TECHNICAL MEMORANDUM

736

The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice reanalysis ORAP5. Part 1: Implementation

Hao Zuo, Magdalena A. Balmaseda, Kristian Mogensen

Research Department

February 2015

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

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

Abstract

A new eddy permitting global ocean and sea-ice reanalysis has been developed in the ECMWF as a contribution to the EU FP7 MyOcean2 project. It is called ORAP5 (Ocean ReAnalysis Pilot 5), and it spans the period 1979 to 2012. Data from ORAP5 became publicly available from MyOcean portal from August 2014. ORAP5 is based on NEMO ocean model and NEMOVAR data assimilation system. A series of system developments have been carried out to upgrade the current low resolution ORAS4 to high resolution ORAP5, including new NEMO version and configuration, revised surface fluxes scheme, new version of satellite sea surface height data and treatment, revised NEMOVAR horizontal correlation length-scale scheme, coupling LIM2 sea-ice model with NEMO ocean model, and introduction of sea-ice concentration assimilation in NEMOVAR, among others. This technical memorandum describes the new components of ORAP5 in detail, and offers a preliminary evaluation of its performance as well as results from a range of sensitivity experiments.

1 Introduction

Ocean reanalyses 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. At ECMWF, ocean reanalyses are primarily used for initialization of the coupled forecasting systems (medium range, extended range and seasonal). Currently, the operational ocean reanalysis is ORAS4 (Ocean Reanalyses System 4, [Balmaseda et al., 2013a]), which was implemented operationally in 2010 and includes 5 members. Since then, ORAS4 has also been used extensively by the climate community to initialize seasonal and decadal forecasts (Pohlmann et al. [2013],Guemas et al. [2012]) and for exploring climate signals (Balmaseda et al. [2013b], Mayer et al. [2014], England et al. [2014], Chen and Tung [2014], Drijfhout et al. [2014], among others). However, ORAS4 has two important limitations for its use by future generation of coupled models: it has low horizontal resolution (about 1°), and it does not have an interactive sea-ice model. Here we present the first ECMWF implementation of a higher resolution ocean and sea-ice reanalyses. It is called ORAP5 (Ocean Reanalyses Prototype 5 (ORAP5), which has been produced as a contribution to the multi-ocean-reanalyses program within the EU FP7 MyOcean-2 project.

The MyOcean global ocean reanalysis activities include efforts from different European institutions. ECMWF joined two years ago, when MyOcean2 started. There are five different eddy-permitting (0.25°) ocean reanalyses, as well as one reference simulation (no assimilation). Summary of these products can be found in Table 1. They normally cover the recent period during which satellite altimeter data are available (from 1993 onwards) and represent state of the art on ocean reanalysis. ORAP5 is the ECMWF contribution to MyOcean, and covers the period 1979-2012. Monthly means of three-dimensional fields and daily two-dimensional fields from ORAP5 have been delivered and are publicly available in the MyOcean portal
http://www.myocean.eu/web/69-myocean-interactive-catalogue.php?option=com_csw&view=details&product_id=GLOBAL_REANALYSIS_PHYS_001_017

This technical memorandum presents the key technical characteristics of ORAP5 and provides a preliminary assessment of its quality. The overview of the configuration and setup of ORAP5 is described in Section 2.1, which also specifies the main upgrades between ORAS4 and ORAP5. Details of the ocean and sea-ice model are given in Section 2.2, while the data assimilation scheme and observations are described in Sections 2.3 and 2.4 respectively. Assimilation statistics for ORAP5, as well as an evaluation using independent observations are presented in Section 3. Section 4 shows a number of sensitivity experiments carried out for specifications of parameters in ORAP5. The comparison of ORAP5 with other

MyOcean reanalyses will be presented in a follow-up report. The ORAP5 sea-ice has been thoroughly evaluated by Tietsche et al. [2014], and it will not be shown here.

Table 1: Overview of MyOcean global physical ocean reanalyses

product name	product centre	data range	data type
GLORYS2V3	Mercator Ocean	1993-2011	reanalysis
CGLORS	CMCC	1993-2011	reanalysis
UR025.4	Uni. of Reading	1989-2010	reanalysis
ORAP5	ECMWF	1979-2012	reanalysis
MJM105B	LGGE	1993-2011	reference simulation

2 ORAP5 system configuration

2.1 Overview

The ECMWF-MyOcean2 high resolution global ocean reanalysis ORAP5 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 an active sea-ice model (LIM2, Fichefet and Maqueda 1997). The reanalysis is conducted with NEMOVAR [Mogensen et al., 2012] in its 3D-Var FGAT configuration. NEMOVAR is used to assimilate subsurface temperature, salinity, sea-ice concentration and sea-level anomalies, using a 5 day assimilation window with 1200 seconds model step. The observational information is also used via an adaptive bias correction scheme [Balmaseda et al., 2013a]. 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. 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].

In addition to resolution and sea-ice model, ORAP5 also contains several system upgrades compared with ORAS4. The main differences between ORAS4 and ORAP5 system settings are summarized in Table 2. Details of most system upgrades are discussed in the following sections.

Selected variables from this high resolution reanalysis are available through MyOcean2 portal, including, monthly means of Temperature, Salinity, Currents, Sea Surface Height, Sea Surface Temperature and sea-ice parameters, 5-day mean surface currents, and also daily values of derived 2-dimensional fields (mixed layer depth, depth of 20/26/28 $^{\circ}C$ isotherms, heat content in upper 300m/700m/total column, steric height, and bottom pressure among others).

Table 2: Overview of differences between ORAS4 and ORAP5 ocean reanalysis settings

		ORAS4	ORAP5
grids		~1°, 42 vertical levels	~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
assimilation	Background Error	constant meridional length-scale	Rossby-radius dependent meridional length-scale
	sea-ice	no assimilation	sea-ice concentration assimilation
	prior-1993 global freshwater 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

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

ORCA is a common NEMO ocean model global configuration that includes tri-polar grid with the three poles located over Antarctic, Central Asia and North Canada. The DRAKKAR consortium focuses on developing different resolution ocean configurations of NEMO for their used in European Earth System models. The 2012 reference version of DRAKKAR high resolution ORCA configurations (ORCA025.L75, see Barnier et al. 2006) has been used together with the NEMO ocean model (version 3.4.1). Input files from DRAKKAR configurations can be found in `http://servdap.legi.grenoble-inp.fr/meom/ORCA025.L75-REF-VERSION/`. The ORCA025.L75 configuration is a grid with 0.25° resolution at the equator, increasing to up to 12 *km* in some areas of the Arctic

Ocean. There are 75 vertical levels with a resolution varying 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. The vertical diffusion coefficient is determined using the Turbulent Kinetic Energy (TKE) scheme. Solar penetration in the ocean is calculated using a 2 bands scheme with two depths of extinction and unified Chlorophyll concentration (0.05 $mg\ m^{-3}$). The LIM2 sea-ice model is coupled with NEMO in every 3 model steps and uses a the Visco-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 the input of TKE from breaking waves is represented in the standard NEMO by a constant parameter. In ORAP5, this 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 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 ECMWF WAM model. The transfer coefficient for momentum is defined directly from the wave model drag coefficient, and it is used to derive transfer coefficients for sensible/latent heat and evaporation computation via the CORE bulk formula [Large and Yeager, 2009]. 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].

Surface forcing fields are extracted from ECMWF ERA-Interim atmospheric reanalysis product [Simmons et al., 2007, Dee et al., 2011], including:

- 6-hourly wind velocity (u/v), temperature and specific humidity at 10 m above mean sea-level
- 24-hourly radiation fluxes (downward short-wave and long-wave radiation), total and snow precipitation
- Additional wave fields required (available from ERA-Interim): Significant wave height, Mean wave period, Surface Stokes drift velocity, Energy flux, Water-side stress, Neutral 10-m wind speed, Wave drag coefficient

The initial conditions for the ORAP5 were produced in two phases:

- A 12-year (1979-1990) model spin up from the cold start given by the climatology from World Ocean Atlas 2009 (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.
- A 5-year assimilation period (1975-1979), starting from the end of the previous spin up conditions.

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 $W\ m^{-2}\ K^{-1}$. Different SST datasets are used for SST damping:

- ERA-40 reanalysis SST (Uppala et al. 2005) from 19790101 to 19810831.

- U.K Met Office Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) SST [Donlon et al., 2012] from 19850101 to 20071231 (reanalysis) and from 20090101 to 20121231 [Roberts-Jones et al., 2012] (operational).
- NOAA Optimal Interpolation 0.25° daily SST analysis (OIv2d, Reynolds et al. 2007) whenever the OSTIA dataset is not available (19810901-19841231 and 2008).

The monthly climatological river runoff [Dai and Trenberth, 2002] also contributes to the fresh water flux, and it is applied at the location of the river mouths. A sea surface salinity relaxation to WOA09 monthly climatology is applied to constrain salinity, with the strength set to -33.3 $mm\ day^{-1}$, also acts as a fresh water flux term. Finally, the global fresh water 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]). More details of this adjustment are given in Section 2.3). In addition, a very weak (with a time-scale of about 20 years) global 3D relaxation to temperature and salinity climatological value from WOA09 is also applied through the water column.

2.3 Data assimilation scheme

NEMOVAR is a variational data assimilation system developed based on OPAVAR (OPA VARiational 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 the First-Guess at Appropriate Time (FGAT) approach. The analysis cycle consists of a single iteration of 3D-Var FGAT with observational Quality Control (QC) and bias correction steps. This iteration constains three disctinctive phases. In the first phase (or first outer loop) the NEMO model is integrated forward and used for calculation of the model equivalent of each available observation at the time step closest to the observation time, after which the QC of the observations is performed. The background state and the quality-controlled observations are passed to the inner loop part of 3D-Var FGAT where the incremental cost function is minimized using an observation space conjugate gradient (RPCG) method [Gürol et al., 2014] with 40 RPCG iterations to produce the assimilation increment. 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 spanning the same time window as for the assimilation window. Temperature and salinity profiles, sea-level anomalies and sea-ice concentration observations are assimilated using 5-day assimilation cycle in ORAP5 and share the outer loop model integrations.

Assimilation of SIC data from OSTIA gridded product (see Section 2.4) is also included in ORAP5. The background state of ocean and sea ice states are produced from coupled NEMO-LIM2 run, but the minimization of the sea ice cost-function is separated from all ocean state variables in a different loop. The sea-ice minimization can be separated from the other variables since it is assumed that there is no cross covariance between sea-ice and other variables. Variables which are physically related are divided into *balanced* and *unbalanced* components. The *balanced* components are linearly dependent (related by the multi-variate relationships), while the unbalanced components are independent and uncorrelated with other variables. The ORAP5 balance relations are the same as for ORAS4 where the details can be found in Mogensen et al. [2012].

The specification of the background error (BGE) consists of multivariate relationships and the values of the error for the unbalanced components. In ORAP5, the multivariate relationships are the same as ORAS4, but with some modifications to the beta-plane geostrophic balance. As in ORAS4, the

Table 3: Summary of the values used in ORAP5 for calculation of background-error standard deviations

	temperature	**Unbalanced salinity**	**Unbalanced SSH**	**SIC**
ORAP5	$\sigma_T^{max} = 1.5\ ^\circ C$, $\sigma_T^{ml} = 0.5\ ^\circ C$, $\sigma_T^{do} = 0.07\ ^\circ C$, $\delta_z = 10\ m$	$\sigma_{Su}^{max} = 0.25\ PSU$, $\sigma_{Su}^{do} = 0.01\ PSU$	$\phi_{ex} = \pm 20^\circ$, $\sigma_{\eta u}^{ex} = 0.01\ m$, $\sigma_{\eta u}^{eq}$=0	$\overline{\sigma}_{ice}^{b} = 0.05$

σ_X^{max}: *maximum allowed value for BGE standard deviation*
σ_X^{ml}: *lower bounds in the mixed layer for BGE standard deviation*
σ_X^{do}: *minimum value for BGE standard deviation in the deep ocean*
X can be Temperature (T) or Unbalanced Salinity (Su)
δ_z: *vertical displacement error for temperature*
ϕ_{ex}: *latitude beyond which unbalanced SSH BGE standard deviation* = $\sigma_{\eta u}^{ex}$
$\sigma_{\eta u}^{eq}$: *unbalanced SSH BGE standard deviation directly at equator*
$\overline{\sigma}_{ice}^{b}$: *constant BGE standard deviation for SIC*

BGE standard deviations for both temperature and unbalanced salinity in ORAP5 contain some flow-dependent aspects. The standard deviation of the temperature BGE is parameterized in terms of the vertical gradient of the background temperature fields so flow-dependent aspects could be captured. A latitude dependent function is used for construction of BGE standard deviation for unbalanced Sea Surface Height (SSH), to account for the importance of this barotropic component in extra-tropical regions. A constant BGE standard deviation (0.05) for SIC is used everywhere. Parameter specification for calculation of BGE standard deviations used in ORAP5 are summarized in Table 3. These are very similar to those used in ORAS4 (Table 2 in Mogensen et al. [2012]). Arguably, BGE in ORAP5 and ORAS4 should be different, since the first guess has been produced with different model version, resolution, forcing fields and assimilation window. Sensitivity experiments to some of these BGE parameters have been carried out. In particular, multi-year reanalyses were conducted with modified values of vertical displacement error (δ_z), and minimum BGE standard deviation in the deep ocean (σ_T^{do}). Results showed best performance with the parameter values in Table 3.

A revised scheme for calculating background error horizontal correlation length-scales has been implemented. This is a variant of the scheme developed by Waters et al. [2014], by which the horizontal background correlation length-scales are set-up by the Rossby radius of deformation to represent the mesoscale processes. Outside of the Equator, the first baroclinic Rossby radius of deformation (R) at latitude ϕ can be determined as

$$R = \frac{c}{|f(\phi)|} \tag{1}$$

where $f(\phi)$ is the Coriolis parameter and c is the gravity wave phase speed. In Waters et al. [2014] c is spatially varying, with values derived from a climatology of the ocean density field. In ORAP5 only the latitudinal variations of the Coriolis parameter are included in the Rossby radius of deformation, while a constant climatological mean gravity wave speed $c = 2.7\ m\ s^{-1}$ [Chelton et al., 1998] has been used, in a compromise between computational efficiency and complexity. The horizontal length scales of BGE

for any given ocean state variable X are then parameterized as a function of R. Thus, the meridional length-scale (L_X^{ϕ}) is parameterized as follows:

$$L_X^{\phi} = MAX\Big(MIN\Big(R,\ R_{max} \Big),\ R_{min} \Big) \tag{2}$$

where the value of R is capped to a minimum value of 50 km (R_{min}) and maximum value of 150/250 km (R_{max}) near the equator for the meridional/zonal length-scales respectively, to take into account the variations in grid size and the equatorial beta-plane approximation. The zonal correlation length scales L_X^{λ} are elongated near the equator, to represent the distanced travelled by a Kelvin wave during the duration of the assimilation window:

$$L_X^{\lambda} = C(\phi;\delta) L_X^{\phi} \tag{3}$$

where $C(\phi;\delta)$ is the horizontal stretching factor as defined in Section 4.6.2 in Mogensen et al. [2012]. Fig. 1 shows the BGE zonal correlation length-scales used for temperature in ORAP5: the capped Rossby radius of deformation and zonal stretching factor are shown in the top panel, and the corresponding zonal length-scale appears in the bottom panel.

