Available online at www.sciencedirect.com
ScienceDirect
Physica D 230 (2007) 197–208
PHYSICA D
www.elsevier.com/locate/physd

Practical global oceanic state estimation

Carl Wunsch*, Patrick Heimbach

Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Available online 22 November 2006

Abstract

The problem of oceanographic state estimation, by means of an ocean general circulation model (GCM) and a multitude of observations, is described and contrasted with the meteorological process of data assimilation. In practice, all such methods reduce, on the computer, to forms of least-squares. The global oceanographic problem is at the present time focussed primarily on smoothing, rather than forecasting, and the data types are unlike meteorological ones. As formulated in the consortium Estimating the Circulation and Climate of the Ocean (ECCO), an automatic differentiation tool is used to calculate the so-called adjoint code of the GCM, and the method of Lagrange multipliers used to render the problem one of unconstrained least-squares minimization. Major problems today lie less with the numerical algorithms (least-squares problems can be solved by many means) than with the issues of data and model error. Results of ongoing calculations covering the period of the World Ocean Circulation Experiment, and including among other data, satellite altimetry from TOPEX/POSEIDON, Jason-1, ERS-1/2, ENVISAT, and GFO, a global array of profiling floats from the Argo program, and satellite gravity data from the GRACE mission, suggest that the solutions are now useful for scientific purposes. Both methodology and applications are developing in a number of different directions.
© 2006 Elsevier B.V. All rights reserved.

Keywords: State estimate; Data assimilation; Ocean circulation; Adjoint method; Method of Lagrange multipliers

1. Introduction

In physical oceanography, the problem of combining observations with numerical models differs in a number of significant ways from its practice in the atmospheric sciences. It is these differences that lead us to use the terminology "state estimation" to distinguish the oceanographers' problems and methods from those employed under the label "data assimilation" in numerical weather prediction. "Data assimilation" is an apt term, and were it not for its prior use in the meteorological forecast community, it would be the terminology of choice. But meteorologists, faced with the goal of daily weather forecasting, have developed sophisticated techniques directed at their own particular problems, along with an opaque terminology not easily penetrable by outsiders. Because much of oceanography has goals distinct from forecasting, the direct application of meteorological methods is often not appropriate.

This paper is primarily based upon experience with calculating the oceanic state within an academic consortium "Estimating the Circulation and Climate of the Oceans" (ECCO). The consortium was established in 1998 to exploit the new global data sets that became available during the World Ocean Circulation Experiment (WOCE) as well as the much more skillful ocean general circulation models that had also been developed.

In general terms, anyone faced with trying to calculate a system state by combining a physical model with a set of observations has an estimation problem. Estimation theory is a large subject with a wide range of methods available. Among the techniques used are maximum likelihood, minimum variance, and Bayesian ones. In a somewhat different context, these are all methods for solution of "inverse problems", although there has been a tendency by some authors to define inverse problems as referring only to steady-state situations. (For a more general point of view, see [6].) In the geophysical sciences, inverse methods were introduced by Backus and Gilbert [3] (see especially [37]) for situations in which the parameters sought represented a static earth.

Much of the theory underlying meteorological data assimilation, as in the original inverse theory of Backus and

* Corresponding address: Department of Earth, Atmospheric & Planetary Sciences, Massachusetts Institute of Technology, Room 54-152477 Massachusetts Ave., Cambridge, MA 02139, USA. Fax: +1 617 2534464.
E-mail addresses: cwunsch@mit.edu (C. Wunsch), heimbach@mit.edu (P. Heimbach).

0167-2789/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2006.09.040

Gilbert [3], formulates the problems in terms of continuous time and space. The combination of continuous, infinite domain, fields with finite, discrete observations leads to an interesting set of mathematical issues in functional analysis (see [37,48, 5,6]). Oceanographers, coming later to this problem, enter it at a time when computational resources are widespread and far more powerful than they were in the early days of the meteorological problem. Thus, the point of view taken within the ECCO Consortium is, both for practical and pedagogical reasons, that the problems are discrete and finite, *ab initio*, and that in practice, they can all be regarded as versions of classical non-linear least-squares.

The fundamental justifying assumption is that an adequate discretization of the fluid dynamical and thermodynamical equations governing the fluid flow is available. "Adequate" means that with sufficient grid refinement or increase in the number of basis functions, one could not distinguish the numerical solution from that of the underlying continuous formulation that generated it. Whether the discretization exists when required is a separate problem—the assumption here is that it *could*, and if important enough, *will* eventually exist. Alternatively, one is simply asserting that a finite dimensional discrete scheme, no matter how large that dimension might prove to be in practice, can always be found that is sufficient for the purpose at hand. Is there a physically meaningful system that is not in principle susceptible to rendering in finite dimensions? We are assuming, without proof, that the answer is "no". (The expression "physically meaningful" is intended to exclude analogues of the pathological functions used by mathematicians as counter-examples to physical intuition. A simple representative is the curve of unbounded variation, $\sin(1/t)$, $|t| \to 0$, which probably does not occur anywhere in nature.)

There is another aspect to this problem. All atmospheric and oceanic circulation models contain semi-empirically determined sub-gridscale processes. An example is the invocation of biharmonic friction operators, ∇^4, to mimic small-scale turbulent dissipation. Many of the problems involved in solving inverse problems in continuous formulations are generated by these high derivatives e.g., the singular behavior of an analytical Green function—see [5]. What is unclear is whether the purely discrete form of the biharmonic operator is any less convincing as a representation of what is basically an unknown process? If the discrete rendering of this and other terms is regarded as adequate, all issues of convergence, continuity, differentiability, and existence, essentially disappear. From a pedagogical point of view, many of the most important estimation methods (Kalman filters, adjoints, etc.) can be reduced to the practice of conventional least-squares [54,55]. That is, we have deliberately removed the most intriguing mathematical issues from the problem—in the interest both of practical utility, and as a way of making the methodologies accessible to scientists interested in phenomenology rather than mathematics. (Analogous reductions are widespread in practical problems. A particularly well-known one is Levinson's [26] conversion of Norbert Wiener's mathematically challenging filter theory—based on spectral factorization by Wiener–Hopf methods—to the least-squares form widely exploited in exploration geophysics.)

The data sets employed by oceanographers differ considerably from those employed by meteorologists. As discussed below, these differences do not create any fundamental obstacles, but require changes in the details of representation.

2. The goal

Meteorological data assimilation has been driven by the compelling need to forecast the weather. The dominant problem in oceanography at the present time is quite different—it is to understand how the system works. Observing the ocean is technically far more difficult than observing the atmosphere—fundamentally because of its opacity to electromagnetic radiation. Thus, until comparatively recently, oceanographers of necessity had to treat the ocean as an unchanging, essentially static, system. With the data sets that have accumulated since approximately the beginning of WOCE, about 1992, marked by the launch of the high precision altimetric mission TOPEX/POSEIDON, it becomes possible to contemplate calculating the time-evolving ocean state—to begin the process of answering the questions of whether, how, and by how much and why, the ocean circulation changes? In the terminology of control theory, the meteorological problem has been primarily one of filtering and prediction (extrapolation); the oceanographic one is primarily one of "smoothing" (interpolation). Both of these emphases can ultimately be expected to change and overlap, but they describe the present situation. Note in particular that ocean models are major components of climate models. Such models are run out into the indefinite future, and without any general understanding of ocean model errors, it is almost impossible to interpret the result. State estimation as practiced here becomes in part a serious attempt to understand which elements of ocean models are likely skillful, and capable of integration with bounded errors for finite times.

The ECCO consortium thus has undertaken the smoothing problem of using a stored data set over a finite time interval, for the purpose of making best estimates of the oceanic state during that time in such a way that the final estimate would be dynamically and kinematically self-consistent, accompanied by an understanding of the structure of residual data/model misfits. Solution of the smoothing problem makes the result differ fundamentally from what is sometimes called, in the meteorological literature "reanalysis" (e.g., [22]). Such calculations maintain a fixed numerical model configuration over a finite time interval, but produce an analysis at time, t_a, using data only for $t \leq t_a$, in an approximate "filtering" calculation. The omission of data in the formal future, $t > t_a$, means that the system state estimate can undergo jumps, implying implicit non-physical sources, and rendering very difficult the physical interpretation of the time-evolving state. Methods have been employed (e.g., [7]) to smooth out the discontinuities over finite times, but still leaving artificial imbalances in the solution. Reanalyses done sequentially in the forward direction alone can leave important global imbalances: for example [47], the Kalnay et al. [22] estimates do not conserve freshwater or enthalpy. Such artifacts were deemed

unsuitable for ECCO purposes, and are one reason that conventional meteorological practice is not obviously suitable for the oceanographic problem.

