Combining altimetric and all other data with a general circulation model *

P. Heimbach[1], R.M. Ponte[2], C. Evangelinos[1], G. Forget[1], M. Mazloff[1], D. Menemenlis[3], S. Vinogradov[2] and C. Wunsch[1]

(1) Massachusetts Institute of Technology, EAPS, Cambridge (MA), USA
`{heimbach, ce107, gforget, mmazloff, cwunsch} AT mit DOT edu`
(2) Atmospheric and Environmental Research, Lexington (MA), USA
`{rponte, svinogra} AT aer DOT com`
(3) Jet Propulsion Laboratory, Pasadena (CA), USA
`{menemenlis} AT jpl DOT nasa DOT gov`

May 17, 2006

Abstract

The TOPEX/POSEIDON mission was formulated in conjunction with the World Ocean Circulation Experiment (WOCE). Although altimetric data are by far the largest ocean data set that emerged from that experiment, it was recognized from the outset that the best estimate of the ocean circulation and its variability would be made by combining all the data with a good general circulation model. The US ECCO-GODAE consortium has now produced useful estimates of the three-dimensional time evolving ocean circulation at 1 degree lateral resolution, 23 layers in the vertical, over the time period 1992-2004. At the present time, the solution represents a least-squares misfit to about 410 million separate observational constraints. A large variety of fields can be analyzed in the results, but the focus here will be on the overturning circulation and its changes over this time period. The residual model/data misfits raise difficult questions about remaining errors in the altimetric/geoid fields.

1 Introduction

The era of satellite altimetry has seen a shift in the paradigm of the ocean as a steady, laminar fluid to a turbulent one exhibiting a plethora of processes on various space and time scales with consequences for observational requirements (e.g. *Munk and Wunsch* [1982a], *Wunsch* [2002b]). We rely today on a 15-year continuous record of sea surface height measurements of quasi-global coverage from several altimetric missions [*Fu and Cazenave* 2001]. Maintaining this record indefinitely, as is done in the case of many meteorological observations, is an obvious priority, but is at present far from guaranteed. Direct, quasi-continuous in space and time measurements of the ocean interior have remained even more challenging. The World Ocean Circulation Experiment (WOCE) provides mostly (occasional) snapshots of hydrographic properties along several longitudinal and latitudinal sections, with consequent problems in interpretation due to the unsteady nature of the ocean circulation (e.g. [*Wunsch* 2006b]). Similarly, serious sampling issues beset most of the so-called “climatological” data sets such as the NODC World Ocean Atlas 2001 [*WOA01, Conkright et al.* 2002]. Over the past five years or so autonomous floats (initially P-ALACE, now ARGO) have taken on the role of augmenting altimetric with in-situ observations (but limited to 2000 m depth), at a reasonable quasi-global continuous coverage (e.g. [*Roemmich and Owens* 2000, *ARGO* 2002]). These data sets, combined with various other sources (XBT’s, TOGA/TAO array, tide gauges, surface drifters, satellite-borne measurements of sea surface temperature, wind stress, the time-varying gravity data from GRACE), while appearing plentiful, do in reality confront the oceanographer, with serious issues: (1) disparity of the data (how does one assess consistency of a sea surface height observation with a salinity measurement at 500 m depth?),

* *15 Years of Progress in Radar Altimetry Symposium*, Venice, 13–18 March 2006, ESA Special Publication SP-614

(2) data sparsity and sampling (3) quantification of uncertainties (4) dynamical consistency among various data sets (budgets). These become major stumbling blocks when attempting to address climate-related issues. Examples are

• estimates of decadal changes in the meridional overturning circulation, e.g. *Bryden et al.* [2005] which are based on only five transatlantic sections taken over the course of almost 50 years,

• estimates of change in ocean heat content that don't include contributions below 2000m depths (e.g. *Willis et al.* [2004]), or that have major sampling issues, and virtually no data at all in most of the Southern Ocean (e.g. *Antonov et al.* [2005]); the same is true, if not worse for attempts to derive salinity changes using the existing observational record [*Curry et al.* 2003],

• open questions regarding the nature and causes of global sea-level rise, see e.g. the review of *Munk* [2002], *Cazenave and Nerem* [2004],

• inferences from model-only studies, such as the one by *Marsh et al.* [2005] which do provide a dynamically consistent time-evolving state, but no systematic assessment of misfit to observations, or

• inferences from dynamically inconsistent assimilation products, such as the one by *Carton et al.* [2005], which, while providing "snapshots" of the time-evolving state, do not permit to address dynamical links between the state at different instances in time.

These issues will be taken up in Section 3.

A way forward is to combine, in an optimal manner, the available observations with a state-of-the-art ocean general circulation model to produce a complete, dynamically consistent ocean state estimate (OSE) which fits the observations within known or estimated uncertainties (comprising both observational errors as well as representation errors of the model). Various methods are available to achieve what reduces to a least-square fit (or optimization) of the model to the data. While generally referred to as "data assimilation" (e.g. [*Lermusiaux et al.* 2006]) we distinguish its heavy usage in the extrapolation (forecasting) context by the (mostly) meteorological community from our goal of interpolating, describing and understanding the observed system. We thus prefer the term "state estimation" to make this distinction explicit [*Wunsch and Heimbach* 2005].

Taking on the scope of producing, in a pre-operational fashion, an OSE of the time-evolving ocean circulation beginning with the ERS-1 and TOPEX/POSEIDON altimetric record, the Consortium "Estimating the Circulation and Climate of the Ocean" (ECCO) [*Stammer et al.* 2002b] set out in the late 1990s, adopting what is often called the adjoint method, but best known in a wider community as the method of Lagrange multipliers (e.g. *Wunsch* [1996, 2006a]). One of this method's key advantages is that it fullfills the requirement of dynamical consistency. That is, the best-estimate ocean trajectory obeys the underlying model equations exactly over the full interval of the state estimate without introducing artificial sources or sinks in the ocean interior to "nudge" the model to the data. This distinguishes the obtained estimate markedly from so-called "rea-analysis" efforts undertaken in the atmospheric community (e.g. NCEP/NCAR [*Kalnay and 21 others* 1996], ECMWF ERA-40 [*Uppala and 46 others* 2005]), or oceanographic products such as SODA [*Carton et al.* 2005]. It is widely known (e.g. *Taylor* [2000], *Stammer et al.* [2004]) that these reanalysis products suffer from significant imbalances in some of the fluxes provided and are of limited use when attempting time-dependent budget calculations. They also pose serious problems when used in driving ocean GCMs (e.g. *Large and Yeager* [2004]). Dynamical consistency is a key requirement to address science-related questions on the nature of oceanic variability. Another strength of the adjoint method is its ability to propagate information contained in observations backward in time, thus making optimal use of each element.

