CrossMark

The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals

Hao Zuo[1] · Magdalena A. Balmaseda[1] · Kristian Mogensen[1]

Abstract A new eddy-permitting ocean reanalysis has been recently completed at ECMWF. It is called Ocean ReAnalysis Pilot 5 (ORAP5), and it spans the period 1979–2012. This work describes the new system, evaluates its performance, and investigates how the estimation of climate indices are affected by the assimilation system settings. ORAP5 introduces several upgrades with respect to its predecessor ORAS4, including increased horizontal and vertical resolution, an prognostic sea-ice component, new versions of the ocean and data assimilation system, revised surface fluxes, new version and treatment of satellite sea surface height data, and assimilation of sea-ice concentration, among others. ORAP5 shows similar performance to ORAS4, with improvements in the northern extratropics (especially in salinity), and slight degradation in the Southern Ocean, probably because the observations are insufficient to constrain the increased level of variability in ORAP5. The sensitivity experiments show that superobbing of altimeter data and correlation length-scales of the background errors have a visible impact on the time evolution of global steric height and its partition into thermo/halo-steric contributions. The sensitivities are especially large in the pre-Argo period, when there is the risk of producing unrealistic steric height variations by overfitting the altimeter data. Compared with a control run without data assimilation, all the assimilation experiments also show stronger variability in the halosteric component in the pre-Argo period. The results highlight the importance of sub-surface observations to assist the assimilation of altimeter data, and the need of using a variety of metrics for evaluating ocean reanalysis systems.

Keywords Ocean reanalyses · Sea level · Sensitivity experiment

✉ Hao Zuo
haozuo6@googlemail.com

[1] European Centre for Medium-Range Weather Forecasts, Reading, UK

1 Introduction

Ocean reanalyses (ORAs in what follows) are historical reconstructions of the ocean states, obtained by using an ocean model driven by atmospheric forcing fluxes, and constrained by ocean observations (surface and profiles) via data assimilation methods. These reconstructions are used for climate studies (Balmaseda et al. 2013b; Mayer et al. 2014; England et al. 2014; Chen and Tung 2014; Drijfhout et al. 2014, among others) and also often used for the initialization of seasonal and decadal forecasts (e.g. Zhu et al. 2012, 2013; Pohlmann et al. 2013; Guemas et al. 2012; Balmaseda et al. 2010).

Evaluating and quantifying the uncertainty in current estimates of climate indicators is crucial for the understanding and prediction of climate internal and forced variability. This is one of the objectives of the current Ocean Reanalyses Intercomparison Project (ORA-IP, Balmaseda et al. 2015). Reliable estimation of the ocean state and associated uncertainty 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 needs to provide 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 method, which are often flow-dependent, and not easy to estimate. The uncertainty of ORAs can be estimated from comparison with observational data, assimilated or not. Another 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, and this is theme of several of papers submitted to this special issue within the context of the ORA-IP project (Chevallier et al. 2015; Shi et al. 2015; Palmer et al. 2015; Toyoda et al. 2015a, b; Karspeck et al. 2015).

In order to make progress towards reducing the uncertainties, it is important to identify where the uncertainty stems from. To this end, the multi-system approach needs to be assisted by other focused studies, since the uncertainty comes from a variety of aspects in the reanalysis systems. In addition, improved estimations are only possible by continuous system upgrades, based on the lessons learnt from previous evaluation studies.

The work presented here aims at evaluating uncertainty on important climate indices arising from the choice of assimilation parameters within a single ocean-reanalysis system. To this end we use the new ECMWF eddy-permitting ocean and sea-ice reanalysis system ORAP5 (Zuo et al. 2015b), which has been produced as a contribution to the multi-ocean-reanalyses program within the EU FP7 MyOcean-2 project. It is also the baseline for the next ECMWF operational reanalysis system. The performance of ORAP5 is evaluated by comparing it with the current operational ocean reanalysis system ORAS4 (Balmaseda et al. 2013a) and an equivalent ocean-only simulation constrained only by sea surface temperature (CNTL). A series of controlled sensitivity experiments within the ORAP5 framework are conducted by modifying some parameters of the assimilation system, mainly associated with the assimilation of altimeter data. The impact of these parameters on the estimated global mean sea-level (GMSL) trend attribution is presented and discussed.

This article is organised as follows. The configuration and setup of ORAP5 is described in Sect. 2. Assimilation statistics for ORAP5, as well as some preliminary evaluation are presented in Sect. 3. Sensitivity experiments are presented in Sect. 4, with the focus on evaluation of satellite altimeter data assimilation. The ORAP5 sea-ice has been thoroughly evaluated by Tietsche et al. (2014).

2 ORAP5 system configuration

2.1 Overview

The ORAP5 ocean reanalysis covers the period 01-Jan-1979 to 31-Dec-2012. It has been produced using the V3.4.1 of the NEMO ocean model (Madec 2008) at a resolution of 0.25° in the horizontal and 75 levels in the vertical, with variable spacing (the top level has 1 m thickness). It also includes a prognostic thermodynamic-dynamic sea-ice model (LIM2, Fichefet and Maqueda 1997). ORAP5 surface forcing comes from ERA-Interim (Dee et al. 2011), and includes the impact of surface waves in the exchange of momentum and turbulent kinetic energy (Janssen et al. 2013) (see Sect. 2.2). The reanalysis is conducted with NEMOVAR (Mogensen et al. 2012) in its 3D-Var FGAT (the first-guess at appropriate time) configuration (see Sect. 2.4). NEMOVAR is used to assimilate subsurface temperature, salinity, sea-ice concentration (SIC) and sea-level anomalies, using a 5 day assimilation window with 1200 s model step. In addition, sea surface temperature (SST), sea surface salinity (SSS), and global mean sea-level trends are used to modify the surface fluxes of heat and freshwater (see Sects. 2.3 and 2.5). The observational information is also used via an adaptive bias correction scheme (Balmaseda et al. 2013a) (see Sect. 2.6). The main differences between ORAS4 and ORAP5 system settings are summarized in Table 1. Details of system upgrades can be found in Zuo et al. (2015b).

2.2 Ocean and sea-ice model, spin-up and forcing fields

ORCA is a common NEMO ocean model global configuration that uses a tri-polar grid with the three poles located over Antarctic, Central Asia and North Canada. The 2012 reference version of DRAKKAR high resolution ORCA configurations (ORCA025.L75, see Barnier et al. 2006) has been used here for the NEMO ocean model (version 3.4). The ORCA025.L75 configuration is a grid with 0.25° resolution at the equator and increases to 12 km in some areas in the Arctic Ocean. There are 75 vertical levels with resolution increasing from 1 m near the surface to 200 m in the deep ocean. The vertical discretization scheme uses partial steps to have better representation of the flow over steep topography. The bathymetry is derived from ETOPO1 (Amante and Eakins 2009) with a minimum depth set to 3 m. Vertical diffusion coefficients are determined using a turbulent kinetic energy (TKE) scheme (Madec 2008). Solar penetration in the ocean is calculated using the 2 bands scheme and constant Chlorophyll concentration (0.05 mg m^{-3}). The LIM2 sea-ice model has been used

Table 1 Overview of differences between ORAS4 and ORAP5 ocean reanalysis settings

	ORAS4	ORAP5
Grids	$\sim 1°$, 42 Vertical levels	$\sim 0.25°$, 75 Vertical levels
Model	NEMO 3.0, no sea-ice model	NEMO 3.4.1, LIM2 ice model
Forcing	Direct surface fluxes from ERA40 and ERA-Interim	ERA-Interim with bulk formula + wave forcing (Breivik et al. 2015)
Assimilation		
Background error	Constant meridional length-scale	Rossby-radius dependent meridional length-scale
Sea-ice	No assimilation	Sea-ice concentration assimilation
Prior-1993 global fresh-water closure	Altimeter sea level climatology	GRACE bottom-pressure climatology
Assimilation window	10 days	5 days
Observations		
SST	OIv2 SST until 2010, then OSTIA (NWP)	OSTIA reanalysis + OSTIA(NWP) + Reynolds OIv2d
T/S prof	EN3 until 2010, then GTS	EN3 to 2012
SL	AVISO altimeter (ECMWF version 2-3-4-5)	AVISO altimeter (ECMWF version 4) with revised MDT
Sea-ice	N/A	OSTIA gridded sea-ice concentration data
Period	1959-Present	1979–2012
Ensemble	5 Ensemble members	1 Ensemble member

here coupled hourly with NEMO and with the viscous-plastic (VP) rheology.

Several modifications have been made to the standard NEMO version in order to represent the impact of surface waves in the ocean mixing and circulation. The enhanced mixing due to a flux of TKE from breaking waves, represented in NEMO by means of a constant parameter, has been modified to use instead the spatially and time varying TKE flux derived from the surface waves, which can be obtained from ERA-Interim reanalysis (Janssen et al. 2013). Surface wave information from ERA-Interim is also used to modify the momentum flux. The Stokes–Coriolis forcing, a term arising from the interaction of the wave momentum and the rotation of the Earth, is computed from the Stokes drift and other wave parameters computed by the ECMWF WAM model (ECMWF 2013). This term is added as a tendency to the horizontal momentum equation in NEMO. Finally, the stress on the water-side will differ slightly from the air-side stress due to storage and release of momentum in the wave field. This momentum flux is also computed by the ECMWF WAM model. The transfer coefficient for momentum is defined directly from the wave model drag coefficient, and it is used as input to the CORE bulk formula (Large and Yeager 2009) to derive transfer coefficients for sensible/latent heat and evaporation computation. The implementation of these processes in the NEMO ocean model, as well as their impact on the ocean mean state and variability is described in Breivik et al. (2015). All the surface forcing fields are from ECMWF ERA-Interim atmospheric reanalysis product (Simmons et al. 2007; Dee et al. 2011).

The initial conditions for the ORAP5 were produced in two phases. First, a 12-year (1979–1990) model spin up was performed from rest and with temperature and salinity fields defined from the World Ocean Atlas 2009 climatology (WOA09, see Locarnini et al. 2010; Antonov et al. 2010), forced with ERA-Interim fluxes, and using a 3-year relaxation to the same WOA09 climatology. This is followed by a 5-year assimilation period (1975–1979), starting from the end of the previous spin up conditions and forced with ERA-40 fluxes (Uppala et al. 2005).

2.3 Observations

In-situ profiles of temperature and salinity data from the quality-controlled EN3 data set (Ingleby and Huddleston 2007) are assimilated in ORAP5. EN3 version 2 with XBT depth corrections (Wijffels et al. 2008) is used from 1979 to 2011 and a standard version EN3 (version EN3_v2a) is used for year 2012. These are subjected to the NEMOVAR automatic quality control (QC) procedure, which includes a duplicate check, background check and stability check, among others. The same shallow water rejection scheme as used in ORAS4 is applied to ORAP5 to reject all observations in regions where model depth is less than 500 m, so that observations on the continental shelves are not assimilated. Arguably this is a choice that may need revisiting since ORAP5 has higher resolution than ORAS4. A horizontal thinning scheme is applied to CTD and XBT data with a minimum distance requirement of 25 km and time gap set to 1-day. A vertical thinning scheme with no more

than two observations per model level is also applied to all in-situ profiles (see Mogensen et al. 2012 for details). The thinning procedure is a pragmatic way to reduce the impact of spatial observation error correlations. This is different from ORAS4, which allowed three observations per model level for vertical thinning, and uses 100 km as minimum distance for horizontal thinning (Balmaseda et al. 2013a).

ORAP5 also assimilates along-track altimeter-derived sea-level anomalies (SLA) data from AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic data) delayed mode data set. (The altimeter products were produced by Ssalto/Duacs and distributed by AVISO, with support from Cnes-http://www.aviso.altimetry.fr/duacs/). It includes observations from ERS-1, ERS-2, Envisat, Jason-1, Jason-2 and Topex/Poseidon. The most up-to-date AVISO SLA at the time of production was used when producing ORAP5. In comparison, ORAS4 uses different versions of AVISO data: prior to operational implementation it used the delayed AVISO product available in 2010; during its operational phase, ORAS4 has been using the near-real-time product from subsequent AVISO releases.

To filter out the correlation on the SLA observation error, a super-observation scheme (hereafter superob) as implemented in ORAS4 is also used in ORAP5 for SLA data. A grid with approximately 100 km resolution is defined (superob grid). Altimeter observations are then binned in time and space: observations within the same day and within each area of the superob-grid are averaged to create a superob observation (see Mogensen et al. 2012 for details). Experiments show that applying superob on SLA data has a large impact in the ocean subsurface (see Sect. 4). An alternative solution is to thin the SLA observations instead of superobbing, which is not yet implemented.

