

# The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4

Kristian Mogensen, Magdalena Alonso  
Balmaseda, Anthony Weaver<sup>1</sup>

Research Department

<sup>1</sup>CERFACS, Toulouse

February 2012

*This paper has not been published and should be regarded as an Internal Report from ECMWF.  
Permission to quote from it should be obtained from the ECMWF.*



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

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: library@ecmwf.int

©Copyright 2012

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

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Abstract

NEMOVAR is a variational data assimilation software for the NEMO ocean model. Its development is a collaborative project among different partners, ECMWF included. NEMOVAR is the basis of the new ECMWF operational Ocean analysis System 4 (Ocean-S4). It is the first time that NEMOVAR is used operationally.

Ocean-S4 uses NEMOVAR in its multivariate 3D-Var FGAT configuration. It assimilates temperature and salinity profiles and altimeter-derived sea level anomalies. In its current configuration Ocean-S4 includes an ensemble of NEMOVAR-derived reanalyses, consisting of 5 members. No attempt has been made to use the ensemble information in the specification of the background-error covariances. In Ocean-S4, the specification of the background-error covariances is somewhat heuristic, with computational efficiency and simplicity being important considerations. This report describes NEMOVAR and how it is used in both the reanalysis and real-time suites of Ocean-S4, called ORAS4 and ORTS4, respectively.

The performance of NEMOVAR is evaluated using a set of objective metrics, including comparison with independent data and the impact on seasonal forecast skill. NEMOVAR improves the mean state and inter-annual variability of the ocean estimation in most regions, including the Equatorial Atlantic, a traditionally challenging area. NEMOVAR also improves the skill of seasonal forecasts. While the overall impact of NEMOVAR on the ocean state estimation is positive, some problem areas have been identified such as the representation of the Atlantic Meridional Overturning Circulation. It is expected that ongoing developments to include more versatile formulations of the background-error covariances will result in an improved data assimilation system.

## 1 Introduction

The ocean component of the ECMWF forecasting system has been recently changed. At ECMWF the ocean model is coupled to the atmosphere in the extended range forecasts: in the Ensemble Prediction System (EPS) from day 10 ([Vitart et al., 2008](#)) and in seasonal forecasts from day 0 ([Stockdale et al., 2011; Molteni et al., 2011](#)). In 2011 both the ocean model and ocean data assimilation system have been changed: the HOPE (Hamburg Ocean Primitive Equation) ocean model ([Wolff et al., 1997](#)) has been replaced with the NEMO (Nucleus for European Modelling of the Ocean) model ([Madec, 2008](#)). The Optimal Interpolation (OI) of Ocean Reanalysis System 3 (ORAS3) ([Balmaseda et al., 2008](#)) has been replaced by the newly developed data assimilation system NEMOVAR ([Mogensen et al., 2009](#)). The purpose of this report is to introduce NEMOVAR and to describe its specific implementation in the operational ocean analysis system 4 (Ocean-S4). Ocean-S4 provides the ocean initial conditions for the extended-range forecasting systems at ECMWF.

Ocean-S4 is a global assimilation system based on the ORCA1 global configuration used by the NEMO model. The ORCA1 configuration has an approximate horizontal resolution of  $1^\circ$  with meridional refinement reaching  $0.3^\circ$  in the equatorial region. This resolution matches the resolution used by the coupled modelling systems mentioned above.

NEMOVAR is a variational data assimilation system adapted to the NEMO model. It is designed as a four-dimensional variational assimilation (4D-Var) algorithm. Three-dimensional variational assimilation (3D-Var) is also supported, using the First-Guess at Appropriate Time (FGAT) approach. NEMOVAR is based on the OPAVAR system ([Weaver et al., 2003; Vialard et al., 2003; Ricci et al., 2005; Weaver et al., 2005; Daget et al., 2009](#)) developed mainly at CERFACS for an earlier and now obsolete version of NEMO/OPA (version 8.1/8.2, [Madec et al. \(1998\)](#)). The development of NEMOVAR is a collaborative enterprise involving a number of partners (CERFACS, ECMWF, Met Office, INRIA/LJK). Ocean-S4 is the first operational implementation of NEMOVAR and is based on 3D-Var FGAT.

This technical memorandum has the following sections. Section 2 is a generic but not exhaustive description of

the NEMOVAR system. Here, the general four-dimensional (4D) version of NEMOVAR is presented first, and the simplifying assumptions made specifically for Ocean-S4 are outlined. Section 3 describes in more detail the components of NEMOVAR used in Ocean-S4. Features of NEMOVAR that are not used in Ocean-S4 are generally not discussed. Section 4 describes the specifications made for Ocean-S4 in its reanalysis (ORAS4) and real-time (ORTS4) streams. Section 5 assesses the performance of NEMOVAR by evaluating the quality of ORAS4 according to a set of objective metrics, including the impact on the initialization of seasonal forecasts.

## 2 The NEMOVAR system: general formulation

### 2.1 Nonlinear formulation

NEMOVAR is a four-dimensional data assimilation system that is designed to produce an ocean analysis by approximately minimizing the cost function

$$\mathcal{J}[\mathbf{w}] = \frac{1}{2} (\mathbf{w} - \mathbf{w}^b)^T \mathbf{B}^{-1} (\mathbf{w} - \mathbf{w}^b) + \frac{1}{2} (G(\mathbf{w}) - \mathbf{y}^o)^T \mathbf{R}^{-1} (G(\mathbf{w}) - \mathbf{y}^o) \quad (1)$$

where  $\mathbf{w}$  is the vector of analysis or control variables,  $\mathbf{w}^b$  is the background estimate of the control vector,  $\mathbf{y}^o$  is the observation vector, and  $\mathbf{B}$  and  $\mathbf{R}$  are symmetric, positive-definite matrices containing estimates of the background- and observation-error covariances, respectively. On a given assimilation window  $t_0 \leq t_i \leq t_N$ , the observation vector can be written as  $\mathbf{y}^o = ((\mathbf{y}_0^o)^T \cdots (\mathbf{y}_i^o)^T \cdots (\mathbf{y}_N^o)^T)^T$  where  $\mathbf{y}_i^o$  is the observation vector at time  $t_i$ . The nonlinear operator  $G(\mathbf{w})$  is a generalized observation operator that maps the control vector into the space of the observation vector. It has the general form

$$G(\mathbf{w}) = \begin{pmatrix} \vdots \\ G_i(\mathbf{w}) \\ \vdots \end{pmatrix} = \begin{pmatrix} & & \vdots \\ H_i(M(t_i, t_0)(K(\mathbf{w}))) & & \\ & & \vdots \end{pmatrix} \quad (2)$$

where  $K(\mathbf{w}) = \mathbf{x}(t_0)$  is a balance operator that transforms the control vector, defined at  $t_0$ , into initial conditions for the ocean model;  $M(t_i, t_0) = M(t_i, t_{i-1}) \cdots M(t_1, t_0)$  is the nonlinear model operator that propagates the initial state vector to the observation time  $t_i$ ,  $\mathbf{x}(t_i) = M(t_i, t_0)[\mathbf{x}(t_0)]$ ; and  $H_i$  is the observation operator that transforms the state vector at  $t_i$  into measured quantities and locations,  $\mathbf{y}_i = H_i[\mathbf{x}(t_i)]$ . The global minimizing solution of (1) is the analysis  $\mathbf{w}^a$ . The optimal trajectory through the assimilation window  $t_0 \leq t_i \leq t_N$  is then  $\mathbf{x}^a(t_i) = M(t_i, t_0)(\mathbf{x}^a(t_0))$  where  $\mathbf{x}^a(t_0) = K(\mathbf{w}^a)$ .

In NEMO, the state variables comprising  $\mathbf{x}(t_i)$  are potential temperature ( $T$ ), salinity ( $S$ ), sea-surface height (SSH) ( $\eta$ ) and the horizontal components of velocity ( $\mathbf{u}^h = (u, v)$ )<sup>1</sup>. The control variables comprising  $\mathbf{w}$  are constructed in a way to render their cross-covariances much weaker than those of the state variables. Formally, this is done using the inverse of the balance operator which attempts to remove from each state variable those components that are physically or statistically related to the other variables. In NEMOVAR, the control variables are taken to be  $T$  and so-called *unbalanced* components of salinity ( $S_U$ ), SSH ( $\eta_U$ ), and velocity ( $\mathbf{u}_U^h = (u_U, v_U)$ ). In practice, the cross-covariances between these transformed variables are neglected altogether, so that  $\mathbf{B}$  is assumed to have a simplified, univariate structure involving block-diagonal components associated with  $T$ ,  $S_U$ ,  $\eta_U$  and  $\mathbf{u}_U^h$ . The balance operator relating these variables to the state variables is described in section 3.1, while the univariate covariance model is described in section 3.2.

<sup>1</sup>In this notation, the individual components of the state and control vectors are scalar functions of the spatial coordinates.

## 2.2 Incremental formulation

Define  $\mathbf{w}$  as the sum of the background state and an increment  $\delta\mathbf{w}$ :

$$\mathbf{w} = \mathbf{w}^b + \delta\mathbf{w}.$$

The operator (2) can then be approximated as

$$G(\mathbf{w}^b + \delta\mathbf{w}) \approx G(\mathbf{w}^b) + \mathbf{G}\delta\mathbf{w} \quad (3)$$

where

$$\mathbf{G} = \begin{pmatrix} \vdots \\ \mathbf{G}_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \mathbf{H}_i \mathbf{M}(t_i, t_0) \mathbf{K} \\ \vdots \end{pmatrix} \approx \begin{pmatrix} \vdots \\ \partial G_i / \partial \mathbf{w} |_{\mathbf{w}=\mathbf{w}^b} \\ \vdots \end{pmatrix},$$

$\mathbf{H}_i$ ,  $\mathbf{M}(t_i, t_0)$  and  $\mathbf{K}$  being linear operators that approximate the tangent-linear (TL) of the observation operator, model propagator, and balance operator, respectively. Substituting the approximation (3) into (1) yields a quadratic cost function in terms of the increment:

$$J[\delta\mathbf{w}] = \frac{1}{2} \delta\mathbf{w}^T \mathbf{B}^{-1} \delta\mathbf{w} + \frac{1}{2} (\mathbf{G}\delta\mathbf{w} - \delta\mathbf{y}^o)^T \mathbf{R}^{-1} (\mathbf{G}\delta\mathbf{w} - \delta\mathbf{y}^o) \quad (4)$$

where

$$\delta\mathbf{y}^o = \begin{pmatrix} \vdots \\ \delta\mathbf{y}_i^o \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \mathbf{y}_i^o - G_i(\mathbf{w}^b) \\ \vdots \end{pmatrix}$$

is a 4D vector of observation-minus-background state differences (innovations), which in compact form can be written as

$$\delta\mathbf{y}^o = \mathbf{y}^o - G(\mathbf{w}^b).$$

The minimization of  $J$  is performed using a conjugate gradient (CG) method, which requires evaluating the gradient of the cost function on each iteration. From (4) the gradient is computed as

$$\nabla_{\delta\mathbf{w}} J = \mathbf{B}^{-1} \delta\mathbf{w} + \mathbf{G}^T \mathbf{R}^{-1} (\mathbf{G}\delta\mathbf{w} - \delta\mathbf{y}^o) \quad (5)$$

where

$$\mathbf{G}^T = (\cdots \mathbf{G}_i^T \cdots) = (\cdots \mathbf{K}^T \mathbf{M}(t_i, t_0)^T \mathbf{H}_i^T \cdots)$$

is the adjoint of the generalized observation operator.

On the first CG iteration, the increment is initialized to zero:

$$\delta\mathbf{w}_{(0)} = \mathbf{0}$$

where the subscript between parentheses denotes the CG iteration counter. After the final CG iteration  $m$ , the background state is updated with the increment:

$$\mathbf{w}_{(m)} = \mathbf{w}^b + \delta\mathbf{w}_{(m)}.$$

The control vector  $\mathbf{w}^a = \mathbf{w}_{(m)}$  defines the *analysis* and the increment  $\delta\mathbf{w}^a = \delta\mathbf{w}_{(m)}$  defines the *analysis increment*. The analysis and the analysis increment in the space of the model initial conditions can be obtained by applying the balance operator:

$$\begin{aligned} \mathbf{x}^a &= K(\mathbf{w}^a) \\ \text{and } \delta\mathbf{x}^a &= K(\mathbf{w}^b + \delta\mathbf{w}^a) - K(\mathbf{w}^b) \approx \mathbf{K}\delta\mathbf{w}^a. \end{aligned}$$

## 2.3 Multi-incremental formulation

An outer loop is added to the incremental algorithm to allow the reference state of the linearized operators to be updated with the most recent estimate of the control vector obtained during minimization. The algorithm is the basis of (multi-)incremental 4D-Var ([Courtier et al., 1994](#)) and is a variant of the classical Gauss-Newton method in optimization theory ([Gratton et al., 2007](#)). Let  $\mathbf{w}^{(k-1)}$  denote the reference state and let  $\delta\mathbf{w}^{(k)}$  be an increment to the reference state such that

$$\mathbf{w}^{(k)} = \mathbf{w}^{(k-1)} + \delta\mathbf{w}^{(k)}.$$

The superscript between parentheses denotes the outer iteration counter,  $k = 1, \dots, N_k$ , where the initial value  $\mathbf{w}^{(0)}$  is taken to be the background state  $\mathbf{w}^b$ . In terms of  $\mathbf{w}^{(k-1)}$  and  $\delta\mathbf{w}^{(k)}$ , the generalized observation operator on outer iteration  $k$  can be approximated as

$$G(\mathbf{w}^{(k-1)} + \delta\mathbf{w}^{(k)}) \approx G(\mathbf{w}^{(k-1)}) + \mathbf{G}^{(k-1)}\delta\mathbf{w}^{(k)}$$

where  $\mathbf{G}^{(k-1)}$  is an approximation of the TL of  $G(\cdot)$  about the reference state  $\mathbf{w}^{(k-1)}$ . With this approximation, the minimization of the non-quadratic cost function (1) is transformed into a sequence,  $k = 1, \dots, N_k$ , of minimizations of *quadratic* cost functions:

$$\begin{aligned} J^{(k)}[\delta\mathbf{w}^{(k)}] &= \frac{1}{2} \left( \delta\mathbf{w}^{(k)} - \delta\mathbf{w}^{b,(k-1)} \right)^T \mathbf{B}^{-1} \left( \delta\mathbf{w}^{(k)} - \delta\mathbf{w}^{b,(k-1)} \right) \\ &\quad + \frac{1}{2} \left( \mathbf{G}^{(k-1)}\delta\mathbf{w}^{(k)} - \delta\mathbf{y}^{o,(k-1)} \right)^T \mathbf{R}^{-1} \left( \mathbf{G}^{(k-1)}\delta\mathbf{w}^{(k)} - \delta\mathbf{y}^{o,(k-1)} \right) \end{aligned}$$

where

$$\delta\mathbf{w}^{b,(k-1)} = \mathbf{w}^b - \mathbf{w}^{(k-1)} \tag{6}$$

can be interpreted as the background estimate of  $\delta\mathbf{w}^{(k)}$ , and

$$\delta\mathbf{y}^{o,(k-1)} = \mathbf{y}^o - G(\mathbf{w}^{(k-1)}) \tag{7}$$

is a 4D vector of observation-minus-reference state differences. The standard incremental formulation described in the previous section corresponds to the special case  $N_k = 1$ . In this outer-loop framework, the quadratic minimization problem is called the inner loop and the iterations of the CG solver are called inner iterations.

On the first outer iteration

$$\delta\mathbf{w}^{b,(0)} = \mathbf{0}$$

since

$$\mathbf{w}^{(0)} = \mathbf{w}^b. \tag{8}$$

On the first inner iteration of each outer iteration, the increment is initialized to zero,

$$\delta\mathbf{w}_{(0)}^{(k)} = \mathbf{0}.$$

The total number of inner iterations may be different from one outer iteration to the next. After the final inner iteration on each outer iteration, the reference state is updated with the increment,

$$\mathbf{w}^{(k)} = \mathbf{w}^{(k-1)} + \delta\mathbf{w}_{(m_k)}^{(k)} \tag{9}$$

where  $m_k$  is the number of inner iterations performed on outer iteration  $k$ . In what follows, we will drop the subscript  $m_k$  for clarity of notation. For the next outer iteration  $k$ , the background estimate of the increment can be written, using (6), (8) and (9), in terms of the sum of the increments from all previous outer iterations,

$$\delta\mathbf{w}^{b,(k-1)} = -\sum_{l=1}^{k-1} \delta\mathbf{w}^{(l)}.$$

The reference state after the final outer iteration is

$$\mathbf{w}^{(N_k)} = \mathbf{w}^{(N_k-1)} + \delta\mathbf{w}^{(N_k)} = \mathbf{w}^b + \sum_{k=1}^{N_k} \delta\mathbf{w}^{(k)}.$$

The control vector  $\mathbf{w}^a = \mathbf{w}^{(N_k)}$  is the analysis, and

$$\delta\mathbf{w}^a = \sum_{k=1}^{N_k} \delta\mathbf{w}^{(k)} = -\delta\mathbf{w}^{b,(N_k)}$$

is the analysis increment.

A generalization of incremental 4D-Var commonly used in atmospheric data assimilation involves the application of a simplification or filtering operator to reduce the resolution of the increment in the inner loop relative to the resolution of the state vector in the outer loop. This feature is used to reduce the computational cost of incremental 4D-Var, notably in the integration of the TL and adjoint models where most of the cost is incurred. The state vector is updated with the lower resolution increment using an interpolation operator which is typically an approximation of a generalized inverse of the simplification operator. While this algorithmic feature has been developed for NEMOVAR, it has not been applied in Ocean-S4 which employs identical resolution in the inner and outer loops, and an alternative simplification as described in section 2.5.2.

## 2.4 Inner-loop minimization algorithm and preconditioner

The CG minimization software used in NEMOVAR is a close variant of the CONGRAD routine (Fisher, 1998) used in the Integrated Forecast System (IFS). CONGRAD is based on a Lanczos implementation of a standard preconditioned CG algorithm (Golub and Van Loan (1996), p. 431). To reduce the effects of roundoff error, CONGRAD includes a complete reorthogonalization of the Lanczos vectors using a Gram-Schmidt procedure, as discussed in Fisher (1998).

In NEMOVAR the **B** operator is used as a first-level preconditioner for the CG minimization. The optimal preconditioner would be the inverse of the Hessian matrix of  $J^{(k)}$  since this would result in convergence of the CG minimization in a single iteration. The **B** preconditioner corresponds to the inverse of the Hessian of the background term only. In the **B**-preconditioned space, the eigenvalues of the Hessian matrix are bounded below by 1 and have a cluster of eigenvalues at 1 with size at least  $\max(0, n - p)$  where  $n$  is the size of the control vector and  $p$  is the size of the observation vector (Tshimanga *et al.*, 2008). When  $p \ll n$ , which is typically the case in ocean data assimilation, this preconditioner can significantly improve the convergence properties of the minimization. For example, with a single observation ( $p = 1$ ) the minimization will converge in one iteration.

With CONGRAD, **B** preconditioning is achieved through a change of variables

$$\mathbf{v}^{(k)} = \mathbf{U}^{-1} \left( \delta\mathbf{w}^{(k)} - \delta\mathbf{w}^{b,(k-1)} \right)$$

where  $\mathbf{B} = \mathbf{U}\mathbf{U}^T$ . In terms of  $\mathbf{v}^{(k)}$  the cost function becomes

$$J^{(k)}[\mathbf{v}^{(k)}] = \frac{1}{2} \left( \mathbf{v}^{(k)} \right)^T \mathbf{v}^{(k)} + \frac{1}{2} \left( \mathbf{G}^{(k-1)} \delta \mathbf{w}^{(k)} - \delta \mathbf{y}^{o,(k-1)} \right)^T \mathbf{R}^{-1} \left( \mathbf{G}^{(k-1)} \delta \mathbf{w}^{(k)} - \delta \mathbf{y}^{o,(k-1)} \right)$$

where

$$\delta \mathbf{w}^{(k)} = \mathbf{U} \mathbf{v}^{(k)} + \delta \mathbf{w}^{b,(k-1)}. \quad (10)$$

The CG minimization is performed directly in  $\mathbf{v}^{(k)}$ -space. At the end of the minimization, the optimized  $\mathbf{v}^{(k)}$  is transformed back into control space using (10).

For multiple outer iteration problems, the NEMOVAR version of CONGRAD also includes second-level preconditioners based on the Ritz and LBFGS limited-memory preconditioners ([Tshimanga et al., 2008](#)).

## 2.5 Simplifications

### 2.5.1 Linearized balance

The nonlinear balance operator  $K(\cdot)$  is required by the generalized observation operator  $G(\cdot)$  in (7) to compute the model counterpart of the observation vector given the reference state  $\mathbf{w}^{(k-1)}$ . Through successive linearizations about  $\mathbf{w}^{(l)}$ ,  $l = 0, \dots, k-2$ , this operator can be approximated by

$$\mathbf{x}^{(k-1)} = K\left(\mathbf{w}^{(k-1)}\right) \approx K\left(\mathbf{w}^{(0)}\right) + \sum_{l=1}^{k-1} \mathbf{K}^{(l-1)} \delta \mathbf{w}^{(l)} = \mathbf{x}^b + \sum_{l=1}^{k-1} \delta \tilde{\mathbf{x}}^{(l)}$$

where  $\mathbf{x}^b = K\left(\mathbf{w}^{(0)}\right)$  is the background estimate of the initial state vector, and

$$\delta \tilde{\mathbf{x}}^{(k)} = \mathbf{K}^{(k-1)} \delta \mathbf{w}^{(k)} \quad (11)$$

is an approximation of the increment in model space. With this approximation, only the sequence of *linearized* balance operators  $\mathbf{K}^{(k-1)}$  are required to iterate the incremental algorithm. This approximation has been adopted in NEMOVAR as discussed further in Section 3.1.

### 2.5.2 Linearized model propagator

An approximate TL propagator  $\mathbf{M}^{(k-1)}(t_i, t_0)$  and its adjoint  $(\mathbf{M}^{(k-1)}(t_i, t_0))^T$  have been developed for NEMO at INRIA/LJK ([A. Vidard, personal communication](#)). The main simplifications are in the physical parameterizations (vertical and isopycnal mixing) as in the simplified TL and adjoint codes developed by [Weaver et al. \(2003\)](#) for an earlier version of NEMO (OPA). The forward integration of  $\mathbf{M}^{(k-1)}(t_i, t_0)$ , which is needed to evaluate  $J^{(k)}$ , and the backward integration of its adjoint  $(\mathbf{M}^{(k-1)}(t_i, t_0))^T$ , which is needed to compute  $\nabla_{\delta \mathbf{x}^{(k-1)}} J^{(k)}$ , are the most expensive steps of each inner iteration in 4D-Var. A computationally efficient variant of incremental 4D-Var can be obtained by setting  $\mathbf{M}^{(k-1)}(t_i, t_0)$  to the simplest possible linear model, the identity operator  $\mathbf{I}$ . The same increment is thus used to compute the model equivalent of the observation-minus-reference state differences at all times in the assimilation window, but the reference state is still evolved with the full nonlinear model in the outer loop to compute the observation-minus-reference state differences themselves. The resulting algorithm, which is known as 3D-Var FGAT, has the characteristics of 3D-Var in the inner loop but retains features of 4D-Var in the outer loop. The Ocean-S4 implementation of NEMOVAR is based on 3D-Var FGAT.

### 2.5.3 Initialization

Only one outer iteration is used in Ocean-S4. For clarity, we can then drop the outer iteration counter ( $k$ ) on vectors and matrix operators and adopt the simpler notation of section 2.2. Given the analysis at initial time,

$$\mathbf{x}^a = \mathbf{x}^b + \delta\tilde{\mathbf{x}}^a$$

where  $\delta\tilde{\mathbf{x}}^a$  is given by (11) (with  $k = 1$ ), the analyzed state at any time in the assimilation window  $t_0 \leq t_i \leq t_N$  can be defined as

$$\mathbf{x}^a(t_i) = M(t_i, t_0)[\mathbf{x}^a(t_0)]$$

where  $\mathbf{x}^a(t_0) = \mathbf{x}^a$ . Directly initializing the model with an analysis computed from 3D-Var, however, can lead to unphysical fast adjustment processes that can contaminate  $\mathbf{x}^a(t_i)$ . The balance operator can alleviate this problem to some extent but is generally not sufficient. In 4D-Var, the linear dynamical propagator provides an additional (4D) balance constraint which makes direct initialization much less problematic than in 3D-Var. To reduce the initialization shock from 3D-Var, the Incremental Analysis Updates (IAU; Bloom *et al.* (1996)) procedure is employed in NEMOVAR. With IAU, the analysis increment is included as an additional forcing term in the prognostic equation. The resulting analysis trajectory can thus be expressed as

$$\mathbf{x}^a(t_i) = M(t_i, t_{i-1})[\mathbf{x}^a(t_{i-1}), F_i \delta\tilde{\mathbf{x}}^a] \quad (12)$$

where  $\mathbf{x}^a(t_0) = \mathbf{x}^b$  and  $F_i$  is a weighting function such that  $\sum_{i=1}^N F_i = 1$ , with  $N$  denoting the number of time-steps in the assimilation window. The forcing term acts as a low-pass time filter.

## 3 The NEMOVAR system: components

### 3.1 Balance operator

Central to the general formulation of the cost function (1) is the balance operator  $K(\cdot)$  that transforms the control vector  $\mathbf{w} = (T, S_U, \eta_U, u_U, v_U)^T$  of assumed mutually uncorrelated variables into the initial state vector  $\mathbf{x} = (T, S, \eta, u, v)^T$ . (For notational convenience, we have dropped the initial time coordinate  $t_0$  from  $\mathbf{x}$ ). The balance operator in NEMOVAR uses potential temperature  $T$  as the starting point to establish the balanced parts for the other variables. It is based on the formulation described in Weaver *et al.* (2005) with some modifications to the density and SSH balance as detailed below. As in the previous section, the superscript ( $k$ ) will be dropped for clarity.

The set of balance relationships that define  $\mathbf{x} = K(\mathbf{w})$  can be summarized by the sequence of equations

$$\left. \begin{aligned} T &= T \\ S &= K_{S,T}(T) + S_U = S_B + S_U \\ \eta &= K_{\eta,\rho} \rho + \eta_U = \eta_B + \eta_U \\ u &= K_{u,p} p + u_U = u_B + u_U \\ v &= K_{v,p} p + v_U = v_B + v_U \end{aligned} \right\} \quad (13)$$

where

$$\left. \begin{aligned} \rho &= K_{\rho,TS}(T, S) \\ p &= K_{p,\rho} \rho + K_{p,\eta} \eta \end{aligned} \right\} \quad (14)$$

are relations for the diagnostic variables of density and pressure, respectively. The operators  $K_{S,T}(\cdot)$  and  $K_{\rho,TS}(\cdot)$  in italic font are *nonlinear* balance transformations from temperature into salinity, and from temperature and salinity into density. The operators  $K_{Y,X}$  in roman font are *linear* balance transformations from the variable  $X$  into the variable  $Y$ . Variables with a subscript B (U) denote the balanced (unbalanced) component of that variable. It is important to remark that the unbalanced component of a given variable can contribute to the balanced component of another variable following it in the balance sequence; it is only strictly independent of the variables *preceding* it in the balance sequence. For example, in (13) and (14),  $S_U$  is independent of  $T$  but is related to  $\eta$  through  $\rho$ .