A first milestone of ECCO was the demonstration that this approach is feasible through the publication by *Stammer et al.* [2002a, 2003] of a quasi-global ocean state estimate at 2° horizontal resolution, comprising 23 vertical layers, covering the globe between 80°N/S for the period 1992 and 1997. In a second phase, the horizontal resolution was improved ro 1° (vertical resolution maintained), and the period extended to 2002 [*Köhl et al.* 2005]. The result of this optimization, version 1, which comprised 69 iterations, has been made publicly available via the central ECCO data server at [*ECCO* URL] and is now termed v1.69.

In 2004 the follow-on project ECCO-GODAE was put into place with the stated purpose to sustain and improve the ECCO OSE. At this point the OSE system was transferred from SIO to MIT, and production was resumed from v1.69 under what is now called version 2, covering the period 1992–2004 (soon to be updated to 2005) with computing resources provided by GFDL. Since then, efforts have gone into (1) improving various aspects of the model (resolution, numerical schemes, parameterizations), (2) improving the quality of the observations, (3) and addressing the issue of uncertainty and sampling of the various data sets. All these aspects, despite appearing technical, bear serious science questions, and ultimately affect the quality and interpretation of the estimate obtained. The present paper gives an overview of the work undertaken by the ECCO-GODAE group leading to the more recent product v2.177. The product is served in various forms (LAS, DODS, GDS, Dapper, Ingrid) to the community; for more details visit [*ECCO-GODAE* URL]. Example results based on this product showcase the power of having a complete state estimate for science purposes. An outlook is given into a future OSE system.

2 The ECCO-GODAE Ocean State Estimation system

The problem is formulated in terms of a least-square cost function J, given here in a very simplified form to focus on the overarching structure (specifics on the dimensionality of the concrete OSE, version 2, are provided in footnotes):

$$\begin{aligned} J = & \frac{1}{2}\Big(\mathbf{L}(\mathbf{x}(t),\mathbf{u}(t)) \,-\, \mathbf{H}(\mathbf{y}(\mathbf{t}))\Big)^T \\ & \mathbf{R}^{-1}\Big(\mathbf{L}(\mathbf{x}(t),\mathbf{u}(t)) \,-\, \mathbf{H}(\mathbf{y}(\mathbf{t}))\Big) \end{aligned} \quad (1)$$

$\mathbf{L}$ denotes the nonlinear ocean model which steps forward in time the model state $\mathbf{x}(\mathbf{t})$ [1], subject to a set of independent (or control) variables $\mathbf{u}(t)$ such as initial conditions and time-varying air-sea buoyancy and momentum fluxes [2], $\mathbf{y}(t)$ is a set of observations which are mapped from observation to model space via a data operator $\mathbf{H}$ [3], and $\mathbf{R}^{-1}$ is a weight matrix, under ideal circumstances an inverse error covariance, but in practice (absence of knowledge of the full covariance, and treatment of a $10^8 \times 10^8$-dimensional matrix) a diagonal matrix containing prior uncertainty estimates.

The minimum of J is sought, subject to the requirement that the model equations be fulfilled exactly. To do so, the gradient of J with respect to the set of controls $\mathbf{u}(t)$ is needed to iteratively decrease J via gradient-based descent algorithms (here a quasi-Newton variable storage method of *Gilbert and Lemaréchal* [1989]). The method of Lagrange multipliers turns this constrained optimization problem into an unconstrained one. Propagating the Lagrange multipliers or adjoint variables backward in time via the adjoint model (in essence the transpose of the tangent linear model) yields the gradient $\nabla_{\mathbf{u}} J$ in question. We omit details here and refer to e.g. *Wunsch* [1996, 2006a], but write down the general structure of the gradient $\nabla_{\mathbf{u}} J$ in terms of the adjoint:

$$\nabla_{\mathbf{u}} J \;=\; \nabla \mathbf{L}^{\mathbf{T}} \cdot \nabla \mathbf{H}^{\mathbf{T}} \cdot \mathbf{R}^{-1} \cdot \Big(\mathbf{L}(\mathbf{x}(t),\mathbf{u}(t)) - \mathbf{H}(\mathbf{y}(\mathbf{t}))\Big) \quad (2)$$

The interpretation of eqn. (2) is as follows: The gradient $\nabla_{\mathbf{u}} J$ is driven by the misfit of the model to the data misfit, which after weighting ($\mathbf{R}^{-1}$) is projected onto the model adjoint variable space via the data operator's adjoint, and propagated onto the control ajoint variable space via the model's adjoint operator (the model Jacobian's transpose).

The qualitative forms of eqns. (1), (2) serve to highlight the elements which affect the OSE system and determine its solution: (1) the observations, (2) the model and data operator (and their adjoints), (3) the weights or uncertainties. Problems in any of the above elements may result in problems in the optimization, and ultimately failure to produce an acceptable solution.

2.1 The ocean model and its adjoint

One of the cornerstones of the OSE system is the ocean model, here the MIT general circulation model (MITgcm) [*Marshall et al.* 1997a,b, *Adcroft et al.* 2002]. Configured in height coordinates, it solves the Boussinesq form of the Navier-Stokes equations for an incompressible fluid, hydrostatic or fully non-hydrostatic, in a curvilinear framework. The horizontal assembly of the finite volume grid cells is based on a domain decomposition to enable efficient parallelization across a variety of high performance compute (HPC) architectures. The model is endowed with state-of-the art physical parameterization schemes for sub-grid-scale horizontal and vertical mixing of momentum and tracer properties, as well as a sophisticated dynamic/thermodynamic sea-ice model, plus atmospheric boundary layer scheme over the open ocean. It is currently being used for high-resolution global-scale ocean simulations [*Menemenlis et al.* 2005a].

