Dynamically and Kinematically Consistent Global Ocean Circulation and Ice State Estimates

DRAFT

Carl Wunsch and Patrick Heimbach
Department of Earth, Atmospheric and Planetary Science
Massachusetts Institute of Technology
Cambridge MA 02139 USA
email: cwunsch@mit.edu, heimbach@mit.edu

September 25, 2012

Contents

Abstract

The World Ocean Circulation Experiment (WOCE) drove the development of estimates of the decadal scale time evolving general circulation that are dynamically and kinematically consistent. A long time-scale, and a goal of estimation rather than prediction, preclude the use of meteorological methods called "data assimilation (DA)." Instead, "state estimation" methods are reviewed here and distinguished from DA. Results from the dynamically-consistent ECCO family of solutions based upon least-squares Lagrange multipliers (adjoints) are used to discuss the determination of the dominant elements of the circulation in the period since 1992—which marked the beginning of the satellite altimetric record.

1 Introduction

The goal of what we call "state estimates" of the oceans arose directly out of the plans for the World Ocean Circulation Experiment (WOCE). That program, out of necessity, employed in a pragmatic way observational tools of a very wide diversity of type—including classical hydrography, current meters, tracers, satellite altimeters, floats, and drifters. The designers of WOCE realized that to obtain a coherent picture of the global ocean circulation approaching a time-scale of a decade, they would require some form of synthesis method: one capable of combining very disparate observational types, but also having greatly differing space-time sampling, and geographical coverage.

Numerical weather forecasting, in the form of what had become known as "data assimilation" (DA), was a known analogue of what was required: a collection of tools for combining the best available global numerical model representation of the ocean with any and all data, suitably weighted to account for both model and data errors (e.g., Talagrand, 1997; Kalnay, 2003; Evensen, 2009). Several major, and sometimes ignored, obstacles existed in employing meteorological methods for the oceanic problem. These included the large infrastructure used to carry out DA within the national weather forecast centers—organizations for which no oceanographic equivalent existed or exists. DA had developed for the purposes of *forecasting* over time scales of hours to a few days, whereas the climate goals of WOCE were directed at time scales of years to decades with a goal of understanding and *not* forecasting. Another, more subtle, difficulty was the WOCE need for state estimates capable of being used for global-scale energy, heat, and water cycle budgets. Closed global budgets are of little concern to a weather forecaster—as their violation has no impact on short-range prediction skill, but they are crucial to the understanding of climate change. Construction of closed budgets is also rendered physically impossible by the forecasting goal: solutions "jump" towards the data at every analysis time, usually every six hours, introducing spurious sources and sinks of basic properties.

Because of these concerns, the widespread misunderstanding of what DA usually does, and what oceanographers actually require, the first part of this essay is devoted to a sketch of the basic principles of DA, and the contrast with methods required in practice for use for climate-relevant state estimates. More elaborate accounts can be found in Wunsch (2006), and Wunsch and Heimbach (2007) among others. Within the meteorological literature itself, numerous publications exist (e.g., Trenberth et al., 1995, 2001; Bengtsson et al., 2004; Bromwich and Fogt, 2004; Bromwich et al., 2004, 2007, 2011; Nicolas and Bromwich, 2011; Thorne, 2008) warning against the use of DA and the associated "reanalyses" for the study of climate change. These warnings have been widely disregarded.

A theme of this Chapter is that both DA and state estimation can be understood from elementary principles, ones not going beyond beginning calculus. Those concepts must be distinguished from the far more difficult numerical engineering problem of finding practical methods capable of coping with large volumes of data, large model state dimensions, and a variety of computer architectures. But one can understand and use an automobile without being expert in the manufacture of an internal combustion engine or of the chemistry of tire production.

At the time of the writing of the first WOCE volume, Siedler et al. (2001), two types of large-scale synthesis existed: (1) the time-mean global inverse results of Macdonald (1998) based upon the pre-WOCE hydrography and that of Ganachaud (2003b) using the WOCE hydrographic sections. (2) Preliminary results from the first ECCO (Estimating the Circulation and Climate of the Ocean) synthesis (Stammer et al., 2002) were based upon a few years of data and comparatively coarse resolution models. Talley et al. (2001) summarized these estimates, but little time had been available for their digestion.

In the intervening years, Lumpkin and Speer (2007) produced a revision of the Ganachaud results using somewhat different assumptions, but with similar results, and a handful of other static global estimates (e.g., Schlitzer, 2007) appeared. The ECCO project greatly extended its capabilities and duration for time-dependent estimates. A number of regional, assumed steady-state, box inversions also exist (e.g., Macdonald et al., 2009).

As part of his box inversions, Ganachaud (2003a) had shown that the dominant errors in trans-oceanic property transports of volume (mass), heat (enthalpy), salt, etc. arose from the temporal variability. Direct confirmation of that inference can be seen in the ECCO-based time-varying solutions, and from in situ measurements (Rayner et al., 2011). So-called synoptic sections spanning ocean basins and which had been the basis for most global circulation pictures at best produce "blurred" snapshots of transport properties. We are now well-past the time in which they can be labelled and interpreted as being the time-average. A major result of WOCE was to confirm the conviction that *the ocean must be observed and treated as a fundamentally* *time-varying system*, especially for any property involving the flow field. Gross scalar properties such as the temperature or nitrogen concentrations have long been known to be stable on the largest scales: that their distributions are nonetheless often dominated by intense temporal fluctuations, sometimes involving very high wavenumbers, represents a major change in understanding of the classical ocean properties. That understanding inevitably drives one towards state estimation methods.

2 Definition

Consider any model of a physical system satisfying known equations, written generically in discrete time as,

$$\mathbf{x}(t) = \mathbf{L}\left(\mathbf{x}(t-\Delta t), \mathbf{q}(t-\Delta t), \mathbf{u}(t-\Delta t)\right), \quad 1 \leq t \leq t_f = M\Delta t, \tag{1}$$

{model1}

where $\mathbf{x}(t)$ is the "state" at time t, discrete at intervals Δt, and includes those prognostic or dependent variables usually computed by a model, such as temperature or salinity in an advection-diffusion equation or a stream function in a flow problem. $\mathbf{q}(t)$ denotes known forcings, sources, sinks, boundary and initial conditions, and internal model parameters, and $\mathbf{u}(t)$ is any such elements that are regarded as only partly or wholly unknown, hence subject to adjustment and termed independent or control variables (or simply "controls"). $\mathbf{L}$ is an operator and can involve a large range of calculations, including derivatives, or integrals or any other mathematically defined function. In practice, it is usually a computer code working on arrays of numbers. (Notation is approximately that of Wunsch, 2006.) Time, $t = m\Delta t$, is assumed to be discrete, with $m = 0, \ldots, M$, as that is almost always true of models run on computers.[1] Note that the steady-state situation is a special case, in which one writes an additional relationship, $\mathbf{x}(t) = \mathbf{x}(t-\Delta t)$ and $\mathbf{q}, \mathbf{u}$ are then time-independent. For computational efficiency, steady models are normally rewritten so that time does not appear at all, but that step is not necessary. Thus the static box inverse methods and their relatives such as the beta-spiral are special cases of the ocean estimation problem.

Useful observations at time t are all functions of the state and, in almost all practical situations, are a linear combination of one or more state vector elements,

$$\mathbf{y}(t) = \mathbf{E}(t)\,\mathbf{x}(t) + \mathbf{n}(t), \quad 0 \leq t \leq t_f, \tag{2}$$

{data1}

[1] An interesting mathematical literature surrounds state estimation carried out in continuous time and space in formally infinite dimensional spaces. Most of it proves irrelevant for calculations on computers which are always finite dimensional. Digression into functional analysis can be needlessly distracting.

where $\mathbf{n}(t)$ is the inevitable noise in the observations and $t_f = M\Delta t$. $\mathbf{y}(t)$ is a vector of whatever observations of whatever, diverse, type are available at t. (Uncertain initial conditions are included here at $t = 0$, representing them as noisy observations.) Standard matrix-vector notation is being used. In a steady-state formulation, parameter t would be suppressed. (On rare occasions, data are a nonlinear combination of the state vector: an example would be a speed measurement in terms of two components of the velocity, or a frequency spectrum for some variable is known. Methods exist, not discussed here, for dealing with such observations. Observations relating to the control vector may exist, and one easy approach to using them is to redefine elements of $\mathbf{u}(t)$ as being part of the state vector.) The "state estimation problem"[2] is defined as determining $\tilde{\mathbf{x}}(t)$, $0 \le t \le t_f$, $\tilde{\mathbf{u}}(t)$, $0 \le t \le t_f - \Delta t$, exactly satisfying both Eqs. (1), and (2). Tildes here denote estimates to distinguish them from the true values.

Important Note: "exact" satisfaction of Eq. (1) must be understood as meaning the model *after adjustment* by $\tilde{\mathbf{u}}(t)$. Because $\mathbf{u}(t)$ can represent, if necessary, very complex, nonlinear, and large changes to the original model, which is usually defined with $\mathbf{u}(t) = \mathbf{0}$, the adjusted model can be very different from the initial version. *But the adjusted model is known, fully specified, and exactly satisfied,* and is what is used for discussion of the physics or chemistry, etc. It thus differs in a fundamental way from other types of estimate rendered discontinuous by "data injection," or forcing to data, during the final forward calculation.

Typically, one must also have some knowledge of the statistics of the controls, $\mathbf{u}(t)$, and observation noise, $\mathbf{n}(t)$, commonly as the first and second-order moments,

$$\langle \mathbf{u}(t) \rangle = \mathbf{0}, \quad \left\langle \mathbf{u}(t)\, \mathbf{u}(t')^T \right\rangle = \mathbf{Q}(t)\, \delta_{tt'} \quad 0 \le t \le t_f - \Delta t = (M-1)\,\Delta t, \tag{3a}$$ {stat1}

$$\langle \mathbf{n}(t) \rangle = \mathbf{0}, \quad \left\langle \mathbf{n}(t)\, \mathbf{n}(t')^T \right\rangle = \mathbf{R}(t)\, \delta_{tt'} \quad 0 \le t \le t_f = M\Delta t \tag{3b}$$ {stat2}

The brackets denote expected values and superscript T is the vector or matrix transpose.

In generic terms, the problem is one of *constrained estimation/optimization*, in which, usually, one seeks to minimize both the normalized quadratic model-data differences,

$$\left\langle \left(\mathbf{y}(t) - \mathbf{E}(t)\,\mathbf{x}(t)\right)^T \mathbf{R}^{-1}(t) \left(\mathbf{y}(t) - \mathbf{E}(t)\,\mathbf{x}(t)\right) \right\rangle \tag{4}$$

and the normalized independent variables ("controls"),

$$\left\langle \mathbf{u}(t)^T \mathbf{Q}^{-1}(t)\, \mathbf{u}(t) \right\rangle. \tag{5}$$

— subject to the exact satisfaction of the *adjusted model* in Eq. (1).

[2] A terminology borrowed from control theory (e.g., Gelb, 1974).

For data sets and controls that are Gaussian or nearly so, the problem as stated is equivalent to weighted least-squares minimization of the scalar,

$$J = \sum_{m=0}^{M} \left(\mathbf{y}\left(t\right) - \mathbf{E}\left(t\right)\tilde{\mathbf{x}}\left(t\right)\right)^{T} \mathbf{R}^{-1}\left(t\right)\left(\mathbf{y}\left(t\right) - \mathbf{E}\left(t\right)\tilde{\mathbf{x}}\left(t\right)\right) + \sum_{m=0}^{M-1} \tilde{\mathbf{u}}\left(t\right)^{T} \mathbf{Q}^{-1}\left(t\right)\tilde{\mathbf{u}}\left(t\right), \quad t = m\Delta t, \tag{6}$$

{data3}

subject to Eq. (1). It is *a PDE-constrained least-squares problem*, and nonlinear if the model or the observations are nonlinear. The uncertain initial conditions, contained implicitly in Eq. (6), are readily written out separately if desired.

In comparing the solutions to DA, note that *the latter problem is different.* It seeks to minimize,

$$\mathrm{diag}\left\langle \left(\tilde{\mathbf{x}}\left(t_{now}+\tau\right) - \mathbf{x}\left(t_{now}+\tau\right)\right)\left(\tilde{\mathbf{x}}\left(t_{now}+\tau\right) - \mathbf{x}\left(t_{now}+\tau\right)\right)^{T}\right\rangle, \tag{7}$$

{varmin}

that is the variance of the state about the true value at some time *future* to t_{now}. Brackets again denote the expected value. The role of the model is to make the forecast, by setting $\mathbf{u}\left(t\right) = 0$, $t_{now} + \Delta t \leq t \leq t_{now} + \tau$, because it is unknown, and starting with the most recent estimate $\tilde{\mathbf{x}}\left(t_{now}\right)$ at t_{now}. Eq. (7) is itself equivalent to a requirement of minimum square deviation at $t_{now} + \tau$. A bit more will be said about this relationship below.

Model error deserves an extended discussion by itself. A consequence of exact satisfaction of the model equations is that we assume the discretized version of Eq. (1) to be error-free. Model errors comes in roughly three flavors: (a) the equations are incomplete or an approximated form of the real system; (b) errors are incurred in their discretization (e.g., numerical diffusion); (c) sub-grid scale parameterizations are incomplete, and/or their parameter choices sub-optimal. Methods exist to quantify these errors in an estimation framework. For example, an explicit error term may be introduced in Eq. (1) and whose estimation would become part of the least-squares optimization. Problems arise in practice when observational coverage is insufficient to achieve adequate partition of errors between those in the explicit error terms and those in the initial and boundary conditions. Furthermore, the error terms are effectively source terms in the tendency equations that violate dynamic and kinematic consistency. The approach taken in ECCO is to move toward estimating three-dimensional fields for the most important parameters as part of the gradient-based optimization (e.g., Ferreira et al., 2005; Stammer, 2005), thus rendering the problem one of combined state and parameter estimation. Adjustments required to compensate for model errors may be projected into the parameter estimates.

Most of the fundamental principles of practical state estimation and of DA can be understood from the common school problem of the least-squares fitting of lines and curves to data in one dimension. The central point is that the *concepts* of state estimation and data assimilation are very simple; but it is equally simple to surround them with an aura of mystery and complexity that is needless for anyone who wishes primarily to understand the meaning of the results.

3 Data Assimilation and the Reanalyses

Despite the technical complexities of the numerical engineering practice, DA and what are called "reanalyses" should be understood as approximate methods for obtaining a solution of a least-squares problem. Using the same notation as in Eq. (7), consider again an analysis time, $t_{now} = t_{ana} + \tau$, when data have become available, and where t_{ana} is the previous analysis time, $\tau > 0$, typically 6 hours earlier. The weather forecaster's model has been run forward to make a prediction, $\tilde{\mathbf{x}}\left(t_{now}, -\right)$, with the minus sign denoting that newer observations have *not yet* been used. The new observations are $\mathbf{E}\left(t_{now}\right)\mathbf{x}\left(t_{now}\right) + \mathbf{n}\left(t_{now}\right) = \mathbf{y}\left(t_{now}\right)$. With some understanding of the quality of the forecast, expressed in the form of an uncertainty matrix (2nd moments about the truth) called $\mathbf{P}\left(t_{now}, -\right)$, and of the covariance matrix of the observational noise, $\mathbf{R}\left(t_{now}\right)$, the best combination in the L_2-norm of the information of the model and the data is the minimum of,

$$\begin{aligned} J_1 = & \left(\tilde{\mathbf{x}}\left(t_{now}\right) - \tilde{\mathbf{x}}\left(t_{now}, -\right)\right)^T \mathbf{P}\left(t_{now}, -\right)^{-1} \left(\tilde{\mathbf{x}}\left(t_{now}\right) - \tilde{\mathbf{x}}\left(t_{now}, -\right)\right) + \\ & \left(\mathbf{y}\left(t_{now}\right) - \mathbf{E}\left(t_{now}\right)\mathbf{x}\left(t_{now}\right)\right)^T \mathbf{R}\left(t_{now}\right)^{-1} \left(\mathbf{y}\left(t_{now}\right) - \mathbf{E}\left(t_{now}\right)\mathbf{x}\left(t_{now}\right)\right), \end{aligned} \tag{8}$$

{kalman2}

and whose least-squares minimum for a linear model is given rigorously by the Kalman filter. In DA practice, only some very rough approximation to that minimum is sought and obtained. True Kalman filters are never used for prediction in real geophysical fluid flow problems as they are computationally overwhelming (for more detail, see e.g., Wunsch, 2006). *Notice that J_1 assumes that a summation of errors is appropriate, even in the presence of strong nonlinearities.*

A brief excursion into meteorological "reanalyses" is worthwhile here for several reasons: (1) They are often used as an atmospheric "truth" to drive ocean, ice, chemical, and biological models. (2) A number of ocean circulation estimates have followed their numerical engineering methodology. (3) With the long history of the atmospheric data assimilation effort, many have been unwilling to believe that any alternatives exist.

Note that the "analysis" consists of an operational weather model run in conventional prediction mode, analogous to the simple form described in the previous section, adjusted, and thus displaying discontinuities at the analysis times, by attempts to approximately minimize J_1. Because of the operational/real-time requirements, only a fraction of the global operational meteorological observations are relayed and quality-controlled in time to be available at the time of analysis. Furthermore, because models have changed so much over the years, the stored analyses are inhomogeneous in the underlying physics[3] and model codes. Oceanographers have no such products at this time; *"analyses" in the meteorological sense do not exist,* and thus the jargon "reanalysis" for ocean state estimates is inappropriate.

