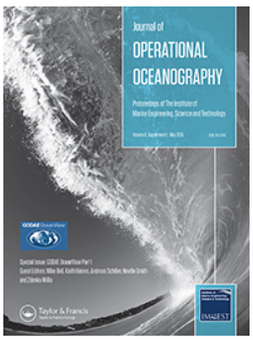

Taylor & Francis
Taylor & Francis Group

Journal of Operational Oceanography

ISSN: 1755-876X (Print) 1755-8778 (Online) Journal homepage: http://www.tandfonline.com/loi/tjoo20

The Ocean Reanalyses Intercomparison Project (ORA-IP)

M.A. Balmaseda, F. Hernandez, A. Storto, M.D. Palmer, O. Alves, L. Shi, G.C. Smith, T. Toyoda, M. Valdivieso, B. Barnier, D. Behringer, T. Boyer, Y-S. Chang, G.A. Chepurin, N. Ferry, G. Forget, Y. Fujii, S. Good, S. Guinehut, K. Haines, Y. Ishikawa, S. Keeley, A. Köhl, T. Lee, M.J. Martin, S. Masina, S. Masuda, B. Meyssignac, K. Mogensen, L. Parent, K.A. Peterson, Y.M. Tang, Y. Yin, G. Vernieres, X. Wang, J. Waters, R. Wedd, O. Wang, Y. Xue, M. Chevallier, J-F. Lemieux, F. Dupont, T. Kuragano, M. Kamachi, T. Awaji, A. Caltabiano, K. Wilmer-Becker & F. Gaillard

To cite this article: M.A. Balmaseda, F. Hernandez, A. Storto, M.D. Palmer, O. Alves, L. Shi, G.C. Smith, T. Toyoda, M. Valdivieso, B. Barnier, D. Behringer, T. Boyer, Y-S. Chang, G.A. Chepurin, N. Ferry, G. Forget, Y. Fujii, S. Good, S. Guinehut, K. Haines, Y. Ishikawa, S. Keeley, A. Köhl, T. Lee, M.J. Martin, S. Masina, S. Masuda, B. Meyssignac, K. Mogensen, L. Parent, K.A. Peterson, Y.M. Tang, Y. Yin, G. Vernieres, X. Wang, J. Waters, R. Wedd, O. Wang, Y. Xue, M. Chevallier, J-F. Lemieux, F. Dupont, T. Kuragano, M. Kamachi, T. Awaji, A. Caltabiano, K. Wilmer-Becker & F. Gaillard (2015) The Ocean Reanalyses Intercomparison Project (ORA-IP), Journal of Operational Oceanography, 8:sup1, s80-s97, DOI: 10.1080/1755876X.2015.1022329

To link to this article: http://dx.doi.org/10.1080/1755876X.2015.1022329

Published online: 09 Jun 2015.

Journal of Operational Oceanography, 2015
Vol. 8, No. S1, s80–s97, http://dx.doi.org/10.1080/1755876X.2015.1022329

The Ocean Reanalyses Intercomparison Project (ORA-IP)

M.A. Balmaseda[a*], F. Hernandez[b,c], A. Storto[d], M.D. Palmer[e], O. Alves[e], L. Shi[f], G.C. Smith[g], T. Toyoda[h], M. Valdivieso[i], B. Barnier[j], D. Behringer[k], T. Boyer[l], Y-S. Chang[m,n], G.A. Chepurin[o], N. Ferry[c], G. Forget[p], Y. Fujii[h], S. Good[e], S. Guinehut[q], K. Haines[i], Y. Ishikawa[r], S. Keeley[a], A. Köhl[s], T. Lee[t], M.J. Martin[e], S. Masina[d,u], S. Masuda[v], B. Meyssignac[w], K. Mogensen[a], L. Parent[c], K.A. Peterson[e], Y.M. Tang[a], Y. Yin[f], G. Vernieres[x], X. Wang[y], J. Waters[e], R. Wedd[f], O. Wang[t], Y. Xue[k], M. Chevallier[z], J-F. Lemieux[g], F. Dupont[g], T. Kuragano[h], M. Kamachi[h], T. Awaji[r], A. Caltabiano[aa], K. Wilmer-Becker[ab] and F. Gaillard[ac]

[a]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, United Kingdom; [b]Institut de Recherche pour le Développement (IRD), Toulouse, France; [c]Mercator Océan, Ramonville Saint-Agne, France; [d]Centro Euro mediterraneo sui Cambiamenti Climatici (CMCC), Bologna, Italy; [e]Met Office, Exeter, United Kingdom; [f]Centre for Australian Weather and Climate Research, Bureau of Meteorology (BOM), Melbourne, Australia; [g]Environment Canada, Québec, Canada; [h]Meteorological Research Institute, Japan Meteorological Agency (MRI/JMA), Tsukuba, Japan; [i]University of Reading (U-Reading), Reading, United Kingdom; [j]Centre National de Recherche Scientifique (CNRS), Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), Grenoble, France; [k]Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, Maryland, USA; [l]NOAA/NODC, College Park, Maryland; [m]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration (GFDL/NOAA), Princeton, New Jersey; [n]Department of Earth Science, Kongju National University, Kongju, South Korea; [o]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA; [p]Program in Atmosphere, Ocean, and Climate, Massachusetts Institute of Technology; [q]Collecte Localisation Satellites (CLS), Ramonville Saint-Agne, France; [r]Center for Earth Information Science and Technology, Japan Agency of Marine-Earth Science and Technology (CEIST/JAMSTEC), Yokohama, Japan; [s]Universität Hamburg (U-Hamburg), Hamburg, Germany; [t]Jet Propulsion Laboratory (JPL), California Institute of Technology, Pasadena, California; [u]Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy; [v]Research and Development Center for Global Change (RCGC), JAMSTEC, Yokosuka, Japan; [w]Laboratoire d'Etudes en Géophysique et Océanographie Spatiale' (LEGOS), Centre National d'Etudes Spatiales (CNES) in Toulouse, France; [x]Goddard Space Flight Center, National Aeronautics and Space Administration (GSFC/NASA), Greenbelt, Maryland; [y]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California; [z]CNRM-GAME, Météo-France, CNRS UMR3589, Toulouse, France; [aa]International CLIVAR Global Project Office, First Institute of Oceanography, State Oceanic Administration, China.; [ab]GODAE OceanView Project Office, Met Office, Exeter, United Kingdom; [ac]Laboratoire de Physique des Océans (LPO/IFREMER), France

Uncertainty in ocean analysis methods and deficiencies in the observing system are major obstacles for the reliable reconstruction of the past ocean climate. The variety of existing ocean reanalyses is exploited in a multi-reanalysis ensemble to improve the ocean state estimation and to gauge uncertainty levels. The ensemble-based analysis of signal-to-noise ratio allows the identification of ocean characteristics for which the estimation is robust (such as tropical mixed-layer-depth, upper ocean heat content), and where large uncertainty exists (deep ocean, Southern Ocean, sea ice thickness, salinity), providing guidance for future enhancement of the observing and data assimilation systems.

Introduction

There is increasing demand for historical records of the ocean climate (Balmaseda et al. 2010; Dee 2014). These are needed as a reference for monitoring the current state of the climate, and also to initialise and validate long-range (e.g. seasonal and decadal) forecasts. Observations alone are often inadequate to generate the required estimate of the ocean variables. Ocean model simulations can provide some insight on the ocean variability, but they are affected by biases due to errors in model formulation, specification of initial states and forcing, and are not directly constrained by observations. Ocean reanalyses are the combination of ocean models, atmospheric forcing fluxes and ocean observations via data assimilation methods and have the potential to provide more accurate information than observation-only or model-only based ocean estimations.

The production of ocean reanalyses (ORAs hereafter) is now an established activity in several research and operational centres. ORAs are revisited every so often, and new 'vintages' are produced at intervals of about five years, as improvements in ocean models, data assimilation

*Corresponding author. Email: maleni_alonso@yahoo.co.uk

methods, forcing fluxes or ocean observations become available. The previous vintage of ORAs (produced around 2006) has already been documented (Stammer et al. 2010; Lee et al. 2009). A new vintage has recently been generated, which has come about through the availability of new surface forcing fluxes (from new atmospheric reanalyses), improved qualitycontrolled ocean datasets, including important corrections to the observations (Lyman et al. 2010; Wijffels 2009), as well as the steady improvement in the ocean models and data assimilation methods. There are lower resolution reanalyses (~1 degree horizontal resolution), spanning a long time-period of typically 50 years, as well as higher resolution products (about ¼ of degree), available for shorter records, usually the altimeter period 1993-onwards.

Although new reanalysis vintages are produced relatively infrequently, some of the ORAs are continuously updated in quasi-real-time, with the model and data assimilation methodology kept fixed. This is the case for the ORAs produced in operational centres to initialize coupled forecasts. These real-time ORAs have the additional advantage that they allow monitoring of relevant climate variables (Xue et al. 2010). The monitoring of the tropical Pacific conditions with a multi ocean reanalysis system (multi-ORA) is now a reality, as can be seen in the NCEP ocean monitoring pages [http://www.cpc.ncep.noaa.gov/products/GODAS/multiora_body.html]

In spite of the continuous improvements in methodology, the estimation of the historical ocean state with reliable error estimates is a major challenge. In addition to the estimation of the three-dimensional ocean state at a given time (the analysis problem), an ocean reanalysis also provides an estimation of the time evolution. The time evolution represented by an ORA will be sensitive to the temporal variations of the observing system, to the errors of the ocean model, atmospheric fluxes and assimilation system, which are often flowdependent, and not easy to estimate (Masina et al. 2011). All these factors contribute to the so-called structural uncertainty, i.e. the uncertainty associated with the methodology and that cannot be sampled with a single system. A crude but pragmatic way of estimating the current uncertainty in our ability to measure key ocean variables is to carry out an intercomparison of ORAs within the framework of a multi-reanalysis ensemble approach. For it to work, it is necessary that the individual components are sufficiently distinct while at the same time have similar levels of error (i.e. equally likely). The multi-analysis ensemble approach has already been successfully used to study the ocean heat content (Xue et al. 2012; Zhu et al. 2011), and to initialize seasonal (Zhu et al. 2012; Zhu et al. 2013) and decadal (Pohlmann et al. 2013; Bellucci et al. 2013) forecasts. The ensemble approach is also used in the framework of the EU funded MyOcean project (Ferry et al. 2012) using eddy-permitting reanalyses over the satellite period (1993-onwards).

The operational oceanographic community continuously carries-out coordinated inter-comparison of ocean forecasting systems (Crosnier & Le Provost 2006, 2007; Xie et al. 2008; Hernandez et al. 2009, 2014; Oke et al. 2012). In the same way, there is also need for routine coordinated evaluation of ORAs, which would exploit the existing information for a variety of purposes, namely (i) quantifying uncertainty, (ii) measuring progress in the quality of the reanalyses and (iii) producing indices for ocean monitoring with associated error estimates. These are the motivations for the current Ocean Reanalyses Intercomparison Project (ORA-IP). This paper offers just a first glimpse of the emerging results, with focus on the benefits of the ensemble approach both to improve the estimation of the signals and to provide uncertainty ranges.

The current ORA-IP project

The joint GODAE OceanView/CLIVAR-GSOP (Global Synthesis and Observation Panel) workshop in Santa Cruz (13–17 June 2011) (Oke et al. 2011) called for a community action on exploitation of the latest ORAs for real time climate monitoring and intercomparison. Although the ultimate goal is the near real-time monitoring of the ocean through indices based on an ensemble of reanalyses, the first stage was to complete an ORA-IP. A viable proposal was put forward in Santa Cruz. The reanalyses producers were to provide relevant information (gridded fields of basic primary variables) in agreed formats and grids (where applicable), to enable the agreed intercomparison procedure to be carried out. A 'processing centre' would take responsibility for the intercomparison of a particular variable in which they had a strong interest and expertise. The processing centres would analyse ensemble statistics based on the input from the individual reanalyses, and create relevant indices, metrics or graphics that could be directly compared.

Table 1 provides a list of the variables chosen for intercomparison. Table 2 lists the ORAs included in the study,

Table 1. List of ocean variables inter-compared and responsible processing institution.

Variable	
Ocean Heat Content	MetOffice
Steric Height	CMCC
Sea Level	Mercator Ocean
Surface Heat Fluxes	University Reading
Mixed Layer Depth	MRI/JMA
Salinity	CAWCR
Depth of 20 degree Isotherm	Mercator Ocean
Sea Ice	Env Canada

Table 2. List of Ocean Reanalysis products entering the inter-comparison.

