TECHNICAL MEMORANDUM

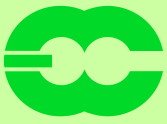
508

The ECMWF System 3 ocean analysis system

Magdalena Balmaseda, Arthur Vidard and David Anderson

Research Department

February 2007



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

Series: ECMWF Technical Memoranda

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

Contact: library@ecmwf.int

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

Abstract

A new operational ocean analysis system (system 3 or S3) has been implemented at ECMWF. It consists of two analysis streams: (i) a historical reanalysis from 01/01/1959 which is continuously maintained up to 11 days behind real time and is used to initialize seasonal forecasts, and (ii) an early delivery ocean analysis, produced daily in real time, used to initialize the monthly forecasts.

The S3 ocean analysis has several innovative features, including an on-line bias correction algorithm, the assimilation of salinity data on T-surfaces and assimilation of altimeter-derived sea level anomalies and global trends. The new S3 ocean analysis system outperforms the previous operational system (S2) in the tropics; the biases in both temperature and salinity are reduced, and the representation of the interannual variability is improved. In the extratropics S3 has larger interannual and decadal variability than S2, since the relaxation to climatology has been reduced, but the biases are larger.

It is shown that data assimilation has a large impact on the mean state of the first guess, and consistently reduces the bias. In the tropics, the interannual variability is improved, especially the ENSO related variability. In the extratropics, the variability is increased.

Data assimilation has a favourable impact on the skill of seasonal forecasts of SST, especially in the western Pacific, where the forecast skill in terms of RMS error is improved at all lead times. In the first 3 months, the forecast skill of the coupled model is improved by more than 20% by using data assimilation in the initialization of the ocean.

The S3 ocean analysis is a synthesis of available ocean observations, and provides a record of the history of the ocean. It is continued in real time and provides a valuable tool for climate analysis.

1 Introduction

A new operational ocean analysis system has been implemented at ECMWF to provide initial conditions for seasonal and monthly forecasts; it is envisaged that in the near future it will also be used to initialize coupled medium range weather forecasts. In what follows we will refer to this system as System 3 (S3). The S3 ocean analysis extends back to 1959, and can be regarded as a historical ocean re-analysis which is continuously updated on a daily basis. From the period 1981-2005, this ocean re-analysis is used to initialize the calibrating hindcasts of the S3 seasonal forecasting system (Anderson *et al.*, 2007). The earlier period of S3 ocean analysis will be used to initialize seasonal and decadal predictions within the ENSEMBLES project. As well as providing initial conditions for coupled model forecasts, the S3 ocean re-analysis, based on the synthesis of surface and subsurface ocean observations, surface fluxes from atmospheric analyses and reanalyses, and a general circulation ocean model, constitutes an important resource for climate variability studies, as will be shown at the end of this paper.

ECMWF has produced operationally daily global ocean analyses to provide initial conditions for the seasonal forecasting system since 1997. There have been two versions of the ocean analysis, linked to the operational seasonal forecasting system. System 1 (S1) started in 1997 (Alves *et al.*, 2003) and provided the initial conditions for the first ECMWF operational seasonal forecasting system (Stockdale *et al.*, 1998). System 2 (S2) was introduced in 2001 (Balmaseda 2004), and has provided initial conditions for the ECMWF operational seasonal forecasts since 2002 (Anderson *et al.* 2003, Oldenborgh *et al.* 2005a,b, Vialard *et al.* 2005). A comparison between S2 and S1 ocean analyses is given in Balmaseda 2004. In 2004 an extension of S2 was introduced in order to initialize the monthly forecasting system (Balmaseda 2005, Vitart 2005). In summer 2006 the S3 ocean analysis was implemented operationally[1], although at the time of writing the initial conditions for the

[1] S3 ocean analysis will start being used as initial conditions for the operational medium range, monthly and seasonal forecasts during 2007.

operational seasonal forecasting system are still provided by S2. In this paper, the S2 ocean analysis is used as a reference to evaluate the quality of the new system.

Although originally developed solely to provide ocean initial conditions for the seasonal forecast system, the scope of the ocean analysis at ECMWF has been slowly widening with time. Initially the length of the historical record of ocean initial conditions was not too long, since the hindcasts were mainly used to estimate the bias of the coupled model, which, although seasonally dependent, could be robustly estimated with a limited set of integrations. For instance, in the first seasonal forecasting system (S1), only 5 years of calibrating hindcasts were used. However, it was also necessary to provide an estimation of the forecast skill, and that required a larger historical sample, including as wide a range of climate conditions as possible. There was also the realization that calibrating not only the mean, but also the variance of the coupled model forecasts could lead to improved reliability of the seasonal forecasts. Besides, with the advent of multi-model activities it is clear that the robust bayesian combination and calibration of multi-model forecasts needs long records of realizations. All of these applications pointed towards the need for a long historical ocean reanalysis that could provide consistent initial conditions for the "calibrating" coupled hindcasts. And with a long record it is possible to start trying some decadal forecasts, as in the ENACT and ENSEMBLES projects, which try to assess the predictability at the decadal time scales. At the other end of the spectrum, moving towards the shorter time scales, we have the monthly and medium range forecasting activities, where the demand for ocean initial conditions is increasing. Monthly forecasting activities have stirred quite some interest in the last few years, and the benefits of having an active ocean in the forecasting system has been demonstrated (Vitart *et al.* 2006, Woolnough *et al.* 2006). The monthly forecasting system uses the same ocean model as the seasonal forecasting system though there are some differences in the way the ocean initial conditions are produced. In the near future it is planned that the medium range EPS forecasts will also be performed with a coupled model, with the consequent need for real-time ocean initial conditions. The EPS will also have a need for the historical ocean reanalysis, since the reliable prediction of extreme events also requires hindcasts or re-forecasting activities just as is done now for monthly and seasonal forecasting.

In order to accommodate these different demands, the S3 ECMWF ocean analysis has been designed to deliver two kinds of analysis products: a delayed product 11 days Behind Real Time (BRT), used to initialize the seasonal forecasts, and an early delivery product, which runs in Near Real Time (NRT), to be used by the monthly forecasts, and which in the near future is envisaged to be used by the medium range forecasting system as well. The BRT and NRT analysis systems are almost identical: they only differ in the details of the assimilation cycle and in the earlier cut-off of the assimilation windows. The BRT product is a continuation of the historical ocean reanalysis, while the NRT only exists for the recent period, having been introduced to comply with the strong operational constraints imposed by the monthly forecasting system. The details of the implementation of the BRT and NRT systems will be described in section 4. To avoid degradation of the NRT product, the NRT analysis always starts from the most recent BRT analysis and is then brought forward to real time every day. In principle, the design of the S3 ocean analysis system allows the paradigm of a seamless prediction system, since the forecasts for the different time scales (medium range, monthly and seasonal) could all be initialized using the NRT product. Equally, the BRT product would provide initial conditions for the hindcasts needed for the calibration (or re-forecasting) of the seamless products.

This document describes the different components of the S3 ocean analysis system and provides an assessment of its overall performance in comparison with the previous S2. It also investigates the impact of data assimilation in the representation of the ocean mean state, climate variability and seasonal forecast intialization. The impact of data assimilation for initialization of monthly and medium range forecasts is outside the context of this paper. This paper is organized as follows: the main differences between S2 and S3 are summarised in section 2 while in section 3 the data assimilation system in the S3 ocean analysis is described in some detail. In section 4 the differences in the assimilation cycles between the BRT and NRT ocean analyses, which arise

from the operational schedule are described, while in section 5 the differences between S2 and S3 are discussed in terms of representation of the climate, interannual variability and fit to the data. The impact of data assimilation for the description of the climate and for the initialization of seasonal forecasts is investigated in section 6. Finally, some of the climate signals represented by the S3 ocean re-analysis are shown in section 7.

2 Summary of new features: S3 versus S2

As for S2, the ocean data assimilation system for S3 is based on the HOPE-OI scheme: The first guess is given by forcing the HOPE (Hamburg Ocean Primitive Equations) ocean model with daily fluxes of momentum, heat, and fresh water, while the observations are assimilated using an Optimal Interpolation (OI) scheme. The HOPE ocean model (Wolff *et al.*, 1997) uses an Arakawa E grid horizontal discretization. Several modifications took place over the years at ECMWF (Balmaseda 2004, Anderson and Balmaseda 2006). The horizontal resolution was increased to 1^o x 1^o with equatorial refinement, i.e., the meridional resolution increases gradually towards the equator, where it is 0.3^o in the meridional direction. There are 29 levels in the vertical, with a typical vertical thickness of 10 meters in the upper ocean compared to 20 levels and 20m in S1. The vertical mixing is based on Peters *et al.*, 1998. The barotropic solver, originally implicit, was made explicit as described in Anderson and Balmaseda (2006).

In S3 major upgrades have been introduced in the HOPE-OI system. In addition to subsurface temperature, the OI scheme now assimilates altimeter derived sea-level anomalies and salinity data. All the observations in the upper 2000m are assimilated (in S2 only the observations in the upper 400m were used). In S3, the observations come from the quality controlled dataset prepared for the ENACT and ENSEMBLES projects until 2004 (Ingleby and Huddleston 2006), and from the GTS thereafter (ENACT/GTS). The OI scheme is now 3-dimensional, the analysis being performed at all levels simultaneously down to 2000m, whereas in S2, the analysis was carried out on each model level independently and only to 400m. In addition, the decorrelation scales depend on the density gradient, which favours the propagation of information along isopycnals. A pictorial view of the various data sets used in S3 is given in fig 1. The analysis of SST is not produced using the OI-Scheme. Instead, the model SSTs are strongly relaxed to analyzed SST maps. The maps are daily interpolated values derived from the OIv2 SST product (Smith and Reynolds 1998, Reynolds et al 2002) from 1982 onwards. Prior to that date, the same SST product as in the ERA40 reanalysis was used.

The first-guess is obtained from integrating the HOPE ocean model from one analysis time to the next, forced by ERA40/OPS fluxes (ERA40 fluxes from the period January 1959 to June 2002 and NWP operational analysis thereafter). In S2 the fluxes were from ERA15/OPS, but the wind stresses were not directly used: instead, the wind stress was derived from the analyzed winds using an off-line bulk formula. The representation of the upper ocean interannual variability is improved when using the ERA40 wind stress (Uppala *et al.*, 2006), although the stresses are biased weak in the equatorial Pacific. The fresh water flux from ERA-40 (Precipitation - Evaporation, denoted P-E) is known to be inaccurate. S3 uses a better but by no means perfect estimate, obtained by 'correcting' the ERA-40 precipitation values (Troccoli and Kallberg 2004).

When designing a data assimilation system for seasonal forecasts several considerations need to be taken into account. It is important to represent the interannual/decadal variability in the ocean initial conditions, and therefore strong relaxation to climatology is not recommended. On the other hand, in order to avoid spurious trends and signals due to the non-stationary nature of the observing system, the ocean analysis mean state should be close to the observations. It is also important to avoid large initialization shocks in the coupled model, which may damage the forecast skill. In S3 we have tried to strike a balance between the above requirements: the weight to observations has been reduced and the relaxation to climatology is considerably weaker than in S2. This has been possible because an additive bias correction has been included (Balmaseda et al 2007). The bias-

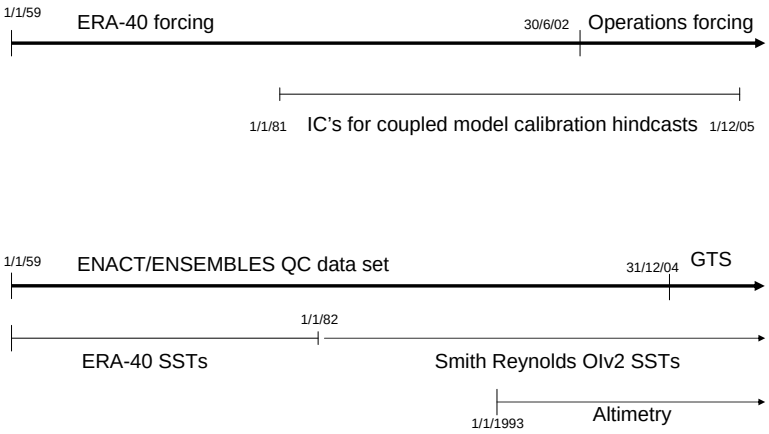


Figure 1: Upper panel shows the surface forcing used in the ocean analysis and the initial conditions for the calibration hindcasts for S3. Lower panel shows the origin of the subsurface data surface temperature fields used.

correction consists of a prescribed a-priori correction to temperature, salinity and pressure gradient, as well as a time-dependent bias term estimated on-line, which acts mainly on the pressure gradient. The on-line bias correction is adaptive and allows for flow-dependent errors.

An important feature of the ECMWF ocean analysis system is that not just a single analysis but several simultaneous analyses are performed. The purpose of the multiple analyses (five in total) is to sample uncertainty in the ocean initial conditions. The ensemble of ocean initial conditions provided by the five analyses contributes to the creation of the ensemble of forecasts for the probabilistic predictions at monthly and seasonal ranges. The five simultaneous ocean analyses are created by adding perturbations, commensurate with the estimated uncertainty, to the wind stress while the model is being integrated forward from one analysis time to the next. The wind perturbations have been revised in S3 to represent the perceived uncertainty in ERA40/OPS wind stress (http://www.ecmwf.int/research/EU_projects/ENSEMBLES/exp_setup/ini_perturb/index.html).

3 S3 Data Assimilation System

In S3, the different data streams are assimilated sequentially, as illustrated by the following scheme:

$$\begin{pmatrix} \tilde{T}^b \\ \tilde{S}^b \\ \tilde{\eta}^b \\ \tilde{\vec{u}}^b \end{pmatrix} \underset{bias}{\overset{remove}{\longrightarrow}} \begin{pmatrix} T^b \\ S^b \\ \eta^b \\ \vec{u}^b \end{pmatrix} \overset{alti}{\longrightarrow} \begin{pmatrix} T_1^a \\ S_1^a \\ - \\ - \end{pmatrix} \overset{Temp}{\longrightarrow} \begin{pmatrix} T_2^a \\ S_2^a \\ - \\ - \end{pmatrix} \overset{Sal}{\longrightarrow} \begin{pmatrix} - \\ S_3^a \\ - \\ - \end{pmatrix} \underset{SL-trend}{\overset{Geostr}{\longrightarrow}} \begin{pmatrix} T^a \\ S^a \\ \eta^a \\ \vec{u}^a \end{pmatrix}$$

First of all, the model background $(\tilde{T}^b, \tilde{S}^b, \tilde{\eta}^b, \tilde{\vec{u}}^b)$ is bias-corrected according to the scheme described in section 3.1. Then, the detrended altimeter-derived sea level anomalies (η'^o) are combined with the bias-corrected model first-guess $(T^b, S^b, \eta^b, \vec{u}^b)$ using the Cooper and Haines 1996 scheme (CH96 hereafter) to produce a first analysis $(T_1^a, S_1^a, \eta_1^b, -)$. This analysis is then used as a first guess for a second assimilation step, where only

subsurface temperature data T^o are assimilated, and salinity is updated by imposing conservation of the model temperature/salinity (T/S) relationship, while the sea level and velocity field remain unchanged. This second analysis is denoted as $(T_2^a, S_2^a, -, -)$. In a third assimilation step, the information provided by the salinity observations S^o is used to modify the model T/S relationship. In this step, the T/S information is spread along isotherms following the scheme of Haines *et al.*, 2006. Only salinity is modified in this step which results in the analysis $(-, S_3^a, -, -)$. After this 3rd assimilation step, velocity updates are derived from the temperature and salinity increments imposing geostrophic balance (Burgers *et al.*, 2002). Finally, the trend in global (area averaged) sea level is assimilated. By combining the altimeter-derived trend in global sea level with the model trend in global dynamic height, it is possible to make the partition between changes in the global volume and changes in the total mass. By doing so, the global fresh water budget is closed and the global surface salinity and sea level adjusted accordingly. Each of the steps is described briefly below.

The analysis is performed every 10 days. All the observations within a centered 10-days window are gathered and quality controlled. Analysis increments in temperature, salinity and velocity are calculated using the procedure outlined above. To avoid exciting gravity waves, and to allow the model dynamics to adjust gradually to the changes in the density field, an Incremental Analysis Update (IAU) method (Bloom *et al.*, 1996) is used: the increment is added slowly over the subsequent 10 days (IAU-10), after which a new background field is available, and the cycle repeated.

3.1 Bias-correction algorithm

The presence of bias in an ocean data assimilation scheme is a serious obstacle to the reliable representation of climate by historical ocean reanalysis. In the equatorial Pacific, the mean temperature assimilation increment in S2 is different from zero, and shows a large scale dipolar structure. Consistent with other assimilation systems, comparison with TAO currents shows that in S2 the equatorial zonal velocity in the Eastern Pacific is degraded when assimilating temperature data, even when salinity is also corrected by imposing preservation of the T-S relationship. The degradation of the zonal velocity is associated with a spurious vertical circulation underneath the thermocline, as pointed out by Bell et al 2005.

The standard procedure to deal with systematic error in a data assimilation system is to augment the model state 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 forecast 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 forecast covariance matrices is not generally appropriate since the bias and the model state vector can have different control variables and/or multivariate balance relationships.

An alternative approximation for the one-step bias correction algorithm for the general case has been implemented in S3, following Balmaseda *et al.*, 2007. The modifications include an explicit multivariate formulation which allows the balance constraints for the bias to be different to those for the state vector. In this context, the correction applied to the pressure gradient can be considered as a particular choice of balance relationship. The control variable of the bias vector is pressure gradient, while the state vector consists of the 3D temperature, salinity, velocity and sea level fields. The pressure gradient correction is derived from the analysis increments in temperature. The final temperature bias is the effect of applying the pressure gradient correction to the momentum equations.

Figure 2 illustrates the sensitivity of the results to the multivariate formulation of the bias covariances. It shows equatorial sections of the mean temperature increment (left column) and mean vertical velocity (right column).

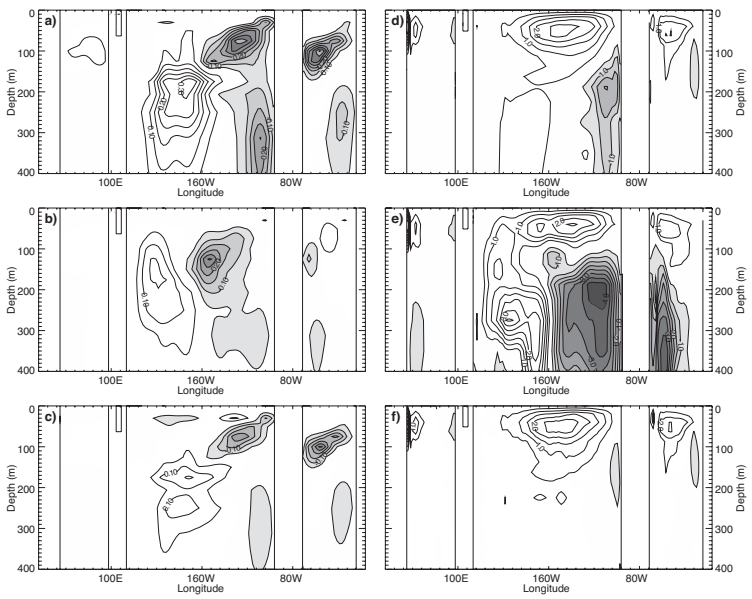


Figure 2: Equatorial longitude-depth section of mean temperature increment (left column) and vertical velocity (right column) for a standard experiment without bias correction (a and d), for an experiment with bias correction directly on temperature (b and e) and for an experiment with bias correction on the pressure gradient(c and f). Contours every 0.5 °C/10-days for the temperature increment and 0.5m/day for the vertical velocity. The zero contour is not plotted, and shading represents negative values. The mean corresponds to the time average during the period 1987-2001.

Results are from a standard data assimilation experiment without bias correction (upper row), an experiment where the bias is corrected directly in the temperature field (middle row) and an experiment where only the bias in the pressure gradient is corrected, following the Bell et al scheme.

The mean increment in fig 2a has the same large-scale dipolar structure in the equatorial Pacific as S2: the data assimilation is correcting the slope of the thermocline, making it deeper in the western Pacific and shallower in the eastern Pacific. Associated with the negative increment in the Eastern Pacific, there is a spurious vertical circulation (fig 2d). In the experiments where the bias has been corrected on-line (fig 2b and c), the resulting mean increment is smaller. However, in the experiment where the bias has been applied directly in the temperature field, the spurious vertical circulation in the Eastern Pacific is even worse (fig 2e). This degradation of the equatorial currents is consistent with that observed in experiments where the weight given to the observations is increased. In the experiment where the bias is treated by applying a correction to the pressure gradient using the Bell et al scheme (fig 2f), the spurious circulation does not appear. This illustrates the large sensitivity of the results to the choice of multivariate relationship: if the bias has its origins in the momentum equation (resulting from inaccuracies in the wind field and in the vertical mixing of momentum among others), the bias correction should be "adiabatic", since it is due to the wrong redistribution of heat.

Modifications have also been introduced in the equation for the time evolution of the bias, which is now described by a simple parametric model as in Balmaseda *et al.*, 2007. The introduction of a memory term limits the influence in time of isolated or sporadic observations. A side effect is that the magnitude of the bias can be underestimated but to compensate for that, a constant term is also introduced. The constant term is not affected by the on-line estimation and has to be estimated *a priori*, preferably with independent information. 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.

Because of the a-priori bias-correction term, the subsurface relaxation to climatology has been weakened: from

a time scale of 18 months in S2 to 10-years in S3. Because of large uncertainties in the fresh water flux, the relaxation to climatology is stronger for surface salinity (approx 3-year time scale), but still weaker than in the S2 analysis system (approx. 6 months).

3.2 Assimilation of altimeter-derived sea level anomalies

In S3, altimeter data are assimilated for the first time in the ECMWF operational ocean analysis. The altimeter information is given by maps of merged satellite product, provided by Ssalto/DUACS and distributed by AVISO, with support from CNES. Twice a week (on Wednesday and on Saturday mornings) $\frac{1}{3}^{o} \times \frac{1}{3}^{o}$ Maps of Sea Level Anomaly (MSLA) for a merged product combining all satellites (Envisat, Jason, Topex/Poseidon, ERS2, GFO) using optimal interpolation and accounting for long wavelength errors are produced (Le Traon *et al.*, 1998, Ducet *et al.*, 2000). Prior to assimilation, these maps are smoothed to remove unrepresented features and interpolated onto the model grid. These are then interpolated in time to produce daily maps. Only the map corresponding to the centre of the assimilation window is assimilated.