Some fundamental assumptions were made, ones that sidestep the underlying mathematical problems:

(1) The ocean state could be usefully represented by the particular GCM of Marshall et al. [32], if run at adequate resolution. This model is an approximate numerical representation of the full non-linear Navier–Stokes and thermodynamical equations on a sphere. There is no proof of the numerical convergence of any known GCM in the limit of arbitrarily fine resolution. (Griffies [18] provides an overview of ocean modelling.)

(2) Oceanic observation errors, control elements, and model state variables are sufficiently close to Gaussian to justify use of a minimum variance estimator (essentially that the probability densities are unimodal). Although there are fragmentary tests of the probability densities of oceanic and atmospheric variables, there is no general description or theory.

(3) It is possible to use oceanographic data of arbitrary type. As will see in the next sections, the data types used in practice vary greatly.

Neither assumptions (1) nor (2) are rigorously correct, but they provide both a starting point and a focus for constant vigilance.

Adopting the notation of [54,55], the algebraic expression of the problem is as follows. Let $\mathbf{x}(t)$ be the model state vector at discrete time t. Although an unnecessary assumption, in practice, all of our data at time t are linearly related to the state vector as,

$$\mathbf{E}(t)\mathbf{x}(t) + \mathbf{n}(t) = \mathbf{y}(t),$$

where $\mathbf{n}(t)$ is a stochastic noise vector such that,

$$\langle \mathbf{n}(t) \rangle = 0, \qquad \langle \mathbf{n}(t)\mathbf{n}(t')^{\mathrm{T}} \rangle = \delta_{tt'}\mathbf{R}(t).$$

The GCM is represented as,

$$\mathbf{x}(t) = \mathbf{L}(\mathbf{x}(t-\Delta t), \mathbf{B}(t-\Delta t)\mathbf{q}(t-\Delta t), \boldsymbol{\Gamma}(t-\Delta t) \times \mathbf{u}(t-\Delta t), t-\Delta t), \tag{1}$$

where $\mathbf{L}$ is a non-linear operator (a FORTRAN code), $\mathbf{q}(t)$ are known sources/sinks/boundary conditions. $\mathbf{u}(t)$ is the unknown control vector with moments,

$$\langle \mathbf{u}(t) \rangle = 0, \qquad \langle \mathbf{u}(t)\mathbf{u}(t')^{\mathrm{T}} \rangle = \delta_{tt'}\mathbf{Q}(t).$$

If $\langle \mathbf{u}(t) \rangle \neq 0$, the known form would be included in $\mathbf{q}(t)$. The matrices $\mathbf{B}$, $\boldsymbol{\Gamma}$ are available to distribute the forcing and control vectors over the state vector in a flexible way (see [54]).

The state vector, $\mathbf{x}(t)$, contains just enough information about the flow to march the model one time step into the future, given externally prescribed boundary conditions and sources or sinks. In the present case, it includes three components of velocity, surface elevation, temperature, and salinity at each grid point. The control vector contains any model variables subject to adjustment, here initial conditions on temperature and salinity, time-dependent corrections to the externally prescribed meteorological forcing, as well as parameters such as mixing coefficients. An important point, commonly only implicitly acknowledged, is that the $\mathbf{u}(t)$ also include the model error. Little is known of model error and an important unsolved mathematical problem concerns its representation as a function of time and space. In practice, structures that are incorrectly rendered by a model are treated as data error with little understanding of the consequences. Initial conditions are represented as $\tilde{\mathbf{x}}(0) = \mathbf{x}_0$, such that $\langle (\mathbf{x}_0 - \mathbf{x}(0))(\mathbf{x}_0 - \mathbf{x}(0))^{\mathrm{T}} \rangle = \mathbf{P}(0)$. (Tildes represent estimates. $\mathbf{x}(t)$ is the hypothetical true value.) That is, there is an uncertain initial condition. (From here, we assume $\Delta t = 1$.)

Adopting the point of view that we are solving a least-squares problem, we seek to minimize the objective function,

$$J = \sum_{t=1}^{t_f} [\mathbf{y}(t) - \mathbf{E}(t)\mathbf{x}(t)]^{\mathrm{T}} \mathbf{R}(t)^{-1} [\mathbf{y}(t) - \mathbf{E}(t)\mathbf{x}(t)] + [\mathbf{x}_0 - \mathbf{x}(0)]^{\mathrm{T}} \mathbf{P}(0)^{-1} [\mathbf{x}_0 - \mathbf{x}(0)] + \sum_{t=0}^{t_f - 1} \mathbf{u}(t)^{\mathrm{T}} \mathbf{Q}(t)^{-1} \mathbf{u}(t), \tag{2}$$

subject to Eq. (1). The second term in J specifically calls out the initial condition uncertainty, but it could equally well be included in the first or last terms. The state and control vectors in Eq. (2) and below could be written with tildes, but these are omitted here as there is little possibility of confusion. Under the Gaussian assumption, implying a linear dynamical model, a solution would coincide with the minimum variance estimate and would simultaneously be a maximum likelihood solution—assuming the error covariance matrices $\mathbf{R}, \mathbf{Q}, \mathbf{P}$ are properly specified. As written, however, because the $\mathbf{R}, \mathbf{Q}, \mathbf{P}$ are poorly known, this problem is best regarded as one of constrained least-squares (curve fitting). One approach to solving it would be to use a sequential method, such as an extended Kalman filter, followed by e.g., an extended form of the RTS smoother [40]. Within ECCO, such an approach in approximate form has been undertaken by Fukumori et al. [13].

Here we proceed to reduce the problem from constrained to unconstrained least-squares by introducing vector Lagrange multipliers, $\boldsymbol{\mu}(t)$ augmenting the objective function,

$$J' = J - 2\sum_{t=1}^{t_f} \boldsymbol{\mu}(t)^{\mathrm{T}} [\mathbf{x}(t) - \mathbf{L}(\mathbf{x}(t-1), \mathbf{B}(t-\Delta t) \times \mathbf{q}(t-1), \boldsymbol{\Gamma}(t-1)\mathbf{u}(t-1), t-1)]. \tag{3}$$

Minimizing J can be regarded as a regularizing requirement. Explicit normal equations can be written, which with nonlinear $\mathbf{L}$, will be a set of nonlinear simultaneous equations, generally well-conditioned:

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{u}(t)} = \mathbf{Q}(t)^{-1}\mathbf{u}(t) + \left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\ \mathbf{Bq}(t),\ \boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{u}(t)} \right)^{\mathrm{T}} \times \boldsymbol{\mu}(t+1) = \mathbf{0}, \quad 0 \leq t \leq t_f - 1 \tag{4}$$

$$\frac{1}{2}\frac{\partial J'}{\partial \boldsymbol{\mu}(t)} = \mathbf{x}(t) - \mathbf{L}[\mathbf{x}(t-1),\ \mathbf{Bq}(t-1),\ \boldsymbol{\Gamma}\mathbf{u}(t-1)] = \mathbf{0}, \quad 1 \leq t \leq t_f \tag{5}$$

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(0)} = \mathbf{P}(0)^{-1}[\mathbf{x}(0)-\mathbf{x}_0] + \left(\frac{\partial \mathbf{L}(\mathbf{x}(0),\,\mathbf{Bq}(0),\,\boldsymbol{\Gamma}\mathbf{u}(0))}{\partial \mathbf{x}(0)}\right)^{\mathrm{T}} \boldsymbol{\mu}(1) = \mathbf{0}, \quad (6)$$

