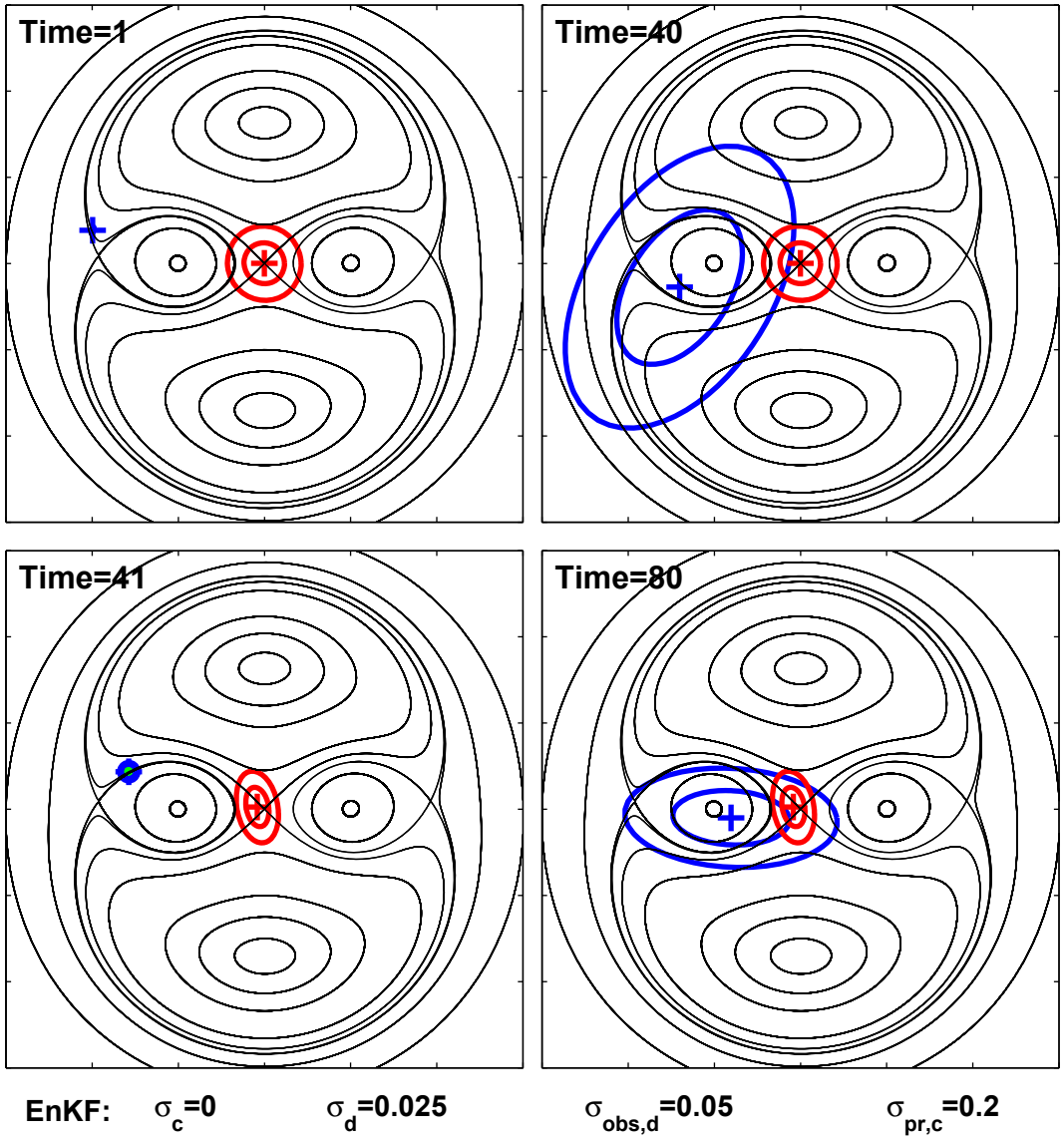

Figure 4.9: Performance of the EnKF. Same format as Figure 4.8.

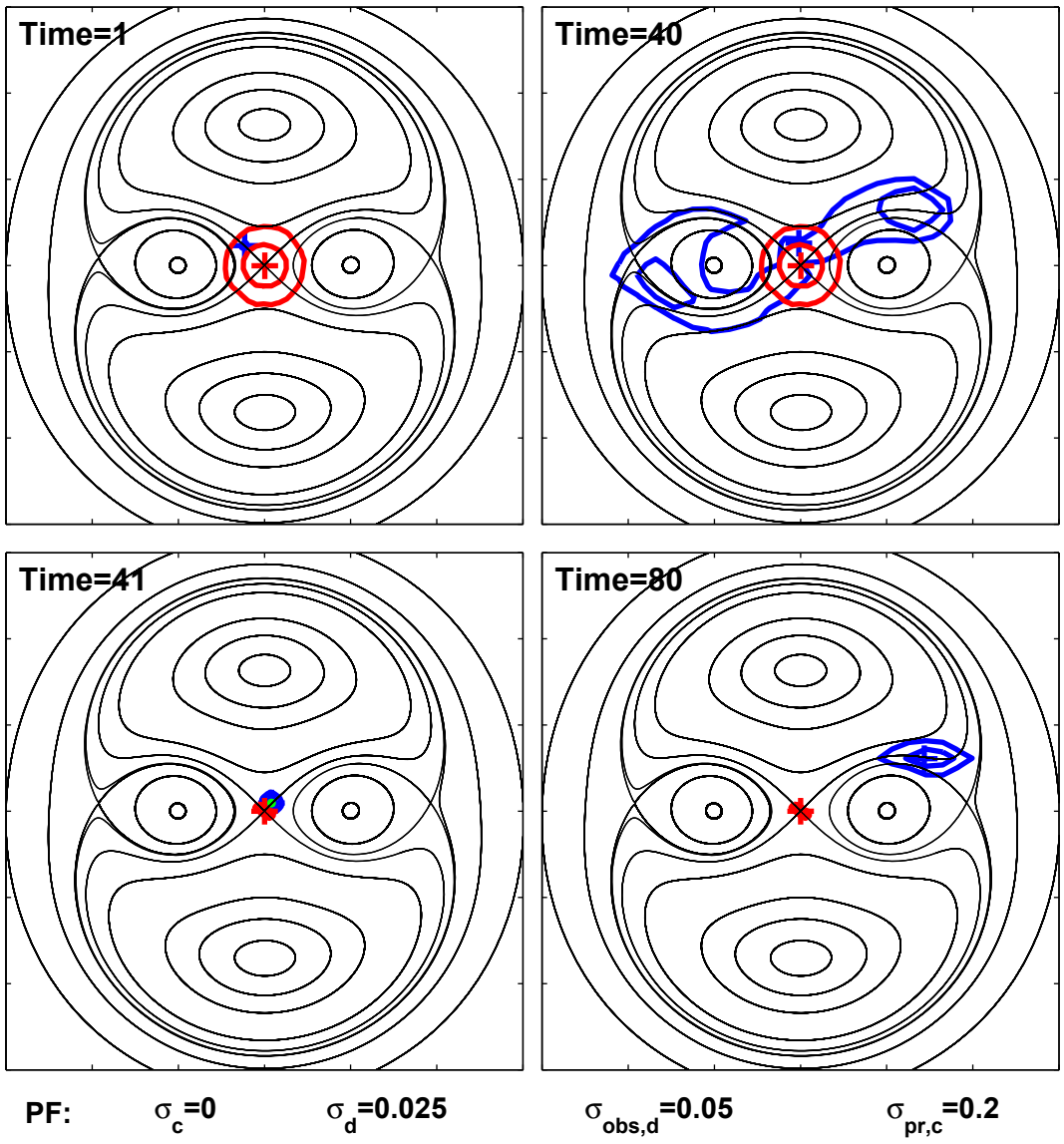

Figure 4.10: Performance of the Particle Filter. Same format for Figure 4.8.

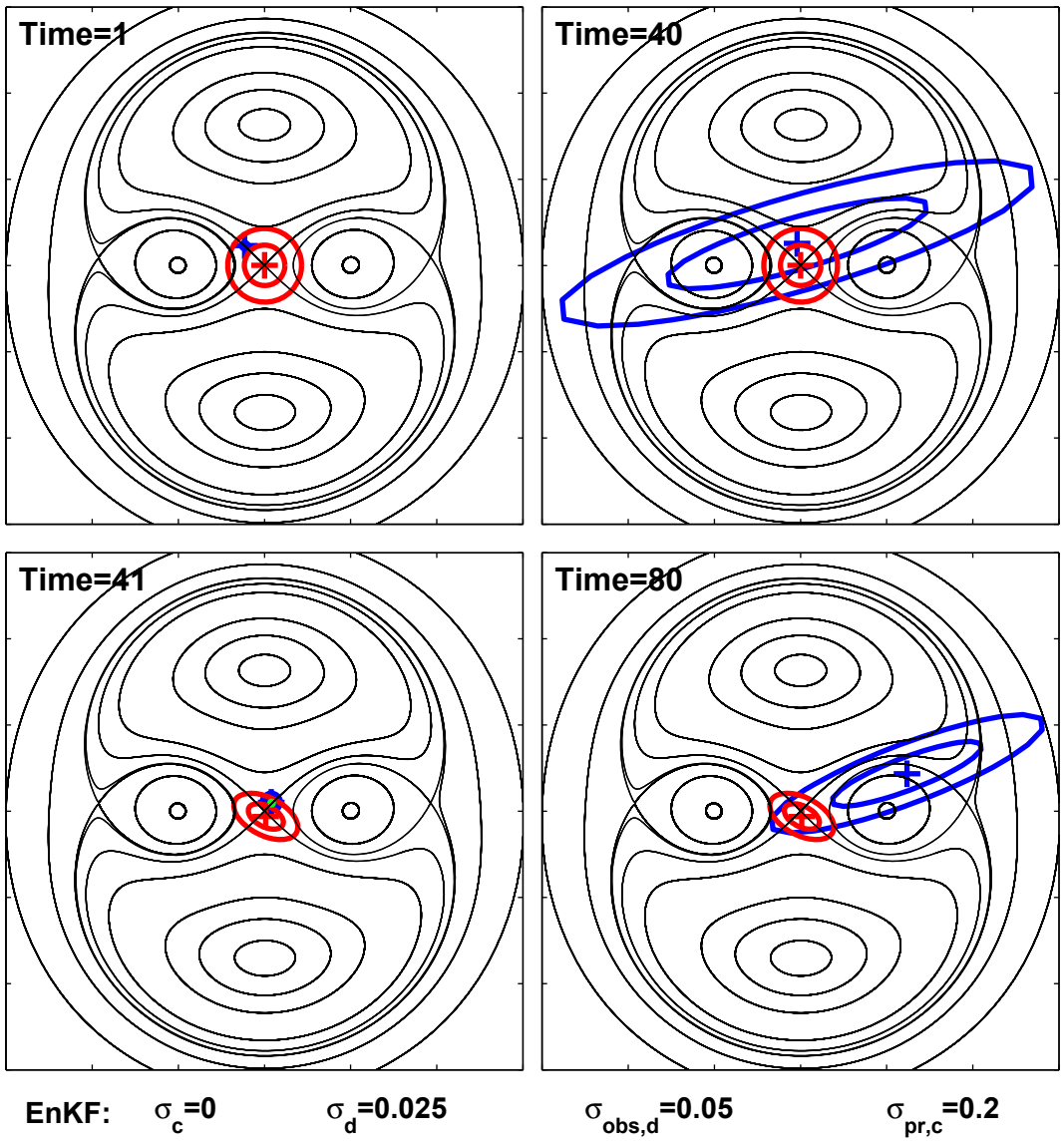

Figure 4.11: Performance of the EnKF. Same format as Figure 4.8.

