Ocean Data Assimilation in Support of Climate Applications

- Status and Perspectives -

D. Stammer¹, M. Balmaseda², P. Heimbach³, A. Köhl¹, A. Weaver⁴

¹ Center für Erdsystemwissenschaften und Nachhaltigkeit (CEN), Universität Hamburg, Hamburg, Germany

² European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, UK

³ Institute for Computational Engineering and Sciences and Jackson School of Geosciences, The University of Texas at Austin, Austin (TX), USA

&

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

⁴ Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique / SUC URA CERFACS/CNRS 1875, Toulouse, France
**Abstract** (max. 150 words)
Ocean data assimilation brings together observations with known dynamics encapsulated in a circulation model to describe the time-varying ocean circulation. Applications of ocean data assimilation are manifold, from marine and ecosystem forecasting to climate prediction and carbon cycle. Here we address only climate applications, which reach from improving our understanding of the 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 have to be used judiciously and need to be selected according to the purpose. Further advances are expected from improved models and methods for estimating and representing error information in the data assimilation system. Ultimately data assimilation into coupled climate system components is required in support of ocean and climate services. However, maintaining the requirements in terms of infrastructure and expertise for sustained data assimilation remains challenging.

Keywords (3-5): Data Assimilation, Observing System, Data Synthesis, Modeling, Climate Predictions, Ocean Circulation
1 Introduction

“Ocean data assimilation” (ODA) encompasses a broad set of mathematical and computational tools, and aims at providing best possible descriptions of the time-varying ocean circulation. Thereby ODA supports studies of ocean dynamics, in particular for estimating unobservable quantities. 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 ocean’s ecosystems, its biogeochemistry, the marine (sea ice) or marine-terminating cryosphere, or the coupled climate system as a whole. One strand of assimilation 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 ODA and “ocean state estimation” (OSE) are often used synonymously, both approaches are succinctly different inverse approaches to yield an ocean synthesis. The term “data assimilation” (DA) was coined initially in numerical weather prediction (NWP) where it referred to the technique of creating initial conditions for atmospheric models for forecasting over timescales of hours to a few days, thereby putting an emphasis on the instantaneous state of the atmosphere. Comprehensive mathematical expositions of the original DA approaches are summarized by Bouttier and Courtier (1999). These approaches were later adopted by the oceanographic and seasonal forecast communities for the purposes of producing nowcasts and for initializing ocean and seasonal forecasts (e.g., Anderson et al., 1996; Talagrand, 1997). In contrast, OSE from the beginning of the World Ocean Circulation Experiment (WOCE) was intended as a means of bringing all ocean surface (including satellite data) and sub-surface 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 and Wunsch, 1982). By definition, OSE thereby focuses on the evolving state and long time scales.

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, 2003). At about the same time it became clear, however, that any ocean synthesis effort has to 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. Respective steps encompassed making “filter” approaches (see below) technically feasible and included the development of “smoother” approaches, such as those that employed “adjoint” representations of modern ocean primitive equation (PE) models (e.g., Thacker and Long, 1988). Applications of the adjoint technique to complex models were made possible through the development of automatic differentiation techniques and software tools (Giering and Kaminski, 1998) and their pilot application to ocean problems (Marotzke et al., 1999). It required the establishment of the computer infrastructure suitable for solving large nonlinear optimization problems.

Traditionally ODA is associated with filter approaches, while smoothers are typically used in OSE efforts. Today both strands have evolved into mature fields, comparable in sophistication and usage to atmospheric reanalysis. Their difference in intention has largely diminished, with both now aiming at supporting climate oriented ocean synthesis. In particular, both approaches are used today to initialize climate forecasts and remain concerned with improving ocean and climate models, with providing uncertainty estimates and with 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, serving then either climate or ocean services. However, significant improvements have to be accomplished before the full potential of DA can be utilized and the goals of ocean synthesis in general can be accomplished.

This paper provides a critical review of the status of ODA in support of climate applications, and lays out the developments that are required to reach its full potential for oceanography and climate science at large. For this purpose we review the strengths and weaknesses of various ODA approaches, provide examples of ongoing applications, summarize the role that ODA plays for analyzing the ocean but also for initializing coupled models with emphasis on climate applications. Results are used to identify improvements required to move toward ocean and climate information systems in support of many applications. Short-term predictions and operational oceanography will not be discussed in detail. Reviews of the marine forecasting applications (both global and regional) were provided recently by Edwards et al. (2014) and Martin et al. (2015).
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 error bars of both). Basic ingredients to such an approach are: (1) a model which is being constrained by (2) quality controlled data, (3) error information about both, (4) a methodology by which data and model results are fused, and (5) a method to estimate uncertainty information about the estimated state.