In ORAP5, the above scheme was used to prescribe temperature, unbalanced salinity and SIC correlation length-scales. For the unbalanced SSH BGE, a constant correlation length scale ($\overline{L}_{\eta}$) of 2° is used everywhere to represent the barotropic background errors. In addition the correlation length-scale for all variables are reduced linearly as a function of distance to the coast. This linear reduction scheme starts at 100 km from distance from the coast (except for SSH, which is 600 km), and reduces all correlation length-scales to a minimum value at the coast, which equals the maximum grid-size (27.8 km for ORAP5). Sensitivity experiments have been carried out to test the impact this revised horizontal correlation length-scale scheme, as well as different combination of parameter values. This revised scheme reduced analysis SST bias in the equatorial regions and improved the SSH correlation with observations in the subtropical Pacific Ocean (not shown).

The same vertical correlation length-scale scheme is applied in ORAP5 as in ORAS4, which is specified as a scalar ($\alpha = 1$) multiple of the local vertical grid-size dz (see Section 4.6.2 in Mogensen et al. [2012]). Larger values of α have also been tested ($\alpha = 2$) but the fit to the observations was degraded, especially by the Mediterranean Ocean outflow (See Section 4). A summary of the parameters used in ORAP5 for calculation of horizontal and vertical BGE correlation length-scales can be found in Table 4. This choice of the horizontal and vertical BGE specifications implies that the volume of ocean potentially affected by an in-situ observation is smaller in ORAP5 than in ORAS4.

Altimeter-derived Global Mean Sea Level (GMSL) variations are also assimilated in ORAP5, following the same scheme as that in ORAS4, and described in [Balmaseda et al., 2013a]. The spatial mean of the sea-level background field and of the input sea-level observations is removed before assimilation, so that the residual can then be used to close the fresh-water budget and thus helping with the attribution of sea-level rise. This scheme uses the fact that the GMSL variations can be decomposed into

$$\Delta\overline{\eta}_o = \Delta\overline{\eta}_s + \Delta\overline{\eta}_m \tag{4}$$

where $\Delta\overline{\eta}_o$ is the observed GMSL change between two given times; $\Delta\overline{\eta}_s$ is the steric component of GMSL change, which is derived from model density fields; $\Delta\overline{\eta}_m$ is the GMSL change due to mass variation. During the altimeter-era, $\Delta\overline{\eta}_o$ can be estimated from the altimeter observations, and the mass

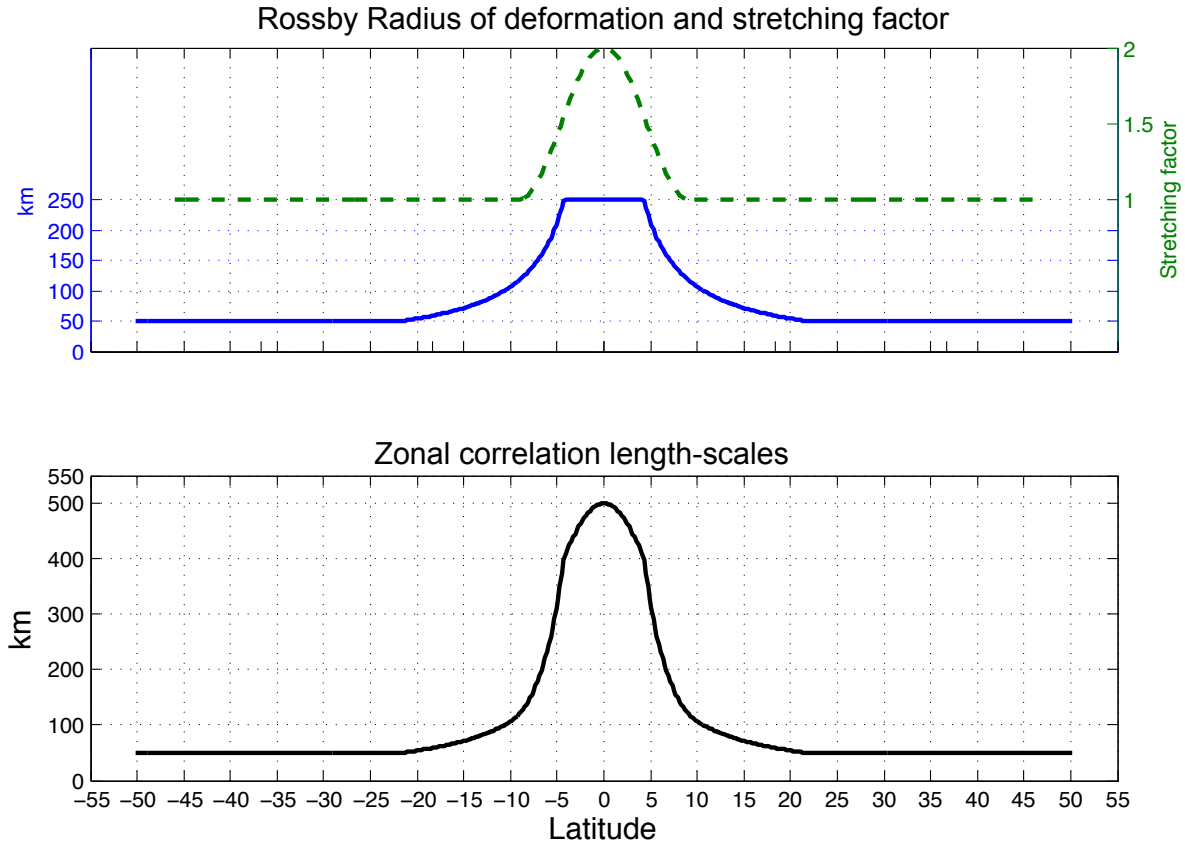


Figure 1: (top): Capped Rossby radius of deformation for temperature zonal length scales (blue solid line) in km as defined in Eq. 2 and zonal stretching factor ($C(\phi;\delta_\lambda)$, green dashed line); (bottom): zonal BGE correlation length-scales (L^λ_X,$X = T$, black) in km as defined in Eq. 3. Please refer to Table 4 for parameter values used for calculation in ORAP5.

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 not information about the GMSL, and additional assumptions are needed. 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 $\Delta\overline{\eta}_m$ are well approximated by the climatology, which is estimated from the GRACE-derived bottom-pressure data for the period 2005-2009. Fig. 2 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 section 2.4).

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 a pressure correction in the tropical regions. The total bias 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. Following the same algorithm as implemented in ORAS4, the online bias

Table 4: Summary of the parameters from ORAP5 regarding to the background-error correlation length-scale calculation

		temperature	Unbalanced salinity	Unbalanced SSH	SIC
ORAP5	L_X^{λ}	$\delta_\lambda = 2$, R_{max}=250 km, R_{min}=50 km, ϕ_L =±10°	same as for T	$\overline{L}_\eta = 2°$	same as for T
	L_X^{ϕ}	R_{max}=150 km, R_{min}=50 km	same as for T	$\overline{L}_\eta = 2°$	same as for T
	L_X^{Z}	$\alpha = 1$	same as for T	N/A	N/A

L_X^{λ} *is zonal BGE correlation length-scale*
L_X^{ϕ} *is meridional BGE correlation length-scale*
L_X^{Z} *is vertical BGE correlation length-scale*
X can be Temperature (T), Unbalanced Salinity (Su) or SIC
δ_λ *is zonal stretching factor at the equator*
ϕ_L *is the latitude band within which horizontal BGE correlation length-scales are modified by* $C(\phi;\delta)$, *see Eq 3*

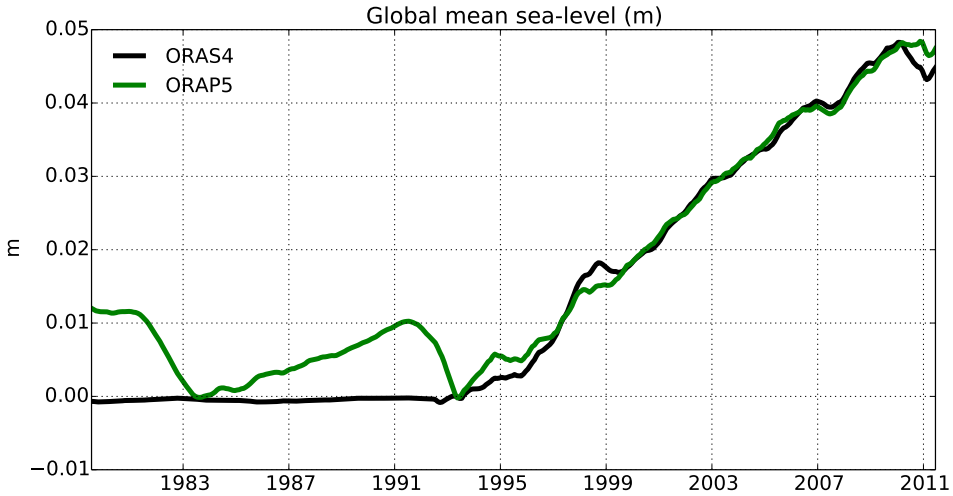


Figure 2: Global Mean Sea Level anomaly from 1979 to 2011 in ORAS4 and in ORAP5, with 12-month running mean and value from 1993 Jan removed. The differences before 1993 are due to the different constraint used (climatological mass in ORAP5 and climatological sea level in ORAS4). The differences after 1993 are due to the different altimeter versions.

correction term on cycle c is determined as

$$b'_c = \alpha_m b'_{c-1} - A\delta x^a_{c-1} \tag{5}$$

where α_m is a memory factor that determines the time evolution of the online-estimated bias term, δx^a

is the state vector increment and A is a linear transformation matrix that includes coefficients which are applied to different bias vector components (temperature, salinity and pressure). These coefficients are expressed as a function of latitude (ϕ):

$$a^{tr,T} = a^{tr,S} = a_1 e^{-(\phi/\phi_c)^2} + a_2(1 - e^{-(\phi/\phi_c)^2}) \tag{6}$$

$$a^{p,T} = a^{p,S} = a_3 e^{-(\phi/\phi_c)^2} \tag{7}$$

The coefficients $a^{tr,T}$ and $a^{tr,S}$ determine the proportion of the assimilation increment affecting the online bias corrections applied directly in the prognostic equations for temperature and salinity, and $a^{p,T}$ and $a^{p,S}$ are coefficients that determine the online bias correction terms acting on the momentum equations. Fig. 3 shows the latitude dependent reduction coefficients as defined in Eq. 6 and 7, combined with parameters taken from Table 5 for ORAS4 and ORAP5, respectively. In one assimilation cycle, only 0.1% of online-estimated bias correction term is applied in the equator in the tracers prognostic equation. This value increases to 0.3% (dashed blue line) in the extra-tropics for ORAS4, and 1% (solid blue line) for ORAP5, respectively. These values ensure that at low latitude the dominant bias term is pressure correction (green solid line). In ORAP5, the bias correction acting on temperature and salinity in the extratropics is higher than in ORAS4.

Fig. 4 shows the annual mean of 300-700 m averaged temperature (upper panels) and salinity (lower panels) offline bias correction applied in ORAS4 and ORAP5, which is added directly to temperature and salinity values in the tracers prognostic equation. In general, the temperature and salinity bias correction patterns are similar between ORAS4 and ORAP5, suggesting common model/forcing errors. The largest 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 in the former. 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.

Table 5: Summary of the parameters for online bias estimation used in ORAS4 and ORAP5

	α_m	a_1	a_2	a_3	ϕ_c
ORAS4	5 years	0.001	0.003	0.015	10°
ORAP5	5 years	0.001	0.01	0.015	10°

α_m is memory factor that determines the time scale for evolution of the online estimated bias terms in Eq. 5.
a_1,a_2,a_3 and ϕ_c are parameters defining the linear transformation matrix A (Eq. 6 and 7).

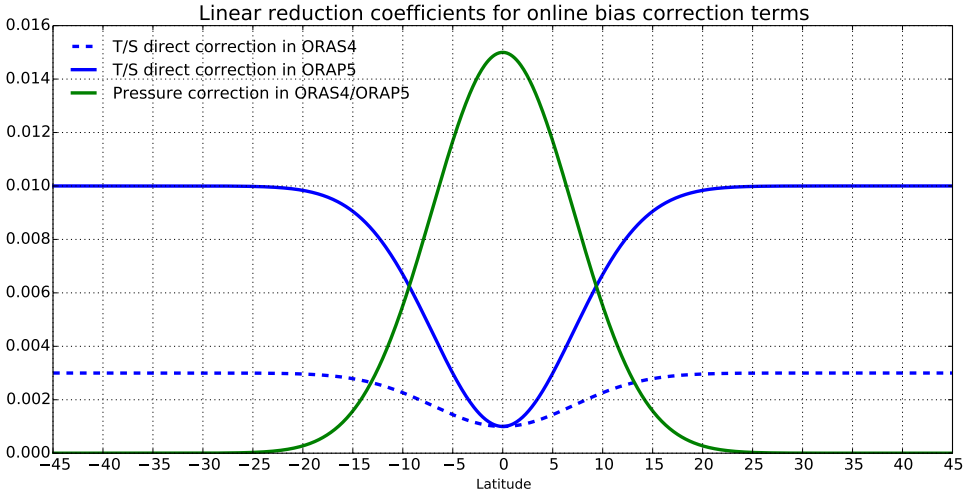


Figure 3: Latitude-dependent linear reduction coefficients as applied on online bias correction terms in equations 6 and 7: blue line - $a^{tr,T/S}$, reduction coefficients that apply to direct temperature and salinity corrections (different for ORAS4 and ORAP5); and green line - $a^{p,T/S}$, reduction coefficients that apply to pressure bias correction.

