The *Journal of Marine Research*, one of the oldest journals in American marine science, published important peer-reviewed original research on a broad array of topics in physical, biological, and chemical oceanography vital to the academic oceanographic community in the long and rich tradition of the Sears Foundation for Marine Research at Yale University.

An archive of all issues from 1937 to 2021 (Volume 1–79) are available through EliScholar, a digital platform for scholarly publishing provided by Yale University Library at https://elischolar.library.yale.edu/.

Requests for permission to clear rights for use of this content should be directed to the authors, their estates, or other representatives. The *Journal of Marine Research* has no contact information beyond the affiliations listed in the published articles. We ask that you provide attribution to the *Journal of Marine Research*.

Yale University provides access to these materials for educational and research purposes only. Copyright or other proprietary rights to content contained in this document may be held by individuals or entities other than, or in addition to, Yale University. You are solely responsible for determining the ownership of the copyright, and for obtaining permission for your intended use. Yale University makes no warranty that your distribution, reproduction, or other use of these materials will not infringe the rights of third parties.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
https://creativecommons.org/licenses/by-nc-sa/4.0/
Reconstructing the recent past ocean variability: Status and perspective

by Simona Masina¹,²,³ and Andrea Storto¹

ABSTRACT

In the last two decades, climate research has benefited from the continuous development of analysis systems dedicated to operational monitoring and forecasting, which opened up the possibility of exploiting the state of the art modeling and data assimilation tools to reconstruct and study the ocean during the past decades. This activity became feasible also thanks to the increasing availability of long time series of high-quality in situ and remotely-sensed observations. Retrospective analyses (or simply reanalyses or ocean syntheses), indeed combine quality controlled reprocessed ocean observations with a state-of-the-art ocean general circulation model (OGCM) using data assimilation methods to estimate the time-varying, three-dimensional state of the ocean. Ocean reanalyses benefit from data assimilation algorithms that are usually inherited from operational oceanography, although they require specific treatment in order to avoid spurious drifts stemming from instrumental or model biases. Unlike observation-only products, ocean reanalyses take advantage of time-varying atmospheric forcing, usually coming from an atmospheric reanalysis, and dynamical and physical balances implied by the OGCM. Here we give an excursus on the availability of global and regional ocean reanalyses, their applications, their strengths and weaknesses, and their future developments foreseen at the present time.

Keywords. Data Assimilation, ocean modeling, ocean observations

1. Introduction

There is an increasing need for estimation of the present and past three-dimensional state of the ocean, in the context of ocean monitoring and short-term forecasting, climate variability assessments, and predictability purposes such as the initialization of long-range forecasts (i.e., seasonal and decadal). However, the oceans remain seriously under-sampled, and observational time series are often of limited usefulness in generating the required ocean estimates and ocean change indicators because of the short periods of coverage and sparse geographical distributions. On the other hand, over the course of the past few decades, considerable advancements have been achieved in the development of ocean data assimilation...
techniques that combine ocean models, atmospheric forcing, and ocean observations. Consequently, several ocean data assimilation systems have been developed to estimate the time-evolving, three-dimensional state of the ocean. The retrospective reconstructions of past observations with a fixed ocean model and a fixed data assimilation scheme in which the best available parameterizations and assimilation options have been prescribed are known as ocean reanalyses (ORAs) or syntheses. Their production is a recent activity that started around the beginning of the 2000s (e.g., Carton, Chepurin, and Cao 2000a, 2000b). Since then, significant progress has been made, and today ORA production is an established reality in several research and operational centers, where reanalyses use advanced, multivariate data-assimilation schemes that allow assimilation of most of the available type of observations. There are low-resolution reanalyses (about 1°), spanning a long time period (typically 50 years) (Balmaseda et al. 2015), as well as higher-resolution products (about 1/4°) (Masina et al. 2017), which exhibit eddy permitting capabilities but are available for shorter records (usually the altimeter period from 1993 onward).

An increasing number of studies utilize the coarse-resolution ORAs available for longer periods to investigate various aspects of ocean variability. For instance, they have been applied to study the nature of sea level variability, mixed-layer heat balance, and trend and variability of upper-ocean heat and salt contents. Furthermore, many examples exist of ORAs’ beneficial impact on the initialization of coupled models for forecasts at seasonal and decadal timescales. They also serve downstream applications such as, for example, driving offline biogeochemical and fishery models, assessing the value of observing networks, and providing lateral boundary conditions for higher-resolution regional ocean general circulation models (OGCMs). Relevant reviews of ocean data assimilation in climate research were recently provided by Schiller, Lee, and Masuda (2013) and Stammer et al. (2016).

Thanks to satellite altimetry, sea level displacements associated with ocean eddies have been observed with a few-centimeter accuracy for more than two decades, and there is evidence, both from observations and modeling studies, of the impact of mesoscale dynamics (e.g., Le Traon and Morrow 2001) on climate and operational oceanography. That is why including mesoscale features in ORAs with constantly-increasing resolution is a fundamental advancement and contributes to our understanding of ocean and climate variability.

Nevertheless, one has to keep in mind that ocean observations before the 2000s are scarce (i.e., prior to the full deployment of Argo floats [Roemmich and Gilson 2009]) and that large uncertainties may still exist in ORAs, making a robust estimation of the ocean history with limited error bars still a major challenge.

A way to have access to both the variability and uncertainty estimates of reanalyses is to perform a multi-system ensemble of ORA. The advantage of dealing with a multi-system ensemble is that it is possible to estimate the uncertainties associated with the reanalysis system used and to gain insight on the signal-to-noise ratio of the ocean state estimations. The ongoing coordinated international efforts on ORAs intercomparison aim to provide recommendations for future reanalyses production by identifying the weaknesses of single existing systems and the suitability of the ensemble approach. Furthermore, they aim to give
feedback on how to improve the ocean observing system, assimilation methods, models, and surface fluxes, to promote interaction with the user community and to encourage the archiving of individual reanalysis products in public data repositories, freely available to all users.

This work is intended to provide a description of the state-of-the-art in ORAs and present some recent results on eddy-permitting reanalyses using, when possible, an ensemble approach. The work will also illustrate examples of possible validation strategies, with the purpose of showing the applicability of these products for a wide range of climate and downstream applications that require the assessment of the oceanic conditions at the global and regional scales in the last decades. Finally, we review strengths and weaknesses of various ORAs and conclude with some recent progress and cutting-edge research topics that will likely be exploited in the near future.

2. Concepts and methods

a. Definition of reanalysis

The traditional methods for monitoring long-term changes in ocean heat and salinity content, steric sea level, and related diagnostics are statistical analyses (or objective analyses) that perform interpolation of the available in situ observations, usually blended with climatology (e.g., Levitus et al. 2012). These methods do not use dynamical ocean models and have the advantage of being relatively simple and computationally cheap. However, objective analyses strongly depend on the correlation functions and data availability (Boyer et al. 2016). For instance, in data-sparse regions such as the Southern Ocean, the analysis can be unrealistically close to the climatology fields. Furthermore, information on the ocean dynamics is completely absent, or, in the best case, approximated by simple balances (e.g., the geostrophic balance in ARMOR3D [Mulet et al. 2012]). On the other hand, the dynamical constraints imposed by physical laws are critical for inferring unobserved quantities, such as currents, or to compute transport of tracers and integrated circulations through sections.

The operational oceanography community has been able to establish an efficient infrastructure for systematic acquisition and processing of in situ and space observations, assimilation of data into global and regional high-resolution ocean circulation models for real-time nowcasts and short-range forecasts, and delivery of products to users, along with advances in numerical models and data assimilation techniques.