The model is continuously undergoing vigorous development to incorporate novel physics, numerical

[1] $\mathbf{x}(t)$ are elements of a $2.1 \cdot 10^7$-dimensional numerical state space per timestep, which is stepped forward in time at hourly intervals over a 13-year period.

[2] When assembled to a vector containing wet point only, the $\mathbf{u}(t)$ form a $3.1 \cdot 10^8$-dimensional control space subject to optimization

[3] Our observational data set consists of currently $4.1 \cdot 10^8$ elements $\mathbf{y}(t)$.

schemes and approaches for treating the horizontal and vertical grid (e.g. *Adcroft et al.* [2004], *Adcroft and Campin* [2004], *Campin et al.* [2004]). In this environment the ability to (re-)generate the adjoint becomes a crucial element to ensure that novel features of the parent model can be carried over to the adjoint model. What has made this possible is the use of so-called automatic (or algorithmic) differentiation (AD) tools [*Griewank* 2000], which generate exact adjoint code of the parent model. Key advantages of this approach are (1) rendering up-to-date versions of the adjoint model, (2) flexible configuration for any type of application (regional vs. global, various resolutions, various packages), (3) preservation of the parallelism of the parent model, and thus scalability on HPC systems, (4) independent test of the parent model and reduction in coding errors (undetected errors in the parent model may show up as apparent errors in the adjoint). We have been relying heavily on the "Tangent linear and Adjont Model Compiler" (TAMC) and in recent years, on its more mature commercial version "Transformation of Algorithms in Fortran" (TAF) by the company Fastopt (see *Giering and Kaminski* [1998], *TAF* [URL] as well as *Marotzke et al.* [1999], *Heimbach et al.* [2005]). We are also involved in the development of a new open source AD tool OpenAD to foster the wider use of such tools [*Utke et al.* 2006, *Naumann et al.* 2006, *OpenAD* URL].

Major changes in the model setup of the OSE are planned for version 3, but some preliminary changes were adopted already in going from version 1 to version 2. These include updating to a more modern equation of state [*Jackett and McDougall* 1995], moving to a third-order upwind scheme for tracer advection, updating the Gent-McWilliams/Redi eddy parameterization scheme, plus other minor technical changes. All updates are subject to the requirement of preserving a stable adjoint solution. The forcing cycle has been shortened to 6-hourly fields, which is the shortest available analysis window of reanalysis products.

2.2 The observations

Listing and describing all the observational data sets used in version 2 would stretch the limits of this proceedings. We refer instread to *Wunsch and Heimbach* [2006] for a full list, and to *ECCO-GODAE* [URL] for further details. Suffice it to say that, compared to version 1, essentially all data sets have been updated to more recent ver-

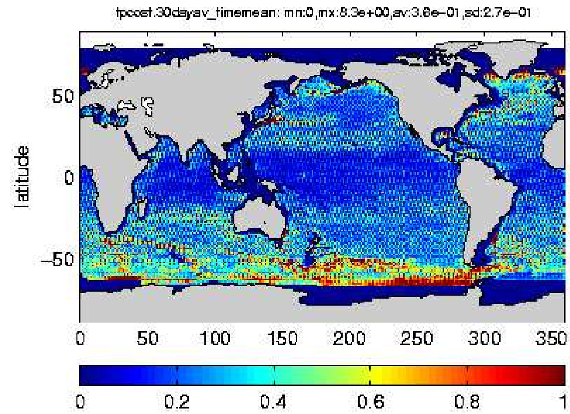

Figure 1: Space- and time-mean cost (normalized misfit) in sea surface height anomaly between model and merged TOPEX-POSEIDON/Jason observations.

sions and periods (the Jason-1 data were added to the T/P altimetric record, missing ERS-1 and ENVISAT data added to the ERS-2 record, GFO data are now being added to the system post-v2.177, newer versions of the ARGO data sets contain significantly more data over the whole observational record, improved estimates of the GRACE geoid have been issued, etc.). As for the climatological set, we have replaced most of the *WOA01, Conkright et al.* [2002] data set, by the WOCE Global Hydrographic Climatology of *Gouretski and Koltermann* [2004] since they provide more details regarding error estimates. Only the top 300 m of the WOA01 atlas is kept, to resolve (at least crudely) a seasonal cycle. It is noteworthy that *Gouretski and Koltermann* [2004], on the basis of their quality control, reject about 90% of the data used in the WOA01 atlas (for more details see *Wunsch et al.* [2006]).

All data were subject to independent quality control. This proved to be important as contamination could potentially hamper the conditioning of the gradient (apparent from eqn. (2)), and ultimately the behaviour of the optimization. Once a relevant norm $\mathbf{R}^{-1}$ had been defined, offline calculations were performed to detect unrealistically large distances relative to large-scale estimates (nondimensional cost values, eqn. (1) exceeding values of 50)). For the altimetric data sets, this lead to the identification of outliers which in most cases could be attributed to regions/periods covered by sea-ice. Coherent patterns in the set of outliers in the ARGO observations pointed to problematic instruments or significant drifts over time [*Forget and Wunsch* 2006].

By way of example, Fig. 1 shows the remaining misfit (the mean cost per element in units of its prior variance estimate or weight, eqn. (1) of sea surface height anomalies between the model and merged TOPEX-POSEIDON/Jason data of the v2.177 product. Good agreement is seen over large parts of the ocean, but remaining deviations point to deficiencies in the model's ability to resolve some relevant processes, particularly in the subpolar Atlantic and in parts of the Southern Ocean. Note that version 2 of the OSE does not include a sea-ice model (but see Section 4.1 for results with an experimental OSE setup that includes a sea-ice model). Note also that the solution is determined by the choices of weights, a crucial issue to be discussed next.