The anomaly maps (η'^{o}) distributed by AVISO are referred to a 7 year mean (1993-2000). To enable comparison with the background field (η^{b}), a reference mean sea level ($\bar{\eta}$) is required:

$$\delta\eta = (\eta'^{o} + \bar{\eta}) - \eta^{b}. \tag{1}$$

In S3, $\bar{\eta}$ is the 7-year mean sea level from an ocean analysis spanning the period 1993-2000. The possibility of using a reference $\bar{\eta}$ derived from the GRACE gravity mission has also been explored, but it was found that sensitivities to the reference mean sea level are large, and could potentially introduce abrupt jumps in the analysis (Vidard *et al.*, 2007, in preparation).

The scheme used to project sea level information into the subsurface temperature and salinity consists of calculating the equivalent vertical displacement of the model water column to a difference in surface elevation $\delta\eta$ between background and observations, subject to the constraint that the bottom pressure is not changed. Effectively, the CH96 vertical displacement of the water column translates into increments $(\delta T_1, \delta S_1)$ to the temperature and salinity field

$$\begin{pmatrix} \delta T_1 \\ \delta S_1 \end{pmatrix} = CH96(T^{b}, S^{b}, \eta^{b}, \eta'^{o}, \bar{\eta}) \tag{2}$$

from which a temporary analysis (T_1^{a}, S_1^{a}) is obtained:

$$\begin{pmatrix} T_1^{a} \\ S_1^{a} \end{pmatrix} = \begin{pmatrix} \tilde{T}^{b} \\ \tilde{S}^{b} \end{pmatrix} + \omega \begin{pmatrix} \delta T_1 \\ \delta S_1 \end{pmatrix}, \tag{3}$$

with ω a weighting factor depending on latitude:

$$\omega = \min\left(\alpha, \beta\left[\cos(lat)\right]^{\gamma}\right). \tag{4}$$

In S3, the values for α, β and γ are set to 0.3, 1 and 8 respectively.

3.3 OI assimilation of Temperature

S3 uses an Optimal Interpolation (OI) scheme for the assimilation for subsurface temperature. The S3-OI scheme, derived from the scheme described by Smith *et al.*, 1991, has evolved substantially: from the original univaratiate 2-dimensional OI scheme, where the analysis was performed on each model level separately, to a

3-dimensional scheme, where the analysis is performed at all levels simultaneously, with isopycnal formulation for the covariance matrices and a-posteriori multivariate updates of velocity and salinity. The OI interpolation is carried out on overlapping sub-domains of the model horizontal grid (in order to reduce the cost of the matrix inversions). Where domains overlap, the analyses are blended together. The subdivisions of the globe into sub-domains depends on the observation distribution and is done so that the maximum number of observations within the domain is less than 500 (for S2, the maximum number of observations was 200).

The horizontal model background errors are represented by Gaussian functions which are anisotropic and inhomogeneous. Within 4 degrees of the equator the correlation length scale in the E/W direction is 1000 km while in the N/S direction it is 150 km. The correlation scales change with latitude, such that polewards of 15 degrees, the correlation length scale is 300 km in all directions. Between the equatorial strip and the sub-tropics there is a smooth transition in correlation scales. Observation errors are assumed to be correlated in space and time, with a spatial correlation function with length scale of 2 degrees and a time correlation scale of 3 days. In general, the observations are given half the weight of the background, although the relative weight given to the data relative to the weight given to the background field varies with depth to account for an increase in uncertainty associated with the large gradients near the thermocline (see Balmaseda 2004). This function has been modified in S3 compared S2 to preserve the vertical structure of the profiles. The net effect is that near the thermocline, the observations in S3 are given less weight than in S2 (about four times less). This reduction of the background weight near the thermocline improves the analysis of the Equatorial Atlantic.

In S3, the assimilation of subsurface temperature is performed with 3D domains down to 2000m. In S3 there are 2 novel aspects: the introduction of vertical scales and the introduction of a density dependent term which favours isopycnal spreading of information. The assimilation of subsurface temperature follows the equation:

$$\begin{pmatrix} T_2^a \\ S_2^a \end{pmatrix} = \begin{pmatrix} T_1^a \\ S_1^a \end{pmatrix} + \begin{pmatrix} \mathbf{K}_{TT} \\ \mathbf{K}_{ST} \end{pmatrix} (T^o - \mathbf{H}T_1^a), \tag{5}$$

where the gain matrix $\mathbf{K}_{TT}$, used to calculate the increments in T is proportional to

$$\mathbf{K}_{TT} \sim e^{-\frac{1}{2}\left(\frac{|\Delta x|}{W_x} + \frac{|\Delta y|}{W_y} + \frac{|\Delta z|}{W_z} + \frac{|\Delta \rho|}{W_\rho}\right)} \tag{6}$$

The terms W_x, W_y in eq 6 represent the weighting given to the observations depending on the x and y distance respectively, and are anisotropic and inhomogeneous as described above. Eqn 6 shows explicitly: a) the introduction of vertical scales (W_z), which weights the influence of the observations depending on the vertical distance to the model grid point ($\mid \Delta z \mid$), and b) the introduction of a density dependent term W_ρ, which weights the observations depending on the background density difference between the observation location and the model grid point ($\mid \Delta\rho \mid$), thus preventing spreading the information across water masses with very different characteristics. In S3, W_ρ is $0.5 Kg/m^3$ and W_z is depth dependent, with the values proportional to the thickness of the model levels.

Fig.3 shows the assimilation increment coming from 3 single observation in temperature along the equator for an experiment using covariances as in S2 (top) and for the S3 analysis (bottom). Contours are background temperature isotherms. One can notice the two differences described above: the information is spread vertically in System 3[2] and the spread is not uniform but depends on density. At 200E/210m where the water-masses are well stratified the increment is almost an ellipsoid while at 120W/100m it has a more complicated shape.

When assimilating temperature data, the salinity field is also updated through the gain matrix $\mathbf{K}_{ST}$ in eq (5). The salinity corrections are derived from the temperature increments by conserving the background water mass properties. The basic idea is to use locally the model background T/S relationship to reconstruct the salinity

[2] It looks like there is a vertical spread in the top panel but it is an artifact of the plotting. It is in fact confined to a model level.

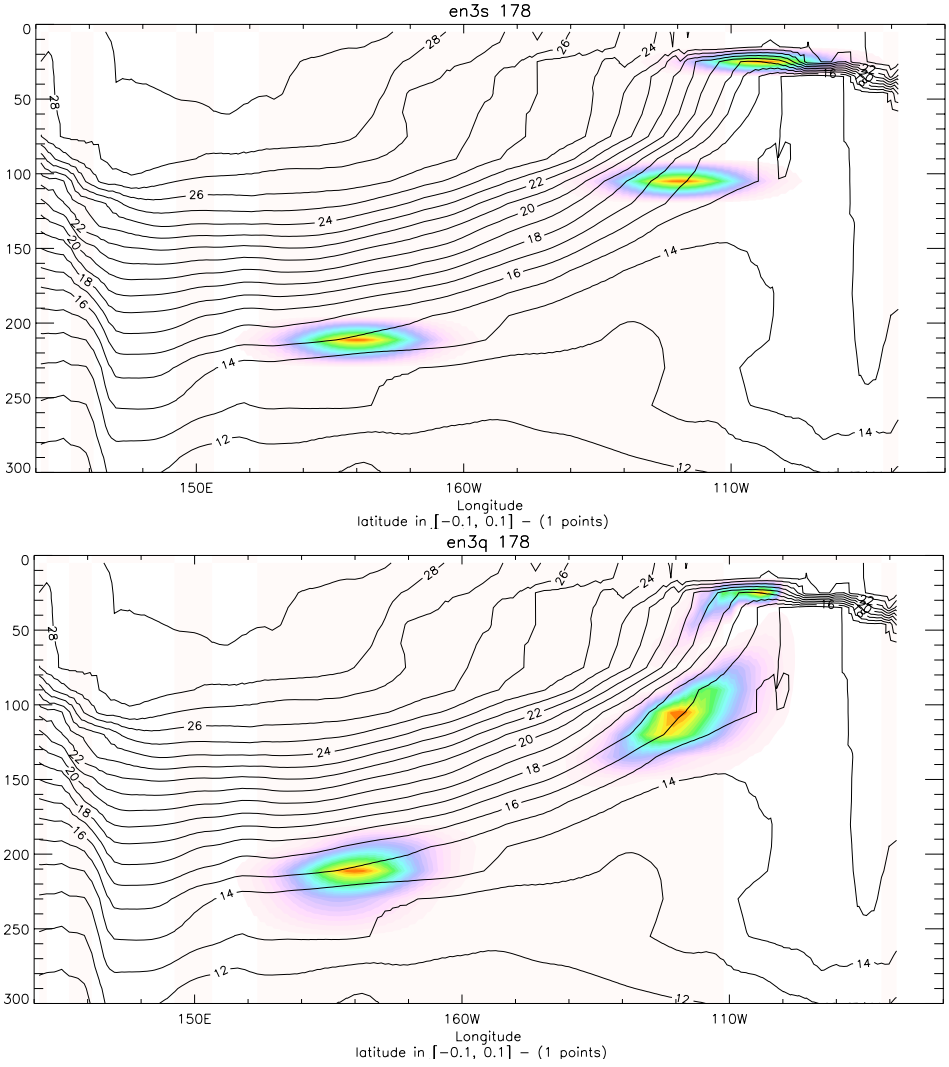


Figure 3: Temperature increment coming from the assimilation of three observation at different locations in the equatorial Pacific using the S2 (top) and S3 (bottom) covariances. Contours represent the background temperature field.

profile from temperature information only (Troccoli and Haines (1999), Troccoli *et al.*(2002)), hereafter TH99). It has been shown that temperature assimilation can improve the salinity field of an ocean model by taking advantage of the large fraction of salinity variance that is strongly correlated with temperature variance. Once the T_1^a analysis has been carried out, then salinity is updated following

$$\begin{aligned} S_2^a(z) &= S_1^a(z) + \delta S_2(z) \\ \delta S_2^a(z) &= \mathbf{K}_{ST}(T_1^a(z), S_1^a(z), T_2^a(z)) \end{aligned} \tag{7}$$

where $\delta S_2(z))$ ensures that the T/S relationship is the same before and after the analysis of temperature, i.e.

$$S_2^a(T_2^a) = S_1^a(T_2^a). \tag{8}$$

The correction on S is not applied in the mixed layer nor at higher latitudes where the hypothesis behind this scheme is less applicable. This is achieved by applying a latitude filter that reduces linearly down to 0 the effect of this scheme from 30^o to 60^o latitude.

3.4 Assimilation of salinity data on temperature surfaces

Another new feature of S3 is the assimilation of salinity data: with the recent development of the ARGO network we now have an unprecedentedly good spatial coverage of salinity observations.

Getting the salinity field right is important in a number of contexts. Salinity has an impact on the density field and hence on ocean currents and transports, e.g. Cooper (1988), Roemmich *et al.*(1994), Vialard and Delecluse (1998). Salinity is also important in certain places in the mixed layer where it controls the stability of the water column and hence to a degree, mixing and air-sea interaction, e.g. in the barrier layer around the western Equatorial Pacific. In addition, the relationship between temperature and salinity contains important information about the nature of the thermocline and subduction rates and areas. It is important therefore to recognize that the correct treatment of salinity data in the context of ocean data assimilation will allow analysed ocean fields to be used for more detailed studies of all of the above phenomena.

The conventional approach to assimilating salinity is to use covariance relationships formulated in (x,y,z) coordinates. However, by doing the analysis in (x,y,z) coordinates we are not taking advantage of the fact that the salinity increments in the TH99 scheme leave the salinity unchanged on T surfaces. Haines *et al.*, 2006 proposed an assimilation scheme by which the temperature and salinity provide two separate pieces of information about the hydrographic structure of the ocean: the temperature information is used to correct the temperature and salinity field by preserving the T/S relationship, and the salinity information can be used to correct the model T/S relationship. They also propose a change of variable when assimilating salinity information: instead of using geographical coordinates, the salinity assimilation will be done in temperature space. We refer to this scheme as S(T) in what follows, and the conventional scheme on geographical coordinates will be denoted as S(z).

Let us assume that at a particular location at level z the temperature analysis has already been performed and has yielded the temperature T_2^a, and salinity analysis S_2^a by preserving the first-guess S/T relationship (CH99). Following the S(T) scheme, the salinity analysis is carried out as follows:

$$S_3^a(T_2^a) = S_2^a(T_2^a) + \mathbf{K}_{SS}[S^o(T_2^a) - \mathbf{H}S_2^a(T_2^a)] \tag{9}$$

Written in temperature coordinates, it is clear that the correction from the S(T) scheme is correcting an entirely different aspect of the salinity field error from the correction by TH99. This elegant property means that the appropriate gain $\mathbf{K}_{SS}$ in the salinity assimilation and $\mathbf{K}_{ST}$ in the temperature assimilation are additive, which is a nice feature.

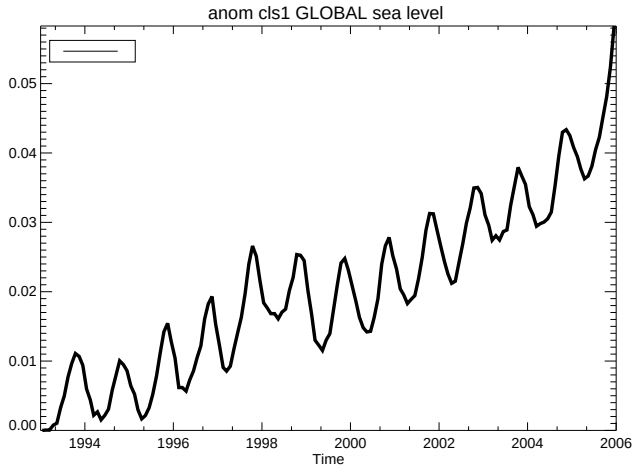


Figure 4: Time series of the global sea level from altimeter data. Units are metres.

In practice we need to express the result of eq (9) in z level coordinates. In z level coordinates eq (9) implies a second salinity increment which is additive to the first salinity increment of eq (7). This can be defined as;

$$\begin{aligned} S_3^a(z) &= S^o(T_2^a(z)) + \delta S_3(T_2^a(z)) \\ \delta S_3(T_2^a(z)) &= \mathbf{K}_{SS}[S^o(T_2^a(z)) - \mathbf{H}S_2^a(T_2^a(z))] \end{aligned} \tag{10}$$

that will be applied in addition to the first salinity increment $\delta S_2(T_2^a(z))$ associated with temperature assimilation. The spreading of the information along isotherms is determined by the gain matrix $\mathbf{K}_{SS}$:

$$\mathbf{K}_{SS} \sim W_R(\Delta r) exp(-(T_2^a - T^o)/W_T). \tag{11}$$

where $W_R(\Delta r) = e^{-\frac{1}{2}\left(\frac{|\Delta x|}{W_x}+\frac{|\Delta y|}{W_y}+\frac{|\Delta z|}{W_z}\right)}$ as in eq (6) and W_T is an additional 'temperature scale' that determines how salinity measurements on one temperature surface should influence salinity on another temperature surface. In S3, $W_T = 1\ ^\circ C$. The assumption that a large fraction of salinity variance is strongly correlated with temperature variance is not valid everywhere, especially at high latitudes. For that reason, using the same latitude filter as the salinity adjustment, the assimilation switches gradually between 30º and 60º from S(T) to S(z). Additionally S(z) is not used in the mixed layer where the search for the right temperature class may fail and produce strange results. Sensitivity experiments show that S(T) produces a better fit to the data than S(z), not only in salinity but also in temperature (Vidard *et al.* 2007, in preparation).

3.5 Assimilation of Sea Level Trends

Figure 4 shows the time series of the global sea level from the altimeter data, for the period 1993-2006. The trend in global sea level dominates the variability. If this trend is produced by thermal expansion due to global warming, it can not be represented by the ocean model as, in common with most ocean models used for climate activities, the Boussinesq approximation is made, which means that ocean models preserve volume. Therefore, if not treated correctly, the trend in sea level can be a problem when assimilating altimeter observations. In S3, the global sea level trend $\Delta\overline{\eta}$ is removed from the altimeter sea level anomalies before they are assimilated via the CH96 scheme. On the other hand there is an open debate about the attribution of sea level trend: how much is due to thermal expansion (steric) and how much is due to mass change over the ocean (Church and White 2006). In principle, ocean reanalyses can help to answer this question, since they use all possible information:

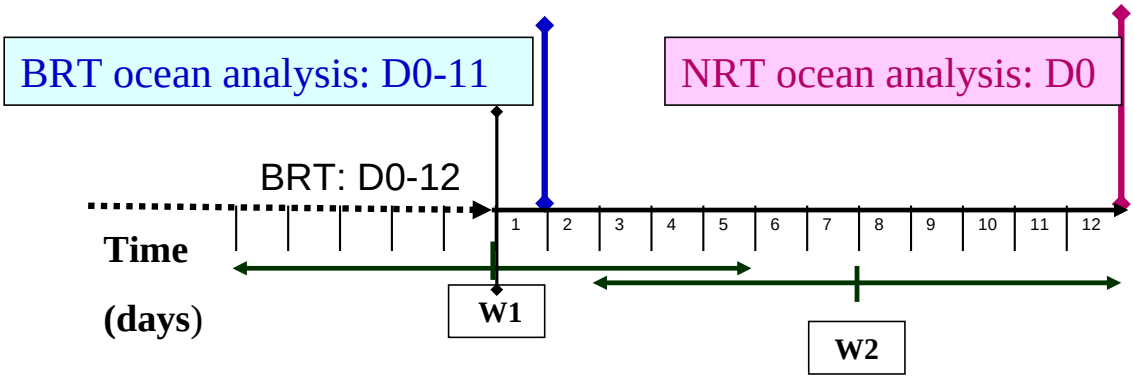


Figure 5: Assimilation cycle in the daily production of the behind-real-time (BRT) and near-real-time (NRT) ocean analyses in S3.

by combining model first guess with subsurface data it should be possible to reduce the error in the estimation of the dynamic height η_s. By comparing the total trend in sea level given by the altimeter data with the trend in steric height given by the ocean analysis, it is possible to estimate the component of the trend due to mass variations. In S3 the information given by the altimeter data about changes in the global mean sea level ($\Delta\overline{\eta}$) is compared every assimilation cycle with the changes in ocean analysis steric height ($\Delta\overline{\eta}_s$). The residual ($\Delta\overline{\eta}_m$) where

$$\Delta\overline{\eta}_m = \Delta\overline{\eta} - \Delta\overline{\eta}_s \tag{12}$$

is applied as a fresh water flux uniform in space. 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 problem in ocean analysis.

4 Daily production of S3 BRT and NRT ocean analyses

When in 2004 the monthly forecasting system became operational at ECMWF, it was necessary to produce near-real-time ocean initial conditions, since the typical 11 days delay of the BRT product was not adequate for the monthly system. To this end, an early delivery ocean analysis system was introduced in May 2004. It generated daily near-real-time (NRT) analyses of the ocean, ran only 8 hours behind-real-time and used the same assimilation scheme as in S2. Details of the timing of the NRT S2 ocean analysis are given in Balmaseda 2005. In S3, the NRT ocean analysis is an integral part of the operational ocean analysis system. In the near future the S3 NRT ocean analysis will provide initial conditions not just for the monthly forecast system but also, on a daily basis, for the Medium Range EPS forecasting system. The NRT assimilation cycle is slightly different from that described in Balmaseda 2005. Figure 5 shows schematically the schedule followed in the production of the NRT and BRT ocean analyses in S3. Every day, the BRT ocean analysis is advanced by 1 day, starting from the BRT analysis from the previous day (D0-12 in fig 5) to produce the BRT analysis at day

D0-11. If the 10-day assimilation cycle is due on that day (D0-12), the OI analysis is performed using a 10-day centered window, and the assimilation increment IAU-10 to be applied during the next 10 days is computed. If, however, no assimilation cycle is due at D0-12, the fraction of the assimilation increment IAU-10 from the previous assimilation cycle is simply applied to the first guess.

Similarly to the BRT analysis, the NRT ocean analysis is also produced daily. Starting from BRT (D0-12), the model is integrated forward 12 days, so bringing it up to real time. All the available observations during that period are used. During the 12 days there are always two assimilation cycles. The first assimilation takes place at (D0-12), using a 10-day window centred at D0-12 (W1), and the assimilation increment using the IAU-10 method is computed (W1-IAU-10). This increment is only applied for the next 7 days, at which time the second assimilation W2 is performed (i.e. at D0-5), using observations from the 10-day window as shown in fig 5. The corresponding IAU-10 increment is then computed (W2-IAU-10), and applied during the next 5 days of the integration. Effectively, W2 only has 9 days worth of data, since the daily averages for day D0 are not available at the time of the operational production. The assimilation windows W1 and W2 overlap for 3 days. The observations within this overlapping period, although used in two different analysis cycles are not given extra weight, since the increment IAU-10 is not applied to completion.

4.1 Observation coverage and Automatic Quality Control

Prior to 2004, the temperature and salinity profiles come from the ENACT/ENSEMBLES quality-controlled data set (Ingleby and Huddleston 2006). From January 2005, the observations come from the GTS (Global Telecommunication System). An automatic quality control procedure is used with several stages: first of all, daily averages are created if applicable: if some site reports more frequently than once per day, daily averages are created (this is the case for the TRITON moorings, which report hourly). Data close to the coast are rejected. A level-by-level quality control is then performed as in Smith *et al.*, 1991, which checks both the distance between model values and observations in relation to the error statistics as well as the consistency between observations by means of a buddy-check. Profiles which are close in space and time are superobbed, following the same criteria as in Smith *et al.*, 1991. In S3 there is an additional check for completeness of the profiles: a profile is considered incomplete, and therefore rejected, if the sparsity of the remaining observations in the vertical is judged insufficient to resolve the vertical temperature gradients. (An observation profile will be rejected if the temperature difference between consecutive levels is larger than 5 $^{\circ}C$ or if it contains a vertical temperature gradient larger than 0.1 $^{\circ}C/m$). Figure 6 shows the typical distribution of profiles in a 10-day window for the BRT ocean analysis (in this case of the window centered around the 22nd of August, 2006). Also shown are the results of the automatic quality control decisions. The figures show that thanks to the ARGO system the coverage of salinity is now comparable to that of temperature, and it is almost global. For comparison with the observation coverage in times prior to ARGO see *http://www.ecmwf.int/products/forecasts/d/charts.* Figure 6 also shows that only the TRITON moorings in the West Pacific and Indian ocean provide salinity in real time. The PIRATA and TAO moorings report only temperature in real time, even when the sensors are also able to measure salinity. Most of the data from XBTs and Mooring are superobbed, whilst the ARGO profiles are often partially rejected by our QC system.