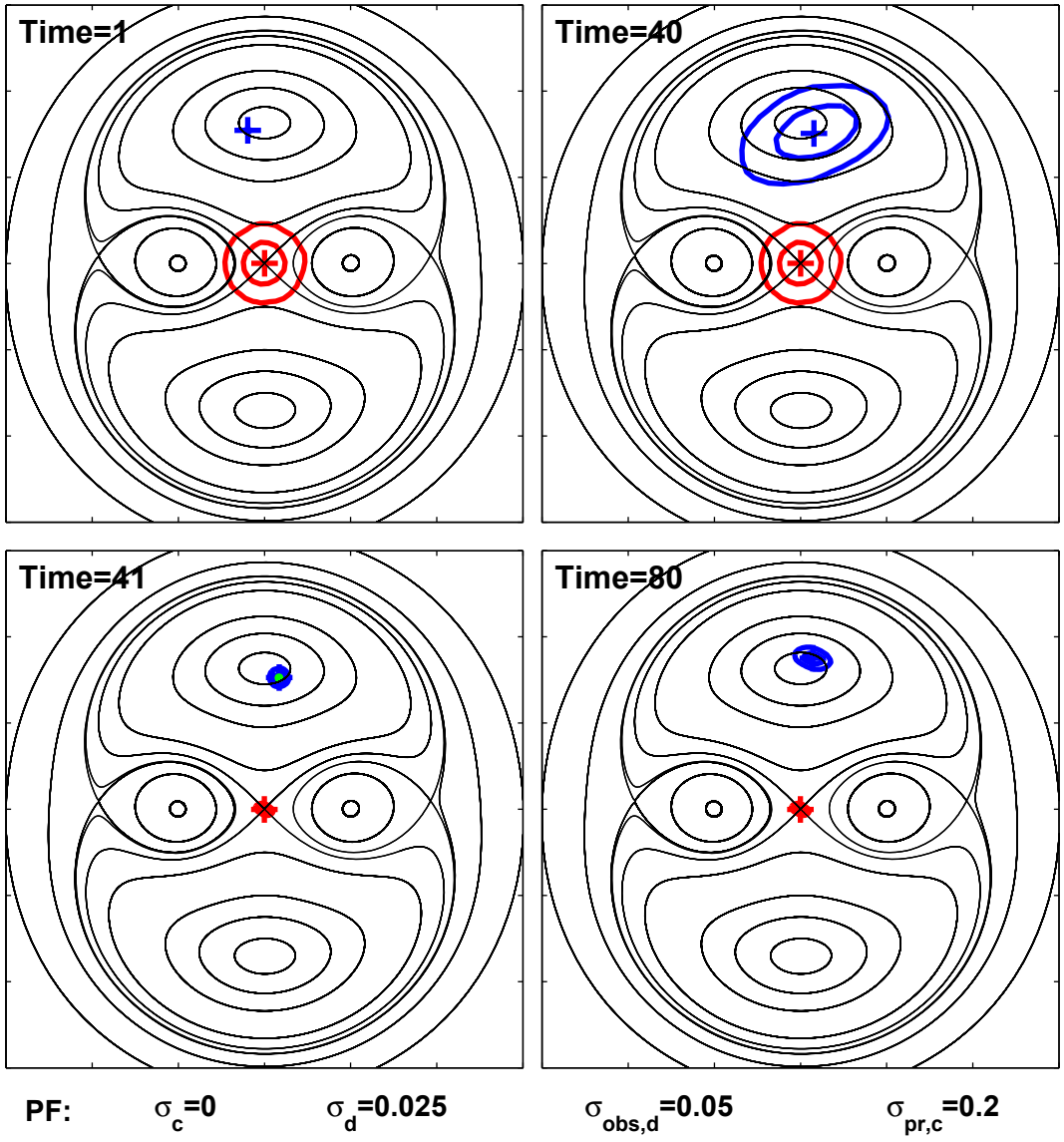

Figure 4.12: Performance of the Particle Filter. Same format for Figure 4.8 .

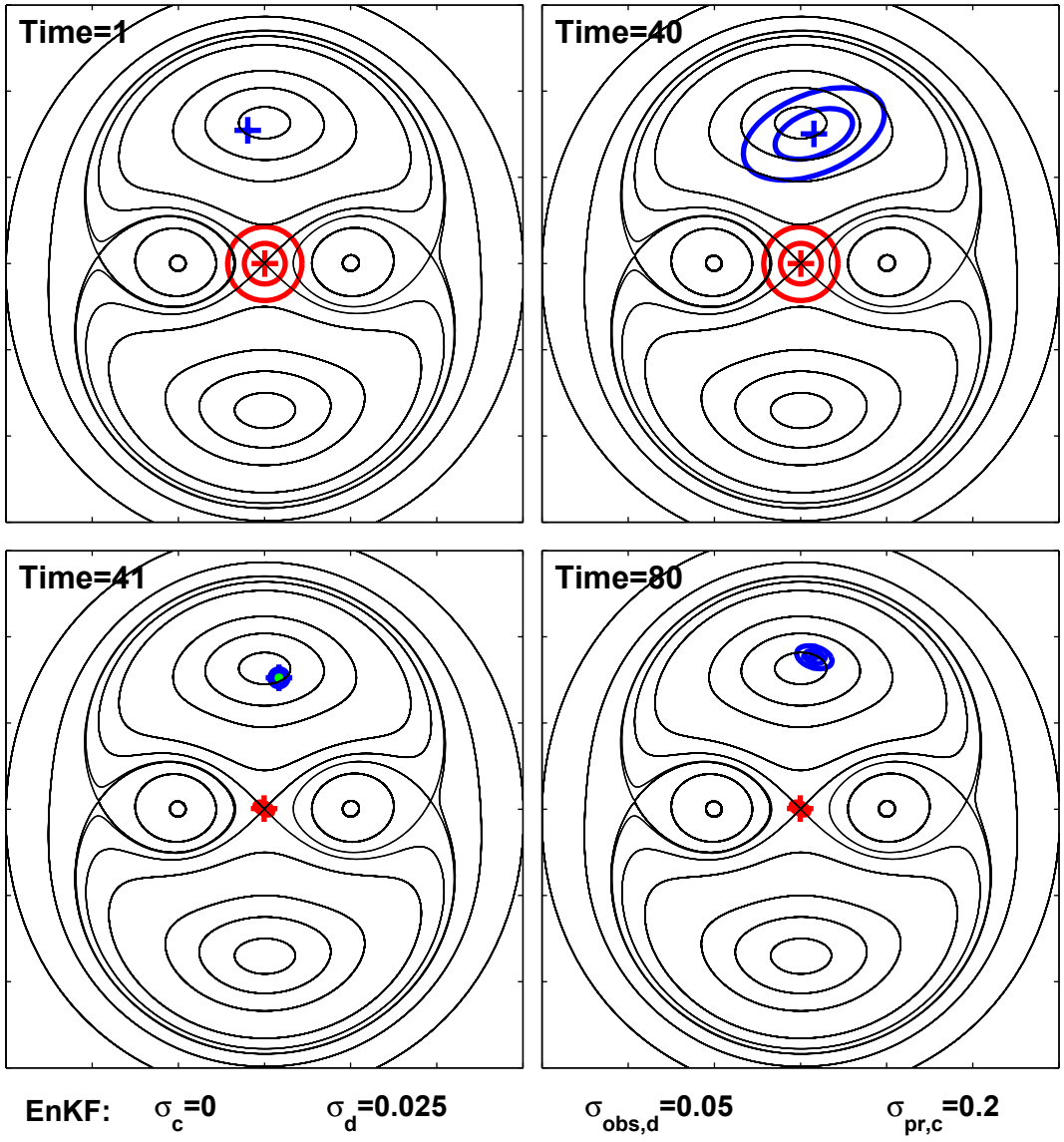

Figure 4.13: Performance of the EnKF. Same format as Figure 4.8.

