CITATION

COPYRIGHT
This article has been published in *Oceanography*, Volume 22, Number 3, a quarterly journal of The Oceanography Society. Copyright 2009 by The Oceanography Society. All rights reserved.

USAGE
Permission is granted to copy this article for use in teaching and research. Republication, systematic reproduction, or collective redistribution of any portion of this article by photocopy machine, reposting, or other means is permitted only with the approval of The Oceanography Society. Send all correspondence to: info@tos.org or The Oceanography Society, PO Box 1931, Rockville, MD 20849-1931, USA.
Ocean State Estimation for Climate Research

BY TONG LEE, TOSHIYUKI AWAJI, MAGDALENA A. BALMASEDA, ERIC GREINER, AND DETLEF STAMMER
ABSTRACT. Spurred by the development of satellite and in situ observing systems, global ocean state estimation has flourished in the past decade. Today, a suite of global ocean state estimates has been generated and is being applied to studies over a wide range of subjects in physical oceanography and climate research as well as other disciplines. This paper highlights some examples of using ocean state estimations for ocean and climate research. Many assimilation groups from different countries participated in a Climate Variability and Predictability program/Global Ocean Data Assimilation Experiment global ocean reanalysis evaluation effort in which intercomparisons were performed for a suite of diagnostic quantities and indices, including evaluations against observations. Examples of the intercomparisons are presented to highlight the consistencies and uncertainties of the estimation products and to examine the ability of these products to detect climate signals. Future challenges for state estimation for climate applications are also discussed.

INTRODUCTION
Since its inception, the Global Ocean Data Assimilation Experiment (GODAE) has maintained three streams of effort: (1) mesoscale ocean analysis and forecast, (2) initialization of seasonal-interannual prediction, and (3) state estimation (reanalysis) for climate research. Some GODAE groups contribute to more than one stream of effort. Separate papers in this issue review the first two aspects (e.g., Hurlburt et al. and Balmaseda et al.). This paper focuses on the third aspect. As satellite and in situ observing systems for the global ocean (e.g., altimetry and Argo) enhance and mature with time, there is an ever-increasing need to synthesize the diverse observations through data assimilation in which observations are used to constrain state-of-the-art ocean general circulation models (OGCMs). The resulting ocean reanalysis products aim to provide estimates of the time-varying, three-dimensional state of the ocean and to help understand the variability of ocean circulation and its relation to climate. They offer a tool to estimate quantities that are difficult to infer from observations alone, such as oceanic heat transport.

The vision of ocean state estimation as a means of synthesizing ocean observations into a dynamically consistent estimate of global ocean circulation was developed first under the World Ocean Circulation Experiment (WOCE) and subsequently as part of the World Climate Research Programme’s (WCRP’s) CLIVAR (Climate Variability and Predictability) and GODAE. As a result, and with the sustained commitment of various funding agencies, climate-oriented ocean reanalysis efforts have flourished in the past decade; the resulting systems routinely produce estimates of the physical state of the ocean that are publicly available through data servers. Examples include (but are not limited to) the following efforts:

• Simple Ocean Data Assimilation (SODA)
  http://www.atmos.umd.edu/~ocean/data.html
• National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory (GFDL) in the United States
  http://nomads.gfdl.noaa.gov/nomads/forms/assimilation.html
• NOAA National Centers for Environmental Prediction (NCEP) in the United States
  http://www.cpc.ncep.noaa.gov/products/GODAS
• European Centre for Medium-Range Weather Forecasting (ECMWF)
  http://www.ecmwf.int/products/forecasts/d/charts/ocean/reanalysis
• Mercator Océan in France
  http://www.mercator-ocean.fr
• CERFACS in France
  http://www.cerfacs.fr/globc
• Istituto Nazionale di Geofisica e Vulcanologia (INGV) in Italy
  http://www.bo.ingv.it/contents/Scientific-Research/Projects/oceans/enact1.html
• K-7 in Japan
  http://www.jamstec.go.jp/frcgc/k7-dbase2
• MOVE-G in Japan
  http://www.jma.go.jp/jma