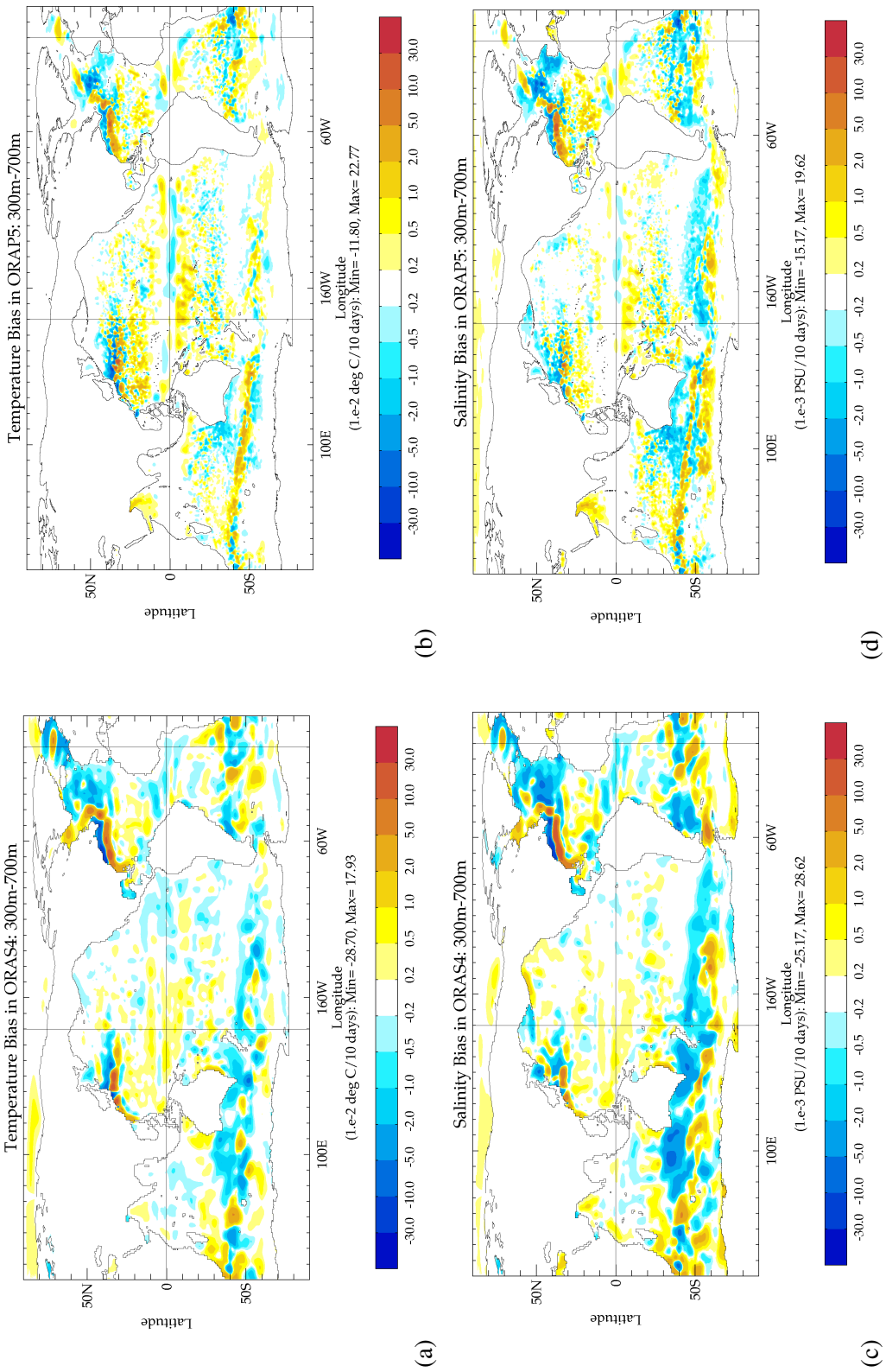


Figure 4: (TOP)Annual mean offline temperature bias (300-700 m) as applied in (a) ORAS4 and (b) ORAP5; (BOTTOM) Annual mean offline salinity bias (300-700 m) as applied in (c) ORAS4 and (d) ORAS5.

2.4 Observations

In-situ profiles of temperature and salinity data from the quality-controlled EN3 dataset [Ingleby and Huddleston, 2007] are assimilated in ORAP5. EN3 version 2 with XBT depth correction [Wijffels et al., 2008] is used from 1979 to 2011 and a standard version EN3 is used for year 2012. Fig. 5-(a) show the locations for EN3 observation profiles available between 2010-July-02 and 2010-July-06, with different colors to separate observation types. 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. 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 (See Mogensen et al. 2012 for details). Red dots in Fig. 5-(b) show rejected observation profiles due to horizontal thinning. A vertical thinning (no more than 2 observations per model level) is also applied, to ensure that data with high vertical resolution (i.e. CTD) are not given too much weight in analysis. This is different from ORAS4, which allows 3 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 dataset (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 (ECMWF version 4) was used when producing ORAP5. This is different from the most recent AVISO version released in mid 2014 (ECMWF version 5). In comparison, ORAS4 uses different versions of AVISO data: prior to operational implementation it used the delayed AVISO product available at 2010; during its operational phase, ORAS4 has been using the near-real-time product from subsequent AVISO releases (ECMWF versions 3, 4 and 5).

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: observation within the same day and within each point 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 section 4).

To assimilate AVISO SLA, a new method was developed which can calculate the model Mean Dynamic Topography (MDT) relative to arbitrary period. The MDT is still derived from a previous assimilation run where T and S are assimilated (SSH_{TS}). But instead of using the same reference period (period 1993-1999, hereafter as P1) as AVISO SLA, the MDT in ORAP5 (MDT_{new}) is estimated by averaging the model sea surface height during the 2000-2009 period (hereafter as P2), when the large scale ocean is adequately sampled by Argo. A spatially dependent correction factor ΔMDT 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 (SLA_{alti}) between the two different periods, as follows:

$$MDT_{old} = SSH_{TS}^{P1} \tag{8}$$

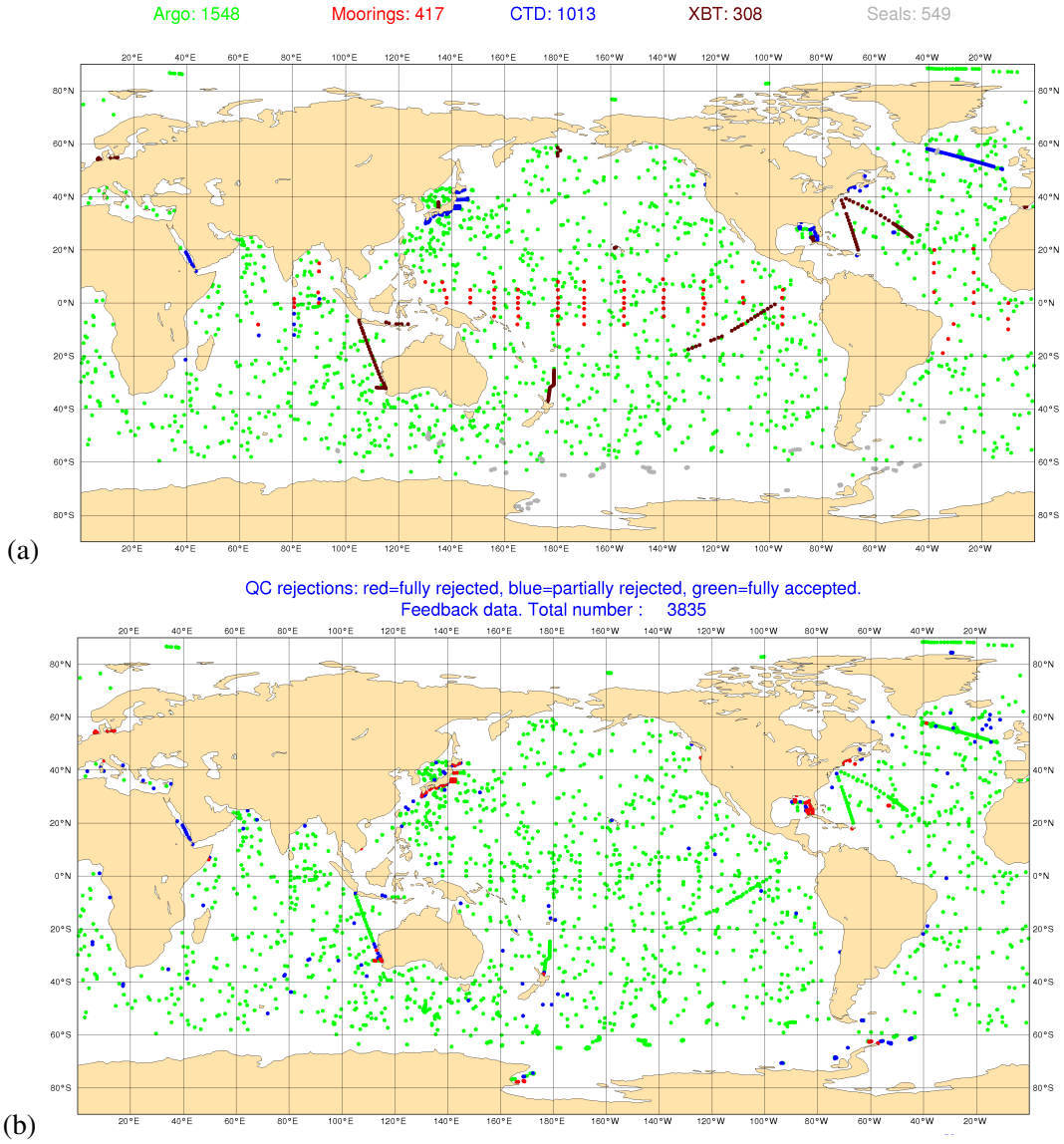


Figure 5: (a) Type and location for EN3 observation profiles from 2010-July-02 to 2010-July-06 for 5 days; (b) observation QC status after applying horizontal thinning for CTD and XBT data profiles. Among 1321 CTD and XBT profiles, 874 are rejected. Note: observation type Seals actually mean mammals here.

$$MDT_{new} = SSH_{TS}^{P2} - \Delta MDT \tag{9}$$

$$\Delta MDT = SLA_{alti}^{P2} - SLA_{alti}^{P1} \tag{10}$$

$$MDT_{new} - MDT_{old} = (SSH_{TS}^{P2} - SSH_{TS}^{P1}) - (SLA_{alti}^{P2} - SLA_{alti}^{P1}) \tag{11}$$

Fig. 6 illustrates the new MDT procedure applied at low resolution (1°) ORAP5-equivalent experiment. The top panel shows the differences in mean model SSH between the 2000-2009 and 1993-1999; the middle panel shows the correction ΔMDT, estimated from the altimeter observations as specified in Eq. 10; the bottom panel shows the differences in MDT between the new and old methods ($MDT_{new} - MDT_{old}$), and equals the top panel minus middle panel. The differences in the tropics have been cancelled out and the main differences are over the Arctic, where altimeter coverage is limited, and over the mid-latitudes oceans, where the in-situ observation coverage has substantially changed. This new MDT estimation method has been validated in low resolution (1°) ORAP5-equivalent experiments, with little impact in 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. The SIC data has been interpolated into the ORCA025 grid and then assimilated in the NEMOVAR system. At the same time a latitude band and thinning algorithm (by a factor of 2) were applied for the SIC data to reduce the data density and to speed up convergence in the cost function.

The specification of OBservation-Errors (OBE) in ORAP5 is similar to ORAS4. All OBE are assumed to be uncorrelated so only OBE standard deviations need to be specified. The OBE standard deviation for temperature and salinity follows an analytical vertical profile [Ingleby and Huddleston, 2007] that only depends on depth (except near the coastlines), being an approximate fit to the vertical profiles of the globally averaged temperature and salinity OBE estimated in EN3. The SLA OBE standard deviation is constructed based on the standard deviation of SLA observations included in the construct of the super-observations, plus an additional term to compensate for those cases with few individual observations (see Section 4.3.1 in Mogensen et al. 2012). For sea-ice concentration, a constant SIC OBE standard deviation (0.2) is used everywhere. Representativeness error near the boundaries are taken into account by inflating the OBE near the coast, and it has been applied for temperature, salinity and along-track SLA observations. No OBE near-boundary inflation is applied to the SIC data at moment. The values of various parameters used in ORAP5 for calculating OBE standard deviations are given in Table 6.

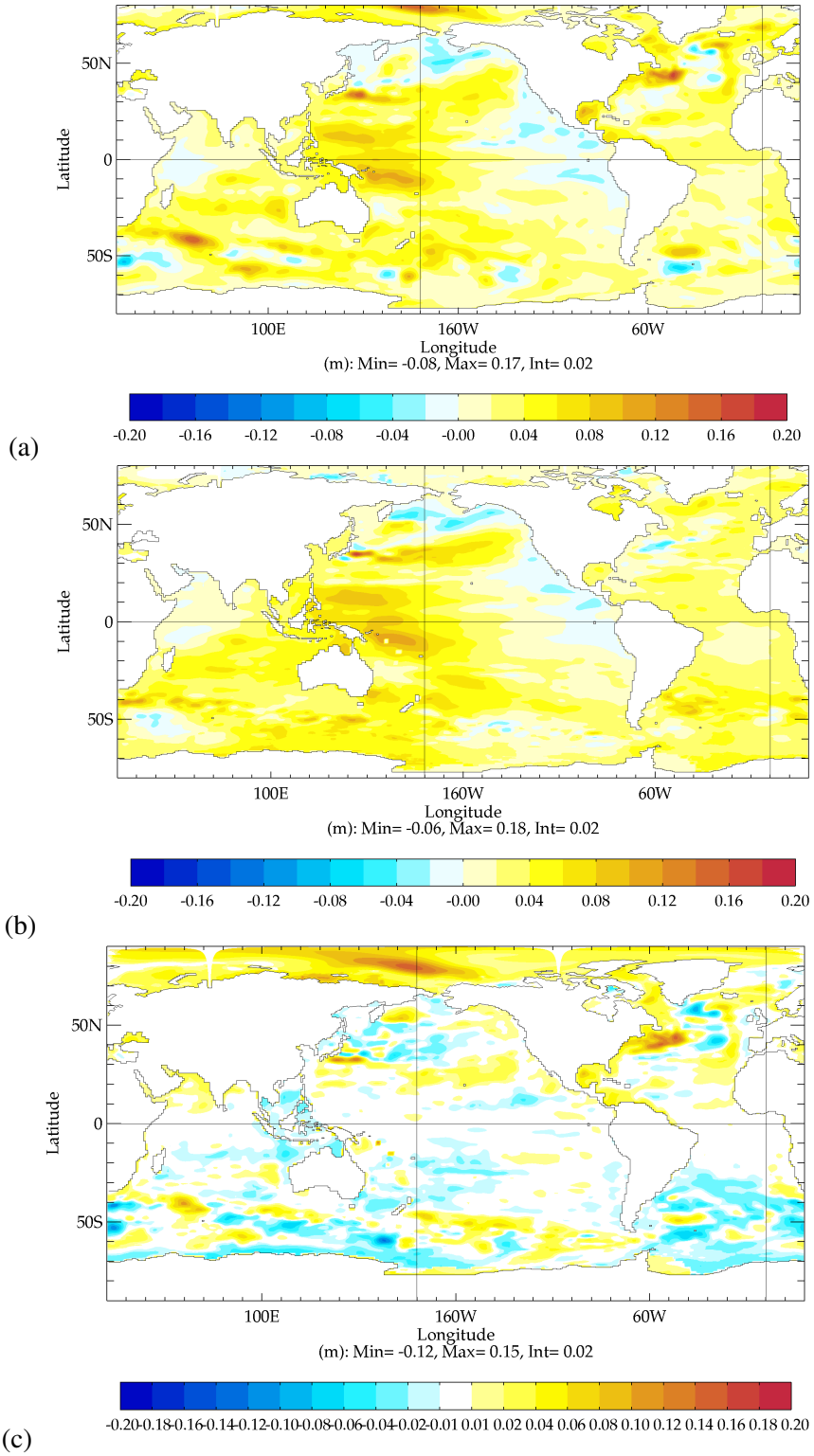


Figure 6: Differences between:(a) mean model SSH in 2000-2009 and 1993-1999 periods; (b) mean altimeter SLA in 2000-2009 and 1993-1999 periods; (c) MDT_{new} - MDT_{old} in m.

