THE UNIVERSITY OF READING
DEPARTMENT OF MATHEMATICS

Data Assimilation for Numerical Weather Prediction Using Control Theory

by

Anne K. Griffith

Thesis submitted for the degree of
Doctor of Philosophy

April 1997

Abstract

Data assimilation is a means of estimating an atmospheric or oceanic state by combining observational data with a prior estimate of the state, usually from a numerical model. We look at application of data assimilation to numerical weather prediction using control theory.

Firstly, we apply observer theory to successive correction methods of data assimilation to show when they converge in time to the true solution. However, we mostly focus on 4D variational data assimilation schemes. Here the approach is to minimize a cost function penalizing distance from observational data over a time interval, subject to the constraint that the model equations are satisfied. The minimization problem can be solved by iterating on the model initial state, which is referred to as "using the initial state as the control vector".

Our aim is to provide a consistent theoretical foundation which allows for model error in variational assimilation. We investigate the "correction term technique" in which a constant correction term approximating model error is added to the model equations and used as a control vector instead of, or as well as, the initial state. We use the concept of complete N-step observability to give conditions for a unique solution of the minimization problem using different control vectors.

We suggest a generalization of the correction term technique in which we use state augmentation to estimate a serially correlated component of model error along with the model state. In particular, we consider using a correction term representing model error that evolves as the model state evolves. We investigate the effectiveness of the constant and the evolving correction term in compensating for different types of model error using simple linear models. We also use the correction term technique for a 1D nonlinear shallow water model in the presence of different types of model error, and find that a constant correction term can compensate for non-constant model error on a significant timescale.

Acknowledgements

First of all, I thank my academic supervisor, Prof. Nancy Nichols, for what I have learnt from her about mathematics and about doing research, and I thank her for further inspiring my interest in the subject. I also thank my industrial supervisor, Mr Andrew Lorenc of the U.K. Meteorological Office, for a chance to learn about data assimilation at the Met. Office, and for valuable discussions about this work.

I would like to thank my family and my friends who encouraged me in my work. I especially thank my "chocaholic" office mates, Alison, Chris and Lee, who helped to make my time as a PhD student particularly enjoyable.

Finally, I acknowledge financial support for my studies from the Engineering and Physical Sciences Research Council and from the U.K. Meteorological Office.

Contents

1 Introduction **1**

1.1 Background on data assimilation 1

1.2 Overview of the thesis 7

2 Mathematical Background **10**

2.1 Introducing the System 10

2.1.1 The nonlinear model 11

2.1.2 The true model state and model error 11

2.1.3 Observational data 12

2.1.4 The linear assimilation system 13

2.1.5 State transition matrix 14

2.2 Controllability and Observability 14

2.2.1 Some definitions 15

2.2.2 Theory for the general linear case 16

2.2.3 Theory for the linear, time-invariant case 18

2.3 Nonlinear optimization theory 20

2.3.1 Preliminaries 20

2.3.2 Unconstrained minimiz 14

2.5 Background on probability theory 29
2.5.1 Definitions

4.2.2 Adjoint model development 76
4.3 The correction term technique 77
4.4 The weak constraint approach 79
4.4.1 The general least squares problem 79
4.4.2 Methods for solving Problem $\mathcal{L}$ 81

5 The correction term technique **85**
5.1 Background 86
5.1.1 Use of the technique in the literature 86
5.1.2 Research issues 88

5Tc1?86–9

6.2.1 A general formulation of model error 140
6.2.2 Examples of how model error can be specified 142
6.2.3 Problem $\mathcal{L}$ for the augmented system 145
6.3 A generalized correction term technique 147
6.4 Using an evolving correction term 150
6.4.1 Introduction 150
6.4.2 Description of the experiments 151
6.5 Results 154
6.5.1 Case a): Imperfect model, known initial state 154
6.5.2 Case b): Imperfect model, unknown initial state 157
6.6 Summary and conclusions 165

7 Experiments with a shallow water model 1 7
7.1 The shallow water model 168
7.2 The data assimilation problem 171
7.2.1 The adjoint model 171
7.2.2 The gradients of $\mathcal{L}$ with respect to the control vectors 173
7.3 Description of the experiments 174
7.3.1 The true model state 174
7.3.2 Observations 174
7.3.3 Model error 175
7.3.4 The descent algorithm 176
7.3.5 The experiments 177
7.4 Results from the experiments 179
7.4.1 Experiment 1: Comparing different control vectors 179
7.4.2 Experiment 2: Fewer observations and observational error . . 184
7.4.3 Experiment 3: The impact of assimilation on a forecast 185
7.5 Summary and Conclusions 210

8 Conclusions 214

Chapter 1

Introduction

We start with an introduction to data assimilation, particularly focusing on its application to numerical weather prediction (NWP). This is followed by an overview

in forecast verification, for archive records and for climate studies. In oceanography, on the other hand, the main use of data assimilation is in studies to increase understanding of ocean circulations, although it is also used in short range ocean forecasting [32].

A wide variety of data assimilation schemes have been proposed and developed over the last 50 years or so, and many of them are taken from state estimation techniques in engineering. We now in

tions over the assimilation interval, and distance from a prior estimate, subject to the constraint that the solution (the analysis) is consistent with the model dynamics. In the *strong constraint* approach to the 4D variational assimilation problem, the constraint is that the solution must satisfy the model equations exactly. In the *weak constraint* approach, the solution is only required to satisfy the model equations approximately, and hence some allowance is made for *model error*.

Meteorological observational data available for assimilation

We now give a brief description of the types of observational data available for use in operational meteorological assimilation to produce analyses for weather forecasting, a fuller description is given in the book by Daley [24].

Meteorological observations are available on a world-wide scale, and there is international cooperation on the data collection and distribution to the various national meteorological centres. A

The quality control of observ

crepancy between noisy data and a prior estimate of the state can produce spurious inertia-gravity waves [32]. Although primitive equation models do in fact exhibit gravity waves which describe a small amount of the flow, atmospheric and oceanic flows at mid-latitudes on the timescale of a forecast are well described by the relatively slow Rossby waves.

Early NWP models were often quasi-geostrophic, and hence avoided the need for initialization, since these models produce only Rossby waves. When primitive equation models became operational for forecasting in the early 1970s, initialization became necessary. Initialization has generally been carried out separately from the data assimilation procedure, by projection of the solution onto the subspace described by the Rossby modes [32], or by the process of nonlinear normal mode initialization introduced by Machenhauer [59]. In "advanced" data assimilation techniques, however, it is possible to incorporate the initialization process in the assimilation. This may be done in Kalman filtering applications by projection of

on, but only as computer power increased was this approach further developed and used extensively in an operational context. The method of *optimal interpolation* (OI) suggested by Gandin in 1963 [30], attempts to provide a statistically optimal estimate of a linear system at a given time. Variants of this method, which are also applicable to nonlinear systems, have been applied widely for operational data assimilation in the 1980s and 1990s. The *three-dimensional variational assimilation* (or 3DVAR) method [73] can be seen as a different approach to solving the same problem as OI, and is currently being developed for operational use at several NWP centres.

In the earlier days of data assimilation, observations were available mainly at the synoptic and sub-synoptic times. Since observations from satellites have become available, however, some observations are available continuously and it has become more important that data assimilation techniques should draw upon the time-tendency information available in the observations. For this reason, there is much interest at present in the design and development of 4D data assimilation methods. Two examples of such methods include the Kalman filter, and 4D variational assimilation.

The *Kalman filter*, proposed by Kalman in 1960 [45] for engineering applications can be used as a sequential 4D assimilation method. For a linear model and under certain assumptions, it provides a statistically optimal solution at a given time taking into account all previous observations. The method in unsimplified form is generally considered too expensive for use with large operational models in meteorology and oceanography [32], but various simplifications ha

Chapter 2

Mathematical Background

Throughout this thesis, we will be looking at data assimilation for meteorology and oceanography using a framework of mathematical control theory. In the first section we introduce the general model system, using control theory notation. Then, in Section 2.2, we give background on some of the basic concepts of control theory which will be useful, and state definitions and theorems which will be referred to later on. In Section 2.3 we give background on nonlinear optimization theory, and in Section 2.4 describe descent algorithms that may be used to iterate to an optimal solution. Sections 2.3 and 2.4 provide the background for the *variational* data assimilation methods. Finally, Section 2.5 gives background on probability theory and the concept of a "most likely" estimate, which is widely used in data assimilation.

In this background chapter, we limit our discussion to *discrete* systems, since this is most convenient for application to numerical models of meteorology and oceanography. Many texts on control theory concentrate on continuous systems, with only brief reference to the discrete case. However, we treat the discrete case since the transition from the continuous to the discrete case is not always immediate [86].

2.1 Introducing the System

To start with, we introduce the general nonlinear model system which we will use throughout the thesis, and explain what we mean by the true model state and model

where $\boldsymbol{\psi}_k : \mathcal{U} \to \mathcal{U}$ is a uniquely defined nonlinear solution operator [17]. We now define the *true model state* $\mathbf{x}$

$\mathbb{R}^{p_k}$ represents the *observational error* at time t_k. In the context of control theory, the observations are generally referred to as model *outputs*. If observation times do not coincide with the timesteps t_k, then $\mathbf{h}_k$ will include temporal interpolation. The number of observations p_k varies with time, and this includes the possibility of no observations at some timesteps.

Observational error has two components usually referred to as *measurement errors* and *representational errors*. The measurement errors are due to errors in the measurement instruments and in the transmission of information, and the representation errors are due to errors in $\mathbf{h}_k$. More detail on the form of observational error is given in [24].

2.1.4 The linear assimilation system

In some cases we will limit our attention to linear theory, and so we consider the discrete, linear, time-varying model

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k, \tag{2.7}$$

with $\mathbf{x}_k$ and $\mathbf{u}_k$ defined as in (2.1), and with $A_k \in \mathbb{R}^{n \times n}$, $B_k \in \mathbb{R}^{n \times m}$. We assume that A_k is nonsingular and that B_k has rank m for all k, so that specification of $\mathbf{x}_0$ and the $\mathbf{u}_j$, $j = 0, .., k-1$ uniquely determines $\mathbf{x}_k$ for $k > 0$. We suppose that the evolution of the true model state $\mathbf{x}_k^t$ satisfies

$$\mathbf{x}_{k+1}^t = A_k \mathbf{x}_k^t + B_k \mathbf{u}_k + \boldsymbol{\varepsilon}_k, \qquad k = 0, .., N-1, \tag{2.8}$$

where $\boldsymbol{\varepsilon}_k \in \mathbb{R}^n$ is the model error as defined in Subsection 2.1.2.

We now suppose that the observations are related linearly to the true model state as follows,

$$\mathbf{y}_k = C_k \mathbf{x}_k^t + \boldsymbol{\delta}_k, \qquad k = 0, .., N-1, \tag{2.9}$$

with $\mathbf{y}_k$ and $\boldsymbol{\delta}_k$ defined as in (2.6) and $C_k \in \mathbb{R}^{p_k \times n}$.

If the assimilation system (2.8),(2.9) is a linearization of the system (2.4),(2.6) about some reference state $\mathbf{x}_k^o$ and input $\mathbf{u}_k^o$, then A_k and C_k are the Jacobians of $\mathbf{f}_k$ and $\mathbf{h}_k$ respectively with respect to $\mathbf{x}_k$, and B_k is the Jacobian of $\mathbf{f}_k$ with respect to $\mathbf{u}_k$, all evaluated at $(\mathbf{x}_k^o, \mathbf{u}_k^o)$. In this case, the model (2.8) is often referred to as the *tangent linear model* of (2.4) in data assimilation literature [48], [20].

2.1.5 State transition matrix

For some applications of the linear system, it will be useful to relate the state at a given time to the state at any earlier time. We therefore introduce the *state transition matrix* $\Phi(k, j)$, for the unforced system

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k, \tag{2.10}$$

which relates the state at time t_k to the state at an earlier time t_j as follows, [2],

$$\mathbf{x}_k = \Phi(k, j)\mathbf{x}_j \qquad \forall k \geq j, \tag{2.11}$$

with

$$\Phi(j, j) = I \qquad \forall j. \tag{2.12}$$

For the system (2.10) the state transition matrix is given uniquely by

$$\Phi(k, j) = \prod_{i=j}^{k-1} A_i. \tag{2.13}$$

Clearly we have

$$\Phi(l, j) = \Phi(l, k)\Phi(k, j) \qquad \forall l \geq k \geq j, \tag{2.14}$$

and since the matrices A_i are assumed to be nonsingular, we also may define

$$\Phi(j, k) = \Phi^{-1}(k, j), \qquad \forall j \leq k. \tag{2.15}$$

For the forced model (2.7), we now have [2]

$$\mathbf{x}_k = \Phi(k, j)\mathbf{x}_j + \sum_{i=j}^{k-1} \Phi(k, i+1)B_i \mathbf{u}_i \qquad \forall k > j. \tag{2.16}$$

The relationship (2.16) will be important later on in the thesis.

2.2 Controll bility nd Observ bility

The general aim of control theory is to regulate the state to some desired state by a suitable choice of the inputs which we are free to choose. The variables we use to manipulate the state are known as *control variables*. Generally, the model inputs are used as control variables. In some cases we might be free to choose the initial

Definition 2.1 The system (2.17),(2.18) is *completely μ-step controllable at time t_j* if for any arbitrary state $\mathbf{x}_j$ at time t_j and any desired state $\mathbf{x}^d$, there is an admissible control sequence $\mathbf{u}_j, .., \mathbf{u}_{j+\mu-1}$ on the discrete time interval $[t_j, t_{j+\mu-1}]$ which drives the system to the desired state $\mathbf{x}^d$ at time $t_{j+\mu}$.

If the system is completely μ-step controllable for any time t_j, it is *completely μ-step controllable.*

If the system is completely μ-step controllable (at time t_j) for some μ, we might simply say that the system is *completely controllable (at time t_j).*

Definition 2.2 The system (2.17),(2.18) is *completely ν-step observable at time t_j* if and only if knowledge of the outputs $\mathbf{y}_j, \mathbf{y}_{j+1}, .., \mathbf{y}_{j+\nu-1}$ and of the inputs $\mathbf{u}_j, \mathbf{u}_{j+1}, .., \mathbf{u}_{j+\nu-2}$ is sufficient to determine the state $\mathbf{x}_j$.

If the system is completely ν-step observable for any time t_j, it is *completely ν-step observable.*

If the system is completely ν-step observable (at time t_j) for some ν, we might simply say that the system is *completely observable (at time t_j).*

2.2.2 Theory for the general linear case

For the linear system (2.17),(2.18), the following theorems can be used to determine whether the system is controllable or observable. We first introduce the *μ-step controllability matrix $\mathcal{C}_\mu^j$ for time t_j* and the *ν-step observability matrix $\mathcal{O}_\nu^j$, for time t_j* as follows.

$$\mathcal{C}_\mu^j = (B_{j-1}, \Phi(j, j-1)B_{j-2}, \ldots, \Phi(j, j-\mu+1)B_{j-\mu}), \tag{2.19}$$

$$\mathcal{O}_\nu^j = \begin{pmatrix} \end{pmatrix}$$

(ii) To show that Rank $(\mathcal{O}_\nu^j) = n$ is a necessary condition for complete ν-step observability at time t_j, we suppose that the system is completely ν-step observable at time t_j, but that Rank $(\mathcal{O}_\nu^j) < n$, and let $\mathbf{u}_k = 0$, $k = j, .., j + \nu - 2$.
Then there exists a *nonzero* vector $\mathbf{v} \in \mathbb{R}^n$, such that

$$\mathcal{O}_\nu^j \mathbf{v} = 0. \tag{2.25}$$

Putting $\mathbf{x}_j = \mathbf{v}$ in (2.23) with zero input, we have $\mathbf{z} = 0$, which violates complete μ-step observability (since we have zero output over the whole time interval $[t_j, t_{j+\nu-1}]$ although the state at time t_j is not zero). □

W

Later in the thesis, we will want to apply theoretical results involving the concept of complete ν-step observability to the special case of a time invariant system, and where it is possible, to express the results in terms of the more familiar concept of complete observability. The following theorem enables us to do this.

Theorem 2.3 *a) If the linear time-invariant system is not completely observable, then it is not completely ν-step observable for all positive integers ν.*
b) If $\nu \geq n$ then the linear time-invariant system is completely ν-step observable if and only if it is completely observable.

Proof

a) We must show that $\text{Rank}(\mathcal{O}_n^*) < n$ implies $\text{Rank}(\mathcal{O}_\nu^*) < n$ for all positive integers ν, and we do this by showing

$$\text{Rank}(\mathcal{O}_\nu^*) \leq \text{Rank}(\mathcal{O}_n^*) \tag{2.29}$$

for all ν.
This is clearly true for $\nu \leq n$. We suppose that $\nu = n + 1$. By the Cayley Hamilton theorem [2], we have

$$A^n = \sum_{j=0}^{n-1} \gamma_j A^j \tag{2.30}$$

for some $\gamma_j \in \mathbb{R}$, and so CA^n can be be written as a linear combination of the rows of $\mathcal{O}_n^*$, and hence (2.29) holds. Similarly, for any $\nu > n$,

$$A^\nu = (\sum_{j=0}^{n-1} \gamma_j A^j) A^{\nu-n}, \tag{2.31}$$

and hence CA^ν is still a linear combination of the rows of $\mathcal{O}_n^*$, and so (2.29) holds for all positive integers ν.

b) It follows from part a) that for any positive integer ν, the linear time-invariant system is completely ν-step observable only if it is completely observable. We now suppose that the linear time invariant system is completely observable, and hence is completely n-step observable. As noted earlier, complete n-step observability

implies complete ν-step observability for any $\nu \geq n$, and so part b) of the theorem holds. □

One further result which will be useful when considering the time invariant case is the Hautus condition [28], which is given in Theorem 2.4.

Theorem 2.4 *The linear time-invariant system (2.26),(2.27) is completely observable if and only if,* $\forall\ \lambda \in \mathbb{C}$ *and* $\forall\ \mathbf{s} \in I\!R^n$,
$(A - \lambda I)\mathbf{s} = 0$ *and* $C\mathbf{s} = 0 \Leftrightarrow \mathbf{s} = 0$.

2.3 Nonlinear optimization theory

The theory we give here provides background for the variational methods of data assimilation which we investigate in this thesis. Useful texts for this material include [36], [29], [43], [10], and [89].

2.3.1 Preliminaries

A *Hilbert space* is a complete, linear inner-product space. All the properties of Hilbert spaces are important for our purposes [43]. We denote the inner product defined on a Hilbert space $\mathcal{V}$ by $< \mathbf{x}, \mathbf{y} >_{\mathcal{V}}$, for any two elements $\mathbf{x}$ and $\mathbf{y} \in \mathcal{V}$. We note that real, n-dimensional Euclidean space $\mathbb{R}^n$ with the Euclidean inner product (or "dot product") is a Hilbert space, and throughout the thesis use the notation

$$< \mathbf{x}, \mathbf{y} >:= \mathbf{x}^T\mathbf{y} \tag{2.32}$$

to refer to this inner product.

Later in the thesis, we refer to the *adjoint* of a linear operator. For a linear operator A from a Hilbert space $\mathcal{U}$ to a Hilbert space $\mathcal{V}$, the *adjoint operator* A^* is the linear operator from $\mathcal{V}$ to $\mathcal{U}$ for which, for all $\mathbf{u} \in \mathcal{U}$ and $\mathbf{v} \in \mathcal{V}$

$$< \mathbf{v}, A\mathbf{u} >_{\mathcal{V}} = < A^*\mathbf{v}, \mathbf{u} >_{\mathcal{U}} . \tag{2.33}$$

In the case where $\mathcal{U}$ is $\mathbb{R}^m$ and $\mathcal{V}$ is $\mathbb{R}^n$, both with the Euclidean product (or dot product), $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is an $n \times m$ matrix and we have

$$< \mathbf{v}, A\mathbf{u} >= \mathbf{v}^T A\mathbf{u} = (A^T\mathbf{v})^T\mathbf{u} = < A^T\mathbf{v}, \mathbf{u} >, \tag{2.34}$$

so that $A^T : \mathbb{R}^n \to \mathbb{R}^m$ is the adjoint of A.

We consider a nonlinear, real valued function $\mathcal{J}$ on $\mathcal{V}$. We suppose that $\mathcal{J}$ is three times differentiable at $\mathbf{v}_0 \in \mathcal{V}$, and that $\mathbf{v}_0 + \theta\boldsymbol{\delta}\mathbf{v} \in \mathcal{V}$ represents a perturbation of size $\theta \in [-1, 1]$ in a direction $\boldsymbol{\delta}\mathbf{v}$ from $\mathbf{v}_0$. The Taylor series expansion of $\mathcal{J}$ about $\mathbf{v}_0$ can be written as follows [43]

$$\mathcal{J}(\mathbf{v}_0 + \theta\boldsymbol{\delta}\mathbf{v}) = \mathcal{J}(\mathbf{v}_0) + \theta < \nabla_{\mathbf{v}}\mathcal{J}(\mathbf{v}_0), \boldsymbol{\delta}\mathbf{v} >_{\mathcal{V}} + \theta^2 < \boldsymbol{\delta}\mathbf{v}, \mathcal{H}_{\mathbf{v}}(\mathbf{v}_0)\boldsymbol{\delta}\mathbf{v} >_{\mathcal{V}} + O^3(\theta), \tag{2.35}$$

where the vector $\nabla_{\mathbf{v}}\mathcal{J}(\mathbf{v}_0) \in \mathcal{V}$ is the *gradient* of $\mathcal{J}$ with respect to $\mathbf{v}$ at $\mathbf{v}_0$, and the linear operator $\mathcal{H}_{\mathbf{v}}(\mathbf{v}_0) : \mathcal{V} \to \mathcal{V}$ is the *Hessian* of $\mathcal{J}$ with respect to $\mathbf{v}$ at $\mathbf{v}_0$. Throughout the thesis, we use this notation to denote the gradient and the Hessian of a real valued function.

2.3.2 Unconstrained minimization

We suppose that we wish to minimize a real valued function $\mathcal{J}$, usually referred to as a *cost function*, which is defined on a Hilbert space $\mathcal{V}$. The unconstrained minimization problem we consider is

Problem $\mathcal{U}$:

Minimize $\mathcal{J}$; ie, find $\mathbf{v}^* \in \mathcal{V}$ such that

$$\mathcal{J}(\mathbf{v}^*) \leq \mathcal{J}(\mathbf{v}) \tag{2.36}$$

for all $\mathbf{v}$ in some neighbourhood $\mathcal{N} \subseteq \mathcal{V}$ of $\mathbf{v}^*$.

If such a $\mathbf{v}^*$ exists, it is called a *local minimum* of $\mathcal{J}$. If the inequality in (2.36) is strict, then $\mathbf{v}^*$ is a *unique* local minimum. If $\mathcal{N}(\mathbf{v}^*) = \mathcal{V}$, then $\mathbf{v}^*$ is also a *global* minimum.

Since we have no constraints, the following is a necessary condition for $\mathbf{v}^*$ to minimize $\mathcal{J}$

$$\nabla_{\mathbf{v}}\mathcal{J}(\mathbf{v}^*) = 0. \tag{2.37}$$

In the special case that the cost function is quadratic in $\mathbf{v}$,

$$\mathcal{J} = \frac{1}{2} < \mathbf{v}, A\mathbf{v} >_{\mathcal{V}} + < \mathbf{b}, \mathbf{v} >_{\mathcal{V}} + c, \tag{2.38}$$

where $\mathbf{b} \in \mathcal{V}$ and $c \in \mathbb{R}$ are constants and $A : \mathcal{V} \to \mathcal{V}$ is a linear operator, if A is a positive definite operator, then a minimum $\mathbf{v}^*$ exists, is unique, and is given by $\mathbf{v}^* = -A^{-1}\mathbf{b}$, [43]. If, however, A is only positive semi-definite, a minimum $\mathbf{v}^*$ exists but is not unique, since $\mathbf{v}^* + \mathbf{z}$ is also a minimum for any $\mathbf{z}$ satisfying $< \mathbf{z}, A\mathbf{z} >= 0$. Further, if A is indefinite, then there is no minimum.

We now return to the general case where $\mathcal{J}$ is not necessarily linear or quadratic. We suppose $\mathcal{J}$ is three times differentiable, and so can be expanded in a Taylor series of the form (2.35). Then, for $\|\theta\boldsymbol{\delta}\mathbf{v}\|$ small enough, the quadratic part of the expansion dominates, so if $\nabla_{\mathbf{v}}(\mathbf{v}^*) = 0$, and $\mathcal{H}_{\mathbf{v}}(\mathbf{v}^*)$ is a positive definite operator, then $\mathbf{v}^*$ is a unique local minimum of $\mathcal{J}$ [43]. If $\mathcal{H}_{\mathbf{v}}(\mathbf{v}^*)$ is only a positive semi-definite operator, we can draw no conclusions about $\mathbf{v}^*$, because of the influence of the higher order terms in the expansion. However, if $\mathcal{H}_{\mathbf{v}}(\mathbf{v}^*)$ is indefinite, then $\mathbf{v}^*$ cannot be a minimum.

2.3.3 Constrained minimization

In this subsection, we consider constrained minimization of a real valued function $\mathcal{J}$ over $\mathbb{R}^n$, which with the Euclidean inner product (2.32) is a Hilbert space.

