The ECMWF operational ensemble reanalysis-analysis system for ocean and sea-ice: a description of the system and assessment

Hao Zuo[1], Magdalena Alonso Balmaseda[1], Steffen Tietsche[1], Kristian Mogensen[1], and Michael Mayer[1]

[1]The European Centre for Medium-Range Weather Forecasts, Shinfield Rd, Reading U.K. RG2 9AX

Correspondence: Hao Zuo (Hao.Zuo@ecmwf.int)

Abstract. The ECMWF OCEAN5 system is a global ocean and sea-ice ensemble of reanalysis and real-time analysis. This manuscript gives a full description of the OCEAN5 system, with the focus on upgrades of system components with respect to its predecessors ORAS4 and ORAP5. An important novelty in OCEAN5 is the ensemble generation strategy that includes perturbation of initial conditions, and a generic perturbation scheme for observations and forcing fields. Other upgrades include revisions to the a-priori bias correction scheme, observation quality control and assimilation method for sea-level anomaly. The OCEAN5 historical reconstruction of the ocean and sea-ice state is the ORAS5 reanalysis, which includes 5 ensemble members and covers the period from 1979 onwards. Updated version of observation data sets are used in ORAS5 production, with special attention devoted to the consistency of sea surface temperature (SST) and sea-ice observations. Assessment of ORAS5 through sensitivity experiments suggests that all system components contribute to an improved fit to observation in reanalyses, with the most prominent contribution from direct assimilation of ocean in-situ observations. Results of observing system experiments further suggest that Argo float is the most influential observation type in our data assimilation system. Assessment of ORAS5 has also been carried out for several key ocean state variables and verified against reference climate data sets from ESA CCI project. With respect to ORAS4, ORAS5 has improved ocean climate state and variability in terms of SST and sea-level, mostly due to increased model resolution and updates in assimilated observation data sets. In spite of the improvements, ORAS5 still underestimates the temporal variance of sea level, and continue exhibiting large SST biases in the Gulf Stream and extension regions which are possibly associated with misrepresentation of front positions. Overall, the SST and sea-ice uncertainties estimated using five ORAS5 ensemble members have spatial patterns consistent with those of analysis error. The ensemble spread of sea-ice is commensurable with the sea-ice analysis error. On the contrary, the ensemble spread is under-dispersive for SST.

1 Introduction

Ocean and sea-ice reanalyses (ORAs, or ocean syntheses) are reconstructions of the ocean and sea-ice states using an ocean sea-ice coupled model driven by atmospheric surface forcing, and constrained by ocean observations via a data assimilation

method (Balmaseda et al., 2015). Therefore, improvements in model physics, resolution, atmospheric forcings, observation data sets and data assimilation methods all contribute to the advancing the quality of successive generations of ORAs. The primary purpose of ORAs includes climate monitoring, initialization/verification of both seasonal forecasts and long term prediction such as decadal or climatic projection. These require that the ocean model and data assimilation method are kept frozen during the production of the reanalysis. In addition, a Real-Time (RT) extension of the ORAs is also produced in operational centres to initialize coupled forecasts (Xue et al., 2011; Waters et al., 2015; Balmaseda et al., 2013a), as well as for routinely monitoring of ENSO (Xue et al., 2017). For this purpose consistency between the ORA and its RT extension is crucial. This can be obtained by keeping a tight link between the RT extension and the ORA system (Mogensen et al., 2012). In this study, we describe OCEAN5, a new operational ocean and sea-ice ensemble reanalysis-analysis system at ECMWF, with the focus on description of system components, ensemble generation and assessment of some several key ocean state variables in the ORA produced using this system. Climate signals and uncertainties estimation using ORAs are important application of ORAs (Balmaseda et al., 2013b). However, relevant discussions are not included in this manuscript for the sake of concise, and will be included in a second paper (in preparation).

Ocean Reanalyses with a Real-Time extension have been produced routinely at ECMWF since 2002, when the OCEAN2 system was implemented (Balmaseda, 2005) as integral part of the seasonal forecasting system. It was with the implementation of OCEAN3 (Balmaseda et al., 2008) that the ocean reanalysis was run independently of seasonal forecasts, since it was also used to initialized the extended range (re-)forecast; this was the first time that ocean reanalysis at ECMWF were used to monitor the ocean climate. OCEAN4 (Balmaseda et al., 2013a) followed the same structure as OCEAN3, but it was a major upgrade: it was the first time that the NEMO ocean model was used at ECMWF, and the variational data assimilation NEMOVAR (Mogensen et al., 2012) was introduced.

OCEAN5 is the fifth generation of ocean reanalysis-analysis system at ECMWF. It comprises a Behind-Real-Time (BRT) component, that was used for production of Ocean ReAnalysis System 5 (ORAS5); and a Real-Time (RT) component, that is used for generating daily ocean analysis for Numerical Weather Prediction (NWP) applications. The ORAS5 has been developed at ECMWF based on ORAP5 (Zuo et al., 2015). As a successor to ORAS4 (Balmaseda et al., 2013a), ORAS5 benefits from many upgrades in both model and data assimilation method, as well as in source/use of observation data sets. The ocean model resolution has been increased to $0.25°$ in the horizontal and 75 levels in the vertical, compared to $1°$ and 42 layers in ORAS4. ORAS5 also includes a prognostic thermodynamic-dynamic sea-ice model (LIM2, see Fichefet and Maqueda (1997)) with assimilation of sea-ice concentration data. Another important novelty in ORAS5 is the explicit inclusion of surface waves effects in the exchange of momentum and turbulent kinetic energy (Breivik et al., 2015). The NEMOVAR data assimilation scheme has been updated with a new Rossby-radius-dependent spatial correlation length-scale (Zuo et al., 2015) and a new generic ensemble generation scheme which accounts for both representativeness errors in observation and structure/analysis errors in surface forcing (Zuo et al., 2017a). The OCEAN5-RT component includes all upgrades developed for ORAS5. It is initialized from ORAS5, and runs once a day to provide ocean and sea-ice initial conditions for all ECMWF coupled forecasting system.

The aim of this document is to describe ORAS5 as the ocean reanalyses component of the OCEAN5 system. Details of system upgrades after ORAP5 are discussed. This includes updates in the surface forcing and initialization (in Section 2.2), updates in in-situ observation and assimilation (in Section 2.3); updates in altimeter observation and assimilation (in Section 2.4); generation of the ensemble perturbations (in Section 2.5). The OCEAN5-RT analysis is presented in Section 3. Sensitivity experiments and assessment of ORAS5 system components can be found in Section 4. Section 5 presents evaluation results with selected ocean Essential Climate Variables.

2 The ORAS5 system

ORAS5 is a global eddy-permitting ocean and sea-ice ensemble reanalysis produced via the OCEAN5 system in its BRT stream. ORAS5 provides historical ocean and sea-ice conditions from 1979 onwards. And a spin-up period between 1958 to 1978 is also provided (INI1 in Table. 2), which can be treated as a backward extension by users that are interested in a longer reanalysis period. Here we give a brief overview of the model and methods used, with emphasis on the differences between ORAS5 and its predecessor ORAP5. This includes different observation data sets of SST, SIC, and in-situ observations; updates in bias estimation and observation quality controls; and a new method in ensemble generation and initialization. Impacts of these updates have been assessed with data assimilation experiments, normally in a reduced resolution in order to reduce computing cost. It is worth pointing out that improvements from these updates presented in this section may not add up to an accumulative "sum" of improvements in ORAS5, and an optimized best configuration is not always guaranteed if it is based on results from a low resolution system. However, this is the standard and only possible procedure to test many components in a complex system such as ORAS5.

2.1 Ocean-sea ice model and data assimilation

ORAS5 uses the same ocean model and spatial configuration as ORAP5 (Table 1). The NEMO ocean model version 3.4.1 (Madec, 2008) has been used for ORAS5 in a global configuration ORCA025.L75 (Barnier et al., 2006), a tripolar grid which allows eddy to be represented approximately between 50°S and 50°N (Penduff et al., 2010). Model horizontal resolution is approximately 25 km in the tropics, and increases to 9 km in the Arctic. There are 75 vertical levels, with level spacing increasing from 1 m at the surface to 200 m in the deep ocean. NEMO is coupled to the Louvain-la-Neuve sea-ice model version 2 (LIM2, see Fichefet and Maqueda (1997)) implemented with the viscous-plastic (VP) rheology. The wave effects introduced since ORAP5 (Breivik et al., 2015) were also implemented in ORAS5, with updated ocean mixing terms for wind. Given that the wave field is not defined under sea-ice, the wave impact in the Turbulent Kinetic Energy (TKE) scheme is not used under sea-ice. Instead, a constant value of 20 is used under sea-ice as coefficient of the surface input of TKE in ORAS5.

The reanalysis is conducted with NEMOVAR (Weaver et al., 2005; Mogensen et al., 2012) in its 3D-Var FGAT (First-Guess at Appropriate Time) configuration. NEMOVAR is used to assimilate subsurface temperature, salinity, sea-ice concentration (SIC) and sea-level anomalies (SLA), using a 5 day assimilation window with a model time step of 1200 s. The observa-

Table 1. Overview of differences between ORAP5 and ORAS5 in production system settings

	ORAP5	ORAS5
period	1979-2013	1979-present + a spin-up from 1958-1978
ensemble	1 member	5 members with perturbations in initial conditions, forcings and observations
spin-up	recursive spin-up, 1 member	spin-up with 5 ensemble members and different parameter choices
grids	~0.25°, 75 vertical levels	as ORAP5
model	NEMO 3.4, LIM2 ice model, wave effects	as ORAP5 + TKE mixing in partial ice cover + updated wave effects
forcing	ERA-Interim bulk formula + wave forcing	ERA-40 (before 1979) ERA-Interim (1979-2015) ECMWF NWP (2015-present) bulk formula + wave forcing
assimilation	3D-Var FGAT with 5-day window	as ORAP5 + revised observation QC + revised MDT for altimeter data assimilation
bias corre.	adaptive bias correction scheme	as ORAP5 + ensemble-based bias estimation + stability check
observations		
SST	ERA40 + Reynolds OIv2d (Reynolds et al., 2007) + OSTIA reprocessed + OSTIA operational	HadISST2 + OSTIA operational (Donlon et al., 2012)
T/S prof	EN3 with XBT/MBT correction (Wijffels et al., 2008)	EN4 with XBT/MBT correction (Gouretski and Reseghetti, 2010) + NRT
SLA	AVISO DT2010 (Dibarboure et al., 2011)	AVISO DT2014 (Pujol et al., 2016) + NRT
sea-ice	Same as SST	as ORAP5

tional information is also used via an adaptive bias correction scheme (Balmaseda et al., 2013a), which will be explained in Section 2.3.
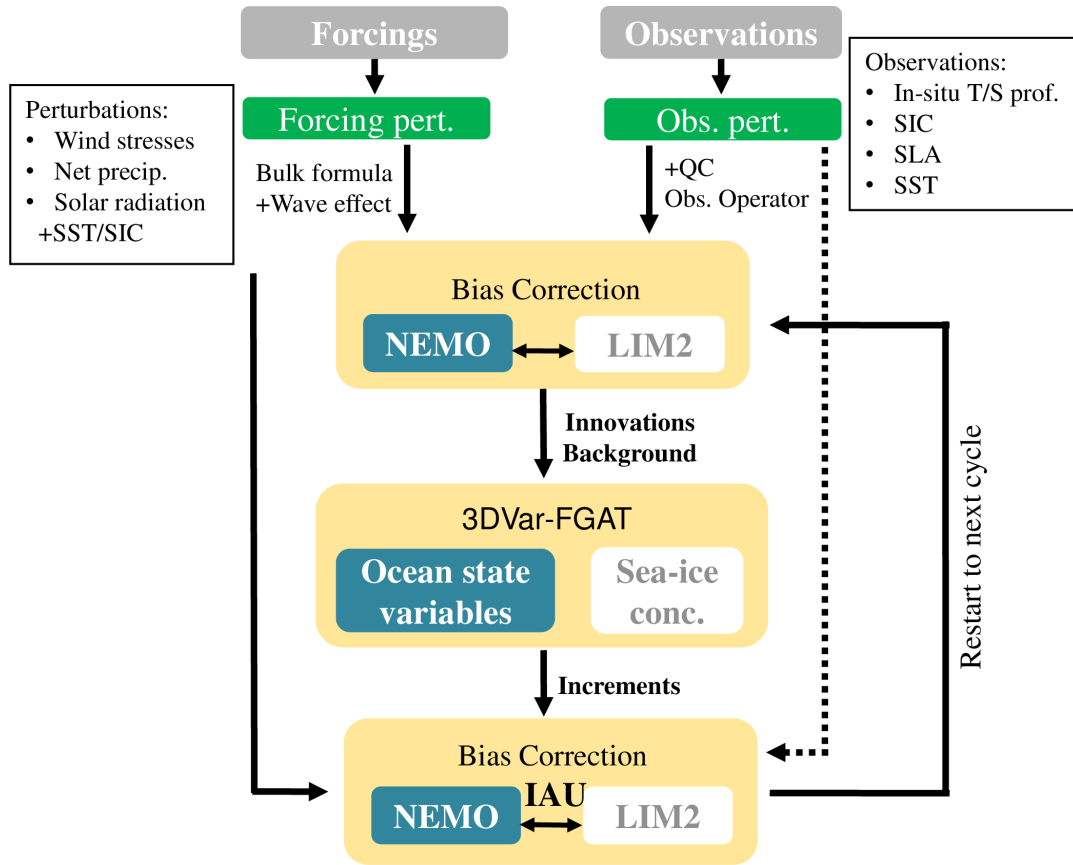
Figure 1. Schematic diagram of the ORAS5 system

A schematic diagram of the ORAS5 system can be found in Fig 1. The analysis cycle consists of one outer iteration of 3D-Var FGAT with observational QC and bias correction steps. In the first step (also called the 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 quality-controlled observations and model background state are passed to the so-called inner loop, where the 3D-Var FGAT method minimizes the linearized cost function to produce the assimilation increment. The increment is applied during a second forward integration of the model (the second outer loop) using the incremental analysis updates method with constant weights (IAU; Bloom et al. (1996)). Both SIC and other observations are assimilated using a 5-day assimilation cycle in ORAS5 and share the outer loop model integrations.

As in ORAP5, assimilation of SIC data is also included in ORAS5. The background state of ocean and sea ice is produced from a coupled NEMO-LIM2 run, but the minimization of the SIC cost function is separated from the minimization of the cost function for all other ocean state variables. The separation of the sea-ice minimization assumes that there are no covariances

between SIC 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 ORAS5 balance relations are the same as for ORAS4 (Mogensen et al., 2012) and ORAP5. The observation and background errors specifications are the same as in ORAP5 (Zuo et al., 2015),

5   except for sea level (see Section 2.4).

2.2   Model initialization and forcing fields

2.2.1   Initialization

As for the previous ocean reanalysis system ORAS4, perturbing the ocean initial conditions at the beginning of the reanalysis period is considered paramount. In ORAS4 different initial states in 1958 were given by sampling a 20 year ocean integration.

10   ORAS5 had a longer spin-up using reanalyses for the period 1958-1979, conducted using either ERA40 (Uppala et al., 2005) or ERA20C (Poli et al., 2016) forcing and assimilating in-situ data. ORAS5 starts in 1979, so it is in principle possible to have initial conditions representative of that given date. A series of ocean reanalyses assimilating in-situ profiles using different surface forcing, data sets and parameters was conducted from the period 1958 to 1975 (Table 2), as an attempt to account for the uncertainty of ocean state at a given point in time. This approach gives a set of 5 initial conditions (INI1-5) to start each

15   of the ensemble member of ORAS5, thus generating the ORAS5 initial perturbations. The control member of ORAS5 was initialized from INI1 with a similar configuration to ORAP5, and is unperturbed: neither the forcing fields nor the observations perturbations are applied (see Section 2.5 for details). A second spin-up from 1975 to 1979 was then conducted with the same settings as used for ORAS5, and the integrations are then continued after 1979. The impact of the initial perturbations is illustrated in Fig. 2, which shows the evolution of the global ocean heat content (OHC) from the 5 spin-up ocean reanalyses

20   listed in Table 2, and ORAS5 with its 5 ensemble members.

The initial uncertainty of ORAS5 OHC is illustrated by OHC spread (here we define the spread as the maximum value minus minimum value in OHC, taking into account all ORAS5 ensemble members at a given time) in Fig. 2. The initial spread inherited from the 5 spin-up remains high especially for the first 5 years between 1975-1979. There is a constant reduction of OHC for all members during 1975-1982, with rapid cooling for two warm members initialized from INI4 and INI5. This

25   OHC spread reduces gradually and reaches a relatively stable state after 2000, suggesting a robust uncertainty maintained by the other components of the perturbation scheme (See Section 2.5).

2.2.2   Forcing, SST and SIC

Forcing fields for ORAS5 are derived from the atmospheric reanalysis ERA-Interim (Dee et al., 2011) until 2015, and from the ECMWF operational NWP thereafter (see Fig. 3), using revised CORE bulk formulas (Large and Yeager, 2009) that include

30   the impact of surface waves on the exchange of momentum and turbulent kinetic energy (Breivik et al., 2015). Compared to ORAP5, the wind enhanced mixing due to surface waves is updated with a revised spatial distribution scheme. In addition, observational data of SST, sea surface salinity (SSS), global mean sea level trends and climatological variations of the ocean mass
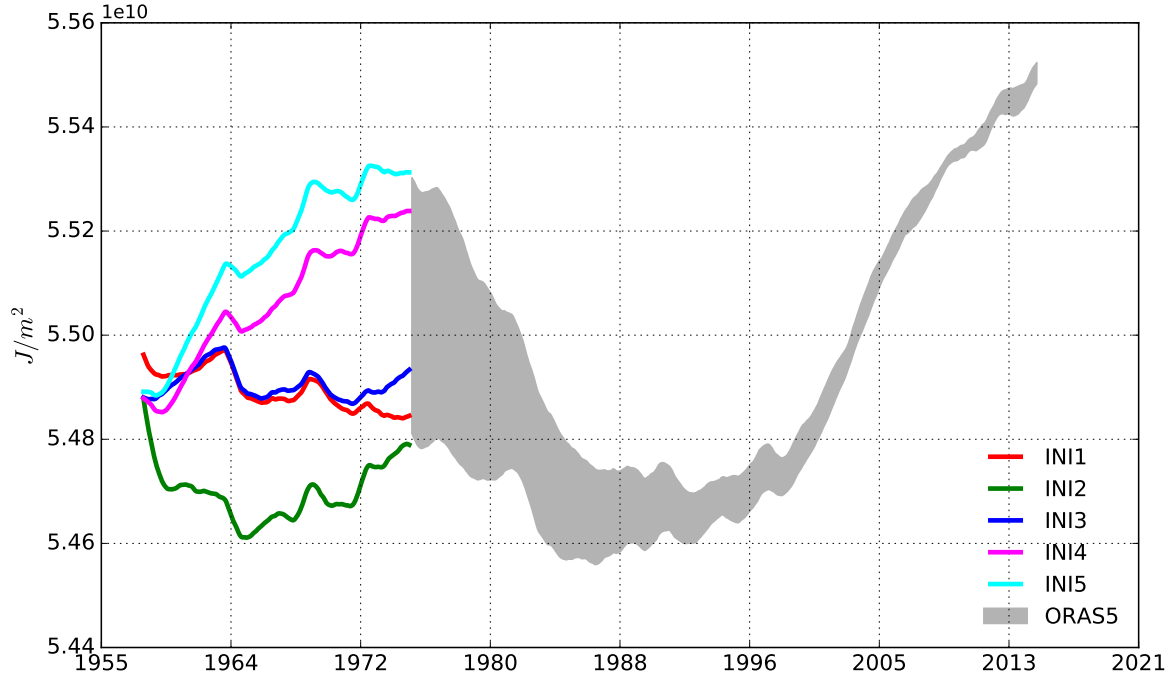
Figure 2. Time series of global ocean heat content (in $10^{10}\,\mathrm{Jm^{-2}}$) integrated for the whole water column, from 5 spin-up runs (INI1-5, 1958-1974) and ORAS5 from 1975 onwards. The shaded areas encompass the spread of all ORAS5 ensemble members. A 12-month running mean has been applied.

are used to modify the surface fluxes of heat and freshwater. Readers, however, should note that ORAS5 will be re-processed with ERA-Interim forcing and reprocessed observation data set (e.g. EN4 (Good et al., 2013)) from 2015 onwards. This re-processed ORAS5 product will be extended annually with consistent forcing and observation data set whenever possible. This should produce consistent time series that are suitable for climate monitoring applications beyond 2014. The reprocessed
5   ORAS5 will be available as part of the ensemble of global reanalyses distributed by the Copernicus Marine Environmental Monitoring Services (CMEMS).