Meteorological reanalysis is the recomputation, using the same prediction methodology as previously used in the analysis, but with the differences that (1) the model code and combination methodology are held fixed over the complete time duration of the calculation (e.g., over 50 years) thus eliminating artificial changes in the state from model or method improvements and, (2) including many data that arrived too late to be incorporated into the real-time analysis (see Kalnay, 2003; Evensen, 2009).

Estimated states still have the same discontinuities at the analysis times when the model is forced towards the data. Of even greater significance for oceanographic and climatic studies are the temporal shifts induced in the estimates by the major changes that have taken place in the observational system over several decades—most notably, but not solely, the appearance of meteorological satellites. Finally, no use is made of the information content in the observations of the *future* evolution of the state.

Although as already noted above, clear warnings have appeared in the literature—that spurious trends and values are artifacts of changing observation systems (and see for example, Marshall, et al., 2002; Elliott and Gaffen, 1992; Thompson et al., 2008)—the reanalyses are rarely used appropriately, meaning with the recognition that they are subject to large errors. In Fig. 1 for example, the jump in precipitation minus evaporation ($P - E$) with the advent of the polar orbiting satellites implies either that the unspecified error estimates prior to that time must, at a minimum, encompass the jump, and/or that computation has been erroneous, or that a remarkable coincidence has occurred. But even the smaller transitions in $P - E$, e.g., over the more recent period 1992 onward, are likely too large to be physical; see Table 1.

Figure 2 and other, similar ones, are further disquieting, showing that reanalyses using essentially the same data, and models that have been intercompared over decades, have significant qualitative disagreements on climate time scales. Differences in the reanalyses in the northern hemisphere are not so large, and are generally agreed to be the result of a much greater data density. They remain, nevertheless, significant, as evidenced in the discussion of analysis increments over the Arctic by Cullather and Bosilovich (2012). *Evidently, considerations of data density and types dominate the reanalyses, with the models being of secondary importance.*

For climate studies, another major concern is the failure of the reanalyses to satisfy basic global conservation requirements. So for example, Table 1 shows the global imbalances on a

[3] We employ "physics" in its conventional meaning as encompassing all of dynamics and thermodynamics.

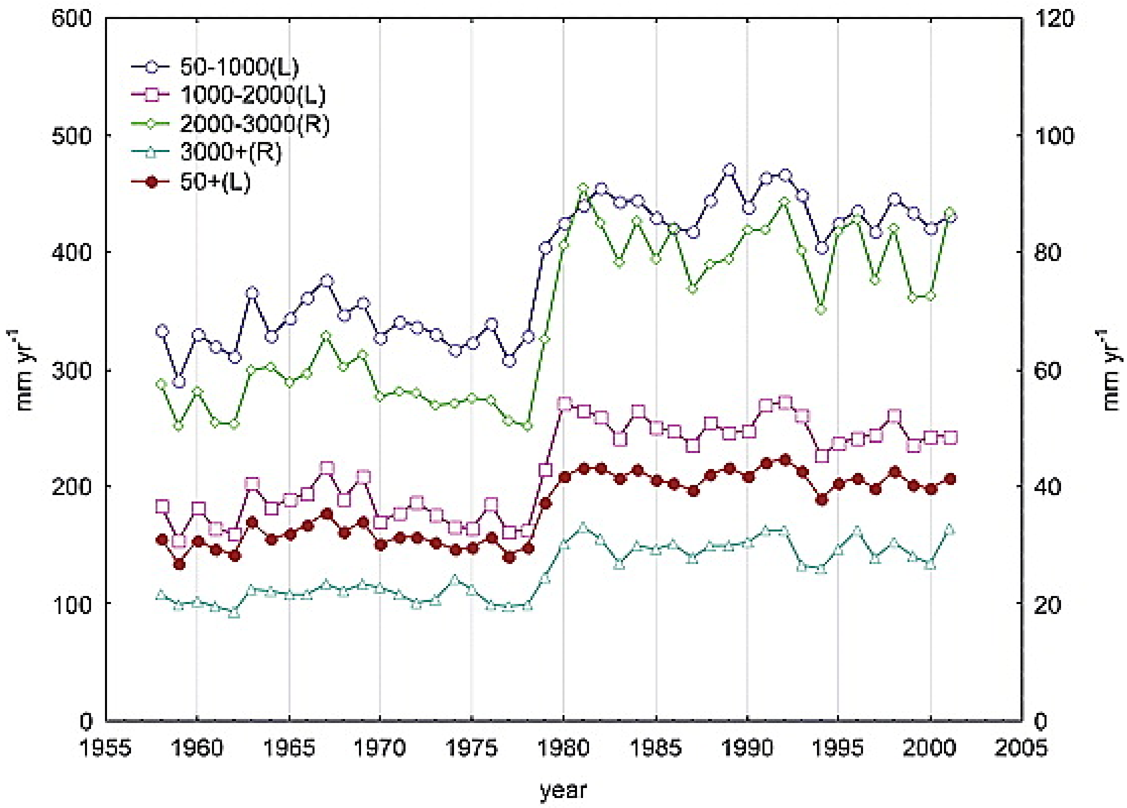


Figure 1: Mean annual precipitation minus evaporation over the Antartctic as a function of time in the ECMWF reanalysis ERA-40 showing the impact of new observations, in this case, the arrival of the polar orbitting satellites (from Bromwich et al., 2007). Different curves are for different elevations. The only simple inference is that the uncertainties must exceed the size of the rapid transition seen in the late 1970s. L and R identify whether the left or right axis is to be used for that curve. {bromwich_etal

per year basis of several reanalysis products in apparent heating of the oceans and in the net freshwater flux from the atmosphere. Such imbalances can arise either because global constraints are not implied by the model equations, and/or because biassed data have not been properly handled, or most likely, some combination of these effects is present. Trenberth and Solomon (1994) for example, note that the NCEP/NCAR reanalysis implies a meridional heat transport within continental land masses. "User beware" is the best advice we can give.

State estimation as defined in the ECCO context is a much more robust and tractable problem than is, for example, prediction of future climate states. As is well known even to beginning scientists, extrapolation of very simple models can be extremely unstable, with interpolation [4],

[4]The commonplace term "interpolation," is used in numerical analysis to imply that fitted curves pass exactly through data points—an inappropriate requirement here.

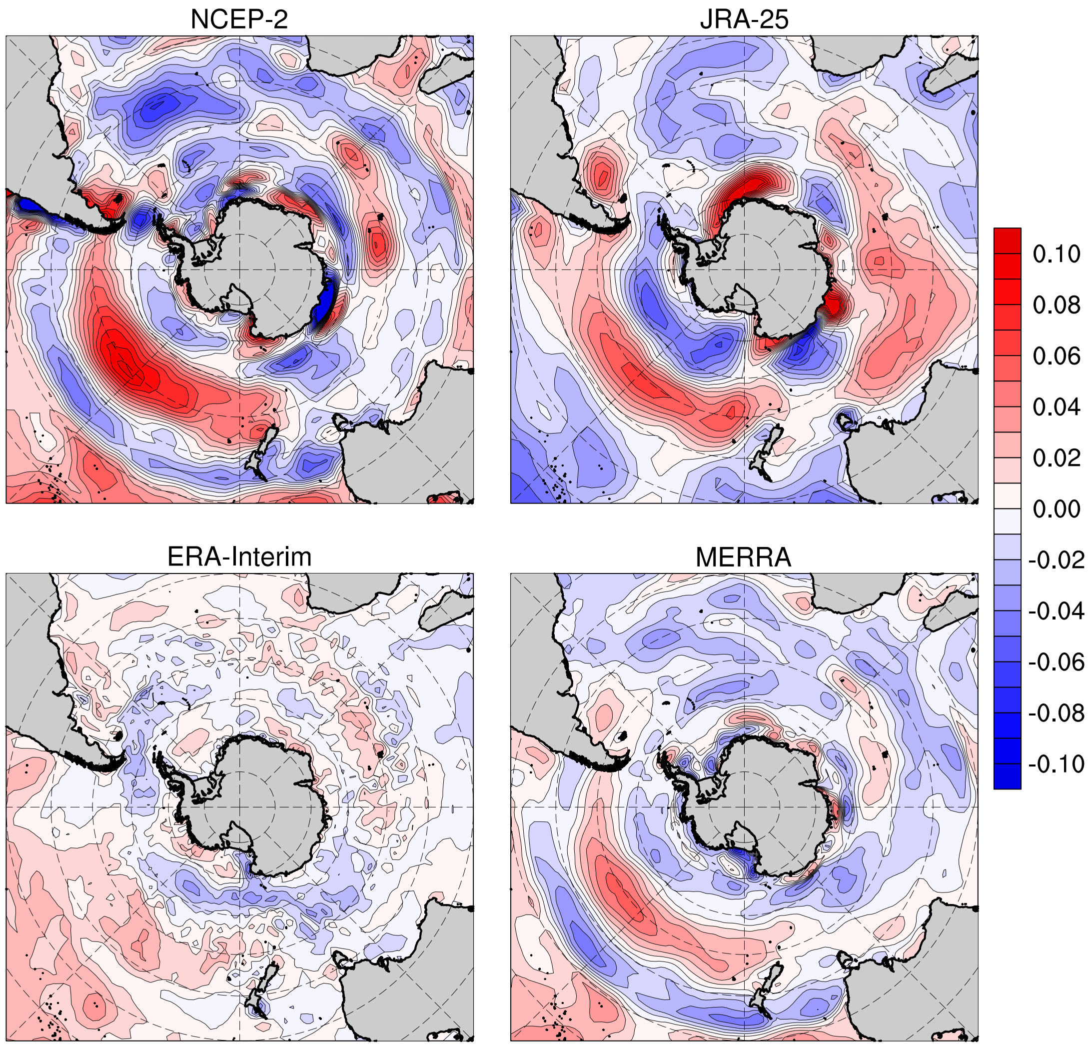


Figure 2: Calculated trends (meters/second/year) in the 10meter zonal wind fields at high southern latitudes from four different atmospheric reanalyses (D. Bromwich and J. P. Nicolas, of Ohio State University, private communication, 2010). Note particularly the different patterns in the Indian Ocean and the generally discrepant amplitudes. Because of the commonality of data sets, forecast models, and methodologies, the differences here must be lower bounds on the true uncertainties of trends. See Bromwich et al. (2011) for a description of the four different estimates. {bromwich&nico

or curve-fitting, remaining robust. (A classical example is the use of a cubic polynomial to fit some noisy data, and which can be very effective. But one is advised *never* to use such a fit to extrapolate the curve; see Fig. 3). The ECCO process is effectively a temporal curve-fit of the WOCE-era data sets by a model and which, with some care to avoid data blunders, produces a robust result. It is the interpolating ("smoothing") character, coupled with the expectation

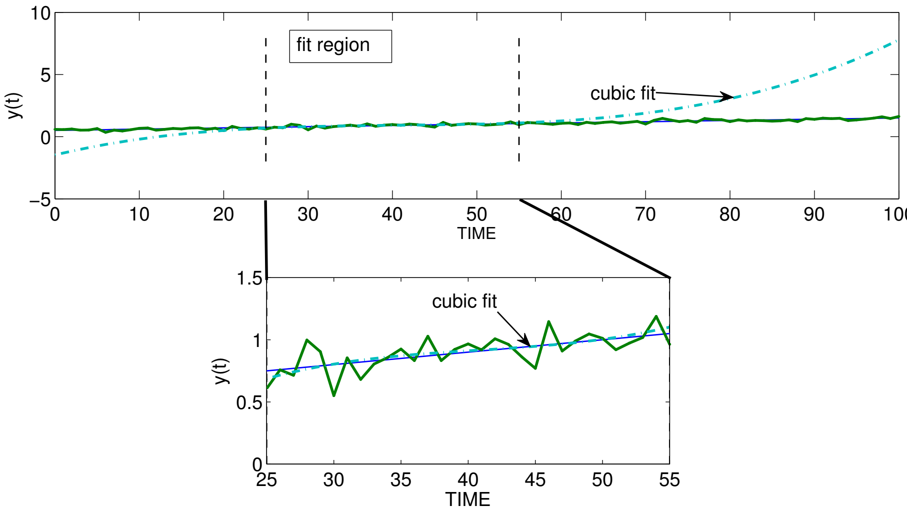


Figure 3: A textbook example of the robust interpolation of noisy data by a cubic polynomial and its gross instability when used to extrapolate. This analogue is a very simplified representation of the problem of extrapolating a GCM state into unobserved time spans.

{cubic_fit.eps

of thermal wind balance over most of the domain, that produces confidence in the basic system products. As is well-known, least-squares methods tend to generate meaningless structures in unconstrained parts of the domain. Some regions of spatial extrapolation do exist here, depending upon the time-varying distribution of observations, and although they tend to be limited in both space and time, detailed values there should be regarded skeptically.

Terminological Note. The observational community has lost control of the word “data,” and which has come to be used, confusingly, for the output of models, rather than having any direct relationship to instrumental values. In the context of reanalyses and state estimates involving both measurements and computer codes, the word generally no longer conveys any information. For purposes of this essay, “data” always represents instrumental values of some sort, and anything coming out of a GCM is a “model-value” or “model-datum,” or similar label. We recognize that models are involved in all real observations, even in such familiar values as those coming from e.g., a mercury thermometer, in which a measured length is converted to a temperature. Most readers can recognize the qualitative difference between conventional observations and the output of a 100,000+ line computer code.

4 Ocean State Estimates

The remainder of this paper is primarily devoted to a summary description and discussion of some results of the Estimating the Circulation and Climate of the Ocean (ECCO) groups which, beginning with Stammer et al. (2002), were directed at decadal and longer state estimates satisfying known equations and using as much of the WOCE-era-and-beyond data as possible. No claim is made that these estimates are definitive, nor that the discussion is comprehensive. A number of other, superficially similar, estimates exist (Carton et al. 2000; Martin et al., 2007; Hurlburt et al., 2009), but these generally have had different goals e.g., a fast approximate estimate primarily of the upper ocean, or prediction of the mesoscales over ocean basins. Some weather forecasting centers have undertaken "operational oceanography" products closely resembling atmospheric weather forecasts. To our knowledge, however, the ECCO estimates are today the only ones specifically directed at physically continuous, dynamically consistent, top-to-bottom estimates from a comprehensive data set.

A number of review papers exist that attempt to compare different such solutions (e.g., Carton and Santorelli, 2008; Lee et al., 2010) as though they were equivalent. But as the above discussion tries to make clear, estimates are not equally reliable for all purposes and comparisons make no sense unless their individual purposes are well understood. Although one could compare a crop-dusting airplane to a jet fighter, and both have their uses, few would regard that effort as helpful except as a vehicle for discussion of the highly diverse applications of aero-physics. Thus a numerical scheme directed primarily at mesoscale prediction, and using a model not conserving energy, may well be a useful tool for forecasting the trajectory of the Gulf Stream over a few weeks, but it would be unsuited to a discussion of global ocean heat transports—a useful model of which is, in turn, unsuitable for mesoscale interests. These other applications are discussed in this volume by Schiller et al. (2012).

Originally, ECCO was meant primarily to be a demonstration of the practicality of its approach to finding the oceanic state. When the first ECCO estimates did become available (Stammer et al., 2002) they proved sufficiently useful even with that short duration and coarse resolution, that a decision was made to continue with a gradually improving data set and computer power. This review summarizes mainly what has been published thus far, but as optimization is an asymptotic process, the reader should be aware that newer, and likely better, solutions are being prepared continuously and the specific results here will have been refined in the intervals between writing, publishing, and reading.

4.1 Basic Notions

As described above, most state estimation problems in practice are generically those of constrained least-squares, in which one seeks to minimize objective or cost or misfit functions similar to Eq. (6) subject to the solution (including both the estimated state $\mathbf{x}(t)$, and the controls, $\mathbf{u}(t)$) of the model-time stepping equations.[5] One approach, among many, to solving such problems is the method of Lagrange multipliers (MLM) dating back 200 years. This method is discussed at length in Wunsch (2006) and the references there. In a very brief summary, one "adjoins" the model equations using vectors of Lagrange multipliers, $\boldsymbol{\mu}(t)$, to produce a new objective function,

$$\begin{aligned}
J' = &\sum_{m=0}^{M} \left(\mathbf{y}(t) - \mathbf{E}(t)\,\mathbf{x}(t)\right)^T \mathbf{R}(t)^{-1} \left(\mathbf{y}(t) - \mathbf{E}(t)\,\mathbf{x}(t)\right) \\
&+ \sum_{m=0}^{M-1} \mathbf{u}(t)^T \mathbf{Q}(t)^{-1} \mathbf{u}(t) \\
&- 2\sum_{m=1}^{M-1} \boldsymbol{\mu}(t)^T \left[\mathbf{x}(t) - \mathbf{L}\mathbf{x}(t-\Delta t), \mathbf{B}\mathbf{q}(t-\Delta t), \boldsymbol{\Gamma}\mathbf{u}(t-\Delta t)\right], \\
& t = m\Delta t, \quad m = 0, \ldots, M
\end{aligned} \tag{9}$$

{Jp}

Textbooks explain that the problem can now be treated as a conventional, unconstrained least-squares problem in which the $\boldsymbol{\mu}(t)$ are part of the solution. In principle, one simply does vector differentiation with respect to all of $\mathbf{x}(t)$, $\mathbf{u}(t)$, $\boldsymbol{\mu}(t)$, sets the results to zero, and solves the resulting "normal equations" (they are written out in Wunsch, 2006). J and J' are very general, and one easily adds e.g., internal model parameters such as mixing coefficients, water depths, etc. as further parameters to be calculated, thus rendering the problem one of combined state and parameter estimation.