The application of an ocean prediction system for the purpose of describing the evolution of the ocean state over a multi-decadal period for climate applications, however, is neither trivial nor immediate, although it has been historically fostered by long-range (i.e., seasonal to decadal) coupled prediction systems requiring ocean initial conditions. Different changes (e.g., changes in model resolution and implementation, data provision, or assimilation schemes) introduced intermittently in operational systems affect the accuracy and homogeneity of analyses and make it impossible to use the series of 3D ocean analyses
produced in real time for the exploration of ocean variability aspects, such as trends and low-frequency signals. Instead, a reanalysis is a 4D retrospective reconstruction of the ocean state obtained by a data-assimilative experiment that uses the same numerical model, the same forcing from atmospheric reanalysis, and the same assimilation scheme throughout the considered period. Reanalyses are, in practice, created via a “frozen” system that ingests all available observations over the period being analyzed. A reanalysis system differs from an operational system in that it does not aim to produce the best analysis of today, but the most coherent and consistent state of the past ocean. Operational analyses are usually subject to online model and assimilation updates, bugs fixing, or variations of atmospheric forcing, observations, and other inputs (e.g., bathymetry, runoff data, and ancillary files). Those variations might determine space-time inconsistency and undesired transient phenomena that undermine the final product’s quality and limit its possible applications. Reanalyses can therefore contribute to more precise attribution of the processes responsible for the observed variability, as well as the geographical patterns and magnitudes of the responses. The usefulness of reanalysis products to estimate climate changes and trends over several decades is, however, nontrivial to achieve, owing to the unavoidable evolution of the observing system in terms of sampling technology, sampling accuracy, coverage, resolution, and data quality. Even in the atmospheric community, the use of reanalysis for climate change assessment is not uncontroversial (e.g., Thorne and Vose 2010; Dee et al. 2011a). Over the duration of each reanalysis product, changes in the observation network or the introduction of new types of observations into the assimilation system can produce artificial variability and spurious trends. Because of this, the atmospheric community, for instance, adopted the strategy of assimilating only surface data within centennial reanalyses (e.g., Compo et al. 2011); the same idea has also been followed by Simple Ocean Data Assimilation (SODA) (Giese and Ray 2011) for the global ocean.

b. Observational requirements

Prior to the Array for Real-time Geostrophic Oceanography (ARGO) network being deployed, observing networks typically did not provide continuous coverage of the ocean above 2,000 m, despite the fact that some regions have benefited from specific monitoring instrumentation (e.g., the tropical Pacific through the Tropical Atmosphere Ocean/Triangle Trans-Ocean Buoy Network [TAO/TRITON] array) (McPhaden et al. 1998). Special measurement campaigns have offered nearly global coverage over selected time periods, notably the World Ocean Circulation Experiment (WOCE) (Ganachaud and Wunsch 2000). Specific satellite missions equipped with radiometers and altimeters—since 1981 and 1992, respectively—provided fundamental information about the sea surface. In the early 1990s the U.S./French mission Topex/Poseidon (T/P) provided, for the first time, a very precise altimeter system optimized for large-scale sea level observations. T/P revolutionized our vision and understanding of the ocean circulation at high resolution in near-real time. However, altimeter data can constrain only barotropic-type motions and the first baroclinic
mode, so they are complementary, rather than alternative, to in situ observing networks. In this situation, integrated quantities that are key climate indexes, such as transports, heat, and freshwater content, can be inferred only indirectly. The advancement in the field of observation and measurement collection, observation processing, and data exchange within many international projects and organizations (e.g., the Global Climate Observing System/Global Ocean Observing System [GCOS/OOOS], Ocean Sustained Interdisciplinary Time Series Environmental Observation System [OceanSITES], Global Sea Level Observing System [GLOSS], SeaDataNet, ARGO, the European Marine Observation and Data Network [EMODNET], among many others) made available long-time series of quality-checked in situ and remotely-sensed data that encouraged ORAs’ production and potential usage. Analogous progress in the atmospheric observing system, and in the field of atmospheric reanalyses, over the last decade made available qualified forcing data starting from the late seventies, without artificial shifts and spurious trends, that are continuously updated to present time. These products have allowed researchers to run ORAs without the need to merge different atmospheric datasets. However, the current ocean observing network does not yet seem to meet the requirements highlighted by many recent studies (e.g., Purkey and Johnson 2010; Balmaseda, Trenberth, and Källén 2013; Wunsch and Heimbach 2014), which stress the significant deep and abyssal ocean contribution to total ocean heat content variability, thus influencing the Earth’s climate.

There is a close relationship between the quality of reanalyses and the availability of consistent historical observations; thus, one of the key scientific issues is to produce consistent time series of reprocessed satellite observations of sea level (SL), sea surface temperature (SST), and sea ice (SI) coming from different types of instruments with different error characteristics. Such an activity is progressing in the framework of initiatives such as the European Space Agency (ESA) Climate Change Initiative (CCI), the Making Earth System Data Records for Use in Research Environments (MEaSUREs) programs through the NASA Earth Science Data Systems (ESDS), and the Copernicus Marine Environment Monitoring Service (CMEMS). Also, in situ observations need to be reprocessed to limit the impact of different data-set errors and biases on reanalysis estimation. Data reprocessing is still a developing field that necessitates the perfection of delayed mode (DM) quality-check procedures based on multi-data approach and cross checking. European projects, like Euro-Argo Improvements for the Global Monitoring for Environment and Security (GMES) Marine Service (E-AIMS) and SeaDataNet, funded efforts such as EN4 (Good, Martin, and Rayner 2013), and the worldwide-known World Ocean Database (Boyer et al. 2013) contributed to the development of DM quality-assessment strategies and data recovery. DM quality control is based on statistical methods that take into account prior knowledge (climatology) and comparison with other observations when available (ship based measurements or other nearby floats) or, eventually, with ORAs themselves. Furthermore, in situ measurements from most platforms (e.g., XTBs, Argo floats, CTDs, and moorings) are often affected by systematic biases whose optimal correction requires posterior calibration of the DM observations. This applies, for instance, to fall-rate biases in bathythermographs (e.g.,
c. Methods

In essence, the systems that are producing 4D historical reconstructions should be different from those used in real-time. An obvious difference when operating an assimilation system in retrospective mode lies in the possibility of using both “past” and “future” observations. This characteristic is often referred to as a “smoother,” as opposed to a “filter,” that uses observations measured before the analysis time, which is a typical configuration of operational systems. Adjoint methods (or ocean state estimation methods, e.g., Thacker and Long 1988; Wunsch and Heimbach 2013) are examples of smoothers that are used by the Estimating the Circulation and Climate of the Ocean (ECCO) consortium (see Forget et al. 2015 for the latest release of the ECCO reanalysis) among others, and are specifically conceived for reanalyses. Ocean state estimations minimize analysis errors over the whole reanalysis periods, resolving the all 4D ocean state at once through iterations involving long tangent-linear and adjoint integrations. An advantage of these methods is their ability to provide a temporally consistent state of the ocean, allowing conservation of the water properties throughout the reanalysis period, with the additional possibility of constraining the long-term climatology. Their computational cost and the use of linearized physics during the optimization procedure make them more difficult to implement in high-resolution configurations.

Four-dimensional variational data assimilation (4DVAR) is an approximation of the adjoint method in which the reanalysis period is subdivided into short time-windows, within which 4DVAR finds the optimal state estimation. Its computational demand is thus generally lower than adjoint methods. Simpler methods like 3DVAR or Optimum Interpolation (OI), which are still widely used thanks to their relatively low computational demand, neglect the temporal dimension within individual assimilation time-windows.

Kalman-filter and derived assimilation schemes (used either in filter or smoother mode, depending on the definition of assimilation time-window) have the advantage of propagating the error covariance matrixes forward in time. For now, only the Ensemble Kalman filter has a computational cost that makes it affordable for reanalysis applications (e.g., Yin, Alves, and Oke 2011; Chang et al. 2013). A clear advantage of Ensemble Kalman filter schemes is the modulation of the background-error covariance matrix with time, unlike 3DVAR and OI, consistently with the “error of the day,” which, from the reanalysis perspective, depends also on time-varying observation sampling. The development of ensemble square-root filters (e.g., Whitaker et al. 2004) has also been fostered by the need to avoid observation perturbation in the ensemble-generation design, which was found to be particularly critical for multi-decadal atmospheric reanalysis applications (Compo et al. 2006). Meanwhile, the ocean modeling community is recognizing, since recently, the potential of using stochastic physics for probabilistic simulations (e.g., Bessières et al. 2017 and references therein), which could provide a more robust framework for ensemble generation in ORAs.
Hybrid ensemble-variational methods originally developed in the numerical weather prediction (NWP) context (Hamill and Snyder 2000), in which the static variational and ensemble-derived, flow-dependent error covariances are combined, are likewise appealing for reanalysis applications spanning several decades. Alternatively, simpler approaches that inflate the background-error covariance matrix depending on the observation coverage (Yang et al. 2016) should be considered when variational schemes are applied to multi-decadal reanalyses.

Variational (especially 3DVAR), OI, and Kalman filter–derived data assimilation schemes generally do not protect the reanalysis system from model initialization shocks or prevent conservation problems, although practical techniques, such as the incremental analysis update (i.e., applying the analysis corrections as a forcing, spread over a certain period [Bloom et al. 1996]), the use of overlapping assimilation time-windows (Fisher et al., 2011), or the inclusion of additional penalty terms to conserve properties (e.g., Takacs, Suárez, and Todling 2016) may limit these problems.