Table 6: Summary of the values used in ORAP5 for calculation of observation-error standard deviations

	temperature	**salinity**	**sea level**	**SIC**
ORAP5	$\sigma_T^{sur} = 0.78\ ^\circ C$, $\sigma_T^{max} = 1.0\ ^\circ C$, $\sigma_T^{do} = 0.07\ ^\circ C$, $D_1 = 75\ m$, $D_2 = 300\ m$, $D_3 = 450\ m$, $D_4 = 1000\ m$, $\delta_c = 6$ $R_c = 800\ km$	$\sigma_S^{sur} = 0.18\ PSU$, $\sigma_S^{do} = 0.02\ PSU$, $D_5 = 750\ m$, $\delta_c = 6$ $R_c = 800\ km$	$\overline{\sigma}_\eta = 0.05\ m$, $N_{min} = 10$, $\delta_c = 6$ $R_c = 800\ km$	$\overline{\sigma}_{ice}$=0.2, No coast inflation

σ_X^{sur} *is OBE standard deviations at the surface*

σ_X^{do} *is the minimum OBE standard deviation at deep ocean*

X can be temperature (T) or salinity (S)

σ_T^{max} *is the maximum temperature OBE standard deviation located at depth=*D_1

D_1,D_2,D_3,D_4 *and* D_5 *are depth parameters for constructing analytical profiles for OBE standard deviation*

δ_c *is the inflation factor at the coastline*

R_c *is the distance to coastline from which the inflation factor starts to apply*

$\overline{\sigma}_\eta$ *is the constant SLA OBE standard deviation*

N_{min} *is the minimum number for SLA super-observation sample size*

$\overline{\sigma}_{ice}$ *is the constant SIC OBE standard deviation*

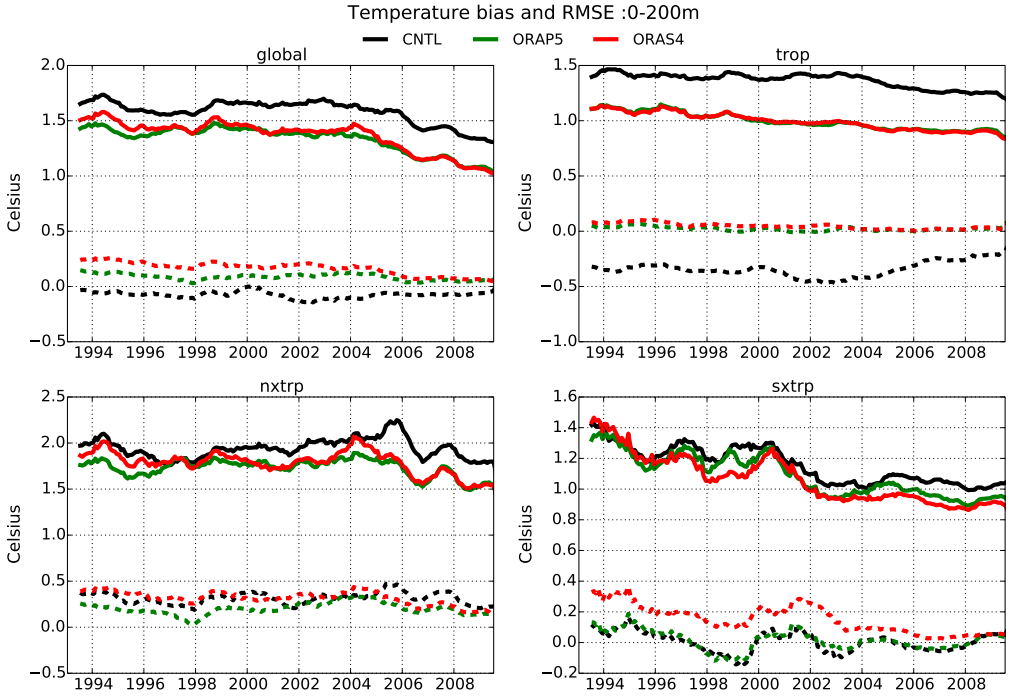
3 Preliminary Evaluation of ORAP5

3.1 Assimilation Statistics in Observation Space

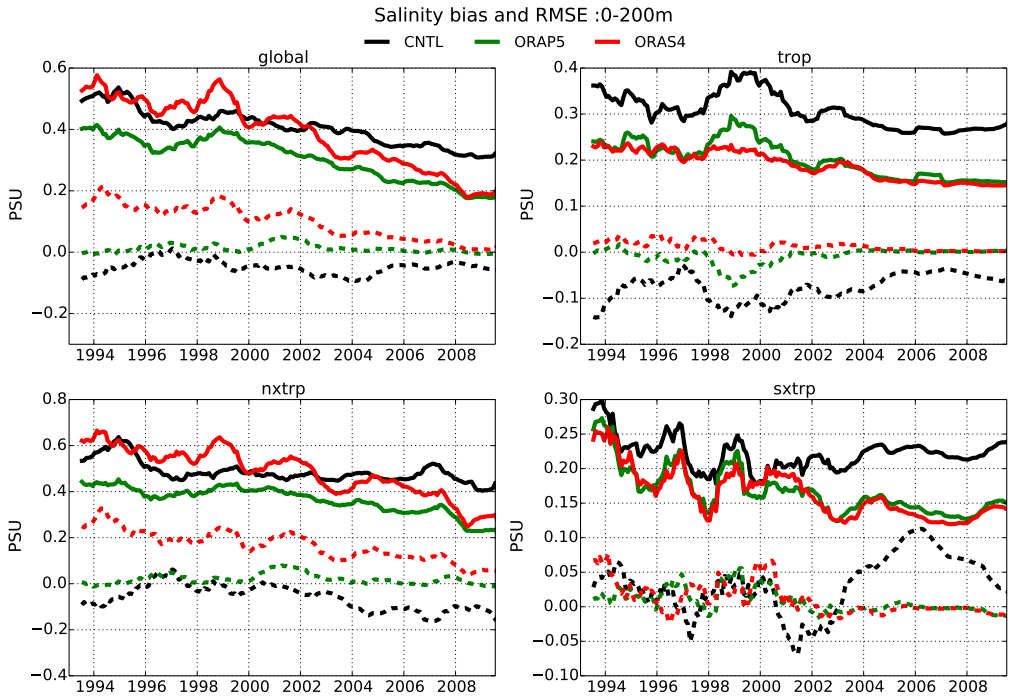
The quality-controlled EN3 dataset has been used for evaluation of model fit to observations. Bias and Root Mean Square Error (RMSE) statistics of the first guess minus observations are presented. The first guess is effectively the model background value from the first outer loop before updating the model variables with IAU. ORAP5 statistics are compared with those of ORAS4 and a control integration (hereafter CNTL). CNTL is an ocean-only simulation using the same initial condition, forcing fields, SST relaxation and 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 Section 2.3, 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 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 observation data (See Section 2.4 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 Fig. 7 and Fig. 8 respectively. Fig. 7 shows time series of bias (dashed line) and RMSE (solid line), averaged over the upper 200 m for temperature (Fig. 7-(a)) and salinity (Fig. 7-(b)). Different regions appear in separate panels. Fig. 8 shows maps of the bias in temperature (Fig. 8-(a),(c),(e)) and salinity (Fig. 8-(b),(d),(f)) for CNTL, ORAP5 and ORAS4, respectively. The statistics in observation space have been gridded by averaging over 5° by 5° boxes.

Fig. 7, shows time series for the global ocean and for three separate regions: tropics (trop: -30°S to 30°N), northern extratropics (nxtrp: 30°N to 70°N) and southern extratropics (sxtrp: -70°S to -30°S). Globally, the time series of temperature RMSE from ORAP5 and ORAS4 are very similar (Fig. 7-(a)), and both show improvement over the CNTL, with a mean RMSE reduced by ~0.25 $^{\circ}C$. 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. However, the left panels in Fig. 8 show that both assimilation experiments (ORAP5 and ORAS4 in Fig. 8-(c) and (e), respectively) exhibit significant smaller temperature bias than the CNTL (Fig. 8-(a)). This is the case for the large scale cold biases (~0.5 $^{\circ}C$) 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. 7-(b)) among three integrations. The salinity RMSE in ORAS4 is larger than CNTL before 2000, but reduced quickly following the introduction of Argo observations, suggesting a 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. 7-(b)), including CNTL, which does not assimilate any data. This is most likely the result of evaluating model in observation space, since the observation coverage is continuously evolving over time, and only with the Argo data reaches an uniform spatial sampling. Since the model errors in the open ocean are usually smaller than close to the coast, the uniform spatial sampling provided by Argo results in reduced global RMSEs. Fig. 8-(b) 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. 8-(d) and (f) for ORAP5 and ORAS4 respectively, due to assimilation of observations. Comparing to CNTL, the least improvement was found at data-sparse polar regions, and at coastal regions where most observations were either rejected or associated with large prescribed OBE variances.



(a)



(b)

Figure 7: 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 first guess minus observation and in the same observation space of EN3 data, after averaged over the upper 200 m in different regions: global (-90°S to 90°N), trop (-30°S to 30°N), nxtrp (30°N to 70°N), sxtrp (-70°S to -30°S).

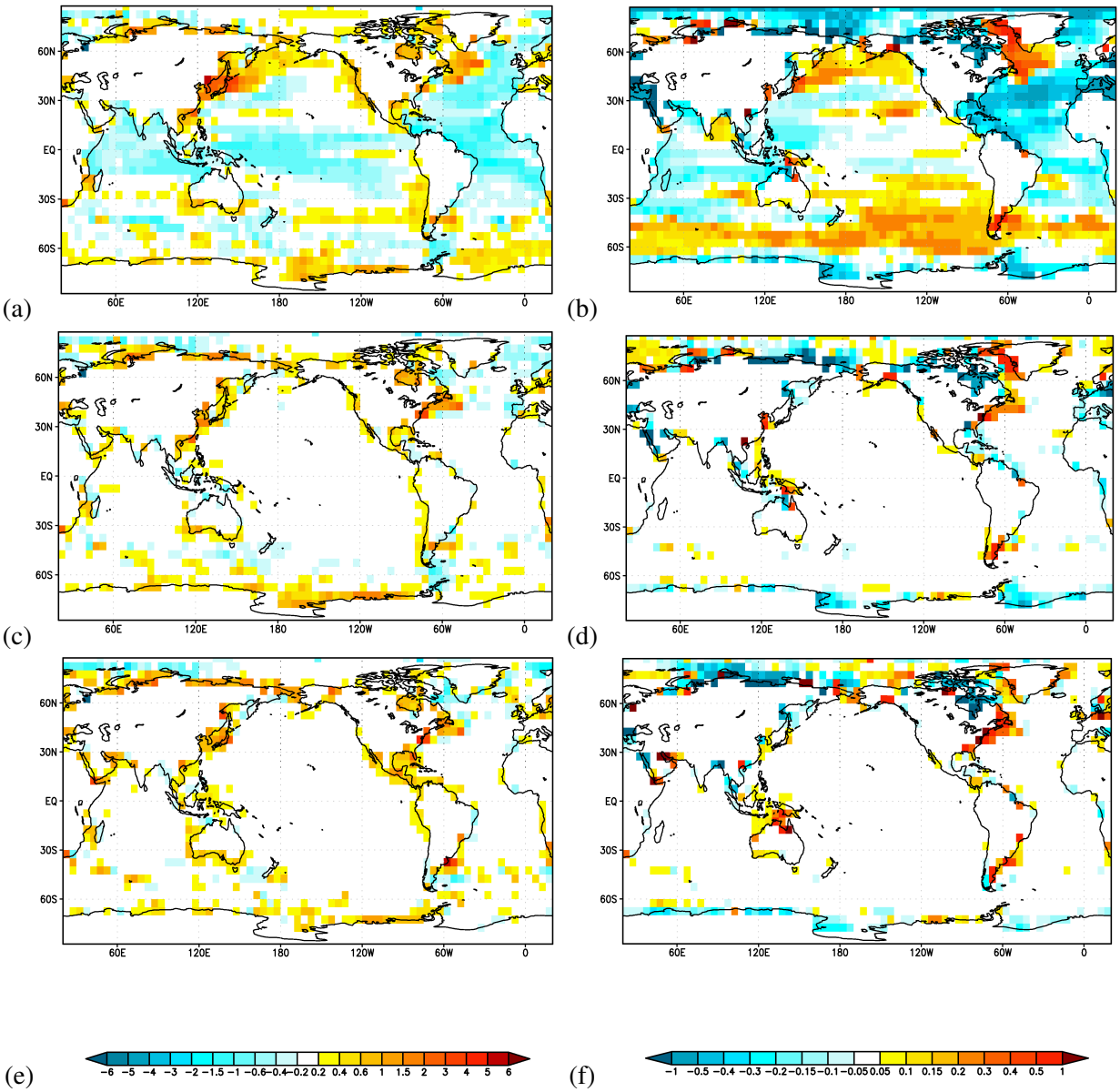


Figure 8: Maps of model temperature (left panels) and salinity (right panels) biases (first guess-observation) for CNTL (a,b); ORAP5 (c,d) and ORAS4 (e,f) after averaged from 0 to 200 m and over the period 1993-2009. Statistics are computed using model background value from the first outer-loop and against 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.

In the northern extratropics, ORAP5 shows smaller errors than ORAS4, especially in salinity (Fig. 7), which is likely due to increased model resolution. Here the RMSE is reduced by ~0.15 *PSU* in ORAP5 (Fig. 7-(b)). The strong positive salinity bias of ~0.2 *PSU* in ORAS4 is related to increased salinity errors at the Gulf Stream region (Fig. 8-(f)) . This suggests that an eddy-permitting model (i.e. horizontal resolution $\leq 0.25°$) helps the representation of ocean regions with strong eddy activities. ORAP5 also shows improved salinity bias (Fig. 8-(d)) in the Arctic region, but degrades 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. 8-(c) for ORAP5.

The ORAS4 and ORAP5 temperature errors are very similar in the tropical ocean (between -30°S and 30°N, Fig. 7-(a)), with improved temperature bias at the Gulf of Mexico and Indonesian Archipelago for ORAP5 (Fig. 8-(c)). Both assimilation experiments are substantially better than CNTL. When looking at the time series of tropical salinity statistics, a rising salinity RMSE between 1998 and 2001 appears in ORAP5 and CNTL, which is absent in ORAS4 (Fig. 7-(b)). This errors are caused by localized negative salinity bias around the Gulf of Mexico and north coast of Australia (Fig. 8-(d)), and correspond to statistics from shallow water CTD observations that account for over 80% of the total salinity observations in the Tropical Ocean during this period. In ORAP5 these observations have been rejected by QC process and therefore were not assimilated. The high resolution CNTL presents the same behaviour as ORAP5. Thus, the increased error in the salinity statistics for the high resolution simulations may be due to more detailed but largely unconstrained spatial structure present in these regions of large errors.

In the southern extratropics, ORAS4 shows slightly reduced RMSE in both temperature and salinity relative to ORAP5 (Fig. 7). ORAS4 also have reduced temperature biases near the Drake Passage and when near the coast of Antarctic (Fig. 8-(e)), and reduced salinity biases at Ross and Weddell Seas (Fig. 8-(f)). Again, this may be a consequence of the higher variability in ORAP5, which remains insufficiently constrained.