SST is assimilated in ORAS5 by modifying the surface non-solar total heat flux using the product of a globally uniform restoration term of $-200\,\mathrm{Wm^{-2}K^{-1}}$ and the difference between modelled and observed SST (see Haney (1971)). The effect of this restoration can be illustrated as follows: assuming a constant mismatch to observations of 1 K within a well mixed upper
10   50 m layer of water, the relaxation term will restore the water temperature in this mixed layer by 1 K in about 12 days. The numerical value is unchanged from previous ECMWF ocean reanalyses-ORAS4; the original choice was motivated to keep SST errors within 0.2 K in the global ocean. The same value is used in other ocean reanalysis systems with similar horizontal resolution as ORAS5 (Masina et al., 2017). However, given that ORAS5 has finer vertical resolution, this term may need

Table 2. Ensemble of ORAS5 initial conditions

Name	Year of initialization	Forcing	SST/SIC	In-situ	Bias capping	ϕ_c
INI1	ORAP5-1990	ERA40	ERA40	EN3	No	$10°$
INI2	ORAP5-1980	ERA20C	HadISST2	EN4	No	$10°$
INI3	INI1-1970	ERA40	ERA40	EN4	Yes	$10°$
INI4	INI1-1970	ERA40	HadISST2	EN4	Yes	$2°$
INI5	INI1-1970	ERA40	HadISST2	EN4	No	$10°$

Bias capping is the switch to cap the minimum value of salinity bias correction term to prevent static instability, see Section 2.3.4.

ϕ_c is a constant value of latitudinal bands (in degrees) used to define a reduction coefficient for the pressure gradient component of bias correction, see Eq. 6 and 7 in Zuo et al. (2015).

All spin-ups are carried out in ORCA025.L75 configuration.

revision. Besides, it has also been found that ocean circulation in climate model is sensitive to the strength of SST restoration (Servonnat et al., 2014). More discussion of SST nudging and associated impact on ocean state can be found in Section 5.1. A similar global uniformed SSS restoration term of -33.3 mm/day to climatology has been applied by adding a term to the surface freshwater fluxes equation. This is equivalent to a restoration time-scale of about 1 year for a well mixed upper 10 m layer of water with a mean model surface salinity of 35 psu.

Temporal consistency in the SST analysis product employed is important for both ocean and atmospheric reanalysis. Hirahara et al. (2016) found that the OSTIA SST reanalysis product has a noticeably different global mean with respect to its homonymous real-time product; they recommended to use SST from Titchner and Rayner (2014) in combination with the real-time OSTIA for production of the atmospheric reanalysis ERA5. HadISST2.1 is a new pentad SST product with a spatial resolution of $0.25°$ resulting from the EU FP7 project ERA-CLIM2. The bias correction and data homogenization in this product is superior to its predecessor HadSST3 (Kennedy et al., 2011a, b), and more importantly, the resulting SST are consistent with those delivered operationally by OSTIA (Donlon et al., 2012). ORAS5 has adopted the same SST as ERA5. Therefore, SST in ORAS5 prior to 2008 comes from HadISST2.1, and from operational OSTIA thereafter.

The SIC data assimilated in ORAS5 comes from the OSTIA reanalysis before 2008. This is the same as in ORAP5. Sea-ice data in HadISST2.1 includes both re-processed sea-ice concentration data from the EUMETSAT Ocean and Sea Ice Satellite Application Facilities (OSI-SAF) and polar ice charts data from National Ice Center (NIC). SIC in HadISST2.1 is calibrated against NIC sea-ice charts in order to ensure consistency with chart analyses prior to the satellite era. However, sea-ice concentration in sea-ice charts has large uncertainties itself (Karvonen et al., 2015). Moreover, some sea-ice charts are biased towards

Table 3. Sensitivity experiments to inform the choice of SST and SIC observation data sets

Name	SST	SIC
ASM-HadI	HadISST2	HadISST2
ASM-OST	OSTIA	OSTIA
ASM-HadI-OST	HadISST2	OSTIA

All experiments are carried out at ORCA1.L42 resolution and in OP5-LR configuration.

ERA40 and Reynolds OIv2d data was used before 1985 when OSTIA product is not available.

high SIC. As a result, sea-ice concentration in the HadISST2.1 data is substantially higher than in the OSI-SAF data (Titchner and Rayner, 2014) and OSTIA analysis .
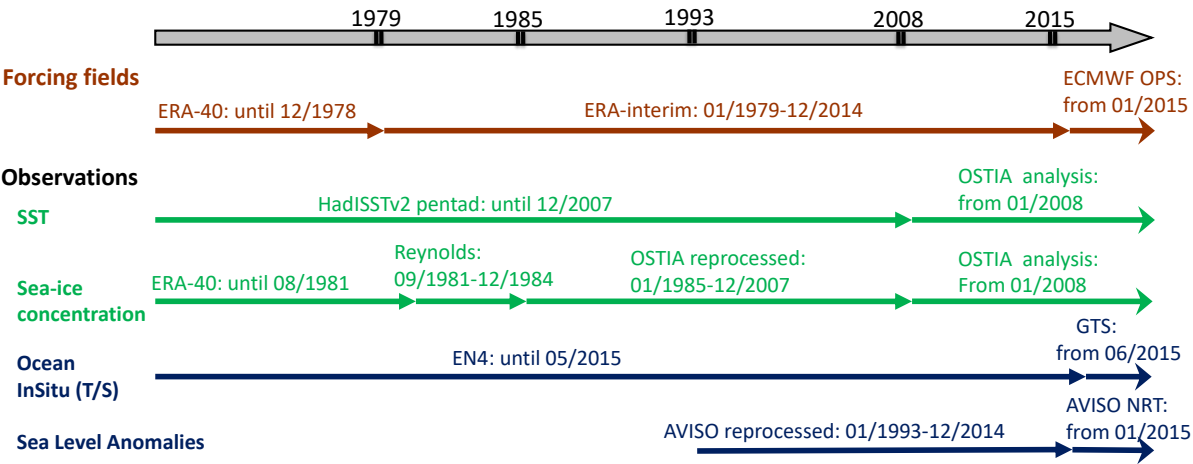
Figure 3. Time line of changes to the reanalysis forcing and assimilation data sets for ORAS5

In order to assess the impact of assimilating different SST and SIC products in our system, sensitivity experiments have been carried out at ORCA1.L42 resolution (approximately 1° at tropics with 42 vertical levels) with ORAP5-equivalent Low-Resolution configuration (hereafter referred to as OP5-LR). SST and SIC data used in these experiments are listed in Table 3, together with the experiment names. Global mean SST from these experiments are shown in Fig. 4, together with the SST analysis products that were assimilated. For verification, the latest European Space Agency Surface Temperature Climate Change Initiative (ESA SST CCI) multi-year SST record (Merchant et al., 2014) (version 1.1) is also included here as a reference. This data set is generated from satellite observations only and is independent from in situ observations.

Despite the discrepancy in the early period, HadISST2 and OSTIA SST analyses are very similar after 2008, suggesting that HadISST2 is more consistent with the operational OSTIA SST product than the OSTIA reanalysis SST itself, as already pointed out by Hirahara et al. (2016). OSTIA reanalysis SST is systematically colder than both HadISST2 and ESA CCI SST before 2008, by approximately 0.1°C and 0.16°C in the global mean, respectively. Unlike HadISST2 and OSTIA, both of which define SST as the night-time temperature, ESA CCI SST are defined as the daily-mean temperature at 0.2 m depth, and thus provides the warmest SST among these three products. Time-series of global mean SST from ASM-HadI and ASM-HadI-OST are almost indistinguishable from each other, or from HadISST2 itself. ASM-OST, on the other hand, generates a global mean SST which lies in-between OSTIA reanalysis and HadISST2 SST. This result indicates that assimilated near-surface in-situ observations agree better with HadISST2 SST than with OSTIA SST, thus pull the analysed SST towards the warmer side. This lack of consistency between near-surface in-situ observations and OSTIA reanalysis, and between operational OSTIA SST and OSTIA reanalysis, determined the final choice of SST product for ORAS5.

The above experiments were also used to inform the choice of the SIC data set. Departure of sea-ice thickness (SIT) from the three sensitivity experiments (Table 3) against laser altimeter freeboard measurements from ICESat (Kwok et al., 2009) (data downloaded from http://nsidc.org/data/nsidc-0393) for October 2007 are shown in Fig. 5. Among the three, ASM-HadI-OST clearly shows the smallest SIT discrepancy, especially for the thick ice in the Beaufort Gyre and at the north coast of Greenland and the Canadian Archipelago. Assimilating HadISST2 SIC data results in profoundly overestimated SIT in ASM-HadI as verified against ICESat observations. This is mainly due to assimilation of HadISST2 SIC that is in general higher than those of Reynolds/OSTIA data. In fact, assimilation of HadISST2 SIC during 1979-1984 implies strong positive sea-ice volume increments with respect to ERA40/Reynolds data, which are equivalent to adding approximately 3 meters of SIT per year in most of the Arctic basin during this period (not shown). This effect has also been discussed by Tietsche et al. (2013) in their sea-ice assimilation experiments. As a result, ASM-HadI exhibits unrealistic sea-ice conditions in both the Arctic and the Antarctic (not shown). Therefore, we chose to use the OSTIA reanalysis SIC in ORAS5 until 2008, together with SST observation from HadISST2.
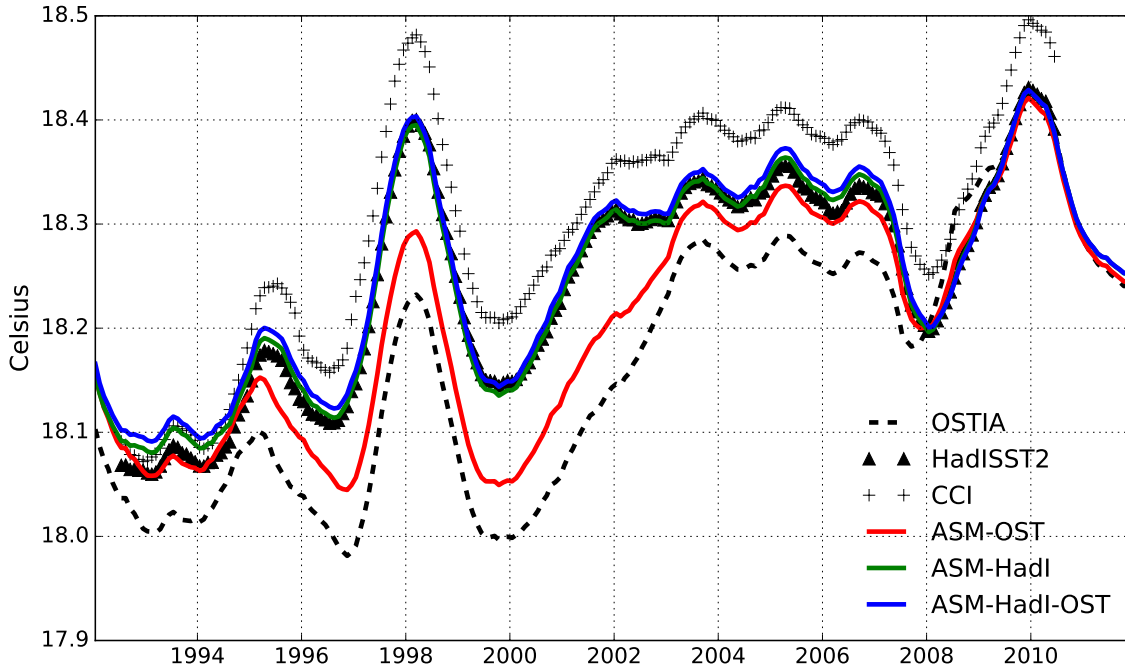
Figure 4. Time series of global mean SST (°C) from ocean reanalyses when assimilating different SST and SIC analysis products. A 12-month running mean filter has been applied.

2.3 Assimilation of in-situ observations

2.3.1 In-situ observation data set

The in-situ temperature and salinity (T/S) profiles in ORAS5 come from the recently released quality controlled data set EN4 (Good et al., 2013) with Expendable BathyThermograph (XBT) and Mechanical bathythermograph (MBT) depth corrections from Gouretski and Reseghetti (2010) until May 2015. EN4 is a re-processed observational data set with globally quality-controlled ocean T/S profiles. It includes all conventional oceanic observations (Argo, XBT/MBT, Conductivity-Temperature-Depth (CTD), moored buoys, ship and mammal-based measurements). Data from the Arctic Synoptic Basin Wide Oceanography (ASBO) project was also included in EN4, therefore improves data coverage in the Arctic. Compared to its predecessor EN3 (used in ORAS4 and ORAP5), EN4 has increased vertical resolution, improved QC and duplication check, and extends farther back in time. For the latest years, EN4 also contains a more complete and cleaned record of the Argo data, with bias-corrected data whenever possible. After May 2015, ORAS5 starts using the operational data from the Global Telecommunications System (GTS), which consists of data received in near-real-time at ECMWF. The same quality control procedures as

a) ASM-OST **b) ASM-HadI** **c) ASM-HadI-OST**
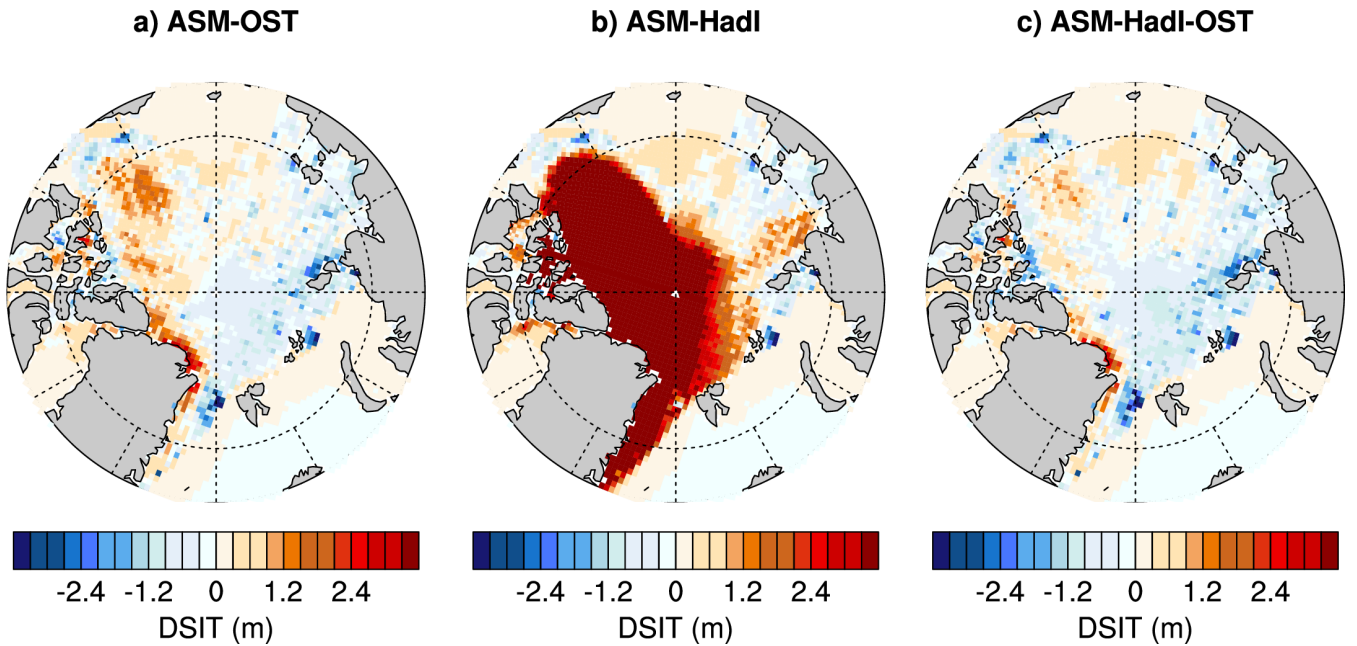
Figure 5. Departure of sea-ice thickness (DSIT) in meters for (a) ASM-OST, (b) ASM-HadI and (c) ASM-HadI-OST. The departure is computed with respect to ICESat observations for October 2007.

described in Section 2.3.3 are applied to all GTS data, to ensure that only good quality observations similar to EN4 data are assimilated in ORAS5.

The new EN4 data set has been evaluated against the EN3 data set using twin experiments carried out in the OP5-LR configuration at ORCA1.L42 resolution. Twin experiments comprise a reference run EXP3 that assimilates EN3 data, and another run EXP4 that assimilates EN4 data but are otherwise identical. For verification purpose, a group of CTD mooring arrays in the Barents Sea was withdrawn from data assimilation in either EXP3 or EXP4. Mean bias and root-mean-square departure of model background with respect to these CTD moorings are shown in Fig. 6 for both experiments. The EXP4 has reduced temperature and salinity RMS errors in the Barents Sea. This better estimation of mean ocean state in EXP4 can be attributed to an improved observation coverage of EN4. After 2005, the Arctic ocean observation almost doubled in EN4 with respect to EN3. As a results, EXP4 also show freshening (up to 0.2 psu) near the Greenland coast, at the edge of East Siberian Sea and across the Baffin Bay, which are directly related to discrepancies between the EN3 and EN4 data sets (not shown).