The constrained minimization problem we consider is

Problem $\mathcal{C}$:

Minimize $\mathcal{J}$ subject to the r constraints

$$g_k(\mathbf{v}) = 0, \qquad k = 1, .., r, \tag{2.39}$$

or equivalently

$$\mathbf{g}(\mathbf{v}) = 0, \tag{2.40}$$

where $r \leq n$ and $\mathbf{g}$ is a vector of r real valued functions $g_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, .., r$ which are continuously differentiable. We further assume that the ve

Lagrangian function associated with Problem $\mathcal{C}$ is defined to be

$$\mathcal{L}(\mathbf{v}, \boldsymbol{\lambda}) = \mathcal{J}(\mathbf{v}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{v}), \tag{2.41}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^r$ is a vector of r Lagrange multipliers λ_k. A solution of Problem $\mathcal{C}$, if it exists, can be found by extremizing the (unconstrained) Lagrangian function $\mathcal{L}$ with respect to $\mathbf{v}$ and $\boldsymbol{\lambda}$. Necessary conditions for an extremal are [29],

$$\nabla_{\mathbf{v}} \mathcal{L} = 0, \tag{2.42}$$

$$\nabla_{\boldsymbol{\lambda}} \mathcal{L} = 0. \tag{2.43}$$

Any vector $\mathbf{v} \in \mathbb{R}^n$ satisfying (2.40) can be written in the form

$$\mathbf{v} = \begin{pmatrix} \mathbf{u} \\ \mathbf{x} \end{pmatrix} \tag{2.44}$$

with $\mathbf{u} \in \mathbb{R}^{n-r}$ and $\mathbf{x} \in \mathbb{R}^r$, where the $n-r$ components u_j may be chosen independently, and the r components are determined from the choice of the u_j through (2.40) [89]. We refer to the $n-r$ variables u_j as *control variables*, and the vector $\mathbf{u}$ as a *control vector*.

2.3.4 Solving Problem $\mathcal{C}$ by reducing the control vector

We now describe an iterative method for finding $\mathbf{v}^*$ satisfying necessary conditions for a solution of Problem $\mathcal{C}$ by iterating on the control variables. Since this involves iterating on the control vector $\mathbf{u}$ rather than on the full vector $\mathbf{v}$, this technique is referred to as "reduction of the control vector". This method was suggested for application to 4D variational assimilation by Le Dimet and Talagrand [51], who used the optimal control approach of Lions [53] rather than the Lagrange multiplier approach we use here.

Necessary conditions for an extremal of $\mathcal{L}$ are given by

$$\nabla_{\mathbf{u}} \mathcal{L} = \nabla_{\mathbf{u}} \mathcal{J}(\mathbf{v}) + G_{\mathbf{u}}^T(\mathbf{v}) \boldsymbol{\lambda} = 0, \tag{2.45}$$

$$\nabla_{\mathbf{x}} \mathcal{L} = \nabla_{\mathbf{x}} \mathcal{J}(\mathbf{v}) + G_{\mathbf{x}}^T(\mathbf{v}) \boldsymbol{\lambda} = 0, \tag{2.46}$$

$$\nabla_{\boldsymbol{\lambda}} \mathcal{L} = \mathbf{g}(\mathbf{v}) = 0, \tag{2.47}$$

where $G_{\mathbf{u}} \in \mathbb{R}^{r\times(n-r)}$ and $G_{\mathbf{x}} \in \mathbb{R}^{r\times r}$ are the Jacobian matrices of $\mathbf{g}$ with respect to $\mathbf{u}$ and $\mathbf{x}$ respectively. Since the vectors $\nabla_{\mathbf{v}} g_k(\mathbf{v})$ for $k = 1, .., r$ are linearly independent, the Jacobian $G_{\mathbf{x}}(\mathbf{v})$ is invertible.

From a guess $\mathbf{u}$ fof1Iheguess–?i–cTJ1oΦΦTΦTD1?u–Tj1rol?a–n

2.4.1 The steepest descent algorithm

In this case, the direction $\mathbf{d}^k$ in (2.49) is simply the direction $\nabla_{\mathbf{u}}\mathcal{J}(\mathbf{u}^k)$, G_k is the identity, and ρ^k is chosen to ensure

$$\mathcal{J}(\mathbf{u}^{k+1}) < \mathcal{J}(\mathbf{u}^k). \tag{2.50}$$

In practice, this might be done by setting $\rho^k = 1$ initially on each iteration, and halving ρ^k until (2.50) holds. Alternative step-length choices are given in [36].

The advantage of the steepest descent method lies in its simplicity, but the rather ad-hoc method of finding the step-length can render it very inefficient since it involves many evaluations of $\mathcal{J}$ and $\nabla_{\mathbf{u}}\mathcal{J}$. Further, choosing the direction $\mathbf{d}^k$ with no consideration of the previous directions used is not the most efficient approach. The conjugate gradient method provides a more sophisticated approach to calculating ρ^k and $\mathbf{d}^k$, and we describe this next.

2.4.2 The conjugate gradient method

The aim of the conjugate gradient method (CGM) is to choose the k^{th} descent direction $\mathbf{d}^k$ to be a projection of the gradient $\nabla_{\mathbf{u}}\mathcal{J}(\mathbf{u}^k)$ onto a subspace of $\mathbb{R}^n$ which is orthogonal to $\mathbf{d}^j$ for $j = 0, 1, .., k-1$. Primarily, the CGM addresses an unconstrained minimization problem with quadratic cost function,

$$\mathcal{J} = \frac{1}{2} < \mathbf{u}, A\mathbf{u} > + < \mathbf{b}, \mathbf{u} >, \tag{2.51}$$

where $A \in \mathbb{R}^{n\times n}$ is symmetric, positive definite, and $\mathbf{b} \in \mathbb{R}^n$. The method calculates the optimal step-length ρ^k forhbhdcTΦΦTTΦ?i–?gΦTf1aii–?ts–acpnb

where

$$\rho^k = \frac{< \mathbf{r}^k, \mathbf{d}^k >}{< \mathbf{d}^k, A\mathbf{d}^k >}, \qquad \beta^k = < \mathbf{r}^k$$

The INRIA N1QN3 minimization algorithm

This minimization algorithm uses the the quasi-Newton update formula (2.62), in which G_k, the current approximation of the inverse Hessian, is calculated using a limited memory BFGS update. It is based on an algorithm by Nocedal, [65], with an added preconditioning option, and is described in the documentation [34] and in the paper by Gilbert and Lamaréchal [33].

The general *inverse BFG formula*, for approximating the new inverse Hessian G_{k+1} from G_k is as follows

$$G_{k+1} = (I - \frac{\mathbf{s}^k(\mathbf{y}^k)^T}{(\mathbf{y}^k)^T\mathbf{s}^k})G_k(I - \frac{\mathbf{y}^k(\mathbf{s}^k)^T}{(\mathbf{y}^k)^T\mathbf{s}^k} + \frac{\mathbf{s}^k(\mathbf{s}^k)^T}{(\mathbf{y}^k)^T\mathbf{s}^k}), \tag{2.63}$$

where

$$\mathbf{s}^k = \mathbf{u}^{k+1} - \mathbf{u}^k, \qquad \mathbf{y}^k = \nabla_{\mathbf{u}}\mathcal{J}(\mathbf{u}^{k+1}) - \nabla_{\mathbf{u}}\mathcal{J}(\mathbf{u}^k). \tag{2.64}$$

The matrix G_k is not stored explicitly in memory, but the product $G_k\nabla_{\mathbf{u}}\mathcal{J}(\mathbf{u}_k)$ is calculated from a diagonal matrix D_k and $\hat{m}$ pairs of vectors

$$\{(\mathbf{y}_j, \mathbf{s}_j) : k - \hat{m} \leq j \leq k - 1\} \tag{2.65}$$

if $k \geq \hat{m} + 1$, or just k pairs otherwise. In this way, at the k^{th} iteration with $k \geq \hat{m} + 1$, the oldest pair is discarded and a new pair added. The matrix G_k can be represented using $(2\hat{m} + 1)$ n-vectors, where $\hat{m}$ is an integer supplied by the user, and this is all that need be stored in memory.

The form of the starting matrix $D_{\mathbf{k}}$

The step-length ρ^k in (2.62) is chosen to satisfy *Wolfe's conditions*

$$\mathcal{J}(\mathbf{u}^{k+1}) \leq \mathcal{J}(\mathbf{u}^k) + \alpha_1 \rho^k < \nabla_{\mathbf{u}} \mathcal{J}(\mathbf{u}^k), G_k \nabla_{\mathbf{u}} \mathcal{J}(\mathbf{u}^k) >, \tag{2.68}$$

$$< \nabla_{\mathbf{u}} \mathcal{J}(\mathbf{u}^{k+1}), G_k \nabla_{\mathbf{u}} \mathcal{J}(\mathbf{u}^k) > \geq \alpha_2 < \nabla_{\mathbf{u}} \mathcal{J}(\mathbf{u}^k), G_k \nabla_{\mathbf{u}} (\mathbf{u}^k) >, \tag{2.69}$$

where the constants α_1 and α_2 must be set in the ranges $0 < \alpha_1 < \frac{1}{2}$ and $\alpha_1 < \alpha_2 < 1$. In the algorithm, these are set at the values $\alpha_1 = 10^{-4}$, $\alpha_2 = 0.9$.

The algorithm provides the option of preconditioning by altering the way in which D_k is specified in (2.66). In the preconditioned version, D_k is calculated from D_{k-1} using a diagonal update formula, and the matrix is now diagonal with respect to a new inner product to be specified by the user. This change of inner product is equivalent to a change of orthonormal basis from the canonical basis for $\mathbb{R}^n$, and this change of basis forms the preconditioning. If the usual inner product is the Euclidean product, as assumed in the above, then a new inner product could be of the form

$$< \mathbf{a}, \mathbf{b} >_L = \mathbf{a}^T L^T L \mathbf{b}, \tag{2.70}$$

where L is nonsingular, and the Canonical basis is altered by this change from the basis $\{\mathbf{e}_j\}$, $j = 1, .., n$ to the basis $\{L^{-1}\mathbf{e}_j\}$. Rather than storing the matrix L, or the new basis, the user provides a subroutine which specifies how the inner product is to be calculated.

2.5 Background on probability theory

This section gives a brief overview of probability theory. The aim is to introduce the concept of a statistically "most likely estimate", which is a very important concept in data assimilation. Before this, we give necessary definitions and background on the Gaussian distribution. References for this theory include [3], [14], [55] and [44].

2.5.1 Definitions

Random variables and probability density functions

A *random variable* can be thought of as a numerical value associated with a random event. The *range* of a random variable X, denoted R_X, is the set of all possible values

of X. We consider here only *continuous random variables*, or random variables with an uncountable range.

An n-vector $\mathbf{X}$ of random variables X_j, $j = 1, .., n$ we refer to as a *random n-vector*, or simply as a *random vector* if its dimension is not to be specified. The range of a random n-vector we denote $R_{\mathbf{X}}$, where $R_{\mathbf{X}} = R_{X_1} \times R_{X_2} \times ... \times R_{X_n}$.

Associated with any random variable X is a *probability density function* (abbreviated to pdf), $p_X : R_X \rightarrow \mathbb{R}$. The pdf of a continuous random variable X describes how the unit of probability of X is distributed on the real line. The probability $P(a \leq X \leq b)$ that X takes a value between a and $b \in \mathbb{R}$ is given by

$$P(a \leq X \leq b) = \int_a^b p_X(x)dx. \tag{2.71}$$

The other fundamental properties of a pdf are

$$p_X(x) \geq 0 \quad \text{for all} \quad x \in R_X, \tag{2.72}$$

$$\int_{R_X} p_X(x)dx = 1. \tag{2.73}$$

We write the pdf of a random n-vector as $\mathbf{p_X} : R_{\mathbf{X}} \rightarrow \mathbb{R}^n$, where $\mathbf{p_X}$ is the vector of the pdfs of the random variables X_j, $j = 1, .., n$.

The *joint pdf* of two random variables X and Y is given by $p_{XY} : R_X \times R_Y \rightarrow \mathbb{R}$, with

$$((a \leq X \leq b) \bigcap (c \leq Y \leq d)) = \int_a^b \int_c^d p_{XY}(x, y)\ dx\ dy, \tag{2.74}$$

$$p_{XY}(x, y) \geq 0 \quad \text{for all} \quad x \in R_X, y \in R_Y \tag{2.75}$$

$$\int_{R_X} \int_{R_Y} p_{XY}\ dx\ dy = 1 \tag{2.76}$$

The random variables X and Y are *independent* if

$$p_{XY}(x, y) = p_X(x)p_Y(y). \tag{2.77}$$

The *conditional* pdf of X, given that Y has taken a value y^0 (so y^0 is a *realisation* of the random variable Y) is defined to be

$$p_{X|Y=y^0}(x) = \frac{p_{XY}(x, y)}{p_Y(y^0)}. \tag{2.78}$$

This relation is from *Bayes theorem*, and it can also be written in the form

$$p_{X|Y=y}$$

The *standard deviation* of X is defined as

$$\sigma(X) = (\mathrm{Var}\{X\})^{\frac{1}{2}}. \tag{2.86}$$

The *covariance* of two random variables X and Y is defined as

$$\mathrm{Cov}\{X, Y\} = \mathcal{E}\{(X - \mathcal{E}\{X\})(Y - \mathcal{E}\{Y\})\}, \tag{2.87}$$

and the *correlation* between X and Y is

$$\mathrm{Cor}\{X, Y\} = \mathcal{E}\{XY\}. \tag{2.88}$$

If $\mathrm{Cov}\{X, Y\} = 0$, then X and Y are *uncorrelated*. The *correlation coefficient* of X and Y is

$$\rho(X, Y) = \frac{\mathrm{Cov}\{X, Y\}}{\sigma(X)\sigma(Y)}, \tag{2.89}$$

where $-1 \leq \rho(X, Y) \leq 1$.

By linearity of the expectation operator, we have for matrices A and B of suitable dimensions

$$\mathrm{Cov}\{A\mathbf{X}, B\mathbf{Y}\} = A\mathrm{Cov}\{\mathbf{X}, \mathbf{Y}\}B^T. \tag{2.90}$$

We also note that if $\mathbf{X}$ and $\mathbf{Y}$ are unbiased, then

$$\mathrm{Cov}\{\mathbf{X}, \mathbf{Y}\} = \mathcal{E}\{\mathbf{X}\mathbf{Y}^T\}. \tag{2.91}$$

By the covariance matrix of a random vector $\mathbf{X}$, we mean the covariance matrix $\mathrm{Cov}\{\mathbf{X}, \mathbf{X}\}$.

2.5.2 The Gaussian distribution

We now suppose that a random variable X represents random error. The *Gaussian* distribution, also called the *Normal* distribution, has the following characteristics that make it suitable for representing errors:

1. Continuity

2. An unbounded range

3. Symmetry about the mean (so positive and negative errors are equally likely)

4. A "bell-shaped" distribution, which gives small probability to large errors and largest probability to the smallest errors,

5. Tractability, ie a pdf that is easy to work with.

The Gaussian pdf for a random variable X with mean μ and variance σ^2 is given by

$$p_X(x) = \frac{1}{\sqrt{(2\pi)\sigma}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} \right. . \tag{2.92}$$

For a random n-vector with mean $\boldsymbol{\mu} \in \mathbb{R}^n$ and nonsingular covariance matrix R, the Gaussian pdf is

$$\mathbf{p_X}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}(\det(R))^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu})). \tag{2.93}$$

2.5.3 "Most likely" estimates

The following development broadly follows that of the paper by Lorenc [55].

We suppose that $\mathbf{x}^f$ is a "prior"estimate of a random n-vector $\mathbf{X}$. If we know that

$$\mathbf{X} = \mathbf{x}^f + \mathbf{e}^f, \tag{2.94}$$

where $\mathbf{e}^f$ is a random n-vector of the error $\mathbf{X} - \mathbf{x}^f$, and we know that $\mathbf{e}^f$ is Gaussian and unbiased with nonsingular covariance matrix P^f, then the pdf of $\mathbf{X}$ is given by

$$\mathbf{p_X}(\mathbf{x}) = k_1 \exp(-\frac{1}{2}(\mathbf{x}-\mathbf{x}^f)^T (P^f)^{-1}(\mathbf{x}-\mathbf{x}^f)), \tag{2.95}$$

where k_1 is a constant. We now suppose that we have a random p-vector $\mathbf{Y}$ that satisfies

$$\mathbf{Y} = C\mathbf{X} + \boldsymbol{\delta}, \tag{2.96}$$

where $C \in \mathbb{R}^{p \times n}$, and $\boldsymbol{\delta}$ is a random p-vector of the error $\mathbf{Y} - C\mathbf{X}$, which is Gaussian and unbiased, with nonsingular ΦTf1ΦΦTD1?is–?a–?random–TJ1TΦΦTf1Φ,f1ΦΦiΦΦΦ

To do this we need to know $\mathbf{p_{X|Y=y^0}(x)}$, which by (2.79) is given by

$$\mathbf{p_{X|Y=y^0}(x)} = \frac{\mathbf{p_{Y|X=x}(y^0)p_X(x)}}{\mathbf{p_y(y^0)}}, \tag{2.98}$$

hence

$$\mathbf{p_{X|Y=y^0}(x)} = \frac{k_1 k_2}{k_3}$$)#1TTf1ΦΦTD1TΦΦPΦΦTf1TΦΦyf1f1ΦΦT1?(–Tj

Chapter 3

Sequential data assimilation

Sequential data assimilation schemes treat observations as they become available in time, and then discard them. If a 3D data assimilation method, which is designed to produce an analysis at a single time, is applied repeatedly, this can be seen as sequential data assimilation. 4D sequential data assimilation methods, however, are designed so that an analysis should gradually draw closer to the true model state, as more observations are processed. In control theory, dynamic observers are designed for this very purpose, and so observer theory is very relevant to sequential data assimilation. An example of an observer originally designed for engineering applications which is being investigated for use in data assimilation, is the Kalman

are available frequently. Using observer theory, we are able to give conditions for the linear, time invariant case under which the successive correction analysis will converge *in time* to the true solution. In Section 3.5, we compare the Cressman successive correction scheme with a robust observer in data assimilation for a simple example. These experiments serve to illustrate how an observer which is designed for temporal convergence to the true solution can perform much better than successive correction scheme designed for an analysis at a single time.

3.1 Background on 3D data assimilation schemes

By "3D" data assimilation schemes we mean schemes that are designed to give an analysis at a single time, and do not attempt to take into account the time-tendency of the observations. This section gives a brief overview of a few 3D data assimilation schemes that have been used in the past and to date, and which we use or refer to in this thesis. We firstly outline successive correction methods, which are some of the earlier schemes to have been proposed and implemented. We then introduce the method of optimal interpolation, on which the schemes currently used in many meteorological centres are based. Finally, we describe the 3D variational assimilation (3DVAR) method which is being developed for operational use at several centres as an intermediate stage in the development of 4D variational assimilation (4DVAR) schemes.

The material in this section is intended to be only a brief outline of the methods discussed. A more in-depth overview of data assimilation methods and further references are given in the review paper by Ghil and Malenotte-Rissoli [32] and the books by Daley [24] and Bennett [4].

3.1.1 Successive correction schemes

Successive correction schemes were introduced to meteorology in the 1950s for operational objective analysis, by Bergthórson and Döös [9], and by Cressman [21]. The *Cressman scheme* [21] was designed for systems with few observations, widely scattered, which are to be fitted as closely as possible. This method was intend-

ed to improve on the earlier polynomial spline methods [68], [35] by being more suitable for use over larger areas with less dense data coverage, and by being com-

$(l,m)^{th}$ element of the matrix $W^{(i)}$ is given by

$$W_{lm}^{(i)} = \frac{(R_m^{(i)})^2 - d_{lm}^2}{(R_m^{(i)})^2 + d_{lm}^2}, \qquad l = 1,..,n, \quad m = 1,..,p_k, \tag{3.3}$$

where $R_m^{(i)}$ is the radius of influence at the i^{th} iteration for observation m (ie, the m^{th} component of $\mathbf{y}_k$), and d_{lm} is the distance between observation m and grid point l. In the original paper introducing this method, 4 iterations were carried out with different radii of influence [21].

3.1.2 Optimal interpolation

The method of optimal interpolation (OI) has been widely used in operational data assimilation for NWP in the 1980s and 1990s [42]. Important references for the method include the papers by Gandin [30] and Lorenc [55]. The OI method was designed for a system in which observations are linearly related to the model state. We suppose that we wish to estimate $\mathbf{x}_k^t$, and that we havm

as discussed in [56]. The OI method is in fact not truly optimal since it does not update the error covariance matrix P_k of the background estimate in a way that takes into account the earlier observations which have already been assimilated. The Kalman filter does this, but the extra cost involved is large. The OI method is sometimes more realistically referred to as *statistical interpolation*, [54]. Although designed for a linear system, the method can be extended for use in a system in which the observations are nonlinearly related to the model state,

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k^t) + \boldsymbol{\delta}_k, \tag{3.8}$$

with $\mathbf{y}_k$, $\mathbf{x}_k^t$ and $\boldsymbol{\delta}_k$ as defined in (3.4), and where $\mathbf{h}_k : \mathbb{R}^n \rightarrow \mathbb{R}^{p_k}$ is a nonlinear operator. This can be done by linearizing (3.8) about $\mathbf{x}_k^b$. We describe this approach in a little more detail in the context of the nonlinear extension to the Kalman filter in Section 3.2.

3.1.3 3D variational assimilation, and the PSAS method

The three dimensional variational assimilation (3DVAR) method takes a different approach to minimizing the function (3.5). Rather than solving equations (3.6) and (3.7), the approach is to iterate to the minimizing solution $\mathbf{x}_k^a$. The gradient of (3.5) with respect to $\mathbf{x}_k$ is

$$\nabla_{\mathbf{x}_k}\mathcal{J} = P_k^{-1}(\mathbf{x}_k - \mathbf{x}_k^b) + C_k^T R_k^{-1}(C_k\mathbf{x}_k - \mathbf{y}_k), \tag{3.9}$$

and this may be used in a gradient method to iterate to the optimal solution. We describe a few such methods in Chapter 2, Section 2.4.

The 3DVAR method is currently being developed for implementation for operational data assimilation at several meteorological centres, with plans for extension to

In general, the dimension p_k of the observation vector is much smaller than the dimension n of the model state $\mathbf{x}_k$; for meteorological applications it might be t-
wo orders of magnitude less, and even less for oceanographic applications. The
advan

MOE :

$$\mathrm{Cov}\{\boldsymbol{\varepsilon}_k, \boldsymbol{\delta}_j\} = 0, \qquad \mathrm{Cov}\{\boldsymbol{\varepsilon}_k, \mathbf{x}_0^f\} = 0, \qquad \mathrm{Cov}\{\boldsymbol{\delta}_k, \mathbf{x}_0^f\} = 0, \quad \forall j, k. \tag{3.15}$$

We note that, from (3.12) and ME2, we have

$$\mathrm{Cov}\{\mathbf{x}_j^t, \boldsymbol{\varepsilon}_k\} = 0, \qquad \forall k \geq j, \tag{3.16}$$

and similarly, from (3.13) and OE2 we have

$$\mathrm{Cov}\{\mathbf{x}_j^t, \boldsymbol{\delta}_k\} = 0, \qquad \forall k > j. \tag{3.17}$$

As mentioned above, we assume that the error covariance matrices Q_k, R_k and P_k^f are nonsingular. All covariance matrices for a random vector with itself are symmetric positive semi-definite, and are positive definite if they do not contain null variances or perfect correlations [81]. We assume that this is so. Assumption ME1 that model error is unbiased is in fact not restrictive; if we have

$$\mathcal{E}\{\boldsymbol{\varepsilon}_k\} = \bar{\boldsymbol{\varepsilon}_k} \neq 0, \tag{3.18}$$

we can define

$$\boldsymbol{\varepsilon}_k$$

A necessary condition for a minimum is that $\nabla_{\mathbf{x}_k} \mathcal{J}(\mathbf{x}_k) = 0$, ie

$$(P_k^f)^{-1}(\mathbf{x}_k - \mathbf{x}_k^f) + C_k^T R_k^{-1}(C_k \mathbf{x}_k - \mathbf{y}_k^0) = 0. \tag{3.21}$$

It can be shown [60] that the best estimate $\mathbf{x}_k^a$ is a unique, global minimum of (3.21), and satisfies

$$\mathbf{x}_k^a = \mathbf{x}_k^f + K_k(\mathbf{y}_k^0 - C_k \mathbf{x}_k^f) \tag{3.22}$$

where $K_k \in \mathbb{R}^{n \times p_k}$ is the *Kalman gain matrix* given by

$$K_k = P_k^f C$$

It can be verified that $\mathbf{e}_k^a$ and $\boldsymbol{\varepsilon}_k$ are uncorrelated because of the relation (3.16). Further, $\mathbf{e}_{k+1}^f$ is unbiased since $\mathbf{e}_k^a$ and $\boldsymbol{\varepsilon}_k$ are, and we have

$$P_{k+1}^f = A_k P_k^a A_k^T + Q_k. \tag{3.31}$$

We can now apply the Kalman filter equations (3.22) and (3.23) to find

The other assumptions on model error and observational error are as before, including the assumption that model error and observational error are uncorrelated.

The analysis given for the standard Kalman filter in equations (3.22) and (3.23) still holds for producing the best estimate or analysis $\mathbf{x}_k^a$ based on $\mathbf{x}_k^f$, $\mathbf{y}_k^0$ and their error covariance matrices. The expressions (3.27) and (3.28) for the analysis error covariance are also unchanged, since the forecast error and observational error are still uncorrelated. The expression for the new forecast error $\mathbf{e}_{k+1}^f$ given in (3.30) is also unchanged, but because model error is now serially correlated, equation (3.16) no longer holds, and model error and analysis error are now correlated. Hence, the expression (3.31) for the new forecast error covariance matrix must be modified as follows.

$$\begin{aligned} P_{k+1}^f &= \mathrm{Cov}\{(A_k\mathbf{e}_k^a - \boldsymbol{\varepsilon}_k), (A_k\mathbf{e}_k^a - \boldsymbol{\varepsilon}_k)\} \\ &= A_k P_k^a A_k^T - A_k\mathrm{Cov}\{\mathbf{e}_k^a, \boldsymbol{\varepsilon}_k\} - \mathrm{Cov}\{\mathbf{e}_k^a, \boldsymbol{\varepsilon}_k\}A_k^T + Q_k, \end{aligned} \tag{3.34}$$

or, defining the covariance matrix of analysis and model errors as

$$P_k^{aq} = \mathrm{Cov}\{\mathbf{e}_k^a, \boldsymbol{\varepsilon}_k\}, \tag{3.35}$$

we have

$$P_{k+1}^f = A_k P_k^a A_k^T - A_k P_k^{aq} - P_k^{aq} A_k^T + Q_k \tag{3.36}$$

It now remains to specify P_{k+1}^{aq} from P_k^{aq}. This is given by

$$P_{k+1}^{aq} = (I - K_k C_k)(A_k P_k^{aq} - Q_k)G_k^T. \tag{3.37}$$

Finally, Q_{k+1} is calculated from Q_k as follows

$$Q_{k+1} = G_k Q_k G_k^T + S_k. \tag{3.38}$$

So, for serially correlated model error, we have the standard Kalman filter equations, except that the evolution of the forecast error covariance (3.31) is modified to (3.36), and in addition we must propagate the covariance matrix of analysis and model errors as expressed in equation (3.37), and the model error covariance matrix (3.38).