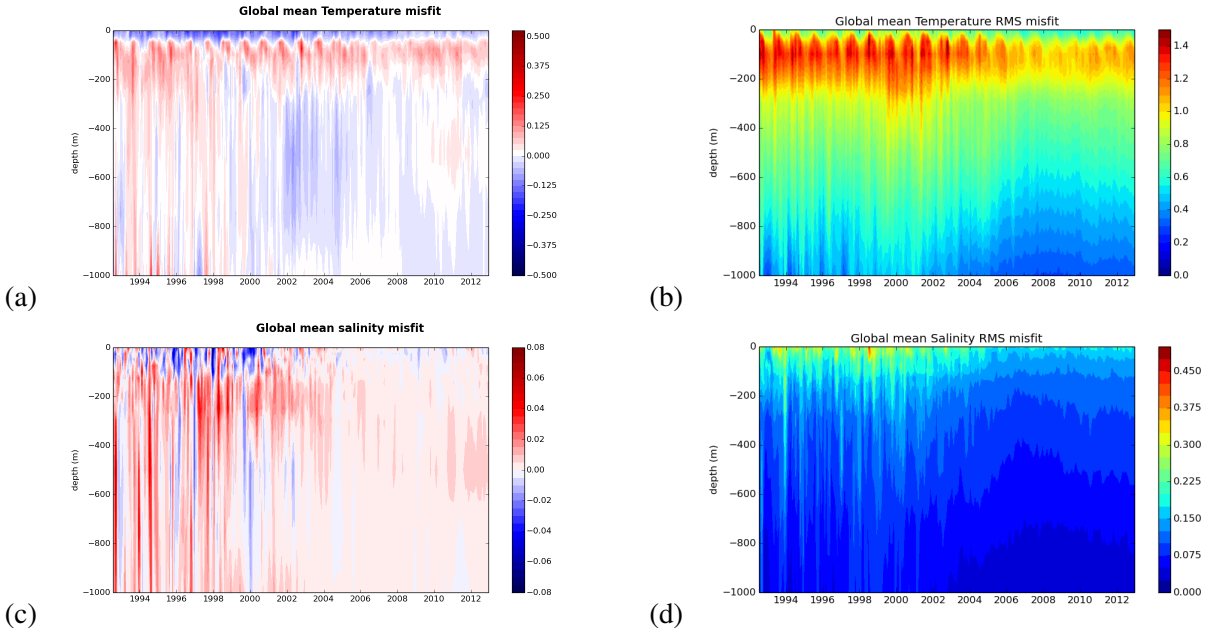
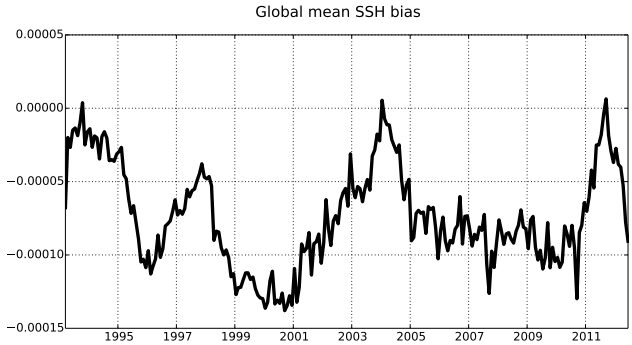


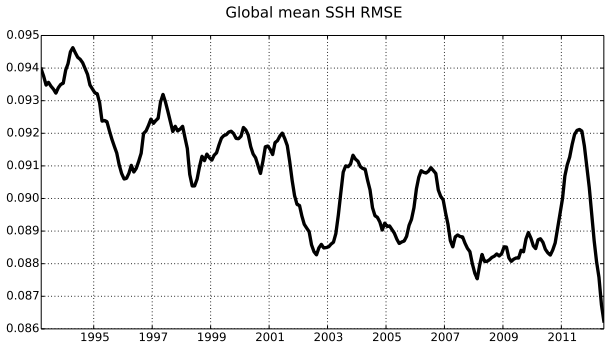
Figure 9: Global mean ORAP5 fit to observation errors as measured by bias (top panels) and RMSE (bottom panels) for temperature (left panels) and salinity (right panels), respectively. Statistics are computed using model analysis from the second out-loop after corrected by IAU and against EN3 in-situ observations, covered period between 1993 and 2012.

Fig. 9 shows the time evolution and vertical structure of the ORAP5 global temperature and salinity bias and RMSE for the upper 1000 *m*. Statistics are computed using model analysis from the second outer loop after correcting variables with IAU. The temperature RMSE is largest near the depth of 100 *m* and spans the depth range of 0-300 m, with a global mean cold bias (~0.12°C) in the surface and warm bias (~0.1°C) underneath. Seasonal signals are identified from temperature bias and RMSE, with increased errors during boreal summer. It can be seen that there is a general reduction upon to 1000 *m* depth for temperature RMSE during the Argo-era (after 2000s), which is a result of improved spatial sampling of observations, as discussed above. The changing observing system is also reflected in the temperature bias

below 200 *m* depth, as the bias reverses from positive to negative after 2000, which suggests that build-up of model bias at this depth range is only corrected when enough observations become available following the deployment of Argo floats in previous data-sparse areas such as the Southern Ocean. For salinity the maximum RMSE is located near the surface, and reduces quickly with depth. The salinity bias pattern is much noisier than that of temperature, with generally negative bias (~-0.4 *PSU*) in the upper 100 *m* and positive bias (~-0.3 *PSU*) below. A rapid reduction of both bias and RMSE in salinity can be seen in all layers in the Argo-era and suggesting substantial changes in salinity observation coverage. It is because that data collected outside the tropical waters are also restricted to water temperature and for the upper 750 *m* only before the Argo-era.



(a)



(b)

Figure 10: Time evolution of global mean sea level (a) bias and (b) RMSE with 12-month running mean filter from ORAP5 (in m). Statistics are calculated using model SSH analysis from the second-outer loop and against AVISO observations.

The time evolution of global mean SSH bias and RMSE (against AVISO observation) from ORAP5 are shown in Fig. 10 from 1993-2012. A mean SSH bias of -7·10^{-5} *m* is obtained after averaging the whole period, with a strong inter-annual variability. The global mean SSH RMSE varies between 8.5-9.5 *cm*, with a decreasing trend for the whole analysis period (1993-2012). A rapid reduction of RMSE from 2011 onward as shown in Fig. 10-(b) is related with switch of observation dataset from delayed to near-real-time AVISO products.

Desroziers et al. [2005] proposed specific assimilation statistics to diagnose the a-posteriori background and observation error covariances, assuming unbiased errors. Weaver et al. [2013] extended the Desroziers et al. [2005] to a biased system, by removing the spatial mean error. Following Weaver et al. [2013], the observation-space assimilation statistics can be characterized as

$$D^2 = \overline{(d^b - \overline{d^b})^2} \tag{12}$$

$$A^2 = \overline{(d^a - \overline{d^a})^2} \tag{13}$$

$$R_d^2 = \overline{(d^b - \overline{d^b})(d^a - \overline{d^a})} \tag{14}$$

$$B_d^2 = D^2 - R_d^2 \tag{15}$$

where $d^b = (observation - background)$ is the innovation; $d^a = (observation - analysis)$ is the analysis residual (here analysis refer to the background field plus the increment obtained after the inner-loop minimization); the overbar operator denotes spatial averaging; D^2 and A^2 are spatial variances for innovation and analysis residual. R_d^2 and B_d^2 are Desroziers diagnosed OBE and BGE variances.

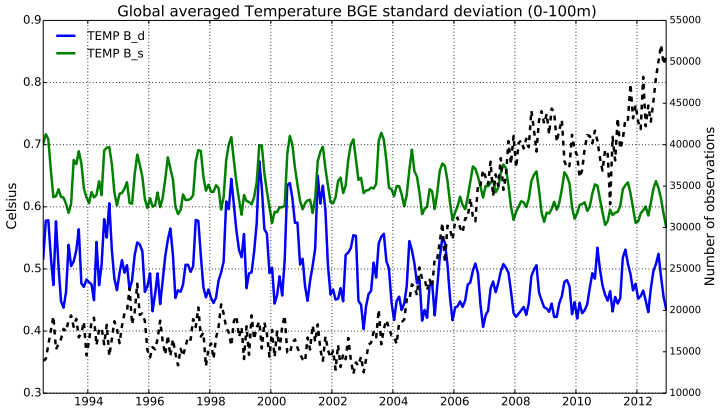
In order to check the statistical consistency of BGE covariance specifications in ORAP5, the globally averaged observation-space representation of the specified (B_s) and Desroziers diagnosed (B_d) BGE standard deviation in the upper 100 meters are calculated and shown as 20-year time series in Fig. 11-(a) and (b), for temperature and salinity, respectively. The specified BGE standard deviation in ORAP5 are parameterized in a way that depends on the vertical gradient of background state (see Section 2.3). For temperature BGE standard deviation, consistent seasonal variation can be found in both B_s and B_d, while both time series display a decreasing trend following the increase of observation number (dashed black line in Fig. 11). It is suggesting that temperature background state in ORAP5 improves due to assimilation of Argo float data. Specified salinity BGE standard deviation, however, does not present the same decreasing trend in B_d (blue line in Fig. 11-(b)) following the changing observation system. More work is needed to understand these results and parameterization of salinity BGE variances need to be reviewed.

The specified BGE standard deviations (B_s) for both temperature and salinity in the global domain are larger than that diagnosed from Desroziers method (B_d) in ORAP5. Ratios of diagnosed over specified standard deviation of BGE (B_d/B_s) in ORAP5 for temperature and salinity are shown as globally-averaged time series in Fig. 11-(c), with mean values of ~0.6 and ~0.3, for temperature and salinity, respectively. The BGE standard deviation is better specified in temperature (with B_d/B_s closer to 1) than in salinity. Both ratios are less than 1, which suggests that BGE covariances could be overestimated for temperature and salinity in ORAP5. However, these global ratios should be interpreted with care, since they assume an unbiased system. Although the bias has been removed over the global domain, the statistics will still include the spatial structure of the biases.

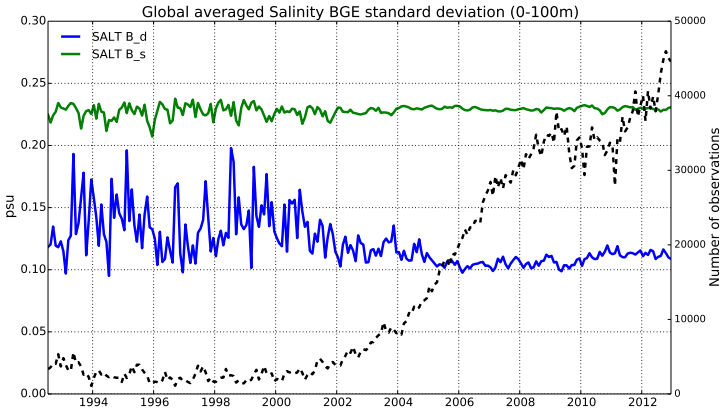
Mean ratio of B_d/B_s as diagnosed from ORAP5 are shown in Fig. 12 as spatial maps. Large temperature ratio ($B_d/B_s > 2$) can be found in the boundary current regions and Antarctic Circumpolar northern front. It is indicating that the specified temperature BGE is underestimated in these regions, or that large biases still exist. In contrast, the salinity B_d/B_s ratio rarely exceeds 1, and there are large areas (Southern Ocean, North-Eastern Atlantic, among others) where B_s is one order of magnitude larger than B_d. To interpret the implications of these diagnostics it would be necessary to compare similar statistics and ratios for the OBE.

3.2 Comparison with independent observations and estimates

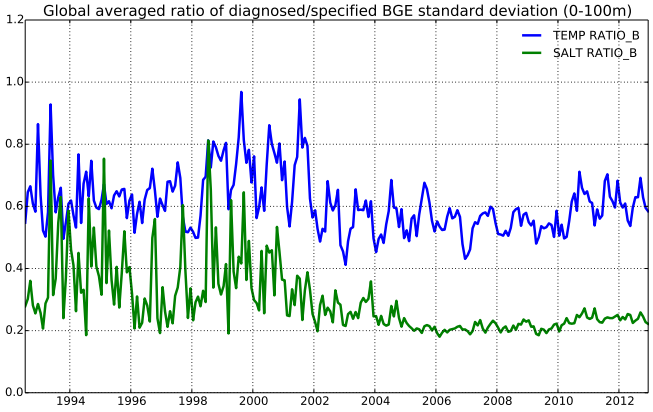
A preliminary validation against independent observations and ocean estimates has been carried out for ORAP5 reanalysis including using AVISO, tide gauges sea-level data and Ocean Surface Current Analysis - Real time (OSCAR) data. The same diagnostics from CNTL and ORAS4 reanalysis are also included.



(a)



(b)



(c)

Figure 11: Time series of global-averaged background error standard deviations as specified (B_s) in ORAP5 reanalysis and diagnosed (B_d) using Desroziers method [Weaver et al., 2013] for (a) temperature and (b) salinity; and (c) ratios of B_d/B_s. The black dashed line shows the number of observations as function of time, beware that axes for the number of observations between temperature and salinity are different. BGE standard deviations are averaged over the upper 100 m using monthly mean feedback files.

A useful metric to evaluate the fidelity of the interannual variability in the ocean reanalyses is the correlation with externally analyzed maps of sea level anomalies derived from altimeter data. These maps

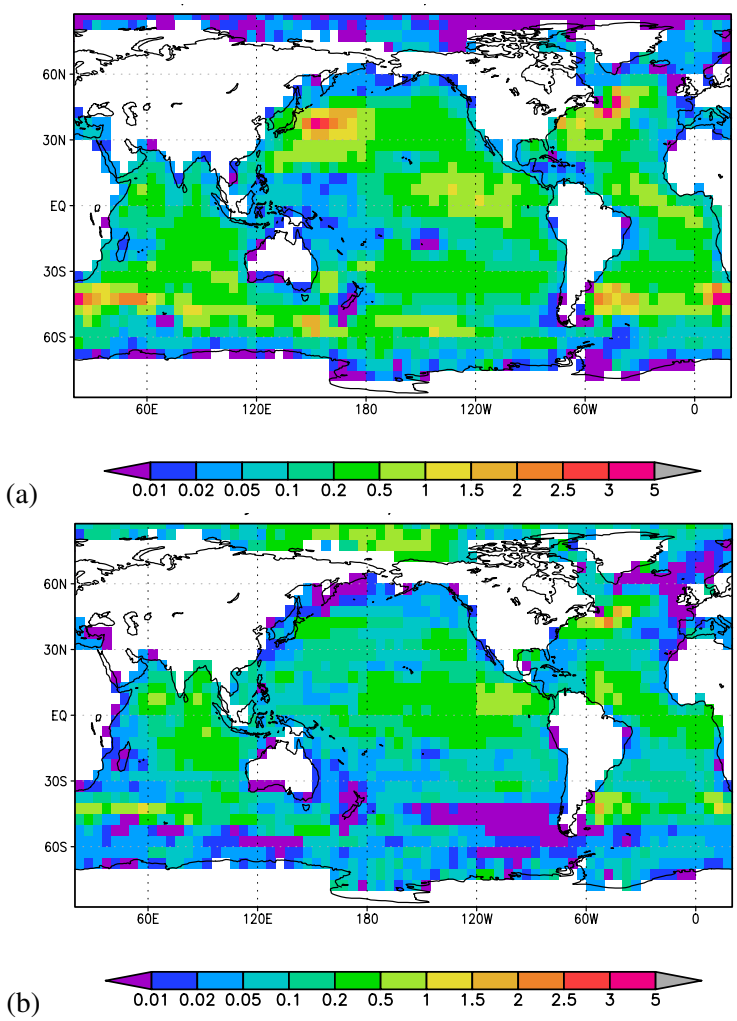


Figure 12: Ratio of diagnosed Desroziers BGE standard deviation (B_d) to specified BGE standard deviation (B_s) in ORAP5 for (a) temperature and (b) salinity. BGE standard deviations are averaged for the upper 100 m and over the period 1993–2012

are routinely distributed by AVISO. Although the data is not necessary independent (it might have been assimilated in the reanalyses), the gridding algorithm used by AVISO in the production of this maps is independent from the method in ORAP5, and it certainly does not contain information from any ocean dynamical model or in-situ observations. Maps of temporal correlation of monthly mean sea-level between three model estimates (CNTL, ORAP5 and ORAS4) and AVISO altimetry for the period 1993–2008 are shown in Fig. 13. Compared to CNTL, the correlation with AVISO altimetry is improved in ORAP5, and particularly in the tropical regions. ORAS4 shows an even higher AVISO-correlation than ORAP5. One possible explanation is that the ratio altimeter super-observations to model grid points is higher in ORAS4 than in ORAP5. The Superobbing scheme for pre-treatment of altimeter data is the same in ORAP5 and ORAS4, both using 1° super-observations for sea-level assimilation. This choice, suitable for ORAS4, may not be appropriate for ORAP5, which has higher horizontal resolution. See Section 4 for further discussion.