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(t)} = \mathbf{E}(t)^{\mathrm{T}}\mathbf{R}(t)^{-1}[\mathbf{E}(t)\mathbf{x}(t)-\mathbf{y}(t)] - \boldsymbol{\mu}(t) + \left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\,\mathbf{Bq}(t),\,\boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{x}(t)}\right)^{\mathrm{T}} \times \boldsymbol{\mu}(t+1) = \mathbf{0}, \quad 1 \le t \le t_f - 1 \quad (7)$$

$$\frac{1}{2}\frac{\partial J'}{\partial \mathbf{x}(t_f)} = \mathbf{E}(t_f)^{\mathrm{T}}\mathbf{R}(t_f)^{-1}[\mathbf{E}(t_f)\mathbf{x}(t_f)-\mathbf{y}(t_f)] - \boldsymbol{\mu}(t_f) = \mathbf{0}. \quad (8)$$

The use of Lagrange multipliers has become known as the "adjoint method", and, opaquely, as "4DVAR", in meteorology. The special case in control theory of a terminal constraint problem (see e.g., [53]) is widely known as the "Pontryagin Principle". In the interests of making clear to the widest possible community what is going on, we will call it the "method of Lagrange multipliers", MLM, which has meaning to anyone who has encountered classical mechanics, the calculus of variations, or constrained least-squares.[1]

Conceptually, the oceanic problem is perhaps best regarded as a form of control problem—one seeks to find those controls (typically the wind, freshwater, and heat exchanges with the atmosphere) that drive the ocean (actually the model) through a trajectory within error bars of *all* of the observations. The problem is somewhat more complex than a standard discrete control problem in that the estimated initial conditions, and often internal parameters of the model, are subject to change as part of the computation. In analogy with the classical problem of robotic control, a robot arm is required to pass within observational error of a series of known positions, velocities, accelerations, etc. on its way to a terminal state. But the numerical description of the robotic arm (the model) is believed to contain inaccuracies that are to be removed as data are employed. Similarly, the starting conditions for the robotic arm are believed erroneous in part as well and are to be re-estimated. The goal is to deduce as best possible what the state of the robotic arm was over the duration of observation and what the actual controls were, while simultaneously improving the estimates of the model and of the initial conditions. A useful general survey of many aspects of control theory, written for the physics community, can be found in [4].

If $\mathbf{L}$ were a linear operator, one might contemplate simply solving the normal equations by Gaussian elimination or other algorithm. If non-linear, a variety of search/descent algorithms are available. The main issue for the oceanographic problem is one of dimension. In the existing ECCO configuration (July 2005), with the model written at one-degree spatial resolution and 23 vertical layers, the state vector, $\mathbf{x}(t)$ is a 5.3 million dimension vector of three components of velocity, temperature and salinity and seasurface elevation (u, v, w, T, S, η). The model is time-stepped every hour for 13 years (1992–2004) producing a total state vector of 6.1×10^{11} elements. The control vector, $\mathbf{u}(t)$, has 310 elements over 13 years, and the data dimension is 2.1×10^9 million elements, also over 13 years. Data numbers are dominated by the 6-hourly meteorological forcing—but they are adjusted only at two-day intervals to provide partial smoothing over atmospheric synoptic scales, and to reduce the size of the control vector. (To restart the optimization requires a numerical state vector at time t, of about 21 million values, including tendency terms.) Thus the number of equations (albeit they are sparse) is much too large to be solved directly even with a linear model. In a sequential system such as the Kalman filter plus smoother, one is faced with the computation and storage of covariance matrices square of the dimension of $[\mathbf{x}(t), \mathbf{u}(t)]$. It is for this reason that Fukumori et al. [13] and others have resorted to a series of approximations to the filter/smoother equations so as to reduce the effective state and control vector dimensions.

The MLM method, if a solution can be found, does not require the use of covariance matrices for the state vector—rather it is a whole domain method in which no averaging of interim solutions takes place—in the present context, that is its great attraction. Consider how an iterative solution might work: Set, initially, $\mathbf{u}(t) = \mathbf{0}$, as the first guess at the controls. Integrate Eq. (5) forward in time to produce a first estimate of $\mathbf{x}(t)$. Then integrate Eq. (7) backwards in time to produce a first estimate of $\mu(t)$. To do these calculations, the partial derivatives,

$$\left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\,\mathbf{Bq}(t),\,\boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{u}(t)}\right)^{\mathrm{T}}, \quad \left(\frac{\partial \mathbf{L}(\mathbf{x}(t),\mathbf{Bq}(t),\,\boldsymbol{\Gamma}\mathbf{u}(t))}{\partial \mathbf{x}(t)}\right)^{\mathrm{T}}, \quad (9)$$

must be known. But these define the partial derivatives of J, J' with respect to the problem parameters. Thus we are in a position to invoke a quasi-Newton method to reduce the value of J. One then modifies the problem parameters, integrates forward and backward again (an iteration) and re-evaluates the derivatives along the way. This iterative procedure was apparently first laid out in an oceanographic context by Thacker and Long [49]

[1] A useful way to gain some insight into these equations is to make the connection to classical mechanics and differential geometry. Adjoint operators can be defined independent of any optimization problem (e.g., [34,24]) and are intimately connected to the theory of Green functions; for matrix operators; they are just the transpose. Consider too (e.g., [51]) that the partial differential operators appearing in Eqs. (4)–(8), before transposition, define a tangent linear operator (or model) δL acting on a tangent bundle, TM, and mapping control vector perturbations, i.e. elements of the tangent space $T_{x(0)}(M)$ at $(\mathbf{x}(0))$ into the tangent space of the model state $T_{x(t)}(M)$ at time t via the model Jacobian. Their transposes, as they appear in Eqs. (4)–(8), act on the corresponding cotangent bundle, T^*M. They define a mapping from the cotangent space $T^*_{x(t)}(M)$ into $T^*_{x(0)}(M)$. (This description applies to the special case where the control vector consists of the initial conditions only. The concept is readily extended to time varying boundary controls with a bit more algebra.) The spaces T^*_x are dual spaces to T_x, and their elements are called "co-vectors" (note that the gradient is, strictly speaking, a co-vector rather than a vector). It can readily be shown that the transpose $\delta^* L$ is the adjoint operator of the tangent linear operator, and thus Eq. (7) defines an adjoint model operating on co-vectors along $\mathbf{x}(t)$, and which are often called "adjoint variables". Following along this route leads into differential geometric optimal control theory, a subject we will not pursue here.

and has been used many times since then. The favored descent algorithm is that of Gilbert and Lemaréchal [16].

3. Making it work

There are a number of major issues. In some ways, the most interesting is the problem of finding the partial derivatives (9). The transposed partial derivatives define the "adjoint model". In the discrete formulation we are using, $\mathbf{L}$ is defined by a computer code (hundreds of thousands of lines of FORTRAN) and there is no analytical expression available. Our approach has been to define the adjoint of the FORTRAN-coded model. What has made this a practical possibility is the existence of automatic (or algorithmic) differentiation (AD) tools (e.g., [39, 17]), beginning with software such as ADIFOR. But the most important advance was that of Giering and Kaminski [15] who, taking account of the special structures of oceanic GCMs, provided a semi-automatic AD tool for generating code for the tangent linear model (so-called forward mode—the non-transposed partial derivatives in Eq. (9)) and the adjoint model (the so-called reverse mode) corresponding to the transposed partial derivatives. In its present configuration, the particular AD tool used is called TAF—Transformation of Algorithms in FORTRAN. The tools can generate higher-order derivative code, e.g., for the Hessians. The latter are of use in assessing posterior uncertainties; Hessian eigenvalues determine the principal curvature structures of the objective function, with large curvatures indicating small posterior uncertainties in the control estimates (e.g. [50]). AD is a powerful tool even in conventional sequential methods such as the Kalman filter, because for non-linear systems one requires the state transition matrix, that is, the tangent linear model, and which is obtainable in this fashion.

The availability of the adjoint model greatly reduces the computational load of the minimization iterations (see [31, 19]). The GCM that is used in ECCO was constructed from the beginning with its use with an AD tool in mind, as part of the ATOC (1999) [1] acoustic tomography project. Certain FORTRAN77 and 90 structures remain problematic in TAF, but are avoidable. AD tools permit the adjoint code to be updated and maintained in a practical way in the presence of continuing vigorous development and improvement of the parent non-linear model. (One can, as has been done, manually code the analytical form of the adjoint partial differential equations. See [49,30,42]. The adjoint code then differs from the adjoint of the forward code. The difference may be immaterial for some applications, but the maintenance of a hand-coded adjoint for a model undergoing development can be a major task.)