We note that serially correlated observational errors can be dealt with in a similar way, if we assume model error is uncorrelated in time [23]. In this case the equation

for the analysis error covariance propagation would be modified, and we would need to work out the propagation of the covariance matrix of forecast and observational errors. It is also possible to allow for serial correlations in both model error and observational error [44], but at greater complication still.

3.2.4 The extended Kalman filter

We now consider the extension of Kalman filtering theory to the nonlinear, stochastic dynamic system

$$\mathbf{x}_k^t = \mathbf{f}_k(\mathbf{x}_k^t, \mathbf{u}_k) + \boldsymbol{\varepsilon}_k, \tag{3.39}$$

with observations nonlinearly related to the state as follows

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k^t) + \boldsymbol{\delta}_k, \tag{3.40}$$

with observations

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k^t) \tag{3.45}$$

defined as in (2.1),(2.6), assuming no model error or observation error. A dynamic observer for (3.44),(3.45) may be written in the form

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) + G_k(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k)), \tag{3.46}$$

where $\mathbf{x}_k$ is an estimate of the true model state $\mathbf{x}_k^t$, and the *feedb*

and the error equation corresponding to (3.48) is

$$\mathbf{e}_{k+1} = (A - GC)\mathbf{e}_k, \tag{3.52}$$

and therefore

$$\mathbf{e}_k = (A - GC)^k \mathbf{e}_0. \tag{3.53}$$

Hence, to satisfy the condition $\mathbf{e}_k \to 0$ as $t_k \to \infty$, the eigenvalues of (A-GC), denoted by $\lambda_i(A - GC)$, must satisfy the condition

$$|\lambda_i(A - GC)| < 1 \qquad \forall i = 1, .., n. \tag{3.54}$$

In certain cases it is possible to choose the feedback matrix G so that the matrix $(A - GC)$ has any specified eigenvalues, and in particular, it is possible to ensure that the condition (3.54) holds. Theorem 3.1 gives sufficient conditions for this to hold [90].

Theorem 3.1 *If the system (3.49),(3.50) is completely observable, then it is possible to choose the matrix G in (3.51) so that the eigenvalues of $(A - GC)$ take prescribed values.*

3.3.2 Eigenstructure assignment

The *inverse eigenvalue problem* of assigning eigenvalues to the system (3.51) allows some freedom in choosing the corresponding eigenvectors in the case $p > 1$, and since we have some freedom in choosing the eigenvectors also, our problem now is one of *eigenstructure assignment* [28]. We now describe how we can choose the eigenstructure of the dynamic observer (3.51).

We suppose that the conditions of Theorem 3.1 hold, and that the set of eigenvalues we wish to assign is

$$\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_n\}; \tag{3.55}$$

where

$$\lambda_i \in \mathbb{C}, \quad |\lambda_i| < 1, \quad \text{and} \quad \lambda \in \Lambda \Rightarrow \bar{\lambda} \in \Lambda \qquad \text{for} \quad i = 1, ..., n. \tag{3.56}$$

We let $D = diag\{\lambda_i\}$ and let X be the modal matrix of right eigenvectors of $(A-GC)$ and Y be the modal matrix of $(A^T - C^T G^T)$. Then our problem is to choose G and X to satisfy

$$(A - GC)X = XD, \tag{3.57}$$

or, equivalen

is still some freedom to choose the $\boldsymbol{\eta}_i$ if $p > 1$. We can use this freedom to ensure that our selected eigenvalues are as insensitive as possible to perturbations in A, C and G and thus that the system is *robust*

This method for improving the robustness of the system can not be guaranteed to converge to the minimum possible value of $\|(Y^T)^{-1}\|_F$, but in practice it has been found to reduce its value significantly.

An algorithm for a robust observer

1) Calculate the QR decomposition of C^T into

$$C^T = \left[\tilde{Q}_c, Q_c\right] \begin{bmatrix} R_o \\ 0 \end{bmatrix}. \tag{3.72}$$

2) For each $i = 1, .., n$,
calculate the QR decomposition of $(A - \lambda_i I)Q_c$ into

$$(A - \lambda_i I)Q_c = \left[\tilde{S}_i, S_i\right] \begin{bmatrix} R_i \\ 0 \end{bmatrix}. \tag{3.73}$$

3) Choose columns from each of the S_i to be columns of the first guess Y, in such a way that Y is invertible.

4) For $i = 1, ..., n$, modify the columns $\boldsymbol{\eta}_i$ of Y as follows:

4a) calculate the QR decomposition of $Y_{-i} = \{\boldsymbol{\eta}_1, .., \boldsymbol{\eta}_{i-1}, \boldsymbol{\eta}_{i+1}, .., \boldsymbol{\eta}_n\}$ into

$$Y_{-i} = [\tilde{Z}_i, \mathbf{z}_i] \begin{bmatrix} \tilde{Y}_i \\ 0 \end{bmatrix}. \tag{3.74}$$

4b) project the vector $\mathbf{z}_i$ into space S_i to satisfy condition **a**) and then normalize:

$$\boldsymbol{\eta}_i = S_i S_i^T \mathbf{z}_i / \|S_i S_i^T \mathbf{z}_i\|. \tag{3.75}$$

5) Repeat Step 4 until $\|(Y^T)^{-1}\|_F$ reaches a local minimum.

6) Using the Y found, let the feedback matrix be G where

$$G^T = R_o^{-1} \tilde{Q}_c^T (A^T Y - Y D) Y^{-1}. \tag{3.76}$$

a theoretical extension of successive correction schemes to 4D. This could provide a way to make successive correction schemes more appropriate for use in a modern application of data assimilation in which observations are available more frequently.

3.4.2 Successive correction schemes as observers

Here we suppose that a successive correction method is to be used for data assimilation with observations available frequently, and we show how it may be regarded as an observer. If an observer is applied over an assimilation interval, then the analysed solution over that interval does not satisfy the model dynamics, but the observer dynamics. Hence, considering a sequential data assimilation method as an observer gives a different way of understanding some of the properties of the analysed solution.

In particular, by considering a successive correction scheme using observer theory, we are able to consider theoretical convergence *in time* of the scheme to the true model solution. In the data assimilation literature, the issue of whether a successive correction scheme *converges* generally refers to the question of whether the successive iterations or corrections (at a single analysis time) bring the analysis close to the true solution at that time [24]. In general, however, observations from more than one time are needed to determine the true state uniquely. Here, we consider whether the successive correction technique converges in time to the true model state.

We suppose that the evolution of the true model state is given by the linear, time invariant system

$$\mathbf{x}_{k+1}^t = A\mathbf{x}_k^t + B\mathbf{u}_k, \qquad k = 0, .., N-1 \tag{3.77}$$

as defined in (2.17) and we suppose that we have observations available at every timestep, related to the true model state by

$$\mathbf{y}_k = C\mathbf{x}_k^t, \qquad k = 0, .., N-1 \tag{3.78}$$

as defined in (2.18).

The successive correction method, with a constant number s of corrections, finds an analysis $\mathbf{x}_k^a$ from a prior estimate $\mathbf{x}_k^b$ using the following iteration

$$\mathbf{x}_k^{(i+1)} = \mathbf{x}_k^{(i)} + W^{(i+1)}(\mathbf{y}_k - C\mathbf{x}_k^{(i)}), \qquad i = 0, .., s-1, \tag{3.79}$$

with $\mathbf{x}_k^{(0)} = \mathbf{x}_k^b$, and $\mathbf{x}_k^a = \mathbf{x}_k^{(s)}$, and where the $W^{(i)}$, $i = 1, .., s$ represent the weighting matrices used in the successive corrections. After manipulation, the method can be written for theoretical purposes in the form

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \tilde{W}^{(s)}(\mathbf{y}_k - C\mathbf{x}_k^b), \tag{3.80}$$

where the matrix $\tilde{W}^{(s)}$ is given by the recursion

$$\tilde{W}^{(i+1)} = W^{(i+1)}(I - C\tilde{W}^{(i)}) + \tilde{W}^{(i)}, \qquad i = 1, .., s-1, \tag{3.81}$$

with $\tilde{W}^{(1)} = W^{(1)}$.

From an analysis $\mathbf{x}_k^a$ at time t_k, the prior estimate for the next timestep, $\mathbf{x}_{k+1}^b$, is found using the model equations

$$\mathbf{x}_{k+1}^b = A\mathbf{x}_k^a + B\mathbf{u}_k, \qquad k = 0, .., N-1. \tag{3.82}$$

Substituting (3.82) into (3.80) gives

$$\mathbf{x}_{k+1}^a = A\mathbf{x}_k^a + B\mathbf{u}_k + \tilde{W}^{(s)}(\mathbf{y}_{k+1} - C\mathbf{x}_{k+1}^b), \tag{3.83}$$

which expresses the successive correction method in the form of a dynamic observer. Subtracting (3.77) from (3.83) and using (3.78) gives the following equation for the evolution of the error $\mathbf{e}_k = \mathbf{x}_k^a - \mathbf{x}_k^t$,

$$\mathbf{e}_{k+1} = A\mathbf{e}_k + \tilde{W}^{(s)}(C(A\mathbf{x}_k^t + B\mathbf{u}_k) - C(A\mathbf{x}_k^a + B\mathbf{u}_k)), \tag{3.84}$$

or

$$\mathbf{e}_{k+1} = (A - \tilde{W}^{(s)}CA)\mathbf{e}_k. \tag{3.85}$$

Hence, the successive correction scheme converges to the true model state $\mathbf{x}_k^t$ in time if the eigenvalues of $(A - \tilde{W}^{(s)}CA)$ have modulus less than unity. The weighting matrix $\tilde{W}^{(s)}$ plays a similar rôle as the feedback matrix G in the observer (3.51). The matrix C in (3.52) has been replaced by the matrix product CA in (3.85), because (3.83) uses observations at time t_{k+1}, rather than at time t_k as the observer (3.51) does.

If observations are available less frequently, then $\mathbf{x}_{k+1}^a$ is specified by (3.83) when observations are available, and is equal to $\mathbf{x}_{k+1}^b$ given in (3.82) when observations

are not available. If observations are available every r^{th} timestep, then the error $\mathbf{e}_k$ satisfies

$$\mathbf{e}_{r(k+1)} = (A - \tilde{W}^{(s)}CA)A^{r-1}\mathbf{e}_{rk}, \tag{3.86}$$

and hence the successive correction scheme converges to the true model sate in time if the eigenvalues of $(A - \tilde{W}^{(s)}CA)A^{r-1}$ have modulus less than unity.

Discussion

In Section 3.3 on dynamic observer theory we discussed how the feedbac

The theta method for the 1D heat equation

The 1D heat equation on $z \in [0,1]$, $t \in [0,T]$, with a point heat source of strength $\frac{1}{3}$ at $z = \frac{1}{4}$ is,

$$v_t = \sigma v_{zz} + \frac{1}{3}\delta(z - \frac{1}{4}), \tag{3.87}$$

where δ is the Dirac delta function. For this equation, with initial condition

$$v(z,0) = \alpha(z), \tag{3.88}$$

and zero boundary conditions

$$v(0,t) = 0, \quad v(1,t) = 0, \tag{3.89}$$

the "theta method" discretisation for some $\theta \in [0,1]$ is

$$x_j^{k+1} - x_j^k = \frac{\sigma \Delta t}{\Delta z^2}\left\{(1-\theta)\delta^2 x_j^k + \theta \delta^2 x_j^{k+1}\right\} + s_j \Delta t \tag{3.90}$$

with initial condition

$$x_j^0 = \alpha(j\Delta z), \tag{3.91}$$

and zero boundary conditions

$$x_0^k = 0, \quad x_J^k = 0, \tag{3.92}$$

where $x_j^k \approx v(j\Delta TD\theta cc\theta, Tj\theta pcTj\theta TTf\theta TD\theta\phi\theta\eta Tj\theta \ TD\theta and\ j\theta \ mcfr\alpha TJ\theta TTf\theta \ TD\theta Tj\theta zT$

where the state $\mathbf{x}_k$ at time t_k

$$
\begin{aligned}
C_{i,j} &= \frac{(j+1)\Delta z - obs_i}{\Delta z}, \\
C_{i,j+1} &= {}^{obs_i}
\end{aligned}
$$

assimilation is, but it does not have the spatial shape of the true solution. Fig. 3.3 shows the case $p = 1$, where only one observation is available right near one of the boundaries of the domain, using $R = 0.3$. In this case, the data assimilation has only a small impact on the results. Increasing R to 0.9, so that the radius of influence extends all the way across the spatial domain, convergence is still slow, as Fig. 3.4 shows.

The robust observer

Very good results are achieved using eigenvalue set Λ_b, in which we assign to the observer system (ie, to the matrix $(A - CG)$) the system eigenvalues multiplied by 0.75. In this case convergence to the true solution is achieved in fewer than 20 timesteps using 5 observations, as Fig. 3.5 shows. Using eigenvalue set Λ_a, in which the eigenvalues to be assigned are evenly distributed between -0.5 and 0.5 gives less pleasing results. From this it seems that it is important for good convergence to reduce the modulus of all the eigenvalues, as we do when assigning eigenvalue set Λ_b

produces much faster convergence in data sparse areas. This serves as an example of how designing the feedback matrix of an observer to ensure temporal convergence to the true solution can improve on the empirical spatial smoothing of a successive correction method. The robust observer design itself, involving eigenstructure assignment, however, would be too expensive for systems with very large dimension, and hence for application to operational data assimilation.

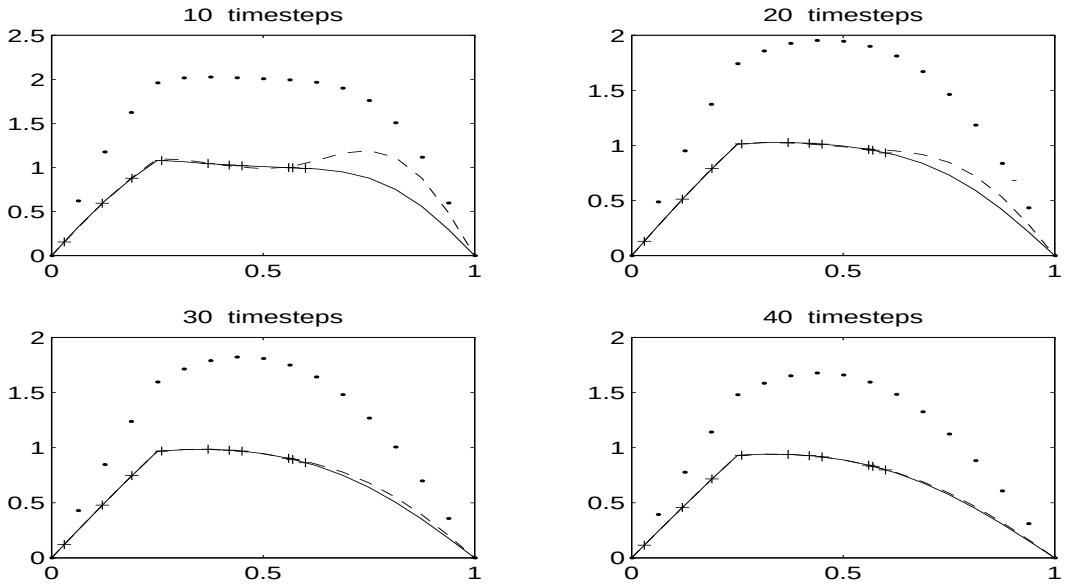
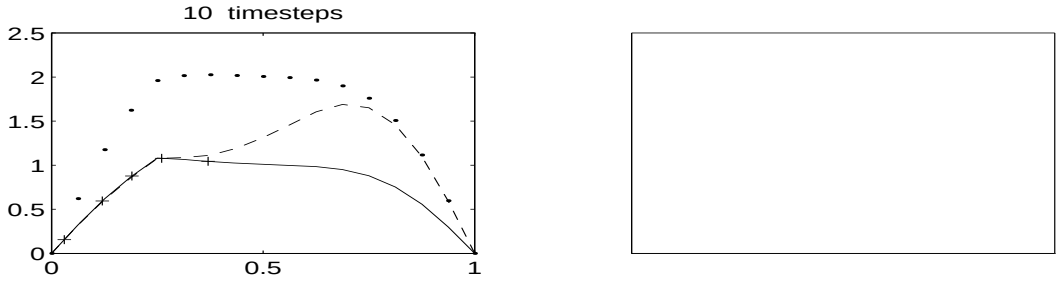
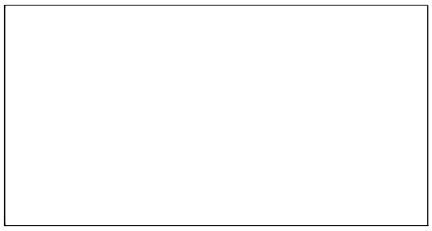


Figure 3.1: Data assimilation using the Cressman scheme with $R = 0.3$ using 10 observations. Solid line: true solution; dotted line: solution with no assimilation; dashed line: solution with assimilation, crosses: observations.





o

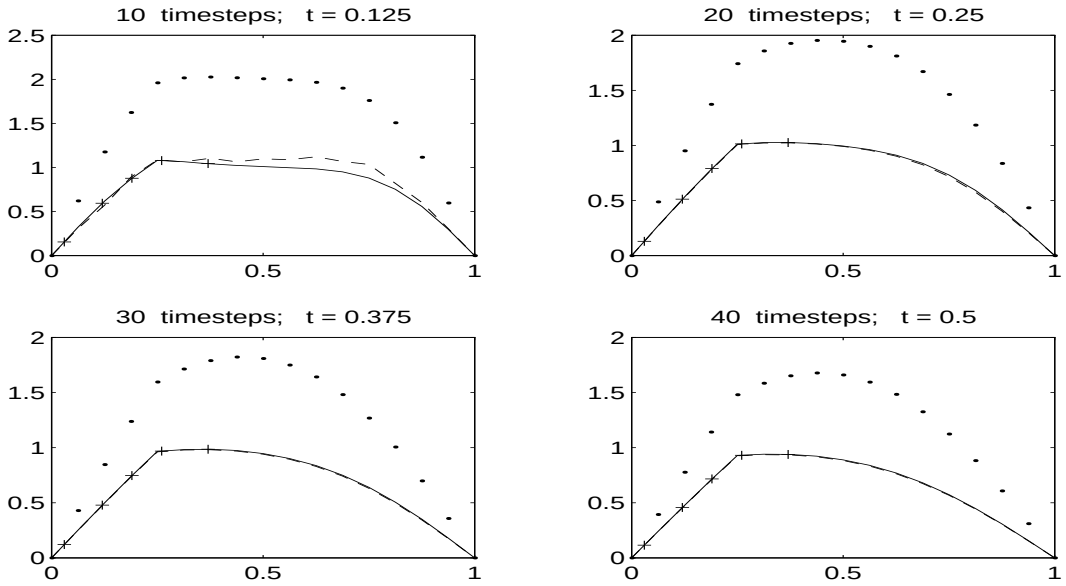
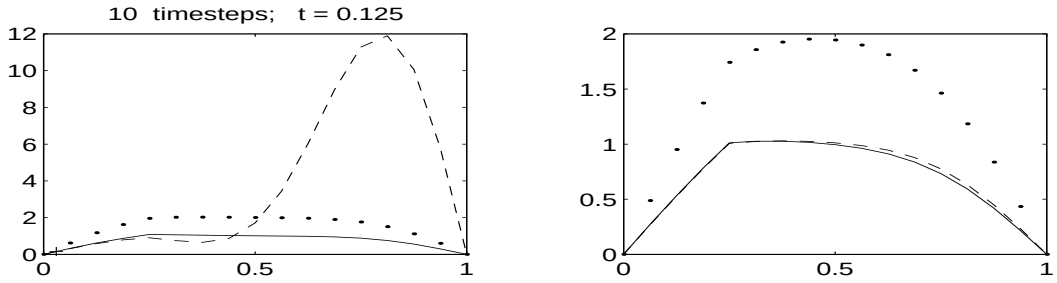


Figure 3.5: Data assimilation using the robust observer with eigenvalue set Λ_b


10 timesteps; t = 0.125
12
10
8
6
4
2
0
0
0.5
1
2
1.5
1
0.5
0
0
0.5
1

Chapter 4

4D Variational assimilation

4D variational methods of data assimilation were introduced to meteorology by Sasaki in his paper of 1958 [75]. These schemes seek to find the model state which minimizes some cost function over a particular assimilation interval, subject to constraints on the model state. Most typically the constraints require that the model state should satisfy the dynamical model equations over the assimilation time period.

Sasaki put forward two approaches to variational assimilation. In the *strong constraint* approach, the solution is constrained to satisfy the model equations exactly. In the *weak constraint* approach, the model equations are required to hold only approximately, allowing for model error. Sasaki's papers [75], [76], [77], [78], deal with methods of solving these minimization problems analytically for simple, continuous models.

Various methods for solving the strong constraint problem are outlined in [32]. One method is to iterate on the model initial state rather than on the model state over the whole assimilation interval. This technique of "reducing the control vector" which we outlined in Chapter 2, Section 2.3, significantly reduces the cost of variational assimilation. In this case the initial state is the control vector. The method was introduced to meteorology in the mid 1980s in the papers by Le Dimet and Talagrand [51], Lewis and Derber [52], Lorenc [55], and Courtier and Talagrand [18], [80]; and to oceanography by Thacker and Long [83]. It is currently being developed for implementation as an operational data assimilation scheme at several national meteorological centres [73].

Derber [26] suggested carrying out 4D variational assimilation, adding to the

4.1 The strong constraint approach

In the strong constraint approach, model error is neglected and we work with the model

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k), \qquad k = 0, .., N-1. \tag{4.4}$$

We suppose that we have observations given by (4.2) and a background estimate of the initial state given by (4.3).

4.1.1 The method

In the strong constraint approach to variational assimilation, we aim to minimize a cost function $\mathcal{J}$ with three components,

$$\mathcal{J} = \mathcal{J}_b + \mathcal{J}_o + \mathcal{J}_c, \tag{4.5}$$

where $\mathcal{J}_b$ penalizes distance from the background estimate, $\mathcal{J}_o$ penalizes distance from the observations, and $\mathcal{J}_c$ ensures that the solution has required smoothness properties, so that the process of initialization (or part of it) mentioned in Chapter 1 can be incorporated in the optimization procedure. The work in this thesis does not include the component $\mathcal{J}_c$ for simplicity, although it is important in operational applications of data assimilation [91], [94], [19], that use more complex models and fewer observations than we use in our idealized experiments.

The strong constraint problem we address is

Problem

Minimize, with respect to $\mathbf{x}_0, .., \mathbf{x}_N$

$$\mathcal{J} = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{j=0}^{N-1}(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j)^T R_j^{-1}(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j), \tag{4.6}$$

subject to (4.4)

The matrices $P_0^{-1} \in \mathbb{R}^{n\times n}$ and $R_j^{-1} \in \mathbb{R}^{p_j \times p_j}$ are symmetric positive definite weighting matrices which reflect the accuracies of $(\mathbf{x}_0 - \mathbf{x}_0^b)$ and of $(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j)$. If the inverse covariance matrices of the errors $\boldsymbol{\beta}_0$ and $\boldsymbol{\delta}_j$ are known and are nonsingular, these can be used as weighting matrices, and under certain assumptions this choice

leads to a statistically optimal analysis. We discuss this further in Section 4.4 where we also consider model error. Some detail on how the matrices P_0^{-1} and R_j^{-1} are prescribed in practice is given in [13], [57].

We now use the theory of Chapter 2, Section 2.3, on reducing the control vector to give a method for solving the constrained minimization problem Problem . We

We note that, since $\mathbf{x}_1, .., \mathbf{x}_N$ have been calculated from the current guess of the control vector, the vectors $\boldsymbol{\lambda}_k$ of Lagrange multipliers can be calculated recursively backwards from the condition (4.12).

The system of equations (4.11), (4.12) is known as the system of *adjoint equations* for the model (4.4), or as the *adjoint model*. In this context, the Lagrange multipliers are known as *adjoint variables*, and we refer to the vectors $\boldsymbol{\lambda}_k$ of adjoint variables as *adjoint vectors*.

We can now evaluate the gradients of $\mathcal{L}$ with respect to the control vectors. The gradient with respect to the initial state is given by

$$\begin{aligned} \nabla_{\mathbf{x}_0}\mathcal{L} &= \nabla_{\mathbf{x}_0}\mathcal{J} - F_0^T(\mathbf{x}_0)\boldsymbol{\lambda}_1 && (4.13)\\ &= P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + H_0^T(\mathbf{x}_0)R_0^{-1}(\mathbf{h}_0(\mathbf{x}_0) - \mathbf{y}_0) - F_0^T(\mathbf{x}_0)\boldsymbol{\lambda}_1 && (4.14)\\ &= P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) - \boldsymbol{\lambda}_0, && (4.15) \end{aligned}$$

where the additional adjoint vector $\boldsymbol{\lambda}_0 \in \mathbb{R}^n$ is defined via the relation (4.11) with $k = 0$. This gradient can be used in a descent algorithm, such as one outlined in Chapter 2, Section 2.4, to improve our guess of the control vector. We summarize this procedure in the following algorithm.