Finally, ORA formulations should include strategies to avoid or limit drifts and biases that may compromise the long-term reanalysis diagnostics, which are generally neglected in operational ocean analysis systems. Whereas biases in atmospheric reanalyses mostly arise from changes in the observing network characteristics and have been treated for a long time, structural biases mostly concern ORAs. These are typically due to biases in the atmospheric forcing propagated onto the ocean, unresolved processes, or uncertainty in the parameterizations and are difficult to correct. Besides the obvious improvements in the formulation and implementation of ocean general circulation models, from which ORAs will clearly benefit, methods have been proposed to include bias-correction methods in ORAs. These attempts rely on characterizing the model biases, possibly with independent observations, and correcting the reanalysis system accordingly, either prior to data assimilation (Balmaseda et al. 2007; Storto, Masina, and Navarra 2016) or through inclusion of model bias parameters in the data-assimilation formulation (see Dee 2005 for a general discussion and Lea et al. 2008 for an ORA application). Weak-constraint 4DVAR (e.g., Tremolet 2006) may also be conceived in this context. An intrinsic limitation of all these bias-correction methodologies is the obvious need to estimate the bias, which requires a set of unbiased (“anchoring”) observations. The global ocean observing networks does not allow such a procedure for reanalyses spanning multi-decadal periods, and suboptimal schemes relying either on climatology or present-time bias estimates should be considered. The improvement of bias-correction strategies will certainly be pursued in the future production of ORAs, and they remain among the most critical issues for the applicability of ORAs to climate-oriented studies.

d. Uncertainty assessment

Quantification of the uncertainties of ORAs is a fundamental goal for qualifying the reanalysis quality and the feasibility of their adoption in climate studies. This is particularly
relevant because observational sampling—including not only the oceanic observing network but also the one that drives the accuracy of the atmospheric forcing—has changed dramatically over the last decades, generally increasing the accuracy of ORAs with time (see Storto and Masina 2017 for a more detailed discussion). This seems particularly crucial for the ocean, where poorly observed areas, periods, or vertical regions of the global ocean may in turn foster low confidence in the reanalysis results. Within many international initiatives (e.g., Ocean Reanalyses Intercomparison Projects [ORA-IP] in the see next section), quantification of the reanalysis accuracy is often driven by a fit-for-purpose strategy, where process-oriented metrics are built to determine whether the reanalysis is suitable for a certain application.

In terms of the uncertainty assessment per se, ensemble data-assimilation methods, despite their limited application in global ORAs owing to their high computational demand for eddy-resolving, or even eddy-permitting, horizontal resolutions, directly provide an estimation of the analysis-error covariances; thus, they are readily suitable for uncertainty assessment. Alternatively, variational data assimilation can be run in ensemble mode (small ensemble size, with members generated through perturbation of relevant input dataset; e.g., Balmaseda, Mogensen, and Weaver 2013) to provide information on the uncertainty deducted from the ensemble spread. However, perturbation-generation methods, as long as localization procedures limit spurious covariances, may have a large impact on the analysis-error estimates themselves (Bowler 2006). The exploitation of multi-model reanalysis systems to estimate the ensemble mean (EM) accuracy from the ensemble spread seems to be a sensible way to diagnose the accuracy of reanalyses in a general sense (Balmaseda et al. 2015), as exemplified later in Section 3. However, practical issues linked to the availability of individual reanalyses may limit their routine application, along with the fact that, due to the different resolution of the reanalyses participating the multisystem, one can question the true spatial scale of the signal contained in the error-uncertainty estimates.

Promising results in variational data assimilation arise from approximating, somehow, the inverse of the Hessian matrix of the variational cost function, which is known to equal the analysis-error covariance matrix (e.g., Navon 2009). This equivalence can be exploited in a variety of ways, e.g., through stochastic (randomization) procedures or by retaining the dominant modes of the inverse Hessian matrix (see Bousserez et al. 2015 for a comparison of possible strategies).

Minimization algorithms can also be exploited in order to provide a reduced-space approximation (i.e., approximated by the leading eigenvectors). This is the case, for instance, for the conjugate gradient minimizer formulated as a sequence of Lanczos vectors (Fisher and Courtier 1995) or the modified version of the limited-memory quasi-Newton minimizer proposed by Veerse (1999). All these approaches rely on a reduced-space definition of the Hessian matrix, whose size within realistic geophysical applications would not be possible to store nor to compute explicitly, meaning that the estimate cannot span the full space of the analysis-error covariance matrix. Alternatively, Moore et al. (2012) proposed the combined use of the adjoint model and perturbed 4DVAR analyses to assess analysis error
for linear scalar functions such as transports and vertical integrals. A similar strategy of applying reduced-space error estimation on selected metrics is also proposed by Kalmikov and Heimbach (2014). Within ocean applications, the ROMS model implements the online estimation of analysis-error covariance matrix (Moore et al. 2011; Moore, Arango, and Broquet 2012), although evaluation of the approximate analysis errors in long-term global reanalyses has not been accomplished yet.

3. Global products: the ensemble approach

There is an increasing need to understand the consistency of ORAs and to provide uncertainties. In most cases, however, the theoretical estimation of posterior or analysis-error covariance still has a computational cost that limits the applications to limited-domain GCM applications and to only few parameters. Using an ensemble of ORAs of limited size, from the same system (Masina et al. 2011; Balmaseda, Mogensen, and Weaver 2013) or multiple systems (Stammer et al. 2010; Balmaseda et al. 2015; Masina et al. 2017), is an alternative, more-affordable way to access uncertainty. The spread among the ensemble members for a diagnostic variable (e.g., heat content change or mixed layer depth) can potentially be used as a measure of the uncertainty estimate. At present, there are two ongoing efforts aimed at exploiting existing and new ORAs for a variety of purposes, such as quantifying improvements in reanalyses quality, quantifying uncertainty, and defining indices for ocean monitoring:

1) The Ocean Reanalyses Intercomparison Projects (ORA-IP), undertaken by the Global Ocean Data Assimilation Experiment (GODAE) Ocean View (GOV) and Climate Variability and Predictability (CLIVAR) Global Synthesis and Observations Panel (GSOP) communities. The intercomparison includes mainly ocean reanalyses produced after 2010 and benefits from the latest atmospheric reanalysis products and quality-controlled data sets. Rather than following a fixed protocol, ORA-IP exploits existing reanalysis products, which are available mostly at coarse resolution (typically $1^\circ \times 1^\circ$), taking advantage of the spontaneous diversity to gain insight into how robust our knowledge of the ocean is. As part of this effort, a large suite of indices and diagnostic quantities obtained from various ORA products are compared and evaluated, using observations when available.

2) The CMEMS, continuing the E.U.-funded MyOcean/MyOcean2 project, which provided a series of validated eddy-permitting and/or eddy-resolving global and European Seas ORAs covering the recent “altimetry era” (namely 1 January 1993 onward). These products are targeted toward not only the climate community, but also to the fishery and offshore industry, and they foster intermediate and downstream services for the benefit of agencies with environmental assessment responsibilities.

Both coordinated efforts try to deal with the issues related to the uncertainties that still exist in the ORAs. In general, there are many sources of uncertainty in ocean modeling (e.g., parameterized processes, initialization, atmospheric forcing and numerical implementation, etc.) that lead to differences between the true values (unknown) and the measured or modeled
values of the physical properties. Data assimilation can, in principle, reduce the uncertainties by combining dynamical models and observations. Thanks to the available global ocean coarse-resolution and eddy-permitting reanalyses, an EM can be computed, from which the spread is used to infer the “reliability envelop,” or the uncertainty, for each ocean indicator. The EM of ORAs can indeed be evaluated as one single system regarding its reliability in reproducing the climate signals, where both uncertainty and its variability are assessed through the ensemble spread (ES). The ocean state can then be analyzed and discussed based on the ORAs ensemble compared to the observation-only-based products, rather than on the same comparison done using each single reanalysis. This evaluation can help to establish the strategy of a single, versus an ensemble, reanalysis system.

a. The ORA-IP

The ORA-IP initiative aims to quantify the degree of maturity of the latest vintage of ORAs, focusing on the identification of their strengths and weaknesses, with the twofold objective of providing recommendations to the user community and highlighting priorities for the next generation of ORAs. The initiative is a follow-up to the intercomparisons carried out during the second half of the 2000s and documented by Stammer et al. (2010). Previous comparisons provided indications of the quality of reanalyses with regards to only a few parameters and included fewer products. For instance, Xue et al. (2012) concentrated their analysis on the upper-ocean heat content, finding robust heat-content climate signals in only the upper ocean (top 300 m), to the exclusion of the Southern Ocean before the Argo period.