The linearized balance equations that transform the control vector increment  $\delta\mathbf{w} = (\delta T, \delta S_U, \delta \eta_U, \delta u_U, \delta v_U)^T$  into a state vector increment  $\delta\mathbf{x} = (\delta T, \delta S, \delta \eta, \delta u, \delta v)^T$  can be summarized as

$$\left. \begin{aligned} \delta T &= \delta T \\ \delta S &= K_{S,T}^b \delta T + \delta S_U = \delta S_B + \delta S_U \\ \delta \eta &= K_{\eta,\rho} \delta \rho + \delta \eta_U = \delta \eta_B + \delta \eta_U \\ \delta u &= K_{u,p} \delta p + \delta u_U = \delta u_B + \delta u_U \\ \delta v &= K_{v,p} \delta p + \delta v_U = \delta v_B + \delta v_U \end{aligned} \right\}$$

where

$$\left. \begin{aligned} \delta \rho &= K_{\rho,T}^b \delta T + K_{\rho,S}^b \delta S \\ \delta p &= K_{p,\rho} \delta \rho + K_{p,\eta} \delta \eta \end{aligned} \right\}.$$

The superscript ‘b’ on  $K_{S,T}^b$ ,  $K_{\rho,T}^b$  and  $K_{\rho,S}^b$  indicates that these operators have been linearized about the background state.

### 3.1.1 Salinity balance

The linearized transformation between increments of temperature and salinity is a local adjustment that is designed, following [Troccoli and Haines \(1999\)](#), to preserve approximately the water-mass properties of the background state by making vertical displacements ( $\delta z$ ) of salinity in response to changes in temperature. The expression for the balanced part of the salinity increment is given by [\(Ricci et al., 2005\)](#)

$$\delta S_B = \gamma_S^b \left( \frac{\partial S}{\partial z} \right)^b \delta z \quad (15)$$

where

$$\delta z = \left( \frac{\partial z}{\partial T} \right)^b \delta T. \quad (16)$$

The coefficient  $\gamma^b$  is set to one except at points where temperature and salinity are expected to be weakly correlated, where it is set to zero. In particular,

$$\gamma_S^b = \begin{cases} 0 & \text{if } z < -D_{ml}^b \\ 0 & \text{if } |(\partial S / \partial z)^b| / |(\partial T / \partial z)^b| > 1 \text{ psu } ^\circ\text{C}^{-1} \\ 0 & \text{if } (\partial T / \partial z)^b < 0.001^\circ\text{C m}^{-1} \\ 1 & \text{otherwise} \end{cases}, \quad (17)$$

where the first condition defines the mixed layer region ( $D_{ml}$  is the depth of the mixed layer defined according to a density difference criterion); the second condition defines regions, such as the barrier layer, where salinity

is strongly stratified but temperature is well mixed; and the third condition defines regions of weak temperature stratification in order to avoid problems in evaluating  $\delta z$  in (16) where division by  $(\delta T / \delta z)^b$  is required. When  $\gamma^b = 0$ , the salinity increment is described entirely by its unbalanced component. The vertical derivatives in (15) and (16) are estimated using a cubic spline. Finally, a local two-gridpoint Shapiro filter is used to smooth the  $T$ - $S$  balance coefficients in each model level. The impact of this linear  $T$ - $S$  balance in 3D-Var has been evaluated in detail by [Ricci et al. \(2005\)](#). In particular, they showed that the  $T$ - $S$  constraint had a significant positive impact on velocity as well as salinity compared to a 3D-Var experiment in which no  $T$ - $S$  constraint was applied. Note that since the  $T$ - $S$  balance coefficients depend on the background state, they evolve from one assimilation cycle to the next.

### 3.1.2 SSH balance

The SSH increment is partitioned into a *baroclinic* component  $\delta\eta_B$  that is balanced with  $\delta T$ ,  $\delta S$  and a *depth-dependent* geostrophic velocity increment  $(\delta\hat{u}_B, \delta\hat{v}_B)$ , and a *barotropic* component  $\delta\eta_U$  that is balanced with a *depth-independent* geostrophic velocity increment  $(\delta\bar{u}_B, \delta\bar{v}_B)$  (see later). The baroclinic SSH component is estimated by computing the dynamic height at the surface  $z = 0$  relative to a reference depth  $z_{\text{ref}}$ :

$$\delta\eta_B = - \int_{z_{\text{ref}}}^0 (\delta\rho/\rho_0) dz \quad (18)$$

where  $\rho_0 = 1020 \text{ kg m}^{-3}$  is a reference density, and

$$\frac{1}{\rho_0} \delta\rho = -\alpha^b \delta T + \beta^b \delta S \quad (19)$$

is a linearized equation of state. Equation (18) can be viewed as an approximation of a more accurate expression for the baroclinic component of SSH involving the solution of an elliptic equation ([Fukumori et al., 1998](#)). The alternative expression accounts for variations in topography and is independent of a reference depth, but is more complicated and costly than the simpler balance (18). The simpler expression has been chosen for Ocean-S4.

Coefficients  $\alpha^b$  and  $\beta^b$  in (19) are the thermal expansion and saline contraction coefficients, respectively, functions of both the background potential temperature, salinity and depth. (Recall that here  $T$  stands for potential temperature, and therefore the above coefficients are not the ordinary thermal expansion and saline contraction coefficients defined as functions of *in situ* temperature and salinity, as discussed in [Gill \(1982\)](#)). Two formulations have been tested for the estimation of the coefficients  $\alpha^b$  and  $\beta^b$  in (19). The first formulation derives these coefficients through a direct tangent-linearization of the equation of state of [Jackett and McDougall \(1995\)](#):

$$\alpha^b = -\frac{1}{\rho_0} \left( \frac{\partial \rho}{\partial T} \right)^b \quad \text{and} \quad \beta^b = \frac{1}{\rho_0} \left( \frac{\partial \rho}{\partial S} \right)^b.$$

In combination with the SSH balance (18), this formulation tends to favour density increments at deeper levels, where the levels are thicker, when assimilating altimeter data. The vertical stratification is not considered explicitly in this formulation, although some provision can be made for it indirectly by specifying depth-dependent background-error standard deviations (see the discussion in section 4.6). This formulation was proposed by [Weaver et al. \(2005\)](#) and used in the global-ocean assimilation experiments of [Daget et al. \(2009\)](#) (which did not assimilate altimeter data as in Ocean-S4).

The second formulation attempts to take into account the vertical stratification of the background density field. This is done using the background value of the Brunt Väisälä frequency ([Gill, 1982](#)),

$$(N^b)^2 = g\beta^b \left\{ \frac{\alpha^b}{\beta^b} \left( \frac{\partial T}{\partial z} \right)^b - \left( \frac{\partial S}{\partial z} \right)^b \right\} \quad (20)$$

where the ratio  $\alpha^b/\beta^b$  and  $\beta^b$  are defined from the polynomial expressions of McDougall (1987), evaluated using the background temperature and salinity fields. Rearranging (20) gives

$$\alpha^b = \gamma_T^b \beta^b \left\{ \frac{(N^b)^2}{g \beta^b} + \left( \frac{\partial S}{\partial z} \right)^b \right\} \left( \frac{\partial z}{\partial T} \right)^b \quad (21)$$

where the coefficient  $\gamma_T^b$  has been introduced to account for the case when  $(\partial T/\partial z)^b$  is small. In particular,

$$\gamma_T^b = \begin{cases} 0 & \text{if } (\partial T/\partial z)^b < 0.02^\circ\text{C m}^{-1} \\ 1 & \text{otherwise} \end{cases},$$

which is a stricter variation of the third condition in (17).

Substituting (21) into (19), and writing  $\delta S = \delta S_B + \delta S_U$  with  $\delta S_B$  given by (15) and (16), yields

$$\frac{1}{\rho_0} \delta \rho = -\gamma_T^b \frac{(N^b)^2}{g} \left( \frac{\partial z}{\partial T} \right)^b \delta T - (\gamma_T^b - \gamma_S^b) \beta^b \left( \frac{\partial S}{\partial z} \right)^b \left( \frac{\partial z}{\partial T} \right)^b \delta T + \beta^b \delta S_U. \quad (22)$$

The second term in the above vanishes at most points. It can be made to vanish identically by using the same threshold conditions for  $\gamma_T^b$  and  $\gamma_S^b$ , which would be reasonable but has not been done. Both this term and the third term which involves the unbalanced salinity increment have been neglected in the current formulation. The normalized density increment is thus approximated by

$$\frac{1}{\rho_0} \delta \rho \approx -\gamma_T^b \frac{(N^b)^2}{g} \delta z \quad (23)$$

where  $\delta z$  is given by (16). The above expression illustrates that this formulation follows the Cooper and Haines (1996) scheme, by which changes in the SSH correspond to vertical displacements of the water column. More generally, the explicit use of  $(N^b)^2$  in the linearized density (22) encourages density increments in strongly stratified regions when assimilating altimeter data, and preserves static stability. Some work is still needed in NEMOVAR to streamline the choice of parameters in (22).

### 3.1.3 Velocity balance

Away from the equator, the balanced part  $(\delta u_B, \delta v_B)$  of the horizontal velocity increment is assumed to be geostrophic; i.e., proportional to the horizontal gradient of pressure divided by the Coriolis parameter  $f$ . The unbalanced part  $(\delta u_U, \delta v_U)$  corresponds to the ageostrophic component of the velocity increment. The pressure increment at a given depth  $z$  can be computed by integrating the hydrostatic equation from  $z$  to the free surface  $\delta \eta$ . After substituting  $\delta \eta = \delta \eta_B + \delta \eta_U$  where  $\delta \eta_B$  is given by (18), this yields

$$\delta p = - \int_{z_{\text{ref}}}^z \delta \rho g ds + \rho_0 g \delta \eta_U. \quad (24)$$

The reference level is set to the ocean bottom,  $z_{\text{ref}} = -H$ , which is spatially varying. The horizontal gradient of the first term in (24) is associated with the baroclinic geostrophic velocity increment  $(\delta \hat{u}_B, \delta \hat{v}_B)$ , while the horizontal gradient of the second term is associated with the barotropic geostrophic velocity  $(\delta \bar{u}_B, \delta \bar{v}_B)$ , where  $(\delta u_B, \delta v_B) = (\delta \hat{u}_B + \delta \bar{u}_B, \delta \hat{v}_B + \delta \bar{v}_B)$  is the total balanced velocity increment.

Special care is required near the equator where the  $f$ -plane geostrophic relation becomes singular. Following Burgers *et al.* (2002), the zonal component  $\delta u_B$  is taken to be geostrophically balanced at the equator while

the meridional component is reduced to zero. Geostrophic balance for  $\delta u_B$  is imposed near the equator using a  $\beta$ -plane approximation combined with a modified pressure increment  $\delta \tilde{p} = \delta p + \delta p_c$  to enforce a symmetric pressure increment about the equator (Picaut and Tournier, 1991; Lagerloef *et al.*, 1999). The correction term is given by

$$\delta p_c = -\phi \left( \frac{\partial \delta p}{\partial \phi} \right)_{\phi=0} \exp(-\phi^2/L_p^2) \quad (25)$$

where  $\phi$  is latitude, and  $L_p$  is an e-folding scale whose size is of the order of the equatorial Rossby radius of deformation. Equation (25) is independent of longitude  $\lambda$  (and thus does not affect  $\delta v_B$ ) and is negligible far from the equator. The modified pressure increment satisfies both the  $\beta$ -plane constraint,

$$\left( \frac{\partial^2 \tilde{p}}{\partial \phi^2} \right)_{\phi=0} = \left( \frac{\partial^2 p}{\partial \phi^2} \right)_{\phi=0}$$

and the necessary condition for geostrophic balance at the equator,

$$\left( \frac{\partial \tilde{p}}{\partial \phi} \right)_{\phi=0} = 0.$$

To allow for a smooth transition between the equatorial ( $\beta$ -plane) geostrophic velocity and the standard ( $f$ -plane) geostrophic velocity away from the equator, weighting functions  $W_\beta = \exp(-\phi^2/L_\beta^2)$  and  $W_f = 1 - W_\beta$  are introduced, where  $L_\beta$  is an e-folding scale analogous to  $L_p$  in the pressure correction term. Here,  $L_\beta$  is set to  $2.2^\circ$  following Lagerloef *et al.* (1999). The same value is used for  $L_p$ . At the equator,  $W_\beta = 1$  and  $W_f = 0$ , while far away from the equator,  $W_\beta \approx 0$  and  $W_f \approx 1$ . The complete expression for the balanced velocity increments in geographical coordinates is given by

$$\begin{aligned} \delta u_B &= -\frac{1}{\rho_0} \left( \frac{W_f}{f} + \frac{W_\beta}{\beta} \frac{1}{a} \frac{\partial}{\partial \phi} \right) \left( \frac{1}{a} \frac{\partial \delta \tilde{p}}{\partial \phi} \right), \\ \delta v_B &= \frac{1}{\rho_0} \frac{W_f}{f} \frac{1}{a \cos \phi} \frac{\partial \delta \tilde{p}}{\partial \lambda} \end{aligned}$$

where  $a$  is the radius of earth,  $f = 2\Omega \sin \phi$  is the Coriolis parameter,  $\Omega$  is the rotation rate of earth, and  $\beta = \partial f / \partial(a\phi) = (2\Omega \cos \phi)/a$ . The  $\beta$ -plane term is sensitive to sharp gradients in the density field which can arise from the on/off nature of the salinity balance in regions of weak stratification. To alleviate this problem, the salinity increment is not included in the evaluation of the second-order derivative of the pressure increment.

It is questionable whether the simple balance relationships described above are appropriate close to boundaries where nonlinear and frictional effects are important. In these regions, it may be better to apply no balance at all rather than an erroneous one. In Ocean-S4, the  $T-S$  and SSH balance are applied everywhere, whereas the velocity balance is modified near boundaries according to

$$\left. \begin{aligned} \delta u_B &\mapsto A_c(r) \delta u_B \\ \delta v_B &\mapsto A_c(r) \delta v_B \end{aligned} \right\}$$

where

$$A_c(r) = \begin{cases} 1 + \frac{1}{2}(\delta_c - 1)(1 + \cos(\pi r/R_c)) & \text{if } 0 \leq r \leq R_c \\ 1 & \text{if } r > R_c \end{cases}, \quad (26)$$

$r = r(z)$  being the Cartesian distance between the model grid-point and the nearest model coastline point, which depends on the depth  $z$  of each model level,  $\delta_c = A_c(0)$  the value of the weight directly at the coastline, and  $R_c$  the distance from the coastline beyond which the weight is set to one. In Ocean-S4,  $\delta_c = 0$  and  $R_c = 200$  km.

### 3.2 Background-error covariances

The Ocean-S4 implementation of NEMOVAR is based on 3D-Var FGAT, as already mentioned. The observations that are assimilated are temperature and salinity profiles, and altimeter-derived sea-level anomalies. There is no information on velocity other than that provided by the balance operator, so the unbalanced components  $\delta u_U$  and  $\delta v_U$  of the control vector can be ignored. The active control variables are thus  $\delta T$ ,  $\delta S_U$  and  $\delta \eta_U$ .

The background-error covariance matrix is assumed to be block-diagonal with respect to  $\delta T$ ,  $\delta S_U$  and  $\delta \eta_U$ . It can be factored in standard form as

$$\mathbf{B} = \mathbf{D}^{1/2} \mathbf{C} \mathbf{D}^{1/2} \quad (27)$$

where  $\mathbf{D}^{1/2}$  is a diagonal matrix of standard deviations, and  $\mathbf{C}$  is a correlation matrix. For convenience, the correlation matrix is constructed as a symmetric product of a ‘square-root’ matrix  $\mathbf{C}^{1/2}$  and its transpose  $(\mathbf{C}^{1/2})^T = (\mathbf{C}^T)^{1/2}$ . The preconditioning transformation in (10) can then be identified as

$$\mathbf{U} = \mathbf{D}^{1/2} \mathbf{C}^{1/2}.$$

Each block component of  $\mathbf{U}$  can be written as  $\mathbf{U}_X = \mathbf{D}_X^{1/2} \mathbf{C}_X^{1/2}$  where  $X = T, S_U$  or  $\eta_U$ . The specification of the standard deviation matrices  $\mathbf{D}_X^{1/2}$  for Ocean-S4 is described in section 4.6. The correlations are assumed to be approximately Gaussian and modelled in grid-point space using a normalized diffusion operator (Weaver and Courtier, 2001). For each variable, the square-root factor has the form

$$\mathbf{C}_X^{1/2} = \mathbf{\Gamma}_X^{1/2} \mathbf{L}_X^{1/2} \mathbf{W}_X^{-1/2}$$

where  $\mathbf{L}_X^{1/2}$  is a centred finite-difference representation of an explicitly-formulated diffusion operator  $\mathcal{L}_X^{1/2}$ ,  $\mathbf{W}_X^{-1/2}$  is a diagonal matrix containing the inverse of the square-root of the volume or area elements represented by  $X$ , and  $\mathbf{\Gamma}_X^{1/2}$  is a diagonal matrix of normalization factors to ensure that the diagonal elements of  $\mathbf{C}_X = \mathbf{C}_X^{1/2} (\mathbf{C}_X^{1/2})^T$  are approximately equal to one. These factors are estimated using a randomization technique (Weaver and Courtier, 2001). For  $T$  and  $S_U$ ,  $\mathcal{L}_X^{1/2}$  is a 3D operator that is formulated as a product  $\mathcal{L}_X^{v1/2} \mathcal{L}_X^{h1/2}$  of a horizontal diffusion operator  $\mathcal{L}_X^{h1/2}$  acting along geopotential surfaces ( $z$  levels) in the spherical-shell domain described by geographical coordinates  $(\lambda, \phi, z)$ , and a vertical diffusion operator  $\mathcal{L}_X^{v1/2}$  acting perpendicular to geopotential surfaces. For  $\eta_U$ ,  $\mathcal{L}_X^{1/2}$  is a 2D operator defined by  $\mathcal{L}_X^{h1/2}$ . Specifically,

$$\mathcal{L}_X^{h1/2} = (1 + \nabla_h \cdot \boldsymbol{\kappa}_X^h \nabla_h)^{M_X^h/2} \quad (28)$$

$$\text{and} \quad \mathcal{L}_X^{v1/2} = \left( 1 + \frac{\partial}{\partial z} \kappa_X^v \frac{\partial}{\partial z} \right)^{M_X^v/2}. \quad (29)$$

where  $\nabla_h \cdot$  and  $\nabla_h$  denote the horizontal divergence and gradient operators, respectively. The length-scales of the quasi-Gaussian kernels of the operators (28) and (29) are assumed to be spatially dependent and are controlled by the horizontal diffusion tensor  $\boldsymbol{\kappa}_X^h$  and vertical diffusion coefficient  $\kappa_X^v$ . At each point  $(\lambda, \phi)$ , the horizontal diffusion tensor  $\boldsymbol{\kappa}_X^h$  is a  $2 \times 2$  matrix. For simplicity, this matrix is taken to be diagonal. In geographical coordinates the Laplacian term can thus be written as

$$\nabla_h \cdot \boldsymbol{\kappa}_X^h \nabla_h = \frac{1}{a^2 \cos^2 \phi} \frac{\partial}{\partial \lambda} \left( \kappa_X^\lambda \frac{\partial}{\partial \lambda} \right) + \frac{1}{a^2 \cos \phi} \frac{\partial}{\partial \phi} \left( \kappa_X^\phi \cos \phi \frac{\partial}{\partial \phi} \right).$$

The diagonal elements  $\kappa_X^\lambda$  and  $\kappa_X^\phi$  can have different values in order to accommodate local anisotropic length-scale ratios different from one. This feature is used in Ocean-S4 to account for anisotropy near the equator, as described in section 4.6.

The length-scale of a Gaussian correlation function is defined in the usual way (e.g., see p. 110 in [Daley \(1991\)](#) or [Weaver and Mirouze \(2012\)](#) for a more general discussion). The local horizontal length-scales  $L_X^\lambda$  and  $L_X^\phi$  can be assembled in a diagonal tensor  $\mathbf{L}_X^h$  which is a spatially-varying,  $2 \times 2$  matrix whose diagonal elements are the square of these length scales,  $(L_X^\lambda)^2$  and  $(L_X^\phi)^2$ . The local vertical length-scales can be represented by a spatially-varying coefficient  $L_X^z$ . Given estimates of these length-scales, the diffusion parameters can be adjusted according to the relationships

$$\left. \begin{aligned} \kappa_X^h &= \frac{1}{2M_X^h} \mathbf{L}_X^h \\ \kappa_X^v &= \frac{1}{2M_X^v} (L_X^z)^2 \end{aligned} \right\} \quad (30)$$

where the iterations  $M_X^h$  and  $M_X^v$  are chosen to be large enough to keep the explicit scheme stable. The boundary conditions used at solid boundaries and at the ocean surface are of Neumann type. This results in a slight increase of the effective length-scale of the correlation functions near the boundaries, as illustrated in [Mirouze and Weaver \(2010\)](#).

### 3.3 Use of observations in NEMOVAR

NEMOVAR has been designed to assimilate *in situ* temperature and salinity profiles, sea surface temperature (SST) and altimeter-derived sea-level anomalies (SLA) referenced to a model-derived or external Mean Dynamic Topography (MDT). These data-types are all direct observations of the model state variables so the only action of the observation operator  $H_i = \mathbf{H}_i$  is to interpolate the model variables to the measurement locations. The observation times are approximated by the model time step closest to the measurement time. One exception is for temperature data from moored buoys (TAO, TRITON, PIRATA, RAMA) which in the EN3 data-base<sup>2</sup> used by Ocean-S4 are available as daily-averaged quantities. For these data, the observation operator involves averaging the model temperature fields over one day before interpolating them to the observation location. An observation operator to compare with horizontal Eulerian velocity observations has also been developed for diagnostic purposes.

For SST and SLA data, a 2D horizontal interpolation operator is needed. For *in situ* profiles, a 1D vertical interpolation operator is needed in addition to the 2D interpolation operator. Vertical interpolation is achieved using a cubic spline interpolation where the top and bottom boundary conditions for the second derivative of the interpolating polynomial in the spline are set to zero (natural conditions). At the bottom boundary, this is done using the land-ocean mask.

Horizontal interpolation requires special care on the orthogonal curvilinear (ORCA) grid of the global version of NEMO. For horizontal interpolation a bilinear remapping technique is used ([Jones, 2001](#)). The technique is an iterative scheme that determines the weights for interpolation by mapping a quadrilateral model cell into a cell with coordinates (0,0), (1,0), (0,1) and (1,1). Since the ORCA grids are non-trivial, it is not always straightforward to determine which grid-points surround any given observational position, so before the interpolation can be performed, a search algorithm is required to determine the corner points of the quadrilateral cell in which the observation is located. This is the most difficult and time consuming part of the horizontal interpolation procedure. A robust test for determining if an observation falls within a given quadrilateral cell is as follows. Let  $P(\lambda, \phi)$  denote the observation point, and let  $A(\lambda_A, \phi_A)$ ,  $B(\lambda_B, \phi_B)$ ,  $C(\lambda_C, \phi_C)$  and  $D(\lambda_D, \phi_D)$  denote the bottom left, bottom right, top left and top right corner points of the cell, respectively. To determine if  $P$  is inside

<sup>2</sup>Available from <http://www.metoffice.gov.uk/hadobs/en3/>.

the cell, we verify that the cross-products

$$\begin{aligned}\mathbf{r}_{PA} \times \mathbf{r}_{PC} &= [(\lambda_A - \lambda)(\phi_C - \phi) - (\lambda_C - \lambda)(\phi_A - \phi)] \hat{\mathbf{k}} \\ \mathbf{r}_{PB} \times \mathbf{r}_{PA} &= [(\lambda_B - \lambda)(\phi_A - \phi) - (\lambda_A - \lambda)(\phi_B - \phi)] \hat{\mathbf{k}} \\ \mathbf{r}_{PC} \times \mathbf{r}_{PD} &= [(\lambda_C - \lambda)(\phi_D - \phi) - (\lambda_D - \lambda)(\phi_C - \phi)] \hat{\mathbf{k}} \\ \mathbf{r}_{PD} \times \mathbf{r}_{PB} &= [(\lambda_D - \lambda)(\phi_B - \phi) - (\lambda_B - \lambda)(\phi_D - \phi)] \hat{\mathbf{k}}\end{aligned}$$

point in the opposite direction to the unit normal  $\hat{\mathbf{k}}$  (i.e., that the coefficients of  $\hat{\mathbf{k}}$  are negative), where  $\mathbf{r}_{PA}$ ,  $\mathbf{r}_{PB}$ , etc. correspond to the vectors between points P and A, P and B, etc.. In order to speed up the grid search, a lookup table is used which contains the lower and upper bounds of the model coordinates to be searched for on a  $0.5^\circ \times 0.5^\circ$  regular grid ([D. Lea, personal communication](#)).

In numerical models it is common to divide the model grid into subgrids (or domains) where each subgrid is executed on a single processing element with explicit message passing for exchange of information along the domain boundaries when running on a massively parallel processor system. This approach is used by the NEMO ocean model ([Madec, 2008](#)) and has been used in the parallelization of the balance and background-error covariance operators described earlier. For observations there is no natural distribution since the observations are not evenly distributed on the globe. In NEMOVAR there are two options for parallel distribution of observations: 1) a grid-point domain decomposition; or 2) a round-robin distribution with equal number of observations per processing elements.

The first is the simplest option where the observations are distributed according to the domain of the grid-point parallelization. The top panel of Figure 1 shows an example of the distribution of *in situ* data on processing elements in a  $4 \times 2$  decomposition. A different colour has been used to distinguish observations on different processors. The grid-point domain decomposition is clearly visible in the plot. The advantage of this approach is that all information needed for horizontal interpolation is available without any communication between processing elements. Since the NEMO model has only one halo point in all directions, we are limited to using a  $2 \times 2$  grid-point stencil for the interpolation (which is the case for bilinear interpolation). This approach cannot be used for higher order interpolation schemes requiring a  $4 \times 4$  grid-point stencil or for observations which have a footprint covering several horizontal grid-points.

