Review in Advance first posted online on October 14, 2015. (Changes may still occur before final publication online and in print.)

Ocean Data Assimilation in Support of Climate Applications: Status and Perspectives

D. Stammer,[1] M. Balmaseda,[2] P. Heimbach,[3,4] A. Köhl,[1] and A. Weaver[5]

[1]Centrum für Erdsystemforschung und Nachhaltigkeit (CEN), Universität Hamburg, 20148 Hamburg, Germany; email: detlef.stammer@uni-hamburg.de

[2]European Centre for Medium-Range Weather Forecasts, RG2 9AX Reading, United Kingdom

[3]Institute for Computational Engineering and Sciences and Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas 78712

[4]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

[5]Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), SUC URA, CNRS 1875, 31100 Toulouse, France

Annu. Rev. Mar. Sci. 2016. 8:18.1–18.28

The *Annual Review of Marine Science* is online at marine.annualreviews.org

This article's doi:
10.1146/annurev-marine-122414-034113

Keywords

data assimilation, observing system, data synthesis, modeling, climate predictions, ocean circulation

Abstract

Ocean data assimilation brings together observations with known dynamics encapsulated in a circulation model to describe the time-varying ocean circulation. Its applications are manifold, ranging from marine and ecosystem forecasting to climate prediction and studies of the carbon cycle. Here, we address only climate applications, which range from improving our understanding of ocean circulation to estimating initial or boundary conditions and model parameters for ocean and climate forecasts. Because of differences in underlying methodologies, data assimilation products must be used judiciously and selected according to the specific purpose, as not all related inferences would be equally reliable. Further advances are expected from improved models and methods for estimating and representing error information in data assimilation systems. Ultimately, data assimilation into coupled climate system components is needed to support ocean and climate services. However, maintaining the infrastructure and expertise for sustained data assimilation remains challenging.

1. INTRODUCTION

Ocean data assimilation (ODA) encompasses a broad set of mathematical and computational tools aimed at providing the best possible descriptions of the time-varying ocean circulation. It thereby supports studies of ocean dynamics, in particular for estimating unobservable quantities. The results are used to describe the impact of the changing ocean circulation on various quantities of societal relevance, such as the interaction of the ocean with its ecosystems, its biogeochemistry, the marine (sea ice) or marine-terminating cryosphere, and the coupled climate system as a whole. One strand of ODA activities is to produce useful descriptions of the ocean's flow field as the basis for deriving products in the context of ocean services. Ultimately, ODA aims to improve the skill of climate predictions by providing accurate descriptions of the present climate state as initial conditions for coupled climate models in support of climate services.

Although the terms ODA and ocean state estimation (OSE) are often used synonymously, they are in fact different inverse approaches to an ocean synthesis, describing the time-varying ocean circulation based on all available observations and the underlying dynamics as embedded in circulation models. The term data assimilation (DA) was coined initially in the field of numerical weather prediction, where it referred to the technique of creating initial conditions for atmospheric models designed to forecast over timescales of hours to a few days, thereby emphasizing the instantaneous state of the atmosphere. [Bouttier & Courtier (1999) summarized the comprehensive mathematical expositions of the original DA approaches.] These approaches were later adopted by the oceanographic and seasonal forecast communities for the purposes of producing nowcasts and initializing ocean and seasonal forecasts (e.g., Anderson et al. 1996, Talagrand 1997). By contrast, from the beginning of the World Ocean Circulation Experiment (WOCE), OSE was intended to bring all ocean surface (including satellite data) and subsurface observations into a dynamically consistent description of the past and recent time-varying ocean circulation for the purpose of studying ocean dynamics and variability as well as global-scale and regional energy, heat, and water budgets (Munk & Wunsch 1982). By definition, OSE therefore focuses on the evolving state of the ocean and on long timescales.

The first box inverse applications in oceanography were introduced during the 1970s (see Wunsch 1978) to describe the steady ocean circulation using different hydrographic databases (Macdonald 1998, Ganachaud & Wunsch 2003). At around the same time, however, it became clear that any ocean synthesis effort must address the ocean circulation as a time-varying problem. This insight fostered the development of modern DA methods, which began in the late 1980s (Bennett 1992, Anderson et al. 1996, Malanotte-Rizzoli 1996, Wunsch 1996). Important milestones during this evolution included the development of inverse methods that can be applied to ocean circulation models using supercomputers. The steps encompassed making filter-based approaches (see below) technically feasible and developing smoother-based approaches, such as those that employed adjoint representations of modern primitive-equation models (e.g., Thacker & Long 1988). Applications of the adjoint technique to complex models were made possible by the development of automatic differentiation techniques and software tools (Giering & Kaminski 1998) and their pilot application to ocean problems (Marotzke et al. 1999). This approach required establishing a computer infrastructure suitable for solving large nonlinear optimization problems.

Traditionally, ODA is associated with filters, whereas OSE typically uses smoothers. Today, both approaches have evolved into mature fields, comparable in sophistication and usage to atmospheric reanalysis, i.e., a repetition of the operational analysis system of a numerical weather forecast center over a historic period using exactly the same model setup (the term reanalysis in the ocean is often used synonymously with the term synthesis). Their difference in intention has largely diminished, with both now aiming to support climate-oriented ocean synthesis. In particular, both approaches are used today to initialize climate forecasts and remain concerned with improving ocean and climate models, providing uncertainty estimates, and helping to improve the ocean observing system. To deal with all these requirements properly, ultimately DA will need to target the coupled atmosphere-ocean (and marine cryosphere) climate system and the coupled physical-biological-biogeochemical ocean so that it can be used to support both climate and ocean services. However, significant improvements are needed before the full potential of DA can be reached and the goals of ocean synthesis in general can be accomplished.

This article provides a critical review of the status of ODA in support of climate applications and lays out the developments necessary to reach its full potential for oceanography and climate science at large. We review the strengths and weaknesses of various ODA approaches, provide examples of ongoing applications, and summarize the role that ODA plays not only in analyzing the ocean but also in initializing coupled models, with an emphasis on climate applications. We then identify the improvements required to move toward ocean and climate information systems in support of many applications. We do not discuss short-term predictions and operational oceanography in detail. Reviews of marine forecasting applications (both global and regional) were provided recently by Edwards et al. (2014) and Martin et al. (2015). Relevant reviews of ocean assimilation in the context of operational oceanography and ocean state estimation have also been published by Schiller et al. (2013) and Wunsch & Heimbach (2013).

2. ASSIMILATION FRAMEWORK

ODA and OSE are general frameworks for finding the solution to ocean inverse problems by converting information available in ocean observations into estimates of the ocean state, including uncertain physical parameters such as surface forcing, mixing, and viscosity coefficients that are not directly observable and therefore are not well determined from observations alone. In practical terms, this entails bringing an ocean circulation model into consistency with the observed ocean state (within the error bars of both). The basic ingredients for such an approach are (*a*) a model that is being constrained by (*b*) quality-controlled data, (*c*) error information about both the model and data, (*d*) a methodology by which data and model results are fused, and (*e*) a method to estimate uncertainty information about the estimated state.

2.1. Models and Data

In the following, we first define a model in the context of data assimilation and then describe data issues.

2.1.1. Models. In the context of ODA, a model can be any mathematical description of an ocean parameter (or variable in the widest sense) that is being estimated through the DA approach. Such a description can be a simple statistical or dynamical relationship between the parameter of interest and observables. However, almost all present-day physical applications resort to comprehensive general circulation models (GCMs) of the ocean or the fully coupled climate system. Griffies & Adcroft (2008) have reviewed ocean model formulations, and the remaining model deficiencies can be summarized as follows (e.g., Griffies et al. 2001): missing physics not embedded in the underlying equations, structural errors in the formulation of numerical algorithms, unresolved sub-grid-scale physical processes and uncertainties in their parameterization, and uncertain model parameters (e.g., mixing and diffusion). Uncertainties also arise from inaccurate initial or boundary conditions, the latter of which include surface forcing fields and interactions with the ocean floor and the terrestrial hydrology. A specific goal of state estimation is to improve those uncertain model parameters, either individually or in combination. However, depending on the approach, success may be limited, and large uncertainties in the estimation remain that are not always easy to quantify.

2.1.2. Data. ODA fundamentally depends on the availability of quality-controlled observations provided by an ocean or climate observing system. Through the experience gained during WOCE and subsequent efforts such as OceanObs'99 (Smith & Koblinsky 2001) and OceanObs'09 (Hall et al. 2010), the ocean observing system has evolved into a multitude of in situ and satellite-based measurement platforms, communication components, and data analysis centers. Satellite observations, in particular altimetry, scatterometry, and passive microwave radiometry, have proven indispensable for observing ocean variability (Fu & Cazenave 2001). The Argo network (Roemmich et al. 2001) enables continuous monitoring of the temperature and salinity of the upper ocean on basin scales down to 2,000-m depths. Merging satellite observations, other ocean observations, and an ocean circulation model into a description of the ocean flow field through DA is important to maximizing the use of existing observations for oceanography and climate studies and should be considered part of a complete observing strategy, equivalent to the one used in numerical weather prediction.

Substantial and often unknown uncertainties remain in existing observations, with the expendable bathythermograph (XBT) fall-rate errors being but one prominent example (Abraham et al. 2013). Uncertainties in surface fluxes are usually unknown, and continued data reanalysis and quality-control efforts must be part of any sustained ocean and climate observing efforts. In addition, significant gaps remain in the ocean observing system, such as the lack of large-scale and sustained observations in the deep ocean below 2,000 m and observations of currents. Evaluating past climate variability and change from an observing system and forcing fields that have changed markedly in quality and quantity over time remains a major challenge; ODA efforts can support this process and can also be a valuable tool for optimizing the future ocean observing system.

2.2. Methodologies and Approaches

Most ODA approaches are variants of the classical least squares method of combining models with data, assuming that errors are Gaussian. The best solutions ideally encompass dynamically consistent state fields, uncertain model parameters such as mixing coefficients and sub-grid-scale closure, and error estimates of these fields and parameters. The resulting states, along with the inferred uncertain parameters, minimize an objective function, J, measuring the weighted squared norm of the vector of differences between observations and their model equivalents. The term observations is used here in a general sense and includes prior estimates of the adjustable fields or parameters as well as the ocean observations proper. The weighting matrix is defined as an estimate of the inverse of the error covariance matrix of the observations.

Major differences remain in the underlying assimilation schemes, which range from simple but computationally efficient [e.g., optimal interpolation (OI)] to rigorous but computationally intensive [e.g., Kalman filters (KFs), four-dimensional variational assimilation (4D-VAR), adjoint approaches, and other smoothers]. Applied DA schemes (e.g., Wunsch 1996) vary in the way the individual DA components are defined and in the extent to which the optimum values of J are subjected to additional conditions. This concerns, for example, the details of how DA schemes assimilate available observations, whether a solution to a constrained or unconstrained optimization problem is sought, and the level of accuracy with which prior error estimates of observations and the model dynamics are described.

As a result of model structural errors, obtaining realistic and dynamically consistent solutions with reliable and formal error information is not yet possible. Different DA methodologies

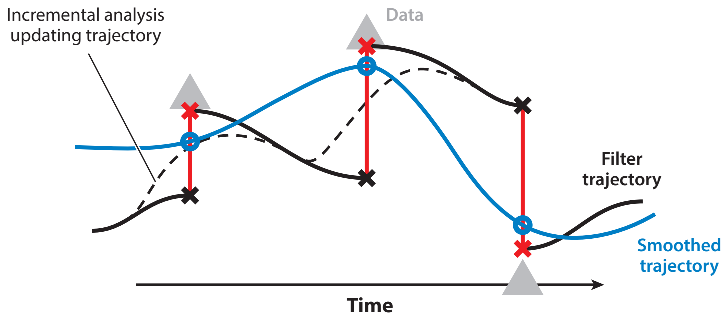

Figure 1

Schematic of the differences between filters and smoothers in producing an estimated state.

make different compromises between the fidelity and range of temporal and spatial scales to be represented and the degree of dynamical consistency sought in the solution. Understanding the substantial difference in the resulting solutions (**Figure 1**) is essential for their appropriate use. As can be inferred from the figure and described in more detail below, DA usually minimizes (in a least square sense) the prediction error, whereas SE minimizes an error over the entire time (see also Sorensen 1970). In the following, we describe two types of approaches typical for ODA and OSE: filters and smoothers.

2.2.1. Filters. Filter approaches sequentially estimate the ocean state at discrete points in time (so-called analysis steps) by merging present observations with the model forecast (or background) state, which, as a result of previous assimilation cycles, implicitly contains information from past observations. The introduction of the analysis increment that corrects the model state may violate conservation principles (as embedded in the first principles of the ocean circulation) and often may introduce discontinuities in the time evolution of the model trajectory. The use of incremental analysis updating (Bloom et al. 1996) can remedy discontinuities to some extent by transforming the increment into a forcing that distributes the correction over a particular period; the corrections remain dynamically unbalanced, however. Nevertheless, the resulting fields are consistent with the prescribed model forecast and data error covariances at this moment, and applications (e.g., for skillful forecasting) usually justify this approach. Approaches used in oceanography encompass three major avenues: OI, three-dimensional variational assimilation (3D-VAR), and various forms of the KF (Kalman 1960); the first two approaches can be shown to be approximations of the latter.