The entire problem of state estimation thus reduces to finding the stationary values of J'. The large literature on what is commonly called the "adjoint method" ("4DVAR" in weather forecasting, where it is used only incrementally over short time-spans) reduces to coping with a very large set of simultaneous equations (and some are nonlinear). But as an even larger literature deals with solving linear and nonlinear simultaneous equations by many methods, ranging from direct solution, to downhill search, to Monte Carlo, etc., most of the discussion

[5] Advantages exist to using norms other than L_2 including those such as one and infinity norms commonly regarded as robust. These norms are not normally used in ocean and atmosphere state estimation or data assimilation systems because software development for parallel computers permitting computation at super-large dimensions has not yet occurred.

of adjoint methods reduces to technical details, many of which are complex, but which are primarily of interest to computer-code constructors (Heimbach et al., 2005). Within the normal equations, the time-evolution of the Lagrange multipliers is readily shown to satisfy a set of equations usually known as the "adjoint" or "dual" model. This dual model can be manipulated into a form having time run "backwards," although that interpretation is unnecessary; see the references.

One very interesting complication is worth noting: the description in the last two paragraphs assumes one can actually differentiate J, J'. In oceanographic practice, that implies differentiating the computer code which does everything. The "trick" that has made this method practical for GCMs is so-called automatic (or algorithmic) differentiation (AD), in which a software tool can be used to produce the partial derivatives and their transposed values—in the form of another software code (see, e.g., Giering and Kaminski, 1998; Griewank and Walther, 2008; Utke et al., 2008). This somewhat bland statement hides a complex set of practical issues; see e.g., Heimbach et al. (2005) for discussion in the context of the MIT general circulaition model (MITgcm). Most of the difficult problems are of no particular concern to someone mainly interested in the results.[6]

As discussed in more detail by Wunsch and Heimbach (2007), the central ECCO estimates are based upon this Lagrange multiplier method, with the state estimates obtained from the adjusted, but then *freely running*, MITgcm, as is required in our definition of state estimation. At the time of this writing, most of the estimates have restricted the control variables (the adjustable parameters) to the initial conditions and the meteorological forcing, although following exploratory studies by Ferreira et al. (2005), Stammer (2005) and Liu et al. (2012a), state estimates are beginning to become public that also adjust internal model parameters, such as isopycnal, thickness or vertical diffusion.

A full modern oceanic general circulation model (GCM or OGCM) such as that of Marshall et al. (1997) as modified over subsequent years (e.g., Adcroft et al., 2004; Campin et al., 2004), is a complex machine consisting of hundreds thousands of lines of code encompassing the Navier-Stokes equations, the relevant thermodynamics, sea ice and mixed-layer sub-codes, various schemes to represent motions below the model resolution (whatever it may be), and further subsidiary codes for overflow entrainment, etc. Understanding such a model is a difficult proposition, in part because different elements were written by different people over many years, sometimes without full understanding of the potential interactions of the existing or future

[6]The situation is little different than that in ordinary ocean GCM studies. Technical details of time-stepping, storage versus recomputation, re-starts, etc. are very important and sometimes very difficult, but not often of consequence to most readers, except where the author necessarily calls attention to them.

subcomponents. Furthermore, various studies have shown the inevitability of coding errors (e.g., Basili et al., 1992) and unlike the situation with the real ocean, one is faced with determining if some interesting or unusual behavior is real or an artifact of interacting, possibly very subtle, errors. (Nature presumably never solves the incorrect equations; but observational systems do have their own mysteries that must be understood: recent examples include the discovery of systematic errors in fall rates to infer the depth of XBT data e.g., by Wijffels et al. (2008), and calibration errors of pressure sensors onboard some of the Argo floats (Barker et al., 2011)).

By recognizing that most algorithms can be regarded as directed at the approximate solution of a least-squares problem, one can exploit the two-hundred-year history of methodologies that have emerged (e.g., Björck, 1996), substituting differing numerical algorithms where necessary. For example Köhl and Willebrand (2002) and Lea et al. (2002) suggested that the Lagrange multiplier method would fail when applied at high resolution to oceanic systems that had become chaotic. Although such behavior has been avoided in oceanographic practice (Gebbie et al., 2006; Hoteit et al., 2006; Mazloff et al. 2010), one needs to separate the *possible* failure of a particular numerical algorithm to find a constrained minimum from the inference that no minimum exists. If local gradient descent methods are not feasible in truly chaotic systems, one can fall back on variations of Monte Carlo or other more global methods. Obvious failure of search methods using local derivatives has had limited importance in oceanographic practice. This immunity is likely a consequence of the observed finite time interval in the state estimation problem, in which structures such as bifurcations are tracked adequately by the formally future data, providing adequate estimates of the algorithmic descent directions. Systematic failure to achieve an acceptable fit to the observations can lead to accepting the hypothesis that the model should be rejected as an adequate representation. Potential model falsification is part of the estimation problem, and is the pathway to model improvement.

Modern physical oceanography is largely based upon inferences from the thermal wind, or geostrophic-hydrostatic, equations. Scale analyses of the primitive equations (e.g., Pedlosky, 1987; Vallis, 2006; Huang, 2010) all demonstrate that apart from some very exceptional regions of small area and volume, deviations from geostrophic balance are slight. This feature is simultaneously an advantage and a liability. It is an advantage because any model, be it analytical or numerical must, to a first approximation satisfy the linear thermal wind equations. It is a liability because it is only the deviations which define the governing physics of the flow maintenance and evolution, and which are both difficult to observe and to compute with adequate accuracy. In the present context, one anticipates that over the majority of the oceanic volume, any plausible model fit to the data sets must be, to a good approximation, a rendering of the ocean circulation in geostrophic, hydrostatic, balance, with Ekman forcing, and volume or mass conservation imposed regionally and globally as an automatic consequence of the model configuration. The most visible ageostrophic physics are the variability, seen as slow, accumulating, deviations from an initial state.

4.2 The Observations

Data sets used for many (not all) of the ECCO family of solutions are displayed in Table 2. As noted in the Introduction, they are of very diverse type, geographical and temporal distribution, and with very different accuracies and precisions.

As is true of any least-squares solution, no matter how it is obtained, the results are directly dependent upon the weights or error variances assigned to the data sets. An over-estimate of the error corresponds to the suppression of useful information; an under-estimate to imposing erroneous values and structures. Although an unglamorous and not well-rewarded activity, a quantitative description of the errors is essential and is often where oceanographic expertise is most central. Partial discussions are provided by Stammer et al. (2007), Ponte et al. (2007), Forget and Wunsch (2007), and Ablain et al. (2009). Little is known about the space-time covariances of these errors, information, which if it were known, could improve the solutions (see Weaver and Courtier, 2001, for a useful direction now being used in representing spatial covariances). Model errors, which dictate how well estimates should fit to hypothetical perfect data, are extremely poorly known and are generally added to the true data error—as in linear problems the two types of error are algebraically indistinguishable.

5 Global Scale Solutions

Solutions of this type were first described by Stammer et al. (2001, 2002, 2003) and were computed on a $2° \times 2°$ grid with 22 vertical levels. As the computing power increased, a shift was made to a $1° \times 1°$, 23-level solution and that, until very recently, has remained the central vehicle for the global ECCO calculations. Although some discrepancies continue to exist in the ability to fit certain data types, these solutions (Wunsch and Heimbach, 2007) based as they are on geostrophic, hydrostatic balance over most of the domain, were and are judged adequate for the calculation of large-scale transport and variability properties. The limited resolution does mean that systematic misfits were expected, and are observed, in special regions such as the western boundary currents. Often the assumed error structures of the data are themselves of doubtful accuracy.

As noted above, Ganachaud (2003a) inferred that the dominant error in trans-oceanic transport calculations of properties arose from the temporal variability. Perhaps the most important lesson of the past decade has been the growing recognition of the extent to which temporal aliasing is a serious problem in calculating the oceanic state. For example, Figs. 4, 5 display the global meridional heat and fresh water transport as a function of latitude along with their standard errors computed from the monthly fluctuations. The figures suggest that errors inferred from hydrography are under-estimated (and error estimates of the non-eddy resolving ECCO estimates are themselves lower bounds of the noise encountered in the real ocean). The classical oceanographic notion that semi-synoptic sections are accurate renderings of the time-average properties, while having some qualitative utility, has now to be painfully abandoned—an essential step if the subject is to be a quantitative one. Temporal effects are most conspicuous at low latitudes, but in many ways, the difficulty is greatest at high latitudes: the long time scales governing behavior there mean that the hydrographic structure is very slowly changing, requiring far longer times to produce an accurate time-mean. In other words, a 10-year average at 10°N will be a more accurate estimate of the longer-term mean than one at 50°N. Even this comment begs the question of whether a stable long-term mean exists, or whether the system drifts over hundreds and thousands of years? This latter is a question concerning the frequency spectrum of oceanic variability, and which is very poorly known at periods beyond a few years.

For the 19+ years now available in the global state estimates, most of the large-scale properties, including the time variations, are stable from one particular set of assumptions to others, probably as a consequence of the dominance of overall geostrophic balance and the comparatively well-sampled hydrography and altimetric slopes. They are thus worth analyzing in detail. The intricacies of the global, time-varying ocean circulation are a serious challenge to the summarizing capabilities of authors. A full discussion, however, of the global state estimates becomes a discussion of the complete three-dimensional time-varying ocean circulation, a subject requiring a book, if not an entire library, encompassing distinctions amongst time and space scales, geographical position, depth, season, trends, the forcing functions (controls). No such synthesis is attempted here! Instead we can only give a bit of the flavor of what can and has been done with the estimates along with enough references for the reader to penetrate the wider literature.

Note too, as discussed e.g., by Heimbach et al. (2011) and other, earlier efforts, the Lagrange multipliers are the solution to the dual model. As such, they are complete solutions in three spatial dimensions and time, and convey the sensitivity of the forward model to essentially any parameter or boundary or initial condition in the system. The information content of the dual solutions is very large—representing not only the sensitivities of the solution to the data and model parameters and boundary and initial conditions, but also the flow of information through the system. Analyzing the dual solution does, however, require the same three-dimensional time-dependent representations of any full GCM, and these elements of the state estimates remain

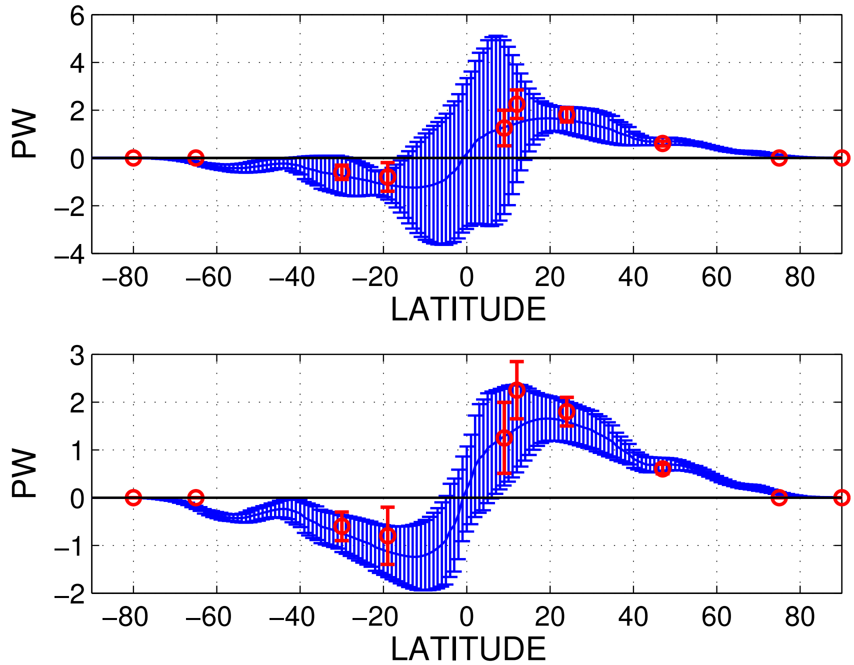


Figure 4: Global meridional heat transport in the ocean from ECCO-Production version 4 (G. Forget, private communication, 2011). Upper panel shows the standard error including the annual cycle and the lower one, with the annual cycle removed—as being largely predictable. Possible systematic errors are not included. Red dots with error bars are estimates from Ganachaud and Wunsch (2002). Note that the WOCE-era hydrographic survey failed to capture the southern hemisphere extreme near 10°S, thus giving an exaggerated picture of the oceanic heat transport asymmetry about the equator. {heat_transpor

greatly under-exploited at the present time.

5.1 Summary of Major, Large-Scale Results

None of the results obtained so far can be regarded as the final state estimate: obtaining fully consistent misfits by the model to the observations has never been achieved (see the residual misfit figures in the references). Misfits linger for a variety of reasons, including the sometimes premature termination of the descent algorithms before full optimization, mis-representation of the true model or data errors, or selection of a local rather than a global minimum in the major nonlinear components of the model. As with all very large nonlinear optimization problems, approach to the "best" solution is asymptotic. With these caveats, we describe some of the more salient oceanographic features of the recent solutions, with no claim to being comprehensive. Note that results from a variety of ECCO-family estimates are used, largely dictated by the

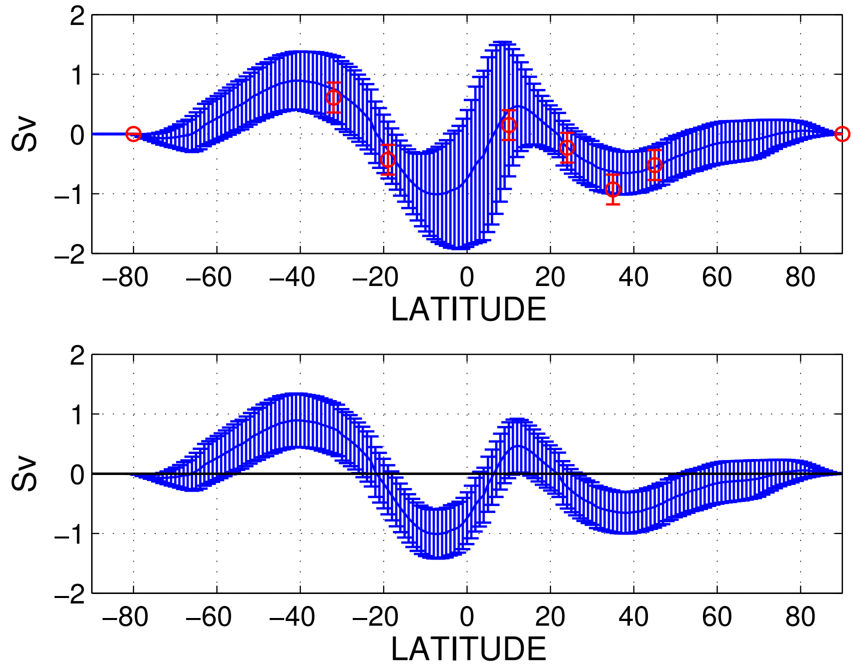


Figure 5: Same as Fig. 4 except for the freshwater transport (G. Forget, private communication, 2011). Upper panel shows standard errors that include the seasonal cycle, and the lower without the seasonal cycle. Red dots are again from Ganachaud and Wunsch (2002).

{fw_transport_

particular problem that was the focus of the calculation.

Volume, Enthalpy, Freshwater Transports and Their Variability

The most basic elements describing the ocean circulation and its large-scale variability are usually the mass (or volume, which is nearly identical) transports. Stammer et al. (2001, 2002, 2003) depicted the basic global-scale elements of the mass transport as averaged over the duration of their estimates. A longer duration estimate (v3.73) has been used (Fig. 6) to compute the vertically integrated volume stream function. We reiterate that diagrams such as this one are finite duration averages whose relationship to hypothetical hundred year or longer climatologies remains uncertain.