ORA-IP includes a vintage of reanalyses produced as for 2012. The target period for the comparisons is from 1993 to 2010, corresponding to the satellite altimetry era. The general strategy is to assess the ORAs performance with respect to variables of recognized importance for ocean climate. Table 1 reports the parameters that have been compared, along with the relevant reference. A summary of the intercomparison at an initial stage of completion is available in Balmaseda et al. (2015). Emphasis is also given to the assessment

<table>
<thead>
<tr>
<th>Ocean variable</th>
<th>Ocean Reanalysis Inter-comparison Project (ORA-IP) references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth of the 20°C isotherm</td>
<td>Hernandez et al. (2014b)</td>
</tr>
<tr>
<td>Mixed layer depth</td>
<td>Toyoda et al. (2017a)</td>
</tr>
<tr>
<td>Ocean heat content</td>
<td>Palmer et al. (2017)</td>
</tr>
<tr>
<td>Salinity</td>
<td>Shi et al. (2017)</td>
</tr>
<tr>
<td>Sea ice</td>
<td>Chevallier et al. (2017)</td>
</tr>
<tr>
<td>Sea level</td>
<td>Hernandez et al. (2014a)</td>
</tr>
<tr>
<td>Steric sea level</td>
<td>Storto et al. (2017)</td>
</tr>
<tr>
<td>Surface heat fluxes</td>
<td>Valdivieso et al. (2017)</td>
</tr>
</tbody>
</table>
of the EM as a product itself, following the ideas proposed by Stammer et al. (2010), and to
the evaluation of the signal-to-noise ratio—or signal-to-spread ratio—defined as the ratio
between the variability, or signal, of the EM and the ensemble spread, i.e., the deviation of
the members from the mean (Balmaseda et al. 2015). This approach allows for detection of
robust climate indexes and areas of significant disagreement among the products. A total
of 19 ORAs participated in the intercomparison, along with a few (2 to 4, depending on the
parameter) objective analyses. Noticeably, three products comprise the ocean component
of a coupled ocean-atmosphere model simulation that assimilates oceanic observations.

In general, the ORA-IP assessments indicate good agreement and reliability of reanalysis
products for the ocean variables directly constrained by oceanic observing networks, such
as upper-ocean heat content, sea level, and sea-ice concentration, especially at the seasonal
scale. Regions and parameters that are observed only partly, indirectly, or discontinuously
generally suffer from a lack of consistency among the products, dramatically visible at
the inter-annual scale. This applies to the deep ocean (i.e., below 700 m), the Southern
Ocean, salinity estimates before the development of the Argo observing network, and sea-
ice thickness. Other parameters, such as the mixed layer depth (MLD) and sea-ice velocity,
are simulated reasonably well, although they bear the spread from the atmospheric and
model formulations that they strongly depend upon. As an additional remark, the EM and
its temporal variability proves accurate and reliable within all the comparisons, fostering
the adoption of a multi-model ensemble approach for the reconstruction of the ocean climate in
past decades. For some parameters (steric sea level, salinity variability, MLD), the ensemble
of reanalyses is found to outperform the objective analyses with respect to independent
estimates.

To evaluate steric sea level from reanalyses and objective analyses, satellite-derived inde-
dependent estimates for a recent period (2003–2010) are constructed (Storto et al. 2017). Dur-
ing this validation period, the EM of reanalyses proves more skillful than the ensemble of
objective analyses, especially in the Southern Ocean, suggesting that the lack in the South-
ern Ocean observing system may be partly recovered in reanalyses, through atmospheric
forcing and model dynamics and physics. Although linear trends for both thermosteric and
halosteric sea levels from the EM are in agreement with recent estimates (e.g., Hanna et al.
2013), the discrepancy between products is high. Regional results for either the halosteric
sea level or the steric sea level below 700 m indicate scarce agreement between the reanal-
yses; hence, the poor robustness of the estimates. As an example, Figure 1 shows the 1993
to 2010 linear trends of thermo- and halosteric sea level from the reanalyses and objec-
tive analyses participating in the intercomparison of Storto et al. (2017), along with their
respective EMs (REAENS and OAENS). Such a comparison illuminates the significantly
large spread of the halosteric trends from reanalyses that exceeds the mean trend, unlike
the thermosteric trend.

The air-sea heat fluxes study within the ORA-IP initiative (Valdivieso et al. 2015) compares net, radiative (short and long wave), and turbulent (sensible and latent) fluxes calculated by the ORAs for use in the surface-boundary heat–flux conditions. Unlike
Figure 1. Linear trends (1993–2010) of (a) thermosteric and (b) halosteric sea level from the ocean reanalyses (ORAs) and objective analyses included in the Ocean Reanalyses Intercomparison Projects (ORA-IP) comparison (Storto et al. 2017). Uncertainty of the trends estimated from bootstrap is represented by dashed lines. Ensemble means of the objective analyses (OAENS) and reanalyses (REAENS) are also shown, with the corresponding ensemble spread.

atmospheric reanalyses (Liang and Yu 2016), ORAs are generally not designed for heat flux studies, and their validation is not straightforward owing to the scarcity of in situ measurements (from only a few buoys), whereas global estimates may be inferred only indirectly, by means of Earth’s energy budget considerations (e.g., Allan et al. 2014). Nonetheless, the global-mean net heat flux from ORAs offers slightly less disagreement with respect to the heat flux from atmospheric reanalyses, or blended products, that use satellite estimates of radiative fluxes (Figure 2). This result is strengthened when data-assimilation analysis increments are summed to the net heat flux, because of the compensating effect of data assimilation and atmospheric forcing biases. Ocean state estimates like ECCO and German ECCO (GECCO) (e.g., Forget et al. 2015) already include flux adjustment in response to ocean data assimilation, providing results somehow closer to indirect estimates. ORAs show significant agreement in the reproduction of the seasonal cycle of air-sea fluxes. However, signal-to-noise ratios smaller than one were found in the analysis of the heat flux inter-annual variability, although most of the reanalyses are able to reproduce anomalous events, such as the negative Pacific decadal oscillation (PDO) in 2008 or the persistent negative North Atlantic oscillation (NAO) period in 2009. This comparison also indicates that errors in the latent heat flux at Stratus Buoy dominate the net heat-flux errors, suggesting that local inaccuracies in the air-sea fluxes may be partly ascribed to systematic errors in ocean-surface winds from atmospheric products.

The upper-ocean salinity comparison (Shi et al. 2017) stresses the agreement of reanalyses at low latitude, in contrast to the large spread in the Southern Ocean attributable to the poor
oceanic and atmospheric observation network, especially in the 1990s. Although regional trends and patterns of inter-annual variability show some degree of robustness, especially in the tropical oceans, globally-averaged values of salinity and their trends are quite dispersed, highlighting that the salinity performance in reanalyses is still far from being suitable for climate applications.

Although the upper-ocean (top 700 m) heat-content seasonal cycle and inter-annual variability seem well-captured by the majority of ORAs, Palmer et al. (2017) show that little agreement is found at depths below 700 m. This is likely the result of different representations of deep-water formation processes, and of model adjustments and drifts that lead to the decrease in signal-to-noise ratio with depth.

MLD comparison (Toyoda et al. 2017a) suggests the superiority of reanalyses, as an ensemble, to objective analyses, partly because of the positive impact of the higher horizontal and vertical resolutions of the former. Although individual products show significant
biases, the seasonal cycle and inter-annual variability of the EM are in good agreement with independent MLD estimates (MIxed Layer data set of Argo, Grid Point Value [MILA-GPV]) (Hosoda et al. 2010). Reanalyses seem problematic at high latitudes, owing to not only the scarcity of observations, but also to sea-ice modeling deficiencies.

The Arctic sea-ice intercomparison (Chevallier et al. 2017) provides guidelines for the evolution of sea-ice modeling and data assimilation. It appears that sea-ice cover, at least when assimilated into the reanalyses, presents high reliability, and part of the spread may be explained by differences in the assimilated sea-ice concentration dataset. On the contrary, the skill of sea-ice thickness compared with the few existing validating data (i.e., IceSat) is generally poor, because of a combination of factors. Among these, the sea-ice dynamics seem crucial, especially to avoid the accumulation of ice in the Beaufort gyre, which is typically present in most reanalyses jointly with too-thin ice near the North Pole. To this end, both modeling deficiencies—among them sea ice rheology, air-ice and ice-ocean stress parameterizations—and atmospheric forcing accuracy at high latitudes may play a role. It is also shown that data assimilation systems are not able to reasonably constrain the sea-ice thickness when assimilating sea-ice concentration, motivating future research toward properly-established, multivariate sea-ice assimilation (e.g., Massonnet, Fichefet, and Goosse 2015). The different maturity of ORAs in capturing sea-ice concentration and thickness in the Arctic Ocean is sketched in Figure 3. The figure depicts the September evolution (1993–2010) of sea-ice extent and volume from reanalyses, their EMs, and validation datasets, suggesting that, although the ensemble spread of the sea-ice extent is generally limited, that of the sea-ice volume is significantly large.