2.3.2 Observing system experiments

Observing System Experiments (OSEs) are widely used as a method to evaluate the impact of existing observations, and are routinely carried out at ECMWF for assessment of previous operational ocean reanalysis systems and seasonal forecast (Balmaseda and Anderson, 2009). To understand the impact of individual observation types in EN4, a series of OSEs have been carried out using the OP5-LR configuration at ORCA1.L42 resolution, except that bias correction and SLA assimilation were
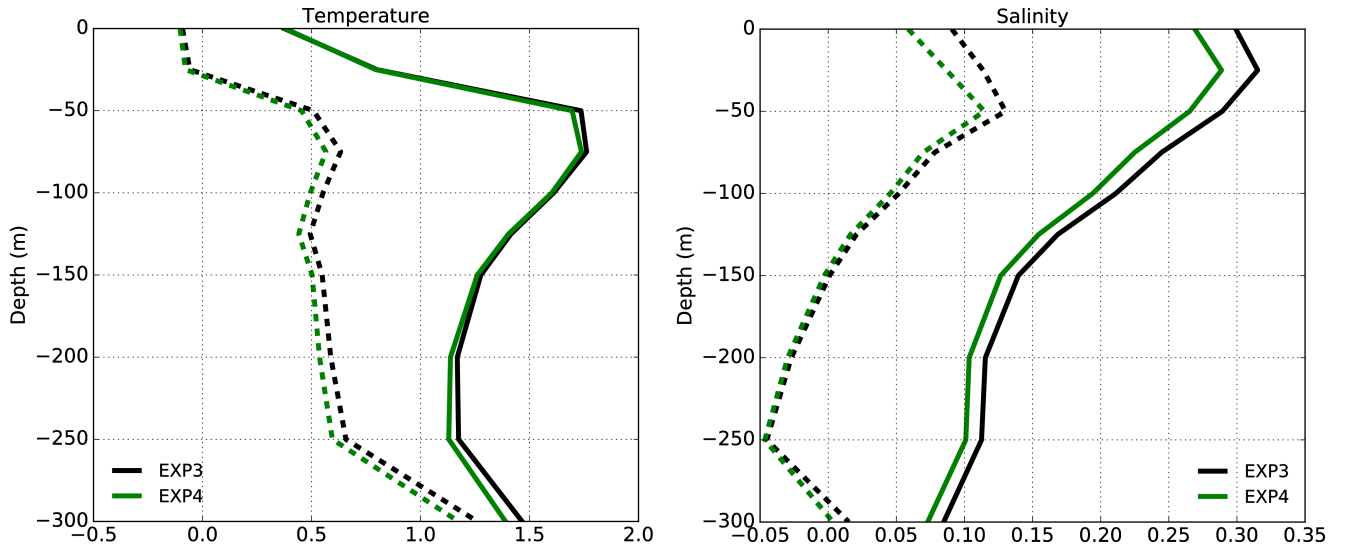
Figure 6. Profiles of model bias (dashed lines) and RMSE (solid lines) for (left) temperature (K) and (right) salinity (psu) for the upper 300 m. Statistics are calculated using the misfit of the model background value from (black) EXP3 and (red) EXP4 with respect to CTD profiles in the Barents Sea and for September 2009.

switched off. First, a reference experiment (ORA-ALL) has been carried out by assimilating all in-situ observations from the quality-controlled EN4 data set. Four OSE-ORAs experiments were then carried out based on ORA-ALL, by withdrawing individual in-situ observation types from the global data assimilation system: 1) NoArgo – removing Argo floats; 2) NoMooring - removing Moored buoys data; 3) NoShip – removing XBT, MBT and CTD data; 4) NoInsitu – removing all in-situ
5   observations. All OSE-ORAs have been driven by the same forcing from ERA-Interim.

   To illustrate impacts from withholding different observation types from the global ocean observing system (GOOS), maps of normalized RMS departure (RMSD) of upper 700 m temperature inter-annual anomalies between these four OSE-ORAs and the OSE-ALL are shown in Fig. 7. Diagnostics were computed over the 2005-2015 period, when Argo floats reaches a relatively homogeneous global coverage. Results suggest that removal of moored buoy data mostly affects the tropical regions (Fig. 7a),
10 with visible increased RMSD at locations of global tropical moored buoy array: that is the Tropical Atmospheric Ocean (TAO), Triangle Trans-Ocean Buoy Network (TRITON), Prediction and Research Moored Array in the Atlantic (PIRATA) and Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA). The degradation resulting from the removal of PIRATA is slightly larger than that coming from TAO/TRITON and RAMA. This can be attributed to a more realistic ocean state in the tropical Pacific and Indian oceans constrained by surface observations (SST) and forcings
15 (winds and surface fluxes) in our system, but is also likely to associated with the drastic reduction in the observation number from TAO/TRITON since 2012. Removal of moored buoys data also shows some remote effects in the North Atlantic Ocean, i.e. in some eddy-dominated regions with large uncertainties in the ocean reanalyses.
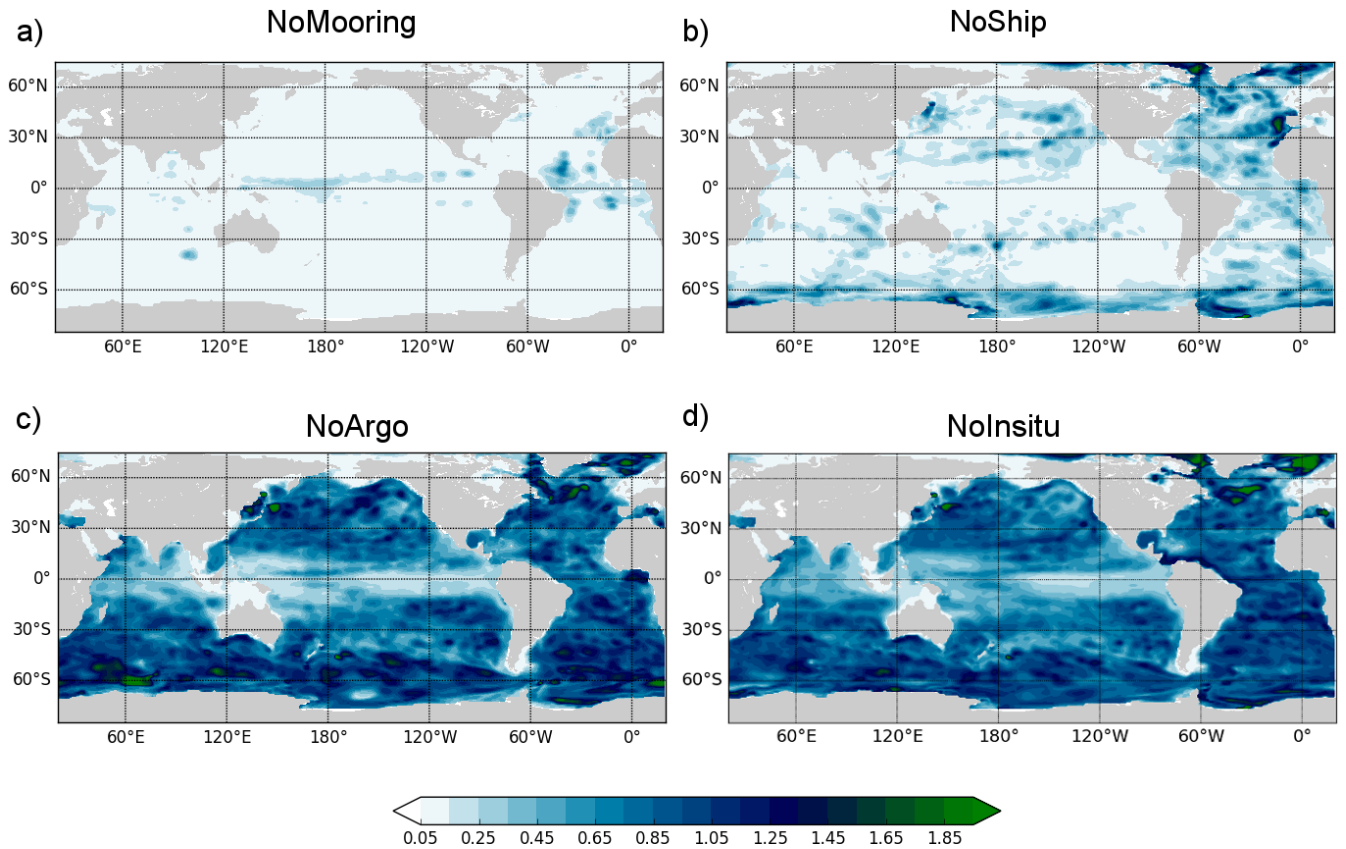
Figure 7. Maps of normalized RMSD of upper 700 m column-averaged temperature between the OSE-ALL and (a) NoMooring, (b) NoShip, (c) NoArgo, and (d) NoInsitu. Statistics are computed using monthly-mean anomaly data over the 2005-2015 period after removal of the seasonal cycle information, then normalized against the temporal standard deviation of temperature in OSE-ALL over the same period.

Ship-based observations (Fig. 7b) have visible impact along most frequent commercial shipping routes carried out by voluntary observing ships and ships of opportunity, but also show important contribution in the high-latitude through dedicated scientific campaigns, where Argo floats normally are not available. Removal of Argo floats (Fig. 7c) degrades the ocean state almost everywhere except for the tropical Pacific and Indian oceans, again due to an already well constrained ocean state from surface in these regions.

Removal of all ocean in-situ observations (Fig. 7d) gives an estimation about the total impact of GOOS, which is not a simple linear combination of individual observation type. Note that in the Southern Ocean the RMSD is sometimes larger in NoArgo than in NoInsitu, which indicate some inadequacy of the data assimilation process. Overall, the weak impact of removal of observations in the Indian Ocean is possibly related to the comparatively sparse observing system in that region. Generally, the tropical Atlantic seems to be more sensitive to the removal of in-situ observations than the other tropical ocean basins.

2.3.3 Quality control of in-situ data

All input observation are subject to a global quality control procedures similar to those employed in EN4. Among these are checks on duplication, background, stability, bathymetry, and using the Argo grey list (from ftp://ftp.ifremer.fr/ifremer/argo/etc/ar_grey-list/). In addition, a new temperature–salinity pair check has been introduced in ORAS5, in which salinity observation will be rejected whenever the corresponding temperature observation at the same location is not available. This pair check has been designed to avoid assimilating salinity observation alone, considering that temperature is the primary variable in the multivariate balance operator (Weaver et al., 2005) of NEMOVAR. This implementation has been tested using twin experiments in the OP5-LR configuration. The twin experiments comprise a reference experiment without the new T/S pair check (PC-OFF), and an otherwise identical experiment except that uses the pair check (PC-ON). Figures 8a,b highlight an inverse temperature bias pattern in the eastern North Atlantic Ocean in PC-OFF, with cold bias up to 0.8 K at 1500 m and a warm bias of ~1 K at 2000 m. This error pattern is also visible in the previous ECMWF ocean reanalyses (ORAS4 and ORAP5) and is associated with spurious vertical convection following the Mediterranean outflow waters. This was improved in PC-ON as shown in Figures 8c,d with a small compensating temperature difference (~0.3 K) defined as PC-ON minus PC-OFF, which also leads to reduced RMSE in PC-ON (not shown) between 1000–2000 m. This new pair-check mostly affected the North Atlantic Ocean between 1000–2000 m and rejected ~3% of salinity observations in this region.

2.3.4 Bias correction scheme

Model bias correction is essential for the ocean data assimilation system, especially for dealing with irregular and inhomogeneous ocean observations. A similar multi-scale bias correction scheme as described in Balmaseda et al. (2013a) has been implemented in ORAS5 to correct temperature/salinity biases in the extra-tropical regions. A pressure correction for the tropical regions has been implemented as well in this bias correction scheme. This is an important method for mitigation of suspicious climate signals that could be introduced due to the assimilation of an evolving observation network. Compared to ORAP5, the ORAS5 bias correction scheme includes two major upgrades. First, the a-priori bias term (offline bias) in ORAS5 has been estimated using an ensemble of five realizations of assimilation runs (only temperature and salinity) during the Argo era (2003–2012) with different forcing and model parameters (See Table. 4). The sampling period starts a few years after the Argo floats, when a relatively homogeneous global ocean observing network becomes available. The equivalent term in ORAP5 was estimated from a single realization of reanalysis from a shorter period (2000–2009). The ensemble approach allows uncertainties of model errors to be estimated, and could provide, in some regions, a more robust estimation of the systematic model error. In ORAS5 only the ensemble mean of a-priori biases estimated from these five realizations (BIAS1-5) was used in order to account for seasonal variations of the model/forcing errors.

To help readers understanding about relative contributions of offline bias correction in different systems, Fig. 9 shows the mean vertical profiles of the a-priori bias correction applied to temperature and salinity in ORAS5 and two previous ECMWF ocean reanalyses (ORAS4 and ORAP5). It is worth noting that the value shown in Fig. 9 has been added in the reanalysis system to correct model background errors, therefore it is opposite to model biases. In general, the two high-resolution reanalyses
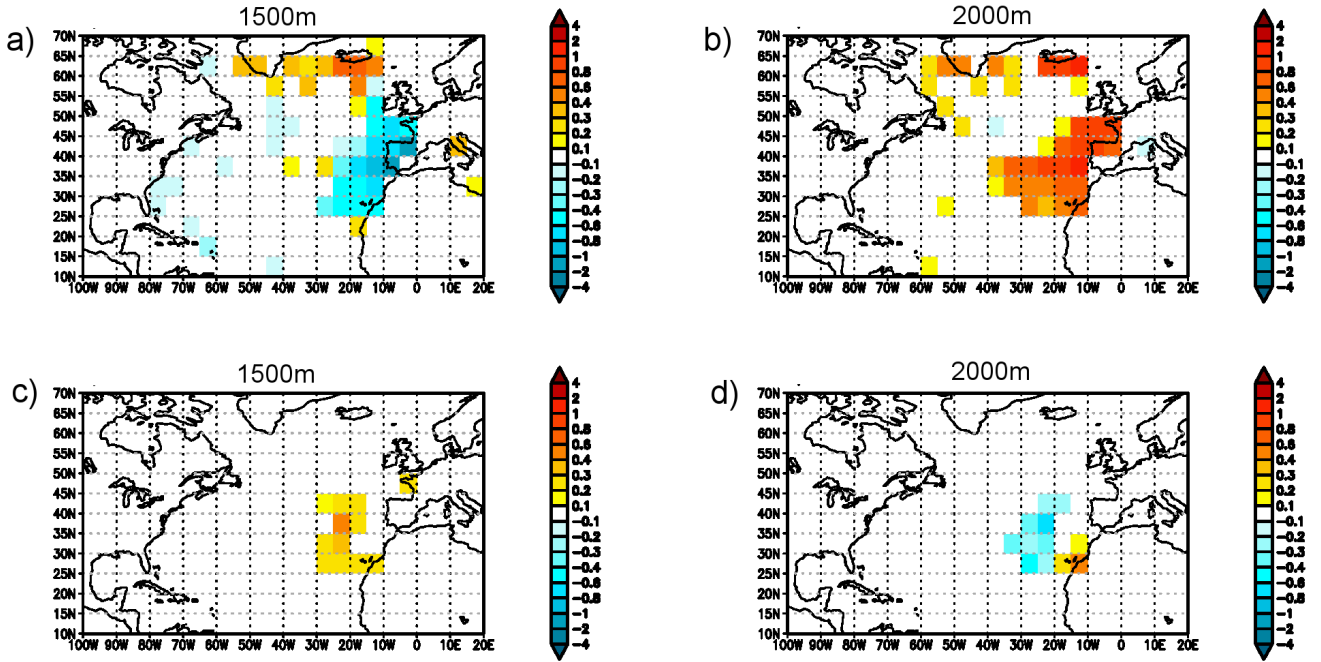
Figure 8. (a,b) PC-OFF mean temperature biases (K) with respect to observations at (a) 1500 m and (b) 2000 m; (c,d) PC-ON temperature departures (K) with respect to PC-OFF at (c) 1500 m and (d) 2000 m. Statistics are computed based on June data over the period 2005–2010, after binning and averaging the observation-space departures over $5° \times 5°$ latitude/longitude boxes.

Table 4. Summary of ORAS5 off-line bias correction ensemble estimations

Name	SST	SIC	H. Thin. Dist.	σ_T^{do}
BIAS1	HadISST2	HadISST2	100 km	0.07
BIAS2	HadISST2	HadISST2	100 km	0.07
BIAS3	OSTIA	OSTIA	25 km	0.07
BIAS4	OSTIA	OSTIA	25 km	0.098
BIAS5	HadISST2	OSTIA	25 km	0.098

σ_T^{do} *is the minimum temperature observation error standard deviation at deep ocean, see Zuo et al. (2015).*

H. Thin. Dist. is the length scale for horizontal thinning of in-situ observations

All BIAS runs are carried out in ORCA025.L75 configuration, assimilate EN4 data set but without SLA assimilation.

(ORAP5 and ORAS5) have opposite and weaker temperature biases to ORAS4. Considering that all three reanalyses use the same ERA-Interim forcing, the different sign of bias terms is likely a result of model physics/resolution rather than forcing. However, both the SST observational data set and the surface flux formulation have changed substantially between ORAS4 and ORAS5, and therefore, the effect of surface fluxes and SST cannot be neglected. Compared to ORAP5, ORAS5 has slightly

5 increased cold bias around 100 m, but with reduced cold bias below 200 m. All three reanalyses show fresh biases in salinity for the upper 100 m, with ORAS5 bias stays in-between ORAP5 and ORAS4. The same offline bias correction terms in maps are shown in Fig. 10 for ORAP5 and ORAS5. Both ORAP5 and ORAS5 show very similar spatial patterns in temperature and salinity biases, suggesting common model or forcing errors. However, temperature bias in ORAS5 are clearly weaker than in ORAP5 between 300–700 m, especially for the Tropics. On the contrary, Upper 100 m salinity bias in ORAS5 is larger

10 than ORAP5 almost everywhere. This bias term is the systematic model/forcing errors estimated using in-situ observations, therefore the result is subject to the temporal and spatial coverage of global ocean observing system. The differences between ORAS5 and ORAP5 as seen in Fig. 9 and Fig. 10 are results from (a) improved temporal and spatial coverage in the new EN4 data set with increased vertical resolution; (b) a different climatological period used for ORAS5 bias estimation; and (c) the ensemble bias estimation method used in ORAS5.
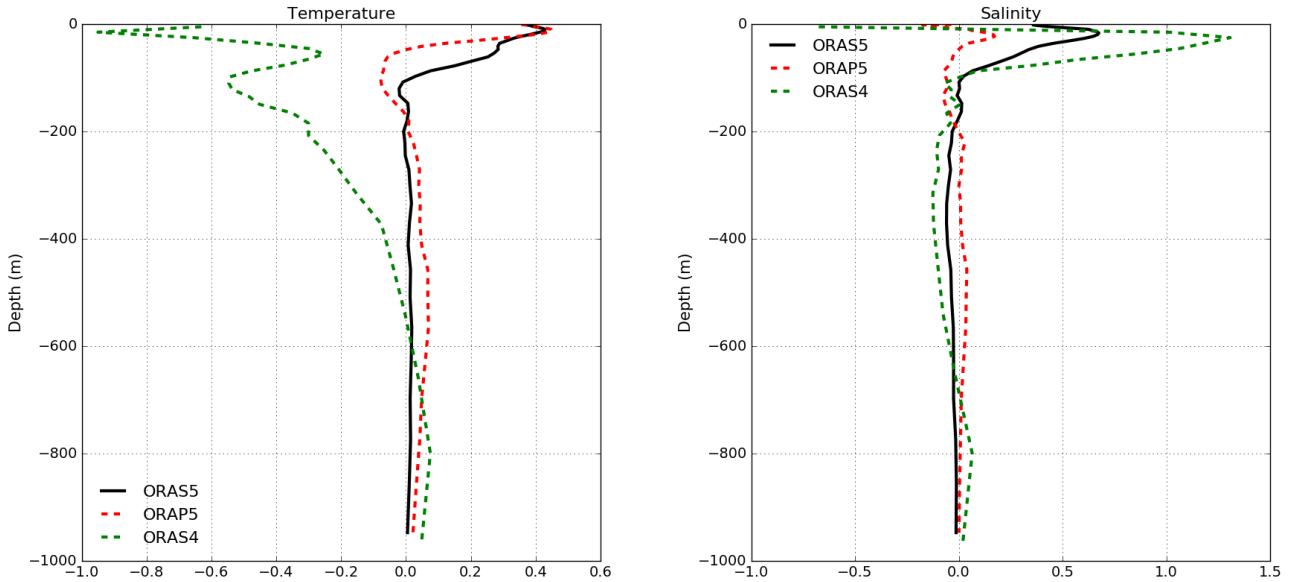
Figure 9. Vertical profile of global mean a-priori bias corrections applied to (left) Temperature (units are 0.01 K per 10 days) and (right) Salinity (units are 0.001 psu per 10 days) for ORAS5 (black solid), ORAP5 (red dashed) and ORAS4 (green dashed).