OI is the simplest form of an optimal least squares estimator (e.g., Gandin 1963). For each observation, a correction of the model by observations is defined based on the difference between the observation and the corresponding model simulation (referred to as the innovation). Interpolated values are then calculated from a linear combination of the innovations weighted by the inverse of the sum of the estimated observation error variance and the background error variance at observation points. OI provides an optimal instantaneous estimate for a particular set of constant weights; however, the OI solution is suboptimal over the entire measurement period because a time dimension is absent from the problem it solves (e.g., Fukumori 2002).

The KF, which is likewise a minimum variance estimator developed for solving prediction problems, has the advantage that it evolves the model state error covariance matrix in time according to the underlying dynamics of the numerical model and the assumed error covariance matrix of the numerical model. In practice, propagating the model state error covariance matrix is associated with a large computational burden, which makes the complete KF unfeasible for assimilating observations into full ocean GCMs. Several approximations of the KF have been devised; among these is the so-called partitioned KF, which solves the larger estimation problem by partitioning it into a series of smaller calculations (Fukumori 2002), thereby limiting errors to small correlation distances and their regional approximations. An extended KF (Gelb 1974) can be applied to weakly nonlinear problems under the tangent-linear approximation but still suffers from excessive computational costs. For stronger nonlinear problems, Evensen (1994) proposed a different extension of the KF, called the ensemble KF (EnKF), to estimate the model forecast error covariance matrix by means of a limited number of Monte Carlo simulations from a set of parallel analyses. In contrast to other realizations of the linear KF, the EnKF is suitable for high-resolution global eddy-permitting DA.

Several variants and extensions followed to deal with large dimensions. Among them, the singular evolutive extended Kalman (SEEK) filter and its interpolated variant, the singular evolutive interpolated Kalman (SEIK) filter developed by Pham et al. (1998), use empirical orthogonal functions to reduce the rank of the covariance matrix. To overcome problems associated with using small sample sizes in ensemble methods and the undesirable impact of the analysis step on the properties of the ensemble, Anderson (2001) proposed the ensemble adjustment KF, which is based on the ensemble transformation (Bishop & Toth 1999) and does not require adding perturbations to the observations.

3D-VAR is a maximum likelihood estimator that treats the elements in J independently in time and seeks an approximate solution through iterative minimization (e.g., Derber & Rosati 1989, Courtier et al. 1998). Its implementation requires the existence of the adjoint of the observation operators, not of the full GCM. In contrast to a normal sequential approach, 3D-VAR eliminates the need to split the analysis domain into subsections (so-called data selection, a source of noise in OI-type analyses) and provides a more general framework for including complex (including nonlinear) constraints in the cost function, such as nonlinear observation operators, dynamical balance constraints, and physically motivated conservation relationships (Ricci et al. 2005, Weaver et al. 2005). It allows for full-rank, nondiagonal formulations of the background error covariance matrix (Weaver & Courtier 2001).

2.2.2. Smoothers. Smoother-based approaches use observations from the future and the past to constrain the ocean circulation in a retrospective analysis. They differ from filter-based methods in that they estimate an ocean state not by changing the prognostic model state at analysis times but rather by changing model-independent parameters (as opposed to elements of the prognostic state) such that the simulated state best matches, to within uncertainty measures, the observed ocean state over an extended time period (years to several decades). The solution thereby obeys the ocean dynamics as embedded in the underlying GCM, is dynamically self-consistent, and guarantees the conservation of heat, freshwater, and momentum over the estimation period. Estimation efforts are typically targeted at reconstructions and descriptions of the time-varying ocean circulation.

The development of two major smoother approaches was essential for making OSE practical: the optimal Rauch-Tung-Striebel (Rauch et al. 1965) smoother and the adjoint method. These methods have different algorithmic properties but are equivalent, at least for linear systems, as long as they make the same assumptions about the data and model dynamic constraint errors (e.g., Bennett 2002, Lee et al. 2009; comprehensive mathematical expositions of the original smoother formulations are provided by Bouttier & Courtier 1999 and Wunsch 1996). 4D-VAR is a variant of the adjoint method that is applied over shorter time windows and offers substantial benefits over 3D-VAR (e.g., Weaver et al. 2003).

The optimal Rauch-Tung-Striebel smoother is a minimum variance estimator and thus recursive algorithm that seeks estimates of the state vector and associated uncertainty at each point in time based on all observations from both the past and the future (e.g., Cohn & Dinovitzer 1994). The use of observations from the future leads to uncertainties that are smaller than those associated with filtered results (e.g., Fukumori 2002). This approach is complementary to the KF in that it acts to smooth the filtered results by estimating model parameters required to reduce the temporal discontinuities that result from the sequential input of data.

By contrast, the whole-domain adjoint or Lagrange multiplier approach, which originated from Pontryagin's minimum principle, estimates the ocean state in an iterative way by changing model parameters, using observations that are distributed in time (e.g., Sasaki 1970, Talagrant & Courtier 1987, Thacker & Long 1988). This method is based on the assumption that model equations are correct (sometimes referred to as strong-constraint formalism). It can deal with weakly nonlinear problems but might fail for turbulent (i.e., highly nonlinear) systems (Tanguay et al. 1995).

Bennett (1985) revised the 4D-VAR approach by introducing a weak-constraint formalism that allows departures from model dynamics while obtaining an objective state estimate. The so-called representer method, which is one algorithm for solving the weakly constrained 4D-VAR problem, seeks the solution in the observation space (e.g., Bennett 2002). However, for large observational data sets, it can represent an even larger computational demand above the already computationally demanding strong-constraint adjoint formulation. Hybrid ensemble-variational methods have been devised that aim to combine the strengths of variational and ensemble methods in sequential DA (Hamill & Snyder 2000). Variational methods have algorithmic advantages for solving the analysis problem and for including complex analytical constraints, whereas the sequential ensemble methods provide an appropriate statistical mechanism for generating flow-dependent estimates of the background error covariances.

3. STATUS OF OCEAN DATA ASSIMILATION

3.1. Existing Ocean Syntheses

The first pilot large-scale attempts to use OSE to estimate time-varying ocean states took place in the 1990s (Fukumori et al. 1993, Stammer et al. 1997), which was also when the first multiyear ODA products in support of seasonal forecasts were created (Derber & Rosati 1989, Ji et al. 1995). Since then, expanding technical capabilities have led to a demand for more sophistication, leading to higher spatial resolution and longer estimation periods but also to more complex applications, including biogeochemical investigations. Today, several global synthesis systems exist that are being used across several research and operational institutions to support a variety of applications. **Table 1** summarizes these existing global ODA and OSE efforts, which differ in their goals and assimilation methods, data used, formulation of constraints, model numerics and resolution, surface boundary conditions (forcing), uncertainty estimates, and assimilation window size. Short-term operational ocean analysis involves timescales of days to weeks, requires high spatial resolution, and is produced in quasi-real time; climate-oriented state estimation involves monthly to decadal timescales. By contrast, initialization of monthly and seasonal forecasts involves long timescales but has the operational constraint of prompt real-time delivery.

Based on selected examples, the following sections review the status of ocean synthesis separately for climate and (operational) high-resolution applications. The results of various ocean synthesis efforts can differ substantially because of the specific underlying models and assimilation approaches used by each, and analyzing them indiscriminately might be misleading, as not all related inferences would be equally reliable for the purpose.

Table 1 Existing ocean syntheses, including the names of the systems, their institutions, their intent, the ocean model configurations and forcings, the data assimilation method, and the observations used

System[a]	Institutions[b]	Intent[c]	Configuration[d]	Data assimilation method (observations)[e]
CFSR	NOAA National Centers for Environmental Prediction	INI, ORAM (1979–2010)	1/2° MOM4 coupled DA	3D-VAR (T/SST/SIC)
C-GLORS	Centro Euro-Mediterraneo sui Cambiamenti Climatici	ORA (1993–2010)	1/2° NEMO3.2 forcing EI	3D-VAR (SLA/T/S/SST/SIC)
ECCO-NRT	NASA Jet Propulsion Laboratory	ORAM (1992–present)	1° MITgcm forcing NCEP-R1	KF-FS (SLA/T)
ECCO	NASA Jet Propulsion Laboratory, MIT, and AER	State estimation (1992–2011)	1° × 1/3° MITgcm forcing EI	Adjoint (SLA/SSH/T/S/SST/SIC)
GECCO	Universität Hamburg	State estimation (1948–2014)	1° × 1/3° MITgcm forcing NCEP-R1	Adjoint (SLA/T/S/MDT/SST)
ECDA	NOAA Geophysical Fluid Dynamics Laboratory	INI, ORAM (v3.1, 1961–2012)	1/3° MOM4 coupled DA	EnKF (T/S/SST)
GloSea	UK Met Office	INI, ORAM (1996–2009)	1/4° NEMO3.2 forcing EI/UKMO NWP	3D-VAR (SLA/T/S/SST/SIC)
MERRA Ocean	NASA Global Modeling and Assimilation Office	INI, ORAM (1960–present)	1/2° MOM4 forcing MERRA	EnOI (SLA/T/S/SST/SIC)
GODAS	NOAA National Centers for Environmental Prediction	INI, ORAM (1979–present)	1° × 1/3° MOM3 forcing NCEP-R2	3D-VAR (T/SST)
GLORYS	Mercator Ocean (CNRS, Ifremer, IRD, Meteo-France, and SHOM)	ORA (1993–present)	1/4° NEMO3.1 forcing EI	KF + 3D-VAR (SLA/T/S/SST/SIC)
K7-ODA (ESTOC)	JAMSTEC Research and Development Center for Global Change	ORA (1957–2011)	1° MOM3 forcing NCEP-R1	4D-VAR (SLA/T/S/SST)
K7-CDA	JAMSTEC Center for Earth Information Science and Technology	ORA (1980–2006)	1° MOM3 coupled DA	4D-VAR (SLA/SST)
PEODAS	Centre for Australian Weather and Climate Research	INI, ORAM (1982–2006)	1° × 2° MOM2 forcing E40-NCEP-R2-NWP	EnKF (T/S/SST)
ORAS	European Centre for Medium-Range Weather Forecasts	INI, ORAM (s4, 1958–2014)	1° NEMO3 forcing E40-EI-NWP	3D-VAR (SLA/T/S/SST)
MOVE-C	JMA Meteorological Research Institute	ORA (1950–2011)	1° MRI.COM2 coupled DA	3D-VAR (SLA/T/S/SST)
MOVE-G	JMA Meteorological Research Institute	INI, ORAM	0.5° × 1° MRI.COM3 forcing JRA-55/NWP	3D-VAR (SLA/T/S/SST)
MOVE-CORE	JMA Meteorological Research Institute	ORA (1948–2007)	0.5° × 1° MRI.COM3 forcing CORE.2	3D-VAR (T/S)

SODA	University of Maryland and Texas A&M University	ORA (1997–2004)	1/4° POP2.1 E40-EI	OI (T/S/SST)
UR025.4	University of Reading	ORA (1989–2010)	1/4° NEMO3.2 forcing EI	OI (SLA/T/S/SST/SIC)
FOAM	UK Met Office	MF (January 2007–August 2009)	1/4° NEMO UKMO NWP	3D-VAR (SLA/T/S/SST/SIC)
Bluelink	Australian Bureau of Meteorology	MF (1992–2006)	1/10° MOM4 BoM NWP	EnOI (SLA/T/S/SST)
GOFS	Naval Oceanographic Office	MF	1/12° HYCOM NWP	3D-VAR (SLA/T/S/SST/SIC)
CONCEPTS	Environment Canada	MF	1/4° NEMO NWP	SEEK (SLA/T/S/SST/SIC)
MERCATOR	Mercator Ocean (CNRS, Ifremer, IRD, Meteo-France, and SHOM)	MF (1993–present)	1/4° and 1/12° NEMO ECMWF NWP	SEEK (SLA/T/S/SST)

[a]System abbreviations: C-GLORS, Community Coordinated Modeling Center Global Ocean Physical Reanalysis System; CFSR, Climate Forecast System Reanalysis; CONCEPTS, Canadian Operational Network of Coupled Environmental Prediction Systems; ECCO, Estimating the Circulation and Climate of the Ocean; ECDA, ensemble coupled data assimilation; ESTOC, European Ocean Observatory Network; FOAM, Fast Ocean Atmosphere Model; GECCO, German ECCO; GLORYS, Global Ocean Reanalysis and Simulation; GODAS, Global Ocean Data Assimilation System; GOFS, Navy Global Ocean Forecast System; K7-CDA, K7 Coupled Data Assimilation; K7-ODA, K7 Ocean Data Assimilation; MOVE, Multivariate Ocean Variational Estimation; MOVE-C, coupled MOVE; MOVE-G, global MOVE; NRT, near real time; ORAS, Ocean Reanalysis System; PEODAS, Predictive Ocean Atmosphere Model for Australia; SODA, Simple Ocean Data Assimilation; UR025.4. University of Reading 025.4.