2.1 Models and Data

Models: A model in the context of ODA can be any mathematical description of an ocean parameter (or variable in the widest sense) that is being estimated though 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 and Adcroft (2008) provide a recent review of ocean model formulations. Remaining model deficiencies can be summarized as (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, as well as uncertain model parameters (e.g., mixing and diffusion). Uncertainties also arise from inaccurate initial or boundary conditions, the latter including 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, individually or in combinations. However, depending on the approach, the success may be limited and large uncertainties in the estimation remain that are not always easy to quantify.

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 through subsequent efforts such as OceanObs’99 (Smith and 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 to be indispensable for observing the ocean variability (Fu and 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 2000 m depths. Merging satellite observations, other ocean observations and an ocean circulation model into a description of the ocean flow field through DA is an important process of making maximum use of existing observations for oceanography and climate studies, and should be considered part of an observing strategy, equivalent to what is pursued in NWP.

Substantial, often unknown, uncertainties remain in existing observations, the XBT fall-rate errors just being one prominent example (Abraham et al., 2013). Uncertainties in surface fluxes are usually unknown and continued data reanalysis and quality control efforts have to 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 2000 m and current observations. Evaluating past climate variability and change from an observing system and forcing fields that changed markedly in quality and quantity over time remains a major challenge; ODA efforts can support this process and are also a valuable tool for informing the optimization of the future ocean observing system.

2.2 Methodologies and Approaches

Most ODA approaches are variants of classical “least squares” of combining models with data, assuming that errors are Gaussian. “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) to rigorous but computationally intensive (e.g., Kalman filters, 4D variational/adjoint or other smoothers). Applied DA schemes (e.g., Wunsch, 1996; Hólm, 2003) 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 detail of how DA schemes assimilate available observations, whether a solution to a constrained or unconstrained optimization problem is
being sought, and the level of accuracy by which priori error estimates of observations and the model dynamics are being 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 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 (Fig. 1) is essential for their appropriate use. In the following, approaches typical for ODA and OSE will be introduced, subdivided into filters and smoothers.

**Filter approaches** estimate sequentially 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 first principles of the ocean circulation and often may introduce discontinuities in the time evolution of the model trajectory. Discontinuities can be remedied to some extent via “incremental analysis updating” (IAU; Bloom et al., 1996) by transforming the increment into a forcing that distributes the correction over a period; corrections remain dynamically unbalanced, though. Nevertheless, 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, notably optimal interpolation (OI), three-dimensional variational assimilation (3D-VAR), and the Kalman Filter (KF; Kalman, 1960) in various forms (the first two approaches can be shown to be approximations of the latter).

In this list, 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 observation and corresponding model simulation (called 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, likewise a minimum variance estimator, 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 to the KF have been devised; among them is the so-called partitioned KF. It 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 (EKF; Gelb, 1974) can be applied to weakly nonlinear problems under the tangent-linear approximation but still suffers from excessive computational costs. For stronger non-linear 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) reduce the rank of the covariance matrix by empirical orthogonal functions (EOF). 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 Kalman Filter, which is based on the ensemble transformation (Bishop and Toth, 1999) and does not require adding perturbations to the observations.

3D-VAR is a maximum likelihood estimator, which treats the elements in $J$ independently in time and seeks an approximate solution through iterative minimization (e.g., Derber and 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, non-diagonal formulations of the background error covariance matrix (Weaver and Courtier, 2001).
**Smoothen Approaches** use observations from the future and the past to constrain the ocean circulation in a retrospective analysis. They differ from filter methods in that they estimate an ocean state, not by changing the prognostic model state at analysis times, but 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. Respective 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 ocean state estimation practical: the optimal Rauch-Tung-Striebel (RTS, Rausch et al., 1965) smoother and the adjoint method. These methods have different algorithmic properties but are equivalent as long as assumptions about data and model dynamic constraint errors are the same (e.g., Bennett, 2002; Lee et al., 2009; comprehensive mathematical expositions of original smoother formulations are provided by Bouttier and Courtier, 1999, or Wunsch, 1996). 4D-VAR is a variant of the adjoint method, applied over shorter time windows with substantial benefits over 3D-VAR (e.g., Weaver et al., 2003).