Product	Forcing	Configuration	Data Assim. Method
ARMOR3D[a,b] CLS	N/A	1/3° Obs-Only (T/S/SSH/U/V)	*OI (SLA/MDT/T/S/SST)*
CFSR[c,d] NOAA NCEP	Coupled DA	1/2° MOM4 coupled	*3DVAR (T/SST/SIC)*
C-GLORS05V3[e] CMCC	ERAi corr+Bulk	1/2° NEMO3.2	*3DVAR (SLA/T/S/SST/SIC)*
ECCO-NRT[f] JPL/NASA	NCEP-R1 +CORE Bulk	1° MITgcm	*KF-FS (SLA/T)*
ECCO-v4[g,h] MIT/AER/JPL	ERAi+CORE Bulk	1° MITgcm	*4DVAR (SLA/SSH/T/S/SST)*
EN3 v2a[i] Hadley Center	N/A	1° Obs-Only (T/S)	*OI (T/S)*
GECCO2[j] U. of Hamburg	NCEP-R1+Bulk	1°×1/3° MITgcm	*4DVAR (SLA/T/S/MDT/SST)*
ECDA[k,l] GFDL/NOAA	Coupled DA	1/3° MOM4 coupled	*EnKF (T/S/SST)*
GloSea5[m,n] UK MetOffice	ERAi+CORE Bulk	1/4° NEMO3.2	*3DVAR (SLA/T/S/SST/SIC)*
MERRA Ocean GSFC/NASA/GMAO	Merra +Bulk	1/2° MOM4	*EnOI (SLA/T/S/SST/SIC)*
GODAS[o] NOAA NCEP	NCEP-R2 Flux.	1°×1/3° MOM3	*3DVAR (SST/T)*
GLORYS2V1(G2V1) Mercator Océan	ERAi corr+CORE Bulk	1/4° NEMO3.1	*KF+3DVAR (SLA/T/S/SST/SIC)*
GLORYS2V3(G2V3) Mercator Océan	ERAi corr+CORE Bulk	1/4° NEMO3.1	*KF+3DVAR (SLA/T/S/SST/SIC)*
K7-ODA(ESTOC)[p] JAMSTEC/RCGC	NCEP-R1 corr. Flux	1° MOM3	*4DVAR (SLA/T/S/SST)*
K7-CDA[q] JAMSTEC/CEIST	Coupled DA	1° MOM3 coupled	*4DVAR (SLA/SST)*
LEGOS[r] LEGOS	N/A	1/4° Obs-Only (SL)	*OI+EOF (SLA/SSH)*
NODC[s] **NODC/NOAA**	N/A	1° Obs-only (T/S)	*OI (T/S)*
PEODAS[t] CAWCR(BoM)	ERA40 to 2002; NCEP-R2 thereafter.	Flux 1°×2° MOM2	*EnKF (T/S/SST)*
ORAS4[u,v] ECMWF	ERA40 to 1988; ERAi thereafter. Flux.	1° NEMO3	*3DVAR (SLA/T/S/SST)*
MOVE-C[w] MRI/JMA	Coupled DA	1° MRI.COM2 coupled	*3DVAR (SLA/T/S/SST)*
MOVE-G2[x] MRI/JMA	JRA-55 corr+Bulk	0.5°×1° MRI.COM3	*3DVAR (SLA/T/S/SST)*
MOVE-CORE[y,z] **MRI/JMA**	CORE.2 Bulk	0.5°×1° MRI.COM3	*3DVAR (T/S)*
SODA[aa] U. of Maryland and TAMU	ERA40 to 2002; ERAi thereafter. Bulk	1/4° POP2.1	*OI (T/S/SST)*
UR025.4[bb] U. of Reading	ERAi +CORE Bulk	1/4° NEMO3.2	*OI (SLA/T/S/SST/SIC)*
AVISO[cc] CLS	N/A	1/4° Obs-Only (SSH/SLA)	*OI (SLA)*
SlCCI[dd] ESA	N/A	1/4° Obs-Only (SSH/SLA)	*OI (SSH)*

[a]Guinehut et al. 2012; [b]Mulet et al. 2012; [c]Saha et al. 2010; [d]Xue, 2011; [e]Storto et al. 2011; [f]Fukumori, 2002; [g]Wunsch & Heimbach, 2013; [h]Speer & Forget, 2013; [i]Ingleby & Huddleston, 2007; [j]Köhl, 2014; [k]Zhang et al. 2007; [l]Chang et al. 2013; [m]Blockley et al. 2013; [n]Waters et al. 2014, [o]Behringer, 2007; [p]Masuda et al. 2010; [q]Sugiura et al. 2008; [r]Meyssignac et al. 2012; [s]Levitus et al. 2012; [t]Yin et al; 2011; [u]Balmaseda et al. 2013; [v]Mogensen et al. 2012; [w]Fujii et al. 2009; [x]Toyoda et al. 2013; [y]Tsujino et al. 2011; [z]Danabasoglu et al. 2013; [aa]Carton & Giese, 2008; [bb]Haines et al. 2012; [cc][(http://www.aviso.oceanobs.com/fileadmin/documents/data/tools/hdbk_duacs.pdf]; [dd]Ablain et al. 2013.

and provides some details about the product name, associated institution, surface forcing, the ocean model, its resolution[1], assimilation method and observations assimilated. The realtime ORAs are shown in italic. The data assimilation column lists the observation types used for their estimation (T/S for temperature and salinity; SLA:

altimeter-derived sea level anomalies; SSH: sea surface height -from tide gauges; SST: sea surface temperature, MDT: mean dynamic topography, SIC: sea ice concentration), as well as assimilation techniques used for reanalysis: Optimal Interpolation (OI), Ensemble Kalman Filter (EnKF), Kalman Filters and Smoothers (KF-FS), Ensemble OI (EnOI), variational methods (3D-var and 4D-var). Some of the observational products also use statistical techniques such as Empirical Orthogonal Functions (EOFs). In addition to ORAs, Table 2 also lists products named Obs-only (OO in what follows), meaning that they are observation-only products that do not include a dynamical ocean model. The OOs provide sea surface height (SSH) or its anomaly (SLA), and/or temperature and salinity (T/S) estimates, and sometimes 3D velocities (U,V), as in the case of ARMOR3D. The atmospheric surface forcing is usually provided by atmospheric reanalyses, using either direct daily fluxes, or different bulk formulations. Sometimes the atmospheric reanalysis forcing is corrected (suffix *corr* in Table 2), using a variety of methodologies. There are also systems that use fluxes from coupled data assimilation systems (Coupled DA), which come in multiple flavours (parameter estimation, EnKF, weakly coupled). The section on 'Surface Heat Fluxes' below provides additional discussion. The detailed description of the analysis systems joining ORA-IP and their differences is beyond the scope of this paper. However, more details about the products can be found in the references given in the Table.

The production centres provided monthly mean fields interpolated to the standard 1 × 1 degree latitude-longitude grid used by the World Ocean Atlas 2009 (WOA09) (Locarnini et al. 2010). Heat and salinity content, their steric contribution, and assimilation increments of temperature were provided as vertically integrated quantities from the surface down to a number of depths: 0–100 m; 0–300 m; 0–700 m; 0–1500 m; 0–3000 m; and 0–4000 m.

The ORAs can be exploited, among other purposes, to assess the strengths and weakness of the different systems, to identify gaps in the observing systems, and to identify robust quantities to use in climate monitoring. The focus of the results presented is to identify the commonalities and differences among the existing reanalyses. To this end, a multisystem ensemble approach is followed, where the signal and its associated uncertainty are measured by the ensemble mean EM(t) and the ensemble standard deviation (ESD(t)) respectively, defined as:

$$\mathrm{EM(t)} = \frac{1}{N}\Sigma_{K=1}^{N} X_K(t) \quad \text{and}$$
$$ESD(t) = \sqrt{\frac{1}{N-1}\Sigma_{K=1}^{N}(X_K(t) - EM(t)^2)} \qquad (1)$$

where $X_k(t)$ denotes an individual reanalysis product. The different signals in a time series (mean, seasonal cycle, interannual variability, etc.,) are the result of a temporal filter F, and here EMF(s) and ESDF(s) denote the ensemble mean and ensemble standard deviation of the filtered signal[2]. The temporal standard deviation of the filtered ensemble mean EMF(s) is defined as σ^F_{EM}, while σ^F_{ESD} is the quadratic mean of ensemble spread of the filtered $\mathrm{ESD^F(s)}$ as follows:

$$\sigma^F_{EM} = \sqrt{\frac{1}{M^F - 1}\Sigma_{S=1}^{S=M^F}\left(EM^F(s) - \overline{EM^F}\right)^2} \quad \text{and}$$
$$\sigma^F_{ESD} = \sqrt{\frac{1}{M^F}\Sigma_{S=1}^{S=M^F}\left(ESD^F(s)\right)^2} \qquad (2)$$

with $\overline{EM^F}$ the time mean of the filtered EM, and $\mathrm{M^F}$ is the number of independent temporal samples in the filtered timeseries. The signal-to-noise ratio, defined as the ratio $\sigma^F_{EM}/\sigma^F_{ESD}$, provides guidance on whether the estimation is robust. For instance, estimations with signal-to-noise less than unity are usually not considered robust.

In what follows, the term EM-ORA and EM-OO will be used to refer respectively to the ensemble mean of ORAs and OOs. The rest of the article presents a brief overview of the preliminary results of the intercomparison of the variables listed in Table 1.

Heat content

Monthly mean depth integrated potential temperatures (K m) were used in this study. The vertically integrated temperature was converted to ocean heat content (OHC) per unit of area by multiplying by reference values for density (1025 kg $\mathrm{m^{-3}}$) and specific heat capacity (3985 J $\mathrm{Kg^{-1}\,K^{-1}}$). This quantity, further integrated in the horizontal global domain, and computed relative to a common reference period of 1993–2007, has been used to estimate changes in the global OHC. Note that when the timeseries are dominated by trends, the choice of reference period impacts the time evolution of the spread among the timeseries[3]. The apparent increase in spread among analyses during the 2000s is substantially reduced if one chooses the 2003–2007 reference period (not shown).

Time series of global OHC change (Figure 1) show best agreement for the upper levels and the products start to diverge as the integration is carried out to deeper levels. The largest rates of 0–4000 m OHC rise during the 1990s exceed 3 $\mathrm{Wm^{-2}}$ (expressed relative to Earth's surface area) for some products initialized in the early 1990s and cannot be considered physical. They are most likely artefacts of system spin-up or 'shocks' related with introduction of the altimeter data. Trends over the period 2000–2009 for 0–4000 m OHC give values between about 0.1 and 0.8 $\mathrm{Wm^{-2}}$. The OO products ARMOR3D and EN3 are both near the upper end of this range. Ocean heat uptake below 300 m appears to increase markedly in the

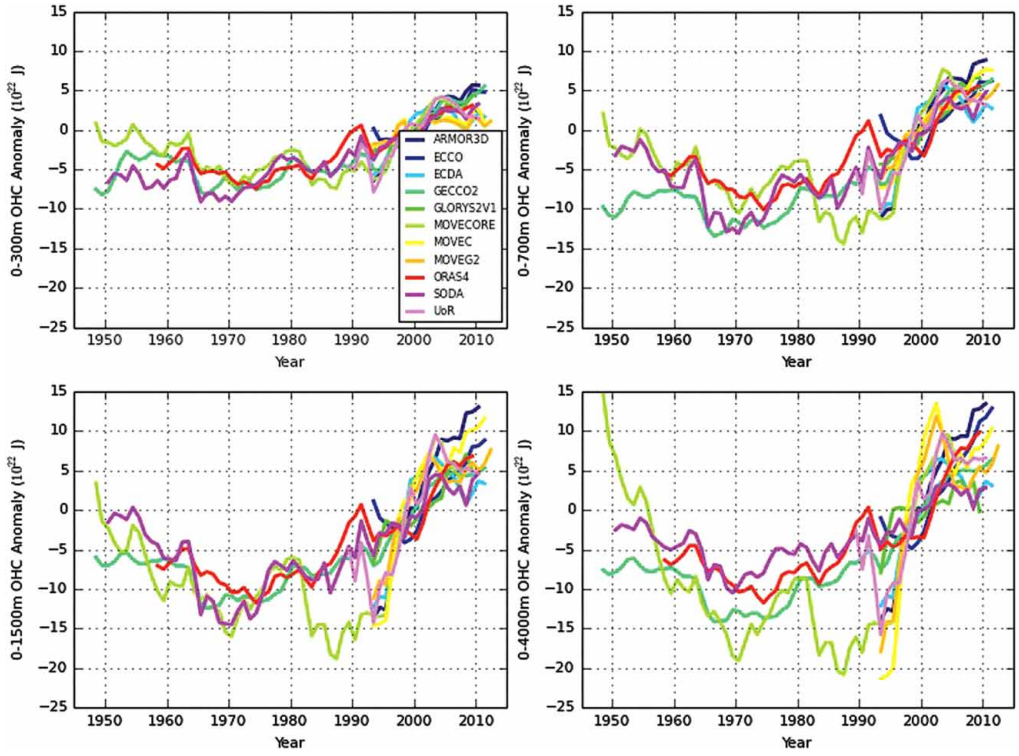

Figure 1. Time series of global ocean heat content anomaly, relative to a baseline period of 1993–2007. Note that SODA only includes grid boxes that span the full column and therefore will tend to underestimate OHC changes as the depth of integration increases. ARMOR3D and EN3 are obs-only analyses and do not include a dynamic model component. [UoR in legend corresponds to the URO25.5 in Table 2].

early 2000s for most products (Palmer et al. 2014), qualitatively supporting the results from the ORAS4 system (Balmaseda et al. 2013), although there is still a large spread in the amplitude of the OHC, and spatial patterns of change below 300 m vary among ocean data assimilation products (not shown). The OO products ARMOR3D and EN3 both show a similar signal of deep ocean heat uptake to ORAS4, illustrating that this signal is inherent to the observations (Palmer 2014).

