

# Adjoint Data Assimilation Methods

**Andrew M. Moore**

**Ocean Sciences Department**

University of California Santa Cruz

Santa Cruz CA 95064 U.S.A.

**Abstract** The use of adjoint methods in data assimilation is reviewed, and illustrative examples are presented.

## 1 Introduction

Adjoint operators are central to many operational data assimilation systems used for numerical weather prediction, and are gaining popularity in oceanography also. In this chapter we shall review the use of adjoint methods for data assimilation. We begin in section 2 with an exploration of the concept of the adjoint of a linear operator, and the important properties that make it an indispensable tool for data assimilation. Familiar illustrative examples are used throughout to highlight the important ideas. The fundamental concepts underpinning 4-dimensional variational data assimilation (4D-Var) are reviewed in section 3, and in section 4 example 4D-Var calculations for the California Current System are presented using the Regional Ocean Modeling System (ROMS).

## 2 What is an adjoint operator?

Adjoint operators exist *only* for linear operators. The concept of an adjoint operator is best illustrated by considering first the discrete form of linear operators and functions, namely matrices and vectors. The following is a brief exposé on adjoint operators, but an excellent in-depth description can be found in the classic text by Lanczos (1961).

## 2.1 Spaces

Any continuous linear operator in function space has a discrete analog in the form of a matrix. Similarly, any continuous function has a discrete analog in the form of a vector. With this in mind, consider the  $N \times M$  rectangular matrix  $\mathbf{A}$ . The matrix  $\mathbf{A}$  operates on the set of vectors  $\mathbf{u}$  of length  $M$  and yields a set of vectors  $\mathbf{w}$  of length  $N$ , so that  $\mathbf{w} = \mathbf{A}\mathbf{u}$ . So we say that  $\mathbf{A}$  maps from a space of dimension  $M$  (“ $M$ -space”) to a space of dimension  $N$  (“ $N$ -space”). The adjoint of the operator  $\mathbf{A}$  can be identified with the matrix transpose, namely  $\mathbf{A}^T$ . The formal connection between the matrix transpose and an adjoint operator will be made in section 2.2, but for now identifying the adjoint as the matrix transpose will suffice. The adjoint  $\mathbf{A}^T$  is a  $M \times N$  matrix and operates on the set of vectors  $\mathbf{v}$  of length  $N$  to yield the set of vectors  $\mathbf{z}$  of length  $M$ , namely  $\mathbf{z} = \mathbf{A}^T\mathbf{v}$ . So we say that the adjoint maps from  $N$ -space to  $M$ -space.

Suppose that we wish to solve the system of linear equations  $\mathbf{y} = \mathbf{A}\mathbf{x}$  given  $\mathbf{A}$  and  $\mathbf{y}$ . This represents a system of  $N$  equations for the  $M$  unknown elements of  $\mathbf{x}$ . If  $N < M$ , the system is said to be underdetermined since there are fewer equations than unknowns. In this case we might ask whether a unique solution exists for  $\mathbf{x}$ ? The answer is yes, and is given by the so-called “natural solution” for which the adjoint operator plays a critical role. Suppose we search for a solution of the form  $\mathbf{x} = \mathbf{A}^T\mathbf{s}$ . While the vector  $\mathbf{x}$  resides in  $M$ -space, the vector  $\mathbf{s}$  resides in  $N$ -space, so we are effectively restricting our search for solutions to  $N$ -space, the space in which the known vector  $\mathbf{y}$  resides. We have now reduced the problem to solving  $\mathbf{y} = \mathbf{A}\mathbf{A}^T\mathbf{s}$  which is a well determined system since  $\mathbf{A}\mathbf{A}^T$  is a  $N \times N$  matrix that maps  $\mathbf{s}$  to the  $N$  known elements of  $\mathbf{y}$ . The solution  $\mathbf{x} = \mathbf{A}^T\mathbf{s}$  is called the natural solution, and  $\mathbf{s}$  is referred to as the generating function. As we shall see, generating functions are important players in some approaches to data assimilation.

Let us consider a familiar geophysical example in which the natural solution plays a critical role. The vertical component of relative vorticity of a fluid element is given by  $\zeta = \partial v / \partial x - \partial u / \partial y$  where  $(u, v)$  are the  $x$  and  $y$  components of velocity. Suppose we are given a field of values of  $\zeta$  at discrete points  $(x, y)$  such as on the grid of a numerical model. This is the discrete analog of a function which we will denote by the vector  $\boldsymbol{\zeta}$  where each element of  $\boldsymbol{\zeta}$  represents a grid point value of  $\zeta$ . Given the field  $\boldsymbol{\zeta}$  in  $N$ -space, how do we find the corresponding velocity components  $(\mathbf{u}, \mathbf{v})$  in  $M$ -space, where  $M = 2N$  in this example, and  $\mathbf{u}$  and  $\mathbf{v}$  are the vectors of grid point values of  $u$  and  $v$  respectively? This is an underdetermined linear system for which a natural solution exists of the form:

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \mathbf{A}^T \mathbf{s}$$

where  $\mathbf{A} = \begin{pmatrix} -\partial/\partial y & \partial/\partial x \end{pmatrix}$ , and  $\boldsymbol{\zeta} = \mathbf{A}\mathbf{A}^T \mathbf{s} = -(\partial^2/\partial y^2 - \partial^2/\partial x^2)\mathbf{s}$ . If we identify  $\mathbf{s} = -\boldsymbol{\psi}$  we recover the familiar equation  $\boldsymbol{\zeta} = \nabla^2 \boldsymbol{\psi}$  relating vorticity  $\boldsymbol{\zeta}$  to stream function  $\boldsymbol{\psi}$  arising from Helmholtz theorem for a horizontally non-divergent flow. This example reveals that stream function is the generating function for vorticity for a horizontally non-divergent flow.

Recall that the  $N \times M$  matrix  $\mathbf{A}$  has an operator equivalent  $A$  in function space, so for the case  $N < M$  the function  $A$  acts only on part of the function space. We say that only some of the dimensions of the function space are activated by  $A$ . In the discrete case only  $N$  of the possible  $M$  dimensions are activated by the matrix  $\mathbf{A}$ . The remaining  $(M-N)$  non-activated dimensions are referred to as the “null space.” The  $M \times N$  adjoint operator  $\mathbf{A}^T$  identifies the activated part of the  $M$ -space and ignores the null space. The natural solution  $\mathbf{y} = \mathbf{A}\mathbf{A}^T \mathbf{s}$  therefore represents the set of solutions that exist only in the activated dimensions of  $\mathbf{A}$ . In the theory of linear differential equations, the natural solution is also called the particular integral. Solutions that reside in the null space satisfy the equation  $\mathbf{A}\mathbf{x} = \mathbf{0}$  and in the theory of linear differential equations are referred to as the complementary function. The general solution of any linear differential equation or equivalent discrete linear system is the sum of the particular integral (natural solution) and the complementary function. In both function space and discrete space the adjoint operator identifies the space in which these two parts of the general solution reside. As we shall see later, we can use this important property of adjoint operators for data assimilation.

Before concluding this discussion of vector spaces, it is instructive to consider the  $N \times M$  matrix  $\mathbf{A}$  where now  $N > M$ . The corresponding linear system  $\mathbf{y} = \mathbf{A}\mathbf{x}$  is now over determined since there are more equations than there are unknown elements of  $\mathbf{x}$ . Recall that the adjoint  $\mathbf{A}^T$  maps vectors from  $N$ -space to  $M$ -space where the solution for  $\mathbf{x}$  resides, so it is tempting to solve the system  $\mathbf{A}^T \mathbf{y} = \mathbf{A}^T \mathbf{A}\mathbf{x}$ . It is easy to show that the solution of this system minimizes  $(\mathbf{A}\mathbf{x} - \mathbf{y})^T (\mathbf{A}\mathbf{x} - \mathbf{y})$  and is the familiar least squares solution for an over determined system. Thus we see that the adjoint operator plays a critical role in identifying the least squares solution of over determined systems also.

## 2.2 Operator adjoints

To complete the connection between operator adjoints and matrices, it is instructive to return to function space. In function space, we will denote a linear operator as  $A$  and the adjoint of the operator as  $A^+$ . Consider the functions  $u$  and  $w$  so that  $w = Au$  which is the continuous analog of the discrete case considered in section 2.1. For any two functions  $u$  and  $v$ , there will in general exist an inner-product and an associated norm which we will denote by  $\{v, u\}$ . An adjoint operator is always

associated with a particular inner-product and is defined by  $\{v, Au\} = \{A^\dagger v, u\}$ , often referred to as the Greens identity. The adjoint operators of different inner-products are in fact linearly related. To illustrate, suppose we let  $\{v, u\}$  represent the inner-product of the Euclidean norm, and define a different inner-product as  $(v, u) = \{v, Mu\}$  where for now  $M$  is a linear, self-adjoint (i.e.  $M^\dagger = M$ ), invertible operator. The adjoint of  $A$  with respect to the new inner-product will be denoted  $A^\ddagger$  and is defined by the Greens identity  $(v, Au) = (A^\ddagger v, u) = \{M^\dagger A^\dagger M v, Mu\}$ , which shows that  $A^\ddagger = M^\dagger A^\dagger M$ .

In the discrete case, the inner-product of function space is replaced by a dot-product, so that for the Euclidean norm we have  $\{v, u\} = \mathbf{v}^T \mathbf{u}$ . Similarly the Greens identity for the adjoint becomes  $\mathbf{v}^T \mathbf{A} \mathbf{u} = (\mathbf{A}^T \mathbf{v})^T \mathbf{u} = \mathbf{u}^T \mathbf{A}^T \mathbf{v}$ , showing that for the Euclidean norm the discrete equivalent of the adjoint operator  $A^\dagger$  is the matrix transpose  $\mathbf{A}^T$ . If a different norm is used, such that  $(v, u) = \mathbf{v}^T \mathbf{M} \mathbf{u}$  where  $\mathbf{M}$  is a symmetric, invertible matrix, then the corresponding adjoint  $\bar{\mathbf{A}} = \mathbf{M}^{-1} \mathbf{A}^T \mathbf{M}$ .

Exercise 1: If  $\mathbf{A}^T$  is the adjoint of the operator  $\mathbf{A}$  with respect to the Euclidean norm  $\mathbf{v}^T \mathbf{u}$ , derive an expression for the adjoint operator  $\bar{\mathbf{A}}$  with respect to the norm  $\mathbf{v}^T \mathbf{M} \mathbf{u}$  where  $\mathbf{M}$  is a symmetric, invertible matrix, and show that  $\bar{\mathbf{A}} = \mathbf{M}^{-1} \mathbf{A}^T \mathbf{M}$ .

### 2.3 An illustrative example

The ideas of sections 2.1 and 2.2 are best illustrated using a simple yet familiar geophysical example. Consider a rectangular, homogeneous, flat bottomed ocean of undisturbed depth  $H$ , in the form of a rotating channel that spans the Cartesian domain  $0 \leq x \leq l$ ,  $0 \leq y \leq l$ , that is periodic in  $x$ , and subject to zero normal flow boundary conditions on the circulation at  $y=0$  and  $y=l$ .