Fig. 7 shows the zonally integrated and vertically accummulated meridional transport as a function of depth and ocean. The very large degree of temporal variability can be seen in Fig. 4 from a new fully global solution which is about to become available online at the time of writing (ECCO-Production version 4; see Table 3) with error bars derived from the temporal variances. These time averages have been an historically important goal of physical oceanography, albeit estimates derived from unaveraged data were commonly assumed without basis to accurately

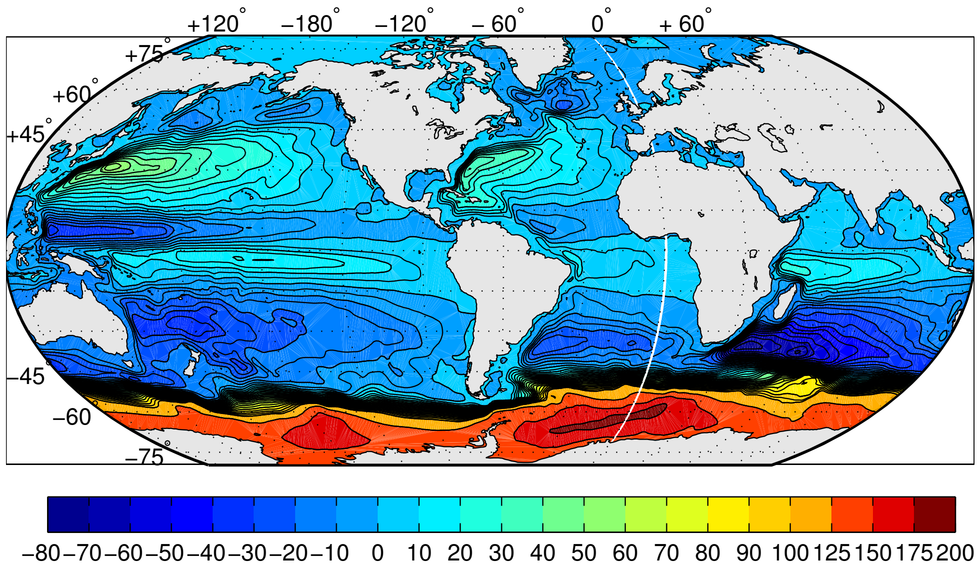


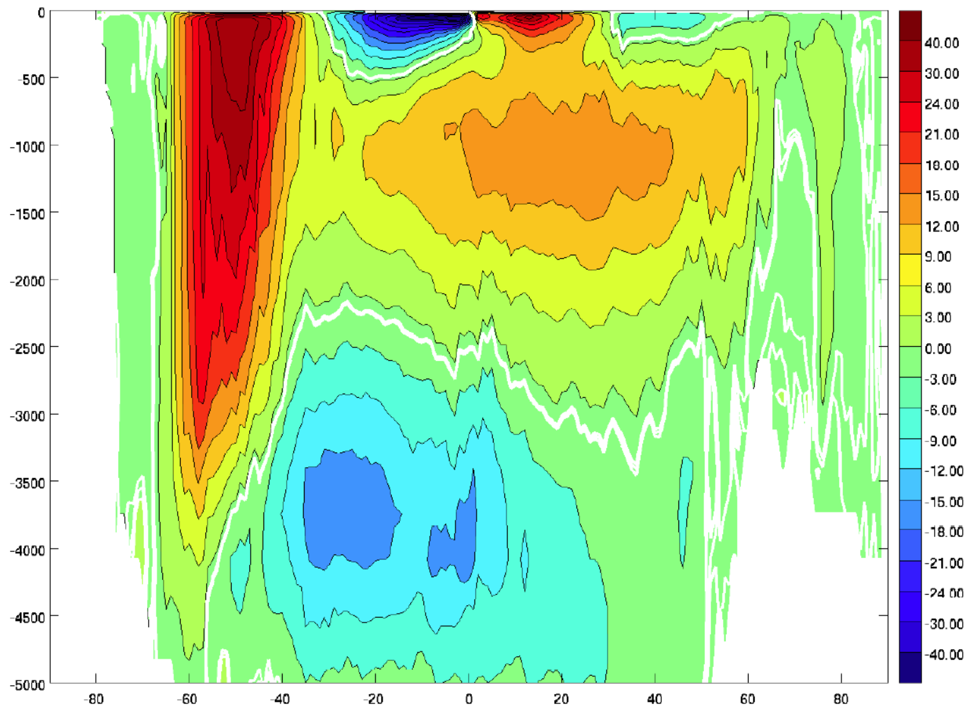
Figure 6: The top-to- bottom transport stream function from ECCO v3.73 (Wunsch, 2011). Qualitatively, the wind-driven gyres dominate the result, with the intense transports in the Southern Ocean particularly conspicuous.

{transportstre

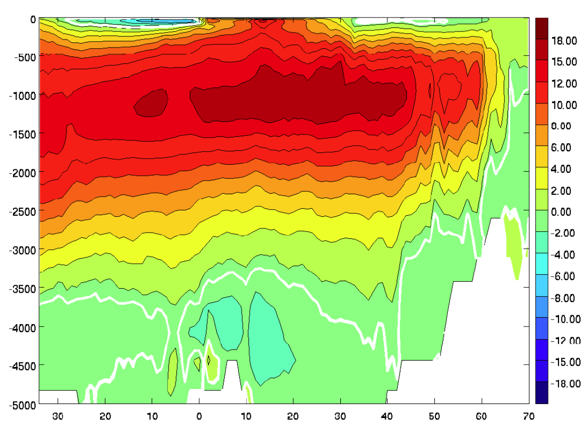
depict the true time-average. Perhaps the most important utility of the existing state estimates has been the ability, at last, to estimate the extent of the time-variability taking place in the oceans (Wunsch and Heimbach, 2007, 2009, 2012). Withheld, direct in situ observations in a few isolated regions (Kanzow et al., 2009; Baehr, 2010) are consistent with the inference that even volume transports integrated across entire ocean basins have a large and qualitative temporal variability. More generally, mooring data and the now almost 20-year high resolution high accuracy altimetric records all show the intense variations that exist everywhere. With ECCO-like systems, syntheses of these data sets are now possible.

The Annual Cycle

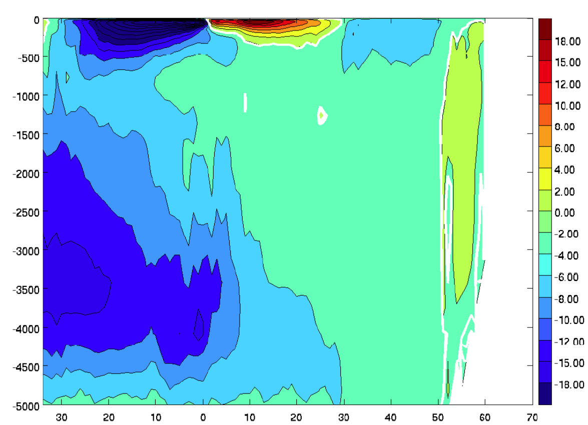
The annual cycle of oceanic response is of interest in part because the ultimate forcing function (movement of the sun through the year) is very large and with very accurately known structure. In practice, that forcing is mediated through the very complex atmospheric annual changes, and understanding how and why the ocean shifts seasonally on a global basis is a difficult problem. Using the ECCO state estimates, Vinogradov et al. (2008) mapped the amplitude and relative contributions for salt and heat of the annual cycle in sea level (Figs.



(a) Global



(b) Atlantic Ocean



(c) Pacific & Indian Ocean

Figure 7: Mean (1992-2010) of the meridional volume transport stream function in Sverdrups (Sv-$10^6\mathrm{m}^3/\mathrm{s}$) from ECCO-Production version 4 (Wunsch and Heimbach, 2012; Forget et al., in prep. 2012). Panel (a) is the global result; panels (b,c) are the Atlantic, and the combined Indo-Pacific, respectively. Note the complex equatorial structure, and that this representation integrates out a myriad of radically different dynamical sub-regimes. In the Southern Ocean, interpretation of zonally integrated Eulerian means requires particular care owing to the complex topography and relatively important eddy transport field.

{mean_merid.ep

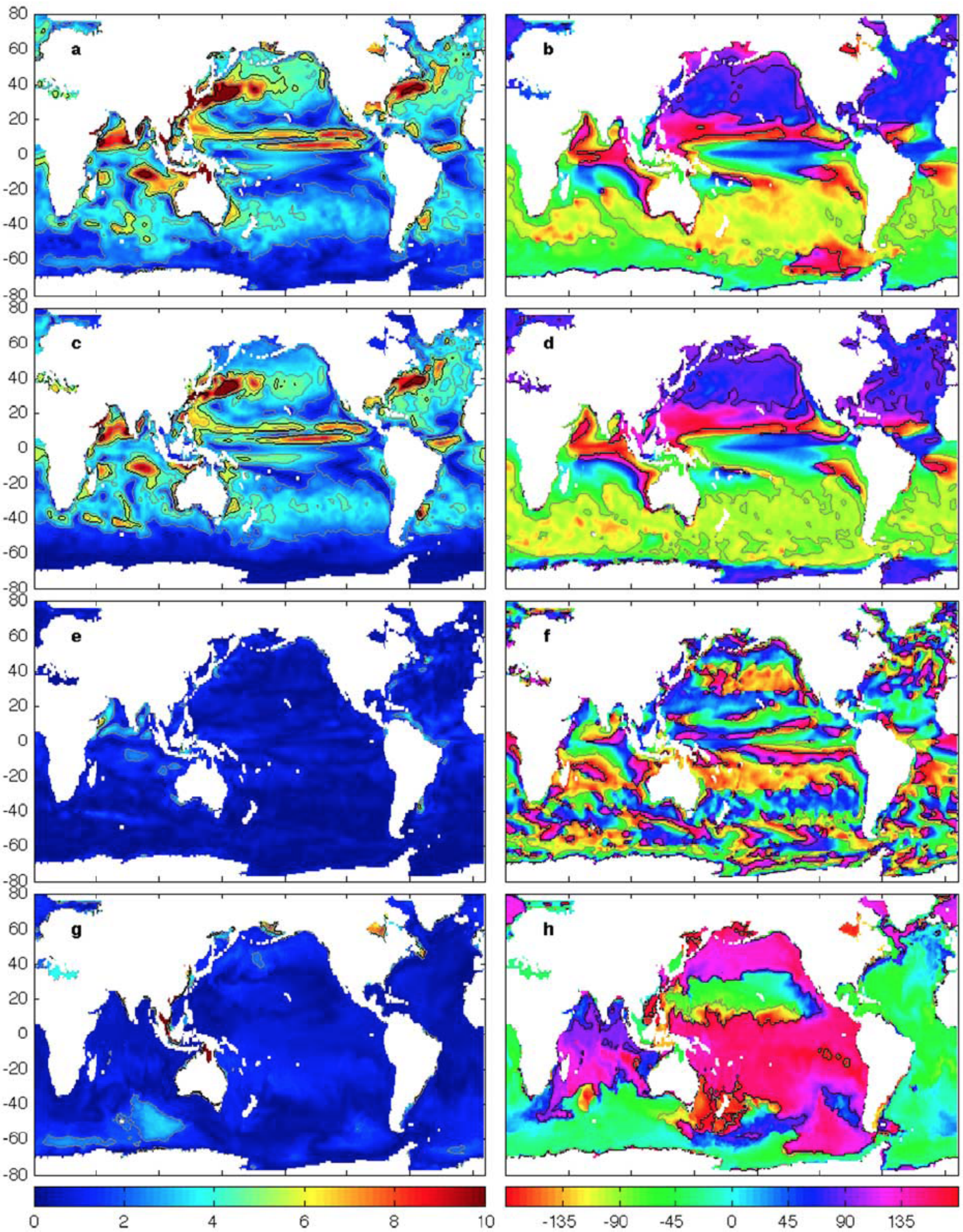


Figure 8: From Vinogradov et al. (2008) showing the annual cycle in sea level from ECCO Climate State v2.177. Left column is the amplitude in cms and the right column the phase. From top-to-bottom, they are the surface elevation (a,b), the thermosteric component (c,d), the halosteric component (e,f), and at bottom, the bottom pressure (g,h).

{vinogradov_et

8. The importance of the annual cycle, more generally, is visible in Fig. 4, 5 as the large contribution to the standard errors.

Sea Level Change

The sea surface height is simultaneously a boundary condition on the oceanic general circulation and a consequence of that circulation. Because of the intense interest in possible large-scale changes in its height and the potential shifts in vulnerability to storm surges, and associated issues such as ecosystem and freshwater reservoir declines, the ECCO state estimation system has been used to estimate the shifts taking place in the era since 1992 (Wunsch et al., 2007). A summary of a complex subject is that sea level change is dominated by regional variations more than an order of magnitude larger than the putative global average, and arising primarily from wind field shifts. Varying spatial contributions from competing exchanges of freshwater and heat with the atmosphere and the extremely inhomogeneous (space and time) in situ data sets render the global mean and its underlying causes far more uncertain than some authors have claimed.

At the levels of accuracy appearing to be required, very careful attention must now be paid to modeling issues such as water self-attraction and load (Vinogradova et al., 2011, Kuhlmann et al., 2011) not normally accounted for in OGCMs. Conventional approximations to the moving free-surface boundary conditions generate systematic errors no longer tolerable (e.g., Huang, 1993; Wunsch et al., 2007). Usefully accurate sea level estimation over multiple decades may be the most demanding requirement on both models and data sets now facing oceanographers (Griffies and Greatbatch, 2012). The global means are claimed by some to have accuracies approaching a few tenths of a millimeter per year—an historically extraordinary requirement on any ocean estimate. Despite widely publicized claims to the contrary (e.g., Cazenave and Remy, 2011; Church et al., 2011), state estimate results suggest that at the present time, the global observing system appears to be insufficient to provide robust partitioning amongst heat content changes, land and ice sheet runoff, and large-scale shifts in circulation patterns. A particular difficulty pertains to the deep ocean, below depths measured by the Argo array, where the distinction between apparent changes occuring (Kouketsu, 2011) and the significant deep eddy variability (Ponte, 2012) remains obscure due to poor observational coverage. Claims for closed budget elements involve accuracies much coarser than are stated for the total value. [7]

Biogeochemical Balances

[7] We have omitted here the distinction between absolute sea level with respect to the geoid, and relative sea level measured by tide gauges, and ignored processes associated with the unloading of the solid Earth from ice sheet shrinkage. None of these is represented in current ocean or climate models (e.g., Munk, 2002; Milne et al., 2009).

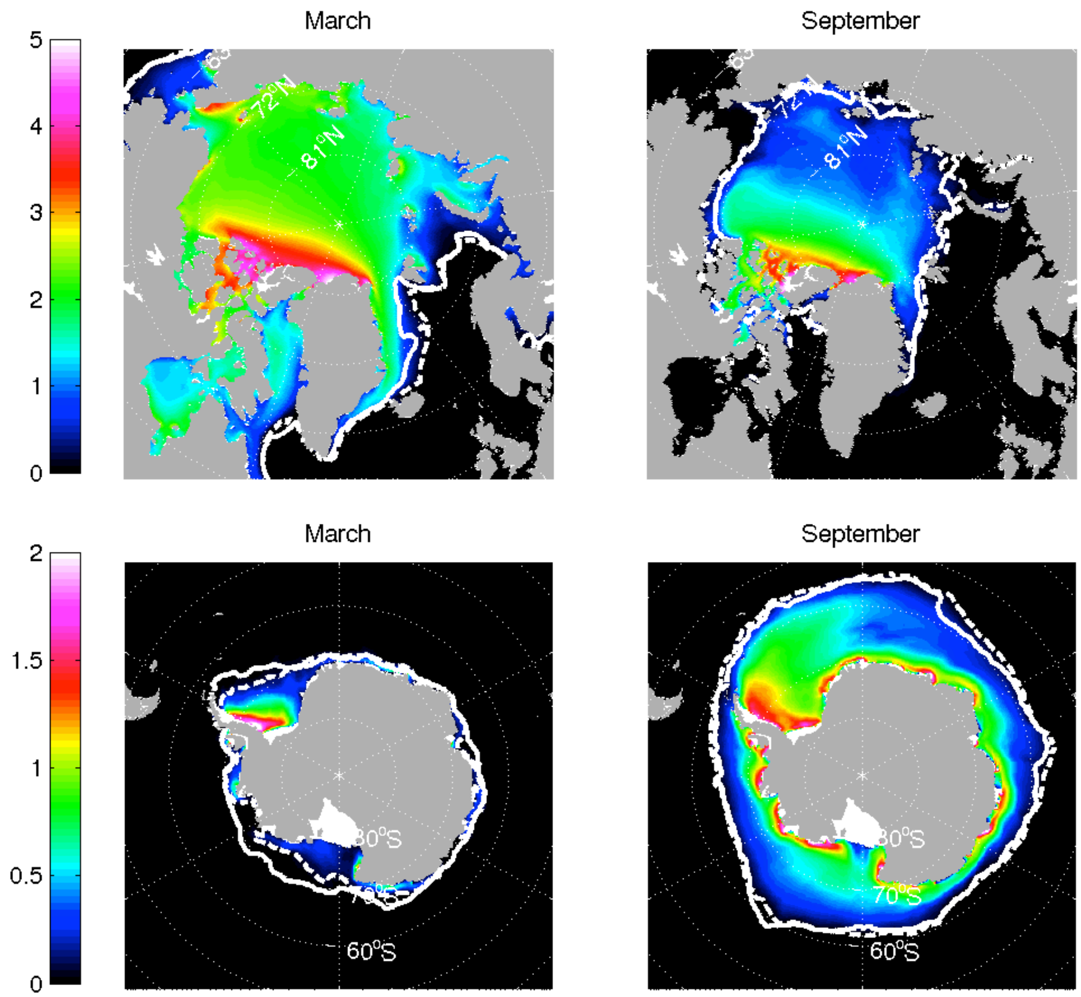


Figure 9: 1992-2002 mean March (left) and September (right) effective ice thickness distributions (in meters) for northern (top) and Southern (bottom) hemispheres. Obtained from a global eddy-permitting ECCO2 simulation, for which a set of global parameters has been adjusted. Also indicated are the ice edge (15% ice concentration isoline) inferred from the model (dashed line) and from satellite-retrieved passive microwave radiometry (solid line). From Losch et al. (2010). {cube76marsepi

From the adjoint of the tracer concentration sub-model of the ECCO system Dutkiewicz et al. (2006) calculated the sensitivity of the nutrient production in the system to iron enrichment. This work is representative of the use of dual solutions to probe large complex models in any field. They found a strong dependence upon the available light, and that the tropical ocean had the greatest sensitivity to iron limitation. Among other considerations, these inferences are important in the erstwhile debate over whether iron fertilization makes any sense for control of atmospheric CO_2.