A hierarchy of assimilation methods has been adopted by various groups to produce the ocean reanalyses, ranging from sequential methods such as optimal interpolation (OI), three-dimensional variational (3DVAR), and Kalman filter and smoother, to an adjoint method
(also known as Lagrange multiplier, four-dimensional variational, or 4DVAR method). The sequential methods as implemented by various ocean reanalysis groups (e.g., Fukumori and Manlanotte-Rizzoli, 1995; Cooper and Haines, 1996) are typically computationally more efficient than the adjoint method (e.g., Stammer et al., 2003; Sugiura et al., 2008). In particular, the Kalman filter approach can provide estimated errors of ocean state more readily. The sequential approaches allow the estimated state to deviate from an exact solution of the underlying physical model by applying statistical corrections to the state that could be based on simple physical constraints such as geostrophy. Such corrections act as internal sources/sinks of, for example, heat, salt, and momentum. These “internal controls” often render the estimated state closer to the observations that are being assimilated (depending on the treatment of the model and data errors). Different from sequential filtering, the adjoint method is often implemented in such a way that there is no internal statistical correction of ocean state. In other words, model physics is satisfied exactly without any internal/statistical source/sink. In such an implementation of the adjoint method, the state optimization is accomplished by adjusting the control variables, such as the initial state of the entire model trajectory, surface forcing, and model parameters. The absence of an internal source/sink in the adjoint approach makes it more difficult for the model to fit certain aspects of the observations over a long integration (unless time-distributed, internal controls such as mixing coefficients are included in the estimation). However, the adjoint-based estimation products are characterized by physical consistency such that property budgets are closed, and the estimated forcing is consistent with the estimated state. The adjoint method is the main approach used by the ECCO consortium and Japan’s K-7 project. Physical consistency is an important requirement for many aspects of ocean and climate research, such as heat budget analysis, and diagnosing the roles of different forcings on the ocean.

“...SIGNIFICANT ADVANCES HAVE BEEN ACHIEVED IN OCEAN STATE ESTIMATION EFFORTS THAT ARE GEARED TOWARD CLIMATE RESEARCH.”

more efficient than the adjoint method (e.g., Stammer et al., 2003; Sugiura et al., 2008). In particular, the Kalman filter approach can provide estimated errors of ocean state more readily. The sequential approaches allow the estimated state to deviate from an exact solution of the underlying physical model by applying statistical corrections to the state that could be based on simple physical constraints such as geostrophy. Such corrections act as internal sources/sinks of, for example, heat, salt, and momentum. These “internal controls” often render the estimated state closer to the observations that are being assimilated (depending on the treatment of the model and data errors).

Different from sequential filtering, the adjoint method is often implemented in such a way that there is no internal statistical correction of ocean state. In other words, model physics is satisfied exactly without any internal/statistical source/sink. In such an implementation of the adjoint method, the state optimization is accomplished by adjusting the control variables, such as the initial state of the entire model trajectory, surface forcing, and model parameters. The absence of an internal source/sink in the adjoint approach makes it more difficult for the model to fit certain aspects of the observations over a long integration (unless time-distributed, internal controls such as mixing coefficients are included in the estimation). However, the adjoint-based estimation products are characterized by physical consistency such that property budgets are closed, and the estimated forcing is consistent with the estimated state. The adjoint method is the main approach used by the ECCO consortium and Japan’s K-7 project. Physical consistency is an important requirement for many aspects of ocean and climate research, such as heat budget analysis, and diagnosing the roles of different forcings on the ocean.

**APPLICATIONS FOR CLIMATE RESEARCH**

CLIVARs and GODAE’s global ocean reanalysis products and tools have been applied to studies over a wide range of topics in physical oceanography, including the nature of sea level variability (e.g., Carton et al., 2005; Wunsch et al., 2007; Fukumori et al., 2007; Köhl and Stammer, 2008a), water-mass pathways (e.g., Fukumori et al., 2004; Wang et al., 2004; Masuda et al., 2006), mixed-layer heat balance (e.g., Kim et al., 2004, 2007; Du et al., 2008; Halkides and Lee, 2009), estimating surface fluxes and river runoff (e.g., Stammer et al., 2004), and interannual and decadal variability of the upper ocean and heat content (e.g., Masina et al., 2004; Capotondi et al., 2006; Köhl et al., 2007; Carton and Santorelli, 2009). They have also been applied to research in other disciplines such as biogeochemistry (e.g., McKinley et al., 2004; Dutkiewicz et al., 2006) and geodesy (e.g., Ponte et al., 2001; Dickey et al., 2002). Due to limited space, here we only highlight a few examples of ocean circulation studies, focusing on meridional overturning circulations (MOCs), and discuss the implications for observing systems.