#### 2.3.1 The linear shallow water equations.

We will consider first the case of linear waves in an ocean in which the circulation is described by the linear shallow water equations:

$$\partial u / \partial t - fv = -g \partial h / \partial x \quad (1)$$

$$\partial v / \partial t + fu = -g \partial h / \partial y \quad (2)$$

$$\partial h / \partial t + \partial(Hu) / \partial x + \partial(Hv) / \partial y = 0 \quad (3)$$

where  $(u, v)$  are the components of velocity in the  $x$  and  $y$  directions,  $h$  is the sea surface displacement,  $f=f(y)$  is the Coriolis parameter, and  $g$  is the acceleration due to gravity. The zero normal flow boundary conditions correspond to  $v=0$ , at  $y=0$  and  $y=l$ , while periodicity in  $x$  requires that  $u(0,y)=u(l,y)$ ,  $v(0,y)=v(l,y)$ , and  $h(0,y)=h(l,y)$ . Recall that the adjoint of equations (1)-(3) depends on the choice of an inner-product. A natural inner-product for the shallow water equations is that which yields the energy norm:

$$E = \int_0^1 \int_0^1 \frac{1}{2} H (u^2 + v^2) + \frac{1}{2} g h^2 dx dy. \quad (4)$$

If we introduce the shorthand notation  $s=(u,v,h)$  then equations (1)-(3) can be written as  $s_t + A s = 0$ , where the subscript denotes differentiation with respect to time, and the operator  $A$  is given by:

$$A = \begin{pmatrix} 0 & -f & g \partial / \partial x \\ f & 0 & g \partial / \partial y \\ H \partial / \partial x & H \partial / \partial y & 0 \end{pmatrix}. \quad (5)$$

The adjoint of (1)-(3) is defined by the Greens identity, which for the inner-product associated with the energy norm can be written as:

$$\int_0^1 \int_0^1 s^+ M A s dx dy = \int_0^1 \int_0^1 s M A^+ s^+ dx dy \quad (6)$$

where  $s^+=(u^+, v^+, h^+)$  is a function in the space on which  $A^+$  operates, and  $M=\text{diag}(H, H, g)$ . To apply (6) to the shallow water equations, consider  $I = \int_0^1 \int_0^1 H u^+ \times (1) + H v^+ \times (2) + g h^+ \times (3) dx dy = 0$ . After integration by parts, it is easy to show that the adjoint of (1)-(3) is given by  $-s_t^+ - A^+ s^+ = 0$ , where the negative time derivative indicates that time is reversed. The adjoint operator  $A^+$  is given by:

$$A^+ = \begin{pmatrix} 0 & -f & g \partial / \partial x \\ f & 0 & g \partial / \partial y \\ H \partial / \partial x & H \partial / \partial y & 0 \end{pmatrix} \quad (7)$$

where the adjoint variables are periodic in  $x$  and satisfy the boundary conditions  $v^+=0$ , at  $y=0$  and  $y=l$ . In addition,  $s$  and  $s^+$  must satisfy the condition:

$$\frac{\partial}{\partial t} \left( \int_0^1 \int_0^1 H(uu^+ + vv^+) + gh h^+ dx dy \right) = 0 \quad (8)$$

which is equivalent to time invariance of the inner-product  $\{s^+, Ms\}$ .

Comparing (5) and (7) shows that the operator  $A$  and its adjoint  $A^+$  are identical with respect to the energy norm. In addition, the adjoint equation  $-s_t^+ - A^+ s^+ = 0$  also satisfies the zero normal flow and periodic boundary conditions that are identical to those imposed on (1)-(3). Therefore, taken as a whole, the operator  $A$  and adjoint operator  $A^+$  are identical and have identical boundary conditions, and the system is said to be self-adjoint.

Exercise 2: Using  $I = \int_0^1 \int_0^1 Hu^+ \times (1) + Hv^+ \times (2) + gh^+ \times (3) dx dy = 0$ , derive the adjoint shallow water operator given by (7), and show that as a consequence of  $I=0$ , the adjoint equation  $-s_t^+ - A^+ s^+ = 0$  must satisfy (a) zero normal flow boundary conditions at  $y=0$  and  $y=1$ , and (b) the condition given by equation (8).

Wave solutions of  $s_t + As = 0$  given by (1)-(3) take the form of eastward and westward propagating inertia-gravity waves and Rossby waves (Gill, 1982). Similarly wave solutions exist for  $-s_t^+ - A^+ s^+ = 0$  but with opposite phase and group velocities because of the reversal of time. For example, long wavelength Rossby waves that carry energy westward in the shallow water equations  $s_t + As = 0$  will carry energy eastward in the adjoint equations  $-s_t^+ - A^+ s^+ = 0$ .

The adjoint equation can also be written as  $s_t^+ + A^+ s^+ = 0$  showing that it is mathematically equivalent to equations (1)-(3) in the case of self-adjoint  $A$ . Therefore, for the same initial conditions the solutions  $s=(u,v,h)$  and  $s^+=(u^+,v^+,h^+)$  will be identical, in which case equation (8) is an expression of energy conservation.

### 2.3.2 The linear shallow water equations in the presence of a mean circulation.

Consider now the case of linear waves in the same periodic rectangular ocean which is now in motion with a mean geostrophic circulation  $\bar{u} = (-g/f)\partial\bar{h}/\partial y$ . The linearized shallow water equations in this case are:

$$\partial u / \partial t + \bar{u} \partial u / \partial x + v \partial \bar{u} / \partial y - fv = -g \partial h / \partial x \quad (9)$$

$$\partial v / \partial t + \bar{u} \partial v / \partial x + fu = -g \partial h / \partial y \quad (10)$$

$$\partial h / \partial t + \partial(\tilde{H}u) / \partial x + \partial(\tilde{H}v) / \partial y = 0 \quad (11)$$

where  $\tilde{H} = H + \bar{h}$ , and  $\bar{h}$  is the sea surface displacement associated with the geostrophic circulation.

Using the same compact form as before, we can express (9)-(11) as  $s_t + As = 0$ , where the operator  $A$  is now given by:

$$A = \begin{pmatrix} \bar{u} \partial / \partial x & -f + \partial \bar{u} / \partial y & g \partial / \partial x \\ f & \bar{u} \partial / \partial x & g \partial / \partial y \\ \tilde{H} \partial / \partial x & \tilde{H} \partial / \partial y & 0 \end{pmatrix}. \quad (12)$$

Applying the Greens identity and using the energy norm, the adjoint equation is of the form  $-s_t^+ - A^+ s^+ = 0$  where now:

$$A^+ = \begin{pmatrix} \bar{u} \partial / \partial x & -f & g \partial / \partial x \\ f - \partial \bar{u} / \partial y & \bar{u} \partial / \partial x & g \partial / \partial y \\ \tilde{H} \partial / \partial x & \tilde{H} \partial / \partial y & 0 \end{pmatrix}. \quad (13)$$

As before, equation (8) is a necessary condition, and the adjoint variables are periodic in  $x$  and satisfy the boundary condition  $v^+ = 0$ , at  $y=0$  and  $y=I$ . In this case  $A$  and  $A^+$  describe different operators, and the system is no longer self-adjoint. The self-adjoint nature of the  $A$  has been destroyed by the  $y$ -gradient of the mean circulation  $\bar{u}$ , the relative vorticity  $-\partial \bar{u} / \partial y$  of the mean circulation. The circulation gradient in  $\bar{u}$  acts as a source of energy for linear waves and is the familiar source of barotropic instability if the waves can undergo sustained exponential growth. In this case, solutions of  $s_t + As = 0$  and  $s_t^+ + A^+ s^+ = 0$  with the same initial conditions will no longer be identical, and energy is no longer conserved. Sustained exponential growth will occur if the potential vorticity gradient  $\partial f / \partial y - \partial^2 \bar{u} / \partial y^2$  changes sign anywhere within the channel (Pedlosky, 1987).

Exercise 3: Using  $I = \int_0^1 \int_0^1 Hu^+ \times (9) + Hv^+ \times (10) + gh^+ \times (11) dx dy = 0$ , derive the adjoint shallow water operator given by (13), and show that as a consequence of  $I=0$ , the adjoint equation  $-s_t^+ - A^+ s^+ = 0$  must satisfy (a) zero normal flow boundary conditions  $y=0$  and  $y=I$ , and (b) the condition given by equation (8).

### 3 Variational Data Assimilation

#### 3.1 Notation

Before proceeding to describe the important role of adjoint operators in variational data assimilation methods, it is necessary to introduce some notation. Ocean models solve the discrete equations of motion on grids in both space and time. There are a wide range of ocean models available, and many use different kinds of grid configurations and coordinate systems (e.g., staggered grids, terrain following vertical coordinates, isopycnal vertical coordinates, etc). Nonetheless, all models solve for a standard set of prognostic variables, typically temperature, salinity, velocity, and free surface displacement, and can be expressed in a generic symbolic form using the concept of a state-vector. To this end we introduce a state-vector  $\mathbf{x}(t_i)$  which represents a vector of *all* grid point values of the prognostic state variables in space at a given time  $t_i$ . The ocean state  $\mathbf{x}(t_i)$  will depend on the state at some earlier time  $\mathbf{x}(t_{i-1})$  and upon the ocean surface forcing  $\mathbf{f}(t_i)$  and boundary conditions  $\mathbf{b}(t_i)$  over the time interval  $[t_{i-1}, t_i]$ . The discrete equations of motion will in general be nonlinear, and can be represented by the discrete nonlinear operator  $M$ . Thus the time evolution of the ocean state by an ocean model can be expressed in a convenient and compact form as:

$$\mathbf{x}(t_i) = M(t_i, t_{i-1})(\mathbf{x}(t_{i-1}), \mathbf{f}(t_i), \mathbf{b}(t_i)) \quad (14)$$

where  $M(t_i, t_{i-1})$  denotes a forward integration of the nonlinear ocean model from time  $t_{i-1}$  to  $t_i$  and for convenience  $\mathbf{f}(t_i)$  and  $\mathbf{b}(t_i)$  denote the forcing and boundary conditions over the entire interval  $[t_{i-1}, t_i]$ . This notation is fairly standard in numerical weather prediction and ocean modeling (Ide et al, 1997).

#### 3.2 The Incremental Formulation

As discussed in Chapter ??, the aim of data assimilation is to construct an estimate of the ocean circulation by combining *prior* information from an ocean model with observations. The solution of an ocean model over the interval  $t=[t_0, t_N]$  is uniquely determined by the initial conditions,  $\mathbf{x}(t_0)$ , the surface forcing,  $\mathbf{f}(t)$ , and the boundary conditions,  $\mathbf{b}(t)$ , collectively referred to as control variables. Therefore, *prior* estimates of all control variables are required, and will be denoted  $\mathbf{x}_b(t_0)$ ,  $\mathbf{f}_b(t)$  and  $\mathbf{b}_b(t)$  respectively. Our hypothesis about the uncertainty associated with each *prior* is embodied in the *prior* error covariance matrices for the initial conditions,  $\mathbf{B}$ , the surface forcing,  $\mathbf{B}_f$ , and boundary conditions,  $\mathbf{B}_b$ . For conven-

ience, we will denote by  $\mathbf{D}$  the block diagonal covariance matrix comprised of  $\mathbf{B}$ ,  $\mathbf{B}_f$ , and  $\mathbf{B}_b$ , namely  $\mathbf{D}=\text{diag}(\mathbf{B},\mathbf{B}_f,\mathbf{B}_b)$ . Similarly, we will denote by  $\mathbf{y}$  the vector comprised of all observations in the interval  $t=[t_0,t_N]$  with the associated observation error covariance matrix  $\mathbf{R}$ . The goal of data assimilation is then to identify the control vector  $\mathbf{z} = (\mathbf{x}^\top(t_0), \mathbf{f}^\top(t_0), \mathbf{f}^\top(t_1), \dots, \mathbf{f}^\top(t_N), \mathbf{b}^\top(t_0), \mathbf{b}^\top(t_1), \dots, \mathbf{b}^\top(t_N))^\top$  that maximizes the conditional probability  $p(\mathbf{z} | \mathbf{y}) \propto e^{-J_{NL}}$ , where:

$$J_{NL}(\mathbf{z}) = \frac{1}{2} \mathbf{z}^\top \mathbf{D}^{-1} \mathbf{z} + \frac{1}{2} (\boldsymbol{\phi} - \mathbf{y})^\top \mathbf{R}^{-1} (\boldsymbol{\phi} - \mathbf{y}) \quad (15)$$

and  $\mathbf{z}_b = (\mathbf{x}_b^\top(t_0), \mathbf{f}_b^\top(t_0), \mathbf{f}_b^\top(t_1), \dots, \mathbf{b}_b^\top(t_0), \mathbf{b}_b^\top(t_1), \dots)^\top$  is the vector of *priors*, and  $\boldsymbol{\phi}$  is the vector of the model equivalent of the observations at the observation times and locations. Each element of  $\boldsymbol{\phi}$  is of the form  $H_j(\mathbf{x}(t_j))$  where  $H_j$  is the operator that transforms or interpolates the state-vector  $\mathbf{x}(t)$  to the observation points at time  $t_j$ . The scalar  $J_{NL}$  is called the penalty function or cost function.