Figure 1 also shows that the ensemble spread of the multi- ORA is larger than the ensemble spread of the ORAS4 system (Balmaseda et al. 2013b). Whether this holds for other individual ensemble-based ORA would need to be evaluated. A more difficult question is whether the multi-ORA spread is a good estimator of the existing uncertainty. It appears similar to the spread obtained with observation-only estimations (Palmer 2010).

Steric sea level

Steric sea level (SSL) refers to the change of sea level due to ocean density variations associated with thermal and haline expansion or contraction of sea-water. SSL rise is responsible for about 30% to 40% of the total sea level rise during the last decades, according to recent estimates (Cazenave & Llovel 2010; Church et al. 2011). The ORA-IP is being used to investigate the steric sea level variability, by: (i) quantifying the global SSL, its uncertainty and consistency with respect to independent estimates; (ii) assessing the regional SSL change and the agreement among ocean reanalyses; (iii) quantifying the relative contributions of the thermal and haline components and (iv) quantifying the relative contributions of different vertical depth ranges (Storto et al. 2014). Some of these aspects are clearly related with the ocean heat content variations and with the attribution of sea-level changes, but are not discussed here. Instead, this section focuses on the performance of the EM-ORA compared to EM-OO.

SSL can be diagnosed in two different ways: (i) as normalized vertical integration of density anomalies (SSL-density), and (ii) as the differences between sea-level and bottom pressure anomalies (SSL-residual). The latter is not easy to infer from models, which are volume-preserving by virtue of the Boussinesq approximation (Greatbatch 1994). Instead, temperature and salinity monthly means from the ORA and OO products, containing information from *in-situ* observations, are used to diagnose SSL-density. Satellite-products are used to derive SSL-residual, thus providing an independent validation data set. Here monthly means of altimetric sealevel

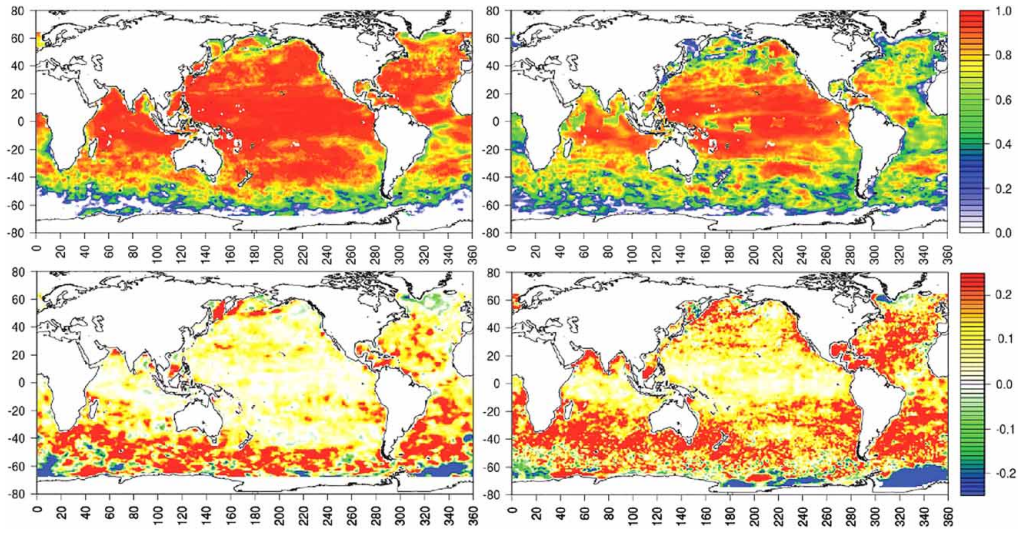

Figure 2. 2005–2009 Steric Sea Level anomaly correlation of EM-ORA with respect to the validation dataset (altimetry minus gravimetry) described in the text, for the full (top-left) and the inter-annual signal (top-right). Correlations higher than 0.25 are significant (at the 95% confidence level). The bottom panels show the map of differences between the EM-ORA anomaly correlation and the EM-OO anomaly correlation for the full (bottom-left) and the inter-annual signal (bottom-right). Positive (negative) values indicate that the correlation is higher (lower) with EM-ORA than with EM-OO.

anomaly (from AVISO) minus gravimetric ocean bottom pressure anomaly are used; (from GRACE RL05 (Chambers & Bonin 2012) available from 2005).

The top-left panel of Figure 2 depicts the 2005–2009 map of temporal anomaly correlations between SSL-density from EM-ORA and SSL-residual (altimetry minus gravimetry). The high values of the correlation suggest high consistency in SSL between two independent estimates of SSL over most of the Global Ocean. In the Southern Ocean, south of approximately 60S, where the availability of *in-situ* observations is poor, the correlation is lower. The top-right panel shows the temporal anomaly correlations, calculated after the seasonal signal has been removed (i.e. inter-annual signal retained). Although removing the seasonal cycle decreases the correlation value (especially in the Atlantic Ocean and at high latitudes), EM-ORA still exhibits high correlations for the inter-annual signal in the tropical areas and at mid-latitudes.

The correlation between SSL-density and SSL-residual is higher for EM-ORA (0.84) than for any individual product (0.77 at the maximum), and also higher than for EM-OO (0., not shown). The latter is especially evident in areas where the *in-situ* observing network is poor and/or where there is impact of deep and bottom waters. The bottom panels of Figure 2 show the difference of the anomaly correlation with respect to the validation dataset between the EM-ORA and the EM-OO for the full (left) and inter-annual (right) signals. The high correlation obtained by EM-ORA emphasizes the added value of the dynamical constraints and atmospheric forcing included in the ORAs. This is evident in the full fields (in the Southern Ocean, in the South Atlantic and just south of the Bering Strait), and especially noticeable for the inter-annual signal.

Although the EM-ORA proves to be a good estimator of total steric height, uncertainty still remains regarding the partition into thermal and haline components, and the contribution of different depth ranges. Preliminary results over a longer intercomparison period (1993–2009) (Storto et al. 2014) show a large spread regarding the contribution of deep layers (below 700 m of depth) to SSL trends, with a low signal-to-noise ratio in the trend estimation of less than 1.

Sea level

The sea levels from the ORAs in Table 2 and two OO products (ARMOR3D and LEGOS) have been evaluated. (The sea level ARMOR3D is effectively the delayed gridded AVISO [http://www.aviso.oceanobs.com/fileadmin/documents/data/tools/hdbk_duacs.pdf] product, also called DUACS). This comparison focuses mainly on the interannual variability and regional distribution of the trend, and it uses globallydetrended monthly means of sea level anomalies. For each product, the seasonal cycle was removed at each location of the ocean domain. The global mean sea level for each month was also removed.

Two reference data sets have been used for the evaluation: sea level from tide gauges, and the newly reprocessed altimeter-derived sea level from the ESA Climate Change Initiative (SLCCI) (Ablain et al. 2013). The latter is a gridded dataset where the original altimeter data has been reprocessed with improved algorithms (orbit, wet tropospheric corrections, among others) and ancillary data

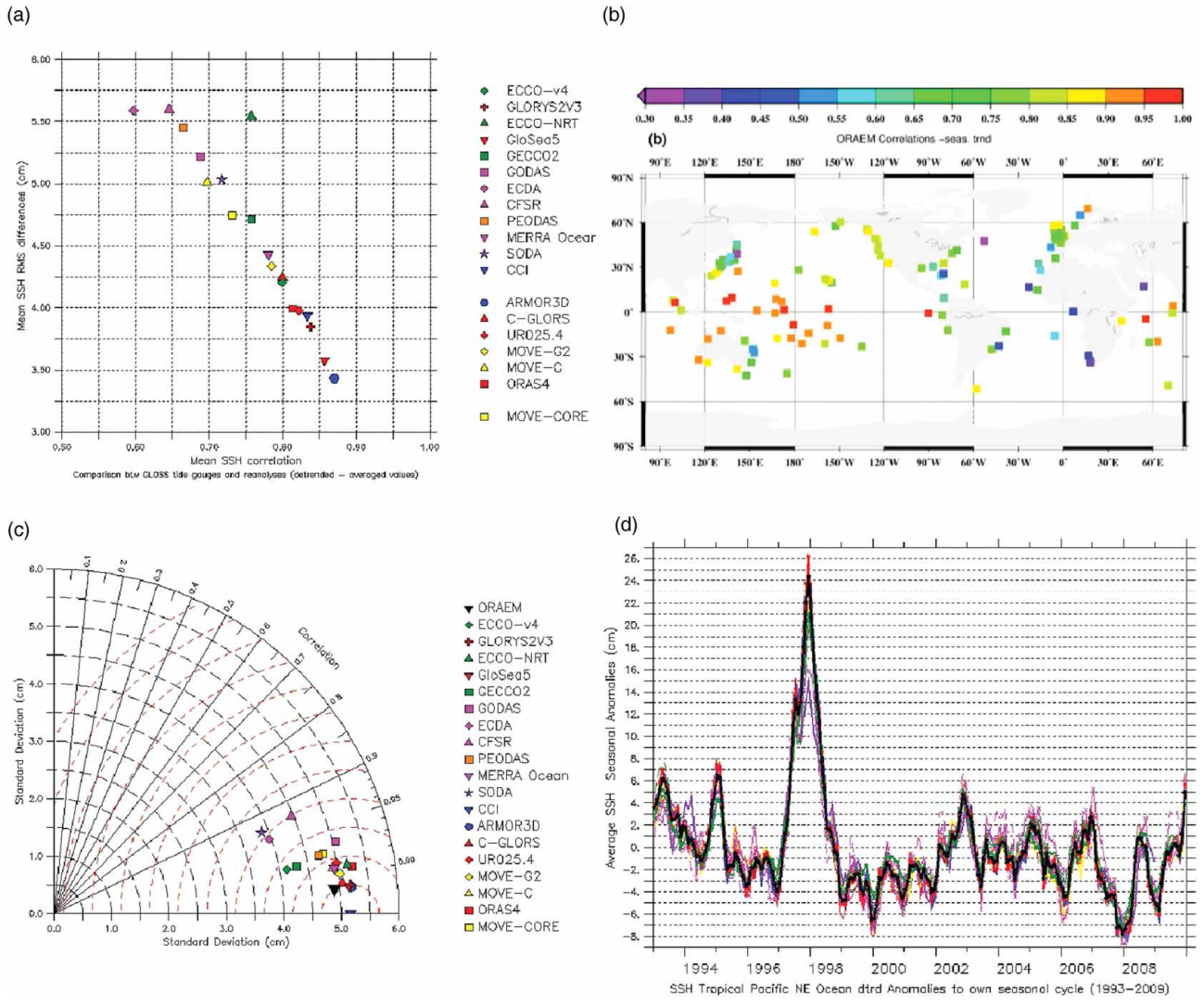

Figure 3. Top: Comparison between tide gauges and ORAs and OOs, after detrending and removing of seasonal cycle: (a) RMS/Correlation diagram for the individual products using GLOSS tide gauge data as reference; (b) correlations between EM-ORAalti and tide gauges time series, at tide gauge locations. Bottom: Evaluation of a sea level index: (c) Taylor diagram using SLCCI as verification (red contours interval is 0.5 cm, centered in the verification); (d) Index time series, defined as the area-averaged sea level anomalies over the North-East Tropical Pacific region (0–12°N, 84–108°W). Red contours in the Taylor diagram in (c) represent the RMS error, plotted at 0.5 cm intervals centered on the verification value. Anomalies and statistics have been computed over the 1993–2009 period.