In order to disentangle intrinsic observational deficiencies from inadequacies of ocean models, data-assimilation systems, and atmospheric forcing, it is highly desirable that ORA-IP intercomparison activities be sustained and repeated in the future, with a new vintage of ORAs. This will guide the reanalysis community in documenting and, eventually, understanding and attributing the advances in the representation of key climate parameters within ORAs.

b. Eddy-permitting reanalyses

Since the start of the twenty-first century, there has been an increasing interest in understanding the mesoscale processes, supported by the evidence that eddies play a role in the meridional transport of heat (e.g., Souza et al. 2011; Smith et al. 2000; Valdivieso et al. 2014) and that ocean mesoscale features have an impact on atmospheric winds (e.g., Chelton et al. 2004; Maloney and Chelton 2006). Eddy variability is still poorly represented in global ocean models, despite its acknowledged important contribution to oceanic variability and its expected impact on climate variability in the upcoming generation of coupled models, which include eddying oceans. The next generation of operational long-range climate-prediction systems will implement eddy-permitting ocean models, and it is therefore urgent...
Figure 3. Arctic Ocean sea ice extent (a) and volume (b) for the period of 1993 to 2010 from the ORAs included in the Ocean Reanalyses Intercomparison Projects (ORA-IP) comparison (Chevallier et al. 2017). Single products are shown with thin lines; ensemble means and spreads are shown with red lines and orange shades, respectively; reference observation–derived time series from National Snow and Ice Data Center (NSIDC) sea-ice concentration and Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) sea-ice thickness datasets, respectively, are shown in blue.

to assess the capability of the global-ocean ORAs to be able to provide good-quality initial conditions for such systems.

Here, we present a coordinated intercomparison effort, carried out within the MyOcean global ORA activity (Masina et al. 2017), that provided a series of eddy-permitting global–ocean ORAs at 1/4° horizontal resolution, constrained by assimilation of observations, and covering the recent period during which altimeter data are available (the period starting with the launch of TOPEX POSEIDON and ERS-1 satellites at the end of 1992). Four available global eddy–permitting ORAs (from CMCC, University of Reading, Mercator Océan, and
European Center for Medium range Weather Forecasting (ECMWF) are built using state-of-the-art ocean data assimilation systems. They assimilate, in different ways, reprocessed observations of sea surface temperature (SST), in situ temperature and salinity profiles, sea level anomaly (SLA), and sea-ice concentration (SIC). All the ORAs cover at least the “altimetric era” (namely 1 January 1993 until 31 December 2011), but some are also available for longer periods into the past and will be continuously updated in the framework of CMEMS. One of main goals of the reanalysis community should be to harmonize the validation methods used for the ORA assessment, following user requirements and feedback with the objective of ensuring that the accuracy of products is adequately monitored and that information on product quality is scientifically sound, consistent, useful, and effectively communicated. The adoption of shared validation strategies is very important for the user community because it enhances confidence in products and encourages their potential adoption. Each individual global eddy-permitting ORA is assessed using the standard set of diagnostics, called “Marine EnviRonment and Security for the European Area (MERSEA)-GODAE metrics,” which were designed to present an overview of the ocean and sea-ice dynamics, and to evaluate prediction systems’ quality, consistency, and performance (Crosnier and Le Provost 2007). Each eddy-permitting ORA is consistently verified with the current knowledge of ocean circulation and climatologies, and performance and quality are measured through comparison with reference observational data sets identified by validation experts. This strategy has been proven to provide an efficient intercomparison of global ORAs and, if applied to different regional systems, it promotes products’ harmonization and advancement. It helps to identify system weaknesses and determine the necessary improvements for future updates.

Eddy-permitting reanalysis systems differ from data assimilation schemes implemented to control the ocean state (observational dataset, data assimilation method itself, background error covariances, observation errors, bias correction schemes, etc.) mainly in that they share a similar OGCM configuration (ORCA025 with 75 levels for three ORAs and 50 levels for one reanalysis), the same model (Nucleus for European Modelling of the Ocean [NEMO]) (Madec 2008), and the same surface-forcing data (ERA-Interim; Dee et al. 2011b). Therefore, the spread of the ensemble of ORAs is assumed to be mainly representative of our ability to gauge uncertainty in the assimilation methods. However, part of the uncertainty comes also from the assimilated observations themselves, the spin up, and the surface forcing being treated in different ways (bulk formula and corrections vary from one system to the other). In summary, although this is not a fully multisystem approach because of the same ocean model, the ensemble does partly span the uncertainty linked with ocean model parameterizations and initial conditions, even if it is not possible to assess the contribution of the individual sources of uncertainty to the ensemble spread.

As an example of this approach, we show here the analysis on sea surface salinity (SSS) (Masina et al. 2017). The evaluation of the SSS trends, unlike SST, is more difficult because there is no reliable reprocessed SSS product available for the period of 1993 to 2011, owing to the scarcity of data. According to the EM, the positive SSS trend over most of the
Atlantic and tropical Pacific Oceans seems robust, at least considering the signal-to-noise ratio, even if the amplitude of the trend is overestimated with respect to the observation-based products EN3 (Ingleby and Huddleston 2007; Figure 4). The only large regions where the trend is significantly negative are the Southern Ocean south of 40°S and the North Pacific Ocean. At high latitudes in the Northern Hemisphere, the trend is also negative but the high spread among the ORAs does not allow the drawing of any robust conclusion in this region, which is covered by sea-ice, and the disagreement might be caused by sea-ice misplacement. In general, the patterns of the regional trends are consistent with those derived from an observation-only estimate of SSS over the period from 1950 to 2008 by Durack and Wijffels (2010, fig. 5). We note, in particular, increasing salinity in the Atlantic and a freshening in the extratropical Pacific and Indonesian region, consistent with an amplification of the existing inter-basin salinity contrast and, perhaps, corresponding with an amplification of the hydrological cycle. However, the analysis of our ensemble indicates that these trends are associated with higher uncertainties (higher spread) in the South Pacific Intertropical Convergence Zone (ITCZ) and in the Atlantic ITCZ, i.e., in regions with high precipitation rates.

The eddy-permitting capability of the estimates presented here allows us to address other important aspects of the “ocean climate,” like large-scale turbulence. One of the major reasons for producing global ORAs at higher resolution than those typically used for climate purposes is indeed to try to resolve currents, in particular boundary currents, in a more accurate way and, therefore, allow more accurate estimates of mass and heat transports, not only in a zonal mean sense, but also with spatial variation. We have therefore analyzed the time-mean eddy–variability of the ORAs by means of the comparison between their eddy kinetic energy (EKE) and values obtained from the Ocean Surface Current Analyses Real-time (OSCAR) surface velocities (Bonjean and Lagerloef 2002). EKE is defined here as the difference between the total kinetic energy and the mean kinetic energy derived using

Figure 4. Sea surface salinity (SSS) linear trends (1993–2011) from (a) the ensemble of the global ocean reanalyses and (b) SSS gridded observations from EN3_v2. The left panel also shows signal-to-noise contours, i.e., dashed contours correspond to negative trends and solid contours correspond to positive trends greater than the ensemble spread (ES).
the velocity, averaged in time, over the period of 1993 to 2011. The EM shows realistic patterns of the mesoscale variability at large scale, with quite a good representation of the eddy energy associated with both the Gulf Stream and the Kuroshio pathways, the equatorial currents, and also with the highest EKE regions of the Southern Ocean (Masina et al. 2017). The EKE structure can also be used as a good indicator of the mean flow structure, and the EM also shows good agreement in terms of the representation of the longitudinal extent and the separation from the coast of the mean western boundary currents. However, some differences characterize each product. The zonal mean average suggests a lower level of eddy variability in the subtropical gyres from some products, likely owing to the application of a super-obbing procedure to the altimeter data, in order to reduce the weight given to the altimeter in relation to the in situ observations (not shown). However, when compared to a control run (C-GLORS4-CTRL, i.e., as C-GLORS4 but without data assimilation), it is clear that the amplitude of the EKE zonal average of each reanalysis is significantly higher at all latitudes between 60°S and 60°N (Fig. 5). We can therefore conclude that the EM gives a realistic representation of both the spatial pattern and strength of the small-scale variability at the global scale, and that the improvement is largely because of the assimilation of observations that can introduce variability that otherwise might only appear in models with much higher spatial resolution.