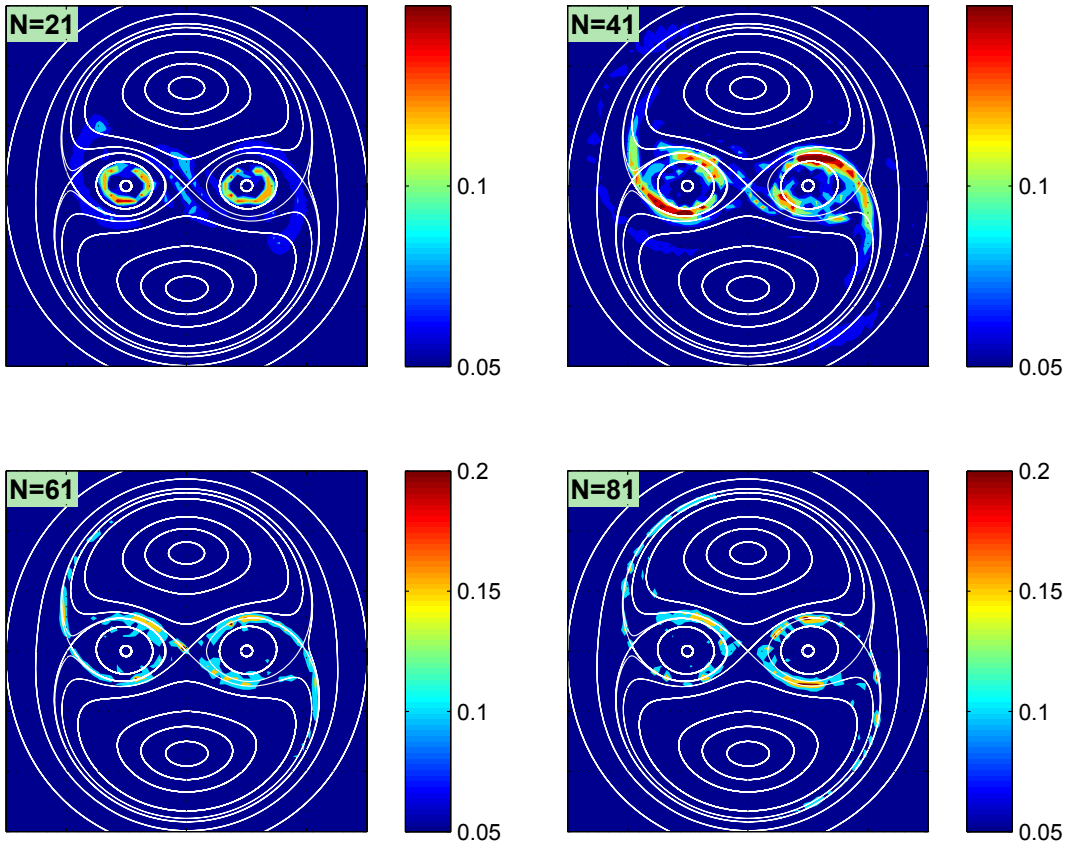

Figure 4.14: Error in the EnKF as a function of drifter release point. The shading shows the distance between the mean position of the center calculated by the EnKF and PF. Each panel corresponds to a different time, N_{obs}, just before a single drifter observation is assimilated. The following model parameters were used for all of the four experiments: model error for the drifter and center locations are $\sigma_d = 0.025$ and $\sigma_c = 0$, respectively, and the observation error for the drifter is $\sigma_{obs,d} = 0.05$, no observations are available for the location of the center. The initial spread (variance) of the center locations for each state vector is $\sigma_{pr,c} = 0.2$.

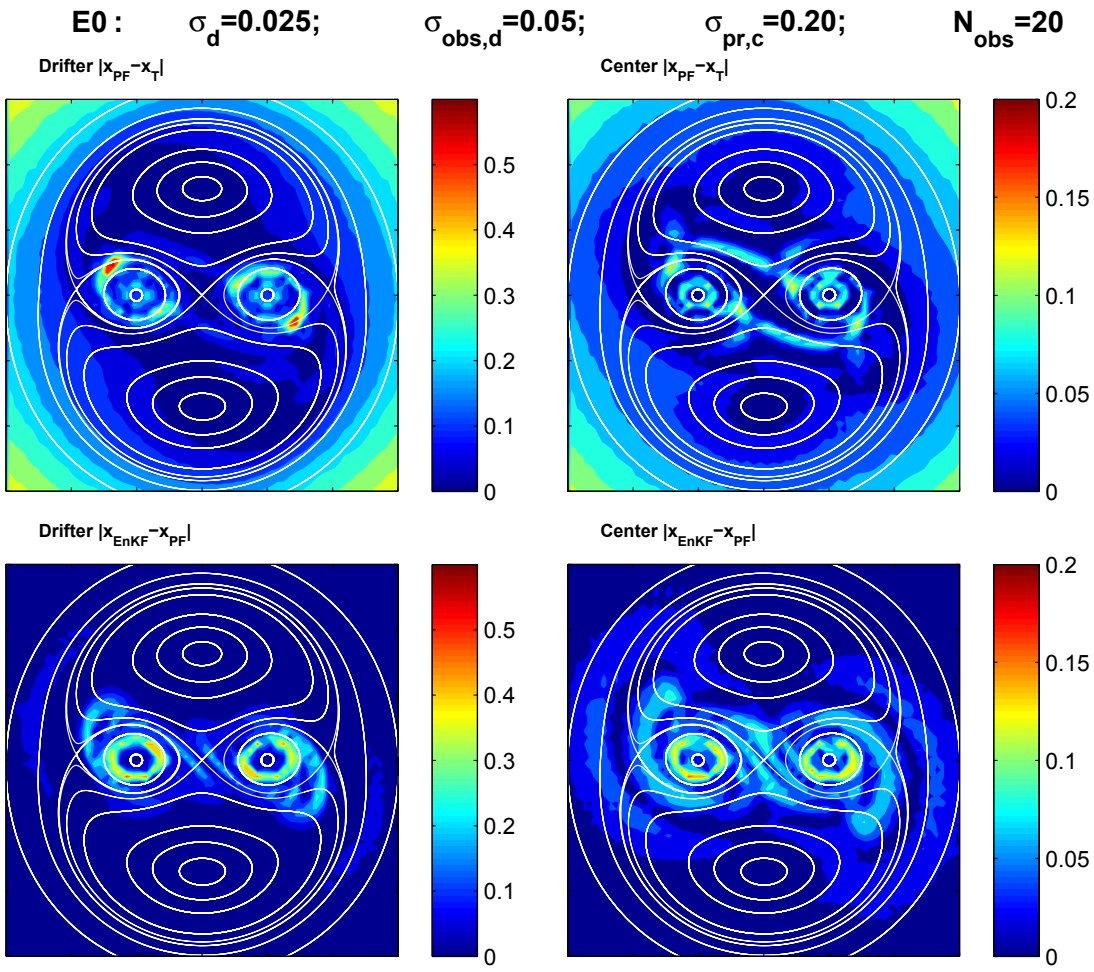

Figure 4.15: Error maps for experiment E0 (see Table 4.1 for parameter settings). The white isolines show the steady particle trajectories. The upper panels show the distance between the true location (of drifter and center) and the mean locations computed using the PF; the lower panels show the distance between positions estimated by the EnKF and the PF.

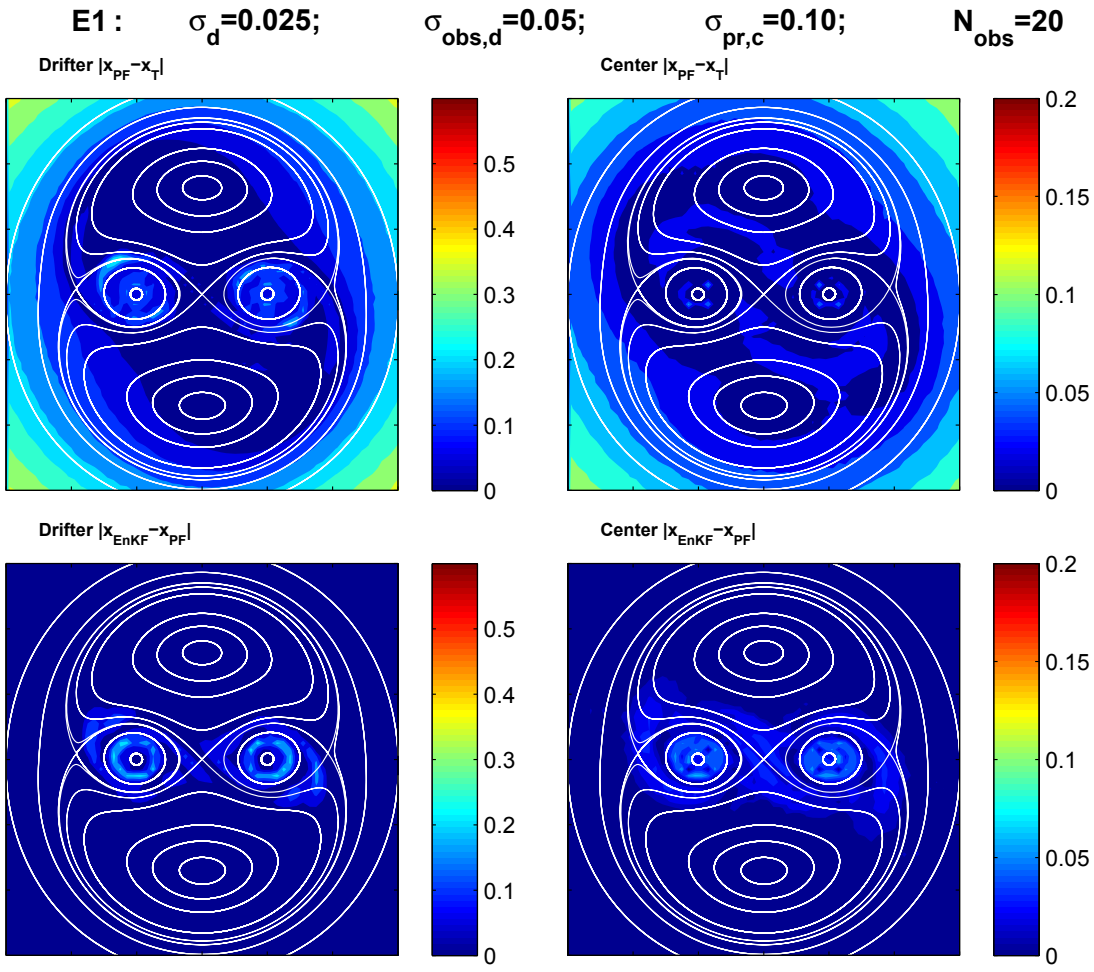

Figure 4.16: Performance metrics for experiment E1. Same format as Figure 4.15.