It is common to linearize (15) by considering small increments  $\delta \mathbf{z}$  to the *prior*  $\mathbf{z}_b$ , so that  $\mathbf{z} = \mathbf{z}_b + \delta \mathbf{z}$  (Courtier et al., 1994). The assumption underlying this approximation is that  $\mathbf{z}_b$  does not lie too far from the true state of the ocean, in which case the goal of data assimilation then becomes one of finding  $\delta \mathbf{z} = (\delta \mathbf{x}^\top(t_0), \delta \mathbf{f}^\top(t_0), \delta \mathbf{f}^\top(t_1), \dots, \delta \mathbf{b}^\top(t_0), \delta \mathbf{b}^\top(t_1), \dots)^\top$ , the vector of control variable increments, that minimizes the linearized form of (15), namely:

$$J(\delta \mathbf{z}) = \frac{1}{2} \delta \mathbf{z}^\top \mathbf{D}^{-1} \delta \mathbf{z} + \frac{1}{2} (\mathbf{G} \delta \mathbf{z} - \mathbf{d})^\top \mathbf{R}^{-1} (\mathbf{G} \delta \mathbf{z} - \mathbf{d}) \quad (16)$$

where the matrix  $\mathbf{G}$  is the operator that maps the model increments to the observation points. The vector  $\mathbf{d} = \mathbf{y} - \boldsymbol{\phi}_b$  is called the innovation vector, where  $\boldsymbol{\phi}_b$  is the vector ( $H_j(\mathbf{x}_b(t_j))$ ). The increment  $\delta \mathbf{z}_a$  that minimizes  $J$  corresponds to the maximum value of  $p(\mathbf{z} | \mathbf{y})$ , and is often called a maximum likelihood estimate. The maximum likelihood ocean state estimate is then given by the analysis or so-called *posterior*  $\mathbf{z}_a = \mathbf{z}_b + \delta \mathbf{z}_a$ . In the event that the *prior* hypotheses embodied in  $\mathbf{D}$  are correct, then the theoretical minimum value of the cost/penalty function is  $J_{min} = N_{obs}$ , the total number of observations (Bennett, 2002).

The *prior* circulation estimate  $\mathbf{x}_b(t)$  during the interval  $t=[t_0,t_N]$  is assumed to be a solution of the model equations (14) forced by the *prior*  $\mathbf{f}_b(t)$  and subject to the *prior* boundary conditions  $\mathbf{b}_b(t)$ . Under the assumption that the *prior* is already a good estimate of the circulation, the increments  $\delta \mathbf{z}$  will be small compared to  $\mathbf{z}_b$ , in which case a good approximation for  $\delta \mathbf{x}(t)$  will be the first-order Taylor expansion of (14), namely:

$$\delta \mathbf{x}(t_i) = \mathbf{M}(t_i, t_{i-1}) \delta \mathbf{u}(t_{i-1}) \quad (17)$$

where  $\mathbf{M}(t_i, t_{i-1})$  represents the linearization of  $M$  in (14) about the time evolving prior  $\mathbf{x}_b(t)$ , and  $\delta \mathbf{u}(t_{i-1}) = (\delta \mathbf{x}^\top(t_{i-1}), \delta \mathbf{f}^\top(t_{i-1}), \delta \mathbf{b}^\top(t_{i-1}))^\top$ . Equation (17) is referred to as the tangent linear model, since solutions  $\delta \mathbf{x}(t_i)$  are locally tangent to the solution  $\mathbf{x}_b(t_i)$  of (14). The operator  $\mathbf{G}$  in (16) is a combination of the tangent linear model  $\mathbf{M}$  and  $\mathbf{H}$  (the linearization of the observation operator  $H$ ), and represents solutions of (17) evaluated or mapped to the observation points.

The increment  $\delta \mathbf{z}_a$  corresponding to the most likely state estimate satisfies the condition  $\partial J / \partial \delta \mathbf{z} = 0$ , and is given by  $\delta \mathbf{z}_a = \mathbf{K} \mathbf{d}$  where  $\mathbf{K}$  is called the gain matrix and given by:

$$\mathbf{K} = (\mathbf{D}^{-1} + \mathbf{G}^\top \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{R}^{-1}. \quad (18)$$

Equivalently, the gain matrix can also be written as:

$$\mathbf{K} = \mathbf{D} \mathbf{G}^\top (\mathbf{G} \mathbf{D} \mathbf{G}^\top + \mathbf{R})^{-1}. \quad (19)$$

In both (18) and (19), evaluation of the gain matrix involves a matrix inverse. In (18), the matrix to be inverted is  $(\mathbf{D}^{-1} + \mathbf{G}^\top \mathbf{R}^{-1} \mathbf{G})$  and has the dimension of  $\delta \mathbf{z}$  which will generally be greater than the number of model grid points and may be very large and a challenge to invert. The dimension of  $\delta \mathbf{z}$  is  $N_m = (N_x + N_f + N_b)$  where  $N_x$  is the dimension of  $\mathbf{x}$ , and  $N_f$  and  $N_b$  are the dimensions of  $\mathbf{f}$ , and  $\mathbf{b}$ , respectively, multiplied by the number of model time steps  $N$  in the interval  $t = [t_0, t_N]$ . In addition the expression in (18) in parentheses involves  $\mathbf{D}^{-1}$  which may also be difficult to evaluate. Equation (18) is often referred to as the primal form of the gain matrix.

Alternatively the matrix to be inverted in (19) is  $(\mathbf{G} \mathbf{D} \mathbf{G}^\top + \mathbf{R})$  which has a dimension equal to that of  $\mathbf{y}$ , the number of observations,  $N_{obs}$ . In general,  $N_{obs} \ll N_m$ , so (19) may be more convenient to use than (18). Equation (19) is often referred to as the dual form of the gain matrix. In practice, both (18) and (19) are used for ocean data assimilation and in either case the most likely state of the ocean circulation is given by  $\mathbf{z}_a = \mathbf{z}_b + \mathbf{K} \mathbf{d}$ .

Exercise 4: Prove that  $\mathbf{z}_a = \mathbf{z}_b + \mathbf{K} \mathbf{d}$  when  $\partial J / \partial \delta \mathbf{z} = 0$ , and that  $\mathbf{K}$  is given by (18).

**Exercise 5:** Using the identity  $(\mathbf{A}+\mathbf{B})^{-1}=\mathbf{A}^{-1}(\mathbf{A}^{-1}+\mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}$ , show that  $\mathbf{K}$  can also be expressed in the dual form (19).

Regardless of whether the primal or dual form of  $\mathbf{K}$  is used, the matrix inverse in (18) and (19) is never explicitly evaluated. Instead, an equivalent system of linear equations is solved, and the adjoint  $\mathbf{G}^T$  of the tangent linear model sampled at the observation points,  $\mathbf{G}$ , plays a crucial role in this process. The adjoint of the tangent linear model can be expressed symbolically as:

$$\delta \mathbf{u}^+(t_{i-1}) = \mathbf{M}^T(t_{i-1}, t_i) \delta \mathbf{x}^+(t_i) \quad (20)$$

where the reversed order of the time arguments of  $\mathbf{M}^T$  compared to  $\mathbf{M}$  in (17) indicates that time is integrated backwards in time as in the shallow water examples of section 2.3. Since identification of the most likely increment  $\mathbf{Kd}$  is equivalent to identifying the condition  $\partial J / \partial \delta \mathbf{z} = 0$ , the resulting data assimilation methods are collectively referred to as 4-dimensional variational (4D-Var) data assimilation, where the four dimensions are space and time.

As in section 2.1, there are two fundamental spaces; primal space of dimension  $N_m$ , and dual space of dimension  $N_{obs}$ . The tangent linear operator  $\mathbf{G}$  maps a vector from primal space to dual space, while the adjoint of the tangent linear operator  $\mathbf{G}^T$  maps a vector from dual space to primal space.

### 3.3 Primal Space 4D-Var

It is important to realize that while (18) and (19) are written in terms of matrix products, the matrices involved are never explicitly computed, and all matrix manipulations are performed using models, including models for  $\mathbf{D}$  and  $\mathbf{R}$ . This leads to very useful iterative algorithms which can be used to identify the minimum of  $J(\delta \mathbf{z})$  or  $J_{NL}(\mathbf{z})$ . With this in mind, consider the derivative of  $J(\delta \mathbf{z})$  in (16) with respect to  $\delta \mathbf{z}$ :

$$\partial J / \partial \delta \mathbf{z} = \mathbf{D}^{-1} \delta \mathbf{z} + \mathbf{G}^T \mathbf{R}^{-1} (\mathbf{G} \delta \mathbf{z} - \mathbf{d}). \quad (21)$$

Equation (21) shows that the gradient of the cost/penalty function with respect to  $\delta \mathbf{z}$  can be evaluated by (i) running the adjoint of the tangent linear model,  $\mathbf{G}^T$ , forced by  $\mathbf{R}^{-1}(\mathbf{G} \delta \mathbf{z} - \mathbf{d})$ , the weighted difference between the solution of the tangent linear model sampled at the observation points,  $\mathbf{G} \delta \mathbf{z}$ , and the innovation vector,  $\mathbf{d}$ , and (ii) adding  $\mathbf{D}^{-1} \delta \mathbf{z}$  to the result of (i). Therefore, the cost/penalty function (16) can be minimized iteratively as follows:

1. Run the nonlinear model (14) with the *prior* initial conditions,  $\mathbf{x}_b(t_0)$ , *prior* forcing,  $\mathbf{f}_b(t)$ , and *prior* boundary conditions,  $\mathbf{b}_b(t)$ , and compute the *prior* circulation estimate  $\mathbf{x}_b(t)$  for the interval  $t=[t_0, t_N]$ .
2. Run the tangent linear model (17) linearized about  $\mathbf{x}_b(t)$  from step 1 for the interval  $t=[t_0, t_N]$ , compute  $J(\delta\mathbf{z})$  from (16), and compute and save  $\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{z}-\mathbf{d})$ .
3. Run the adjoint of the tangent linear model (20) *backwards* in time linearized about  $\mathbf{x}_b(t)$  from step 1 and forced by  $\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{z}-\mathbf{d})$  from step 2 for the interval  $t=[t_N, t_0]$ .
4. Add to the adjoint solution from step 3 the vector  $\mathbf{D}^{-1}\delta\mathbf{z}$  to yield  $\partial J/\partial\delta\mathbf{z}$  according to (21).
5. Using the cost function gradient  $\partial J/\partial\delta\mathbf{z}$  from step 4, use a conjugate gradient method to identify a new  $\delta\mathbf{z}$  that will reduce the value of  $J$  resulting from a subsequent run of the tangent linear model as in step 2.
6. Using the new  $\delta\mathbf{z}$  from step 5, repeat steps 2-5 until the minimum of  $J$  has been identified.

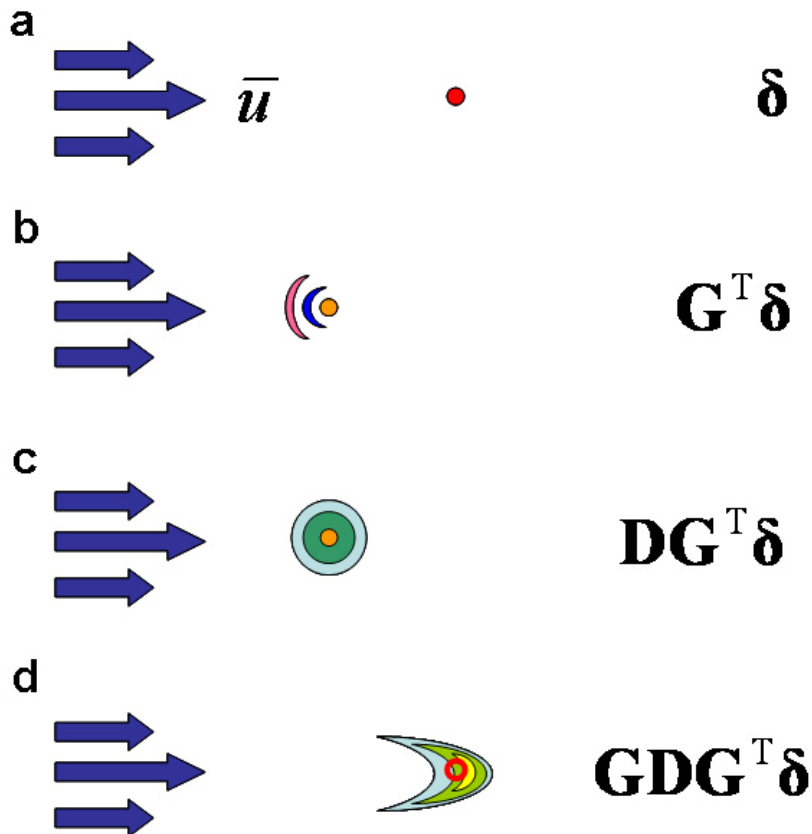
At the minimum of  $J$ , the gradient  $\partial J/\partial\delta\mathbf{z}=0$  and the iterative procedure described by steps 1-6 is equivalent to evaluating  $\mathbf{Kd}$  using the primal form of the gain matrix in (18).