The fundamental ECCO principle, borne out in practice, is that adequate numerical methods exist for reducing the misfit between a model and data: numerical methods can be found that solve the least-squares problem. That there are challenging numerical and storage problems goes without saying, but the principle that least-squares "works" is not in doubt. (Much of the literature in this subject, done using synthetic "data", really reduces to the uninteresting conclusion that least-squares solutions can be obtained.)

We postulate that numerical means are available for finding the stationary point(s) of Eq. (3), be it by direct solution of the normal equations, by deterministic or stochastic search, or by pure, skillful guess. Such a solution will have a physical validity directly dependent upon the weight matrices $\mathbf{Q}(t), \mathbf{R}(t), \mathbf{P}(0)$ appearing in J. The latter are conventionally supposed to be the error covariances of the data (because in the linear limit, the least-squares solution would be the maximum likelihood or minimum variance solutions as well—ones which have a ready physical interpretation). Part of $\mathbf{Q}(t)$ represents the covariances of the control vector, describing e.g., the extent to which the prior windfield is subject to modification during the calculation. ($\mathbf{Q}(t)$ also formally includes model errors.) The solution, however it is obtained, can be no better than the choice of these matrices.

The number of elements summed in the combined,

$$[\mathbf{y}(t) - \mathbf{E}(t)\mathbf{x}(t)]^{\mathrm{T}}\mathbf{R}(t)^{-1}[\mathbf{y}(t) - \mathbf{E}(t)\mathbf{x}(t)],$$
$$(\mathbf{x}_0 - \mathbf{x}(0))^{\mathrm{T}}\mathbf{P}(0)^{-1}(\mathbf{x}_0 - \mathbf{x}(0)), \quad \mathbf{u}(t)^{\mathrm{T}}\mathbf{Q}(t)^{-1}\mathbf{u}(t),$$

is approximately 2.1×10^9 as of this writing. Every one of these terms requires a weight! Although of minor interest to mathematicians, the choice of the weights dominates the current effort. At this stage, we have a problem primarily of oceanography and meteorology rather than one of mathematics. Fig. 6 provides a summary of the different data types now being fit.

With some minor exceptions, the only estimates of $\mathbf{R}(t), \mathbf{Q}(t), \mathbf{P}(0)$ refer to the diagonal values alone—that is, almost nothing is known quantitatively (or more precisely, usefully) of error covariances in the observed fields or in the controls.[2] Obtaining useful, truly independent, estimates of the non-diagonal elements of these matrices is an important, ongoing, but barely underway, effort (this problem is one primarily of oceanography and meteorology, not of state estimation, per se). An interim practice is to use, where regarded as reasonable, smoothing requirements on some fields, which are equivalent to specifying off-diagonal covariances. Currently, however, even the diagonal elements represent a great deal of subjective judgement and interpretation of the vast oceanographic and meteorologic literature referring to all of the observations used. A number of papers [11,38,47] begin to describe the estimated data errors.

As noted above, because the relationship of the weight matrices to the true error covariances remains uncertain, the solutions described below should be regarded as least-squares solutions, rather than as maximum likelihood or minimum variance ones—much of the judgement as to acceptability lies with aesthetics, rather than with rigorous statistics.

4. Sample results

A few papers have appeared (e.g., [27,21,9,28]) exploiting the adjoint model interpretation as the sensitivity of the

[2] Because the control vector includes adjustments to the meteorological fields, and which can be regarded as in part "observed", the distinction between state vector and controls and what is observed, is largely arbitrary.

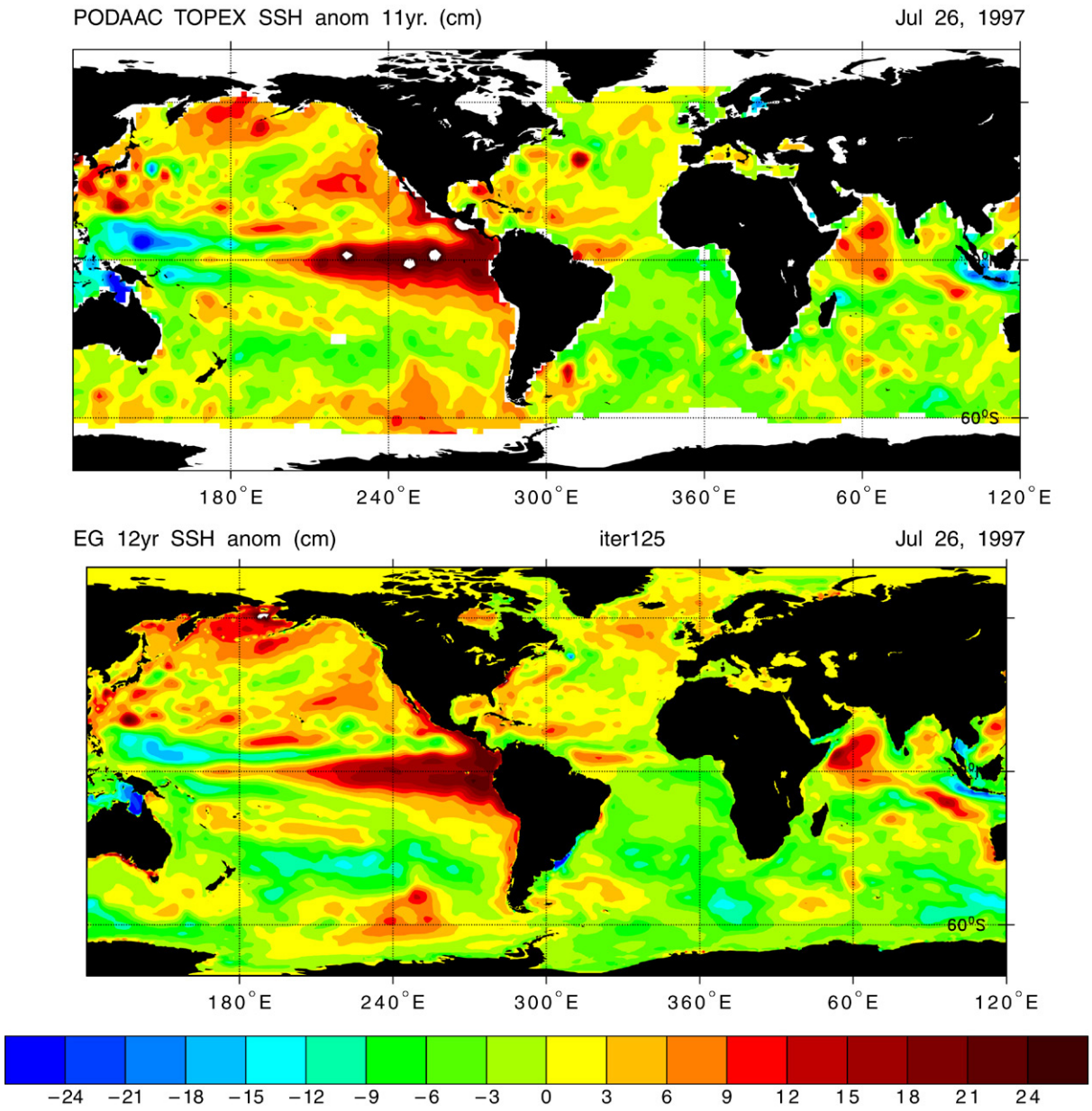


Fig. 1. Upper panel shows the seasurface elevation as inferred from 10 days of altimetric data during July 1997. Lower panel shows the estimate made from the model constrained both to the upper panel (with data used daily) and to many other data types. Main visual signal is the large El Niño event that occurred during this time.

objective function J to perturbations in data and parameters. We will focus instead on some examples of the results, as yet unpublished, of the most recent optimization effort in ECCO (for an overview, see [20]). The main points to be made here are that the numerical system "works" with real data, and that there is much more to be done.