15 Furthermore, a stability check was introduced in the ORAS5 bias correction that caps the minimum value of salinity bias correction term to prevent static instability. We define a minimum value for the squared buoyancy frequency as N^2_{min}. In
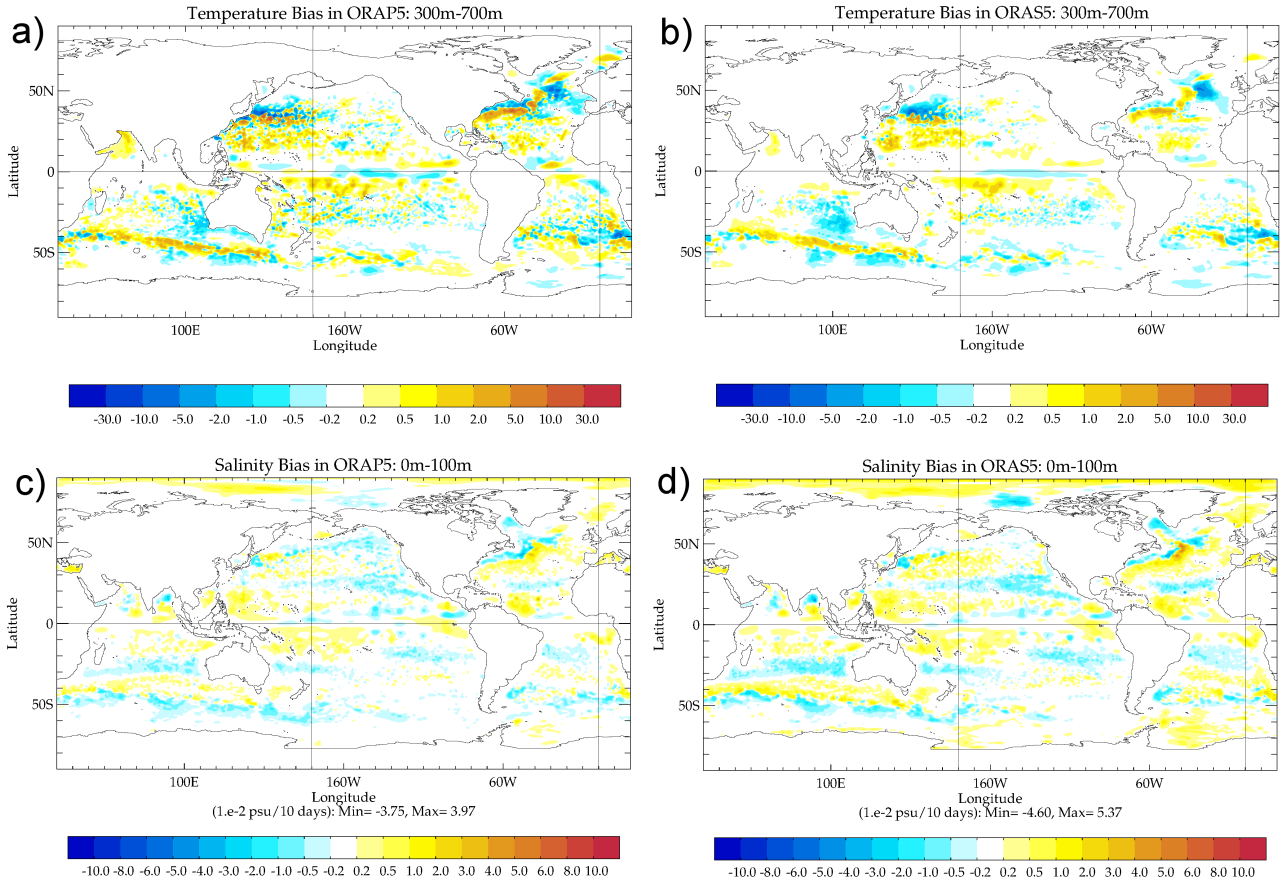
Figure 10. Maps of annual mean a-priori bias correction term applied to (a,c) ORAP5 and (b,d) ORAS5 as (a,b) temperature (units are 0.01 K per 10 days) and (c,d) salinity (units are 0.01 psu per 10 days). The reader should note that temperature bias is averaged over 300–700 m, and salinity bias is averaged over 0–100 m.

every model grid cell where N^2 as defined by the model background potential density profile (ρ_σ) is close to static instability ($N^2 <= N^2_{min}$, $N^2_{min} = 1e^{-10}$), we modify the salinity bias to ensure that δN^2 due to total bias (both temperature and salinity) is 0. In this way, the salinity bias correction is prevented from introducing instability in the water column, which could otherwise induce spurious vertical convection, thought to be the cause of large reanalysis biases in regions around the Mediterranean

5   outflow waters in the Northern Atlantic Ocean (Zuo et al., 2017b). Results of model fit-to-observation errors from a set of twin assimilation experiments testing the impact of the bias capping can be found in Fig. 11. The twin experiments were set up in the OP5-LR configuration – but assimilating EN4 data set instead of EN3. The reference run (NoCap) does not activate salinity bias capping, while the other run (CP10) adds salinity bias capping and has otherwise exactly the same configuration. Both temperature and salinity RMSE profiles of NoCap show a local maximum at 1000 m, which is associated with the spurious

convection between 1000 and 2000 m due to warm and salty Mediterranean outflow. The new salinity bias capping in CP10 successfully reduces bias and RMSE for both temperature and salinity at this depth range. As a result, CP10 also exhibits improved sea level correlation with altimeter data compared to NoCap (not shown). Further assessment of this bias correction method with respect to in-situ observations can be found in Section 4.2.
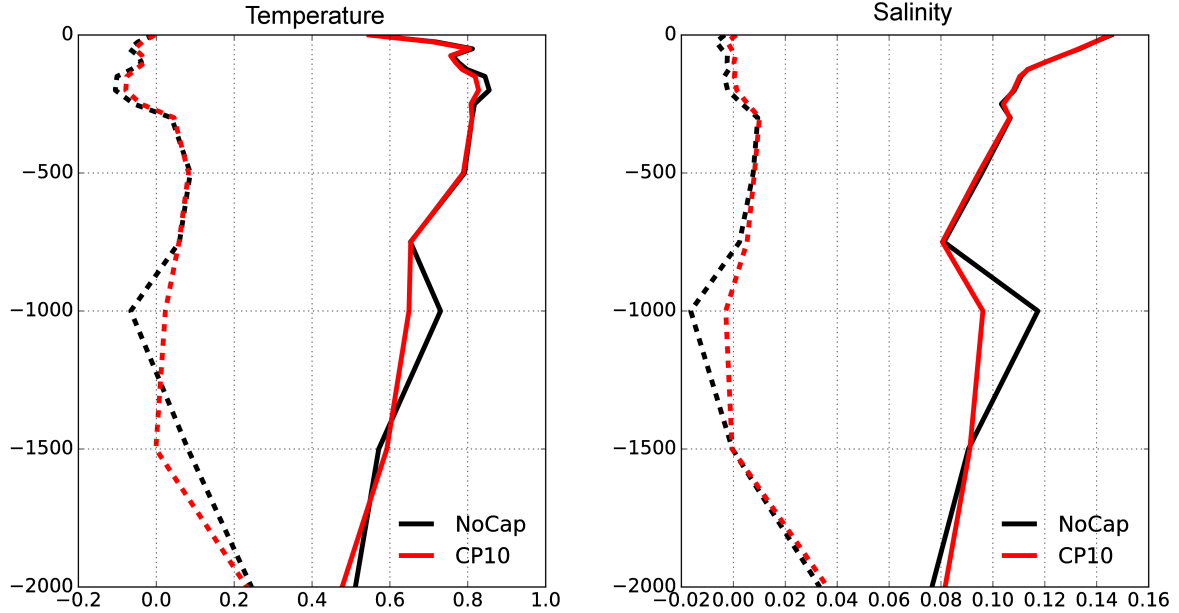
Figure 11. Profiles of model mean bias (dashed lines) and RMSE (solid lines) for (left) temperature in K and (right) salinity in psu. Statistics are calculated using the model background value from NoCap (in black) and CP10 (in red), with respect to the quality-controlled EN4 data set, after averaging over the 1996-2011 period and the eastern North Atlantic Ocean.

2.4 Assimilation of satellite altimeter sea-level anomalies

The sea-level anomaly (SLA) observations produced by AVISO (Archiving Validation and Interpretation of Satellite Oceanographic data) DUACS (Data Unification and Altimeter Combination System) has been updated to the latest version DT2014 (Pujol et al., 2016) in ORAS5 for both filtered along-track and gridded SLA data. Compared to the previous version DT2010 (Dibarboure et al., 2011) that has been used in ORAS4 and ORAP5 reanalyses, the DT2014 data set has received a series of major upgrades, including a new 20-year altimeter reference period (1993–2012) and increased spatial resolution (14 km in low latitudes), among others. Another important change in ORAS5 w.r.t. ORAS4/ORAP5 is that SLA thinning is now done by stratified random sampling (Zuo et al., 2017a) instead of creating superobbing SLA observations, as a method to account for observation representativeness errors from along-track SLA data. As a result, ORAS5 ingests SLA observations with increased

local variability but reduced observation error standard deviations (OBE STD). Compared to ORAS4, the SLA OBE STD in ORAS5 is reduced by approximately 20% in the Tropics due to increased spatial resolution of DT2014 data set. ORAS5 also assimilates more along-track SLA data whenever newly available satellite missions (i.e. GeoSat Follow-On, HaiYang-2A, Topex New, Jason-1 Geodetic, Jason-1 New, Saral/AltiKa) are available in DT2014. Other parts of the scheme, e.g. a reduced-

5   grid construction (typical 1° by 1° in latitude/longitude) and a method for diagnosing OBE STD (Mogensen et al., 2012), remain unchanged. SLA observation has not been assimilated in ORAS5 outside the latitudinal band from 50°S to 50°N, nor in regions shallower than 500 m. Assessment of this change in SLA assimilation can be found in Section 5.2.
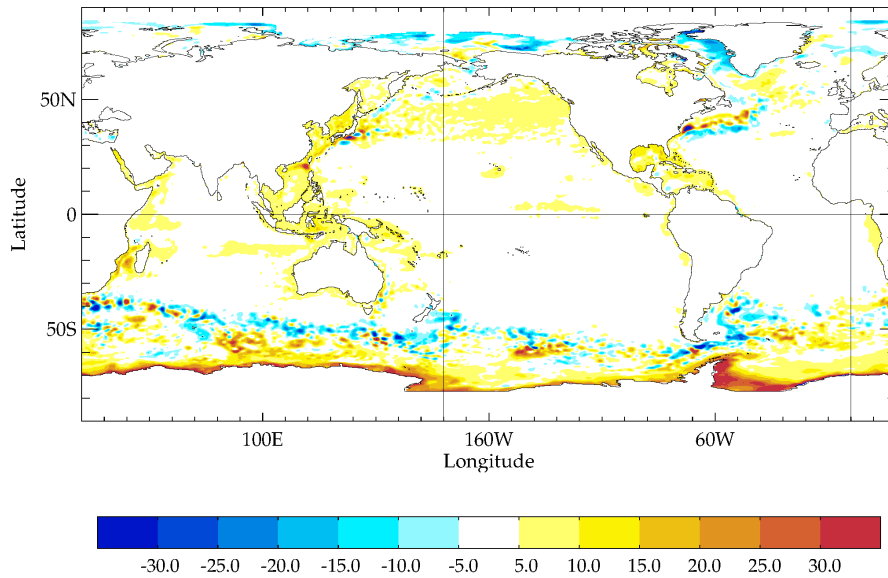
Figure 12. Difference (in cm) in MDT used by ORAS5 and by FOAM system. the MDT in FOAM is constructed using CNES CLS2013 (Rio et al., 2014) plus error adjustment term.

A reference mean dynamic topography (MDT) is required in order to assimilate SLA along-track data in an ocean general circulation model. This is necessary because altimeter measurement and the state variable in the ocean model are with respect

10   to different reference surfaces. There are several approaches to tackle this problem. One approach consists on using an external MDT (Rio et al., 2014), which is further corrected by using cumulative SLA innovation terms (Lea et al., 2008). This is the approach followed in the Met Office's global Forecasting Ocean Assimilation Model (FOAM, Waters et al. (2015)) and in the CMEMS global ocean monitoring and forecasting system (Lellouche et al., 2018). A different approach is used at ECMWF, and consists on estimating the MDT from a multi-year pre-reanalysis run assimilating T/S observations; this is the so-called model

15   MDT approach, and it is described in (Balmaseda et al., 2013a). The MDT in ORAS5 follows this model MDT approach, except that the pre-reanalysis run, which assimilate only in-situ observations and with bias correction, was produced using two parallel streams instead of one sequential integration, in order to accelerate the process of computing the MDT. The MDT was then constructed by averaging the resulting sea-surface height over a reference period 1996–2012, with additional correction

term to account for the different averaging period w.r.t the DT2014 data set as done in ORAP5 (Zuo et al., 2015). In this way, the assimilation of SLA constrains the temporal variability of the reanalysis without affecting the reanalysis mean state. However, it also means that the assimilation of SLA will not further correct model mean state. The difference in MDT used by ORAS5 and by FOAM system is shown in Fig. 12. Large differences can be found in regions with strong meso-scale eddy activities
5 (e.g. along the Western Boundary Currents and the ACC currents), and along the Antarctic coasts. A dipole of positive-negative MDT departures along the Gulf Stream and extensions is of particular interest. This is consistent with the estimated a-priori temperature and salinity biases in ORAS5 (Fig.10), suggesting some model/forcing errors in this regions.

Prior to 1993, mass variation that contributes to the change of Global Mean Sea Level (GMSL) in ORAS5 was constrained using the GRACE-derived climatology. The total GMSL was then constrained by assimilating altimeter-derived GMSL after
10 1993. This is the same as in ORAP5 (Zuo et al., 2015). The GMSL was derived from altimeter observations, firstly using reprocessed DT2014 gridded SLA data up to 2014, then using AVISO NRT gridded SLA from 2015 onwards. A systematic offset of GMSL between these two data sets is expected, due to slightly different data processing methods (e.g. multi-mission and mapping method). This offset is corrected for, in order to avoid introducing spurious GMSL discontinuities in the system. Assuming that sources of error do not change over time, this GMSL offset between delayed and NRT gridded SLA products
15 can be derived using GMSL difference averaged over their overlapping period. This period covers from May 2014 to November 2014 at the time of ORAS5 production. This value was then added for bias correction of GMSL derived from NRT data from 2015 onwards.

2.5   Ensemble generation

A new generic ensemble generation scheme developed by perturbing both observations and surface forcings has been imple-
20 mented in ORAS5. Here, we give a brief summary of the scheme. Preliminary assessments of ORAS5 temperature and salinity ensemble spread are also presented here. The reader should refer to Zuo et al. (2017a) for details about this ensemble generation scheme.

ORAS5 has employed a stratified random sampling method for pre-processing of both surface and sub-surface observations. As a result, the different members of the ensemble see different observations. This is a way to optimize the number of the
25 observations, since more observations are used in the ensemble. The in-situ observation profiles are perturbed in ORAS5 in two ways: by perturbing the longitude/latitude locations, and vertical perturbation by applying vertical stratified random thinning. The latitude/longitude locations of ocean in-situ profiles are perturbed so that the resulting locations are uniformly distributed within a circle of radius 50 km around the original location. This radius is chosen primarily considering observation representativeness error with respect to model horizontal resolution. The vertical thinning is applied by assuming a uniform
30 distribution of possible observation location within any given vertical range, and a maximum of 2 observations within each model level, if available, are then randomly selected for data assimilation. A similar stratified random thinning method is also applied to perturbing ORAS5 surface observations (SIC and SLA). In all cases some pre-defined reduced grids are constructed in order to carry out thinning, where observations within a given stencil in the reduced grid are randomly selected. As a result, each ensemble member assimilates slightly different observations. For SIC observation, this reduced grid is constructed with a

length scale of approximately 30 km in the Arctic region. For SLA observation, this reduced grid is constructed with a length scale of approximately 100 km in the Tropics. These values were chosen to ensure a reasonable sample size within the reduced grid. Altimeter observations from different satellite missions are treated separately. This method ensures that the number of observation assimilated in each of the perturbed ORAS5 members is comparable to that in the unperturbed member.

5   A new method has also been developed to perturb surface forcing fields used to drive ORAS5. This method preserves the multivariate relationship between different surface flux components, and has been used to perturb SST, SIC, wind stress, net precipitation (precipitation minus evaporation) and solar radiation. ORAS5 forcing perturbation takes into account both structural errors, which are derived from differences between separate analyses data sets (e.g. wind stress differences between NCEP and ERA-40); and analysis errors, which are derived from differences between ensemble members within the same

10   ensemble analysis (e.g. the 10 ensemble members of ERA20C (Poli et al., 2016)). The forcing in the ORAS5 control member remains unperturbed.

    Assessment of the ORAS5 temperature and salinity ensemble spreads has been carried out with respect to specified model BackGround Error (BGE) standard deviation (σ_b^s) and the BGE standard deviation diagnosed with the Desroziers method (σ_b^d), following the same procedure described in (Zuo et al., 2017a). Readers are reminded that the salinity σ_b^s shown here is for

15   unbalance component only. Fig. 13 shows spatial map of these diagnosed values at 100 m depth, after binning and averaging in $5°\times5°$ lon/lat boxes. Here, the ORAS5 temperature ensemble spread (Fig. 13a) shows a spatial pattern that is very similar to the diagnosed value using the Desroziers method (Fig. 13e), except its amplitude is weaker, especially in the Tropics. The salinity ensemble spread in ORAS5 (Fig. 13b) is in general under-dispersive when verified against diagnosed σ_b^d (Fig. 13f). The spatial patters between salinity ensemble spread and diagnosed σ_b^d are reasonably consistent. On the contrary, the specified σ_b^s in

20   ORAS5 are clearly overestimated almost everywhere for both temperature and salinity (Fig. 13c,d), suggesting that the current method of specifying temperature and salinity background error standard deviations using analytical functions (Mogensen et al., 2012) may be sub-optimal, especially for the tropical regions. Fig. 14 shows tropical averaged vertical profile of the same variables. Specified values for temperature and salinity (grey shaded area in Fig. 14a,b) are both larger than those of diagnosed σ_b^d (cyan dashed) values from surface to 2000 m. Estimations from ORAS5 ensemble spread (red solid), on the

25   other side, are more consistent with σ_b^d profiles, except for the upper 300 m where ensemble spreads are underestimated, by a factor of approximately 2. In order to assess impact of model resolution, we include here results from a ORAS5-equivalent Low-Resolution experiment (O5-LR, see Table 5) as well. The ensemble spreads in O5-LR (green solid) are almost always smaller than those of ORAS5. However, there are with noticeable variations in that difference depending on region and depth range.