4. High-resolution regional products

Thanks to recent advances in computational resources, several national centers have developed high-resolution ocean analysis and forecasting systems that also operate on regional scales and provide initial estimates of the mesoscale ocean state using the
eddy-permitting and eddy-resolving Global Data Assimilation Experiment (GODAE), and now GODAE Ocean View (GOV), operational framework. Some of these operational systems have recently started to also provide ORAs that share much of the data-assimilation methodology, including the resolution, which is in most cases eddy-resolving. The advent of MyOcean and the following CMEMS, which led to the consolidation of operational oceanography in Europe, concurred with the advance of ORA activities at the regional scale over the European Seas (http://www.copernicus.eu).

Similar projects are Australia’s Bluelink analysis and forecasting system (http://wp.csiro.au/bluelink) and the Japanese Multivariate Ocean Variational Estimation system/Meteorological Research Institute (MOVE/MRI). The Bluelink ReANalysis (BRAN) system (Oke et al. 2013) uses a near-global configuration with 1/10° grid spacing around Australia with an Ensemble Optimal Interpolation (EnOI) system to produce eddy-resolving reanalyses that, despite the chaotic nature of the eddy-scales in the ocean, are suitable for a range of applications and process studies.

The Japanese system, which reaches 1/10° resolution in the Western North Pacific (Usui et al. 2015), uses a 4DVAR assimilation scheme and has been shown to improve mesoscale variability, in particular short-scale variability, such as small-scale Kuroshio fluctuations. The eddy-resolved (1/10°) long-term (1982–current) ORA, FORA-WNP30 (Four-dimensional Variational Ocean ReAnalysis for the Western North Pacific over 30 years), covering the Western North Pacific Ocean (117°E–160°W, 15°N–65°N), is newly produced using cycled 4DVAR data-assimilation system with quasi-10-day assimilation period. The reanalysis project was led by a joint research group of the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) and Meteorological Research Institute, Japan (MRI). Validation studies against independent data indicate that temporal variation of the Kuroshio meandering path off the southern coast of Japan is well-reproduced throughout the entire analysis period (presented at the CLIVAR/JAMSTEC Workshop on the Kuroshio Current and Extension System: Theory, Observations, and Ocean Climate Modelling, 12–13 January 2016).

Regarding the European marginal seas, around 2012, the first generation of regional reanalysis products covered a heterogeneous time-period, mostly less than 20 years, and were produced without a coordinated approach. The different capabilities of the regional monitoring and forecasting centers (MFCs) to implement reanalysis systems depended on the availability of the necessary input data like atmospheric forcing fields, remote and in situ observations, and model data for assimilation and boundary conditions, respectively. One of the main outcomes at the end of the comprehensive MyOcean project was the release of regional ORA products covering all the European marginal seas and sharing similar validation methodology (Simoncelli et al. 2016). The regional products are being updated with a delay of about one year. This is mainly because of ORA system dependencies, including the availability and time needed to retrieve all the input fields (atmospheric forcing, boundary conditions, observations for assimilation), the product quality assessment, and the dissemination time.
The extensive assessment of regional ORAs through common metrics allowed the identification of their criticalities, but, at the same time, it promoted user uptake, increasing user understanding and confidence. Regional ORAs are, in fact, among the most required and downloaded products by users because of their wide range of applicability.

The main criticality characterizing regional ORAs is time-consistency, for reasons such as changes in lateral boundary conditions, introduction of new observational data sets into the assimilation line, and change in quality and quantity of observations over time. The challenge is to ameliorate product consistency in space and time, in particular, through the assimilation of reprocessed datasets on one hand, and through the consolidated use of lateral boundary conditions from global ORAs on the other.

So far, the assessment of regional product quality has been performed principally on a monthly basis, but higher-frequency diagnostics (daily and hourly) have been introduced, for example, in Iberian-Biscay-Irish (Sotillo et al. 2015) and Mediterranean ORAs. The quality of monthly and higher–frequency products could be very different for each released parameter, and users need to know both skills in order to maximize the data exploitation. In the meantime, MFCs carry out automated diagnostic procedures that enable users to assess the quality of the ORA with a short delay to the production time, assuring continuous monitoring of ORA status and accuracy. Another frequent user requirement is the assessment of ORA versus operational analyses, especially in recent years, to orient the user in choosing the best products for specific purposes.

As an example of the application of a regional reanalysis, a mesoscale-permitting reconstruction of recent-past ocean variability in the Mediterranean Sea was used to analyze the surface circulation (Fig. 6) and confirmed that the Northern Ionian Sea’s circulation reversed in sign over two 10-year periods, the first in 1987 to 1996 and the second in 1997 to 2006. The phenomenon is known as the Northern Ionian reversal (Pinardi et al. 2015), and it is the first time that a clear picture of this phenomenon has emerged, owing to the high space-time consistency of the reanalysis field.

Outside Europe, some other regional ORAs have been produced for the coastal waters of China and adjacent Seas (Han et al. 2013), the Indian and West Pacific oceans (Yan, Zhu, and Xie 2015), and the Tasman Sea (Sakov and Sandery 2015), showing promising results at eddy-resolving resolution, but also showing that the open boundary conditions are an obvious source of model error absent in the global products and contribute to initialization shocks.

More efforts should be invested into harmonizing the validation strategy, in order to enable the comparison between products covering the same region, but also the comparison between global and regional products as demanded by the user community. The intercomparison between global and regional ORAs is crucial to proving their consistency and to demonstrating the advantage of higher-resolution regional products for specific downstream applications. Moreover, this is an obliged step toward the future nesting of regional into global systems, assuring continuity at the boundaries.
Figure 6. The decadal-mean circulation in the Mediterranean basin at 15-m depth (Simoncelli et al. 2014). (a) mean of the period of 1987 to 2014. (b) mean of the period of 1987 to 1996. (c) mean of period of 1997 to 2006. Velocity amplitudes greater than 0.1 ms$^{-1}$ are depicted by light blue shaded areas.
5. Applications

Estimating the state of the ocean is a primary target in the context of both climate variability assessments and in the production of ocean initial conditions for seasonal and longer-timescale climate forecasts. The time-evolving, three-dimensional ocean state provided by the ORAs is especially useful for analyzing unobserved quantities, among them the meridional overturning circulation in the North Atlantic (e.g., Pohlmann et al. 2013; Karspeck et al. 2017) and oceanic transports such as, for example, the Indonesian Throughflow (Lee et al. 2010), as well as transports at different WOCE sections (Masina et al. 2017). The number of studies that utilize the products from reanalysis systems to investigate various aspects of ocean circulation and climate variability is increasing. However, because of the limited space, we can only mention a few examples. For instance, ORAs have been applied to study a wide range of topics in physical oceanography and climate research, like the nature of sea level variability (e.g., Carton et al. 2005; Wunsch, Ponte, and Heimbach 2007; Köhl and Stammer 2008; Storto et al. 2017), mixed-layer heat balance (e.g., Halkides and Lee 2009; Toyoda et al. 2017b), heat transports (e.g., Haines et al. 2012), water masses (Masuda et al. 2006; Toyoda et al. 2011), and trend and variability of upper-ocean heat content (Carton and Santorelli 2008; Masina et al. 2011; Zhu et al. 2012; Balmaseda et al. 2013b; Palmer et al. 2017). Reviews on the status of ORAs in support of climate-oriented applications are given by Lee et al. (2010a), Schiller, Lee, and Masuda (2013), and Stammer et al. (2016). The two most recent ones (Schiller, Lee, and Masuda 2013; Stammer et al. 2016) also offer critical overviews on the use of ORAs to improve the initialization of climate predictions. In several operational centers, the use of the upper-ocean thermal structure coming from ocean data-assimilation products has become a routine practice for initializing seasonal-to-interannual forecasting systems, mostly focusing on the El Niño–Southern Oscillation (ENSO) phenomenology (Balmaseda et al. 2010). Beneficial effects on predictability have also been reported, increasing the space-time coverage of the observational network (Balmaseda, Anderson, and Vidard 2007). These studies have often showed reductions in model errors and possibly increased skill in ENSO predictions. However, results are substantially dependent on the model, the geographical region, and the year and season under consideration (e.g., Alessandri et al. 2010).