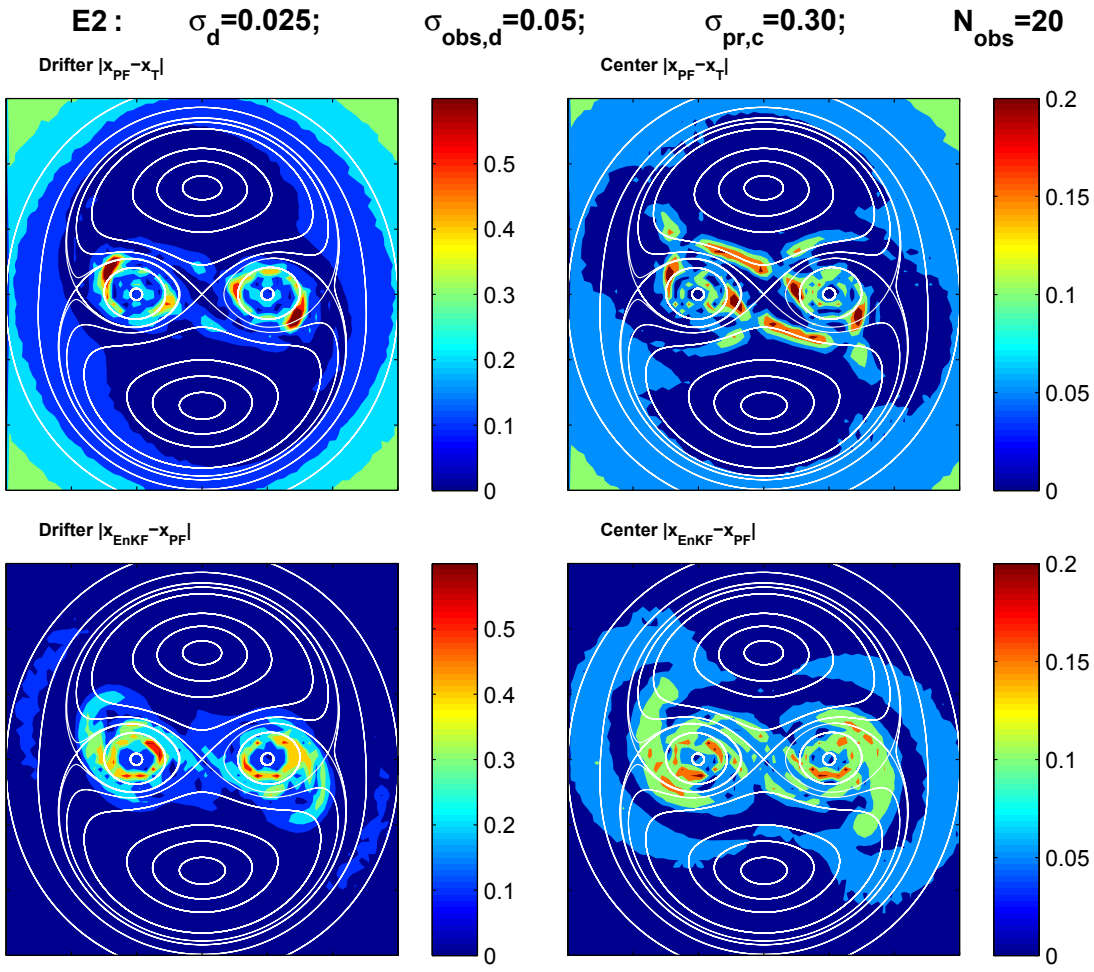

Figure 4.17: Performance metrics for experiment E2. Same format as Figure 4.15.

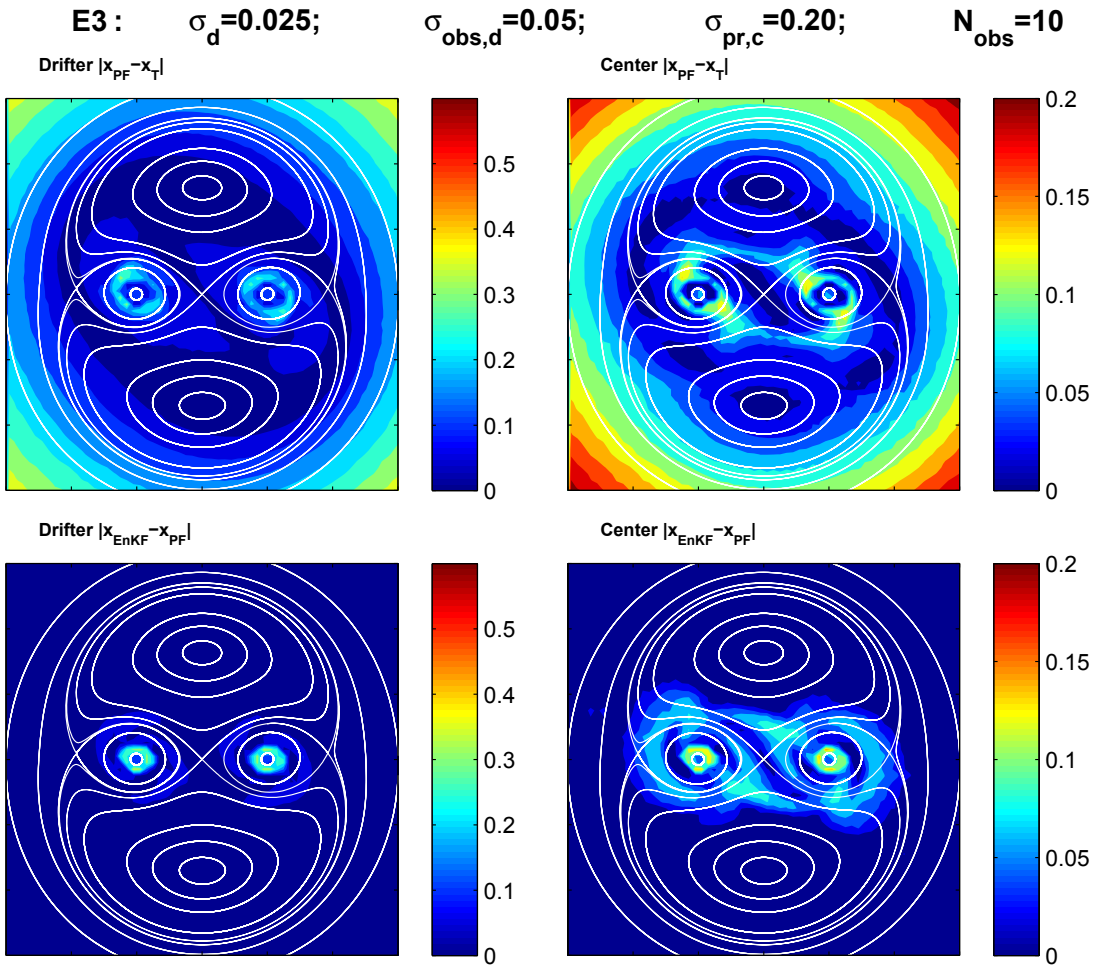

Figure 4.18: Performance metrics for experiment E3. Same format as Figure 4.15.

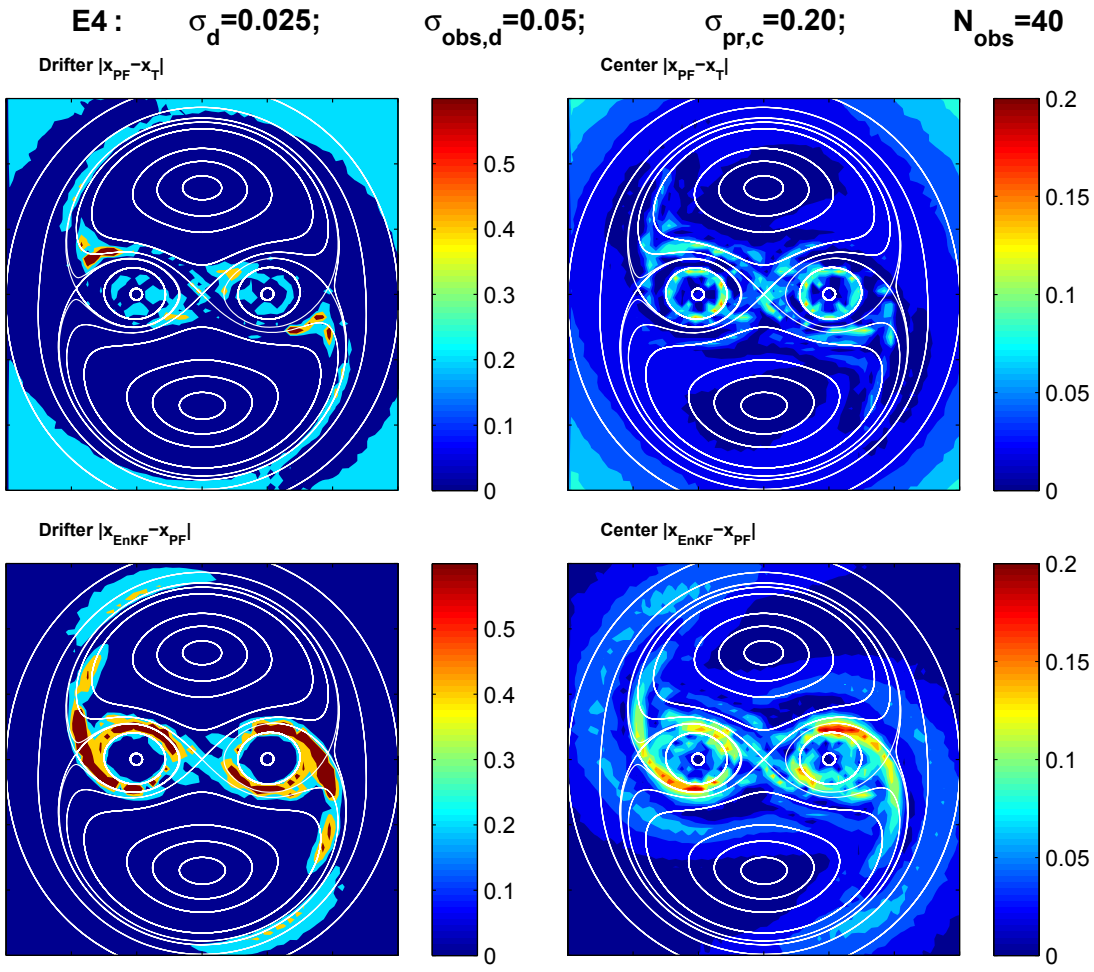

Figure 4.19: Performance metrics for experiment E4. Same format as Figure 4.15.

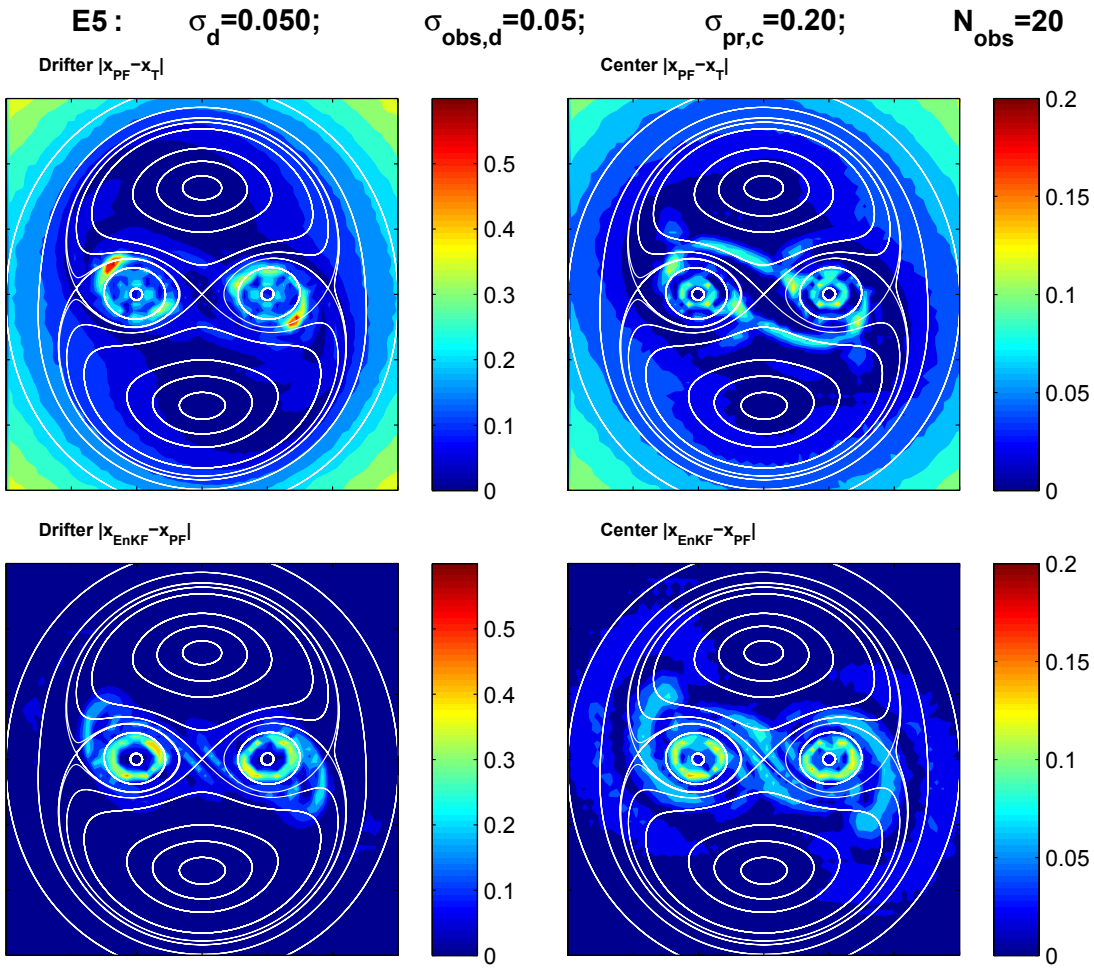

Figure 4.20: Performance metrics for experiment E5. Same format as Figure 4.15.