The results shown here are the product of hundreds of iterations of the combined forward and adjoint models over the duration of the data sets. A truly quantitative statement concerning the optimization pathway is impossible as the current configuration builds upon some years of calculation with earlier, shorter data durations, different model configurations (e.g., changed resolution and sub-components such as mixed-layer models), new data types (e.g., the Argo floats), changed weights in J, modifications to the line-search algorithm, changes in machine architecture, etc. Earlier results have been described at length by Stammer et al. [43–45], Köhl et al. [23] among others. The calculations strain the largest available high performance computers.

Fig. 1 is intended to show the extent to which an early solution (which is still evolving) over the interval 1992–2004 is capable of reproducing one of the major constraining data sets—in this case the seasurface height anomaly (relative to a long-term mean) during July 1997. (An animation is available at http://ocean.mit.edu/king/mpegs/12iter125.mpeg.) The TOPEX/POSEIDON and related altimetric missions provide remarkably accurate measurements of the absolute global ocean surface elevation every 10 days (absolute

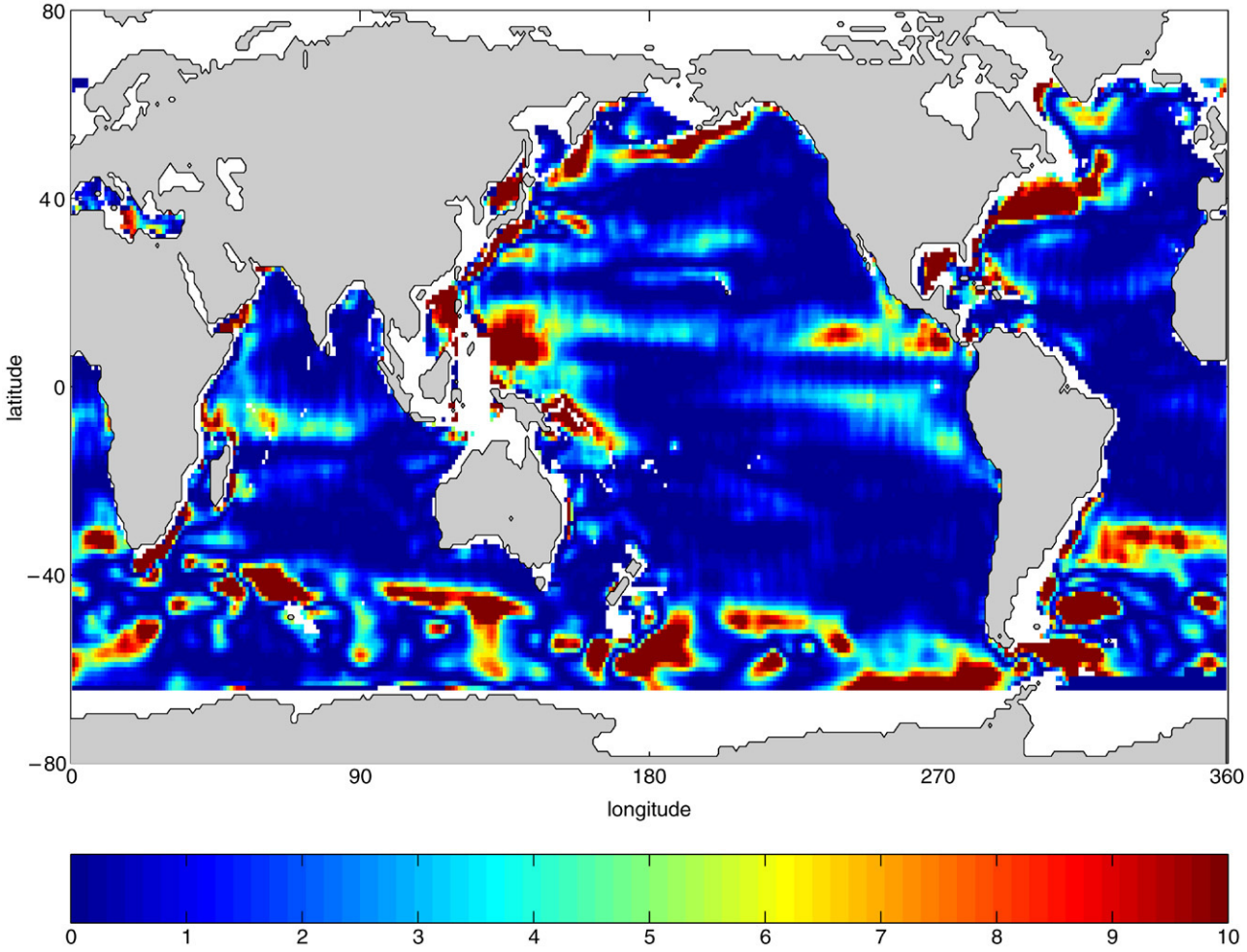


Fig. 2. Misfit of the time mean over 13 years of the constrained model to the time-average altimetric seasurface height in an early stage of iteration. An ideal solution would be everywhere of order 1. Here there are strong regional misfits whose explanation is in part misweighted data, model error, and failure thus far to have achieved complete convergence.

accuracies in the range of 2–3 cm; see [58,12]). The misfit terms in J corresponding to the time-mean seasurface height (or mean dynamic topography, MDT) over the 13-year model run are displayed in Fig. 2. Ideally, all these values would be order 1 with outliers consistent with a χ^2 distribution. There is a reasonable fit over the bulk of the open ocean. Serious misfits persist in coastal regions generally, in much of the Southern Ocean, and in some regions of strong variability. The scientific issue is now whether these misfits represent (A) a mis-weighting of the data; (B) model (representation) error; or (C) simply that the iterative minimization has not had time to reduce these terms, or all of these things. As noted above, the issue of data weighting takes one into the intricacies of the underlying observations and their errors. In this particular case, calculation of the data errors involves the determination of Earth gravity (providing the reference gravitational surface called the "geoid"; see [58]), the myriad error terms in altimetric determination of seasurface topography (see e.g., [8]), and the expected temporal variability as a function of position and duration in the ocean [38]. Model error quantification involves determining the extent to which a 1° horizontal resolution, 23 vertical-layer model would be expected to reproduce the mean seasurface topography—were it perfectly known. (MDT estimates are expected to improve as the GRACE gravity-measuring satellite mission accumulates a longer record.)

A representative misfit to an entirely different data type is shown in Fig. 3. These data are obtained by so-called Argo floats (see [41]) drifting freely in the ocean, and which periodically profile temperature and salinity on a vertical ascent to the seasurface where the data are transmitted ashore by satellites. Similar questions apply here as to the altimetry—are the data being properly weighted, and is the model being forced toward structures which are physically impossible for it?

Numerous other applications related to ECCO have appeared. Examples include [2,14], which are regional estimates, the latter using open boundary conditions and high resolution. Stammer [46] and Ferreira et al. [10] discuss estimation of internal eddy-stress and mixing parameters as part of the optimization. Many others will be forthcoming.

A reviewer has inquired about system "validation". In general (see [36]) no system can ever be validated or "verified". Within the system itself, there is a very powerful test of the model against the data, as there is no guarantee that