Figure 7 shows the number of observations in a typical 10-day window from the BRT and NRT ocean analysis, for both temperature and salinity. ARGO is by far the most important contributor to the ocean observing system, followed by the Moorings and then by the XBTs (these latter only report temperature data). The ratio between the number of observations received in the NRT and BRT is close to 0.9 for ARGO and Moorings (0.9 is the expected ratio if all the observations arrived with only 24 hour delay, which would amount to W2 in the NRT having only 9 days worth of observations). The NRT/BRT ratio is smaller for XBTs (0.7).

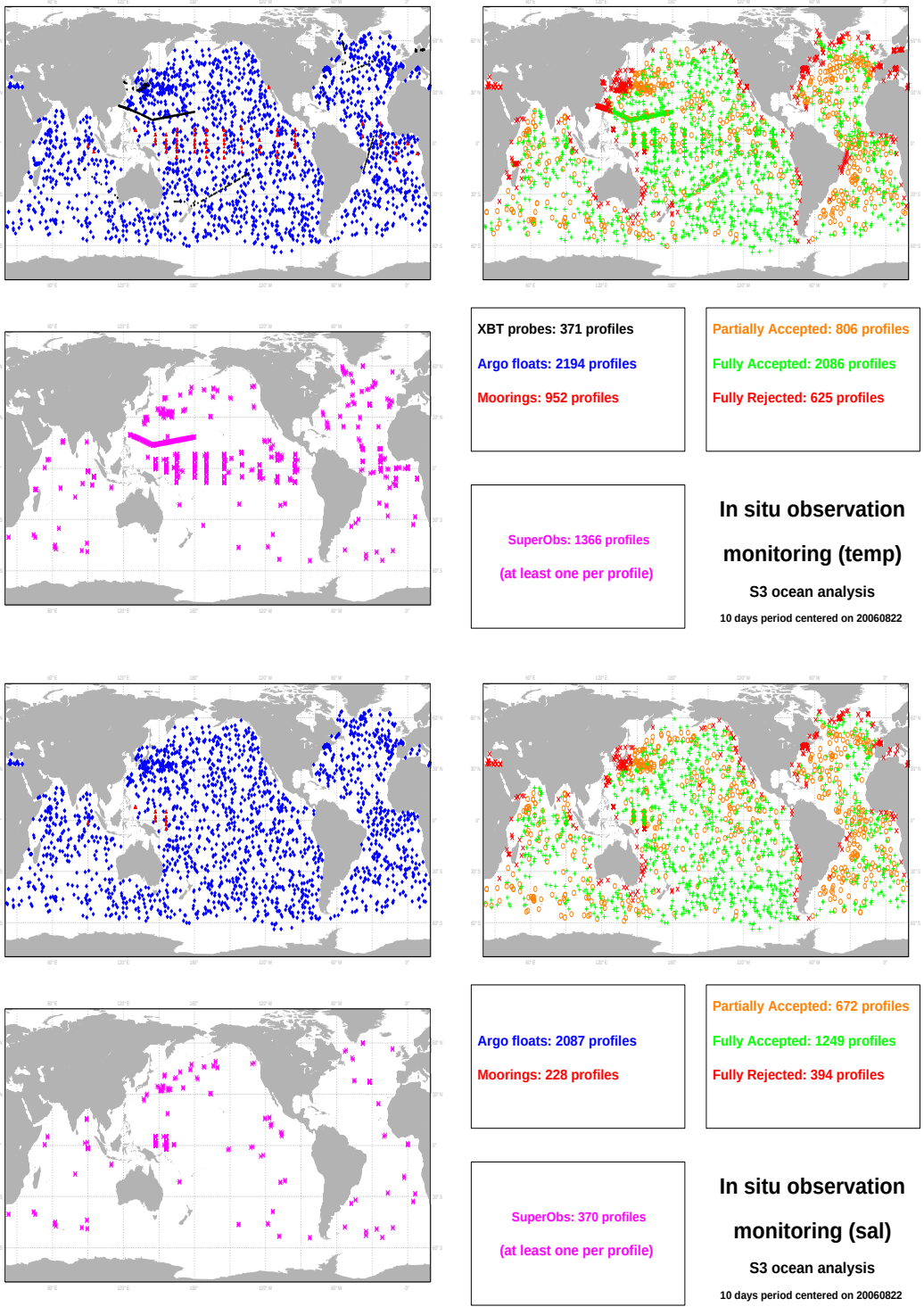


Figure 6: Typical 10-day distribution of temperature and salinity profiles in the S3 BRT ocean analysis. The upper block of panels is for T and the lower for S. The upper left panel of each block shows the coverage; the other two panels show the quality control and super-obbing decisions. Source: http://www.ecmwf.int/products/forecasts/d/charts.

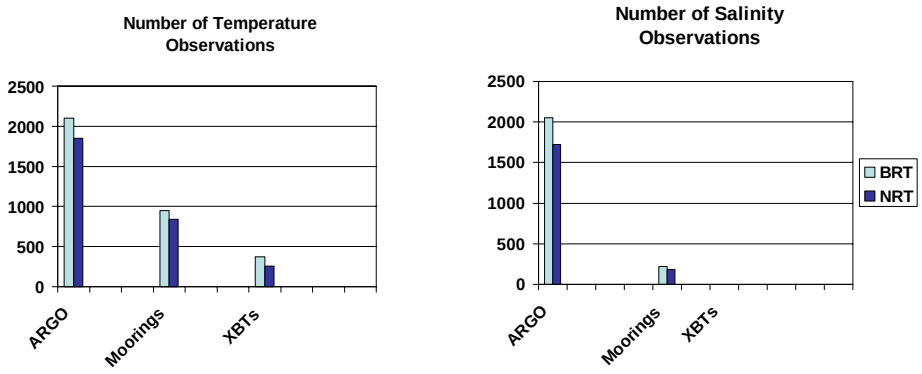


Figure 7: Number of temperature (left) and salinity (right) observations according to instrument within a typical 10-day window in the BRT and NRT S3 ocean analysis.

4.2 SST and altimeter in the NRT analysis

The arrival of SST information is one of the major factors for introducing a delay in the analysis. Global SST maps from NCEP are received every Monday at midday, representing the average of the previous week SST values. For the delayed ocean analysis, daily SST maps are obtained by interpolation of the weekly products, which requires the existence of two consecutive weekly values. This can introduce a delay of up to 12 days. The NRT analysis does not wait for the second map of SST to be available. Instead, a daily SST product is created by adding the latest SST anomaly to the daily climatology. Figure 8a shows the rms differences between the SST maps from the BRT and NRT ocean analyses, calculated for the period for which the two products overlap (August 2006 to October 2006). The difference between the NRT and BRT product is indicative of the error introduced by persisting the SST anomaly. For this time of year the largest differences (of the order of 0.5 $^{\circ}C$)

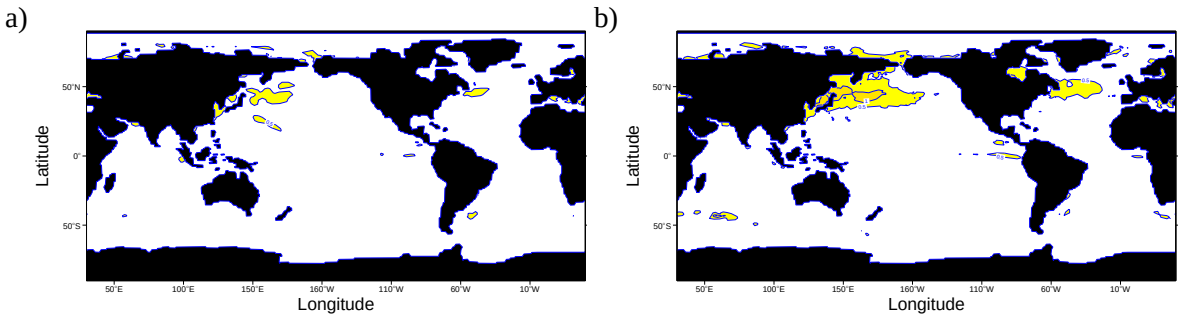


Figure 8: a) RMS difference between the SST maps from the NRT and BRT ocean analyses, indicative of the error introduced by persisting the SST anomaly. b) RMS difference between SST maps 7 days apart, indicative of the error introduced by persisting the full SST. In both cases, the rms differences have been calculated for the period August-October 2006. Contour interval 0.5 $^{\circ}C$

occur in the northern hemisphere, in the northern extensions of the Kuroshio and Gulf streams. The differences are about half the size of the typical 7-day persistence error, shown in fig 8b, suggesting that for SST the persistence error associated with the seasonal cycle is comparable with the persistence error of the anomaly.

The procedure of persisting the latest anomaly is also used in the preparation of the altimeter product that goes into the NRT analysis. The altimeter maps from AVISO are weekly maps. They are received twice per week, centered on Sundays and Wednesdays, with a delay of 7 days. For the BRT analysis, two consecutive maps are interpolated in time to produce daily maps. Only the daily map corresponding to the center of the assimilation window will be assimilated. In the NRT analysis, the daily map at the center of the assimilation window is produced either by time interpolation between two consecutive weekly maps where these are available, or by persisting the sea level anomaly from the latest map available. The anomaly is then added to the daily 1993-2001 climatology. The typical rms difference between the altimeter maps assimilated in the NRT and BRT does not exceed 5 cm (not shown), being largest in the Antarctic Circumpolar Current area. In the case of sea level, the error in persisting the anomaly is as large as persisting the full field.

5 Comparison with S2 ocean analysis

This section documents the changes between the S3 and S2 ocean analysis systems, both in mean state and variability. The attribution of changes to specific components of the new ocean analysis system is not always possible. To this end, more targeted studies would be necessary. An evaluation of the two systems is carried out by comparing both analyses with observations. The period 1987-2001 has been chosen for the comparison since this is the period that has been used to compute the anomalies for the S2 operational products. The period 2001-2006 will be subject to a more detailed study based on observing system experiments.

5.1 Mean State

The left column of figure 9 shows horizontal maps of the 1987-2001 climatology of the S3 analysis for selected variables. The sea level (SL), average temperature in the first 300m (T300), and average salinity in the first 300m (S300) are shown in the upper, middle and bottom row respectively. The right column of fig 9 shows the differences between the S3 and S2 analyses (S3-S2). The figure shows that the relative SL difference between tropics and extratropics is more pronounced in S3 than in S2: S3 has higher SL in the extratropics and lower SL within the band +/-10 degrees of the equator. The SL in S3 is also lower than in S2 along the western boundary of the north Atlantic basin. Both of these differences (i.e., SL high in the extratopics and low in the tropics and low SL along the Atlantic North American coast) are related to salinity changes: figure 9f shows a decrease in salinity in mid-latitudes in S3 with respect to S2, and an increase in the salinity content in the waters of the Western Atlantic. The SL difference in fig 9b also shows a deepening in the NEC/NECC trough, specially in the eastern part of the Pacific, which results in a strengthening of the equatorial current system (not shown). The differences in T300 shown in fig 9d are small in the tropics, although there is an overall impression that S3 has a stronger zonal temperature gradient than S2, with the Western Pacific warmer and the Eastern Pacific colder. Differences are larger in mid latitudes, on the northern extensions of the Kuroshio and Gulf streams.

Figure 10 shows equatorial sections of the 1987-2001 mean temperature (upper row) and salinity (lower row) for the S3 analysis (left) and differences S3 minus S2 (right). Figure 10b clearly shows that the large scale zonal temperature gradient in the Pacific is more pronounced in S3 than in S2. The temperature differences occur at thermocline depth and below, except for the far eastern equatorial Pacific, where S3 is warmer than S2 in a very narrow near coastal region above the thermocline. In the Indian ocean there is also a change in the zonal temperature gradient, with a deeper thermocline in the west and a shallower thermocline in the east. In the

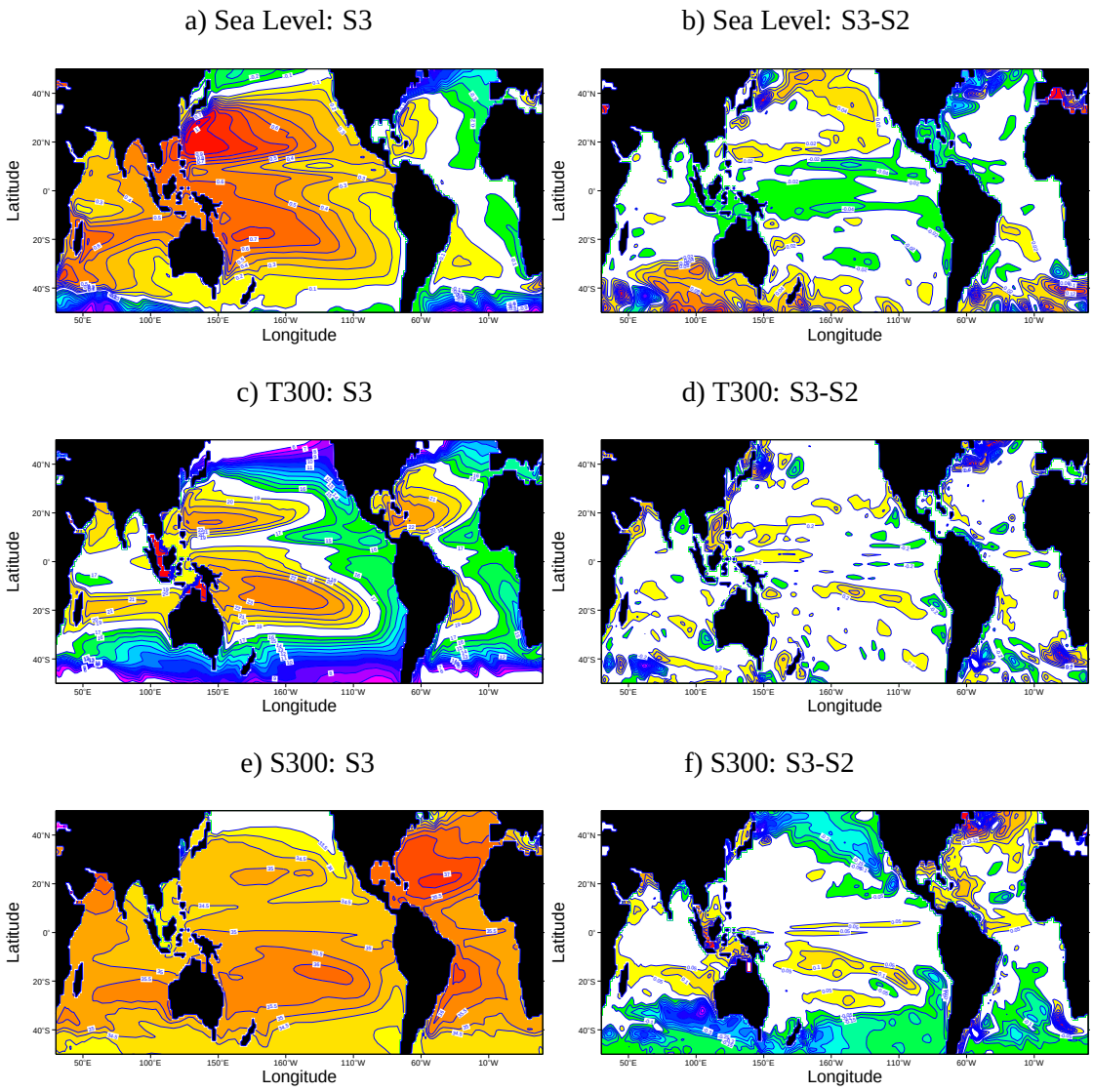


Figure 9: Mean sea level (upper row), T300 (middle) and S300 (bottom) for S3 analysis (left) and differences S3 minus S2. The averaging period is 1987-2001. The contour interval (reference level, blanked) for the full fields in the left column is 0.1m (0m) for SL, 1 °C (18 °C) for T300 and 0.5 psu (33 psu) for S300. The contour interval for the difference fields in the right column is 2cm for sea level, 0.2 °C for T300 and 0.05 psu for S300, and the zero contour is not plotted.

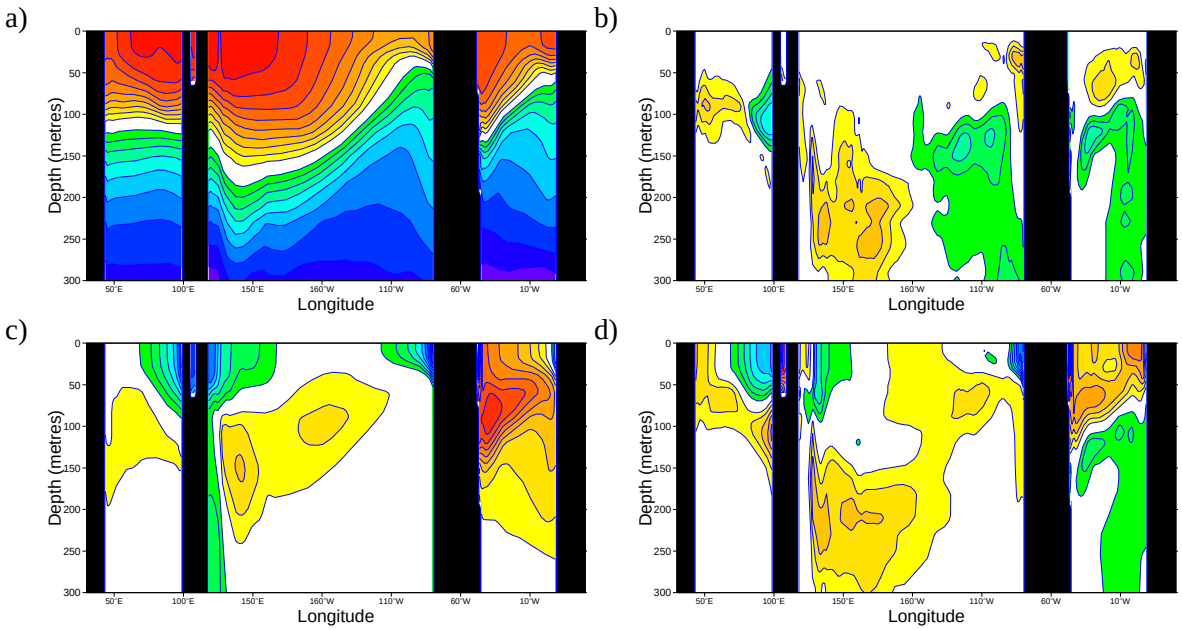


Figure 10: Equatorial zonal sections of the mean temperature (upper row), and salinity (bottom) for S3 analysis (left) and differences S3 minus S2 (right). The averaging period is 1987-2001. The contour interval (reference level) for temperature is 1 °C (20 °C) for the full fields and 0.2 °C (0 °C) for the difference fields. For salinity, the contour interval (reference level) is 0.2psu (35psu) for the full fields and 0.05psu (0psu) for the difference fields. The reference level is not plotted, and therefore, the two consecutive contours below and above the reference level are 2 contours apart. Values above (below) the reference level are in yellow/red (green/blue).

Atlantic ocean there are changes in the vertical temperature structure, with S3 more "stably" stratified than S2 (warmer waters above the thermocline and colder waters underneath). The stronger temperature stratification in S3 compensates for the stronger salinity inversion (fig 10c and d), i.e, saltier water above the thermocline and fresher waters underneath. In the Western Pacific and Eastern Indian ocean the barrier layers are more pronounced in S3 than in S2: water is fresher in the upper ocean and saltier underneath. In all three basins, changes in the stratification due to temperature are compensated by changes due to salinity.