(using improved atmospheric fields for instance) in order to produce consistent time series of sea level for climate studies. SLCCI has not been assimilated by any of the products in Table 2, although many of these products assimilate along-track satellite altimetry (usually AVISO). Only two products (ECCO-v4 and LEGOS) use information from tide gauges.

The tide gauges used for the evaluation are the same as the subset from the Global Sea Level Observing System [GLOSS, see http://www.gloss-sealevel.org/] chosen for evaluation of sea level reconstructions (Meyssignac et al. 2012).

Monthly means of sea-level anomaly at the tide gauge locations were created after removing the effects of tides and inverse barometer from the original tide gauge data. This allows a relevant comparison with sea level anomalies from the reanalysis products because tides and inverse barometer are not represented in the reanalysis products. The ORAs and OOs were spatially interpolated to the tide gauge locations. All the time series involved in the analyses were detrended at each location by removing the product-specific local linear trend.

The comparison with tide gauges appears in the top panels of Figure 3. The statistics are for the period 1993–2009. Figure 3(a) shows the scatter diagram for the individual products (top-left), with the temporal correlation (x-axis) and the rms error (cm, in the y-axis). A large scatter in the scores is seen among different products, with the

best fit generally obtained by the products assimilating SLA, and in particular by those with higher horizontal resolutions, with scores comparable to those obtained by the altimeter-derived SLCCI and AVISO. This result indicates that not all the ORAs are equally likely, and therefore the grand ensemble mean may not be appropriate to estimate coastal sea-level variations. In this case, the ensemble approach is limited to those products that assimilate altimeter EM-ORAalti. Figure 3b shows the correlation map between tide gauges and EM-ORAalti. Even with the reduced ensemble, the correlation is higher in the open ocean than in the continental shelves, and it appears higher in the tropics than at higher latitudes.

A different application of the multi-ORA ensemble is the definition of climate indices relevant for regional climate monitoring, which is illustrated in the following (although more work is needed to define relevant indices). The sea level variability averaged over the Eastern North Tropical Pacific region (0–12°N, 84–108° W) has been chosen as an example, because although different from the traditional equatorial El Niño index, it reflects the impact the El Niño in the Western Coast of Mexico related to the coastal propagation of Kelvin waves. In this case, ESACCI is used as validation data set. All products show a coherent interannual variability (Figure 3d), even when altimeter data are not assimilated, and there is very small spread around EM-ORA (black). The variability is dominated by the El Niño 1997–98, and a significant negative trend of ~3–4 mm/y, consistent with the lack of Eastern Pacific ENSO in the last decade, and with the recently reported strengthening of the Pacific trade winds (Balmaseda et al. 2013b; England et al. 2014; de Boisséson 2014). Figure 3c shows the corresponding Taylor diagram for the different ORAs and OOs averaged over this region. In contrast with the tide-gauge evaluation, here the scores of the ORAs versus SLCCI are quite similar (sixteen of the nineteen products show correlations higher than 0.95 and rms differences lower than 1.5 cm). The smallest rms error (around 0.5 cm rms error, y-axis in Taylor diagram) is achieved by the EM-ORA. The EM-ORA score is comparable to that achieved by the best members (which assimilate satellite altimetry) and by AVISO, and is larger than 0.99. EM-ORA has a weaker signal than SLCCI (4.8 instead of 5.2 cm of standard deviation, x-axis in Taylor diagram), a natural consequence of the ensemble averaging.

In other areas, like the North Atlantic (not shown), there are more discrepancies among the reanalysis products and weaker signal-to-noise ratios. Discrepancies can arise from different choices in the assimilation systems (Hernandez et al. 2014). It has been shown that products assimilating altimeter data can be distinguished from those that do not. The different methods used to assimilate altimeter information can also introduce spread. For instance, the altimeter can be used to constrain only the baroclinic mode, or only the barotropic mode, or to constrain the fresh water budget, or the three aspects simultaneously. The altimeter data can be assimilated in anomaly mode (using anomaly values relative to a reference period) or using the absolute values (which implies the use of an external mean dynamic topography (MDT), which differs between systems). The ORA-IP can be used to gain insight into the sensitivity arising from the assimilation methods, but this is beyond the scope of this paper.

Surface heat fluxes

The purpose of this comparison is to assess the global heat closure in ORAs, the consistency of the seasonal cycle and interannual variability between the products, and to compare with other heat fluxes from a variety of sources (primarily satellite, ships, buoys and atmospheric reanalysis). These other sources are not completely independent (with the exception of satellite based radiative fluxes) because they may also use SST or near surface meteorological data to generate products. Nevertheless, they enable some assessment of the uncertainty introduced by the reanalysis methods themselves. Additional datasets include the OAFlux latent and sensible heat flux product (Yu et al. 2008) combined with ISCCP satellite based radiation (Zhang et al. 2004), the ship-based NOC2.0 product (Berry & Kent 2009), the Large and Yeager (2009) hybrid flux dataset CORE.2, and two atmospheric reanalysis products, the ECMWF ERAInterim reanalysis (Dee et al. 2011) (referred to as ERAi) and the NCEP/DOE reanalysis R2 (referred to as NCEP-R2) (Kanamitsu et al. 2002).

Most ORAs are forced with bulk formulae using an atmospheric dataset taken from an atmospheric reanalysis product. In these cases, assimilation of sea surface temperature (SST) observations directly influences the net surface heat flux, as the turbulent latent and heat fluxes, computed from bulk formulae, and the outgoing long wave radiation, computed using the Stefan-Boltzmann Law, depend on the SSTs. The ORAs can also close their heat budget through the temperature assimilation increments, since the vertically integrated temperature assimilation increments, with the appropriate unit transformation, are equivalent to a heat flux (Balmaseda et al. 2013b).

Figure 4 shows the 17 yr mean globally integrated heat fluxes for 15 individual ORAs and for the ensemble mean, as well as for the other global flux products. The interannual variability over the same period is shown by the error bars. Most ocean reanalyses have a positive surface imbalance (mean net surface heat flux into the ocean), usually considerably smaller than for the observational products, e.g. ISCCP/ OAFlux and NOC2.0, and smaller than for atmospheric reanalyses in some cases. The largest interannual variability is seen for the PEODAS product which uses ERA-40 (Uppala et al. 2005) forcing fields until 2002, and NCEP-R2 based forcing thereafter. Interannual variations

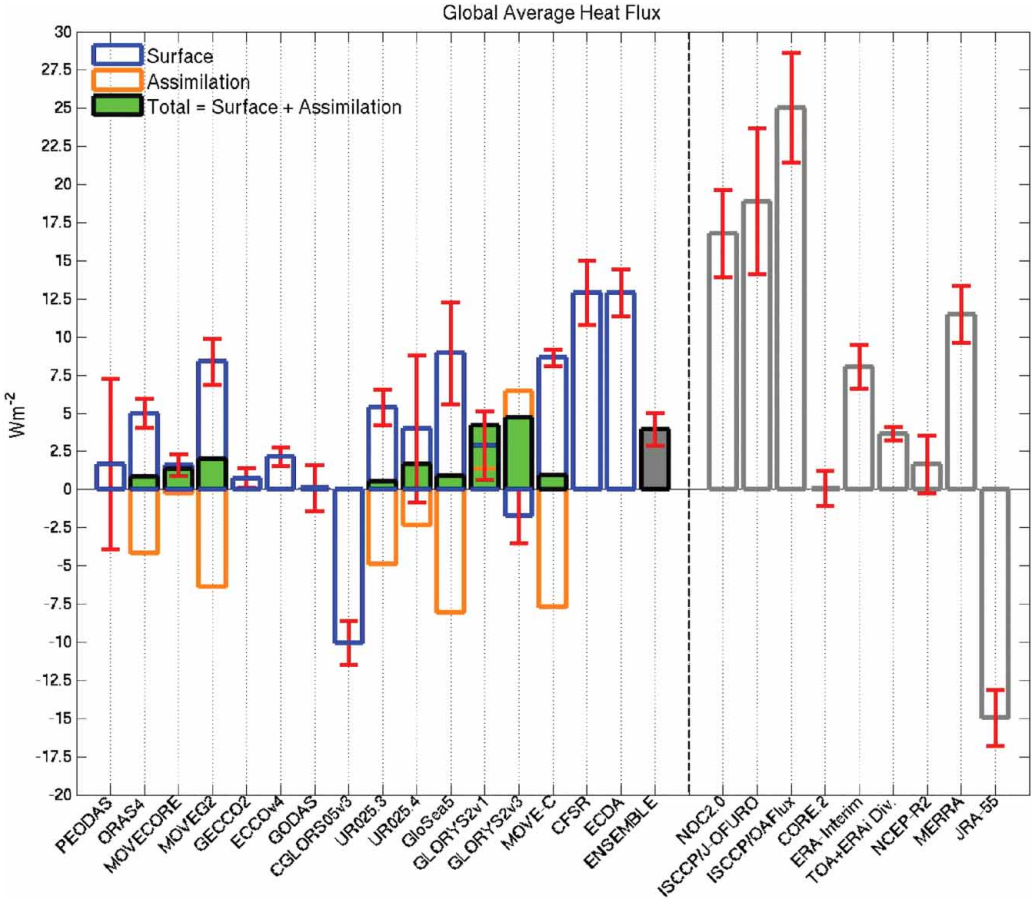

Figure 4. Time mean global net 'Surface' heat fluxes (grey bars) and their interannual standard deviations (red error bars) over the 17 years (1993 – 2009) spanned by all data sets. The 15 member ensemble of 'Surface' flux products is also shown (dark grey bar), along with observation based on atmospheric reanalysis products to the right hand side (orange bars). Eight products also have 'Assimilation' fluxes (blue bars) computed by integrating the temperature increments from the surface down to the bottom, along with 'Total' -fluxes, i.e. 'Surface' +'Assimilation' fluxes (green bars). Positive is heat flux into the ocean. Units are in Wm^{-2}.

over 1993–2009 are only ~1 W/m^2 for the ensemble of 15 flux estimates. The contributions from the assimilation increments are mostly negative (removing heat from the ocean on the global average), resulting in a reduction of the net heat flux. The total net heat flux applied (i.e. surface plus assimilation) is still positive and mostly smaller than ±2 W/m^2, consistent with reported warming in global ocean heat content (Levitus et al. 2012; Balmaseda et al. 2013b; Palmer et al. 2010).

The seasonal cycle in surface heat fluxes closely agrees in most regions between the reanalysis products (not shown), with monthly spreads generally being smaller than 10 Wm-2 over most of the global ocean, exceptions being the subpolar gyres, the Southern Ocean and some eastern subtropical basin areas (Valdivieso et al. 2014). Interannual signal-to-noise ratios for the surface heat fluxes over the period 1993–2009 show strong signals (2+) in the ENSO affected regions and perhaps some signals at higher latitudes, but with signal/ noise ~1, longer analysis periods may be needed to identify this variability more clearly. Regional comparisons are being extended to include individual flux components (representing radiative and turbulent transfers), and also validation against in situ flux measurements at a number of OceanSITES moorings (Valdivieso et al. 2014), which provide an independent check that is not reliably gained from any other source.

Mixed layer depth

Mixed layer depth (MLD) is one of the most important variables for both the dynamical process of climate variation and for biogeochemistry. Intercomparison of the seasonal to interannual variability in the global MLD provides a useful gauge of the value of ORAs for the study of climate variability.