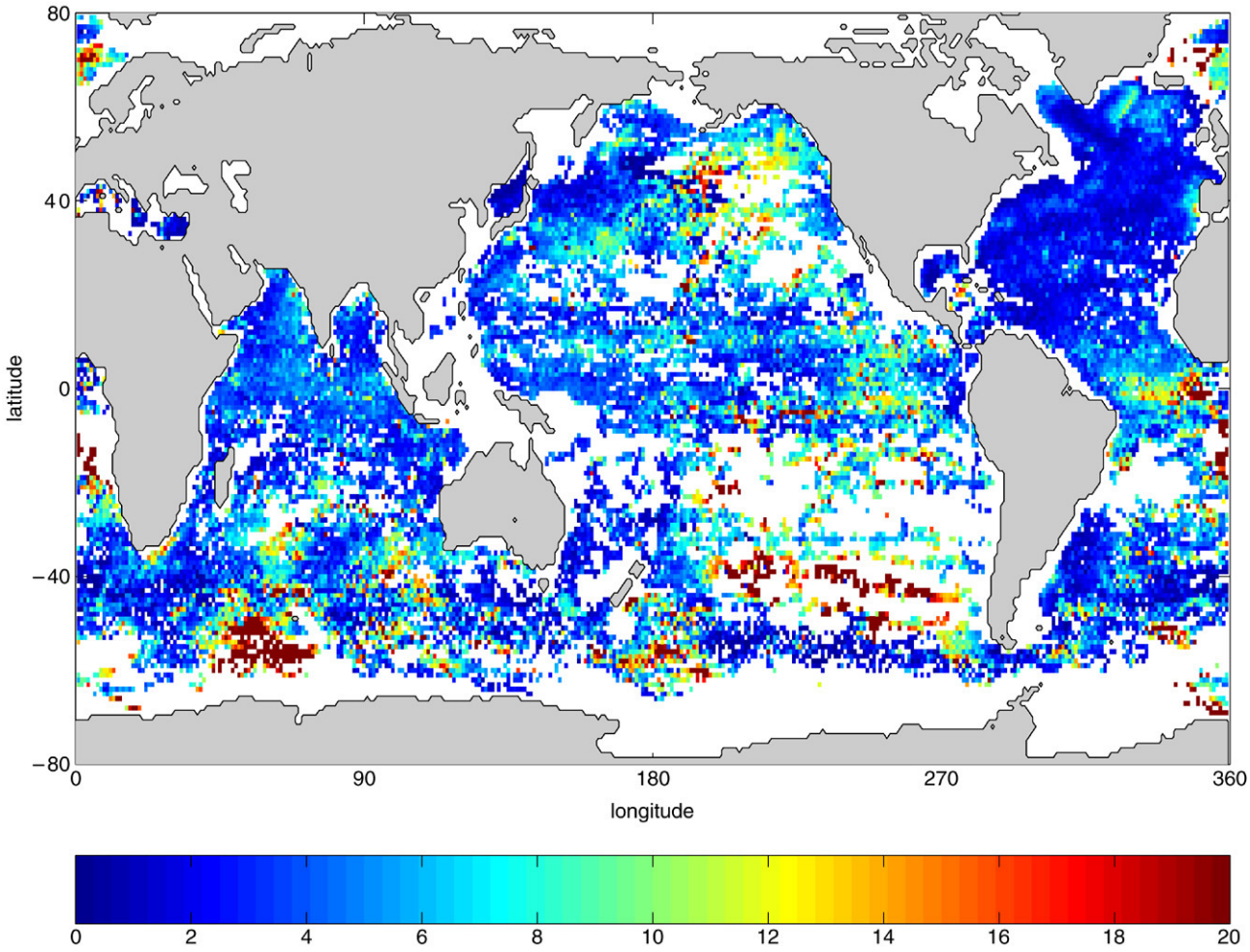


Fig. 3. Misfit over 13 years to the so-called ARGO temperature data, vertically averaged. (These data do not become available until about halfway through the calculation and are dominated in numbers by the last two years, 2003, and 2004.) Again, a nominal value of one would denote an acceptable solution.

the model can reproduce the observations within error (that is, there may be no solution). The solutions discussed here appear to be close to consistency with the data, but the optimization calculations are still incomplete. The present system has been tested against a certain amount of "withheld" data (what is sometimes meant by "validation", and which is related to cross-validation and split-sampling; see Fig. 6). As expected, there are similarities and differences, but a determination of whether the differences imply system failure has no readily available answer.

5. Concluding discussion

A practical system exists for solving a very large, global-scale, oceanographic least-squares problem using the method of Lagrange multipliers. The methodology relies heavily on automatic differentiation (AD) tools specially constructed for use with oceanic general circulation models (GCMs). When written in the finite, discrete form required by computers, most of the interesting and challenging mathematical issues are side-stepped. Many of the issues in practice are oceanographic rather than mathematical, and focus on the error covariances of a multiplicity of data types.

Where is this effort going? ECCO and related projects are developing in several directions. Understanding of the data errors and acquisition of new data types is ongoing, and likely will never be completed; it is an asymptotic process. AD tools (http://www-unix.mcs.anl.gov/OpenAD) are under development in a number of places, directed at greater efficiency, an open-source, and new computer languages; see [35,52]. Model errors of certain types can be eliminated by augmenting the control vector by uncertain model elements, e.g., as Losch and Heimbach [28] did for bottom topography. The quest for greater model resolution (see [33]) continues; this too, is likely an asymptotic process, as is the search for rules for parameterizing the indefinitely present sub-gridscale motions. At bottom, of course, the purpose of all of this effort is to understand the ocean and the practical implications of that understanding. As one example of that use, Fig. 4 displays the estimated mass, heat and salt fluxes in the North Atlantic across 26°N as a function of time from the least-squares fit [56]. This result is part of the continuing debate over whether the North Atlantic circulation is undergoing some kind of important shift. Another example of the application of the results is in picking

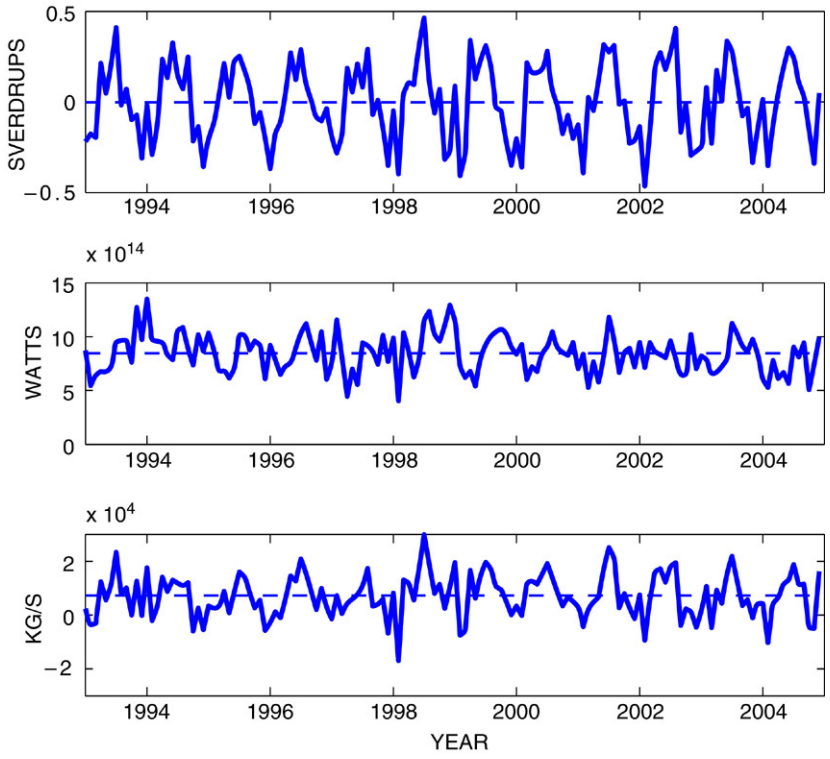


Fig. 4. Zonal and vertical integrals across the constrained model at 26°N in the North Atlantic as a function of time, showing (upper panel) the total meridional mass flux (in Sverdrups—10^6 m^3/s), the meridional heat flux in W (middle panel), and the salt flux (in kg/s) in the lowest panel. This solution is not the final one.

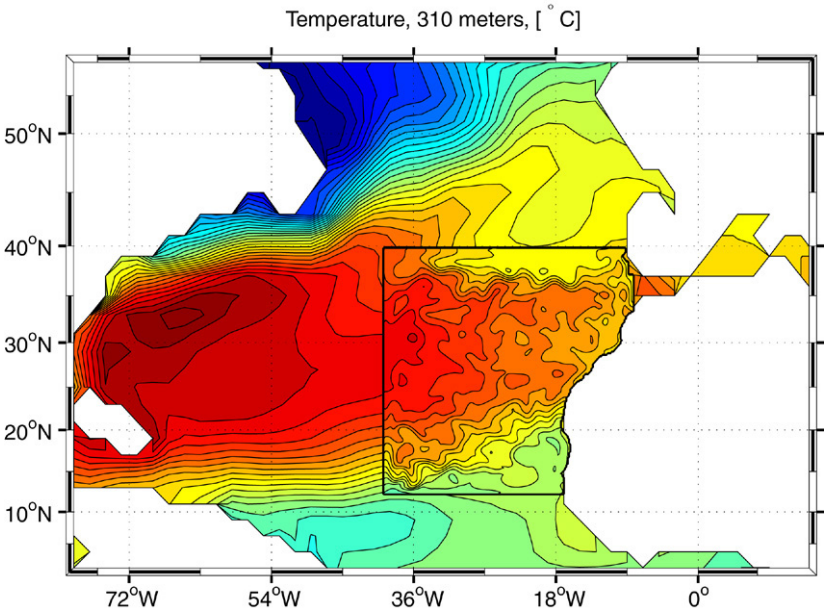


Fig. 5. Temperature field [14] showing nesting of an open boundary high resolution model into the global one. Open boundary properties, including flows, become part of the control vector. The temperature ranges from 3.5 °C (blue) to 18 °C (red).

apart the heat and freshwater contributors to global sea level rise [57].