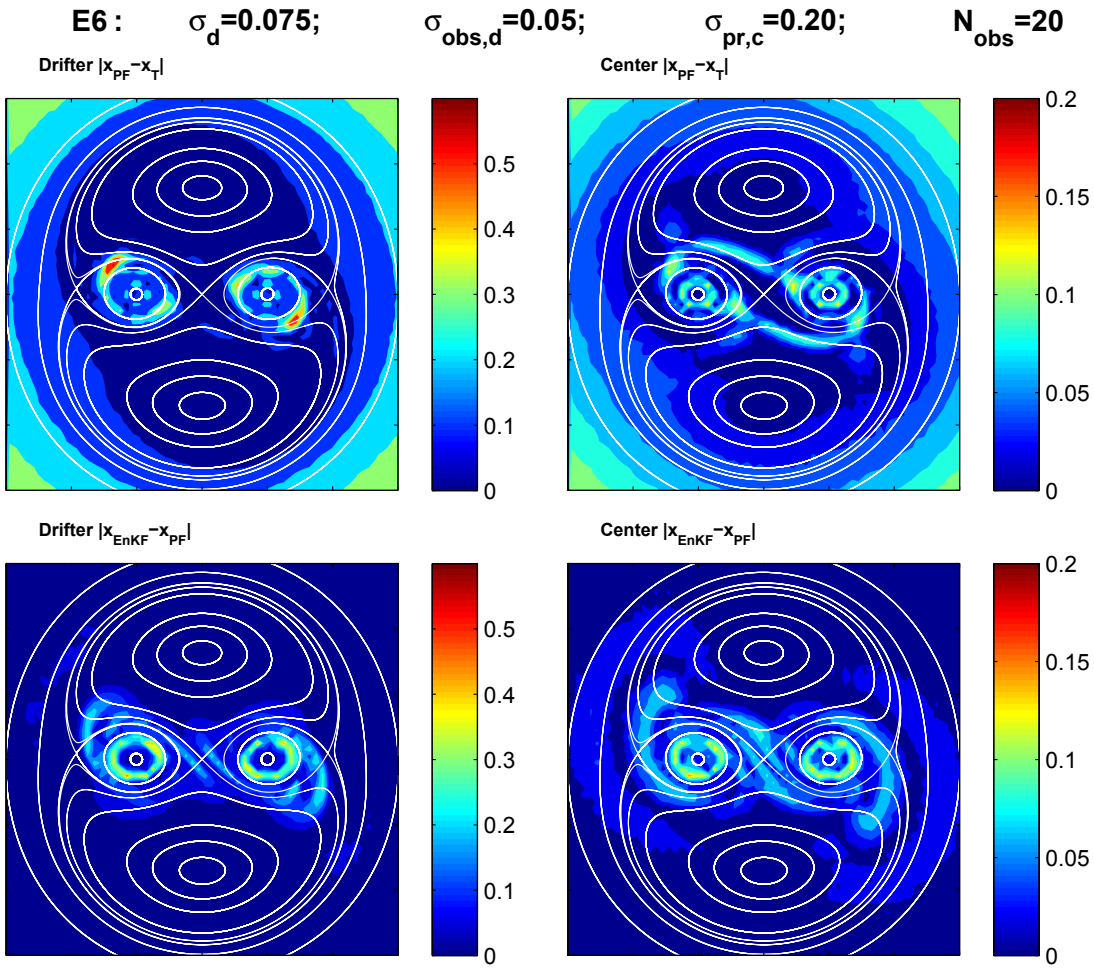

Figure 4.21: Performance metrics for experiment E6. Same format as Figure 4.15.

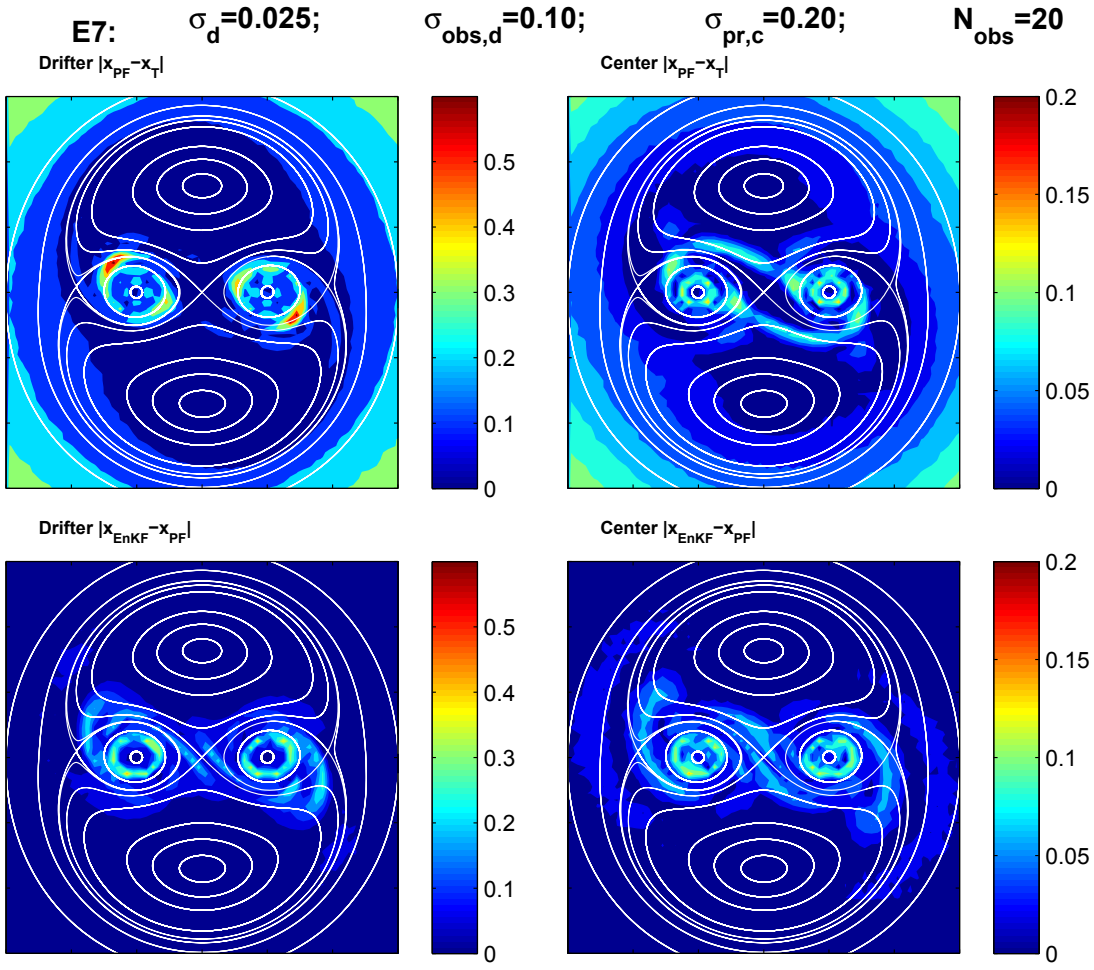

Figure 4.22: Performance metrics for experiment E7. Same format as Figure 4.15.

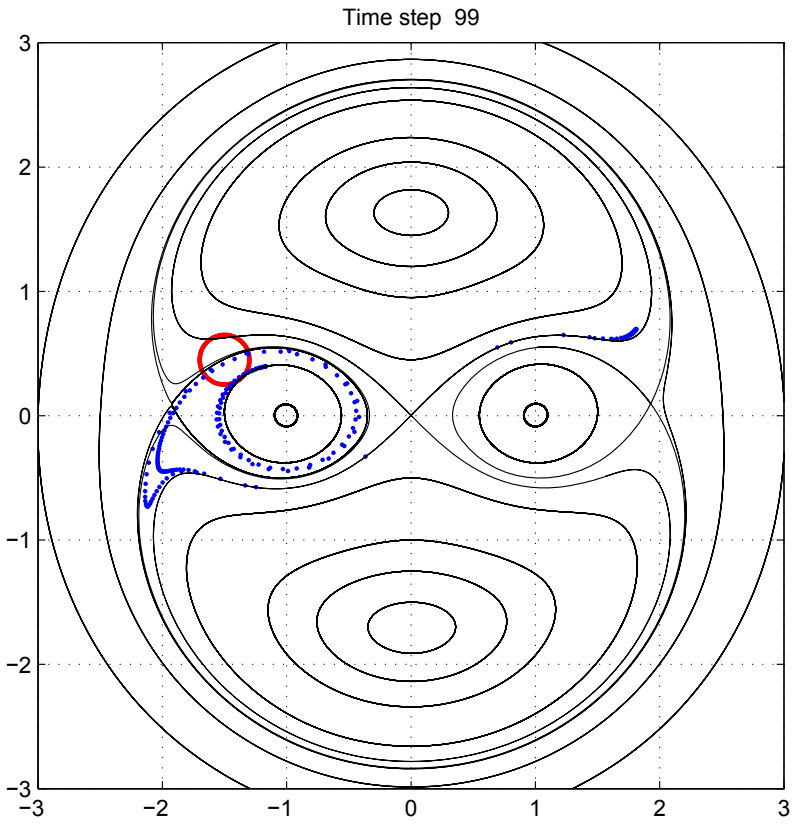

Figure 4.23: Final distribution (blue dots) of 200 drifters after 100 time steps. The figure shows that the drifters follow three distinct paths, clockwise around the Rankine vortex and along the outflowing manifolds of the left and center hyperbolic points, respectively. The drifters were released on a circle (red), centered on $(x, y) = (-1.5, 0.45)$, of radius=0.2 (corresponding to $\sigma_{pr,c} = 0.2$ as used in most experiments).