### 3.4 Dual Space 4D-Var

In the dual formulation of the optimal increment  $\mathbf{Kd}$ , the matrix in parentheses in (19) is inverted by introducing an intermediate variable  $\mathbf{w}=(\mathbf{GDG}^T+\mathbf{R})^{-1}\mathbf{d}$ , where  $\mathbf{Kd}=\mathbf{DG}^T\mathbf{w}$ . In practice,  $\mathbf{w}$  is identified by solving the linear system  $(\mathbf{GDG}^T+\mathbf{R})\mathbf{w}=\mathbf{d}$  using an iterative conjugate gradient method to minimize the function  $I=1/2\mathbf{w}^T(\mathbf{GDG}^T+\mathbf{R})\mathbf{w}-\mathbf{w}^T\mathbf{d}$ , and  $\mathbf{w}$  plays the role of a generating function as discussed in section 2.1. The increment  $\mathbf{Kd}$  is then identified according to  $\mathbf{DG}^T\mathbf{w}$ . The following iteration algorithm is typical of those commonly used:

1. Run the nonlinear model (14) with the *prior* initial conditions,  $\mathbf{x}_b(t_0)$ , *prior* forcing,  $\mathbf{f}_b(t)$ , and *prior* boundary conditions,  $\mathbf{b}_b(t)$ , and compute the *prior* circulation estimate  $\mathbf{x}_b(t)$  for the interval  $t=[t_0, t_N]$ .
2. Run the adjoint of the tangent linear model (20) *backwards* in time linearized about  $\mathbf{x}_b(t)$  from step 1 and forced by  $\mathbf{w}$  for the interval  $t=[t_N, t_0]$  to yield  $\mathbf{G}^T\mathbf{w}$ .
3. Apply the *prior* covariance  $\mathbf{D}$  to the adjoint model solution from step 2 at  $t=0$  to yield  $\mathbf{DG}^T\mathbf{w}$ .
4. Run the tangent linear model (17) linearized about  $\mathbf{x}_b(t)$  from step 1 for the interval  $t=[t_0, t_N]$  using the result of step 3 as the initial condition to yield  $\mathbf{GDG}^T\mathbf{w}$ .

5. Add  $\mathbf{R}\mathbf{w}$  to the result of step 5, and evaluate the gradient  $\partial I / \partial \mathbf{w} = (\mathbf{GDG}^T + \mathbf{R})\mathbf{w} - \mathbf{d}$ .
6. Using the gradient  $\partial I / \partial \mathbf{w}$  from step 5, use a conjugate gradient method to identify a new  $\mathbf{w}$  that will reduce the value of  $I$  resulting from a subsequent reevaluation of steps 2-4.
7. Using the new  $\mathbf{w}$  from step 6, repeat steps 2-6 until the minimum of  $I$  has been identified.
8. Having identified the  $\mathbf{w} = \mathbf{w}_a$  that minimizes  $I$ , evaluate the increment  $\delta \mathbf{z}_a = \mathbf{Kd} = \mathbf{DG}^T \mathbf{w}_a$  by repeating steps 2 and 3 using  $\mathbf{w}_a$ .



**Fig. 1:** A schematic illustrating the action of the adjoint, *prior* error covariance, and tangent linear operators on a  $\delta$  function located at the site of a single observation in a steady zonal shear flow represented by the blue arrows. (a) The initial location of the  $\delta$  function in dual space. (b) The action of the adjoint operator  $\mathbf{G}^T$  which propagates the  $\delta$  function backwards in time “upstream” and maps it into primal space. (c) The *prior* error covariance matrix smoothes  $\mathbf{G}^T \delta$  in primal space. (d) The final action of tangent linear operator  $\mathbf{G}$  propagates the smoothed field from step (c) forward in time and maps the field back to dual space, indicated by the red open circle located at the observation point.

The connection between primal and dual space, and the role that  $\mathbf{G}$  and  $\mathbf{G}^T$  play in transforming between one space and the other is probably best illustrated by considering the operations represented by steps 2, 3 and 4 when applied to a *single* observation. Consider a  $\delta$ -function at the time and location of the single observation, in which case steps 2-4 yield  $\mathbf{G}\mathbf{D}\mathbf{G}^T\delta$ . The sequence of operations that lead to this result are illustrated schematically in Fig. 1 for the case of a mean geostrophic flow in the form of a jet in the zonal channel considered in section 2.3.2. In this example, advection is the dominant dynamical process, although there will be some wave propagation as well as illustrated in Fig. 1b. Hence the actions of  $\mathbf{G}^T$  and  $\mathbf{G}$  primarily advect information upstream and downstream respectively.

### 3.5 Computation of $z_a$

Having identified the increment control vector  $\delta z_a$  using either the primal or dual form of 4D-Var, it then remains to compute the most likely circulation estimate  $\mathbf{x}_a(t) = \mathbf{x}_b(t) + \delta \mathbf{x}_a(t)$  over the interval  $t = [t_0, t_N]$ . Two approaches are generally used: (i) using the non-linear model (14) to advance the circulation  $\mathbf{x}(t_0) = \mathbf{x}_b(t_0) + \delta \mathbf{x}_a(t_0)$  forward in time using  $\mathbf{f}_b(t) + \delta \mathbf{f}_a(t)$  and  $\mathbf{b}_b(t) + \delta \mathbf{b}_a(t)$ , or (ii) using the tangent linear model (17) forced by  $\delta \mathbf{f}_a(t)$  and subject to  $\delta \mathbf{b}_a(t)$  to advance the increments  $\delta \mathbf{x}(t)$  in time. Approach (i) is generally used in primal space applications of 4D-Var (Courtier et al, 1994), while both (i) and (ii) are used in dual space formulations (Da Silva et al; 1995; Egbert et al, 1994). To the author’s knowledge there are primal formulations that use (ii).

In the dual formulation, each element of the innovation vector  $\mathbf{d}$  is assumed to be a linear combination of the elements of the state-vector increment  $\delta \mathbf{x}$ , according to the tangent linear operator  $\mathbf{G}$ , and it is the generating function  $\mathbf{w}$  of section 3.4 that identifies the activated part of primal space into which  $\mathbf{d}$  maps. All linear functions of  $\delta \mathbf{x}$  are known as the dual of  $\delta \mathbf{x}$ , and possess the important property that the tangent linear model equivalent of each element  $d_j$  of  $\mathbf{d}$  can be expressed as  $\mathbf{r}_j^T \delta \mathbf{x}$ , the so-called Riesz representation theorem. The vectors  $\mathbf{r}_j$  are called representer functions, and approach (ii) in the dual formulation is equivalent to ex-

pressing the increment  $\delta \mathbf{x}_a(t)$  as  $\mathbf{S}\mathbf{c}$ , where  $\mathbf{S}$  is a  $N_m \times N_{obs}$  matrix, and each column of  $\mathbf{S}$  is a representer function  $\mathbf{r}$ . The vector  $\mathbf{c}$  is comprised of the weights assigned to each of the  $N_{obs}$  representers (Bennett, 2002). In the schematic of Fig. 1d, the pattern of colored contours is a representer function, and the open red circle is  $\mathbf{GDG}^T \boldsymbol{\delta}$ , the representer function sampled at the observation point.

### 3.6 Strong constraint versus weak constraint

In the formulations of 4D-Var presented so far, it has been implicitly assumed that the most likely circulation estimate  $\mathbf{x}_a(t)$  is an exact solution of the nonlinear model equations (14). This is tantamount to assuming that the model is perfect and free of errors, and the increments  $\delta \mathbf{z}_a(t)$  that minimize the cost/penalty function  $J$  in (16) are said to be subject to the ‘‘strong constraint’’ imposed by model dynamics (Sasaki, 1970). Of course, *all* models possess errors and uncertainties, and to account for these it is necessary to augment the control vector of increments  $\delta \mathbf{z}$  so that:

$$\delta \mathbf{z} = \left( \delta \mathbf{x}^T(t_0), \delta \mathbf{f}^T(t_0), \delta \mathbf{f}^T(t_1), \dots, \delta \mathbf{b}^T(t_0), \delta \mathbf{b}^T(t_1), \dots, \boldsymbol{\eta}(t_0), \boldsymbol{\eta}(t_1), \dots \right)^T \quad (22)$$

where  $\boldsymbol{\eta}(t)$  represents the corrections for model error at each grid point and time step. The *prior* for model error is assumed to be  $\mathbf{0}$  with an associated error covariance matrix  $\mathbf{Q}$ . In the presence of model error, the development of 4D-Var in primal and dual space of sections 3.1-3.5 is unchanged, except that now the *prior* error covariance matrix is given by the block diagonal matrix  $\mathbf{D} = \text{diag}(\mathbf{B}, \mathbf{B}_f, \mathbf{B}_b, \mathbf{Q})$ . The most likely circulation estimate  $\mathbf{x}_a(t)$  in this case is no longer an exact solution of the non-linear model equations (14), and the increments  $\delta \mathbf{z}_a(t)$  that minimize the cost/penalty function  $J(\delta \mathbf{z})$  in (16) are said to be subject to the ‘‘weak constraint’’ imposed by model dynamics (Sasaki, 1970).

Under the weak constraint, the dimension,  $N_m$ , of primal space increases by  $N_x N_t$ , and in general the weak constraint 4D-Var problem becomes intractable. However, the dimension of dual space is unchanged by the imposition of the weak constraint, so weak constraint 4D-Var is often performed using the dual formulation.

### 3.7 Inner- and Outer-loops

Following the incremental formulation of 4D-Var described in section 3.2, the most likely circulation that minimizes the cost function  $J(\delta \mathbf{z})$  is given by (16). However, the assumption underlying (16) is that the *prior* is close to the true cir-

ulation. This of course is a big assumption, and in all likelihood will not always be true, if ever. Therefore, it is preferable to identify the most likely circulation estimate that minimizes instead the cost function  $J_{NL}(\mathbf{z})$  in (15). In practice,  $J_{NL}(\mathbf{z})$  is a non-quadratic function of  $\mathbf{z}$  because the state-vector is a solution of the nonlinear model (14). As a result,  $J_{NL}(\mathbf{z})$  may possess multiple minima, and the global minimum corresponding to the most likely circulation may be difficult to identify. However, a common technique for identifying the minima of (15) is by solving a sequence of linear minimizations of the form (16), where each member of the sequence is referred to as an ‘‘outer-loop.’’ During each outer-loop,  $J(\delta\mathbf{z})$  given by (16) is minimized using the iterative algorithms of section 3.3 or section 3.4, and each iteration is called an ‘‘inner-loop.’’

During the first outer-loop, the tangent linear model (17) and adjoint model (20) are linearized about the *prior* circulation estimate  $\mathbf{x}_b(t)$  over the interval  $t=[t_0, t_N]$ . At the end of the first outer-loop, the circulation estimate is updated using approach (i) or (ii) of section 3.5, to yield the state vector,  $\mathbf{x}_1(t)=\mathbf{x}_b(t)+\delta\mathbf{x}_1(t)$ , forcing,  $\mathbf{f}_1(t)=\mathbf{f}_b(t)+\delta\mathbf{f}_1(t)$ , boundary conditions,  $\mathbf{b}_1(t)=\mathbf{b}_b(t)+\delta\mathbf{b}_1(t)$ , and in the weak constraint case the corrections for model error,  $\boldsymbol{\eta}_1(t)$ , where the subscript ‘‘1’’ refers to the outer-loop number. During the second outer-loop, the tangent linear and adjoint models are linearized about  $\mathbf{x}_1(t)$ . Repeating the sequence of outer-loops, it is easy to see that during outer-loop  $n$ , the tangent linear and adjoint models are linearized about  $\mathbf{x}_{n-1}(t)$ , where  $\mathbf{x}_{n-1}(t)$  is the updated circulation estimate forced by  $\mathbf{f}_{n-1}(t)$ , and subject to  $\mathbf{b}_{n-1}(t)$  and  $\boldsymbol{\eta}_{n-1}(t)$ . It is important to note, however, that the innovation vector never changes, and  $\mathbf{d}=\mathbf{y}-(H_j(\mathbf{x}_b(t_j)))$  is *always* computed using the *prior* circulation estimate  $\mathbf{x}_b(t_j)$ .