Woloszyn et al. (2011) used the ECCO higher-resolution Southern Ocean State Estimates

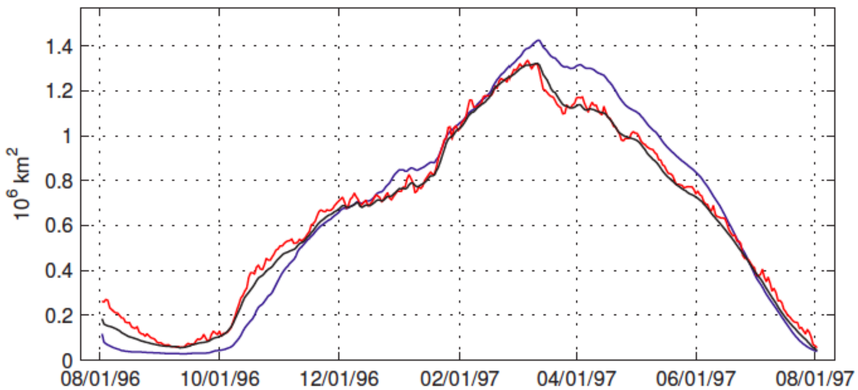
(SOSE) of Mazloff et al. (2010) to demonstrate the great importance of adequate resolution in calculating carbon exchange between the atmosphere and ocean. The same configuration was adopted by Ito et al. (2010) to describe the Ekman layer contribution to the movement of carbon dioxide.

The emerging field of microbial oceanography seeks a zeroth-order understanding of the biogeography and diversity of marine microbes. Coupling between ocean physics and ecology is being explored through the use of ECCO state estimates which drive models of marine ecology (e.g., Follows et al., 2007; Follows and Dutkiewicz, 2011). Crucial requirements of the estimates are (1) to be in sufficiently close agreement with the observed physical ocean state such as to reduce uncertainties in the coupled models from the physical component, and (2) to furnish an evolution of the physical state in agreement with conservation laws.

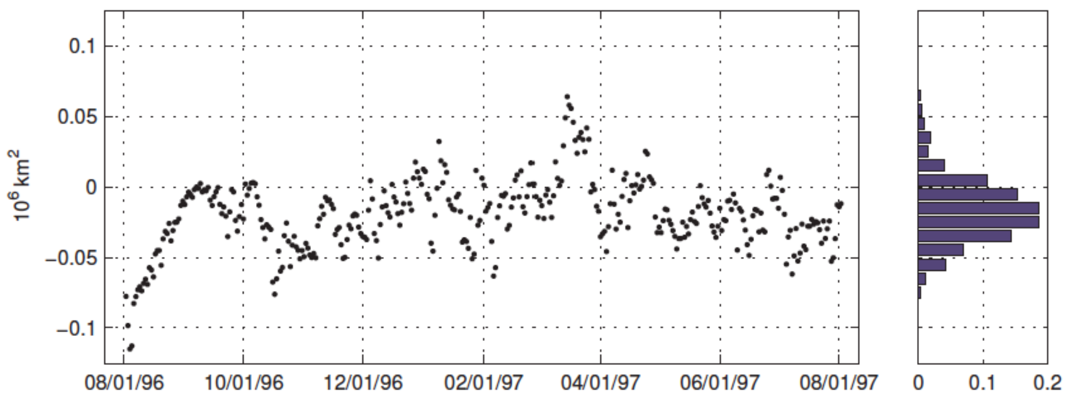
Sea Ice

The importance of sea ice to both the ocean circulation and climate more generally has become much more conspicuous in recent years. Sea ice models have been developed within the state estimation framework as fully coupled sub-systems influenced by and influencing the ocean circulation (Menemenlis et al., 2005; Losch et al., 2010). By way of example, Fig. 11, taken from Losch et al. (2010) depicts 1992-2002 mean March and September effective ice thickness distributions representing the months of maximum and minimum ice cover in both hemispheres. Also shown are the modeled and observed ice edge, represented as 15% isolines of the fractional sea ice concentrations (0 to 100%). The results were obtained from an early version of the ECCO2 eddy-permitting alternative optimization method using Green functions (Menemenlis et al., 2005a,b) on the cubed-sphere grid at 18 km horizontal resolution (see Table 3). More detailed studies focussing on the Arctic were carried out with similar and higher-resolution (4 km) configurations (Nguyen et al., 2011, 2012), but with a very limited control space available for parameter adjustment via the Green function approach.

A comprehensive step toward full coupled ocean-sea ice estimation, in which both ocean and sea ice observation were synthesized, was made by Fenty and Heimbach (2012a,b) for the limited region of the Labrador Sea and Baffin Bay. Fig. 10a shows an annual cycle of total sea ice area in the domain from observations, the state estimate, and the unconstrained model solution. Also shown are the remaining misfits, as evidence of the random nature of the residuals, as required by theory, Eqns. (2) and (3b). An important result of that study is the demonstration that adjustment well within their prior uncertainties in the high-dimensional space of uncertain surface atmospheric forcing, patterns can achieve an acceptable fit between model and observation, placing stringent requirements on process studies which aim to discriminate between model errors and forcing deficiencies.



(a) Total sea ice area



(b) Total sea ice area residuals

Figure 10: (Top): Annual cycle from August 1996 to July 1997 of daily-mean total sea ice area in the Labrador Sea and Baffin Bay from observations (red), a regional state estimate (black), and the unconstrained model solution (blue). (Bottom): Residual misfits between estimated and observed sea ice area and its frequency of occurrence histogram (right panel). Taken from Fenty et al. (2012a). {fenty-labsea-

As in the discussion of biogeochemical balances above, the adjoint or dual solution of the coupled ocean-sea ice model can provide detailed sensitivity analyses. Heimbach et al. (2010) used the dual solution to study sensitivities of sea ice export through the Canadian Arctic Archipelago to changes in atmospheric forcing patterns in the domain. Kauker et al. (2009) investigated the causes of the 2007 September minimum in Arctic sea ice cover in terms of sensitivities to atmospheric forcing over the preceding months. A similar sensitivity study on longer time scales is shown in Fig. 11 of solid (sea ice and snow) freshwater export through Fram Strait for two study periods, January 1989 to September 1993, and January 2003 to September 2007 (unpublished work). The objective function was chosen to be the annual sea ice export between October 1992 and September 1993, and October 2006 and September 2007. Export

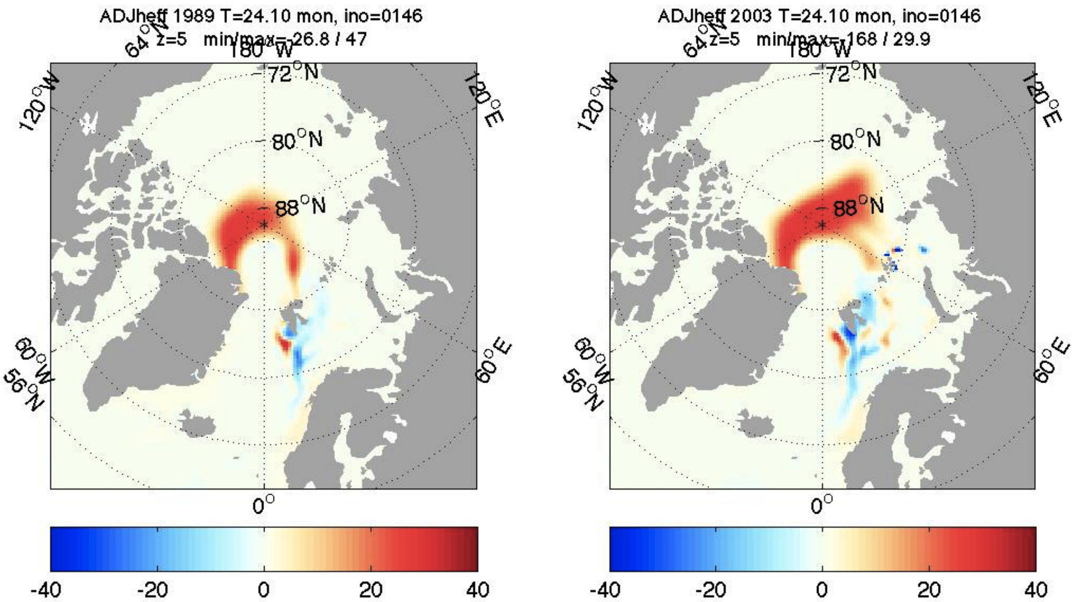


Figure 11: Sensitivity of sea ice export through Fram Strait to changes in effective sea ice thickness 24 months back in time. Two integration periods were considered, Jan. 1989 to Sep. 1993 (left) and Jan. 2003 to Sep. 2007 (right). The objective function is annual sea ice export between October 1992 and September 1993 (left), and October 2006 and September 2007 (right). {fram_export_2

sensitivities to changes in effective sea ice thickness, 24 months prior to September 1993 and 2007, respectively, are shown. The dominant patterns are positive sensitivities upstream of Fram Strait, and for which an increase in ice thickness will increase ice export at Fram Strait 24 months later. (Spurious patterns south of Svalbard have been attributed to masking errors in the sea ice adjoint model and were corrected in Fenty and Heimbach, 2012a.) Sensitivities are linearized around their respective states, and depend on the state trajectory. The extended domain of influence for the 2007 case compared to 1993 suggests more swift transport conditions in the central Arctic, possibly due to favorable atmospheric conditions, or to weaker sea ice, or both.

Ice Sheet-Ocean Interactions

The intense interest in sea level change and the observed acceleration of outlet glaciers spilling into narrow deep fjords in Greenland and ice streams feeding vast ice shelves in Antarctica (e.g., Payne et al., 2004; Alley et al., 2005; Shepherd and Wingham, 2007; Pritchard et al., 2009; Rignot et al., 2011) has led to inferences that much of the ice response may be due to regional oceanic warming at the glacial grounding lines, an area termed by Munk (2011) *"this last piece of unknown ocean"*. One such region is the Amundsen Sea Embayment in West Antarctica (Fig. 12, taken from Schodlok et al., 2012), where the ocean is in contact with

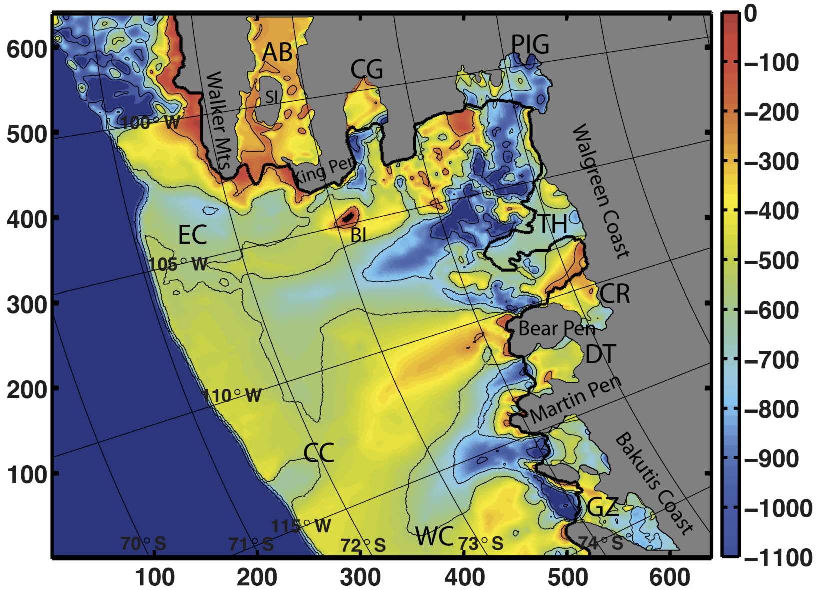
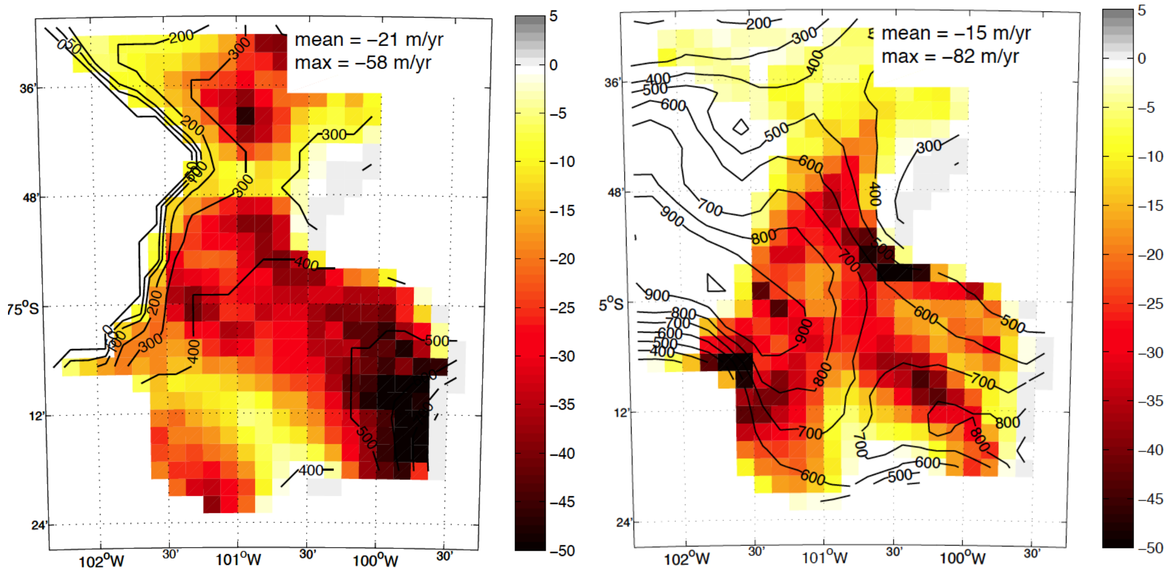


Figure 12: (From Schodlok et al., 2012): bottom topography (in meters) of the Amundsen Sea Embayment, West Antarctica, with thick black lines delineating the edge of several large ice shelves which buttress the following glaciers grounded deep below sea level: Abbot (AB), Cosgrove (CG), Pine Island Glacier (PIG), Thwaites (TH), Crosson (CR), Dotson (DT), and Getz (GZ). Also indicated are prominent topographic features, such as Sherman Island (SI), Burke Island (BI), Eastern Channel (EC), Central Channel (CC), and Western Channel (WC).

{pig_schodlok_

several large shelves, among which Pine Island Ice Shelf (PIIS) and Glacier (PIG) exhibits one of the largest changes in terms of ice sheet acceleration, thinning, and mass loss. Recent, and as yet incomplete model developments have been directed at determining the interactions of changing ocean temperatures and ice sheet response, and for the purpose of inclusion into the coupled state estimation system (Losch, 2008). Simulated melt rates under PIIS are depicted in Fig. 13, but cannot be easily measured directly. A first step toward their estimation in terms of measured hydrography has been undertaken by Heimbach and Losch (2012) who developed an adjoint model complementing the sub-ice shelf melt rate parameterization. By way of example, Fig. 14 depicts transient sensitivities of integrated melt rates (Fig. 13) to changes in ocean temperatures. The spatial inhomogeneous pattens have implications for the interpretation of isolated measurements and optimal observing design.

The critical dependence of sub-ice shelf cavity circulation and melt rates to details of the bathymetry and grounding line position noted by Schodlok et al. (2012) revives the issue of bottom topography as a dominant control on ocean circulation and the necessity for its inclusion into formal estimation systems (Losch and Wunsch, 2003; Losch and Heimbach, 2007).



(a) velocity-independent transfer coefficient

(b) velocity-dependent transfer coefficient

Figure 13: Simulated melt rates (colors, in meters/year) under Pine Island Ice Shelf (PIIS) derived from variants of the Holland and Jenkins (1999) melt rate parameterization, using either velocity-independent (a) or velocity-dependent transfer coefficients (from Dansereau, 2012). Large melt rates correspond to either locations deep inside the cavity where the ice shelf is in contact with the warmest Circumpolar Deep Waters, or to locations of highest flow at the ice shelf-ocean interface. Direct measurement of melt rates is challenging, making robust inferences difficult. {fig:pig-forwa

Air-Sea Transfers and Property Budgets

By definition, state estimates permit calculations up to numerical accuracies of global budgets of energy, enthalpy, etc. Many of these budgets are of interest for the insight they provide into the forces powering the ocean circulation. Josey (2012, this volume) discusses estimates of the air-sea property transfers using the ECCO estimates. As an example, Fig. 15 is an estimate by Stammer et al. (2004) of the net air-sea transfers of fresh water. That paper compares this estimate to other more ad hoc calculations and evaluates its relative accuracy.

As examples of more specific studies using the state estimates, we note only Piecuch and Ponte (2011, 2012) who examined the role of transport fluctuations on the regional sea level and oceanic heat content distribution, and Roquet et al. (2011) who used them to depict the regions in which mechanical forcing by the atmosphere enters into the interior geostrophic circulation. Many more such studies are expected in the future.