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 gauges records as daily

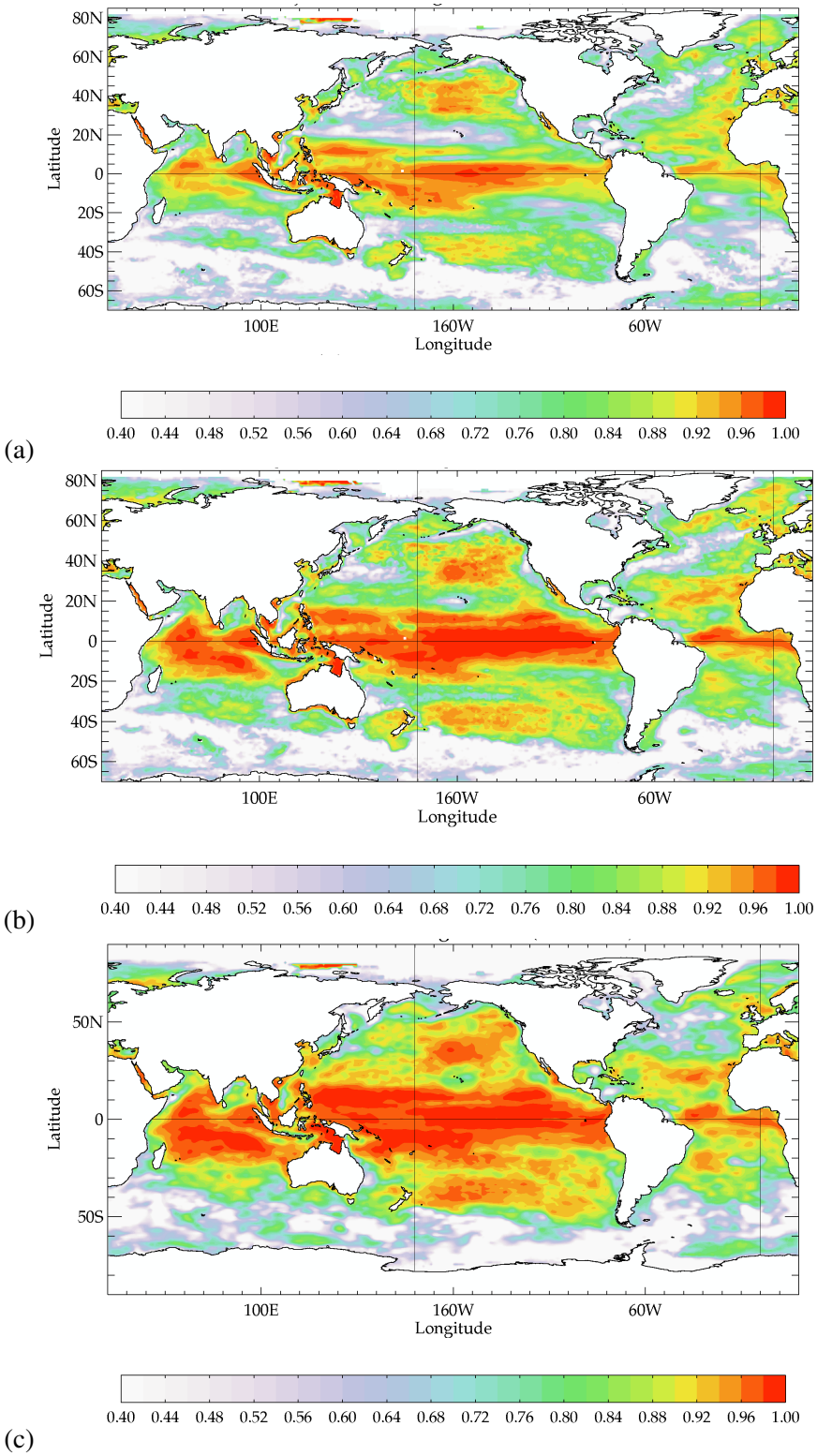


Figure 13: Maps of temporal correlation between AVISO sea level and (a) CNTL, (b) ORAP5 and (c) ORAS4 monthly sea level results. The statistics have been computed with monthly mean sea level for the period 1993-2008. Only values above 0.4 are shown.

averaged sea level and has been used for altimeter calibration [Lefèvre et al., 2005]. A reduced data set of BADOMAR data including 72 tide gauges records as monthly mean sea level after being corrected from inverse barometer effect and tides by Mercator Ocean are used here for independent validation of ORAP5 simulation of SL variations. RMSE and correlation between ORAP5 monthly sea level and BADOMAR tide gauges data over the period 1993–2011 are shown in Fig. 14 and 15-(a), respectively. For each tide gauge station, the sea-level values at the nearest sea model points are used for comparison. Values of RMSE are generally satisfactory except along some coasts where tidal effect is important, with only 10% >8 *cm*. 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% has correlation value >0.6, and only 10% has correlation value <0.4. Fig. 15-(b) shows the differences in correlation values between ORAP5 and CNTL (ORAP5-CNTL) regarding to BADOMAR tide gauge data. Comparing to CNTL, whose global mean sea level trend is also constrained, ORAP5 is in general better correlated 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. 15-(c)) suggest that ORAP5 performed better for sea level variability in the Atlantic Ocean, but slightly worse in the Indian Ocean. 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 .

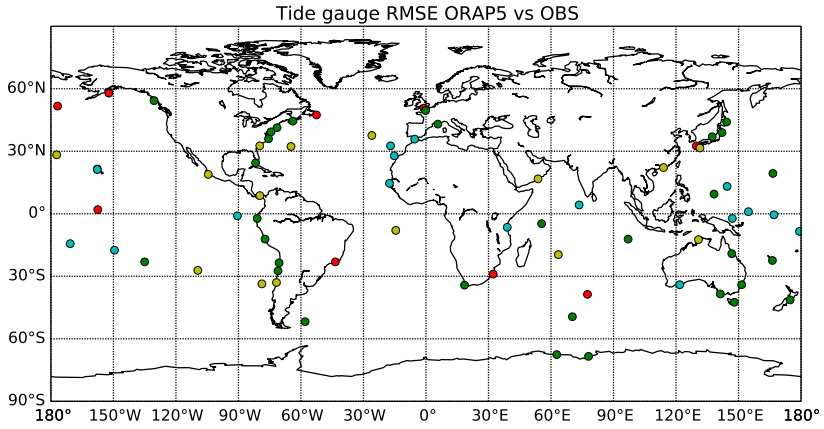
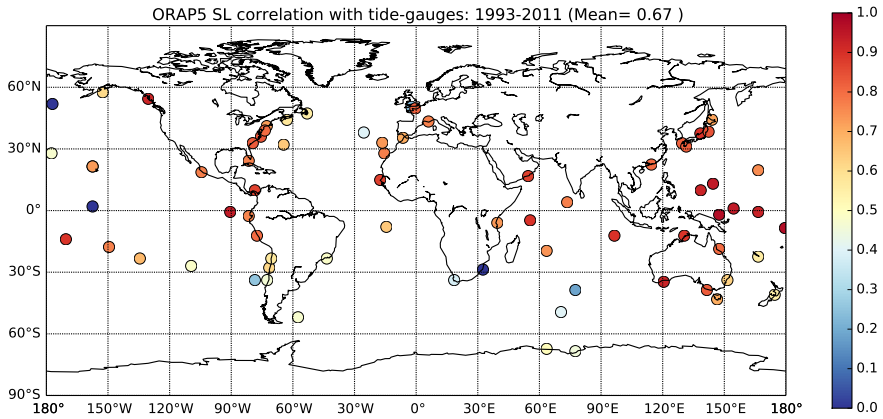
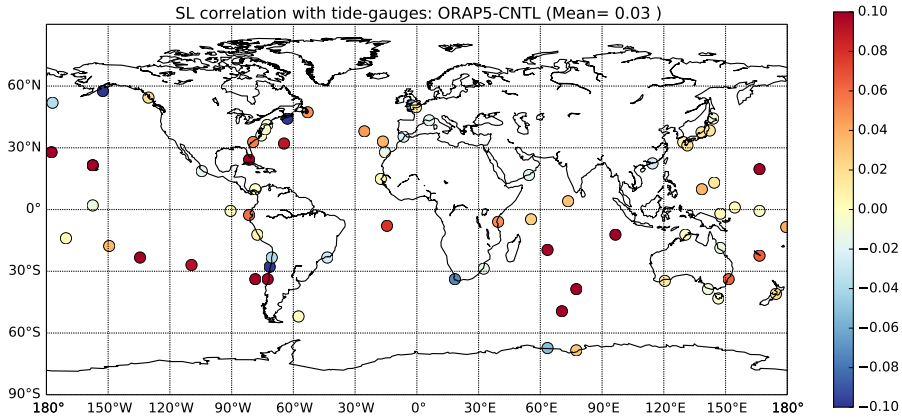


Figure 14: RMSE (cm) of ORAP5 sea level as calculated using monthly mean value and BADOMAR tide-gauge stations data from 1993 to 2011. Color meaning: 8 cm<RMSE (red), 6 cm<RMSE <8 cm (yellow), 4 cm<RMSE <6 cm (green), 2 cm<RMSE <4 cm (cyan), RMSE <2 cm (blue)

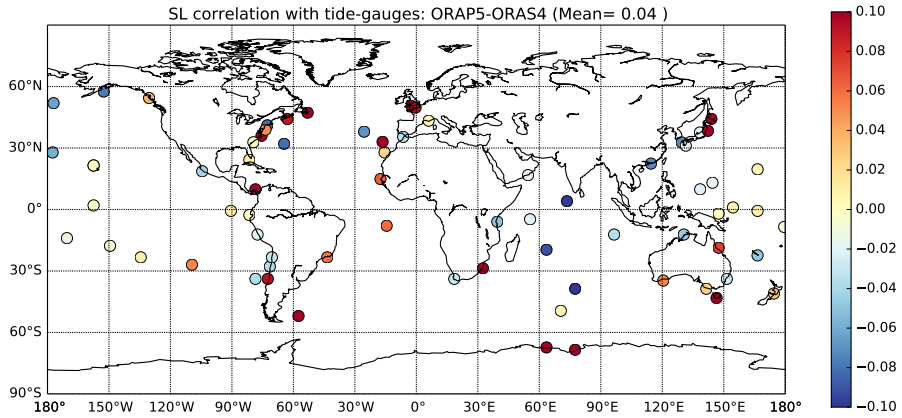
The OSCAR oceanic surface currents data is derived from satellite altimeter and scatterometer data [Bonjean and Lagerloef, 2002]. Since no oceanic velocity data is assimilated in ORAP5, the OSCAR data can be used as an independent source for evaluation of ORAP5 reanalysis. Fig. 16-(a) shows the correlation between surface zonal velocities from OSCAR monthly means and ORAP5. The correlation between ORAP5 and OSCAR surface currents over the period 1993–2008 is generally larger than 0.6 in the tropical regions, while the minimum correlation is found in the regions with strong currents, i.e. in the Gulf Stream region of the Northern Atlantic Ocean and in the Southern Atlantic and Indian Oceans. Difference in surface velocity correlation with OSCAR monthly data between ORAP5 and CNTL is shown in Fig. 16-(b). It is clear that ORAP5 has higher correlation value than CNTL almost everywhere in the global ocean, with some maximum increases up to 0.5 in the extra-tropical regions. Comparing to CNTL, the surface current in ORAP5 have been significantly improved due to assimilation of satellite



(a)



(b)



(c)

Figure 15: (a) Temporal correlation between ORAP5 sea level and BADOMAR tide-gauge stations data; (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 values at the nearest model point to each tide gauge station and over the period 1993–2011.

altimeter data as well as in-situ temperature and salinity data.

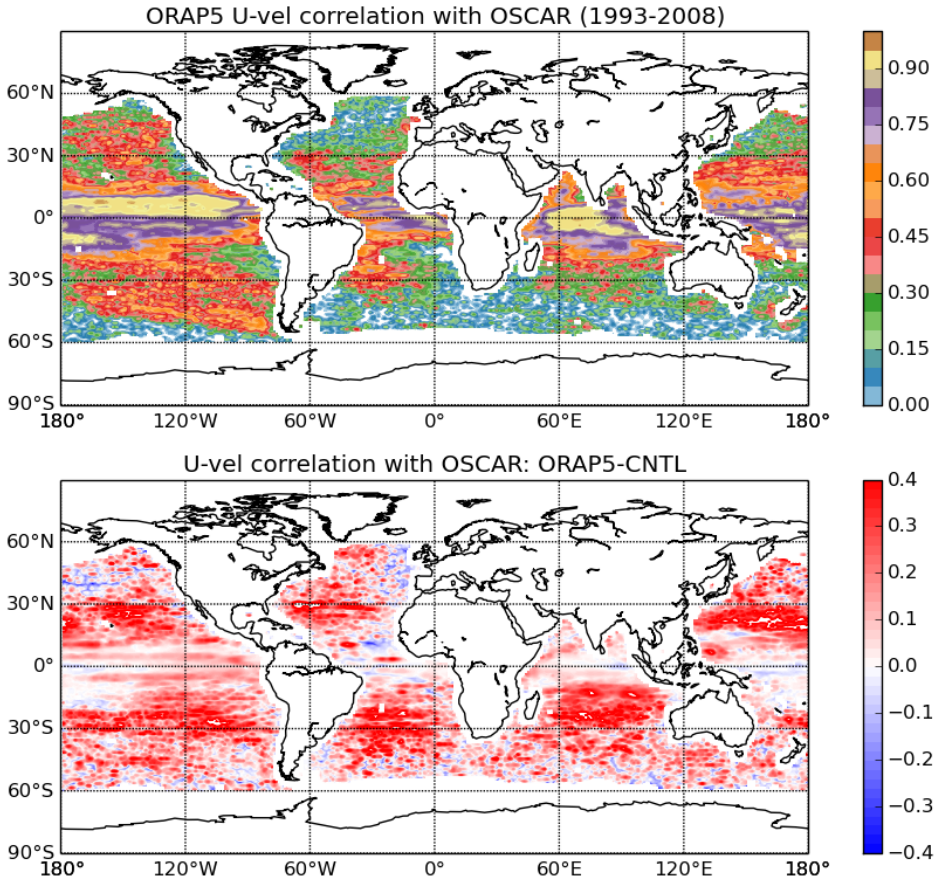


Figure 16: (Top) Temporal correlation between ORAP5 and OSCAR surface zonal velocity; (bottom) Correlation with OSCAR surface zonal velocity: ORAP5 - CNTL. Statistics are computed using monthly mean value from ORAP5 and CNTL over the period 1993–2008.