The optimal RTS 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 both before and after (e.g., Cohn et al., 1994). The use of observations from the future thereby leads to uncertainties that are smaller than those associated with filtered results (e.g., Fukumori, 2002). It 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.

In contrast, the “whole domain” adjoint or Lagrange multiplier approach, originating 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; Talagrand and Courtier, 1987; Thacker and Long, 1988). The method is based on the assumption that model equations are correct (“strong constraint” formalism). The approach can deal with weakly non-linear problems; however, it 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 allowed departures from model dynamics while obtaining an optimal state
estimate. The so-called “Representer Method”, which is one algorithm for solving the weakly constrained 4D-VAR problem, seeks the solution in 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 were devised which aim to combine the strengths of variational and ensemble methods in sequential DA (Hamill and Synder, 2000). Variational methods have algorithmic advantages for solving the analysis problem and for including complex analytical constraints, while 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

First pilot large-scale OSE attempts of the time-varying ocean states came into existence during the 1990s (Fukumori et al., 1992; Stammer et al., 1997), at a time when also the first multi-year ODA products in support of seasonal forecasts were created (Derber and Rosati, 1989, Ji et al., 1995). Since then, with expanding technical capabilities the demand for more sophistication grew, leading to higher spatial resolution, longer estimation periods, but also more complex applications, including biogeochemical investigations. Today several global synthesis systems exist which are being used across several research and operational institutions, supporting a variety of applications. Table 1 provides a summary of all 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 involving timescales of days to weeks, requires high spatial resolution, and are produced in quasi-real time; climate-oriented state estimation involves monthly to decadal timescales. In contrast, initialization of monthly and seasonal forecasts, involve long time scales but have the operational constraints of prompt real-time delivery.

Based on selected examples, we will review in the following the status of ocean synthesis separately for climate and (operational) high-resolution applications. Because underlying models and underlying assimilation approaches differ, results from individual ocean synthesis efforts will be seen to differ substantially and analyzing them indiscriminately might be misleading.
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 extending a few decades back in time, preceding 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 respective uncertainties. Examples include studies of sea level variability and change (e.g., Stammer et al., 2002a, 2004; Carton and Giese, 2008; Wunsch et al., 2007; Köhl and Stammer, 2008; Balmaseda et al., 2013a, Piecuch and Ponte, 2014, Storto et al., 2015), water-mass analysis (e.g., Fukumori et al., 2004; Wang et al., 2004; Masuda et al., 2006; Toyoda et al., 2009, 2015, Speer and Forget, 2013), mixed-layer heat balance (e.g., Kim et al., 2007; Halkides and Lee, 2009; Buckley et al., 2015), or changes in the ocean’s heat content (Carton and Santorelli, 2008, Balmaseda et al., 2013b, Wunsch and Heimbach, 2014).

Ocean heat content 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 ocean heat content (OHC) benefits from the combination of observations and models via dynamical constraints provided by the DA system. Results lead to more obvious variations in OHC related to El Niño Southern Oscillation (ENSO) than present in observations-only syntheses, especially prior to the Argo period. Recent comparison of ocean reanalyses (Balmaseda et al., 2015; Palmer et al., 2015) suggest that while the upper OHC is relatively well constrained in the recent period, substantial uncertainty exists among the existing estimates regarding the vertical penetration of heat, and large uncertainty remains in the period prior to Argo. As can be seen from Fig. 2, the OHC increase is not monotonic and smooth but shows significant variations on all time scales. We expect similar variability to exist in future OHC changes and all other climate variables for that same 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.

While global indicators of climate change in the ocean are important, it is usually the regional changes that are of largest consequence and therefore 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 potentially can 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) compare linear trends over the period 1993 – 2010 of steric height from different ocean syntheses and find that large variations exist among individual products on regional scale, largely arising from uncertainties in the deep ocean and discrepancies in the halo-steric component.