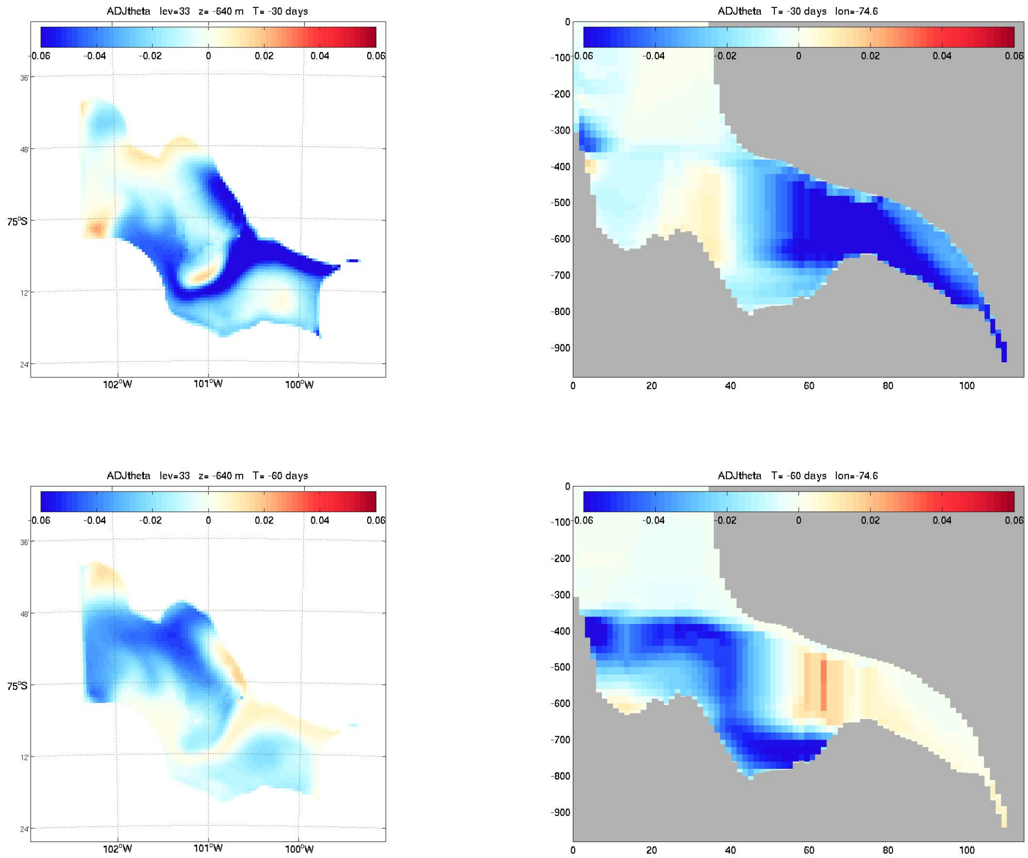


Figure 14: Transient sensitivities, $\delta^* T = (\partial J/\partial T)^T$, of integrated melt rates J under PIIS (from Fig. 11b) to changes in temperature T at times $t = \tau_f - 30$ days (upper row) and -60 days (lower row) prior to computing J. Left panels are horizontal slices at 640 m depth, right panels are vertical slices taken along the dashed line depicted in Fig. 13. Units are in $\mathrm{m^3\,s^{-1}\,K^{-1}}$, where $0.1\ \mathrm{m^3\,s^{-1}\,K^{-1}} \approx 3\ \mathrm{Mt\,a^{-1}\,K^{-1}} \approx 3\ \mathrm{mm\,a^{-1}\,K^{-1}}$.

{fig:pig-adjoi

5.2 Longer Duration Estimates

Although the original ECCO estimates were confined to the period beginning in the early 1990s with the improved observational coverage that became available in association with WOCE, the intense interest in decadal scale climate change has led to some estimates of the ocean state emulating the meteorological reanalyses, extending 50 years and longer into the past. Some of these estimates are based essentially on the reanalysis methods already described (e.g., Rosati et al., 1995; Hurlburt et al., 2009), and having all of their known limitations.

Köhl and Stammer (2008), Wang et al. (2010) have pioneered the application of the ECCO least-squares methods to an oceanic state estimate extending back to 1960. Their estimates have

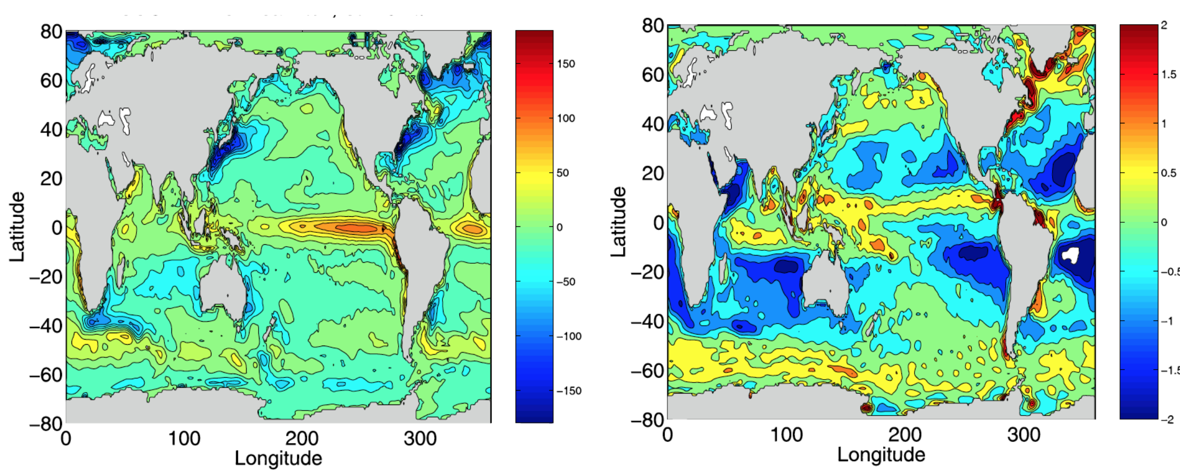


Figure 15: From Stammer et al. (2004) showing an estimate of the multi-year average heat (left, in W/m^2) and fresh water (right, in m/y) transfers between ocean and atmosphere. {stammer_josey

the same virtue as the wider ECCO family of solutions, in satisfying known model equations of motion and dynamics and with known misfits to all data types. The major problem is the extreme paucity of data in the ocean preceding the WOCE era; see e.g., Figs. 1 and 2 of Forget and Wunsch (2007), and the accompanying very limited meteorological forcing observations in the early days. Note that polar orbiting meteorological satellites do not exist prior to 1979—see Fig. 2 and Bromwich and Fogt (2004). Useful altimetry appears only at the end of 1992. "Whole domain" methods such as smoothers or Lagrange multipliers do carry information backwards in time, and in the estimates for the underconstrained decades prior to about 1992, the gross properties of the ocean circulation are better determined because of the later, denser, data sets. But the memory of the upper ocean, which is most prominent e.g., in climate forecasting attempts, appears to be restricted to a few years, and one expects considerable near-surface uncertainty to occur even as recently as the 1980s.

A preliminary step of assessing the impact of observational assets in constraining the ECCO solutions has been taken through observing system experiments in the context of short-duration optimizations during the Argo array period (Heimbach et al., 2009; Zhang et al., 2010). Results suggest that the impact of altimetry and Argo floats in constraining, e.g. the MOC is drastic, compared to the pre-WOCE period when only hydrographic sections were available.

The published solutions for the interval prior to about 1992 are best regarded as physically possible, but whose uncertainty estimates, were they known, would surely be very much greater than they are in the later times, but diminishing as the WOCE-era is approached. These long-duration estimates, decades into the past, thus present a paradox: if they are quantitatively useful—other than as examples of *possible* solutions—then the relatively large investment in observation systems the community has made since the early 1990s was unnecessary. If that investment has been necessary, then one cannot readily quantitatively interpret the early estimates. We leave the subject here as one awaiting the necessary time-dependent uncertainty estimates.

5.3 Short-Duration Estimates

Finding a least-squares fit over 19+ years is computationally very demanding and for some purposes, estimates over shorter time intervals can be useful. In particular, Forget (2010) used the same model and methodology as that of the ECCO Climate State 1° system (Wunsch and Heimbach, 2007), but limited the calculation of three overlapping 18-months periods in the years 2004-2006. In his estimate, the model-data misfit is considerably reduced compared to that in the 16+year solution. The reasons for that better fit are easy to understand from the underlying least-squares methodology: The number of adjustable parameters (the control vector) has the same number of degrees-of-freedom in the initial condition elements as does the decade+ calculations, but with many fewer data to fit, and with little time to evolve away from the opening state. (Meteorological elements change over the same time scales in both calculations.) It is much more demanding of a model and its initial condition controls to produce fits to a 16-year evolution than to an 18-month one. Although both calculations have time-scales short compared to oceanic equilibrium times of hundreds to thousands of years, in an 18 month interval little coupling exists between the meteorological controls and the deep data sets—which are then easily fit by the estimated initial state.

Solutions of this type are very useful, particularly for upper ocean and regional oceanographic estimates (see the water mass formation rate application in Maze et al., 2009). An important caveat, however, is that one must resist the temptation to regard them as climatologies. They *do* bring us much closer to the ancient oceanographic goal of obtaining a synoptic "snapshot."

5.4 Global High Resolution Solutions

Ocean modelers have been pursuing ever-higher resolution from the very beginning of ocean modelling and the effort continues. In classical computational fluid dynamics, one sought "numerical convergence": the demonstration that further improvements in resolution did not qualitatively change the solutions, and preferably that they reproduced known analytical values. Such demonstrations with GCMs are almost non-existent, and thus a very large literature has emerged attempting to demonstrate the utility of "parameterizations"—constructs intended to mimic the behavior of motions smaller than the resolution capability of any particular model. A recent review is by Ringler and Gent (2011). Absent fully-resolved solutions with which to compare the newer parameterizations, the question of their quantitative utility remains open. They do represent clear improvements on older schemes.

Despite the parameterization efforts, considerable evidence exists (e.g., Hecht and Smith, 2008; Lévy et al., 2010) that qualitative changes take place in GCM solutions when the first baroclinic deformation radius, at least, is fully resolved. From the state estimation point of view, one seeks as much skill as possible in the model—which is meant to represent the fullest possible statement of physical understanding. On the other hand, state estimation, as a curve-fitting procedure, is relatively immune to many of the problems of prediction. In particular, because of the dominant geostrophic balance, its mass transport properties are insensitive to unresolved spatial scales—bottom topographic interference being an exception. In data dense regions, away from boundary currents, one anticipates robust results even at modest resolution.

Ultimately, however, the boundary current regions particularly must be resolved (no parameterizations exist for unresolved boundary currents) so as to accurately compute transport properties for quantities such as heat or carbon that depend upon the rendering of the second moments, $\langle C\mathbf{v}\rangle$, where C is any scalar property, and $\mathbf{v}$ is the velocity. Thus a major effort has been devoted to producing global or near-global state estimates from higher resolution models (Menemenlis et al., 2005a, b). The same methodologies used at coarser resolution are also appropriate at high resolution—as has been demonstrated in the regional estimates taken up next, but the computational load rapidly escalates with the state and control vector dimensions. Thus available globally constrained models have used reduced data sets, and have been calculated only over comparatively short time intervals (see Table 3).

Because of the short-duration, much of the interest in these high resolution models lies with the behavior of the eddy field rather than in the large-scale circulation (e.g., Wortham, 2012). As with ordinary forward modelling, how best to adjust the eddy flux parameterizations when parts of the eddy field have been resolved, is a major unknown.

5.5 Regional Solutions

Because the computational load of high resolution global models is so great, efforts have been made to produce regional estimates, typically embedded in a coarser resolution global system. Embedding, with appropriate open boundary conditions is essential, because so much of the ocean state in any finite region is directly dependent upon the boundary values. Implementing open boundary conditions is technically challenging, particularly where the velocity field is directly involved—with slight barotropic imbalances producing large volume imbalances (Ayoub, 2006).

Gebbie et al. (2006) discussed estimates in a small region of the North Atlantic, and their results were used to calculate (Gebbie, 2007) the eddy contribution to near-surface subduction processes. In a much-larger region, the Mazloff et al. (2010) Southern Ocean State Estimate (SOSE), was computed initially over the restricted time interval 2005-2006 (now being extended) at $1/6°$ horizontal and 42 vertical-level resolution.

6 The Uncertainty Problem

From the earliest days of least-squares as used by Gauss and Lagrange, it was recognized that an important advantage of the methods is their ability to produce uncertainty estimates for the solutions, generally as covariances about the expected solution or the underlying true solution. The art of calculating those errors in historically large systems (especially in geodesy and orbit estimation—the fields where the method originated) is highly refined. Unhappily, large as those systems are, their dimensions pale in comparison to the state and control vector sizes encountered in the oceanographic problem. This dimensionality issue renders impractical any of the conventional means that are useful at small and medium size. Numerous methods have been proposed, including direct calculation of the coefficients of the normal equations (the matrix $\mathbf{A}$, defining any system of simultaneous equations) and inversion or pseudo inversion, of $\mathbf{A}^T\mathbf{A}$ (the Hessian); the indirect calculation of the lowest eigenvalues and eigenvectors of the inverse Hessian from algorithmic differentiation (AD) tools; to solutions for the probability density through the Fokker-Planck equation; to the generation of ensembles of solutions. Mostly they have been applied to "toy" problems—somewhat similar to designing a bridge to span the Strait of Gibraltar, and then pointing at a local highway bridge as a demonstration of its practicality. Serious efforts, more generally, to calculate the uncertainties of any large model solution are continuing, but when a useful outcome will emerge is unknown at this time.

In the interim, we generally have only so-called standard errors, representing the temporal variances about the mean of the estimate (Figs. 4, 5). These are useful and helpful. Sensitivities, derived from the adjoint solutions (e.g., Heimbach et al., 2011; and see Figs. 11, 16), are computationally feasible and need to be more widely used. In the meantime, the quest of ocean and climate modelers and for the state estimation community more specifically, for useful understanding of reliability, remains a central, essential, goal. One should note that conventional ocean GCMs or coupled climate models, run without state estimation are almost never accompanied by uncertainty estimates—a serious lack—particularly in an era in which "prediction skill" is being claimed.

Some authors compare their solutions to those inferred from more conventional methods e.g., transport calculations from box inversions of hydrographic sections. These comparisons are

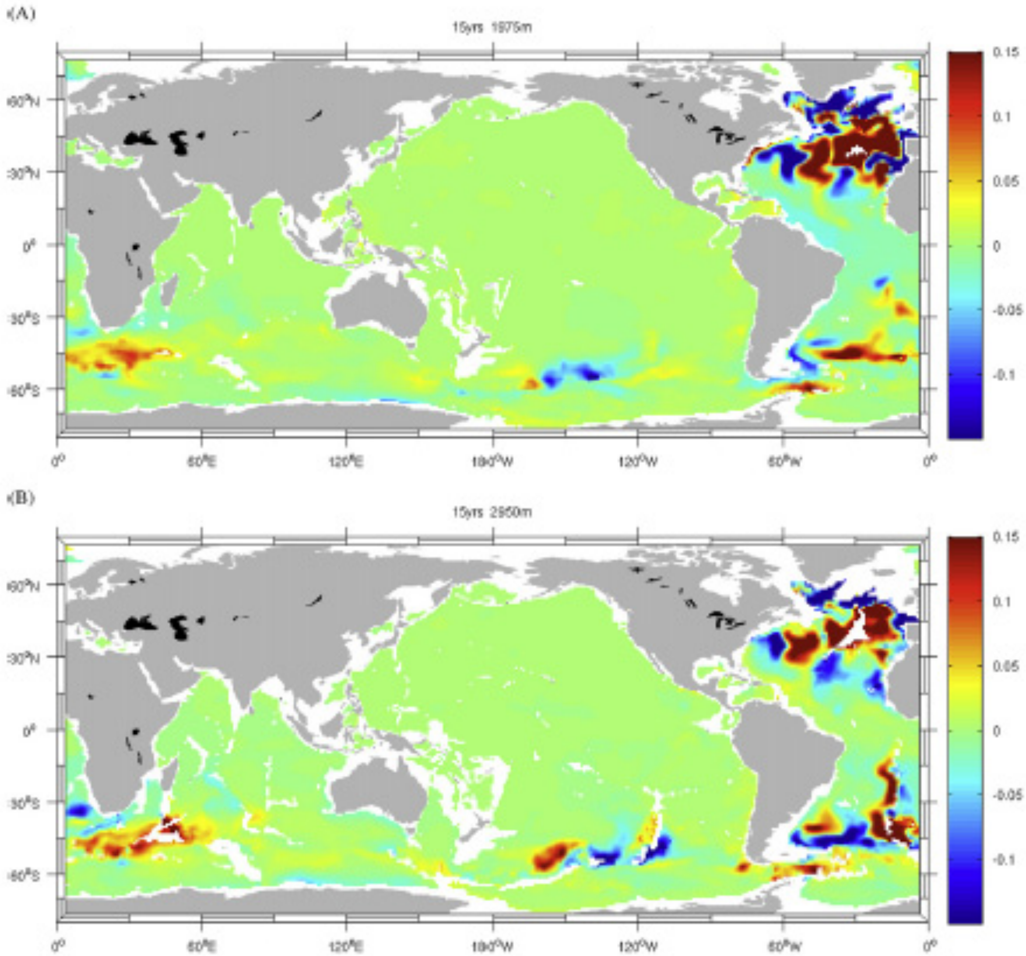


Figure 16: Sensitivities (from Heimbach et al., 2011) of the meridional heat transport across 26°N in the North Atlantic from temperature perturbations at two depths, *15 years earlier.* Top panel is for 1875m, and lower panel is for 2960m.

{heimbach_etal

worthwhile, but it is a major error to treat the hydrographic solutions as if they were true time-averages or climatologies. It is now possible to compare a state estimate from data obtained over a short interval (e.g., March 2003) with a state estimate for that time, sampled in the same way. Differences will appear in the objective function, J. Inevitable discrepancies raise all of the fundamental questions of allocating errors amongst the data, the model, and external controls. In the decadal+ prediction problem (not discussed here), by definition there are no data, and measures of error and skill are far more difficult to obtain. Divergence of IPCC (2007) models over time (e.g., Schmittner et al., 2005; Stroeve et al., 2007), even where fitted to the historical observations, is a strong indicator of the fundamental difficulties involved in extrapolating even systems that appear to give an apparently good fit to historical data, and they are reminiscent of the parable above of fitting cubics to data.