To assimilate AVISO SLA, a new method was developed which can calculate the model mean dynamic topography (MDT) file relative to an arbitrary period. The MDT is still derived from a previous assimilation run where T and S are assimilated. But instead of using the same reference period (period 1993–1999) as AVISO SLA, the MDT in ORAP5 is estimated by averaging the model sea surface height during the 2000–2009 period, when the large scale ocean is adequately sampled by Argo. A spatially dependent correction factor is then added to take into account the different reference periods used by model and observations. The correction factor is estimated as the differences in the altimeter SLA means between the two different periods. This new MDT estimation method has been validated in low resolution (1°) ORAP5-equivalent experiments, with little impact on the analysis results.

The daily mean gridded SIC data are now assimilated in NEMOVAR. As for SST, this comes from a combination of NOAA and OSTIA products (Zuo et al. 2015b). A thinning algorithm was applied to the SIC data to reduce the data density to a grid resolution of ∼0.5°. Both SIC and other observations are assimilated using the 5-day assimilation cycle. A thorough evaluation of SIC assimilation in ORAP5 has been carried out by Tietsche et al. (2014).

As in ORAS4, the sea surface temperature data in ORAP5 is used to correct the turbulent surface heat fluxes. This is done via a restoring term, with the strength set to $-200\,\mathrm{Wm^{-2}K^{-1}}$. Instead of the low resolution OIv2 SST used in ORAS4, ORAP5 SST are based on the high-resolution OSTIA SST reanalysis (Donlon et al. 2012), available for the period 1985–2007. The operational real-time OSTIA SST are used from 2008 onwards (Roberts-Jones et al. 2012). For the periods when OSTIA is not available, the NOAA optimal interpolation 0.25° daily SST analysis (OIv2d, Reynolds et al. 2007) is used. Before 1982 the SST are from then ERA-40 reanalysis (Uppala et al. 2005). See Table 1 and Zuo et al. (2015b) for details.

A sea surface salinity data from WOA09 monthly climatology is applied as a freshwater fluxes in ORAP5, with a relaxation constant equivalent to a 1-year timescale. The monthly climatological values from river runoff (Dai and Trenberth 2002) is also applied along the land mask and treated as a freshwater flux. Volume of the global freshwater fluxes is constrained using observation data (see Sect. 2.5). In addition, a global 3D relaxation with a timescale of about 20 years to temperature and salinity climatological value from WOA09 is applied through the water column to avoid long-term model drifts.

2.4 Data assimilation scheme

NEMOVAR is a variational data assimilation system developed from the OPAVAR data assimilation system (Weaver et al. 2005) for the NEMO ocean model by Mogensen et al. (2012). For our analysis NEMOVAR is applied as an incremental three-dimensional variational assimilation (3D-Var) using FGAT approach. In FGAT the observations and model states are compared at the appropriate time during the first model integration covering the assimilation window (sometimes called first outer loop). The resulting misfits are the input for a cost function that is minimized to produce the assimilation increment. This minimization takes place in the "inner loop". In ORAP5 the minimization is done in observation space using the restricted preconditioned conjugate gradient (RPCG) algorithm implemented in NEMOVAR (Gürol et al. 2014). In the final phase of the analysis cycle, the assimilation increment resulting from the inner-loop minimization is applied using incremental analysis updates (IAU; Bloom et al. 1996), with constant weights, during a second model integration (second outer loop) spanning the same time window as for the assimilation window. The assimilation cycle in ORAP5 is 5 days.

The specification of the background error (BGE) covariance in ORAP5 is similar to the one in ORAS4, in that it uses the same balance relationships between variables and similar flow dependency. The temperature background error depends on the stratification and on the distance to the coast. Temperature and salinity increments are linked by the so-called "T-S" relationship, that aims at preserving hydrostatic balance. The part of salinity that can not be explained by the "T-S" relationship is called the unbalanced component, and is corrected independently. Density and currents are related via geostrophy (in the new NEMOVAR version there is a small modification to the beta-plane approximation). The sea level variations are separated into balanced and unbalanced parts, which approximately correspond to baroclinic and barotropic components. The balanced component is associated with vertical displacements of the water column whenever the stratification is strong; the unbalanced component does not result in direct temperature and salinity increments (it will in turn change the barotropic velocities). More explicit formulation of the NEMOVAR BGE can be found in Mogensen et al. (2012) and in Balmaseda et al. (2013a).

The parameters of the BGE had to be revised to take into account the higher resolution of the ocean model. In particular, the model for the BGE horizontal correlation scales has been changed using a simplification of the scheme proposed by Waters et al. (2014). Changes have also been made to the vertical decorrelation scales, which are specified in ORAP5 as a scalar ($\alpha = 1$) multiple of the local vertical grid-size dz (see Section 4.6.2 in Mogensen et al. 2012). The impact of this parameter will be discussed in Sect. 4.5. The details on the implementation of this revised correlation scheme in ORAP5, as well as the specification of model background-error (BGE) and observation-error (OBE) can be found in Zuo et al. (2015b). The choice of these BGE specifications in ORAP5 implies that the volume of ocean potentially affected by an in-situ observation is smaller in ORAP5 than in ORAS4.

2.5 Global fresh water closure

The model freshwater flux is adjusted by constraining the global model sea-level changes to the changes given by the altimeter data after 1993. Before that, the globally-averaged fresh-water variations are constrained by the bottom-pressure climatology derived from GRACE (Gravity Recovery and Climate Experiment, Tapley et al. 2004). This latter is an upgrade to the existing scheme in ORAS4.

After 1993, Altimeter-derived Global Mean Sea Level (GMSL) variations is available and can be assimilated in ORAP5 following the same scheme as that in ORAS4, and described in Balmaseda et al. (2013a). This scheme uses the fact that the GMSL variations can be decomposed into

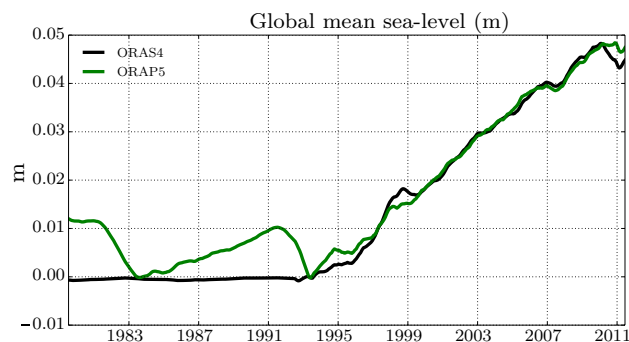$$\Delta \bar{\eta}_o = \Delta \bar{\eta}_s + \Delta \bar{\eta}_m \tag{1}$$

Fig. 1 Global mean sea level anomaly from 1979 to 2011 in ORAS4 and in ORAP5, with 12-month running mean and value from 1993 Jan removed. The differences before 1993 are due to the different constraint used (climatological mass in ORAP5 and climatological sea level in ORAS4). The differences after 1993 are due to the different altimeter versions

where $\Delta \bar{\eta}_o$ is the observed GMSL change between two given times; $\Delta \bar{\eta}_s$ is the steric component of GMSL change, which is derived from model density fields; $\Delta \bar{\eta}_m$ is the GMSL change due to mass variation. During the altimeter-era, $\Delta \bar{\eta}_o$ can be estimated from the altimeter observations, and the mass contribution is then estimated as the residual between the total GMSL variations and the model-derived steric component. This is then applied as a spatially uniform fresh-water flux.

Before the altimeter era however, there is very limited information about the GMSL, and additional assumptions are needed. One possibility is to use alternative GMSL reconstructions from tide gauge information (Church and White 2011; Hay et al. 2015). But in ORAP5 we have followed a different approach, that consist on constraining the global mass variations using a climatology of global mass changes from the GRACE gravity mission. In ORAS4, the interannual variations of total GMSL were neglected, and the daily climatology of GMSL from the altimeter for the period 1993–1999 was used. In ORAP5, this has been modified, and we assume that the mass variations of GMSL $\Delta \bar{\eta}_m$ are well approximated by the climatology, which is estimated from the GRACE-derived bottom-pressure data for the period 2005–2009. Figure 1 shows the time series of the resulting GMSL in ORAS4 and ORAP5. It shows that the modified scheme in ORAP5 allows for interannual variations in GMSL due to changes in the steric height. The differences after 1993 are due to the different versions of the AVISO product used (see Sect. 2.3).

2.6 Bias correction scheme

A bias correction scheme (Balmaseda et al. 2007) has been implemented in NEMOVAR to correct temperature/salinity biases in the extra-tropical regions, as well as a pressure

correction in the tropical regions. The total bias term contains two terms: (i) a-priori bias (offline bias), which is estimated based on a pre-production run from 2000 to 2009 with only assimilation of in-situ temperature and salinity and accounts for the seasonal variations; and (ii) an online bias, which is updated each analysis cycle using assimilation increments. The scheme is basically the same as that applied in ORAS4 (Balmaseda et al. 2013a) but with minor modifications. For example, the off-line bias term was estimated using the climatology of 3D relaxation terms and increments, instead of online bias term from a pre-production run used in ORAS4. The time evolution of the online term has also been modified in such a way that the temperature and salinity bias can be updated faster at mid latitudes. See Fig. 3 in Zuo et al. (2015b) for specific details.

Figure 2 shows the annual mean of 300–700 m averaged temperature (upper panels) and salinity (lower panels) offline bias correction applied in ORAS4 and ORAP5. In general, the temperature and salinity bias correction patterns are similar between ORAS4 and ORAP5, suggesting common model/forcing errors. The largest value corrections are found along the western boundary currents for both ORAS4 and ORAP5, although differences along the North-Atlantic drift and Labrador Sea are visible. The bias correction pattern along the Kuroshio Current in ORAP5 has much finer structure than that in ORAS4, reflecting the large amount of eddy variability. Comparing with ORAS4, the temperature bias term in ORAP5 is reduced significantly in the Southern Ocean, Labrador Sea and over the whole North-Eastern Atlantic basin. Along the Northern edge of the Antarctic Circumpolar Current, both temperature and salinity bias corrections have a continuous and sharp frontal structure in ORAP5, while ORAS4 does not have any clear sign of front. In general, the bias terms have finer structure in ORAP5 than in ORAS4. This is a consequence of the higher model resolution, but it also reflects the smaller spatial scales used for the assimilation increments, and the modified strategy for the bias estimation.
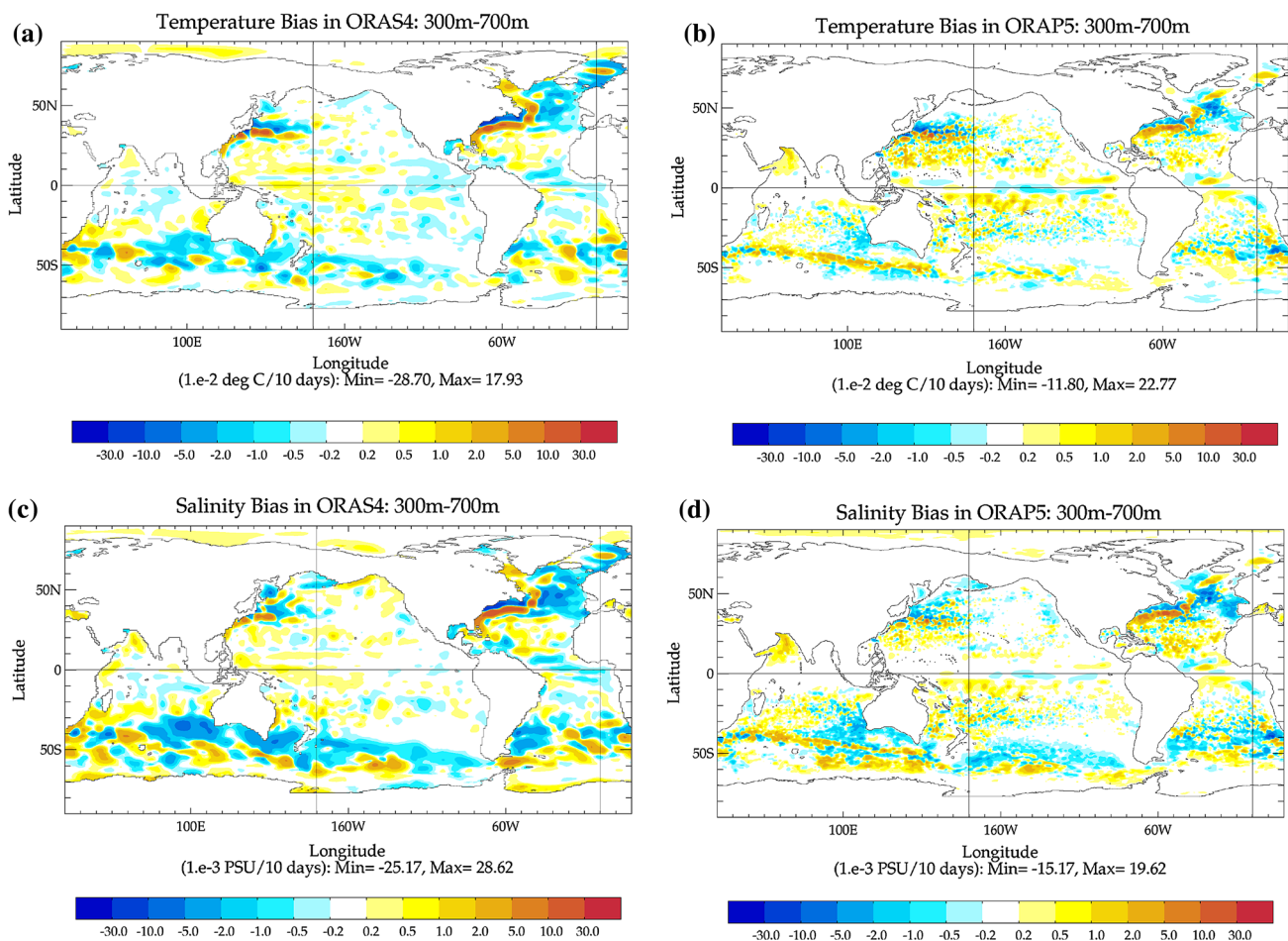