2.3 The uncertainty estimates

Version 2 remains limited to diagonal forms of the weight matrices $\mathbf{R}^{-1}$ in eqn. (1). Even determining those diagonal elements or prior errors can be complicated. Main issues lie with (1) the instrument errors (sometimes unkown, in particular for historical observations), (2) processing errors (e.g. corrections made to the altimetric data), systematic errors (biases, drift), (3) representation errors, and related sampling issues.

Most of the weights have been changed in version 2 of the OSE as a result of revisiting the above-mentioned issues and analyzing the nature of variability of the observed quantities. In addition, suitability of the weights for preconditioning the gradient for the descent algorithm (line search) needed re-considering. *Ponte et al.* [2005] provide an analysis for the combined TOPEX/POSEIDON and Jason record of sea surface height anomaly. They provide an error budget based on estimated errors in each of the elements entering the processing (tidal, electromagnetic bias, inverse barometer, sampling). The weights used for all altimetric data sets were updated accordingly. Improvements in the mean dynamic ocean topography (DOT) estimate provided by the new GRACE geoid over previous EGM96-based estimates were also assessed [*Stammer et al.* 2006].

For in-situ observations in particular, limitations imposed by the sampling on inferences of time-varying aspects of the data became apparent [*Forget and Wunsch* 2006, *Wunsch et al.* 2006]. Closely related and of equal significance is the assessment, required globally for the present application, of the hydrographic variability in relation to in-situ observations provided through the climatological atlases, hydrographic sections, and other isolated-in-time measurements. *Forget and Wunsch* [2006] provide a detailed analysis focusing on (1) the actual sampling of the climatological atlas, and the ARGO data, (2) standard deviations provided with the atlas and computed from ARGO measurements, (3) estimates of hydrographic variability from high-frequency in-situ observations as well as high-resolution model calculations of *Menemenlis et al.* [2005a], (4) related to (3), representation errors incured by the model's failure to resolve eddy-scale processes. They devise an interpolation method to obtain a global, three-dimensional uncertainty estimate which attempts to account for aspects (1) to (4). This estimate provides a basis for computing weights for all in-situ hydrographi observations in the version 2 cost function (CTD, XBT, ARGO, climatology).

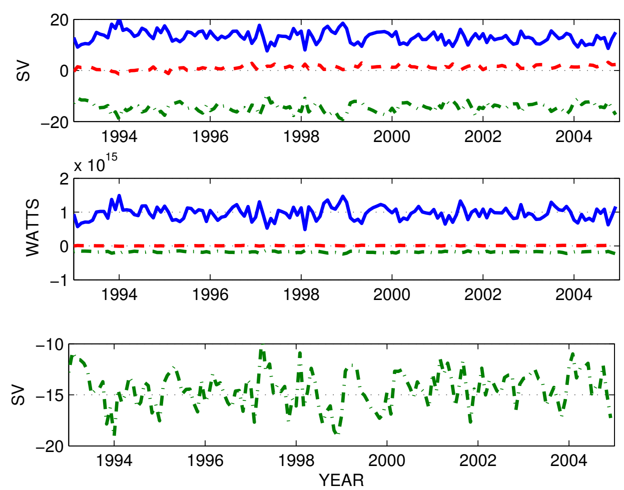

Figure 2: Top: monthly mean zonally and vertically integrated volume fluxes above 1,165 m (blue/solid), between 1,165 m and 4,450 m (green/dash-dot), and 4,450 m to bottom (red/dashed). Middle: same as top, but for associated heat transport. Bottom: expanded scale version of the 1,165m–4,450m volume transport (top panel).

3 Some science results

3.1 Decadal variation in mass and heat fluxes in the North Atlantic

Efforts to quantify and understand decadal changes in mass and property transports are faced with many of the problems alluded to in the introduction. Fig. 2 taken from a study by *Wunsch and Heimbach* [2006] of the North Atlantic meridional overturning circulation (MOC) and its associated heat flux encapsulates some of the issues. The top panel shows monthly mean zonally and vertically integrated volume fluxes above 1,165 m (blue/solid), between 1,165 m and 4,450 m (green/dash-dot), and 4,450 m to bottom (red/dashed). The middle panel shows the associated heat transport (same color coding). Things to note are (1) a clear compensation between upper-ocean and mid-depth volume transport, (2) an increase in mid-depth and abyssal trends in volume transport countering the slight decrease of the upper ocean branch (at -0.19±0.05 Sv/yr marginally significant), (3) a strong correlation between upper-ocean heat and volume transport (but the mean heat transport of 0.84±0.18 PW has no significant trend associated). A conclusion is that upper ocean heat transport variability is largely driven by variability in the volume transport. Clearly, changes in upper-ocean transport cannot be understood in isolation, but require knowledge of the state through the entire water column.

The sampling issue is highlighted by the bottom panel of 2 which is an expanded scale version of the 1,165m-4,450m volume transport in the top panel. The strong month-to-month variability ranging between -9 and -19 Sv raises serious concerns of aliasing in sub-sampled (e.g. section-based) estimates.

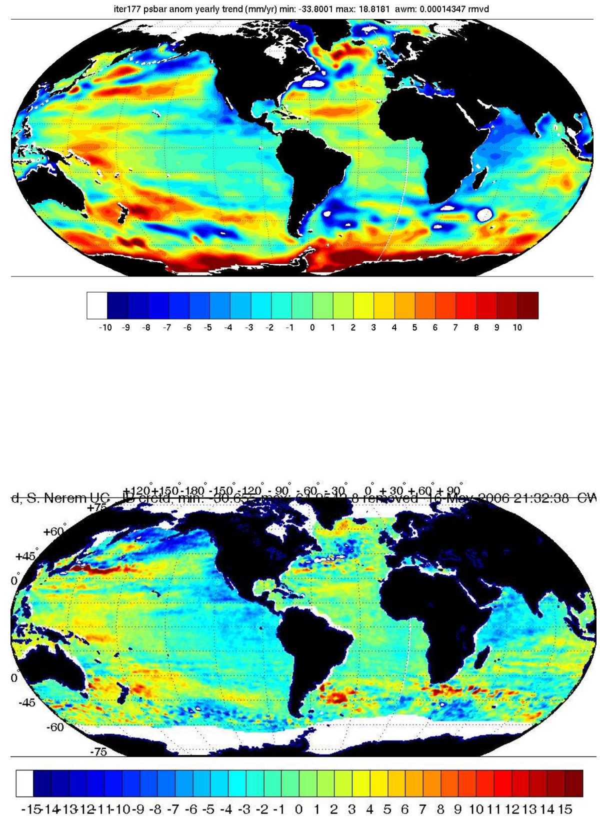

Figure 3: (top): Linear sealevel trends (mm/yr) over the period 1993–2004 computed from v2.177 solution. Patterns have a zero spatial mean given that the model solution conserves total volume. Trends in globally averaged steric height, not included in the estimate, amount to $\sim$ 3 mm/yr. (bottom): Linear sealevel trends (mm/yr) from altimeter data after *Cazenave and Nerem* [2004]. A spatial mean of 2.8 mm/yr has been removed for consistency with the top panel.