The MLD used in this study is defined as the depth where potential density exceeds the 10 m depth value by $\Delta\rho = 0.03$ or 0.125 kg/m^3 (MLDr003/MLDr0125). Similarly, the isothermal layer depth (ILD) is defined as the

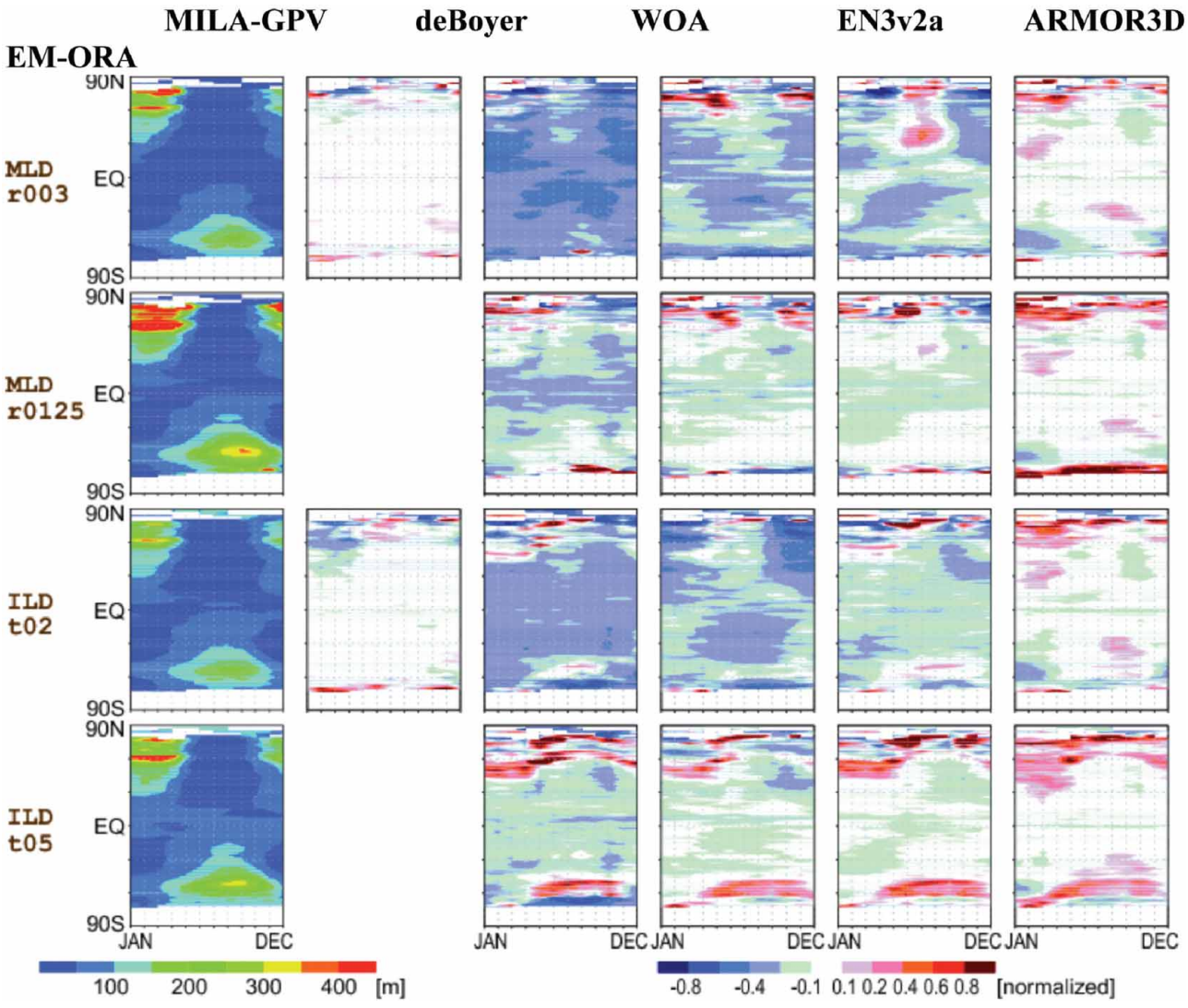

Figure 5. Zonal mean monthly MLDs and ILDs from MILA-GPV averaged over 2005–2011 (left column). Others: Differences from MILA-GPV, normalized by the MILA-GPV values.

depth where potential temperature differs from the 10 m depth value by $\Delta T = 0.2$ or 0.5°C (ILDt02/ILDt05). Different criteria are used because it is not easy to find a unique threshold that defines the mixed layer depth at all latitudes. For MLD/ILD verification the MILAGPV (Hosoda et al. 2010) and deBoyer (deBoyer Montegut et al. 2004) datasets are used, also estimated from the individual TS profiles, following the definitions above. In particular, MILA-GPV uses only the Argo profiles without interpolation between grid points, although the spatio-temporal coverage of the dataset is limited. deBoyer provides the monthly climatological fields (MLDr003 and ILDt02).

The MLD/ILD are calculated from monthly means of temperature-salinity (TS) fields on the individual native grids of three OO products (EN3v2a, ARMOR3D, WOA09) and 16 ORAs; these are then interpolated to the regular global longitude-latitude common grid. EM-ORA is estimated as the ensemble average of individual MLD/ILD on the common grid (this will differ from the MLD/ILD calculated from the ensemble mean of TS). The MLD/ILD from the individual ORAs exhibit various biases in the mean fields depending on the diversity of model configurations and assimilation systems (not shown). Here the evaluation of EM-ORA is focussed on rather than on the detailed representations of the individual reanalysis fields, which will be described in future work.

Figure 5 presents the zonal mean monthly MLD/ILD normalized differences of EN3v2a, ARMOR3D, deBoyer, WOA09 and EM-ORA with respect to the MILA-GPV as reference. Note that values averaged over the Argo-rich 2005–2011 period are plotted for MILA-GPV, EN3v2a, ARMOR3D and EM-ORA, while the climatological

fields are provided by deBoyer and WOA. The differences between deBoyer and MILA-GPV (MLDr003 and ILDt02) are generally small, since MILA-GPV and deBoyer are comparable datasets that use individual TS profiles. The larger differences appear in high latitudes, where the availability of ocean observations is limited. MLDs/ILDs for EN3v2a and ARMOR3D are biased-shallow due to the use of gridded and monthly mean TS fields. ILDt02s in WOA are 20 to 40% shallower than MILA-GPV globally, due to the use of the climatological TS field (de Boyer Montegut et al. 2004). Using larger values for criterion ($\Delta\rho$ = 0.125 kg/m^3 and ΔT = 0.5°C) reduces the shallow biases. The shallow biases in MLDr0125 and ILDt05 for EN3v2a and ARMOR3D are generally less than 20% except at high latitudes. We found that a large portion of these shallow biases result from the coarser vertical resolution of the OO gridded TS products at relevant depths compared with the model based reanalyses (Toyoda et al. 2014). Model biases do cancel in most areas in the EM, although large positive biases remain in regions where common biases are well known from coarse resolution models (Hasumi et al. 2010) (e.g. the Kuroshio Extension and Antarctic Circumpolar Current regions). In addition, ILDt05 values from WOA, EN3v2a, ARMOR3D and EM-ORA are commonly larger than those from MILA-GPV in the subarctic regions and Southern Ocean around spring. This is likely due to the fact that MILA-GPV is the only product that does not use monthly means of TS when deriving MLD/ILD. This specific topic will be described in future work.

Salinity in the top 700 m

Salinity variability has a significant impact on the density structure and dynamics of the ocean. However, it is only in the past few years that assimilation of salinity has received attention, largely because of the advent of Argo [see http://argo.jcommops.org], which has significantly improved the sampling of the global ocean salinity), and because of its importance in obtaining balanced ocean states. For instance, recent studies on seasonal forecasts (Zhao et al. 2013a, b) demonstrate that the assimilation of salinity observations results in improving ocean states density and T/S properties, resulting in better ENSO prediction.

This study evaluates the averaged salinity in the top 700 m of the ocean (S700) as represented by the EM-ORA and compared it with the EM-OO. As discussed, the ESDORA gives an indication of uncertainty, and the signalto- noise ratio provides guidance on where the signal measured by the ensemble mean dominates over the noise measured by the ensemble spread.

Figure 6a shows the difference of annual mean S700 between EM-ORA and EM-OO in Table 2 over the period 1993-2010. The difference is largest (~0.2 psu) in regions of strong frontal variability such as the Gulf Stream, Southern Ocean along the Antarctic Circumpolar Current (ACC) region, and to a lesser extent the Kuroshio region. In the tropics the difference is generally less than 0.05 psu.

Figure 6(b) shows the ESD-ORA of the S700 1993–2010 mean (or σ^M_{ESD}, where M denotes 1993–2010 temporal mean). In general the largest spread, up to 0.15 psu, is also associated with the areas of strong variability or greatest mean difference compared to the EM-OO analyses. Around most of the ACC, the ESD-ORA is just large enough to encompass the large differences between EM-ORA and EM-OO. The spread is relatively large in the eastern equatorial Atlantic and the western equatorial Indian Ocean, where the spread reaches up to 0.1 psu

Figure 6(c) shows the correlation of S700 interannual anomalies between the EM-ORA and EM-OO for the period 1993–2010. Correlations are relatively high, greater than 0.75, in the equatorial and sub-equatorial Pacific, particularly in the centre and west. They are also high in the eastern equatorial Indian Ocean, and throughout parts of the sub-tropical and mid-latitude oceans. Correlations are relatively low, less than 0.5, around the northern edge of the ACC, Western Indian Ocean and parts of the sub-tropical Atlantic, particularly downstream of the Mediterranean outflow. Each individual ORA can be correlated with the EM-OO. Then the spread in this correlation gives an indication of the disagreement in the estimate of variability between the different systems. This is shown in Figure 6(d). There is some correspondence between areas with large spread and low correlation in Figure 6(c), e.g. the northern edge of the ACC in the Pacific Sector and the northern part of the tropical Atlantic. Equally, the high correlation in the Tropical Pacific, Eastern Indian Ocean, North East Pacific and North East Atlantic, where the spread is low, is indicative of consistency between the different estimates. The Southern Ocean is an exception, presenting relatively large values of the correlation and large values of spread.

Figure 6e shows the standard deviation of the interannual S700 anomalies (seasonal cycle removed) of EM-ORA (σ^I_{EM}, where I stands for 'interannual'). This gives an estimate of the amplitude and geographical distribution of the S700 interannual signal, which appears highest in subduction areas close to the edge of strong boundary currents. It is also high in the western equatorial Pacific and central Indian Ocean, probably associated with changes in the fresh-water fluxes. Figure 6(f) shows the spread in the S700 monthly anomalies of the ORAs (or σ^I_{ESD}) I. The spread is largest in the sub-tropics and midlatitudes, particularly associated with western boundary currents and the Southern Ocean. In the western boundary current regions and parts of the Southern Ocean it exceeds 0.1 psu. The signal-to-noise ratio is greater than 1 in the equatorial west Pacific, central Indian Ocean and small regions

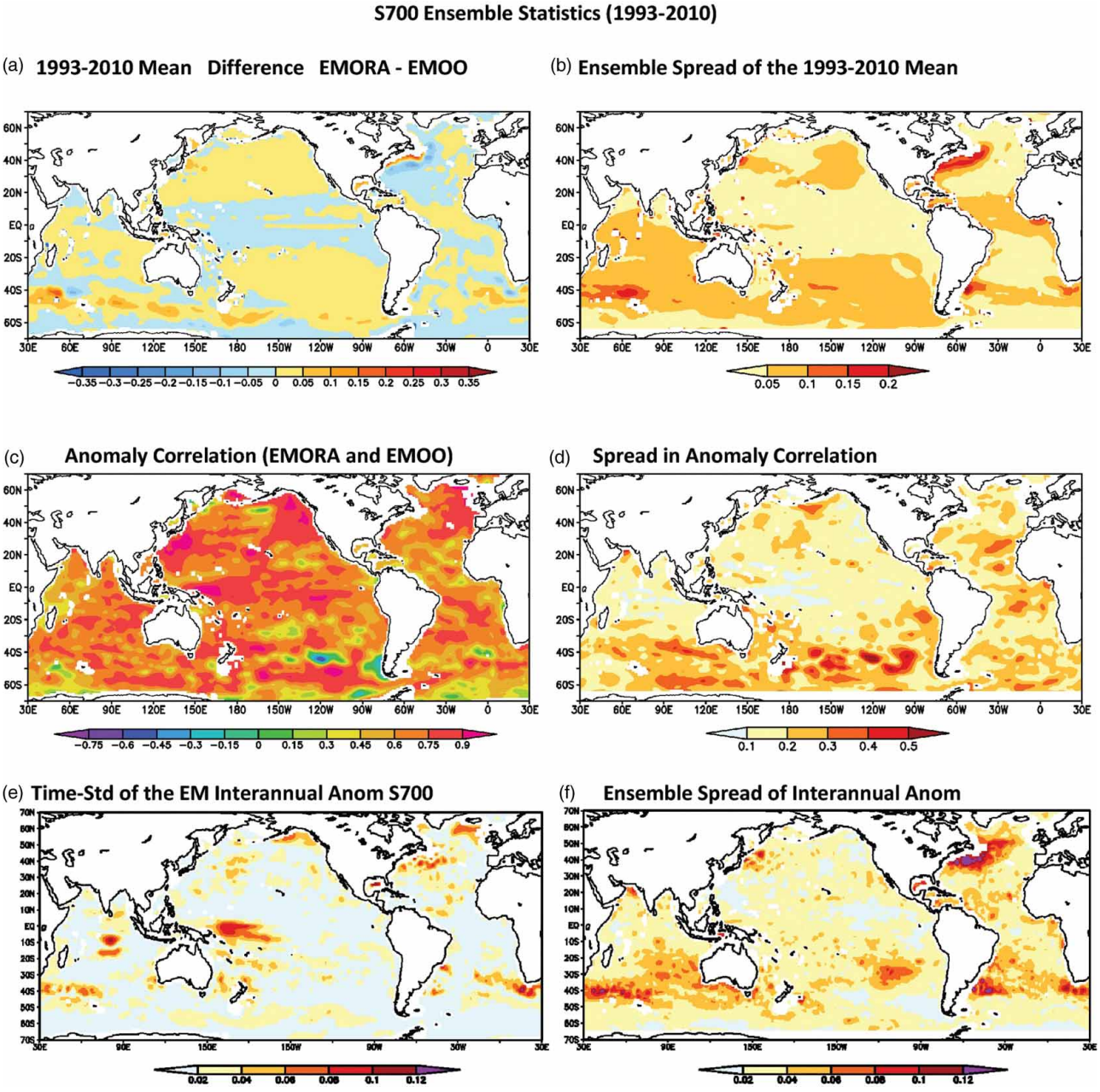

Figure 6 (a) 1993–2010 mean difference of S700 (Depth-averaged salinity over 0–700 m) between EM-ORA and EM-OO. The interval of colour bar is 0.05 psu. (b) The ensemble spread of the mean S700 (ESD-ORA). The interval of colour bar is 0.05 psu. (c) Temporal correlation of S700 monthly interannual anomalies between the EM-ORA and EM-OO, for the period 1993–2010. (d) Spread of correlation coefficients of S700 anomalies from the individual ORAs. (e) The inter-annual standard deviation (1993–2010) of EM-ORA S700, representative of the interannual 'signal' (σ^I_{EM})I. The interval of colour bar is 0.02 psu. (f) The average ensemble spread of the interannual anomalies of S700 (σ^I_{ESD} representing the uncertainty or 'noise'). The interval of colour bar is 0.02 psu.

of the mid-latitude ocean. However, over most of the oceans the signal-to-noise ratio is less than 1 (Alves et al. 2014).