A disadvantage with the above scheme is that the number of observations on each processor can be very different. This can lead to load imbalance if the cost of the actual interpolation is expensive relative to the communication of data needed for interpolation. An alternative approach is to distribute the observations equally among processors (*round-robin* scheme) and use message passing to retrieve the local stencil for interpolation. The bottom panel of Figure 1 shows the distribution of *in situ* data using the round-robin scheme with a  $4 \times 2$  domain decomposition. The observations on a given processing element are now clearly randomly distributed on the globe. Performance testing has shown that for a typical NEMOVAR application the round-robin scheme is faster than the domain-decomposition scheme.

### 3.4 Accounting for model bias

The standard procedure to deal with systematic error in a data assimilation system is to augment the control vector with a set of systematic error or bias variables. In sequential data assimilation, this approach requires two analysis steps: one for the bias estimation and a second for the state vector. Assuming that the bias is nearly constant in time, and that the bias-error covariance matrix is proportional to the background-error covariance matrix, with the proportionality constant small, the algorithm can be approximated so it only requires one analysis step, and thus the bias term can be updated at little extra cost ([Dee, 2005](#)). However the requirement of proportionality between the bias- and background-error covariance matrices is not generally appropriate since

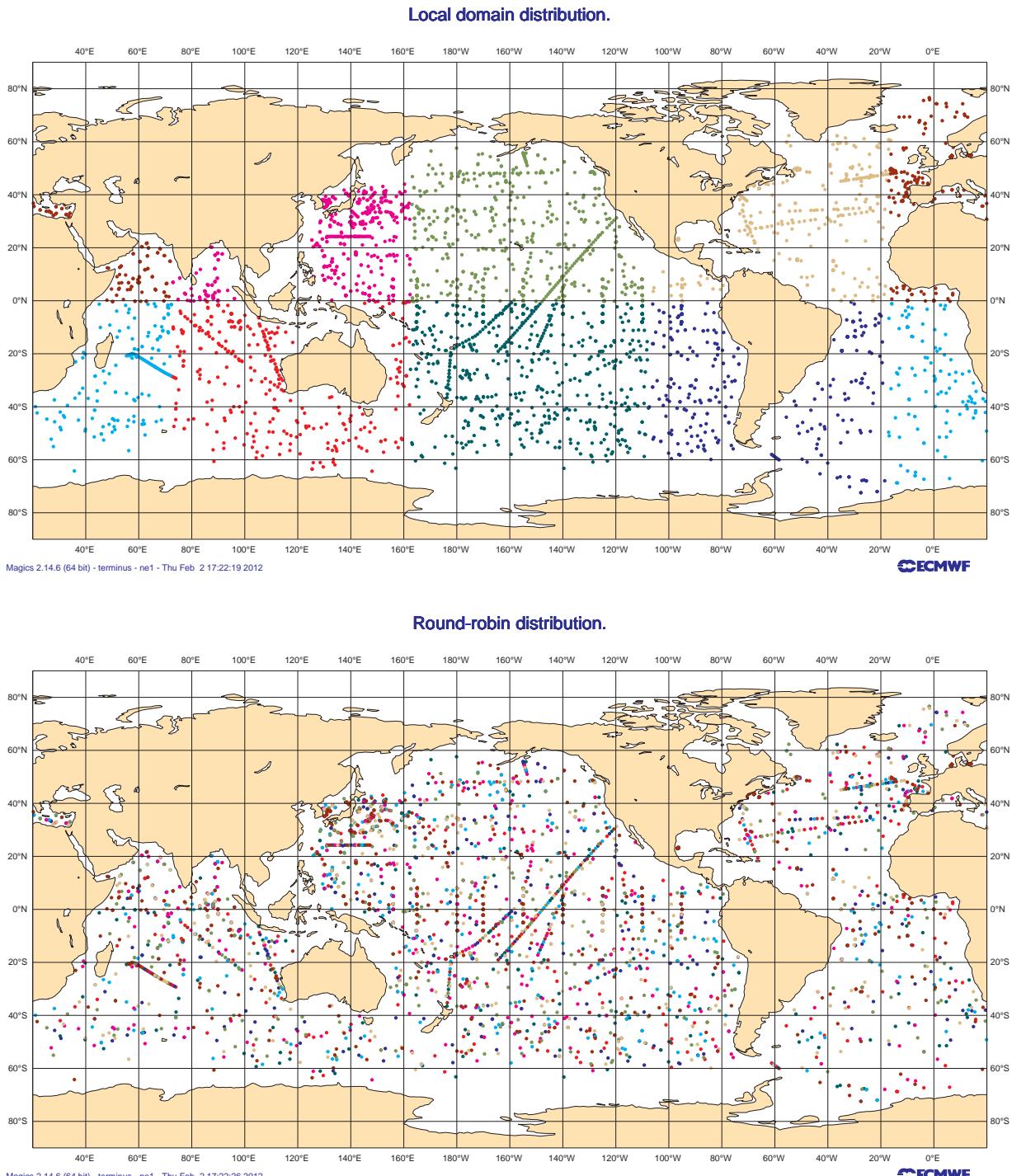


Figure 1: Example of distribution of observations on processing elements with the geographical distribution of observational data for geographical distribution (top) and round-robin (bottom).

the bias and model state vector can have different control variables and/or multivariate balance relationships. A more general approximation of the one-step bias correction algorithm, which allows the bias variables to be different from the state vector variables, was proposed by [Balmaseda et al. \(2007a\)](#) and has been implemented in NEMOVAR. The bias control variables are used to perturb the model tendencies in the tracer (temperature and salinity) and momentum equations. The modifications also include a model for the time evolution of the bias, which now consists of two terms: a bias term estimated *a priori*  $\bar{\mathbf{b}}$ , which can account for seasonal variations, and a bias term estimated online  $\mathbf{b}'_c$ , updated each analysis cycle  $c = 1, 2, \dots$ , so the total bias  $\mathbf{b}_c$  is given by:

$$\mathbf{b}_c = \bar{\mathbf{b}} + \mathbf{b}'_c. \quad (31)$$

In NEMOVAR, the bias vector has four distinct components  $\mathbf{b} = (b^{tr,T}, b^{tr,S}, b^{p,T}, b^{p,S})^T$ , where  $b^{tr,T}$  and  $b^{tr,S}$  are the bias correction terms used to adjust the tendencies in the tracer equations, and  $b^{p,T}$  and  $b^{p,S}$  are the bias corrections terms used to adjust the horizontal pressure gradient terms in the momentum equations.

The bias control vector on cycle  $c$  is estimated from the analysis increments as:

$$\mathbf{b}'_c = \alpha \mathbf{b}'_{c-1} - \mathbf{A} \delta \mathbf{x}_{c-1}^a \quad (32)$$

where  $\alpha$  is a memory factor that also determines a time-scale for the time evolution of the online bias,  $\mathbf{A}$  is a linear transformation from the state vector increment  $\delta \mathbf{x} = (\delta T, \delta S, \delta \eta, \delta u, \delta v)^T$  to the bias control vector, and  $\mathbf{b}'_0 = \mathbf{0}$ .

The memory factor  $\alpha$  is introduced to limit the influence in time of isolated or sporadic observations. A side effect is that the magnitude of the bias can be underestimated, and to compensate for that, an *a priori* bias term is also introduced, which is not affected by the online estimation. The *a priori* term  $\bar{\mathbf{b}}$  can be seasonally varying. The *a priori* term has the potential to provide a smoother analysis by preventing abrupt changes in the analysis associated with the introduction of new observing systems. More details about its estimation and the implementation of the bias in Ocean-S4 are given in section 4.5.

The transformation  $\mathbf{A}$  between assimilation increments and bias components can be written as

$$\mathbf{A} = \begin{bmatrix} a^{tr,T} & 0 & 0 & 0 & 0 \\ 0 & a^{tr,S} & 0 & 0 & 0 \\ a^{p,T} & 0 & 0 & 0 & 0 \\ 0 & a^{p,S} & 0 & 0 & 0 \end{bmatrix} \quad (33)$$

where the individual elements are weights that determine how much an increment component contributes to the different bias variables. These weights are always positive and less than one, and in principle can be spatially dependent: in NEMOVAR, the coefficients in matrix  $\mathbf{A}$  are allowed to change in latitude. It is important to note that in this formulation, the bias correction term modifying the momentum equations is calculated directly from the temperature and salinity increments, instead of being linearly derived from the velocity increments. This allows us to impose a balance between density and momentum tendency increments different from geostrophy, which may be more relevant for the slow time scales (i.e., Sverdrup balance at the equator).

The bias correction is only updated at the beginning of the first outer loop in the assimilation cycle, using the analysis increment from the previous cycle. The bias correction is used to modify the tendencies of the nonlinear model used in the outer loop, so the time evolution of the background and analysis states can be expressed as

$$\begin{aligned} \mathbf{x}_c^b(t_i) &= M(t_i, t_{i-1}) [\mathbf{x}_c^b(t_{i-1}), \mathbf{b}_{c-1}], \\ \mathbf{x}_c^a(t_i) &= M(t_i, t_{i-1}) [\mathbf{x}_c^a(t_{i-1}), \mathbf{b}_{c-1}, F_i \delta \tilde{\mathbf{x}}_c^a] \end{aligned} \quad (34)$$

where the second equation in the above expression is a generalization of (12). More explicitly, the tracer equations for the time evolution of temperature and salinity are modified by adding the respective correction terms

$$\begin{aligned}\Delta T_c &= -\frac{1}{N} b_{c-1}^{tr,T}, \\ \Delta S_c &= -\frac{1}{N} b_{c-1}^{tr,S}\end{aligned}\quad (35)$$

where  $N$  is the number of time steps in the assimilation window. The negative sign multiplying the bias variables highlights that the bias is being removed. The correction in the momentum equations takes place via a nonlinear transformation: modified temperature and salinity fields ( $T_c^p(t_i), S_c^p(t_i)$ ) are used in the computation of the density ( $\rho(T_c^p, S_c^p)$ ) which in turn is used to derive the horizontal pressure gradients:

$$\begin{aligned}T_c^p(t_i) &= T_c(t_i) + b_{c-1}^{p,T}, \\ S_c^p(t_i) &= S_c(t_i) + b_{c-1}^{p,S}.\end{aligned}\quad (36)$$

This correction is only applied in the density calculation used to derive the horizontal baroclinic pressure gradients. The density is modified so as to preserve the original balance in the momentum equations, and therefore attempts to make the changes in density caused by assimilation transparent to the momentum equations (Bell *et al.*, 2004). This explains why there is a positive sign multiplying the bias variables in the above equations.

## 4 Specifications for Ocean-S4

The ECMWF Ocean-S4 consists of two streams: an ocean reanalysis stream (ORAS4), which spans the period September 1957 to present, with a certain delay, and a near-real-time stream (ORTS4), which produces daily and timely initial conditions for the monthly and seasonal forecasting systems. The delay for the ORAS4 stream is determined by the arrival of high quality observational data and has been chosen to be 6 days to allow for retrieval of updated sea surface height products from AVISO<sup>3</sup>. ORTS4 only exists for the operational period, i.e., it does not cover the historical record. These two streams are related, as explained later in this section.

The NEMO model used is based on version 3.0 of the NEMO reference model (Madec, 2008), and the global ORCA1 configuration with 42 vertical levels (ORCA1\_L42), developed jointly by the National Oceanography Centre, Southampton (NOCS) and the Met Office. ORCA1\_L42 uses the same tripolar grid like the other ORCA configurations. It has approximately 1° resolution in the extratropics and refined meridional resolution in the tropics with a minimum value of 0.3° directly at the equator. Of the 42 vertical levels, 18 are in the upper 200 m. The first level has a 10 m thickness. The vertical discretization scheme uses partial steps to have better representation of the flow over steep topography. The Caspian Sea has been removed from the model domain due to problems with the water budget there. There is no prognostic sea-ice model. Instead, the sea-ice concentration is prescribed from an external input file. The surface fluxes (daily fluxes of solar radiation, total heat flux, evaporation-minus-precipitation and surface wind-stress) calculated by an atmosphere model (typically from an atmospheric reanalysis) are read from input files, instead of computed using a bulk-formula within NEMO. The heat flux closure is provided by a Newtonian relaxation to analysed values of sea surface temperature (SST), as described in the next section. The fresh water flux is also adjusted via a relaxation to monthly climatology of surface salinity from World Ocean Atlas 2005 (WOA05, Locarnini *et al.* (2006); Antonov *et al.* (2006)), with a time scale of approximately 1 year. A global adjustment of the fresh-water flux

<sup>3</sup><http://www.aviso.oceanobs.com/en/data/products/sea-surface-height-products/global/sla/index.html>.

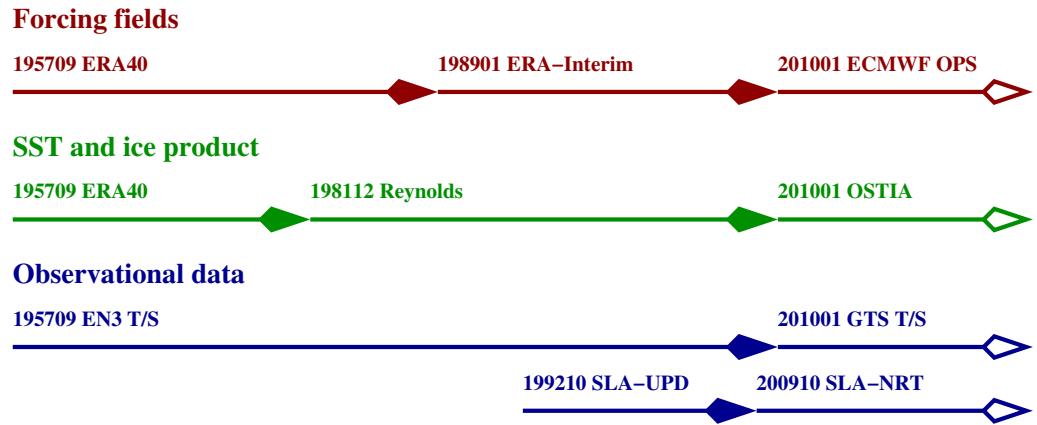


Figure 2: Timeline of changes to the reanalysis forcing and assimilation data-sets for ORAS4.

in order to control global sea-level changes is also applied, as described below. In addition to the bias correction scheme, a weak (20-year time scale) relaxation to temperature and salinity climatological values from WOA05 is applied throughout the water column.

The analysis cycle consists of a single outer iteration of 3D-Var FGAT with the addition of observational quality control (QC) and bias correction steps. The background state trajectory is produced by integrating forward the NEMO model forced by daily fluxes, relaxing to SST analyses and using bias correction as described in the previous section. The model equivalent of each available observation is calculated and a QC of the observations performed. The background state and the quality-controlled observations are passed to the inner loop part of 3D-Var FGAT where the incremental cost function is minimized using the CG algorithm described in section 2.4 to produce the assimilation increment. For Ocean-S4 we have chosen to use a fixed number of 40 CG iterations on each cycle, which typically reduces the cost function and the Euclidean norm of the gradient of the cost function to 0.7 and  $10^{-4}$  of their respective initial values. In the final phase of the analysis cycle, the assimilation increment resulting from the inner-loop minimization is applied using IAU with constant weights ( $F_i = 1/N$ ) during a second model integration spanning the same time window as for the background trajectory (see section 2.5.3). The reanalysis stream ORAS4 uses a 10-day assimilation cycle. The assimilation cycle for the real-time stream ORTS4 varies in time.

For simplicity, we first describe the different elements of ORAS4, which are often shared by ORTS4. The specific changes needed for ORTS4 are summarized in section 4.7. Figure 2 shows schematically the different data-sets used for the production of ORAS4.

#### 4.1 Forcing fields, SST and Sea Ice

The surface fluxes are taken from the ERA40 reanalysis (Uppala *et al.*, 2005) from September 1957 until December 1989, the ERA-Interim reanalysis (Dee *et al.*, 2011) from January 1989 until December 2009, and the ECMWF operational archive from January 2010 onwards, as indicated in Figure 2.

Rather than assimilating SST observational data, a strong relaxation to gridded SST products is applied. The relaxation coefficient is  $-200 \text{ W/m}^2/\text{ }^\circ\text{C}$ , equivalent to about a 2-3 day time scale over a depth of 10 m. Neither of the LIM (The Louvain-la-Neuve-sea Ice Model) ice models embedded in the NEMO system are used with ORAS4. Instead, a gridded sea-ice product is prescribed and the model SST adjusted to be consistent with the sea-ice concentration. If the sea-ice concentration is higher than a certain threshold (55%), the model SST is set

to freezing point. On the other hand, if sea-ice concentration is below the given threshold (i.e., there is no ice in the observations) but the model SST is below freezing point, the strength of the relaxation term to observed SST is increased.

In the early period of the ocean reanalysis (from September 1957 until November 1981) the SST/sea-ice are taken from the ERA40 archive. From December 1981 until December 2009, SST/sea-ice are taken from the NCEP OI\_v2 weekly product ([Reynolds et al., 2002](#)), and from January 2010 onwards the OSTIA SST/sea-ice products ([Stark et al., 2007](#)) are used. There was some choice to be made regarding the SST/sea-ice product to use, and several SST products were compared. Figure 3 shows time series of globally-averaged SST from the following different products.

- OIv2\_1×1: NOAA  $1^\circ \times 1^\circ$  weekly SST product; global coverage; available from late 1981 onwards (<http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html>). This is the SST product currently used in the operational seasonal/monthly forecasting system. Black curve in Figure 3.
- OIv2\_025\_d1: NOAA  $0.25^\circ \times 0.25^\circ$  daily SST; available from late 1981 onwards (<http://www.ncdc.noaa.gov/oa/climate/research/sst/oi-daily.php>). It uses only one instrument, AVHRR. Red line in Figure 3.
- OIv2\_025\_d2: NOAA  $0.25^\circ \times 0.25^\circ$  daily SST; available from late 2002 onwards (<http://www.ncdc.noaa.gov/oa/climate/research/sst/oi-daily.php>). It uses two instruments: AVHRR and AMSR. This is roughly the same observational information as used with the current OSTIA product. Blue line in Figure 3.
- Ersstv3b: NOAA Extended Reconstruction  $2^\circ \times 2^\circ$  monthly SST (<http://www.ncdc.noaa.gov/oa/climate/research/sst/ersstv3.php>); available since 1880. Green line in Figure 3.
- CMIP5\_proto: SST monthly maps recommended by the CMIP5 protocol (<http://pcmdi-cmip.llnl.gov/cmip5/forcing.html#amip>). It is a blend of HadISST1 and the NOAA OIv2\_1×1 weekly product. Purple curve in Figure 3.

The most obvious feature in Figure 3 is the cold bias of the high resolution SST products relative to the low resolution products. The bias is especially pronounced in the case of OIv2\_025\_d1. The operational OSTIA product (not shown) has similar values to the high resolution NOAA SST from 2008 onwards. Note that the curves for OIv2\_1 × 1 and cmip5\_PROTO are virtually indistinguishable.

## 4.2 Temperature and Salinity profiles

For the period 01/09/1957 to 01/01/2010, ORAS4 assimilates temperature and salinity (T/S) profiles from the quality-controlled EN3\_v2a data-set available from <http://www.metoffice.gov.uk/hadobs/en3/>, with XBT depth corrections from [Wijffels et al. \(2009\)](#) (labelled EN3\_v2a\_NoCWT\_WijffelsTable1XBTCorr on the website). These data consist of the following data-types: eXpendable BathyThermographs (XBTs, T only); Conductivity-Temperature-Depth sensors (CTDs, T/S); TAO/TRITON/PIRATA/RAMA moorings (T/S); Argo profilers (T/S); and Autonomous Pinniped Bathymeter (APBs or elephant seals, T/S). In the EN3 data-set the APBs are coded as an Argo data-type, but they can be distinguished via their WMO code number (Q9900XXX). From 01/01/2010, operational data from the Global Telecommunications System (GTS) are used.

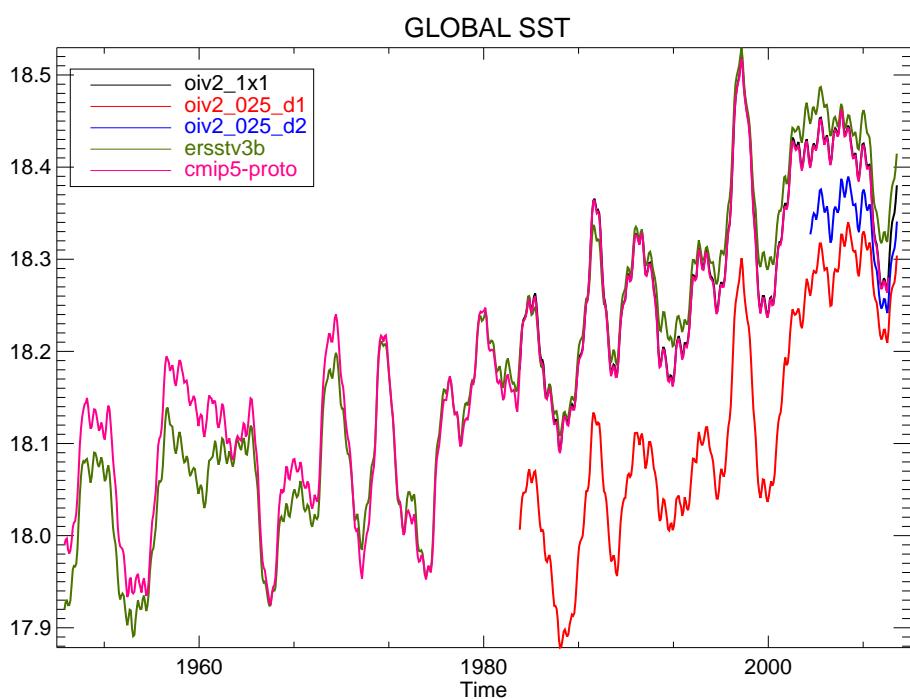


Figure 3: Time series of globally-averaged SST from different products as described in the text. The NOAA OIv2\_025\_d1 is cold with respect to the other products. The OIv2\_025\_d2 is still colder than the low resolution products, but the difference is not so pronounced. The curves for OI\_v21x1 and cmip5-proto are virtually indistinguishable. Units are  $^{\circ}\text{C}$ .

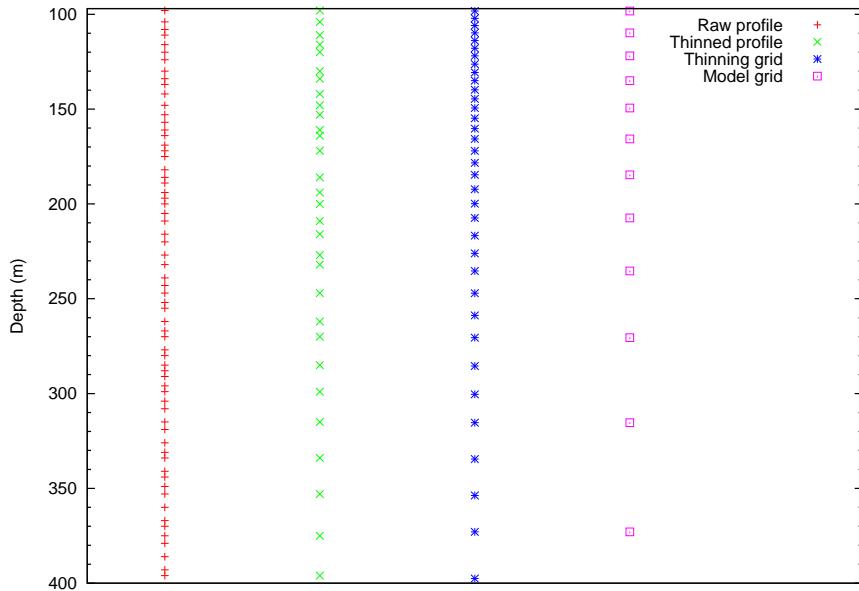


Figure 4: Example of the vertical thinning of observations with high vertical density illustrating that less observations are used where the model levels are further apart.

Both EN3 and GTS are quality controlled by the NEMOVAR QC procedure. The main component of the QC procedure is described by [Ingleby and Huddleston \(2007\)](#), except that the background check here is performed against the model background interpolated T/S values rather than climatology. For the EN3 input data the TAO/TRITON/PIRATA/RAMA moorings are already available as daily mean super-observations. For the GTS data, daily mean super-observations are constructed for each platform in the same way as for the EN3 data. Some model-dependent QC checks and procedures have been implemented in addition to the main QC procedure.

Among these procedures is a vertical thinning of observations to allow no more than three observations per model level. This has been implemented to ensure that data with high vertical resolution (typically CTDs) are not given too much weight in the analysis. In order to maintain the vertical structure of a given profile a thinning level set is constructed which contains the model levels and two additional levels equally distributed between each model level pair. An illustration of the vertical thinning procedure is given in Figure 4.

Another model-dependent QC check is to reject observations in regions where the total model depth is less than 500 m in order to avoid assimilating data on the continental shelves where the model has poor representativeness. For GTS data, different levels can be indicated in different reports for a given platform at a specific observation time. To avoid any ambiguities all reports with the same position, time and WMO identifier are merged into a single report with duplicate levels discarded.