[b]Institution abbreviations: AER, Atmospheric and Environmental Research; CNRS, Centre National de la Recherche Scientifique; IRD, Institut de Recherche pour le Développement; JAMSTEC, Japan Agency for Marine-Earth Science and Technology; JMA, Japan Meteorological Agency; MIT, Massachusetts Institute of Technology; NASA, National Aeronautics and Space Administration; NOAA, National Oceanic and Atmospheric Administration; SHOM, Service Hydrographique et Océanographique de la Marine.

[c]These independent efforts employ different methods for different purpose. The intent therefore has been classified as follows: INI, initialization of coupled model forecasts at monthly, seasonal, or decadal timescales; ORA, ocean reanalysis for a finite period of time; ORAM, ocean reanalysis with real-time extension (used for monitoring); MF, marine forecasting.

[d]Configuration abbreviations: BoM, Bureau of Meteorology; CORE, Coordinated Ocean-Ice Reference Experiments; DA, data assimilation; E40, ECMWF reanalysis; ECMWF, European Centre for Medium-Range Weather Forecasting; EI, ECMWF reanalysis interim; HYCOM, Hybrid Coordinate Ocean Model; JRA, Japanese reanalysis; MERRA, Modern Era Retrospective-Analysis for Research; MITgcm, MIT General Circulation Model; MOM, Modular Ocean Model; MRI, Meteorological Research Institute; NCEP, National Centers for Environmental Prediction; NEMO, Nucleus for European Modelling of the Ocean; NWP, Numerical Weather Prediction; POP, Parallel Ocean Program; UKMO, United Kingdom Meteorological Office.

[e]Method abbreviations: 3D-VAR, three-dimensional variational assimilation; 4D-VAR, four-dimensional variational affiliation; EnKF, ensemble Kalman filter; EnOI, ensemble optimal interpolation; KF, Kalman filter; KF-FS, Kalman filter–filter smoother; SEEK, singular evolutive extended Kalman. Observation abbreviations: MDT, mean dynamic topography; S, salinity; SIC, sea ice concentration; SLA, sea level anomaly; SSH, sea surface height; SST, sea surface temperature; T, temperature.

3.2. Climate Applications

Historically, ocean observations are very sparse, making it difficult to extract climate signals in the ocean from the limited observations extending more than a few years into the past. This problem is exacerbated for studies that began several decades ago, before the altimeter and Argo era. Much of the ongoing use of ocean syntheses for climate science is therefore devoted to a quantitative understanding of ocean variability (especially regionally) and its associated uncertainties. Examples include studies of sea level variability and change (e.g., Stammer et al. 2002, 2004; Wunsch et al. 2007; Carton & Giese 2008; Köhl & Stammer 2008; Balmaseda et al. 2013a; Piecuch & Ponte 2014; Storto et al. 2015), water masses (e.g., Fukumori et al. 2004; Wang et al. 2004; Masuda et al. 2006; Toyoda et al. 2011, 2015; Speer & Forget 2013), mixed-layer heat balance (e.g., Kim et al. 2007, Halkides & Lee 2009, Buckley et al. 2015), and changes in ocean heat content (OHC) (Carton & Santorelli 2008, Balmaseda et al. 2013b, Wunsch & Heimbach 2014).

OHC and sea level are important indicators of climate change, and there is hope that ocean syntheses produce simultaneous analyses of both quantities. It appears that the estimation of the global OHC benefits from the combination of observations and models via dynamical constraints provided by the DA system. The results have shown more obvious variations in OHC related to the El Niño–Southern Oscillation (ENSO) than are present in observation-only syntheses, which holds especially before the Argo period. Recent comparisons of ocean reanalyses (Balmaseda et al. 2015, Palmer et al. 2015) suggested that although the upper-ocean heat content is relatively well constrained in the recent period, substantial uncertainty remains in existing estimates of the vertical penetration of heat especially prior to the pre-Argo period. As shown in **Figure 2**, the OHC increase is not monotonic and smooth but rather shows significant variation on all timescales. We expect similar variability to exist in future OHC changes (and in all other climate variables, for that matter). The figure also shows that during the spin-up phase (a few years), all ocean syntheses should be treated with great caution or not used at all.

Although global indicators of climate change in the ocean are important, regional changes are usually of the largest consequence, and therefore these changes are of major interest. In this context, ocean syntheses can provide valuable estimates of climate-relevant indices or quantities not easily assessable from data alone. A quantity of considerable concern is regional sea level and its variability, which integrates many individual aspects of the ocean state and the climate system at large. Changes in sea level can potentially have a substantial impact on society; understanding ongoing and past changes as well as their regional character is therefore of specific importance. Storto et al. (2015) compared linear trends in steric height over the period 1993–2010 from different ocean syntheses and found that large variations exist among individual products on the regional scale, largely arising from uncertainties in the deep ocean and discrepancies in the halosteric component.

The Atlantic meridional overturning circulation (AMOC), a measure of zonally and vertically integrated poleward volume transports, is another important climate index because it is associated with poleward heat and freshwater transports that play an important role in the coupled climate system (Wunsch & Heimbach 2006, Cunningham et al. 2007). Major challenges remain in the use of ocean syntheses for accurate inferences of the AMOC. Karspeck et al. (2015) investigated the variability and trends in several multidecadal ocean synthesis products. As an example, **Figure 3*a*** documents the diversity of the solutions in terms of the 1960–2007 time-mean AMOC stream function in depth/latitude space. The structural AMOC features are broadly similar, with net northward flow above a depth of approximately 1,000 m and southward flow below this level. However, all products except GECCO2 have more than one distinct positive maximum at different

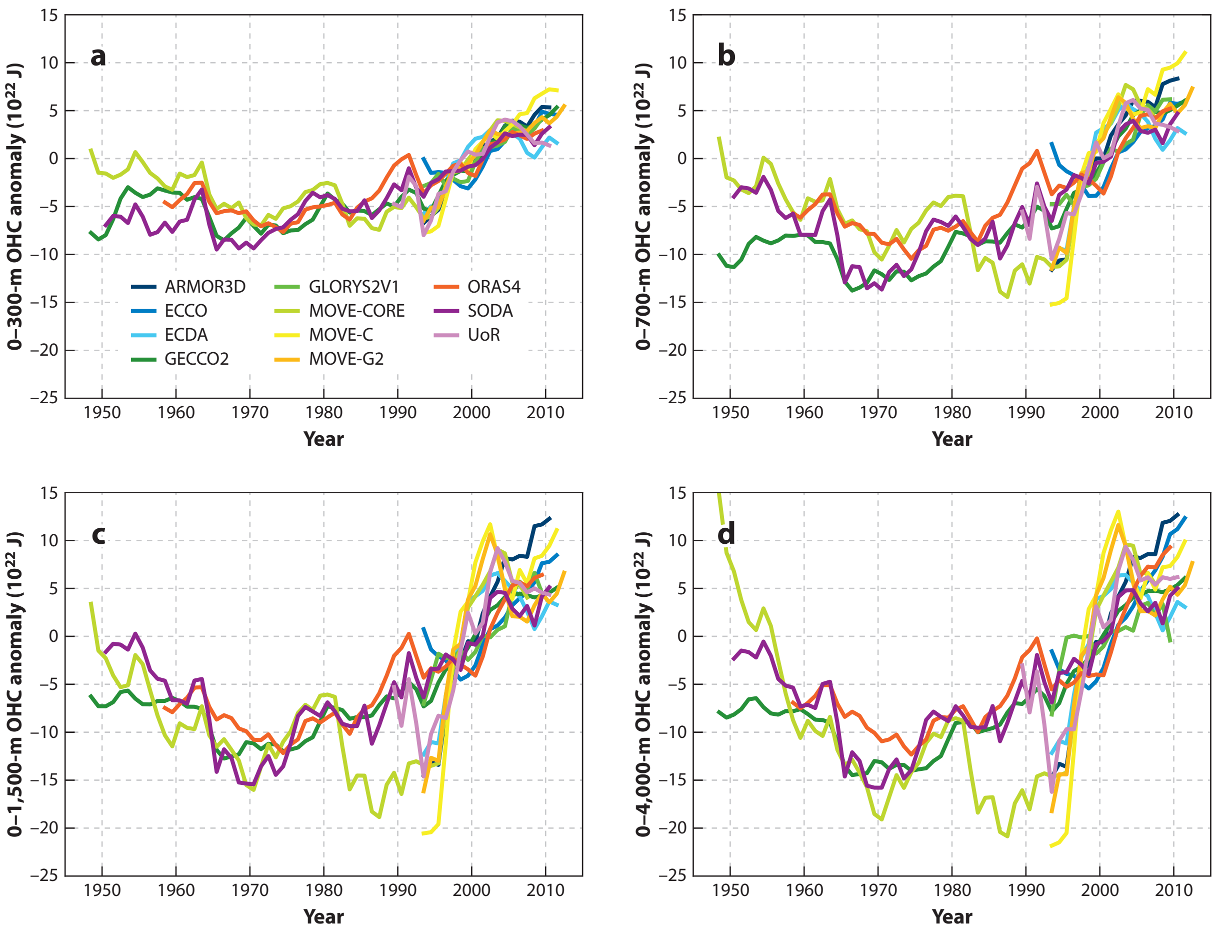

Figure 2

Estimated ocean heat content (OHC) from several ocean syntheses at different depth ranges: (*a*) 0–300 m, (*b*) 0–700 m, (*c*) 0–1,500 m, and (*d*) 0–4,000 m. For an explanation of the synthesis products, see **Table 1**. Figure provided by Matthew Palmer and modified with permission from Balmaseda et al. (2015).

latitudes, with DEPRESYS, SODA, and MOVE-CORE showing localized circulations near the equator. Even though all reanalysis products were constrained by roughly the same in situ data sets, there are substantial differences in the strength and meridional structures, with some showing opposite trends over significant periods. **Figure 3*b*** compares time series of the AMOC anomaly at 1,000-m depth at 45°N and 26.5°N. Visual inspection suggests very little agreement in the year-to-year changes and trends in the synthesis set, implying that even in relatively well-observed areas like the North Atlantic, the different ocean syntheses fail to provide a consistent estimate of AMOC variability, but instead might be strongly influenced by the assimilation approaches and/or the underlying models, including differences in forcing.

Using various ocean syntheses, Toyoda et al. (2015) investigated seasonal-to-decadal variations of mixed-layer depth in the Pacific. The authors found two coherent dominant modes of variability, one related to changes in the Pacific Decadal Oscillation and one suggesting the existence of a