Algorithm IS

1. From a guess of the control vector $\mathbf{x}_0$,
   calculate the model states $\mathbf{x}_1, .., \mathbf{x}_N$ using the model equations (4.4).

2. From the end condition (4.12), calculate the adjoint vectors $\boldsymbol{\lambda}_{N-1}, .., \boldsymbol{\lambda}_0$ using the model states calculated in Step 1.

3. From $\boldsymbol{\lambda}_0$, calculate $\nabla_{\mathbf{x}_0}\mathcal{L}$ using (4.15)

4. Use the gradient $\nabla_{\mathbf{x}_0}\mathcal{L}$ in a gradient algorithm to obtain a better guess of the control vector $\mathbf{x}_0$, and repeat until convergence criteria are satisfied.

4.1.2 The incremental approach for Problem $\mathcal{S}$

Although the method of reducing the control vector significantly reduces the expense of 4D variational assimilation, further reductions in its expense are still required to

make it feasible for operational implementation. The incremental approach [20] was suggested to allow flexibility to incorporate simplifications which will reduce the expense of the method. We first describe the incremental approach, and then describe the approximations that can be made to reduce expense.

Expanding the nonlinear model (4.4) in a Taylors series about the "background" state $\mathbf{x}_k^b$ obtained from a model run using the model with $\mathbf{x}_0^b$ as the initial state, we have for a small perturbation $\boldsymbol{\delta}\mathbf{x}_k$ of $\mathbf{x}_k$,

$$\mathbf{x}_{k+1}^b + \boldsymbol{\delta}\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k^b) + F_k(\mathbf{x}_k^b)\boldsymbol{\delta}\mathbf{x}_k + \mathbf{o}(\boldsymbol{\delta}\mathbf{x}_k), \qquad k = 0, .., N-1, \tag{4.16}$$

where $F_k(\mathbf{x}_k^b)$ is the Jacobian of $\mathbf{f}_k$ with respect to $\mathbf{x}_k$ evaluated at $\mathbf{x}_k^b$, and $\mathbf{o}(\boldsymbol{\delta}\mathbf{x}_k)$ represents the higher order terms in the expansion. Since (4.4) holds at $\mathbf{x}_k^b$, we have, after neglecting higher order terms,

$$\boldsymbol{\delta}\mathbf{x}_{k+1} = F_k(\mathbf{x}_k^b)\boldsymbol{\delta}\mathbf{x}_k, \tag{4.17}$$

which is referred to as the *tangent linear model* (TLM). For mid-latitude meteorological models, Lacarra and Talagrand [48] have shown that the the TLM is a fair approximation to the full nonlinear model for periods of up to around 48 hours.

The observations $\mathbf{y}_k$ are related to the perturbation $\boldsymbol{\delta}\mathbf{x}_k^t = (\mathbf{x}_k^t - \mathbf{x}_k^b)$ as follows

$$\begin{aligned} \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k^t) + \boldsymbol{\delta}_k = \mathbf{h}_k(\mathbf{x}_k^b + \boldsymbol{\delta}\mathbf{x}_k^t) + \boldsymbol{\delta}_k & (4.18)\\ &\approx \mathbf{h}_k(\mathbf{x}_k^b) + H_k(\mathbf{x}_k^b)\boldsymbol{\delta}\mathbf{x}_k^t + \boldsymbol{\delta}_k, \qquad k = 0, .., N-1. & (4.19) \end{aligned}$$

The incremental approach to solving Problem proceeds as follows. Firstly, a background run is performed from a background guess $\mathbf{x}_0^b$ of the initial state, using the full nonlinear model (4.4) to calculate the terms $\mathbf{x}_k^b$. The minimization problem is then to find the optimal *increment* or perturbation $\boldsymbol{\delta}\mathbf{x}_0$ to $\mathbf{x}_0^b$, by minimizing a cost function of the form

$$\mathcal{J}(\boldsymbol{\delta}\mathbf{x}_0) = \frac{1}{2}\boldsymbol{\delta}\mathbf{x}_0^T P_0^{-1}\boldsymbol{\delta}\mathbf{x}_0 + \frac{1}{2}\sum_{j=0}^{N-1}(H_j(\mathbf{x}_j^b)\boldsymbol{\delta}\mathbf{x}_j - \mathbf{d}_j)^T R_j^{-1}(H_j(\mathbf{x}_j^b)\boldsymbol{\delta}\mathbf{x}_j - \mathbf{d}_j) \tag{4.20}$$

subject to the constraints (4.17), $k = 0, .., N-1$, where

$$\mathbf{d}_k = \mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k^b). \tag{4.21}$$

Since all the constraints (4.17) are linear, we now have a cost function which is quadratic in the control vector $\boldsymbol{\delta}\mathbf{x}_0$, and so a unique global minimum to this problem exists if the Hessian of $\mathcal{J}$ with respect to $\boldsymbol{\delta}\mathbf{x}_0$ (satisfying the constraints (4.17)) is positive definite. We give conditions for this to hold in Chapter 5, Section 5.2. The adjoint equations are, in this case,

$$\boldsymbol{\lambda}_k = F_k^T(\mathbf{x}_k^b)\boldsymbol{\lambda}_{k+1} - H_k^T(\mathbf{x}_k^b)R_k^{-1}(H_k(\mathbf{x}_k^b)\boldsymbol{\delta}\mathbf{x}_k - \mathbf{d}_k), \quad k = 0, .., N-1, \tag{4.22}$$

$$\boldsymbol{\lambda}_N = 0, \tag{4.23}$$

but using (4.21) and (4.19) we see these are the same as the adjoint equations (4.11) of the full, nonlinear system, except that the higher order terms of (4.19) have been neglected in the forcing in (4.22), and that the Jacobian F_k is evaluated at $\mathbf{x}_k^b$.

If $\boldsymbol{\delta}\mathbf{x}_0^*$ is the control vector which solves the incremental problem, then $\mathbf{x}_0^* := \mathbf{x}_0^b + \boldsymbol{\delta}\mathbf{x}_0^*$, (and hence the model states $\mathbf{x}_1^*, .., \mathbf{x}_N^*$ found from this initial state) is a good approximation to the solution of the full, nonlinear mimimization problem, provided the TLM is a "valid" approximation to the full nonlinear model.

The incremental approach can be used to further reduce the cost of 4D variational assimilation by performing the background run to calculate the $\mathbf{x}_k^b$ using the full nonlinear model (4.4), but carrying out the iteration on $\boldsymbol{\delta}\mathbf{x}_0$ at lower resolution. The iteration at lower resolution could also be performed using a simplified (and hence less expensive) version of the TLM. Research is also being carried out on variants of the incremental approach. These include applying the incremental approach at lower resolution, perhaps using multi-grid strategies (this is the so-called multi-incremental approach), and interspersing several "inner loop" iterations with an "outer loop" nonlinear run, in which a new background field for the next inner loop iterations is obtained. One question being looked at both theoretically and practically, is whether the low resolution inner loop iterations give improvements which correspond to an improvement at full resolution, and whether these methods do converge to a solution of the full nonlinear problem.

Several centres planning to implement the adjoint method for large models are developing simplified or modified tangent linear models for the minimization using the incremental approach. This gives a way of overcoming some of the problems of

the adjoin

Since $\Phi^T(k,j)$ is the *adjoint* operator of $\Phi(k,j)$ with respect to the Euclidean inner product, we see one reason why the adjoint equations are so-called.

Hence we have the following property,

$$< \tilde{\boldsymbol{\lambda}}_k, \boldsymbol{\delta}\mathbf{x}_k > = < \tilde{\boldsymbol{\lambda}}_k, \Phi(k,j)\boldsymbol{\delta}\mathbf{x}_j > = < \Phi^T(k,j)\tilde{\boldsymbol{\lambda}}_k, \boldsymbol{\delta}\mathbf{x}_j > \tag{4.30}$$

$$= < \tilde{\boldsymbol{\lambda}}_j, \boldsymbol{\delta}\mathbf{x}_j >, \qquad \text{for all} \quad k \geq j \tag{4.31}$$

where $\tilde{\boldsymbol{\lambda}}_k$ and $\tilde{\boldsymbol{\lambda}}_j$ solve the unforced adjoint equations (4.26), and where $\boldsymbol{\delta}\mathbf{x}_j$ and $\boldsymbol{\delta}\mathbf{x}_k$ solve the tangent linear model (4.17).

4.2.2 Adjoint model development

Clearly, the derivation of the adjoint model is a major part in the setting up of the adjoint method. Here we outline a few different approaches to the derivation of the adjoint model which might be used in the wider context of optimization problems.

One approach is to work with a continuous, rather than a discrete version of the model. In this case, the calculus of variations is the appropriate theory for finding conditions for extrema of an optimization problem (we give some background on the calculus of variations, and an overview of this approach in the report [40]). Using the method of Lagrange multipliers (which in this case are functions) to deal with the model constraints, the adjoint model is given by the *Euler Lagrange equations*, and an expression analogous to equation (4.15) can be found for the gradient of the Lagrange functional with respect to the initial state. This leads us to a continuous analogue of Algorithm IS, involving the model equations, the adjoint equations and the gradient of the Lagrangian functional with respect to the initial state. In general, these will have to be discretized in some appropriate way so that the problem can be solved numerically. In application to data assimilation, this approach has been used in [6], for example.

A disadvantage of this approach, however, is that in general, the discretized adjoint equations will not in fact be the true adjoint equations of the discretized model.

The approach most generally taken in the development of adjoint models for data assimilation is that it is better to find the adjoint of the discrete model, so that theoretically at least, we can obtain an exact expression for the required gradient, [19]. We use this approach in the work described in this thesis, and work out the adjoint of the discrete models which we wish to use "by hand". For very large and complex models, however, this would be a much more difficult task.

A different approach to deriving the adjoint model is to work directly from the model computer code. Each assignment statement in the computer code can be treated as a constraint to be multiplied by a Lagrange multiplier. Differen

The other advantage of the correction term technique is that it can account for *schematic* model errors, and it is suggested in [26] that by application to many cases, this method could yield an estimate of the model's systematic error in each timestep. Further, the correction term found for an assimilation interval could be used in a subsequent forecast to counteract model error here too.

In the correction term technique, the model error is approximated by

$$\boldsymbol{\varepsilon}_k = s_k \mathbf{e}, \qquad k = 0, .., N-1 \tag{4.32}$$

where the $s_k \in \mathbb{R}$ are predetermined constants, and $\mathbf{e} \in \mathbb{R}^n$ is a constant *correction term* to be determined. Hence we work with the model

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + s_k$$

Enforcing $\nabla_{\boldsymbol{\lambda}_k} \mathcal{L} = 0$ for $k = 1$

Problem $\mathcal{L}$

Minimize, with respect to $\mathbf{x}_0, .., \mathbf{x}_N, \boldsymbol{\varepsilon}_0, .., \boldsymbol{\varepsilon}_{N-1}$

$$\begin{aligned} \mathcal{J} &= \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{j=0}^{N-1}(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j)^T R_j^{-1}(\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j) \\ &\quad + \frac{1}{2}\sum_{j=0}^{N-1}\boldsymbol{\varepsilon}_j^T Q_j^{-1}\boldsymbol{\varepsilon}_j \end{aligned} \tag{4.44}$$

subject to (4.41),

where the symmetric, positive definite weighting matrices *TDθ*,*Tjθ*T

In the nonlinear case, however, a cost function $\mathcal{J}$ with multiple minima corre-

covariance matrix P_0 for the next assimliation interval, and also gives us a way of calculating the accuracy of a forecast initiated at time t_N.

In this advantage of the Kalman filter lies also (arguably) the biggest drawback for its application in operational assimilation in meteorology and oceanography, compared to other methods of finding the same statistically optimal solution. This drawback is the h

The adjoint method

We consider here the technique of reducing the control vector for solving Problem $\mathcal{L}$. In this case, we have $N+1$ state vectors plus N model error vectors as variables, and N sets of model constraints (4.41). If we use $\mathbf{x}_0$ and $\boldsymbol{\varepsilon}_0, .., \boldsymbol{\varepsilon}_{N-1}$ as $N+1$ control vectors, we can uniquely determine the remaining variables, the N state vectors $\mathbf{x}_1, .., \mathbf{x}_N$ from the N constraints (4.41). The constrained minimization problem Problem $\mathcal{L}$ is equivalent to the unconstrained optimization problem of extremizing the Lagrangian function

$$\mathcal{L} = \mathcal{J} + \sum_{j=0}^{N-1} \boldsymbol{\lambda}_{j+1}^T (\mathbf{x}_{j+1} - \mathbf{f}_j(\mathbf{x}_j) - \boldsymbol{\varepsilon}_j), \tag{4.45}$$

where $\boldsymbol{\lambda}_1, .., \boldsymbol{\lambda}_N \in \mathbb{R}^n$ are N vectors of Lagrange multipliers.

Using the method of reducing the control vectors described in Chapter 2, Section 2.3, the solution can be found by iterating on the control vectors $\mathbf{x}_0$ and $\boldsymbol{\varepsilon}_0, .., \boldsymbol{\varepsilon}_{N-1}$, as follows. From a guess of the control vectors, the model states $\mathbf{x}_1, .., \mathbf{x}_N$ are calculated using the constraints (4.41). As before, this ensures that $\nabla_{\boldsymbol{\lambda}_k} \mathcal{L} = 0$ for $k = 1, .., N$, and again, enforcing $\nabla_{\mathbf{x}_k} \mathcal{L} = 0$ for $k = 1, .., N$, yields the same adjoint equations

$$\boldsymbol{\lambda}_k = F_k^T(\mathbf{x}_k)\boldsymbol{\lambda}_{k+1} - H_k^T(\mathbf{x}_k) R_k^{-1}(\mathbf{h}_k(\mathbf{x}_k) - \mathbf{y}_k), \qquad k = 1, .., N-1, \tag{4.46}$$

$$\boldsymbol{\lambda}_N = 0. \tag{4.47}$$

As before, the gradient with respect to the initial state is given by

$$\nabla_{\mathbf{x}_0} \mathcal{L} = P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) - \boldsymbol{\lambda}_0, \tag{4.48}$$

where the additional adjoint vector $\boldsymbol{\lambda}_0 \in \mathbb{R}^n$ is defined via the relation (4.46) with $k = 0$. The gradient with respect to the model error vector $\boldsymbol{\varepsilon}_k$ for $k = 0, .., N-1$ is given by

$$\nabla_{\boldsymbol{\varepsilon}_k} \mathcal{L} = Q_k^{-1} \boldsymbol{\varepsilon}_k - \boldsymbol{\lambda}_{k+1}, \qquad k = 0, .., N-1. \tag{4.49}$$

These gradients can be used in a descent algorithm, such as one outlined in Chapter 2, Section 2.4, to improve our guess of the control vectors. We summarize this procedure in the following algorithm.

Algorithm ISME

1. From a guess of the control vectors $\mathbf{x}_0$ and $\boldsymbol{\varepsilon}_0, .., \boldsymbol{\varepsilon}_{N-1}$,
   calculate the model states $\mathbf{x}_1, .., \mathbf{x}_N$ using the model equations (4.41).

2. From the end condition (4.47), calculate the adjoint vectors $\boldsymbol{\lambda}_{N-1}, .., \boldsymbol{\lambda}_0$ using the model states calculated in Step 1.

3. From $\boldsymbol{\lambda}_{N-1}, .., \boldsymbol{\lambda}_0$, calculate $\nabla_{\boldsymbol{\varepsilon}_{N-1}}\mathcal{L}, .., \nabla_{\boldsymbol{\varepsilon}_0}\mathcal{L}$ and $\nabla_{\mathbf{x}_0}\mathcal{L}$ using (4.49) and (4.48).

4. Use these gradients in a gradient algorithm to obtain a better guess of the control vectors $\mathbf{x}_0$,and $\boldsymbol{\varepsilon}_0, .., \boldsymbol{\varepsilon}_{N-1}$, and repeat until convergence criteria are satisfied.

In practice, this algorithm is expensive for operational data assimilation, since it involves iterating on N control vectors of dimension n. A second problem is that the conditioning of the problem minimizing the cost function simultaneously with all these control vectors could be very poor, [5], [83].

We mention here, however, an attempt to solve a similar problem by Thacker and Long [83] in the context of a simple oceanographic model. Rather than trying to recover the model error vectors $\boldsymbol{\varepsilon}_k$, they attempt to recover unknown model forcing terms (wind stresses) from the data, although they mention that model error is accounted for via uncertainty in the forcing. They look at the question of data sufficiency, and find that if forcing is to be recovered with the initial state, a huge amount of extra data is needed, much more than they could expect to be available in practice. They also mention that the problem of recovering model forcing with the initial state is ill-conditioned and requires many iterations of the descent algorithm.

Chapter 5

The correction term technique

We described the "correction term technique" for 4D variational assimilation introduced by Derber [26] in Chapter 4 Section 3. Here we take a further look at both theoretical and practical aspects of using a correction term as a control vector, instead of or as well as using the initial state as a control vector.

In Section 5.2, we look for conditions for uniqueness of the solution to the 4D variational assimilation problem using the initial state, the correction term and both together as control vectors. Using the initial state as the control vector, uniqueness depends on the condition of complete N-step observability of the system. We show, however, that in general conditions for a unique solution using the correction term as a control vector are different, and so it might be possible to determine uniquely the initial state from the data and not the correction term, or vice versa. In each case adding a background estimate of the control vector to the cost function guarantees uniqueness in the case of data insufficiency. This point has been considered in data assimilation using the initial state as the control vector [8], but not in published work using the correction term technique. We look at uniqueness of the solution using both control vectors together by using the technique of state augmentation, and by relating conditions for observability of the augmented system with conditions for observability of the original system.

In a practical context, we compare the performance of the different control vectors using a simple linear model. We compare the ability of each control vector to compensate for errors in the initial state and for model error which is constant in

time. We also examine the impact of the number of observations on the results, and the use of a background estimate of the correction term. The experiments are described in Section 5.3, and the results are presented and discussed in Section 5.4. In Section 5.5 we summarize the theoretical and practical results of the chapter.

correction term as the control vector w

This approach is justified because of the validity of the tangent linear model over the assimilation length scales and is necessary because of limitations of computational resources (Chapter 4, Section 4.1). Further, and importantly, the incremental approach gives us a minimization problem with a unique solution under certain conditions which w

Problem $\mathcal{A}_{ISCT}$

Minimize with respect to $\mathbf{x}_0, .., \mathbf{x}_N; \mathbf{e}$,

$$\mathcal{J} = \frac{1}{2} \sum_{j=0}^{N-1} (C_j \mathbf{x}_j - \mathbf{y}_j)^T R_j^{-1} (C_j \mathbf{x}_j - \mathbf{y}_j) \tag{5.5}$$

subject to (5.3).

In (5.5) the matrices $R_j^{-1} \in \mathbb{R}^{p_j \times p_j}$ are assumed to be symmetric positive definite, and to represent the relative accuracies of the observational data. Ideally, they should be the inverse observational error covariance matrices. We consider modifications of Problem $\mathcal{A}_{ISCT}$ involving background terms later.

Since $\text{Rank}(A_k) = n$ and $\text{Rank}(B_k) = m$ for all k by assumption, specification of $\mathbf{x}_0$ and $\mathbf{e}$ uniquely determines the model state at all subsequent times. Hence $\mathbf{x}_0$ and $\mathbf{e}$ can be used as control vectors, and we view Problem $\mathcal{A}_{ISCT}$ as that of finding an optimal initial state (IS) and correction term (CT).

If we consider the correction term to be fixed (for example, if we assume that there is no model error, or that model error is represented by a known bias), the initial state is the control variable and we have the familiar strong constraint 4D variational assimilation method outlined in Chapter 4, Section 4.1. Here we will refer to this as Problem $\mathcal{A}_{IS}$.

Problem $\mathcal{A}_{IS}$

Minimize $\mathcal{J}$ *defined in (5.5) with respect to* $\mathbf{x}_0, ...\mathbf{x}_N$, *subject to (5.3), with* $\mathbf{e}$ *fixed.*

The problem addressed by correction term technique, using the correction term only as the control vector, we refer to here as Problem $\mathcal{A}_{CT}$. In this case we assume the initial state is known.

Problem $\mathcal{A}_{CT}$

Minimize $\mathcal{J}$ *defined in (5.5) with respect to* $\mathbf{x}_1, ...\mathbf{x}_N; \mathbf{e}$, *subject to (5.3), with* $\mathbf{x}_0$ *fixed.*

One of our objectives in this chapter is to give conditions under which Problems $\mathcal{A}_{ISCT}$, $\mathcal{A}_{IS}$ and $\mathcal{A}_{CT}$ have a unique solution.

5.2.1 Using the initial state as the control vector

In this subsection we consider Problem $\mathcal{A}_{IS}$, which consists of constrained minimization of (5.5) subject to (5.3), with $\mathbf{e}$ fixed, so throughout this subsection $\mathbf{e}$ is assumed to be given.

Our aim is to find conditions for a unique minimum $\mathbf{x}_0$ of Problem $\mathcal{A}_{IS}$. In terms of the initial state $\mathbf{x}_0$, (5.14) can be written

$$\tilde{\mathcal{J}} = \frac{1}{2}\mathbf{x}_0^T \tilde{A}_{IS}\mathbf{x}_0 + \tilde{\mathbf{b}}_{IS}^T\mathbf{x}_0 + \tilde{c}_{IS}, \tag{5.16}$$

where $\tilde{\mathbf{b}}_{IS} \in \mathbb{R}^n$, $\tilde{c}_{IS} \in \mathbb{R}$, and the Hessian matrix $\mathcal{A}_{IS} \in \mathbb{R}^{n\times n}$ is given by

$$\tilde{A}_{IS} = \sum_{j=0}^{N-1} \Phi_j^T C_j^T R_j^{-1} C_j \Phi_j. \tag{5.17}$$

From the theory of Chapter 2, Section 2.3, a necessary condition for $\mathbf{x}_0$ to be a minimum is that $\nabla_{\mathbf{x}_0}\tilde{\mathcal{J}}$ vanishes, ie,

$$\tilde{A}_{IS}\mathbf{x}_0 + \tilde{\mathbf{b}}_{IS} = 0. \tag{5.18}$$

Any $\mathbf{x}_0$ satisfying (5.18) is a minimum since $\tilde{A}_{IS}$ is positive semi-definite, and it is unique if and only if $\tilde{A}_{IS}$ is positive definite or equivalently if and only if $\text{Rank}(\tilde{A}_{IS}) = n$. We now link this condition of uniqueness to the observability of the system.

The observations (5.4) may be related to $\mathbf{x}_0$ and $\mathbf{e}$ using (5.9) as follows.

$$\mathbf{y}_k = C_k(\Phi_k\mathbf{x}_0 + \tilde{B}_k\mathbf{e}) + \boldsymbol{\delta}_k, \qquad k = 0, .., N-1, \tag{5.19}$$

which may be written as

$$\mathcal{O}_N\mathbf{x}_0 = Y_N - \mathcal{T}_N\mathbf{e} + D_N, \tag{5.20}$$

where

$$\mathcal{O}_N = \begin{pmatrix} C_0\Phi_0 \\ C_1 \\ \vdots \\ \end{pmatrix}$$

and

$$Y_N = \begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N-1} \end{pmatrix}, \qquad D_N = \begin{pmatrix} \boldsymbol{\delta}_0 \\ \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_{N-1} \end{pmatrix}. \tag{5.22}$$

We note that $\mathcal{O}_N = \mathcal{O}_N^0$, the N-step observability matrix at time t_0 defined in equation (2.20).

If there is no observational error, then the right hand side of (5.20) is known. In Chapter 2, Section 2.2, we defined *complete N-step observability at time t_0* as the ability to determine uniquely the state $\mathbf{x}_0$ from the observations $\mathbf{y}_k$ and specified inputs $B_k\mathbf{e}$, $k = 0, .., N-1$. It was proved in Theorem 2.2 that the system is completely N-step observable at time t_0 if and only if $\text{Rank}(\mathcal{O}_N) = n$, and therefore N-step observability is a necessary and sufficient condition to determine a unique initial state $\mathbf{x}_0$ in this case. We note, however, that only if the model (5.3) is a perfect representation of the evolution of the true model state $\mathbf{x}_k^t$, will the solution $\mathbf{x}_0$ of (5.20) represent the true initial state $\mathbf{x}_0^t$.

In practice, the observational error is not negligible. It is therefore not possible in general to estimate $\mathbf{x}_0$ exactly from the data, since the observational errors are unknown. Hence we attempt to find a least-squares estimate for $\mathbf{x}_0$, which can be done by solving Problem $\mathcal{A}_{IS}$. If the observational errors $\boldsymbol{\delta}_k$ are unbiased, Gaussian and uncorrelated in time with covariance matrices $\text{Cov}\{\boldsymbol{\delta}_k, \boldsymbol{\delta}_k\} = R_k$, and if the model (5.3) is a perfect representation of the evolution of the true model state, then this least squares estimate of $\mathbf{x}_0$ is the *most likely* estimate of $\mathbf{x}_0^t$. If these assumptions on the observational error statistics and on the model accuracy do not hold to a good approximation, then our estimate $\mathbf{x}_0$ will not be a good approximation to the most likely estimate.

The question of whether Problem $\mathcal{A}_{IS}$ has a unique solution depends on complete N-step observability, as we now show.

Definition 5.1 *We say that the linear time varying system (5.3),(5.4) containing observational error is completelyt*

Theorem 5.1 *Problem $\mathcal{A}_{IS}$ has a unique solution if and only if the system (5.3),(5.4) is completely N-step observable at time t_0.*

Proof

From the previous discussion, it suffices to show that Rank $(\tilde{A}_{IS}) = n$ if and only if

Hence, Rank$(\tilde{A}_{IS}) = n$ if and only if Rank$(\mathcal{O}_N) = n$. $\square$

Theorem 5.1 is a recasting for our data assimilation problem of a known result in filtering theory that the time-varying system (5.3),(5.4) is completely observable if and only if $\tilde{A}_{IS}$ is positive definite for some N, [44].