MOCs play an important role in regulating meridional heat transport in

---

Tong Lee (tong.lee@jpl.nasa.gov) is Principal Research Scientist, Science Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA. Toshiyuki Awaji is Professor, Kyoto University, Kyoto, Japan, and Director General, Data Research Center for Marine-Earth Science and Technology, Yokohama, Japan. Magdalena A. Balmaseda is Research Scientist, European Centre for Medium-Range Weather Forecasting, Reading, UK. Eric Greiner is Scientific Supervisor, Mercator Océan, Ramonville-Saint-Agne, France. Detlef Stammer is Professor and Director, Institut für Meereskunde, KlimaCampus, Universität Hamburg, Hamburg, Germany.
the ocean that affects climate variability. Direct measurement or inference of MOCs and meridional heat transport is difficult. Again, ocean reanalysis provides an important tool for characterizing and quantifying these variables and understanding the mechanisms that produce their variations. For instance, estimates of volume, heat, and freshwater transports of the global ocean have been obtained by fitting an OGCM to WOCE data (Stammer et al., 2003), which facilitates the analysis of interannual to decadal changes of these transports in the ocean (e.g., Köhl et al., 2007).

Many studies have used various ocean reanalysis products to study regional MOCs and their relationships to heat transport and heat content changes. Lee and Fukumori (2003) used an ECCO assimilation product to study the shallow MOC that connects the tropical and subtropical Pacific Ocean, the so-called subtropical cell (STC). The interannual-decadal variability of the lower branches of the STC, the pycnocline flow, exhibits a zonal structure characterized by anticorrelated variability of the western boundary currents and interior flow. As such, western boundary and interior flows play opposite roles in charging and discharging upper-ocean heat content in the tropical Pacific, with the interior flow being more dominant. This process has important implications for understanding the El Niño Southern Oscillation (ENSO) and its decadal modulation, and indicates the need to have sustained in situ measurements near the low-latitude western boundary, a region not well resolved by existing observing systems. The partial compensation of boundary and interior flow associated with the Pacific STC has been confirmed by several subsequent studies, for example, by Schott et al. (2007) based on the German ECCO (G-ECCO) product. The multidecadal G-ECCO and Simple Ocean Data Assimilation (SODA) products have also been used to study the structure and variability of STCs in the Atlantic and Indian oceans as well as the Pacific Ocean (e.g., Shoenefeldt and Schott, 2006; Rabe et al., 2008; Schott et al., 2007, 2008).

Ocean reanalyses have also been applied to MOC studies away from the tropics. For example, Cabanes et al. (2008) used an ECCO reanalysis product to study the mechanism of interannual variability in North Atlantic MOC. The estimated MOC variability in the subtropics correlates well with the west-east difference in pycnocline depth, which reflects the zonal pressure gradient that drives the meridional flow associated with MOC. Variability in the west-east difference in pycnocline depth is primarily associated with fluctuations of pycnocline depth near the western boundary due to substantially larger Ekman pumping in the western part of the basin. The mechanism identified by this study would be helpful to the interpretation of dominant interannual variability captured by in situ monitoring systems for MOC (e.g., the RAPID [Rapid Climate Change] array). Köhl and Stammer (2008b) investigated decadal MOC changes as estimated by G-ECCO. Changes in the estimated MOC strength at 25°N are strongly influenced by southward communication of density anomalies along the western boundary originating from the subpolar North Atlantic that are related to changes in the Denmark Strait overflow, but are only marginally influenced by water mass formation in the Labrador Sea. The influence of density anomalies propagating along the southern edge of the subtropical gyre associated with baroclinically unstable Rossby waves is equally important. Wind-driven processes such as local Ekman transport explain a smaller fraction of the variability on those long time scales.