The method used to compute the circulation estimate  $\mathbf{x}_n(t)$  at the end of each outer-loop varies. The most common approach is to use the non-linear model as in approach (i) of section 3.5 to advance  $\mathbf{x}_n(t_0)=\mathbf{x}_b(t_0)+\delta\mathbf{x}_n(t)$  forward in time. However, in the dual formulation in which the method of representers is used, approach (ii) is modified, and the finite-amplitude tangent linear model is used to advance  $\mathbf{x}_n(t_0)=\mathbf{x}_b(t_0)+\mathbf{S}\mathbf{c}$  forward in time. The finite-amplitude tangent linear model can be expressed symbolically as:

$$\mathbf{x}_n(t_i) = M(t_i, t_{i-1})(\mathbf{x}_{n-1}(t_{i-1}), \mathbf{f}_{n-1}(t_i), \mathbf{b}_{n-1}(t_i)) + \mathbf{M}_{n-1}(t_i, t_{i-1})(\mathbf{g}_n(t_{i-1}) - \mathbf{g}_{n-1}(t_{i-1})) \quad (23)$$

where  $\mathbf{g}_n(t_i) = (\mathbf{x}_n^\top(t_i), \mathbf{f}_n^\top(t_i), \mathbf{b}_n^\top(t_i))^\top$ , and  $\mathbf{M}_{n-1}$  denotes the tangent linear model linearized about  $\mathbf{x}_{n-1}(t)$ . During the first outer-loop, the solution of (23) reduces to the sum of the prior,  $\mathbf{x}_b(t)$ , and the tangent linear model solution (17). The representer approach to 4D-Var in dual space using the algorithm of section 3.4 for the inner-loops, and updating the circulation estimates in the outer-loops using (23) is equivalent to minimizing  $J_{NL}(\mathbf{z})$  in (15) by solving the non-linear Euler-Lagrange equations. Full details of this approach are beyond the scope of this article but can be found in Bennett (2002).

## 4 Examples of 4D-Var for the California Current

We will present here some illustrative examples of 4D-Var using both the primal and dual formulations applied to the California Current circulation using the Regional Ocean Modeling System (ROMS).

### 4.1 *The Regional Ocean Modeling System (ROMS)*

ROMS is a primitive equation ocean model that has gained considerable popularity in recent years because of the great flexibility that it affords for modeling different regions of the world ocean. ROMS uses a curvilinear orthogonal coordinate system in the horizontal, and terrain-following coordinates in the vertical, both of which allow for increased resolution in regions where it is most needed (e.g., in regions of complex topography and bathymetry, in shallow water, and near the ocean surface). ROMS is a hydrostatic model, and employs a wide range of user-controlled options for the numerics and physical parameterizations, as well as a range of options for prescribing open boundary conditions. A detailed description of ROMS is beyond the scope of this article, and the reader is referred to Shchepetkin and McWilliams (2005) and Haidvogel et al. (2000) for more information. ROMS is a community ocean model and is freely available from <http://myroms.org>.

### 4.2 *ROMS 4D-Var*

There are many practical aspects of 4D-Var that will not be discussed here, but which nonetheless are very important. The main features of the ROMS 4D-Var system are listed below, along with references that provide more information. In addition, the ROMS 4D-Var system is described in detail by Moore et al. (2009a,b). The main features and attributes of ROMS 4D-Var can be summarized as follows:

1. Incremental formulation (Courtier et al., 1994)
2. Primal and dual formulations (Courtier, 1997)
3. Primal formulation is referred to as Incremental 4D-Var, hereafter I4D-Var
4. Dual formulations following the Physical-space Statistical Analysis System of Da Silva et al. (1995), hereafter 4D-PSAS, and the indirect representer method of Egbert et al. (1994), hereafter R4D-VAR
5. Strong constraint formulation for both primal and dual formulations
6. Weak constraint for dual formulations only

7. Inner-loops use a preconditioned Lanczos formulation of the conjugate method (Golub and van Loan, 1989; Lorenc, 2003; Fisher and Courtier, 1995; Tshimanga et al., 2008)
8. *Prior* covariances  $\mathbf{D}$  are modeled using a pseudo-heat diffusion equation approach (Derber and Bouttier, 1999; Weaver and Courtier, 2001), and a multivariate balance operator (Weaver et al., 2005)
9. MPI parallel architecture

### 4.3 The California Current System

The California Current System (CCS) is an archetypal eastern boundary current regime that is dominated by mesoscale eddies, and characterized by a pronounced seasonal cycle of coastal upwelling and primary productivity. A comprehensive reviews of the CCS can be found in Hickey (1998).

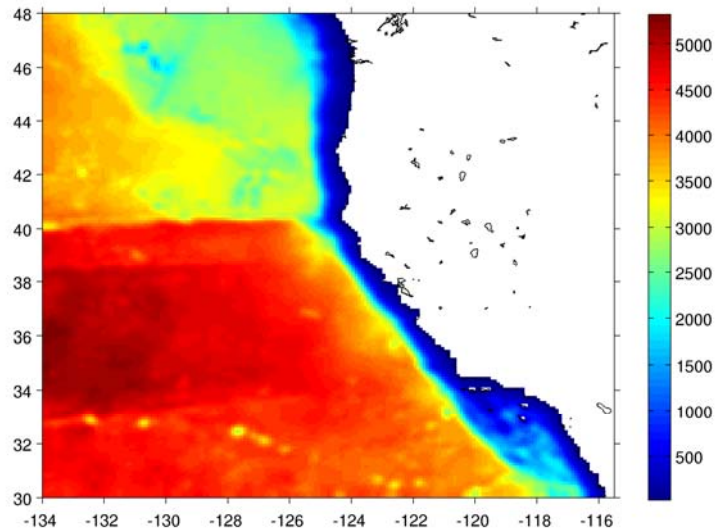
ROMS has been configured for the CCS (hereafter referred to as ROMS-CCS) and spans the domain 116W-134W, 30N-48N shown in Fig. 2. Several configurations of ROMS-CCS exist, with horizontal resolutions ranging from 3-30km, and with 30-42 levels in the vertical. ROMS-CCS is forced with near surface air data from the Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) of Doyle et al. (2009) which are converted to surface wind stress and surface fluxes of heat and freshwater using the bulk flux formulations of Fairall et al. (1996a,b) and Liu et al. (1979). The model domain has open boundaries at the northern, western, and southern edges, and at these boundaries the temperature, salinity, and velocity fields are prescribed using data from the Estimating the Circulation and Climate of the Ocean project (ECCO) (Wunsch and Heimbach, 2007) which is a global ocean data assimilation product at 1 degree resolution. The radiation conditions of Chapman (1985) and Flather (1976) are also imposed at the open boundaries on the free surface and vertically integrated velocity respectively. ROMS-CCS produces very good simulations of the CCS circulation as documented by Veneziani et al (2009).

### 4.4 ROMS-CCS 4D-Var Configuration

ROMS-CCS has been used extensively for 4D-Var as described by Broquet et al. (2009a,b,c) and Moore et al. (2009b). Some example 4D-Var calculations using ROMS-CCS are presented here to illustrate the ideas and concepts introduced in section 2 and 3.

Recall that for data assimilation we require *prior* estimates for the elements of the control vector  $\mathbf{z}$ , namely for the initial conditions,  $\mathbf{x}_b(t_0)$ , the surface forcing,

$\mathbf{f}_b(t)$ , and the boundary conditions,  $\mathbf{b}_b(t)$ , for the interval  $t=[t_0, t_N]$ . In the case of weak constraint 4D-Var, the *prior* for the corrections for model error is the null vector  $\mathbf{0}(t)$ . ROMS-CCS 4D-Var is usually run sequentially using data assimilation windows in time that span time intervals  $t_N-t_0$  that are typically 4-14 days long. Each data assimilation window is referred to as a “cycle,” and the initial condition *prior*,  $\mathbf{x}_b(t_0)$ , is the best circulation estimate from the end of the previous cycle. The *priors* for surface forcing,  $\mathbf{f}_b(t)$ , are the surface fluxes derived from the COAMPS air data, and the *priors* for the boundary conditions,  $\mathbf{b}_b(t)$ , are the open boundary data from ECCO.



**Fig. 2: The ROMS-CCS domain and bathymetry in meters on a grid with 10km horizontal resolution.**

In addition to the *prior* fields, *prior* error covariance matrices are also required, namely,  $\mathbf{B}$  for the initial conditions,  $\mathbf{B}_r$  for the surface forcing,  $\mathbf{B}_b$  for the boundary conditions, and  $\mathbf{Q}$  for the model errors in the case of weak constraint 4D-Var. In general, the *prior* error covariance matrices vary in time, but in practice they are assumed to be time invariant. Following Weaver et al. (2005), each of the *prior* error covariance matrices are factorized into a block diagonal, univariate correlation matrix, a diagonal matrix of standard deviations, and a multivariate dynamical balance operator. The univariate correlation matrix is modeled as the solution of a pseudo-heat diffusion equation (Weaver and Courtier, 2001). This is an involved procedure, and the prescription of each *prior* error covariance matrix is beyond the

scope of this presentation. However, a full description for ROMS-CCS can be found in Broquet et al. (2009). For the present purpose, it suffices to say that the *prior* error covariance matrices can be prescribed for each of the *prior* components of the control vector  $\mathbf{z}_b$ .

Observations from various platforms are assimilated into ROMS-CCS, including satellite derived sea surface temperature (SST) and sea surface height (SSH), and sub-surface hydrographic measurements of temperature,  $T$ , and salinity,  $S$ , collected from shipboard CTDs, XBTs, drifting Argo profiling floats, and tagged elephant seals. The data used were extracted from the rigorously quality controlled global ocean data archive of Ingleby and Huddleston (2007). The observation error covariance matrix,  $\mathbf{R}$ , is assumed to be time invariant and diagonal (i.e. spatially and temporally uncorrelated observation errors), and the error variances are instrument and variable dependent. The elements of  $\mathbf{R}$  are a reflection of several sources of error: instrument error, interpolation errors introduced by the observation operator  $\mathbf{H}$  of section 3.2, and errors of representativeness. The largest source of error is typically the error of representativeness, which is a measure of the uncertainty associated with the ability of a single ocean observation to describe the circulation in a single ocean model grid cell. The following observation error standard deviations were used in ROMS-CCS:  $\sim 2$  cm for SSH;  $\sim 0.4$ C for SST;  $\sim 0.1$ C for  $T$ ;  $\sim 0.01$  for  $S$ .

#### 4.4.1 Primal versus Dual 4D-Var

The first example considered will contrast the performance of the primal and dual formulations of ROMS 4D-Var in the CCS, and results are shown from the ROMS-CCS configuration with a horizontal resolution of 30 km and 30 levels in the vertical. The cost function from two strong constraint calculations using 1 outer-loop and 75 inner-loops, is shown Fig. 3 for a 4 day assimilation window spanning the period 1-4 July, 2000, during which time there were  $\sim 1 \times 10^4$  observations available, most in the form of satellite data.