We now specialize to the time-invariant case, which is given by

$$\mathbf{x}_{k+1}$$

does not tell us whether complete observability is sufficient for uniqueness if $N < n$, but Theorem 5.1 does. Since in most applications of data assimilation in meteorology and oceanography, the number of timesteps over which observations are available is far less than the dimension of the system (ie, $N << n$), this is an important point. Hence, using the concept of complete N-step observability is important for our application not only because it allows generalization to a time-varying system, but also because it can be used to give a condition for uniqueness of the data assimilation problem that depends on the length of the assimilation interval, and the particular time t_0 at which the assimilation is started.

We next $\mathbf{x}_N$ t . i s 1 ? e n – T D 1] ? ? i s – f 1 i n – T J 1 Φ i ? w h i c ? ? p a

$\mathcal{J} =$ 1

$$\frac{}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2}\sum_{j=0}^{N-1}(C_j\mathbf{x}_j - \mathbf{y}_j)^T R_j^{-1}(C_j\mathbf{x}_j - \mathbf{y}_j)$$

Proof

The Hessian matrix $(P_0^{-1} + \tilde{A}_{IS})$ is the sum of a positive definite and a positive semi-definite matrix, and hence is positive definite. □

5.2.2 Using the correction term as the control vector

Here we consider the case where $\mathbf{x}_0$

given in [70], [71], and [64]. Our problem is rather different, however, since we only look for a constant input, and because we consider the discrete, time-varying case.

Since observational errors are not negligible in reality, we seek a least squares estimate of the correction term $\mathbf{e}$ from the observational data by solving Problem $\mathcal{A}_{CT}$. We now give conditions under which this problem has a unique solution.

Theorem 5.3 *Problem $\mathcal{A}_{CT}$ has a unique solution if and only if Rank$(\mathcal{T}_N) = m$.*

Proof

With the matrix $\tilde{U}$ defined as in (5.24), we have

$$\tilde{A}_{CT} = \mathcal{T}_N^T \tilde{U}^T \tilde{U} \mathcal{T}_N. \tag{5.39}$$

Since $\tilde{U}$ is positive definite, we have, by the same argument as in Theorem 5.1, that

$$\text{Rank}(\tilde{A}_{CT}) = m \quad \text{if and only if} \quad \text{Rank}(\mathcal{T}_N) = m. \tag{5.40}$$

□

It is interesting to note, however, that complete N-step observability is neither a necessary nor a sufficient condition for a unique solution of Problem $\mathcal{A}_{CT}$. Hence, given the same set of observations, it is possible that Problem $\mathcal{A}_{IS}$ has a unique solution but Problem $\mathcal{A}_{CT}$ does not, and vice versa. We show this by means of simple counter-examples, in the case where $m = n$ and the matrices B_k are equal to the identity matrix. We note that the result does not rely on the fact that the observations $\mathbf{y}_0$ contain information about the initial state, but not the correction term.

Theorem 5.4 *Complete N-step observability at time t_0 is neither a necessary nor a sufficient condition for a unique solution of Problem $\mathcal{A}_{CT}$.*

Proof

We consider the system (5.3) with

$$A_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \tag{5.41}$$

and $B_k = I$ for $k = 0, 1$.

To show that complete N-step observability is not a necessary condition for a unique solution of Problem $\mathcal{A}_{CT}$, we give an example in which

$$\text{Rank}(\mathcal{O}_3) < 2, \qquad \text{Rank}(\mathcal{T}_3) = 2. \tag{5.42}$$

We suppose that the data set is given by (5.4) with

$$C_0 = (0, 1), \qquad C_1 = (0, 1), \qquad C_2 = (1, 1). \tag{5.43}$$

In this case

$$\mathcal{O}_3 = \begin{pmatrix} C_0 \\ C_1 A_0 \\ C_2 A_1 A_0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 2 \end{pmatrix}, \tag{5.44}$$

$$\mathcal{T}_3 = \begin{pmatrix} 0 \\ C_1 \\ C_2 A_1 + C_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix}, \tag{5.45}$$

and so (5.42) holds.

To show that complete N-step observability is not a sufficient condition for a unique solution of Problem $\mathcal{A}_{CT}$, we give an example in which

$$\text{Rank}(\mathcal{O}_3) = 2, \qquad \text{Rank}(\mathcal{T}_3) < 2. \tag{5.46}$$

If the data set is now given by

$$C_0 = (0, 0), \qquad C_1 = (1, 3), \qquad C_2 = (1, 1), \tag{5.47}$$

we have

$$\mathcal{O}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 4 \\ 0 & 2 \end{pmatrix}, \qquad \mathcal{T}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 3 \\ 1 & C_3 \end{pmatrix},,$$

o1),1TΓ1?B–Tj1aΦTD1B2

and so (5.46) holds. □

We have, however, the following special case as an example where complete $(N-1)$-step observability is a necessary and sufficient condition for a unique solution of Problem $\mathcal{A}_{CT}$. We consider again the time invariant system (5.32),(5.33). The experiments we describe in Section 5.3 are for a time invariant system, and so Theorem 5.5 is relevant for this case.

Theorem 5.5 *For the time invariant system (5.32),(5.33) with $m = n$ and B non-singular, Problem $\mathcal{A}_{CT}$ has a unique solution if and only if the system (5.32),(5.33) is completely $(N-1)$-step observable.*

Proof

We need to show that $\text{Rank}(\mathcal{T}_N) = n$ if and only if $\text{Rank}(\mathcal{O}_{N-1}) = n$.
We let $\mathcal{T}' = \mathcal{T}_N B^{-1}$, and note that since B is nonsingular, $\text{Rank}(\mathcal{T}_N) = n$ if and only if $\text{Rank}(\mathcal{T}') = n$.
For the time invariant system, we have

$$\mathcal{O}_{N-1} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{N-2} \end{pmatrix}, \qquad \mathcal{T}' = \begin{pmatrix} 0 \\ C \\ C(A+I) \\ \vdots \\ C(A^{N-2} + A^{N-3} + .. + I) \end{pmatrix} \tag{5.49}$$

Since each row of $\mathcal{T}'$ can be written as a linear combination of rows of $\mathcal{O}_{N-1}$, we have that $\text{Rank}(\mathcal{T}') \leq \text{Rank}(\mathcal{O}_{N-1})$. Further, since each row of $\mathcal{O}_{N-1}$ can be written as a linear combination of the rows of $\mathcal{T}'$, we have that $\text{Rank}(\mathcal{O}_{N-1}) \leq \text{Rank}(\mathcal{T}')$. Hence $\text{Rank}(\mathcal{O}_{N-1}) = \text{Rank}(\mathcal{T}') = n$ if and only if $\text{Rank}(\mathcal{T}_N) = n$, which proves the result. □

Finally, we show that a unique solution to Problem $\mathcal{A}_{CT}$ can be guaranteed provided we add a background term to the cost function. In this case, Problem $\mathcal{A}_{CT}$ is modified to

Problem $\mathcal{B}_{CT}$

Minimize with respect to $\mathbf{x}_1, .., \mathbf{x}_N; \mathbf{e}$

$$\mathcal{J} = \frac{1}{2}(\mathbf{e} - \mathbf{e}^b)^T Q^{-1}(\mathbf{e} - \mathbf{e}^b) + \frac{1}{2}\sum_{j=0}^{N-1}(C_j\mathbf{x}_j - \mathbf{y}_j)^T R_j^{-1}(C_j\mathbf{x}_j - \mathbf{y}_j) \qquad (5.50)$$

subject to (5.3) with $\mathbf{x}_0$ *fixed.*

In (5.50), $\mathbf{e}^b \in \mathbb{R}^m$ is a background estimate for $\mathbf{e}$, and $Q^{-1} \in \mathbb{R}^{m\times m}$ is a symmetric, positive definite matrix, ideally representing the inverse covariance matrix of $(\mathbf{e} - \mathbf{e}^b)$.

Although it is known in the data assimilation literature that a background term is needed in some cases to give uniqueness to Problem $\mathcal{A}_{IS}$, the applications of Problem $\mathcal{A}_{CT}$ (ie, applications of the correction term technique) mentioned in Section 5.1 did not use a background term for the control vector. Wergen [87] added an extra term to the cost function which acted as a background term which constrained just one of the three model fields to be close to zero.

In analogy to the working of the previous subsection, the Hessian of $\mathcal{J}$ with respect to $\mathbf{e}$ is $(Q^{-1} + \tilde{A}_{CT})$, and the following result holds.

Theorem 5. *Problem $\mathcal{B}_{CT}$ has a unique solution.*

Proof

The Hessian matrix $(Q^{-1} + \tilde{A}_{CT})$ is the sum of a positive definite and a positive semi-definite matrix, and so is positive definite. Hence Problem $\mathcal{B}_{CT}$ has a unique solution. □

It could also be deduced from the results in Bennett and Miller's paper [8] on the importance of a background estimate of the initial state, that when terms representing model error are estimated, a background estimate for these terms is sufficient for uniqueness in the linear case.

5.2.3 Using both the initial state and the correction term as control vectors

We notice that the system (5.3),(5.4) can equivalently be written as

$$\mathbf{x}_{k+1} = A_k\mathbf{x}_k + B_k\mathbf{e}_k, \tag{5.51}$$

$$\mathbf{e}_{k+1} = \mathbf{e}_k, \qquad k = 0,..,N-1, \tag{5.52}$$

with observations

$$\mathbf{y}_k = C_k\mathbf{x}_k^t + \boldsymbol{\delta}_k, \qquad k = 0,..,N-1. \tag{5.53}$$

It is helpful to rewrite this system as the *augmented system*

$$\mathbf{w}_{k+1} = M_k\mathbf{w}_k, \tag{5.54}$$

$$\mathbf{y}_k = \tilde{C}_k\mathbf{w}_k^t + \boldsymbol{\delta}_k, \tag{5.55}$$

where $\mathbf{w}_k \in \mathbb{R}^{n+m}$, $M_k \in \mathbb{R}^{(n+m)\times(n+m)}$ and $\tilde{C}_k \in \mathbb{R}^{p_k\times(n+m)}$ are given by

$$\mathbf{w}_k = \begin{pmatrix} \mathbf{x}_k \\ \mathbf{e}_k \end{pmatrix}, \qquad M_k = \begin{pmatrix} A_k & B_k \\ 0 & I \end{pmatrix}, \qquad \tilde{C}_k = \begin{pmatrix} C_k & 0 \end{pmatrix}. \tag{5.56}$$

In this augmented system, $\mathbf{w}_k$ is the *augmented state vector* and $\mathbf{w}_0$ is the *augmented control vector*. The augmented state transition matrix is $\tilde{\Phi}(k,j)$, and we have

$$\tilde{\Phi}(k,0) = \tilde{\Phi}_k = \begin{pmatrix} \Phi_k & \tilde{B}_k \\ 0 & I \end{pmatrix}. \tag{5.57}$$

Problem $\mathcal{A}_{ISCT}$ can now equivalently be written

Problem $\mathcal{A}_{ISCT}$

Minimize with respect to $\mathbf{x}_0,..,\mathbf{x}_N;\mathbf{e}$

$$\mathcal{J} = \frac{1}{2}\sum_{j=0}^{N-1}(\tilde{C}_j\mathbf{w}_j - \mathbf{y}_j)^T R_j^{-1}(\tilde{C}_j\mathbf{w}_j - \mathbf{y}_j) \tag{5.58}$$

subject to (5.54).

In this form, we see that Problem $\mathcal{A}_{ISCT}$ is just Problem $\mathcal{A}_{IS}$ applied to the system (5.54), and so the theory of Section 5.2.1 applies here. From Theorem 5.1 we

know that Problem $\mathcal{A}_{ISCT}$ has a unique minimum if and only if the N-step observability matrix $\tilde{\mathcal{O}}_N$ has rank $(n+m)$, where

$$\tilde{\mathcal{O}}_N = \begin{pmatrix} \tilde{C}_0 \\ \tilde{C}_1\tilde{\Phi}_1 \\ \vdots \\ \tilde{C}_{N-1}\tilde{\Phi}_{N-1} \end{pmatrix}. \tag{5.59}$$

We now consider observability of the augmented system (5.54),(5.55) in terms of observability of the original system (5.3),(5.4). Observability of the augmented system concerns the ability to determine $\mathbf{w}_0$ or equivalently $\mathbf{x}_0$ and $\mathbf{e}$ from the observations, and observability of the original system concerns the ability to reconstruct $\mathbf{x}_0$ from the same observations. We now show that the following result holds.

Theorem 5.7 *Necessary conditions for Problem $\mathcal{A}_{ISCT}$ to have a unique solution are that the original system (5.3),(5.4) is completely N-step observable at time t_0 and that Rank$(\mathcal{I}_N) = m$.*

Proof

From Theorem 5.1, Problem $\mathcal{A}_{ISCT}$ has a unique solution if and only if Rank$(\tilde{\mathcal{O}}_N) = (n+m)$, with $\tilde{\mathcal{O}}_N$ given in (5.59).

Since

$\tilde{C}$ *y that,*

Hence, Rank$(\mathcal{O}_N) = n$, or equivalently by Theorem 2.2, complete N-step observability is a necessary condition for a unique solution of Problem $\mathcal{A}_{ISCT}$.

By a similar argument, we have that Rank$(\mathcal{T}_N) = m$ is also a necessary condition for a unique solution of Problem $\mathcal{A}_{ISCT}$. □

One simple example for which Problem $\mathcal{A}_{ISCT}$ has a unique solution is the case where C_0 and C_1 both have rank n, (as is the case in our experiments where we use the full set of observations), as we now show.

We suppose $\mathbf{v} \in \mathbb{R}^{n+m}$ is arbitrary and that $\tilde{\mathcal{O}}_N \mathbf{v} = 0$. Then by (5.62)

$$\mathcal{O}_N \mathbf{v}_1 + \mathcal{T}_N \mathbf{v}_2 = 0, \tag{5.63}$$

and by equation (5.60)

$$C_k \Phi_k \mathbf{v}_1 + C_k \tilde{B}_k \mathbf{v}_2 = 0, \qquad k = 0, .., N-1. \tag{5.64}$$

With $k = 0$ we have

$$C_0 \mathbf{v}_1 = 0, \tag{5.65}$$

and hence $\mathbf{v}_1$

where

$$M = \begin{pmatrix} A & B \\ 0 & I \end{pmatrix}, \qquad \tilde{C} = \begin{pmatrix} C & 0 \end{pmatrix}. \tag{5.70}$$

The following result is also applicable to the system we use in the experiments we describe in Section 5.3, and we use it later.

Theorem 5.8 *For the time invariant system (5.68), (5.69) with $m = n$ and B nonsingular, Problem $\mathcal{A}_{ISCT}$ has a unique solution if and only if Rank$(C) = n$.*

Proof

By Theorem 5.1, Problem $\mathcal{A}_{ISCT}$ has a unique solution if and only if the augmented system (5.68), (5.69) is completely N-step observable. We show that if Rank$(C) < n$, the system (5.68), (5.69) is not completely observable, and hence not completely ν-step observable for any ν by Theorem 2.3 Part a.

By the negation of the Hautus condition (Theorem 2.4), the system is not completely observable if there exists a non-zero vector $\mathbf{v} \in \mathbb{R}^{2n}$ and $\lambda \in \mathbb{C}$ such that

$$(M - \lambda I)\mathbf{v} = 0, \tag{5.71}$$

$$\tilde{C}\mathbf{v} = 0, \tag{5.72}$$

or equivalently

$$(A - \lambda I)\mathbf{v}_1 + B\mathbf{v}_2 = 0, \tag{5.73}$$

$$(I - \lambda I)\mathbf{v}_2 = 0, \tag{5.74}$$

$$C\mathbf{v}_1 = 0, \tag{5.75}$$

where $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$.

Suppose that Rank$(C) < n$. Then there exists a non-zero $\mathbf{v}_1$ such that $C\mathbf{v}_1 = 0$. Let $\mathbf{v}_2$ be given by

$$\mathbf{v}_2 = B^{-1}(A - \lambda I)\mathbf{v}_1, \tag{5.76}$$

and $\lambda = 1$. Then (5.73)– (5.75) or equivalently (5.71), (5.72) hold for a non-zero vector $\mathbf{v}$, and $\lambda = 1$. Hence, if Rank$(C) < n$ the augmented system is not completely observ

The fact that Problem $\mathcal{A}_{ISCT}$ has a unique solution if $\text{Rank}(C) = n$ follows from our previous remarks which showed that this is true for the general time-varying system with $\text{Rank}(C_0) = n$, $\text{Rank}(C_1) = n$. □

By Theorem 5.2, we know that we can guarantee a unique solution to Problem $\mathcal{A}_{ISCT}$ by adding a background estimate of the augmented control vector $\mathbf{w}_0$ to the cost function, and so we formulate the following problem.

Proof

We note that (5.78) can equivalently be written

$$\mathcal{J} = \frac{1}{2} \sum_{j=0}^{N-1} (D_j \mathbf{w}_j - \mathbf{z}_j)^T S_j^{-1} (D_j \mathbf{w}_j - \mathbf{z}_j), \tag{5.79}$$

where $D_j \in \mathbb{R}^{(p_j+m)\times(n+m)}$, $S_j^{-1} \in \mathbb{R}^{(p_j+m)\times(p_j+m)}$ and $\mathbf{z}_j \in \mathbb{R}^{p_j+m}$ are given by

$$D_j = \begin{pmatrix} C_j & 0 \\ 0 & I \end{pmatrix}, \qquad$$

$N\,1$

or equivalently

$$\mathcal{O}_N \mathbf{v}_1 = 0. \tag{5.87}$$

Hence, there exists a non-zero vector $\mathbf{v} \in \mathbb{R}^{n+m}$ such that $\hat{\mathcal{O}}_N \mathbf{v} = 0$ if and only if there exists a non-zero vector $\mathbf{v}_1 \in \mathbb{R}^n$ such that $\mathcal{O}_N \mathbf{v}_1 = 0$. It follows that $\hat{\mathcal{O}}_N$ has rank $(n+m)$ if and only if $\mathcal{O}_N$ has rank n, which proves the result. □

In Section 5.5, we present a summary of the theoretical results in this chapter. We now compare the performance of the initial state, correction term and both together as control vectors practically, in experiments with a simple model.

5.3 Description of the experiments

The aim of these experiments is to compare the performance of variational assimilation using the different control vectors: the initial state, the correction term and the augmented control vector containing the initial state and the correction term.

The true model state

We suppose that the true model state $\mathbf{x}_k^t$ satisfies (5.88) started from the true initial state $\mathbf{x}_0^t$ is given by

$$(x_j^0)^t = 1, \qquad j = 1, .., n. \tag{5.89}$$

Model error

As a source of model error, we suppose that the constant source term is omitted from the model equations. Hence, in the "imperfect" model, $\mathbf{s}$ is set to zero.

Observations

We suppose that we have error free observations at p of the 15 grid points at every timestep on the interval $[0, 0.$

subject to (5.91), where $R^{-1} = \frac{2}{N} I \in \mathbb{R}^{p \times p}$, and $Q^{-1} = qI \in \mathbb{R}^{n \times n}$. The matrices R^{-1} give equal weight to all observations, and are not supposed to represent error covariances. The value of q is sometimes taken to be zero, in which case we do not constrain the size of the correction term to be small.

The adjoint model is

$$\boldsymbol{\lambda}_k = A^T \boldsymbol{\lambda}_{k+1} - C^T R^{-1} (C\mathbf{x}_k - \mathbf{y}_k), \qquad k = \frac{N}{2} - 1, .., 0, \tag{5.93}$$

with

$$\boldsymbol{\lambda}_{\frac{N}{2}} = 0. \tag{5.94}$$

The gradients of the Lagrange function $\mathcal{L}$ associated with $\mathcal{J}$ with respect to the control vectors are

$$\nabla_{\mathbf{x}_0} \mathcal{L} = -\boldsymbol{\lambda}_0, \tag{5.95}$$

$$\nabla_{\mathbf{e}} \mathcal{L} =$$

where the control vector $\mathbf{u}$ might be the initial state, correction term or the aug-

Calculating $< \mathbf{r}^{k+1}, \tilde{A}\mathbf{d}^k >$

From (5.99) and (5.101) we have

$$\tilde{A}\mathbf{d}^k = (\mathbf{r}^k - \mathbf{r}^{k+1})/\rho^k. \tag{5.104}$$

Everything on the right hand side of (5.104) is known by the time we need to evaluate $< \mathbf{r}^{k+1}, \tilde{A}\mathbf{d}^k >$, so $\tilde{A}\mathbf{d}^k$ can be evaluated easily.

Calculating $< \mathbf{d}^k, \tilde{A}\mathbf{d}^k >$

This expression can be built up using the following iteration. Starting from

$$\boldsymbol{\xi}_0 = \mathbf{d}^k, \tag{5.105}$$

$$\boldsymbol{\zeta}_0 = R^{-1}C\boldsymbol{\xi}_0, \tag{5.106}$$

$$\boldsymbol{\sigma}_0 = \boldsymbol{\zeta}_0^T\boldsymbol{\zeta}_0, \tag{5.107}$$

for $i = 1, .., \frac{N}{2} - 1$ we let

$$\boldsymbol{\xi}_i = \boldsymbol{\psi}_i(\boldsymbol{\xi}_{i-1}), \tag{5.108}$$

$$\boldsymbol{\zeta}_i = R^{-1}C\boldsymbol{\xi}_i, \tag{5.109}$$

$$\boldsymbol{\sigma}_i = \boldsymbol{\sigma}_{i-1} + \boldsymbol{\zeta}_i^T\boldsymbol{\zeta}_i, \tag{5.110}$$

where

$$\boldsymbol{\psi}_i(\boldsymbol{\xi}_{i-1})$$

Case a: Perfect model, unknown initial state

In this case, the first guess of the initial state is

$$x_j^0 = 2, \qquad j = 1, .., n, \tag{5.115}$$

rather than (5.89).

Case b: Imperfect model, known initial state

The source term **s** is assumed to be zero in this imperfect model, but this time the true initial state is known.

Case c: Imperfect model, unknown initial state

Here we use the imperfect model of Case **b** with the first guess initial state specified in Case **a**.

Data assimilation is carried out over the time interval $[0, 0.5]$, where observations are available. We then suppose that no more observ

Experiment 1b: Imperfect model, known initial state

In this case we still expect that the minimization problem has a unique solution, but the optimal initial state found will not be the true one, but that which gives the solution of the imperfect model which is closest to the observations.

The true initial state is known, and the minimization iterations are started from this value. If a different start guess for the initial state is used, the same results are found, but this generally takes a couple of iterations more. Fig. 5.3 shows that when the full set of observations are used ($p = 15$), a smooth initial state is found with a higher value at the position of the source point. If fewer observations are used (Fig. 5.4 shows the results when $p = 5$), the initial state obtained is no longer smooth, but matches the true solution at the observation positions. However, the model dissipation soon acts to smooth out the solution.

The best results in each case are in the middle of the assimilation interval, at $t = 0.25$. This tendency of the variational assimilation method to give a closest fit to observations in the middle of the assimilation interval has often been noted [19]. In this case of an imperfect model, the method does not produce the true initial state but finds one for which the ensuing solution throughout the assimilation interval is closer to the observations. In this way, the effects of model error, rather than building up in time, have been spread throughout the assimilation interval, as noted in [87]. Although the assimilation has improved the results at $t = 0.5$, the benefits at $t = 1$ are much smaller, because the forecast has been carried out with an uncorrected imperfect model.

5.4.2 Experiments using the correction term as the control vector

Here we carry out data assimilation using the correction term as a control vector, and with $\mathbf{x}_0$ fixed. Comparison of these results with those using the initial state as the control vector gives a comparison of the efficiency of the two control vectors in each situation. We also examine the use of the background term $\frac{1}{2}\mathbf{e}^T q I \mathbf{e}$ in the cost function, with different values of q. We look first at the situation of the imperfect

model with known initial state, since this problem is more naturally treated by using the correction term as a control vector.

Experiment 2b: Imperfect model, known initial state

Since the system is completely observable and $\frac{N}{2} > n+1$, we know that Problem $\mathcal{A}_{CT}$ has a unique solution. Since in Case **b** the initial state is known and model error is constant in time, this method should find a correction term which exactly represents the model error, and hence be able to reproduce the true solution.

Fig. 5.5 shows this to be so when $p = 5$ using $q = 0$. If the correction term found is used in the forecast started at the end of the assimilation interval ($t = 0.5$), the forecast using this corrected model matches the true solution exactly. The number of iterations taken in this case is 17, very similar to the number taken in Experiment 1a using the initial state as the control vector; and as in that case, the number of iterations decreases when fewer observations are used. When only one observation is used with $q = 0$, (Fig. 5.6), the results are still good.

Experiments were also carried out using the background term with $q = 1$. For a given value of p, fewer iterations were needed to satisfy the stopping criterion than using $q = 0$, but the results were very slightly less accurate.

Experiment 2a: Perfect model, unknown initial state

Again, we expect a unique solution, but not the true solution. The method will find the correction term for which the model started from a wrong initial state is as close as possible to the observations.

In these experiments the correction term is not included in the forecast following the assimilation. This is because the correction term is supposed to compensate for the errors in the initial conditions throughout the assimilation period, and since the model is perfect, an improved solution at the end of the assimilation interval (at $t = 0.5$) will give an improved forecast.