4 Sensitivity Experiments

Sensitivity experiments have been carried out in the process of specifying parameters for ORAP5 system configuration, including sensitivities to horizontal and vertical thinning of in-situ profile observations, bias correction scheme, assimilation of sea-ice concentration data, horizontal and vertical correlation length-scale scheme for specifying BGE. In this report we only present a subset of them, with focus on the evaluation of the satellite altimeter assimilation. These include sensitivity to superobbing of the satellite altimeter data and correlation length-scales of the BGEs. System settings for these sensitivity experiments are summarized in Table 7. 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, as well as the bias correction (Fig. 4). AVISO altimetry data are assimilated from 1993 onwards (except for NoSLA), and it is also used for constraining GMSL variations in all sensitivity experiments (see Section 2.3).

Table 7: Summary of the sensitivity experiments for sea-level assimilation

experiment name	Assim SLA	altimetry superobbing	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

4.1 Correlation with SLA maps

Maps of temporal correlation (period 1993–2008) of monthly mean sea-level between three sensitivity experiments (NoSLA, NoSuperob, Superob1) and AVISO gridded maps of SLA, are shown in Fig. 17. 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. 13-(a)) is considerable. Still, the NoSLA correlation map shows sharp minima in areas of step thermocline remain: along the North Equatorial Counter Current in the Pacific, and roughly along the paths of the North and South Equatorial currents in the Indian ocean. Along the Equatorial Pacific, correlation maxima collocated over the tropical mooring array are visible. Compared with the CNTL (Fig. 13-(a)) the assimilation of T/S also increases the correlation with the altimeter over the poleward side of the Pacific subtropical convergence areas, on both hemispheres.

The assimilation of AVISO along-track altimeter (in NoSurperob and Superob1) substantially increases the correlation with the AVISO altimeter maps (Fig. 17-(b) and (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, see Table 6). Superob1 is equivalent to NoSuperob but with the superobbing scheme (See Section 2.4) applied to altimeter data before assimilation. This practically implies a reduced 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 AVISO-correlation in the tropical regions (between 20°S and 20°N) is reduced slightly with respect to NoSuperob. Interestingly, the AVISO-correlation in the extra-tropical Pacific Ocean is higher in Superob1 than in NoSuperob. The correlation in Superob1 is very similar to ORAP5 (Fig. 13-(b)) and to Superob2 (not shown).

According to this metric (i.e., fit to the altimeter) the best estimation 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 whether the assimilation of altimeter improves the fit to the in-situ observations. The impact of altimeter assimilation on climate indices (which are not always observable) is also evaluated.

4.2 Fit to in-situ observations

The fit to the EN3 in-situ observations is shown in Fig. 18. The mean vertical profiles of model misfits to observations (as measured by RMSE) over the tropical oceans are shown in Fig. 18-(a) for temperature, and (b) for salinity, respectively. Statistics are calculated using the first-guess 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. 18-(a) and (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 in Superob1 is able to reduce the temperature RMSE between 50 and 200 m by ~0.08 $^{\circ}C$ when comparing 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 can improve the fit to in-situ observations, this is not guaranteed, and careful treatment of the altimeter data and evaluation of the results is 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 turned be decomposed into thermosteric and halosteric, i.e, changes in volume due to temperature/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, since often 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 area average of the vertical integral of the model density. Mass variations in the ocean is specified as equivalent bottom pressure (EBP) and esti-

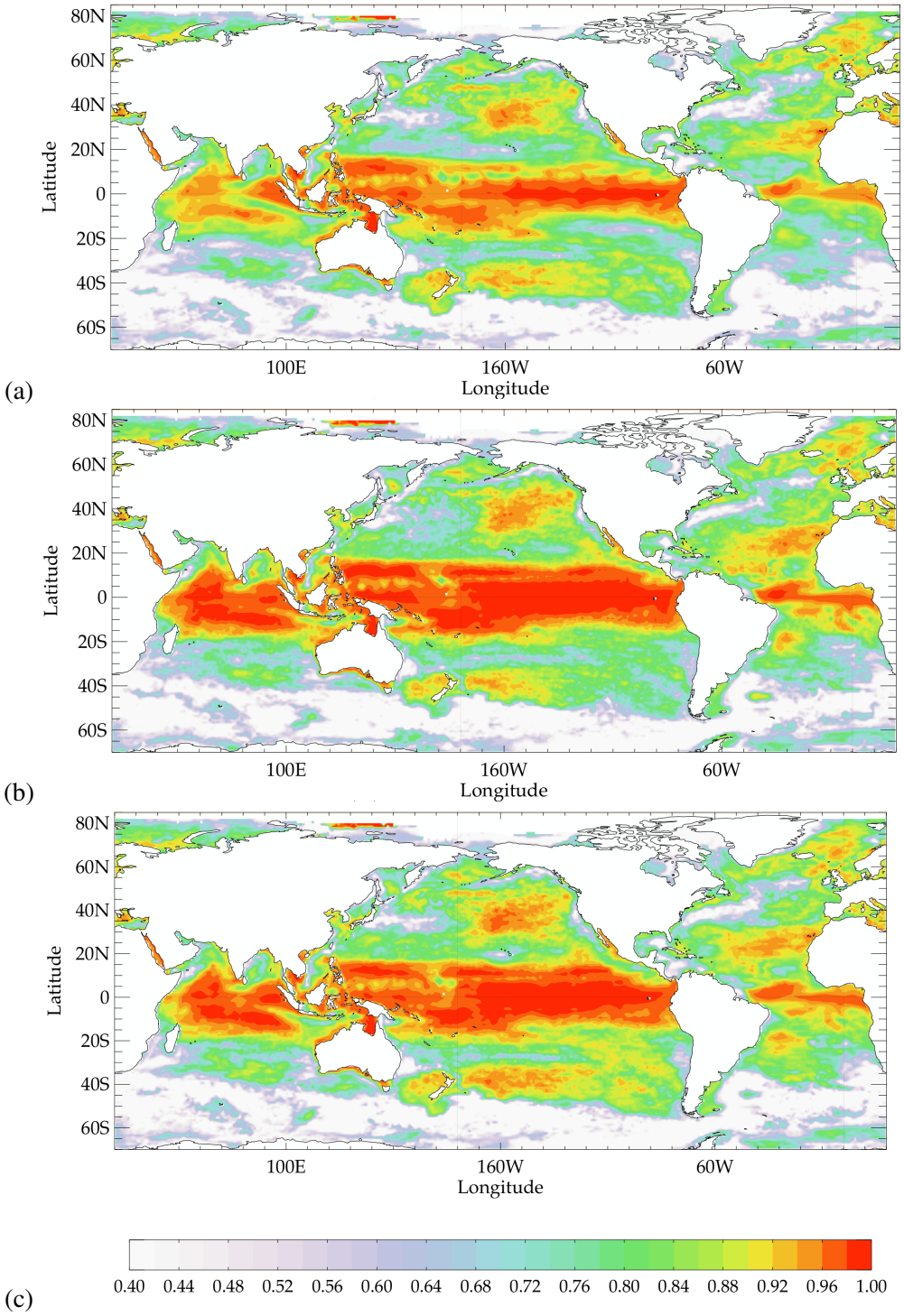


Figure 17: Maps of temporal correlation between analysis and AVISO sea level, with analysis from (a) NoSLA, (b) NoSuperob and (c) Superob1. The statistics have been computed with monthly mean sea level for the period 1993-2008, with only value above 0.4 are shown in the map.

mated as residual between GMSL and GSH [Balmaseda et al., 2013a]. Linear trend of GMSL and its components GSH and EBP are computed for all sensitivity experiments over the period 1993-2012, with results shown in Table 8. The linear trend in the GMSL for this period is about 2.8 $mm\ year^{-1}$ and is the

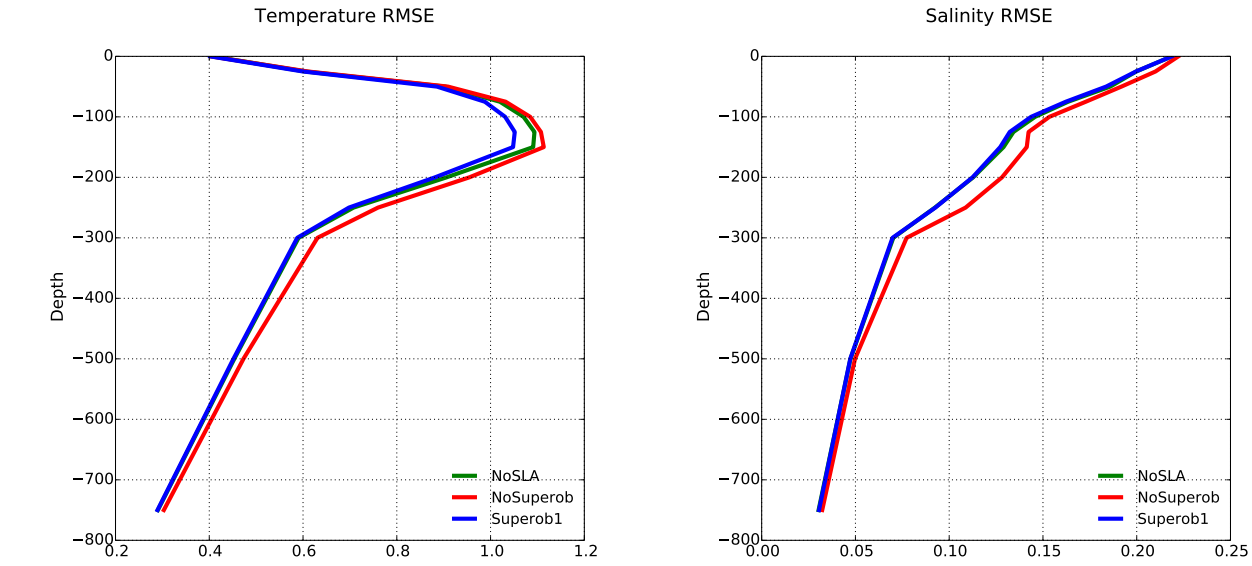


Figure 18: Profiles of mean model fit to observations RMSE for (a) temperature (°C) and (b) salinity (PSU) for the upper 800 m at the tropical oceans; RMSE are computed using background value from the first outer loop for NoSLA(green), NoSuperob(red) and Superob1(blue), after being averaged between 1993 and 2012. Here tropical oceans are defined as between 0°E to 360°E and 30°S to 30°N.

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 8 reflect only one aspect of the sensitivities. Perhaps more interesting is the time evolution of the GMSL partition. Fig. 19 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. 19-a) 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. 19-(c)), which assimilates temperature and salinity, shows a slow but steady increase of the steric contribution during 1993-2000. This increase accelerates from 2000 until 2004 (probably an artefact of the build-up of the 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 experiments shows a very different behaviour to CNTL and NoSLA (Fig. 19-(d)), 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 non-realistic. 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. The partition of GMSL changes into steric/mass in Superob1 is more even than in NoSuperob, as shown in Fig. 19-(d) and (e), respectively). In Superob1 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. 19-(b)-(e) and (f) respectively). ORAP5 and Superob2 have the same superobbing scheme as Superob1, but the horizontal correlation scale for the barotropic component of the altimeter $\bar{L}_\eta$ 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.

Table 8: 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

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

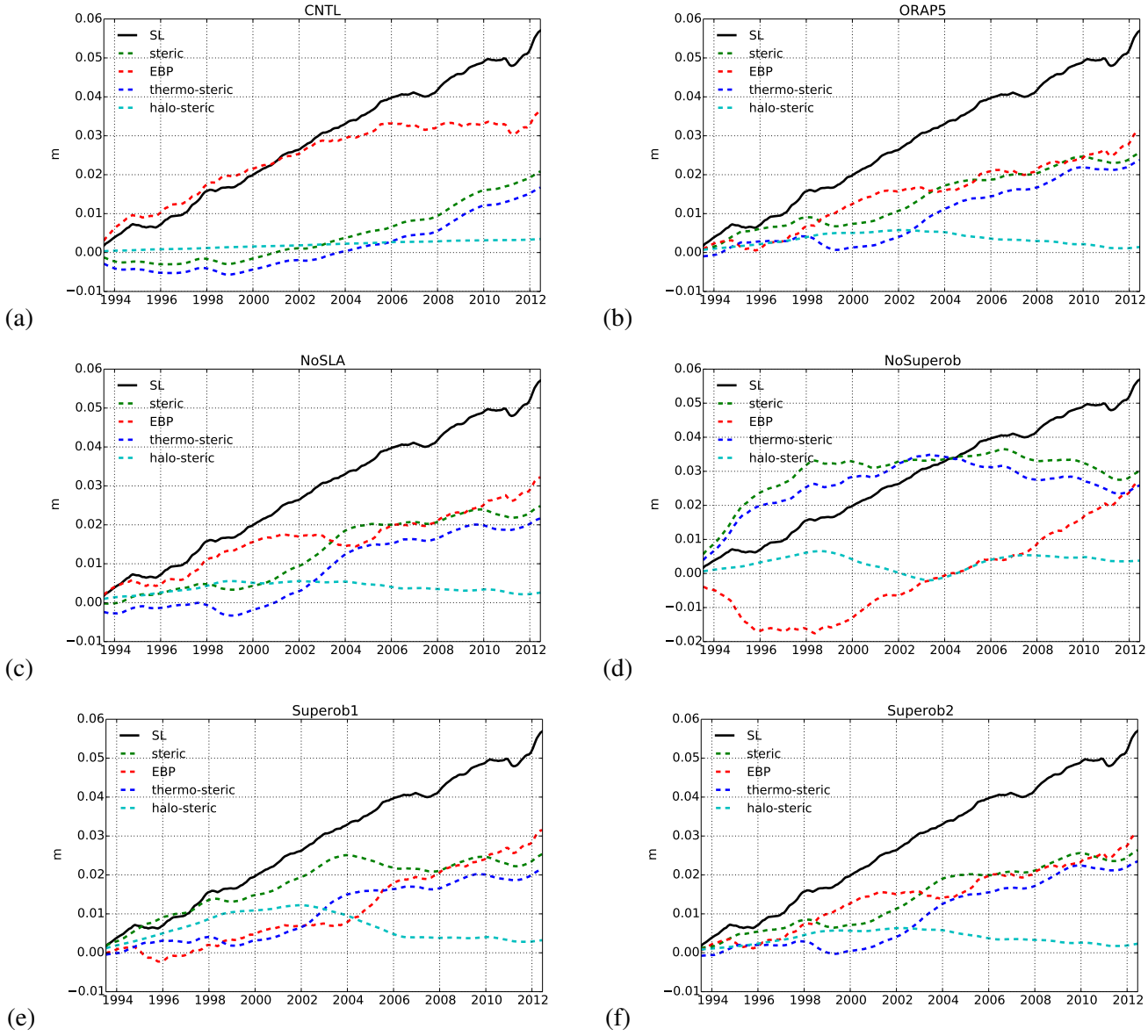


Figure 19: Time series of global mean sea-level anomalies (m) (solid line) and its components (dashed lines) as steric changes (green), EBP changes (red), thermo-steric changes (blue) and halo-steric changes (cyan) for (a) CNTL, (b) ORAP5 and sensitivity experiments: (c) NoSLA, (d) NoSuperob, (e) Superob1 and (f) Superob2. Time series calculated using monthly mean fields with 12-month running mean and values from 1993 January removed.

of data and parameter choice. Fig. 19 also shows the time series of thermosteric changes (blue dashed line) and halo-steric changes (cyan dashed line) for 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. 19-(a)), the halosteric component is almost constant, although a slight increase can be appreciated. This is more likely related with the vertical 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 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 BGE horizontal correlation length-scales for unbalanced SSH ($\overline{L}_\eta$ in Table 7). Superob1 exhibiting the largest halosteric variations (1.2 *mm year*$^{-1}$) with $\overline{L}_\eta = 4^\circ$, which account for about 2/3 of the total steric contribution for the period 1993-2002 (Table 9). The Superob1 halosteric component decreases after 2002, probably due to assimilation of Argo data. Superob2 is equivalent to Superob1 but with $\overline{L}_\eta$ reduced from 4° to 2°. Superob2 reduced the linear trend in halo-steric term to 0.6 *mm year*$^{-1}$ before 2002, and the sea level changes due to halosteric term is consistent with those that derived from NoSLA experiment (cyan dashed line in Fig. 19-(f) and (c), respectively). ORAP5 is equivalent to Superob2 but with vertical correlation length-scales factor (α in Table 7) 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. 19-(b) and Fig. 19-(f), respectively). Linear trend of total steric height changes and its components for all sensitivity experiments over the period 1993-2002 are shown in Table 9.