3.2 Global sealevel pattern changes

With more than a decade of near-global altimeter measurements available, one can start to examine regional patterns of low frequency variability in sealevel since the early 1990's. As an example, Figure 3, top panel, shows linear sealevel trends based on v2.177 solution, computed over the 12-year period 1993–2004. Strong spatial variations in trend are apparent, including changes in sign. Intricate patterns of sealevel rise and fall can be seen, for example, in the North Atlantic and North Pacific oceans. Amplitudes range from a few mm/yr to more than 1cm/yr in several regions, substantially larger than the global trend of $\sim$ 2.8 mm/yr inferred by *Cazenave and Nerem* [2004] from altimeter observations. The patterns in v2.177 solution resemble to a large extent the observed altimeter regional trends, bottom panel of Fig. 3, but with noticeable differences particularly in western boundary regions (e.g., Agulhas retroflection or Brazil-Malvinas confluence regions). These differences are expected because, in the optimization, relatively weak weights are given to the altimeter data in regions of strong eddy variability and consequently large representation errors *Ponte et al.* [2005]. Having a 3-dimensional, physically consistent solution provided by the optimization procedure permits a full exploration of the relation between sealevel trends, subsurface temperature and salinity fields, and relevant surface atmospheric forcing fluxes. Such a detailed study of the decadal sealevel trends is inpreparation [*Wunsch et al.* 2006].

4 Towards a next-generation OSE system

The quality of any state estimate is ultimately limited by the ability of the model to faithfully mimic the dynamical processes which govern the oceanic circulation. Efforts are currently under way to move to a next generation model configuration which will form the basis of an OSE, version 3. Major ingredients are planned to be (1) a truly global grid, (2) increased horizontal and vertical resolution, (3) improved numerical schemes and avoidance of "legacy approximations" (in the sense of *Griffies* [2005]), (4) a sea-ice model, (5) improved model forcing through the use of (simple) atmospheric boundary layer schemes, (6) improved parameterization of mixing processes, (7) an augmented or modified control space, (8) improved injection of observations, in particular those available at isolated instances in space and time (e.g. profiles), (9) move towards a mass-conserving, rather than volume-conserving model configuration. Addressing all these issues, while preserving full adjointability of the system is a major endeavour, and progress will have to come in discrete steps.

4.1 COSIE: An OSE coupled to a sea-ice model

Here we report on an experimental solution of a coupled ocean sea-ice estimation (COSIE) system similar in many aspects to v2.177, but with novel aspects addressing the above items (4), (5) and (7): The configuration uses the dynamic/thermodynamic sea-ice model of *Zhang and Hibler III* [1997], *Zhang et al.* [1998] to compute sea-ice concentration and volume, as well as modified air-sea fluxes over ice-covered regions of the ocean. Over open water bulk formulae are used (here the implementation of *Large and Pond* [1981, 1982]) to compute air-sea fluxes from the atmospheric state provided by NCEP/NCAR (surface air temperature, specific humidity, precipitation, wind velocity vector, net shortwave radiation). The computed heat flux are thus consistent with

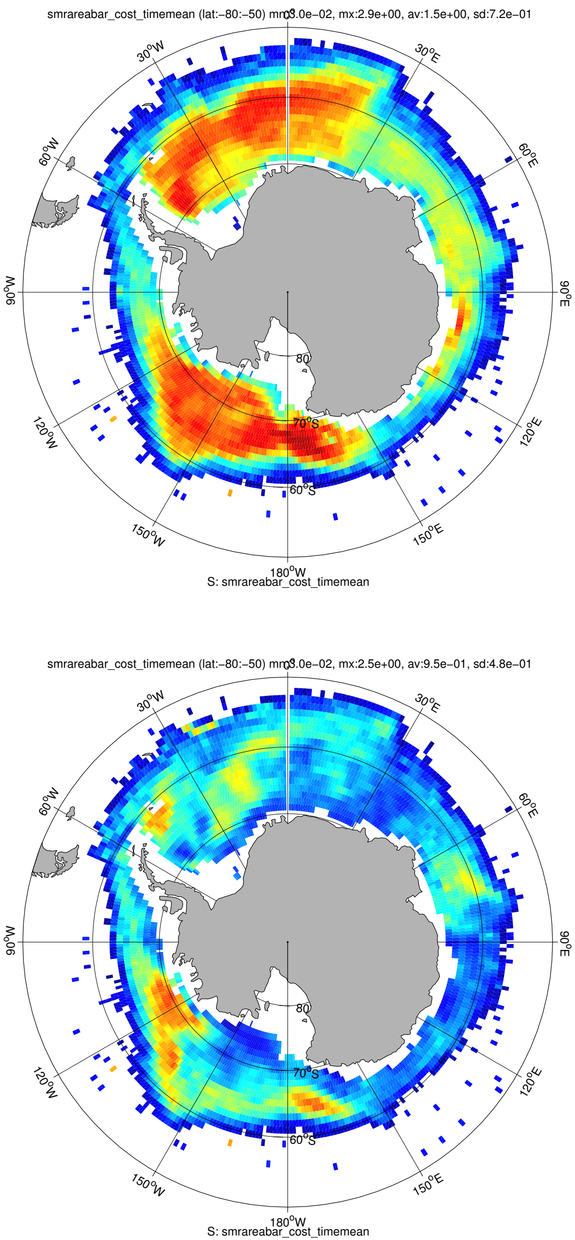

Figure 4: Mean cost (normalized misfit) between daily model and NSIDC sea-ice concentration for iteration 0 (top) and 27 (bottom) of COSIE. Color range is 0 to 3.