The Atlantic meridional overturning circulation (AMOC), a measure of zonally and vertically integrated poleward volume transports, is another important climate index since it is associated with poleward heat and freshwater transports that play an important role in the coupled climate system (Cunningham et al., 2007; Wunsch and Heimbach, 2006). 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 multi-decadal ocean synthesis products. As an example, Fig. 3 in its left part 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 approximately 1000 m and southward flow below this level. However, all products except GECCO2 have more than one distinct positive maximum at different latitudes, with DEPRESYS, SODA and MOVE-CORE showing localized circulations near the equator. Despite the fact that 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. In the right part of the figure, time series of the AMOC anomaly at 1000 m depth at 45°N (top panels) and 26.5°N (bottom panels) are compared. Visual inspection suggests very little agreement in the year-to-year changes and trends in the synthesis set, implying that even in areas relatively well observed 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, the underlying model, including differences in forcing.

Using various ocean syntheses, Toyoda (2015) investigated seasonal-to-decadal variations of mixed layer depth in the Pacific (MLD). The authors found two coherent dominant modes of MLD variability, one related to Pacific Decadal Oscillation (PDO) changes, and the second one suggesting the existence of a coupled mode between mixed-layer induced SST anomalies and variations in atmospheric sea level pressure related to the West Pacific Index. Taking advantage of property conservation of state estimates, Buckley et al. (2015) attributed SST and upper-ocean heat content changes in ocean syntheses to local buoyancy versus wind
forcing, as well as processes involving ocean dynamics (advection versus subgrid-scale mixing). The transient nature of the ocean circulation with its long-term memory also implies 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

Computing uncertainty estimates for ocean syntheses in practice remains challenging due to 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 Kalman filter, the analysis error covariance matrix is required by the solution algorithm; 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 combination of parameters that are best and least well determined by the observations. Using Hessian information for inferring posterior error covariances is being explored within limited-domain GCM applications (e.g., Sapsis and Lermusiaux, 2009; Moore et al., 2011; Kalmikov and Heimbach, 2014). For most applications, however, strict application of the theory has been limited to applications in which only a limited number of parameters are estimated or 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 path of western boundary
currents stand out as the areas with 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 and Brassington, 2010); they can also provide initial conditions for short-term high-resolution ocean forecasting. However, progress has been hindered by the fact that ODA methods fundamentally rely on linearized model dynamics. Techniques, such as the EnKF (Evensen 1994) or approximate adjoint models (Köhl and Willebrand, 2002; Hoteit et al., 2005), were developed to deal with exponential error growth associated with non-linear 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 periods of time to resolutions of 1/4° (Table 1). Nevertheless, on 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 are provided by MyOcean (http://www.myocean.eu/) and the Copernicus program (http://www.copernicus.eu/). Similar projects exist as part of the U.S. Integrated Ocean Observing System (IOOS) or 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 of them provide daily short-term forecasts. By way of example, Fig. 4 shows a comparison of results obtained with a 1/36° version of the Operational Mercator Ocean analysis and forecast system with MODIS SST data, demonstrating the amount of detail current systems resolve. An example of high-resolution state estimates for climate science is the Southern Ocean State Estimate (SOSE; 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, for performing climate sensitivity studies to understand climate dynamics and to optimize the observing system. All 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. Early on, it was realized in the atmospheric community (e.g. Hall et al., 1986) that this gradient – also called adjoint sensitivity - provides a comprehensive tool to explore model sensitivities to parameters. In ocean modeling, adjoint sensitivities were long ignored, however. 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. Stammer et al., 2008; Galanti and Tziperman, 2003; Masuda et al., 2010).

Different from conventional sensitivity calculations via perturbation sensitivities that infer how the climate system responds to changes of individual parameters at a time, adjoint sensitivities provide the answer to 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 Fig. 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. Related, the use of adjoint sensitivities for reconstructing the full circulation using known perturbations (either time-varying, or mean) and explaining mechanistic causes in terms of dominant perturbations was further explored by Fukumori et al. (2007), Czeschel et al. (2010), and Heimbach et al. (2011). Over the years, the use of adjoint sensitivities to ocean circulation was extended to ocean biogeochemical processes (Dutkiewicz et al., 2006), to coupled ocean-sea ice processes in the Arctic component (e.g., Kauker et al., 2009), and to melt rates in sub-ice shelf cavities (Heimbach and Losch, 2012).

A variant of the sensitivity analysis, the optimal observations defined by Köhl and Stammer (2004), combine 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, for instance anomalous overturning at a certain time and place, to the past and future changes in the ocean. The technique was used by Köhl (2005) 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 change significantly the probability of occurrence of weather patterns (Palmer and
Anderson 1994). While not the initial motivation, the subject of climate predictions on seasonal to decadal and longer time scales has steadily moved into the central focus of several synthesis efforts, largely fostered through the World Climate Research Program (WCRP) and its core-project CLIVAR.