In one calculation the primal formulation I4D-Var was used, while in the second calculation the dual formulation R4D-Var was employed. Figure 3 shows that in both cases the cost function decreases with an increasing number of inner-loop iterations. In the case of I4D-Var,  $J(\delta\mathbf{z})$  decreases monotonically, and asymptotes to a near constant value after about 40 or so inner-loops. R4D-Var on the other hand exhibits very quite different behavior in which  $J(\delta\mathbf{z})$  undergoes quite large fluctuations until after about 60 inner-loops it asymptotes to a similar near constant value as that reached by I4D-Var. The difference in behavior of  $J(\delta\mathbf{z})$  in the primal and dual formulations of 4D-Var is associated with the different approach employed to identify the cost function minimum. In the primal case,  $J(\delta\mathbf{z})$  is minimized directly as described in section 3.3, while in the dual formulation  $J(\delta\mathbf{z})$  is minimized indirectly using a generating function  $\mathbf{w}$  according to section 3.4. El Akkraoui and Gauthier (2009) have explored the different behavior of the primal

and dual formulations and suggest some effective remedies for increasing the rate of convergence of  $J(\delta\mathbf{z})$  in the dual case.

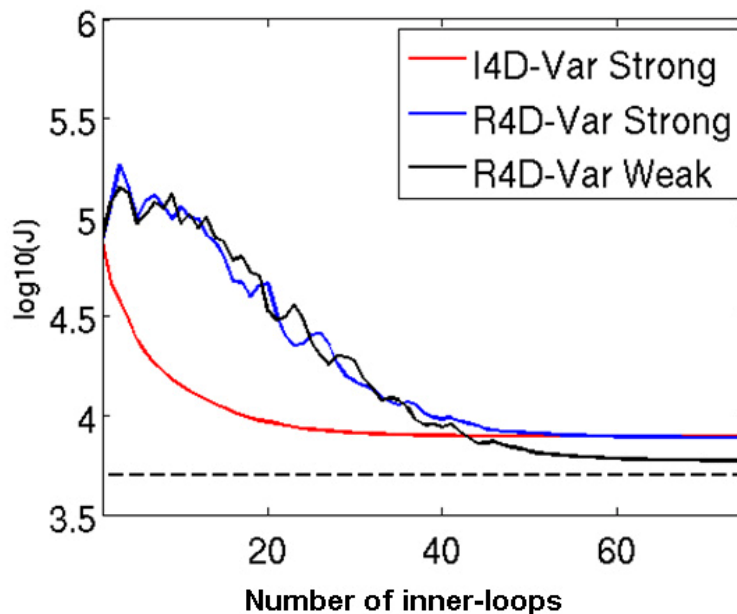


Fig. 3:  $\text{Log}_{10}(J(\delta\mathbf{z}))$  (from (16)) versus the number of inner-loops for three 4D-Var assimilation calculations for the period 1-4 July, 2000 for the case of 1 outer-loop. The three experiments shown are from a strong constraint calculation in primal space (I4D-Var, red curve), a strong constraint calculation in dual space (R4D-Var, blue curve), and a weak constraint calculation in dual space (R4D-Var, black curve). The dashed curve shows the theoretical minimum value of the cost function,  $J_{min}$ , that will be reached only if all of the *prior* hypotheses embodied in  $\mathbf{D}$  are correct.

Nonetheless, despite the remarkable difference between the primal and dual formulations in the rate of convergence of  $J(\delta\mathbf{z})$  to its asymptotic minimum value, Fig. 3 indicates that as expected from the equivalence of (18) and (19), both calculations yield the same solution. A comparison of the most likely CCS circulation estimates from 4D-Var in both cases (not shown) confirms that the primal and dual form ultimately yield the same circulation after 75 inner-loops. In both cases, however, the minimum value reached by  $J(\delta\mathbf{z})$  is larger than  $J_{min}$  indicating that either the *prior* hypotheses  $\mathbf{D}$  are incorrect, or that the global minimum value of  $J_{NL}(\mathbf{z})$  has not been reached.

$n \times m$	$1 \times 100$	$2 \times 50$	$3 \times 33$	$4 \times 25$	$10 \times 10$
$J_{NL}(\mathbf{z})$	$1.56 \times 10^4$	$1.18 \times 10^4$	$0.93 \times 10^4$	$0.87 \times 10^4$	$0.71 \times 10^4$

**Table 1:** Values of  $J_{NL}(\mathbf{z})$  from (15) that are reached at the end of a 4 day strong constraint I4D-Var data assimilation experiment spanning the period 1-4 July, 2000 for different combinations of the number of outer-loops,  $n$ , and number of inner-loops,  $m$ . In each case, the total number of iterations,  $n \times m$ , is the same. The value of  $J_{min}$  in each case is  $0.52 \times 10^4$ .

The influence of different combinations of the number of outer-loops and the number of inner-loops on the final value of  $J_{NL}(\mathbf{z})$  in (15) that is reached at the end of the 1-4 July 2000 strong constraint I4D-Var cycle is shown in Table 1. Table 1 reveals that updating the circulation estimate about which the tangent linear and adjoint models are linearized is clearly beneficial for reducing the value of  $J_{NL}(\mathbf{z})$  for the same computational effort (i.e. when the total combined number of outer-loops and inner-loops is the same). In addition, increasing the number of inner-loops yields a value of  $J_{NL}(\mathbf{z})$  that is closer to  $J_{min}$ .

An example of the overall performance of 4D-Var when applied sequentially is shown in Fig. 4 which shows the root mean square (rms) difference between ROMS-CCS and the observations for SST and SSH. The result from three calculations are shown corresponding to: (i) no data assimilation, when ROMS-CCS is forced only by the *prior* forcing  $\mathbf{f}_b(t)$  and subject to the *prior* boundary conditions  $\mathbf{b}_b(t)$ ; (ii) primal I4D-Var applied sequentially using 14 day assimilation windows; and (iii) the *prior* circulation estimates  $\mathbf{x}_b(t)$  which arise from runs of the model initialized from the most likely circulation estimate at the end of the previous assimilation cycle. The period considered in each case is 1 Jan 1999-31 Dec 2004. Figure 4 shows that I4D-Var has a positive impact on the agreement of the model to the observations, both during each data assimilation cycle, and on the *prior*.

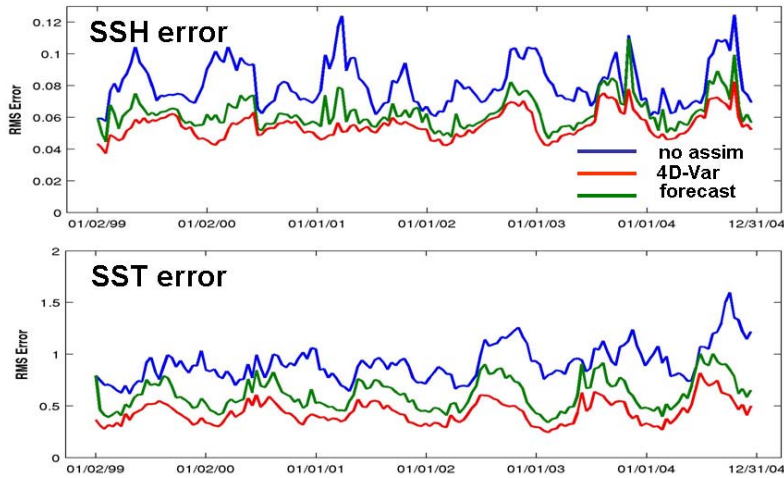


Fig. 4: Time series of the root mean square (rms) difference between the model and observations for SST and SSH for period Jan 1999 – Dec 2004. The blue curve shows the rms differences averaged over 14 day intervals from the model forced only by the *prior* forcing and subject to the *prior* boundary conditions. The red curve shows results from a case where primal I4D-Var is applied sequentially over 14 day assimilation windows, and the rms differences are the average over each 14 day window. The green curve shows the rms difference associated with the *prior* circulation estimate  $x_b(t)$ , and is also a 14 day average. Since the *prior* circulation is the result of a model run started from the most likely circulation estimate at the end of each I4D-Var window, it can also be thought of as a “forecast” of the circulation for the subsequent 14 day period.

#### 4.4.2 The Control Vector Influence

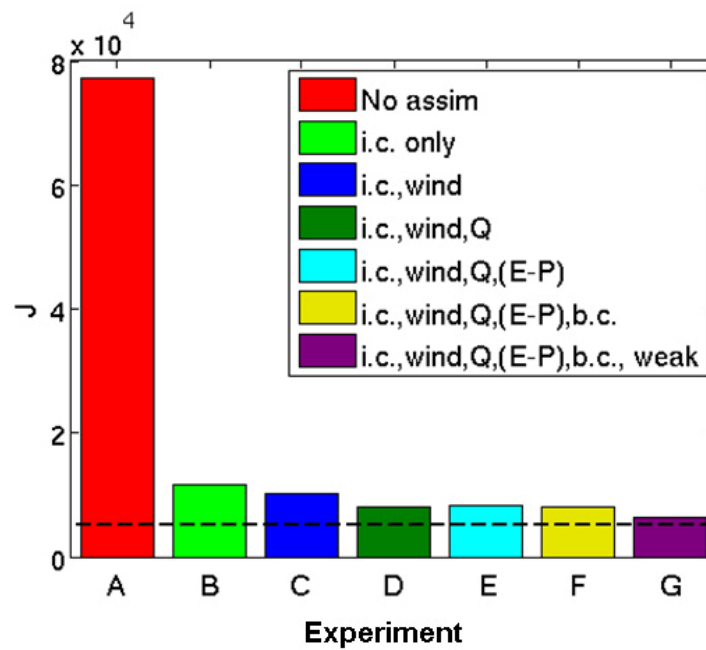


Fig. 5: The cost function  $J(\delta z)$  in (16) after 1 outer-loop and 75 inner-loops from several dual R4D-Var experiments for the 4 day period 1-4 July, 2000. Expt. A: no data assimilation; Expt. B:  $\delta z$  comprised of the initial condition increments  $\delta x(t_0)$  only; Expt. C:  $\delta z$  comprised of only  $\delta x(t_0)$  and the surface wind stress components of  $\delta \mathbf{f}(t)$ ; Expt. D:  $\delta z$  comprised of only  $\delta x(t_0)$  and the surface wind stress and heat flux components of  $\delta \mathbf{f}(t)$ ; Expt. E:  $\delta z$  comprised of  $\delta x(t_0)$  and all components of  $\delta \mathbf{f}(t)$ ; Expt. F:  $\delta z$  comprised of  $\delta x(t_0)$ ,  $\delta \mathbf{f}(t)$ , and  $\delta \mathbf{b}(t)$ ; Expt. G: same as F except using the weak constraint as described in section 4.4.3. The dashed line indicates the theoretical minimum value of the cost function,  $J_{min}$ .

In the strong constraint examples of 4D-Var presented in section 4.4.1, the control vector  $\mathbf{z}$  was comprised of the model initial conditions,  $\mathbf{x}(t_0)$ , the surface forcing,  $\mathbf{f}(t)$ , and the boundary conditions,  $\mathbf{b}(t)$ , for the interval  $t=[t_0, t_N]$ . However, it is instructive to explore the relative impact of each component of  $\mathbf{z}$  on the cost function. To this end, Fig. 5 shows the results of a sequence of 4D-Var calculations in which the control vector was successively augmented with different control variables.

Figure 5 reveals that as control vector increment  $\delta\mathbf{z}$  is augmented with more components of the surface forcing and boundary conditions, the minimum value reached by the cost function  $J(\delta\mathbf{z})$  progressively decreases. This is because the length of  $\delta\mathbf{z}$  is increasing at the same time meaning that there are more degrees of freedom available to 4D-Var with which to fit the observations. It is important to realize, however, that the 4D-Var process is not linear, so if the order of the calculations in Fig. 5 is changed, the same progression of changes in  $J(\delta\mathbf{z})$  will not necessarily be obtained. Nonetheless, Fig. 5 does suggest that accounting for the errors and uncertainties in the initial conditions has by far the largest impact on the efficacy of the circulation estimate measured in terms of  $J(\delta\mathbf{z})$ . Significant further reductions in  $J(\delta\mathbf{z})$  are afforded when uncertainties in surface forcing are accounted for also. However, Fig. 5 suggests that uncertainties in the boundary conditions have the least impact on  $J(\delta\mathbf{z})$ .

#### 4.4.3 Weak Constraint 4D-Var

Model errors are a significant source of uncertainty in model-derived circulation estimates, and it is important to account for these uncertainties during 4D-Var. However, quantifying and identifying the sources of model error is one of the greatest challenges in ocean data assimilation. In ROMS-CCS, we have made some progress towards identifying some of the most significant impacts of model error, but our understanding is far from complete.