7 Discussion

The history of fluid dynamics generally, and of complex model use in many fields, all support the inference that models unconstrained by data can and do often go wildly wrong (in the wider sense, see e.g., May, 2004; Post and Votta 2005). Readers will recognize the strong point of view taken by the present authors: that models unaccompanied by detailed, direct, comparisons with and constraints by data are best regarded as a kind of science novel.

As we go forward collectively, the need to develop methods describing GCM and state estimate uncertainties is compelling: how else can one combine the quantitative understanding of oceanographic, meteorological and cryospheric physics with the diverse sets of system observations? Such syntheses are the overarching goal of any truly scientific field. Existing state estimates have many known limitations, some of which will be overcome by waiting for the outcome of Moore's Law over the coming years. Other problems, including the perennial and difficult problem of oceanic mixing and dissipation (Munk and Wunsch, 1998; Wunsch and Ferrari, 2004) are unlikely simply to vanish with any forseeable improvement in computer power. Further insight is required.

Lack of long-duration, large-scale, observations generates a fundamental knowledge gap. Without the establishment and maintenance of a comprehensive global ocean observing system, which satisfies the stringent requirements for climate research and monitoring, progress over the coming decades will remain limited (Baker et al., 2007).

Oceanographers now also directly confront the limits of knowledge of atmospheric processes. Until about 20 years ago, meteorological understanding so greatly exceeded that of the ocean circulation that estimated state errors for the atmosphere were of little concern. The situation has changed emphatically with the global observations starting in WOCE, along with the development of oceanic state estimates.[8] These estimation systems are better suited for the purposes of climate research than those developed for numerical weather prediction. What is needed now for climate change purposes, are useful state estimation systems including simultaneously, coupled oceanic, atmospheric, and sea ice physics, and the entirety of the relevant observations in those fields. Thus atmospheric precipitation and evaporation pattern changes can be constrained tightly by changes in the oceanic state. ECCO and related programs have demonstrated how to carry out such recipes. Conventional weather forecast methods are not

[8] The authors have been asked repeatedly at meetings "Why don't oceanographers adopt the sophisticated methods used by meteorologists?" The shoe, however, is now firmly on the other foot.

appropriate, and implementation of a fully coupled state estimation system that will be ongoing is a challenge to governments, university, and research organizations alike. (cf. Bengtsson et al., 2007, who propose a limited step in this direction. Sugiura et al., 2008, and Mochizuki et al., 2009, have made some tentative starts on it.) Surely we must have the capability.

Acknowledgements.

Supported by the US National Ocean Partnership Program with contributions from the National Aeronautics and Space Administration, the National Science Foundation, and the National Oceanographic and Atmospheric Administration. The collaboration of our many ECCO partners is gratefully acknowledged. CW also thanks AOPP and Balliol College, Oxford for support and hospitality through the George Eastman Visiting Professorship. Detailed comments by R. Ponte and the anonymous reviewers were very helpful.

References

Ablain, M., A. Cazenave, G. Valladeau, and S. Guinehut, 2009: A new assessment of the error budget of global mean sea level rate estimated by satellite altimetry over 1993-2008. Ocean Sci, 5, 193-201.

Adcroft, A., C. Hill, J-M. Campin, J. Marshall and P. Heimbach, 2004: Overview of the formulation and numerics of the MIT GCM. in: Proceedings of the ECMWF Seminar on Recent developments in numerical methods for atmospheric and ocean modelling, 6-10 September 2004, Shinfield Park, Reading, UK, pp. 139-150.

Alley, R.B., P.U. Clark, P. Huybrechts, and I. Joughin, 2005: Ice-sheet and sea-level changes. Science, 310, 456-460.

Ayoub, N., 2006: Estimation of boundary values in a North Atlantic circulation model using an adjoint method. Ocean Modelling, 12(3-4), 319-347.

Baehr, J., 2010: Influence of the 26 degrees N RAPID-MOCHA Array and Florida Current Cable Observations on the ECCO-GODAE State Estimate. Journal of Physical Oceanography, 40, 865-879.

Baker, D.J., R.W. Schmitt, and C. Wunsch, 2007: Endowments and new institutions for long-term observations. Oceanography, 20(4), 10-14.

Barker, P.M., J.R. Dunn, C.M. Domingues, S.E. Wijffels, 2011: Pressure sensor drifts in Argo and their impacts. J. Atmos. Oceanic Technol., 28, 1036-1049.

Basili V, C. G., McGarry F, Pajerski R, Page G, Waligora S., 1992: The software engineering laboratory: an operational software experience factory. Proc. Int. Conf. Software Eng., 14th, ACM, 370-381.

Bengtsson, L., S. Hagemann, and K. I. Hodges, 2004: Can climate trends be calculated from reanalysis data? Journal of Geophysical Research-Atmospheres, 109, D11111.

Bengtsson, L., and Coauthors, 2007: The need for a dynamical climate reanalysis. Bulletin of the American Meteorological Society, 88, 495-501.

Björck, A. 1996: Numerical Methods for Least Squares Problems. Society for Industrial Mathematics, Philadelphia, 408pp.

Bromwich, D. H., and R. L. Fogt, 2004: Strong trends in the skill of the ERA-40 and NCEP-NCAR reanalyses in the high and midlatitudes of the southern hemisphere, 1958-2001. Journal of Climate, 17, 4603-4619.

Bromwich, D. H., J. P. Nicolas, and A. J. Monaghan, 2011: An Assessment of Precipitation Changes over Antarctica and the Southern Ocean since 1989 in Contemporary Global Reanalyses. Journal of Climate, 24, 4189-4209.

Bromwich, D. H., Z. C. Guo, L. S. Bai, and Q. S. Chen, 2004: Modeled antarctic precipitation. Part I: Spatial and temporal variability. Journal of Climate, 17, 427-447.

Bromwich, D. H., R. L. Fogt, K. I. Hodges, and J. E. Walsh, 2007: A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. Journal of Geophysical Research-Atmospheres, 112, D10.

Campin, J.-M., A. Adcroft, C. Hill, and J. Marshall, 2004: Conservation of properties in a free surface model. Ocean Modelling, 6, 221-244.

Carton, J. A., and A. Santorelli, 2008: Global Decadal Upper-Ocean Heat Content as Viewed in Nine Analyses. Journal of Climate, 21, 6015-6035.

Carton, J. A., G. Chepurin, X. H. Cao, and B. Giese, 2000: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950-95. Part I: Methodology. Journal of Physical Oceanography, 30, 294-309.

Cazenave, A. and F. Remy, 2011: Sea level and climate: measurements and causes of changes. Wiley Interdisciplinary Reviews-Climate Change, 2, 647-662.

Church, J. A., N. J. White, L. F. Konikow, C. M. Domingues, J. G. Cogley, E. Rignot, J. M. Gregory, M. R. van den Broeke, A. J. Monaghan, and I. Velicogna, 2011: Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. Geophys. Res. Lett., 38, L18601

Cullather, R.I. and M. Bosilovich, 2012: The energy budget of the polar atmosphere in MERRA. J. Clim., 25, 5-24.

Dansereau, V., 2012: Ice shelf-ocean interactions in a general circulation model: melt-rate modulation due to mean flow and tidal currents. Master's Thesis, MIT/WHOI Joint Progam, Cambridge, MA.

Dutkiewicz, S., M. Follows, P. Heimbach, and J. Marshall, 2006: Controls on ocean productivity and air-sea carbon flux: An adjoint model sensitivity study. Geophys. Res. Letts., 33, doi:10.1029/2005GL024987.

Elliott, W. P., and D. J. Gaffen, 1991: On the utility of radiosonde humidty archives for climate studies. Bulletin of the American Meteorological Society, 72, 1507-1520.

Evensen, G., 2009: Data Assimilation: The Ensemble Kalman Filter. Springer Verlag.

Fenty, I. and P. Heimbach, 2012a: Coupled sea ice-ocean state estimation in the Labrador Sea and Baffin Bay. J. Phys. Oceanogr., in revision.

Fenty, I.G. and P. Heimbach, 2012b: Hydrographic Preconditioning for Seasonal Sea Ice Anomalies in the Labrador Sea. J. Phys. Oceanogr., in revision.

Ferreira, D., J. Marshall, and P. Heimbach, 2005: Estimating eddy stresses by fitting dynamics to observations using a residual-mean ocean circulation model and its adjoint. J. Phys. Oceanogr., 35, 1891-1910.

Follows, M.J. and S. Dutkiewicz, 2011: Modeling diverse communities of marine microbes. Ann. Rev. Marine Sci., 3, 427-451.

Follows, M.J., S. Dutkiewicz, S. Grant and S.W. Chisholm, 2007: Emergent biogeography of microbial communities in a model ocean. Science, 315, 1843-1846.

Forget, G., 2010: Mapping ocean observations in a dynamical framework: A 2004-06 ocean atlas. Journal of Physical Oceanography, 40,1201-1221.

Forget, G., and C. Wunsch, 2007: Estimated global hydrographic variability. Journal of Physical Oceanography, 37, 1997-2008.

Ganachaud, A., 2003a: Error budget of inverse box models: The North Atlantic. Journal of Atmospheric and Oceanic Technology, 20, 1641-1655.

Ganachaud, A., 2003b: Large-scale mass transports, water mass formation, and diffusivities estimated from World Ocean Circulation Experiment (WOCE) hydrographic data. Journal of Geophysical Research, 108, 3213.

Ganachaud, A., and C. Wunsch, 2002: Large-scale ocean heat and freshwater transports during the World Ocean Circulation Experiment. J. Clim., 16, 696-705.

Gebbie, G., 2007: Does eddy subduction matter in the Northeast Atlantic Ocean? J. Geophys. Res., 112, C06007.

Gebbie, G., P. Heimbach, and C. Wunsch, 2006: Strategies for nested and eddy-permitting state estimation. Journal of Geophysical Research-Oceans, 111, C10073.

Gelb, A. (Ed.), 1974: Applied Optimal Estimation. 382 pp., MIT Press, Cambridge, MA.

Giering, R. and T. Kaminski, 1998: Recipes for adjoint code construction. ACM Transactions on Mathematical Software, 24, 437-474.

Griewank, A. and A. Walther, 2008: Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation. 442 pp., SIAM, Philadelphia.

Griffies, S. M. and Greatbatch, R. J., 2012: Physical processes that impact the evolution of global mean sea level in ocean climate models. Ocean Modelling, 51(C), 3772. doi:10.1016/j.ocemod.2012.04.003.

Hecht, M. W., R. D. Smith, 2008: Towards a physical understanding of the North Atlantic: a review of model studies. Ocean Modeling in an Eddying Regime, AGU Geophysical Monograph 177, M. W. Hecht and H. Hasumi, Eds... 213-240.

Heimbach, P., C. Wunsch, R. M. Ponte, G. Forget, C. Hill, and J. Utke, 2011: Timescales and regions of the sensitivity of Atlantic meridional volume and heat transport magnitudes: Toward observing system design.. Deep-Sea Research-II, 58, 1858-1879.

Heimbach, P., C. Hill, and R. Giering, 2005: An efficient exact adjoint of the parallel MIT General Circulation Model, generated via automatic differentiation. Future Generation Computer Systems, 21, 1356-1371.

Heimbach, P. and M. Losch, 2012: Adjoint sensitivities of sub-ice shelf melt rates to ocean circulation under Pine Island Ice Shelf, West Antarctica. Ann. Glaciol., 53(60), 5969, doi:10.3189/2012/AoG60A025.

Heimbach, P., G. Forget, R. Ponte, and C. Wunsch (lead authors), 2009: Observational Requirements for global-scale ocean climate analysis: Lessons from ocean state estimation. Community White Paper. In: Hall, J., D.E. Harrison, and D. Stammer (Eds.), 2010: Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society, Venice, Italy, 21-25 September 2009, ESA Publication WPP-306, Vol. 2, doi:10.5270/OceanObs09.cwp.42.

Heimbach, P., D. Menemenlis, M. Losch, J.M. Campin, and C. Hill, 2010: On the formulation of sea-ice models. Part 2: Lessons from multi-year adjoint sea ice export sensitivities through the Canadian Arctic Archipelago. Ocean Modelling, 33(1-2), 145-158.

Hoteit, I., B. Cornuelle, A. Köhl, and D. Stammer, 2006: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. Quarterly Journal of the Royal Meteorological Society, 131 (613), 3659-3682.

Hoteit, I., B. Cornuelle, and P. Heimbach, 2010: An Eddy-Permitting, Dynamically Consistent Adjoint-Based Assimilation System for the Tropical Pacific: Hindcast Experiments in 2000. J. Geophys. Res., 115, C03001.

Huang, R. X., 1993: Real freshwater flux as a natural boundary condition for the salinity balance and thermohaline circulation forced by evaporation and precipitation. J. Phys. Oc., 23, 2428-2446.

Huang, R. X., 2010: Ocean circulation: wind-driven and thermohaline processes. Cambridge University Press, xiii, 791 p. pp.

Hurlburt, H. E., and Coauthors, 2009: High-resolution global and basin-scale ocean analyses and forecasts. Oceanography, 22, 110-127.

IPCC Intergovernmental Panel on Climate Change, 2007: Climate Change 2007 - The Physical Science Basis. Cambridge Un. Press, Cambridge, 1009pp.

Josey, S. A. and others, 2012: Unknown Title. This Volume.

Ito, T., M. Woloszyn, and M. Mazloff, 2010: Anthropogenic carbon dioxide transport in the Southern Ocean driven by Ekman flow. Nature, 463, 80-83.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. J. Basic Eng., 82, 35-45.

Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation, and Predictability. Cambridge University Press, xxii, 341 p., 344 p.

Kanzow, T., and Coauthors, 2009: Basinwide Integrated Volume Transports in an Eddy-Filled Ocean. Journal of Physical Oceanography, 39, 3091-3110.

Kauker, F. and 5 others, 2009: Adjoint analysis of the 2007 all time Arctic sea-ice minimum.

Geophys. Res. Lett., 36, L03707.

Köhl, A., and D. Stammer, 2008: Decadal Sea Level Changes in the 50-Year GECCO Ocean Synthesis. J. Clim., 21, 1876-1890.

Köhl, A. and J. Willebrand, 2002: An adjoint method for the assimilation of statistical characteristics into eddy-resolving models. Tellus, 54A, 406-425.

Kouketsu, S., et al., 2011: Deep ocean heat content changes estimated from observation and reanalysis product and their influence on sea level change, J. Geophys. Res., 116, C03012, doi:10.1029/2010JC006464.

Kuhlmann, J., H. Dobslaw, and M. Thomas, 2011: Improved modeling of sea level patterns by incorporating self-attraction and loading. J. Geophys. Res., 116, C11036.

Lea, D. J., T. W. N. Haine, M. R. Allen, and J. A. Hansen, 2002: Sensitivity analysis of the climate of a chaotic ocean circulation model. Quarterly Journal of the Royal Meteorological Society, 128, 2587-2605.

Lee, T., and Coauthors, 2010: Consistency and fidelity of Indonesian-throughflow total volume transport estimated by 14 ocean data assimilation products. Dynamics of Atmospheres and Oceans, 50, 201-223.

Lévy, M., P. Klein, A. M. Treguier, D. Iovino, G. Madec, S. Masson, and K. Takahashi, 2010: Modifications of gyre circulation by sub-mesoscale physics. Ocean Modelling, 34, 1-15.

Losch, M. and P. Heimbach, 2007: Adjoint sensitivity of an ocean general circulation model to bottom topography. J. Phys. Oceanogr., 37(2), 377-393.

Losch, M. and C. Wunsch, 2003: Bottom topography as a control variable in an ocean model. J. Atmos. Ocean. Technol., 20, 1685-1696.

Losch, 2008: Modeling ice shelf cavities in a z-coordinate ocean general circulation model. J. Geophys. Res., 113, C08043.

Losch, M., D. Menemenlis, J.M. Campin, P. Heimbach, and C. Hill, 2010: On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations. Ocean Modelling, 33(1-2), 129-144.

Lumpkin, R., and K. Speer, 2007: Global ocean meridional overturning. Journal of Physical Oceanography, 37, 2550-2562.

Macdonald, A. M., 1998: The global ocean circulation: a hydrographic estimate and regional analysis. Progress in Oceanography, 41, 281-382.

Macdonald, A. M., and Coauthors, 2009: The WOCE-era 3-D Pacific Ocean circulation and heat budget. Progress in Oceanography, 82, 281-325.

Marshall, G. J., V. Lagun, and T. A. Lachlan-Cope, 2002: Changes in Antarctic Peninsula tropospheric temperatures from 1956 to 1999: A synthesis of observations and reanalysis data.

Int. J. Climatol., 22, 291-310.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. Journal of Geophysical Research-Oceans, 102, 5753-5766.

Martin, A. J., A. Hines, and M. J. Bell, 2007: Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact. Quarterly Journal of the Royal Meteorological Society, 133, 981-995.

May, R., 2004: Uses and abuses of mathematics in biology. Science, 303 (5659) 790-793.

Maze, G., G. Forget, M. Buckley, J. Marshall, and I. Cerovecki, 2009: Using transformation and formation maps to study the role of air-sea heat fluxes in North Atlantic Eighteen Degree Water formation. Journal of Physical Oceanography, 39, 1818-1835.