30   Despite the fact that ORAS5 does not include stochastic model perturbations and has a small ensemble of only 5 members, its ensemble spread is still considered to be a better estimation of BGE than the specified values used in the current ocean analysis system. This indicates that specified model background error standard deviation can be improved by including this ensemble information, possibly in a hybrid way, in order to achieve better statistical consistency. This would also introduce a flow-dependent component into the NEMOVAR BGE covariances matrix through combing the ensemble-based and climatological

35   estimation of BGE covariances.
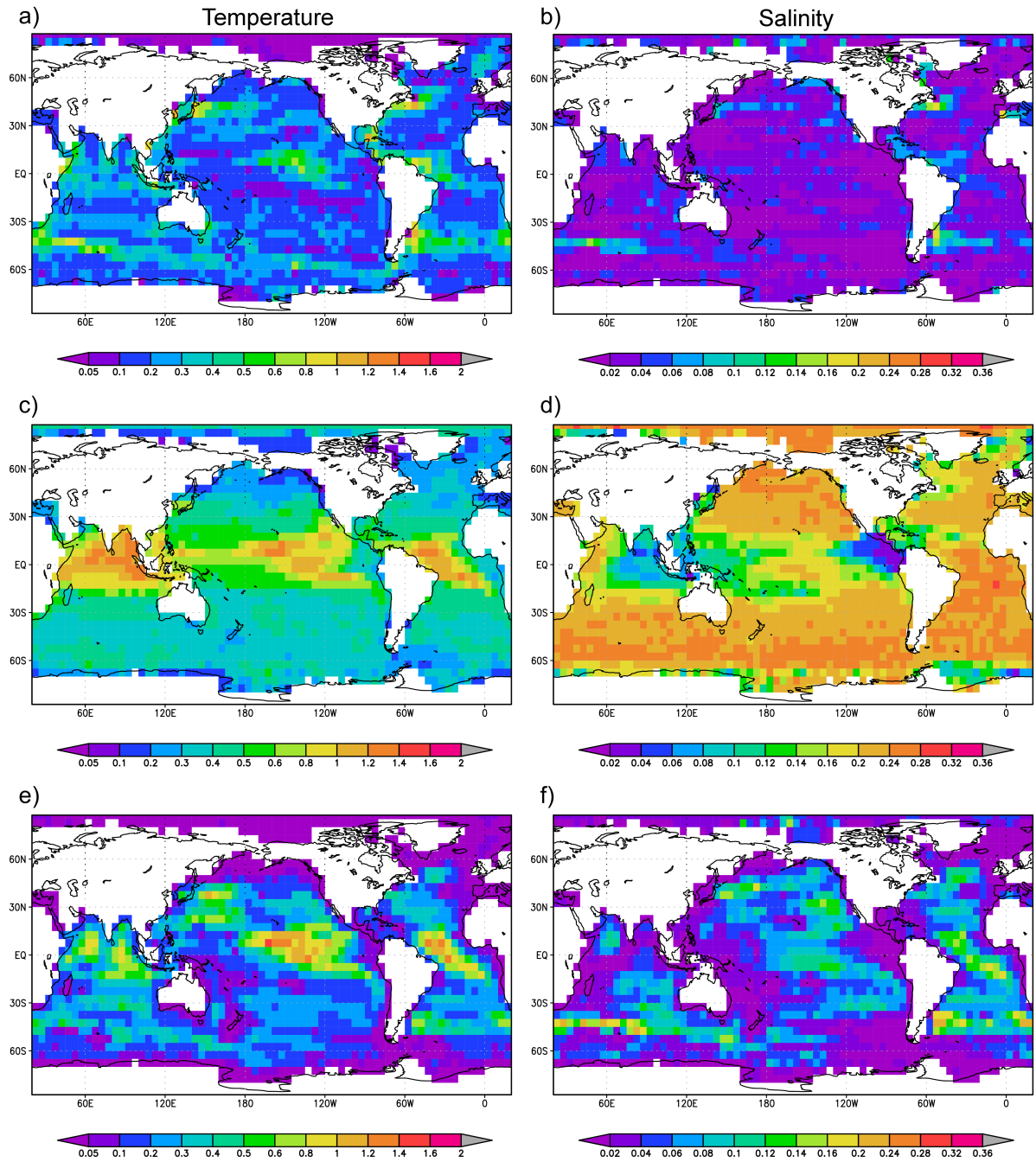
Figure 13. Maps of ORAS5 (a,c,e) temperature in K and (b,d,f) salinity in psu at 100 m, and as (a,b) ensemble spread, (c,d) specified BGE standard deviations and (e,f) diagnosed BGE standard deviations; Ensemble spread is calculated using model background values from all five ORAS5 ensemble members. All diagnostics are averaged over the 2010–2013 period, and binned and averaged into 5°×5° lon/lat boxes.
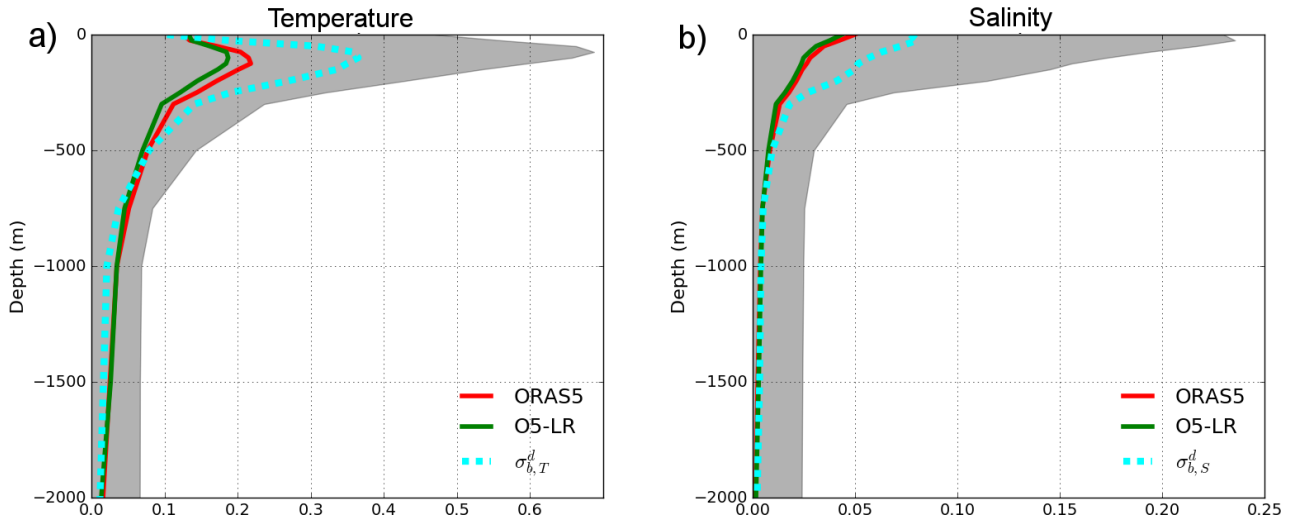
Figure 14. Vertical profiles of ensemble spread of (a) temperature (in K) and (b) salinity (in psu) from ORAS5 (red solid) and O5-LR (green solid). Ensemble spread is calculated using model background values, temporally averaged over the 2010–2013 period, and spatially averaged over the tropics (30°S to 30°N). σ_b^d is the diagnosed ORAS5 BGE standard deviation (cyan dashed) using Desroziers method (Desroziers et al., 2005); The specified BGE standard deviation (σ_b^s) is shown as the grey shaded area for reference.

3 The OCEAN5 Real-time analysis system

Based on ORAS5, a real-time ocean analysis system has been developed that forms the OCEAN5-RT component. This development has been done following a similar strategy as OCEAN4-RT (See Mogensen et al. (2012)). Now this OCEAN5-RT analysis provides the ocean and sea-ice initial conditions for all ECMWF coupled forecasting systems, including the ECMWF
5   medium-range and monthly ensemble forecast (ENS) since November 2016 (Buizza et al., 2016); the long-range forecasting system SEAS5 since November 2017 (Stockdale et al., 2017); and the high-resolution deterministic forecast (HRES) since June 2018 (Buizza et al., 2018). Work is on-going at ECMWF for coupling the lower boundary conditions of atmospheric analysis system to OCEAN5-RT analysis with SST and SIC (Browne et al., 2018)). Now the OCEAN5 system is a major component needed for ECMWF's Earth system approach, with an ever stronger coupling between the atmosphere, land, waves, ocean and
10   sea-ice components.

Fig. 15 shows schematically how the OCEAN5 suite, with its BRT and RT components, is implemented at ECMWF. The OCEAN5-BRT uses a 5-day assimilation window and is updated every 5 days with a delay D of 7 to 11 days. A minimum delay period of 7 days has been chosen in order to avoid a large degradation of the sea level analysis caused by delays in receiving NRT altimeter observations from CMEMS. The OCEAN5-RT analysis is updated daily using a variable assimilation window
15   of 8 to 12 days (equal to $D+1$): starting from the last BRT analysis, it brings the RT analysis forward up to current conditions, to produced ocean states suitable to initialise the coupled forecast. This RT extension contains 2 assimilation cycles (Chunk) with a variable second assimilation window. The RT extension is always initialized from the last day of the BRT analysis and

synchronically switches to the new initialization whenever the BRT analysis updates, hence the variable assimilation window. Taking current model day in its Year-Month-Day format (YMD), then in Fig. 15 the RT assimilation window length for YMD is 10 days, and is initialized from YMD-10 BRT analysis. In practice, the OCEAN5 RT analysis is launched every day at 14Z (same as ORTS4) to produce a daily analysis valid for 0Z for the following day (YMD+1).
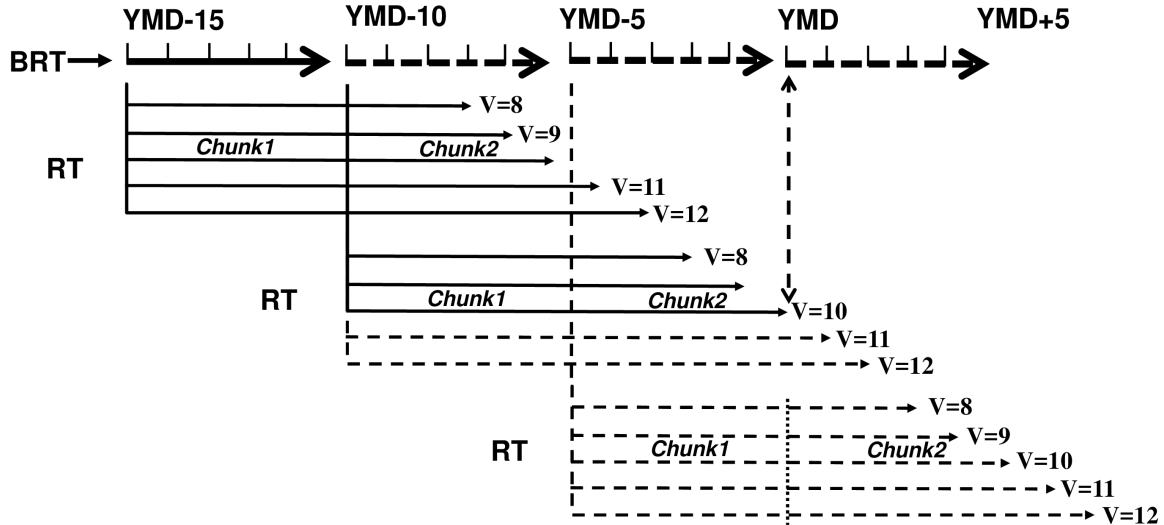
Figure 15. Schematic plot of OCEAN5 BRT and RT components: YMD=Current model date, V=Variable assimilation window length in the RT component. Solid lines denote analyses already produced in either BRT or RT component; dashed lines denote analyses not yet produced.

5   Unlike the historical ocean reanalysis, which is driven by atmospheric reanalysis forcing (e.g. ERA-Interim) and assimilates re-processed observation data sets whenever possible; the OCEAN5-RT component relies on ECMWF NWP forcings and NRT observation data input. The surface forcing fields that drive the OCEAN5-RT component come from ECMWF operational atmospheric analysis, except for the last day (YMD) when forcing is provided by ECMWF operational long forecast. Observations assimilated in OCEAN5-RT analysis come from GTS (ocean in-situ observations), CMEMS operational service (NRT

10   sea-level anomalies) and daily-mean SIC and SIC data from OSTIA operational analysis. However, these may be different from the BRT. In the case of in-situ observations, not all observations will be available at the start time or during the run time of the RT stream. SST and SIC data for the last day (YMD) are persisted from the previous day (YMD-1), since they are not available by the time the RT analysis is produced.

Table 5. Summary of ORAS5 sensitivity experiments

Name	Assim. SIC	Assim. Insitu	SST Nudging	Assim. SLA	Bias Corr.	notes
ORAS5	YES	YES	YES	YES	YES	
CTL-NoSST	NO	NO	NO	NO	NO	control run without SST
CTL-HadIS	NO	NO	YES	NO	NO	control run with SST
O5-NoAlt	YES	YES	YES	NO	YES	ORAS5 without SLA
O5-NoBias	YES	YES	YES	YES	NO	ORAS5 without bias correction
O5-LR	YES	YES	YES	YES	YES	ORAS5-equivalent low resolution run

All experiments are in ORCA025.L75 configuration, except for O5-LR which is in ORCA1.L42 configuration.

4 Assessment of ORAS5 system components

4.1 Sensitivity experiments

Additional experiments have been conducted within the ORAS5 framework to help with assessment of different system components. These include sensitivities to SST nudging, bias correction, assimilation of in-situ and satellite altimeter data. Studies of other system parameters, e.g. sensitivity to OBE STD specification, have been carried out but are not discussed here for the sake of conciseness. A summary of system configurations of these sensitivity experiments can be found in Table 5. All sensitivity experiments cover the period 1979 to 2015, and are driven by the same surface forcing fields from ERA-Interim. For all experiments except CTL-NoSST, SST are nudged to the HadISST2 product before 2008, and OSTIA operational analysis after 2008 (see Fig. 3). All diagnostics presented in this section focus on the unperturbed member only and ORAS5 always refers to the unperturbed member of the reanalysis in all of the following discussions.

4.2 Verification in observation space

Assessment of ORAS5 performance in observation space is carried out using model background errors with respect to all assimilated observations. We compute the model RMSE based on discrepancy between model background and observation for ORAS5 and all sensitivity experiments in Table 5. This approach allows to assess contributions from different system components and the performance of ORAS5 as an integrated reanalysis system. The reader should note that error statistics in CTL-NoSST and CTL-HadIS were computed in a observation space slightly differently (without vertical thinning of in-situ profiles) from other assimilation runs (ORAS5, O5-NoAlt, O5-NoBias). Assuming that there is no significant change in model

error characteristics within some small vertical depth range (e.g. within 100 m), then this comparison between control runs and assimilation runs is still valid. Time series of global mean RMSE in temperature and salinity from different sensitivity experiments are shown in Fig. 16, together with the total number of assimilated observations of various types shown with the right y-axis. Mean vertical profiles of these model RMSEs can be found in Figures 17, after temporally averaged over a period

5 (2005–2014) that is with near-homogeneous global Argo distribution.

Overall, all components of the ocean reanalysis system (SST nudging, bias correction, assimilation of in-situ observation and altimeter data) contribute to reducing the model error, both in temperature (Fig. 16(top)) and salinity (Fig. 16(bottom)). However, by construction, some components have a more profound impact on the improvement of the ocean state, e.g. the assimilation of in-situ observations. The magnitude of RMSE reduction due to direct T/S assimilation can be derived from

10 departure between O5-NoBias (red lines) and CTL-HadIS (green lines). The error reduction due to assimilation of in-situ data varies over time and is loosely proportional to the total number of observations assimilated. Over the Argo period 2005–2014, assimilation of in-situ data accounts for 65% of total RMSE reduction in temperature, and for nearly 90% of total RMSE reduction in salinity. These values are normalized against the total RMSE reductions derived from departures between ORAS5 (black lines) and CTL-NoSST (blue lines). Note that CTL-NoSST also shows a declining trend in its fit-to-observation errors,

15 especially following the introduction of the Argo floats (Fig. 16). It is important to point out that this trend in CTL-NoSST does not represent a change of model errors over time, but is mainly a result of the evolving GOOS. For instance, most southern extra-tropical ocean regions are only sampled with Argo floats after 2005. Therefore, global mean RMSE reduces after including these extra regions, because (a) by construction, observation errors are larger near the coast than in the open ocean (see Zuo et al. (2015)), and (b) there is much less land in the Southern than in the Northern Hemisphere. As a result, observations were

20 given more weight in these regions. The readers should note that results in Fig. 16 are also subject to changes in the surface driving forcings, e.g. improvement in ERA-Interim forcings over time due to better atmospheric observation coverage could result in reduced CTL-NoSST error as well.

After 2015 a noticeable drop in the available Argo observations is due to switching from re-processed EN4 to the NRT GTS data stream, leading to small rise of ORAS5 temperature and salinity RMSEs in Fig. 16. A disruption in TAO/TRITON

25 mooring array between 2012–2014 is also visible in Fig. 16top, which caused slightly increased ORAS5 RMSE in the Tropics during this period (not shown).

Differences between the CTL-NoSST and CTL-HadIS in Fig. 16 and Fig. 17 give an estimate of surface SST nudging contributions. This component contributes about 18% to the global temperature error reduction (Fig. 16(top)). However, it leads to an increase of salinity errors between 1985 and 2005 (Fig. 16(bottom)). This deterioration can be as large as 10% in

30 the mid 1990s. SST nudging is the dominant term in temperature error reduction for the upper 200 m in the northern extra-tropics (Fig. 17b), but also leads to slightly increased temperature error in the Tropics below 300 m (Fig. 17a). This degradation may be linked with the inappropriate partition of surface non-solar heat fluxes above and below the tropical thermocline, which is normally shallow than 200 m. During the Argo period, SST nudging also reduces the salinity RMSE for the upper 1000 m in the southern extra-tropics (Fig. 17f). For the upper 200 m of the southern extra-tropics, SST nudging accounts for nearly 40%
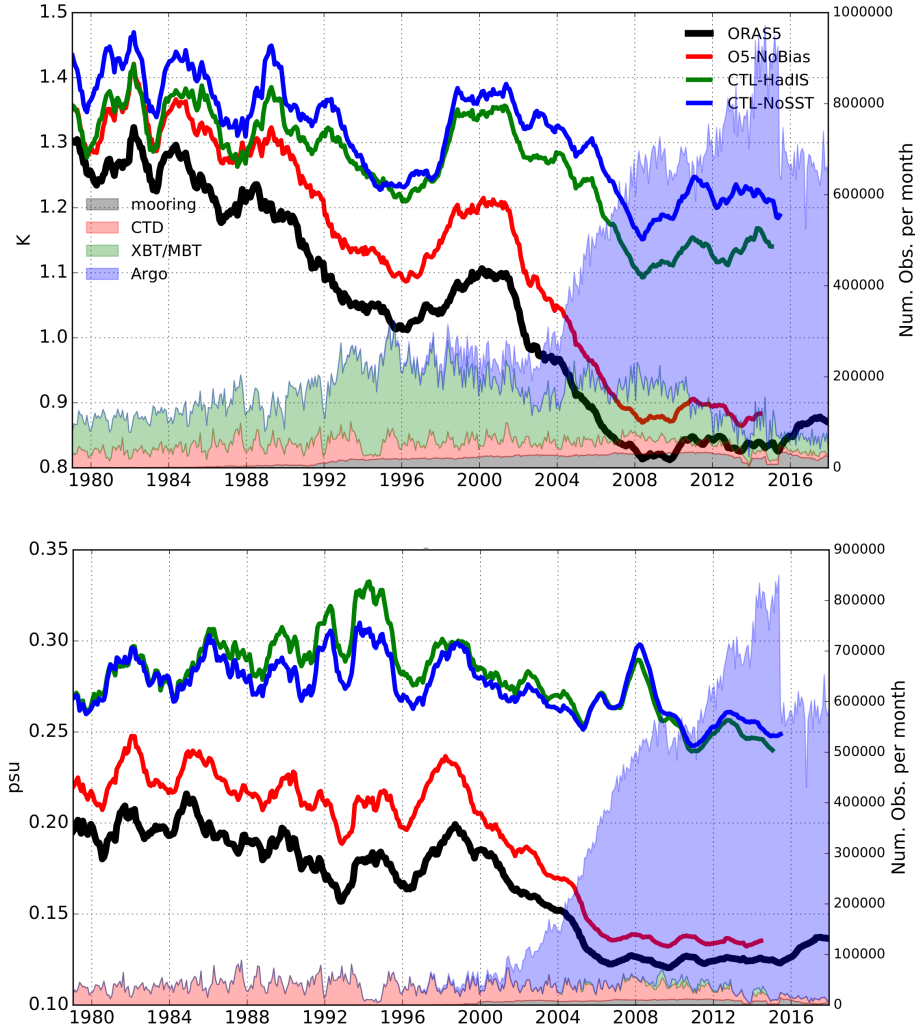
Figure 16. Time-series of global mean model fit-to-observation RMSE in (top) temperature (K) and (bottom) salinity (psu) from ORAS5 (black), O5-NoBias (red), CTL-HadIS (green) and CTL-NoSST (blue). Diagnostics are computed using model background departures from EN4 in-situ observations before June 2015, and departures from GTS observations from June 2015 on, and averaged over the upper 1000 m after smoothed with a 12-month running mean filter. Coloured patches and right y-axes show number of observations from different sources assimilated per month in ORAS5, accumulated for the upper 1000 m.