An interesting question arising from this study is why the spread appears to be larger in the Gulf Stream than in other western boundary currents. One possible explanation is related to the stronger salinity fronts in this region, such that small variations in the Gulf Stream path can produce strong salinity anomalies. But other factors can contribute as well, such as the uncertainty associated with deep-water formation, sea ice, and a larger uncertainty in the representation of the Gulf Stream path itself (compared with other western boundary currents). The uncertainty introduced by the assimilation method cannot be discounted

either, and it would be interesting to evaluate the uncertainty pattern of ocean model simulations, as well as that of individual ensemble-based data assimilation systems.

Depth of 20°C isotherm

Variations in the thermocline depth are associated with major modes of tropical climate variability. The depth of the 20°C isotherm (D20) has been considered as part of this intercomparison project as a proxy for thermocline depth and variability in the tropical oceans. D20 monthly means from the different ORAs in Table 2 and from two OO products (EN3v2a and ARMOR3D) have been used.

The absolute value of D20 depends on the vertical discretisation of the model used in each reanalysis. Most of the products have between 16–25 levels in the upper 200 m depth. There is the small group of eddy-permitting, NEMO based reanalyses, characterized by high vertical resolution of the upper ocean (1 m in the first level, then 31 levels for the first 200 m depth). There is also some ambiguity regarding the definition of D20 monthly means included in the evaluations: these can be either 'monthly means of D20 from instantaneous values' or 'the D20 from the monthly means of the temperature field'. In this preliminary diagnostics, different groups have used different methods (Hernandez et al. 2014).

Figure 7 shows the spatial pattern of D20 in the EM-ORA (Figure 7(a)), the differences between the two OO products (Figure 7(b)), and the difference between EM-ORA and each of the OO products (Figure 7(c) and Figure 7 (d)). On average, EM-ORA is shallower than the OO products in the centre of gyres, and deeper on both eastern and western boundaries of the ocean basins. There are also large differences at the western boundaries, especially along the Gulf Stream, which may be related with the misrepresentation of the path of western boundary currents by the models. However, differences along the western boundary currents are also large between the OO products EN3v2a and ARMOR3D (Figure 7(b)). Compared to the OO products, the D20 EM-ORA is slightly deeper in the Equatorial

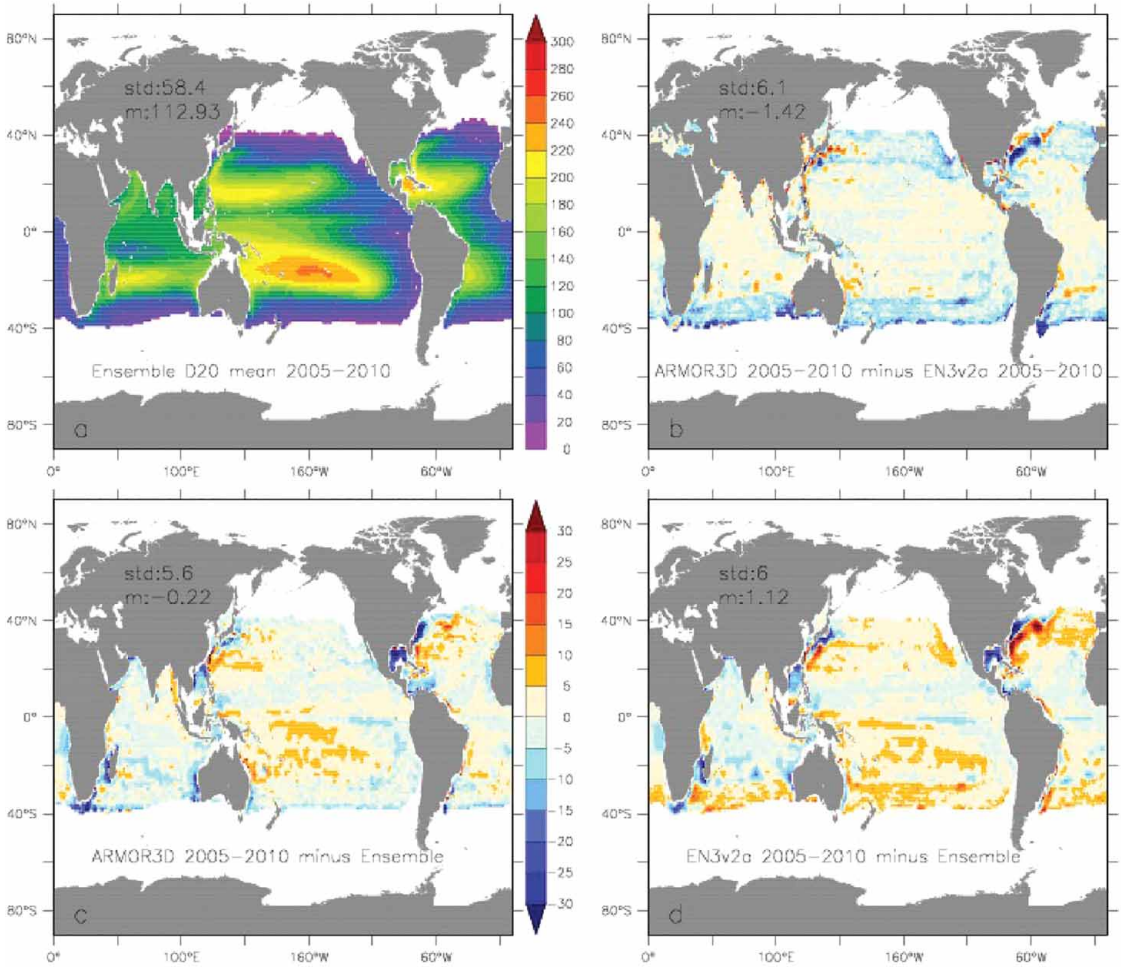

Figure 7. (a) Global map of mean of D20 from EM-ORA. Differences in mean D20 between (b) the two OO products ARMOR3D and EN3v2, (c) EM-ORA and ARMOR3D and (d) EM-ORA and EN3v2. Units are m. The mean fields have been calculated over the 2005–2010 periods.

Indian, Atlantic and Eastern Pacific Oceans, and shallower in the Pacific Warm Pool. The reasons for this unexpected difference will be investigated in future work.

Sea ice

Several studies have suggested that sea ice thickness may be a predictor for seasonal sea ice extent (Kauker et al. 2009; Chevallier et al. 2013). This highlights a weakness in almost all ice forecasting systems in that they don't include the explicit assimilation of ice thickness observations. Moreover, it remains to be seen how the predictability of the seasonal ice cover depends on the representation of various physical processes and model details, such as spatial resolution and the inclusion of an ice thickness distribution. By intercomparing various properties of the sea ice cover in existing ice-ocean reanalyses, it may be possible to highlight deficiencies and best practises in these systems toward answering the question: Are current ice-ocean reanalyses suitable for initializing seasonal forecasts of the ice cover? Here we present preliminary results from this intercomparison.

The ice-ocean reanalyses considered here use a variety of model resolutions, physics and analysis methods. Reanalysis details are provided in Table 2. For the ECMWF reanalysis system, two additional versions of the system were considered whereby only the method of ice assimilation was varied (ERAL-linear, ERAN-non-linear (Tang et al. 2013); note that these products do not appear in Table 2).

Sea ice models used here include two community models, the Los Alamos Community Ice model (CICE (Hunke & Lipscomb 2010)) and the Louvain sea Ice Model (LIM (Fichefet & Maqueda 1997)), as well as independently developed models. While these models and their particular implementation details may vary widely, an important distinction is the representation of the ice thickness distribution. Some models include a sophisticated multi-category approach, while others use a single ice category. This different treatment of ice thickness impacts

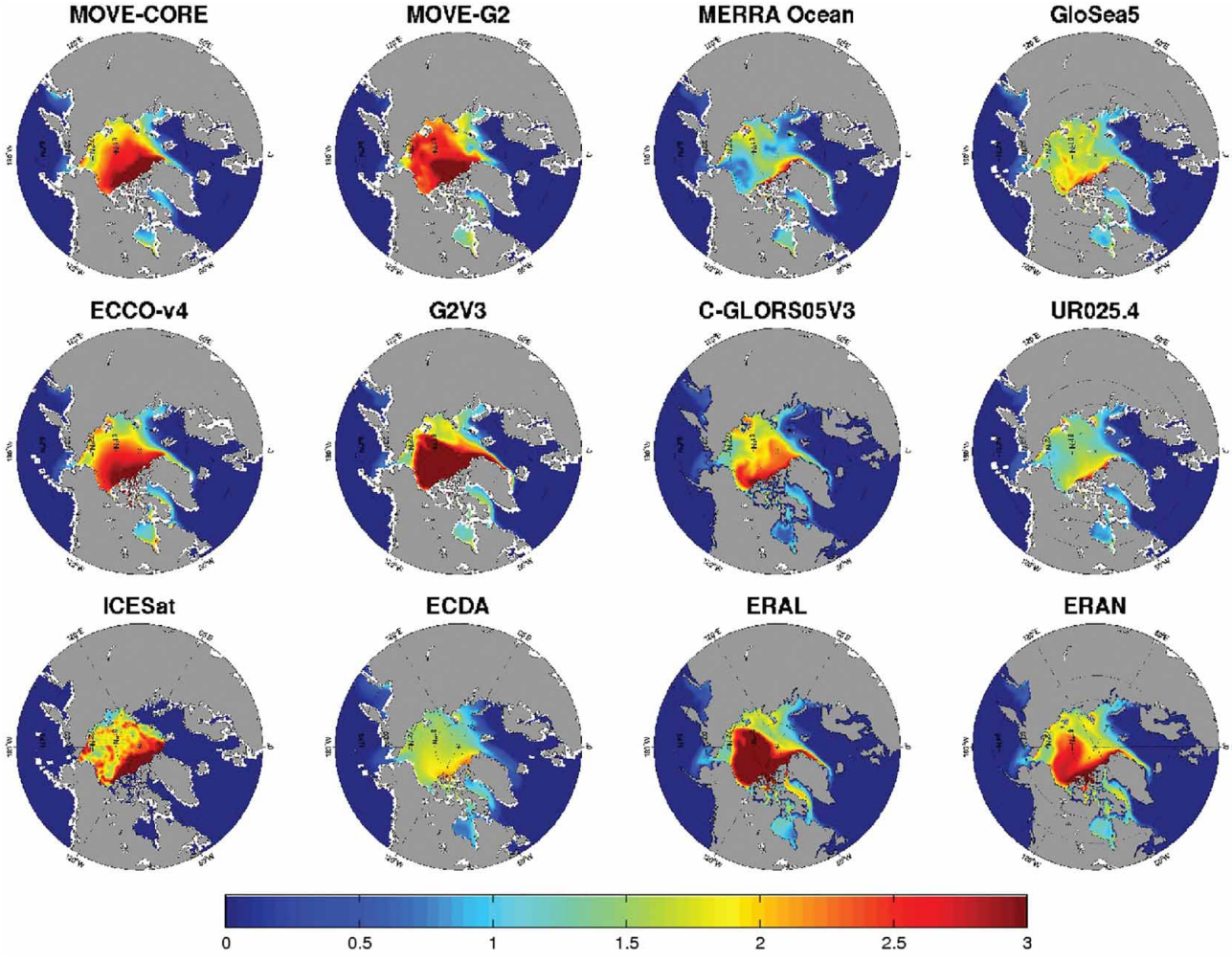

Figure 8. Example of mean sea ice thickness for the various ice-ocean reanalyses for March 2007. Also shown is a satellite estimate of sea ice thickness from ICESat (bottom left).

both ice dynamics and thermodynamics (Martensson et al. 2012).