model's prognosed SST field. The control space spanning the time-varying surface boundary conditions is shifted from adjusting the air-sea fluxes to adjusting the atmospheric state. This allows for some physics-based coupling between the control adjustments for heat (air temperature) and momentum (wind speed) through the adjoint of the bulk formulae. The sea-ice model is currently used in the forward integration only, but switched off in adjoint mode. Misfits are diagnosed between prognosed daily sea-ice concentrations and combined Scanning Multichannel Microwave Radiometer (SMMR) and Special Sensor Microwave/Imager (SSM/I) data from the National Snow and Ice Data Center [*NSIDC* URL]. Corresponding mean (normalized) cost maps are depicted in Fig. 4 for iteration 0 and 27 of the experimental OSE. The significant reduction in cost is readily apparent, underlining the ability of the system to improve the coupled ocean/sea-ice state. The NSIDC data have not been used so far (but are planned) to explicitly constrain the model.

Fig. 5 depicts mean adjustments in surface air temperature (SAT) and precipitation after 27 iterations. Several aspects are apparent. The most prominent one is a significant reduction in precipitation over the tropics. This result is in agreement with detailed investigation of the re-analysis products which point to significant overestimation of precipitation in these regions [*Andersson and 13 others* 2005, *Uppala and 46 others* 2005]. As in the case of air-sea flux controls, dipole patterns in most variables of the adjusted atmospheric state are visible over the Gulf Stream region, indicating that the optimization attempts to reposition a mis-aligned Gulf stream through adjustment of the controls. A significant decrease in SAT over the seasonally ice-covered regions of the Southern Ocean has a clear beneficial impact as it reduces the misfit between modeled and observed sea-ice concentration. The origin of the strong positive SAT adjustments near the Antarctic coast is unclear, and likely of different origin for different regions. Note that these areas coincide with areas for which no observations are available. The shelf of the Weddell Sea near the Antarctic peninsula appears as one of the few regions in the Southern Ocean supporting multi-year sea-ice. In contrast, NSIDC time series suggest that sea-ice near the coast in the Ross Sea disappears much faster at the end of the Austral winter than in many other places of the Southern Ocean. The adjusted SAT point to these regions as interesting places to investigate in more detail.

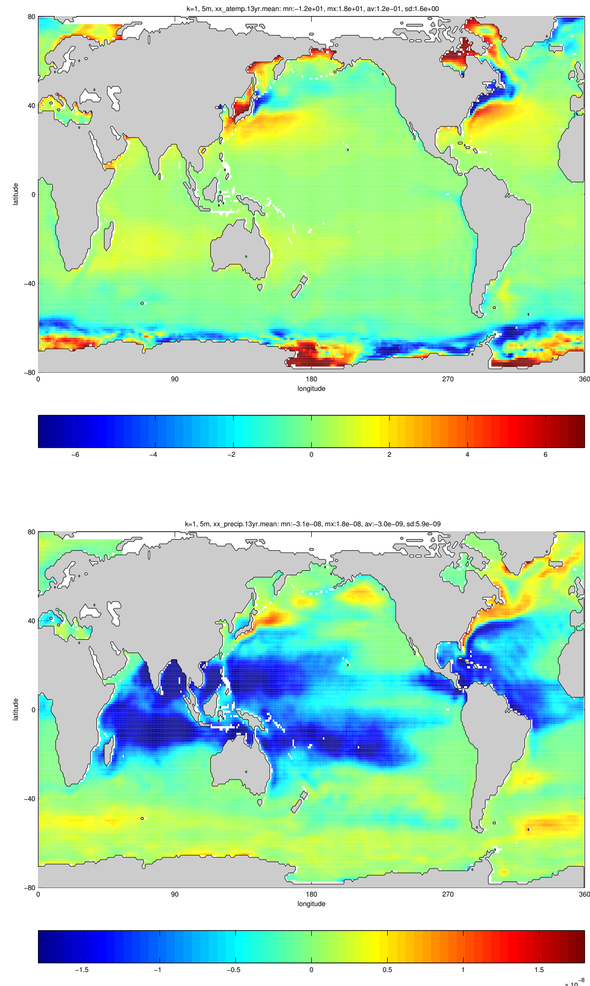

Figure 5: 13-year mean adjustment in atmospheric controls for surface air temperature (top) and precipitation (bottom) at iteration 27

4.2 High-resolution COSIE

In an effort to increase horizontal and vertical resolution, *Mazloff* [2006] is expanding the regional, high-resolution, adjoint-based state estimation work of *Ayoub* [2006] and *Gebbie et al.* [2006] to the dynamically much more complex Southern Ocean domain with 1/6° horizontal resolution and 42 vertical levels. This is a huge technological and computational challenge. Nevertheless, four iterations of the adjoint-method optimization have been already been successfully carried out (see *SDSC Threads* [February 2006]). This effort will ultimately merge into global high-resolution calculations.

Within NASA's Modeling, Analysis and Prediction Program (MAP) the ECCO2 project has embarked on the development and production of global, high-resolution, full-depth ocean and sea-ice state estimation [*ECCO2* URL]. To accomplish its goal, a symbiosis of state-of-the-art ocean modeling development, [*Menemenlis et al.* 2005a, *Hill et al.* 2006], advanced parameter estimation methods [*Menemenlis et al.* 2005b], and high-resolution adjoint-based estimation techniques as described above will be required.

Much remains to be done. For the present system, as more iterations bring the cost further down, more data types are added, the record extended in time, and the model improved, the solution will unavoidably change. A "*perfect*" solution won't exist in the foreseeable future, rather a series of estimates of hopefully improving quality over time, and with increasing details regarding their error bars. Each "intermediate" solution disseminated to the community should thus be perceived as a best present estimate, subject to change as our knowledge of the oceans evolves.