Fig. 2 (*Top*) Annual mean offline temperature bias (300–700 m) as applied in **a** ORAS4 and **b** ORAP5; (*Bottom*) Annual mean offline salinity bias (300–700 m) as applied in **c** ORAS4 and **d** ORAS5

3 Preliminary evaluation of ORAP5

3.1 Assimilation statistics in observation space

The quality-controlled EN3 data set has been used for evaluation of model fit to observations. Bias and root mean square error (RMSE) statistics of the background minus observations are presented. The background is the model state from the first outer loop before updating the model variables with IAU. ORAP5 statistics are compared with those of ORAS4 and a control integration (hereafter CNTL). CNTL is a NEMO-LIM2 simulation using the same initial condition, forcing fields, SST relaxation and SSS climatology relaxation as ORAP5 but not assimilating observation data. The CNTL global mean sea level is also constrained using the same scheme as described in Sect. 2.5, but without bias correction since it is considered as part of the NEMOVAR assimilation system. In all the cases, the same observations from EN3 have entered in the statistics, independently on whether or not they were actually assimilated. The statistics are effectively computed before horizontal/vertical thinning or any additional quality control has been applied on the observations (see Sect. 2.3 for details), so the statistics from ORAS4 can be compared with these from ORAP5 and CNTL in the same observation space. Time series and spatial patterns of these statistics averaged in the upper 200 m are presented in Figs. 3 and 4 respectively. Figure 3 shows time series of bias (dashed line) and RMSE (solid line), averaged over the upper 200 m for temperature (Fig. 3a) and salinity (Fig. 3b). Different regions appear in separate panels. Figure 4 shows maps of the bias in temperature (Fig. 4a, c, e) and salinity (Fig. 4b, d, f) for CNTL, ORAP5 and ORAS4, respectively. The statistics in observation space have been gridded by averaging over 5° by 5° boxes.

Figure 3, shows time series for the global ocean, northern extratropics (nxtrp: 30°N to 70°N), southern extratropics (sxtrp: −70°S to −30°S) and tropics (trop: −30°S to 30°N). Globally, the time series of temperature RMSE from ORAP5 and ORAS4 are very similar (Fig. 3a), and both show improvement over the CNTL, with a mean RMSE reduced by ∼0.25°C. Additional experiment with data assimilation but without bias correction has also been carried out and shows similar but slightly increased temperature RMSE over ORAP5 (not shown). These results are very similar to Fig. 7 in Balmaseda et al. (2013a), and suggesting that the bias correction is still insufficient to remove all the model biases. The global mean temperature biases from these three integrations are not so different, but it is difficult to interpret global biases since cancellation of errors can occur with the spatial averaging. Indeed, the left panels in Fig. 4 show that both assimilation experiments (ORAP5 and ORAS4 in Fig. 4c, e), respectively)

exhibit significant smaller temperature bias than the CNTL (Fig. 4a). This is the case for the large scale cold biases (∼0.5 °C) in the Tropics and warm biases around Japan. The warm biases along the Gulf Stream separation region are reduced by the assimilation, but they are not eliminated in ORAP5 or ORAS4.

For salinity, ORAP5 has the smallest global mean error as measured by bias and RMSE (Fig. 3b) among three integrations. The salinity RMSE in ORAS4 is larger than CNTL before 2000, but reduced quickly following the introduction of Argo observations, suggesting relatively large salinity errors in ORAS4 before the Argo-era. There is an obvious declining trend of the salinity global RMSE in all three experiments (Fig. 3b), including CNTL, which does not assimilate any data. This is most likely the result of evaluating the statistics in observation space, since the observation coverage is continuously evolving over time, and only with the Argo data reaches a uniform spatial sampling. Since the model errors in the open ocean are usually smaller than near the coast, the uniform spatial sampling provided by Argo results in reduced global RMSEs. Figure 4b shows the CNTL salinity bias map for the upper 200 m, with strong negative bias over the northern Atlantic Ocean and positive bias in the Labrador Sea, North Pacific subpolar gyre and South Pacific Gyre. Improvements can be seen in Fig. 4d, f for ORAP5 and ORAS4 respectively, due to assimilation of observations. Compared to CNTL, the least improvement was found in data-sparse polar regions, and in coastal regions where most observations were either rejected or associated with large prescribed OBE variances.

In the northern extratropics, ORAP5 shows smaller errors than ORAS4, especially in salinity (Fig. 3), which is likely due to increased model resolution. Here the RMSE is reduced by ∼0.15 PSU in ORAP5 (Fig. 3b). The strong positive salinity bias of ∼0.2 PSU in ORAS4 is related to increased salinity errors in the Gulf Stream region (Fig. 4f). This suggests that the higher model resolution (i.e. horizontal resolution ≤0.25°) helps the better representation of sharp salinity front in ORAP5. It is also possible that this improvement is due to better resolved Labrador Current. ORAP5 also shows reduced salinity bias (Fig. 4d) in the Arctic region, but increased salinity bias in the Barents Sea and along the west coast of Greenland relative to ORAS4. In temperature, the most noticeable improvement over ORAS4 is over the sea of Japan, as shown in Fig. 4c for ORAP5.

For the tropical ocean, only mooring observations are used for computation of model bias and RMSE. This choice is motivated by the fact that the network of moorings is a relatively stationary observing system (in both time and space), although during the period spanned in this evaluation the mooring array was expanded by the
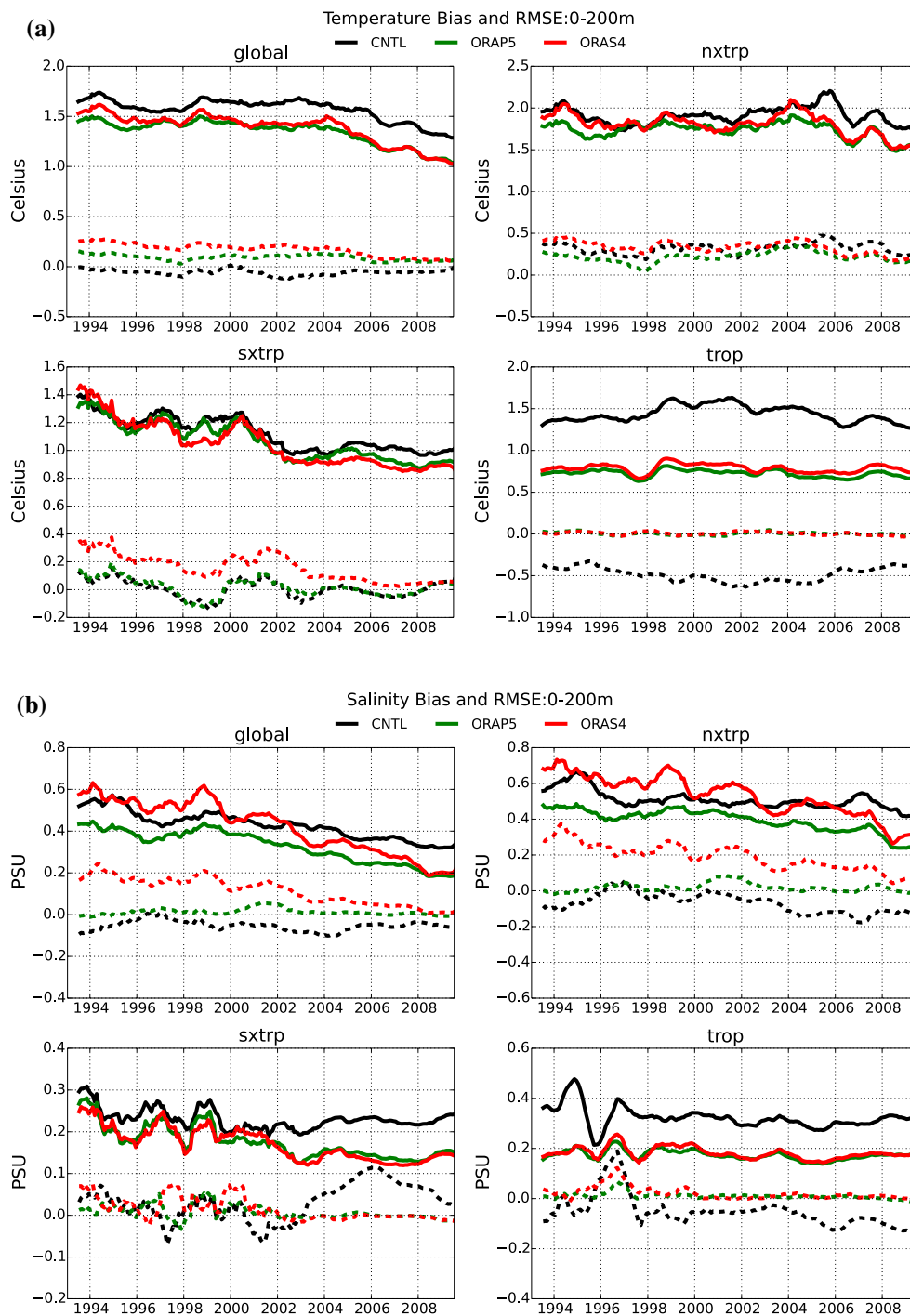
Fig. 3 Time series of model misfit to **a** temperature and **b** salinity observations as bias (*dashed line*) and RMSE (*solid line*) for CNTL (*black*), ORAP5 (*green*) and ORAS4 (*red*) with 12-month running mean filter. Statistics are computed using model background minus observation and in the same observation space of EN3 data, after

averaging over the upper 200 m in different regions: global (−90°S to 90°N), nxtrp (30°N to 70°N), sxtrp (−70°S to −30°S), trop (−30°S to 30°N). *Note* for the tropical regions only mooring observations are used for computation of these statistics

inclusion of the TRITON moorings in the Western Pacific, the PIRATA moorings in the Atlantic, and more recently the RAMA array in the Indian Ocean. Compared with the other regions, the temperature bias and RMSE in the tropics