A dominant feature of the CCS circulation is the occurrence of coastal upwelling in the spring and summer along much of the California, Oregon and Washington coasts which is driven by equatorward alongshore winds. ROMS-CCS simulates the seasonal cycle of upwelling very well (Veneziani et al., 2009) although in the absence of data assimilation, comparisons of the model with observations indicates that the model SST is biased towards temperatures that colder than observed during the peak of the upwelling season. Independent investigations of the quality of the surface wind forcing by Doyle et al. (2009) have revealed that the COAMPS *priors*,  $\mathbf{f}_b(t)$ , used to drive ROMS-CCS agree well with wind observations from satellite scatterometers and ocean buoys. Therefore, it seems likely that the cold bias in ROMS-CCS SST is associated with model errors rather than errors in the surface forcing. Further evidence for this hypothesis is provided by the work of Broquet et al. (2009b,c) using strong constraint 4D-Var. They found that data assimilation led to a reduction in strength of the upwelling favorable alongshore

surface winds during spring and summer, and a general degradation of the wind *prior* when compared to satellite-derived surface wind observations. An important conclusion of their work, therefore, is that in the absence of corrections for model error, data assimilation may yield undesirable and non-physical corrections to potentially all elements of the control vector in an attempt to minimize the cost function.

Based on the findings of Broquet et al. (2009b,c) attempts are currently underway to account for model error in ROMS-CCS during 4D-Var using the weak constraint. The impact of imposing a weak constraint during R4D-Var on the cost function  $J(\delta\mathbf{z})$  is also shown in Fig. 3. Based on the known bias of the ROMS-CCS SST in spring and summer, it was assumed that the model error affecting SST is present mainly in the model temperature equation, and confined to within a short distance of the coast. To account for such errors during 4D-Var, a *prior* for the model error covariance matrix of the form  $\mathbf{Q}=\mathbf{W}\mathbf{B}$  was assumed, where  $\mathbf{B}$  is the initial condition *prior* error covariance matrix, and  $\mathbf{W}$  is a diagonal rescaling matrix. All elements of  $\mathbf{W}$  are zero except those corresponding to ocean temperature grid points within 300 km of the North American coast. The non-zero elements of  $\mathbf{W}$  are of the form  $\alpha(1-d/300)$ , where  $\alpha=0.05$  is a variance scaling factor, and  $d$  is the distance measured from the coast in km. The choice for  $\alpha$  represents a standard deviation for the *prior* model error in temperature that is  $\sim 22\%$  of that of the error in the initial condition *prior*.

Figure 3 shows that the weak constraint R4D-Var cost function  $J(\delta\mathbf{z})$  asymptotes to a lower value compared to the strong constraint case, and is closer to the theoretical minimum value  $J_{min}$ . Additional experiments have revealed that the minimum value of  $J(\delta\mathbf{z})$  reached during weak constraint R4D-Var decreases with increasing  $\alpha$ . However, for  $\alpha>0.05$  the resulting most likely circulation estimates possess features that are very unphysical. Nonetheless, these preliminary attempts at accounting for uncertainties due to model during 4D-Var are encouraging.

Figure 5 also shows the impact of a weak constraint on  $J(\delta\mathbf{z})$  in relation to the sequence of experiments described in section 4.4.2 where the control vector is successively augmented. Experiment G is the same as that shown in Fig. 3, but only the minimum value reached by  $J(\delta\mathbf{z})$  is shown. Figure 5 indicates that augmenting the increment control vector  $\delta\mathbf{z}$  with  $\boldsymbol{\eta}(t)$  leads to a further significant reduction in  $J(\delta\mathbf{z})$ , and  $J(\delta\mathbf{z})$  is closer still to  $J_{min}$ .

#### 4.5 4D-Var Diagnostics

Having successfully computed the most likely circulation estimate using 4D-Var, there are a number of diagnostic calculations that can be performed using the 4D-Var output that of considerable interest. Two examples are presented here that provide information about the accuracy of the resulting circulation in the form of the *posterior* error covariance, and that provide valuable information about the

impact of each individual observation on different physical aspects of the circulation.

The calculations presented below are greatly facilitated by the Lanczos formulation of the conjugate gradient algorithm that is used to minimize  $J(\delta\mathbf{z})$  in the ROMS 4D-Var algorithms (Fisher and Courtier, 1995). Specifically, the primal form of the gain matrix (18) can be expressed as  $\mathbf{K} = \mathbf{V}_p \mathbf{T}_p^{-1} \mathbf{V}_p^T \mathbf{G}^T \mathbf{R}^{-1}$ , where each column of the matrix  $\mathbf{V}_p$  is a primal Lanczos vector (Golub and van Loan, 1989), and  $\mathbf{T}_p$  is a tridiagonal matrix. Each inner-loop yields one additional member of the Lanczos sequence which form an orthonormal basis, and after  $m$  inner-loops the matrix  $\mathbf{V}_p$  has dimension  $N_m \times m$ . Similarly, the dual form of the gain matrix (19) can be expressed as  $\mathbf{K} = \mathbf{D} \mathbf{G}^T \mathbf{V}_d \mathbf{T}_d^{-1} \mathbf{V}_d^T$  where  $\mathbf{V}_d$  is a matrix of dual Lanczos vectors of dimension  $N_{obs} \times m$ .

#### 4.5.1 Posterior Error

If  $\mathbf{z}_t$  represents the control vector describing the true state of the ocean, then the error covariance matrix of the *posterior* estimate  $\mathbf{z}_a = \mathbf{z}_b + \delta\mathbf{z}_a$  is given by  $\mathbf{E}_a = E((\mathbf{z}_a - \mathbf{z}_t)^T (\mathbf{z}_a - \mathbf{z}_t))$  where  $E$  is the expectation operator. It is easy to show that in the case where the observation errors and *prior* errors are uncorrelated, the *posterior* error covariance matrix can be expressed as:

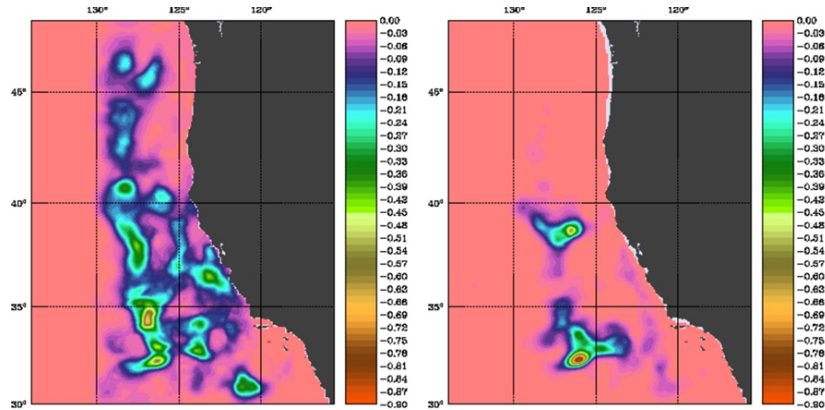
$$\mathbf{E}_a = (\mathbf{I} - \mathbf{K}\mathbf{G})\mathbf{D}(\mathbf{I} - \mathbf{K}\mathbf{G})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T \quad (24)$$

where  $\mathbf{K}$  is the gain matrix of section 3.2.

Exercise 6: Using the definitions  $\mathbf{E}_a = E((\mathbf{z}_a - \mathbf{z}_t)^T (\mathbf{z}_a - \mathbf{z}_t))$  and  $\mathbf{D} = E((\mathbf{z}_b - \mathbf{z}_t)^T (\mathbf{z}_b - \mathbf{z}_t))$  and the dual form of  $\mathbf{K}$  in (19), derive equation (24) for the *posterior* error covariance matrix assuming uncorrelated *prior* errors and observation errors.

Figure 6 shows an example of the reduction in the uncertainty in SST and subsurface temperature as a result of data assimilation using R4D-Var. In this case, ROMS-CCS with 10km horizontal resolution and 42 vertical levels was used. The diagonal elements of  $\mathbf{D}$  and  $\mathbf{E}_a$  represent the *prior* and *posterior* error variance respectively of each control variable. Therefore, the diagonal of the difference matrix  $\Delta = (\mathbf{E}_a - \mathbf{D})$  represents the change in error variance of the *prior* due to 4D-Var. Negative values of the diagonal of  $\Delta$  represent *posterior* grid point variables that are more certain than the *prior*, while zero values indicate that the *posterior* and *prior* are equally uncertain. Figure 6 shows the diagonal elements of  $\Delta$  corresponding to SST and 75m temperature, and indicates that the *posterior* upper ocean temperature is more certain than that of the *prior* over large areas of the

ROMS-CCS domain, particular in the region of high eddy kinetic energy along the California central coast identified in satellite observations by Kelly et al. (1998).



**Fig. 6:** Posterior variance-prior variance for SST (left) and 75m temperature (right) for a typical R4D-Var calculation in ROMS-CCS with 10km resolution.

#### 4.5.2 Observation Impacts

It is of considerable interest to know which observations or which types of observations exert the greatest influence on particular physical aspects of the most likely circulation estimate derived from 4D-Var. For example, Fig. 7 shows a vertical section of the time average of the alongshore velocity over the upper 500m of the water column along 37N for the period 1 Jan 1999-31 Dec 2004 from two ROMS-CCS calculations: one in which no data were assimilated, and one in which data were assimilated sequentially every 14 days using strong constraint I4D-Var. In the latter case, the increment control vector is comprised only of  $\delta\mathbf{x}(t_0)$  as described in Broquet et al. (2009a). Clearly the assimilation of observations has a significant impact on the structure of the CCS at this latitude.

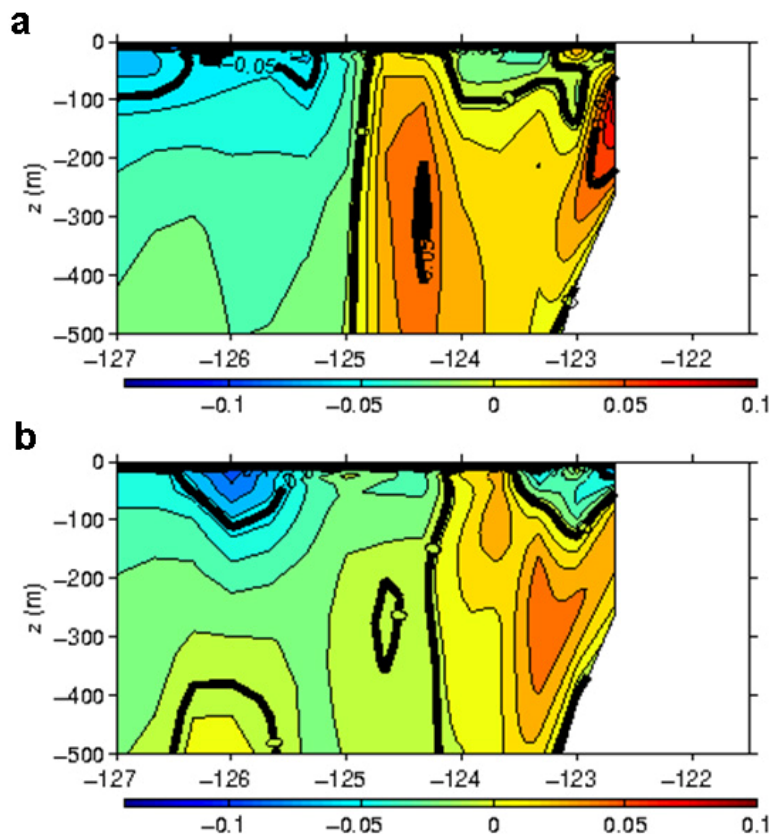


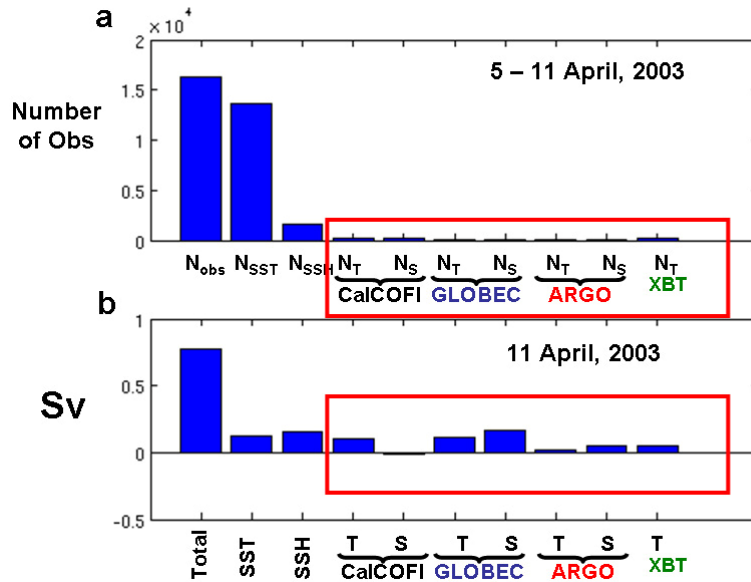
Fig. 7: Time average vertical sections of the alongshore component of velocity from ROMS-CCS during the period 1 Jan 2000-31 Dec 2004 for the case (a) where there is no data assimilation, and (b) where data are assimilated sequentially every 14 days using strong constraint I4D-Var.