Decadal change in the North Atlantic MOC has been a topic of extensive discussion since the analysis by Bryden et al. (2005) using several synoptic hydrographic sections that suggest a substantial slowdown of the North Atlantic MOC at 26°N. Based on estimates by an ECCO-GODAE product, Wunsch and Heimbach (2006) found a much weaker but statistically significant decrease in the strength of MOC at 26°N as described by the northward volume transport above 1200 m. However, it is accompanied by a strengthening of the deeper meridional circulations (i.e., the southward outflow of North
Atlantic Deep Water and northward inflow of abyssal water). The decrease in meridional heat transport is not statistically significant.

An analysis of the ECMWF ORA-S3 operational ocean reanalysis (Balmaseda et al., 2007) shows that assimilation significantly improves the time-mean estimates of MOC strength (Figure 1). The resulting estimate of MOC strength at 26°N shows only a weak decreasing trend. No decrease in heat transport is found because the enhanced vertical temperature gradient due to near-surface warming counteracts the decrease in MOC strength. Both Wunsch and Heimbach (2006) and Balmaseda et al. (2007) discussed the large, high-frequency fluctuations in estimated MOC strength and the potential for their fluctuations to be aliased into low-frequency changes if sampling is infrequent (e.g., synoptic measurements). These studies indicate the need to enhance the scope of MOC observations (or ocean circulation in general) both in space and time in order to: (1) capture high-frequency signals that avoid significant aliasing, (2) have a more complete picture of MOC changes at different depths and latitudes, and (3) accurately estimate climatically important quantities such as meridional heat transport.

**Signal and Uncertainty in Ocean Reanalyses**

As part of the ongoing CLIVAR/GODAE global ocean reanalysis evaluation effort, many ocean reanalysis projects have participated in an intercomparison that involves several diagnostic quantities, including a comparison among reanalysis products and a comparison with observations. The objectives of such intercomparison are to: (1) examine the consistency of reanalyses (though multi-product comparison), (2) evaluate the accuracy of these products (by comparison with observations), (3) estimate uncertainty, (4) identify areas where improvement is needed, and (5) define the observational accuracy necessary to distinguish the quality of the reanalyses and identify future observational needs.

The following are examples of such intercomparisons as a way to estimate the uncertainty of a quantity estimated from different reanalysis products. Estimating uncertainty is a major challenge. In addition to the three-dimensional estimation of ocean state at a given time (analysis problem), the estimation of time evolution is also required in a reanalysis. The time evolution represented by an ocean reanalysis will be sensitive to time variations in the observing system, to the errors of the ocean model, to atmospheric fluxes, and to the assimilation system, which are all often flow dependent, and not easy to estimate. In cases where the complexity is too large to be tackled analytically, ensemble methods can offer a helping hand. So, in the same way as the ensemble of multimodel forecasts is used to get an estimation of the uncertainty of future climate projections, an ensemble of different reanalysis products can be used to provide a first glance at

---

Figure 1. Atlantic meridional overturning circulation (MOC) variability at 26°N. The MOC time evolution for the ocean reanalysis done by the European Centre for Medium-Range Weather Forecasting (black) and for the no-assimilation run (blue) is shown using monthly values (thin lines) and annual means (thick lines). Over-plotted are the annual mean MOC values from Bryden et al. (2005) based on synoptic hydrographic sections (red circles), and Cunningham et al. (2007) based on RAPID (Rapid Climate Change) mooring data (green circle). After Balmaseda et al. (2007)
the time evolution and uncertainty of the climate reconstructions, although the caveats of the ensemble methodology should be taken into account when interpreting the results.