appear relatively stable (Fig. 3a), except for a small variation around 1998–1999, coinciding with the implementation of TRITON and PIRATA. In this region, the assimilation shows a clear improvement with respect to the CNTL. The
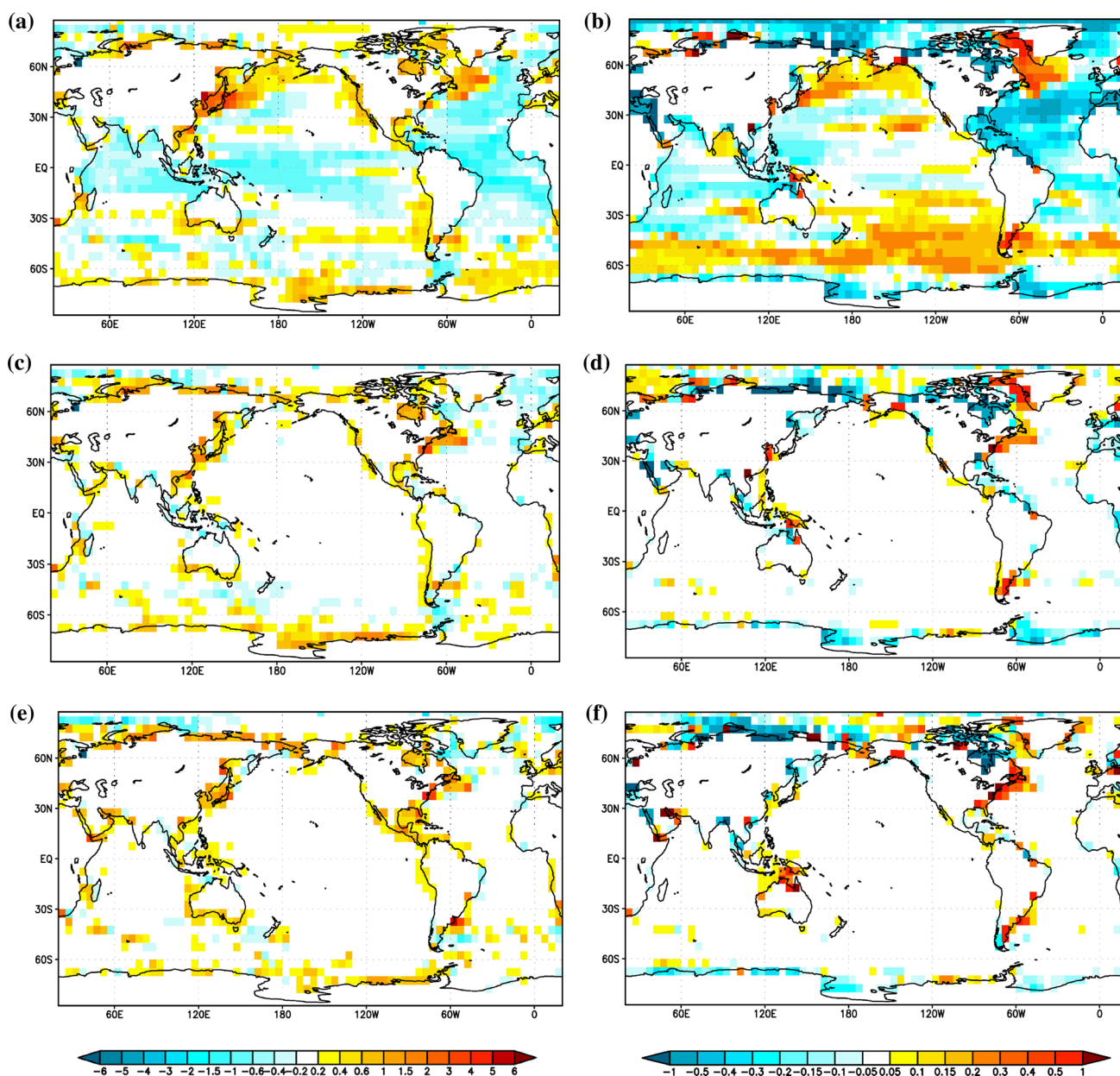
Fig. 4 Maps of model temperature (*left panels*) and salinity (*right panels*) biases (*background-observation*) for CNTL (**a**, **b**); ORAP5 (**c**, **d**) and ORAS4 (**e**, **f**) after averaging from 0 to 200 m and over the period 1993–2009. Statistics are computed using the model back-ground value from the first outer-loop and EN3 in-situ observations before thinning or any additional quality control was applied (i.e. shallow water rejection), and averaged over 5° by 5° *boxes*

ORAS4 and ORAP5 temperature errors are very similar in this tropical area, with reduced temperature bias at the Gulf of Mexico and Indonesian Archipelago for ORAP5 (Fig. 4c). The salinity bias and RMSE in the tropics stabilizes after the introduction of salinity observations in the moorings, roughly around 1998 (Fig. 3b), and again, both assimilation experiments show similar statistics, and a substantially improved performance with respect to the CNTL.

In the southern extratropics, the assimilation statistics are quite variable until approximately 2002, coinciding with the implementation of Argo (Fig. 3a, b). After 2002 the statistics stabilize, and the assimilation experiments show a clear reduction of the bias and RMSE, both with respect to the CNTL and with respect to the period before Argo. ORAS4 shows slightly reduced RMSE in both temperature and salinity relative to ORAP5. ORAS4 have reduced temperature biases near the Drake Passage and when near the coast of Antarctica (Fig. 4e), and reduced salinity biases in the Ross and Weddell Seas (Fig. 4f). This may be a consequence of the higher variability in ORAP5, which remains insufficiently constrained.

3.2 Comparison with other observational estimates

A validation against observations and other ocean state estimates has been carried out for ORAP5 reanalysis including the use sea-level data from satellite altimeters and tide gauges. The same diagnostics from CNTL and ORAS4 reanalysis are also included.

The gridded SLA maps from the ESA sea level essential climate variable (ECV) products (Cazenave et al. 2014) are used here to evaluate the fidelity of the temporal variability in the ocean reanalysis. This sea level maps are produced and validated in ESA Climate Change Initiative (Ablain et al. 2015) project, and is calculated after merging all the altimetry mission measurements together into monthly grids with a spatial resolution of 0.25°. This is a gridded data set where the original altimeter data has been reprocessed with improved algorithms (orbit, wet tropospheric corrections, among others) and ancillary data (using improved atmospheric fields for instance), in order to produce consistent time series of sea level for climate studies. These data has not been directly assimilated in ORAP5 or ORAS4, which use along-track sea level anomalies from AVISO. In what follows, we refer to this product as CCI1.

Figure 5 shows the temporal correlation of the sea level anomalies from these two products. The correlation is high in the tropics and over the poleward side of the Pacific subtropical convergence areas, on both hemispheres. There are large areas in the subtropics and Southern Ocean, where the correlation between these two observational products is relatively low. This result should be interpreted as a reference for the level of temporal correlation between the SLA from ocean reanalysis and the AVISO or CCI1 SLA maps.

Maps of temporal correlation of monthly mean sealevel between three model estimates (CNTL, ORAP5 and ORAS4) and CCI1 maps for the period 1993–2010 are

shown in Fig. 6. This correlation includes both the seasonal and interannual signals. Compared to CNTL, the correlation with CCI1 maps is improved in ORAP5, particularly in the tropical regions. The correlation patterns among ORAS4 are very similar to ORAP5, with slightly improved correlation in the Western Tropical Pacific, along the edge of the North Equatorial Counter Current. However, the CCI1-correlation in the extra-tropical Pacific Ocean and north of 60° N in the Atlantic Ocean are higher in ORAP5 than in ORAS4. It is interesting to note that the spatial patterns of correlation between both ORAP5 and ORAS4 reanalyses and CCI1 are not qualitatively different from the correlation between AVISO and CCI in Fig. 5, although the values are higher in the latter. The same correlation map between the sea level from these three model estimates and AVISO maps are computed and discussed in Zuo et al. (2015b). Comparing with CCI1, the correlation between sea level from these three model estimates and AVISO maps increases almost everywhere. One possible reason is that the same AVISO altimeter data (along-track) has been assimilated in ORAP5 and ORAS4. However the CNTL correlations with AVISO are also higher than with CCI. This is probably because the gridded AVISO MLSA map uses different spatial smoothing. We will return to this point later in Sect. 4.1.

BADOMAR is a specific processed tide gauges database developed and maintained at collecte localisation satellites (CLS) and consists of filtered tide gauge data from the GLOSS/CLIVAR "fast" sea level data tide gauge network. The full BADOMAR data set contains 286 tide-gauge records as daily averaged sea level and has been used for altimeter calibration (Lefèvre et al. 2005). A reduced data set of BADOMAR data including 72 tide-gauge records as monthly mean sea level after being corrected from inverse barometer effect and tides by Mercator Ocean are used here

Fig. 5 Maps of temporal correlation between CCI1 and AVISO monthly gridded sea level data. The statistics have been computed with monthly mean sea level for the period 1993–2010. Only values above 0.4 are shown
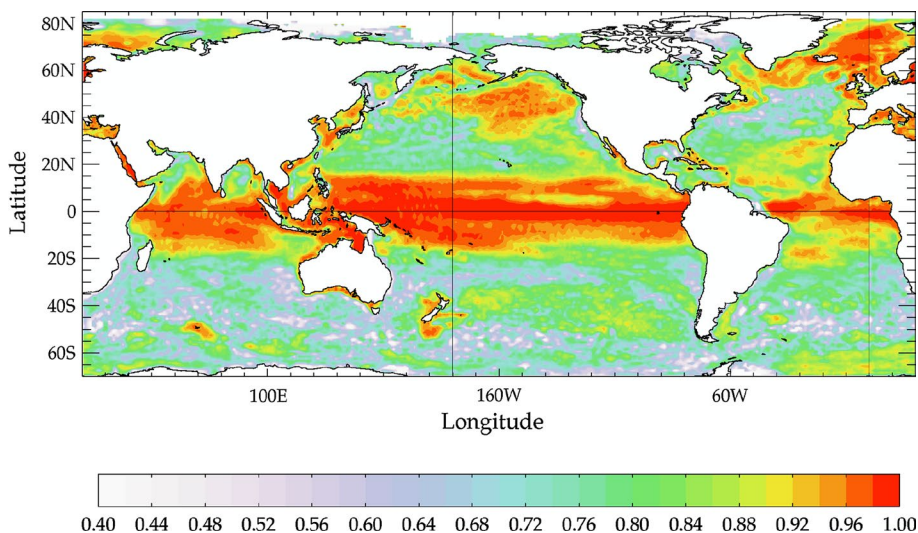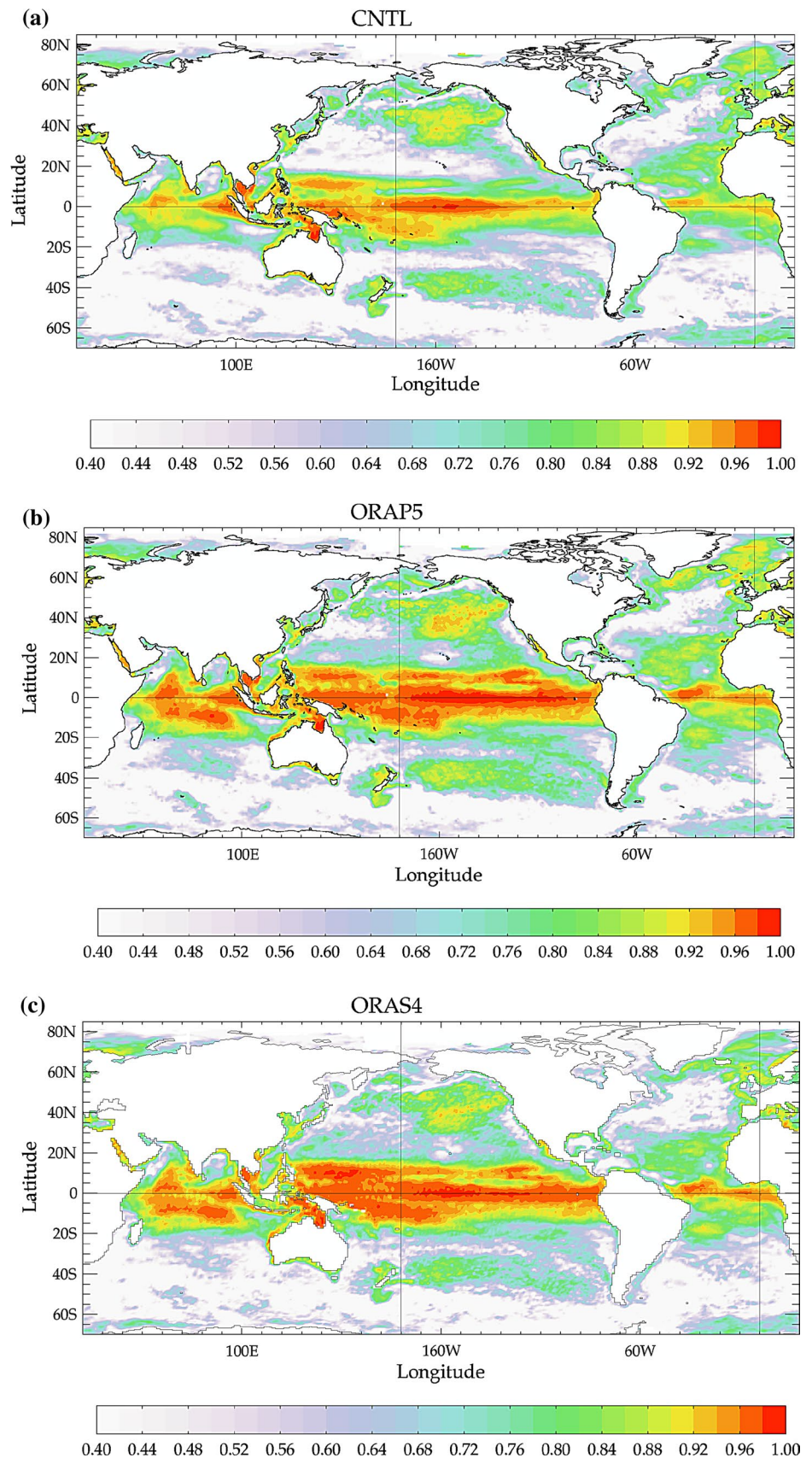
Fig. 6 Maps of temporal correlation between CCI1 sea level and **a** CNTL, **b** ORAP5 and **c** ORAS4 monthly sea level results. The statistics have been computed with monthly mean sea level for the period 1993–2010. Only values above 0.4 are shown. This correlation includes both the seasonal and interannual signals