Chapter 5

Discussion

The overall objective of this study is to quantify how well the ocean mesoscale can be mapped in an Eulerian frame of reference by assimilating Lagrangian float trajectories. An important related objective is to compare the performance of two very different assimilation methods, the EnKF and PF, and assess their strengths and weaknesses using a highly idealized model of the ocean mesoscale.

Chapter 2 provides an overview of ocean data assimilation using a Bayesian framework and a consistent mathematical notation. Several assimilation schemes are introduced including the linear Kalman Filter, its non-linear extension (Extended Kalman Filter), a version based on ensemble methods (Ensemble Kalman Filter), and the Particle Filter. Following their description, the various assimilation methods are tested using a simple but relevant ocean model, i.e., the linear Stommel model. Although simple, the Stommel model supports some of the most important features found in more complex models of the large scale ocean circulation including westwards propagating Rossby waves, intensification of western boundary currents, and Sverdrup balance in the interior ocean. The model is subsequently used in a series of identical twin experiments in order to illustrate the various assimilation schemes. It is shown that the linear and Extended Kalman Filter perform very well as expected (not surprising given the linearity of the Stommel model). It is also shown the Ensemble Kalman Filter performs well. The Particle Filter also recovers the true state accurately except for the very easterly border region of the model domain; the addition of an extra MCMC step in the PF improves the assimilation results and the eastern boundary is better recovered. All experiments show that the quality of the estimated true state improves with the number of observations as expected.

In chapter 3 a new idealized model is introduced based on two co-rotating, self advecting vortices: the Self Advecting Vortex Model (SAVM). It is meant to represent, in the simplest way possible, the interaction of two ocean eddies and their impact on the trajectories of passively advected particles. It is based in part on the ideas of Aref (1984) who modeled "chaotic advection" using a system of point-vortices that could move or "blink" thus making the Eulerian flow unsteady and the trajectories of advected particles potentially chaotic. One attraction of Aref's approach is that there is no error in the integration of $dx/dt = u(x,t)$ because the displacements over each "blink" are specified analytically. This is one reason the SAVM is an improvement over the more complex, continuous-time model of Kuznetsov et al. (2003). Another reason the SAVM is more attractive than the Kuznetsov et al. model is that it is very cheap to run, making it ideal when carrying out a large number of sensitivity studies involving parameters such as model noise and drifter release points. To increase the relevance of the SAVM, the evolution of the saddle point are allowed to evolve stochastically. This random component to the model was shown to have a significant effect on the trajectories of the drifters. The position of each drifter, following advection by the vortices, was also been assumed to be changed by the addition of a random forcing which represents unknown and unresolved dynamical processes of the model.

The SAVM supports a rich array of flow fields. By tracing the trajectories of passively advected particles it is possible to show that the flow field has 4 elliptic points and 3 hyperbolic points. The hyperbolic points each have inflowing and outflowing manifolds. It is also possible to define separatrices that partition the flow domain into separate regions that do not allow exchange of advected particles. Ensembles of pseudo drifters released in the model's flow field were used to detect and evaluate situations where assimilation methods might break down. Two different types of experiment are performed, i.e. experiments with uncertain drifter release locations and experiments with uncertain model dynamics. The objective of these experiments was to determine the probability density functions (pdf) of the location of the drifters

and the (moving) center, and to determine when they take on a non-Gaussian shape. To quantify the onset of non-Gaussianity a new metric, Δ, was introduced. The new metric is both simple to calculate and interpret. The experiments gave two main results. The first result is that increasing uncertainty in the model's dynamics causes the pdfs of the drifter and center to become more non-Gaussian, as expected. The second, and more important, result is that the pdf of drifter position depends strongly on the release location and the growth of non-Gaussianity can be effectively summarized in a map of Δ.

In Chapter 4 the performance of the Ensemble Kalman Filter and Particle Filter were evaluated in a series of identical twin experiments based on the SAVM. The results of Chapter 3 were used to design the assimilation experiments i.e., where drifters are released in the SAVM, the parameter settings, the locations and timing of observations, and motion of the moving saddle point. The goal of the assimilation experiments was to evaluate the limits of the Ensemble Kalman Filter or, equivalently, when it is outperformed by the Particle Filter. Using the Δ metric it was shown that the Ensemble Kalman Filter typically performs poorly in regions of high Δ and drifter release locations located close to, or on, inflowing manifolds. The latter are regions where the probability distributions of drifter position typically become very non-Gaussian. A suite of performance metrics gave some insight into the best regions in which to deploy drifters in order to estimate the flow field of the SAV. For example it was shown that deploying the drifters in the ghost zones generally gave good estimates of the position of the center (and thus the positions of the self advecting vortices).

Sensitivity experiments were also performed to understand how the performance of the data assimilation schemes depend on the model parameters and the way the observations are collected. Overall the parameters that have the strongest impact on the performance of the data assimilation are (i) the release point of the drifter, (ii) the time between the drifter observations, and (iii) the prior pdf for the flow field. For the studies undertaken here, the observation error for drifter position, and also model error for drifter evolution did not have much impact on the performance of the

schemes.

The computationally cheaper Ensemble Kalman Filter only fails in limited regions of the SAVM flow field while performing at least as well as the Particle Filter elsewhere. However, the regions where the Ensemble Kalman Filter does fail are important because they are usually around inflowing manifolds. In order to form an efficient and effective assimilation scheme, and at the same time minimize unnecessary computational load, a new assimilation method is needed, possibly based on a modified or hybrid approach based on a combination of the Ensemble Kalman Filter and the Particle Filter, or the Ensemble Kalman Filter augmented by MCMC. The development of this type of hybrid in more realistic ocean models (e.g., a quasi-geostrophic model) should be considered for a future project or thesis.

Bibliography

A Apte, C. K. R. T Jones, and A. M. Stuart. A Bayesian approach to Lagragian data assimilation. *Tellus A, published online*, n/a(n/a):n/a, 2008.

H. Aref. Stirring by chaotic advection. *J. Fluid. Mech.*, 143:1–21, 1984.

R. G. Brown and P. Y. C. Hwang. *Introduction To Random Signals and Applied Kalman Filtering*. John Wiley and Sons, 1997.

G. Chen and Y. Chen. Some aspects of spatial aliasing in satellite altimetry. *Int. J. Remote Sensing*, 25(23):5525–5535, 2004.

Y. Chuang. Metropolis Hastings Markov Chain Monte Carlo. Talk, McMaster University, 2006.

M. Cooper and K. Haines. Altimetric assimilation with water property conservation. *Journal of Geophysical Research-Oceans*, 101(C1):1059–1077, 1996.

R. Daley. *Atmospheric Data Assimilation*. Cambridge University Press, 1992.

P. De Mey and M. Benkiran. A multivariate reduced-order optimal interpolation method and its application to the mediterranean basin-scale circulation. In N. Pinardi and J. Woods, editors, *Ocean Forecasting*, chapter 15, pages 281–306. Springer, 2002.

T. Delcroix, M. J. McPhaden, A. Dessier, and Y. Gouriou. Time and space scales for sea surface salinity in the tropical oceans. *Deep-Sea Research Part I-Oceanographic Research Papers*, 52(5):787–813, 2005.

A. Doucet. Frequently asked questions on particle filtering and our subjective answers. http://www-sigproc.eng.cam.ac.uk/smc/FAQ.html, 2006.

A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Pratice*, chapter 1, pages 3–14. Springer, 2001a.

A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Pratice*, pages 1–26. Springer, 2001b.

N. Ducet and P. Y. Le Traon. A comparison of surface eddy kinetic energy and Reynolds stresses in the gulf stream and the kuroshio current systems from merged topex/poseidon and ers-1/2 altimetric data. *Journal of Geophysical Research-Oceans*, 106(C8):16603–16622, 2001.

F. Durand, L. Gourdeau, T. Delcroix, and J. Verron. Assimilation of sea surface salinity in a tropical oceanic general circulation model (ogcm): A twin experiment approach. *Journal of Geophysical Research-Oceans*, 107(C12), 2002. 8004.

G. Evensen. Using the extended kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research-Oceans*, 97(C11):17905–17924, 1992.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forcast error statistics. *J. Geophys. Res.*, 99(C5): 10143–10162, 1994.

G. Evensen. The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation. *Ocean Dynamics*, 53:343–367, 2003.

G. Evensen and J. P. van Leeuwen. An Ensemble Kalman Smoother for Nonlinear Dynamics. *Mon. Wea. Rev.*, 128(6):1852–1867, 2000.

G. Evensen and P. J. van Leeuwen. Assimilation of Geosat altimeter data for the Agulhas Current using the Ensemble Kalman Filter with a quasi-geostrophic model. *Mon. Wea. Rev.*, 124:85–96, 1996.

S. J. Fan, L. Y. Oey, and P. Hamilton. Assimilation of drifter and satellite data in a model of the northeastern gulf of mexico. *Continental Shelf Research*, 24(9): 1001–1013, 2004.

Q. R. Gilks and C. Berzuini. Following amoving target - Monte Carlo Inference for dynamic Bayesian models. *J. R. Statist. Soc. B*, 63:127–146, 2001.

S. T. Gille. Aliasing of high-frequency variability by altimetry: Evaluation from bottom pressure recorders. *Geophysical Research Letters*, 28(9):1755–1758, 2001.

N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Proc.-F*, 140(2):107–113, 1993.

S. M. Grewal and P. A. Andrews. *Kalman Filtering*. John Wiley and Sons, 2001.

L. Guan and H. Kawamura. Merging satellite infrared and microwave ssts: Methodology and evaluation of the new sst. *Journal of Oceanography*, 60(5):905–912, 2004.

S. Guinehut, P. Y. Le Traon, G. Larnicol, and S. Philipps. Combining argo and remote-sensing data to estimate the ocean three-dimensional temperature fields - a first approach based on simulated observations. *Journal of Marine Systems*, 46 (1-4):85–98, 2004.

S. Hermann. *A particle filtering approach to joint passive radar tracking and target classification*. PhD thesis, University of Illinois, 2002.

I. Hoteit, Dinh-Tuan Pham, and J. Blum. Efficient Reduced Kalman Filtering and Application to Altimetric Data Assimilation in Tropical Pacific. Technical Report 3937, Institute National de Recherche en Informatique et en Automatique, 655, avenue de'l Europe, 38330 MONTBONNOT ST MARTIN (France), May 2000.