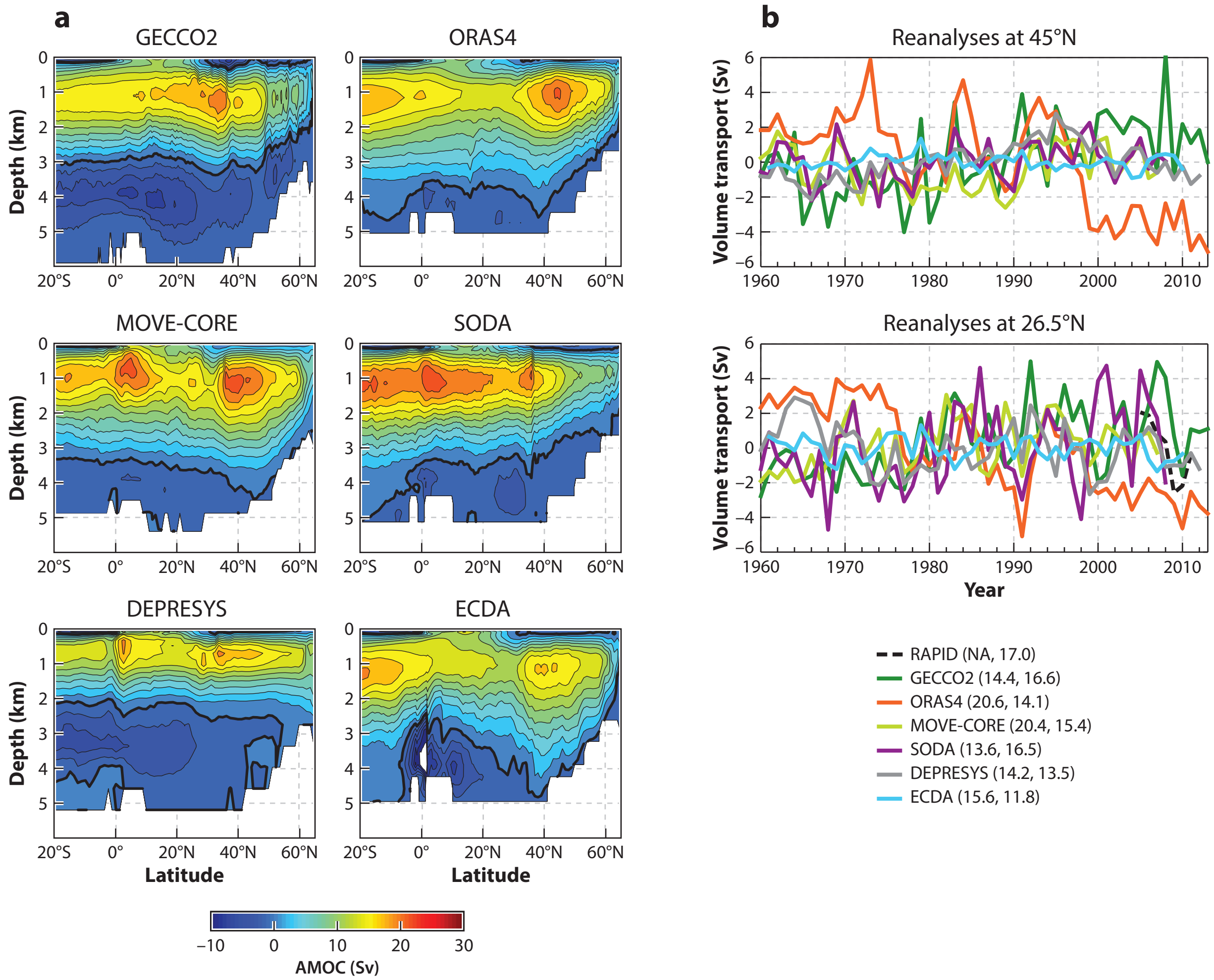

Figure 3

(*a*) Time-mean Atlantic meridional overturning circulation (AMOC) stream functions from 1960 to 2007 in depth/latitude space for a set of ocean syntheses. Positive and negative contours indicate clockwise and counterclockwise circulations, respectively. The bold lines are the zero contours, and the contour interval is 2 sverdrups (Sv). (*b*) Time series of the AMOC anomaly at 1,000-m depth at 45°N (*top*) and 26.5°N (*bottom*) for the same set of ocean syntheses. The time mean has been removed from each time series. The key shows the means (in sverdrups) at 45°N and 26.5°N, respectively, in parentheses. Time series from RAPID are included for comparison. Additional abbreviation: NA, not applicable. Modified with permission from Karspeck et al. (2015).

coupled mode between mixed-layer-induced anomalies in sea surface temperature (SST) and variations in atmospheric sea level pressure related to the West Pacific Index. Taking advantage of the property conservation of state estimates, Buckley et al. (2015) attributed SST and upper-ocean heat content changes in ocean syntheses to local buoyancy as opposed to wind forcing and to processes involving ocean dynamics (advection as opposed to sub-grid-scale mixing). The transient nature of the ocean circulation and its long-term memory also imply that vertical exchanges with the ocean interior, whose proper accounting requires closed property budgets, may play an important role in near-surface thermal property changes (Liang et al. 2015).

3.3. Dealing with Uncertainties

In practice, computing uncertainty estimates for ocean syntheses remains challenging because of the large dimension of the state vector in ODA. The theoretical estimate of the posterior or analysis error covariance matrix can be used to quantify uncertainty. In the KF, the solution algorithm requires the analysis error covariance matrix; in the adjoint method, the inverse of the Hessian matrix (the inverse of the matrix of second derivatives of the cost function) approximates the analysis error covariance matrix but is not directly computed as part of the solution algorithm. Nevertheless, useful information about the Hessian matrix can be diagnosed, albeit at a computational cost. For example, the eigenpairs associated with the extreme eigenvalues provide information about the combinations of parameters that are best and least well determined by the observations. The use of Hessian information to infer posterior error covariances is being explored within limited-domain GCM applications (e.g., Sapsis & Lermusiaux 2009, Moore et al. 2011, Kalmikov & Heimbach 2014). In most cases, however, strict use of the theory has been limited to applications that estimate only a few parameters or those in which a limited number of observations effectively constrain the problem.

Using ensembles of reanalyses from the same system (Balmaseda et al. 2013a) or multiple systems (Stammer et al. 2010, Karspeck et al. 2015) is another way to assess uncertainty in reanalyses. The ensemble spread among ocean syntheses is frequently used as a measure of the uncertainty (e.g., Corre et al. 2012). However, this measure does not quantify whether ocean syntheses have common biases or other limitations that would give the appearance of artificial consistency. Nevertheless, a recent intercomparison by Balmaseda et al. (2015) has shown that the ensemble mean is usually a better estimate than any individual ocean reanalysis, although there are exceptions where a subset of best products is better than the grand ensemble. Their work also identified specific geographical areas where the uncertainty is large, thus providing a focus for future developments in the observing system, modeling, or DA method. The global ocean below the top few hundred meters, the Southern Ocean (Antarctic Circumpolar Current region), coastal areas, and the paths of western boundary currents stand out as the areas with the largest uncertainty in the density, temperature, and salinity fields.

3.4. High-Resolution Applications

High-resolution ocean syntheses can provide important first-order insights into basin-scale ocean current systems (e.g., Maximenko et al. 2008, Divakaran et al. 2010) as well as initial conditions for short-term, high-resolution ocean forecasting. However, progress has been hindered because ODA methods fundamentally rely on linearized model dynamics. Techniques such as the EnKF (Evensen 1994) and approximate adjoint models (Köhl & Willebrand 2002, Hoteit et al. 2005) were developed to deal with exponential error growth associated with nonlinear dynamics. Besides these technical and scientific problems, the extra cost involved in performing the assimilation step has so far limited global ocean syntheses for extended time periods to resolutions of 1/4° (**Table 1**). Nevertheless, on a regional to basin-wide scale, applications of much higher resolution exist (Edwards et al. 2014, Martin et al. 2015).

Examples for the European marginal seas, the North Atlantic, and global applications include MyOcean (**http://www.myocean.eu**) and the Copernicus program (**http://www.copernicus.eu**). Similar projects exist as part of the US Integrated Ocean Observing System and Australia's Bluelink analysis and forecasting system (**http://wp.csiro.au/bluelink**). Within these projects, national centers have developed high-resolution systems that operate on regional and global scales and have fostered the development and improvement of operational ocean analysis and forecast systems worldwide. Most of these systems assimilate real-time observations, and more than half provide daily short-term forecasts. By way of example, **Figure 4** compares SST measured by a Moderate Resolution Imaging Spectroradiometer (MODIS) with the results of a 1/36° version of the operational Mercator Ocean analysis and forecast system, demonstrating the amount of detail current systems resolve. An example of high-resolution state estimates for climate science is the Southern Ocean State Estimate (Mazloff et al. 2010), with various applications for the Southern Ocean now being published (**http://sose.ucsd.edu**).

3.5. Adjoint Sensitivity Studies

Beyond performing state estimation, an adjoint model is valuable for estimating uncertain model parameters and for performing climate sensitivity studies in order to understand climate dynamics and optimize the observing system. All of these fundamental applications are based on the fact that the adjoint model provides an efficient means to compute the derivative of scalar-valued functions with respect to a large number of parameters. The atmospheric community realized early on (e.g., Hall 1986) that this gradient—also called adjoint sensitivity—provides a comprehensive tool to explore model sensitivities to parameters; however, the ocean modeling community long ignored adjoint sensitivities. Only in recent years, with the availability of adjoint codes for full realistic ocean GCMs (e.g., Marotzke et al. 1999), has sensitivity analysis become popular (e.g., Galanti & Tziperman 2003, Stammer et al. 2008, Masuda et al. 2010).

In contrast to conventional sensitivity calculations via perturbation sensitivities, which infer how the climate system responds to changes to individual parameters, adjoint sensitivities reveal how a specific scalar-valued target quantity of interest (e.g., climate index) is affected by many different kinds of model parameters. As an example, the sensitivities calculated in **Figure 5** show that Rossby waves traveling in the baroclinically unstable region of the subtropical gyres are most relevant for affecting the equatorial temperature because perturbations are amplified in these regions. Fukumori et al. (2007), Czeschel et al. (2010), and Heimbach et al. (2011) further explored the use of adjoint sensitivities to reconstruct the full circulation using known perturbations (either time varying or mean) and to explain mechanistic causes in terms of dominant perturbations. Over the years, the use of adjoint sensitivities to ocean circulation has been extended to ocean biogeochemical processes (Dutkiewicz et al. 2006), coupled ocean–sea ice processes in the Arctic component (e.g., Kauker et al. 2009), and melt rates in sub-ice-shelf cavities (Heimbach & Losch 2012).

A variant of the sensitivity analysis, the optimal observations defined by Köhl & Stammer (2004), combines classical and adjoint-derived sensitivities to estimate distributions of observations that are optimally suited for their use in variational DA. As such, this technique explores the relation of an event (e.g., anomalous overturning at a certain time and place) to the past and future changes in the ocean. Köhl (2005) used this technique to describe mechanisms that affect the overturning variability in the Atlantic.

4. INITIALIZING FORECASTS

An important motivation for ODA has long been to provide initial conditions for seasonal-to-interannual (SI) forecast systems. SI forecasting is concerned with atmospheric circulation changes up to a few months ahead of time in response to anomalous boundary forcing, which can significantly change the probability of occurrence of specific weather patterns (Palmer & Anderson 1994). Although not their initial motivation, climate prediction on seasonal to decadal and longer timescales has steadily become a central focus of several synthesis efforts, largely fostered by the

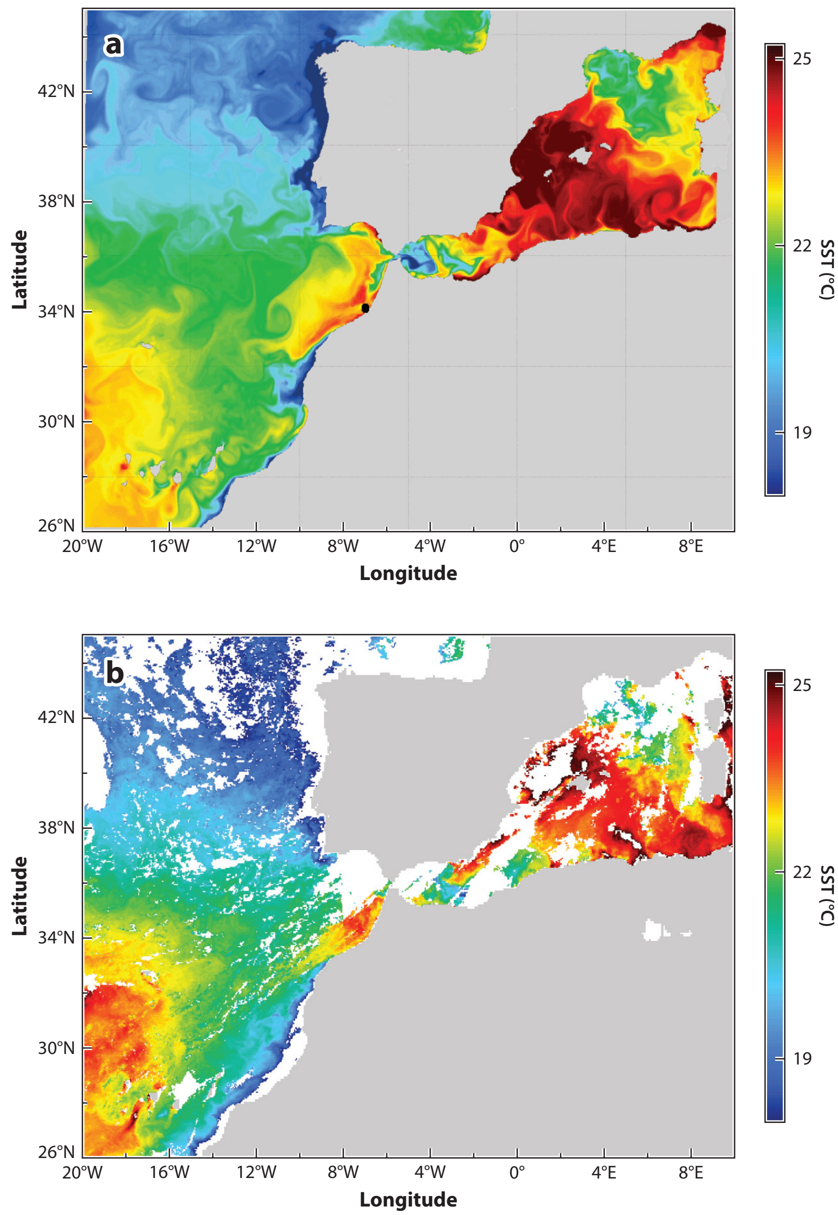

Figure 4

Sea surface temperature (SST) on July 19, 2014, from (*a*) Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data (obtained from **http://podaac-ftp.jpl.nasa.gov**) and (*b*) the corresponding four-day forecast from the 1/36° Daily Iberian Biscay Irish Physical Bulletin created by Mercator Ocean (obtained from **http://bulletin.mercator-ocean.fr**).