If $p = 15$, ie, the full set of observations are used, (Fig. 5.7), the solution is closest to the true solution about halfway through the assimilation interval. At the end of the assimilation interval, it is hard to judge whether the assimilation has

which gives the benefits obtained using each of the control vectors separately, while keeping the extra cost at an acceptable level. As illustrated above, there are some situations in which using the initial state as the control vector, and other situations in which using the correction term as the control vector works particularly well. If we do not know before starting the assimilation which control vector is better for the situation, we would like to be able to use both together and obtain the same benefits as if the preferable control vector had been chosen. This is examined here by looking at Cases **a** and **b** using the augmented control vector. We then look at Case **c**, the more general situation of an imperfect model with unknown initial state, to see whether by using the augmented control vector we can obtain the true solution in this case.

From the theory of Section 5.2, we kno

perform many more iterations of the descent algorithm than using just one control vector. In this case, 90 iterations are used in the case $p = 5$, though fewer iterations are needed when fewer observations are used, (26 iterations for $p = 1$).

Increasing q reduces the number of iterations required, however. These results show that when both control vectors are used, the background term is essential for sensible results, and that an appropriate choice of q is important to save extra cost.

not have a unique solution using $q = 0$ and $p < 15$.

Adding the background term (using $q > 0$), ensures uniqueness and results in good solutions, although for smaller values of p the results are not completely smooth at the beginning of the assimilation interval. Figs. 5.12 and 5.14 show the results for $p = 5$ and $p = 10$, respectively, with $q = 1$. In all cases the match to the true solution is very good at the end of the assimilation, and the forecast initiated at this time using the correction term found in the assimilation maintains a perfect match to the true solution.

Increasing the value of q to 10 gives a smoother solution for smaller values of p, and reduces the number of iterations needed. How

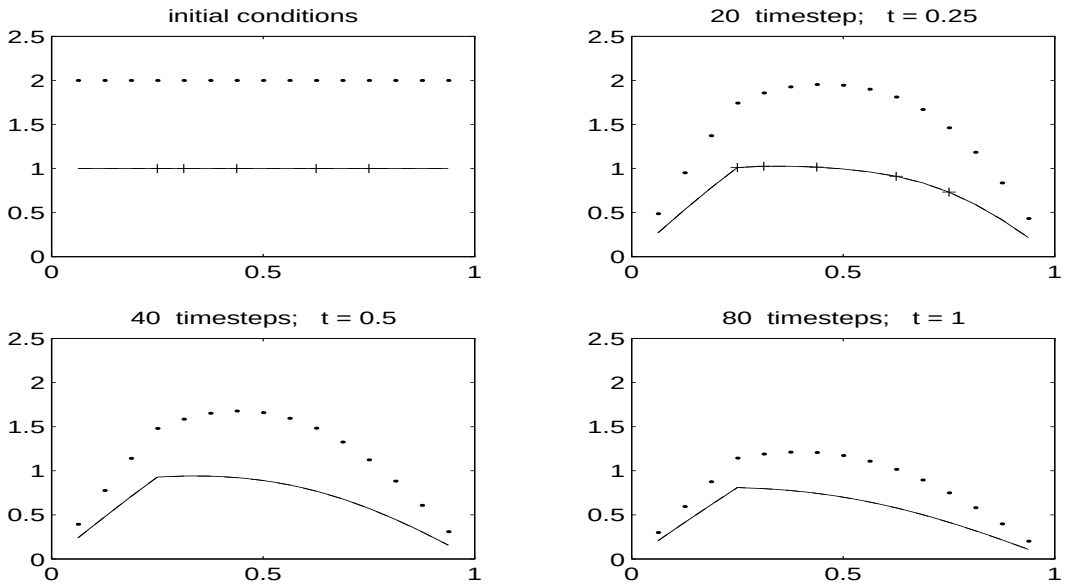
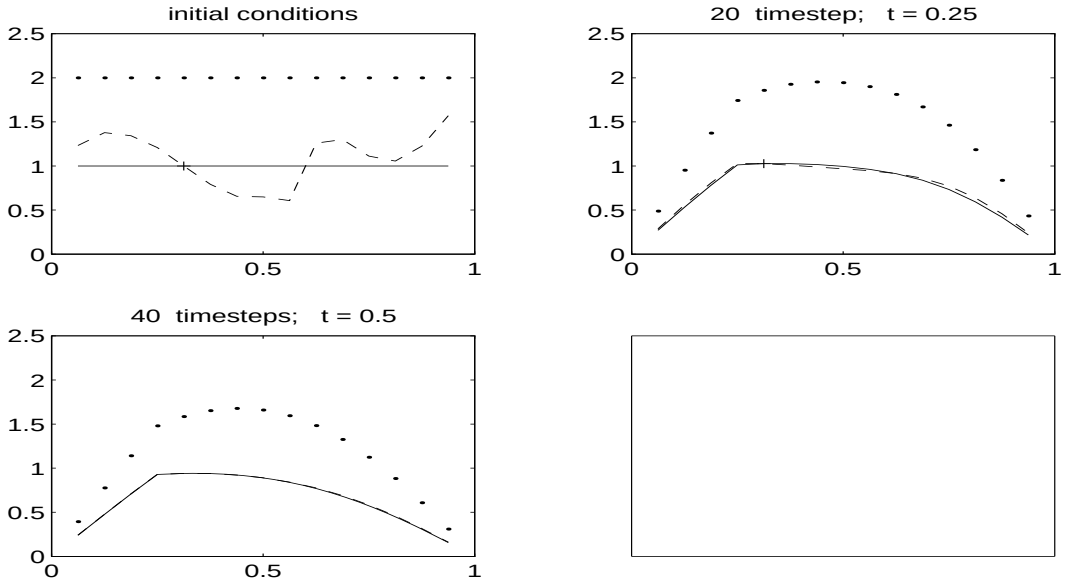
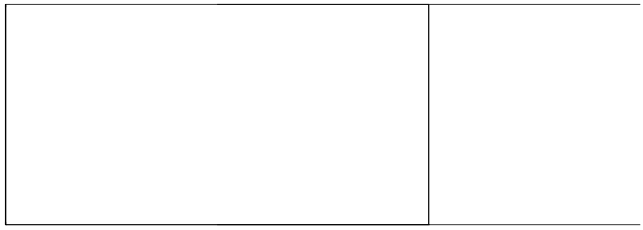
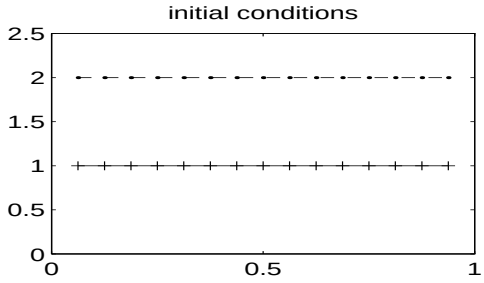


Figure 5.1: Variational assimilation using the initial state as the control vector. Assimilation on the interval $t \in [0, \frac{1}{2}]$ using 5 observations, followed by a forecast on the interval $t \in [\frac{1}{2}, 1]$. Solid line: true solution; dotted line: background solution (no assimilation); dashed line: solution with assimilation; crosses: observations.

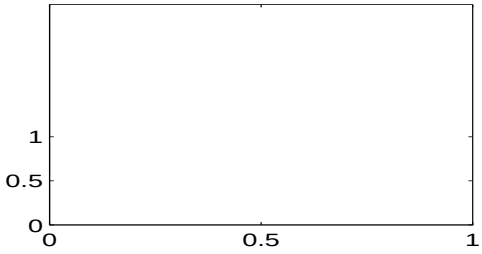


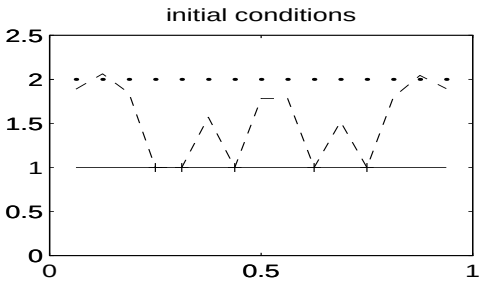


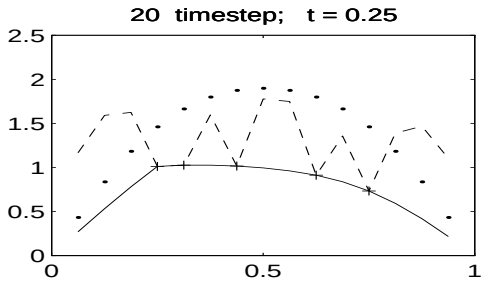
o

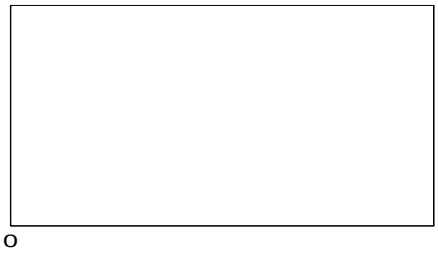

initial conditions
2.5
2
1.5
1
0.5
0
0
0.5
1


0


1
0.5
0
0
0.5
1


initial conditions
2.5
2
1.5
1
0.5
0
0
0.5
1


20 timestep; t = 0.25
2.5
2
1.5
1
0.5
0
0
0.5
1


0

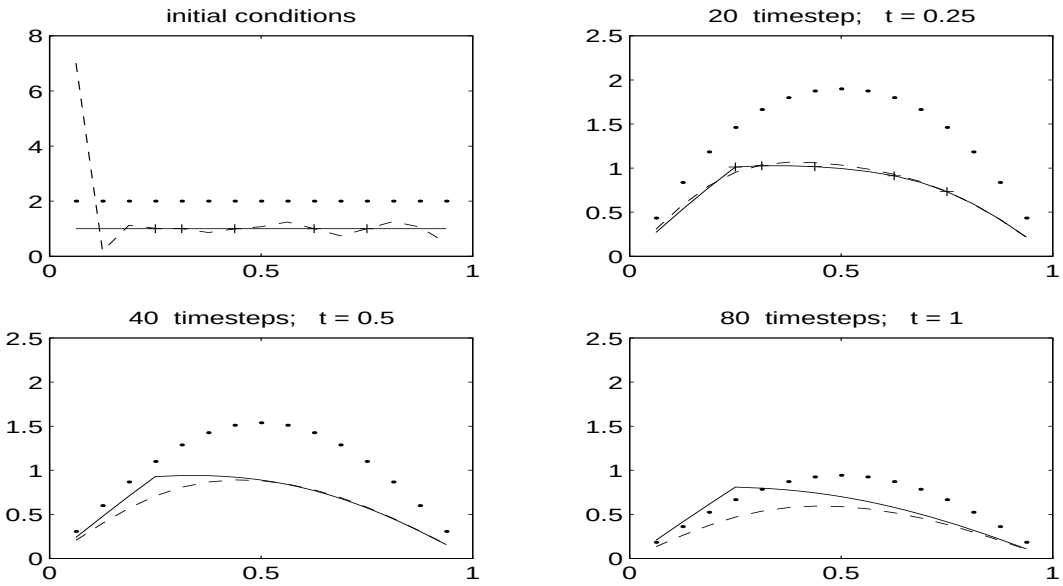
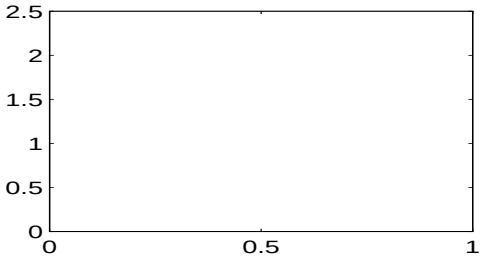


Figure 5.13: As Fig. 5.11, but using $q = 50$.



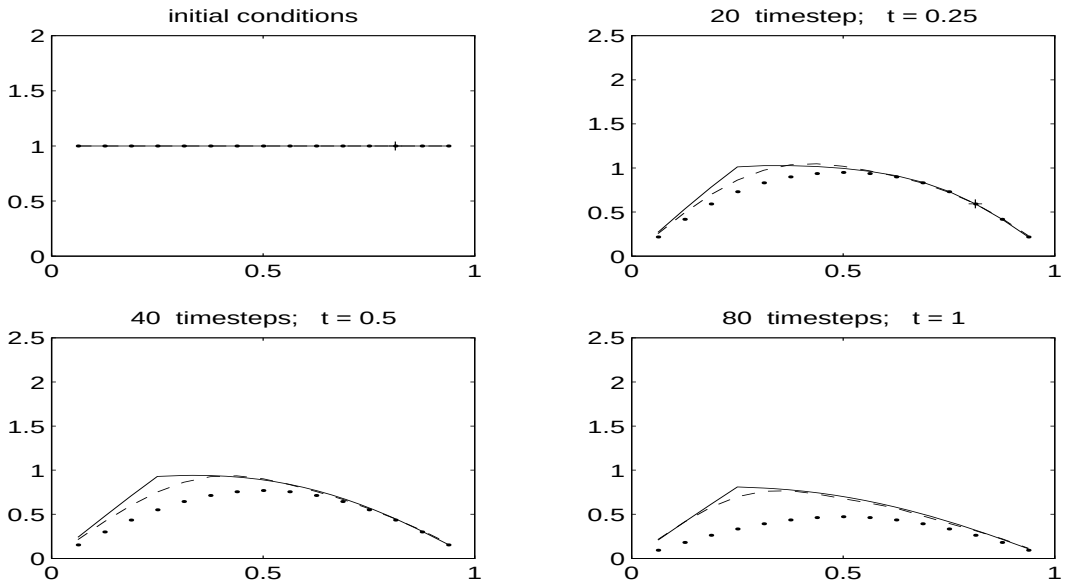


Figure 5.15: Variational assimilation using the correction term as the control vector, with $q = 0$. Assimilation on the interval $t \in [0, \frac{1}{2}]$ using 1 observation, followed by a forecast on the interval $t \in [\frac{1}{2}, 1]$. Solid line: true solution; dotted line: background solution (no assimilation); dashed line: solution with assimilation; crosses: observations.

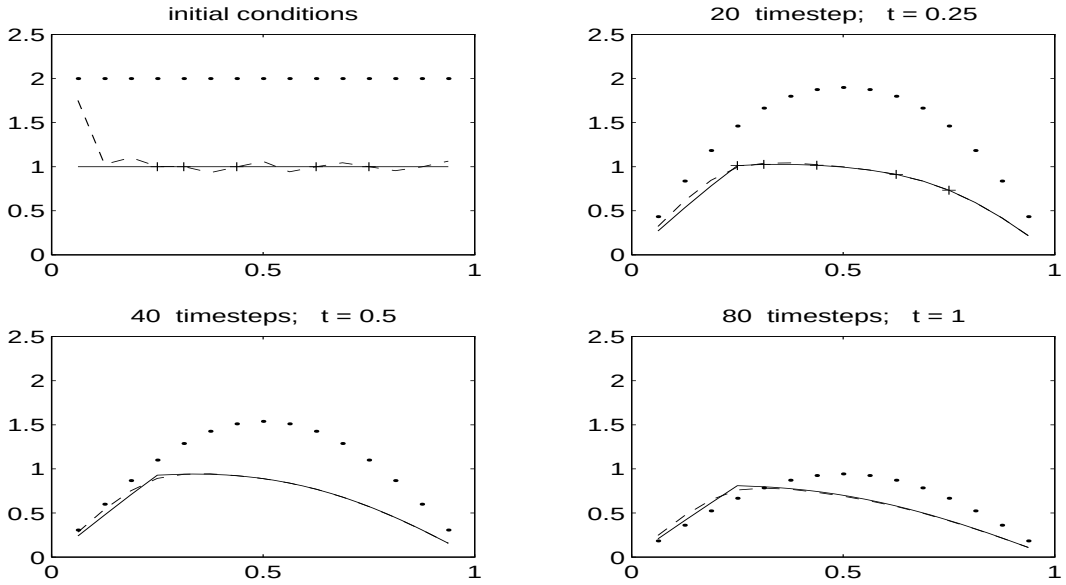


Figure 5.17: Variational assimilation using both the initial state and the correction term as the control vectors, with $q = 1$. Assimilation on the interval $t \in [0, \frac{1}{2}]$ using 5 observations, followed by a forecast on the interval $t \in$ terval $t \in$

5 observations,f1bΦΦΦΦoΦ:ΦTDΦions, follow

5.5 Summary and conclusions

In this chapter we have looked at both theoretical and practical aspects of using the correction term as the control vector compared with using the initial state as the control vector, and also of using both control vectors together. Here we briefly summarize the theoretical results of Section 5.2, and giv

the correction term should not be included in an ensuing forecast.

In all experiments with unknown initial state in which the correction term was used as a control vector, using a background term with a large weight constraining the correction term to be small (ie, a large value of q) was vital for sensible results

tion. The first is the problem of reducing the number of iterations needed when both control vectors are used. This could be achieved by suitable preconditioning of the descent process, but we do not take this any further in this work. Secondly, the question arises of ho

Chapter 6

ccounting for model error in variational assimilation

We start this chapter with a discussion on the problem of how to account for model error in data assimilation. In particular, we note that the assumptions made on

evolving correction term as a control variable produces a significant improvement in the results.

The theoretical part of our work described in this chapter has been published in a shorter form in [41].

6.1 Background on representing model error

In 4D variational assimilation using the strong constraint approach, which is currently being developed for operational application in meteorological centres, model error is neglected. Recently, however, the problem of how to account for model error in variational assimilation in a cost effective way has begun to receive more attention [61].

Studies in predictability which explicitly attempt to represent the effects of model error on forecast error, [12],[22],[85] show that the impact of model error on forecast error in meteorological models is indeed significant. The study in [22] leads to the conclusion that the predictability limit of a forecast might be extended by two or three days if model error were eliminated. However, there is a lack of quantative information on model error in such forecast models, even of its size relative to that of the model state. Hence, the problem arises of how to represent model error in data assimilation.

The Kalman filter does account for model error, and in the Standard Kalman filter model error is treated as serially uncorrelated, unbiased random error. In Chapter 4, Section 4.4 we also discussed other approaches to weak constraint variational assimilation which make the same assumptions about model error. An interesting paper by Dee [25] however, questions the validity of this representation of model error. Using an analysis of model error similar to that given in Chapter 2, Subsection 2.1.2 here, he argues that since model error in general depends on the model state, it is likely to be serially correlated. In Chapter 3, Section 3.2 we gave background on how the Kalman filter can be modified to deal with serially correlated model error, but this involves a large increase in the expense of the method, which (unless in simplified form) is already thought to be too expensive for operational

a serially correlated component of model error along with the model state. In Section 6.3 we giv

which is made up of serially correlated errors and serially uncorrelated random errors. We therefore write

$$\boldsymbol{\varepsilon}_k = B_k \mathbf{e}_k + \mathbf{q}'_k, \tag{6.3}$$

where the vectors $\mathbf{q}'_k$ are serially uncorrelated, random n-vectors, the matrices $B_k \in \mathbb{R}^{n\times m}$ are prescribed matrices as before, and the vectors $\mathbf{e}_k \in \mathbb{R}^m$ represent the *serially correlated component of model error*. We suppose that we know how the error $\mathbf{e}_k$ evolves in time, and for now write this in a very general form,

$$\mathbf{e}_{k+1} = \mathbf{g}_k(\mathbf{x}_k, \mathbf{e}_k) + \mathbf{q}''_k, \tag{6.4}$$

where $\mathbf{g}_k : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is some function to be specified, and the vectors $\mathbf{q}''_k$ are serially uncorrelated random m-vectors.

As we discussed in Section 6.1, we know very little about the form of the model error, and in practice will have to specify (6.3), (6.4) in a very simple form which reflects any knowledge of model error we do have. We give

6.2.2 Examples of how model error can be specified

i) Serially uncorrelated model error

Setting all the $\mathbf{e}_k$ in (6.3) to zero we have

$$\boldsymbol{\varepsilon}_k = \mathbf{q}'_k, \tag{6.8}$$

in which case model error is a serially uncorrelated random vector, as assumed in the Standard Kalman filter.

ii) Constant model bias error

Setting

$$\boldsymbol{\varepsilon}_k = B_k \mathbf{e}_k + \mathbf{q}'_k, \tag{6.9}$$

$$\mathbf{e}_{k+1} = \mathbf{e}_k, \tag{6.10}$$

allows for a constant vector of unknown "dynamical parameters" as discussed in [44]. If this form of model error is purely deterministic (ie, $\mathbf{q}'_k = 0$), this represents the correction term technique of Chapter 5. In Derber's paper [26] introducing the correction term technique, the matrices B_k were the $n \times n$ identity multiplied by a time-varying scalar and by the time-step length Δt, to reflect the rôle of this form of model error as a correction to the time derivative of the model equations. As discussed in Chapter 5, we expect this form of model error to be appropriate for representing constant errors in the forcing or in the boundary conditions.

iii) Model error evolving with model evolution

In Section 6.1, we discussed that model error is likely in general to depend on the true model state, and hence to change with the flow. In this case model error evolution might be approximated by

$$\boldsymbol{\varepsilon}_k = B_k \mathbf{e}_k + \mathbf{q}'_k, \tag{6.11}$$

$$\mathbf{e}_{k+1} = G_k \mathbf{e}_k, \tag{6.12}$$

where $G_k \in \mathbb{R}^{m \times m}$ represents a simplified form of the model state evolution. This might be an appropriate approximation to model error evolution if model error

represents truncation error. This is similar to the form of serially correlated model error which was suggested in the paper by Daley [23] in formulating a Kalman filter allowing for serially correlated model error, which we discussed in Chapter 3, Section 3.2.

The matrices B_k in (6.3) allow for a serially correlated component of model error with dimension m which may be less than or greater than the dimension n of the model state. In Chapter 5 we showed how using $m < n$ can lead to greater efficiency in the correction term technique if the source of model error is known to be localized. We now consider how including the possibility that $m > n$ can allow for greater flexibility in the specification of model error.

We may partition the serially correlated component of model error in r subvectors of dimension s (where $rs = m$), and write

$$\boldsymbol{\varepsilon}_k = B_k \mathbf{e}_k + \mathbf{q}'_k, \tag{6.13}$$

$$\mathbf{e}_{k+1} = \mathbf{g}_k(\mathbf{w}_k) + \mathbf{q}''_k, \tag{6.14}$$

where this time

$$B_k = (B_k^{(1)}, B_k^{(2)}, .., B_k^{(r)}), \qquad \mathbf{e}_k = \begin{pmatrix} \mathbf{e}_k^{(1)} \\ \vdots \\ \mathbf{e}_k^{(r)} \end{pmatrix}, \tag{6.15}$$

where $\mathbf{e}_k^{(1)}, .., \mathbf{e}_k^{(r)} \in \mathbb{R}^s$ and $B_k^{(1)}, .., B_k^{(r)} \in \mathbb{R}^{n \times s}$. The following examples illustrate how this generalization might be useful.

iv) Model error growing in time

Here, rather than using a constant correction term to represent model error as in the correction term technique, we allow for a correction term which can increase or decrease linearly in time. In this case model error has the form

$$\boldsymbol{\varepsilon}_k = B_k \mathbf{e}_k + \mathbf{q}'_k, \tag{6.16}$$

$$\mathbf{e}_{k+1} = \mathbf{e}_k, \tag{6.17}$$

vii) Piecewise constant model error

Here we suppose that the assimilation interval $[t_0, t_N]$ is broken into r subintervals over whic

the evolution of the augmented state vector. We suppose that as before, w

of including a background term for $\mathbf{e}_0$ in the cost function.

where $\boldsymbol{\lambda}_k \in \mathbb{R}^n$, $\boldsymbol{\mu}_k \in \mathbb{R}^m$, and where $G_k \in \mathbb{R}^{m\times n}$ is the Jacobian of $\mathbf{g}_k$ with respect to $\mathbf{x}_k$, and $\Gamma_k \in \mathbb{R}^{m\times m}$ is the Jacobian of $\mathbf{g}_k$ with respect to $\mathbf{e}_k$. With

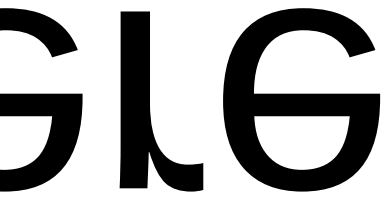
$$\tilde{P}_0^{-1} = \begin{pmatrix} P_0^{-1} & 0 \\ 0 & Q_0^{-1} \end{pmatrix}, \tag{6.42}$$

where $P_0 \in \mathbb{R}^{n\times n}$ and $Q_0 \in \mathbb{R}^{m\times m}$ are the covariance matrices of $(\mathbf{x}_0 - \mathbf{x}_0^b)$ and $(\mathbf{e}_0 - \mathbf{e}_0^b)$ respectively, equation (6.37) becomes

$$\nabla_{\mathbf{x}_0}\mathcal{L} = P_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) - \boldsymbol{\lambda}_0, \tag{6.43}$$

$$\nabla_{\mathbf{e}_0}\mathcal{L} = Q_0^{-1}(\mathbf{e}_0 - \mathbf{e}_0^b) - \boldsymbol{\mu}_0, \tag{6.44}$$

with $\boldsymbol{\lambda}_0$ and $\boldsymbol{\mu}_0$ defined bΦ??x–Tj11?P–Tj1TΦΦsΦTD1?fi–Tj1Tad)ΦTD1?g–Φa–ΦTf11Ta91TΦΦ



extra control vector $\mathbf{e}_0$ and its gradient, although the dimension m of these might be much less than n. The increased dimension of the augmented control vector also means that the part of the descent algorithm which uses the gradient information to improve a guess of the control vector will be more expensive. A larger problem, however, is that the conditioning of the problem using the augmented control vector approach will be altered, and as a result, more iterations and hence more model and adjoint runs may be needed, as we found in the experiments of Chapter 5 using both control vectors.

6.4 Using n evolving correction term

6.4.1 Introduction

In the experiments of Chapter 5, we saw that the correction term technique is successful in correcting for model error which behaves like a constant forcing term. Here, we consider the upwind discretization of the linear advection equation in which model error is present due to dissipation. The model error can be expressed as truncation error, and since this depends on the true model state, it will change in time with the model state. In this section, we consider the generalized correction term technique supposing that the correction term representing model error evolves with the model equations.