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 time scales. 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. It reduces the large uncertainty (error) due to the forcing fluxes and improves forecast skill (Alves et al., 2004; Balmaseda et al., 2010). Respective SI forecasting systems are based on coupled ocean-atmosphere general circulation models 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 (few decades); these hindcasts are also needed for skill assessments. The realism of their interannual variability will determine 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 currently 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 (Piers 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 (Fig. 6a), being far from the real world is detrimental for the forecast skill (Fig. 6b). The best skill is obtained by using empirical corrections of model error (blue and green line), which reduce the initialization shock and decrease model drift. It is expected that a more balanced “coupled” initialization is desirable, but it 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 at successfully predicting 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, Zanchetti et al., 2013). Nevertheless, initial decadal prediction efforts in recent years show 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 multi-model ensembles exist that suggest 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, to be better represented in initialized hindcasts, compared to non-initialized simulations. Many of the recent decadal prediction studies find enhanced predictive skill notably in the North Atlantic region associated with AMOC variability and predictability (Meehl et al., 2014). 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 time scales of seasons to a year and longer. A key difference between initialized decadal predictions and initialized predictions on shorter time-scales is the need for observations in the deeper ocean (below 500 m depth); even observations below 2000 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 time scales; e.g., initializing large-scale modes of
decadal variability (e.g., PDO) 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, the anomaly initialization was therefore more frequently used in decadal forecasts, but shows weaker performance than the full initialization that is currently favored, especially on seasonal time scales. Decadal forecasting is a rapidly evolving field (Meehl et al., 2014), which 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 NWP (e.g., Buizza and 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 and Sardesmuk, 1995). Recently Zanna et al. (2011) have shown that predictability studies using optimal perturbation techniques reveal ocean dynamical mechanisms that can limit predictability horizons of climate indices such as the AMOC (Fig. 7). Important implications for prediction are that (1) ensemble generation plays a key role with perturbations to include the initial state of the ocean, and (2) the need for ocean observations that reach to significant depths to be able to constrain prediction models.

5 outlook: The way forward

With an ever increasing diversity and heterogeneity of ocean observations, 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, required to provide ocean information on a regular basis and for many applications. In particular we envision that ocean syntheses will be used increasingly by other disciplines, e.g., for carbon or nutrient cycle studies, or for investigating the dependence of biodiversity on the physical climate state. To further increase the value of ocean synthesis products for all those applications, much effort is needed to characterize uncertainties in each product, to improve the products by including better/more observations as constraints, to improve models, and to advance assimilation approaches. On the other hand, we expect ODA to become an integral part of seamless climate prediction system including seasonal,
interannual and decadal timescales that will allow investigation of multi-scale interactions. It is expected that the best forecasts will be produced by coupled models that are directly constrained by climate data (i.e., **coupled data assimilation**). Ultimately, every ocean or coupled synthesis should be accompanied by formal uncertainty measures provided on a geographic grid for any estimated parameter. All these aspects are cutting edge research topics which to address in full detail is not permitted due to space limitations. Nevertheless, some thoughts will be provided below.

### 5.1 Improved uncertainty measures

Given the large remaining differences between individual ocean syntheses, one important step forward is to provide ensemble mean estimates and their uncertainties, a step akin to what is now common practice in NWP. 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 realized in existing measures. Much effort is required to realize the computation of 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., Ponte et al., 2007; Forget and Wunsch, 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 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 for treatment of 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 be a very high dimensional (typically order of $10^{18}$ or higher) covariance matrix is impractical. Approximate approaches or
projection methods onto low-dimensional (scalar) climate indices/quantities of interest are required; existing approximations inferring leading eigenvectors of the posterior error covariance matrix are a promising approach for capturing at least dominant uncertainty structures (Moore et al., 2011; Kalmikov and Heimbach, 2014). In connection with singular vector approaches, they could also reveal what observations (types and spatial distribution) would have the most impact on estimation and forecasting.

5.2 Coupled Data Assimilation

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 thickness measurement 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 modelling is a field that is rapidly evolving (e.g., Feltham, 2008; Hunke et al., 2010).