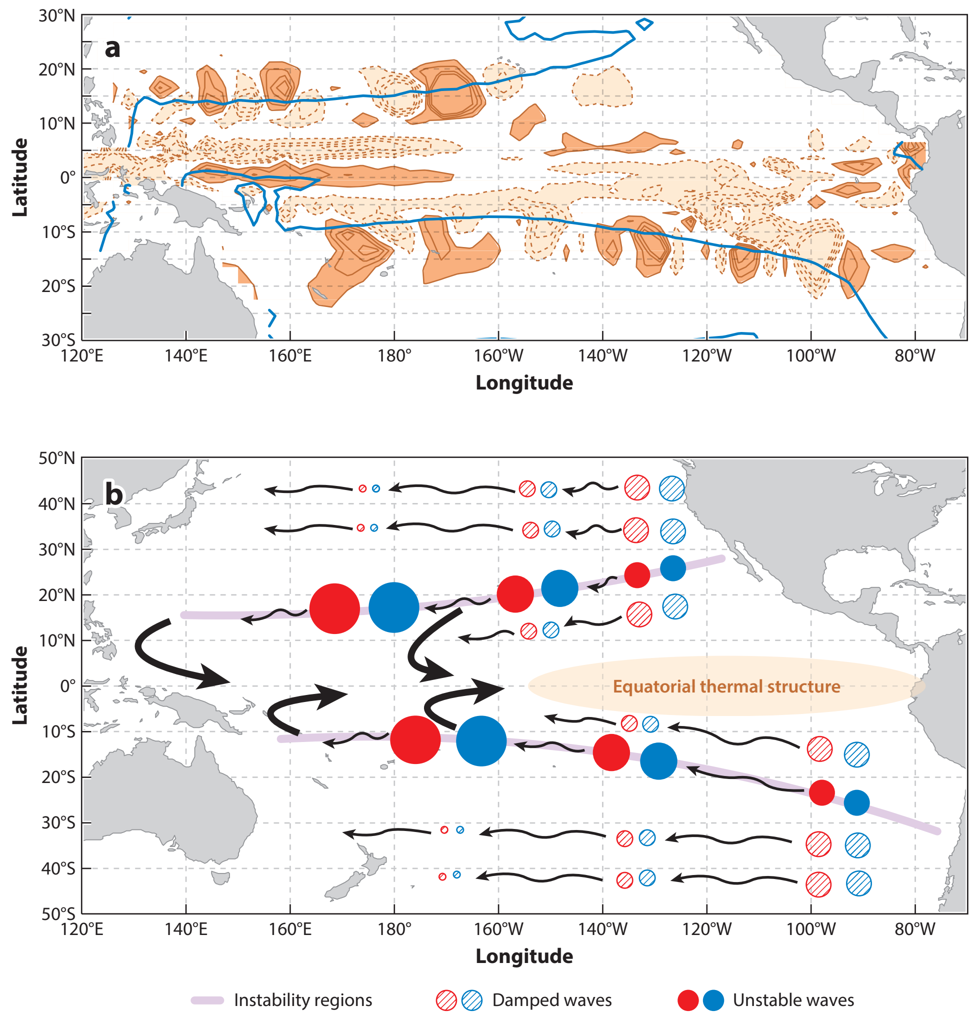

Figure 5

(*a*) Sensitivity to temperature perturbations at a depth of 200 m four years before the cost function evaluation, which is the near-surface temperature at 100°W, 0°N. Values above 0.005 are shaded with dark orange, and values below −20.005 are shaded with light orange. The thick blue line denotes the 16.8°C isotherm. (*b*) Schematic of the mechanism for a wave teleconnection from the midlatitude Pacific to the equator. Midlatitude planetary Rossby waves travel westward at all latitudes and are damped except for those amplified in baroclinically unstable regions of the subtropics. Modified with permission from Galanti & Tziperman (2003).

World Climate Research Programme and its CLIVAR (Climate and Ocean: Variability, Predictability, and Change) core project.

4.1. Seasonal-to-Interannual Forecast Applications

Several operational centers worldwide provide seasonal forecasts initialized with ocean and atmospheric analyses (Balmaseda et al. 2010). The initialization of the ocean subsurface is key for successful predictions of SST at seasonal timescales. Of special importance is the proper representation of tropical SST variations associated with ENSO, which have the potential to alter the large-scale atmospheric circulation associated with tropical convective cells. Using information from SST, surface fluxes from atmospheric reanalyses, subsurface temperature and salinity, and altimeter-derived sea level anomalies is instrumental to initialize the upper-ocean thermal structure, thereby reducing the large uncertainty (error) resulting from the forcing fluxes and improving forecast skill (Alves et al. 2004, Balmaseda et al. 2010). SI forecasting systems are based on coupled atmosphere-ocean GCMs that predict both the surface boundary forcing and their impact on the atmospheric circulation, and require near-real-time knowledge of the state of the climate. The chaotic nature of the atmosphere is taken into account by issuing probabilistic forecasts from an ensemble of coupled integrations. To cope with deficiencies in coupled models, the forecasts need calibration before the forecast is issued. The calibration is performed by conducting a series of past seasonal hindcasts starting from synthesis-based initial conditions for a historical period (a few decades); these hindcasts are also needed for skill assessments. The realism of their interannual variability determines the forecast quality.

The most common SI initialization strategy is the so-called full-state initialization, where the DA corrects the ocean model time-mean state as well as the variability. In the presence of model biases, changes in the observing system can lead to spurious variability in the ocean estimate. Thus, consistent ocean reanalysis requires an explicit treatment of the model bias during the initialization procedure (Balmaseda et al. 2007). The model bias estimation obtained during the initialization procedure could in principle be used to correct model errors during the forecasts. This is not yet possible when the full initialization is conducted in uncoupled mode, which is the common practice. The separate initialization of the ocean and atmosphere systems can also lead to initialization shock during the forecasts. An alternative approach is the so-called anomaly initialization, first introduced to initialize decadal forecasts (Pierce et al. 2004), in which the observations are used only to estimate the anomalous state (Smith et al. 2007). This approach reduces the initialization shock but leads to a biased mean state. **Figure 6** shows that although the initialization shock is larger with the full field initialization (**Figure 6*a***), being far from the real world is detrimental for the forecast skill (**Figure 6*b***). The best skill is obtained by using empirical corrections of model error, which reduces the initialization shock and decreases model drift. A more balanced coupled initialization is presumably desirable but remains challenging.

4.2. Decadal and Long-Term Climate Forecast Systems

Early applications of ocean syntheses in the context of decadal prediction include those by Smith et al. (2007), Keenlyside et al. (2008), and Pohlmann et al. (2009). The predictive skill of such a system is usually tested, and initialization techniques are optimized in hindcasts that aim to successfully predict the past, assuming that forecasts with the same system of the future will be skillful. This can be misleading because of errors in the climate sensitivity of the model, e.g., in the case of a major volcanic eruption, when different strategies are required to model the response (e.g., Driscoll et al. 2012, Zanchettin et al. 2013). Nevertheless, initial decadal prediction efforts in recent years have shown predictive skill in global average temperature up to a decade in advance from both initial conditions and the climate change signal related to the known emission of greenhouse gases.

Today, initialized multimodel ensembles exist that suggest that some aspects of decadal variability—such as the mid-1970s shift in the Pacific, the mid-1990s shift in the western Pacific, and the early-2000s hiatus—are better represented by initialized hindcasts than by noninitialized simulations. Many recent decadal prediction studies find enhanced predictive skill notably in the North Atlantic region associated with AMOC variability and predictability (Meehl et al. 2014).

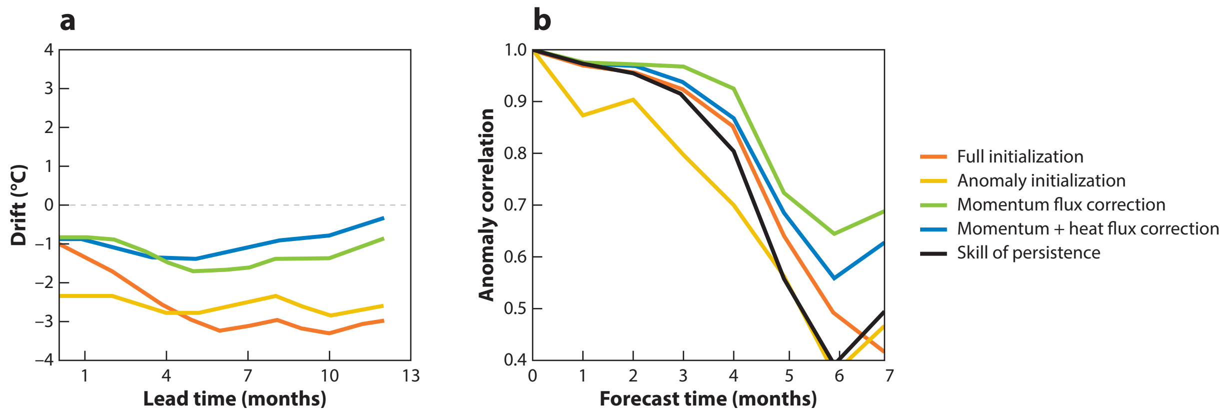

Figure 6

Forecast drift in (*a*) sea surface temperature and (*b*) skill in precipitation in the central Pacific from different forecast strategies: full initialization (*orange*), anomaly initialization (*yellow*), momentum flux correction (*green*), and momentum plus heat flux correction (*blue*). The momentum flux correction exhibits the best skill. Modified with permission from Magnusson et al. (2013).

However, it remains unclear how errors in the ocean initial state affect the predictive skill of the forecast and what the impact is of the initialization of different aspects of the climate system, such as sea ice extent, soil moisture, snow cover, and the state of surface vegetation over land, on timescales of seasons to a year and longer. A key difference between initialized decadal predictions and initialized predictions on shorter timescales is the need for observations in the deeper ocean (below 500 m); even observations below 2,000 m are likely to play a significant role, e.g., in the prediction of the AMOC (Zanna et al. 2012). The ocean syntheses used to initialize, calibrate, and verify decadal forecasts should span longer time records (several decades) and should attempt to initialize the process relevant at decadal timescales; for example, initializing large-scale modes of decadal variability (such as the Pacific Decadal Oscillation) may be important. This is a real challenge for current DA systems.

Until fully coupled DA approaches are developed, dynamical forecasting systems will rely on separate assimilation approaches in the ocean and initialization methods for the coupled system. In the past, anomaly initialization was therefore more frequently used in decadal forecasts, but it shows weaker performance than the full initialization that is currently favored, especially on seasonal timescales. Decadal forecasting is a rapidly evolving field (Meehl et al. 2014) that now also includes full-field initialization and even flux corrections (Magnusson et al. 2013, Polkova et al. 2014).

Understanding which perturbations have the largest impact on uncertainty growth in chosen forecast norms or indices, and therefore understanding limits to predictability, has become a well-developed branch of numerical weather prediction (e.g., Buizza & Palmer 1995). The use of singular vectors, which characterize optimal perturbation and error growth and which can be computed using an adjoint model, has been adopted by the oceanographic community for ENSO prediction studies (e.g., Penland & Sardeshmukh 1995). Zanna et al. (2011) showed that predictability studies using optimal perturbation techniques reveal ocean dynamical mechanisms that can limit predictability horizons of climate indices such as the AMOC (**Figure 7**). The important implications for prediction are that (*a*) ensemble generation mechanisms need to include

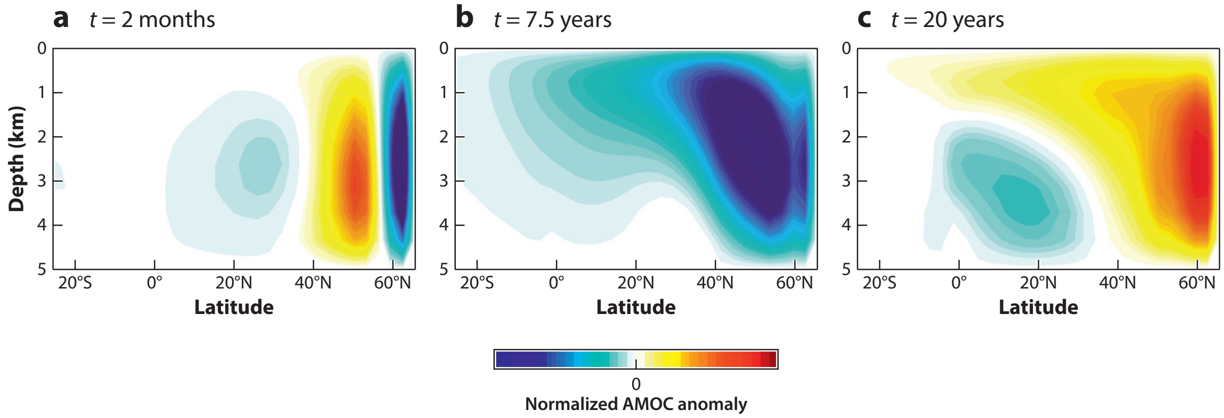

Figure 7

Latitude-depth section of time-mean Atlantic meridional overturning circulation (AMOC) anomalies at (*a*) $t = 2$ months, (*b*) $t = 7.5$ years, and (*c*) $t = 20$ years. The results represent normalized anomalies obtained from an idealized Atlantic-like rectangular basin model. Modified with permission from Zanna et al. (2011).

perturbations of the initial state of the ocean (not just the atmosphere) and (*b*) ocean observations that reach significant depths are needed in order to constrain prediction models.