Extension of the primarily fluid physical model used here is being undertaken to include a more adequate ice dynamics and thermodynamics model. Dutkiewicz et al. [9] have employed the same AD tools, and the background physical state for parallel study of oceanic biogeochemical cycles. Further extension of the models and methods to the study of past climate states is now getting underway.

Inadequate model resolution is likely to be a major problem for the foreseeable future. Some fluid phenomena in some regions (e.g., boundary jets, fronts, etc.) cannot be

DATA TYPE	Source	Spatial Extent	Variable(s)	Duration	Number of values
Altimetry: TOPEX/POSEIDON	PODAAC	Global, equatorward of 65 degrees	height anomaly, temporal average	1993-2004	4500/day
Altimetry: Jason	PODAAC	Global equatorward of 65 degrees	height anomaly, temporal average	2002-2004	4500/day
Altimetry: Geosat-followon	US Navy, NOAA	Global, equatorward of 65 degrees	height anomaly	2001-2004	4300/day (gridded)
Altimetry: ERS-1/2, ENVISAT	AVISO	Global, equatorward of 81.5 degrees	height anomaly	1995-2004	3800/day
Hydrographic climatology	Gouretski and Koltermann (2004)	global, 300m to seafloor	temperature, salinity	1950-2002 inhomogeneous average	16 million
Hydrographic climatology	World Ocean Atlas (2001), Conkright et al. (2002)	global to 300m	temperature, salinity	multidecadal average seasonal cycle	included above
CTD synoptic section data	Various, including WOCE Hydro. Prog.	global, all seasons, to 3000m	temperature, salinity	NA	17 thousand
XBTs	D. Behringer (NCEP)	global, but little So. Ocean	temperature	1992-2004	1.4 million
ARGO Float profiles	IFREMER	global, above 2500m	temperature, salinity	1997-2004	2.1 million
Sea Surface Temperature	Reynolds and Smith (1999)	global	temperature	1992-2004	5.3 million
Sea Surface Salinity	Etudes Climatiques de l'Oce'an Pacifique (ECOP)	tropical Pacific	salinity	1992-1999	24,238
TMI	NASA/NOAA	global	temperature	1998-2003	1.5 million
Geoid (GRACE mission)	GRACE Project. (GGM02C_DOT_cel6. Tapley et al. (2003)	global	mean dynamic topography	NA	1 degree resol.
Bottom Topography	Smith&Sandwell(1997)+ETOP05	Smith/Sandwell to 72.006, ETOP05 to 79.5	water depth	NA	1 degree resol.
FORCING:					
Windstress-scatterometer	PODAAC	global	stress	1992-99, 7/1999-2004	9.4 million
Windstress	NCEP/NCAR reanalysis Kalnay et al. (1996)	global	stress	1992-2004	192x94 Gaussian grid (approx. 1.875 deg) 6-hourly
Heat Flux	NCEP/NCAR reanalysis	global	sensible + latent heat	1992-2004	"
Freshwater Flux	NCEP/NCAR reanalysis	global	evap-precip	1992-2004	"
Short/long Wave Radiation (experimental)	NCEP/NCAR reanalysis	global		1992-2004	"
WITHHELD (as of April 2006):					
tide gauges		global, sparse	sealevel		
TOGA-TAO array		equatorial oceans	velocity/temperature		
Tomographic integrals		N. Pacific	heat content		
Florida Current transport		Florida Straits	mass flux		
Float and Drifter Velocities		Global	velocity		

Fig. 6. Table 1, showing the main data sources, coverage, and some information about the originator. These were the main data types and coverage used as of July 2005. At the bottom, the table lists some of the withheld data (but which are intended ultimately for inclusion).

parameterized in any easy or known way, and the estimation of their properties demands a model resolution that is impossible on a global scale. To this end, developments are underway permitting embedding of high resolution models into lower resolution global estimates such as the one described here. The major issue is that the sub-model boundaries must be taken as "open", so as to permit exchange of fluid and information between the two models. Although there are some mathematical questions about the posedness of such approaches (e.g., [6]), no particular problems have been encountered in practice [2,14] when the numerics are once again viewed as least-squares problems in which regularization is straightforward. The Gebbie et al. [14] embedded model is depicted in Fig. 5. It has been argued (e.g., [25]) that the system may become so chaotic that tangent linear gradients fail to be useful. Real systems demonstrating a lack of useful numerical differentiability would direct one toward Monte Carlo methods (e.g. [29]), but no such situation has so far been met.

Acknowledgements

This work was supported by the National Ocean Partnership Program, the National Science Foundation ITR program, and the National Aeronautics and Space Administration through the Jason and GRACE missions. The bulk of the ECCO computations reported here were carried out at the Geophysical

Fluid Dynamics Laboratory, NOAA, and at the National Center for Atmospheric Research. Thanks are due our many ECCO partners. Comments from two anonymous reviewers were helpful.

References

[1] ATOC Consortium, Ocean climate change: Comparison of acoustic tomography, satellite altimetry, and modeling, Science 281 (1998) 1327–1332.

[2] N. Ayoub, Estimation of boundary values in a North Atlantic circulation model using an adjoint method, Ocean Model. 12 (3–4) (2006) 319–347, doi:10.1016/j.ocemod.2005.06.003.

[3] G.E. Backus, J.F. Gilbert, Numerical applications of a formalism for geophysical inverse theory, Geophys. J. Roy. Astron. Soc. 13 (1967) 247–276.

[4] J. Bechhoefer, Feedback for physicists: A tutorial essay on control, Rev. Modern Phys. 77 (2005) 783–836.

[5] A.F. Bennett, Inverse Methods in Physical Oceanography, Cambridge University Press, Cambridge, 1992, 346 pp.

[6] A.F. Bennett, Inverse Modeling of the Ocean and Atmosphere, Cambridge University Press, Cambridge, 2002, 234 pp.

[7] S.C. Bloom, L.L. Takacs, A.M. Da Silva, D. Ledvina, Data assimilation using incremental analysis updates, Mon. Weather Rev. 124 (1996) 1256–1271.

[8] D.B. Chelton, J.C. Ries, B.J. Haines, L.-L. Fu, P.S. Callahan, Satellite altimetry, in: L.-L. Fu, A. Cazenave (Eds.), Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications, Academic, 2000, pp. 1–132.

[9] S. Dutkiewicz, M. Follows, P. Heimbach, J. Marshall, Controls on ocean productivity and air–sea carbon flux: An adjoint model sensitivity study, Geophys. Res. Lett. 33 (2005) L02603, doi:10.1029/2005GL024987.

[10] D. Ferreira, J. Marshall, P. Heimbach, Estimating eddy stresses by fitting dynamics to observations using a residual-mean ocean circulation model and its adjoint, J. Phys. Oceanogr. 35 (10) (2005) 1891–1910, doi:10.1175/JPO2785.

[11] G. Forget, C. Wunsch, Estimated global hydrographic variability, J. Phys. Oceanogr. (2006) (in press).

[12] L.-L. Fu, A. Cazenave (Eds.), Satellite Altimetry and Earth Sciences. A Handbook of Techniques and Applications, Academic, San Diego, 2001, 463 pp.

[13] I. Fukumori, R. Raghunath, L.-L. Fu, Y. Chao, Assimilation of TOPEX/Poseidon altimeter data into a global ocean circulation model: How good are the results? J. Geophys. Res. 104 (1999) 25647–25665.