K. Ide, L Kuznetsov, and C. K. R. T. Jones. Lagrangian data assimilation for point vortex systems. *J. Turbulence (on-line journal)*, 3(053):1–7, 2002.

R.E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *J. of Basic Engineering*, 82(D):35–45, 1960.

E. Kalnay. *Atmospheric Modelling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

C. L. Keppenne. Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter. *Mon. Wea. Rev.*, 128(6):1971–1981, 2000.

C. L. Keppenne and M. M. Rienecker. Initial Testing of a Parallel Ensemble Kalman Filter for the Poseidon Isopycnal Ocean General Model. *Mon. Wea. Rev.*, 130(12): 2915–2965, 2002.

C. L. Keppenne and M. M. Rienecker. Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble Kalman filter. *JMS.*, 40-41:363–380, 2003.

C. L. Keppenne, M. M. Rienecker, N. P. Kurkowski, and D.A. Adamec. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlinear Processes in Geophysics*, 12: 491–503, 2005.

A. Kim, G. L. Eyink, J. M. Restrepo, F.J Alexander, and G. Johnson. Ensemble Filtering for Nonlinear Dynamics. *Mon. Wea. Rev.*, 131(11):2586–2594, 2003.

G. Kitagawa. Monte Carlo and Smoother for Non-Gaussiam Nonlinear State Space Models. *Journal of Computational and Graphical Statistics*, 5:1–25, 1996.

G. Kitagawa. A Self-Organizing State-Space Model. *Journal of the American Statical Association*, 93(443):1203–1215, 1998.

G. A. Kivman. Sequential parameter estimation for stochastic systems. *Nonlinear Processes in Geophysics*, 10:253259, 2003.

A. Kong, J. S. Liu, and W. H. Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, 1994.

L. Kuznetsov, C. R. K. T. Jones, and K. Ide. A method for the assimilation of Lagrangian data. *Mon. Wea. Rev.*, 131(10):22472260, 2003.

P.-Y. Le Traon. Satellite oceanography for ocean forecasting. In N. Pinardi and J. Woods, editors, *Ocean Forecasting*, chapter 2, pages 19–36. Springer, 2002.

P. Y. Le Traon and G. Dibarboure. Velocity mapping capabilities of present and future altimeter missions: The role of high-frequency signals. *Journal of Atmospheric and Oceanic Technology*, 19(12):2077–2087, 2002.

P. Y. Le Traon, F. Nadal, and N. Ducet. An improved mapping method of multi-satellite altimeter data. *Journal of Atmospheric and Oceanic Technology*, 15(2): 522–534, 1998.

P. Y. Le Traon, Y. Faugere, F. Hernandez, J. Dorandeu, F. Mertz, and M. Ablain. Can we merge geosat follow-on with topex/poseidon and ers-2 for an improved description of the ocean circulation? *Journal of Atmospheric and Oceanic Technology*, 20(6):889–895, 2003.

E. N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963a.

E. N. Lorenz. Predictability of hydrodynamic flow. *Transactions of the New York Academy of Sciences*, 25(4):409–432, 1963b.

E. N. Lorenz. Atmospheric predictability experiments with a large numerical-model. *Tellus*, 34(6):505–513, 1982.

D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2005.

H. Madsen and R. Canizares. Comparison Of Extended and Ensemble Kalman Filters For Data Assimilation In Coastal Area Modelling. *International Journal For Numerical Methods In Fluids*, 31:961–981, 1999.

A. M. Mancho, D. Small, and S. Wiggins. A tutorial on dynamical systems concepts applied toLagrangian transport in oceanic flows defined as finite time data sets:Theoretical and computational issues. *Physics Reports*, 437:55–124, 2006.

A. Molcard, L. I. Piterbarg, A. Griffa, T. M. Özgökmen, and A. J. Mariano. Assimilation of drifter observations for the reconstruction of the eulerian circulation field. *Journal of Geophysical Research-Oceans*, 108(C3), 2003. 3056.

A. Molcard, A. Griffa, and T. M. Özgökmen. Lagrangian Data Assimilation in Multilayer Primitive Equation Ocean Models. *Journal of Atmospheric and Oceanic Technology*, 22(1):70–83, 2005.

A. Molcard, A. C. Poje, and T. M. Özgökmen. Directed drifter launch strategies for Lagrangian data assimilation using hyperbolic trajectories. *Ocean Modelling*, 12: 268–289, 2006.

T. M. Özgökmen, A. Griffa, A. J. Mariano, and L. I. Piterbarg. On the predictability of lagrangian trajectories in the ocean. *Journal of Atmospheric and Oceanic Technology*, 17(3):366–383, 2000.

A. Pascual, Y. Faugere, G. Larnicol, and P. Y. Le Traon. Improved description of the ocean mesoscale variability by combining four satellite altimeters. *Geophysical Research Letters*, 33(2), 2006. L02611.

B.S. Powell, A.M. Moore, H.G. Arango, E. Di Lorenzo, R.F. R.F. Milliff, and R.R. Leben. Near real-time ocean circulation assimilation and prediction in the intra-americas sea with roms. *Dynamics of Atmosphere and Oceans*, 48(1-3):46–68, 2009.

B. Ristic, A. Arulampalam, and Gordon N. *Beyond the Kalman Filter*. Artech House, 2004.

C. P. Robert and G. Casella. *Monte Carlo Methods*. Springer, 1999.

H. Salman, L. Kuznetsov, and C. K. R. T. Jones. A Method for Assimilating Lagrangian Data into a Shallow-Water-Equation Ocean Model. *Mon. Wea. Rev.*, 134 (4):1081–1101, 2006.

A. J. Sergers. *Data assimilation in atmospheric chemistry models using Kalman filtering*. Delft University Press, 2002.

K. R. Thompson and E. Demirov. Skewness of sea level variability of the world's oceans. *Journal of Geophysical Research-Oceans*, 111(C5), 2006. C05005.

J. P. van Leeuwen. Particle filtering in geophysical systems. *Mon. Wea. Rev.*, 137: 4089–4114, 2009.

J. P. van Leeuwen and G. Evensen. Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation. *Mon. Wea. Rev.*, 124:2898–2913, 1996.

P. J. van Leeuwen. A Variance-Minimizing Filter for Large-Scale Applications. *AMS*, 131:2071–2084, September 2003.

G. Welch and G. Bishop. *SIGGRAPH 2001*, chapter 8. University of North Carolina at Chapel Hill, 2001.

S. Wiggins. The Dynamical Systems Approach to Lagrangian Transport in Oceanic Flows. *Annu. Rev. Fluid Mech*, 37:295–328, 2005.

P. Zarchan and H. Musoff. *Fundamentals of Kalman Filtering*. American Institute of Aeronautics and Astronautics, 2002.

Appendix A

The Linear Stommel Model

One of the simplest ocean models that includes a western boundary current is the linear Stommel model. The time dependent form of the Stommel model is based on the following equations:

$$u_t - fv = -g\eta_x + \tau^x - \lambda u \tag{A.1}$$

$$v_t + fu = -g\eta_y + \tau^y - \lambda v \tag{A.2}$$

$$\eta_t + H(u_x + v_y) = 0 \tag{A.3}$$

where (u, v) denote the velocities in zonal and meridional directions respectively, η is the sea surface elevation, f is Coriolis parameter, (τ^x, τ^y) are the wind stresses in zonal and meridional directions respectively, and λ is the friction parameter. Subscripts denote derivatives.

Taking the curl of (A.1) and (A.2), and assuming a rigid lid, leads to the following vorticity equation:

$$\frac{\partial}{\partial t}(v_x - u_y) + f(u_x + v_y) + \beta v = \Phi - \lambda(v_x - u_y) \tag{A.4}$$

where $\beta = f_y$ and Φ is the curl of the wind stress. Introducing a stream function ψ such that $v = \psi_x$ and $u = -\psi_y$, it follows

$$v_x - u_y = \nabla^2\psi. \tag{A.5}$$

Assuming variations in the meridional direction are periodic with wave number l, the vorticity equation can be written

$$\begin{array}{ccccccc} \frac{\partial}{\partial t}(\psi_{xx} - l^2\psi) & + & \beta\psi_x & = & \Phi & - & \lambda(\psi_{xx} - l^2\psi) \\ I & & II & & III & & IV \end{array} \tag{A.6}$$

The following balances can be interpreted:

- The balance between II and III holds for the interior of the ocean in steady state and is the so called Sverdrup balance.
- The balance between I and II is the unforced, non dissipative form of (A.6) and describes propagating barotropic Rossby waves with the following dispersion relation
$$\omega = \frac{-\beta k}{k^2 + l^2} \tag{A.7}$$
where k is the zonal wave number.
- For a steady state (or slowly varying) solution the balance of II and IV (creation and dissipation of vorticity) holds for the western boundary current region where ψ_{xx} varies rapidly. The width L_s of the western boundary current can be derived from this steady state balance and is given by $L_s = \lambda/\beta$. This is called the Stommel width.

Before discretizing, (A.6) is non-dimensionalized. This is done by scaling such that $x = Lx^*$ where L is the ocean width and $t = (1/\beta L)t^*$. The non-dimensionalized form of (A.6) becomes

$$\frac{\partial}{\partial t^*}(\psi_{x^*x^*} - (lL)^2\psi) + \psi_{x^*} = \frac{L}{\beta}\Phi - \frac{\lambda}{\beta L}(\psi_{x^*x^*} - (lL)^2\psi) \tag{A.8}$$

A.1 Discretization Of The Vorticity Equation

Let ψ_t denote the spatially discretized form of $\psi(t, x)$:

$$\psi_t = (\psi(\Delta x, t)\ \psi(2\Delta x, t)\ ...\ \psi(N\Delta x, t)) \tag{A.9}$$

where N is the number of interior grid points. First and second derivatives of ψ with respect to x can be approximated by the forms $M_x\psi_t$ and $M_{xx}\psi_t$ where M_x and M_{xx} are tridiagonal matrices depending on the grid spacing. Substituting the discretized derivatives of $\psi(t, x)$ into (A.8) gives

$$\frac{\partial}{\partial t}[M_{xx} - l^2]\psi_t + M_x\psi_t + \lambda[M_{xx} - l^2]\psi_t = \Phi \tag{A.10}$$

with the following changes in notation: $\frac{\partial}{\partial t^*} \to \frac{\partial}{\partial t}$, $\frac{\lambda}{\beta L} \to \lambda$, $(lL)^2 \to l^2$ and $\frac{L}{\beta}\Phi \to \Phi$.