Another important distinction is in the application of ice assimilation. Many systems employ a simple nudging of ice concentration toward a gridded ice analysis product (e.g. OSISAF, NSIDC), while a few systems use more sophisticated ice assimilation methods (e.g. 3DVar, SEEK). However, perhaps the most important aspect of ice assimilation is in how the increments to concentration affect ice thickness. Two systems (ECMWF and Mercator) supplied different versions of their reanalyses with/without ice assimilation and the impact on ice thickness is non-negligible, albeit unconstrained (not shown).

Figure 8 shows an example highlighting the large range of mean ice thicknesses found for the various ice-ocean reanalyses in March 2007. Also shown is a satellite estimate of the ice thickness derived from ICESat (Kwok & Rothrock 2009) for February/March 2007. In general, the reanalysis products all exhibit the basic feature of thicker ice cover north of the Canadian Arctic Archipelago and Greenland as seen in the observations, albeit to a varying degree. However, the thickness of ice in the central Arctic and along the Siberian coast varies widely. In particular, the reanalyses tend to cluster toward either overly thin ice (~1 m) or overly thick ice (>3 m), with perhaps only one or two showing realistic thicknesses of about 2 m. These differences are larger than interannual variations and are on the scale of the decadal thinning of the ice cover (not shown). The relative contribution of the various factors (e.g. model physics and resolution, atmospheric forcing, data assimilation) that may be contributing to these differences is a topic of on-going study. Such large biases may limit the usefulness of these products for seasonal forecasting.

Summary

This paper presents the first results of the ORA-IP, which aims at exploiting the diversity of existing ocean reanalyses to identify those aspects that are robustly represented by the different products and those where there is a large level of discrepancy. The agreement can be exploited to define indices for monitoring or verification, while the discrepancies point towards areas for future enhancement of assimilation and observing systems. The paper also illustrates the use of independent evaluation metrics to measure the quality of the ensemble mean and individual products, thus providing guidance on the adequacy of the ensemble approach.

The intercomparison has focused on a small set of ocean variables, interpolated into a common horizontal grid, and for a limited set of vertical levels (when applicable). The intercomparison period is mainly 1993–2010, although shorter periods are also used. Where relevant (mixed layer, ocean heat content, steric height, sea level, salinity and thermocline depth) the ensemble mean of the ocean reanalyses was compared with observation-only estimates, to assess if the model-derived estimates show any systematic differences from the observation-only estimates. The ensemble spread is also used as a measure of the existing uncertainty.

It is shown that in general the ensemble mean is usually a better estimation than any individual ocean reanalyses. However, in the case of coastal sea level variability, the evaluation with tidegauge data indicates that ORAs with high-resolution models and assimilation of altimeter are more skilful, and the scores are better when using a sub-ensemble including the subset of best ORAs instead of the grand ensemble.

Systematic differences between OOs and ORAs are largest in the tropics, where model-physics and the wind variability are key assets for the ORAs. These differences are seen in the thermocline and mixed layer depth. In addition, the ensemble of ORAs performs better than the OO products in the estimation of steric height variability at seasonal and interannual time scales in the Atlantic and outside the tropics.

The surface heat flux estimates from ocean reanalyses were compared with other products, mostly based on atmospheric reanalyses. Although large uncertainty still exists, the ocean reanalyses global surface heat fluxes appear more balanced than the atmospheric-based products, especially when the contribution of the assimilation increments is taken into account. The results suggest that data assimilation methods and ocean observations can contribute to the estimation of surface heat fluxes.

The estimation of interannual variability of salinity continues to be a challenge. Signal-to-noise ratios larger than one are confined to the tropical western Pacific, dominated by the ENSO signal. This was the case in the previous intercomparison of reanalyses (circa 2006) (Stammer et al. 2010; Lee et al. 2009), and continues to be so now, in spite of the increased salinity observations in recent years. More work is needed to establish the source of uncertainty (changing observing system; e.g. differences before and after Argo, forcing fields, assimilation methods and error specification).

The intercomparison of sea ice showed a large uncertainty in the estimation of sea ice thickness, which is largely unconstrained by the assimilation methods, highlighting the need for observations of ice thickness for both assimilation and validation.

This ORA-IP has also identified areas where the uncertainty is large, thus providing a focus for future developments in the observing system and modelling/data assimilation. The deep ocean (below the top few hundred metres), the Southern Ocean (Antarctic Circumpolar Current region), coastal areas and the path of western boundary currents appear as the areas with largest uncertainty in the density, temperature and salinity fields. Not

only are the differences between ORAs and OO products the largest, but there is also a large spread among ORAs (as expected from model error), and among OOs (likely because observation representativeness errors are large). These are also important areas for climate.

It is clear that we are still a long way from providing ocean estimations that can answer satisfactorily many fundamental questions, and that continuous development of the assimilation and observing system is needed. In the meantime, the multi model ensemble strategy is a pragmatic approach to exploit the current resources. It is also clear that the evaluation of successive vintages of ocean reanalyses should be a continuous process, since it is needed to assess progress and to identify gaps, thus contributing to setting the directions for future developments.

Acknowledgements

This work has been partially funded by the European Commission funded projects MyOcean, MyOcean2 and COMBINE; by the GEMINA project -funded by the Italian Ministry for Environment; by the NERC-funded VALOR project; by the NERC-funded NCEO program; by the Research Program on Climate Change Adaptation of the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government; by the Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101); by NASA's Modelling Analysis and Prediction Program under WBS 802678.02.17.01.25 and by the NASA Physical Oceanography Program; by the NOAA's Climate Observation Division (COD); by the LEFE/GMMC French national program. The MILA-GPV dataset is provided for this study by RCGC/JAMSTEC. The authors would also like to thanks the constructive suggestions of four anonymous reviewers.

Notes

1. Even the low resolution models resolve the Equatorial Rossby Radius of deformation by including meridional grid refinement close to the Equator.
2. Here *s* is a generic temporal index associated to the temporal filter.
3. In the case of linear trends, the spread will increase with the distance to thecenter of the reference period.

References

Ablain M, co-authors. 2013. Two Decades of Global and Regional Sea Level Observations from the ESA Climate Change Initiative Sea Level Project. ESA Living Planet Symposium, Edinburgh, United Kingdom, 9–13 September, 2013.

Alves O, co-authors. 2014. An Assessment of Upper Ocean Salinity Reanalyses from CLIVAR GSOP/GODAE Systems. CLIVAR EXCHANGES 64. Feb 2014.

Balmaseda MA, co-authors. 2010. Role of the Ocean Observing System in an End-to-End Seasonal Forecasting System. In Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 1), Venice, Italy, 21–25 September 2009, Hall J, Harrison D.E, Stammer D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.pp.03.

Balmaseda MA, Mogensen K, Weaver AT. 2013. Evaluation of the ECMWF ocean reanalysis system ORAS4. Q J R Meteorol Soc. 139: 1132–1161, doi:10.1002/qj.2063.

Balmaseda MA, Trenberth KE, Kallen E. 2013. Distinctive climate signals in reanalysis of global ocean heat content. Geophys Res Lett. 40: 1754–1759, doi:10.1002/grl.50382.

Behringer DW. 2007. The global ocean data assimilation system at NCEP. Preprints, 11th Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, San Antonio, TX, Amer Meteor Soc. 3.3. [Available online at https://ams.confex.com/ams/87ANNUAL/webprogram/Paper119541.html.].

Bellucci A, Gualdi S, Masina S, Storto S, Scoccimarro E, Cagnazzo C, Fogli P, Manzini E, Navarra A. 2013. Decadal climate predictions with a coupled AOGCM initialized with oceanic reanalyses. Clim Dyn. 40: 1483–1497.

Berry DI, Kent EC. 2009. A New Air-Sea Interaction Gridded Dataset from ICOADS with Uncertainty Estimates. Bull Amer Meteor Soc. 90(5): 645–656, doi:10.1175/2008BAMS2639.1.

Blockley E, co-authors. 2013. Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts. Geosci Model Dev Discuss. 6: 6219–6278, doi:10.5194/gmdd-6-6219-2013,2013.

Carton JA, Giese BS. 2008. A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA). Mon Wea Rev. 136: 2999–3017, http://dx.doi.org/10.1175/2007MWR1978.1.

Cazenave A, Llovel W. 2010. Contemporary Sea Level rise. Annu Rev Mar Sci. 2: 145–73.

Chambers DP, Bonin JA. 2012. Evaluation of Release-05 GRACE time-variable gravity coefficients over the ocean. Ocean Sci. 8: 859–868.

Chang YS, Zhang S, Rosati A, Delworth TL, Stern WF. 2013. An assessment of oceanic variability for 1960-2010 from the GFDL ensemble coupled data assimilation. Clim Dyn. 40(3-4): 775–803, doi:10.1007/s00382-012-1412-2.

Chevallier M, Salas y Melia D, Voldoire A, Deque M, Garric G. 2013. Seasonal Forecasts of the Pan-Arctic Sea Ice Extent Using a GCM Based Seasonal Prediction System. J Clim. 26: 6092–6104, doi:10.1175/JCLI-D-12-00612.1.

Church JA, co-authors. 2011. Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. Geophys Res Lett. 38 (L18601), doi:10.1029/2011GL048794.

Crosnier L, Le Provost C. 2006. Internal metrics definition for operational forecast systems inter-comparison: Example in the North Atlantic and Mediterranean Sea. Ocean Weather ForecastingNetherlands, Springer, 455–465.

Crosnier L, Le Provost C. 2007. Inter-comparing five forecast operational systems in the North Atlantic and Mediterranean basins: The MERSEA-strand1 Methodology. J Mar Syst. 65 (1): 354–375.

Danabasoglu G, co-authors. 2013. North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states. Ocean Modell. http://dx.doi.org/10.1016/j.ocemod.2013.10.005.

de Boisséson E, Balmaseda MA, Abdalla S, Källén E, Janssen PAEM. 2014. How robust is the recent strengthening of the Tropical Pacific trade winds? Geophys Res Lett. 41: 4398–4405, doi:10.1002/2014GL060257.

de Boyer Montegut C, Madec G, Fischer AS, Lazar A, Iudicone D. 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile- based climatology. J Geophys Res. 109(C12003), doi:10.1029/2004JC002378.

Dee DP, co-authors. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc. 137: 553–597, doi:10.1002/qj.828.

Dee DP, Balmaseda MA, Balsamo G, Engelen R, Simmons AJ, Thépaut J-N. 2014. Toward a consistent reanalysis of the climate system. BAMS. e-view doi: [http://dx.doi.org/10.1175/BAMS-D-13-00043.1]

England MH, co-authors. 2014. Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. Nat Clim Change. 4: 222–227, doi:10.1038/nclimate2106.

Ferry N, Barnier B, Garric G, Haines K, Masina S, Parent L, Storto A, Valdivieso M, Guinehut S, Mulet S. 2012. NEMO: the modelling engine of global ocean reanalyses. Mercator Ocean Quarterly Newsletter 46: 46–59.

Fichefet T, Maqueda MAM. 1997. Sensitivity of a sea ice model to the treatment of ice thermodynamics and dynamics. J Geophys Res. 102: 12609–12646.

Fujii Y, Nakaegawa N, Matsumoto S, Yasuda T, Yamanaka G, Kamachi M. 2009. Coupled climate simulation by constraining ocean fields in a coupled model with ocean data. J Clim. 22: 5541–5557.

Fukumori I. 2002. A partitioned Kalman filter and smoother. Mon Wea Rev. 130: 1370–1383.

Greatbatch R. 1994. A note on the representation of steric sea level in models that conserve volume rather than mass. J Geophys Res. 99: 12767–12771.

Guinehut S, Dhomps AL, Larnicol G, Le Traon PY. 2012. High resolution 3D temperature and salinity fields derived from in situ and satellite observations. Ocean Sci. 8: 845–857, doi:10.5194/os-8-845-2012.

Haines K, Valdivieso M, Zuo H, Stepanov VN. 2012. Transports and budgets in a 1/4 ° global ocean reanalysis 1989–2010. Ocean Sci. 8(3): 333–344, doi:10.5194/os-8-333-2012.002/qj.2063.