[14] G. Gebbie, P. Heimbach, C. Wunsch, Strategies for nested and eddy-resolving state estimation, J. Geophys. Res. (2006) (in press).

[15] R. Giering, T. Kaminski, Recipes for adjoint code construction, ACM Trans. Math. Softw. 24 (1998) 437–474.

[16] J.C. Gilbert, C. Lemaréchal, Some numerical experiments with variable-storage quasi-Newton algorithms, Math. Program. 45(B) (1989) 407–435.

[17] A. Griewank, Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation, SIAM, Philadelphia, 2000, 369 pp.

[18] S.M. Griffies, Some ocean model fundamentals, in: E.P. Chassignet, J. Verron (Eds.), Ocean Weather Forecasting: An Integrated View of Oceanography, Springer, Berlin, Germany, 2005, pp. 19–74.

[19] P. Heimbach, C. Hill, R. Giering, An efficient exact adjoint of the parallel MIT general circulation model, generated via automatic differentiation, Future Gener. Comput. Syst. 21 (8) (2005) 1356–1371, doi:10.1016/j.future.2004.11.010.

[20] P. Heimbach, R.M. Ponte, C. Evangelinos, G. Forget, M. Mazloff, D. Menemenlis, S. Vinogradov, C. Wunsch, Combining altimetric and all other data with a general circulation model, in: Proceedings of the 15 Years Progress in Radar Altimetry Symposium,Venice, 13–18 March 2006 (ESA Special Publication SP–614).

[21] C. Hill, V. Bugnion, M. Follows, J. Marshall, Evaluating carbon sequestration efficiency in an ocean circulation model by adjoint sensitivity, J. Geophys. Res. 109 (2004) C11005.

[22] E. Kalnay, et al., The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Metereol. Soc. 77 (1996) 437–471.

[23] A. Köhl, D. Stammer, B. Cornuelle, Interannual to decadal changes in the ECCO global synthesis, 2005 (submitted for publication).

[24] C. Lanczos, Linear Differential Operators, Van Nostrand, Princeton, 1961, 564 pp.

[25] D.J. Lea, M.R. Allen, T.W.N. Haine, Sensitivity analysis of the climate of a chaotic system, Tellus A 52 (2000) 523–532.

[26] N. Levinson, The Wiener RMS (root mean square) error criterion in filter design and prediction, J. Math. Phys. 25 (1947) 261–278.

[27] X. Li, C. Wunsch, An adjoint sensitivity study of chlorofluorocarbons in the North Atlantic, J. Geophys. Res. 109 (C1) (2004), doi:10.1029/2003JC002014.

[28] M. Losch, P. Heimbach, Adjoint sensitivity of an ocean general circulation model to bottom topography, J. Phys. Oceanogr. (2006) (in press).

[29] I.W. McKeague, G. Nicholls, K. Speer, R. Herbei, Statistical inversion of South Atlantic circulation in an abyssal neutral density layer, J. Mar. Res 63 (2005) 683–704.

[30] G.I. Marchuk, Adjoint Equations and Analysis of Complex Systems, Kluwer, Dordrecht, 1995, 466 pp.

[31] J. Marotzke, R. Giering, K.Q. Zhang, D. Stammer, C.N. Hill, T. Lee, Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity, J. Geophys. Res. 104 (1999) 29529–29547.

[32] J. Marshall, A. Adcroft, C. Hill, L. Perelman, C. Helsey, A finite-volume, incompressible Navier–Stokes model for studies of the ocean on parallel computers, J. Geophys. Res. 102 (1997) 5753–5766.

[33] D.C. Menemenlis, et al., NASA supercomputer improves prospects for ocean climate research, EOS, Trans. Amer. Geophys. Un. 86 (89) (2005) 95–96.

[34] P.M. Morse, H. Feshbach, Methods of Theoretical Physics, 2 volumes, McGraw-Hill, New York, 1953, 1978 pp.

[35] U. Naumann, J. Utke, P. Heimbach, C. Hill, D. Ozyurt, C. Wunsch, M. Fagan, N. Tallent, M. Strout, Adjoint code by source transformation with OpenAD/F, in: P. Wesseling, E. Oñate, J. Périaux (Eds.), European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2006, TU Delft, The Netherlands, 2006.

[36] N. Oreskes, K. Shrader-Frechette, K. Belitz, Verification, validation, and confirmation of numerical models in the earth sciences, Science 263 (1994) 641–646.

[37] R.L. Parker, Geophysical Inverse Theory, Princeton University Press, Princeton, 1994, 377 pp.

[38] R.M. Ponte, C. Wunsch, D. Stammer, Spatial mapping of time-variable errors in TOPEX/POSEIDON sea surface height measurements, J. Atmos. Ocean. Technol. (2006) (in press).

[39] L.B. Rall, Automatic Differentiation: Techniques and Applications, Springer-Verlag, Berlin, 1981, 165 pp.

[40] H.E. Rauch, F. Tung, C.T. Striebel, Maximum likelihood estimates of linear dynamic systems, AIAA J. 3 (1965) 1445–1450.

[41] D. Roemmich, W.B. Owens, The Argo project: Global ocean observations for understanding and prediction of climate variability, Oceanography 13 (2) (2000) 5–50.

[42] Z. Sirkes, E. Tziperman, Finite difference of adjoint or adjoint of finite difference, Mon. Weather Rev. 125 (1997) 3373–3378.

[43] D. Stammer, C. Wunsch, R. Giering, C. Eckert, P. Heimbach, C. Hill, J. Marotzke, J. Marshall, The global ocean state during 1992–1997, estimated from ocean observations and a general circulation model. Part I. Methodology and estimate state, J. Geophys. Res. (2002), doi:10.1029/2001JC000888.

[44] D. Stammer, C. Wunsch, I. Fukumori, J. Marshall, State estimation in modern oceanographic research, EOS 83 (27) (2002) 289, 294–295.

[45] D. Stammer, C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C. Hill, J. Marshall, Volume, heat and freshwater transports of the global ocean circulation 1992–1997, estimated from a general circulation model constrained by WOCE data, J. Geophys. Res. (2003), doi:10.1029/2001JC001115.

[46] D. Stammer, Adjusting internal model errors through ocean state estimation, J. Phys. Oceanogr. 35 (2005) 1143–1153.

[47] D. Stammer, A. Köhl, C. Wunsch, Estimating the ocean circulation and the geoid through a combined analysis of altimetric and GRACE geoid information, 2006 (submitted for publication).
[48] O. Talagrand, Assimilation of observations, an introduction, J. Meteorol. Soc. Japan 75 (1997) 191–209.
[49] W.C. Thacker, R.B. Long, Fitting dynamics to data, J. Geophys. Res. 93 (1988) 1227–1240.
[50] W.C. Thacker, The role of the Hessian matrix in fitting models to measurements, J. Geophys. Res. 94 (1989) 6177–6196.
[51] W. Thirring, Classical Mathematical Physics, Springer-Verlag, New York, 1997, 543 pp.
[52] J. Utke, U. Naumann, M. Fagan, N. Tallent, M. Strout, P. Heimbach, C. Hill, D. Ozyurt, C. Wunsch, OpenAD/F: A modular open source tool for automatic differentiation of Fortran codes, ACM Trans. Math. Softw. (2006) (submitted for publication).
[53] C. Wunsch, Transient tracers as a problem in control theory, J. Geophys. Res. 93 (1988) 8099–8110.
[54] C. Wunsch, The Ocean Circulation Inverse Problem, Cambridge University Press, Cambridge, 1996, 437 pp.
[55] C. Wunsch, Discrete Inverse and State Estimation Problems. With Geophysical Fluid Applications, Cambridge University Press, Cambridge, 2006, 384 pp.
[56] C. Wunsch, P. Heimbach, Decadal changes in the North Atlantic meridional overturning and heat flux, J. Phys. Ocenogr. 36 (2006) 2012–2014.
[57] C. Wunsch, R. Ponte, P. Heimbach, Decadal trends in global sealevel patterns and implications for global mean errors, 2006 (in preparation).
[58] C. Wunsch, D. Stammer, Satellite altimetry, the marine geoid and the oceanic general circulation, Annu. Rev. Earth Planet. Sci. 26 (1998) 219–254.