Decadal forecasting is a rapidly-evolving activity (Meehl et al. 2014) that is also starting to suggest that some aspects of decadal variability, notably in the North Atlantic region associated with the AMOC, are better represented by initialized coupled models (Pohlmann et al. 2009; Kröger, Müller, and von Storch 2012; Bellucci et al. 2013, among the most recent). In this context, the goals to initialize and verify decadal forecasts provide a real challenge, imposing the need to provide longer ORAs covering several decades and deeper ocean observations. The relevance of subsurface temperature observations to capture quasi-decadal variability and climate shifts in the North Pacific (Toyoda et al. 2017b) and Atlantic Oceans (Masina, Di Pietro, and Navarra 2004; Vizy and Cook 2014; Yang et al. 2016) support the importance of initializing large-scale modes of decadal variability.
Like the common practice of atmospheric reanalyses, the recent availability of global ORAs has allowed their use as a reference for the evaluation of model performances. For example, ORAs have been used to estimate coupled model biases in simulating the ENSO dynamics (Capotondi et al. 2006; Navarra et al. 2008) and interdecadal changes (Chung and Li 2013). Recently, the NASA-funded Collaborative REAnalysis Technical Environment-Intercomparison Project (CREATE-IP) aimed at homogenizing ORA products for use in validation of the ocean component of future CMIP simulations. However, the exploitation of ORAs for model validation within the ocean modelling community is still not adequate.

Another strand of ORAs is to serve downstream applications such as, for example, driving offline biogeochemical (e.g., Lazzari et al. 2012) and larval dispersal models (e.g., Melià et al. 2013; Blanke et al. 2012).

Furthermore, ORA systems have been successfully used to evaluate the importance and hierarchy of the present and future ocean observing networks. Traditional examples of these applications are the Observing System Experiments (OSEs), which consist of data-denial experiments to evaluate the dependence of validation skill scores on individual observing networks (e.g., Balmaseda, Anderson, and Vidard 2007; Oke, Sakov, and Schulz 2009). Alternatively, Storto et al. (2013) proposed a methodology to assess the impact of observations based on comparing the ensemble spread among ensemble systems, each of them with the withholding of an individual observing network. A dedicated GODAE initiative, the GODAE OceanView Observing System Evaluation Task Team, has also recently been set up to coordinate the observing network assessment efforts.

Studies that aim to assess the potential impact of future observing networks, or to optimize the design and deployment of current observing networks, generally rely on observing system simulation experiments (OSSEs) (e.g., Vecchi et al. 2007; Ballabrera-Poy et al. 2007; Halliwell et al. 2014). OSSEs sample the simulated observations from a model run, usually defined as “truth” or “nature run,” and later assimilate these simulated observations in perturbed experiments whose skill scores are compared with the nature run for validation. Such a methodology has been inherited from meteorological applications and has the advantage that any possible configuration and combination of observing systems can be evaluated, although the definitions of “nature run” and “perturbed simulations” may generally be delicate and influence the assessment.

Finally, with the increased maturity of ORAs, regional reanalyses and regional ocean simulations have started to use data from global reanalyses for the specification of the lateral boundary conditions. This has been the case, for instance, for the South China Sea reanalyses (Zeng et al. 2014) and the CMEMS regional reanalyses for the European North-Western Shelf and the Iberia-Biscay-Ireland regions (Simoncelli et al. 2016), and it is clear that such a strategy will become the standard way of specifying lateral boundary conditions in regional OGCMs in the future.
6. Progress and future outlooks
   a. The 20C reanalyses

   Most of the present ORAs available for climate research cover a relatively short period, generally not before the 1960s, because of the lack of observations. However, for research on long-term changes and decadal variability of global climate, ORAs that cover a longer period (i.e., the whole 20th Century) are an emerging activity, thanks to recent data-rescue efforts. User requirements related to multi-year products are for longer- and finer-resolution time series of data, but going further back in time depends, again, on the availability of atmospheric forcing and in situ data. New global atmospheric-reanalysis products covering the entire 21st century are now accessible, like the 20CR (1871–2012) by Compo et al. (2011) and the ERA20C (1900–2010) by ECMWF (Poli et al. 2016), but they are characterized by low horizontal resolution (2° and 1.125° degrees, respectively) and the assimilation of a restricted set of observations. 20CR assimilates surface and sea-level pressure, whereas ERA20C assimilates surface-pressure and marine-winds observations. Their usage for new ORAs would bring about products with different quality, but such products are necessary to study long-term (inter-decadal) processes, as demanded by the climate community. A crucial effort has to be made by the in situ community to recruit and process in situ historical observations, to be assimilated or used for ORA quality assessment prior to 1990.

   In long-term ORAs, determining the optimal use of observations is still a challenge, owing to abrupt changes in number and accuracy of observations over the 20th century. The observation shocks associated with the introduction of new observing systems may affect the long-term homogeneity of the ORA. Great care must be taken in the research of long-term trends with long-term ORAs (Thorne and Vose 2010). It is clear that identification and reduction of model and observational biases, particularly those changing in time, is needed, as well as further improvements in the data assimilation. It was suggested in the atmospheric community to explore the difference between the full observing system (representing the situation from 1980s onward) and the reduced system (representing the system in the 1960s and 1970s) in order to determine a correction factor to be applied to reestimating the long-term trend (Bengtsson, Hagemann, and Hodges 2004). Other suggestions include the implementation of full four-dimensional data assimilation systems (Dee and Uppala 2009) or ensemble data assimilation, as well as experimental reanalyses using selected sets of observations to assess their relative impact (Bengtsson et al. 2007).

   To attain a quality climate reanalysis, Thorne and Vose (2010) suggested restricting the data input solely to longer-term and/or essential components of the observing system, rather than trying to ingest all data, understand and minimize the impact of large-scale changes to the input data on model behavior ahead of time, and, finally, run an ensemble of reanalyses at reduced resolution to address those methodological choices that have no rigorous basis and understand strengths and limitations by establishing quantitative uncertainty estimates. The ensemble approach that Thorne and Vose (2010) suggest is different from an ensemble Kalman-filter approach in the sense that what they suggest is to build an ensemble that is able to span long-term uncertainties in both model and observations that can affect the
long-term system evolution such as perturbed model physics, different plausible quality control and data assimilation schemes, different realizations of boundary forcings, and different withholding of data. The minimization of the impact of changes in the observing system is already commonly done via bias corrections in both atmospheric reanalyses (Dee and Uppala 2009) and ORAs (Balmaseda et al. 2007 and Storto, Masina, and Navarra 2016), because at present it is still not clear which would be the best procedure to practically optimize model parameterizations for reanalysis. Modern data assimilation schemes are able to combine observations with different biases using variational methods (Dee and Uppala 2009), and the data assimilation process itself provides information on the observation quality integrating all data in a self-consistent framework (Dee et al. 2011a).

In reply to Thorne and Vose (2010) comments, Dee et al. (2011a) reminded that, given the frequency of significant changes in the observing system and the scarcity of some kinds of data, it seems more obvious to reduce trend uncertainties by combining multiple sources of information and using physically-based multivariate relationships between different parameters (e.g., the ones proposed by Weaver et al. 2005 in the oceanic case). It is common practice in reanalyses to apply data selection rules (so-called blacklists) and preprocessing data quality control (for example removing observations that are much too far from model equivalents; e.g., Dee et al. 2001), as well as to use prior estimates of the error characteristics for all datasets in the error covariance matrix.

Nevertheless, in certain cases it might be useful to withhold a limited amount of selected high-quality observations for verification. In order to minimize the problems of time-varying observation quality and quantity, assimilation of only surface observations has been used in long-term atmospheric reanalyses (Compo et al. 2006, 2011; Poli et al. 2016). An ensemble of ORAs, Simple Ocean Data Assimilation with Sparse Input (SODAsi) (Giese et al. 2016) extends back to 1815 and also assimilates only SST. There are, however, attempts to start investigating the feasibility of assimilating hydrographic profiles in long-term reanalyses, unlike previous studies that limited the assimilated network to surface observations. A previous version of SODA (SODA 2.2.4) assimilating both SST and hydrographic data has been produced and used for climate studies (Giese and Ray 2011). Recently, two historical ORAs that cover the period from 1900 to 2010 were performed (Yang, Masina, and Storto 2017). One ORA assimilates both SST and vertical profile data with a 3D-var assimilation scheme, and the other assimilates vertical profile data but is nudged to monthly SST reconstructed fields. The reanalyses are thus designed to investigate the impact of different strategies for constraining the SST in long-term assimilative experiments and the feasibility of assimilating subsurface observations in the same kind of experiments. However, the introduction of observations induces inhomogeneity in the ORAs and makes it difficult to distinguish the real trends from the spurious ones. As an example, the EKE has been observed to increase over time, especially in subtropical regions (Fig. 7). The dual increase of EKE and temperature increment analysis standard deviation, which is a proxy of the energy borne by the data assimilation representing the spatial variability induced by analysis step, indicates the impact of data assimilation on the EKE.
Clearly, the present observing system is not directly suitable for climate monitoring, and the difficulty of detecting long term trends in the climate system is still an issue. It remains a challenge to reconcile the need to exploit the maximum number of observations with the need to achieve long-term homogeneity.