5. OUTLOOK: THE WAY FORWARD

With an ever-increasing diversity and heterogeneity of ocean observations, increasingly including biogeochemical and biological parameters, we expect that over the next decade, ocean synthesis will become an essential part of the infrastructure of ocean and climate service activities and will provide ocean information on a regular basis for many applications. In particular, we envision that ocean syntheses will be used increasingly by other disciplines, e.g., in carbon or nutrient cycle studies or to investigate the dependence of biodiversity on the physical climate state. Further increasing the value of ocean synthesis products for all of these applications will require characterizing the uncertainties in each product, improving the products by including better or more observations as constraints, improving the models, and advancing assimilation approaches.

We further expect ODA to become an integral part of a seamless climate prediction system that includes seasonal, interannual, and decadal timescales, allowing investigation of multiscale interactions. The best forecasts will likely be produced by coupled models that are directly constrained by climate data [i.e., coupled data assimilation (CDA)]. Ultimately, every ocean or coupled synthesis should be accompanied by formal uncertainty measures provided on a geographic grid for any estimated parameter. All of these aspects are cutting-edge research topics that we cannot address in detail owing to space limitations, but we provide some brief thoughts below.

5.1. Improved Uncertainty Measures

Given the large remaining differences between individual ocean syntheses, one important step forward will be to provide ensemble mean estimates and their uncertainties, akin to what is now common practice in numerical weather prediction. However, in the absence of formal posterior error covariance information accompanying the solutions, understanding the mutual consistency among the products and with observations remains difficult. Several steps are involved, most of which are not included in existing measures. Much effort is required to compute realistic uncertainty measures for any practical problem, which involves the specification of prior error information as well as the computation of the a posteriori error covariance for any solution.

In a first step, suitable specification of error covariances (data, background, and model error) is essential to obtain sensible solutions (Fukumori 2002). In reality, several large-scale applications so far resort to simplified expressions of the error covariance operators (e.g., Forget & Wunsch 2007, Ponte et al. 2007). Weaver et al. (2005) implemented a balance operator for large-scale global ODA, which they used to implicitly specify the multivariate component of the background error covariances. The basic technique employs a transformation from the model space, where variables are highly correlated, to a control space, where variables can be considered to be approximately uncorrelated. Balance operators need to be regularly reassessed in response to changes in model resolution and complexity. System bias is another serious obstacle to the reliable representation of climate variability, especially in the realistic case of a time-dependent observing system (e.g., Segschneider et al. 2000). To help suppress artificial variability in the analyses, Balmaseda et al. (2007) implemented a generalized algorithm to treat bias in sequential DA.

Any ocean state estimate should also be associated with an estimate of its error covariance matrix. However, the computation of what amounts to a very-high-dimensional (typically on the order of $10^9 \times 10^9$ or higher) covariance matrix is impractical. Approximate approaches or projection methods onto low-dimensional (scalar) climate indices or quantities of interest are required; existing approximations inferring leading eigenvectors of the posterior error covariance matrix are a promising approach for capturing at least the dominant uncertainty structures (Moore et al. 2011, Kalmikov & Heimbach 2014). In connection with singular vector approaches, these methods could also reveal what observations (types and spatial distribution) would have the most impact on estimation and forecasting.

5.2. Coupled Data Assimilation

Coupled data assimilation can have various degrees of complexity. Common practice now is to use coupled ocean–sea ice models; less developed is the use of full Earth system models.

5.2.1. Coupled ocean–sea ice estimates. The polar regions have received heightened attention in the last decade, in particular the rapid decline in Arctic sea ice cover since the late 1970s (e.g., Meier et al. 2014) and the polar amplification of near-surface temperature changes. The difficulty in determining the ocean's role in these processes is exacerbated by the extreme lack of quasi-continuous observations, in particular of hydrographic changes in the high Arctic and of ice thicknesses that are thought to carry some memory of climate variability. Sea ice models used for assimilation need to produce skillful simulations of thermodynamic and dynamic processes of ice growth, evolution, and melt; sea ice modeling is a rapidly evolving field (e.g., Feltham 2008, Hunke et al. 2010).

Several sea ice DA systems are now available. Sequential systems have been initially targeted at assimilating remotely sensed sea ice concentrations and velocities (e.g., Bertino & Lisaeter 2008, Caya et al. 2010). Adjoint-based coupled ocean–sea ice assimilation has produced initial one-year ocean–sea ice state estimates in a regional domain of the Labrador Sea and Baffin Bay (Fenty & Heimbach 2013a). The dynamical consistency of the state estimates, in turn, has enabled a detailed analysis of what sets maximum winter sea ice extent in that region, the crucial role of ocean dynamics in setting this extent, and the implications for seasonal ice extent predictability (Fenty & Heimbach 2013b). The coupled estimation system is currently being extended to develop a decadal state estimate for the Arctic/North Atlantic domain. Hybrid systems are also being explored with filter and smoother approaches interlaced for the sea ice and ocean, respectively (Panteleev et al. 2010).

5.2.2. Coupled Earth system estimates. Performing initialization in uncoupled mode, as is common practice in most existing climate forecast efforts, leads to initialization shock in the coupled system, which potentially reduces the forecast skill. This suggests that CDA efforts in Earth system models will lead to improved use of ocean information for coupled forecasts ranging from near-term to seasonal and decadal timescales. Coupled Earth system models link modules of the ocean, atmosphere, sea ice, land surface, global carbon cycle and chemistry, and aerosols to simulate changes in the Earth's climate systems.

Some pilot applications of CDA already exist (e.g., Zhang et al. 2007, Sugiura et al. 2008, Fujii et al. 2009, Laloyaux et al. 2015), and several others are spinning up (e.g., Blessing et al. 2014). As an example, Zhang et al. (2007) have applied an EnKF approach to an atmosphere-ocean CDA system with a fully coupled GCM using a super-parallelization technique for ensemble integrations. In perfect model experiments, the assimilation successfully reconstructs the twentieth-century OHC variability and trends in most locations.

Also using an EnKF, Karspeck et al. (2014) applied CDA to the problem of decadal predictions; however, the results were mixed, partly because a state estimated with a coupled EnKF remains dynamically inconsistent with the coupled system if the model parameters are not also improved, as can be done using a smoother-based approach. This technique was pioneered by Sugiura et al. (2008), who took up the challenge of developing a sophisticated CDA system with a fully coupled GCM, using the adjoint method to adjust both the oceanic initial conditions and the drag (coupling) coefficients associated with mass, momentum, and heat exchange at the atmosphere-ocean interface. Their products thus provide dynamically self-consistent coupled fields that are suitable for the initial states in SI prediction experiments. One of the most fascinating elements of their approach is that it filters out chaotic fluctuations that take place on the timescales of weather modes by operating an averaging procedure in order to highlight the representation and forecast of SI variations. In comparison with a hindcast using the same model initialized from ocean-only assimilation, their coupled assimilation demonstrated higher predictive skill, which directly demonstrates the benefit of coupled assimilation. However, those results might not hold in general (Laloyaux et al. 2015).

CDA efforts have now been embraced by several operational centers, instigated by the weakly coupled reanalyses at the National Centers for Environmental Prediction (Saha et al. 2010). For example, the European Centre for Medium-Range Weather Forecasts (ECMWF) has implemented a pilot CDA system for production of coupled reanalyses of the Earth system, known as the Coupled ECMWF Reanalysis (CERA) system (Laloyaux et al. 2015), that is capable of assimilating a wide variety of atmospheric and oceanic observations and produces analyzed states that are consistent with the coupled model at the atmosphere-ocean interface. Compared with an equivalent uncoupled system, CERA shows overall consistency, with slightly improved temperature estimates in the upper ocean and the tropical atmosphere. On a cautionary note, however, a fully coupled GCM inevitably generates rapidly growing modes, particularly in the atmospheric component, which makes it difficult to optimize the simulated state of the atmosphere. The actual coupled phenomena are thought to include a controllable dynamical nature in SI processes because they should contain low-frequency modes generated and controlled by oceanic processes (Palmer et al. 2005). It is therefore possible that CDA could allow better determination of these modes in the coupled system. Forecasts from a single forecasting system would still not be reliable enough, and ensemble generation techniques that sample model uncertainty (multimodel ensemble) are required.

5.3. Model Improvements

ODA procedures require the best possible model representations to maximize performance. In turn, state estimation can contribute substantially to improving models and therefore needs to be tightly coupled to model development and improvement efforts. This holds for CDA as well as ODA. Several avenues are conceivable for ocean and coupled synthesis efforts aimed at improving ocean and coupled climate models. One is to help improve uncertain model parameters through parameter estimations. This avenue might turn out to be the most important one for climate model DA in a coupled context. For example, in ocean-only applications, the estimation of mixing coefficients was one of the main foci for model improvements (e.g., Ferreira et al. 2005, Menemenlis et al. 2005, Stammer 2005, Liu et al. 2012). However, although the estimation of mixing coefficients provided interesting insight into the physical processes, the contribution to improvement of ocean models remained small [e.g., Liu et al. (2012) found that it contributed only 10% of the total model-data misfit]. Model tuning is less crucial for ocean models than for coupled models, where automatic tuning is an active field of research, with the first successful pilot systems in place (e.g., Annan et al. 2005, Liu et al. 2014).

The alternative approach could be to relate innovations in sequential approaches to model errors and attempt to correct them. In fact, recent advances in Earth system modeling have been accompanied by progress in CDA (WMO 2009), which uses observations in more than one component of a coupled model (e.g., atmosphere and ocean) so that the whole coupled model is optimized simultaneously and observations in one subcomponent can influence the estimated state in another component.

Finally, just as ocean models used in climate or Earth system models are improved over time, so are those used in DA. Improved numerics (e.g., advection schemes), vertical discretization [e.g., the z, z^*, and arbitrary Lagrangian Eulerian (ALE) approaches], and representation of kinematic boundary conditions (nonlinear free surface with real water fluxes, compared with linear free surface with virtual salt fluxes) are as pertinent as improvements in the DA schemes (e.g., Forget et al. 2015).

5.4. Closing Remarks

For years to come, it will be essential for the community to recognize the value of ocean synthesis and to expand the applications of ocean synthesis products for research and information services alike. Owing to space limitations, in this article we were able to address only a subset of ODA progress and problems, most of which were related to climate applications. The field of operational oceanography and the importance of ocean syntheses for other fields (such as the evolution of the ocean's ecosystems) and for studies of oceanic tracer constituents (including the transports of biogeochemical substances in general—e.g., carbon uptake by the ocean—and pollutants in coastal regions) are topics of equal importance that require separate reviews. Another topic of significant relevance not addressed here is that of optimizing the ocean observing system. Much more can be done in this context using ODA, following similar examples in numerical weather prediction. With fewer observed parameters in the ocean, there is a need for better DA methods to extract more information from observations. Systems have been tuned to extract information about the mesoscale or tropical climate variability, but currently they appear to be mutually exclusive. Work on high-resolution or biological parameters would be more difficult by an order of magnitude.

Problems often overlooked in many fields are those of expertise, continuity, and especially resources required to further develop the fields. The resources required for technical developments in the various aspects of ODA and OSE in support of ocean syntheses are enormous, and comparable to the requirements of atmospheric forecast centers. There is a need for better software infrastructure that would provide openly available algorithmic differentiation and other assimilation tools and would allow testing of different options and methods. For instance, more research is needed on how to combine ensemble and variational methods effectively and how to improve model bias correction techniques. Furthermore, there is a need to share efficient minimization algorithms and observation operators to avoid duplicating efforts. In almost all cases, reaching infrastructure development milestones required close to a decade of sustained consortium efforts (Stammer et al. 2002). This development has proven to be a large endeavor requiring expertise in ocean observations, modeling, assimilation, and information technology; to be effective, the community must sustain such efforts and maintain a long-term perspective.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Carl Wunsch for helpful comments and Meike Demgen for help with some of the figures. D.S. acknowledges the hospitality provided during a stimulating and pleasant research visit to the Climate and Global Dynamics division at the National Center for Atmospheric Research, where part of this review was written. This work was funded in part through the Deutsche Forschungsgemeinschaft (DFG)–funded Integrated Climate System Analysis and Prediction (CliSAP) initiative of the Universität Hamburg, the Bundesministerium für Bildung und Forschung (BMBF)–funded Regional Atlantic Circulation and Global Change (RACE) project, and the European Union–funded North Atlantic Climate (NACLIM) project. P.H. acknowledges continued support over the years through the Estimating the Circulation and Climate of the Ocean (ECCO) project of the National Aeronautics and Space Administration (NASA) Physical Oceanography program. Additional support comes from NASA (grant NNX13AK88G), the National Science Foundation (grant 1434149), the National Oceanic and Atmospheric Administration (grant NA130AR4310135), and the Department of Energy (grant SC0008060).