If the time derivative in (A.10) is approximated by the standard finite difference form we arrive at the following updating equation for the $N \times 1$ state vector ψ_t:

$$\psi_{t+1} = D\psi_t + q_{t+1} \tag{A.11}$$

where $D = M_1^{-1}M_2$ with

$$M_1 = \frac{1}{\Delta t}[M_{xx} - l^2] + (1 - \theta)(M_x + \lambda[M_{xx} - l^2]) \tag{A.12}$$

and

$$M_2 = \frac{1}{\Delta t}[M_{xx} - l^2] - \theta(M_x + \lambda[M_{xx} - l^2]). \tag{A.13}$$

and Δt is the scaled model time step. The parameter θ controls the extent to which the scheme is explicit ($\theta = 0$) or implicit ($\theta = 1$). The vector q_t represents "forcing".

Appendix B

Pseudo codes for the SIS, generic Particle Filter, SIR Filter and Resample Routine

All listings are taken from Ristic et al. [2004].

B.1 Filtering via SIS

$[\{x_k^i, w_k^i\}_{i=1}^N] = SIS[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^N, z_k]$

- FOR $i = 1 : N$
  - Draw $x_k^i \sim q(x_k|x_{k-1}^i, z_k)$
  - Evaluate the importance weights up to a normalizing constant

    $\tilde{w}_k^i =\propto w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)}$
- END FOR
- Calculate total weight: $t = SUM[\{\tilde{w}_k^i\}_{i=1}^N]$
- FOR $i = 1 : N$
  - Normalize: $w_k^i = t^{-1}\tilde{w}_k^i$
- END FOR

B.2 Generic Particle Filter

$[\{x_k^i, w_k^i\}_{i=1}^N] = PF[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^N, z_k]$

- Filtering via SIS (B.1):

  $[\{x_k^i, w_k^i\}_{i=1}^N] = SIS[\{x_{k-1}^i, w_{k-1}^i\}_{i=1}^N, z_k]$

- Calculate N_{eff} using equation (2.57)

- IF $N_{eff} < N_{thr}$

  - Resample using algorithm (B.4):

    $[\{x_k^i, w_k^i\}_{i=1}^N] = RESAMPLE[\{x_k^i, w_k^i\}_{i=1}^N$

- END IF

B.3 Filtering via SIR

$[\{x_k^i\}_{i=1}^N] = SIR[\{x_{k-1}^i\}_{i=1}^N, z_k]$

- FOR $i = 1 : N$

  - Draw $x_k^i \sim p(x_k|x_{k-1}^i)$
  - Calculate $\tilde{w}_k^i = p(z_k|x_k^i)$

- Calculate total weight: $t = SUM[\{\tilde{w}_k^i\}_{i=1}^N]$

- FOR $i = 1 : N$

  - Normalize: $w_k^i = t^{-1}\tilde{w}_k^i$

- END FOR

- Resample using algorithm (B.4):

  $[\{x_k^i\}_{i=1}^N] = RESAMPLE[\{x_k^i, w_k^i\}_{i=1}^N$

B.4 Resample Algorithm

$[\{x_k^{j*}, w_k^j\}_{i=1}^N] = SIR[\{x_k^i, w_k^i\}_{i=1}^N]$

- Initialize the cumulative sum of weights (CSW): $c_1 = w_k^1$
- FOR i$i = 2 : N$
  - Construct CSW: $c_i = c_{i-1} + w_k^i$
- END FOR
- Start at the bottom of the CSW: $i = 1$
- Draw a starting point: $u_1 \sim \mathcal{U}[0, N^{-1}]$
- FOR $j = 1 : N$
  - Move along the CSW: $u_j = u_1 + N^{-1}(j-1)$
  - WHILE $u_j > c_i$
    - $i = i + 1$
  - END WHILE
  - Assign sample: $x_k^{j*} = x_k^i$
  - Assign weight: $w_k^j = N^{-1}$
- END FOR

Appendix C

The Markov Chain Monte Carlo Move Step

The appropriate choice of the proposal density $q(x)$ is critical for the success of sequential importance sampling (SIS). SIS can only work well if the proposal density $q(x)$ is similar to the true density $\pi(x)$. For complex systems it is in general difficult to define a proposal density $q(x)$ that satisfies this property.

Since the SIR filter is directly derived from the SIS the SIR filter inherits the problem for the need of an appropriate proposal density $q(x)$. In order to overcome this difficulty a method has been developed unknown as the Markov Chain Monte Carlo (MCMC) step. There exist several derivations how to perform this method and one of the most known is the Metropolis-Hastings method.

The main idea of the Metropolis-Hastings method is to simulate the given n-dimensional distribution $\pi(x_k^i)$ that has density $\pi(x)$ with respect to some dominating measure. $x_k^i \in \Re^n$ and $i = 1 \dots N$ where N is the number of ensemble members. The method constructs for a given pdf $\pi(x_k^i)$ a reversible Markov Chain whose stationary distribution is the distribution $\pi(x_k^i)$. To construct the Markov Chain the Metropolis-Hastings method makes use of the proposal density $q(x)$ which depends on the current state, denoted by $q(\cdot|x_k)$, e.g. $q(x^{'}|x_k)$ where x_k denotes the current state and $x^{'}$ denotes an approximation of the current state. Most importantly is that the proposal density $q(x^{'}|x_k)$ can be *any* fixed density from which one can draw samples, i.e. $q(x^{'}|x_k)$ does not have to be similar with the true density $\pi(x)$. Of course there are proposal densities that work better than others for a given type of model.

The description of the Markov Chain Monte Carlo (MCMC) move step follows the theoretical foundations given in (Gilks and Berzuini [2001]). The main idea behind the MCMC move step is that a resampled ensemble member x_k^i is moved to a new

state x_k^{i*}

$$x_k^{i*} = x_k^i + h_{opt} D_k, \tag{C.1}$$

only if $u_i \leq \alpha$, where $u_i \sim \mathcal{U}[0,1]$ and α is the acceptance probability, where $\mathcal{U}$ denotes a uniform distribution. Otherwise the replacement is rejected and x_k^i stays unchanged.

D has the form of a $N \times n_x$ and introduces a form of jittering to the ensemble members where N is the ensemble size and n_x the dimension of the ensemble member (state). The N random vectors of D are picked from a multivariate normal distribution with a mean of zero and a covariance Q of the model error (see chapter 2, 2.12).

The scalar h_{opt} can be viewed as a bandwidth parameter for the jitter. This parameter is actively adapted at every MCMC move step. The more members x_k^i are being replaced by x_k^{i*} by the MCMC move step the larger h_{opt} becomes and vice versa. A large h_{opt} indicates that the distance between the resampled ensemble members x_k^i and the and the corresponding members of the proposal distribution (C.1) is large.

The acceptance parameter α can be derived from the following idea:

Let $X_k = \{x_j, j = 0, \ldots, k\}$, which represents the sequence of all model states up to time k.
The desired density of the MCMC step is $p(X_k|Z_k)$ which can be expressed (see equation (2.53)) as

$$p(X_k|Z_k) = \frac{p(z_k|x_k)p(x_k|x_{k-1})}{p(z_k|Z_{k-1})} p(X_{k-1}|Z_{k-1}) \tag{C.2}$$

The main goal of the MCMC move step is to maximize the density $p(z_k|x_k)p(x_k|x_{k-1})$ in equation (C.2) which is equivalent to maximizing $p(z_k|x_k^{i*})p(x_k^{i*}|x_{k-1}^i)$ in the MCMC.

Assuming

- member x_k^i has been created by the resampling step such that

  $X_k^i = \{x_k^i, X_{k-1}^i\}$

- member x_k^{i*} has been generated by sampling from the proposal distribution $q(\cdot|x_k^i)$, such that

  $X_k^{i*} = \{x_k^{i*}, X_{k-1}^{i*}\}$

Then the Metropolis-Hastings acceptance probability (see Robert and Casella [1999]) is given by

$$\alpha = min\left\{1, \frac{p(X_k^{i*}|Z_k)q(x_k^i|x_k^{i*})}{p(X_k^i|Z_k)q(x_k^{i*}|x_k^i)}\right\}. \tag{C.3}$$

A complete derivation of the acceptance probability can be found in Chuang [2006].

Assuming the proposal density $q(\cdot|x_k^i)$ is symmetric, then $q(x_k^{i*}|x_k^i) = q(x_k^i|x_k^{i*})$ and substitution of (C.2) into (C.3) gives the acceptance probability α

$$\alpha = min\left\{1, \frac{p(z_k|x_k^{i*})p(x_k^{i*}|x_{k-1}^i)}{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}\right\} \tag{C.4}$$

The schematics of a PF with SIR and MCMC move step are shown in figure C.1.

All particle filter based data assimilation experiments and sensitivity studies in this thesis are based on a form of a particle filter configuration based on a combination of the SIR filter and the MCMC step. The resampling algorithm we used in the SIR and MCMC is Kitagawa's form of "deterministic resampling" (Kitagawa [1996],Kitagawa [1998]) rather than the standard form as described above for the SIS procedure.

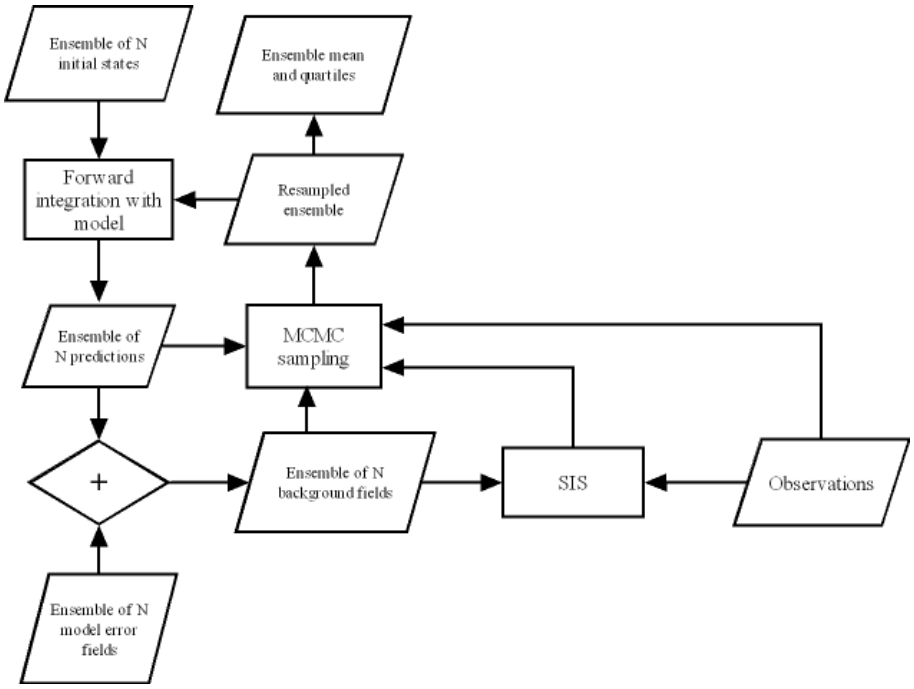

Figure C.1: Schematic of a data assimilation procedure using the MCMC move step.