for independent validation of ORAP5 simulation of SL variations. Correlation between ORAP5 monthly sea level and BADOMAR tide gauges data over the period 1993–2011 are shown in Fig. 7a. For each tide gauge station, the sea-level values at the nearest sea model points are used for comparison. The correlation between ORAP5 SL and BADOMAR tide-gauge records is normally high, with a mean correlation of 0.67. Among all 72 tide gauge stations, over 70 % have correlation value >0.6, and only 10 % have correlation value <0.4. Figure 7b shows the differences in correlation values between ORAP5 and CNTL (ORAP5-CNTL) with respect to BADOMAR tide gauge data. Compared to CNTL, whose global mean sea level trend is also constrained, ORAP5 is in general better correlated than CNTL with the tide-gauge records in most of the locations, except for a few stations (southern tip of Africa, coast of Chile and northern Pacific Ocean) where ORAP5 SL correlation with tide-gauge records is also low. Differences in BADOMAR-correlations between ORAP5 and ORAS4 (Fig. 7c) suggest that ORAP5 performed better for sea level variability in the Atlantic Ocean but slightly worse in the Indian Ocean. This latter is likely a consequence of the larger number of islands resolved in the higher resolution ORAP5. In order to account for representativeness error near the boundaries, observation errors near the coast are inflated in ORAP5 for temperature, salinity and along-track SLA observations. As a result, more islands from high resolution topography used in ORAP5 would lead to increased observation errors due to this inflation strategy. This effect could outweigh the advantage from better resolved island in the higher resolution ORAP5. The performance in the Pacific Ocean is similar between ORAP5 and ORAS4. The mean correlation increases by 0.04 and suggests overall superior performance in ORAP5 relative to ORAS4.

4 Sensitivity experiments

A series of sensitivity experiments has been conducted within the framework of ORAP5, with focus on the evaluation of the satellite altimeter data assimilation. These include sensitivity to superobbing of the satellite altimeter data and specification of the correlation length-scales of the BGEs. System settings for these sensitivity experiments are summarized in Table 2. All sensitivity experiments span the period 19920601–20121232, being initialized at 19920601 from the same ORAP5 initial conditions, and are driven by the same ERA-interim surface fluxes. EN3 in-situ data are assimilated with both horizontal and vertical thinning. AVISO along-track altimeter data are assimilated from 1993 onwards (except for NoSLA). Model bias correction as described in Sect. 2.6 and GMSL constrain as described in Sect. 2.5 are also applied in all sensitivity experiments.

4.1 Correlation with SLA maps

Maps of temporal correlation (period 1993–2010) of monthly mean sea-level between three sensitivity experiments (NoSLA, NoSuperob, Superob1) and CCI1 gridded maps of SLA, are shown in Fig. 8. The lowest correlations are for experiment NoSLA, which shows the level of agreement in sea level that can be reached by assimilating only subsurface temperature and salinity data. In this case, the correlation is largest in the tropics, especially in the Tropical Pacific. The increase of correlation over CNTL (Fig. 6a) is noticeable. Still, the NoSLA correlation map shows that sharp minima in areas of steep thermocline remain: along the North Equatorial Counter Current in the Pacific, and roughly along the paths of the North and South Eqtorial currents in the Indian ocean. Along the Equatorial Pacific, correlation maxima collocated over the tropical mooring array are visible. Compared with the CNTL (Fig. 6a) the assimilation of T/S also increases the correlation with the altimeter data north of 60°N in the Atlantic Ocean.

The assimilation of AVISO along-track altimeter data (in NoSurperob and Superob1) substantially increases the correlation with the CCI1 altimeter maps (Fig. 8b, c). Areas with low correlation remain, especially in regions where the multivariate relationships between altimeter and subsurface are weak (Southern ocean, Western boundary currents), and those with large prescribed OBE variance (along the coast). The areas of low correlation between NoSuperob and CCI1 coincide with the areas where the correlation between CCI1 and AVISO is low. Interestingly, the satellite trajectories can be appreciated in the NoSuperob-CCI1 correlation map. This feature is not visible in the NoSuperob-AVISO correlation maps, and it may be related with the smaller scales used in the mapping of the altimeter data in CCI1.

Superob1 is equivalent to NoSuperob but with the superobbing scheme (see Sect. 2.3) applied to the altimeter data before assimilation. This practically reduces the weight to the altimeter observations (with weaker weights in areas of large representativeness error). Superob1 still shows significantly improved correlation with the altimeter data when comparing with NoSLA, and the pattern and magnitude of the correlation is more similar to NoSuperob than to NoSLA. The Superob1-CCI1 correlation in the tropical regions (between 20°S and 20°N) and subtropical gyres in both Pacific and Atlantic oceans is reduced with respect to NoSuperob. The satellite tracks are not so visible in Superob1. The correlation in Superob1 is very similar to ORAP5 (Fig. 6b) and to Superob2 (not shown).

According to this metric (i.e., fit to the altimeter data) the best estimate is the one obtained with NoSuperob. It is however important to check whether or not this high level of correlation is achieved by over-fitting. A required test is
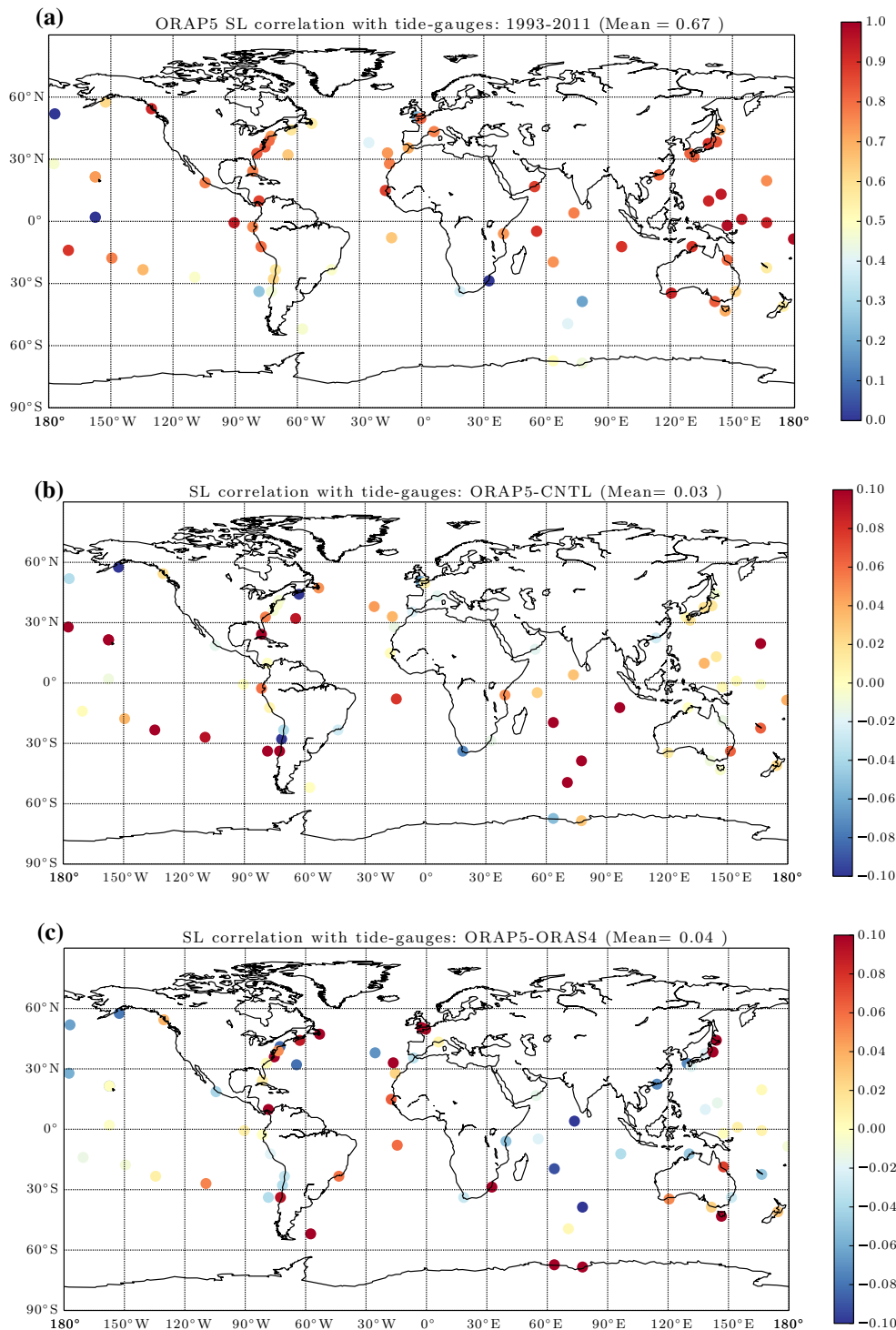
Fig. 7 **a** Correlation between ORAP5 and 72 BADOMAR tide-gauge stations between period 1993–2011; **b** correlation with BADOMAR tide-gauge stations: ORAP5–CNTL; **c** correlation with BADOMAR tide-gauge stations: ORAP5–ORAS4. Statistics are computed using monthly mean sea level analysis from ORAP5 at the nearest model point to each tide gauge station

whether the assimilation of altimeter data improves the fit to the in-situ observations. The impact of altimeter assimilation on climate indices (which are not always observable) is also evaluated next.

4.2 Fit to in-situ observations

The fit to the EN3 in-situ observations is shown in Fig. 9. The mean vertical profiles of model misfits to observations

Table 2 Summary of the sensitivity experiments for sea-level assimilation

Experiment name	Assim SLA	Altimetry superob	Horizontal length-scales	Vertical length-scales
NoSLA	OFF	N/A	$\overline{L}_\eta = 4°$	$\alpha = 2$
NoSuperob	ON	OFF	$\overline{L}_\eta = 4°$	$\alpha = 2$
Superob1	ON	ON	$\overline{L}_\eta = 4°$	$\alpha = 2$
Superob2	ON	ON	$\overline{L}_\eta = 2°$	$\alpha = 2$
ORAP5	ON	ON	$\overline{L}_\eta = 2°$	$\alpha = 1$

$\overline{L}_\eta$ is horizontal correlation length-scales for unbalanced SSH BGE

α is vertical correlation length-scale factor for temperature and unbalanced salinity BGEs

(as measured by RMSE) over the tropical oceans are shown in Fig. 9a for temperature, and Fig. 9b for salinity, respectively. Statistics are calculated using the background value (i.e., the model values are from the first outer loop before correcting the model using IAU) and are averaged over the period 1993–2012. Shown are the profiles of NoSLA, NoSuperob and Superob1. Among three sensitivity experiments, NoSuperob has the largest RMSE (red line in Fig. 9a, b). The discrepancy in both temperature and salinity RMSE between NoSuperob and NoSLA are considered substantial for the upper 800 m water. The degradation of the fit to in-situ observations in NoSuperob is visible in other ocean regions (not shown), although it is only clearly detectable after 2000, with the spin-up of Argo. So it appears that NoSuperob, the assimilation of altimeter data in the model without applying the superobbing scheme, increases the errors in both temperature and salinity. In contrast, the assimilation of altimeter data in Superob1 is able to reduce the temperature RMSE between 50 and 200 m by ∼0.08 °C when compared to NoSLA. The improvement is more obvious in the tropical oceans, being mostly neutral in other regions. These results illustrate that although assimilation of sea level data can improve the fit to in-situ observations, this is not guaranteed. Careful treatment of the altimeter data and evaluation of the results are needed.