Some of the differences between S3 and S2 may be related to changes in the surface fluxes. S3 uses daily fluxes from ERA40/OPS (left column of fig 11), while in S2 the fluxes were from ERA15/OPS. The difference between S3 and S2 wind stress can be seen in the right column of fig 11. In S3, the trade winds seem to be stronger in the tropics, but at the equator the differences between the zonal components of the wind stresses are small. Closer inspection reveals that the equatorial differences do not have coherent spatial scales, and rarely exceed $0.002N/m^2$ (not shown). Separate experimentation shows that while the differences between ERA40/OPS and ERA15/OPS can not explain the deeper thermocline in the Western Pacific, or the changes in the Equatorial Indian ocean, they have some impact on the depth of the Equatorial Atlantic thermocline (deeper with ERA40/OPS)

Figure 12a,b shows an equatorial longitude-depth section of the mean temperature increment from the assimilation in S3 (a) and S2 (b), for the period 1987-2001. The assimilation increment is smaller in S3 than in S2 in the 3 ocean basins and at most depths. This is consistent with the treatment of the bias in the S3 ocean analysis, although it may also reflect the weaker weight given to the observations near the thermocline in S3. In spite of the smaller weight given to the observations, the fit to the data is generally better in S3 than in S2, as will be shown later in this section. The salinity assimilation increment is shown in figures 12c,d for S3 and S2 respectively. In the Eastern Pacific and Atlantic, the salinity increment in S3 has the same structure

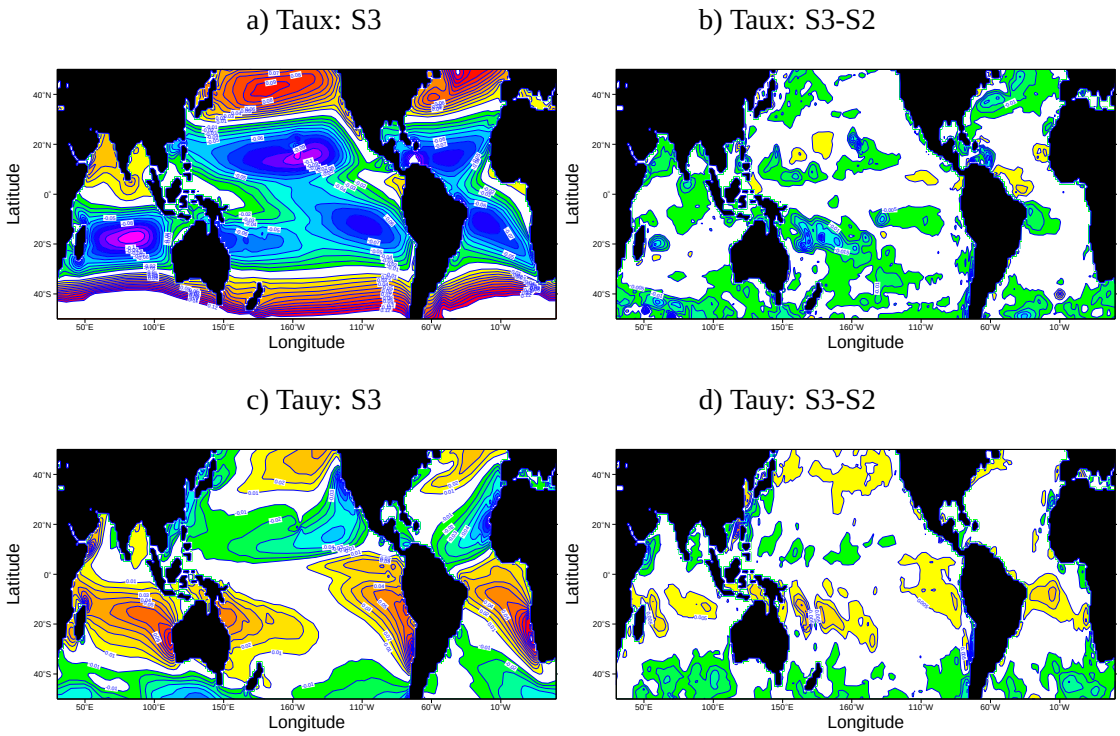


Figure 11: Mean zonal (upper row) and meridional (lower row) wind stress for S3 analysis (left) and differences S3 minus S2. The averaging period is 1987-2001. The contour interval for the full fields in the left column is $0.01N/m^2$*, and* $0.005N/m^2$ *for the difference plots in the right column. As in previous figures, the zero contour is not plotted.*

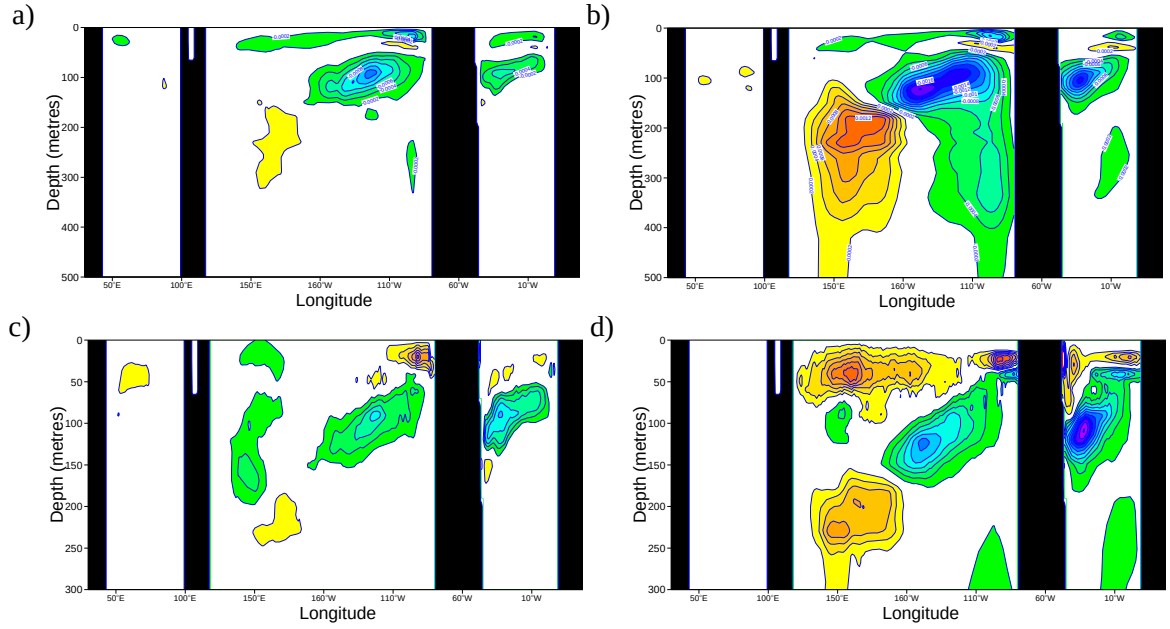


Figure 12: Mean assimilation increment for S3 and in S2. The upper row shows the temperature increment for S3(a) and S2(b). The lower row shows the salinity increment for S3 (c) and S2 (d). The contour interval is $2. \times 10^{-4}$ *°C/hr for the temperature and* 0.2×10^{-5}*psu/hr for the salinity increment. The averaging period is 1987-2001. The reference level is blanked, and therefore, the two consecutive contours below and above the reference level are 2 contours apart. Values above (below) the reference level are in yellow/red (green/blue).*

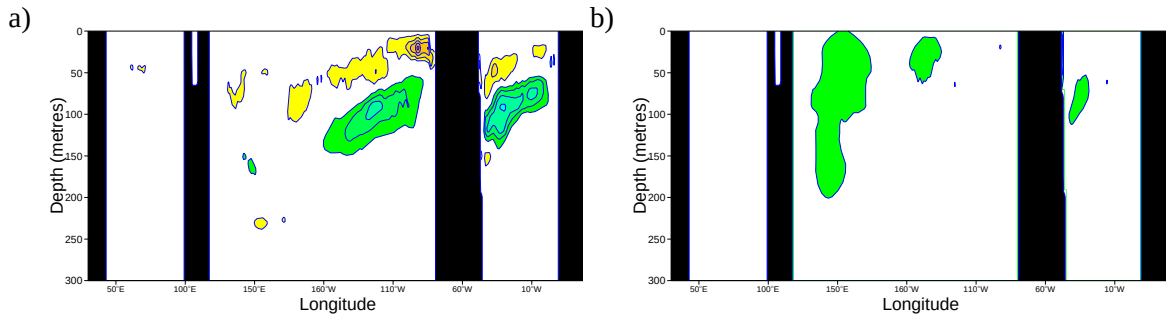


Figure 13: Mean salinity increment in S3 due to the balanced component (a) and to the direct assimilation of salinity observations (b). The contour interval is $0.2X10^{-5}$*psu/hour. The reference level is blanked, and therefore, the two consecutive contours below and above the reference level are 2 contours apart. Values above (below) the reference level are in yellow/red (green/blue). The averaging period is 1987-2001.*

as in S2, but its amplitude is much smaller. The salinity increment is mainly due to the balance operator that preserves the T/S relationship (CH96), which is the only contributor to the salinity increment in S2. (In S3, the "unbalanced" component of the salinity increment is due to the direct assimilation of salinity observations). This being the case, the smaller temperature increment in S3, would result in a smaller salinity increment from the balance operator. This fact is confirmed by figures 13a,b showing respectively the balanced and unbalanced components of the S3 salinity increment. During the period 1987-2001 the direct impact of assimilating salinity observations is more noticeable in the Western and Central Pacific, where the salinity increment is negative. The differences in salinity between S3 and S2 in relation to the difference in the salinity increment is quite remarkable, and points towards a better balanced system: S3 does not need a large salinity increment in order to keep the salinity maximum in the upper ocean. The fact that in S3 the balanced component of the salinity increment during the period 1987-2001 is larger than the unbalanced component may be related to the sparsity of salinity observations during this period. This is no longer the case after both the TRITON moorings and ARGO floats started reporting salinity measurements.

The idea that S3 is a more balanced system than S2 is also supported by fig 14, showing equatorial sections of the zonal and vertical velocities for S3(a,c) and S2(b,d). Changes in the vertical velocity structure are most noticeable in the eastern part of the three basins, where the downwelling that was present in S2 (very strong in the eastern Pacific and Indian oceans) has been substantially reduced in S3. These results are consistent with the effect of correcting the bias in the pressure gradient discussed in section 3. There is also a change in the zonal velocity (figures 14a,b). In S3, the surface current (westward SEC) is stronger and the east-west tilt of the undercurrent more pronounced than in S2.

Figure 15 shows meridional sections of the 1987-2001 mean temperature in the Western Pacific, at 165E (upper row), and in the Atlantic, at 30W (lower row). As in previous figures, the left column shows the fields for S3 and the right column shows the difference S3 minus S2. The contour interval and reference level are 1^oC and 20^oC respectively for the full temperature fields, and 0.5^oC and 0^oC for the difference plots. Fig 15b shows that in the West Pacific, the meridional hemispheric tilt is not as strong in S3 as in S2: in the northern hemisphere and at the equator, the thermocline is deeper in S3 than in S2, whilst in the southern hemisphere the thermocline is shallower in S3 than in S2 (in particular ridges around 2S and 8S-12S are more pronounced in S3). The meridional structure in the tropical Atlantic is also changed in S3 (fig 15d), which now shows a tighter thermocline; the thermocline troughs on both sides of the equator around 2-4S and 2-4N are also more pronounced in S3 than in S2.

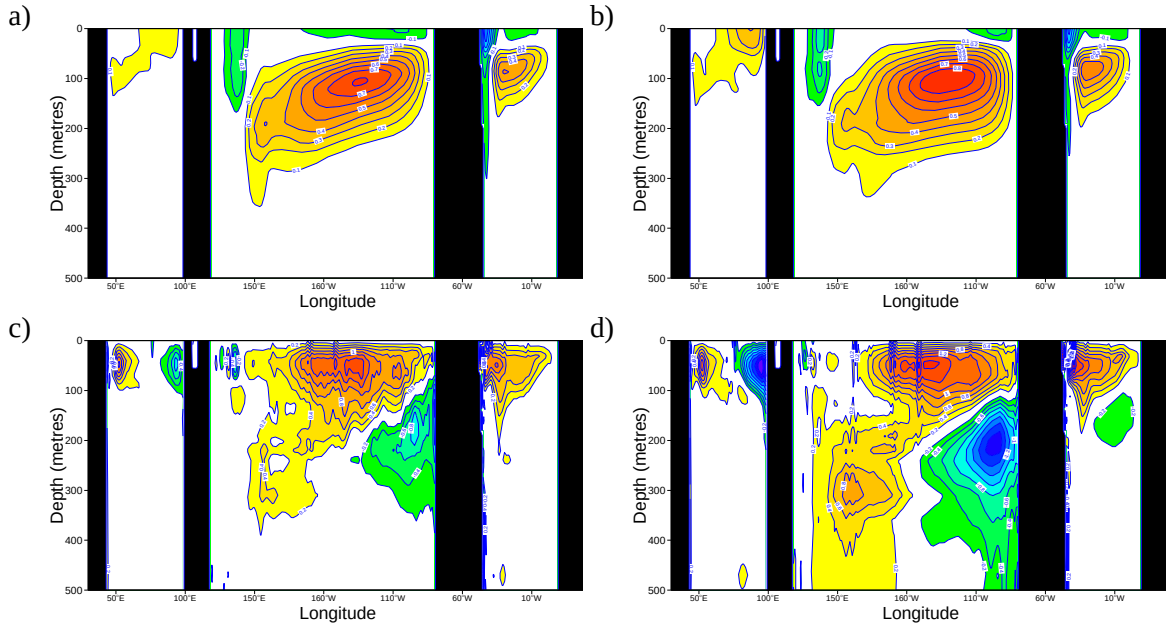


Figure 14: Equatorial section of the mean zonal (upper row) and vertical (lower row) velocity for S3 (left) and S2 (right). Contour is 1m/s for the zonal velocity (a,b) and 0.2 m/day for the vertical velocity (c,d). The reference level is blanked, and therefore, the two consecutive contours below and above the reference level are 2 contours apart. Values above (below) the reference level are in yellow/red (green/blue). The averaging period is 1987-2001.

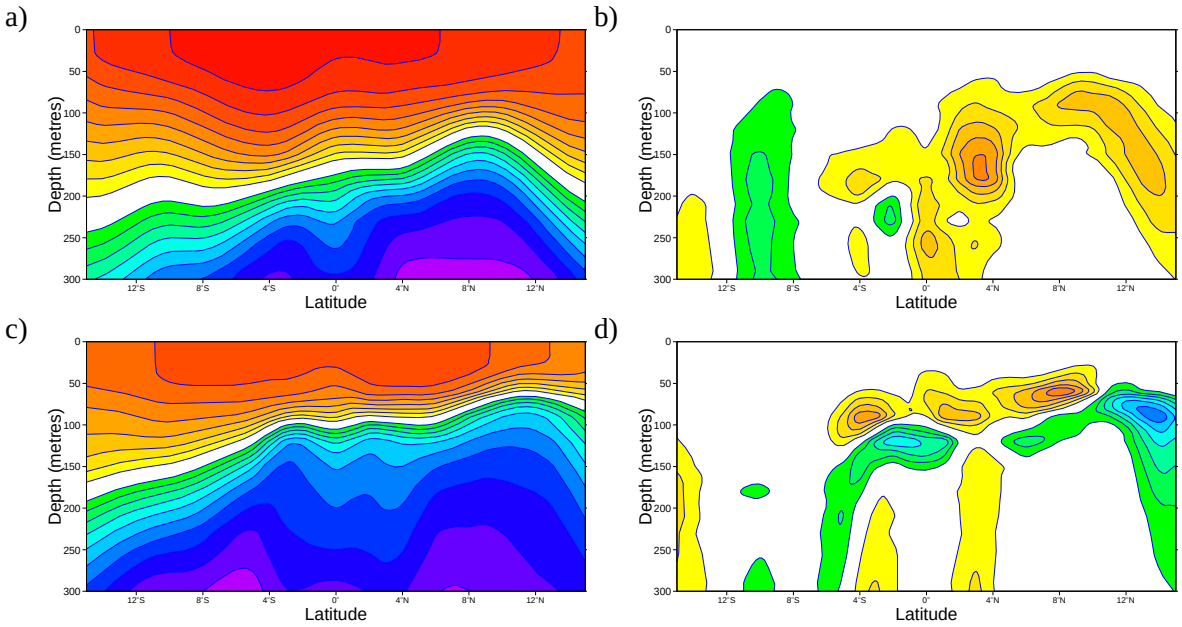


Figure 15: Meridional section of the mean temperature at 165E (upper row), and 30W (bottom) for S3 analysis (left) and differences S3 minus S2 (right). The averaging period is 1987-2001. The contour interval (reference level) is 1 °C (20 °C) for the full temperature fields and 0.5 °C (0 °C) for the difference plots. The reference level is blanked, and therefore, the two consecutive contours below and above the reference level are 2 contours apart. Values above (below) the reference level are in yellow/red (green/blue).

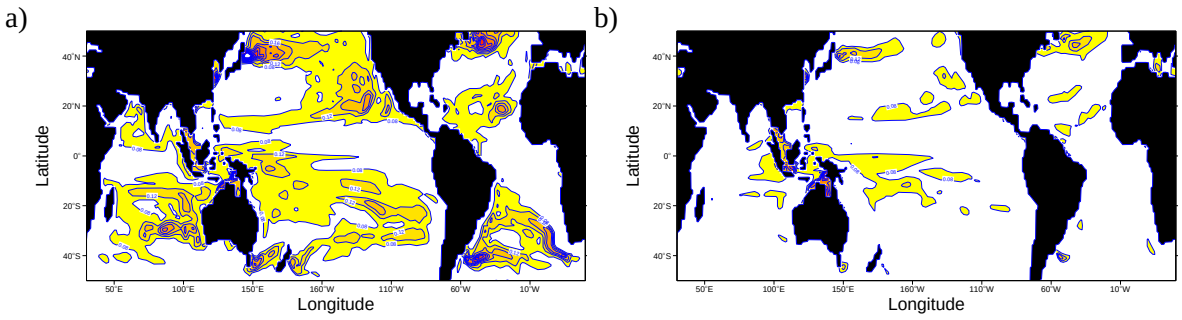


Figure 16: Standard deviation of the interannual anomalies of S300 for S3 analysis (left) and S2 analysis (right). The statistics refer to the period 1987-2001. Contour interval is 0.04psu.

AREA NAME	latitudes	longitudes
NINO3	5N - 5S	150W-90W
NINO4	5N - 5S	160E-150W
EQ3	5N - 5S	150E-170W
EQPAC	5N - 5S	130E-80W
EQATL	5N - 5S	70W-30E
EQIND	5N - 5S	40E-120E
NSTRPAC	30N-10N	105E-90W
SSTRPAC	10S-30S	105E-60W
NPAC	70N-30N	100E-100W
SPAC	30S-80S	130E-70W
NSTRATL	28N-5N	80W-20E
SSTRATL	5N-20S	60W-20E
NATL	70N-30N	70W-15E
SATL	30S-80S	70W-2OE

Table 1: Definition of area-averaged indices

5.2 Variability

The level of variability of the S3 and S2 analysis can be compared by considering the standard deviation of the interannual anomalies (i.e. with the seasonal cycle removed) during the period of study (1987-2001). The spatial structure and amplitude of the variability in SL and T300 are very similar in S2 and S3, and are not shown. This is not the case for the variability in the salinity content. Figure 16 shows horizontal maps of the interannual variability in S300 for S3 (left) and S2 (right). Variations in the salinity content of the upper ocean are much larger in S3 than in S2, where the salinity variability was probably underestimated due to the strong relaxation to climatology. Figure 16 shows that the largest S300 variability in S3 occurs mostly outside the equator (Indian ocean, north-subtropical Atlantic and extratropics), and on the western part of the equatorial Pacific warm pool. The S300 variability is likely to be associated to water mass displacement rather than being directely forced by fresh water fluxes at the surface: the location of S300 variability occurs in areas of large S300 gradients and it does not coincide with the location of the surface salinity variability (not shown), which is directely forced by PME fluxes.

The variability in T300 and SL may have similar spatial structure and amplitude, but there can be large differences regarding the time evolution of the interannual anomalies. As an example, figure 17 shows time series of

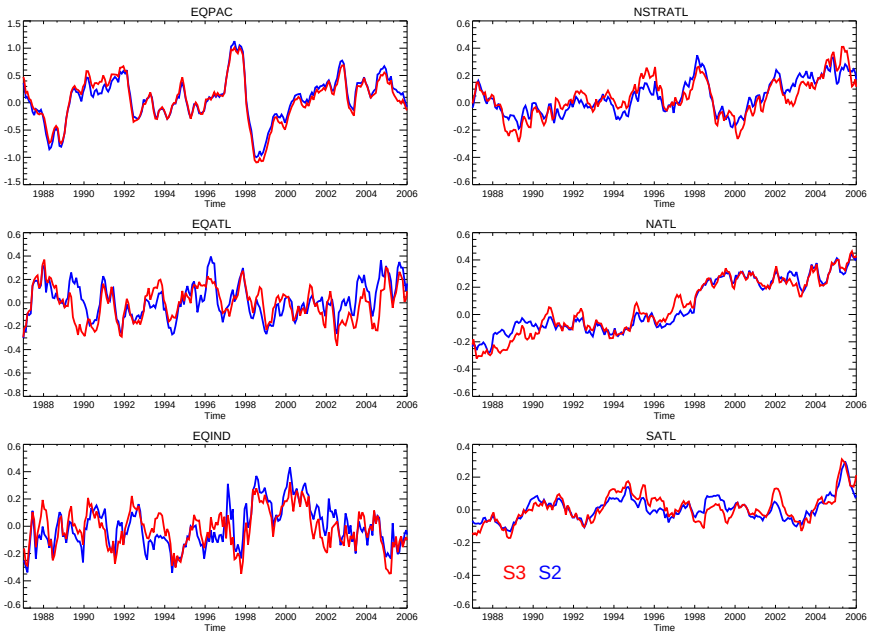


Figure 17: Time series of T300 anomalies from S3 (red) and S2 (blue) in different regions. The anomalies are with respect to the 1987-2001 climatology.

T300 for selected indices (defined in table 1) in blue for S2 and in red for S3. The evolution of T300 anomalies in the equatorial Pacific (EQPAC) is quite similar in S2 and S3, especially after 1993, when the TAO array was consolidated. Even here there is the overall impression that the anomalies in S3 are slighly warmer than in S2 before 1997/8, and sligthly colder after. In the equatorial Atlantic (EQATL), the interannual T300 anomalies are quite different in S3 and S2, especially in 1989/90, and in 1996. After 2002, the anomalies in S3 are systematically lower than in S2. In the equatorial Indian (EQIND) ocean, both S3 and S2 show a rapid warming in the evolution of T300 after 1998, although the details of the higher frequency variability is different in both systems. The variability of the Indian Dipole (not shown) is quite noisy in general, except for the big event of 1997/8. The cold phase of the Indian Dipole in 1996/7 is weaker in S3 than in S2. In the north subtropical Atlantic (NSTRATL), the low-frequency variability of T300 in S2 and S3 is broadly similar, but there are differences in the higher frequencies. In particular S3 shows more pronounced warming during 2005. It is not clear whether the differences in the T300 evolution between S3 and S2 are relevant for the seasonal prediction of SST. The variability of the North Atlantic (NATL) is dominated by an upward trend in both S2 and S3; in this region, differences between the two systems are largest at the begining of the record (1987-1990), during 1997/8 and 2003. The variability in the South Atlantic (SATL) shows also an upward trend, which is more pronounced from 2003 to 2005. Separate experimentation indicates that except for the southern oceans, it is in general not possible to attribute the changes in the T300 variability between S2 and S3 to different wind forcing (ERA40/OPS in S3 and ERA15/OPS in S2).