A number of sea ice DA systems are now available. Sequential systems have been targeted initially at assimilating remotely sensed sea ice concentration and velocities (e.g., Bertino et al., 2008; Caya et al., 2010). Adjoint-based coupled sea ice-ocean assimilation has produced initial one-year sea ice-ocean state estimates in a regional domain of the Labrador Sea and Baffin Bay (Fenty and 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 implications for seasonal ice extent predictability (Fenty and 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 sea ice and ocean, respectively (Panteleev et al., 2010).

5.2.2 Coupled Earth System Estimates

Because in most of the existing climate forecast efforts, the initialization is done in uncoupled mode, this leads to initialization shock in the coupled system, which potentially reduces the forecast skill. This suggests that coupled data assimilation (CDA) efforts in Earth system models will lead to improved use of ocean information for coupled forecasts ranging
from near-term to seasonal to decadal time scales. Coupled Earth system models seamlessly link together models of the oceans, atmosphere, sea-ice, land surface, the global carbon cycle and chemistry, and aerosols, to simulate changes in the Earth’s climate systems with ever-increasing precision (WMO, 2009).

Today, some pilot applications of CDA already exist (e.g., Sugiura et al., 2008; Fujii et al., 2009; Zhang et al., 2007, 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 ocean–atmosphere CDA system with a fully coupled GCM, using a “super”-parallelization technique for ensemble integrations. In perfect model experiments, the assimilation successfully reconstructs the 20th century ocean heat content variability and trends in most locations.

Karspeck et al. (2014), also using an EnKF, applied CDA to the problem of decadal predictions, however with mixed results, partly because a state estimated with a coupled EnKF remains dynamically inconsistent with the coupled system as long as model parameters are not improved as well, as can be done using a smoother approach. This 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 filtering 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 to a hindcast with 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 NCEP (Saha et al., 2010). As an example, ECMWF has implemented a pilot CDA system for production of coupled reanalyses of the Earth-system (CERA; Laloyaux et al., 2015) which is capable of assimilating a wide variety of ocean and atmospheric 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 in this respect, however, a fully coupled GCM inevitably generates rapidly growing modes, particularly in the atmospheric component, which make optimization of the simulated state of the atmosphere difficult. 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 use of 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 are required (multi-model ensemble).

### 5.3 Model Improvements

ODA procedures require best possible model representations for best performances. In turn, state estimation can contribute substantially to improving models and therefore need 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 syntheses efforts for helping to improve ocean and coupled climate models. One is to help improving 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. As an 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; Stammer, 2005; Liu et al., 2012; Menemenlis et al., 2005). 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., only 10% of the total cost reduction was found by Liu et al. (2012). Model tuning is less crucial for ocean models than for coupled models, where automatic tuning is an active field of research with 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 attempting 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., $z$ versus $z^*$ versus ALE), representation of kinematic boundary conditions
(nonlinear free surface with real water fluxes, compared to linear free surface with virtual salt fluxes), are as pertinent as are 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. Given space limitations, in this article we were only able to address 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, 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 all topics of equal importance that require separate review papers. 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 NWP. With less 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. Respective work for high-resolution or biological parameters would be an order of magnitude more difficult.

Problems often overlooked in many fields are those of expertise, continuity and especially resources required to further develop the fields. Resources required for technical developments in the various aspects of ODA and OSE in support of performing ocean syntheses are enormous, comparable to what atmospheric forecast centers require. There is a need for better software infrastructure that provides openly available algorithmic differentiation and other assimilation tools, and that would allow testing different options and methods. For instance, combining ensemble and variational methods in an effective manner, and improving model bias correction techniques are directions that need to be pursued. Furthermore, there is a need to share efficient minimization algorithms and observation operators to avoid duplicating efforts. In almost all cases, reaching accomplishments in the development of the infrastructure required close to a decade of sustained consortium efforts (Stammer et al., 2002a). This development has proven to be a large endeavor that requires expertise in ocean observations, modeling, assimilation as well as information technology, which needs to be sustained and to have a long-term perspective to be effective.
Acknowledgement:
Thanks go to Meike Demgen for help with some of the figures. This work was funded in part through the DFG-funded CliSAP initiative of the University Hamburg. Supported in part through the BMBF funded RACE and the EU funded NaCLIM projects. PH acknowledges continued support over the years through ECCO from NASA’s Physical Oceanography program. Additional support comes from grants NASA #NNX13AK88G, NSF #1434149, NOAA #NA130AR4310135, and DOE #SC0008060.
REFERENCES


34. Divakaran, P., Brassington, G., & Feng, M. (2010). Mesoscale Zonal Features In The Southeast Indian Ocean As Seen In The Bluelink Reanalysis 2.1. Centre for Australian Weather and Climate Research, GPO Box 1289, Melbourne, VIC 3001, Australia, 125.