LITERATURE CITED

Abraham JP, Baringer M, Bindoff NL, Boyer T, Cheng LJ, et al. 2013. A review of global ocean temperature observations: implications for ocean heat content estimates and climate change. *Rev. Geophys.* 52:450–83

Alves O, Balmaseda M, Anderson DLT, Stockdale T. 2004. Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Q. J. R. Meteorol. Soc.* 130:647–68

Anderson DLT, Sheinbaum J, Haines K. 1996. Data assimilation in ocean models. *Rep. Prog. Phys.* 59:1209

Anderson JL. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Weath. Rev.* 129:2884–903

Annan JD, Hargreaves JC, Edwards NR, Marsh R. 2005. Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. *Ocean Model.* 8:135–54

Balmaseda M, Dee D, Vidard A, Anderson DLT. 2007. A multivariate treatment of bias for sequential data assimilation: application to the tropical oceans. *Q. J. R. Meteorol. Soc.* 133:167–79

Balmaseda M, Fujii Y, Aves O, Awaji T, Behringer DW, et al. 2010. Initialization for seasonal and decadal forecasts. See Hall et al. 2010, Vol. 2, chap. 2

Balmaseda M, Hernandez F, Storto A, Palmer MD, Alves O, et al. 2015. The Ocean Reanalyses Intercomparison Project (ORA-IP). *J. Oper. Oceanogr.* 8(Suppl. 1):s80–97

Balmaseda M, Mogensen K, Weaver AT. 2013a. Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q. J. R. Meteorol. Soc.* 139:1132–61

Balmaseda M, Trenberth KE, Källén E. 2013b. Distinctive climate signals in reanalysis of global ocean heat content. *Geophys. Res. Lett.* 40:1754–59
Bennett AF. 1985. Array design by inverse methods. *Prog. Oceanogr.* 15:129–56
Bennett AF. 1992. *Inverse Methods in Physical Oceanography*. Cambridge, UK: Cambridge Univ. Press
Bennett AF. 2002. *Inverse Modeling of the Ocean and Atmosphere*. Cambridge, UK: Cambridge Univ. Press
Bertino L, Lisaeter K. 2008. The TOPAZ monitoring and prediction system for the Atlantic and Arctic Oceans. *J. Oper. Oceanogr.* 1:15–19
Bishop CH, Toth Z. 1999. Ensemble transformation and adaptive observations. *J. Atmos. Sci.* 56:1748–65
Blessing S, Kaminski T, Lunkeit F, Matei I, Giering R, et al. 2014. Testing variational estimation of process parameters and initial conditions of an Earth system model. *Tellus A* 66:22606
Bloom SC, Takacs LL, Da Silva AM, Ledvina D. 1996. Data assimilation using incremental analysis updates. *Mon. Weath. Rev.* 124:1256–71
Bouttier F, Courtier P. 1999. *Data assimilation concepts and methods*. Numer. Weath Predict. Course Notes, Eur. Cent. Medium-Range Weath. Forecasts (ECMWF), Reading, UK. **http://www.ecmwf.int/sites/default/files/Data%20assimilation%20concepts%20and%20methods.pdf**
Buckley M, Ponte RM, Forget G, Heimbach P. 2015. Determining the origins of advective heat transport variability in the North Atlantic. *J. Clim.* 28:3943–56
Buizza R, Palmer TN. 1995. The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.* 52:1434–56
Carton JA, Giese BS. 2008. A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Weath. Rev.* 136:2999–3017
Carton JA, Santorelli A. 2008. Global decadal upper-ocean heat content as viewed in nine analyses. *J. Clim.* 21:6015–35
Caya A, Buehner M, Carrieres T. 2010. Analysis and forecasting of sea ice conditions with three-dimensional variational data assimilation and a coupled ice-ocean model. *J. Atmos. Ocean. Technol.* 27:353–69
Corre L, Terray L, Balmaseda M, Ribes A, Weaver A. 2012. Can oceanic reanalyses be used to assess recent anthropogenic changes and low-frequency internal variability of upper ocean temperature? *Clim. Dyn.* 38:877–96
Courtier P, Andersson E, Heckley W, Vasiljevic D, Hamrud M, et al. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation. *Q. J. R. Meteorol. Soc.* 124:1783–807
Cunningham SA, Kanzow T, Rayner D, Baringer MO, Johns WE, et al. 2007. Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. *Science* 317:935–38
Czeschel L, Marshall DP, Johnson HL. 2010. Oscillatory sensitivity of Atlantic overturning to high-latitude forcing. *Geophys. Res. Lett.* 37:L10601
Derber J, Rosati A. 1989. A global oceanic data assimilation system. *J. Phys. Oceanogr.* 19:1333–47
Divakaran P, Brassington G, Feng M. 2010. *Mesoscale zonal features in the southeast Indian Ocean as seen in the Bluelink reanalysis 2.1*. Tech. Rep. 033, Centre for Australian Weather and Climate Research, Melbourne, Aust.
Driscoll S, Bozzo A, Gray LJ, Robock A, Stenchikov G. 2012. Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. *J. Geophys. Res. Atmos.* 117:D17105
Dutkiewicz S, Follows MJ, Heimbach P, Marshall J. 2006. Controls on ocean productivity and air-sea carbon flux: an adjoint model sensitivity study. *Geophys. Res. Lett.* 33:L02603
Edwards CA, Moore AM, Hoteit I, Cornuelle BD. 2014. Regional ocean data assimilation. *Annu. Rev. Mar. Sci.* 7:21–42
Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res. Oceans* 99:10143–62
Feltham DL. 2008. Sea ice rheology. *Annu. Rev. Fluid Mech.* 40:91–112
Fenty I, Heimbach P. 2013a. Coupled sea ice-ocean-state estimation in the Labrador Sea and Baffin Bay. *J. Phys. Oceanogr.* 43:884–904
Fenty I, Heimbach P. 2013b. Hydrographic preconditioning for seasonal sea ice anomalies in the Labrador Sea. *J. Phys. Oceanogr.* 43:863–83
Ferreira D, Marshall J, Heimbach P. 2005. Estimating eddy stresses by fitting dynamics to observations using a residual-mean ocean circulation model and its adjoint. *J. Phys. Oceanogr.* 35:1891–910

Forget G, Campin JM, Heimbach P, Hill CN, Ponte RM, Wunsch C. 2015. ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Dev. Discuss.* 8:3653–743

Forget G, Wunsch C. 2007. Estimated global hydrographic variability. *J. Phys. Oceanogr.* 37:1997–2008

Fu L-L, Cazenave A. 2001. *Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications*. San Diego, CA: Academic

Fujii Y, Nakaegawa T, Matsumoto S, Yasuda T, Yamanaka G, et al. 2009. Coupled climate simulation by constraining ocean fields in a coupled model with ocean data. *J. Clim.* 22:5541–57

Fukumori I. 2002. A partitioned Kalman filter and smoother. *Mon. Weath. Rev.* 130:1370–83

Fukumori I, Benveniste J, Wunsch C, Haidvogel DB. 1993. Assimilation of sea surface topography into an ocean circulation model using a steady state smoother. *J. Phys. Oceanogr.* 23:1831–55

Fukumori I, Lee T, Cheng B, Menemenlis D. 2004. The origin, pathway and destination of Niño-3 water estimated by a simulated passive tracer and its adjoint. *J. Phys. Oceanogr.* 34:582–604

Fukumori I, Menemenlis D, Lee T. 2007. A near-uniform basin-wide sea level fluctuation of the Mediterranean Sea. *J. Phys. Oceanogr.* 37:338–58

Galanti E, Tziperman E. 2003. A midlatitude-ENSO teleconnection mechanism via baroclinically unstable long Rossby waves. *J. Phys. Oceanogr.* 33:1877–88

Ganachaud A, Wunsch C. 2003. Large-scale ocean heat and freshwater transports during the World Ocean Circulation Experiment. *J. Clim.* 16:696–705

Gandin LS. 1963. *Objektivnyi Analyz Meteorologicheskikh Polei*. Leningrad: Gidrometeoizdat

Gelb A, ed. 1974. *Applied Optimal Estimation*. Cambridge, MA: MIT Press

Giering R, Kaminski T. 1998. Recipes for adjoint code construction. *ACM Trans. Math. Softw.* 24:437–74

Griffies SM, Adcroft AJ. 2008. Formulating the equations for ocean models. In *Ocean Modeling in an Eddying Regime*, ed. M Hecht, H Hasumi, pp. 281–317. Geophys. Monogr. Ser. 177. Washington, DC: Am. Geophys. Union

Griffies SM, Böning C, Bryan FO, Chassignet EP, Gerdes R, et al. 2001. Developments in ocean climate modelling. *Ocean Model.* 2:123–92

Halkides DJ, Lee T. 2009. Mechanisms controlling seasonal-to-interannual mixed-layer temperature variability in the southeastern tropical Indian Ocean. *J. Geophys. Res.* 114:C02012

Hall J, Harrison DE, Stammer D, eds. 2010. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society, Venice, Italy, 21–25 September 2009*. ESA Publ. WPP-306. Paris: Eur. Space Agency. **http://www.oceanobs09.net/proceedings**

Hall MC. 1986. Application of adjoint sensitivity theory to an atmospheric general circulation model. *J. Atmos. Sci.* 43:2644–52

Hamill TM, Snyder C. 2000. A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Weath. Rev.* 128:2905–19

Heimbach P, Losch M. 2012. Adjoint sensitivities of sub-ice shelf melt rates to ocean circulation under Pine Island Ice Shelf, West Antarctica. *Ann. Glaciol.* 53:59–69

Heimbach P, Wunsch C, Ponte RM, Forget G, Hill C, Utke J. 2011. Timescales and regions of the sensitivity of Atlantic meridional volume and heat transport magnitudes: toward observing system design. *Deep-Sea Res. II* 58:1858–79

Hoteit I, Cornuelle B, Köhl A, Stammer D. 2005. Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Q. J. R. Meteorol. Soc.* 131:3659–82

Hunke EC, Lipscomb WH, Turner AK. 2010. Sea-ice models for climate study: retrospective and new directions. *J. Glaciol.* 56:1162

Ji M, Leetmaa A, Derber J. 1995. An ocean analysis for seasonal to interannual climate studies. *Mon. Weath. Rev.* 123:460–81

Kalman RE. 1960. A new approach to linear filtering and prediction problems. *J. Fluids Eng.* 82:35–45

Kalmikov AG, Heimbach P. 2014. A Hessian-based method for uncertainty quantification in global ocean state estimation. *SIAM J. Sci. Comput.* 36:S267–95

Karspeck A, Stammer D, Köhl A, Danabasoglu G, Balmaseda M, et al. 2015. Comparison of the Atlantic meridional overturning circulation between 1960 and 2007 in six ocean reanalysis products. *Clim. Dyn.* In press

Karspeck A, Yeager S, Danabasoglu G, Teng H. 2014. An evaluation of experimental decadal predictions using CCSM4. *Clim. Dyn.* 44:907–23

Kauker F, Kaminski T, Karcher M, Giering R, Gerdes R, Voßbeck M. 2009. Adjoint analysis of the 2007 all time Arctic sea-ice minimum. *Geophys. Res. Lett.* 36:L03707

Keenlyside N, Latif M, Junclaus J, Kornblueh L, Roeckner E. 2008. Advancing decadal climate scale prediction in the North Atlantic. *Nature* 453:84–88

Kim S-B, Lee T, Fukumori I. 2007. Mechanisms controlling the interannual variation of mixed layer temperature averaged over the Niño-3 region. *J. Clim.* 20:3822–43

Köhl A. 2005. Anomalies of meridional overturning: mechanisms in the North Atlantic. *J. Phys. Oceanogr.* 35:1455–72

Köhl A, Stammer D. 2004. Optimal observations for variational data assimilation. *J. Phys. Oceanogr.* 34:529–42

Köhl A, Stammer D. 2008. Variability of the meridional overturning in the North Atlantic from 50-year GECCO state estimation. *J. Phys. Oceanogr.* 38:1913–30