Acknowledgements

We are grateful to Charmaine King and Diana Spiegel for assistance with the model result analysis. Supported in part by NASA, NOAA and NSF through the National Ocean Partnership Program (NOPP), with additional support through NSF's ITR and CMG programs. Computing resources and support by GFDL, NCAR , SDSC and NASA ARC are gratefully acknowledged. Thanks are due to our many MITgcm and ECCO partners.

References

Adcroft, A., J.-M. Campin, C. Hill and J. Marshall [2004], 'Implementation of an atmosphere-ocean general circulation model on the expanded spherical cube', *Mon. Wea. Rev.* **132(12)**, 2845–2863.

Adcroft, A., J.-M. Campin, P. Heimbach, C. Hill and J. Marshall [2002], MITgcm Release1 Manual, (online documentation) , MIT / EAPS, Cambridge, MA 02139, USA. `http://mitgcm.org/sealion/online_documents/manual.html`.

Adcroft, A. and J.M. Campin [2004], 'Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models', *Ocean Modelling* **7(3-4)**, 269–284.

Andersson, E. and 13 others [2005], 'Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system', *Bull. Am. Met. Soc.* **86(3)**, 387–402.

Antonov, J.I., S. Levitus and T.P. Boyer [2005], 'Thermosteric sea level rise, 1955–2003', *Geophys. Res. Lett.* **32**, L12602.

ARGO [2002], 'Voyage of the argonauts', *Nature* **415**, 954–955.

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

Bryden, H.L., H.R. Longworth and S.A. Cunningham [2005], 'Slowing of the Atlantic overturning circulation', *Nature* **438**, 655–657.

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., B.S. Giese and S.A. Grodsky [2005], 'Sea level rise and the warming of the oceans in the Simple Ocean Data Assimilation (SODA) ocean reanalysis', *J. Geophys. Res.* **110**, C09006.

Cazenave, A. and R.S. Nerem [2004], 'Present-day sea level rise: observations and causes', *Rev. Geophys.* **42**, 1–20.

Curry, R.G., B. Dickson and I. Yashayaev [2003], 'A change in the freshwater balance of the Atlantic ocean over the past four decades', *Nature* **426**, 826–829.

Forget, G. and C. Wunsch [2006], 'Global hydrographic variability and the data weights in oceanic state estimates', *in preparation*.

Fu, L.-L. and A. Cazenave, eds [2001], *Satellite Altimetry and Earth Sciences. A Handbook of Techniques and Applications.*, Academic Press, San Diego, 463 pp.

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

Giering, R. and T. Kaminski [1998], 'Recipes for adjoint code construction', *ACM Transactions on Mathematical Software* **24**, 437–474.

Gilbert, J.C. and C. Lemaréchal [1989], 'Some numerical experiments with variable-storage quasi-Newton algorithms', *Math. Programming* **45**, 407–435.

Gouretski, V.V. and K.P. Koltermann [2004], OCE Global Hydrographic Climatology, Berichte des BSH 35/2004, Bundersamt für Seeschifffahrt und Hydrographie, Hamburg und Rostock (Germany).

Griewank, A. [2000], *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation*, Vol. 19 of *Frontiers in Applied Mathematics*, SIAM, Philadelphia, 369 pp.

Griffies, S.M. [2005], Some ocean model fundamentals, *in* E.Chassignet and J.Verron, eds, 'Ocean Weather Forecasting: an integrated view of Oceanography', Springer, Berlin, pp. 19–74.

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(8)**, 1356–1371.

Hill, C., D. Menemenlis, B. Ciotti and C. Henze [2006], 'Investigating solution convergence in a global ocean model using a 2048-processor cluster of distributed shared memory machines', *J. Scientific Programming* **submitted**, .

Jackett, David R. and Trevor J. McDougall [1995], 'Minimal Adjustment of Hydrographic Profiles to Achieve Static Stability', *J. Atmos. Ocean. Technol.* **12**(4), 381–389.

Kalnay, E. and 21 others [1996], 'The NCEP/NCAR 40-Year Reanalysis Project', *Bull. Am. Met. Soc.* **77(3)**, 437–471.

Köhl, A., D. Stammer and B. Cornuelle [2005], 'Interannual to decadal changes in the ECCO global synthesis', **Submitted for publication**, .

Large, W.G. and S. Pond [1981], 'Open ocean momentum flux measurements in moderate to strong winds', *J. Phys. Oceanogr.* **11**, 324–336.

Large, W.G. and S. Pond [1982], 'Sensible and latent heat flux measurements over the ocean', *J. Phys. Oceanogr.* **12**, 464–482.

Large, W.G. and S.G. Yeager [2004], Diurnal to decadal global forcing for ocean and sea-ice models: the data sets and flux climatologies, Technical Note NCAR/TN-460+STR, NCAR, Boulder, CO.

Lermusiaux, P.F.J., P. Malanotte-Rizzoli, D. Stammer, J. Carton, J. Cummings and A.M. Moore [2006], 'Progress and prospects od U.S. data assimilation in ocean research', *Oceanography* **19(1)**, 172–183.

Marotzke, J., R. Giering, K.Q. Zhang, D. Stammer, C. Hill and T. Lee [1999], 'Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport variability', *J. Geophys. Res.* **104, C12**, 29,529–29,547.

Marsh, R., B.A. de Cuevas, A.C. Coward, H.L. Bryden and M. Alvarez [2005], 'Thermohaline circulation at three key sections in the North Atlantic over 1985-2002', *Geophys. Res. Lett.* **32**, L10604.

Marshall, J., A. Adcroft, C. Hill, L. Perelman and C. Heisey [1997b], 'A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers', *J. Geophys. Res.* **102, C3**, 5,753–5,766.

Marshall, J., C. Hill, L. Perelman and A. Adcroft [1997a], 'Hydrostatic, quasi-hydrostatic and nonhydrostatic ocean modeling', *J. Geophys. Res.* **102, C3**, 5,733–5,752.

Mazloff, Matthew [2006], Production and analysis of an eddy-permitting Southern Ocean state estimate, Masters thesis 2006, Massachusetts Institute of Technology, MIT/EAPS, Cambridge (MA), USA.