### 4.3 Global mean sea level attribution: steric and mass

It is also important to evaluate the impact of the assimilation in relevant climate indices. Global mean sea-level (GMSL) can be decomposed into steric changes and mass changes. The steric changes can in turn be decomposed into thermosteric and halosteric, i.e, changes in volume due to temperature and salinity changes respectively. Here we choose to evaluate how this partition of GMSL is affected by the assimilation parameters. Note that these climate indices are not exactly observable quantities, and can involve areas of the ocean poorly constrained by observations (like the deep and Southern Ocean). This fact makes these sensitivity impact studies quite relevant for quantifying the uncertainty of the resulting estimates.

In ORAP5, the global steric height (GSH) is computed as an area average of the vertical integral of the model density. Mass variations in the ocean are specified as equivalent bottom pressure (EBP) and estimated as the residual between GMSL and GSH (Balmaseda et al. 2013a). The linear trend of GMSL and its components GSH and EBP are computed for all sensitivity experiments over the period 1993–2012, with the results shown in Table 3. The linear trend in the GMSL for this period is about 2.8 mm year^{-1} and is the same in all sensitivity experiments due to assimilation of AVISO sea-level trends. The partition of trends into steric and mass variations is, however, very different. In NoSLA, Superob2 and ORAP5, the contributions to global sea-level trend are approximately evenly distributed among steric and mass changes. In NoSuperob, however, the GMSL trend is dominated by the trend in EBP (2.0 mm year^{-1}), i.e, is due to mass variation, which account for over 70 % of the GMSL trend. These different estimates of GMSL trend and its partition in the sensitivity experiments are due to differences in specification of weights given to observations. In particular, these results indicate that climate signals derived from reanalysis could be very sensitive to the treatment of satellite altimeter data, even within the same assimilation system.

The trends in Table 3 reflect only one aspect of the sensitivities. Perhaps more interesting is the time evolution of the GMSL partition. Figure 10 shows time series of GMSL (black solid), EBP (red dashed) and GSH (green dashed) anomalies (respect to January 1993), estimated from CNTL, ORAP5 and sensitivity experiments (NoSLA, NoSuperob, Superob1 and Superob2). The CNTL experiment (Fig. 10a) only shows an increase in steric height from 2004 onwards, with accelerated rates after 2010. This would imply that in this experiment most of the trends in GMSL for the period 1993–2003 are exclusively due to mass increase. It appears that the estimation of steric height in CNTL during this period is underestimated, which may occur if the ocean is not able to absorb heat, either by incorrect surface forcing or by underestimation of the vertical mixing, among other reasons.

Experiment NoSLA (Fig. 10c), which assimilates temperature and salinity data, shows a slow but steady increase

Fig. 8 Maps of temporal correlation between analysis and CCI1 sea level, with analysis from **a** NoSLA, **b** NoSuperob and **c** Superob1. The statistics have been computed with monthly mean sea level for the period 1993–2010, with only values above 0.4 are shown in the map. This correlation includes both the seasonal and interannual signals
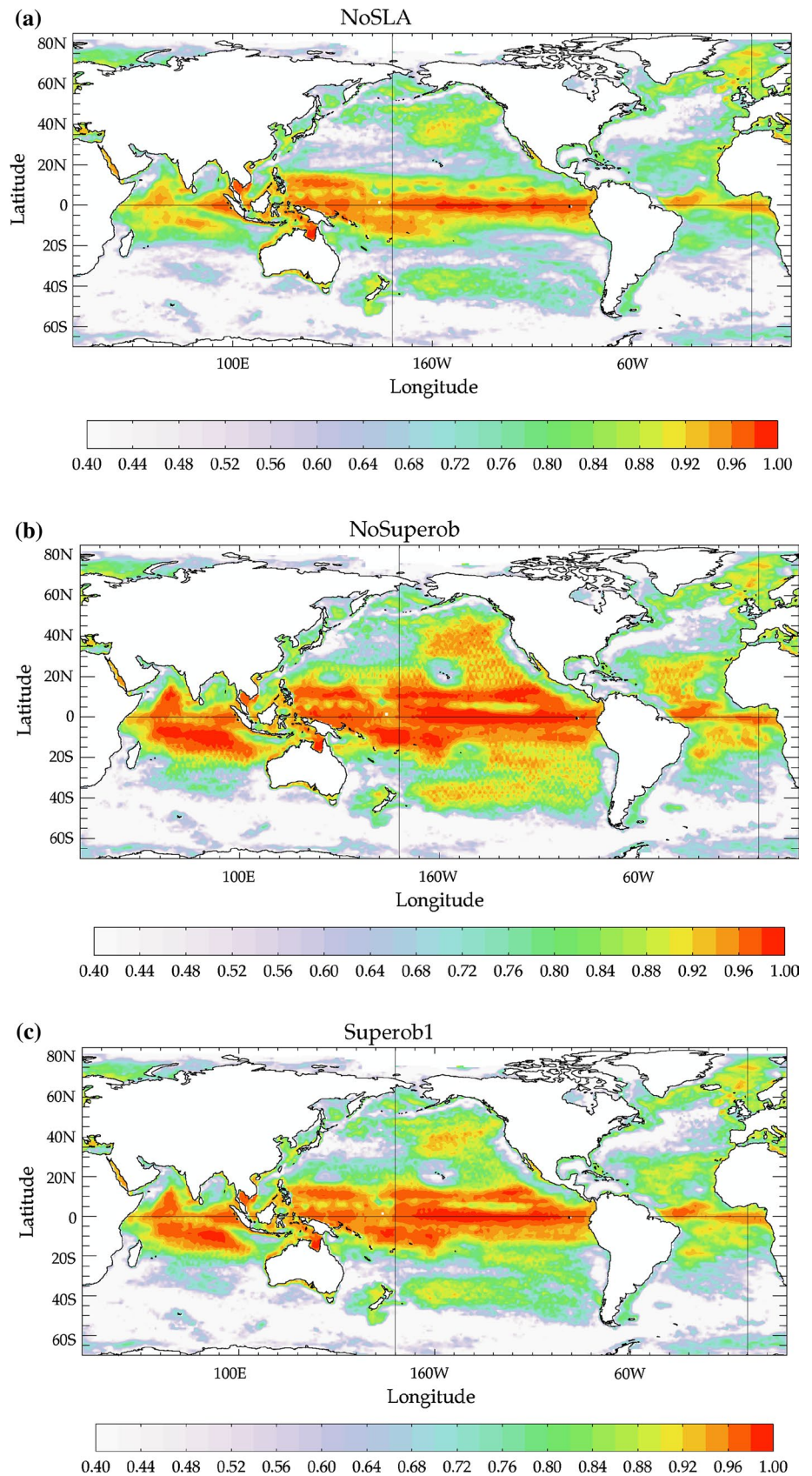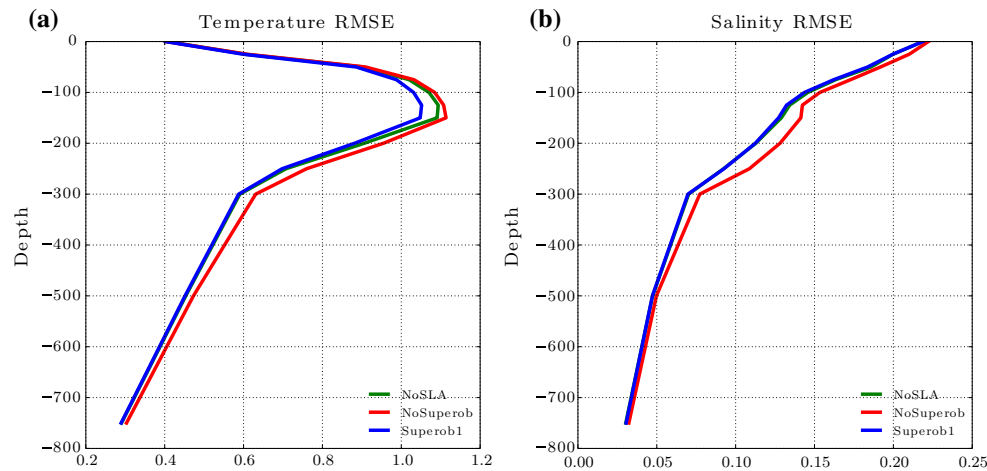
Fig. 9 Profiles of mean model fit to observations RMSE for **a** temperature (°C) and **b** salinity (PSU) for the upper 800 m at the tropical oceans; RMSE are computed using background value from the first outer loop for NoSLA (*green*), NoSuperob (*red*) and Superob1 (*blue*), after being averaged between 1993 and 2012. Here tropical oceans are defined as between 0°E to 360°E and 30°S to 30°N

of the steric contribution during 1993–2000. This increase accelerates from 2000 until 2004 (probably an artifact of the build-up of Argo), after which it continues increasing but at a slower pace. As a result, in experiment NoSLA the increase GMSL is dominated by mass changes during the period 1993–2000, by steric changes during 2000–2004 and partitioned about 1/3 into steric/mass for the period 2005–2012.

The Nosuperob experiment shows a very different behaviour to CNTL and NoSLA (Fig. 10d), with very rapid growth of steric height in the first few years of altimeter assimilation (1993–1998), at a rate of 5.1 mm year^{-1}. This rate exceeds the increase in global sea level, and is arguably unrealistic. This increase in GSH has to be compensated by a strong decrease of EBP (−2.1 mm year^{-1}), achieved by removing ocean mass. After 1998, the steric height in NoSuperob stabilizes; as a consequence, the continuous increase in GMSL is achieved mainly by the increase in ocean mass. The rapid and unrealistic change of steric height in the first 5 years in NoSuperob is probably due to over-fitting the altimeter observations, and illustrates the dangers of assimilating altimeter observations without the anchoring provided by subsurface in-situ observations.

The superobbing scheme in Superob1 effectively reduces the weight to the altimeter observations. Compared to NoSuperob, the partition of GMSL changes into steric/mass in Superob1 is more even (Fig. 10e): the increase in the steric component amounts to about 2/3 of the GMSL for the period 1993–2004, after which the steric increase stabilizes and most of the GMSL from 2005 onwards is due to mass contributions.

Other parameters appear to affect the steric/mass partition, as can be seen by comparing the results of three

experiments that assimilate superobbed altimeter data (ORAP5, Superob1 and Superob2, in Fig. 10b–f respectively). ORAP5 and Superob2 have the same superobbing scheme as Superob1, but the horizontal correlation scale for the barotropic component of the altimeter $\overline{L}_\eta$ has been reduced from 4° (Superob1) to 2° (ORAP5 and Superob2). The reduction in this parameter changes the partition in GMSL, producing a slower increase of the steric component during the period period 1993–2000, compared with Superob1. During this period, the GSH in ORAP5/Superob2 grows slightly faster than in experiment NoSLA. During 2000–2004, it shows an acceleration, which is weaker than that in NoSLA. After 2004, it increases steadily, without any apparent plateau. In these two experiments, the steric/mass partition during the whole period is more even than in any experiment, although the mass contributions appear to dominate for the period 1998–2002, a behaviour also seen in the NoSLA experiment.