Table 9: 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

4.5 Errors in the region of the Mediterranean Outflow

An aspect critically affected by vertical correlation length-scales factor 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. Fig. 20 shows the RMSE of temperature at 1000 *m* in global ocean for Superob2 and ORAP5 after averaged between 2009 and 2012. Both experiments show large errors following the Mediterranean outflow, but the errors are slightly reduced in ORAP5 respect to Superob2. This error grows very quickly between successive assimilation cycles, and it appears mainly after 2009. Additional sensitivity experiments at lower 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 separately, a feature inherit to 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.

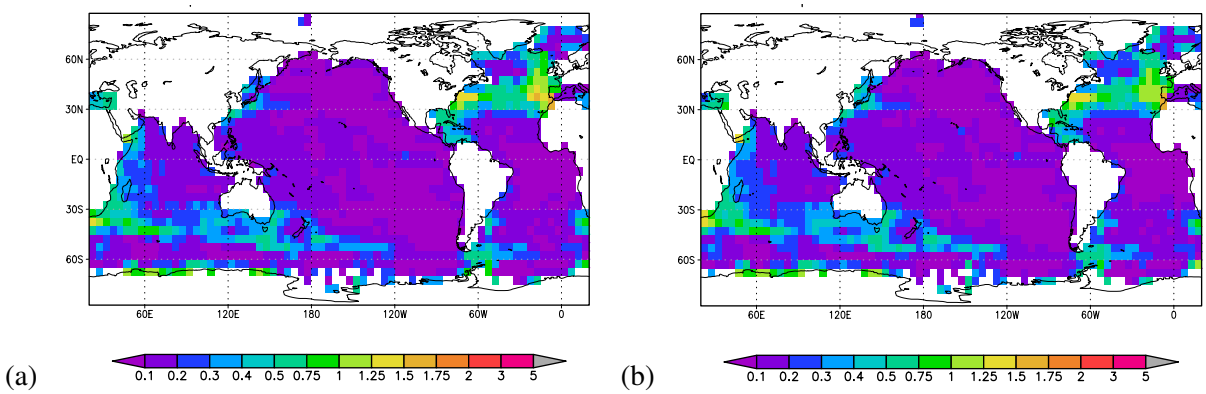


Figure 20: Temperature fit to in-situ observation errors as measured by RMSE (°C) at 1000 m for (a) Superob2 and (b) ORAP5. RMSE are calculated using temperature analysis from the second outer loop after corrected by IAU against EN3 in-situ observations, and averaged over the period 2009-2012.

5 Summary and Discussion

The 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 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 included in ORAP5 relative to ORAS4 have been discussed in detail in this paper.

ORAP5 uses a more recent version of the NEMO ocean model (V3.4.1 instead of V3.0) and an upgraded NEMOVAR. The forcing fields are from ERA-Interim for the whole 1979-2012 period, while ORAS4 only used ERA-Interim for the period 1989-2009. The surface boundary conditions are given by a modified bulk formulation (ORAS4 used fluxes directly from the atmospheric analyses), which include different surface wave effects (TKE input, Stokes-Coriolis, water-side stress and drag coefficient). The specification of the background and observation errors has been revised. ORAP5 assimilates in-situ observations from the same EN3 dataset as ORAS4, but extended to 2012 (ORAS4 started using GTS data in 2010). In addition, ORAP5 uses higher spatial and temporal resolution SST from the OSTIA reanalyses, instead of the weekly low resolution SST used in ORAS4. The sea level data used in ORAP5 is more up-to-date and uniform than that in ORAS4. The freshwater budget closure prior to the altimeter period has also changed in ORAP5. It is now based on a gravity-derived climatology of equivalent bottom pressure variations, allowing for interannual changes in GMSL consistent with changes in the steric height. Modifications were also introduced to reduce computational overhead of estimating the model MDT and bias correction.

The specification of the background errors in ORAP5 has been evaluated using the so-called Desroziers statistics [Desroziers et al., 2005]. These show that some flow dependent aspects of the background covariances are well captured by the NEMOVAR formulation, namely the seasonal cycle and an overall reduction and stabilization of the background error as a function of the number and spatial distribution of the observations. Although results from the globally averaged statistics suggest that the specification of

the BGE variances in ORAP5 are overestimated, this may be a reflection of the presence of spatial structure of the bias. The inspection of BGE maps shows that the temperature BGE can be underestimated in areas with large model errors or large internal variability.

ORAP5 has been evaluated using a variety of metrics, and its performance has been compared with an equivalent non-assimilation experiment (CNTL) and ORAS4. The first-guess from these three integrations is evaluated against quality-controlled EN3 in-situ observations using exactly the same observations. Time evolution and spatial distribution of model bias and RMSE are discussed. Visible declining trends in the temperature/salinity RMSE in all three integrations appear coinciding with the introduction of the Argo observing system. The reduction in RMSE is likely due to change in the observational spatial coverage associated with the implementation of Argo, and it does not necessarily imply an improvement in the ocean state estimation. ORAS4 is more sensitive to the introduction of Argo observations since 2000s and its salinity RMSE is declining faster than ORAP5. 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 but slightly increased errors in the southern extratropics in both temperature and salinity. It may be that the more detailed spatial structure resolved by higher resolution ORAP5 needs to be constrained by more observations. The in-situ observations available in the Southern Ocean are still sparse compared with the available observations in the Gulf Stream region. Care is needed when interpreting these statistics, since the observation coverage is not homogeneous, and quality control decisions are different for ORAS4 and ORAP5.

The temporal correlation with the AVISO 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 than the CNTL experiment. However, the high correlations are not as widespread spatially as in ORAS4. This can be a consequence of the superobbing scheme used in ORAP5 (the same as in ORAS4), which effectively creates a single superob within a 1° radius. Since ORAP5 has much finer resolution than ORAS4, the superobbing translates in giving relatively less weight to observations in ORAP5 than in ORAS4. The differences in temporal correlation with BADOMAR tide gauge stations records between ORAP5 and ORAS4 suggest that ORAP5 performed better for sea level variability in the Atlantic Ocean but slightly worse in the Indian Ocean.

Sensitivity experiments have been conducted in the process of specifying parameters for ORAP5 system configuration. Only a selection of sensitivity experiments are presented in this report, with focus on the assimilation of altimeter data, which constitutes an important pillar of the current ocean observing system. Along track altimeter sea-level has quite uniform spatial and temporal coverage, especially when compared with that of in-situ data, and provides an unique data set to constrain large and small scales. However, extracting information from the altimeter sea-level is not straight forward. The project of the sea level information into the vertical temperature and salinity structure relies heavily on the goodness of the background model profiles. If there are not enough in-situ observations to constrain the background subsurface field, giving too much weight to the altimeter observation can damage the solution. Results show indeed that while satellite altimeter data assimilation can increase the fit to altimeter observations (Fig. 17), it does not always translate into observable improvements in the ocean subsurface.

In ORAP5 the assimilation of altimeter improves slightly the fit to subsurface observations in the tropical ocean, but only after careful choice of the background and observation errors. 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. With-

out altimeter superobbing, these partition exhibited unphysical behaviour before the Argo period, with steric height tendency rates exceeding those of GMSL. The different sensitivity experiments show that the trend in global steric height is dominated by the thermosteric component. All the assimilation experiments show some interannual variability in the halosteric component before the Argo period, that can be considered spurious. This variability is also affected by specific parameters, such as the horizontal correlation length-scales 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, rejection of useful observations in the QC, or it is an inherent problem since assimilation scheme is currently being investigated. It is worth noting that although NEMOVAR preserves hydrostatic stability when assimilating temperature only, it does not guarantee hydrostatic equilibrium when assimilating temperature and salinity simultaneously. The inspection of the quality control decisions in ORAP5 showed that in this region, the temperature data in some profiles were rejected, while the salinity observations at the same location were accepted. This situation is a challenge for NEMOVAR, which has not been designed for the assimilation of salinity-only profiles. There is also a large risk of entering a positive loop of error amplification: since the temperature observations are not used, the background goes further apart from the observations, and the probability of rejection of temperature observations in the subsequent cycles will increase. A conservative approach to avoid this problem is to reject the salinity observations when the associated temperature data have been rejected.

ORAP5 is the basis for the next eddy permitting operational reanalyses ORAS5. Some of the aspects highlighted here will be revised. In particular the adequacy of superobbing versus thinning scheme in the assimilation of altimeter needs to be evaluated. The quality control procedure is also being revised. It is expected that ORAS5 will improve on some of the deficiencies of ORAP5, after the lessons learnt with the evaluation here presented.

Acknowledgements

This work has been carried out under the support of EU MyOcean2 project. Thanks for Jean Marc Molines from LGGE and Andrew Coward from NOCS for providing the input files for the DRAKKAR reference NEMO ORCA025 configurations and assisting with the implementation of the NEMO. We would also like to thanks the members of the NEMOVAR team.

References

C. Amante and B. W. Eakins. *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, 2009.

J. Antonov, D. Seidov, T. Boyer, R. Locarnini, A. Mishonov, H. Garcia, O. Baranova, M. Zweng, and D. Johnson. World Ocean Atlas 2009, Volume 2: Salinity. S. Levitus, Ed. *NOAA Atlas NESDIS*, 69: 184, 2010.

M. A. Balmaseda, D. Dee, A. Vidard, and D. Anderson. A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Quarterly Journal of the Royal Meteorological Society*, 133(622):167–179, 2007.

M. A. Balmaseda, K. Mogensen, and A. T. Weaver. Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, 139(674):1132–1161, 2013a.

M. A. Balmaseda, K. E. Trenberth, and E. Källén. Distinctive climate signals in reanalysis of global ocean heat content. *Geophysical Research Letters*, 40(9):1754–1759, 2013b.

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

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

F. Bonjean and G. S. Lagerloef. Diagnostic model and analysis of the surface currents in the tropical pacific ocean. *Journal of Physical Oceanography*, 32(10):2938–2954, 2002.

Ø. Breivik, K. Mogensen, J.-R. Bidlot, M. A. Balmaseda, and P. A. Janssen. Surface wave effects in the NEMO ocean model: forced and coupled experiments. *Journal of Geophysical Research, submitted*, 2015.

D. B. Chelton, R. A. Deszoeke, M. G. Schlax, K. El Naggar, and N. Siwertz. Geographical variability of the first baroclinic rossby radius of deformation. *Journal of Physical Oceanography*, 28(3):433–460, 1998.

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

A. Dai and K. E. Trenberth. Estimates of freshwater discharge from continents: Latitudinal and seasonal variations. *Journal of hydrometeorology*, 3(6):660–687, 2002.

D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, 2011. ISSN 1477-870X. doi: 10.1002/qj.828.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(613): 3385–3396, 2005.

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

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

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

M. H. England, S. McGregor, P. Spence, G. A. Meehl, A. Timmermann, W. Cai, A. S. Gupta, M. J. McPhaden, A. Purich, and A. Santoso. Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate change*, 4(3):222–227, 2014.

T. Fichefet and M. Maqueda. Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *Journal of Geophysical Research: Oceans (1978–2012)*, 102(C6):12609–12646, 1997.

V. Guemas, F. J. Doblas-Reyes, F. Lienert, Y. Soufflet, and H. Du. Identifying the causes of the poor decadal climate prediction skill over the north pacific. *Journal of Geophysical Research: Atmospheres*, 117(D20), 2012.

S. Gürol, A. Weaver, A. Moore, A. Piacentini, H. Arango, and S. Gratton. B-preconditioned minimization algorithms for variational data assimilation with the dual formulation. *Quarterly Journal of the Royal Meteorological Society*, 140(679):539–556, 2014.

B. Ingleby and M. Huddleston. Quality control of ocean temperature and salinity profiles-Historical and real-time data. *Journal of Marine Systems*, 65(1):158–175, 2007.

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

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

F. Lefèvre, J. Doradeu, E. Sénant, et al. Monitoring of the ENVISAT altimeter drift with a specific tide gauge database. *Workshop on Sea Level Variations Towards an Operational European Sea Level Service*, page 64, 2005.

R. Locarnini, A. Mishonov, J. Antonov, T. Boyer, H. Garcia, O. Baranova, M. Zweng, and D. Johnson. World Ocean Atlas 2009, Volume 1: Temperature. S. Levitus, Ed. *NOAA Atlas NESDIS*, 68:184, 2010.

G. Madec. NEMO ocean engine. (27), 2008.

M. Mayer, L. Haimberger, and M. A. Balmaseda. On the energy exchange between tropical ocean basins related to ENSO. *Journal of Climate*, (2014), 2014.

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

H. Pohlmann, W. A. Müller, K. Kulkarni, M. Kameswarrao, D. Matei, F. S. E. Vamborg, C. Kadow, S. Illing, and J. Marotzke. Improved forecast skill in the tropics in the new miklip decadal climate predictions. *Geophysical Research Letters*, 40(21):5798–5802, 2013.

R. W. Reynolds, T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax. Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, 20(22):5473–5496, 2007.

J. Roberts-Jones, E. K. Fiedler, and M. J. Martin. Daily, Global, High-Resolution SST and Sea Ice Reanalysis for 1985-2007 Using the OSTIA System. *Journal of Climate*, 25(18):6215–6232, 2012.

A. Simmons, S. Uppala, D. Dee, and S. Kobayashi. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter*, 110, 2007.

B. D. Tapley, S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins. GRACE measurements of mass variability in the Earth system. *Science*, 305(5683):503–505, 2004.

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

S. M. Uppala, P. Kållberg, A. Simmons, U. Andrae, V. Bechtold, M. Fiorino, J. Gibson, J. Haseler, A. Hernandez, G. Kelly, et al. The ERA-40 reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012, 2005.

J. Waters, D. J. Lea, M. J. Martin, I. Mirouze, A. Weaver, and J. While. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Quarterly Journal of the Royal Meteorological Society*, pages n/a–n/a, 2014. ISSN 1477-870X. doi: 10.1002/qj.2388.

A. T. Weaver, C. Deltel, É. Machu, S. Ricci, and N. Daget. A multivariate balance operator for variational ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3605–3625, 2005.

A. T. Weaver, L. Correa, and M. A. Balmaseda. Evaluation of the ECMWF Ensemble of Ocean Reanalyses using Assimilation Diagnostics. *Poster for GODAE OceanView Symposium 2013*, 2013.

S. E. Wijffels, J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, and J. A. Church. Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. *Journal of Climate*, 21(21):5657–5672, 2008.