In order to quantify the impact of each individual observation on the circulation, consider the transport across the section shown in Fig. 7 for the *posterior* circulation of a typical 4D-Var cycle, namely  $I_a(t) = \mathbf{h}^T \mathbf{x}_a(t)$ , where  $\mathbf{h}$  is a vector with zero elements everywhere except at the location of the model velocity grid-points along the section in Fig. 7. The non-zero elements take the form of the transformation coefficients required to rotate the total velocity in the alongshore direction, and the appropriate area elements for each grid cell in the vertical. However, recall that the circulation estimate  $\mathbf{x}_a(t) = \mathbf{x}_b(t) + \delta \mathbf{x}_a(t)$ , in which case  $I_a(t) = \mathbf{h}^T \mathbf{x}_b(t) + \mathbf{h}^T \delta \mathbf{x}_a(t) = I_b(t) + \mathbf{h}^T \delta \mathbf{x}_a(t)$ . Therefore, the difference in transport  $\Delta I(t) = I_a(t) - I_b(t)$  between the *posterior* and the *prior* circulation estimates is given by  $\Delta I(t) = \mathbf{h}^T \delta \mathbf{x}_a(t)$ . However, recall that to first-order  $\delta \mathbf{x}_a(t) = \mathbf{M}(t_0, t) \mathbf{K} \mathbf{d}$  where

$\mathbf{M}(t_0, t)$  is the tangent linear model. The difference in transport between the *posterior* and the *prior* can then be expressed as:

$$\Delta I(t) = \mathbf{h}^T \mathbf{M}(t_0, t) \mathbf{K} \mathbf{d} = \mathbf{d}^T \mathbf{K}^T \mathbf{M}^T(t, t_0) \mathbf{h} \quad (25)$$

where  $\mathbf{M}^T(t, t_0)$  is the adjoint model. For each data assimilation cycle, (25) can therefore be used to compute the contribution of each observation, represented by the individual elements of  $\mathbf{d}$ , to the change  $\Delta I(t)$  in the *prior* transport associated with 4D-Var.



**Fig. 8:** Panel (a) shows a histogram of the total number of observations ( $N_{obs}$ ) and the number of observations from different instruments during the period 5-11 April, 2003.  $N_{SST}$  and  $N_{SSH}$  represent the number of satellite measurements of SST and SSH respectively, while  $N_T$  and  $N_S$  represent the number of subsurface temperature and salinity observations from various sources (e.g., the CalCOFI and GLOBEC/LTOP repeat sample arrays along the California and Oregon coasts respectively, Argo profiling floats, and miscellaneous XBTs). Panel (b) shows the difference  $\Delta I$  in transport between the *posterior* and *prior* circulation estimates (labeled “Total”) on 11 April, 2003, and the contribution to  $\Delta I$  of all the observations from the different observation systems. The red box highlights all of the subsurface observations and their contribution to  $\Delta I$ .

Figure 8 shows an example calculation from a 7 day strong constraint R4D-Var calculation for the period 5-11 April, 2003. Figure 8b indicates that change in the *prior* estimate transport along 37N on 11 April as a result of assimilating  $\sim 1.5 \times 10^4$  observations is  $\sim 0.75 \text{ Sv}$ . A comparison of Fig. 8a and Fig. 8b, however, reveals

that even though satellite observations account for about 94% of the total number of observations, about 63% of the change in *prior* transport is associated with the subsurface observations which account for only 6% of  $N_{obs}$ . In this case then, the subsurface observations exert a considerable influence on the change in the circulation along 37N despite the large number of satellite observations.

The impact of each individual observation can be computed for any aspect of the circulation that can be expressed as a differentiable function of the model state-vector  $\mathbf{x}$ . Other examples for ROMS-CCS are described by Moore et al. (2009b).

## 5 Summary

In this chapter, we have attempted to summarize the important ideas and concepts underpinning the use of adjoint methods for variational data assimilation, providing where possible illustrative pedagogic examples and examples from a practical 4D-Var system from ROMS. In the interests of brevity, there are however, many important practical and technical details that we have shamelessly glossed over here or ignored, but interested readers can pursue them further by consulting the references provided herein. This presentation is also very heavily biased towards the ROMS 4D-Var system, but it should be noted that comprehensive 4D-Var systems have developed, or are currently under development, for various other ocean models. Notable examples that are well documented in the scientific literature include efforts in France and Europe (Weaver et al., 2003) and in the U.S. (Stammer et al., 2002).

## References

- Bennett, A.F.: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press (2002).
- Broquet, G., Edwards, C.A., Moore, A.M., Powell, B.S., Veneziani, M., Doyle, J.D.: Application of 4D-variational data assimilation to the California Current System. *Dyn. Atmos. Oceans*, **48**, 69-91 (2009a).
- Broquet, G., Moore, A.M., Arango, H.G., Edwards, C.A., Powell, B.S.: Ocean state and surface forcing correction using the ROMS-IS4DVAR data assimilation system. *Mercator Ocean Quarterly Newsletter*, **34**, 5-13 (2009b).
- Broquet, G., Moore, A.M., Arango, H.G., Edwards, C.A.: Corrections to ocean surface forcing in the California Current System using 4D-variational data assimilation. *Ocean Modelling*, Submitted (2009c).
- Chapman, D.C.: Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model. *J. Phys. Oceanogr.*, **15**, 1060-1075 (1985).

- Courtier, P.: Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.*, **123**, 2449-2461 (1997).
- Courtier, P., Thépaut, J.-N., Hollingsworth, A.: A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367-1388 (1994).
- Da Silva, A., Pfaendtner, J., Guo, J., Sienkiewicz, M., Cohn, S.: Assessing the effects of data selection with DAO's physical-space statistical analysis system. Proceedings of the second international WMO symposium on assimilation of observations in meteorology and oceanography. Tokyo 13-17 March, 1995. WMO.TD **651**, 273-278 (1995).
- Derber, J., Bouttier, F.: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-221 (1999).
- Doyle, J.D., Jiang, Q., Chao, Y., Farrara, J.: High-resolution atmospheric modeling over the Monterey Bay during AOSN II. *Deep Sea Res. II*, **56**, 87-99 (2009).
- Egbert, G.D., Bennett, A.F., Foreman, M.C.G.: TOPEX/POSEIDON tides estimated using a global inverse method. *J. Geophys. Res.*, **99**, 24,821-24,852 (1994).
- El Akkraoui, A., Gauthier, P.: Convergence properties of the primal and dual forms of variational data assimilation. *Q. J. R. Meteorol. Soc.*, In press (2009).
- Fairall, C.W., Bradley, E.F., Godfrey, J.S., Wick, G.A., Ebson, J.B., Young, G.S.: Cool-skin and warm layer effects on the sea surface temperature. *J. Geophys. Res.*, **101**, 1295-1308 (1996a).
- Fairall, C.W., Bradley, E.F., Rogers, D.P., Ebson, J.B., Young, G.S.: Bulk parameterization of air-sea fluxes for tropical ocean global atmosphere Coupled-Ocean Atmosphere Response Experiment. *J. Geophys. Res.*, **101**, 3747-3764 (1996b).
- Fisher, M., Courtier, P.: 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Tech. Memo. **220** (1995).
- Flather, R.A.: A tidal model of the northwest European continental shelf. *Memoires de la Societe Royale des Sciences de Liege* **6 (10)**, 141-164 (1976).
- Gill, A.E.: *Atmosphere-Ocean Dynamics*. Academic Press (1982).
- Ide, K., Courtier, P., Ghil, M., Lorenc, A.C.: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteorol. Soc. Japan*. **75**, 181-189 (1997).
- Golub, G.H., Van Loan, C.F.: *Matrix Computations*. The Johns Hopkins University Press, Baltimore (1989).
- Haidvogel, D.B., Arango, H.G., Hedstrom, K., Beckmann, A., Malanotte-Rizzoli, P., Shchepetkin, A.F.: Model evaluation experiments in the North Atlantic basin: Simulations in nonlinear terrain-following coordinates. *Dyn. Atmos. Oceans*, **32**, 239-281 (2000).
- Hickey, B.M.: Coastal oceanography of western North America from the tip of Baja, California to Vancouver Island. *The Sea*, **11**, 345-393, (1998).
- Ingleby, B., Huddleston, M.: Quality control of ocean temperature and salinity profiles—historical and real-time data. *J. Mar. Systems*, **65**, 158-175 (2007).

- Kelly, K.A., Beardsley, R.C., Limeburner, R., Brink, K.H., Paduan, J.D., Chereskin, T.K.: Variability of the near-surface eddy kinetic energy in California Current based on altimetric, drifter, and moored current data. *J. Geophys. Res.*, **103**, 13,067-13,083 (1998).
- Lanczos, C.: *Linear Differential Operators*. Van Nostrand, New York (1961).
- Liu, W.T., Katsaros, K.B., Businger, J.A.: Bulk parameterization of the air-sea exchange of heat and water vapor including the molecular constraints at the interface. *J. Atmos. Sci.*, **36**, 1722-1735 (1979).
- Lorenc, A.C.: Modelling of error covariances by 4D-Var data assimilation. *Q. J. R. Meteorol. Soc.*, **129**, 3167-3182 (2003).
- Moore, A.M., Arango, H.G., Broquet, G., Powell, B.S., Zavala-Garay, J., A.T. Weaver, A.T.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part I: Formulation. *Ocean Modelling*, In preparation (2009a).
- Moore, A.M., Arango, H.G., Broquet, G.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part II: Formulation. *Ocean Modelling*, In preparation (2009b).
- Pedlosky, J.: *Geophysical Fluid Dynamics*. Springer-Verlag, New York (1987).
- Sasaki, Y.: Some basic formulations in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875-883 (1970).
- Shchepetkin, A.F., McWilliams, J.C.: The regional oceanic modeling system (ROMS): A split explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modeling*, **9**, 347-404 (2005).
- Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., Adcroft, A., Hill, C.N., Marshall, J.: 2002: The global ocean circulation during 1992-1997 estimated from ocean observations and a general circulation model. *J. Geophys. Res.* (2002). doi:10.1029/2001JC000888.
- Tshimanga, J., Gratton, S., Weaver, A.T., Sartenaer, A.: Limited-memory preconditioners with application to incremental variational data assimilation. *Q. J. R. Meteorol. Soc.*, **134**, 751-769 (2008).
- Veneziani, M., Edwards, C.A., Doyle, J.D., Foley, D.: A central California coastal ocean modeling study: 1. Forward model and the influence of realistic versus climatological forcing. *J. Geophys. Res.* (2009). doi:10.1029/2008JC004774.
- Weaver, A.T., Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc.*, **127**, 1815-1846 (2001).
- Weaver, A.T., Vialard, J., Anderson, D.L.T.: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics and consistency checks. *Mon. Wea. Rev.*, **131**, 1360-1378 (2003).
- Weaver, A.T., Deltel, C., Machu, E., Ricci, S., Daget, N.: A multivariate balance operator for variational ocean data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3605-3625 (2005).
- Wunsch, C., Heimbach, P.: Practical global ocean state estimation. *Physica D*, **230**, 197-208 (2007).