This approach was followed by Balmaseda and Weaver (2006) in the comparison of ocean reanalysis organized by the CLIVAR Global Synthesis and Observations Panel (http://www.clivar.org/organization/gsop/synthesis/synthesis.php). The results from 20 global ocean reanalysis products, with different models, assimilation methods, and forcing fluxes, were gathered together to examine the time evolution of upper-ocean temperature and salinity. One of the reanalyses did not include any ocean model. Some of them did not include any ocean observations, being the results of ocean models forced by atmospheric forcing fluxes. For the early years (1960–1980), there were only seven available products. For the period 1993–2002, there were 20 products, and very few were available post 2002. The scarcity of reanalysis products for the last few years hampers assessment of uncertainty on recent climate signals.

Figure 2 shows the signal-to-noise ratio of the interannual variability of temperature (left) and salinity (right) in the upper 300 m (T300 and S300 in what follows) for different regions. The signal is estimated as the temporal standard deviation of the ensemble mean, and the noise as the ensemble spread. For the period 1960–2002, the North Atlantic (NATL, 30°N–60°N) stands out as the region where the temperature signal dominates the noise. The signal is also quite clear in the equatorial Pacific (EQPAC, 5°S–5°N), the tropical Atlantic (TRATL, 20°S–20°N), and the global ocean. In the most recent period (1993–2002), the signal-to-noise ratio in temperature is larger than one in most areas, and the equatorial Pacific clearly stands out. In contrast, the upper ocean salinity signal-to-noise ratio is less than one almost everywhere, except for the equatorial Pacific, where it is close to one. These numbers indicate the need to have enhanced and sustained

![Figure 2. Signal-to-noise ratio for the time evolution of temperature and salinity in the upper 300 m (T300 and S300) in different regions. The estimates have been done separately for the period 1960–2002 and for the most recent period, 1993–2002.](image-url)
observations in many areas of the global ocean, especially for such variables as salinity. They also highlight the need to improve modeling and assimilation systems to achieve better consistency.

SUMMARY AND THE CHALLENGES AHEAD
As part of CLIVAR’s and GODAE’s efforts in the past decade, significant advances have been achieved in ocean state estimation efforts that are geared toward climate research. A suite of global ocean reanalysis products have been produced and updated on a routine basis. There has been an ever-increasing number of applications of these products for oceanographic and climate-related studies over a wide range of topics. The existence of these products also allows a comprehensive intercomparison to evaluate the consistency and fidelity of the various climate-related diagnostic quantities estimated from different products. The intercomparison helps identify areas that need improvement in ocean reanalysis and observing systems. Despite the significant advances in ocean reanalysis, many challenges lie ahead. Ocean data assimilation groups need to work closely with the modeling community to improve model physics, especially that associated with the bias of the mean state.

Estimates of data and model errors dictate the outcome of state estimation for a given model. Therefore, the ocean reanalysis community needs to work closely with the observational community to obtain robust estimates of data errors, an important issue that is often left in the hands of data assimilation groups. The uncertainty in the observations associated with factors such as the changing fall rate of expendable bathythermographs over the past few decades, and some biased Argo data in the past few years, are examples that show that more effort is needed to estimate data error. Moreover, much research is still required to understand model errors. In particular, we need to identify the sources of model errors and determine if appropriate new control variables can be adopted to account for such error sources (e.g., internal model error associated with mixing parameterizations). Improved knowledge about model and data errors used in data assimilation would enhance understanding of the causes for both consistency and discrepancy among global ocean reanalysis products as shown by various intercomparison efforts (e.g., http://www.clivar.org/organization/gsop/gsop.php; Carton and Santorelli, 2009; Gemmell et al., 2008).

Computational resources remain a critical issue for the estimation effort based on ensemble and adjoint methods because they limit the ensemble size and model resolution that one can afford. Finally, the coupled nature of the climate system prompts a coupled approach for state estimation that includes components of the climate system (e.g., the atmosphere, cryosphere, hydrosphere, and biogeochemistry) in order to properly account for the potential feedback among different elements of the system. Currently, coupled ocean-atmosphere, ocean-ice, and ocean physics-biochemistry data assimilations are still in their infancy but are expected to pick up momentum in the coming decade.

ACKNOWLEDGEMENTS
The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA).

REFERENCES
Cunningham, S.A., T. Kanzow, and D. Rayner, M.O. Baringer, W.E. Johns, J. Marotzke, H.R. Longworth, E.M. Grant, J.J.-M. Hirschi,