The differences in the evolution of SL in S3 and S2 are comparatively larger, as it can be seen in figure 18. This figure also shows as a reference the evolution of SL anomalies derived from the altimeter after 1993, which have been assimilated in S3 but not in S2. Not surprisingly, S3 sea level anomalies are closer to the altimeter anomalies than S2 but sometimes there are large departures, such as in the equatorial Indian ocean and in the north subtropical Atlantic after 2002. The large differences between S2 and altimeter sea level anomalies are somehow indicative of the information content of the altimeter data. Observing system experiments withdrawing the altimeter data from S3 indicate that altimeter data, in particular the assimilation of the global trend, is

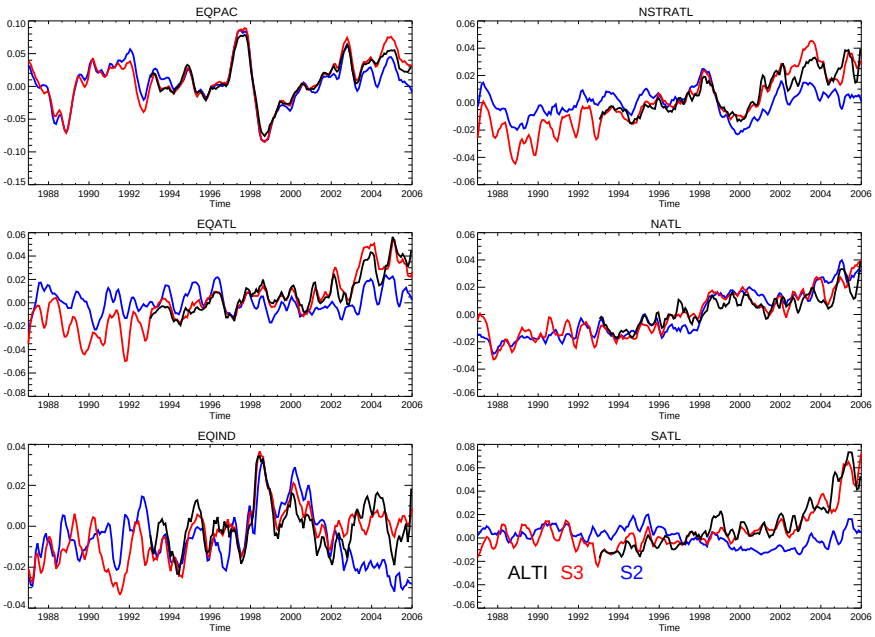


Figure 18: Time series of SL anomalies from S3 (red) and S2 (blue) in different regions. Also shown are the altimeter derived sea level anomalies. The anomalies are with respect to the 1993-2001 climatology, which is different for S3, S2 and altimeter..

instrumental for capturing the sea level trend in the North Atlantic. However, the better representation of the SL in the equatorial Indian ocean in S3 is not only due to the altimeter data, as it will be seen in section 6.

The largest differences between S3 and S2 are in the representation of the S300 anomalies, as can be seen in fig 19. In most of the regions considered, the S300 variability has quite large interannual or decadal components. Particularly noticeable is the increase in S300 after 2002, in the Atlantic, which is related, although not exclusively, to the increase in salinity observations in the ARGO era (see section 6).

5.3 Comparison with observations

Figure 20 shows the vertical profiles of the mean difference between the analysis and observations in the Western Pacific (EQ3) and in the Eastern Pacific (Niño 3) . Positive/negative differences are indicative of warm/cold bias. The bias in both the Eastern and Western Pacific in S2 has been significantly reduced in S3, where the east-west slope of the thermocline is better represented. This result confirms that the increased east-west zonal temperature gradient in S3 shown in fig 10 is an improvement with respect to S2.

A more stringent test of the quality of the analysis is given by the root mean square (RMS) of the difference between analysis and observations. Figure 21 shows the RMS for temperature in different areas of the tropical oceans. In the tropics, the RMS scores of temperature for S3 are systematically better than for S2. The same results hold for salinity, shown in figure 22. In S3, the salinity errors are specially reduced in the Western Pacific (region EQ3). This could be the effect of assimilating salinity data, which is especially beneficial in the Western Pacific (Vidard *et al.*, 2007, in preparation), although during the period for these statistics the salinity data were quite sparse. We believe that the reduction of error in S3 in the tropics, both in temperature and salinity, is probably related to the treatment of the bias in the system, whereby applying the correction in the pressure gradient leads to a more balanced system, eliminating spurious vertical circulations and convection.

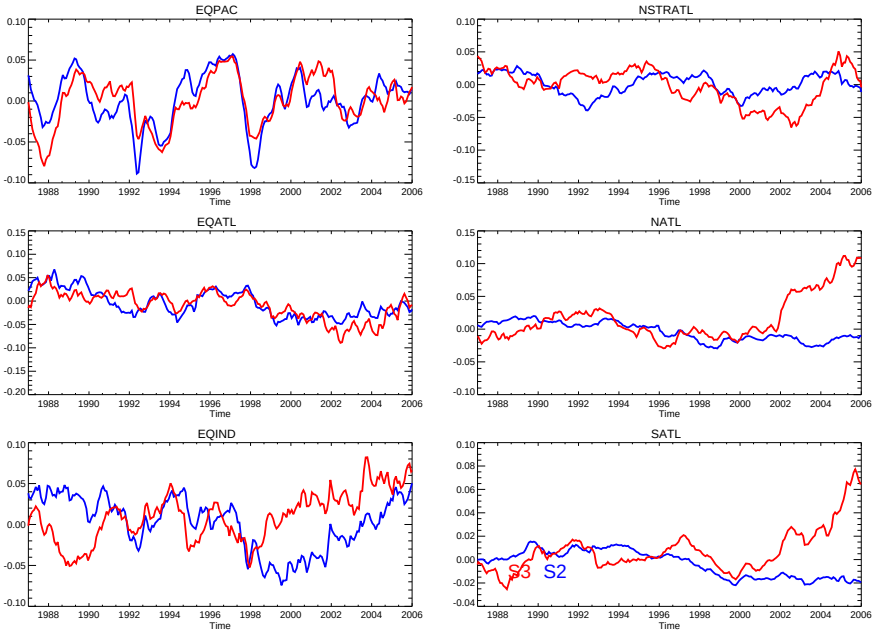


Figure 19: Time series of S300 anomalies from S3 (red) and S2 (blue) in different regions. The anomalies are with respect to their own 1987-2001 climatology.

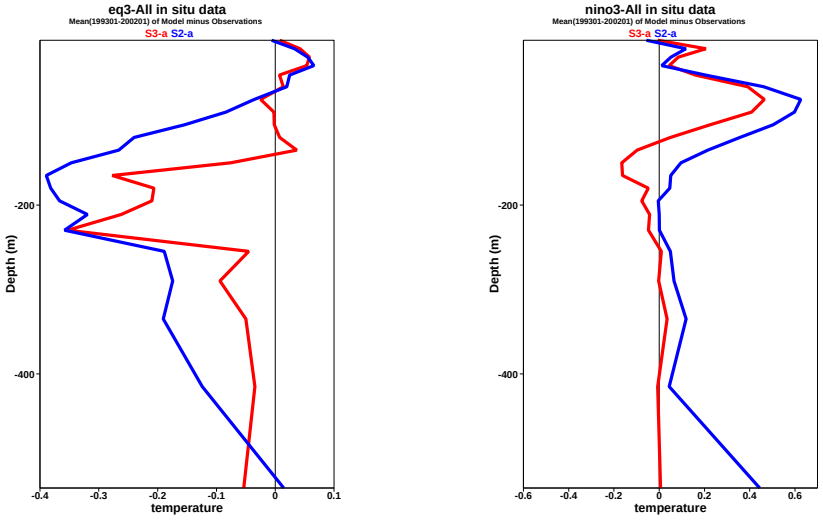


Figure 20: Vertical profiles of the mean difference between analysis and observations in the western (left) and eastern (right) Pacific, for S3 (red) and S2 (blue). Units are °C. In both regions the bias in S3 is smaller than in S2.

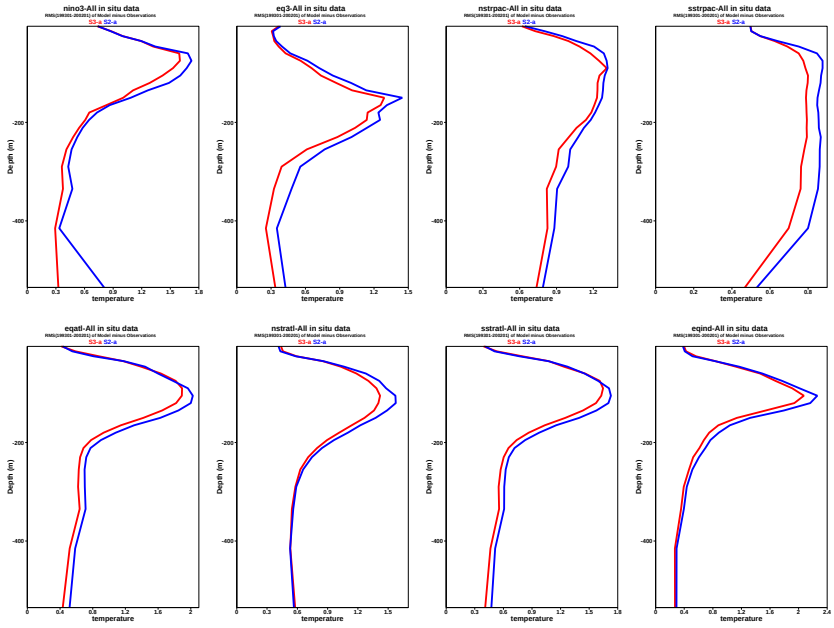


Figure 21: RMS difference between temperature analysis and observations in different regions:S2 (blue) and S3 (red)

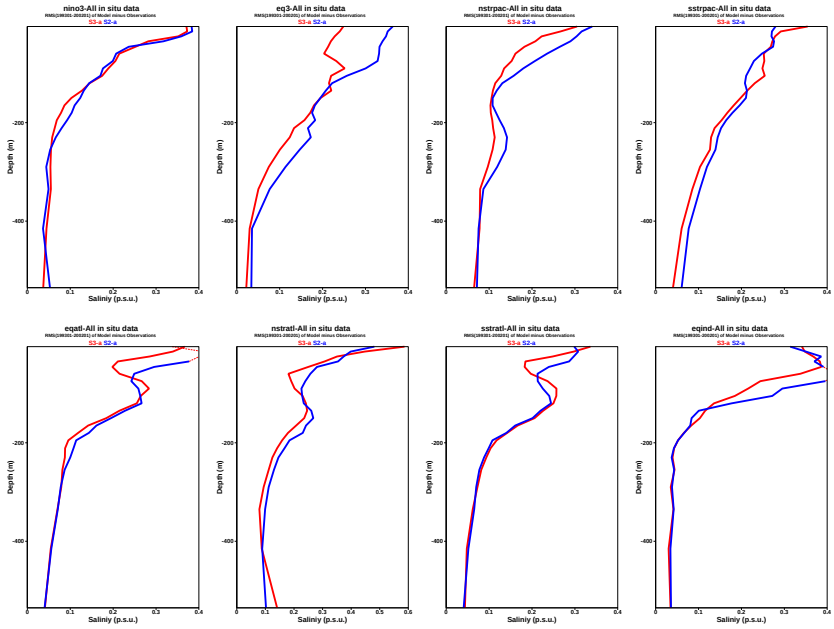


Figure 22: RMS difference between salinity analyses and observations in different regions: S2 (blue) and S3 (red). Units are psu.

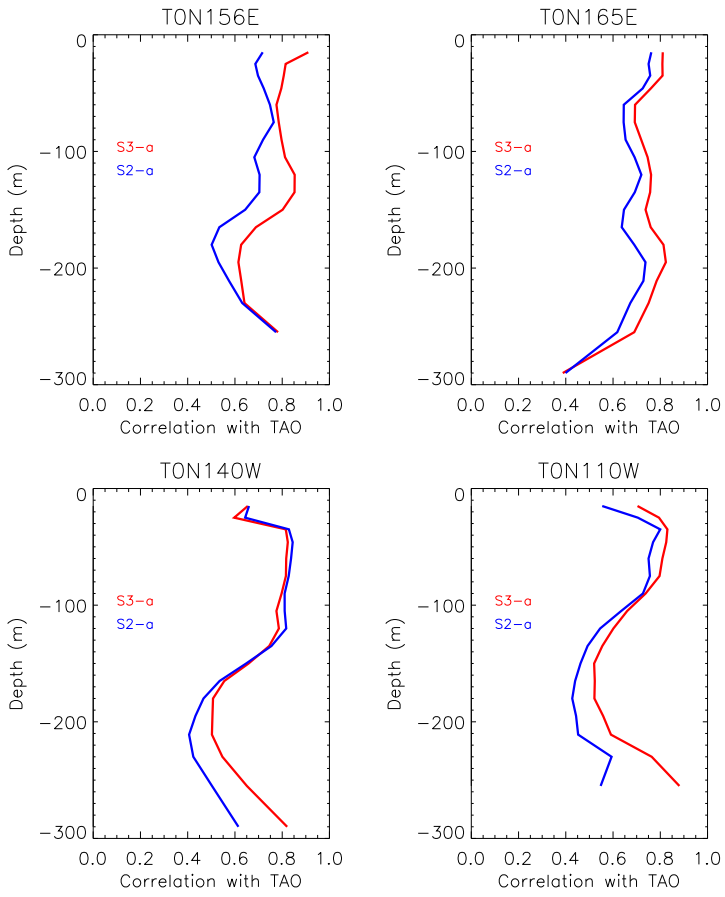


Figure 23: Correlation with TAO currents:S2 (blue) and S3 (red)

The velocity measurements from the TAO moorings provide an independent data set for the validation of ocean analyses, since the currents have not been assimilated in either S2 or S3. Figure 23 shows the vertical profiles of the time correlation of the currents in S2 (blue) and S3 (red) with the TAO currents at different mooring locations. The better correlation shown by S3 indicates that not only is the density field better constrained by observations in S3, but it is also more dynamically consistent.

Outside the tropics, the performance of S3 is not better than S2. The mean bias (analysis minus observations) and rms errors are shown in fig 24 for temperature and in fig 25 for salinity. Both the bias and the rms error are worse for S3 than for S2 in the North and South Pacific, and North and South Atlantic. The worse scores of S3 are probably related to the weaker subsurface relaxation in both temperature and salinity (10 years in S3 versus 1 year in S2). The assimilation scheme used may be better suited for the tropics: for instance the balance relationship for T and S is only fully applied equatorward of 30 degrees, and is not applied at all polewards of 60 degrees. Outside the tropics, the lack of adequate balance relationships between temperature and salinity could be a problem when the historical observations consist mainly of temperature data. There are very few ARGO profiles (with temperature and salinity) in the period used for these statistics. The impact of the ARGO data in the extratopics in the last few years needs to be investigated further.

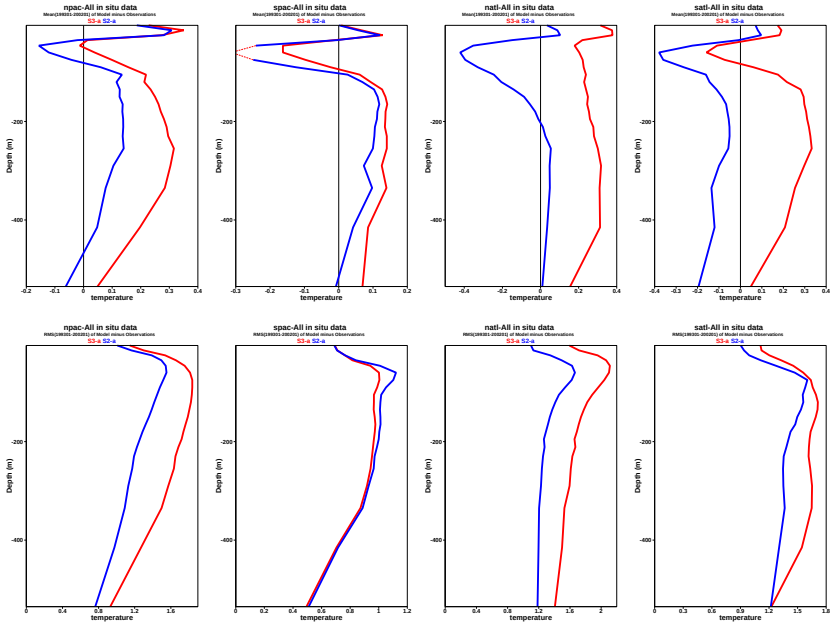


Figure 24: Mean (top) and rms (bottom) difference between temperature analyses and observations in extra-tropical regions. S2 is shown in blue and S3 in red. Units are °C

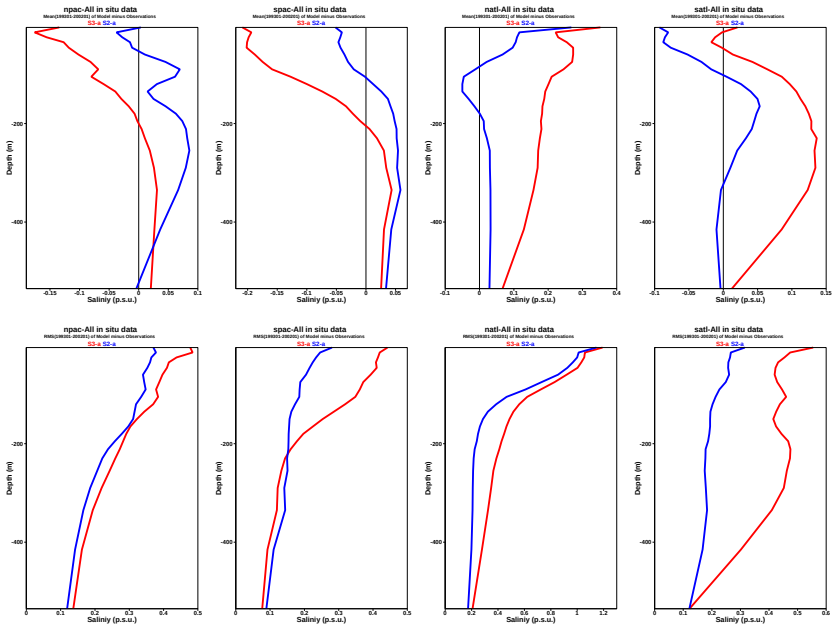


Figure 25: Mean (top) and rms (bottom) difference between salinity analyses and observations in extra-tropical regions. S2 is shown in blue and S3 in red. Units are psu.

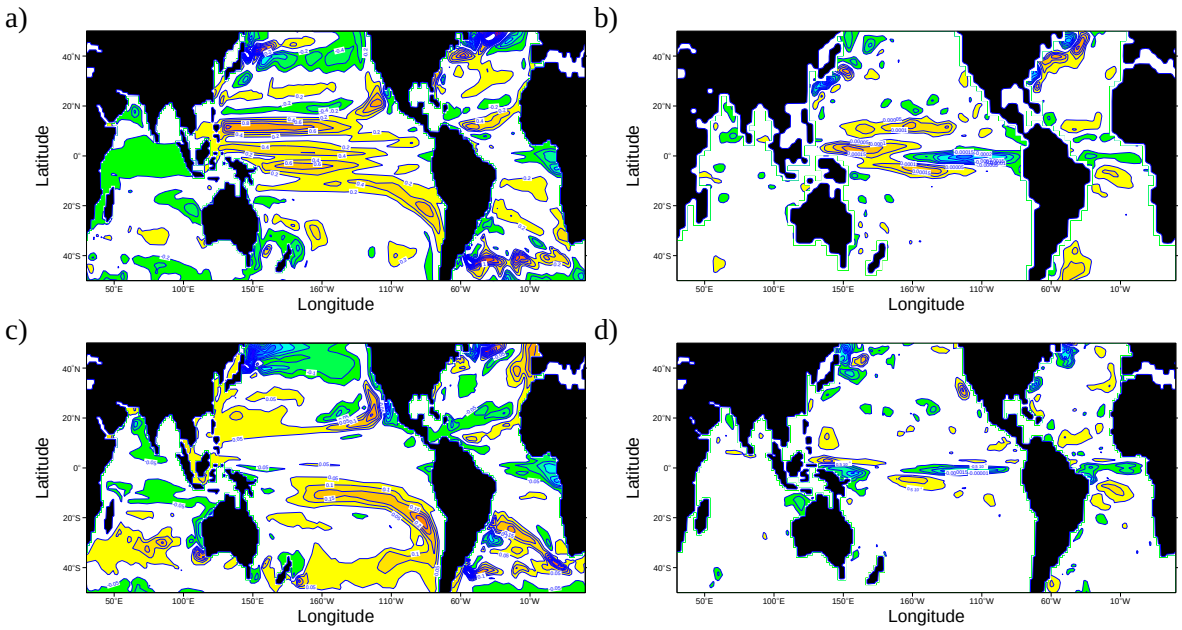


Figure 26: Impact of data assimilation: the left column shows the mean S3 minus S3_nodata differences for T300 (a) and S300 (c). The right column shows the average assimilation increment in the upper 300 meters for temperature (b) and salinity (d). The statistics are for the period 1987-2001. Contour interval is 0.2^oC *for T300 and 0.04psu for T300,* $5.10^{-5}\ ^\circ C/hr$ *for the temperature increment and* $5.10^{-6} psu/hr$ *for the salinity increment.*

6 Assessing the impact of data assimilation

In this section we assess the impact of data assimilation by comparing the S3 analyses with equivalent ocean analyses in which no data have been assimilated. Everything else (spin up, forcing fields, subsurface relaxation, prescribed additive bias correction and relaxation to SST) remains the same. In what follows we refer to this latter experiment as S3_nodata. The impact of data assimilation is evaluated from the point of view of state estimation (e.g. how data assimilation affects the representation of the ocean mean state and variability) and from the point of view of initialization of seasonal forecasts.