40. Fairall C., and co-authors (2010). Observations to Quantify Air-Sea Fluxes and Their Role in Climate Variability and Predictability. In these proceedings (Vol. 2).


82. Lee et al. 2009


116. Piers et al. 19??


154. WMO (2009), 14th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracers Measurement Techniques (Helsinki, Finland, 10-13 September 2007) WMO TD No. 1487, April 2009.


<table>
<thead>
<tr>
<th>System</th>
<th>Intent</th>
<th>Configuration</th>
<th>Data Assim. Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFSR NOAA NCEP</td>
<td>INI/ORA(M[Period?])</td>
<td>1/2’ MOM4 coupled DA</td>
<td>3DVAR (T/SST/SIC)</td>
</tr>
<tr>
<td>C-GLORS CMCC</td>
<td>ORA[1993-2010]</td>
<td>1/2’ NEMO3.2 Forcing E1</td>
<td>3DVAR [SLA/T/SST/SIC]</td>
</tr>
<tr>
<td>ECCO-NRT JPL/NASA</td>
<td>ORAM</td>
<td>1° MIITgcm Forcing NCEP-R1</td>
<td>KF-FS (SLA/T)</td>
</tr>
<tr>
<td>ECCO JPL MIT/AER</td>
<td>state estimation</td>
<td>1’x1/3’ MIITgcm Forcing E1</td>
<td>adjoint [SLA/SST/T/SST/SIC]</td>
</tr>
<tr>
<td>GECCO</td>
<td>state estimation</td>
<td>1’x1/3’ MIITgcm Forcing NCEP-R1</td>
<td>adjoint [SLA/T/S/MDT/SST]</td>
</tr>
<tr>
<td>ECDA</td>
<td>INI/ORA[1979-]</td>
<td>1/3’ MOM4 coupled Coupled DA</td>
<td>EnKF (T/SST)</td>
</tr>
<tr>
<td>GloSea UK MetOffice</td>
<td>INI/ORA (1993-)</td>
<td>1/4’ NEMO3.2 Forcing E1/UKMO NWP</td>
<td>3DVAR [SLA/T/SST/SIC]</td>
</tr>
<tr>
<td>MERRA Ocean</td>
<td>INI/ORA</td>
<td>1/2’ MOM4 Forcing MERRA</td>
<td>EnOI (SLA/T/SST/SIC)</td>
</tr>
<tr>
<td>GODAS NOAA NCEP</td>
<td>INI/ORA[1979-]</td>
<td>1’x1/3’ MOM3 Forcing NCEP-R2</td>
<td>3DVAR [T/SST]</td>
</tr>
<tr>
<td>GLORYS</td>
<td>ORA</td>
<td>1/4’ NEMO3.1 Forcing E1</td>
<td>KF+3DVAR [SLA/T/SST/SIC]</td>
</tr>
<tr>
<td>K7-ODA(ESTOC) JAMSTEC/RCGC</td>
<td>ORA</td>
<td>1° MOM3 Forcing NCEP-R1</td>
<td>4DVAR (SLA/T/SST)</td>
</tr>
<tr>
<td>K7-CDA JAMSTEC/CEIST</td>
<td>ORA</td>
<td>1° MOM3 coupled Coupled DA</td>
<td>4DVAR [SLA/SST]</td>
</tr>
<tr>
<td>PEODAS CAWCR(BoM)</td>
<td>INI/ORA</td>
<td>1’x2’ MOM2 Forcing E40- NCEP-R2-NWP</td>
<td>EnKF (T/SST)</td>
</tr>
<tr>
<td>ORAS ECMWF</td>
<td>INI/ORA</td>
<td>1° NEMO3 Forcing E40-EI-NWP</td>
<td>3DVAR [SLA/T/SST]</td>
</tr>
<tr>
<td>MOVE-C MRI/JMA</td>
<td>ORA</td>
<td>1° MRI.COM2 coupled Coupled DA</td>
<td>3DVAR [SLA/T/SST]</td>
</tr>
<tr>
<td>MOVE-G MRI/JMA</td>
<td>INI/ORA</td>
<td>0.5’x1’ MRI.COM3 Forcing JRA-SS/NWP</td>
<td>3DVAR [SLA/T/SST]</td>
</tr>
<tr>
<td>MOVE-CORE MRI/JMA</td>
<td>ORA</td>
<td>0.5’x1’ MRI.COM3 Forcing CORE.2</td>
<td>3DVAR (T/S)</td>
</tr>
<tr>
<td>SODA U. of Maryland and TAMU</td>
<td>ORA</td>
<td>1/4’ POP2.1 E40-EI</td>
<td>OI (T/SST)</td>
</tr>
<tr>
<td>URO25.4 U. of Reading</td>
<td>ORA</td>
<td>1/4’ NEMO3.2 Forcing E1</td>
<td>OI [SLA/T/SST/SIC]</td>
</tr>
<tr>
<td>FOAM Met Office</td>
<td>MF</td>
<td>1/4’ NEMO UKMO NWP</td>
<td>3DVAR(SLA/T/SST/SIC)</td>
</tr>
<tr>
<td>BLUELINK Bureau of Meteorology Australia</td>
<td>MF</td>
<td>1/10’ MOM4 BoM NWP</td>
<td>EnOI [SLA/T/SST]</td>
</tr>
<tr>
<td>GOFS NAVOCEANO (USA)</td>
<td>MF</td>
<td>1/12’ HYCOM NWP</td>
<td>3DVAR [SLA/T/SST/SIC]</td>
</tr>
<tr>
<td>CONCETPS Environment Canada</td>
<td>MF</td>
<td>1/4’ NEMO NWP</td>
<td>SEEK(SLA/T/SST/SIC)</td>
</tr>
<tr>
<td>MERCATOR France</td>
<td>MF</td>
<td>1/4’ and 1/12’ NEMO ECMWF NWP</td>
<td>SEEK(SLA/T/SST)</td>
</tr>
</tbody>
</table>