(0.05 psu) of salinity RMSE reduction. This suggests that some unstable vertical density structures could persist in the model background for this region.
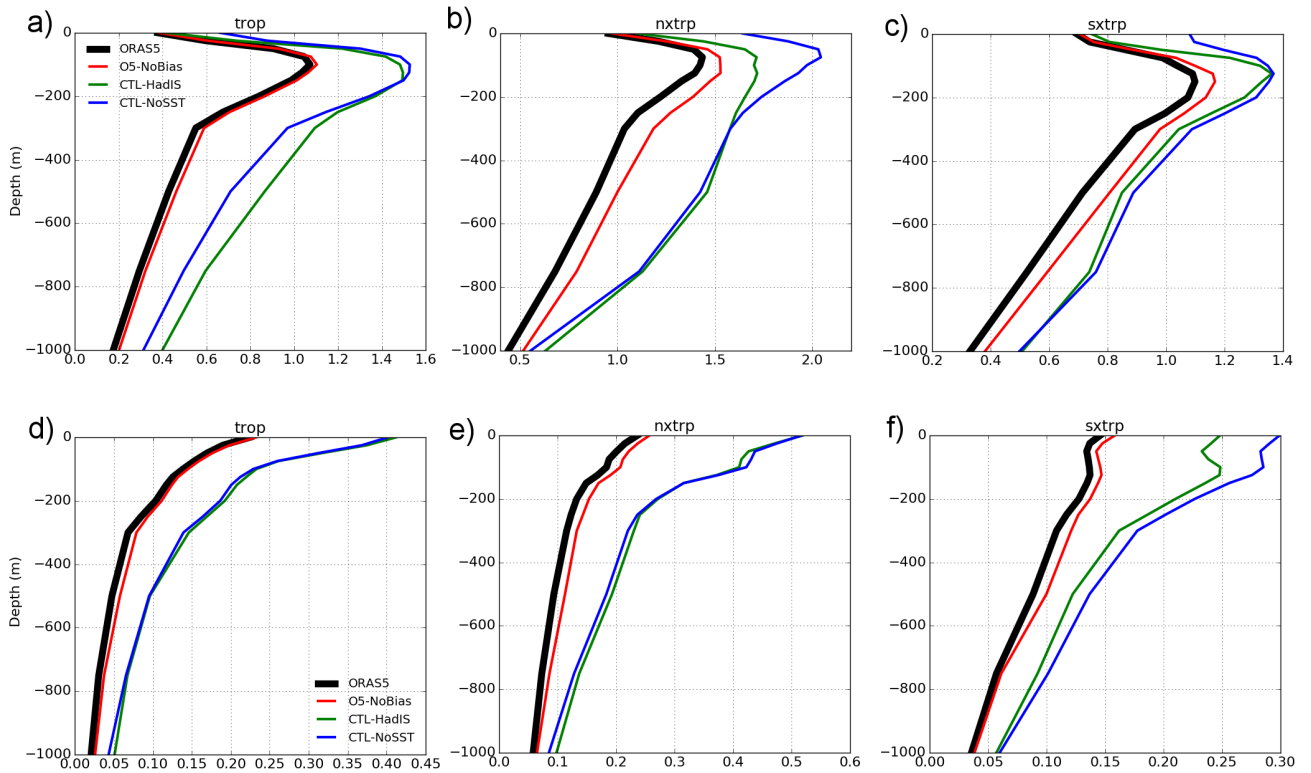
Figure 17. Vertical profiles of model fit-to-observation RMSEs in (a,b,c) temperature (K) and (d,e,f) salinity (psu), averaged over (a,d) Tropics (30°S to 30°N), (b,e) northern extra-tropics (30°N to 70°N), (c,f) southern extra-tropics (30°S to 70°S) and over the 2005–2014 period; for ORAS5 (black), O5-NoBias (red), CTL-HadIS (green) and CTL-NoSST (blue).

Contribution of the multi-scale bias correction implemented in ORAS5 can be derived from differences between O5-NoBias and ORAS5. This component plays an important role in correcting model errors, especially for the extra-tropical regions where the online bias term is applied as a direct correction to the T/S fields (Fig. 17b,c,e,f). In the global ocean, this bias correction contributes to the total RMSE reduction with about 14% for temperature and about 10% for salinity, averaged for the upper 1000 m. This bias correction contribution is also relatively stable over time, and less susceptible to the evolving GOOS (Fig. 16).

Other system components, like the assimilation of the altimeter data, lead to marginal improvements in global temperature (ca. 3%), and have mostly neutral impact on the model salinity errors (not shown). One possible reason for this relatively small impact from assimilation of altimeter data is that, by construction, the assimilation of SLA does not correct mean model biases but only affects the temporal variability of reanalysis. In addition, the altimeter data in the ECMWF reanalyses is perhaps given a weak weight compared with meso-scale applications of ocean data assimilation, as to avoid spurious circulations and degradation of the deep ocean (Zuo et al., 2017b). This result is very similar to ORAP5, which indicates that the new SLA thinning scheme in ORAS5 is as effective as the superobbing scheme in representing observation representativeness error.

Overall, we conclude that all components of the ORAS5 ocean data assimilation contribute to an improved ocean analysis state when verified against in-situ observations.

5   Assessment of ORAS5 ocean essential climate variables

Ocean essential climate variables (ECVs) are ocean variables commonly used for monitoring ocean state and climate signals
5   on decadal or longer time scales. SST, SLA and SIC are three of key ocean ECVs defined by the Global Climate Observing System (GCOS), and they have been selected here for an assessment of ORAS5 for climate applications. The ESA CCI project has developed suitable climate data records of these ECVs, which are generally derived from a combination of satellite and in-situ observations. Here, the latest versions of these ESA CCI climate data records for SST, SLA and SIC were chosen as reference climate data sets to verify ORAS5 and some relevant sensitivity experiments. These observation-only analyses are
10   produced with different production systems (e.g. different satellite missions) and/or processing chains (e.g. bias correction method) compared to the observational data sets that were assimilated in ORAS5. All statistics are computed using monthly-mean fields from ORAS5 and ESA CCI observation data sets interpolated to a common $1° × 1°$ latitude–longitude grid.

5.1   Sea surface temperature

The ESA SST CCI (SST_cci) long-term analysis provides daily surface temperature of the global ocean over the period 1992
15   to 2010. Unlike the HadISST2 and OSTIA SST analyses, both of which are bias-corrected against in-situ observations (e.g. drifting buoys), ESA SST_cci only uses satellites observations (AVHRR and ATSR). Therefore, it provides an reference SST data set of a quality that is suitable for climate research. The latest version 1.1 of the ESA SST_cci (Merchant et al., 2016) data set (referred to as SST_cci1.1 hereafter), has been used here for verification of the performance of ORAS5 at the sea surface. The SST_cci1.1 data set is an update of version 1.0 described by Merchant et al. (2014).
20   Fig. 18c,d show mean bias and normalized RMSE of ORAS5 SST with respect to SST_cci1.1 for the 1993–2010 period. For intercomparison, results from ORAS4 and two other sensitivity experiments are also included here. Compared to ORAS4 (Fig. 18a), ORAS5 SST has reduced warm bias in extra-tropics, especially in the northern North Pacific, the Norwegian sea, the Southern Ocean and in the Brazil/Malvinas current regions. A dipole of positive-negative bias patterns in the Gulf Stream and extension is still visible in ORAS5, though it is with reduced magnitude compared to ORAS4. This suggests that the pathway
25   of Gulf Stream extensions may be misrepresented in ORAS5. Spatial patterns of SST bias and RMSE in ORAS5 (Fig. 18c,d) are consistent with those derived from the difference of HadISST2 and SST_cci1.1 (Fig. 20a,b), with large RMSE normally in regions with strong eddy kinetic energy (EKE). These are also regions where ORAS5 SST has large ensemble spread (>0.5 K in Fig. 19). In general, the SST RMSE in ORAS5 is reduced w.r.t ORAS4 (Fig. 18b), e.g. in the South Indian, the South and western North Pacific, and southern South Atlantic. Readers are reminded that mean differences between ocean syntheses and
30   SST_cci1.1 have been removed before computing RMSE in Fig. 18b,d,f,h. Compared to ORAS4, the global averaged RMSE is reduction by about 10% (30% if taken mean difference into account) in ORAS5.
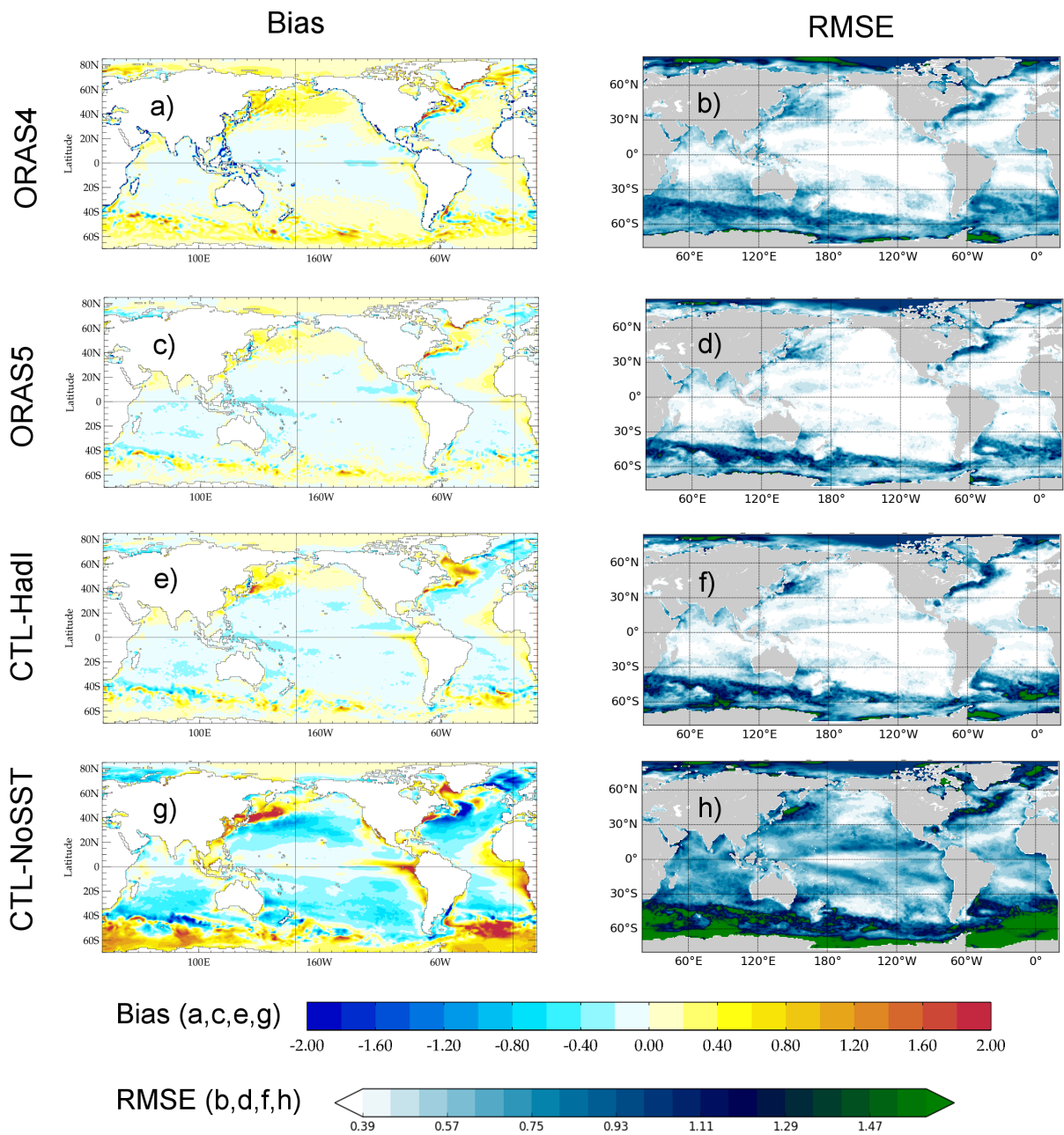
Figure 18. SST (a,c,e,g) bias in K and (b,d,f,h) normalized RMSE for (a,b) ORAS4, (c,d) ORAS5, (e,f) CTL-HadI and (g,h) CTL-NoSST, with respect to SST_cci1.1 data set. SST bias is computed using monthly-mean SST data and averaged over the 1993-2010 period. The RMSE is computed using monthly anomaly SST data after removal of seasonal cycle, and then normalized against the temporal standard deviation of SST_cci1.1 data (also without seasonal cycle) over the same period. Note that RMSE smaller than 0.4 are shown as white.
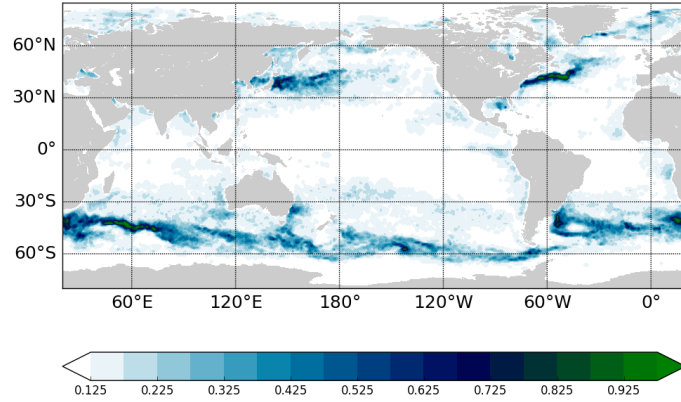
Figure 19. Ensemble spread of ORAS5 SST (K) estimated using five ensemble members of ORAS5, computed using monthly mean SST anomaly in 2010.
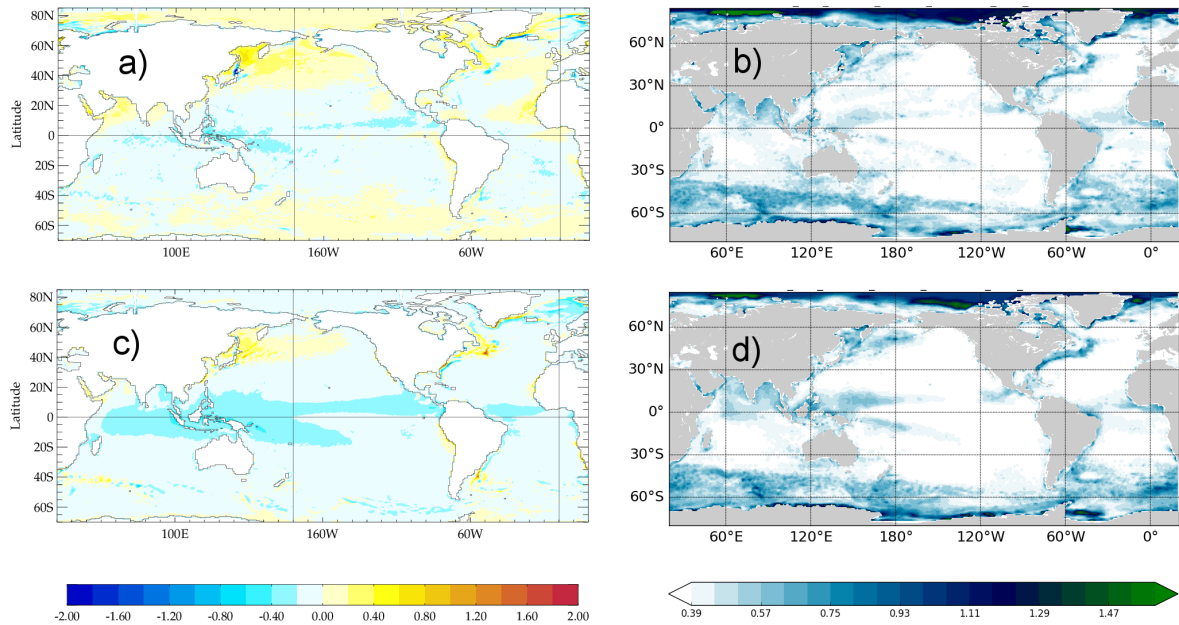
Figure 20. Same as Fig. 18 but for SST from (a,b) HadISST2 and (c,d) OSTIA data sets, as (a,c) mean bias in K and (b,d) normalized RMSE with respect to SST_cci1.1.

It is worth pointing out that different SST data sets were used for constraining SST in these two ocean syntheses before 2008: ORAS4 used OSTIA, and ORAS5 uses HadISST2. However, this improvement in ORAS5 SST can not be attributed to the new HadISST2 data set. To the contrary: w.r.t. SST_cci1.1, SST in HadISST2 has higher RMSE (by about 5%) and increased warm bias than OSTIA in the extra-tropics (Fig. 20). Therefore, improvements in ORAS5 SST should be attributed to increased model resolution and assimilation of updated EN4 in-situ data with improved vertical resolution.

Differences between ORAS5 (Fig. 18c,d) and CTL-HadIS (Fig. 18e,f) are non-trivial, with largely reduced mean biases in ORAS5, especially for the Labrador Sea and East of Japan. These regions also have large SST RMSE due to misrepresentation of mixed layer depth in CTL-HadIS, but are slightly improved in ORAS5 by assimilating in-situ observations. As expected, CTL-NoSST (Fig. 18g,h) has the largest SST biases w.r.t. SST_cci1.1. These biases are associated with systematic model and/or forcing errors, e.g. underestimated upwelling west of South America and South Africa, misrepresentation of mixing in the Southern Ocean, or others. The difference between CTL-NoSST and CTL-HadIS highlights the fact that the SST nudging method is very effective in keeping SST close to observations in the reanalysis system. Further investigation on poor performance in the Gulf Stream and extension is on-going at moment.

5.2 Sea level

The ESA sea-level CCI (SL_cci) project provides long-term along-track and gridded sea-level products from satellites for climate applications. Here, we use the latest reprocessed version 2.0 data from SL_cci (hereafter called SL_cci2) for validation of ocean syntheses sea level. The SL_cci2 sea-level data is an update of version 1.1 (Ablain et al., 2015) and includes data from additional altimeter missions (SARAL/AltiKa and CryoSat-2). Unlike the AVISO DT2014 product, which is dedicated to the best possible retrieval of meso-scale signals, SL_cci2 data focuses on the homogeneity and stability of the sea-level record. It has been produced using a different processing chain, and it also uses new altimeter standards, including a new orbit solution, atmospheric corrections, wet troposphere corrections, and a new mean sea-surface and ocean-tide model (see Quartly et al. (2017)). Therefore, it can be used here as an reference climate data set for validation of ocean syntheses in climate scale.

In order to evaluate the temporal variability of regional sea level in ocean synthesis, the temporal correlation between ORAS5 SLA and SL_cci2 gridded SLA data has been computed over the 2004–2013 period, with its result shown in Fig. 21c. In general, sea level variation of ORAS5 is well reproduced in the tropics, with a temporal correlation normally higher than 0.9. Reduced correlation is visible along the North Equatorial Countercurrent in the Pacific, and is related with the discrepancy between DT2014 and SL_cci2 data sets (Fig. 22a). Poor performance near the coast and in extra-tropics could be attributed partly to no SLA assimilation in these regions. This is similar for ORAS4 (Fig. 21a), except that ORAS4 sea level correlation is lower than ORAS5 almost everywhere, and especially in the tropical Indian, the tropical Atlantic and the Norwegian Sea. This difference can in large parts be attributed to the eddy-permitting model resolution of ORAS5, which accounts for most improvement in the extra-tropics, and the assimilation of the new AVISO DT2014 data set, which accounts for most improvement in the tropics.