6.1 Impact on the ocean mean state

Data assimilation has a significant impact on the mean state and variability of the upper ocean. The left column of figure 26 shows the impact of data assimilation on the mean T300 (top) and S300 (bottom), as measured by the difference between S3 and S3_nodata. In the Pacific, the data assimilation increases the equatorial ocean heat content (fig 26a,c). The heat gain has a banded latitudinal structure, with maxima both sides of the Equator (at around 2-4N/2-4S) and a larger maximum in the area of the NEC (12/15 N). There is also a slight increase in the east-west slope of the thermal gradient. It appears as if the data assimilation is correcting for two kinds of errors in the equatorial Pacific: firstly it corrects the deficit of equatorial heat content in the S3_nodata, due either to an excessive export of heat via the Indonesian Throughflow, too large meridional heat transport or an excessive vertical mixing; secondly, it corrects the slope of the thermocline, which is too flat in the S3_nodata experiment. In the Equatorial Atlantic, data assimilation also steepens the east-west zonal gradient, especially by cooling the eastern part of the basin. In the tropical Indian ocean, data assimilation decreases the heat and salinity content: the excess of heat/salt in the Indian ocean in the S3_nodata experiment is consistent with too

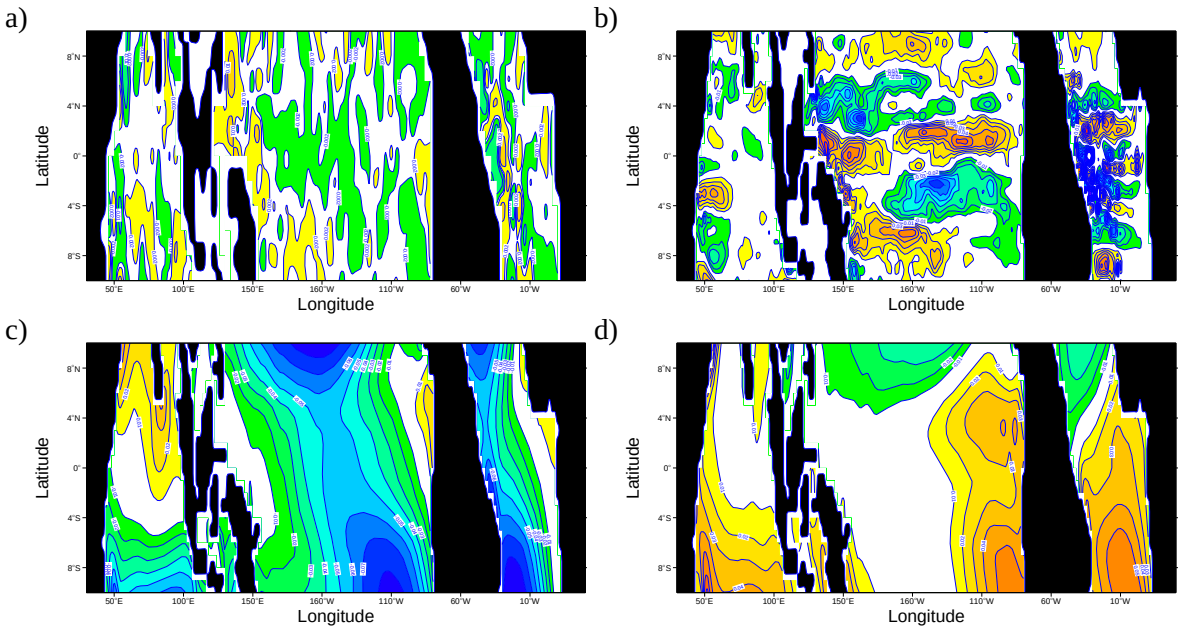


Figure 27: The upper row shows the applied bias in zonal and meridional pressure gradient integrated over the upper 300m (multiplied by -1). For reference, the bottom row shows the applied zonal and meridional component of the wind stress. The contour interval is 0.002N/m^2 for the zonal pressure gradient, and 0.01 N/m^2 for the rest of the figures. The statistics are for the period 1987-2001.

large a transport via the Indonesian Throughflow. In both the tropical Pacific and Atlantic, it appears as if the whole structure associated with the subtropical gyres (both north and south of the equator) is displaced equatorward. In the northern extratopics, the differences in T300 and S300 suggest that data assimilation affects the path of the Kuroshio and Gulf stream currents, which are not as zonal in S3 as in S3_nodata. In the North Pacific data assimilation has a large scale impact on S300, which is much reduced in S3 compared to the S3_nodata experiment. In the South Pacific, the salinity in S3 is higher than in S3_nodata.

The right column of fig 26 shows the average assimilation increment in the upper 300m in temperature (top) and salinity (bottom). By comparing the impact of data in the left column with the applied increment in the right it is obvious that the impact of data is not local. For instance, the largest temperature increment is applied in the eastern Equatorial Pacific (negative), where the differences in T300 between S3 and S3_nodata are not so obvious. Equally, the differences S3 minus S3_nodata in T300 in the Equatorial Atlantic do not require a particularly large negative temperature increment. There are two reasons for the "non-locality" of the increment: the ocean circulation, which can move the increment around, and the fact that between S3 and S3_no data there is another component to the assimilation increment, which is the correction of the bias in the pressure gradient, derived on-line from the accumulation of the assimilation increment.

Figures 27a,b show the average correction in zonal and meridional pressure gradient (reverse sign) resulting from the on-line bias estimation in S3. For consistency with the quantities in figure 26, the (reverse) pressure gradient has been integrated over the first 300m, although most of the contributions are from the top 100m, except for the areas of the western boundaries, where the pressure correction is significant at all depths. The values shown are equivalent to a correction in the zonal and meridional wind stress. For comparison, the zonal and meridional component of the wind stress are shown in the bottom panels. To make possible a detailed comparison, only the region within 10N/10S degrees of latitude is shown, which is one region where the bias correction is large. The bias correction in the pressure gradient is also large in the areas of the western boundary

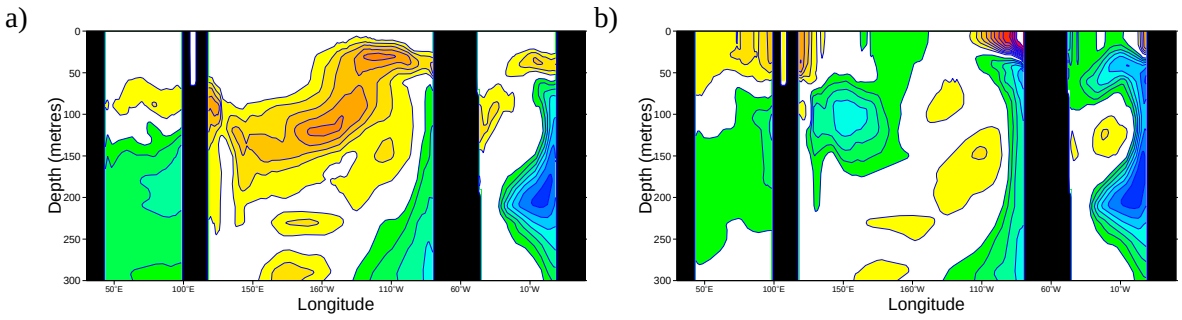


Figure 28: Equatorial zonal section of the mean difference S3 minus S3_nodata in temperature (left), and salinity (right). period is 1987-2001. Contour interval is 0.2^oC *for temperature and 0.05psu for salinity.*

currents, where the correction remains significant at all depths (not shown). In the Pacific, the equivalent correction to the zonal component is negative over most of the equator, and amounts to about 5-10% of the mean zonal wind stress. Its spatial structure is not very coherent though. By contrast, the correction to the meridional component exhibits quite large spatial correlation, and the values are comparable with or larger than the meridional component of the wind stress. On the equator, east of the dateline, the correction to the meridional component consists of an increased meridional divergence, which is probably responsible for the increased westward surface zonal velocity and increased equatorial upwelling. West of the date line the meridional correction is equivalent to an increased northward flow. North of the Equator, there is an increased meridional convergence in the area of the NECC, which is probably responsible for the increased heat content in that area. The structure of the pressure gradient correction in the Atlantic also shows an increased equatorial divergence, but it also contains corrections for the western boundary current.

Figure 28 shows the equatorial cross-section of the differences S3 minus S3_nodata in both temperature (a) and salinity (b). The vertical structure of the differences suggests that in the Pacific, the main effect of the assimilation is a net increase in the heat content by deepening the thermocline. Only in the far eastern Pacific does the data assimilation have a cooling effect below the thermocline, which suggests that it may be correcting problems with too much vertical mixing. The pattern of differences in the Equatorial Indian and Atlantic ocean suggest that the assimilation is correcting for too large vertical mixing at the base of the thermocline. This idea is consistent with the pattern of differences in salinity. As for fig 26, the pattern of the difference between the S3 and S3_nodata is not the same as the assimilation increment (shown in figure 12).

Putting together the information in figures 26, 12 and 27, it seems that the most likely mechanism by which data assimilation increases the equatorial heat content is by warming the water in the Western Pacific, which is then advected eastward by the equatorial undercurrent. The equatorial circulations in both the Pacific and Atlantic are more vigorous with data assimilation: i.e. stronger equatorial undercurrents and increased divergence in the eastern part of the basins (not shown).

The meridional structure of the temperature differences at 165E and 30W between S3 and S3_nodata can be seen in figure 29a and b respectively. At 165E the heat gain observed in fig 26 (at around $2-4^o$ both sides of the Equator, and at around 12^o N) occurs at thermocline depth. In the Atlantic at 30W, there is mainly a tightening of the thermocline.

6.2 Variability

Data assimilation has also a significant impact on the spatial structure, amplitude and phase of the interannual variability. Figure 30 shows the standard deviation of the interannual anomalies of S3(left) and S3_nodata(right)

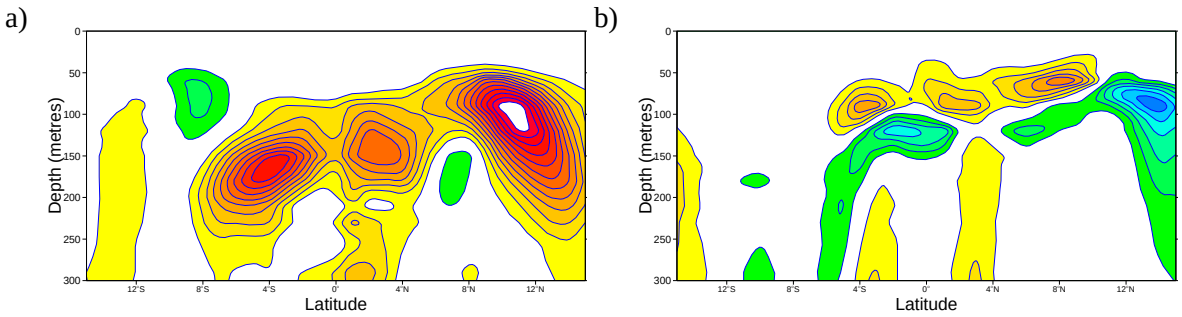


Figure 29: Meridional section mean difference S3 minus S3_nodata temperature anomalies at 165E (left), and 30W (right) during the period 1987-2001. Contour interval is 0.2 °*C.*

in SL, T300 and S300 (upper, middle and bottom row respectively). In the tropical Pacific, the variability in SL and T300 is dominated by ENSO, with a maximum in the Eastern Equatorial Pacific and 2 maxima in the western Pacific off the Equator. In S3, the variability in the Eastern Equatorial Pacific is intensified and shifted towards the coast. In the experiment S3_nodata the maximum in the Western Pacific north of the Equator is split in two, with the upper half extending across the whole basin. With data assimilation (S3) there is a single maximum in the north-western Pacific, which is confined to the western coast. The T300 variability in the north subtropical Atlantic is larger with data assimilation, and so is the variability in mid latitudes. In the Indian ocean, data assimilation reduces the interannual variability. The variability in S300, although broadly similar in S3 and S3_nodata, is larger in S3 than in S3_nodata.

Figure 31 shows equatorial sections of the interannual variability in temperature (upper row) and salinity (lower row), for S3 (left) and S3_nodata (right). The variability in temperature is located around the thermocline, which is deeper in S3 than in S3_nodata. The maximum in the eastern Pacific at thermocline depth is larger in S3. In the Atlantic, the variability in S3 is also larger. The level of salinity variability in S3 is similar to that in S3_no data, but slightly reduced. The variability is mainly concentrated in the mixed layer, above 100m.

Figure 32 shows the standard deviation of the interannual temperature anomalies for S3 (left) and S3_nodata (right) at 165E and 30W. Again, the spatial structures are very similar for both analyses, but S3 shows larger values. In the north-subtropical Atlantic the level of variability in S3 is considerably larger than in S3_nodata.

The interest in the north-subtropical Atlantic region has increased in the last years, especially after the successful seasonal predictions of the cold winter of 2005/6 by the UKMO, which were largely based on the SST of this region. Rodwell *et al.*, 1999 claim that SST anomalies during spring can be used as a predictor of the winter NAO, having the potential to influence the European temperatures the following winter. They hypothesized that the spring SST anomalies remain in the subsurface during the summer, underneath the summer mixed layer. When the mixed layer deepens during autumn and winter, the subsurface temperature anomalies can reach the surface again, and the resulting SST anomalies have an impact on the seasonal forecasts for the European winters by forcing the North Atlantic Oscillation (NAO). Although the precise mechanisms by which the SST in this area can influence the European winters is still subject of debate, the ”re-emergence” of the subsurface anomalies seems to be less controversial. Data assimilation has a significant impact on the amplitude of the interannual variability in this area, and this was the case for the 2005/6 event. Figure 33 shows the meridional sections at 30W of the temperature anomalies for August 2005 (source: *http://www.ecmwf.int/products/forecasts/d/charts.*) for the analysis without data assimilation (right) and for the S3 analysis (left). The size of the temperature anomaly is significantly larger in the assimilation case.

The time evolution of T300 for S3 and S3_nodata are shown in figure 34 for various regions defined in table 1. The differences between S3 and S3_nodata are approximately the same size as between S3 and S2 (shown

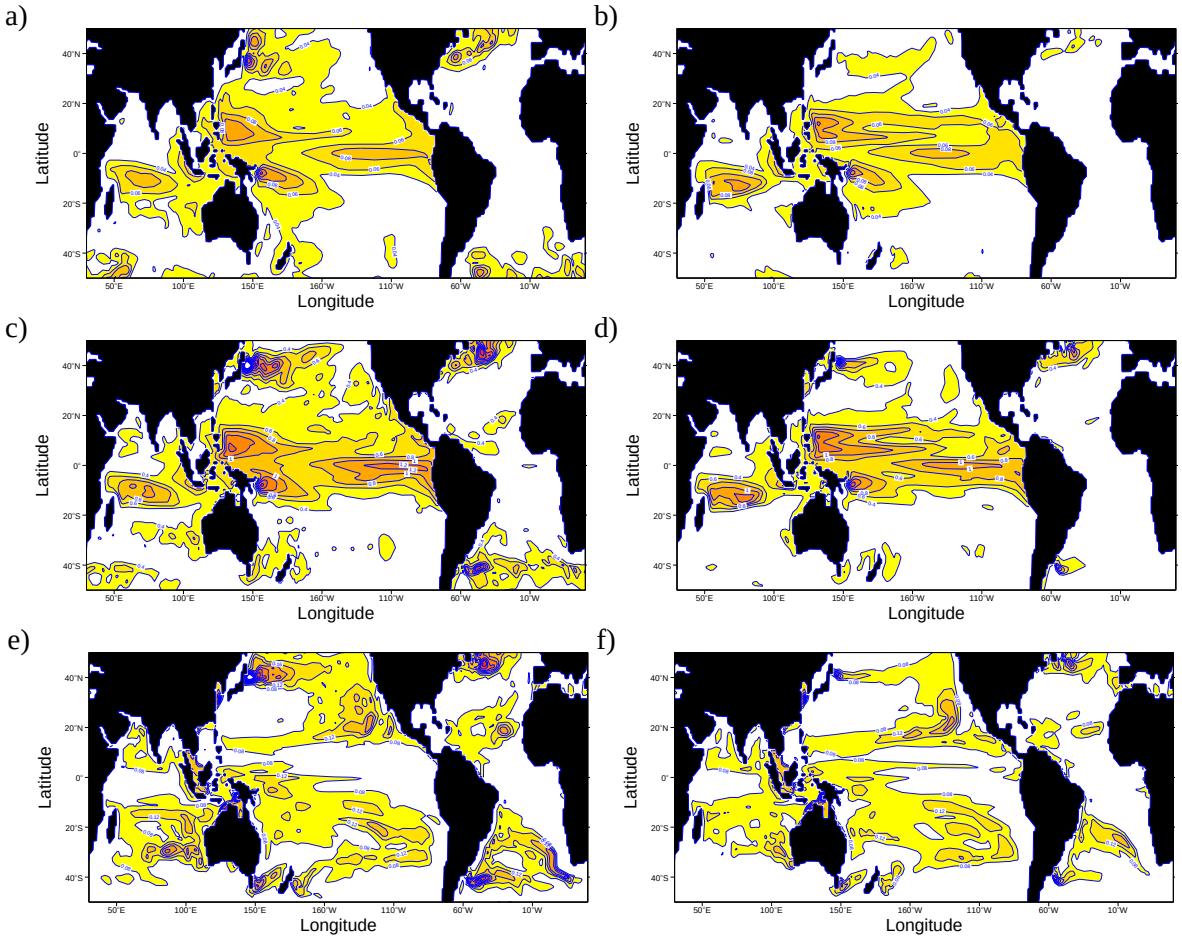


Figure 30: Standard deviation of the interannual anomalies in sea level (upper row), T300 (middle) and S300 (bottom) for S3 analysis (left) and S3_nodata (right). The statistics refer to the period 1987-2001. Contour interval is 0.02m for SL, 0.2°C for T300 and 0.04psu for S300.

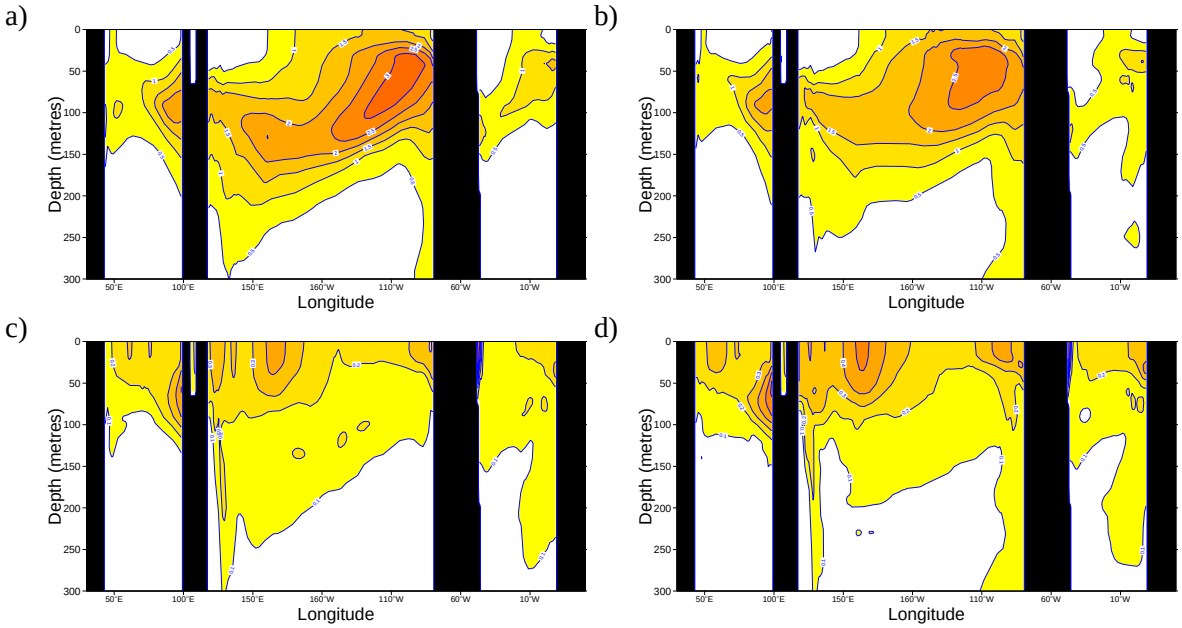


Figure 31: Equatorial zonal section of the standard deviation of the interannual anomalies of temperature (upper row), and salinity (bottom row) for S3 (left) and S3_nodata (right), for the period 1987-2001. Contour interval is 0.5 °*C for temperature and 0.1psu for salinity.*

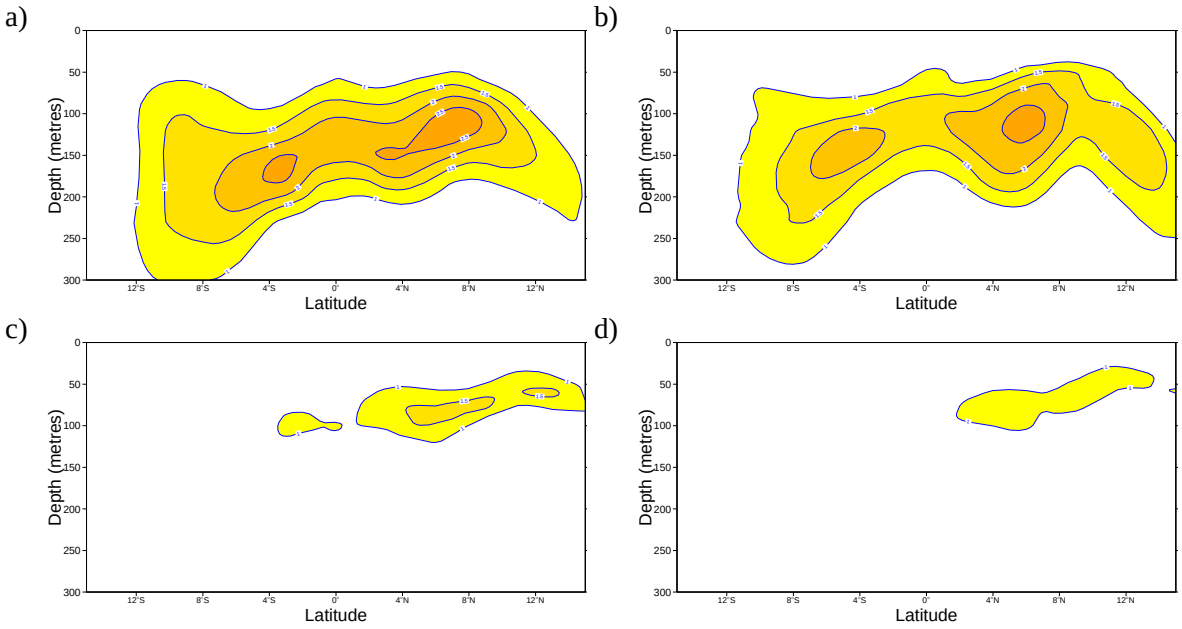


Figure 32: Meridional section of the standard deviation of the temperature anomalies at 165E (upper row), and 30W (bottom row) for S3 (left) and S3_nodata (right) during the period 1987-2001. Contour interval 0.5 °*C.*

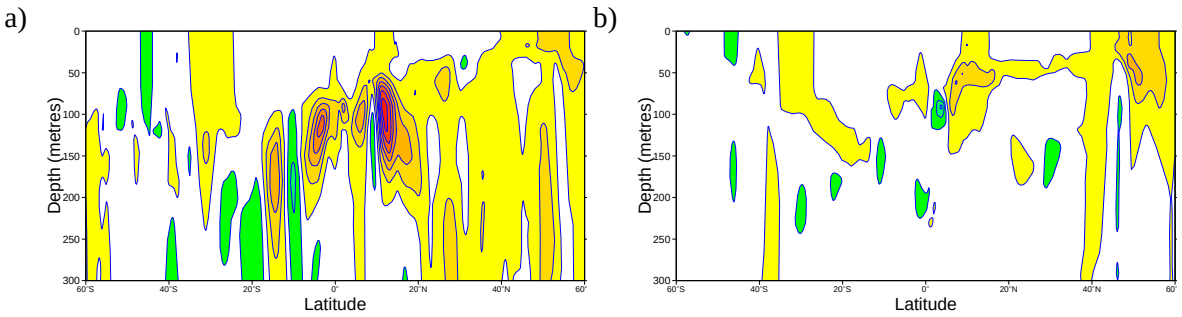


Figure 33: Meridional section at 30W of the temperature anomaly in August 2005 as represented in S3 (left), and S3_nodata (right). The anomalies are computed relative to their respective 1981-2005 climatologies.