All observation errors are assumed to be uncorrelated in Ocean-S4, so that only the observation-error standard deviations need to be specified. It is customary to define observation error as the sum of measurement error and representativeness error ([Lorenc, 1986](#)). The accuracy of temperature and salinity measurements depends on instrument type. For example, as reported by [Ingleby and Huddleston \(2007\)](#), temperature-error standard

$\sigma_T^o$	$\sigma_S^o$	$\sigma_\eta^o$
$\sigma_T^{\text{sur}} = 0.18^\circ\text{C}$	$\sigma_S^{\text{sur}} = 0.18 \text{ psu}$	$\bar{\sigma}_\eta = 0.05 \text{ m}$
$\sigma_T^{\max} = 1.0^\circ\text{C}$	$\sigma_S^{\text{do}} = 0.02 \text{ psu}$	$N_{\min} = 10$
$\sigma_T^{\text{do}} = 0.07^\circ\text{C}$	$D_5 = 750 \text{ m}$	$\delta_c = 6$
$D_1 = 75 \text{ m}$	$\delta_c = 6$	$R_c = 800 \text{ km}$
$D_2 = 300 \text{ m}$	$R_c = 800 \text{ km}$	
$D_3 = 450 \text{ m}$		
$D_4 = 1000 \text{ m}$		
$\delta_c = 6$		
$R_c = 800 \text{ km}$		

Table 1: Summary of the values used in Ocean-S4 for the observation-error standard deviation parameterizations for temperature and salinity profiles, and altimeter data.

deviations from modern CTDs are approximately  $0.002^\circ\text{C}$  whereas those from XBTs are between  $0.1^\circ\text{C}$  and  $0.15^\circ\text{C}$ . However, in a low resolution model like ORCA1, which does not resolve the ocean mesoscale, observation error will generally be dominated by representativeness error. This will be particularly true in boundary current regions such as the Kuroshio and Gulf Stream. In Ocean-S4, the total observation-error standard deviations  $\sigma_X^o$  for temperature and salinity profiles ( $X = T$  and  $S$ ) are specified according to analytical functions that depend, except near coastlines, on depth only. Functions  $\hat{\sigma}_X^o = \hat{\sigma}_X^o(z)$  have been constructed to provide an approximate fit to the vertical profiles of globally-averaged temperature and salinity  $\sigma_X^o$  estimated by [Ingleby and Huddleston \(2007\)](#). To account for poor representativeness near boundaries, the expressions for  $\sigma_X^o$  also include an inflation coefficient near coastlines:

$$\sigma_X^o = A_c(r) \hat{\sigma}_X^o \quad (37)$$

where  $A_c(r)$  has the same general form as (26), with  $r = r(z)$  denoting the Cartesian distance between the observation point and the nearest model coastline point, and  $\delta_c = A_c(0)$  the value of the inflation factor applied directly at the coastline. The expressions for  $\hat{\sigma}_X^o$  for temperature and salinity are

$$\hat{\sigma}_T^o = \begin{cases} (\sigma_T^{\text{sur}} - \sigma_T^{\max}) \left( \frac{z}{D_1} \right) + \sigma_T^{\text{sur}} & \text{if } z \geq -D_1 \\ \sigma_T^{\max} \exp \left( \frac{z+D_1}{D_3} \right) & \text{if } -D_1 > z \geq -D_2 \\ \max \left\{ \sigma_T^{\max} \exp \left( \frac{-D_2+D_1}{D_3} \right) \exp \left( \frac{z+D_2}{D_4} \right), \sigma_T^{\text{do}} \right\} & \text{if } z < -D_2 \end{cases}$$

and

$$\hat{\sigma}_S^o = \max \left\{ \sigma_S^{\text{sur}} \exp \left( \frac{z}{D_5} \right), \sigma_S^{\text{do}} \right\}$$

where  $\sigma_X^{\text{sur}}$  are values for  $T$  and  $S$  at the surface  $z = 0$  (where  $z$  is defined as negative below the surface),  $\sigma_X^{\text{do}}$  are lower bounds for  $T$  and  $S$  in the deep ocean, and  $\sigma_T^{\max}$  is the maximum value for  $T$  which is located at  $z = -D_1$ . The maximum value for  $S$  occurs at the surface. The values of the various parameters used in ORAS4 are given in Table 1.

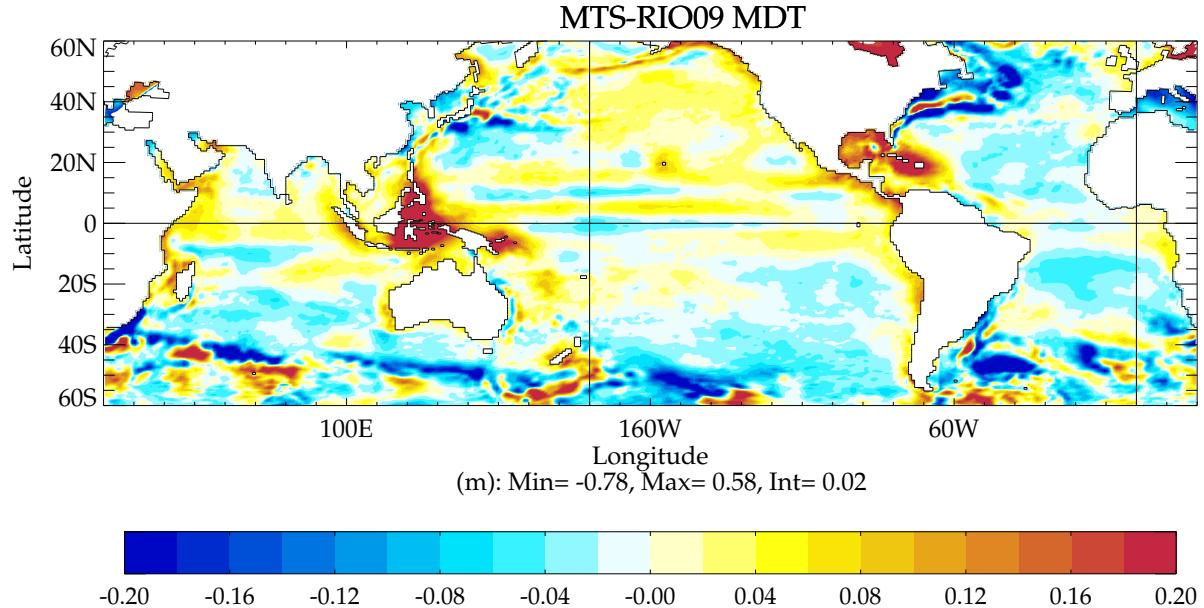


Figure 5: Differences between the MDT used in ORAS4 and RIO09.

### 4.3 Assimilation of altimeter derived sea-level anomalies

Ocean-S4 assimilates along-track sea-level data. This is the first time that along-track data have been used at ECMWF: ORAS3 assimilated gridded maps instead of along-track data. The sea-level data are provided by AVISO<sup>4</sup>. The sea level distributed by AVISO ( $\eta'^o$ ) is an anomaly relative to the 7-year period from 1993 to 1999. To enable comparison with the background field ( $\eta^b$ ) to construct the innovation  $\delta\eta^o$ , a reference mean sea level ( $\bar{\eta}^o$ ) is required:

$$\delta\eta^o = (\eta'^o + \bar{\eta}^o) - \mathbf{H}\eta^b$$

where the observation operator  $\mathbf{H}$  represents horizontal interpolation to the observation point. In Ocean-S4,  $\bar{\eta}^o = \mathbf{H}\bar{\eta}^a$  where  $\bar{\eta}^a$  is the 1993–1999 mean sea level from an ocean reanalysis using Ocean-S4 but assimilating only temperature and salinity data (experiment E-TS). In what follows we refer to it as TS MDT. The possibility of using an alternative reference  $\bar{\eta}^o$  derived from *in situ* observations and data from the GRACE gravity mission has also been explored. In particular, attempts were made to use the RIO09 MDT (Rio *et al.*, 2011), which has large differences relative to TS MDT (Figure 5). As discussed by Vidard *et al.* (2008), without proper treatment of observation bias, this can introduce abrupt jumps in the analysis.

#### 4.3.1 Use of along-track SLA data from altimeter

Assimilation of along-track SLA data with high spatial resolution in a fairly low-resolution model setup like the ORCA1 configuration poses many challenges. One of the challenges is that the altimeter measures features of the ocean which cannot be represented in a low resolution model. This is illustrated in Figure 6 which compares the observations of SLA in a 10-day window (top panel) with the model background equivalent of these data

<sup>4</sup><http://www.aviso.oceanobs.com/en/data/products/sea-surface-height-products/global/sla/index.html>.

(bottom panel). It is obvious that the observations have small scales which are not present in the corresponding model background field.

To avoid assimilating these small-scale features that the model cannot represent, a super-observation scheme has been developed. In this scheme, a reduced grid is constructed, with a specified latitude resolution (typically  $1^\circ$ ) and a specified longitude resolution (also typically  $1^\circ$ ) at the equator with an inverse cosine latitude increase in longitude resolution towards the pole. For a given point on this “superob” grid, the observations for the same day which have this grid point as the closest of all “superob” grid points were collected and the means of the observed values and space/time positions were computed to form super-observations.

It is possible to compute the standard deviation,  $\sigma_\eta^{\text{sup}}$ , of any given SLA super-observation with respect to the sample of SLA observations that are used to construct it. This standard deviation is a measure of the local variability of the SLA measurements at a scale similar to the model resolution. In areas of high local variability we do not want to give too much weight to the observations, so we have chosen to use this standard deviation as a proxy for the observation-error standard deviation. If the super-observation sample is small, however, this error might be underestimated. To compensate for this, an additional term is introduced

$$\sigma_\eta^{\text{min}} = \left( 2 - \frac{\min(N_{\text{sup}}, N_{\text{min}})}{N_{\text{min}}} \right) \bar{\sigma}_\eta$$

where  $\bar{\sigma}_\eta$  is a constant SLA observation-error standard deviation,  $N_{\text{sup}}$  denotes the number of observations in the averaging sample, and  $N_{\text{min}}$  is the minimum number above which  $\sigma_\eta^{\text{min}}$  is set to  $\bar{\sigma}_\eta$ . We then define

$$\hat{\sigma}_\eta^o = \sqrt{(\sigma_\eta^{\text{min}})^2 + (\sigma_\eta^{\text{sup}})^2}. \quad (38)$$

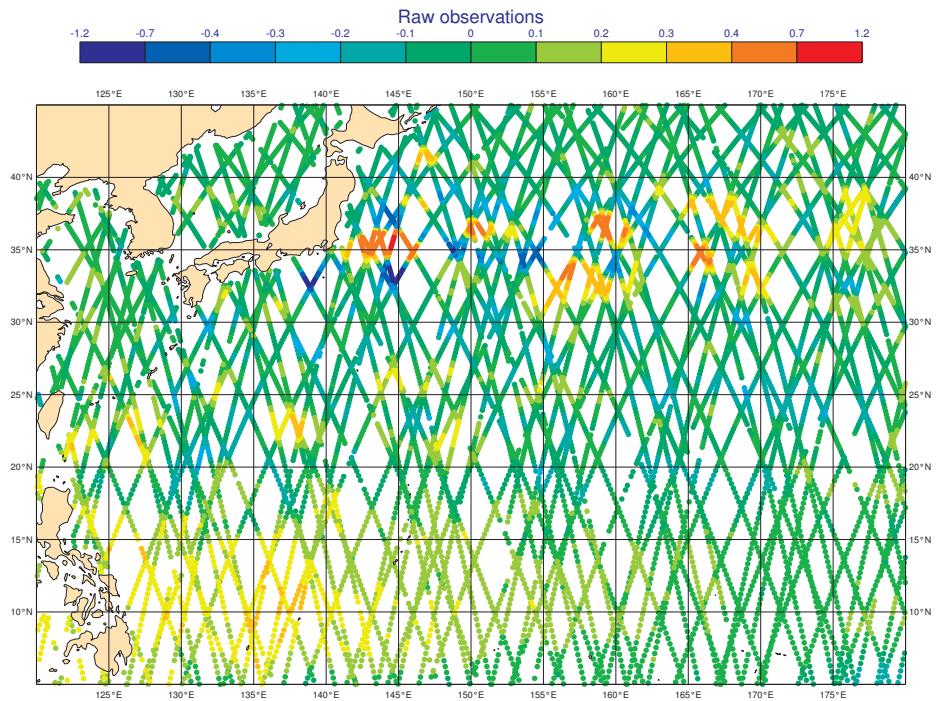
The full expression for the SLA observation-error standard deviation  $\sigma_\eta^o$  also includes the inflation factor  $A_c(r)$  near coastlines, and is thus given by (37) with  $X = \eta$ . Errors in the MDT have not been explicitly accounted for.

Figure 6 shows an example of raw SLA observations (top panel) and the corresponding model (SSH-MDT) values (bottom panel) at observation points with corresponding super-observation values (top panel) and the standard deviation of each super-observation (bottom panel) in Figure 7. The spatial variability is clearly more compatible with that of the model but there is still some strong local spatial variability that could influence the resulting analysis. Equation (38) can thus be used to reduce the influence of the SLA observations in regions with strong variability.

#### 4.3.2 Assimilation of sea-level trends

There is clear evidence that the global sea level is rising, due to the combined effect of thermal expansion (steric) and mass changes over the ocean (Church and White, 2006). The steric component of the global mean sea level cannot be represented by the ocean model since, in common with most ocean models used for climate activities, the Boussinesq approximation is made, which means that the ocean model preserves volume. Therefore, if not treated correctly, the trend in sea level can be a problem when assimilating altimeter observations. To avoid inconsistencies, the spatial mean of the sea-level background field and the input sea-level super-observations is removed before assimilation.

However, the information about the global mean sea level is not neglected, and it is used to close the fresh water budget, thus helping with the attribution of sea-level rise. Although the steric height  $\eta_s$  is not a prognostic variable of the ocean model, it can be diagnosed by vertically integrating the density field of the ocean analysis. By comparing changes in the global mean sea level from the altimeter data with the changes in steric height



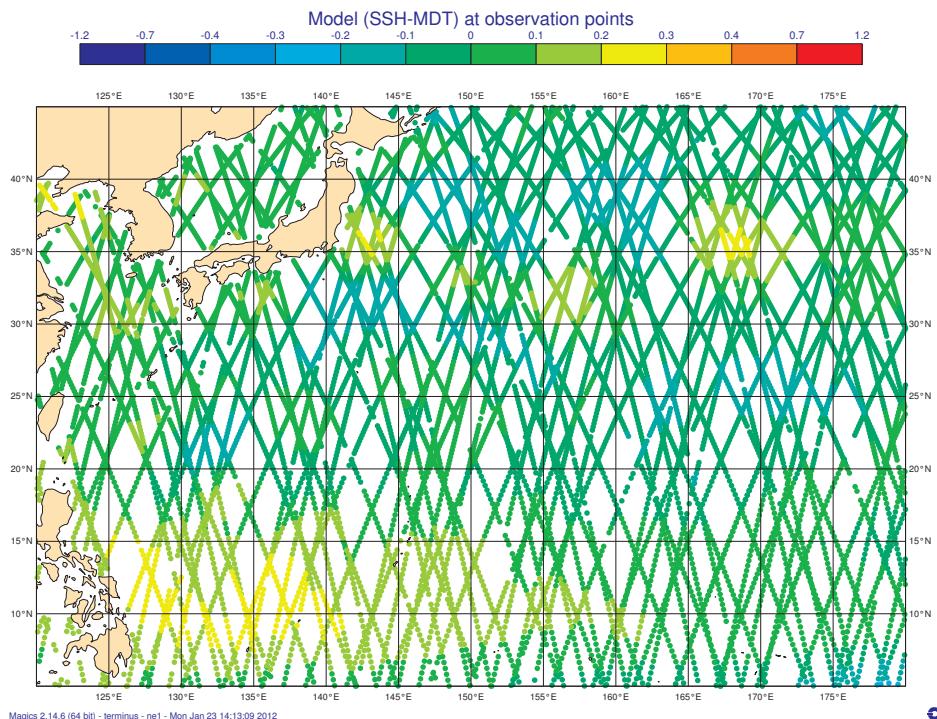



Figure 6: Example of the SLA observations from a 10-day window (top) and the background SLA field at observation points (bottom). Units are metres.

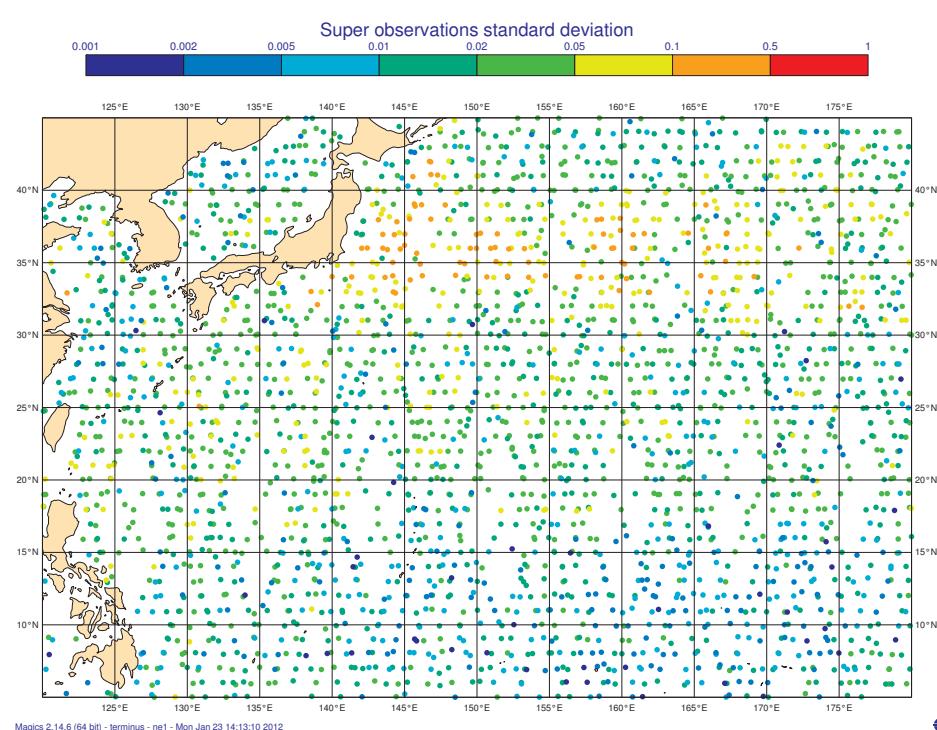
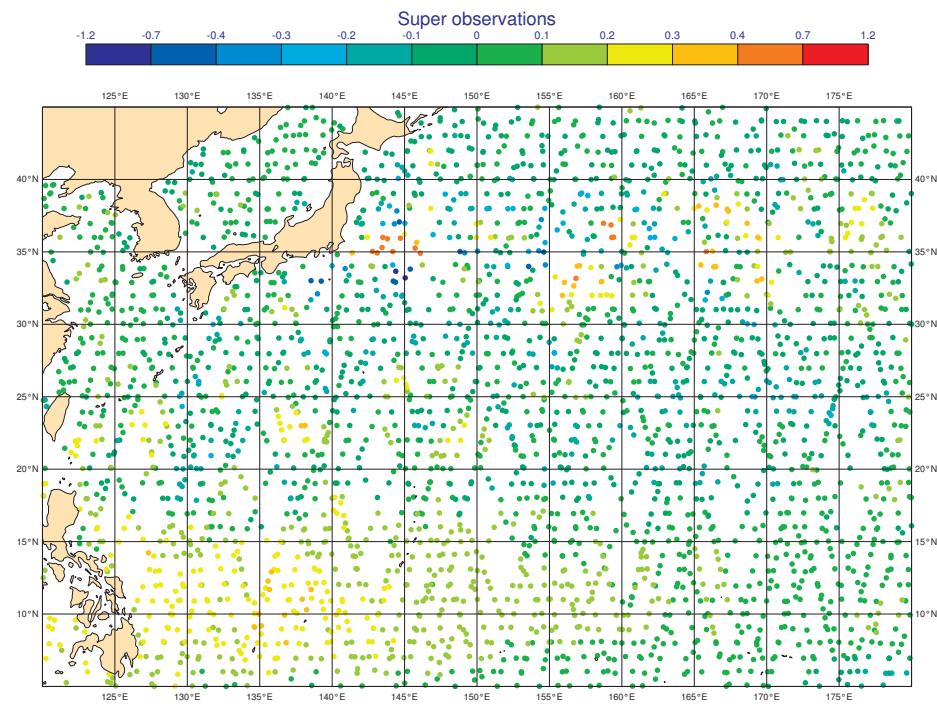


Figure 7: Corresponding super-observations (top) and their standard deviation (bottom). Units are metres.

from the ocean analysis, it is possible to estimate the component of global mean sea level change due to mass variations. This approach, implemented in ORAS3 ([Balmaseda et al., 2008](#)) is also followed in NEMOVAR. The information given by the altimeter data maps<sup>5</sup> about changes in the global mean sea level ( $\Delta\bar{\eta}_o$ ) is compared every assimilation cycle with the changes in the ocean analysis steric height ( $\Delta\bar{\eta}_s$ ). The residual ( $\Delta\bar{\eta}_m$ ) where

$$\Delta\bar{\eta}_m = \Delta\bar{\eta}_o - \Delta\bar{\eta}_s$$

is applied as a spatially uniform fresh-water flux. The partition between volume change and mass change is quite valuable information since it can help to close the fresh-water budget over the oceans, which is currently a large source of uncertainty in the analysis of the ocean. Uncertainty remains on the spatial distribution of the global fresh-water residual though.

#### 4.4 Ensemble generation

As in the previous ECMWF ocean analysis system, the Ocean-S4 consists of 5 ensemble members. These have been generated by adding wind-stress perturbations, the same as those in ORAS3. These perturbations introduce uncertainty mainly in the upper ocean ([Balmaseda et al., 2008](#)). These wind-stress perturbations are added symmetrically to 4 of the 5 ensemble members (the perturbed members), leaving a control ensemble member unperturbed. The ensemble generation in ORAS4 has two new components which, as detailed below, are designed to sample the uncertainty in both the initial conditions at the start of the reanalysis and the observation coverage.

A common strategy to start a reanalysis is to use an ocean state resulting from a spin-up integration. For ORAS4, the spin up consists of a 18-year integration of NEMO, starting from rest and the temperature and salinity fields from January of the monthly climatology of WOA05, and forced with fluxes from the ERA-Interim climatology. A strong relaxation (1-year time scale) to WOA05 is employed. We refer to this spin up as INI0. At the end of the spin up the model state reaches a quasi-equilibrium state which, due to the relaxation to climatology, is not so far from the observed state, and can in principle be used to start the assimilation run. However, the spin-up state is not representative of the state of the ocean at the given point in time. Therefore, there is an error at the initial state of the magnitude of the interannual/decadal variability of the ocean. One would expect the assimilation of ocean data to be efficient in bringing the state close to reality, but this will only be the case if enough observations are available. In the early 1960s, and in the deep ocean, a large uncertainty is likely to remain. In order to sample this uncertainty, we perform a second step in the initialization procedure (INI1), in which a second integration of the ocean model is performed, using time-varying fluxes and assimilating temperature and salinity, for the period 1958 to 1980. Five different ocean restarts from this second integration, sampled at 5-year intervals from 01/09/1960, are used to initialize each of the ensemble members of ORAS4. The impact of this second step in the ensemble generation is apparent in the upper ocean during the first decade of the reanalysis (there is large uncertainty during the 1960s), while the uncertainty (ensemble spread) in the deep ocean remains longer (not shown).

To sample the uncertainty of the observation coverage, a scheme was developed which can reject a given observation with a certain random probability. This probability can be different for each observation type and subtype. In its present implementation, random rejections were applied to T/S profiles for the perturbed ensembles, leaving the unperturbed control analysis using all available observations. For the perturbed members, the probability of rejection is 10% for Argo observations and 5% for other platforms measuring T/S.

In Ocean-S4 no attempt has been made to use the ensemble information in the specification of the background-error covariances, so there is no feedback mechanism between ensemble members. Calibrating the background-

<sup>5</sup><http://www.aviso.oceanobs.com/en/data/products/sea-surface-height-products/global/index.html>.

error covariances based on ensemble information has been explored with the OPAVAR system ([Daget et al., 2009](#)) as well as with the atmospheric data assimilation system at ECMWF ([Bonavita et al., 2011](#)) and will be investigated in the future with NEMOVAR.

#### 4.5 Specification of the bias-correction scheme parameters

Ocean-S4 uses the generalized bias formulation described in section 3.4. The bias vector has four distinct components  $\mathbf{b} = (b^{tr,T}, b^{tr,S}, b^{p,T}, b^{p,S})^T$  where  $(b^{tr,T}, b^{tr,S})$  are the bias components acting directly in the prognostic equations for T and S, and  $(b^{p,T}, b^{p,S})$  are the bias components acting on the momentum equations. The online estimation of these components (primed variables in (31)) follows a simple time evolution in (32), entirely defined by the parameter  $\alpha$ , which determines the memory of the bias estimate, and the elements of the matrix  $\mathbf{A}$ , which specify how much of the assimilation increment is considered to be systematic error. In Ocean-S4 the memory factor  $\alpha$  has been chosen to be equivalent to an e-folding time of 5 years. The coefficients of the matrix  $\mathbf{A}$  in (33) have been chosen to be a function of latitude:

$$\begin{aligned} a^{tr,T} &= a^{tr,S} = a_1 e^{-(\phi/\phi_c)^2} + a_2 \left(1 - e^{-(\phi/\phi_c)^2}\right), \\ a^{p,T} &= a^{p,S} = a_3 e^{-(\phi/\phi_c)^2}. \end{aligned}$$

The latitudinal dependence has been introduced to ensure that at low latitudes the dominant bias term is the pressure correction and that the terms for direct correction in temperature and salinity are weak at the equator. A solid methodology for the optimal estimation of bias parameters is not available. Instead, the parameters of the bias model above have been determined by tuning. To this end, several combination of parameters were evaluated in long reanalysis experiments (about 15 years long), by modifying the values of the memory factor  $\alpha$  as well as the values of  $a_1$ ,  $a_2$ ,  $a_3$  and  $\phi_c$ . Criteria such as the stability of the bias estimate, the fit to the data, the quality of the ocean currents, and the correlation to the altimeter data were used for the evaluation of the results. The final choice of parameters was  $a_1 = 0.001$ ,  $a_2 = 0.003$ ,  $a_3 = 0.015$  and  $\phi_c = 10^\circ$ . This choice is quite conservative, and is likely to underestimate the bias terms, but it was considered safe.

The *a priori* term  $\bar{\mathbf{b}}$  in (31) has been estimated iteratively as follows.

- The climatological spin-up INIO described in section 4.4 was conducted using a 1-year relaxation to the WOA05 climatology. The relaxation terms in the final year of this run, gathered as 12 monthly means, were used to construct the first guess for the *a priori* estimate of the bias. We refer to this first guess as  $\bar{\mathbf{b}}_0$ .
- An assimilation run was then conducted for the period 1989-2009 using ERA-interim fluxes and assimilating temperature and salinity profiles. This run employed the bias correction algorithm with  $\bar{\mathbf{b}}_0$  as the *a priori* offline bias correction term, and the online bias correction  $\mathbf{b}'_c$  estimated from the assimilation increments. The average of the resulting total bias for the period 2000-2008 is the final *a priori* bias  $\bar{\mathbf{b}}$  for ORAS4.

The aim of this second step is to update the bias term estimated from WOA05 with information coming directly from the assimilation of T/S Argo profiles. The term  $\bar{\mathbf{b}}$  thus calculated is applied to ORAS4 from the beginning of the integration in September 1957. This is a way of extrapolating back in time the information from Argo, and it is based on the assumption that the largest part of the error is stationary (except for the seasonal dependence). Arguably, this second step will correlate the background state and the T and S observations in subsequent assimilation experiments, and this correlation is neglected in the system. However, it is important to note that this drawback is not exclusive to this second step of the bias correction. Correlation between background and

observations is also present (and ignored) in any attempt to control the model bias using climatology, either by relaxation (a technique commonly used in most ocean reanalyses) or by additive correction (the correction of model bias is needed to avoid spurious variability in the reanalysis ([Balmaseda et al., 2007a](#))).

Figure 8 shows the spatial distribution of the bias correction terms  $\bar{\mathbf{b}}_0$  (top panel) and  $\bar{\mathbf{b}}$  (bottom panel) for temperature in the depth range 300–700 m. It can be seen that the second iteration produces higher values of the bias, and smaller spatial scales.