As expected, removal of altimeter SLA data significantly degraded system performance, as demonstrated by correlation difference between O5-NoAlt (Fig. 21e) and ORAS5. In addition, assimilation of ocean in-situ observations further improves representation of sea level in the reanalysis due to better representation of meso-scale dynamics. This improvement is relatively
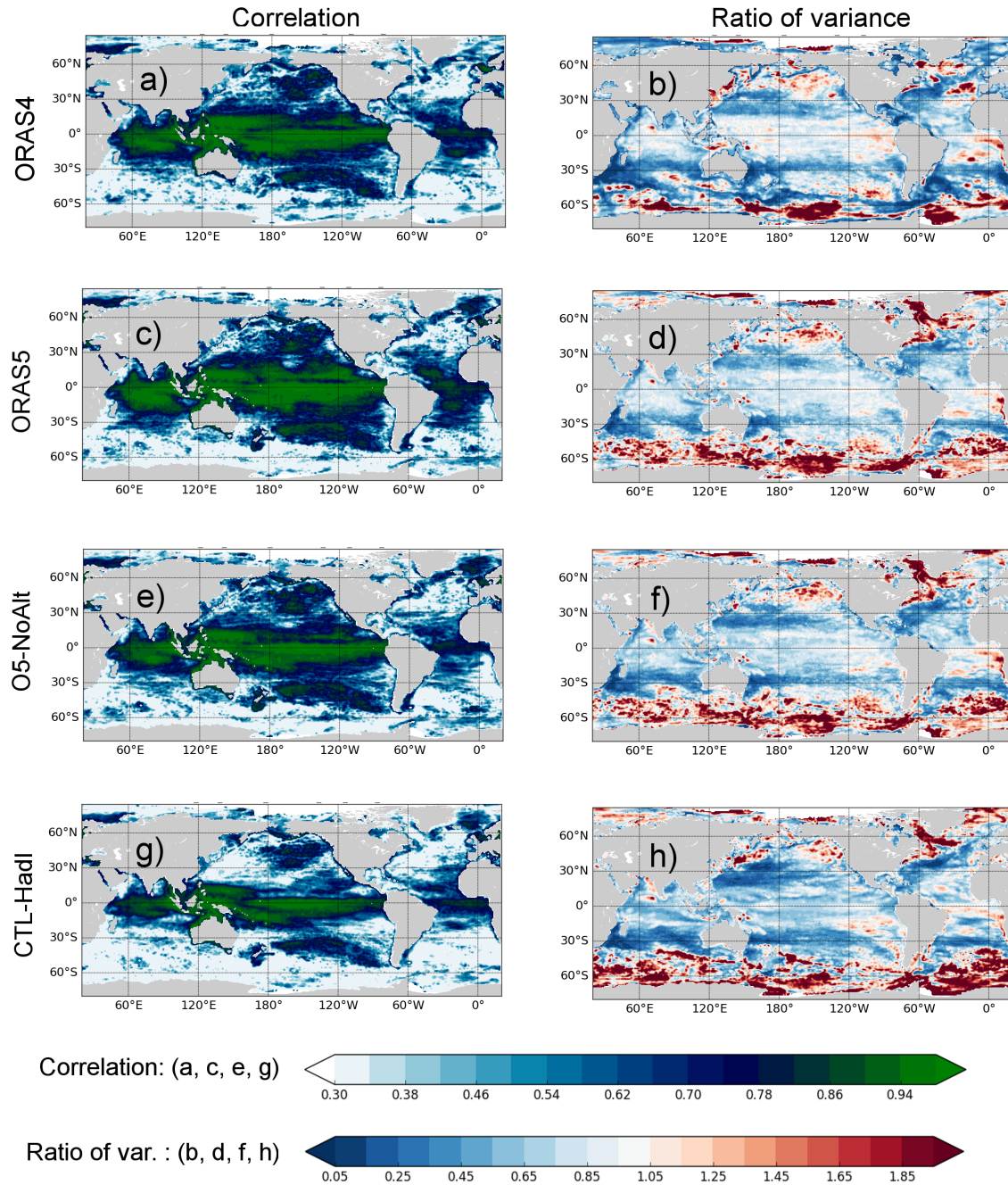
Figure 21. (a,c,e,g) Temporal correlation and (b,d,f,h) ratio of variance between SLA from ocean syntheses and SL_cci2. Ocean syntheses SLAs are from (a,b) ORAS4, (c,d) ORAS5, (e,f) O5-NoAlt and (g,h) CTL-HadIS. Statistics are computed using monthly-mean SLA data over the 2004–2013 period, temporal correlations are diagnosed after removal of the seasonal cycle. Note that correlations smaller than 0.3 are shown as white.
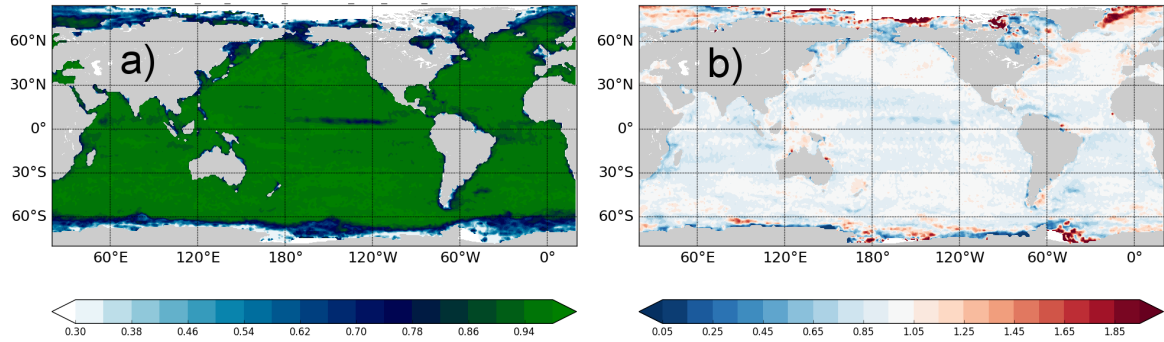
Figure 22. (a) SLA correlation and (b) ratio of variance between AVISO DT2014 and SL_cci2 data. Statistics are computed following the same way in Fig. 21.

homogeneous but most pronounced in the extra-tropical Pacific, as demonstrated by differences between O5-NoAlt and CTL-HadIS (Fig. 21g). We would like to point out that both O5-NoAlt and CTL-HadI performed reasonable well in the tropical Pacific and Indian, suggesting that these regions have the least model/forcing error. This result is consistent with the in-situ observation OSE in Section. 2.3.2.

5   For reference, the same diagnostics have been carried out for AVISO DT2014 data with respect to SL_cci2 (Fig. 22). In general, the temporal correlation between DT2014 and SL_cci2 is very high, indicating excellent agreement of temporal variations between the two data sets. Regions with lower correlation are visible though, e.g. along the North Equatorial Countercurrent in the Pacific between $180°$W and $100°$W (Fig. 22a). This is likely associated with differences in the production chains between DT2014 and SL_cci2, which include different altimeter-mission-dependent orbit solutions and geophysical corrections and

10   different filtering methods in processing along-track SLA. This discrepancy between different observational data sets is also responsible for the low correlation between ORAS5 and SL_cci2 SLA in the same region. Discrepancies in polar sea-level variances between DT2014 and SL_cci2 are likly associated with the new pole tide model (Desai et al., 2015) used in SL_cci2.

In order to evaluate the magnitude of temporal SLA variance in ORAS5, we compute the ratio of SLA variance between ocean syntheses and SL_cci2 for the 2004–2013 period, with results shown in Fig. 21(b,d,f,h). Compared to SL_cci2 data,

15   both ORAS4 (Fig. 21b) and ORAS5 (Fig. 21d) underestimate SLA variance between $50°$S–$50°$N. The domain-averaged SLA variance in ORAS4 is about two third of the SL_cci2 estimate, mostly because ORAS4 is incapable of resolving meso-scale activity and assimilates SLA through a superobbing scheme. This problem has been alleviated by increasing the model resolution and using a new SLA thinning scheme in ORAS5 (see Section. 2.4). However, ORAS5 still underestimates SLA variance by approximately one quarter in the average grid cell. One important reason for this underestimation is that ORAS5 still uses a $1°$

20   reduced grid when applying thinning for SLA observations, which may be sub-optimal considering ORAS5 comprises a $0.25°$ resolution ocean model. However, it remains an open question whether the assimilation should compensate for a deficiency that has its origin in the forward ocean model. The CTL-HadIS experiment clearly exhibits this underestimation (Fig. 21h), it

is likely related to the 0.25° resolution still being insufficient. Some of this underestimation is also attributed to the assimilated DT2014 data set, which has about 10% less variance than SL_cci2 in the average grid cell (see Fig. 22b). This difference between SL_cci2 and DT2014 is mostly due to different geophysical corrections used in production (Jean-François Legeais, personal communication). Removal of altimeter data (O5-NoAlt, Fig. 21f) and in-situ data (CTL-HadIS, Fig. 21h) from the
5   assimilation system further reduces simulated SLA variances, by approximately 3% and 5%, respectively. There are regions where ORAS5 has larger SLA variance though, e.g. in the Baffin Bay, Hudson Bay, and most areas in the Southern Ocean. The readers is referred to Legeais et al. (2018) for a detailed evaluation about ORAS5 sea level trend and its decomposition with respect to AVISO DT2014 and other ESA Sea Level CCI products.

5.3   Sea-ice concentration

10   The ESA Sea-Ice CCI (SI_cci) project has produced a long-term SIC data set based on satellite passive microwave radiances. The latest version 1.1 SIC data from SI_cci (hereafter SI_cci1.1) was produced using a sea-ice concentration algorithm and methodology developed by EUMETSAT Ocean and Sea Ice Satellite Application Facility (Sørensen and Lavergne, 2017). This SI_cci1.1 data set is available from 1993 to 2008 in 25 km resolution, and is used here for the evaluation of the ORAS5 sea ice.

    Fig. 23 shows maps of Arctic SIC RMSE based on departures between ocean syntheses and SI_cci1.1 data, averaged over
15   the 1993–2008 period. Note that coast lines are not drawn on the map. ORAP5 is a pilot reanalysis before ORAS5, and its ability to represent Arctic sea-ice has been documented to be reasonably good (Tietsche et al., 2015; Chevallier et al., 2017; Uotila et al., 2018). Therefore, ORAP5 has been retained here as a reference data set. Overall, ORAS5 SIC (Fig. 23c,d) has the same error characteristics as ORAP5 (Fig. 23a,b), which has already been well documented in Tietsche et al. (2015). The averaged SIC RMSE is normally less than 5% in the Arctic, again comparable with ORAP5. The largest ORAS5 SIC RMSE
20   (up to 20%) appears in the Labrador Sea in Arctic winter (Fig. 23c), which is caused by a mean positive(negative) SIC bias in the western(eastern) part of Labrador Sea. High SIC RMSE is also visible in the east coast of Greenland in both Arctic winter and summer for ORAS5 (Fig. 23d), and is caused by a mean positive(negative) SIC bias in the East Greenland Current north(south) of Iceland. These are also regions identified with large model/forcing errors as shown in CTL-HadIS (Fig. 23e). Like in ORAP5, visible SIC error along the Arctic coastal lines and in the Baltic Sea in ORAS5 can be attributed to observation
25   errors in OSTIA SIC reanalysis. Assimilation of OSTIA SIC has greatly improved sea-ice performance in ORAS5. Compared to CTL-HadIS (Fig. 23e,f), ORAS5 has reduced SIC RMSE almost everywhere in Arctic summer (Fig. 23d). The largest improvement in Arctic winter is located at the east of Greenland along the south edge of the Arctic sea-ice outflow extension, which is associated with model errors in ocean current and/or sea-ice velocity. The SST nudging scheme also contributes to reduction of SIC RMSE in the system (not shown). These improvements are mostly due to correction of thermodynamic errors
30   in the model, which is common in Arctic summer for Arctic surface water, but also in Arctic winter and in the Barent Seas.

    For reference, the ensemble spreads of ORAS5 SIC are shown in Fig. 24, which are estimated using the same monthly mean SIC conditions from the five ensemble members of ORAS5. This is encouraging to see that the spatial patterns of ORAS5 SIC uncertainty match those of RMSE reasonably well, even though that the ORAS5 is over-confident in the Labrador Sea and east coast of Greenland. ORAS5 sea-ice uncertainty has been tested by Richter et al. (2018) in two radiative transfer models
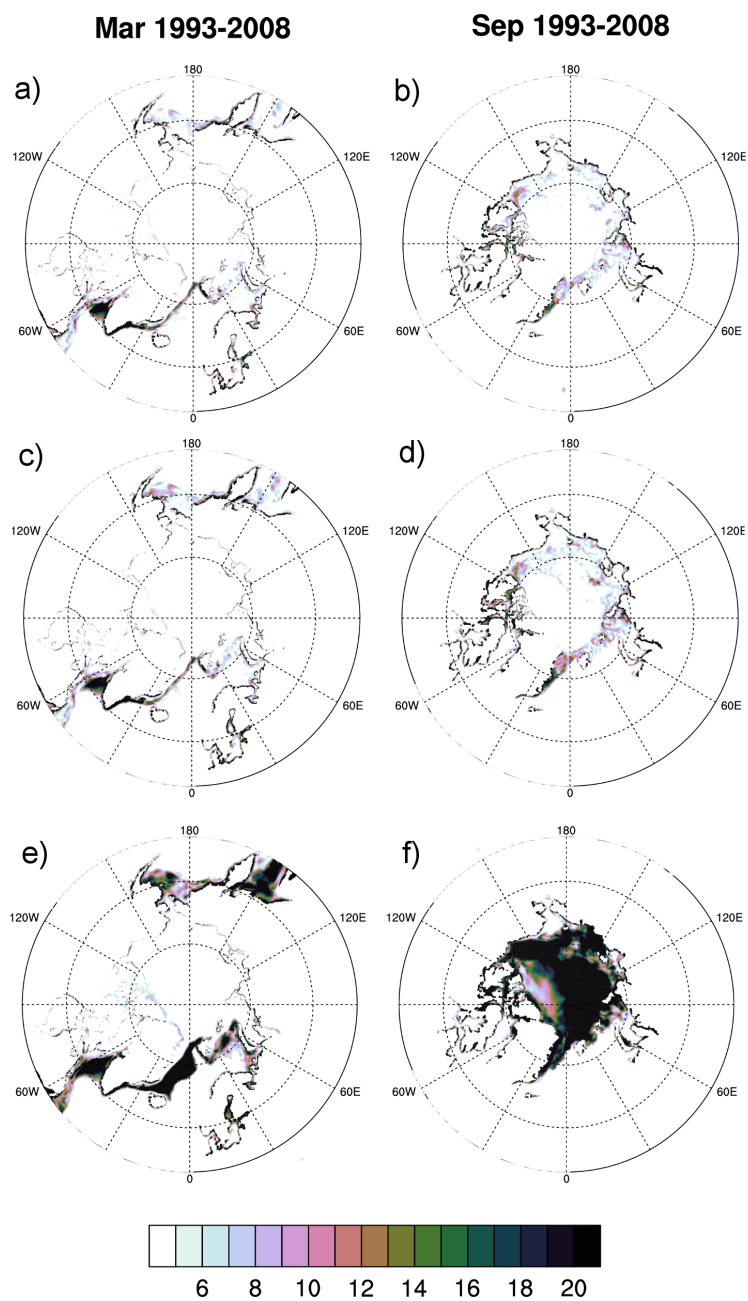
Figure 23. RMS departures (in percent) of ocean syntheses SIC with respect to SI_cci1.1 SIC data in (a,c,e) March and (b,d,f) September; ocean syntheses SIC are from (a,b) ORAP5; (c,d) ORAS5 and (e,f) CTL-HadIS. Statistics are computed using monthly SIC data over the 1993–2008 period.
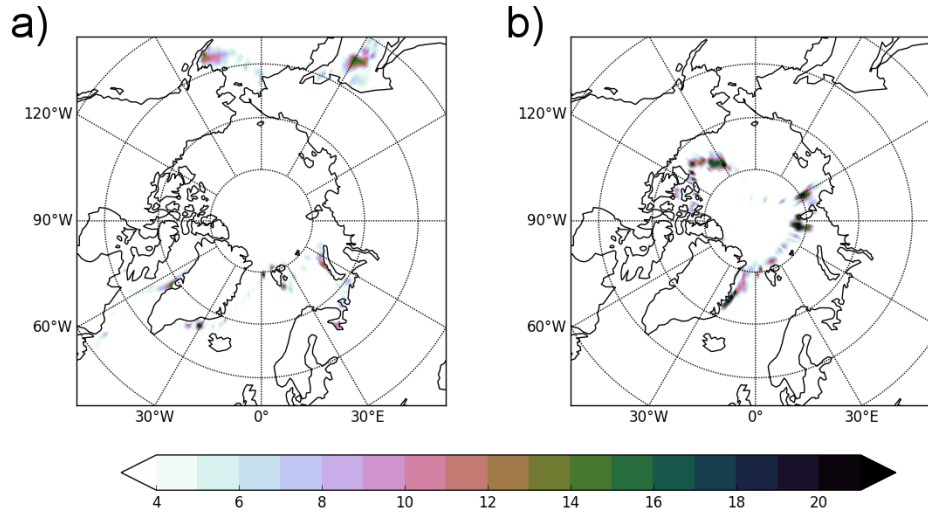
Figure 24. Ensemble spread of ORAS5 SIC (in percent) in (a) March and (b) September. Ensemble spread is estimated as the standard deviation of its five ensemble members, computed using monthly mean SIC in 2007.

to generate atmosphere brightness temperatures. In addition, an evaluation of ORAS5 sea-ice thickness in the Arctic has been carried out by Tietsche et al. (2018) with a focus on thin sea ice with respect to a data set derived from L-band radiances from the SMOS satellite. The interested reader is also referred to Zuo et al. (2018) for a case study about extreme sea-ice conditions derived from ORAS5 in 2016 and possible causes for both Arctic and Antarctic.

6 Conclusions

ORAS5 is a state-of-the-art 0.25° resolution ocean and sea-ice ensemble reanalysis system that covers the period from 1979 to present. ORAS5 and its real-time extension constitute OCEAN5, the fifth generation of ECMWF's ensemble reanalysis-analysis system. Major improvements of ORAS5 w.r.t. ORAS4 are the inclusion of a sea-ice reanalysis, increased resolution in the ocean, improved and up-to-date observational data sets, and improved methods for ensemble generation. ORAS5 also includes a series of system updates w.r.t. ORAP5, a pilot system. These include (a) improved observation pre-processing and quality-control methods; (b) revised bias correction scheme with stability check to prevent static instability; (c) a faster method to estimate the MDT for SLA assimilation. Particular attention is devoted to the consistency of surface observations, e.g. using HadISST2 SST together with OSTIA operational SST, and to an ensemble strategy that includes perturbation of initial condition, bias correction, observation and forcing. These system updates are described in detail in this manuscript, together with an evaluation of system performance in the context of data assimilation.

The OCEAN5 RT analysis is produced daily, and is essential for the timely initialization of the ECMWF coupled forecasts. Initialized from the latest ORAS5 conditions, the RT extension is produced by assimilating all available observational data into the ocean model driven by NWP forcings. Differences to ORAS5 are the variable assimilation window length, the smaller number of observations used, and the atmospheric forcing.

5 A series of sensitivity experiments have been carried out in order to assess ORAS5. It was found that all system components (SST nudging, assimilation of in-situ observation and/or SLA data, bias correction) contribute to an improved ocean state by reducing fit-to-observation errors in ocean syntheses. Among them direct assimilation of in-situ observations accounts for most improvements in both temperature (65%) and salinity (90%). This result suggests that different observation types (multiple altimeters, satellite SST and SIC observations, ocean in-situ) can be effectively assimilated in the ocean and sea-ice

10 model and allow constraining efficiently ocean and sea-ice states. Impact of different in-situ observation types in the current global ocean observing system were tested with global OSEs. Various metrics showed a non-linear degradation of the analysed ocean state for all observation types, with Argo showing the strongest impact. Region-wise, the degradation of the ocean state in the Atlantic was more severe than in the other main ocean basins, indicating the strong need for a dense in-situ network in this region.