Hasumi H, Tatebe H, Kawasaki T, Kurogi M, Sakamoto TT. 2010. Progress of North Pacific modelling over the past decade. Deep-Sea Res. II 57: 1188–1200, doi:10.1016/j.dsr2.2009.12.008.

Herandez F, co-authors. 2014. Performance evaluations, near real-time assessment of operational oceanography forecast products. J Oper Oceanogr. 7(3) (Special)

Hernandez F, co-authors. 2014. Sea Level Inter- Comparison: Initial results. CLIVAR EXCHANGES 64. Feb 2014.

Hernandez F, co-authors. 2014. ORA-IP Depth of the 20°C isotherm: First results. CLIVAR EXCHANGES 64. Feb 2014.

Hernandez F, Bertino L, Brassington G, Chassignet E, Cummings J, Davidson F, Drévillon M, Garric G, Kamachi M, Lellouche J M et al. 2009. Validation and intercomparison studies within GODAE. Oceanography 22(3): 128–143, doi:10.5670/oceanog.2009.71.

Hosoda S, Ohira T, Sato K, Suga T. 2010. Improved description of global mixed-layer depth using Argo profiling floats. J Oceanogr. 66: 773–787, doi:10.1007/s10872-010-0063-3.

Hunke EC, Lipscomb WH. 2010. CICE: the Los Alamos sea ice model documentation and software user's manual version 4.1. Tech. Rep. LA-CC-06-012Los Alamos National Laboratory.

Ingleby B, Huddleston M. 2007. Quality control of ocean temperature and salinity profiles - historical and realtime data. J Mar Syst. 65: 158–175, doi:10.1016/j.jmarsys.2005.11.019.

Kanamitsu M, Ebitsuzaki W, Woolen J, Yang SK, Hnilo JJ, Fiorino M, Potter G. 2002. NCEP-DOE AMIP-II reanalysis (R-2). Bull Amer Meteor Soc. 83: 1631–1643.

Kauker F, Kaminski T, Karcher M, Giering R, Gerdes R, Voßbeck M. 2009. Adjoint analysis of the 2007 all time Arctic sea-ice minimum. Geophys Res Let. 36(3).

Köhl A. 2014. Evaluation of the GECCO2 Ocean Synthesis: Transports of Volume, Heat and Freshwater in the Atlantic. Q J R Meteorol Soc, doi:10.1002/qj.2347.

Kwok R, Rothrock DA. 2009. Decline in Arctic sea ice thickness from submarine and ICESat records: 1958-2008. Geophys Res Let. 36(15).

Large W, Yeager S. 2009. The global climatology of an interannually varying air-sea flux data set. Clim Dyn. 33: 341–364.

Lee T, Awaji T, Balmaseda MA, Grenier E, Stammer D. 2009. Ocean state estimation for climate research. Oceanography 22: 160–167, doi:10.5670/oceanog.2009.74.

Levitus S, co-authors. 2012. World Ocean heat content and thermosteric sea level change (0–2000 m) 1955-2010. Geophys Res Lett 39(L10603). doi:10.1029/2012GL051106.

Locarnini RA, Mishonov AV, Antonov JI, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Johnson DR. 2010. World Ocean Atlas 2009, Volume 1: Temperature. S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.

Lyman JM, Good SA, Gouretski VV, Ishii M, Johnson GC, Palmer MP, Smith DM, Willis JK. 2010. Robust warming of the global upper ocean. Nature. 465: 334–337, doi:10.1038/nature09043.

Martensson S, Meier HEM, Pemberton P, Haapala J. 2012. Ridged sea ice characteristics in the Arctic from a coupled multicategory sea ice model. J Geophys Res. 117(C8).

Masina S, Di Pietro P, Storto A, Navarra A. 2011. Global ocean reanalyses for climate applications. Dyn Atmos Oceans 52((1-2), SI): 341–366, doi:10.1016/j.dynatmoce.2011.03.006.

Masuda S, co-authors. 2010. Simulated Rapid Warming of Abyssal North Pacific Waters. Science. 329: 319–322, doi:10.1126/science.1188703.

Meyssignac B, Becker M, Llovel W, Cazenave A. 2012. An assessment of two-dimensional past sea level reconstructions over 1950-2009 based on tide gauge data and different input sea level grids. Survey in Geophysics, doi:10.1007/s10712-011-9171-x.

Mogensen K, Balmaseda MA, Weaver AT. 2012. The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. Tech. Memo 668, Reading, UK, ECMWF.

Mulet S, Rio MH, Mignot A, Guinehut S, Morrow R. 2012. A new estimate of the global 3D geostrophic ocean circulation based on satellite data and in situ measurements. Deep-Sea Res. II. 77(80): 70–81, doi:10.1016/j.dsr2.2012.04.012.

Oke P, Martin M, Balmaseda MA, Brassington G, Wilmer-Becker K. 2011. Report on the GODAE Ocean View - CLIVAR GSOP Workshop on Observing System Evaluation and Intercomparison https://www.godae-oceanview. org/outreach/meetings-workshops/task-team meetings/godae-oceanview-gsop-clivar-workshop/.

Oke PR, Brassington G, Cummings J, Martin M, Hernandez F. 2012. GODAE Inter-comparisons in the Tasman and Coral Seas. J Oper Oceanogr. 5(2): 11–24.

Palmer M, co-authors. 2010. Future Observations for Monitoring Global Ocean Heat Content In Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21–25 September 2009, Hall J, Harrison D.E, Stammer D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.68.

Palmer M, co-authors. 2014. CLIVAR-GSOP/GODAE intercomparison of ocean heat content: initial results. CLIVAR EXCHANGES 64. Feb 2014.

Pohlmann H, Doug S, Balmaseda MA, Keenlyside NS, Masina S, Matei D, Muller WA, Rogel P. 2013. Predictability of the

mid-latitude Atlantic meridional overturning circulation in a multi model system. Clim Dyn. 41, doi:10.1007/s00382-013-1663-6.

Saha S, co-authors. 2010. The NCEP climate forecast system reanalysis. Bull Am Meteorol Soc. 91: 1015–1057.

Speer K, Forget G. 2013. Global distribution and formation of mode waters. Chapter 9 in "Ocean Circulation and Climate - A 21st century perspective", nternational Geophysics Series, Vol.103. Edited by Sielder G, Church J, Griffes S, Gould J, Church J. Academic Press, Elsevier. ISBN: 978-0-12-391851-2.

SSALTO/DUACS User Handbook: (M)SLA and (M) ADT Near-Real Time and Delayed Time Products. CLSDOS- NT-06-034, SALP mU-P-EA-21065-CLS, v4.0. 69, http://www.aviso.oceanobs.com/fileadmin/documents/data/tools/hdbk_duacs.pdf.

Stammer D, co-authors. 2010. Ocean Information Provided Through Ensemble Ocean Syntheses. OceanObs'09: Sustained Ocean Observations and Information for Society, doi:10.5270/OceanObs09.cwp.85.

Storto A, co-authors. 2014. Comparison of Steric Sea Level from Ocean Reanalyses and Objective Analyses. CLIVAR EXCHANGES 64. Feb 2014.

Storto A, Dobricic S, Masina S, Di Pietro P. 2011. Assimilating along-track altimetric observations through local hydrostatic adjustments in a global ocean reanalysis system. Mon Wea Rev. 139: 738–754.

Sugiura N, Awaji T, Masuda S, Mochizuki T, Toyoda T, Miyama T, Igarashi H, Ishikawa Y. 2008. Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. J Geophys Res. 113(C10017), doi:10.1029/2008JC004741.

Tang YM, Balmaseda MA, Mogensen KS, Keeley SPE, Janssen PAEM. 2013. Sensitivity of sea ice thickness to observational constraints on sea ice concentration. ECMWF Tech Memo Number 707.

Toyoda T, co-authors. 2014. Mixed layer depth intercomparison among global ocean syntheses reanalyses. CLIVAR EXCHANGES 64. Feb 2014.

Toyoda T, Fujii Y, Yasuda T, Usui N, Iwao T, Kuragano T, Kamachi M. 2013. Improved analysis of the seasonalinterannual fields by a global ocean data assimilation system. Theor Appl Mech Jpn. 61: 31–48, doi:10.11345/nctam.61.31.

Tsujino H, Hirabara M, Nakano H, Yasuda T, Motoi T, Yamanaka G. 2011. Simulating present climate of the global ocean-ice system using the Meteorological Research Institute Community Ocean Model (MRI.COM): simulation characteristics and variability in the Pacific sector. Journ Oceanogr. 67: 449–479, doi:10.1007/s10872-011-0050-3.

Uppala SM, co-authors. 2005. The ERA-40 re-analysis. Q J R Meteorol Soc. 131: 2961–3012.

Valdivieso M, co-authors. 2014. Heat fluxes from ocean and coupled reanalyses. CLIVAR EXCHANGES 64. Feb 2014.

Waters J, Lea DJ, Martin MJ, Mirouze I, Weaver A, While J. 2014. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. Q J R Meteorol Soc. doi:10.1002/qj.2388.

Wijffels S, Willis J, Domingues CM, Barker P, White NJ, Gronell A, Ridgway K, Church JA. 2009. Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. J Climate. 21: 5657–5672.

Wunsch C, Heimbach P. 2013. Dynamically and Kinematically Consistent Global Ocean Circulation and Ice State Estimates. Chapter 21 in "Ocean Circulation and Climate - A 21st century perspective", International Geophysics Series, Vol.103. Edited by Sielder G, Church J, Griffes S, Gould J, Church J. Academic Press, Elsevier. ISBN: 978-0-12-391851-2.

Xie J, Zhu J, Li Y. 2008. Assessment and inter-comparison of five high-resolution sea surface temperature products in the shelf and coastal seas around China. Cont Shelf Res. 28(10): 1286–1293.

Xue Y, co-authors. 2012. A Comparative Analysis of Upper-Ocean Heat Content Variability from an Ensemble of Operational Ocean Reanalyses. J Climate. 25: 6905–6929, doi:[http://dx.doi.org/10.1175/JCLI-D-11-00542.1].

Xue Y, co-authors. 2010. Ocean state estimation for global ocean monitoring: ENSO and beyond ENSO. In OceanObs'09: Conference on Sustained Ocean Observations and Information for Society, vol. 2, Venice, 21–25, September 2009. Hall J, Harrison DE, Stammer D. (eds). ESA publication WPP-306, doi:10.5270/OceanObs09.

Xue Y, Huang B, Hu ZZ, Kumar A, Wen C, Behringer D, Nadiga S. 2011. An Assessment of Oceanic Variability in the NCEP Climate Forecast System Reanalysis. Clim Dyn. 37: 2511–2539.

Yin Y, Alves O, Oke PR. 2011. An ensemble ocean data assimilation system for seasonal prediction. Mon Wea Rev. 139: 786–808.

Yu L, Jin X, Weller RA. 2008. Multidecade Global Flux Datasets from the Objectively Analyzed Air–sea Fluxes (OAFlux) Project: Latent and Sensible Heat Fluxes, Ocean Evaporation, and Related Surface Meteorological Variables Woods Hole Oceanographic Institution Technical Report OAFlux Project (OA2008–01).

Zhang S, Harrison MJ, Rosati A, Wittenberg AT. 2007. System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. Mon Wea Rev. 135(10), doi:10.1175/MWR3466.1.

Zhang Y, Rossow WB, Lacis AA, Oinas V, Mishchenk MI. 2004. Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets. J Geophys Res: Atmos. (1984–2012). 109(D19).

Zhao M, Hendon HH, Alves O, Yin Y. 2013b. Impact of improved assimilation of temperature and salinity for coupled model seasonal forecasts. Submitted to Mon Wea Rev.

Zhao M, Hendon HH, Alves O, Yin Y, Anderson DLTA. 2013a. Impact of salinity constraints on the simulated mean state and variability in a coupled seasonal forecast model. Mon Wea Rev. 141: 388–402.

Zhu J, Huang B, Balmaseda MA, Kinter JL. III, Peng P, Hu ZZ, Marx L. 2013. Improved reliability of ENSO hindcasts with multi-ocean analyses ensemble initialization. Clim Dyn. 11/2013 41(9–10), doi:10.1007/s00382-013-1965-8.

Zhu J, Huang B, Balmaseda MA. 2011. An ensemble estimation of the variability of upper-ocean heat content over the tropical Atlantic Ocean with multi-ocean reanalysis products. Clim Dyn. doi:10.1007/s00382-011-1189-8.

Zhu J, Huang B, Marx L, Kinter JL. III, Balmaseda MA, Zhang R-H, Hu Z-Z. 2012. Ensemble ENSO hindcasts initialized from multiple ocean analyses. Geophys Res Lett. 39: L09602, doi:10.1029/2012GL051503.