## 4.6 Specification of the background-error covariance parameters

The background-error covariance model described in section 3.2 requires specifying, for  $X = T, S_U$  and  $\eta_U$ , the standard deviations  $\sigma_X^b$  (the diagonal elements of  $\mathbf{D}_X^{1/2}$  in (27)) and the directional length-scales  $L_X^\lambda, L_X^\phi$  and  $L_X^z$  of the quasi-Gaussian correlation functions represented by the diffusion operators (the right-hand side of the relationships in (30)). For Ocean-S4, the specification of these parameters is somewhat heuristic, being based on physical insight and simplicity rather than on statistical calibration.

### 4.6.1 Standard deviations

As in ORAS3, the background-error standard deviations for  $T$ ,  $\sigma_T^b$ , are parameterized in terms of the vertical gradient of the background temperature field,  $\partial T^b / \partial z$ . [Weaver et al. \(2003\)](#) showed that such a parameterization could capture flow-dependent aspects of variance propagation which are implicit in 4D-Var. Their 4D-Var experiments showed that the model dynamics tend to move the level of maximum  $\sigma_T^b$  to the level of the local thermocline, to the extent that diagnosed profiles of the effective  $\sigma_T^b$  several days into the assimilation window have similar structure to corresponding profiles of  $\partial T^b / \partial z$ . This effect was shown to be particularly pronounced near the equator. From a physical viewpoint, those results from 4D-Var are sensible since largest temperature errors are expected to occur in regions where thermal variability is greatest, that is, at the level of the thermocline where  $\partial T^b / \partial z$  is large.

The following simple analysis leads to a natural formulation of  $\sigma_T^b$  in terms of  $\partial T^b / \partial z$ . Let us assume that the difference between the background temperature  $T^b$  and true temperature  $T^t$  is the result of a vertical displacement error  $\delta z$  in the background profile. That is, we assume that the value of  $T^t$  at some depth  $z$  can be found in the  $T^b$  profile at a different depth  $z' = z + \delta z$ . This is a reasonable assumption below the mixed layer ([Cooper and Haines, 1996](#)). We can then write

$$T^t(z) = T^b(z + \delta z) \approx T^b(z) + \left( \frac{\partial T^b}{\partial z} \right) \delta z,$$

which shows that, to first order, the error in the background temperature state is  $|T^b - T^t| = |(\partial T^b / \partial z) \delta z|$ . It is then reasonable to assume that  $\sigma_T^b \sim |(\partial T^b / \partial z) \delta z|$ . Such a parameterization may lead to unrealistically small values in the mixed layer and deep ocean where  $\partial T^b / \partial z$  is small. Taking this into account, we define

$$\sigma_T^b = \begin{cases} \max \{ \hat{\sigma}_T^b, \sigma_T^{\text{ml}} \} & \text{if } z \geq -D_{\text{ml}} \\ \max \{ \hat{\sigma}_T^b, \sigma_T^{\text{do}} \} & \text{if } z < -D_{\text{ml}} \end{cases}$$

where

$$\hat{\sigma}_T^b = \min \{ |(\partial T^b / \partial z) \delta z|, \sigma_T^{\text{max}} \},$$

$D_{\text{ml}}$  denoting the depth of the mixed layer,  $\sigma_T^{\text{max}}$  the maximum-allowed value of  $\sigma_T^b$ , and  $\sigma_T^{\text{ml}}$  and  $\sigma_T^{\text{do}}$  are lower bounds in the mixed layer and deep ocean, respectively. The specification of  $\sigma_T^b$  is thus transformed into one of choosing appropriate values for  $\sigma_T^{\text{max}}, \sigma_T^{\text{ml}}, \sigma_T^{\text{do}}$  and  $\delta z$ .

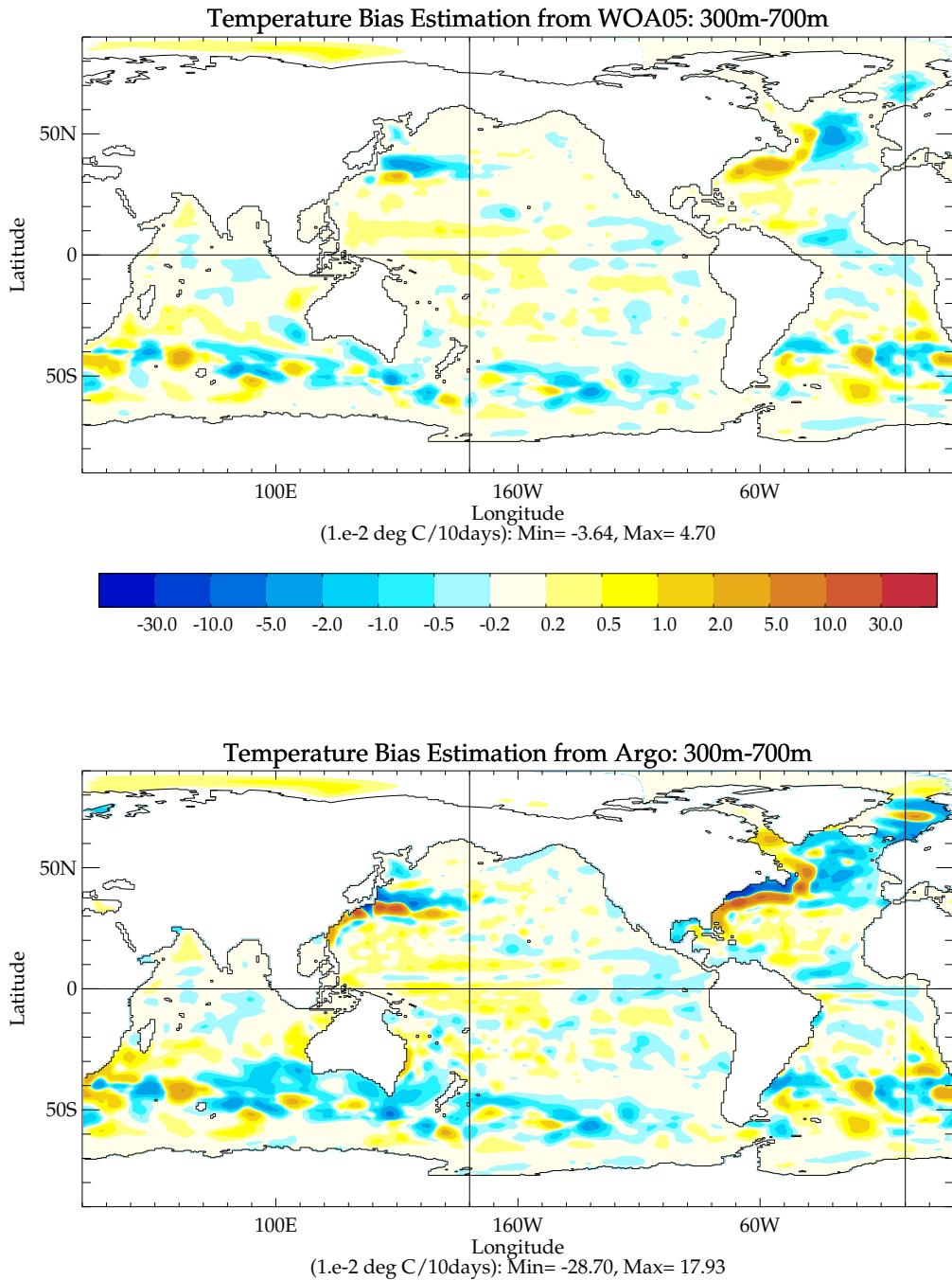


Figure 8: The bias correction terms  $\bar{b}_0$  (top panel) and  $\bar{b}$  (bottom panel) for temperature in the depth range 300-700 m.

The standard deviations  $\sigma_{S_U}^b$  for unbalanced salinity are defined according to the expression

$$\sigma_{S_U}^b = \begin{cases} \sigma_{S_U}^{\max} & \text{if } z \geq -D_{TS} \\ \max \left\{ \hat{\sigma}_{S_U}^b, \sigma_{S_U}^{\text{do}} \right\} & \text{if } z < -D_{TS} \end{cases}$$

where

$$\hat{\sigma}_{S_U}^b = B(z) \sigma_{S_U}^{\max}$$

and

$$B(z) = 0.1 + 0.45 \times \{1 - \tanh(2 \ln(z/D_{TS}))\},$$

$D_{TS}$  denoting the depth of the maximum value of

$$|(\partial S^b / \partial T^b)| = |(\partial S^b / \partial z)(\partial z / \partial T^b)|,$$

$\sigma_{S_U}^{\max}$  the maximum-allowed value of  $\sigma_{S_U}^b$ , and  $\sigma_{S_U}^{\text{do}}$  a lower bound in the deep ocean. The above parameterization thus defines the largest  $\sigma_{S_U}^b$  between the surface and the level of maximum  $S(T)$  gradients, and decreases  $\sigma_{S_U}^b$  monotonically below this level. Large values in the mixed layer are especially important since there salinity is described entirely by its unbalanced component (Ricci *et al.*, 2005). Since  $\sigma_T^b$  and  $\sigma_{S_U}^b$  depend on the state, they will evolve from one assimilation cycle to the next and thus will contain some information about the impact of data assimilation from previous cycles.

For unbalanced SSH, the standard deviations  $\sigma_{\eta_U}^b$  depend on latitude and are defined according to

$$\sigma_{\eta_U}^b = \begin{cases} \sigma_{\eta_U}^{\text{eq}} + \frac{1}{2} (\sigma_{\eta_U}^{\text{ex}} - \sigma_{\eta_U}^{\text{eq}}) (1 - \cos(\pi \phi / \phi_{\text{ex}})) & \text{if } |\phi| \leq \phi_{\text{ex}} \\ \sigma_{\eta_U}^{\text{ex}} & \text{if } |\phi| > \phi_{\text{ex}} \end{cases}$$

where  $\phi_{\text{ex}}$  is the latitude beyond which  $\sigma_{\eta_U}^b$  is set to a constant value  $\sigma_{\eta_U}^{\text{ex}}$ , and  $\sigma_{\eta_U}^{\text{eq}} < \sigma_{\eta_U}^{\text{ex}}$  is the value directly at the equator. This parameterization is designed to account for the greater importance of the barotropic component ( $\eta_U$ ) in extra-tropical regions. The values used in Ocean-S4 for the various constant parameters in these expressions are summarized in Table 2

#### 4.6.2 Length-scales

The horizontal background-error correlations for  $X = T, S_U$  and  $\eta_U$  are assumed to be isotropic poleward of a given latitude  $\phi_L$ , with an identical length-scale  $L_X^\lambda = L_X^\phi = \bar{L}$  used for all variables and at all depths. Equatorward of  $\phi_L$ , the length-scales  $L_X^\lambda$  and  $L_X^\phi$  are modulated by a latitude-dependent function  $C(\phi; \delta)$  where  $\delta = C(0)$  is an inflation or a contraction factor directly at the equator. Specifically,

$$\left. \begin{array}{l} L_X^\lambda(\phi) = C(\phi; \delta_\lambda) \bar{L} \\ L_X^\phi(\phi) = C(\phi; \delta_\phi) \bar{L} \end{array} \right\}$$

where

$$C(\phi; \delta) = \begin{cases} 1 + \frac{1}{2}(\delta - 1)(1 + \cos(\pi \phi / \phi_L)) & \text{if } |\phi| \leq \phi_L \\ 1 & \text{if } |\phi| > \phi_L \end{cases}.$$

This feature is included to allow stretching/shrinking of the zonal/meridional length-scales in the equatorial wave-guide. In particular, with the parameter settings given in Table 2, the respective values of  $L_X^\lambda$  and  $L_X^\phi$  at

$X$	$(L_X^\lambda, L_X^\phi)$	$L_X^z$	$\sigma_X^b$
$T$	$\delta_\lambda = 2$ $\delta_\phi = 0.5$ $\bar{L} = 2^\circ$ $\phi_L = \pm 15^\circ$ $M^h = 200$	$\alpha = 1$ $M^v = 10$	$\sigma_T^{\max} = 1.5^\circ\text{C}$ $\sigma_T^{\text{ml}} = 0.5^\circ\text{C}$ $\sigma_T^{\text{do}} = 0.07^\circ\text{C}$ $\delta z = 10 \text{ m}$
$S_U$	Same as for $T$	Same as for $T$	$\sigma_{S_U}^{\max} = 0.25 \text{ psu}$ $\sigma_{S_U}^{\text{do}} = 0.01 \text{ psu}$
$\eta_U$	Same as for $T$	N/A	$\phi_{\text{ex}} = \pm 20^\circ$ $\sigma_{\eta_U}^{\text{ex}} = 0.01 \text{ m}$ $\sigma_{\eta_U}^{\text{eq}} = 0$

Table 2: Summary of the values of the constant background-error covariance parameters used for Ocean-S4.

the equator have been increased and reduced by a factor of 2 relative to  $\bar{L}$ , which has been set to  $2^\circ$ . Sensitivity experiments were performed using both larger and smaller values of  $\bar{L}$  but produced no obvious overall improvement in terms of the fit to the observations with larger values and some degradation in the fit to the observations with smaller values. The  $\bar{L}$  value of  $2^\circ$  was chosen as a compromise between computational costs, as measured by the number of iterations ( $M^h$ ) required by the explicitly-formulated diffusion operator, and performance, as measured by the fit to the observations only.

The vertical correlation scales for  $T$  and  $S_U$  are specified as a scalar multiple  $\alpha \geq 1$  of the local vertical grid-size  $dz$ ,

$$L_X^z(z) = \alpha dz,$$

in order to ensure adequate smoothing between vertical levels. In Ocean-S4, a minimum smoothing is employed by setting  $\alpha = 1$ . Since the horizontal and vertical length-scales are constant, the normalization factors for the diffusion operators need to be computed only once. To obtain an accurate estimation of these factors, a randomization method has been used with a sample size of  $10^4$ .

#### 4.7 Ocean Real-Time analysis System 4 (ORTS4).

The purpose for the real-time ocean analysis system is to bring the state of the ocean up to date so it can be used as initial conditions for systems which require the ocean state in near-real time. At ECMWF the EPS, including its extension to 32 days for monthly forecasting, requires knowledge of ocean initial conditions on a daily basis.

As previously discussed (see section 4), the reanalysis system runs every 10 days with a 10-day assimilation window (e.g., no overlapping windows). It runs with a delay of 6 days for the retrieval of the SSH products to constrain the global mean of the SSH of the model. This means that initial conditions from the reanalysis system could be up to 16 days old and thus highly inaccurate. In order to create real-time ocean initial conditions on a daily basis, ORTS4 brings the latest ORAS4 state up to real time every day, using the available observations in a variable assimilation window. The length of the window is determined by the time difference between the latest ORAS4 reanalysis and the present day plus 1 day, so the output initial conditions will be valid on 0Z the following morning. The setup is illustrated in Figure 9.

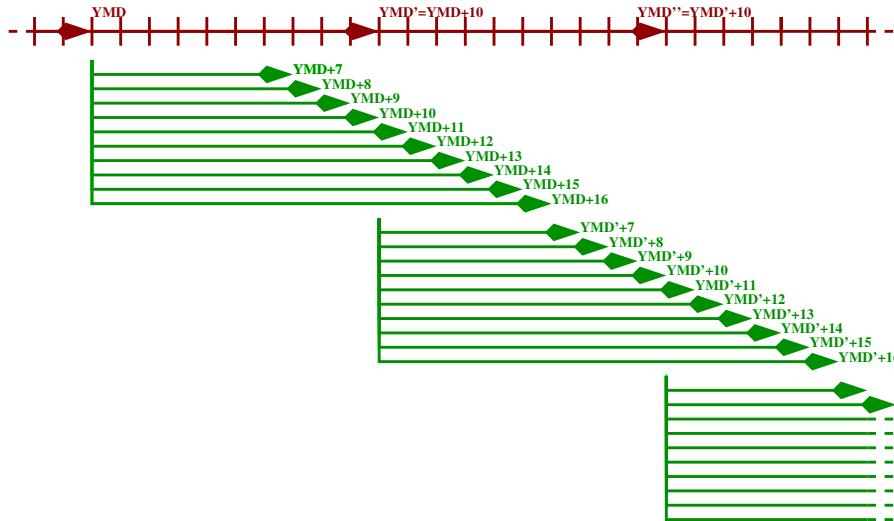


Figure 9: Schedule of the NEMOVAR real-time ocean analysis (ORTS4).

In the current implementation, ORTS4 is launched every day at 14Z to bring the latest reanalysis up to 0Z the following day. In this setup, part of the last day of the window is effectively a forecast since it brings the ocean state into the future. The reason for this set-up is to be able to provide, without delay, initial ocean conditions for the coupled set-up used by the EPS starting from 0Z.

Figure 10 illustrates the difference between the number of observations supplied to ORTS4 and ORAS4. The figure shows an example of the observations used to create ocean states valid for the 9th of August 2011 in ORTS4 (launched the 8th of August 2011 starting from ORAS4 valid at 0Z on 30th of July 2011), and in the ORAS4 delayed reanalysis (launched the 14th of August from the same initial conditions). Both the real-time and the reanalysis run have the same window length for easier comparison. The figure shows that most of the data arrives in time for the real-time system, although some data arrives late. It is of course not surprising that the last day has less observations in the real-time system since it is effectively a short-term forecast. The figure also illustrates that observations (CTDs and XBTs) from either dedicated cruises or ships of opportunity can be very unevenly distributed in time.

## 5 Performance of the ECMWF Ocean Reanalysis System 4 (ORAS4).

Before we discuss the details of the performance of NEMOVAR it is worth recalling that the observational network of temperature and salinity has changed dramatically during the reanalysis period starting in 1957. Figure 11 shows the number of temperature (top panel) and salinity (bottom panel) observations in the EN3 data-set as a function of time, at a depth of around 500 m. The most obvious feature is the dramatic increase in the number of observations (especially of salinity observations) with the build up of the Argo observing system from 2000 onwards. The observation coverage of Argo is quite uniformly distributed in time and space (not shown). The figure also shows the build up of the mooring network (TRITON, PIRATA and RAMA) beginning in the early 1990s. In the EN3 data set the number of actual observations from moorings is quite low due to the fact that the raw data from the moorings are used to construct daily means. These observations are concentrated at low latitudes, and are the backbone of the near-equatorial ocean observing system for seasonal forecasting. They are very important in data assimilation, since they provide long observational time series at fixed positions. CTD and XBT observations from ships (either dedicated in the form of cruises or of opportunity) are available

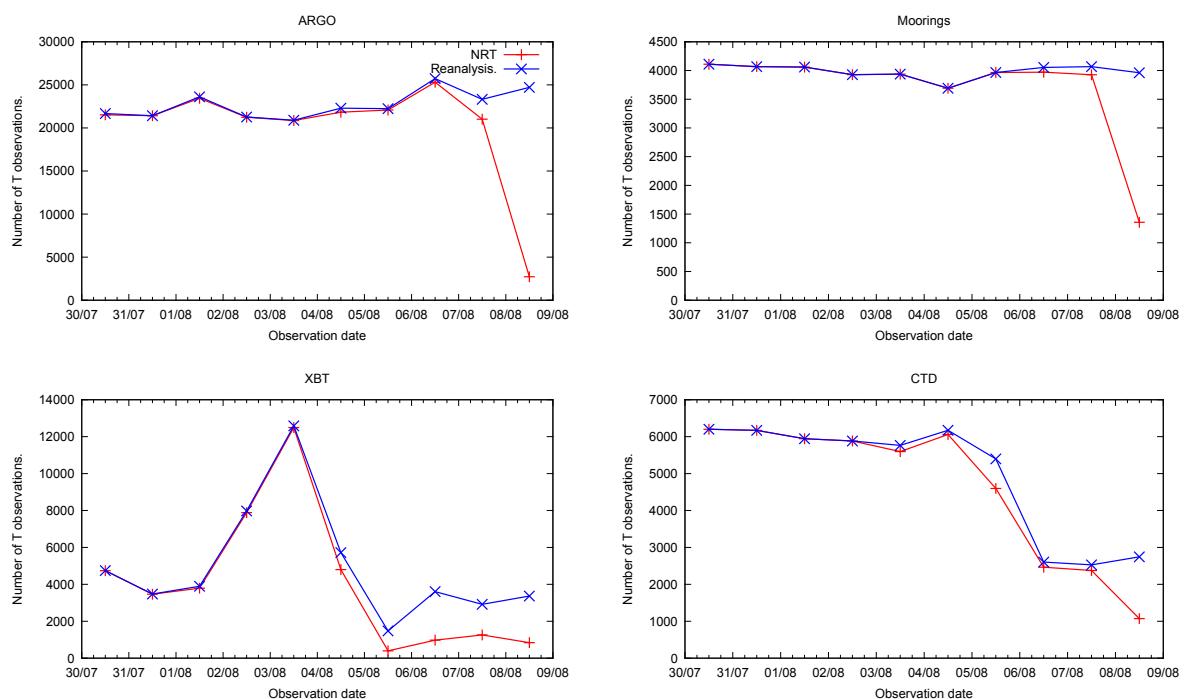


Figure 10: Number of input observations as a function of time for ORTS4 and the 16-day delayed ORAS4. Both analyses were produced with a 10-day window starting the 30/07/2011. The observations are binned into daily time slots.

since the beginning of the record, albeit being very unevenly distributed in time (and in space; not shown).

Figure 11 also shows the total number of observations used in ORAS4 as function of time (orange curves) for both temperature and salinity observations. The differences to the total number of observations (black curves) are mainly due to the vertical thinning (illustrated in Figure 4) which rejects some of the observations in high vertical resolution profiles (mostly from CTDs and XBTs).

The diagnostics presented in the remainder of this section focus on evaluating the unperturbed ensemble member of ORAS4.

## 5.1 Interpreting the fit to the observations

Figure 12 illustrates the fit to the data in the various steps of the assimilation algorithm by showing the root-mean-squared error (RMSE) of the background or first-guess (FG) field, the background field plus the increment obtained after the inner-loop minimization (MIN), and the final analysis (AN) obtained after correcting the model using IAU. The statistics are computed at the observation location and time and averaged in vertical bins to allow for different observation depths. The choice of vertical binning used in this study is based on the typical depths TAO moorings with additional levels in the deep ocean (see Table 3). For a given depth of an observation, the closest level determines the bin in which it is used.

The interpretation of the FG RMSE statistics is straightforward since they just involve a comparison of the observations with the corresponding model background field at the observation position and time (the FGAT approximation). At this stage, the observations in the statistics have not yet been assimilated. The MIN RMSE statistics reflect the ability of the minimization to reduce the model-observation misfit taking into account the constraints imposed by the background- and observation-error covariances. We refer to it as MIN to differentiate it from the final ORAS4 AN. This is the result of a second model integration over the same assimilation window but with the addition of the assimilation increment which is spread in time via IAU. The model values from AN are compared with the observations in the same way as they were for FG.

The consequences of IAU on the statistics shown are not straightforward to interpret. For instance, if all observations were located at the zero time-step of all of the assimilation windows, the AN and FG statistics would be identical, since the increment would not have been applied yet. On the other hand, if the model is quasi-stationary over the assimilation window, and all observations were at the end of the assimilation window, then the AN statistics would be very close to the MIN statistics. Since the observations in reality are more or less equally distributed over the assimilation window we would expect the AN RMSE to be somewhere in between the RMSE of FG and MIN, which indeed is consistent with what is observed in Figure 12. Differences between the FG and AN statistics suggest that, on average, the data assimilation reduces by about 10% the error of the FG in each assimilation window.

## 5.2 Comparison with the assimilated observations

The performance of ORAS4 is evaluated by comparing it with a simulation that does not assimilate data. This simulation, called the control integration (CNTL), uses the same spin-up, forcing fields, SST/sea-ice relaxation and relaxation to climatology (with 20-year time scale) as ORAS4. No additive bias correction is applied, since the bias correction is part of NEMOVAR. The observation operators are applied in the same way as in ORAS4 in order to compare the CNTL fields with the observations that were assimilated in ORAS4.

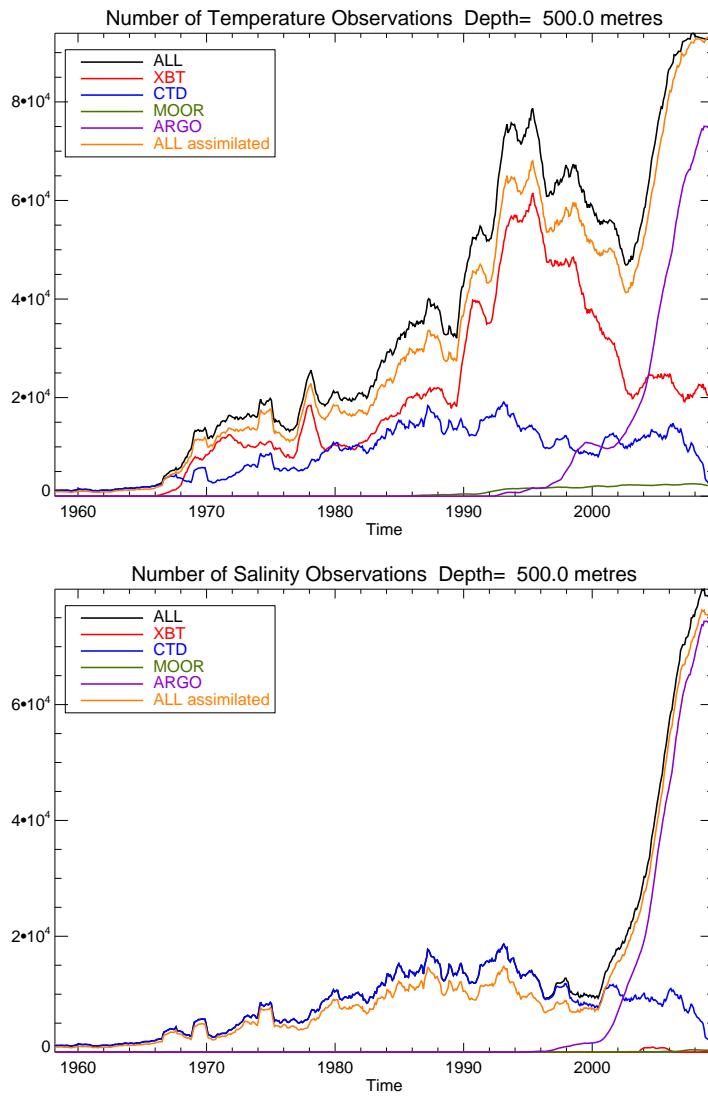


Figure 11: Temperature (top) and salinity (bottom) observations as function of time within the depth range 400 – 600 m. Not all the observations are assimilated, as shown by the difference between the black and orange curves, since there is rejection of data near the coast and vertical thinning.