Table 1: List of existing ocean syntheses. It includes the name of the system, the institution, their intent, the ocean model configuration and forcing, the data assimilation and the observations used. These independent efforts employ different methods for different purpose. The intent therefore has been classified in the Table into INI (initialization of coupled models forecasts - at monthly, seasonal or decadal time scales), ORA (ocean reanalysis for a finite period of time), ORAM (ocean reanalysis with real-time extension, used for monitoring), and MF (for marine forecasting).
Fig 1: Schematic of the differences between filters and smoothers in producing the estimated state.
Fig. 2 Estimated ocean heat content from several ocean syntheses at different depth ranges (from Balmaseda et al. 2015).
Fig. 3: Left: Time-mean AMOC stream function from 1960-2007 in depth/latitude space for a set of ocean syntheses. Positive (negative) contours indicate clockwise (counter-clockwise) circulations, respectively. Bold line is the zero contour, and the contour interval is 2 Sv. Right: Time series of AMOC anomaly at 1000 m depth at 45°N (top panels) and 26.5°N (bottom panels) for the set of ocean syntheses (left panels). The time-mean has been removed from each time series. The means (in Sv) are indicated within brackets at 45N and 26.5N (respectively) in the legends. Time series from RAPID are included for comparison. After Karspeck et al. (2015)
Fig. 4. Sea surface temperature in °C on 19-July-2014 from (right) MODIS satellite obtained from podaacftp.jpl.nasa.gov and (left) the corresponding 4-day forecast from the 1/36° Daily Iberian Biscay Irish Physical Bulletin created by Mercator Ocean and obtained at bulletin.mercator-ocean.fr.
Fig. 5. (Left) Sensitivity to temperature perturbations at a depth of 200 m 4 yr before the time of cost function evaluation, which is the near surface temperature at 100°W, 0°N. Values larger (smaller) than 0.005 (20.005) are shaded with dark (light) gray. The thick black line denotes the 16.8°C isotherm. (Right) 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. From Galanti and Tziperman (2003).
Figure 6. Forecast drift in SST (a) and skill in precipitation (b) in the Central Pacific from different forecast strategies: Full Initialization (red), Anomaly Initialization (pink), Momentum Flux Correction (green) and Momentum + Heat Flux Correction (blue). The best skill is achieved by the Momentum Flux Correction. From Magnusson et al. (2013).
Figure 7. Latitude–longitude section of density anomalies at a depth of 3010 m north of 108N at (top) \( t = \) 2 months, (middle) \( t = \) 7.5 yr, and (bottom) \( t = \) 20 yr. After Zanna et al. (2011).