15 The climate quality of ORAS5 has been evaluated using the three ECVs (SST, SLA and SIC) against reference climate records from the ESA CCI project. Results suggest that ORAS5 has an improved ocean state w.r.t. ORAS4 in the context of reconstructed SST and sea-level, with much reduced warm biases in extra-tropics and better regional sea-level variance between 50°S to 50°N. The performance of SIC in ORAS5 is similar to that of its predecessor ORAP5. In addition, the ORAS5 ensemble of SIC appears to provide a reliable measure of uncertainty in the estimation, being comparable to the

20 RMSE between ORAS5 and ESA CCI SIC observations. It also allows for uncertainty estimation of climate signals, which however is beyond the scope of this document and will be investigated elsewhere (Zuo et al., in preparation). Evaluations of ORAS5 have also been carried out within the framework of ESA SL_cci (Legeais et al., 2018), ESA-SMOS (Tietsche et al., 2018) and CMEMS projects (Zuo et al., 2018).

  The large SST biases in the Gulf Stream and its extensions have improved in ORAS5 compared to ORAS4, as a consequence

25 of increased spatial resolution. However, the bias remains large, and is associated with a fundamental misrepresentation of front positions and overshoot of the northward transport along the coast after Cape Hatteras. The impact of high resolution in ORAS5 is more visible in the area of the sub-polar gyre. Other issues identified in ORAS5 that need improving include the usage of observations in high latitudes, near the coast and on the continental shelf, especially with the recent development of the new ESA CCI sea-level product (Quartly et al., 2017; Legeais et al., 2018) which has reduced uncertainties in these regions. The

30 underestimated SLA variances is thought to associated with sub-optimal parameter specifications in observation errors and data sampling.

  Two clear priorities for developments of the ocean data assimilation system emerge from the experience with ORAS5. One is the treatment of SST observational constraints. The other required improvement is related with the assimilation of altimeter-derived sea level. The current relaxation method to constrain the SST has several shortcomings: i) it lacks the capability to

35 project directly the SST information into the subsurface, relying on the ocean model mixing processes to achieve that; ii) the

strength of the relaxation at high latitudes can have strong impacts on the ocean circulation, introducing process imbalance which damage the coupled forecast. The latter is the subject of a more detailed study (in preparation). It would be possible to optimize the strength of the SST nudging; but a longer-term solution requires investing in the proper assimilation of SST, using an appropriate vertical and horizontal correlation structure function and multivariate relationships. The assimilation of

5 altimeter-derived sea level should also be improved. The current practice of assimilating sea level anomalies (SLA) requires a pre-computed mean dynamic topography (MDT), which is expensive, or even unaffordable in coupled data assimilation, and it is prone to errors. Better solutions should be sought in terms of an online computation of the MDT (Lea et al., 2008), or, preferably, by making direct use of sea surface height and geoid information. The use of altimeter observations should also be optimized by further development of the multivariate background error covariance formulation in NEMOVAR, so as to include

10 constraints between sea surface height and barotropic stream function. This should have a large impact in constraining the position of the Gulf Stream and other oceanic fronts, which should benefit the NWP forecasting activities. Development of the next generation of ocean reanalysis system also requires: a) a better quality atmospheric forcing with increased temporal and spatial resolutions; b) an improved perturbation strategy with stochastic model perturbations; c) a flow-dependent BGE covariance matrix in NEMOVAR; and d) revised parameterizations for both OBE and BGE covariance matrices.

15 *Data availability.* The full ORAS5 data set can be downloaded from the Integrated Climate Data Center portal at http://icdc.cen.uni-hamburg.de/thredds/catalog/ftpthredds/EASYInit/oras5/catalog.html

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Development of the ORAS5 ocean reanalysis system has been supported by ESA through the sea-level CCI project (contract number 4000109872/13/I-NB), by the Copernicus Marine Environment Monitoring Service (CMEMS) through the GLO-RAN

20 project (BDC: 5554, GLO-RAN 23-CMEMS-Lot 2), and by the European Union's Horizon 2020 research and innovation program under grant agreement 633211 (AtlantOS). The production of ORAS5 has been funded by the Copernicus Climate Change Service.

References

Ablain, M., Cazenave, A., Larnicol, G., Balmaseda, M. A., Cipollini, P., Faugère, Y., Fernandes, M. J., Henry, O., Johannessen, J. A., Knudsen, P., Andersen, O., Legeais, J., Meyssignac, B., Picot, N., Roca, M., Rudenko, S., Scharffenberg, M. G., Stammer, D., Timms, G., and Benveniste, J.: Improved sea level record over the satellite altimetry era (1993–2010) from the Climate Change Initiative project, Ocean Science, 11, 67–82, https://doi.org/10.5194/os-11-67-2015, http://www.ocean-sci.net/11/67/2015/, 2015.

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

Balmaseda, M. A.: Ocean analysis at ECMWF: From real-time ocean initial conditions to historical ocean reanalysis, ECMWF Newsletter, 105, 24–42, http://www.ecmwf.int/publications/newsletters/pdf/105.pdf, 2005.

Balmaseda, M. A. and Anderson, D.: Impact of initialization strategies and observations on seasonal forecast skill, Geophysical Research Letters, 36, https://doi.org/10.1029/2008GL035561, 2009.

Balmaseda, M. A., Vidard, A., and Anderson, D. L. T.: The ECMWF Ocean Analysis System: ORA-S3, Monthly Weather Review, 136, 3018–3034, https://doi.org/10.1175/2008MWR2433.1, http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2433.1, 2008.

Balmaseda, M. A., Mogensen, K., and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4, Quarterly Journal of the Royal Meteorological Society, 139, 1132–1161, https://doi.org/10.1002/qj.2063, http://doi.wiley.com/10.1002/qj.2063, 2013a.

Balmaseda, M. A., Trenberth, K. E., and Källén, E.: Distinctive climate signals in reanalysis of global ocean heat content, Geophysical Research Letters, 40, 1754–1759, https://doi.org/10.1002/grl.50382, 2013b.

Barnier, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Le Sommer, J., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J., and De Cuevas, B.: Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution, Ocean Dynamics, 56, 543–567, https://doi.org/10.1007/s10236-006-0082-1, http://link.springer.com/10.1007/s10236-006-0082-1, 2006.

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

Breivik, Ø., Mogensen, K., Bidlot, J.-R., Balmaseda, M. A., and Janssen, P. A.: Surface wave effects in the NEMO ocean model: Forced and coupled experiments, Journal of Geophysical Research: Oceans, 120, 2973–2992, https://doi.org/10.1002/2014JC010565, 2015.

Browne, P., Rosnay, P. d., Zuo, H., Bennett, A., and Dawson, A.: Weakly coupled ocean–atmosphere data assimilation in the ECMWF NWP system, ECMWF Technical Memorandum, 2018.

Buizza, R., Bidlot, J. R., Janousek, M., Keeley, S., Mogensen, K., and Richardson, D.: New IFS cycle brings sea-ice coupling and higher ocean resolution, ECMWF Newsletter, 150, 14–17, 2016.

Buizza, R., Balsamo, G., and Haiden, T.: IFS upgrade brings more seamless coupled forecasts, ECMWF Newsletter, 156, 18–22, 2018.

Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Zuo, H., Balmaseda, M., Chang, Y.-S., Ferry, N., Garric, G., Haines, K., Keeley, S., Kovach, R. M.,

Kuragano, T., Masina, S., Tang, Y., Tsujino, H., and Wang, X.: Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project, Climate Dynamics, 49, 1107–1136, https://doi.org/10.1007/s00382-016-2985-y, https://doi.org/10.1007/s00382-016-2985-y, 2017.

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

Desai, S., Wahr, J., and Beckley, B.: Revisiting the pole tide for and from satellite altimetry, Journal of Geodesy, 89, 1233–1243, https://doi.org/10.1007/s00190-015-0848-7, https://doi.org/10.1007/s00190-015-0848-7, 2015.

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

Dibarboure, G., Pujol, M.-I., Briol, F., Traon, P. Y. L., Larnicol, G., Picot, N., Mertz, F., and Ablain, M.: Jason-2 in DU-ACS: Updated System Description, First Tandem Results and Impact on Processing and Products, Marine Geodesy, 34, 214–241, https://doi.org/10.1080/01490419.2011.584826, 2011.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sensing of Environment, 116, 140–158, https://doi.org/10.1016/j.rse.2010.10.017, 2012.

Fichefet, T. and Maqueda, M. A.: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics, Journal of Geophysical Research: Oceans, 102, 12 609–12 646, https://doi.org/10.1029/97JC00480, 1997.

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, Journal of Geophysical Research: Oceans, 118, 6704–6716, https://doi.org/10.1002/2013JC009067, 2013.

Gouretski, V. and Reseghetti, F.: On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database, Deep-Sea Research Part I: Oceanographic Research Papers, 57, 812–833, https://doi.org/10.1016/j.dsr.2010.03.011, 2010.

Haney, R. L.: Surface Thermal Boundary Condition for Ocean Circulation Models, Journal of Physical Oceanography, 1, 241–248, https://doi.org/10.1175/1520-0485(1971)001<0241:STBCFO>2.0.CO;2, 1971.

Hirahara, S., Balmaseda, M. A., and Hersbach, H.: Sea Surface Temperature and Sea Ice Concentration for ERA5, ERA Report Series, 26, 2016.

Karvonen, J., Vainio, J., Marnela, M., Eriksson, P., and Niskanen, T.: A Comparison Between High-Resolution EO-Based and Ice Analyst-Assigned Sea Ice Concentrations, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 1799–1807, https://doi.org/10.1109/JSTARS.2015.2426414, http://ieeexplore.ieee.org/document/7105862/, 2015.

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., and Saunby, M.: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization, Journal of Geophysical Research, 116, D14 104, https://doi.org/10.1029/2010JD015220, http://doi.wiley.com/10.1029/2010JD015220, 2011a.

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., and Saunby, M.: Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties, Journal of Geophysical Research Atmospheres, 116, https://doi.org/10.1029/2010JD015218, 2011b.

Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, Journal of Geophysical Research, 114, C07 005, https://doi.org/10.1029/2009JC005312, http://doi.wiley.com/10.1029/2009JC005312, 2009.

Large, W. G. and Yeager, S. G.: The global climatology of an interannually varying air–sea flux data set, Climate Dynamics, 33, 341–364, 2009.

Lea, D. J., Drecourt, J.-P., Haines, K., and Martin, M. J.: Ocean altimeter assimilation with observational- and model-bias correction, Quarterly Journal of the Royal Meteorological Society, 134, 1761–1774, https://doi.org/10.1002/qj.320, 2008.

Legeais, J. F., Ablain, M., Zawadzki, L., Zuo, H., Johannessen, J. A., Scharffenberg, M. G., Fenoglio-Marc, L., Joana Fernandes, M., Baltazar Andersen, O., Rudenko, S., Cipollini, P., Quartly, G. D., Passaro, M., Cazenave, A., and Benveniste, J.: An improved and homogeneous altimeter sea level record from the ESA Climate Change Initiative, Earth System Science Data, 10, 281–301, https://doi.org/10.5194/essd-10-281-2018, 2018.

Lellouche, J. M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C. E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., and Le Traon, P. Y.: Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system, Ocean Science, 14, 1093–1126, https://doi.org/10.5194/os-14-1093-2018, 2018.

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

Masina, S., Storto, A., Ferry, N., Valdivieso, M., Haines, K., Balmaseda, M., Zuo, H., Drevillon, M., and Parent, L.: An ensemble of eddy-permitting global ocean reanalyses from the MyOcean project, Climate Dynamics, 49, 813–841, https://doi.org/10.1007/s00382-015-2728-5, 2017.

Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., Good, S., McLaren, A., Rayner, N., Morak-Bozzo, S., and Donlon, C.: Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI), Geoscience Data Journal, 1, 179–191, https://doi.org/10.1002/gdj3.20, http://doi.wiley.com/10.1002/gdj3.20, 2014.

Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E. K., Bulgin, C. E., Corlett, G., Good, S., McLaren, A., Rayner, N., and Donlon, C.: ESA Sea Surface Temperature Climate Change Initiative (ESA SST CCI): Analysis long term product version 1.1, Tech. rep., Centre for Environmental Data Analysis, https://doi.org/10.5285/2262690A-B588-4704-B459-39E05527B59A, http://dx.doi.org/10.5285/2262690A-B588-4704-B459-39E05527B59A, 2016.

Mogensen, K., Balmaseda, M. A., and Weaver, A.: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4, ECMWF Technical Memorandum, 668, 1–59, 2012.

Penduff, T., Juza, M., Brodeau, L., Smith, G. C., Barnier, B., Molines, J.-M., Treguier, A.-M., and Madec, G.: Impact of global ocean model resolution on sea-level variability with emphasis on interannual time scales, Ocean Science, 6, 269–284, https://doi.org/10.5194/os-6-269-2010, http://archimer.ifremer.fr/doc/00003/11386/%5Cnhttp://www.ocean-sci.net/6/269/2010/, 2010.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., Fisher, M., Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, Journal of Climate, 29, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1, http://journals.ametsoc.org/doi/10.1175/JCLI-D-15-0556.1, 2016.

Pujol, M.-I., Faugère, Y., Taburet, G., Dupuy, S., Pelloquin, C., Ablain, M., and Picot, N.: DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years, Ocean Science, 12, 1067–1090, https://doi.org/10.5194/os-12-1067-2016, http://www.ocean-sci.net/12/1067/2016/, 2016.

Quartly, G. D., Legeais, J. F., Ablain, M., Zawadzki, L., Joana Fernandes, M., Rudenko, S., Carrère, L., Nilo Garciá, P., Cipollini, P., Andersen, O. B., Poisson, J. C., Mbajon Njiche, S., Cazenave, A., and Benveniste, J.: A new phase in the production of quality-controlled sea level data, Earth System Science Data, 9, 557–572, https://doi.org/10.5194/essd-9-557-2017, 2017.

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

Richter, F., Drusch, M., Kaleschke, L., Maaß, N., Tian-Kunze, X., and Mecklenburg, S.: Arctic sea ice signatures: L-band brightness temperature sensitivity comparison using two radiation transfer models, The Cryosphere, 12, 921–933, https://doi.org/10.5194/tc-12-921-2018, https://www.the-cryosphere.net/12/921/2018/, 2018.

Rio, M. H., Mulet, S., and Picot, N.: Beyond GOCE for the ocean circulation estimate: Synergetic use of altimetry, gravimetry, and in situ data provides new insight into geostrophic and Ekman currents, Geophysical Research Letters, 41, 8918–8925, https://doi.org/10.1002/2014GL061773, 2014.

Servonnat, J., Mignot, J., Guilyardi, E., Swingedouw, D., Séférian, R., and Labetoulle, S.: Reconstructing the subsurface ocean decadal variability using surface nudging in a perfect model framework, Climate Dynamics, 44, 315–338, https://doi.org/10.1007/s00382-014-2184-7, 2014.

Sørensen, A. and Lavergne, T.: Sea Ice Climate Change Initiative : D3.4 Product User Guide (PUG), Tech. rep., Norwegian Meteorological Institute, https://doi.org/10.5285/70f611b0-ba82-48e6-9190-a62cf9f925f2, http://www.esa-seaice-cci.org, 2017.

Stockdale, T., Johnson, S., Ferranti, L., Balmaseda, M. A., and Briceag, S.: ECMWF's new long-range forecasting system SEAS5, ECMWF Newsletter, 154, 15–20, 2017.

Tietsche, S., Notz, D., Jungclaus, J. H., and Marotzke, J.: Assimilation of sea-ice concentration in a global climate model – physical and statistical aspects, Ocean Science, 9, 19–36, https://doi.org/10.5194/os-9-19-2013, https://www.ocean-sci.net/9/19/2013/, 2013.

Tietsche, S., Balmaseda, M. A., Zuo, H., and Mogensen, K.: Arctic sea ice in the global eddy-permitting ocean reanalysis ORAP5, Climate Dynamics, pp. 1–15, https://doi.org/10.1007/s00382-015-2673-3, http://link.springer.com/10.1007/s00382-015-2673-3, 2015.

Tietsche, S., Alonso-Balmaseda, M., Rosnay, P., Zuo, H., Tian-Kunze, X., and Kaleschke, L.: Thin Arctic sea ice in L-band observations and an ocean reanalysis, The Cryosphere, 12, 2051–2072, https://doi.org/10.5194/tc-12-2051-2018, https://www.the-cryosphere.net/12/2051/2018/, 2018.

Titchner, H. A. and Rayner, N. A.: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations, Journal of Geophysical Research: Atmospheres, 119, 2864–2889, https://doi.org/10.1002/2013JD020316, http://doi.wiley.com/10.1002/2013JD020316, 2014.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V. S., Marnela, M., Massonnet, F., Mignac, D., Peterson, K. A., Sadikni, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J., and Zhang, Z.: An assessment of ten ocean reanalyses in the polar regions, Climate Dynamics, https://doi.org/10.1007/s00382-018-4242-z, https://doi.org/10.1007/s00382-018-4242-z, 2018.

Uppala, S. M., KÅllberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E.,

Hoskins, B. J., Isaksen, L., Janssen, P. a. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, https://doi.org/10.1256/qj.04.176, 2005.

Waters, J., Lea, D. J., Martin, M. J., Mirouze, I., Weaver, A., and While, J.: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, Quarterly Journal of the Royal Meteorological Society, 141, 333–349, https://doi.org/10.1002/qj.2388, http://doi.wiley.com/10.1002/qj.2388, 2015.

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

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

Xue, Y., Huang, B., Hu, Z.-Z., Kumar, A., Wen, C., Behringer, D., and Nadiga, S.: An assessment of oceanic variability in the NCEP climate forecast system reanalysis, Climate Dynamics, 37, 2511–2539, https://doi.org/10.1007/s00382-010-0954-4, https://doi.org/10.1007/s00382-010-0954-4, 2011.

Xue, Y., Wen, C., Kumar, A., Balmaseda, M., Fujii, Y., Alves, O., Martin, M., Yang, X., Vernieres, G., Desportes, C., Lee, T., Ascione, I., Gudgel, R., and Ishikawa, I.: A real-time ocean reanalyses intercomparison project in the context of tropical pacific observing system and ENSO monitoring, Climate Dynamics, 49, 3647–3672, https://doi.org/10.1007/s00382-017-3535-y, 2017.

Zuo, H., Balmaseda, M. A., and Mogensen, K.: The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice reanalysis ORAP5. Part 1: Implementation, ECMWF Technical Memorandum, 736, 2015.

Zuo, H., Balmaseda, M. A., Boisseson, E. D., Hirahara, S., Chrust, M., and Rosnay, P. D.: A generic ensemble generation scheme for data assimilation and ocean analysis, ECMWF Technical Memorandum, 795, 2017a.

Zuo, H., Balmaseda, M. A., and Mogensen, K.: The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals, Climate Dynamics, 49, 791–811, https://doi.org/10.1007/s00382-015-2675-1, 2017b.

Zuo, H., Vidar, L., Sandø, A. B., Garric, G., Bricaud, C., Storto, A., Peterson, K. A., Tietsche, S., and Mayer, M.: Extreme sea-ice conditions. In: Copernicus Marine Service Ocean State Report, Issue 2, Journal of Operational Oceanography, 11, S1–S142, https://doi.org/10.1080/1755876X.2018.1489208, 2018.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mayer, M., Robert, C. D., Mogensen, K., and de Rosney, P.: Evaluation of the ECMWF ensemble ocean and sea-ice reanalysis system ORAS5, in preparation.