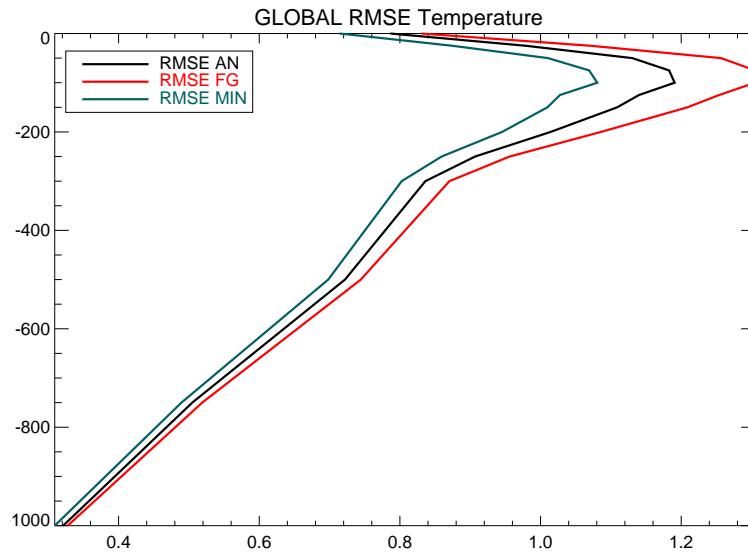


Figure 12: Fit to the temperature observations in the different steps of the assimilation algorithm: the first guess (FG, red curve), the analysis directly after the inner-loop minimization (MIN, blue curve), and the analysis (AN, black curve) after applying the assimilation increment via IAU in the final step. Units are  $^{\circ}\text{C}$ .

Bin number	Depth range (m)
1	0
2	25
3	50
4	75
5	100
6	125
7	150
8	200
9	250
10	300
11	500
12	750
13	1000
14	1500
15	2000
16	2500
17	3000
18	3500
19	4000
20	4500
21	5000
22	5500

Table 3: Depth in metres of vertical levels used for the statistical computations.

Region	Start longitude	End longitude	Start latitude	End latitude
EQUA	0°	360°	-2°	2°
TROP	0°	360°	-30°	30°
NXTRP	0°	360°	30°	70°
SXTRP	0°	360°	-70°	-30°
GLOBAL	0°	360°	-90°	90°

Table 4: Geographical extent of regions used for evaluation of ORAS4.

### 5.2.1 Comparison with temperature and salinity data

As discussed in section 4.2, ORAS4 uses the EN3 data-set which has already been quality controlled. We can use this data set also for comparison of CNTL and FG of ORAS4, effectively computing the model equivalent of exactly the same observations (as long as we ignore any additional quality control decisions made by the ORAS4 system for the FG of ORAS4). In FG the observations have not yet been assimilated, and therefore can be considered as independent observations.

Figures 13 and 14 show vertical profiles of the RMSE (thick lines) and bias (thin lines) of temperature and salinity, respectively, in the different regions listed in Table 4. Shown are the values for ORAS4 FG (red), ORAS4 AN (black) and CNTL (blue). In both figures the full EN3 data-set is used (including observations not assimilated due to vertical thinning *etc.*) to ensure that comparisons between the CNTL and ORAS4 involve exactly the same data. The only observations excluded from the statistics are those at model land-points. The strong relaxation to SST (which is applied to both CNTL and ORAS4) constrains the temperature at the top of the ocean so the fit is quite similar in both simulations. Largest errors are in the thermocline. In all regions, the fit to the observations (in terms of bias and RMSE) is improved in ORAS4 compared to CNTL, the improvement being most visible in the mixed layer and thermocline.

Unlike for temperature, the largest errors in salinity occur at the surface, in both ORAS4 and CNTL, since the surface salinity is poorly constrained, and since the impact of erroneous fresh-water fluxes and ocean mixing manifest themselves in errors in the salinity field. As for temperature, data assimilation improves the fit to the salinity observations, as expected. In principle, this improvement is not only due to the direct assimilation of salinity observations, but also to the assimilation of temperature observations, through the balance relations between temperature and salinity.

Additional information is provided by the time series of the fit of ORAS4 and CNTL to the EN3 data-set as displayed in Figure 15. The time variation in the statistics is due to a mixture of model performance and changes to the observation network, but the relative comparison between CNTL and ORAS4 for any given time should filter the effect of the changes in the observation coverage since the same observations are used. During the first decade of the record, ORAS4 and CNTL are similar, since there are few observations (Figure 11), and have not drifted significantly from the state used to initialize the simulations. From around 1970 onwards, CNTL drifts from the observed climatology, and both its RMSE and bias increase. In comparison, the RMSE and bias of ORAS4 are kept relatively stable at lower values than those of CNTL, since they are constrained by the observations. In the tropics and southern extratropics there is a visible downward trend in the RMSE of both CNTL and ORAS4, which cannot be attributed only to the increase in the number of assimilated ocean observations. The decreasing errors is suggestive of improved quality of the forcing fluxes (transition from ERA40 to ERA-Interim, for instance) and/or SST analyses, since these are common features in ORAS4 and CNTL. The reduced error can also be a consequence of changes in the spatial sampling.

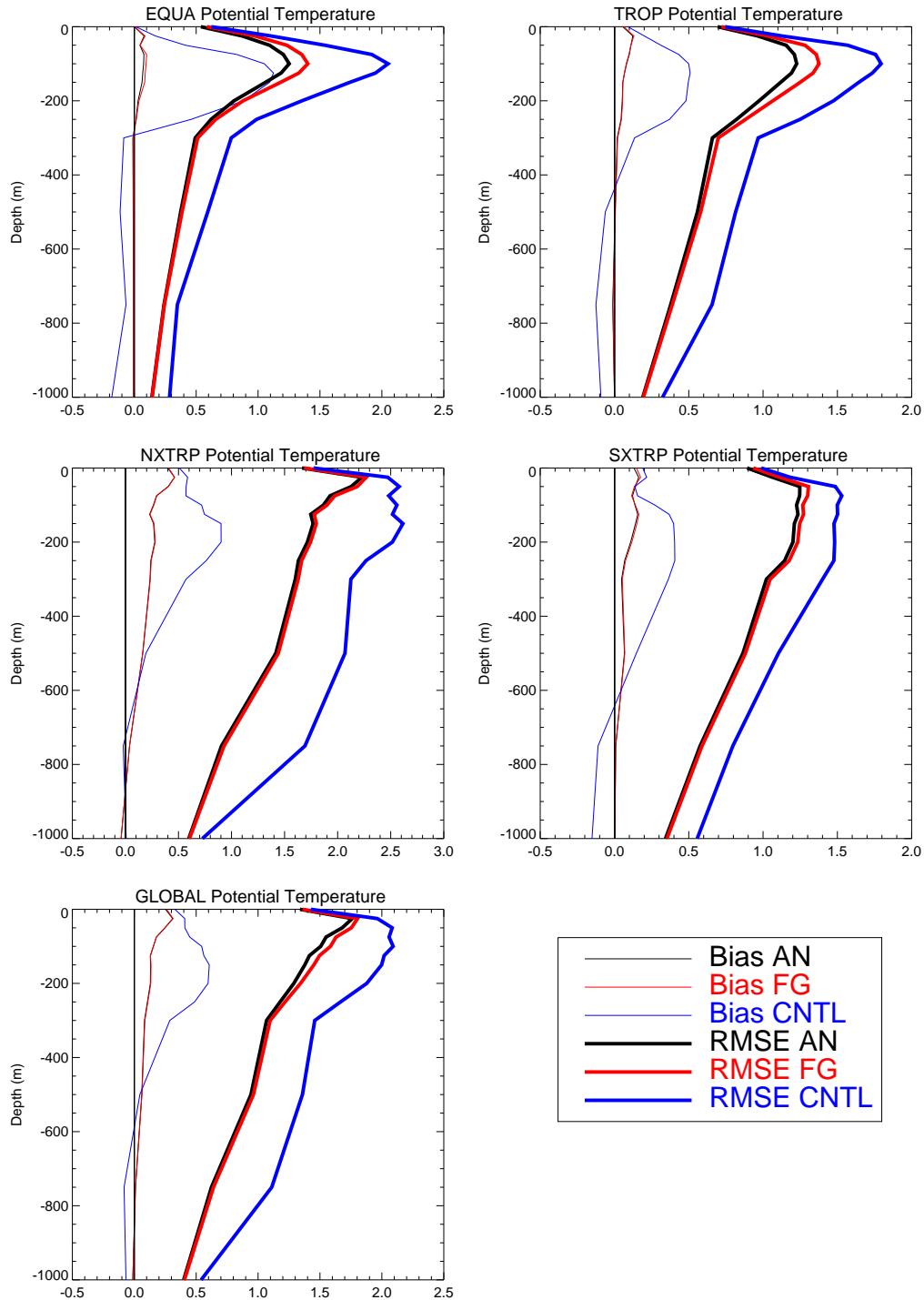


Figure 13: Fit to the temperature observations. Shown are the statistics for the control experiment (CNTL, blue), the ORAS4 analysis (AN, black) and first guess (FG, red). Note that in the case of FG, the observations can be considered independent since they have not yet been assimilated. Units are  $^{\circ}\text{C}$ .

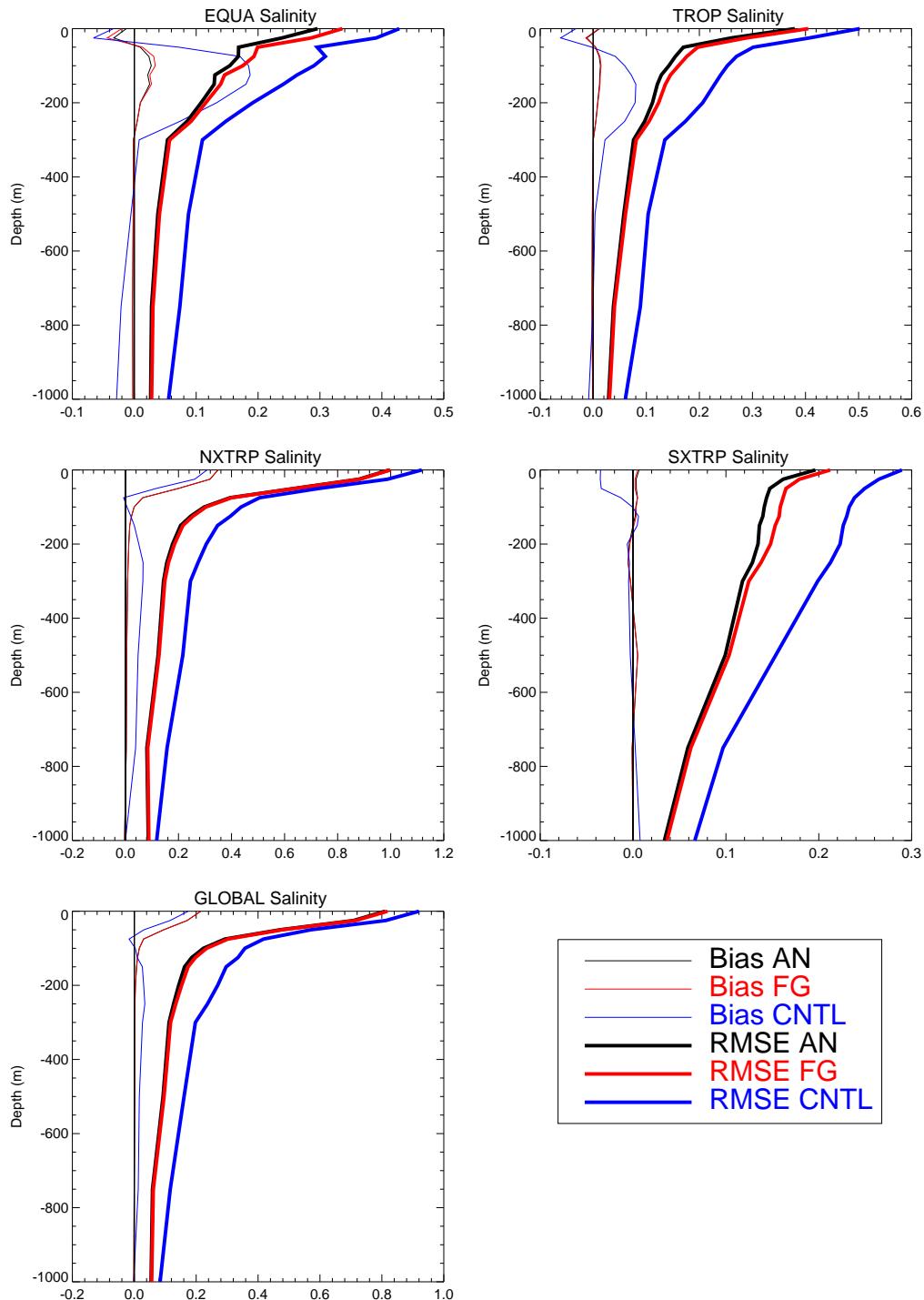


Figure 14: Same as Figure 13 but for salinity. Units are psu.

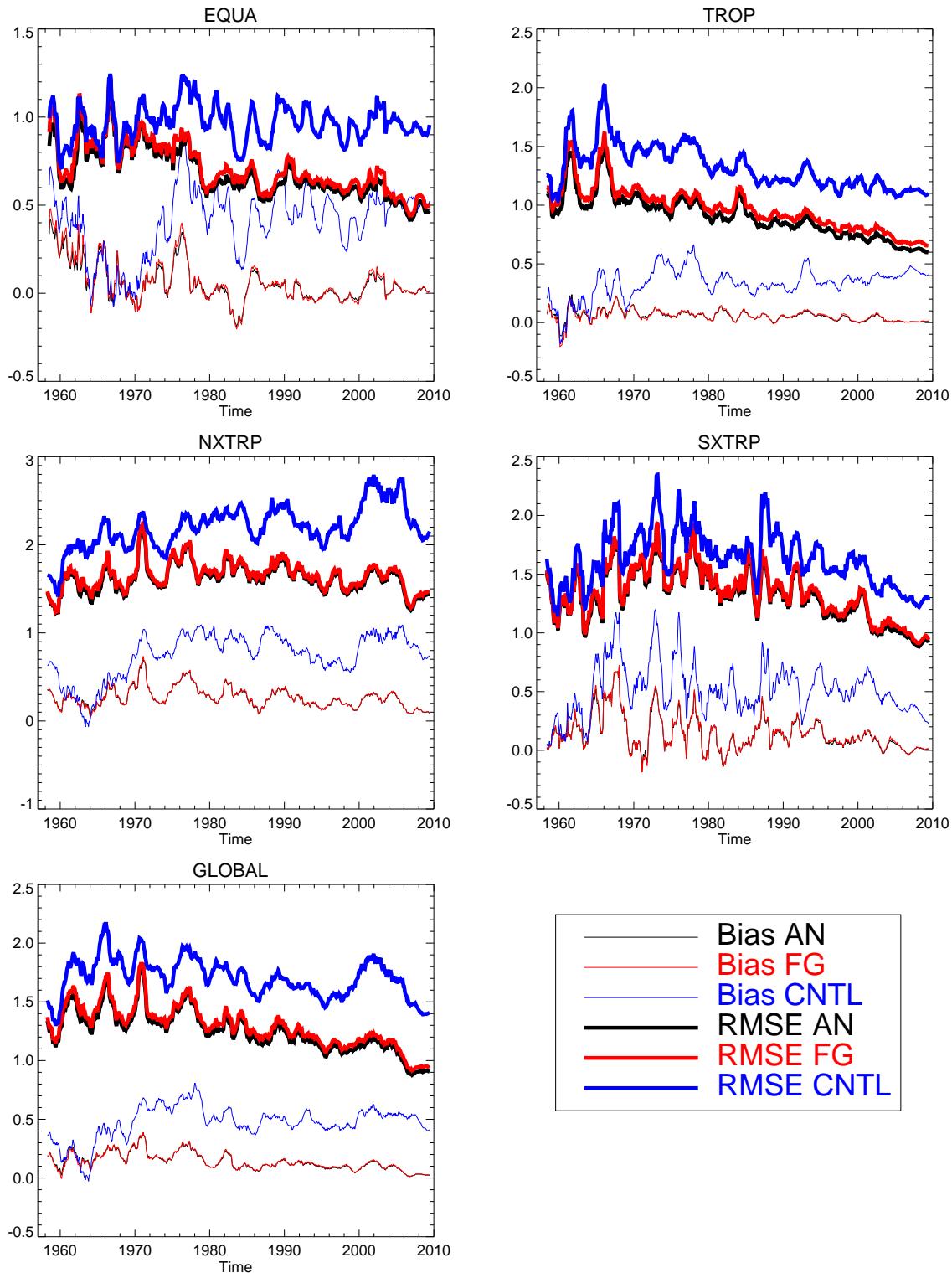


Figure 15: Time series of ORAS4 and CNTL monthly fit to temperature observations at depths between 225 and 275 m for different regions. Units are  $^{\circ}\text{C}$ .

### 5.2.2 Comparison with altimeter data

The quality of the time variability of the ocean state can be gauged by the temporal correlation of the analysis sea level anomalies with the altimeter SLA maps provided by AVISO. Figure 16 shows the correlation of monthly means from the period 1993-2008 for 3 different experiments: CNTL (top), the experiment E-TS (described in section 4.3) used to derive the MDT, which is equivalent to ORAS4 but without altimeter data assimilation, and ORAS4. It can be seen that the assimilation of T and S profiles improves the correlation with the altimeter data in most of the tropical regions, including the Equatorial Atlantic Ocean. There is, however, some degradation in the North Western Atlantic by the Iberian Peninsula, for reasons that are not fully understood, although they seem related with the representation of the water masses from the Mediterranean Outflow. As expected, the inclusion of altimeter data in the assimilation further increases the correlation with the AVISO data. The correlation is not substantially increased in the coastal areas (see for instance the region around Hawaii), where the error variance prescribed to the observations is large, and in high latitudes, where the altimeter representativeness error is larger and the stratification is weak.

## 5.3 Comparisons with independent observations: ocean currents and RAPID transports

Since velocity observations are not assimilated into ORAS4, they provide us with an independent data-set for validation of Ocean-S4. The assimilation process influences the currents of the model both directly via the analysis increments, which through the balance relations between temperature, salinity and SSH (see section 3.1) includes increments for the horizontal velocity components, and indirectly via the model's nonlinear response to IAU.

### 5.3.1 Validation against current-meter data

Some of the moorings in the Pacific (TAO/TRITON), Atlantic (PIRATA) and Indian Ocean (RAMA) possess current meters and Acoustic Doppler Current Profilers (ADCP), providing vertical profiles of the currents. These data are available from the RAMA/TAO Project Office of NOAA/PMEL's website<sup>6</sup> and are used here for validation.

Figure 17 shows the temporal mean profile of the zonal currents from ADCP data for three TAO moorings at positions (0°N, 165°E), (0°N, 170°E) and (0°N, 140°W), a PIRATA mooring at position (0°N, 23°E) and a RAMA mooring at position (0°N, 90°E). Also shown are the corresponding profiles for ORAS4 and CNTL. It can be seen from the figure that the assimilation increases the mean value of the Equatorial undercurrent in the Pacific and Atlantic Ocean, improving the fit to the data there. The strength of the Equatorial undercurrent is noticeably improved by the data assimilation in the Western and Central Pacific. In the Eastern Pacific (0°N, 170°E), even though the undercurrent is slightly stronger with assimilation, it is still too weak compared with the observations. In the Central and Eastern Pacific, the surface westward currents are too strong in both CNTL and ORAS4 compared to the observations. The story is similar for the PIRATA mooring, where again the assimilation strengthens the undercurrent although it is still too weak compared to the observations. The RAMA mooring does not show any obvious undercurrent, and here there is very little difference in the mean current profile with and without assimilation.

Figure 18 allows the visual inspection of the temporal variability of the currents. It shows the zonal current as a function of depth and time for the TAO mooring at the Western Pacific mooring (0°N, 165°E) for the

<sup>6</sup> [http://www.pmel.noaa.gov/tao/data\\_deliv/frames/main.html](http://www.pmel.noaa.gov/tao/data_deliv/frames/main.html).

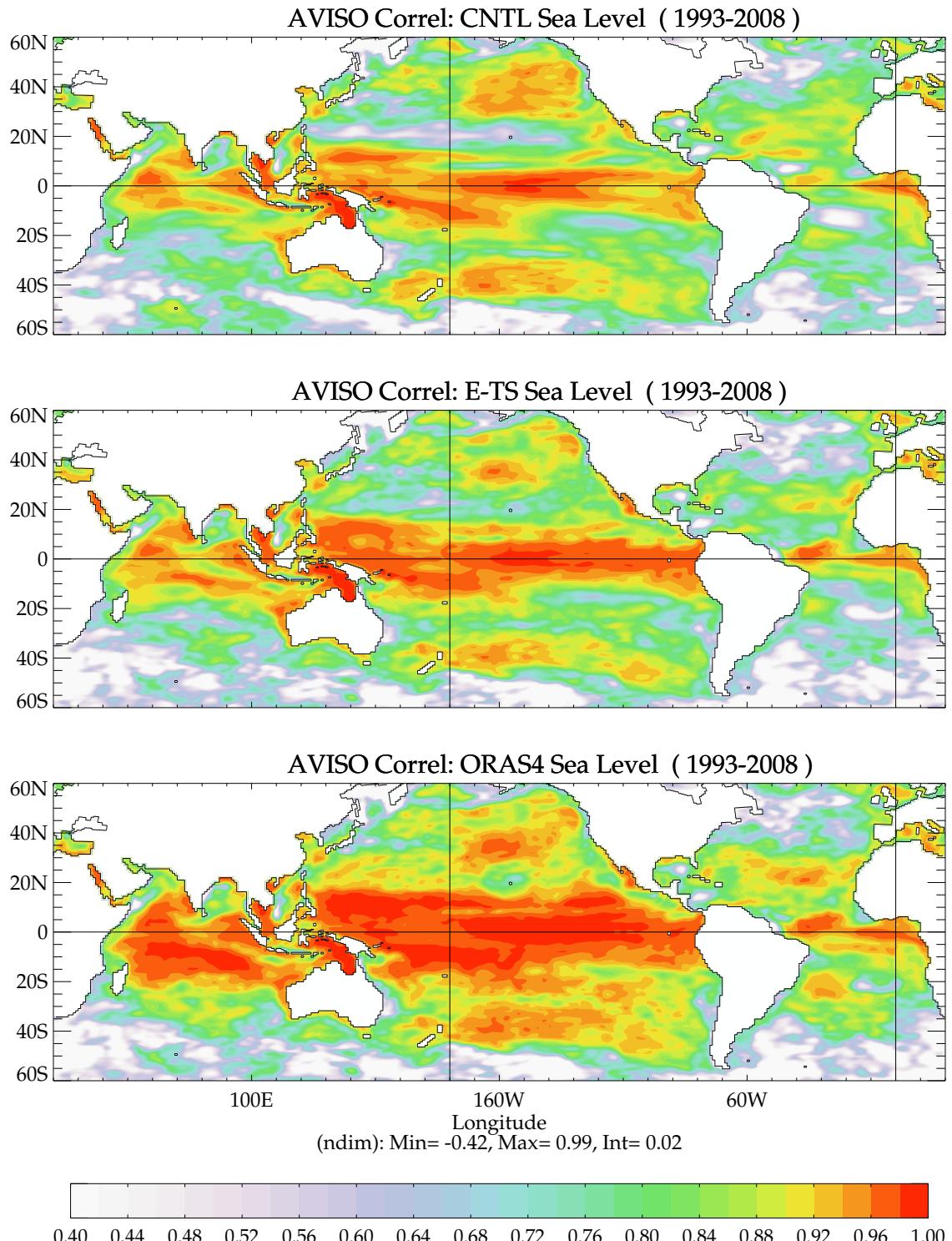


Figure 16: Temporal correlation between analysis and AVISO sea level. Shown are CNTL (top), E-TS, which assimilates T and S but not altimeter data (middle), and ORAS4 (bottom). The statistics have been computed with monthly means for the period 1993-2008. Only values above 0.4 are shown.

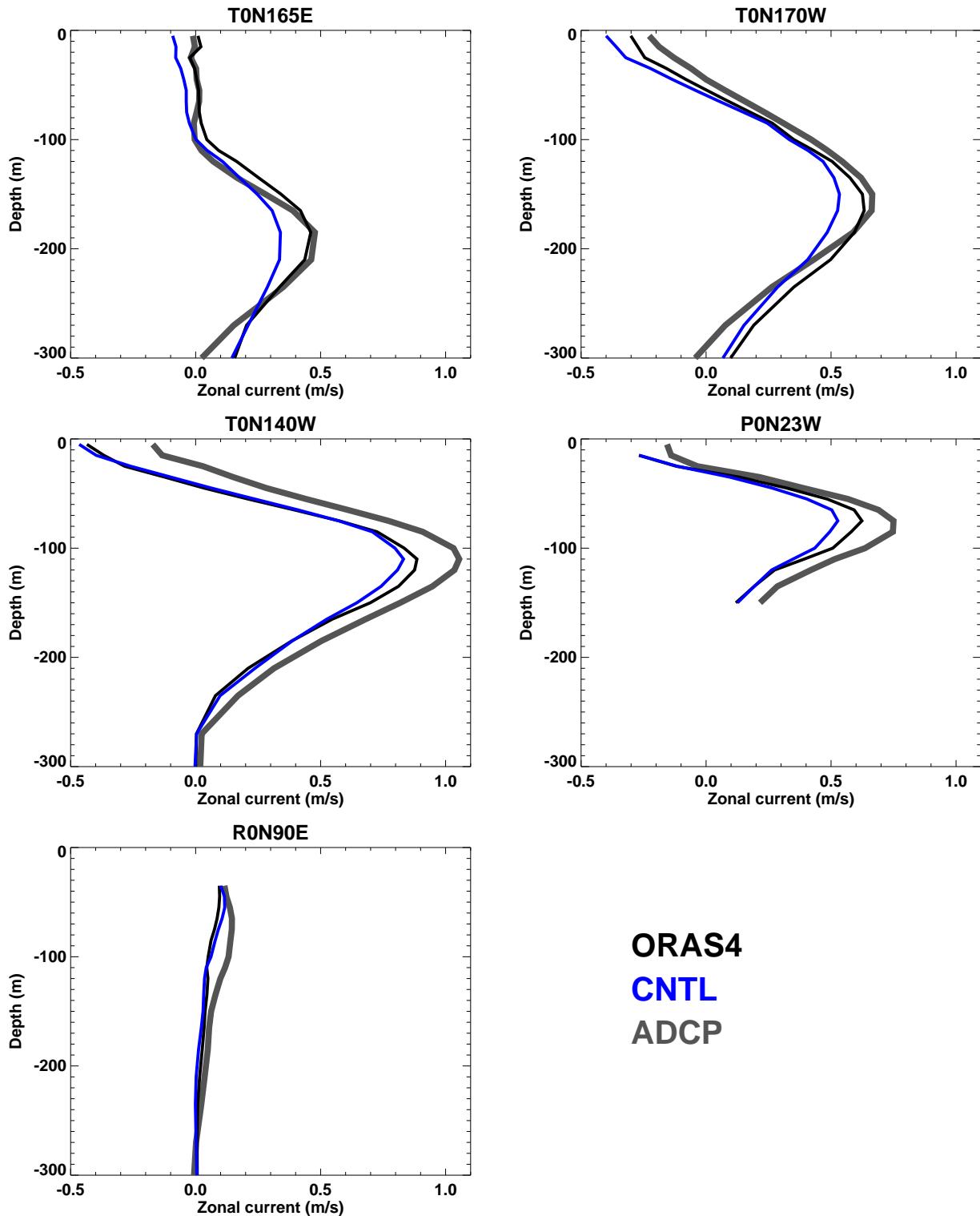


Figure 17: Temporal mean of the zonal current from 5 moorings (TAO, PIRATA, RAMA; thick curves) compared to ORAS4 and CNTL.

observations (top panel), ORAS4 (middle panel) and CNTL (bottom panel). Similar panels for the PIRATA mooring at (0°N, 23°E) appear in Figure 19 and for the RAMA mooring at (0°N, 90°E) in Figure 20.