4.4 Thermosteric and halosteric contributions

The partition of the steric changes into thermosteric and halosteric (i.e, the relative contributions of temperature and salinity variations to the total volume changes) is another aspect sensitive to the assimilation of data and parameter choice. Figure 10 also shows the time series of thermosteric changes (blue dashed line) and halo-steric changes (cyan dashed line) for the different experiments. In all the experiments, the thermosteric component dominates the changes in steric height trends, but interannual variations of the halo-steric component appear to vary among the experiments. In the CNTL experiment (Fig. 10a), the halosteric
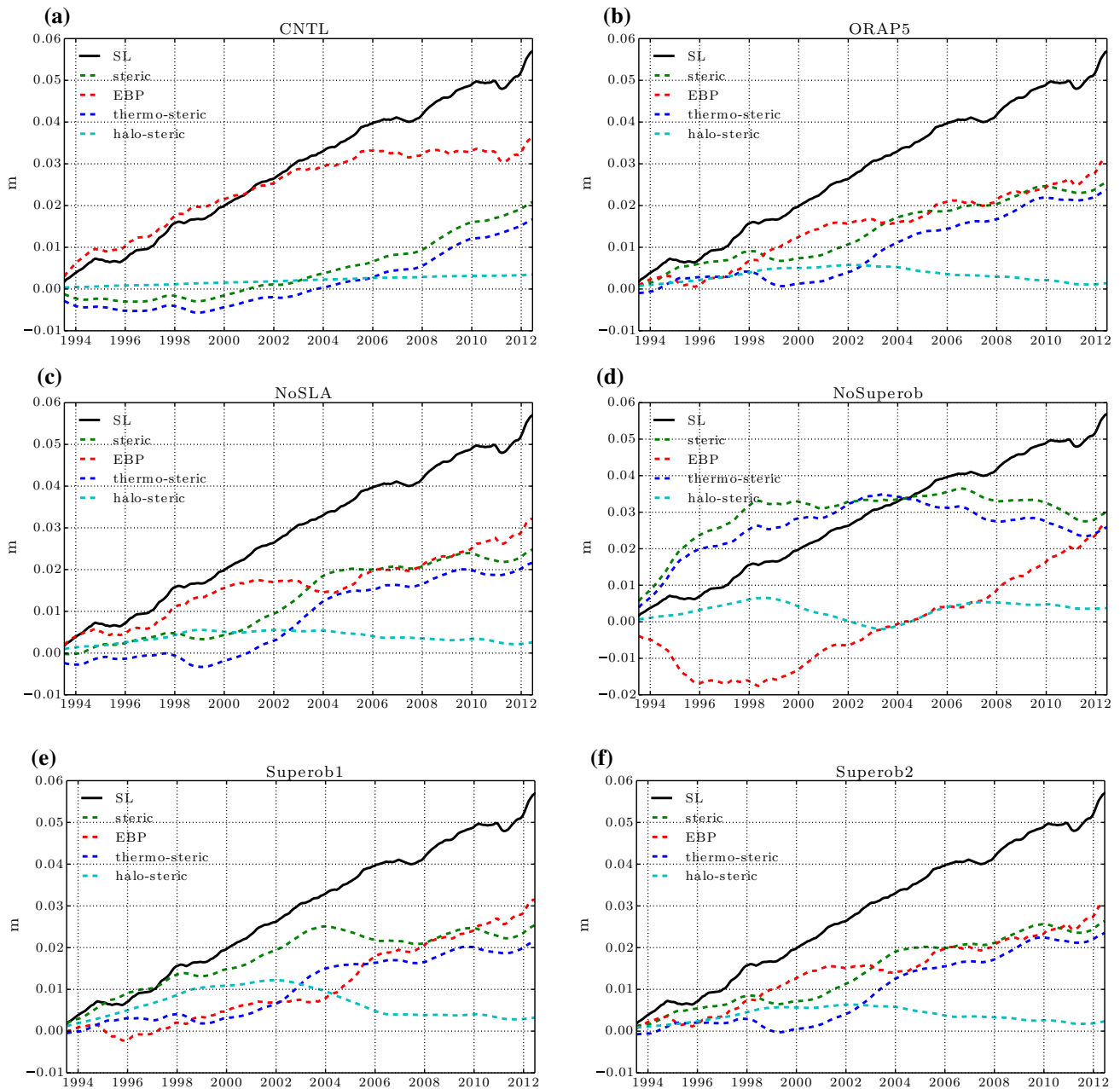
Fig. 10 Time series of global mean sea-level anomalies (m) (*solid line*) and its components (*dashed lines*) as steric changes (*green*), EBP changes (*red*), thermo-steric changes (*blue*) and halo-steric changes (*cyan*) for **a** CNTL, **b** ORAP5 and sensitivity experiments: **c** NoSLA, **d** NoSuperob, **e** Superob1 and **f** Superob2. Time series calculated using monthly mean fields with 12-month running mean and values from 1993 January removed

component is almost constant, although a slight increase can be appreciated. This is more likely related with spatial distribution of salinity rather than the amount of salt, since the integrated salinity remains fairly constant (not shown). Compared with CNTL, all the assimilation experiments show a larger positive contribution of the halosteric component, which is especially noticeable in the experiments with altimeter assimilation in the pre-Argo period. In all the assimilation experiments, the increase in halosteric component is due to a decrease in the integrated salinity of the ocean (not shown), indicating that the assimilation does not preserve salt. This can be a consequence of the multivariate scheme between temperature and salinity, which would make local modifications to the salinity profile when

Table 3 Linear trends (mm year^{-1}) of global mean sea-level changes from 1993 to 2012

Experiment name	Sea level	Steric height	EBP	Thermo-steric	Halo-steric
CNTL	2.8	1.2	1.6	1.1	0.1
NoSLA	2.8	1.5	1.3	1.5	0
NoSuperob	2.8	0.8	2.0	0.8	0
Superob1	2.8	1.1	1.7	1.2	−0.1
Superob2	2.8	1.4	1.4	1.4	0
ORAP5	2.8	1.3	1.5	1.3	0

Table 4 Linear trends (mm year^{-1}) of global mean steric sea-level changes from 1993 to 2002

Experiment name	Steric height	Thermo-steric	Halo-steric
CNTL	0.2	0.1	0.1
NoSLA	1	0.5	0.5
NoSuperob	2.7	2.8	−0.1
Superob1	1.8	0.6	1.2
Superob2	1	0.4	0.6
ORAP5	1	0.4	0.6

assimilating temperature (even in the absence of salinity observations). The advent of Argo appears to put an end to the increase of the halosteric component, which appears to slowly stabilize after 2004.

The contribution of the halosteric component appears sensitive to the horizontal correlation scale, with $\overline{L}_\eta = 4°$ in Superob1 exhibiting the largest halosteric variations, about 2/3 of the total steric contribution for the period 1993–2002 (Fig. 10e). The Superob1 halosteric component decreases

after 2002, probably due to assimilation of Argo data. Superob2 is equivalent to Superob1 but with BGE horizontal correlation length-scales for unbalanced SSH ($\overline{L}_\eta$ in Table 2) reduced from 4° to 2°. Superob2 reduces the linear trend in halo-steric term to 0.6 mm year^{-1} before 2002, and the sea level changes due to halo-steric term is consistent with those that derived from NoSLA experiment (cyan dashed line in Fig. 10c, f, respectively). ORAP5 is equivalent to Superob2 but with vertical correlation length-scales factor (α in Table 2) reduced from 2 to 1 for temperature and unbalanced salinity BGEs. This does not appear to influence the GMSL partition into steric and mass changes, nor the relative contributions of the salinity and temperature to the global volume increase, which show the same behaviour as Superob2 (see Fig. 10b, f, respectively). The linear trend of total steric height changes and its components for all sensitivity experiments over the period 1993–2002 are shown in Table 4.

4.5 Errors in the region of the Mediterranean outflow

An aspect critically affected by the vertical correlation scale α is the representation of the vertical penetration of the Mediterranean outflow waters. The mis-representation of the water mass from the Mediterranean outflow is a well known issue in ORAS4 (Balmaseda et al. 2013b) as well as in ORAP5 (Zuo et al. 2015a). Figure 11 shows the RMSE of temperature at 1000 m in the global ocean for Superob2 and ORAP5 after averaging between 2009 and 2012. Both experiments show large errors following the Mediterranean outflow, but the errors are slightly reduced in ORAP5 with respect to Superob2. This error grows very quickly between successive assimilation cycles, and it appears mainly after 2009. Additional sensitivity experiments at lower
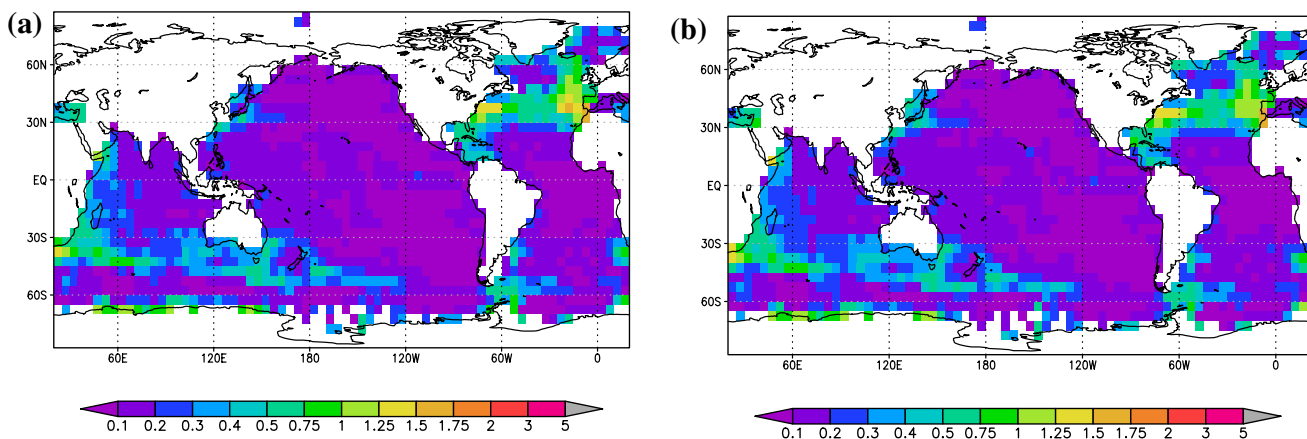

Fig. 11 Temperature fit to in-situ observation errors as measured by RMSE (°C) at 1000 m for **a** Superob2 and **b** ORAP5. RMSE are calculated using temperature analysis from the second outer loop after corrected by IAU against EN3 in-situ observations, and averaged over the period 2009–2012

resolution show that the error is not present if salinity is not assimilated. We speculate that the error is caused by spurious convection arising from the destabilization of the water column that may occurs when assimilating temperature and salinity using two separate control variables with independent errors (temperature and unbalanced salinity), a feature inherent in many assimilation methods. If this is the case, the vertical correlation scale may not be the ultimate reason for this error, but can be amplified by different values of this parameter. Another possibility for this error is the lack of flow dependent vertical correlation scales in ORAP5s implementation of 3D-Var. Note that recent development of 3D-Var in the UK Met Office includes a flow dependent vertical correlation scale that depends on the mixed layer depth (Waters et al. 2014).

5 Summary and discussion

The reduction of uncertainty level in the estimation of climate variability by ocean reanalyses requires continuous system improvements as well as the identification of uncertainty sources. These are the aims of the work presented here, which describes an upgraded ocean reanalysis system, evaluates its performance with respect to its predecessor, and investigates the sensitivity of some climate indices to selected parameters within this recently developed ocean reanalysis system.

ORAP5 is an eddy-permitting ocean reanalysis produced by ECMWF for the MyOcean2 project. ORAP5 is a high resolution (0.25°) global ocean reanalysis based on the NEMO ocean model and NEMOVAR data assimilation system, covering the period 1979–2012. Compared to the current operational ORAS4, ORAP5 increases the model resolution significantly and includes an interactive sea-ice model. A series of system upgrades in ORAP5 has also been discussed in detail in this paper.

ORAP5 performance has been compared with an equivalent non-assimilation experiment (CNTL) and ORAS4. The time evolution of the bias and RMSE against the quality controlled EN3 data set exhibits visible declining trends in the temperature/salinity RMSE in all three integrations, coinciding with the introduction of the Argo observing system. The reduction in RMSE is likely due to changes in the observational spatial coverage associated with the implementation of Argo, and it does not necessarily imply an improvement in the ocean state estimation. Both ORAP5 and ORAS4 show significant improvement over the CNTL due to data assimilation. The differences between ORAP5 and ORAS4 are more noticeable in salinity than in temperature, even though they both assimilate similar observations. ORAP5 shows smaller salinity errors in the northern extratropics relative to ORAS4, particularly over the Gulf

Stream region. ORAP5 shows slightly increased errors in the southern extratropics in both temperature and salinity, probably because the existing observations are unable to constrain the higher variability resolved by ORAP5, at least using our current assimilation method.

The temporal correlation with the CCI1 gridded maps of altimeter-derived SLA has been used to assess the coherence of the interannual variability in the different estimates. ORAP5 shows higher correlations with altimeter data than the CNTL experiment, and very similar to ORAS4, with slight improvements in the extratropics. These findings are supported by the BADOMAR tide gauge data.