Köhl A, Willebrand J. 2002. An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean models. *Tellus A* 54:406–25

Laloyaux P, Balmaseda M, Dee D, Mogensen K, Janssen P. 2015. The ECMWF prototype for a coupled assimilation system. *Q. J. R. Meteor. Soc.* In press

Lee T, Awaji T, Balmaseda M, Greiner E, Stammer D. 2009. Ocean state estimation for climate research. *Oceanography* 22(3):160–67

Liang X, Wunsch C, Heimbach P, Forget G. 2015. Vertical redistribution of oceanic heat. *J. Clim.* 28:3821–33

Liu C, Köhl A, Stammer D. 2012. Adjoint-based estimation of eddy-induced tracer mixing parameters in the global ocean. *J. Phys. Oceanogr.* 42:1186–206

Liu Y, Liu Z, Zhang S, Jacob R, Lu F, et al. 2014. Ensemble-based parameter estimation in a coupled general circulation model. *J. Clim.* 27:7151–62

Macdonald AM. 1998. The global ocean circulation: a hydrographic estimate and regional analysis. *Prog. Oceanogr.* 41:281–382

Magnusson L, Alonso-Balmaseda M, Corti S, Molteni F, Stockdale T. 2013. Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. *Clim. Dyn.* 41:2393–409

Malanotte-Rizzoli P. 1996. *Modern Approaches to Data Assimilation in Ocean Modeling*. Amsterdam: Elsevier

Marotzke J, Giering R, Zhang QK, Stammer D, Hill CN, Lee T. 1999. Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. *J. Geophys. Res.* 104:29529–48

Martin M, Balmaseda M, Bertino L, Brasseur P, Brassington G, et al. 2015. Status and future of data assimilation in operational oceanography. *J. Oper. Oceanogr.* 7:29–46

Masuda S, Awaji T, Sugiura N, Matthews JP, Toyoda T, et al. 2010. Simulated rapid warming of abyssal North Pacific waters. *Science* 329:319–22

Masuda S, Awaji T, Sugiura N, Toyoda T, Ishikawa Y, Horiuchi K. 2006. Interannual variability of temperature inversions in the subarctic North Pacific. *Geophys. Res. Lett.* 33:L24610

Maximenko NA, Melnichenko OV, Niiler PP, Sasaki H. 2008. Stationary mesoscale jet-like features in the ocean. *Geophys. Res. Lett.* 35:L08603

Mazloff MR, Heimbach P, Wunsch C. 2010. An eddy-permitting Southern Ocean state estimate. *J. Phys. Oceanogr.* 40:880–99

Meehl GA, Goddard L, Boer G, Burgman R, Branstator G, et al. 2014. Decadal climate prediction: an update from the trenches. *Bull. Am. Meteorol. Soc.* 95:243–67

Meier WN, Hovelsrud GK, van Oort BEH, Key JR, Kovacs KM, et al. 2014. Arctic sea ice in transformation: a review of recent observed changes and impacts on biology and human activity. *Rev. Geophys.* 52:185–217

Menemenlis D, Fukumori I, Lee T. 2005. Using Green's functions to calibrate an ocean general circulation model. *Mon. Weath. Rev.* 133:1224–40

Moore AM, Arango HG, Broquet G, Powell BS, Weaver AT, Zavala-Garay J. 2011. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part I: system overview and formulation. *Prog. Oceanogr.* 91:34–49

Munk W, Wunsch C. 1982. Observing the ocean in the 1990s. *Philos. Trans. R. Soc. Lond. A* 307:439–64

Palmer MD, Roberts CD, Balmaseda M, Boyer T, Chang Y-S, et al. 2015 . Ocean heat content variability and change in an ensemble of ocean reanalyses. *Clim. Dyn.* In press

Palmer TN, Anderson DLT. 1994. The prospects for seasonal forecasting—a review paper. *Q. J. R. Meteorol. Soc.* 120:755–93

Palmer TN, Doblas-Reyes FJ, Hagedorn R, Weisheimer A. 2005. Probabilistic prediction of climate using multi-model ensembles: from basics to applications. *Philos. Trans. R. Soc. B* 360:1991–98

Panteleev G, Nechaev DA, Proshutinsky A, Woodgate R, Zhang J. 2010. Reconstruction and analysis of the Chukchi Sea circulation in 1990–1991. *J. Geophys. Res.* 115:C08023

Penland C, Sardeshmukh PD. 1995. The optimal growth of tropical sea surface temperature anomalies. *J. Clim.* 8:1999–2024

Pham DT, Verron J, Roubaud MC. 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.* 16:323–40

Piecuch C, Ponte R. 2014. Mechanisms of global mean steric sea level change. *J. Clim.* 27:824–34

Pierce DW, Barnett TP, Tokmakian R, Semtner A, Maltrud M, et al. 2004. The ACPI project, element 1: initializing a coupled climate model from observed initial conditions. *Clim. Change* 62:13–28

Pohlmann H, Jungclaus J, Marotzke J, Köhl A, Stammer D. 2009. Improving predictability through the initialization of a coupled climate model with global oceanic reanalysis. *J. Clim.* 22:3926–38

Polkova I, Köhl A, Stammer D. 2014. Impact of initialization procedures on the predictive skill of a coupled ocean-atmosphere model. *Clim. Dyn.* 42:3151–69

Ponte R, Wunsch C, Stammer D. 2007. Spatial mapping of time-variable errors in Jason-1 and TOPEX/POSEIDON sea surface height measurements. *J. Atmos. Ocean. Technol.* 24:1078–85

Rauch HE, Striebel CT, Tung F. 1965. Maximum likelihood estimates of linear dynamic systems. *AIAA J.* 3:1445–50

Ricci S, Weaver AT, Vialard J, Rogel P. 2005. Incorporating temperature-salinity constraints in the background-error covariance of variational ocean data assimilation. *Mon. Weath. Rev.* 133:317–38

Roemmich D, Boebel O, Desaubies Y, Freeland H, Kim K, et al. 2001. Argo: the global array of profiling floats. In *Observing the Oceans in the 21st Century*, ed. CJ Koblinsky, NR Smith, pp. 248–58. Melbourne, Aust.: Bur. Meteorol.

Saha S, Moorthi S, Pan HL, Wu X, Wang J, et al. 2010. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* 91:1015–57

Sapsis TP, Lermusiaux PF. 2009. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Phys. D* 238:2347–60

Sasaki Y. 1970. Some basic formalisms in numerical variational analysis. *Mon. Weath. Rev.* 98:875–83

Schiller A, Lee T, Masuda, S. 2013. Methods and applications of ocean synthesis in climate research. See Siedler et al. 2013, pp. 581–608

Segschneider J, Anderson DLT, Stockdale TN. 2000. Towards the use of altimetry for operational seasonal forecasting. *J. Clim.* 13:3116–38

Siedler G, Griffies SM, Gould J, Church JA, eds. 2013. *Ocean Circulation and Climate: A 21st Century Perspective*. San Diego, CA: Academic. 2nd ed.

Smith DM, Cusack S, Colman A, Folland CK, Harris GR, Murphy JM. 2007. Improved surface temperature prediction for the coming decade from a global climate model. *Science* 317:796–99

Smith NR, Koblinsky C. 2001. The ocean observing system for the 21st century: a consensus statement. In *Observing the Oceans in the 21st Century*, ed. CJ Koblinsky, NR Smith, pp. 1–25. Melbourne, Aust.: Bur. Meteorol.

Sorensen HW. 1970. Least-squares estimation: from Gauss to Kalman. *IEEE Spectrum* 7:63–68

Speer K, Forget G. 2013. Global distribution and formation of mode waters. See Siedler et al. 2013, pp. 211–26

Stammer D. 2005. Adjusting internal model errors through ocean state estimation. *J. Phys. Oceanogr.* 35:1143–53

Stammer D, Köhl A, Awaji T, Balmaseda M, Behringer D, et al. 2010. Ocean information provided through ensemble ocean syntheses. See Hall et al. 2010, Vol. 2, chap. 85

Stammer D, Park S, Köhl A, Lukas R, Santiago-Mandujano F. 2008. Large-scale hydrographic changes at the Hawaii Ocean Time-series Station. *J. Phys. Oceanogr.* 38:1931–48

Stammer D, Ueyoshi K, Köhl A, Large WB, Josey S, Wunsch C. 2004. Estimating air-sea fluxes of heat, freshwater and momentum through global ocean data assimilation. *J. Geophys. Res.* 109:C05023

Stammer D, Wunsch C, Giering R, Eckert C, Heimbach P, et al. 2002. The global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *J. Geophys. Res.* 107:3118

Stammer D, Wunsch C, Giering R, Zhang QK, Marotzke J, et al. 1997. *The global ocean circulation estimated from TOPEX/POSEIDON altimetry and the MIT general circulation model*. Rep. 49, MIT Cent. Glob. Change Sci., Cambridge, MA

Storto A, Masina S, Balmaseda M, Guinehut S, Xue Y, et al. 2015. Steric sea level variability (1993–2010) in an ensemble of ocean reanalyses and objective analyses. *Clim. Dyn.* In press. doi: 10.1007/s00382-015-2554-9

Sugiura N, Awaji T, Masuda S, Mochizuki T, Toyoda T, et al. 2008. Development of a 4-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res.* 113:C10017

Talagrand O. 1997. Assimilation of observations, an introduction. *J. Meteorol. Soc. Jpn. Ser. 2* 75:81–99

Talagrand O, Courtier P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: theory. *Q. J. R. Meteorol. Soc.* 113:1311–28

Tanguay M, Bartello P, Gauthier P. 1995. Four-dimensional data assimilation with a wide range of scales. *Tellus A* 47:974–97

Thacker WC, Long RB. 1988. Fitting dynamics to data. *J. Geophys. Res. Oceans* 93:1227–40

Toyoda T, Awaji T, Masuda S, Sugiura N, Igarashi H, et al. 2011. Interannual variability of North Pacific eastern subtropical mode water formation in the 1990s derived from a 4-dimensional variational ocean data assimilation experiment. *Dyn. Atmos. Oceans* 51:1–25

Toyoda T, Fujii Y, Kuragano T, Kamachi M, Ishikawa Y, et al. 2015. Intercomparison and validation of the mixed layer depth fields of global ocean syntheses. *Clim. Dyn.* In press. doi: 10.1007/s00382-015-2637-7

Wang O, Fukumori I, Lee T, Cheng B. 2004. On the cause of eastern equatorial Pacific Ocean T-S variations associated with El Niño. *Geophys. Res. Lett.* 31:L15310

Weaver AT, Courtier P. 2001. Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc.* 127:1815–46

Weaver AT, Deltel C, Machu É, Ricci S, Daget N. 2005. A multivariate balance operator for variational ocean data assimilation. *Q. J. R. Meteorol. Soc.* 131:3605–25

Weaver AT, Vialard J, Anderson DLT. 2003. Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part 1: formulation, internal diagnostics, and consistency checks. *Mon. Weath. Rev.* 131:1360–78

WMO (World Meteorol. Organ.). 2009. *14th WMO/IAEA meeting of experts on carbon dioxide, other greenhouse gases and related tracers measurement techniques (Helsinki, Finland, 10–13 September 2007)*. GAW Rep. No. 186, WMO/TD No. 1487, WMO, Geneva, Switz.

Wunsch C. 1978. The North Atlantic general circulation west of 50°W determined by inverse methods. *Rev. Geophys.* 16:583–620

Wunsch C. 1996. *The Ocean Circulation Inverse Problem*. Cambridge, UK: Cambridge Univ. Press

Wunsch C, Heimbach P. 2006. Estimated decadal changes in the North Atlantic meridional overturning and heat flux 1993–2004. *J. Phys. Oceanogr.* 36:2012–24

Wunsch C, Heimbach P. 2013. Dynamically and kinematically consistent global ocean circulation and ice state estimates. See Siedler et al. 2013, pp. 553–79

Wunsch C, Heimbach P. 2014. Bidecadal thermal changes in the abyssal ocean. *J. Phys. Oceanogr.* 44:2013–30

Wunsch C, Ponte RM, Heimbach P. 2007. Decadal trends in sea level patterns: 1993–2004. *J. Clim.* 20:5889–911

Zanchettin D, Bothe O, Graf HF, Lorenz SJ, Luterbacher J, et al. 2013. Background conditions influence the decadal climate response to strong volcanic eruptions. *J. Geophys. Res. Atmos.* 118:4090–106

Zanna L, Heimbach P, Moore AM, Tziperman E. 2011. Optimal excitation of interannual Atlantic meridional overturning circulation variability. *J. Clim.* 24:413–27

Zanna L, Heimbach P, Moore AM, Tziperman E. 2012. Upper-ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability. *Q. J. R. Meteorol. Soc.* 138:500–13

Zhang S, Harrison MJ, Rosati A, Wittenberg A. 2007. System design and evaluation of coupled ensemble data assimilation for global oceanic studies. *Mon. Weath. Rev.* 135:3541–64