The zonal current in the Western Pacific (TAO mooring 0°N, 165°E, Figure 18) exhibits a clear seasonal cycle and interannual variability. Both the amplitude and position of the undercurrent are modulated by ENSO. During the warm episodes the vertical shear is weak, and the surface currents have the same sign as the undercurrent. These weak-shear El Niño episodes are followed by the strong-shear regimes characteristic of La Niña. The most obvious case is the weak-shear regime in 1997, followed by the strong-shear regime in 1998. Both ORAS4 and CNTL capture this variability, which is largely determined by the wind-stress. The undercurrent is more confined in the vertical in ORAS4 than in CNTL, in better agreement with the observations. This is consistent with the data assimilation correcting for an overly diffuse thermocline. Equally, the vertical interannual migrations of the undercurrent seem more coherent in ORAS4 than in CNTL.

The time record of the PIRATA mooring is quite limited, not long enough to show a clear interannual variability, but enough to show the seasonal displacement of the undercurrent, which becomes shallower during the first part of the year, and deepens during the second half. As in the Pacific, this vertical migration of the undercurrent in the Atlantic is more coherent in ORAS4 than in CNTL, probably the result of a more confined undercurrent within a tighter thermocline.

The RAMA mooring does not exhibit any permanent undercurrent, with the largest current values closer to the surface (although there is no data available for the upper 50 m). Both ORAS4 and CNTL seem to capture the seasonal variations. The interannual variability may be overestimated in ORAS4, producing too strong values in 2005 and 2008 (positive and negative, respectively). The negative event in 2008 in ORAS4 is absent in CNTL, and not clear in the observations.

### 5.3.2 Validation against OSCAR current analyses

The Ocean Surface Current Analysis - Real-time (OSCAR) project provides analyses of oceanic surface currents derived from satellite altimeter and scatterometer data (Bonjean and Lagerloef, 2002). They are available from the end of 1992 up to near real-time and now cover the whole ocean from 60°S to 60°N. The OSCAR currents are not completely independent from sea-level data since altimeter data are used in their production. However, we do not use SLA data directly to estimate the currents and therefore there is no guarantee that they should lead to improved velocity analyses. Therefore, it is a good metric to assess the quality of our analysis.

Figure 21 shows the correlation between the zonal component of the surface velocities from OSCAR monthly means and the different analyses: CNTL (top panel), ORAS4 (middle panel) and their difference (bottom panel). The correlation is for the period 1993-2009. The colour scale is nonlinear, and different from that in Figure 16, since the correlation for the currents is lower than for sea level. ORAS4 has higher correlation values than CNTL, except at the equator and along the coasts. Within 10° of the equator, the increased positive correlation is mainly due to the assimilation of *in situ* temperature, while the assimilation of SLA improves the currents almost everywhere.

### 5.3.3 Comparison with RAPID-derived transports

The RAPID observations provide the ocean/climate community with a unique dataset that can be used for model validation as well as for constraining models through data assimilation. Processed datasets are available at National Oceanography Centre, Southampton (NOCS)<sup>7</sup>. The RAPID observations allow comparison of

<sup>7</sup><http://www.noc.soton.ac.uk/rapidmoc>.

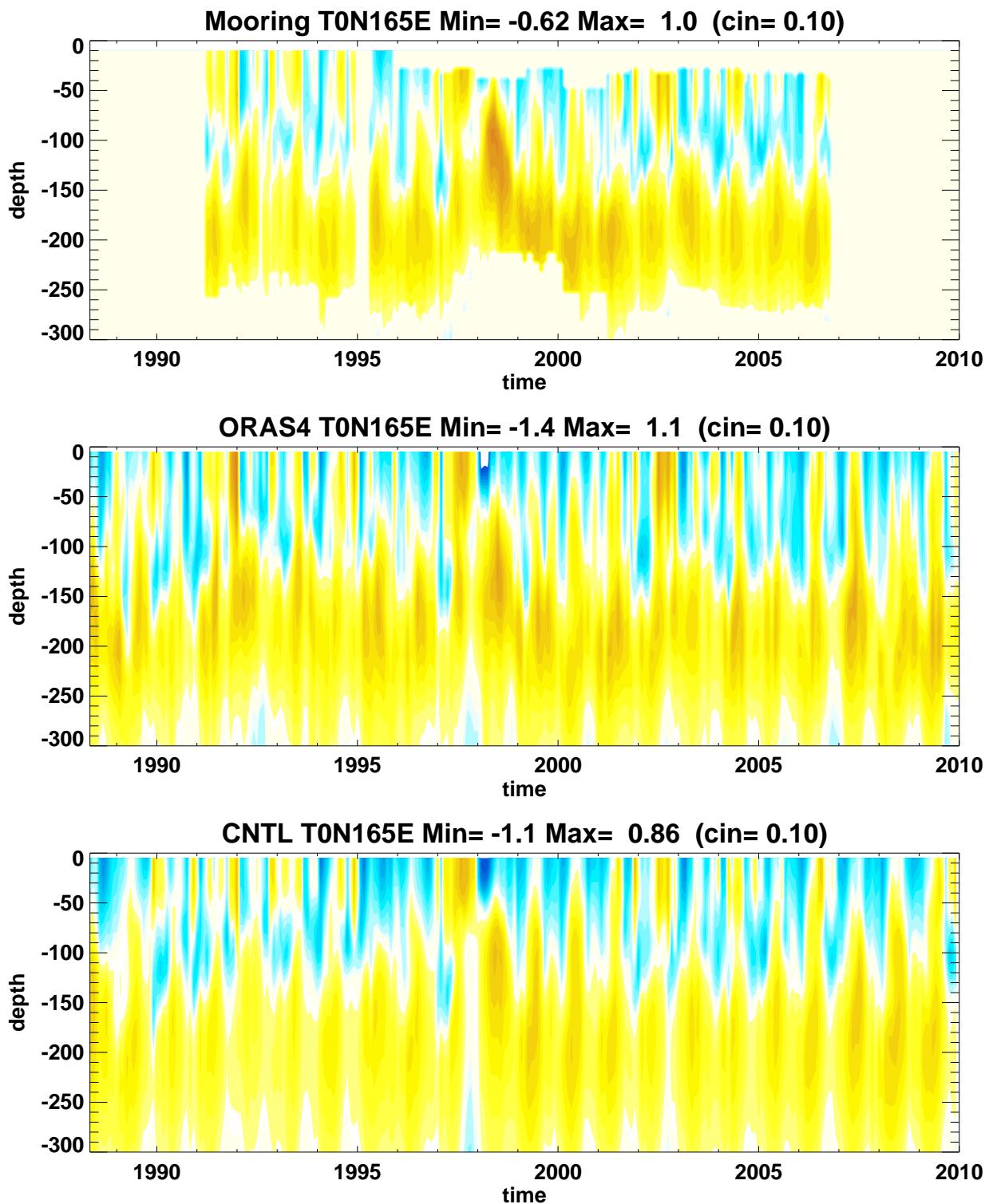


Figure 18: Time evolution of the zonal current (in m/s) of the TAO moorings at position ( $0^{\circ}\text{N}$ ,  $165^{\circ}\text{E}$ ).

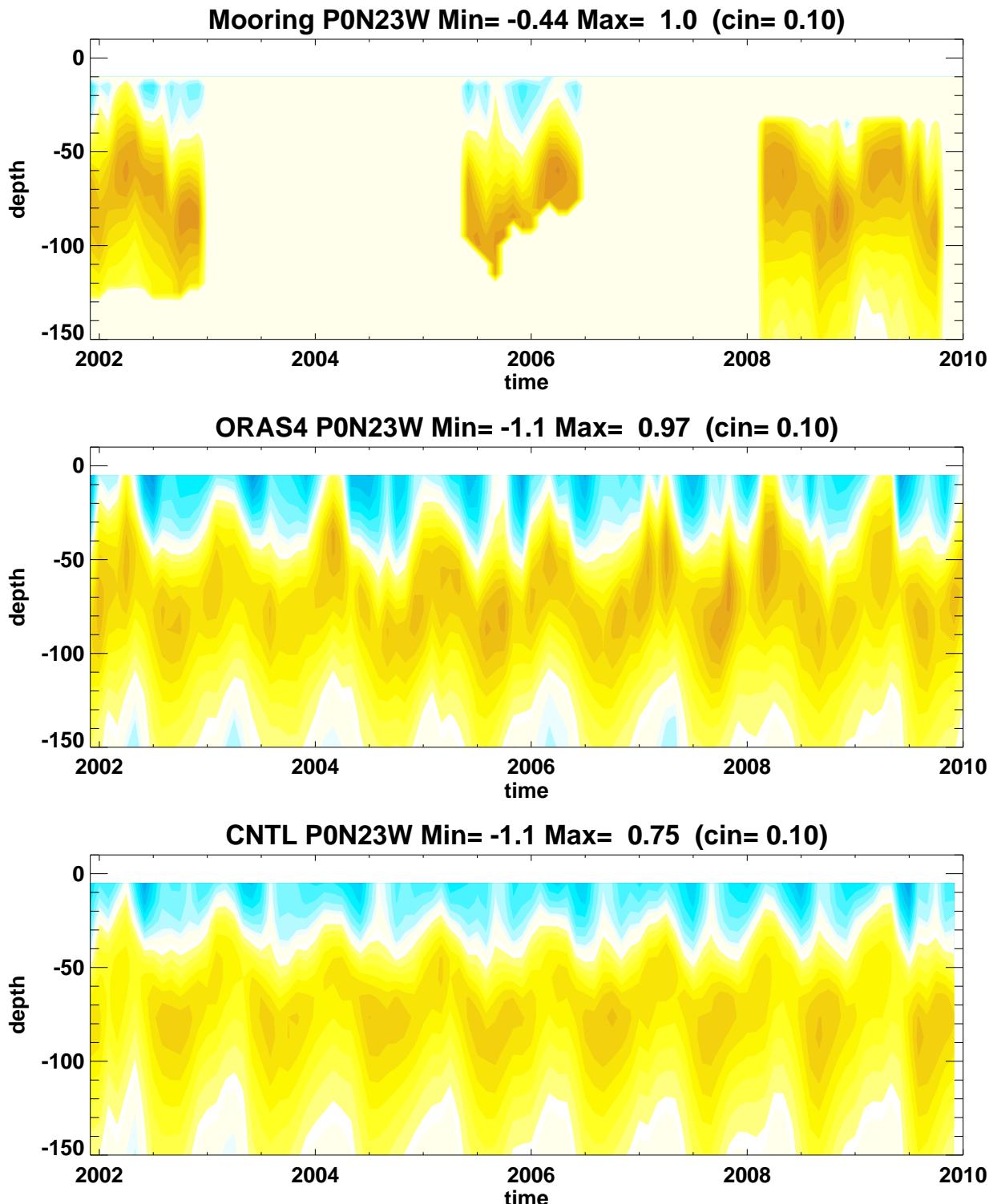


Figure 19: Time evolution of the zonal current (in m/s) of the PIRATA moorings at position (0°N, 23°E).

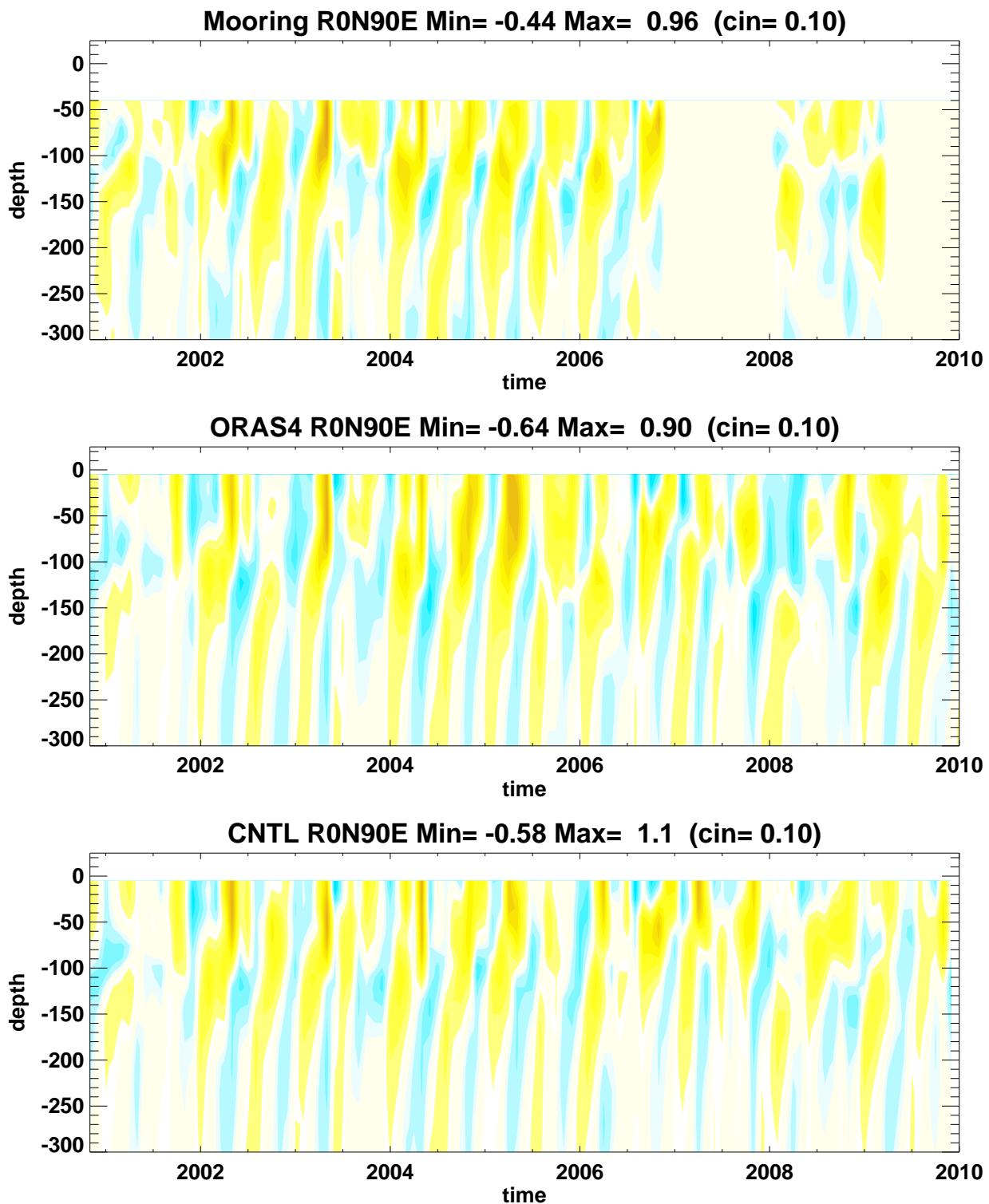


Figure 20: Time evolution of the zonal current (in m/s) of the RAMA moorings at position ( $0^{\circ}\text{N}$ ,  $90^{\circ}\text{E}$ ).

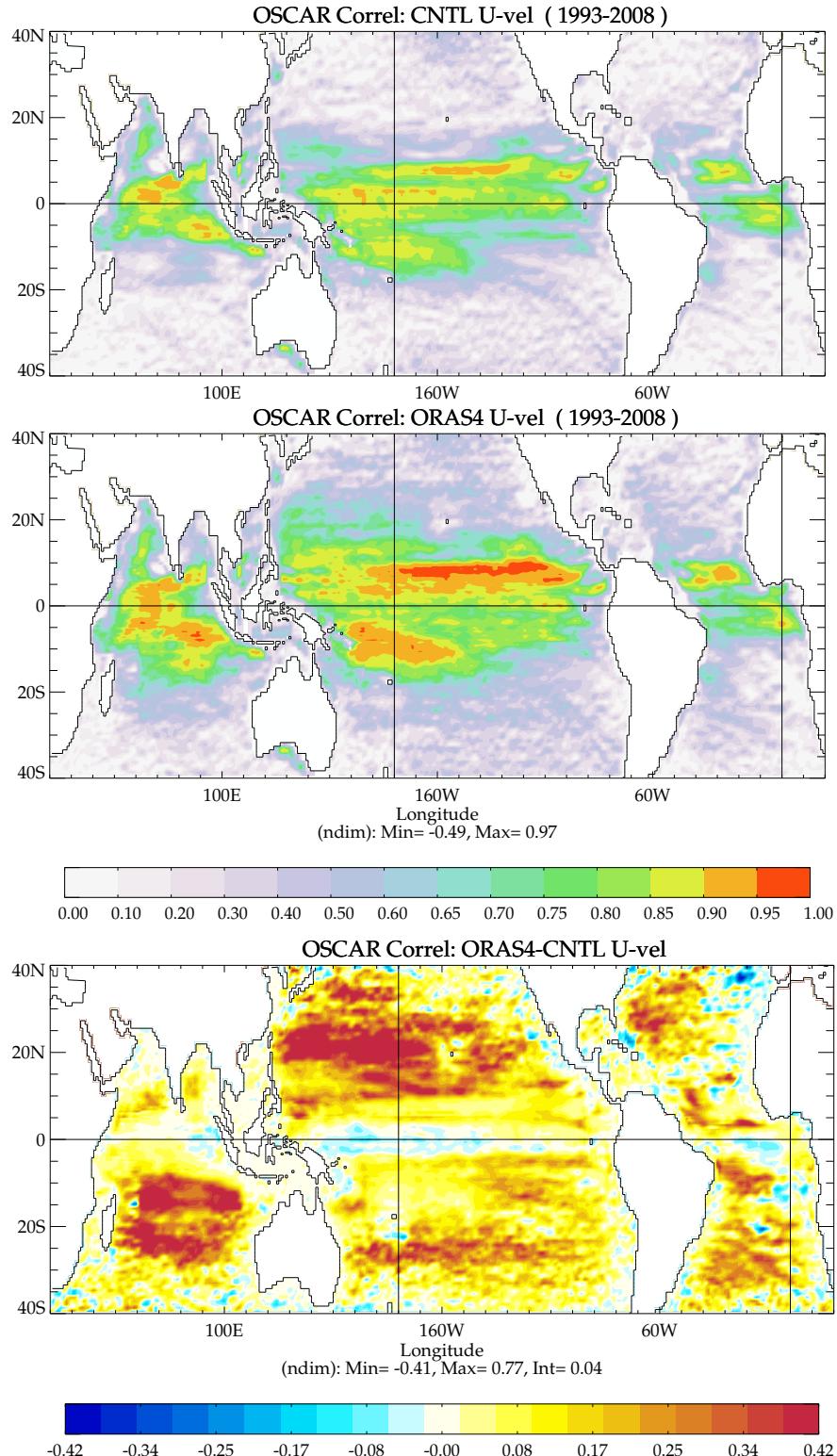


Figure 21: Temporal correlation between analysis and OSCAR surface zonal currents. Shown are the control experiment (top), ORAS4 (middle), and difference ORAS4 minus control (bottom). The statistics have been computed with monthly means for the period 1993-2009. Note the different color scales.

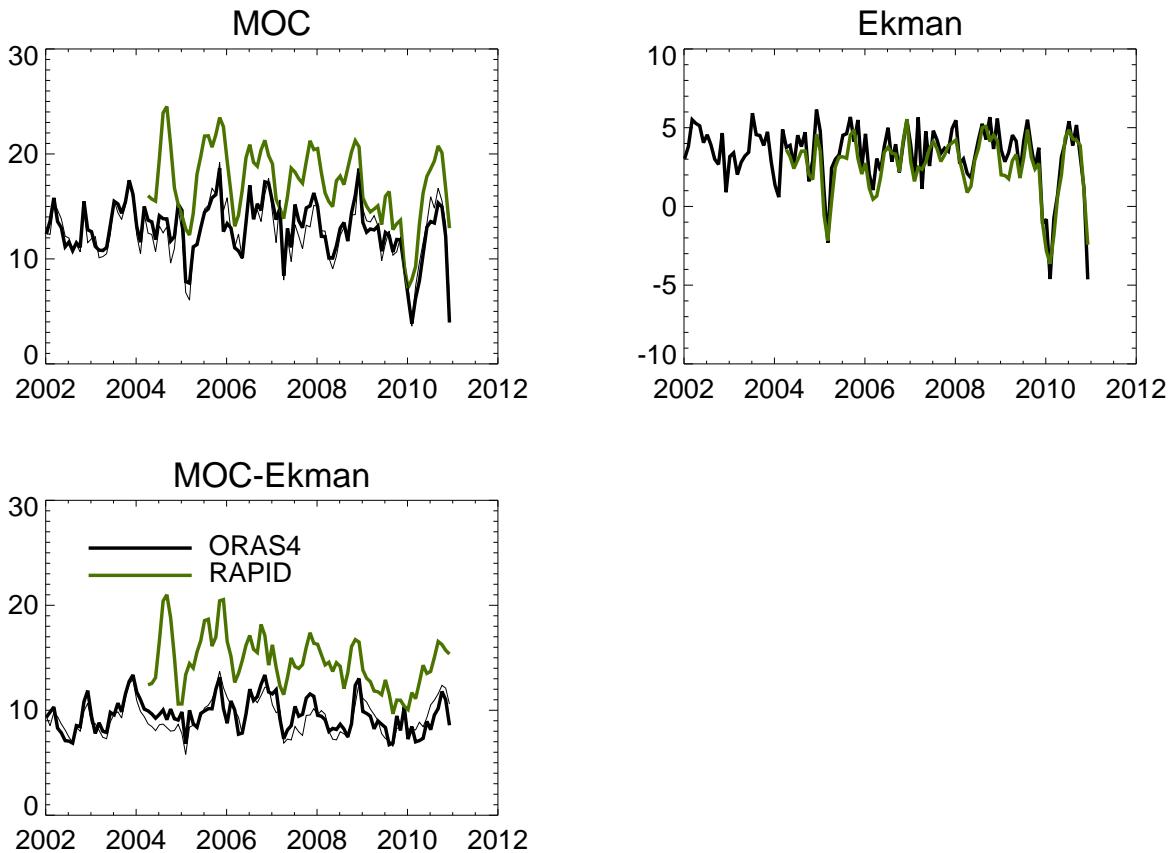


Figure 22: Time series of the Atlantic MOC from ORAS4 (black) and RAPID (green). ORAS4 underestimates the mean MOC value, but shows consistent time variability. The Ekman-driven minimum values at the end of 2009 and 2010 are well captured in ORAS4. Units are Sv.

the strength, variability and vertical structure of the Atlantic Meridional Overturning Circulation (MOC) with observations at 26.5°N.

Figure 22 shows the comparison between the Atlantic MOC at 26°N from ORAS4 (black) and RAPID (dark green). The total MOC in ORAS4 has been estimated following the RAPID method (thick line) and by directly integrating the meridional velocity (thin line). The small difference between the two estimates of the ORAS4 MOC confirms the validity of the RAPID methodology for MOC estimation, which derives the transport from temperature and salinity profiles at the RAPID locations (Cunningham *et al.*, 2007).

The MOC in ORAS4 is substantially lower than RAPID. The time series are not long enough to make strong statements about the time variability of the MOC, which is dominated by the seasonal cycle. Therefore, the time series appear relatively consistent in ORAS4 and RAPID, except for the 2004 peak in RAPID, which is absent in ORAS4. The Ekman components are very similar, since they are determined by the zonal wind-stress, although the Ekman component in ORAS4 corresponds to the ageostrophic residual rather than being directly calculated from wind-stress. ORAS4 captures the low MOC values at the end of 2009 and early 2010, and importantly the total MOC-Ekman difference indicates that this minimum has a geostrophic component as well as an Ekman component.

The vertical profile of the MOC in CNTL and ORAS4, averaged for the period 2004-2009, can be seen in Figure 23. The maximum of the MOC occurs at the correct depth in both CNTL and ORAS4, but it is too

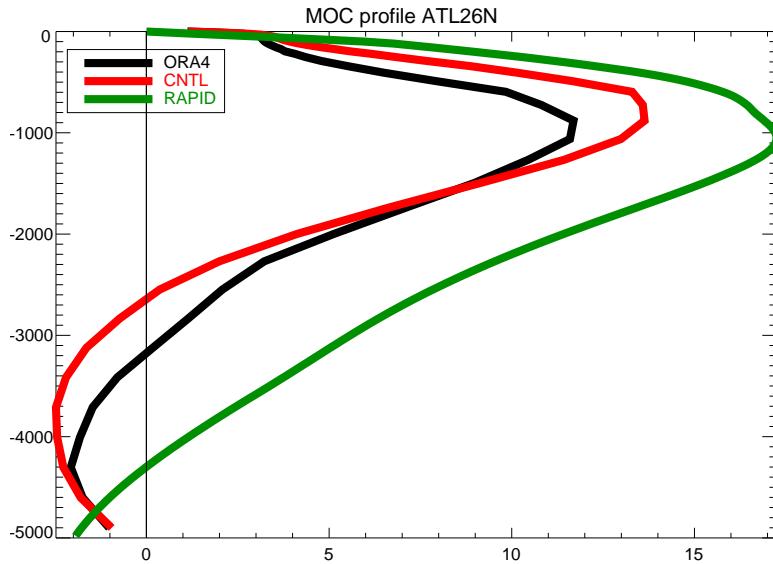


Figure 23: Mean vertical profile of the Atlantic MOC at 26N (Sv) from RAPID (green), ORAS4 (black) and CNTL (red). The y-axis is depth (m) and the x-axis shows the values of the transport (Sv). The average is for the period 2005-2009. Both CNTL and ORAS4 underestimate the MOC peak at around 1000 m and overestimate the AABW cell below 2000 m.

weak. The MOC vertical profiles also show that the deep Antarctic Bottom Water (AABW) cell is shallower in ORAS4 and CNTL than in RAPID, covering a larger depth range. The shallow and strong AABW cell seems to be model related and not caused by the data assimilation. Compared with CNTL, ORAS4 has weaker MOC maximum, but it does not overestimate the AABW cell as much as CNTL.