b. Coupled reanalysis of the Earth System

Emerging attention during recent years is being devoted to the production of Earth-system reanalyses that use atmosphere, ocean, sea-ice, and land data assimilation and general circulation models to reanalyze the recent decades. This general idea may be achieved with different degrees of complexity in the coupling between the assimilation systems of different Earth-system components. Dee et al. (2014) envisage advantages of having a coupled approach to the reanalysis problem. Coupled reanalyses can potentially improve the consistency and conservation of global transports, better exploit near-surface oceanic and atmospheric observations, and reduce imbalances between ocean and atmosphere with consequent initial shocks that may lead to model biases and drift (Mulholland et al. 2015). The first experiments toward fully-coupled reanalyses systems started with the use of coupled, general-circulation models that assimilated ocean observations. This approach was adopted by, for instance, the Geophysical Fluid Dynamics Laboratory (Chang et al. 2013) and the Japan Meteorological Agency (Fujii et al. 2009). Although these setups represented a first
attempt to use a coupled model for climate reanalyses and led to improvements with respect to nonassimilative coupled simulations, the lack of atmospheric data assimilation resulted, in turn, in the lack of reliability of atmospheric variability. A prototypical version of a coupled reanalysis has been produced at the National Centers for Environmental Prediction (NCEP) by Saha et al. (2010) and called Climate Forecast System Reanalysis (CFSR). In this product, the data assimilation in the ocean and atmosphere is “weakly coupled,” meaning that intra-medium propagation of observational information is allowed during the forecast step but not in the data assimilation step; i.e., background fields are consistently created by a coupled atmosphere-ocean general circulation model (AOGCM), although the data assimilation of an Earth component is blind to the observations of another component. A similar approach is also adopted by the U.K. MetOffice in the context of seasonal predictions (Lea et al. 2015). “Strongly coupled” data assimilation for climate reanalyses is currently the subject of active research. For instance, a strongly-coupled reanalysis is being produced at ECMWF. The inter-medium observation propagation stems from the use of outer loops (nonlinear model updates) with the fully coupled model within the 4DVAR scheme. The use of multiple outer loops implies that upper-ocean observations affect the near-surface atmosphere, and conversely for near-surface meteorological observations. This approach successfully shows slight improvements with respect to the system with uncoupled data assimilation (Laloyaux et al. 2016). This strongly coupled configuration will serve the purpose of producing the forthcoming ECMWF coupled XX century reanalysis (CERA-20C) (Dee et al. 2014).

c. Biogeochemical reanalyses

The operational oceanography advent also allowed the extension of ocean prediction to biogeochemistry and the development of integrated modeling systems with data-assimilation capabilities that have been utilized for prototyped biogeochemical reanalysis. Biogeochemical reanalysis systems are still at an early stage and necessitate further efforts to improve the models and their coupling and to refine data assimilation techniques able to assimilate both ocean color data and in situ observations. The scarcity of historical biogeochemical observations, the diversity of biogeochemical measurements, and their reprocessing still constitute limiting factors of reanalysis enhancement (Visinelli et al. 2016). An exhaustive review on building the capacity for monitoring and forecasting marine biogeochemistry and ecosystem dynamics is available from Gehlen et al. (2015).

7. Concluding remarks

ORAs are at the forefront of research development as a result of successful collaborative efforts between different scientific communities leading intercomparison exercises coordinated at the international level. The reanalysis concept was first developed in the atmospheric community in the 1980s and, since then, has become an essential product of major numerical weather prediction systems. Following this experience, today several oceanographic
monitoring and forecasting centers are committed to providing qualified reanalyses to users from different applicative sectors (climate, seasonal and weather forecasting, management of marine resources, coastal environment monitoring and assessment). Different types of reanalyses are needed to serve climate services in the future and might play an important role in advancing our understanding of the climate system. These include extended climate reanalyses reaching back a century or more, to provide a longer record for climate studies and climate model validation, and the comprehensive reanalyses of the recent observing period.

The possibility of estimating climate changes and trends over several decades is one of the most important applications of atmospheric and oceanic reanalyses. The observing system has evolved significantly over the last 50 to 100 years, from surface observations before the 1940s to the present system, which relies on satellite coverage, established in the late 1970s, and the advent of the ARGO subsurface ocean temperature and salinity data with worldwide coverage after 2000s. These continuous changes, despite their fundamental importance, limit the usefulness of reanalyses for the assessment of long-term climate trends because they are likely to introduce artificial variability and spurious trends. Despite these limitations, reanalysis “is by far the most accurate way to interpolate in time and space as well as a superior way to obtain dynamical consistency between different atmospheric variables” (Bengtsson, Hagemann, and Hodges 2004). However, the difficulty in detecting long term trends in the climate system remains a challenge. Coupled model data–assimilation research must be pursued over the coming years. A fully-coupled assimilation system would provide better-constrained air-sea fluxes for climate reanalysis purposes, for example, and would reduce imbalances between ocean and atmosphere with consequent model biases and drift mitigation.

Emerging ORA products of the global and regional ocean/sea-ice system over the recent observing period are now ready to serve downstream users like environmental protection agencies, coastal and marine agencies, coast guards, navies, policy makers, and industry. The number of users is expected to grow in the near future, and their needs will likely become more demanding in terms of length and quality of the products, as well as spatial resolution and output frequency. Some systems have been developed to provide information at the mesoscale level. The increase in resolution is another development of ORA activities linked to the availability of higher-resolution atmospheric forcing, but also to the advancement of the computational infrastructures for running and monitoring ORA systems. In the meantime, the forecasting systems will upgrade at higher resolution, testing new model performance and assimilation technique to be used in future ORAs.

The principal demand from the user community is for a concerted assessment of the quality of the accessible ORAs. Moreover, a clear presentation of the various products available for certain regions and their main differences according to their characteristics (temporal coverage, horizontal and vertical resolution, performance versus reference observational datasets) would provide the user with indications of the possible usage for specific applications. All observations and other data used as input, together with the final reanalysis
products, should be made easily accessible to the user community, as well as guidance regarding the quality of specific products. Provision of reanalyses, observations, and metadata in a common format, for example in the Coupled Model Intercomparison Project (CMIP) archive, would encourage their access, use, and quality assessment.

Technical tools for observation handling, monitoring, and diagnostics have greatly improved. At least in the atmospheric community, it is now relatively straightforward to maintain a reanalysis production in near real-time and to provide regular monthly updates of the dataset to many users. The ocean community is just starting to follow a similar approach. The recent availability of both in situ and remote-sensed data sets can promote the production of new ocean indicators computed from ORA data, complementing those derived from observations only. However, the continuous evaluation of such ocean indicators, like the demanded heat and freshwater trends, is a challenge because ORA might be subject to model drifts of the same order of magnitude as the trends.

There are fundamental limitations in the ability to represent what is, essentially, unobserved. A big challenge for ORA producers is to provide users with better uncertainty across a range of space and time scales that span the model and observation spaces and the responsible processes. An indication of uncertainties can be obtained using ensemble techniques, with the important caveat that it is not practical to sample more than a few selected sources of uncertainty in a reanalysis. However, further efficient developments will require sustained consortium efforts encompassing expertise in ocean observations, modelling, assimilation, and information technology. This effort, in a long-term perspective range, will represent a significant driver for the whole operational and climate oceanography communities in the next decade.

Acknowledgments. We gratefully thank Dr. M. Chevallier (Meteo-France), Dr. M. Valdivieso (University of Reading), and Dr. S. Simoncelli (INGV) for providing the ORA data from the Arctic sea-ice, air-sea fluxes intercomparison, and Mediterranean surface currents, respectively.

REFERENCES


Toyoda, T., Y. Fujii, T. Kuragano, N. Kosugi, D. Sasano, M. Kamachi, Y. Ishikawa et al. 2017b. Interannual-decadal variability of wintertime mixed layer depths in the North Pacific detected by


Received: 29 July 2016; revised: 13 July 2017.

Editor’s note: Contributions to The Sea: The Science of Ocean Prediction are being published separately in special issues of Journal of Marine Research and will be made available in a forthcoming supplement as Volume 17 of the series.