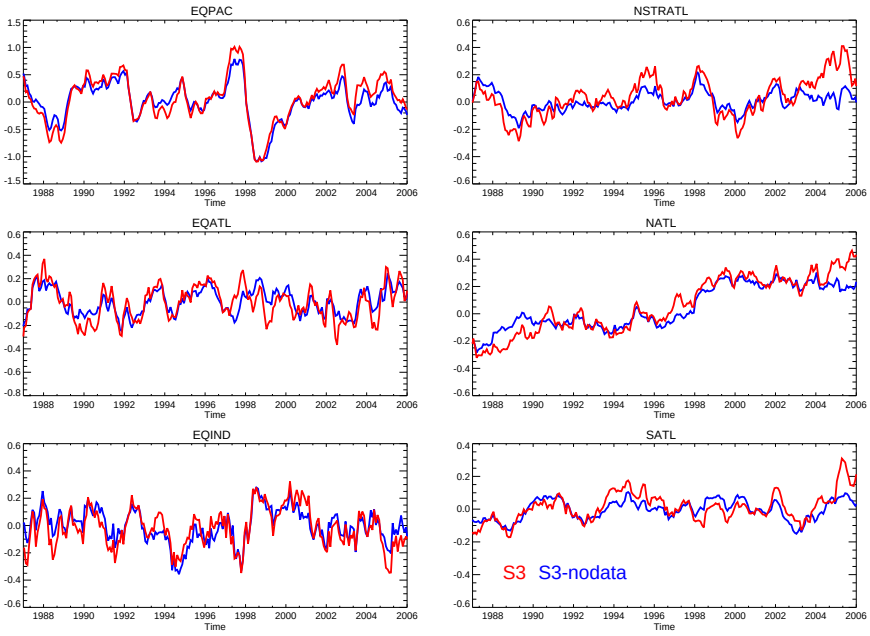


Figure 34: Time series of T300 anomalies from S3 (red) and S3_nodata (blue) in different regions. The anomalies are with respect to their 1987-2001 climatologies.

in figures 17). In the Equatorial Pacific, the amplitude of the interannual variability is slightly larger with data assimilation. The assimilation of data has had a large impact on the Atlantic T300 anomalies in the later years, which may be related to the increase of ARGO data.

Figure 35 shows the time evolution of the sea level anomalies for S3, S3_nodata and altimeter. As expected, S3 is closer to the altimeter product than S3_nodata. The trends in SL are more pronounced in the Atlantic, (equatorial, tropical, north and south), and are better captured by assimilating data. Observing system experiments show that this signal is mainly provided by the assimilation of altimeter data, and in particular by the assimilation of the global sea level trend (not shown, but see Balmaseda *et al.*c, 2007). The difference between S3 and S2 in the equatorial Indian ocean after 2002 discussed in relation to fig 18, seems to be independent of the assimilation of altimeter data: the SL in the experiment S3_nodata shows a close agreement with S3 and the altimeter.

The evolution of S300 anomalies (fig 36), shows large differences between S3 and S3_nodata. The Equatorial Pacific shows large interannual variability in the salinity content in both S3 and S3_nodata, which is broadly

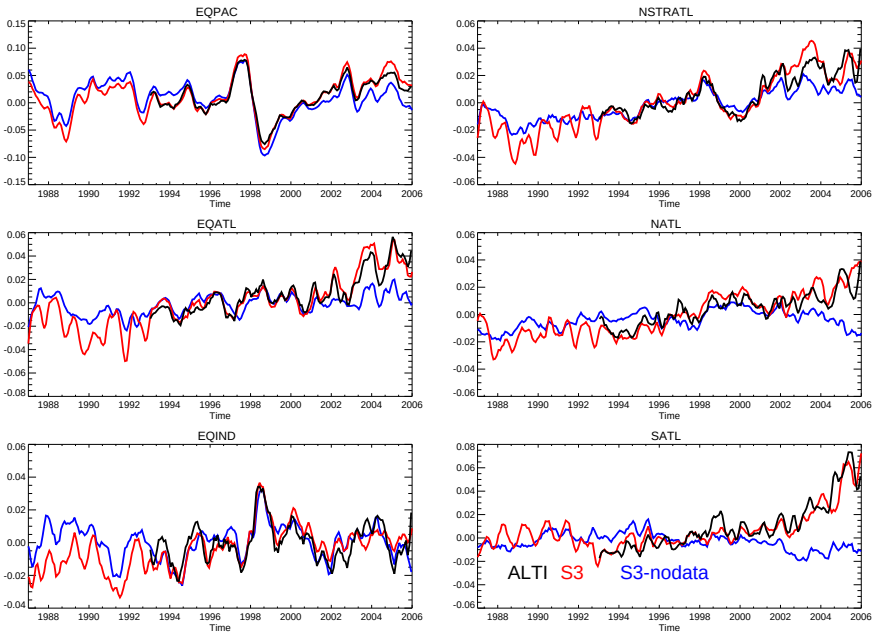


Figure 35: Time series of SL anomalies from S3 (red) and S3_nodata (blue) in different regions. Also shown are the altimeter-derived sea level anomalies (black). The anomalies are relative to their 1993-2001 climatologies.

in phase-opposition with T300 variability: pronounced minima during warm years and smooth maxima during the intervening period. The S300 variability in the Indian Ocean in S3 and S3_nodata is roughly consistent until 2002. After 2002, S3 shows an increased trend in S300 while S3_nodata shows a drop in S300. Results from observing system experiments (OSE) indicate that this difference between S3 and S3_nodata during this period is due to ARGO data (Balmaseda *et al.*, 2007b).

The impact of assimilation in S300 post 2002 is most pronounced in the Atlantic: in the north-subtropical Atlantic (NSTRATL), experiment S3_nodata shows a decreasing trend in S300, while S3 exhibits a marked maximum, peaking around 2005. In the North Atlantic (NATL) and South Atlantic, both S3 and S3_nodata show an increasing trend in S300, but in S3 the trend is larger, with a secondary rise in NATL and SATL after 2005. Results from OSEs also attribute this difference to the assimilation of salinity data comming from ARGO.

6.3 Comparison with observations

Data assimilation improves the fit to the observations, mainly by correcting the ocean mean state. Figure 37 shows the mean difference between the temperature analyses and observations from S3 and S3_nodata. The mean fit to the data for the salinity analysis is shown in figure 38. In general, the improvements in the mean state, in both temperature and salinity extend to all geographical areas and mostly all depths. The eastern Pacific (Nino3) at thermocline depth is an exception, and so is the North Atlantic from 50 to 150 m. In these areas the data assimilation changes the sign of the bias, with S3 being warmer than observations while S3_nodata is colder than observations. For the eastern Pacific, this is a well known problem, which was also present in S2, and which in S3 has been alleviated thanks to the treatment of the bias in pressure gradient, but has not been completely eradicated. In the North Atlantic the stratification in S3_nodata is far too strong in the upper 100m, and very weak below. Data assimilation erodes these vertical gradients, but apparently overdoes it, possibly by

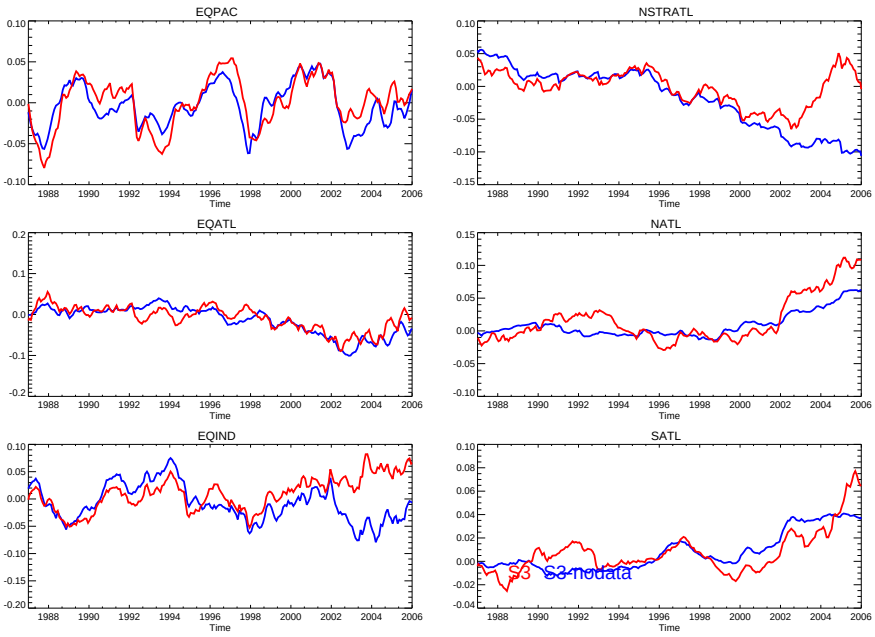


Figure 36: Time series of S300 anomalies from S3 (red) and S3_nodata (blue) in different regions. The anomalies are relative to the 1987-2001 climatology.

interacting with the vertical mixing.

In general, data assimilation reduces the bias in salinity and temperature in the extratropics. This confirms that the deterioration of S3 versus S2 in mid-latitudes seen in figure 20 is related to the weaker subsurface relaxation and not to the assimilation process.

Comparison with the OSCAR currents (Ocean Surface Current Analysis, Bonjean and Lagerloef, 2002) shows that data assimilation improves the representation of surface currents in the system, mainly in the tropical band. 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 end of 1992 up to near real-time and now cover the whole ocean from 60°S to 60°N. Comparisons between OSCAR and the data from the World Wide Drifter Buoy Deployment, and between OSCAR and the TAO/TRITON/PIRATA mooring data shows that OSCAR products are of good quality especially in the tropical areas (*http://www.esr.org/%7Ebonjean/oscar/global_validation/*)

Fig 39 shows the comparison with OSCAR zonal surface velocities: these maps show the correlation between the zonal component of the surface velocities from OSCAR monthly means and S3_nodata and S3 for the period 1993-2005. The colour scale is non linear and correlations below 0.5 are not plotted. OSCAR data are not available along the coasts. Fig 39 shows that data assimilation improves the representation of the surface currents consistently: within 10 degrees of the equator the improvement is mainly due to the assimilation of in situ temperature, while the assimilation of SLA improves the currents almost everywhere, except for some mid latitude areas (Gulf-Stream and Kuroshio regions, southern oceans), where the correlation with Oscar remains below 0.5.

OSCAR currents are not completely independent from sea level data since altimetry data are used in their production. However, we do not use SLA data directly but derive anomalies in T and S from them and therefore there is no guarantee that they should lead to improved velocity analyses. Therefore it is a good metric to assess

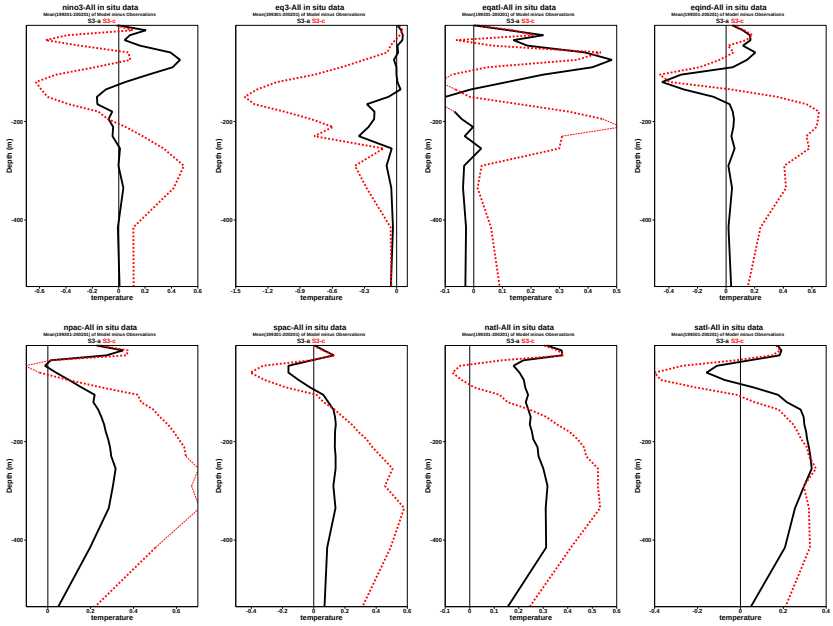


Figure 37: Mean difference between temperature analyses and observations in selected regions. S3 is shown in black and S3_nodata in red. Units are degrees Centigrade.

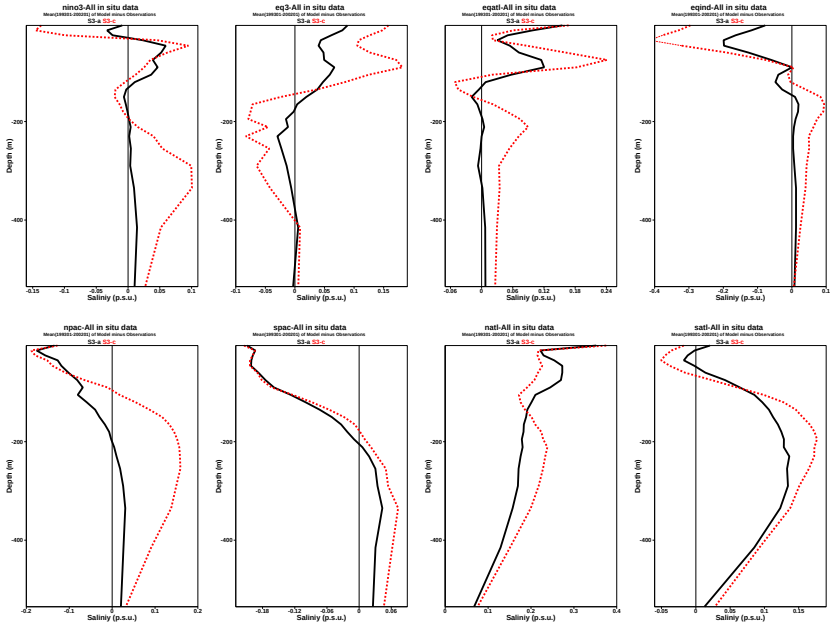


Figure 38: Mean difference between salinity analysis and observations in selected regions. S3 is shown in black and S3_nodata in red. Units are psu.

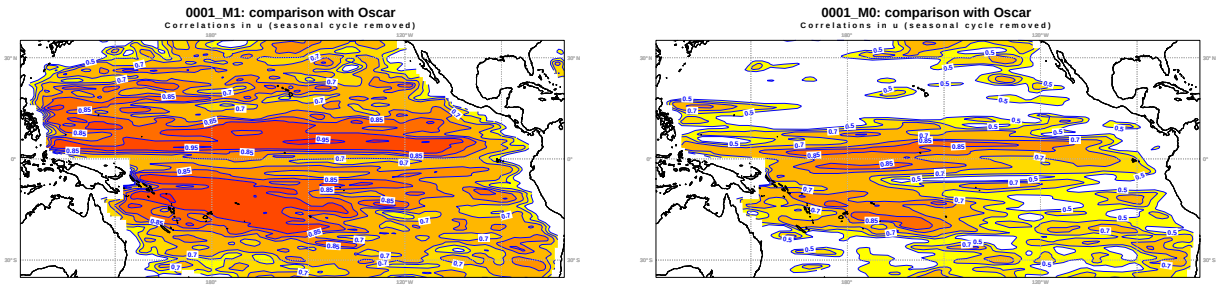


Figure 39: Correlations with Oscar zonal component of the surface currents for S3 (left), and S3_nodata (right)

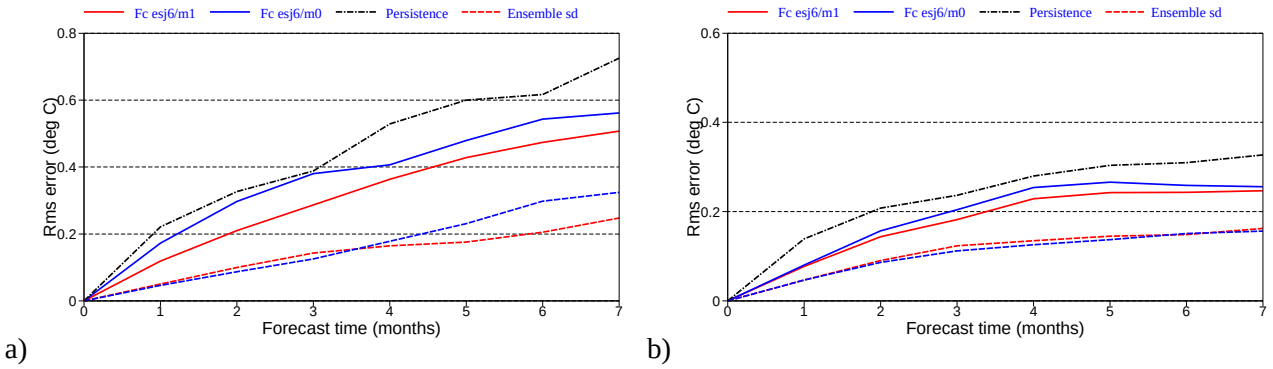


Figure 40: Plot of the rms error in the regions Niño4 (a) and NSTRATL (b) as a function of forecast lead time. Two experiments are shown: forecasts from the S3 analyses (red) and forecasts from the S3_nodata (blue) in which altimeter and subsurface data have been withheld. For reference, the RMS error of persistence is also plotted (black). The forecast spread is shown dashed for the two experiments.

the quality of our analysis.

6.4 Impact on the Forecast Skill

In order to assess the impact of the data assimilation on the forecast skill of the seasonal forecasts, two sets of coupled seasonal forecasts experiments have been conducted, using initial conditions from S3 and S3_nodata. Each set consists of 76 different ensemble forecasts, with initial conditions three months apart (January, April, July and October), spanning the period 1987-2005. For each date, an ensemble of 3 coupled forecasts (with perturbed initial conditions) is integrated up to 6-months lead time. The coupled model is that used by S3 seasonal forecasting system (Anderson *et al.*, 2007), which is based on the IFS cycle 31r1. The forecast SST anomalies are then computed with respect to the model climatology (which depends on the lead time), as described in Stockdale *et al.*, 1998. Fig 40 shows the RMS error in the forecast of SST anomalies as a fuction of lead time in two different areas: Niño4 (a) and NSTRATL (b). The results are indicative of the impact of using data assimilation to initialize seasonal forecasts. The red curves indicate the growth of error (solid) and the ensemble spread (dashed) for the experiment using initial conditions from S3. The blue curve is for the experiment which uses ocean initial conditions from S3_nodata. The reduction of error using data assimilation is significant, especially in the western-central equatorial Pacific (Niño 4, shown in panel a). The reduction of error in the tropical Atlantic is disappointingly small (not shown) though there is some improvement in forecasts for the north-subtropical Atlantic (b).

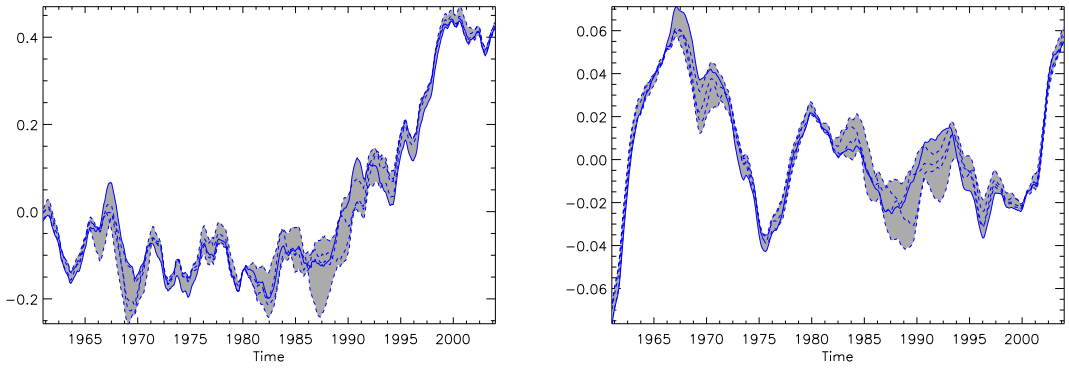


Figure 41: Time series of temperature (left) and salinity (right) anomalies averaged over the upper 300m in the North Atlantic (30N-60N) from the S3 5-member ensemble of ocean analysis. Both curves have been smoothed with a 12-month running mean. Units are °C for temperature and psu for salinity.

7 Some aspects of the climate variability as represented by S3

In this section we will focus only on two specific aspects of the climate variability as represented by the S3 analysis: the decadal variability in the North Atlantic and the trends in global mean sea level.