The model has the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \tag{6.49}$$

and we try to compensate for model error using an *evolving correction term* $\mathbf{e}_k \in \mathbb{R}^n$, where

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{e}_k, \tag{6.50}$$

$$\mathbf{e}_{k+1} = A\mathbf{e}_k. \tag{6.51}$$

We consider using the initial state $\mathbf{x}_0$, the initial correction $\mathbf{e}_0$ and both together as control vectors. We note that when the initial correction $\mathbf{e}_0$ is used as a control vector, the dimension of the augmented model system and its adjoint is twice that of the original system and its adjoint. We noted in the previous section that this can

be av

The scheme can be written as a matrix system as follows,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \tag{6.58}$$

in which $\mathbf{x}_k \in \mathbb{R}^n$ is the state at time t_k, where $n = J$, and $A \in \mathbb{R}^{n\times n}$ is given by

$$A = \begin{pmatrix} (1-\mu) & 0 & & \mu \\ \mu & (1-\mu) & 0 & \\ \ddots & \ddots & \ddots & \\ 0 & & \mu & (1-\mu) \end{pmatrix}. \tag{6.59}$$

The upwind scheme is first order accurate and stable provided $\mu \leq 1$.

We run the model (6.58) with $c = 1$, using $N = 80$, $J = 40$, so $\Delta t = \frac{1}{80}$ and $\Delta z = \frac{1}{40}$. Hence, $\mu = \frac{1}{2}$ and the model state has dimension $n = 40$. Since $c = 1$, the square wave represented by the initial conditions is advected all the way round the model domain to its starting position on the time interval $[0, 1]$.

The true model state

With $\mu = 1$, the upwind discretization yields the true solution of the pde (6.52) on the model grid, ie, there is no model error. So, to compute the true model state $\mathbf{x}_k^t$ on the model grid specified above, we used the model (6.58) with $\mu = 1$, choosing $\Delta t = \frac{1}{80}$, $\Delta z = \frac{1}{80}$, and with initial conditions (6.54).

Observations

We suppose that we have error free observations at p of the 40 grid points at every timestep on the interval $[0, \frac{1}{2}]$, ie for $\frac{N}{2} = 40$ timesteps, and that after this no further observations are available. Hence, the observations are given by

$$\mathbf{y}_k = C\mathbf{x}_k^t, \qquad k = 0, ..., \frac{N}{2} - 1, \tag{6.60}$$

where the observational matrix $C \in \mathbb{R}^{p\times n}$ has a simple form since the observation positions coincide with the grid points. The positions of the observations used in each case are shown in the figures.

The minimization algorithm

The minimization algorithm used is the conjugate gradient descent method, implemented as described in Chapter 5, Section 5.3.

The experiments

We minimize the cost functional

$$\mathcal{J} = \frac{1}{2}\mathbf{e}_0^T Q_0^{-1}\mathbf{e}_0 + \frac{1}{2}\sum_{j=0}^{\frac{N}{2}-1}(C\mathbf{x}_j - \mathbf{y}_j)^T R^{-1}(C\mathbf{x}_j - \mathbf{y}_j), \tag{6.61}$$

subject to

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{e}_k, \tag{6.62}$$

$$\mathbf{e}_{k+1} = A\mathbf{e}_k, \qquad k = 0, .., \frac{N}{2} - 1, \tag{6.63}$$

where $R^{-1} = \frac{2}{N} \in \mathbb{R}^{p\times p}$, and $Q_0^{-1} = qI \in \mathbb{R}^{n\times n}$. As inI(;5and

no erroneous spikes in the data sparse areas. Fig. 6.5 shows the good results obtained using 20 observations, and

6.5.2 Case b): Imperfect model, unknown initial state

Using the initial state as a control vector

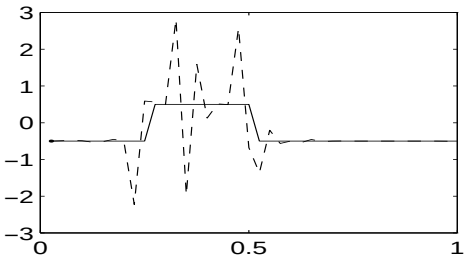
The results using the initial state as the control vector are the same whether the true initial state is known or not, but the minimization procedure requires a few more iterations if it is not known.

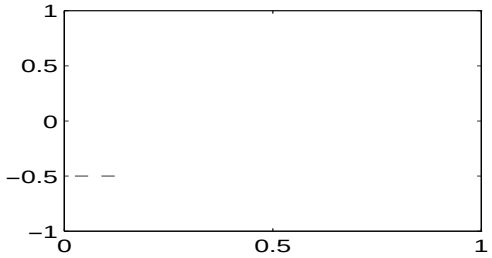
Using an evolving correction term as the control vector

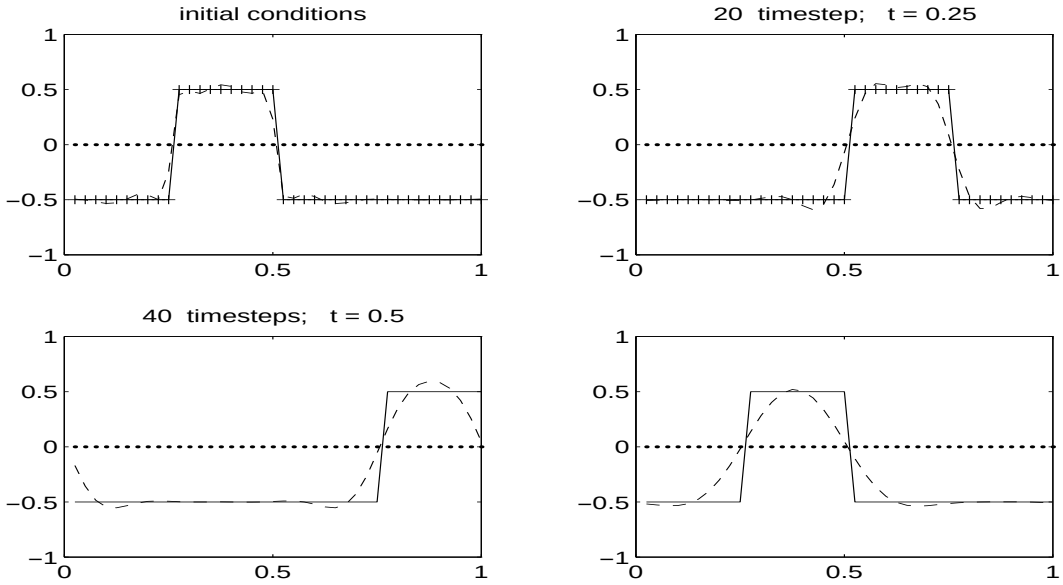
If the first guess of the unknown initial state is taken to be zero, the evolving correction term must make up for very large errors. Fig. 6.9 shows the results using the full set of observations, and Fig. 6.10 shows the results using $p = 20$. The solution produced is the right shape, under-estimating the true solution in the first half of the assimilation interval, and over-estimating it in the second half. In these experiments it is not appropriate to include the evolving correction term in the forecast period.

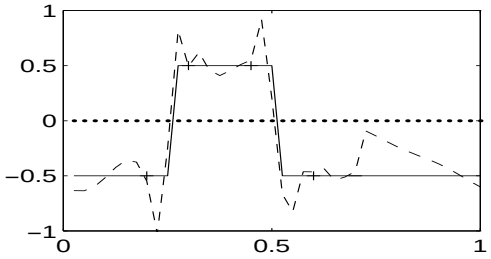
Using both control vectors

Fig. 6.11 shows the solution produced using the full set of observations and $q = 1$. However, the stopping criterion had not been reached when the minimization was terminated after 100 iterations. The solution improves on the solution started from the true initial state, and on the solution obtained using the initial state only as the control vector. Using the evolving correction term in the forecast reduces the effect of model error as before. If the value of q is increased to 10, the initial state control vector appears to have too much influence, and the solution is less accurate at the initial time. Using larger values of q also results in a deterioration of the quality of the forecast, and does not succeed in reducing the number of iterations of descent algorithm to less than 100. Fig. 6.12 shows that fairly good results are still achieved using 20 observations. Fig. 6.13 shows that the results using only 5 observations are poorer, but still an improvement on the background solution.


3
2
1
0
−1
−2
−3
0
0.5
1


1
0.5
0
−0.5
−1
0
0.5
1


initial conditions
20 timestep; t = 0.25
40 timesteps; t = 0.5
1
0.5
0
−0.5
−1
0
0.5
1


1
0.5
0
−0.5
−1
0
0.5
1

6.6 Summary and conclusions

Summary of the theory

We began this chapter with a discussion on why it is important to account for model error in data assimilation, and on some of the limitations of the approach taken to model error in the standard Kalman filter. We then considered a general representation of model error made up of serially correlated and serially uncorrelated components, and gave examples of different representations of model error which might be suitable in different situations. We suggested that the technique of state augmentation could be used in data assimilation to estimate the serially correlated components of model error along with the model state. This leads to a generalization of the least squares problem of Chapter 4 to deal with serially correlated model error.

The correction term technique can be interpreted as giving an optimal solution to this general data assimilation problem in the case that model error is a constant bias error with no sequentially uncorrelated component. We also suggested a generalization of the correction term tec

Chapter 7

Experiments with a shallow water model

Here we describe experiments using a 1D nonlinear shallow water model which includes topography and rotation. In these experiments we aim to investigate to what extent the conclusions of the experiments in Chapters 5 and 6 hold in the context of more complex dynamics. We compare using the initial state, a constant correction term and both together as control vectors. In particular, we aim to see whether the constant correction term can compensate for model error on a significant timescale, when the model error depends on the model state, and hence changes in time. Further, since the correction term technique involves changing the model equations by adding on the correction term, we want to check that the correction term produced in the assimilation does in fact represent an approximation of model error. These experiments are carried out in an idealized context with a full set of observations which are not corrupted by noise. We also begin to look at the situation in which fewer, noisy observations are available. Finally, we check whether assimilation using the correction term technique can result in a better forecast than assimilation using the initial state as the control variable. This is important to check, because in Wergen's study [87] using the correction term technique produced good results during an assimilation interval, but had a detrimental impact on the ensuing forecast. A briefer description of the results from these experiments has been published in [41].

for $x \in [0, 2\pi L]$, $t \in [0, T]$. The discretization is carried out with $\Delta x = \frac{1}{J}$, $\Delta t = \frac{1}{N}$, with discrete variables approximating the continuous variables as follows,

$$\phi_j^k \approx \phi(j\Delta x, k\Delta t), \quad u_j^k \approx u(j\Delta x, k\Delta t), \quad v_j^k \approx v(j\Delta x, k\Delta t), \tag{7.7}$$

for $k = 0, .., N$, $j = 0, .., J-1$. The discretization uses centred time and space differencing, except for the diffusion terms, in which forward time differencing is used for stability. The discrete model is

$$\begin{aligned} m_j^{k+1} = m_j^{k-1} & -\frac{\Delta t}{2\Delta x}\{(u_{j+1}^k + u_j^k)(m_{j+1}^k + m_j^k) - (u_j^k + u_{j-1}^k)(m_j^k + m_{j-1}^k) \\ & +((\phi_{j+1}^k)^2 - (\phi_{j-1}^k)^2)\} \\ & -g\frac{\Delta t}{2\Delta x}\{(\phi_{j+1}^k + \phi_j^k)(H_{j+1} - H_j) + (\phi_j^k + \phi_{j-1}^k)(H_j - H_{j-1})\} \\ & +2\Delta t f n_j^k + 2\frac{\Delta t}{\Delta x^2}K(m_{j+1}^{k-1} - 2m_j^{k-1} + m_{j-1}^{k-1}) \end{aligned} \tag{7.8}$$

$$\begin{aligned} n_j^{k+1} = n_j^{k-1} & -\frac{\Delta t}{2\Delta x}\{(v_{j+1}^k + v_j^k)(m_{j+1}^k + m_j^k) - (v_j^k + v_{j-1}^k)(m_j^k + m_{j-1}^k)\} \\ & -2\Delta t f m_j^k + 2\frac{\Delta t}{\Delta x^2}K(n_{j+1}^{k-1} - 2n_j^{k-1} + n_{j-1}^{k-1}) \end{aligned} \tag{7.9}$$

$$\phi_j^{k+1} = \phi_j^{k-1} \quad -\frac{\Delta t}{\Delta x}(m_{j+1}^k - m_{j-1}^k) + 2\frac{\Delta t}{\Delta x^2}K(\phi_{j+1}^{k-1} - 2\phi_j^{k-1} + \phi_{j-1}^{k-1}) \tag{7.10}$$

for $k = 1, .., N-1$, $j = 0, .., J -$

$$n_j^1 = n_j^0 - \frac{\Delta t}{2\Delta x}\{(v_{j+1}^0 + v_j^0)(m_{j+1}^0 + m_j^0) - (v_j^0 + v_{j-1}^0)(m_j^0 + m_{j-1}^0)\}$$
$$-2\Delta t f m_j^0 + 2\frac{\Delta t}{\Delta x^2}K(n_{j+1}^0 - 2n_j^0 + n_{j-1}^0) \quad (7.15)$$

$$\phi_j^1 = \phi_j^0 - \frac{\Delta t}{\Delta x}(m_{j+1}^0 - m_{j-1}^0) + 2\frac{\Delta t}{\Delta x^2}K(\phi_{j+1}^0 - 2\phi_j^0 + \phi_{j-1}^0) \quad (7.16)$$

where the initial conditions are to be specified.

Since discrete models which include artificial diffusion do not always converge to the correct solution of (7.1)-(7.3), the PC scheme was compared in [69] with a method (Glimm's method) which has been proved to converge to the physically correct solution. The PC scheme was found to give good agreement to corresponding solutions of Glimm's method for several test cases involving hydraulic jumps.

After coding up this model in Fortran 77, we tested it by comparing results with those obtained in two of the examples given in [69]. We describe one of these examples here, since it was modified to provide the example to be used for our experiments. The Coriolis parameter is set at the value for 30^{O} North, ie $f = 7.292 \times 10^{-5}\mathrm{s}^{-1}$. We use a spatial discretization of 100 grid points, so $J = 100$. In the first experiment of the paper, there is no topography, but a hydraulic jump evolves for certain Rossby and Froude numbers from smooth initial conditions given by

$$u(x,0) = U\cos(x/L), \quad (7.17)$$
$$v(x,0) = 0, \quad (7.18)$$
$$\phi(x,0) = \phi_m + U\{(U/8)\cos(2x/L) + (\phi_m - U^2/8)^{\frac{1}{2}}\cos(x/L)\}, \quad (7.19)$$

where the length of the domain is $2\pi L$, $\phi_m = g\eta_m$ where η_m is the mean depth of the fluid, and U is a constant. In our case, we ensured the required Rossby number $Ro = U/fL = 1$ and Froude number $F = U/\phi_m^{\frac{1}{2}} = 1$ were satisfied by choosing the constants L, U and ϕ_m as

$$U = 1\mathrm{ms}^{-1},$$

Since $\Delta x = 2\pi L/J$, we have $\Delta x = 4.58 \times 10^4$m, or approximately 46km. We chose a timestep to satisfy $\frac{\Delta t}{\Delta x} = \frac{1}{10}$, so $\Delta t = 4.58 \times 10^3$s (which is approximately one hour 15 minutes). As in the paper [69], we chose $K = 2.5 \times 10^4 \mathrm{m}^2\mathrm{s}^{-1}$.

Results from this test case were plotted in non-dimensional form, and seen to give good agreement with the corresponding figures in [69]. Our model was also tested on the examples given in [69] which include topography, and found to agree with the results in the paper in these cases, too.

7.2 The d t ssimil tion problem

We define the model state $\mathbf{x}_k \in \mathbb{R}^{3J}$ at time t_k to be the vector

$$\mathbf{x}_k = (m_0^k, .., m_{J-1}^k, n_0^k, .., n_{J-1}^k, \phi_0^k, .., \phi_{J-1}^k)^T, \tag{7.23}$$

and we define the correction term to be the vector $\mathbf{e} \in \mathbb{R}^{3J}$ given by

$$\mathbf{e} = (e_0^{(m)}, .., e_{J-1}^{(m)}, e_0^{(n)}, .., e_{J-1}^{(n)}, e_0^{(\phi)}, .., e_{J-1}^{(\phi)})^T. \tag{7.24}$$

We suppose that we have observations over timesteps t_0 to t_{N-1}. The data assimilation problem we address is to minimize

$$\mathcal{J} = \mathcal{J}_o + \frac{1}{2}\mathbf{e}Q^{-1}\mathbf{e} \tag{7.25}$$

with respect to the control vector or vectors being used, subject to the constraint that the model equations (7.8)-(7.16) hold, where the observational part of the cost function $\mathcal{J}_o$ is to be specified later. The matrix Q^{-1} is given by qI where I is the identity matrix, and different values of q are used in the experiments. We suppose that there are a large number of observations, and so do not include a background of the initial state in the cost function.

7.2.1 The adjoint model

We wish to minimize the cost function $\mathcal{J}$ subject to each of the model equations (7.8)-(7.10) with (7.14)-(7.16) and the relations (7.12),(7.13), and we introduce a Lagrange multiplier for each of these model equations.

for $k = N, .., 1$, $j = 0, .., J-1$ with

$$\lambda_j^{N+1} = 0, \qquad p_j^{N+1} = 0, \qquad \mu_j^{N+1} = 0, \tag{7.31}$$

$$\lambda_j^{N} = 0, \qquad p_j^{N} = 0, \qquad \mu_j^{N} = 0, \tag{7.32}$$

for $j = 0, ..., J-1$.

7.2.2 The gradients of $\mathcal{L}$ with respect to the control vectors

The partial derivatives of the Lagrangian $\mathcal{L}$ with respect to the variables making up the initial state $\mathbf{x}_0$ are given by

$$\frac{\partial \mathcal{L}}{\partial m_j^0} = -\lambda_j^0 - \lambda_j^1 - \frac{\Delta t}{2\Delta x^2} K(\lambda_{j+1}^1 - 2\lambda_j^1 + \lambda_{j-1}^1), \tag{7.33}$$

7.3 Description of the experiments

7.3.1 The true model state

For our experiments, we suppose that the true model state is defined by a run of the model with certain parameters and initial conditions. We use the values of f, J, L, Δx, Δt, and K specified in Section 7.1, but this time we use different initial conditions and a non-zero bottom topography. This low spatial resolution was chosen so that the dimension of the control vector would not be too large. It was noticed, however, that when the model was run at twice the spatial resolution, the results were not significantly different. In Experiments 1 and 2, the model is run for 100 timesteps ($N = 100$), and we take the assimilation interval $[0, T]$ to represent 100 timesteps.

The bottom topography is as given in [69], by

$$H(x) = H_c(1 - (x - \frac{L}{2})^2/a^2) \qquad 0 \leq \ (x - \frac{L}{2}) \ \leq a, \tag{7.41}$$

where H_c is half the initial water depth. We take a to correspond to a length of ten grid points. The shape of the bottom topography is shown in Fig. 7.1.

We define the true model initial state to be given by a fluid depth of 1m, and zero velocities, so we have (taking $g = 10\text{ms}^{-2}$)

$$m_j^0 = 0\text{m}^3\text{s}^{-3}, \tag{7.42}$$

$$n_j^0 = 0\text{m}^3\text{s}^{-3}, \tag{7.43}$$

$$\phi_j^0 = 10\text{m}^2\text{s}^{-2}, \tag{7.44}$$

for $j = 0, ..., J-1$. From this initial state, motion is initiated as fluid flows down from the ridge in the centre of the domain. A wave travels in each direction across the domain. This is illustrated in Fig. 7.1 which shows the true solution at the initial time and also after 50 and after 100 timesteps.

7.3.2 Observations

In Experiments 1 and 3, we suppose that we have a full set of observations, ie, observations of all the model state variables for all 100 timesteps. These observations

are the same as the true model state. In Experiment 3, we also carry out experiments in which we suppose that observ

therefore

$$\varepsilon_j^{(m)k} = g\frac{\Delta t}{2\Delta x}\{(\phi_{j+1}^k + \phi_j^k)(H_{j+1} - H_j) + (\phi_j^k + \phi$$

compromise between a better approximation of the inverse Hessian using a high value of $\hat{m}$, and lower CPU time with a low value of $\hat{m}$.

7.3.5 The experiments

Experiment 1

The aim of this experiment is to compare the performance of the different control vectors in the presence of different types of model error, and of error in the initial state. A full set of observations, uncorrupted by error, is used. The following cases are investigated.

Case a) Perfect model, unknown initial state

In this case there is no model error, but the true initial state (7.42)-(7.44) is unknown. The background estimate of the initial state is

$$m_j^0 = 0\mathrm{m}^3\mathrm{s}^{-3}, \tag{7.56}$$

$$n_j^0 = 0\mathrm{m}^3\mathrm{s}^{-3}, \tag{7.57}$$

$$\phi_j^0 = 15\mathrm{m}^2\mathrm{s}^{-2}, \tag{7.58}$$

for $j = 0, ..., J-1$.

Case b) Omission of topography, known initial state

In this case the model error is type i) above, but the true initial state is known.

Case c) Omission of rotation, known initial state

In this case the model error is type ii) abov

Experiment 2

Experiment 2, Case b) is repeated using observations corrupted by observational noise, and using fewer observations, also corrupted by observational noise. The aim here is to check to what extent the conclusions of Experiment 1 still hold in this more realistic case, rather than to explore the impact of increasing or reducing the number of observations. The following cases are investigated.

Case e) Observations with random error

Experiment 1b is carried out using observations corrupted by noise as described in Subsection 7.3.2.

Case f) Fewer observations with random error

Experiment 1b is carried out using fewer observations corrupted by noise. We suppose that observations (of all the state variables) are available only at every fourth timestep.

Experiment 3

In Experiment 1 we compare the performance of the different co I n s i s e v

Case h) Omission of rotation, known initial state

In this case the model error is type ii) above, but the true initial state at the beginning of the assimilation interval is known.

7.4 Results from the experiments

7.4.1 Experiment 1: Comparing different control vectors

The figures referred to here can be found at the end of this section.

Case a) Perfect model, unknown initial sate

Fig. 7.1 shows the true solution, and the background solution started with the wrong

assimilation interval. The results shown are for $q = 1$. Using $q = 10$, the correction term found is smaller, and so in this case the solution is closest to the true solution at the end rather than in the middle of the assimilation interval. For larger values of q fewer iterations are needed, for $q = 1$, 63 iterations are needed and using $q = 10$, 39 iterations are needed, which is similar to the number of iterations required using the initial state as the control vector. When q is increased further, however, the results are much poorer. The points we make here on the impact of different values of q are consistent with the conclusions w

equation (7.49), the actual model error at each timestep depends on the m-field. Fig. 7.18 shows that the correction term derived in the assimilation is a correction to this field only, and so it is reasonable to assume that the correction term found in the assimilation does indeed represent the temporal average of model error. The errors still existing in the ϕ-field at the end of the assimliation interval are presumably due to the fact that this average does not perfectly represent the actual model error. However, it is significant that the correction term represen

of ϕ is corrected, except at the ridge. The omitted topography and rotation are compensated for as when the initial state is used as the control vector in Cases b) and c). After the assimilation, quite large errors remain in the ϕ-field.

When the correction term is used as the control vector, the ϕ- and n-fields are quite close to the true solution at the end of the assimilation interval, but the m-field has larger errors. The solution no longer underestimates the ϕ-field at the end of the assimilation interval as was the case using the correction term as the control vector in Case a). In this case using the correction term happens to give the best fit to the true solution at the end rather than in the middle of the assimilation interval. This could be explained by the fact that the correction to the ϕ-field, shown in Fig. 7.20, is much smaller than it is in Case a) (Fig. 7.17). The correction to the m-field in (Fig. 7.20) is similar to that obtained in Case b) (Fig. 7.18), but is slightly larger. This might explain why there are larger errors in the m-field at the end time in this case than in Case b). The n-field produced is very similar to that produced in Case c) using the correction term as the control vector.

The results using both control vectors together in this case are very good as can be seen by comparing Fig. 7.14 with the true solution of Fig. 7.13. Fig. 7.16 shows that indeed the errors using both control vectors are much smaller than those using either one of the control vectors, and that these errors are almost zero except for those in the n-field. The results shown in the figures were obtained using $q = 1$, and the iteration was terminated after 300 iterations, before the convergence criterion had been satisfied.

Experiments using other control vectors

We mention briefly an attempt at using a couple of the other control vectors mentioned in Chapter 6, Section 6.2 in Experiment 1. We used the spectral form of model error, and a piecewise constant form using three subintervals. Using the spectral form of model error for Cases b) and c), the results were similar to the results obtained using the correction term which w

Using the piecewise continuous form, problems arose in the iteration process which were probably due to large differences between the three correction terms. It should be possible to rectify this situation, however, and this would be an interesting topic for further work.

7.4.2 Experiment 2: Fewer observations and observational error

The aim of Experiment 2 is to check whether the conclusions of Experiment 1 still hold in the presence of observational error and when there are fewer observations available. We therefore repeat Experimen

term is used as a control vector, approximately three times as many iterations were required. Surprisingly, though, the number of iterations required when both control vectors are used together is about the same as when the whole set of observations are used.

When the correction term is used as the control vector with $q = 0$ (Fig. 7.23), the solution produced by the assimilation is very spiky throughout the assimilation interval, as was found in the experiments of Chapter 5. Also as in Chapter 5, increasing the value of q smoothes the solution, Fig. 7.24 shows the results obtained using $q = 1$. However, in this case the results using $q = 10$ although smooth, were much less accurate. It may be, then, that an alternative method for smoothing the solution is needed when using the correction term as a control vector with fewer observations available.

When both control vectors are used together, the solution obtained is significantly smoother than when either the initial state or the correction term is used alone. Using $q = 0$ produced a smoother initial state than using $q = 1$, but using $q = 1$ produced a smoother solution at later times than using $q = 0$ (Fig. 7.23 and Fig. 7.24).