Sensitivity experiments focussed on the impact of altimeter data on climate indices have been conducted within the framework of ORAP5. Along track altimeter sea-level data has quite uniform spatial and temporal coverage, especially when compared with that of in-situ data, and provides a unique data set to constrain large and small scales. However, extracting information from the altimeter sea-level is not straight forward. The projection of the sea level information onto the vertical temperature and salinity structure relies heavily on the goodness of the background model profiles. Without enough in-situ observations to constrain the background subsurface field, the assimilation of altimeter data can damage the solution. Results show indeed that while the assimilation of altimeter data increases the fit to altimeter observations (Fig. 8), it does not always translate into observable improvements to the ocean subsurface fields.

In ORAP5 the assimilation of altimeter data improves slightly the fit to subsurface observations in the tropical ocean, but only after careful specification of background and observation error covariance parameters, and careful treatment of altimeter data. Of particular importance was the superobbing scheme, without which the fit to the subsurface temperature and salinity observations was degraded. The superobbing and the horizontal correlation scales also affected the estimation of relevant climate indices, such as the partition of GMSL variations into steric and mass changes. Without altimeter superobbing, this partition exhibited unphysical behaviour before the Argo period, with steric height tendency rates exceeding those of GMSL. Whether the problem is due to overfitting to observation or lies somewhere else (like inadequate multivariate relationships) needs further investigation.

The different sensitivity experiments show that the trend in global steric height is dominated by the thermosteric component. All the assimilation experiments, before the Argo period, show some interannual variability in the halosteric component that can be considered spurious. This variability is also affected by specific assimilation parameters, such as the horizontal correlation scale of the barotropic component of sea level.

As with ORAS4, ORAP5 also exhibits large errors in the area following the Mediterranean outflow waters. Errors in this area are sensitive to the vertical correlation scale, and are absent when the salinity observations are not assimilated. Whether this is due to bad observations or rejection of useful observations by the QC, or whether it is an inherent problem with the assimilation scheme, is currently being investigated.

ORAP5 is the basis for the next eddy permitting operational reanalysis ORAS5. Some of the aspects highlighted here will be revised. In particular the adequacy of superobbing versus thinning scheme in the assimilation of altimeter data needs to be evaluated. The quality control procedure is also being revised. The outcome of the results presented in this paper is that the specification of the background and observation error covariance is crucial for the good performance of the reanalysis. We would like to include more flow-dependent aspects in the error correlation scale by, for instance, using ensemble methods. It is expected that ORAS5 will improve on some of the deficiencies of ORAP5, after the lessons learnt with the evaluation presented here.

References

Ablain M, Cazenave A, Larnicol G, Balmaseda MA, Cipollini P, Faugère Y, Fernandes M, Henry O, Johannessen J, Knudsen P et al (2015) Improved sea level record over the satellite altimetry era (1993–2010) from the Climate Change Initiative Project. Ocean Sci Discuss 11(4):2029–2071

Amante C, Eakins BW (2009) ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. US Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Geophysical Data Center, Marine Geology and Geophysics Division

Antonov J, Seidov D, Boyer T, Locarnini R, Mishonov A, Garcia H, Baranova O, Zweng M, Johnson D (2010) World Ocean Atlas 2009, volume 2: salinity. In: Levitus S (ed) NOAA Atlas NESDIS 69, p 184

Balmaseda MA, Dee D, Vidard A, Anderson D (2007) A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. Q J R Meteorol Soc 133(622):167–179

Balmaseda MA, Fujii Y, Alves O, Lee T, Rienecker M, Rosati T, Stammer D, Xue Y, Freeland H, McPhaden MJ et al (2010) Role of the ocean observing system in an end-to-end seasonal forecasting system. In: Proceedings of oceanobs 9

Balmaseda MA, Mogensen K, Weaver AT (2013a) Evaluation of the ECMWF ocean reanalysis system ORAS4. Q J R Meteorol Soc 139(674):1132–1161

Balmaseda MA, Trenberth KE, Källén E (2013b) Distinctive climate signals in reanalysis of global ocean heat content. Geophy Res Lett 40(9):1754–1759

Balmaseda MA, Hernandez F, Storto A, Palmer MD, Alves O, Shi L, Smith GC, Toyoda T, Valdivieso M, Barnier B et al (2015) The Ocean Reanalyses Intercomparison Project (ORA-IP). In: Proceedings of the Institute of Marine Engineering, Science, and Technology, Journal of Operational Oceanography

Barnier B, Madec G, Penduff T, Molines JM, Treguier AM, Le Sommer J, Beckmann A, Biastoch A, Böning C, Dengg J et al (2006) Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. Ocean Dyn 56(5–6):543–567

Bloom S, Takacs L, Da Silva A, Ledvina D (1996) Data assimilation using incremental analysis updates. Mon Weather Rev 124(6):1256–1271

Breivik Ø, Mogensen K, Bidlot JR, Balmaseda MA, Janssen PA (2015) Surface wave effects in the NEMO ocean model: forced and coupled experiments. J Geophys Res

Cazenave A, Larnicol G, Legeais JF, Ablain M, Faugre Y, Mbajon Njiche S, Timms G, Meyssignac B, Balmaseda M, Zuo H, Johannessen J, Scharffenberg M, Stammer D, Andersen O, Knudsen P, Zawadzki L, Thibaut P, Poisson JC, Picard B, Carre L, Mertz F, Lauret O, Rudenko S, Farquhar C, Pechorro E, Roca M, Nilo P, Cipollini P, Calafat F, Fernandes J, Lazaro C, Quartly G, Kurekin A, Nencioli F, Fenoglio-Marc L, Benveniste J, Lucas B, Dinardo S, Ambrzio A (2014) Sea level essential climate variable products, Version 1.1. ESA Sea Level Climate Change Initiative (ESA SL_cci). doi:10.5270/esa-sea_level_cci-1993_2013-v_1.1-201412

Chen X, Tung KK (2014) Varying planetary heat sink led to global-warming slowdown and acceleration. Science 345(6199):897–903

Chevallier M et al (2015) Uncertainties in the Arctic sea ice cover in state-of-the-art ocean reanalyses from the ORA-IP project. Clim Dyn S.I Ocean Reanal

Church JA, White NJ (2011) Sea-level rise from the late 19th to the early 21st century. Surv Geophys 32(4–5):585–602

Dai A, Trenberth KE (2002) Estimates of freshwater discharge from continents: latitudinal and seasonal variations. J Hydrometeorol 3(6):660–687

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597. doi:10.1002/qj.828

Donlon CJ, Martin M, Stark J, Roberts-Jones J, Fiedler E, Wimmer W (2012) The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sens Environ 116:140–158

Drijfhout S, Blaker A, Josey S, Nurser A, Sinha B, Balmaseda M (2014) Surface warming hiatus caused by increased heat uptake across multiple ocean basins. Geophys Res Lett 41(22):7868–7874

ECMWF (2013) IFS documentation CY40R1, part VII: ECMWF Wave Model. ECMWF Tech Memo (79)

England MH, McGregor S, Spence P, Meehl GA, Timmermann A, Cai W, Gupta AS, McPhaden MJ, Purich A, Santoso A (2014) Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. Nat Clim change 4(3):222–227

Fichefet T, Maqueda M (1997) Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. J Geophys Res Oceans (1978–2012) 102(C6):12,609–12,646

Guemas V, Doblas-Reyes FJ, Lienert F, Soufflet Y, Du H (2012) Identifying the causes of the poor decadal climate prediction skill over the north pacific. J Geophys Res Atmos 117(D20)

Gürol S, Weaver A, Moore A, Piacentini A, Arango H, Gratton S (2014) B-preconditioned minimization algorithms for variational data assimilation with the dual formulation. Q J R Meteorol Soc 140(679):539–556

Hay CC, Morrow E, Kopp RE, Mitrovica JX (2015) Probabilistic reanalysis of twentieth-century sea-level rise. Nature

Ingleby B, Huddleston M (2007) Quality control of ocean temperature and salinity profiles-historical and real-time data. J Mar Syst 65(1):158–175

Janssen P, Brevivik O, Mogensen K, Vitart F, Balmaseda MA, Bidlot JR, Keeley S, Leutecher M, Magnusson L, Molteni F (2013) Air-sea interaction and surface waves. ECMWF Tech Memo (712)

Karspeck A et al (2015) Comparison of the Atlantic meridional overturning circulation between 1960 and 2007 in six different ocean reanalysis products. Clim Dyn S.I Ocean Reanal

Large W, Yeager S (2009) The global climatology of an interannually varying air-sea flux data set. Clim Dyn 33(2–3):341–364

Lefèvre F, Doradeu J, Sénant E et al (2005) Monitoring of the ENVISAT altimeter drift with a specific tide gauge database. In: Workshop on sea level variations towards an operational European sea level service p 64

Locarnini R, Mishonov A, Antonov J, Boyer T, Garcia H, Baranova O, Zweng M, Johnson D (2010) World Ocean Atlas 2009, volume 1: temperature. In: Levitus S (ed) NOAA Atlas NESDIS 68, p 184

Madec G (2008) NEMO ocean engine. Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27. ISSN No 1288-1619

Mayer M, Haimberger L, Balmaseda MA (2014) On the energy exchange between tropical ocean basins related to ENSO. J Clim

Mogensen K, Balmaseda MA, Weaver A (2012) The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for system 4. ECMWF Tech Memo (668)

Palmer M et al (2015) Ocean heat content variability and change in an ensemble of ocean reanalyses. Clim Dyn S.I Ocean Reanal

Pohlmann H, Müller WA, Kulkarni K, Kameswarrao M, Matei D, Vamborg FSE, Kadow C, Illing S, Marotzke J (2013) Improved forecast skill in the tropics in the new miklip decadal climate predictions. Geophys Res Lett 40(21):5798–5802

Reynolds RW, Smith TM, Liu C, Chelton DB, Casey KS, Schlax MG (2007) Daily high-resolution-blended analyses for sea surface temperature. J Clim 20(22):5473–5496

Roberts-Jones J, Fiedler EK, Martin MJ (2012) Daily, global, high-resolution SST and sea ice reanalysis for 1985–2007 using the OSTIA system. J Clim 25(18):6215–6232

Shi L et al (2015) Assessment of upper ocean salinity content from the Ocean Reanalyses Inter-Comparison Project (ORA-IP). Clim Dyn S.I Ocean Reanal

Simmons A, Uppala S, Dee D, Kobayashi S (2007) ERA-interim: New ECMWF reanalysis products from 1989 onwards. ECMWF Newslett 110

Tapley BD, Bettadpur S, Ries JC, Thompson PF, Watkins MM (2004) GRACE measurements of mass variability in the earth system. Science 305(5683):503–505

Tietsche S, Balmaseda MA, Zuo H, Mogensen K (2014) Arctic sea-ice in the ECMWF MyOcean2 ocean reanalysis ORAP5. ECMWF Tech Memo (737)

Toyoda T et al (2015a) Interannual-decadal variabilities of wintertime mixed layer depths in the North Pacific observed in an ensemble of ocean syntheses. Clim Dyn S.I Ocean Reanal

Toyoda T et al (2015b) Intercomparison and validation of the mixed layer depth fields of global ocean syntheses/reanalyses. Clim Dyn S.I Ocean Reanal

Uppala SM, Kållberg P, Simmons A, Andrae U, Bechtold V, Fiorino M, Gibson J, Haseler J, Hernandez A, Kelly G et al (2005) The ERA-40 reanalysis. Q J R Meteorol Soc 131(612):2961–3012

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 141(687):333–349. doi:10.1002/qj.2388

Weaver AT, Deltel C, Machu É, Ricci S, Daget N (2005) A multivariate balance operator for variational ocean data assimilation. Q J R Meteorol Soc 131(613):3605–3625

Wijffels SE, Willis J, Dominguez CM, Barker P, White NJ, Gronell A, Ridgway K, Church JA (2008) Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. J Clim 21(21):5657–5672

Zhu J, Huang B, Marx L, Kinter JL, Balmaseda MA, Zhang RH, Hu ZZ (2012) Ensemble ENSO hindcasts initialized from multiple ocean analyses. Geophys Res Lett 39(9)

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 41(9–10):2785–2795

Zuo H, Balmaseda MA, Haines K, Valdivieso MSM, Storto A, Parent L (2015a) The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice reanalysis ORAP5. Part II: comparison with other MyOcean ocean reanalyses. ECMWF Tech Memo (in preparation)

Zuo H, Balmaseda MA, Mogensen K (2015b) The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice reanalysis ORAP5. Part I: Implementation. ECMWF Tech Memo (736)