Menemenlis, D., C. Hill, A. Adcroft, J.M. Campin, B. Cheng, B. Ciotti, I. Fukumori, A. Koehl, P. Heimbach, C. Henze, T. Lee, D. Stammer, J. Taft and J. Zhang [2005a], 'Towards eddy permitting estimates of the global ocean and sea-ice circulations.', *EOS Transactions AGU* **86(9)**, 89.

Menemenlis, D., I. Fukumori and T. Lee [2005b], 'Using Green's functions to calibrate an ocean general circulation model.', *Mon. Wea. Rev.* **133**, 1224–1240.

Munk, W. [2002], 'Twentieth century sea level: An enigma', *Proc. Natl. Acad. Sci. USA* **99(10)**, 6550–6555.

Munk, W. and C. Wunsch [1982a], 'Observing the ocean in the 1990s', *Phil. Trans. Roy. Soc. A* **307**, 439–464.

Naumann, U., J. Utke, P. Heimbach, C. Hill, D. Ozyurt, C. Wunsch, M. Fagan, N. Tallent and M. Strout [2006], Adjoint code by source transformation with OpenAD/F, *in* P. Wesseling, E. O nate and J. Periaux, ed., 'European Conference on Computational Fluid Dynamics ECCOMAS CFD 2006', TU Delft, p. .

Ponte, R.M., C. Wunsch and D. Stammer [2005], 'Spatial mapping of time-variable errors in TOPEX/POSEIDON and Jason-1 seasurface height measurements', *J. Atmos. Ocean. Technol.* **submitted**, .

Roemmich, D. and W.B. Owens [2000], 'The Argo Project: global oceanobservations for understanding and prediction of climate variability', *Oceanography* **13(2)**, 45–50.

Stammer, D., A. Köhl and C. Wunsch [2006], 'Impact of the GRACE geoid on ocean circulation estimates', *Submitted for publication* p. .

Stammer, D., C. Wunsch, I. Fukumori and J. Marshall [2002b], 'State estimation in modern oceanographic research', *EOS Transactions AGU* **83(27)**, 294–295.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C.N. Hill and J. Marshall [2002a], 'The global ocean circulation and transports during 1992 – 1997, estimated from ocean observations and a general circulation model.', *J. Geophys. Res.* **107(C9)**, 3118.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C.N. Hill and J. Marshall [2003], 'Volume, heat and freshwater transports of the global ocean circulation 1993 – 2000, estimated from a general circulation model constrained by WOCE data', *J. Geophys. Res.* **108(C1)**, 3007.

Stammer, D., K. Ueyoshi, W. B. Large, S. Josey and C. Wunsch [2004], 'Global sea surface flux estimates obtained through ocean data assimilation', *J. Geophys. Res.* **109**, C05023.

Taylor, P.K. [2000], Intercomparison and validation of ocean-atmosphere energy flux fields, Final report of the Joint WCRP/SCOR Working Group on Air-Sea Fluxes (WGASF) , SCOR Working Group 110, WCRP, Geneva.

Uppala, S.M. and 46 others [2005], 'The ERA-40 re-analysis', *Q. J. R. Meteorol. Soc.* **131(612)**, 2961–3012.

Utke, J., U. Naumann, M. Fagan, N. Tallent, M. Strout, P. Heimbach, C. Hill, D. Ozyurt and C. Wunsch [2006], 'OpenAD/F: A modular open source tool for automatic differentiation of Fortran codes', *ACM Transactions on Mathematical Software* **in preparation**, .

Willis, J., D. Roemmich and B. Cornuelle [2004], 'Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales', *J. Geophys. Res.* **109(C12)**, C12036.

WOA01, Conkright, M.E., R.A. Locarnini, H.E. Garcia, T.D. O'Brien, T.P. Boyer, C. Stephens and J.I. Antonov [2002], World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures, CD-ROM Documentation , National Oceanographic Data Center (NOAA/NODC), Silver Spring (MD), USA. `http://www.nodc.noaa.gov/OC5/WOA01/`.

Wunsch, C. [1996], *The ocean circulation inverse problem*, Cambridge University Press, Cambridge (UK), 442 pp.

Wunsch, C. [2002b], Ocean observations and the climate forecast problem, *in* R.Pearce, ed., 'Meteorology at the Millenium', Academic Press, pp. 233–245.

Wunsch, C. [2006a], *Discrete Inverse and State Estimation Problems : With Geophysical Fluid Applications*, Cambridge University Press, Cambridge (UK), 400 pp.

Wunsch, C. [2006b], Towards the World Ocean Circulation Experiment and a bit of aftermath, *in* M.Jochum and R.Murtugudde, eds, 'Physical Oceanography: Developments Since 1950', Springer.

Wunsch, C. and P. Heimbach [2005], 'Practical global oceanic state estimation', *Physica D* **under revision**, .

Wunsch, C. and P. Heimbach [2006], 'Estimated Decadal Changes in the North Atlantic Meridional Overturning Circulation and Heat Flux', *J. Phys. Oceanogr.* **in press**, .

Wunsch, C., R. Ponte and P. Heimbach [2006], 'Decadal trends in global sealevel patterns and implications for global mean errors', **in preparation**, .

Zhang, J. and W. D. Hibler III [1997], 'On an efficient numerical method for modeling sea ice dynamics', *J. Geophys. Res.* **102**, 8691–8702.

Zhang, J., W. D. Hibler III, M. Steele and D. A. Rothrock [1998], 'Arctic ice-ocean modeling with and without climate restoring', *J. Phys. Oceanogr.* **28**, 191–217.

ECCO [URL], `http://www.ecco-group.org/`.

ECCO-GODAE [URL], `http://ecco.mit.edu/`.

ECCO2 [URL], `http://ecco2.org/`.

NSIDC [URL], `http://nsidc.org/`.

OpenAD [URL], `http://www-unix.mcs.anl.gov/OpenAD/`.

SDSC Threads [February 2006], `http://www.sdsc.edu/user_services/newsletter/common/newsletter.php?%20issue=2006-02&corner=fstory`.

TAF [URL], `http://www.fastopt.com/`.