7.4.3 Experiment 3: The impact of assimilation on a forecast

Experiments 3g and 3h were first carried out performing an assimilation over just 50 timesteps (rather than 100 timesteps as in Experiments 1 and 2), or on the time interval $t \in [0, \frac{T}{2}]$, using the results to initiate a forecast for the interval $[\frac{T}{2}, T]$. The experiments were then repeated using assimilation and forecast intervals of 100 timesteps each. In this case the assimilation was carried out on the interval $[0, T]$,

demonstrates the effects of the model error over this time interval. Starting a forecast with an imperfect model at time $\frac{T}{2}$ from the true state at that time is equivalent to suddenly removing the topography in the middle of a model run. In addition to the existing motion, there are now also waves travelling towards the centre as the fluid fills the area where ridge used to be. Because of this, the forecast from the true state at time $\frac{T}{2}$ using the imperfect model very quickly diverges from the true solution.

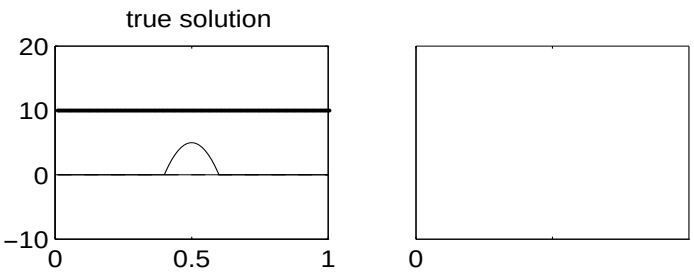
We now describe the results of starting a forecast from the assimilation analysis at time $\frac{T}{2}$, comparing the results using the initial state and using the correction term as the con

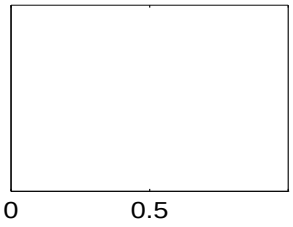
Over this time interval, the waves reac

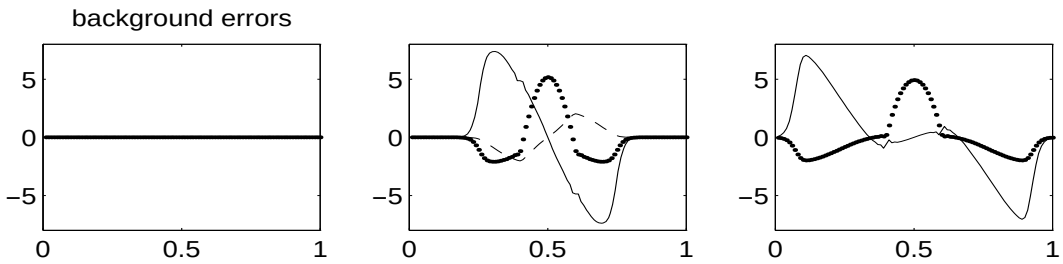
control vectors. If the initial state is used as the control vector in the assimilation interval, the ensuing forecast shows an improvement over the background solution in the middle of the forecast, but not at the end; all the benefit of the assimilation is lost by the end of the forecast.

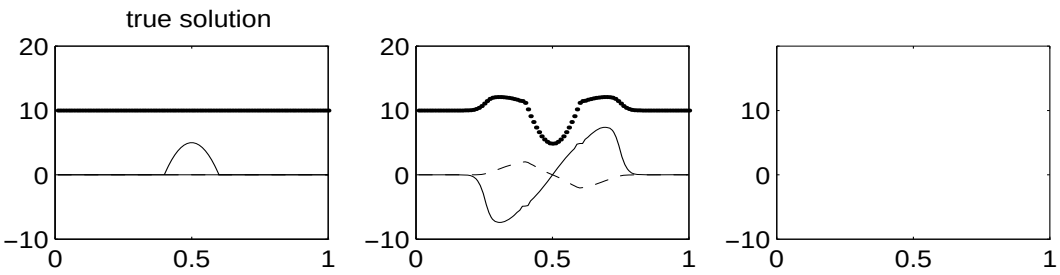
However, when the correction term is used as the control vector during the assimilation interval, the solution at time $\frac{T}{2}$ is closer to the true state, and hence the forecast is better than when the initial state is used as the control vector in the assimilation. This is true whether the correction term is included in the forecast or not, and it is hard to judge whether or not including it is beneficial in this case.

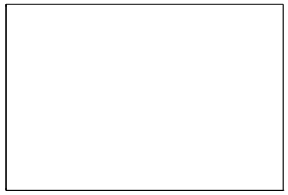
Experiment 3h was repeated using the longer assimilation and forecast intervals. The results for this case are shown in Fig. 7.31 and Fig. 7.32. Here, much the same conclusions hold as for the shorter time interval, except that in this case a better forecast is achieved by using the correction term in the assimilation but not in the forecast.


true solution
20
10
0
−10
0
0.5
1
0

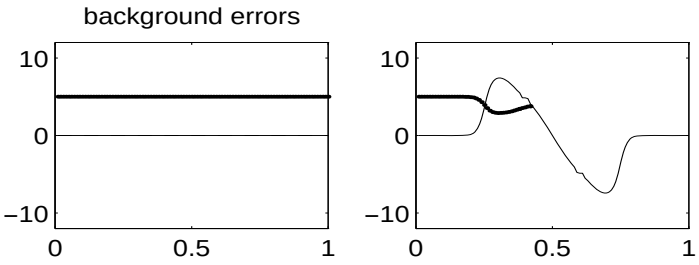

0
0.5

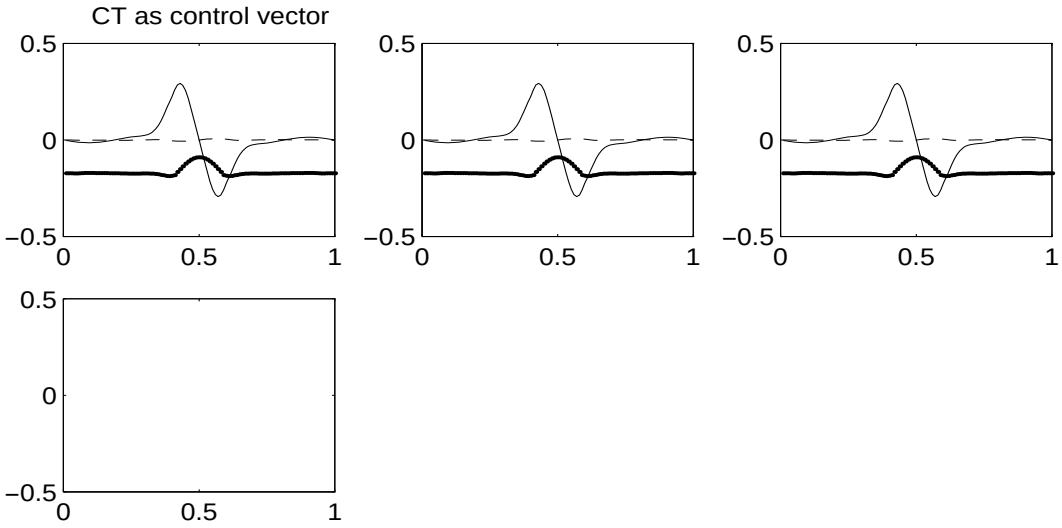

background errors
5
0
−5
0
0.5
1
5
0
−5
0
0.5
1
5
0
−5
0
0.5
1

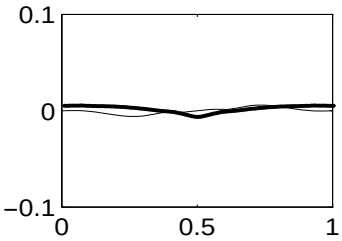

true solution
20
10
0
−10
0
0.5
1
20
10
0
−10
0
0.5
1
10
0
−10
0
0.5
1

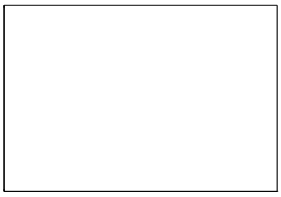


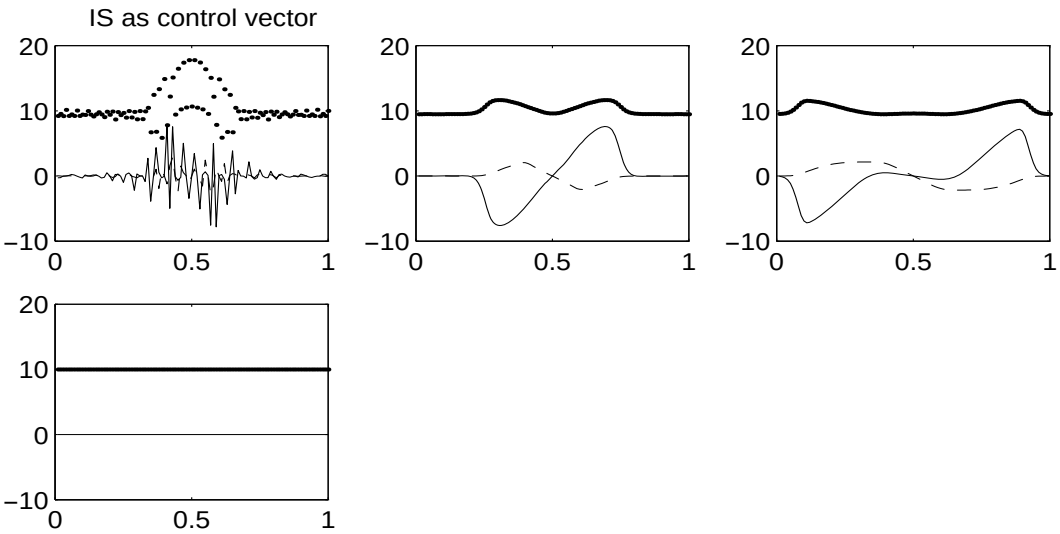
0

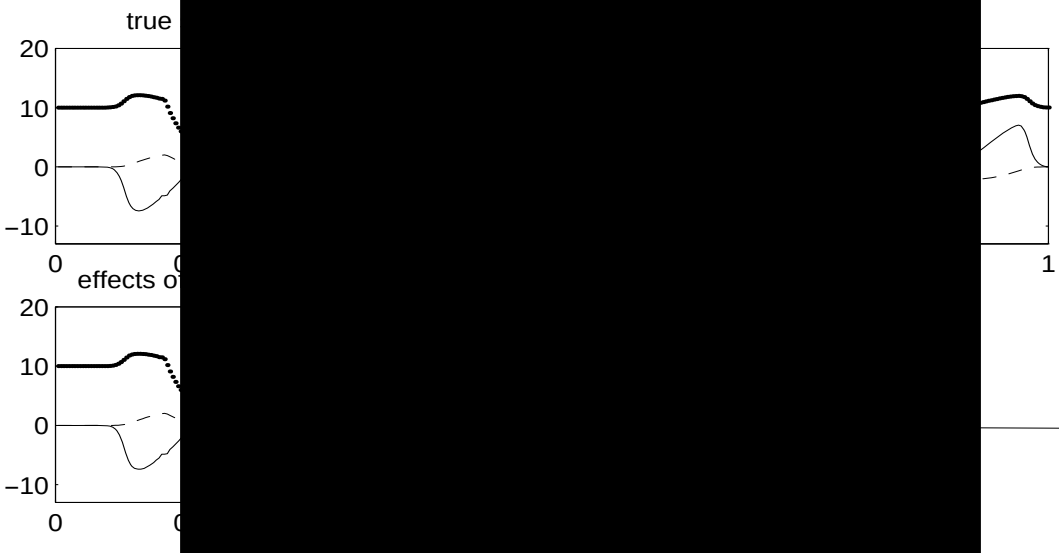

background errors
10
0
−10
0
0.5
1
10
0
−10
0
0.5
1

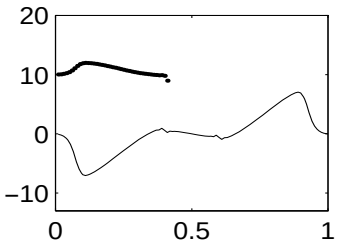

CT as control vector
0.5
0
−0.5
0
0.5
1

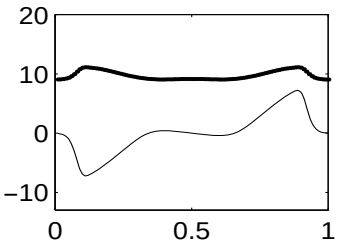

0.1
0
−0.1
0
0.5
1

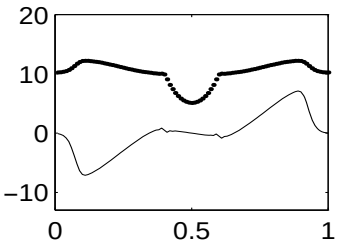



IS as control vector
20
10
0
−10
0
0.5
1


true
20
10
0
−10
0
1
effects o
20
10
0
−10
0


20
10
0
−10
0
0.5
1


20
10
0
−10
0
0.5
1


20
10
0
−10
0
0.5
1

7.5 Summary and Conclusions

Experiment 1: Comparing different control vectors

In Experiment 1 we compared the effectiveness of the initial state, the (constant) correction term and both together as control vectors in compenstating for errors in the initial state, for two types of model error, and for a combination of these errors during an assimilation interval. This extends the experiments of Chapter 5 since the model dynamics are more complex, and because the model error depends on the model state and so is not constant in time. Generally, the conclusions of the experiments of Chapter 5 are found to hold here, too. We found that the in

Experiment 3: The impact of assimilation on the forecast

In Experiment 3 we test whether the improvement in the solution at the end of an assimilation interval, produced by assimilation using each of control vectors, results in an improvement in a subsequent forecast. This is carried out for both examples of model error. In the case that model error is due to the omission of topography, using the initial state as a control vector gives a significant improvement over the background solution not only in the assimilation interval, but also in the forecast.

Using the constant correction term as a control vector gives a good improvement during the assimilation interval, but if the correction term is not included in the forecast, the forecast soon deteriorates, and becomes worse than if no assimilation had been performed. These results can be explained by examining the impact of starting a forecast with an imperfect model from the true solution at the end of the assimilation interval. It is the impact of using a different model for the assimilation and forecast that causes the forecast to deteriorate quickly. We note that this situation could be alleviated by gradually phasing out the correction term during the assimilation interval using the predetermined scalars of equation (4.32), as was done in the original paper on the correction term technique [26]. When we include the correction term in the forecast, however, the forecast is good; better than the forecast produced using the initial state as the control vector in the assimliation.

is used as the control vector, complete N-step observability at time t_0 is a necessary and sufficient condition for uniqueness in the case where the cost function consists of observations from the time interval $[t_0, t_{N-1}]$. We showed however, that complete N-step observability at time t_0 is neither a necessary nor a sufficient condition for uniqueness when the constant correction term is used as a control vector. This means that in some cases the set of observations may contain enough information to specify uniquely the initial state but not the correction term, and vice versa. We showed that if both the initial state and the correction term are used as control vectors, a necessary but not sufficient condition for uniqueness is that conditions for a unique solution using each of the control vectors individually hold. In the time invariant case, we showed that a necessary and sufficient condition for uniqueness using both control vectors is that a full set of observations is available. In each case, adding a background estimate of the control vector to the cost function guarantees uniqueness. These results could be applied more widely in control theory in cases where we wish to determine a constant input from the outputs.

In Chapter 6, we addressed the question of how to allow for a more general form of model error in 4D data assimilation, and in 4D variational assimilation in particular. We considered a general, stochastic representation of model error consisting of serially correlated and serially uncorrelated components. The different representations of model error that have been suggested for use in data assimilation can be expressed using this general form. We considered the technique of state augmentation for estimating the serially correlated component of model error along with the model state in the context of data assimilation, and formulated a general least squares problem for data assimilation allowing for serially correlated model error. This formalism allows us to interpret the correction term technique in a stochastic sense as a method for estimating a constant model bias.

We suggested a "generalized correction term technique" in which the serially uncorrelated part of the model error is neglected, and the augmented initial state is used as an augmented control vector. The generalized correction term technique can therefore allow for various different forms of serially correlated model error. In particular, it can allow for model error which evolves as the model state does. The

theory we present also allows for the dimension m of the correction term to be less than the dimension n of the model state, which could reduce the expense of the assimilation if the effects of model error are known to be localized to a certain area.

As well as considering theoretical aspects of accounting for model error in variational assimilation, we carried out experiments using the correction term technique and generalized correction term technique with simple models exhibiting different types of model error.

In Chapter 5 we compared the initial state, the constant correction term and both together as control vectors in a heat equation model in which model error was due to the omission of a constant source term. Using the correction term as a control vector compensates very well for this model error, and using the correction term in an ensuing forecast gives very good results. We also noted that using the initial state as a control vector partially compensates for the effects of model error, and that using the correction term as the control vector, it is possible to compensate to some extent for errors in the initial state. Using both control vectors together is very effective in this example if we have model error and an unknown initial state. However, this requires about four times as many iterations of the descent algorithm as when only one of the control vectors is used.

In these experiments we also investigated the impact of using a background estimate of zero for the correction term in the cost function, with different values of the weighting q. We found that if the correction term is expected to correct for constant model error, best results are obtained with a small value of q. However, if the correction term is being used to compensate for errors resulting from a wrong initial state, it is important to use a large value of q. We also tried using a correction term with dimension less than that of the model state. Concentrating the correction term around the source point produced good results in far fewer iterations than before.

In Chapter 6, the simple model we used was the linear advection equation with the upwind scheme discretisation. The model error in this example is due to severe dissipation. Using the constant correction term as a control vector has no impact on thesevΦTD?using–ttheas .H

control vector indicate that this problem of reducing the num

[11] M. Berz, C. Bischof, G.F. Corliss, and A. Greiwank. *Computational ifferentiation: Techniques, Applications and Tools*. SIAM, 1996.

[12] G.J. Boer. A spectral analysis of predictability and error in an operational forecast system. *Monthly Weather Review*, 112:1183–1197, Jun 1984.

[13] F. Bouttier. A dynamical estimation of forecast error covariances in an assimilation system. *Monthly Weather Review*, 122:2376–2390, Oct 1994.

[14] D.E. Catlin. *Estimation, Control, and the iscrete Kalman Filter*. Springer-Verlag, 1989.

[15] W.C. Chao and L-P. Chang. Development of a four-dimensional variational analysis system using the adjoint method at GLA. Part 1: Dynamics. *Monthly Weather Review*, 120:1661–1673, Aug 1992.

[16] S.E. Cohn. An introduction to estimation theory. 1995. Submitted to Journal of the Meteorological Society of Japan.

[17] S.E. Cohn and D. P. Dee. Observability of discretized partial differential equations. *IAM J. Numer. Anal.*, 25:586–617, Jun 1988.

[18] P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation. II. Numerical results. *uart. J. R. Met. oc.*, 113:1329–1347, 1987.

[19] P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the direct and adjoint shallow–water equations. *Tellus*, 42A:531–549, 1990.

[20] P. Courtier, J-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-VAR, using an incremental approach. *. J. R. Meteorol. oc.*, 120:1367–1387, 1994.

[21] G.P. Cressman. An operational objective analysis system. *Monthly Weather Review*, pages 367–374, Oct 1959.

[22] A. Dalcher and E. Kalnay. Error growth and predictability in operational ECMWF forecasts. *Tellus*, 39A:474–491, 1987.

[23] R. Daley. The effect of serially correlated observation and model error on

[45] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the A ME series* , 82:35–44, 1960.

[46] R. E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Transactions of the A ME series* , 83:35–44, 1961.

[47] J. Kautsky, N.K. Nichols, and P. van Dooren. Robust pole assignment in linear state feedback. *Int. J. Control*, 41(5):1129–1155, 1993.

[48] J-F. Lacarra and O. Talagrand. Short-range evolution of small perturbations in a barotropic model. *Tellus*, 40A:81–95, 1988.

[49] A. Lawless. A perturbation forecast model and its adjoint. In *American Meteorological ociety* 11^{th} *Conference on Numerical Weather Prediction*, pages 128–129, Norfolk, Virginia, 1996.

[50] F-X. Le Dimet and M. Ouberdous. Retrieval of balanced fields: An optimal control method. *Tellus*, 45A:449–461, 1993.

[51] F-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, 38A:97–110, 1986.

[52] J.M. Lewis and J.C.Derber. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, 37A:309–322, 1985.

[53] J.L. Lions. *Optimal Control of ystems Governed by Partial ifferential Equations*. Springer–Verlag, 1971. (English translation).

[54] A.C. Lorenc. A global three–dimensional multi–variate statistical interpolation scheme. *Monthly Weather Review*, 109:701–721, 1981.

[55] A.C. Lorenc. Analysis methods for numerical weather prediction. *.J.R. Meteorol. oc.*, 112:1177–1194, 1986.

[56] A.C. Lorenc. Optimal nonlinear objective analysis. *.J.R. Meteorol. oc.*, 114:205–240, 1988.

[57] A.C. Lorenc. Iterative analysis using covariance functions and filters. *.J.R. Meteorol. oc.*, 118:569–591, 1992.

[58] A.C. Lorenc and O. Hammon. Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *.J.R. Meteorol. oc.*, 114:515–543, 1988.

[59] B. Machenhauer. On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitraege zur Physik der Atmosphaere*, 50:253–271, 1977.

[60] R Ménard. *Kalman Filtering of Burgers' equation and its application to atmospheric data assimilation.* PhD thesis, McGill University, 1994.

[61] R. Ménard and R. Daley. The application of Kalman smoother theory to the estimation of 4DVAR error statistics. *Tellus*, 48A:221–237, 1996.

[62] H. Michalska. An observer for nonlinear descriptor systems of index one. In *Proceedings of the International ymposium on Implicit and Nonlinear ystems*, pages 372–379, 1992.

[63] R.N. Miller, M. Ghil, and F. Gauthiez. Advanced data assimilation in strongly nonlinear dynamical systems. *Journal of the Atmospheric iences*, 51(8):1037–1056, 1994.

[64] P.J. Moylan. Stable inversion of linear systems. *IEEE Trans. on Autom. Control*, 22:74–78, 1977.

[65] J. Nocedal. Updating quasi-Newton matrices with limited storage. *Mathematics of Computation*, 35(151):773–782, 1980.

[66] J. O'Reilly. *Observers for Linear ystems.* Academic Press, 1983.

[67] S. Oren and E. Spedicato. Optimal conditioning of self-scaling variable metric algorithms. *Mathematical programming*, 10:70–90, 1976.

[68] H.A. Panofski. Objectives

[69] C.A. Parrett and M.J.P. Cullen. Simulation of hydraulic jumps in the presence of rotation and mountains. *uart. J. R. Met. oc.*, 110:147–165, 1984.

[70] R.V. Patel. Minimal-order inverses for linear systems with zero and arbitrary initial states. *Int. J. Control*, 30(2):245–258, 1979.

[71] R.V. Patel. Construction of stable inverses for linear systems. *Int. J. ystems ci.*, 13(5):499–515, 1982.

[72] C. Pires and O. Talagrand. On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus*, 48A:96–121, 1996.

[73] F. Rabier, P. Courtier, J. Pailleux, O. Talagrand, and D. Vasiljevic. A comparison between four-dimensional variational assimilation and simplified sequential assimilation relying on three-dimensional variational analysis. *. J. R. Meteorol. oc.*, 119:845–880, 1993.

[74] J. Rinne and H. Jarvinen. Estimation of the Cressman term for a barotropic model through optimization with use of the adjoint method. *Monthly Weather Review*, 121:825–833, 1993.

[75] Y. Sasaki. An objective analysis based on the variational method. *Journal of the Meteorological ociety Japan*, 36(3):77–88, 1958.

[79] S.M. Stringer. *The Use of Robust Observers in the imulation of Gas upply Networks*. PhD thesis, The University of Reading, Department of Mathematics, 1993.

[80] O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *uart. J. R. Met. oc.*, 113:1311–1328, 1987.

[81] A. Tarantola. *Inverse problem theory. Methods for data fitting and model parameter estimation.* Elsevier, 1987.

[82] W.C. Thacker. Relationships between statistical and deterministic methods of data assimilation. In *Variational Methods in the Geosciences*. Elsevier, 1986.

[83] W.C. Thacker and R.B. Long. Fitting dynamics to data. *Journal of Geophysical Research*, 93(c2):1227–1240, 1988.

[84] R. Todling and S. E. Cohn. Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Monthly Weather Review*, 122:2530–2557, 1994.

[85] J.J. Tribbia and D.P. Baumhefner. The reliability of improvements in deterministic short–range forecasts in the presence of initial state and modeling deficiencies. *Monthly Weather Review*, 108:2276–2288, 1988.

[86] L. Weiss. Controllability, realization and stability of discrete-time systems. *IAM J. Control*, 10(2):230–251, 1972.

[87] W. Wergen. The effect of model errors in variational assimilation. *Tellus*, 44A:297–313, 1992.

[88] J.H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford University Press, 1965.

[89] M.A. Wolfe. *Numerical Methods for Unconstrained Optimization, an Introduction.* Van Nostrand Reinhold Company, 1978.

[90] W. Wonham. On pole assignment in multi–input, controllable linear systems. *IEEE Trans Automatic Control*, AC(12):660–665, 1967.

[91] X. Zou, I.M. Navon, and F-X. Le Dimet. Incomplete observations and control of gravity waves in variational assimilation. *Tellus*, 44A:273–296, 1992.

[92] D. Zupanski. A general weak constraint applicable to operational 4DVAR data assimilation systems. 1996. Submitted to Monthly Weather Review.

[93] D. Zupanski and F. Mesinger. Four–dimensional variational data assimilation of precipitaion data. *Monthly Weather Review*, 123:1112–1127, 1995.

[94] M. Zupanski. Regional four–dimensional variational data assimilation in a quasi–operational forecasting environment. *Monthly Weather Review*, 121:2396–2408, 1993.