Mazloff, M. R., P. Heimbach, and C. Wunsch, 2010: An Eddy-Permitting Southern Ocean State Estimate. Journal of Physical Oceanography, 40, 880-899.

Menemenlis, D., I. Fukumori, and T. Lee, 2005a: Using Green's functions to calibrate an ocean general circulation model. Mon. Wea. Rev., 133, 1224-1240.

Menemenlis, D., and Coauthors, 2005b: NASA supercomputer improves prospects for ocean climate research. EOS Trans. AGU, 86(9), 89.

Milne, G. A., Gehrels, W. R., Hughes, C. W., and Tamisiea, M. E., 2009: Identifying the causes of sea-level change. Nature Geosci., 2(7), 471478. doi:10.1038/ngeo544.

Mochizuki, T., N. Sugiura, T. Awaji, and T. Toyoda, 2009: Seasonal climate modeling over the Indian Ocean by employing a 4D-VAR coupled data assimilation approach. J. Geophys. Res., 114, C11003.

Munk, W., 2002: Twentieth century sea level: An enigma. Proceedings of the National Academy of Sciences, 99(10), 65506555. doi:10.1073/pnas.092704599.

Munk, W., 2011: The sound of climate change. Tellus, 63A, 190-197.

Munk, W. and C. Wunsch, 1998: Abyssal recipes II: energetics of tidal and wind mixing. Deep-Sea Res. I, 45, 1977-2010.

Nicolas, J.P. and D.H. Bromwich, 2011: Precipitation changes in high Southern latitudes from global reanalyses: A cautionary tale. Surv. Geophys, 32(4-5), 475-494.

Payne, A.J., A. Vieli, A.P. Shepherd, D.J. Wingham and E. Rignot, 2004. Recent dramatic thinning of largest West Antarctic ice stream triggered by oceans, Geophys. Res. Lett., 31, L23401.

Pedlosky, J., 1987: Geophysical Fluid Dynamics. 2nd ed. Springer-Verlag, xiv, 710 pp.

Piecuch, C. G., and R. M. Ponte, 2011: Mechanisms of interannual steric sea level variability. Geophys. Res. Lett., 38, L15605.

Piecuch, C. G. and R. M. Ponte, 2012: Importance of circulation changes to Atlantic heat storage rates on seasonal and interannual time scales. J. Climate, 25, 350–362.

Ponte, R. M., 2012: An assessment of deep steric height variability over the global ocean. Geophys. Res. Lett., 39, L04601, doi:10.1029/2011GL050681.

Ponte, R. M., C. Wunsch, and D. Stammer, 2007: Spatial mapping of time-variable errors in TOPEX/POSEIDON and Jason-1 seasurface height mesurements. J. Atm. Oc. Tech,, 24, 1078-1085.

Post, D. E. and L. G. Votta, 2005: Computational science demands a new paradigm,Physics Today, 35-41.

Pritchard, H.D., R.J. Arthen, D.G. Vaughan and L.A. Edwards, 2009: Extensive dynamic thinning on the margins of the Green- land and Antarctic ice sheets, Nature, 461, 971-975.

Rayner, D., and Coauthors, 2011: Monitoring the Atlantic meridional overturning circulation. Deep-Sea Research Part II-Topical Studies in Oceanography, 58, 1744-1753.

Rignot, E., I. Velicogna, M.R. van den Broeke, A. Monaghan and J. Lenaerts, 2011: Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise, Geo- phys. Res. Lett., 38, L05503.

Ringler, T., and P. Gent, 2011: An eddy closure for potential vorticity. Ocean Modelling, 39, 125-134.

Roquet, F., C. Wunsch, G. Madec, 2011: On the patterns of wind-power input to the ocean circulation. J. Phys. Oceanogr., 41, 2328–2342.

Rosati, A., R. Gudgel, and K. Miyakoda, 1995: Decadal analysis produced from an ocean data assimilation system. Monthly Weather Review, 123, 2206-2228.

Schiller A., T. Lee, S. Masuda, 2012: Methods and applications of ocean synthesis in climate research. This volume.

Schmittner, A., M. Latif, and B. Schneider, 2005: Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations. Geophys. Res. Lett., 32, L23710, doi:10.1029/2005GL024368.

Schodlok, M.P., D. Menemenlis, E. Rignot, and M. Studinger, 2012: Sensitivity of the ice shelf ocean system to the sub-ice shelf cavity shape measured by NASA IceBridge in Pine Island Glacier, West Antarctica. Ann. Glaciol., 53(60), 156-162.

Shepherd, A. and D. Wingham, 2007: Recent sea-level contributions of the Antarctic and Greenland ice sheets. Science, 315, 1529-1532.

Siedler, G., Church, J., and W. J. Gould (Eds.), 2001: Ocean Circulation and Climate: Observing and Modeling the Global Ocean. Academic, San Diego, 715pp.

Stammer, D., 2005: Adjusting internal model errors through ocean state estimation. J. Phys.

Oceanogr., 35, 1143-1153.

Stammer, D., and Coauthors, 2001: Transport processes of the global ocean circulation between 1992 and 1997 estimated from global altimeter data, SST fields, daily NCEP surface fluxes, the Levitus climatology and a general circulation model. Fifth Symposium on Integrated Observing Systems, 41-44.

Stammer, D., and Coauthors, 2002: Global ocean circulation during 1992-1997, estimated from ocean observations and a general circulation model. Journal of Geophysical Research-Oceans, 107 (C9), 3118, doi:10.1029/2001JC000888.

Stammer, D., and Coauthors, 2003: Volume, heat, and freshwater transports of the global ocean circulation 1993-2000, estimated from a general circulation model constrained by World Ocean Circulation Experiment (WOCE) data. Journal of Geophysical Research-Oceans, 108 (C1), 3007, doi:10.1029/2001JC001115.

Stammer, D., A. Köhl, and C. Wunsch, 2007: Impact of accurate geoid fields on estimates of the ocean circulation. J. Atmos. Ocean. Tech., 24, 1464-1478.

Stroeve, J., M. M. Holland, W. Meier, T. Scambos, and M. Serreze (2007), Arctic sea ice decline: Faster than forecast, Geophys. Res. Lett., 34, L09501, doi:10.1029/2007GL029703.

Sugiura, N., T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa, 2008: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. J. Geophys. Res., 113 (C10), C10017.

Talagrand, O., 1997: Assimilation of observations, an introduction. J. Meteorol. Soc. Japan, 75, 191-209.

Talley, L. D., D. Stammer, and I. Fukumori, 2001: Towards a WOCE synthesis. International Geophysics Series, 77, 525-546.

Thompson, D. W. J., J. J. Kennedy, J. M. Wallace, and P. D. Jones, 2008: A large discontinuity in the mid-twentieth century in observed global-mean surface temperature. Nature, 453, 646-U645.

Thorne, P.W., 2008: Arctic tropospheric warming amplification?. Nature, 455, E1-E2, doi:10.1038/nature07256.

Trenberth, K. E., and A. Solomon, 1994: The global heat-balance - heat transports in the atmosphere and ocean. Climate Dynamics, 10, 107-134.

Trenberth, K. E., J. W. Hurrell, and A. Solomon, 1995: Conservation of mass in 3-dimensions in global analyses. Journal of Climate, 8, 692-708.

Trenberth, K. E., D. P. Stepaniak, J. W. Hurrell, and M. Fiorino, 2001: Quality of reanalyses in the tropics. Journal of Climate, 14, 1499-1510.

Utke, J., U. Naumann, M. Fagan, N. Thallent, M. Strout, P. Heimbach, C. Hill and C. Wunsch, 2008: OpenAD/F: A modular, open-source tool for automatic differentiation of Fortran codes. ACM Transactions on Mathematical Software (TOMS), 34(4), doi:10.1145/1377596.1377598.

Vallis, G. K., 2006: Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation. Cambridge University Press, xxv, 745 p. pp.

Vinogradov, S. V., and R. M. Ponte, 2010: Annual cycle in coastal sea level from tide gauges and altimetry. Journal of Geophysical Research-Oceans, 115, C02041.

Vinogradov, S. V., R. M. Ponte, P. Heimbach, and C. Wunsch, 2008: The mean seasonal cycle in sea level estimated from a data-constrained general circulation model. Journal of Geophysical Research-Oceans, 113 C03032.

Vinogradova, N. T., R. M. Ponte, M. E. Tamisiea, K. J. Quinn, E. M. Hill, and J. L. Davis, 2011: Self-attraction and loading effects on ocean mass redistribution at monthly and longer time scales, J. Geophys. Res., 116, C08041, doi:10.1029/2011JC007037.

Wang, W. Q., A. Kohl, and D. Stammer, 2010: Estimates of global ocean volume transports during 1960 through 2001. Geophysical Research Letters, 37, L15601.

Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. Quarterly Journal of the Royal Meteorological Society, 127, 1815-1846.

Wijffels, S.E., J. Willis, C.M. Domingues, P. Barker, N.J. White, A. Gronell, K. Ridgway, J.A. Church, 2008: Changing Expendable Bathythermograph Fall Rates and Their Impact on Estimates of Thermosteric Sea Level Rise. J. Climate, 21, 5657-5672.

Woloszyn, M., M. Mazloff, and T. Ito, 2011: Testing an eddy-permitting model of the Southern Ocean carbon cycle against observations. Ocean Modelling, 39, 170-182.

Wortham IV, C. J. L. 2012: A Multi-Dimensional Spectral Description of Ocean Variability with Applications. PhD Thesis, MIT and WHOI, 184 pp.

Wunsch, C., 2006: Discrete Inverse and State Estimation Problems: With Geophysical Fluid Applications. Cambridge University Press, xi, 371 p.

Wunsch, C., 2010: Towards a mid-latitude ocean frequency-wavenumber spectral density and trend determination. J. Phys. Oc., 40, 2264-2281.

Wunsch, C., 2011: The decadal mean ocean circulation and Sverdrup balance. J. Mar. Res., 69, 417-434.

Wunsch, C., and R. Ferrari, 2004: Vertical mixing, energy, and the general circulation of the oceans. Annu. Rev. Fluid Mech., 36(1), 281314. doi:10.1146/annurev.fluid.36.050802.122121.

Wunsch, C., and P. Heimbach, 2006: Estimated decadal changes in the North Atlantic meridional overturning circulation and heat flux 1993-2004. Journal of Physical Oceanography, 36, 2012-2024.

Wunsch, C. and P. Heimbach, 2007: Practical global oceanic state estimation. Physica D-

Nonlinear Phenomena, 230, 197-208.

Wunsch, C. and P. Heimbach, 2009: The global zonally integrated ocean circulation, 1992-2006: Seasonal and decadal variability. Journal of Physical Oceanography, 39, 351-368.

Wunsch, C. and P. Heimbach, 2012: Two decades of the Atlantic Meridional Overturning Circulation: anatomy, variations, extremes, prediction, and overcoming its limitations. J. Clim., submitted.

Wunsch, C., R. M. Ponte, and P. Heimbach, 2007: Decadal trends in sea level patterns: 1993-2004. Journal of Climate, 20, 5889-5911.

Zhang, S., Rosati, A., and Delworth, T., 2010: The adequacy of observing systems in monitoring the Atlantic meridional overturning circulation and North Atlantic climate. Journal of Climate, 23(19), 53115324. doi:10.1175/2010JCLI3677.1.

reanalysis product	net fresh water imbalance [mm/year]		net heat flux imbalance [W/m²]	
	ocean-only	global	ocean-only	global
NCEP/NCAR-I 1992-2010	159	62	-0.7	-2.2
NCEP/DOE-II (1992-2004)	740	-	-10	-
ERA-Interim (1992-2010)	199	53	-8.5	-6.4
JRA-25 (1992-2009)	202	70	15.3	10.1
CORE-II (1992-2007)	143	58		

Table 1: Global net heat and freshwater flux imbalances of atmospheric reanalysis products. {woce_table_im

observation	instrument	product/source	area	period	dT
Mean dynamic topography (MDT)	• GRACE SM004-GRACE3 • EGM2008/DNSC07	CLS/GFZ (A.M. Rio) N. Pavlis/Andersen & Knudsen	global global	time-mean	mean
Sea level anomaly (SLA)	• TOPEX/POSEIDON • Jason • ERS, ENVISAT • GFO	NOAA/RADS & PO.DAAC NOAA/RADS & PO.DAAC NOAA/RADS & PO.DAAC NOAA/RADS & PO.DAAC	65°N/S 82°N/S 65°N/S 65°N/S	1993 – 2005 2001 – 2011 1992 – 2011 2001 - 2008	daily daily daily daily
SST	• blended, AVHRR (O/I) • TRMM/TMI • AMSR-E (MODIS/Aqua)	Reynolds & Smith GHRSST GHRSST	Global 40°N/S Global	1992 - 2011 1998 - 2004 2001 - 2011	monthly daily daily
SSS	Various in-situ	WOA09 surface	Global	climatology	monthly
In-situ T, S	• Argo, P-Alace • XBT • CTD • SEaOS • TOGA/TAO, Pirata	Ifremer D. Behringer (NCEP) various SMRU & BAS (UK) PMEL/NOAA	"global" "gobal" sections SO Tropics	1992 – 2011 1992 – 2011 1992 – 2011 2004 – 2010 1992 – 2011	daily daily daily daily daily
Mooring velocities	• TOGA/TAO, Pirata • Florida Straits	PMEL/NOAA NOAA/AOML	Trop. Pac. N. Atl.	1992 – 2006 1992 – 2011	daily daily
Climatological T,S	• WOA09 • OCCA	WOA09 Forget, 2010	"global" "global"	1950 - 2000 1950 - 2002	mean mean
sea ice cover	• satellite passive microwave radiometry	NSIDC (bootstrap)	Arcitc, SO	1992 - 2011	daily
Wind stress	QuickScat	• NASA (Bourassa) • SCOW (Risien & Chelton)	global	1999 – 2009 climatolggy	daily monthly
Tide gauge SSH	Tide gauges	NBDC/NOAA	sparse	1992 - 2006	monthly
Flux constraints	from ERA-Interim, JRA-25, NCEP, CORE-2 variances	Various	global	1992 - 2011	2-day to 14-day
Balance constraints			global	1992 - 2011	mean
bathymetry		Smith & Sandwell, ETOPO5	global	-	-

Table 2: Data used in the ECCO global 1° resolution state estimates until about 2011. {woce_table_ec

label & version	hor./ver. grid	domain	duration	scope	reference
ECCO-Production Sustained production of decadal climate state estimates (former ECCO-GODAE)					
ver.0 (ECCO-MIT)	2° / 22	80° N/S	1992—1997	first ECCO product – proof of feasibility	Stammer et al. (2002/04)
ver.1 (ECCO-SIO)	1° / 23	80° N/S	1992—2002	begin of 1° sustained production	Köhl et al. (2007)
ver.2 (ECCO-GODAE)	1° / 23	80° N/S	1992—2004	air-sea flux constraints for sea level studies	Wunsch & Heimbach (2006/07)
(OCCA)	1° / 50	80° N/S	2004/5/6/7	Atlas from one-year "synoptic snapshots"	Forget (2010)
(GECCO)	1° / 23	80° N/S	1951—2000	50-year solution covering NCEP/NCAR period	Köhl and Stammer (2008a/b)
ver.3.0 (ECCO-GODAE)	1° / 23	80° N/S	1992—2007	switch to atmos. state controls and sea ice	Wunsch & Heimbach (2009)
revision 1 (ver.3.1)	1° / 23	80° N/S	1992—2010	updates to ver.3.0	Fukumori et al. (in prep.)
ver.4	1° / 50	global	1992—2010	first full-global estimate incl. Arctic	Forget & Heimbach (in prep.)
ECCO-ICES Ocean-Ice Interactions in Earth System Models (former ECCO2)					
ver.1 (CS510 GF)	18 km / 50	global	1992—2002/10	Green's function optim., of eddying model	Menemenlis et al. (2005)
ver.2 (CS510 adjoint)	18 km / 50	global	2004-05/2009-10	adjoint-based global eddying 1-year optim.	Menemenlis et al. (in prep.)
ECCO-JPL near real-time filter & reduced-space smoother					
ver.1 (KF)	1° / 46	80° N/S	1992—present	near-real time Kalman Filter (KF) assimilation	Fukumori et al. (1999)
ver.2 (RTS)	1° / 46	80° N/S	1992—present	smoother update of KF solution	Fukumori (2002)
Regional Efforts (/*/ denote ongoing efforts)					
Southern Ocean (SOSE) /*/	1/6° / 42	25°-80°S	2005—2009	eddy-permitting SO State Estimate	Mazloff et al. (2010)
ECC2 Arctic & ASTE /*/	18 & 4 km / 50	Arctic & SPG	1992—2009	Arctic/subpolar gyre ocean-sea ice estimate	Nguyen et al. (2011/12)
North Atlantic	1° / 23	25°-80°N	1993	experimental 2° vs. 1° nesting	Ayoub (2006)
Subtropical Atialntic	1/6° / 42	…	1992/93	experimental 1° vs. 1/6° nesting	Gebbie et al. (2006)
Tropical Pacific	…	…		experimental 1° vs. 1/3° nesting	Hoteit et al. (2006/2010)
Labrador Sea & Baffin Bay	…	…	1996/97	first full coupled ocean-sea ice estimate	Fenty et al. (2012a/b)

Table 3: Published ECCO family state estimates, divided roughly into categories. The global decade+ estimates are labelled as "ECCO-Production", while others are either regional, or experimental. {woce_table_ec