The RAPID data has been extensively used at ECMWF for the design and validation of ORAS4. The different MOC components (Ekman or ageostrophic, Gulf Stream or Florida Straits Transport (FST), and MidOcean transport) help us to better understand the origin of the differences between numerical models and the RAPID observations. Figure 24 shows time series of the MOC components from ORAS4 (black), CNTL (red) and RAPID (green). The time series have been smoothed with a 12-month running mean for a 50-year record (1960-2009). The long time series show a large degree of low frequency variability, highlighting that the current record of RAPID is not yet sufficient to evaluate the decadal variability of the MOC.

Both CNTL and ORAS4 underestimate the value of the total MOC, with ORAS4 values lower than those of CNTL. The underestimation comes mainly from the FST component, which is too low in the model-based estimates compared to observations, although the geographical separation of FST and MidOcean transport is not so clear in a low resolution model such as ORCA1. For instance, the larger interannual variability in the FST and MidOcean components is seen to be reduced in the total MOC, indicating anti-correlation in the gyre contributions. Assimilation gives a lower and more realistic FST component for the first 25 years or so of the run than that of CNTL, although both decline after the mid 1980s and converge to lower values than observed by the end of the period, the assimilation values remain slightly lower than those of CNTL. This result seems to contrast with results from the previous ECMWF data assimilation system (ORAS3) with HOPE/OI ([Balmaseda et al. \(2007b\)](#)), in which data assimilation was shown to maintain a stronger Atlantic MOC.

Both CNTL and ORAS4 show a decreasing trend in the MOC. Sensitivity experiments with different initial conditions indicate that the strong decreasing trend in CNTL in the 1960s and 1970s is mainly the artifact of model drift, but the decreasing trend after the 1980s seems quite robust in all the integrations performed with

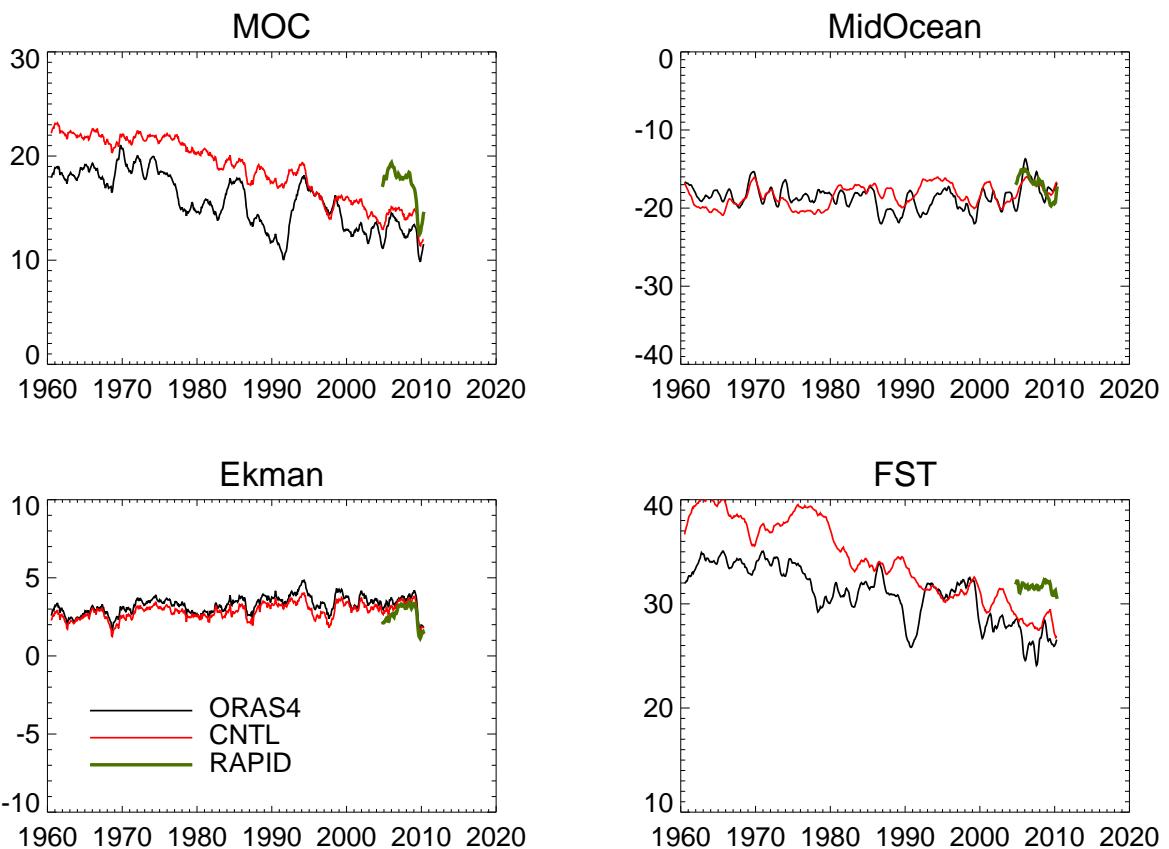


Figure 24: Time series of the Atlantic MOC and its components, at 26°N, from ORAS4, CNTL and RAPID. Both CNTL and ORAS4 underestimate the MOC from 2004-2009 due to the lower values of the FST. The FST is lower in ORAS4 than CNTL. Units are Sv.

NEMO and NEMOVAR. The decreasing trend seems to be associated with the FST, although, as mentioned before, the geographical separation of the FST and MidOcean transport is not so clearcut in a low resolution model such as ORCA1. The figures also show an obvious, albeit small, increasing trend in the Ekman component of the MOC. The Ekman components are of course very similar, since both CNTL and ORAS4 use the same wind-stress forcing. They do not coincide because, as mentioned earlier, the Ekman in ORAS4 and CNTL represent the ageostrophic residual rather than direct estimation from wind-stress.

The reasons for the lower MOC and FST in the data assimilation run compared with CNTL values have been investigated and are partially understood. The reduction of the FST was even larger in previous versions of the ocean data assimilation system, such as the one used to produce the COMBINE-NEMOVAR reanalysis [Balmaseda et al. \(2010\)](#). The FST would maintain its strength only if data in the ocean interior were assimilated. The closer the observations were assimilated near the coast, the lower the value of the FST. This appears counter intuitive at first sight, since one would assume that density gradients closer to observations would produce more realistic ocean currents. However, there are areas where geostrophic balance is not the only balance that should be considered. For instance, it is now well known that at the Equator, the balance between the wind-stress and pressure gradient determines the velocity structure, and its preservation is important in data assimilation. Equally, near large bathymetry gradients, the balance between the density and bathymetry gradients (The Joint Effect of Baroclinicity and Relief; JBAR) is important to determine the velocity structure. Assimilation of density observations may disrupt that balance, inducing undesirable effects on the ocean currents. This could explain why data assimilation can have a very different impact in ocean models if these have different bathymetry gradients. This hypothesis needs to be tested further. There may be several solutions to this problem, but none of them are easy to implement: i) to apply a correction to the momentum equations as to preserve the JBAR balance; ii) to improve the representation of the background-error covariances near coastlines; and iii) to assimilate the transports instead of the direct temperature and salinity observations. Although some of these developments are now ongoing, the operational implementation of NEMOVAR could not wait for them. To alleviate the problem of the FST and western boundary currents in general, the parameters  $\delta_c$  and  $D_c$  controlling the weight to the observations as a function of the distance to the coastline in (37) were retuned to give less weight to observations near coasts (see Table 1). It is possible that this problem will be alleviated when increasing the horizontal resolution of the ocean model.

## 5.4 Impact on coupled forecasts

As mentioned in the Introduction the ocean analysis system at ECMWF is used for initial conditions for the seasonal forecasting system as well as the EPS. It is therefore natural to benchmark the performance of the assimilation system by using the output of ORAS4 as ocean initial conditions for the seasonal forecasting system and comparing the results to those using output from CNTL as ocean initial conditions.

Figure 25 shows comparisons of forecast skill between seasonal forecast experiments initialized with ocean states produced by ORAS4 and CNTL. The experiments were conducted with the atmospheric version C36R4 (the same as the S4 seasonal forecasting system ([Molteni et al., 2011](#))), at a resolution of T159 with 62 levels in the vertical. Each experiment consists of 40 start-dates, 3-month apart, over the period 1989-2008. For each date, an ensemble of 5 members is integrated for 7 months. Figure 25 shows the skill of seasonal forecasts of SST in terms of anomaly correlation as a function of forecast time for different regions. It can be seen from the figure that the ocean initial conditions produced by ORAS4 improve the seasonal forecasts in most regions for all lead times and in all regions for short lead times (less than four months). The most noticeable impact is achieved over the Central Pacific, a region which is important for ENSO prediction. The positive impact of the assimilation in the Atlantic regions, although modest, is worth highlighting, since this is the first time in the history of seasonal forecasts at ECMWF that the ocean data assimilation is able to improve the forecasts in the

Equatorial Atlantic region.

## 6 Summary and conclusions

A new operational ocean analysis system (Ocean-S4) has been implemented at ECMWF. Ocean-S4 is based on the NEMO ocean model and the NEMOVAR ocean data assimilation system. The objective of this technical memorandum is to document the operational implementation of NEMOVAR in the Ocean-S4 system.

In the first part of this technical memorandum we have described NEMOVAR. NEMOVAR is a multi-incremental variational assimilation system which supports both 3D-Var FGAT and 4D-Var. The system is multivariate via a set of balance relations that are embedded in a control variable transformation. The system has a bias correction scheme and is capable of assimilating data from temperature, salinity and SLA observations. The initial implementation of NEMOVAR follows the existing OPAVAR system, with some modifications to the balance relations, the inclusion of a bias correction scheme, and differences in the treatment of observations. It forms the basis for future developments and refinement in the years to come both at ECMWF and at other institutions such as CERFACS, INRIA/LJK and the Met Office.

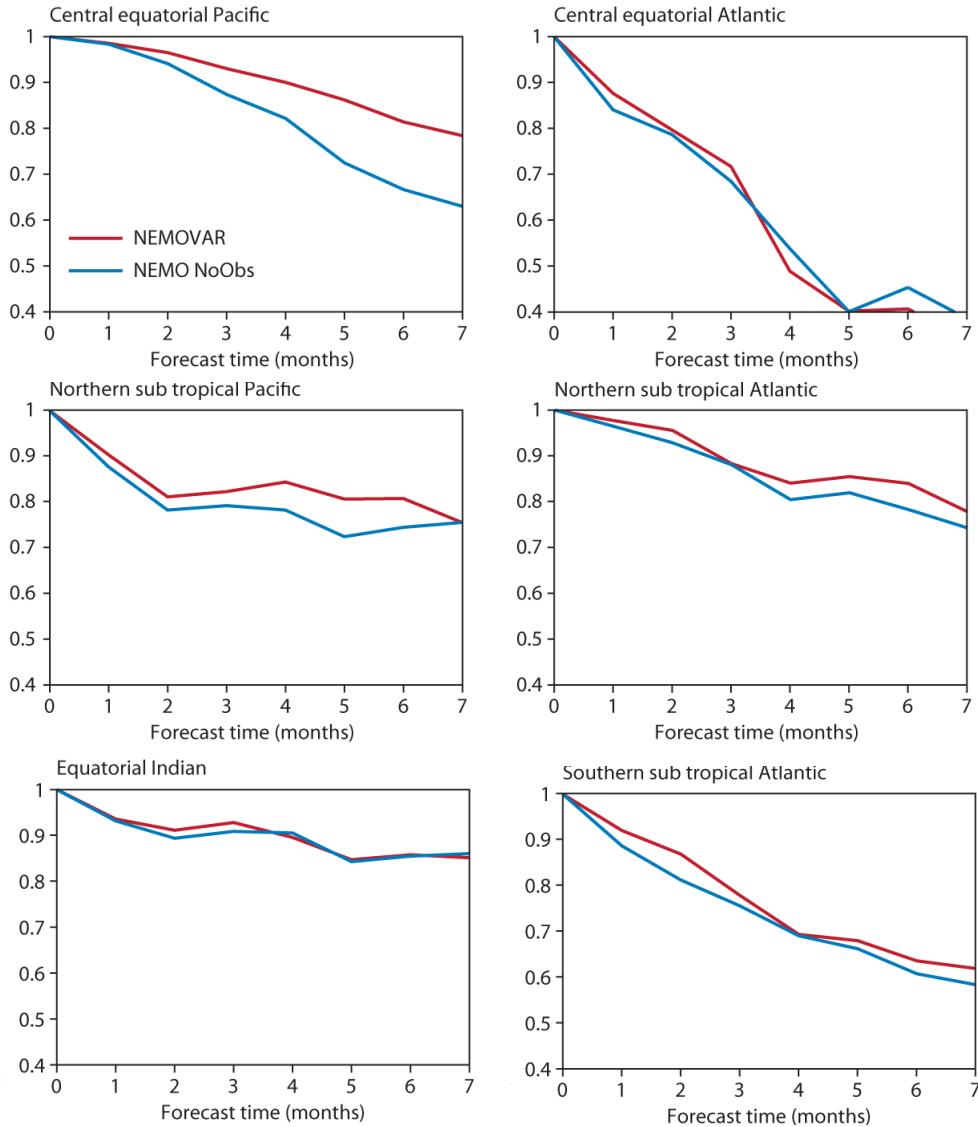
In a second part, we have documented the specific implementation of NEMOVAR in Ocean-S4. Ocean-S4 uses NEMOVAR in its 3D-Var FGAT configuration. The observation handling and quality control procedures, the specification of the background- and observation-error covariances, the balance relationships, the bias correction scheme, and the assimilation cycling procedure have been described.

Ocean-S4 consists of a reanalysis stream (ORAS4) and a real-time stream (ORTS4). The reanalysis stream starts in 1957 and is continuously updated to present. ORAS4 uses a 10-day assimilation window, and it is updated every 10 days with a 6-day delay. The latest ocean estimate from ORAS4 is used as a background state for the real time ORTS4, which is produced daily with a variable window, assimilating the latest observations. The real-time stream can be understood as an early-delivery stream, which, starting from the latest ORAS4 ocean state, brings the ocean analysis to real time on a daily basis. Details about the various components of the Ocean-S4 system have been provided: ocean model configuration, spin-up strategy, forcing fluxes, ensemble generation procedure, and operational schedule.

The performance of NEMOVAR has been evaluated by assessing the quality of ORAS4. An ocean only integration similar to ORAS4 but without assimilating data has been used as a control experiment. Statistics of the model fit to the assimilated observations indicate that NEMOVAR consistently improves the fit to observations of temperature and salinity in terms of bias and RMS. Compared to the control experiment, the bias in ORAS4 is relatively small, and more importantly, quite stationary in time. Time-series of the RMS fit to observations exhibit an interesting decreasing tendency. Several possible reasons that can explain this tendency are discussed. The interannual variability of the estimate also seems to be improved by NEMOVAR. Observations of ocean currents and transports, which have not been assimilated, have been used to provide an independent measure for evaluating the performance of NEMOVAR. The conclusion is that NEMOVAR improves the ocean estimate for most of the parameters investigated. An exception is the Atlantic Meridional Overturning Circulation (MOC), where NEMOVAR tends to underestimate its mean strength.

Initial conditions from ORAS4 have been used to initialize seasonal forecasts with positive impact on the skill of forecasting SST. The positive impact includes Atlantic regions which is a region we were unable to improve with previous ocean analysis systems at ECMWF.

Since the end of 2011, Ocean-S4 ocean initial conditions are used by both the EPS and seasonal forecasting system at ECMWF, both of which use the NEMO model as the ocean component. ORAS4 is used for calibrating hindcasts (or re-forecasting), and the real-time ocean analysis from ORTS4 is used by the actual forecasts. As



Region	Start longitude	End longitude	Start latitude	End latitude
CENTRAL EQ. PACIFIC	190°	230°	-5°	5°
CENTRAL EQ. ATLANTIC	340°	360°	-3°	3°
EQ. INDIAN	40°	120°	-5°	5°
N SubTrop PACIFIC	105°	270°	10°	30°
N SubTrop ATLANTIC	280°	20°	5°	28°
S SubTrop ATLANTIC	300°	20°	-20°	-5°

Figure 25: SST anomaly correlations for 7-month seasonal forecasts for the different regions defined in the table.

ORAS4 is continuously brought up-to present conditions, it constitutes a valuable tool for ocean and climate monitoring, which can complement the atmospheric reanalysis efforts.

## Acknowledgements

NEMOVAR has been a collaborative effort since the beginning of the project, with contributions from numerous people from different institutions. Without their work the operational implementation of NEMOVAR at ECMWF would not have been possible. We would also like to thank Andrew Coward from NOCS for his help in setting up the ORCA1 configuration at ECMWF, and Simon Good from the Met Office for his help in implementing the observational quality control system in NEMOVAR. The RAMA/TAO Project Office of NOAA/PMEI is acknowledged for providing the TAO/TRITON/RAMA currents used for validation of the ORAS4. We acknowledge the RAPID project for providing the MOC transport data. The MOC diagnostics have been done under the VALOR project. A. Weaver would like to acknowledge the financial assistance he received from the French LEFE-ASSIM and GMMC programmes which enabled him to make several visits to ECMWF.

## References

- Antonov JI, Locarnini RA, Boyer TP, Mishonov AV, Garcia HE. 2006. Volume 2: Salinity. In: *World Ocean Atlas 2005*, Levitus S (ed), NOAA Atlas NESDIS 62, U.S. Government Printing Office, Washington, D.C., p. 182 pp.
- Balmaseda M, Mogensen K, Molteni F, Weaver A. 2010. The NEMOVAR-COMBINE ocean re-analysis. COMBINE Technical Report (ISSN 2221-1128) 1.
- Balmaseda MA, Dee D, Vidard A, Anderson DLT. 2007a. A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Q. J. R. Meteorol. Soc.* **133**: 167–179.
- Balmaseda MA, Smith GC, Haines GC, Anderson DLT, Palmer TN, Vidard A. 2007b. Historical reconstruction of the Atlantic meridional overturning circulation from ECMWF operational ocean reanalysis. *Geophys. Res. Lett.* **34**: L23 615, doi:10.1029/2007GL031 645.
- Balmaseda MA, Vidard A, Anderson DLT. 2008. The ECMWF Ocean Analysis System: ORA-S3. *Mon. Weather Rev.* **136**: 3018–3034.
- Bell MJ, Martin MJ, Nichols NK. 2004. Assimilation of data into an ocean model with systematic errors near the equator. *Q. J. R. Meteorol. Soc.* **130**: 873–893.
- Bloom SC, Takacs LL, Da Silva AM, Ledvina D. 1996. Data assimilation using incremental analysis updates. *Mon. Weather Rev.* **124**: 1256–1271.
- Bonavita M, Raynaud L, Isaksen L. 2011. Estimating background-error variances with the ECMWF ensemble of data assimilations system: some effects of ensemble size and day-to-day variability. *Quart. J. Roy. Meteor. Soc.* **137B**(655): 423–434.
- Bonjean F, Lagerloef GSE. 2002. Diagnostic model and analysis of the surface currents in the tropical Pacific Ocean. *J. Phys. Oceanogr.* **32**: 2938–2954.
- Burgers G, Balmaseda M, Vossepoel F, van Oldenborgh GJ, van Leeuwen PJ. 2002. Balanced ocean-data assimilation near the equator. *J. Phys. Oceanogr.* **32**: 2509–2519.

- Church J, White NJ. 2006. A 20th century acceleration in global sea-level rise. *33*: L01602, doi:10.1029/2005GL024826.
- Cooper M, Haines K. 1996. Altimetric assimilation with water property conservation. *J. Geophys. Res.* **101**: 1059–1077.
- Courtier P, Thépaut JN, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**: 1367–1388.
- Cunningham SA, Kanzow T, Rayner D, Baringer MO, Johns WE, Marotzke J, Longworth HR, Grant EM, Hirschi JJM, Beal LM, Meinen CS, Bryden HL. 2007. Temporal variability of the Atlantic meridional overturning circulation at 26°N. *Science* **317**: 935–938. Doi:10.1126/science.1141304.
- Daget N, Weaver AT, Balmaseda MA. 2009. Ensemble estimation of background-error variances in a three-dimensional variational data assimilation system for the global ocean. *Q. J. R. Meteorol. Soc.* **135**: 1071–1094.
- Daley R. 1991. *Atmospheric Data Analysis*. Cambridge Atmospheric and Space Sciences Series, Cambridge University Press. 457 pp.
- Dee DP. 2005. Bias and data assimilation. *Q. J. R. Meteorol. Soc.* **131**: 3323–3343.
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Källberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597.
- Fisher M. 1998. Minimization algorithms for variational data assimilation. In: *Recent Developments in Numerical Methods for Atmospheric Modelling*. ECMWF, pp. 364–385.
- Fukumori I, Raghunath R, Fu L. 1998. Nature of global large-scale sea level variability in relation to atmospheric forcing: a modeling study. *J. Geophys. Res.* **103**: 5493–5512.
- Gill AE. 1982. *Atmospheric-Ocean Dynamics*. Academic Press.
- Golub GH, Van Loan CF. 1996. *Matrix Computations*. Johns Hopkins University Press: Baltimore, 3rd edn.
- Gratton S, Lawless A, Nichols N. 2007. Approximate Gauss-Newton methods for nonlinear least squares problems. *SIAM J. Optimization* **18**: 106–132.
- Ingleby B, Huddleston M. 2007. Quality control of ocean temperature and salinity profiles - historical and real-time data. *J. Mar. Sys.* **65**: 158–175.
- Jackett DR, McDougall TJ. 1995. Minimal adjustment of hydrographic data to achieve static stability. *J. Atmos. Ocean. Tech.* **12**: 381–389.
- Jones PW. 2001. A User's Guide for SCRIP: A Spherical Coordinate Remapping and Interpolation Package. Version 1.4. Technical report, Los Alamos National Laboratory.
- Lagerloef GSE, Mitchum GT, Lukas RB, Niiler PP. 1999. Tropical Pacific near-surface currents estimated from altimeter, wind and drifter data. *J. Geophys. Res.* **104**: 23 313–23 326.

- Locarnini RA, Mishonov AV, Antonov JI, Boyer TP, Garcia HE. 2006. Volume 1: Temperature. In: *World Ocean Atlas 2005*, Levitus S (ed), NOAA Atlas NESDIS 61, U.S. Government Printing Office, Washington, D.C., p. 182 pp.
- Lorenc AC. 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **112**: 1177–1194.
- Madec G. 2008. NEMO reference manual, ocean dynamics component : NEMO-OPA. Preliminary version. Note du Pôle de modélisation 27, Institut Pierre-Simon Laplace (IPSL), France.
- Madec G, Delecluse P, Imbard M, Levy C. 1998. OPA8.1 Ocean General Circulation Model reference manual. Technical note 11, LODYC/IPSL, Paris, France.
- McDougall TJ. 1987. Neutral surfaces. *J. Phys. Oceanogr.* **17**: 1950–1964.
- Mirouze I, Weaver AT. 2010. Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Q. J. R. Meteorol. Soc.* **136**: 1421–1443.
- Mogensen K, Balmaseda MA, Weaver AT, Martin M, Vidard A. 2009. NEMOVAR: A variational data assimilation system for the NEMO model. *ECMWF Newsletter*. **120**: 17–22.
- Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitart F. 2011. The new ECMWF seasonal forecast system (System 4). Tech. Memo. 656, ECMWF.
- Picaut J, Tournier R. 1991. Monitoring the 1979–1985 equatorial Pacific current transports with expendable bathythermograph data. *J. Geophys. Res.* **96**: 3263–3277.
- Reynolds RW, Rayner NA, Smith TM, C SD, Wang W. 2002. An improved in situ and satellite SST analysis for climate. *J. Climate*. **15**: 1609–1625.
- Ricci S, Weaver AT, Vialard J, Rogel P. 2005. Incorporating temperature-salinity constraints in the background error covariance of variational ocean data assimilation. *Mon. Wea. Rev.* **133**: 317–338.
- Rio MH, Guinehut S, G Larnicol G. 2011. New CNES-CLS09 global mean dynamic topography computed from the combination of GRACE data, altimetry, and in situ measurements. *J. Geophys. Res.* **116**: C07018. Doi:10.1029/2010JC006505.
- Stark JD, Donlon CJ, Martin MJ, McCulloch ME. 2007. OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system. In: *Oceans '07 IEEE Aberdeen*. Aberdeen, Scotland.
- Stockdale T, Anderson DLT, Balmaseda MA, Doblas-Reyes F, Ferranti L, Mogensen K, Palmer T, Molteni F, Vitart F. 2011. ECMWF seasonal forecast System 3 and its prediction of sea surface temperature. *Climate Dyn.* **37**: 455–471. 10.1007/s00382-010-0947-3.
- Troccoli A, Haines K. 1999. Use of Temperature-Salinity relation in a data assimilation context. *J. Atmos. Oceanic Technol.* **16**: 2011–2025.
- Tshimanga J, Gratton S, Weaver AT, Sartenaer A. 2008. Limited-memory preconditioners, with application to incremental four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.* **134**: 751–769.
- Uppala SM, Källberg PW, Simmons AJ, Andrae U Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **131**: 2961–3012.

- Vialard J, Weaver AT, Anderson DLT, Delecluse P. 2003. Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. Part 2: Physical validation. *Mon. Weather Rev.* **131**: 1379–1395.
- Villard A, Balmaseda M, Anderson DLT. 2008. Assimilation of altimeter data in the ECMWF ocean analysis system. *Mon. Wea. Rev.* **137**: 1393–1408.
- Vitart F, Buizza R, Balmaseda MA, Balsamo G, Bidlot JR, Bonet A, Fuentes M, Hofstadler A, Molteni F, Palmer T. 2008. The new VAREPS-monthly forecasting system: a first step towards seamless prediction. *Q. J. R. Meteorol. Soc.* **134**: 1789–1799.
- Weaver AT, Courtier P. 2001. Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc.* **127**: 1815–1846.
- Weaver AT, Deltel C, Machu E, Ricci S, Daget N. 2005. A multivariate balance operator for variational ocean data assimilation. *Q. J. R. Meteorol. Soc.* **131**: 3605–3625.
- Weaver AT, Mirouze I. 2012. On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Submitted to Q. J. R. Meteorol. Soc.* .
- Weaver AT, Vialard J, Anderson DLT. 2003. Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks. *Mon. Weather Rev.* **131**: 1360–1378.
- Wijffels S, Willis J, Domingues CM, Barker P, White NJ, Gronell A, Ridgway K, Church JA. 2009. Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. *J. Climate*. **21**: 5657–5672.
- Wolff J, E MR, S L. 1997. The Hamburg Ocean Primitive Equation model. Technical Report 13, Deutsches Klimarechenzentrum, Hamburg, Germany. 98.