Figure 41 shows the time evolution of the 5 ensemble members over the North Atlantic (30N-60N): the temperature and salinity anomalies averaged over the upper 300m are shown in the left and right panels respectively. The spread of the ensemble can be taken as a measure of the uncertainty, which may be underestimated, since only wind error is accounted for, and there is no representation of errors in the heat or fresh water flux, in the model or the analysis

The variability in the upper ocean temperature is dominated by the upward trend, starting around the mid 80's. The presence of this trend should be taken into account when considering the reference climatology for the seasonal forecast products. Salinity variations occur mainly on decadal time scales, and they seem to be correlated with variations in the thermohaline circulation (THC). Figure 42 shows the time evolution of meridional transport in the North Atlantic at 30N in the upper 1000m (upper panel), and at 3000m-5000m depth range (lower panel). The curves have been smoothed with a 2-year running man. For comparison, the values given by Bryden *et al.* 2005 are shown by the stars. Although there is broad agreement with the estimates by Bryden *et al.*, the S3 ocean analysis shows that the decadal variability is large, and therefore sampling is an issue when drawing conclusions about the slowing down of the THC. The S3 ocean analysis will provide initial conditions for the decadal ENSEMBLES forecasts, where the capability of the coupled models to reproduce changes in the THC will be explored.

Figure 43 shows the simultaneous correlation of the meridional transport in the North Atlantic at 30N in the upper 1000m and the zonal component of the wind stress (a) and SST (b), at time scales longer than 4 years; the timeseries have been smoothed by applying a 4-year running mean using triangular window before computing the correlations. The figures shows that an increased (decreased) THC is associated with a weakening (strengthening) of the tradewinds and subtropical gyres. The increased (decreased) THC is also correlated with a cooling (warming) of SST in the Indian and Western Pacific on both sides of the equator. However, statistical correlations do not neccesary imply physical relationships. Figure 43 only indicates that both the Atlantic trade winds, the north/south edges of the warm pool and the Atlantic THC exhibit coherent variability in the long time scales.

Figure 44 shows the time evolution of the global sea level from altimeter data (black line) and from the S3

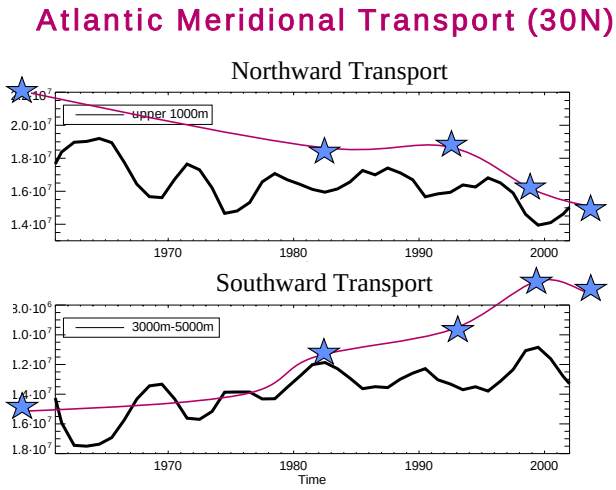


Figure 42: Meridional transport at 30N in the North Atlantic in the upper 1000m (upper figure) and at depths between 3000m -5000m (lower figure) from S3 ocean analysis. The stars show the values given by Bryden et al. *2005. Units are Sverdrup. The model values have smoothed with a 2 year running mean. The error bars in the values by Bryden* et al. *are +/ 6 Sv.*

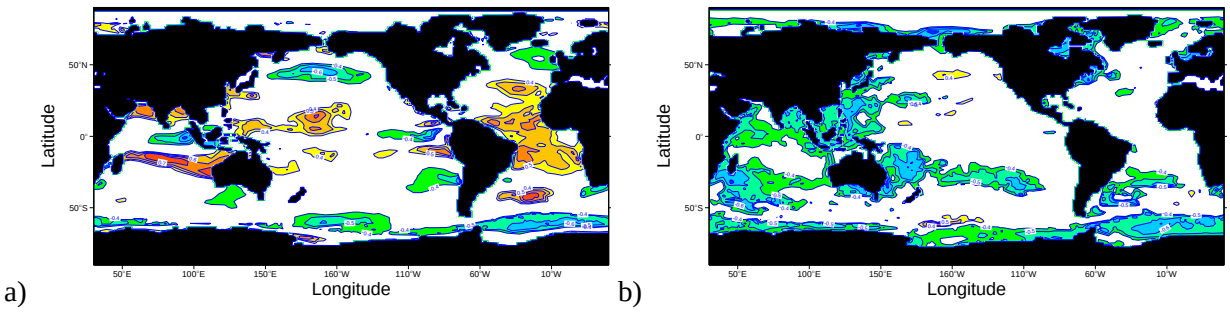


Figure 43: Correlation of the meridional transport in the North Atlantic at 30N in the upper 1000m and the zonal component of the wind stress (a) and SST (b), at time scales longer than 4 years. The figures shows that the increased (decreased) poleward transport is associated with a weakening (strengthening) of the tradewinds and subtropical gyres, and with a cooling/warming of SST in the Indian ocean. Only correlations with absolute value above 0.4 are shown. Contour interval is 0.1.

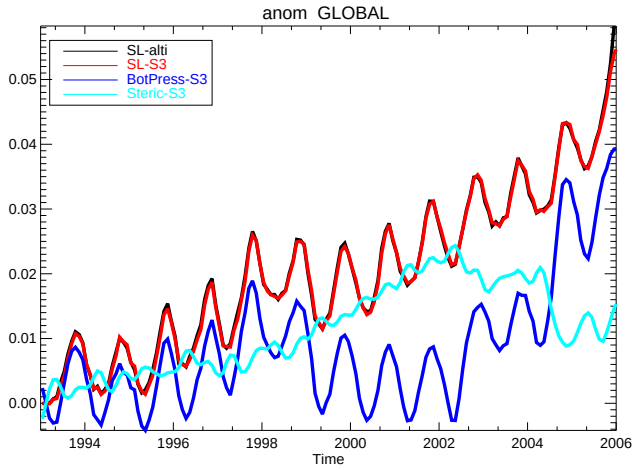


Figure 44: Global trends in sea level (from altimeter data in black, and from S3 in red), in S3 steric height (light blue), representing changes in volume due to thermal expansion, and equivalent bottom pressure (in dark blue). The global equivalent bottom pressure variations are only due to mass changes. Units are metres.

analysis. The similarity between the two curves is not surprising, since the global mean sea level trend has been assimilated. The figure also shows the evolution of the global steric height (SH), associated with changes in volume due to thermal expansion (light blue) and the time evolution of the global bottom pressure (BP), indicative of changes in the global mass. The figure shows that till 2002 the global trend in global mean sea level ($2mm/year$) is mainly due to thermal expansion, except for the seasonal variations due to mass changes (higher sea level during the northern hemisphere winter) and for the period 1997-1999, probably due to the increase of precipitation during El Niño event (the evaporation may be responsible for the delay). In mid 2002 and mid 2004 there are dramatic changes in the mass field. The change in 2002 may be due to the switch from ERA40 to operations, where there is a noticeable increase in the global P-E. As a consequence, the ocean analysis will attribute the changes in sea level to the mass field. The change in mid 2004 is less understood, and experiments are underway to determine if it is due to changes in the observing system (for example, the increased coverage of ARGO floats), or changes in the altimeter product due to the inclusion of JASON. Although the changes observed in the BP and SH changes may be an artifact of changes in the observing system, they are to a certain extent consistent with other reports in the scientific literature: on one hand, there has been reported an acceleration in the melting of the ice over Greenland (Rignold *et al.*, 2006) and Antartica (Velicogna and Wahr, 2006); on the other, Lynman *et al.*, 2006 have reported a decrease in the ocean heat content after 2003, which would imply a decrease of steric height.

As well as trends, figure 44 also shows a pronounced seasonal cycle in the sea level, with an amplitude of about 1cm; the minimum happens during northern hemisphere (NH) winter/spring, and the maximum occurs during NH summer/autumn. The seasonal cycle in the global sea level is dominated by changes in the ocean mass, rather than by temperature changes as traditionally thought. In fact, the amplitude of the seasonal cycle in global steric height is smaller (0.2 cm), it is out of phase with the seasonal cycle of global mean SL and exhibits a semiannual component. The seasonal cycle in global BP (mass of the ocean) is in phase with the seasonal cycle of the global SL. Variations of 1-1.2cm in total mass of the ocean can be explained by the seasonality of the hydrological cycle, which is the combined effect of the seasonality in the precipitation minus evaporation (P-E) and river discharge. The P-E used in S3, although based on ERA40, has been corrected to be consistent with other P-E climatologies, as described in Troccoli and Kallberg, 2004. In average P-E over the oceans is negative (-0.34 mm/day), with peak-to-peak variations analysis of about 0.2 mm/day, with a minimum during southern hemisphere (SH) summer (about $-0.4mm/day$), probably resulting from increased evaporation, and

maximum during the NH summer (about $-0.02 mm/day$).

The seasonality of the river discharge has been estimated from values of the 8 largest catchment areas (Artic rivers, Baltic Sea, Yangtze Kiang, Ganges/Bramaputra, Missisippi, Amazon, Congo and Parana.) using the values given in Hagemann *et al.*, 2005. The seasonal cycle is dominated by the Artic rivers which peaks in NH summer with an amplitude of $200000 m^3/s$, followed by the Ganges/Bramaputra basin, with an amplitude around $80000 m^3/s$, peaking during the NH summer months. The seasonal cycle of large rivers such as the Parana and Congo is by comparison small (one order of magnitude smaller). The estimated seasonal cycle in river discharge has an amplitude of approximately 0.1 mm/day, with maximum values in NH spring/summer, this is to say, roughly in phase with the seasonal cycle of P-E over the oceans. The combined seasonality in P-E and river discharge is consistent with seasonal variations in the mass of the ocean of about 1 cm.

8 Summary and conclusions

A new ocean analysis system (S3) is now operational at ECMWF. The S3 ocean analysis has introduced several improvements with respect to the previous system (S2); the 2D OI of S2 has been upgraded to 3D OI, the subsurface temperature-only assimilation of S2 has been upgraded to a scheme that can assimilate salinity from ARGO, CTD, and moorings as well as the time varying sea level from altimetry. One of the key problems in ocean data assimilation in general is that it mainly acts to correct bias related to the gradient and slope of the thermocline. To deal with this, the analysis of S3 has an explicit bias correction algorithm which allows a better representation of interannual variability than was possible in S2. An extensive ocean reanalysis has been carried out covering the period from 1959 until present. The same analysis system is used throughout. The extended analysis can be used for climate studies as well as providing initial conditions for seasonal forecasts. The S3 seasonal forecasting system will use the analyses from 1981 as initial conditions for the calibration hindcasts. ENSEMBLES will use the earlier analyses for further seasonal and decadal hindcasts.

There are substantial differences in the representation of the ocean mean state between S2 and S3. Overall, S3 outperfoms S2 in the tropical areas, where the fit to the subsurface temperature and salinity improves, in terms of both reducing the rms error and the bias. The correlation with the altimeter also improves, especially in the Atlantic ocean. The representation of the equatorial currents is also improved in S3 compared to S2. In the extratropics however, the fit to the data in S3 is not as good as in S2, probably because S2 was strongly constrained to climatology (1-year time scale), whilst in S3 the climatological constraint was much weaker (10 years time scale).

Data assimilation has a significant impact on the mean state and variability of the upper ocean heat content. In the Equatorial Pacific, it steepens the thermocline and increases the amplitude of the interannual variability. In the Indian ocean it sharpens the thermocline, making it shallower, and increases both the ENSO-related and Indian Dipole variability. The assimilation also corrects for a too difuse thermocline in the Atlantic. The amplitude of the Equatorial Atlantic interannual variability increases with assimilation, and the phase can be significantly modified in several cases (1998, for instance). The interannual variability in both northern and southern tropical Atlantic regions is increased, especially in the later part of the record (after 2000). The assimilation has a large impact on the salinity and sea level fields. Strong changes in the salinity field are associated with the introduction of the ARGO observing system, and are largest in the Atlantic and Indian Oceans. As expected, the fit to the data improves with data assimilation almost everywhere.

S3 ocean analysis has been designed to provide initial conditions to both seasonal forecasts and to monthly or medium range systems. The impact of assimilation on seasonal forecasts of SST is beneficial nearly everywhere, but especially in the west Pacific, where the RMS error is reduced by 20% in the first 3 months. The assimiliation also improves the seasonal forecast skill in the North Subtropical Atlantic.

The S3 historical ocean reanalysis is an important resource for climate studies. This report shows some examples of the climate variability applications: variability and trends in sea level, upper heat content, and meridional overturning circulation. More in depth studies are underway to determine the sensitivity of the S3 climate variability to changes in the observing system.

REFERENCES

Alves O., M. Balmaseda, D Anderson, T Stockdale, 2003, Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Q. J. R. Meteorol. Soc.*,130, Jan 2004, 647-668. Also ECMWF Technical Memorandum 369, available at www.ecmwf.int

Anderson, D. L. T., and M. Balmaseda, 2005: Overview of ocean models at ECMWF. ECMWF Seminar Proceedings. Seminar on Recent developments in numerical methods for atmospheric and ocean modelling, 6-10 September 2004, 103-111.

Anderson, D. L. T., T. Stockdale, M. Balmaseda, L. Ferranti, F. Vitart, P. Doblas-Reyes, R. Hagedorn, T. Jung, A. Vidard, A. Troccoli and T. Palmer, 2003: Comparison of the ECMWF seasonal forecast Systems 1 and 2, including the relative performance for the 1997/8 El Nino. ECMWF Technical Memorandum 404.

Anderson, D. L. T., T. Stockdale, M. Balmaseda, L. Ferranti, F. Vitart, F. Molteni, F. Doblas-Reyes, K. Mogensen and A. Vidard , 2006. Development of the ECMWF seasonal forecast System 3. ECMWF Technical Memorandum 503.

Balmaseda, M.A., 2004: Ocean data assimilation for seasonal forecasts. ECMWF Seminar Proceedings. Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003, 301-326.

Balmaseda, M.A., 2005: Ocean analysis at ECMWF: from real-time ocean initial conditions to historical ocean reanalysis. ECMWF Newsletter No. 105 Autumn 2005.

Balmaseda, M.A., D. Dee, A. Vidard and D.L.T. Anderson, 2007a: A Multivariate Treatment of Bias for Sequential Data Assimilation: Application to the Tropical Oceans. *Q. J. R. Meteorol. Soc.*, In press.

Balmaseda, M.A., D.L.T. Anderson, and A. Vidard 2007b: Impact of ARGO data in the ECMWF ocean analysis. In preparation.

Balmaseda, M.A., A. Vidard and D.L.T. Anderson, 2007c: Impact of Altimeter data in the ECMWF ocean analysis. In preparation.

Bell, M.J., M.J. Martin and N.K. Nichols, 2004. Assimilation of data into an ocean model with systematic errors near the equator. *Q. J. R. Meteorol. Soc.*, 130, 873-893.

Bjerknes J. 1964: Atlantic air-sea interaction. *Adv. Geophys.*, **10**, 1-82.

Bloom, S. C., Takacs, L. L., Da Silva, A. M. and Ledvina, D., 1996: Data assimilation using incremental analysis updates. Mon. Wea. Rev., 124, 1256-1271.

Burgers G., M.Balmaseda, F.Vossepoel, G.J.van Oldenborgh, P.J.van Leeuwen, 2002: Balanced ocean-data assimilation near the equator. *J Phys Oceanogr*, **32**, 2509-2519.

Bonjean F. and G.S.E. Lagerloef, 2002: Diagnostic Model and Analysis of the Surface Currents in the Tropical Pacific Ocean, Journal of Physical Oceanography, Vol. 32, No. 10, pages 2938-2954.

Bryden H.L., H.L. Longworth and C.A. Stuart, 2005: Slowing of the Atlantic Meridional Overturning Circulation at 25°N, *Nature*, 438, 655-657.

Church, J., and N.J. White, 2006: A 20th century acceleration in global sea-level rise. Geophys. Res. Letters, VOL. 33, L01602, doi:10.1029/2005GL024826.

Cooper, M.C. and K. Haines, 1996: Data assimilation with water property conservation, J. Geophys. Res 101, C1, 1059-1077

Dee, D. P.,2005. Bias and data assimilation. *Q. J. R. Meteorol. Soc.* ,**131,** 3323-3343.

Ducet, N., P.-Y. Le Traon, and G. Reverdin, 2000: Global high resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. J. Geophys. Res., 105, 19477-19498.

Drécourt, J.P., K. Haines, M. Martin, 2006: Use of bias aware data assimilation to improve the GOCINA mean dynamic topography. Proceedings of the GOCINA international workshop April 13-15, 2005, Luxembourg.

Haines K., J. Blower, J-P. Drecourt, C. Liu, A. Vidard, I. Astin, X. Zhou, 2006. Salinity Assimilation using S(T) relationships. Mon. Wea. Rev. Vol. 134, No. 3, pages 759–771.

Horel, J. D., and J. M. Wallace, 1981: Planetary-Scale Atmospheric Phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.*, **109**, 813-829.

Ingleby, B and M. Huddleston, 2006. Quality control of ocean temperature and salinity profiles - historical and real-time data. J. Mar. Sys., in press.

Le Traon, P.-Y., F. Nadal, and N. Ducet, 1998: An improved mapping method of multisatellite altimeter data. J. Atmos. Oceanic Technol., 15, 522-534.

Levitus S NODC (Levitus) World Ocean Atlas 1998 data provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/

Lyman J.M., J.K. Willis, G.C. Johnson, 2006: Recent Cooling of the Upper Ocean. Geophys. Res. Letters 33,L18604, doi:10.1029/2006GL0270033.

McPhaden, M. J., A. J. Busalacchi, R. Cheney, J.-R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P. Niiler, J. Picaut, R. W. Reynolds, N. Smith and K. Takeuchi, 1998: The Tropical Ocean-Global Atmosphere observing system: A decade of progress. Journal of Geophysical Research, 103, 14169-14240

van Oldenborgh G. J., M.A. Balmaseda, L. Ferranti, T.N. Stockdale, D.L.T. Anderson: 2005: Did the ECMWF Seasonal Forecast Model Outperform Statistical ENSO Forecast Models over the Last 15 Years? Journal of Climate, 18,No. 16, 3240-3249.

van Oldenborgh G. J., M.A. Balmaseda, L. Ferranti, T.N. Stockdale, D.L.T. Anderson: 2005. Evaluation of atmospheric fields from the ECMWF Seasonal Forecasts over a 15-year period. J. Clim., 18, No. 16, 3250-3269. See also corrections Vol 18,

Pacanowski R. and S G H Philander 1981. Parameterization of vertical mixing in numerical models of tropical oceans. *J. Phys. Oceanogr.*, **11**, 1443-1451.

Peters, H, Gregg, M C and Toole, J M, 1988: On the parameterization of equatorial turbulence. *J. Geophys. Res.*, **93**, 1199-1218.

Reynolds R., N Rayner, T Smith, D Stokes, W Wang 2002: An improved in situ and satellite SST analysis for climate. J Clim , 15, 1609-1625.

Rignot, E. and P. Kanagaratnam, 2006; Changes in the velocity structure of the Greenland ice sheet, Science, 311, 986-990.

Rio, M.-H. and F. Hernandez, 2004: A Mean Dynamic Topography computed over the world ocean from altimetry, in-situ measurements and a geoid model. Journal of Geophysical Research.Vol. 109, C12032

Roemmich, D., M. Morris, W.R. Young and J.R. Donguy, 1994: Fresh equatorial jets. J. Phys. Oceanogr., 24, 540-558.

Smith, Thomas M., Richard W. Reynolds, 1998: A High-Resolution Global Sea Surface Temperature Clima-

tology for the 1961-90 Base Period. *Journal of Climate* **11**, 3320-3323.

Smith N., J Blomley, and G Meyers (1991) A univariate statistical interpolation scheme for subsurface thermal analyses in the tropical oceans. *Prog in Oceanography*, **28**, 219-256.

Stockdale T. N., D. L. T. Anderson, J. O. S Alves, and M. A. Balmaseda, 1998. Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature*, **392**, 370-373.

Stockdale, T.N., M. A. Balmaseda and A. Vidard, 2006: Tropical Atlantic SST prediction with coupled ocean-atmosphere GCMs. *J. Climate*, **19**, No. 23, 6047-6061.

Tapley B.D., D.P. Chambers, S. Bettadpur and J.C. Ries, 2003: Large scale ocean circulation from the GRACE GGM01 Geoid. Geophys. Res. Letters 30.

Troccoli, A., and K. Haines, 1999: Use of the Temperature-Salinity relation in a data assimilation context, J. Atmos. Ocean Tech., 16, 2011-2025.

Troccoli A., M. Balmaseda, J. Schneider, J. Vialard and D. Anderson 2002, Salinity adjustments in the presence of temperature data assimilation. *Mon. Wea. Rev.*, **130**, 89- 102.

Troccoli, A. and Kallberg, P., 2004. Precipitation correction in the ERA-40 reanalysis, ERA-40 Project Report Series, 13.

Uppala, S., and coauthors, 2005. The ERA-40 Reanalysis. *Q. J. R. Meteorol. Soc.***131,Part B,** 2961-3012.

Velicogna I., and J. Wahr, 2006: Measurements of time-variable gravity show mass loss in Antartica, Science, 311, 1754-1756.

Vialard and Delecluse 1998. An OGCM study for the TOGA decade. Part I: Role of the salinity in the Physics of the western Pacific Fresh Pool. J. Phys. Oceanogr., 28, 1071-1088.

Vialard, J., Vitart, F., Balmaseda, M. A., Stockdale, T. N., and Anderson, D. L. T., 2005. An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. *Mon. Wea. Rev.*, **131**, 1379–1395. See also ECMWF Technical Memorandum No 417.

Vidard, A., D.L.T. Anderson, M. Balmaseda, 2006: Impact of ocean observation systems on ocean analysis and seasonal forecasts. *Mon. Wea. Rev.*, In press.

Vidard, A., M. Balmaseda, 2007: Assimilation of altimeter data in the ECMWF ocean analysis system. In preparation

Vidard, A., M. Balmaseda, 2007: Assimilation of in situ observations in the ECMWF ocean analysis system. In preparation.

Vitart, F., 2005: Monthly Forecast and the summer 2003 heat wave over Europe: a case study. Atmospheric Science Letter, 6(2), 112-117.

Vitart, F., S. Woolnough, M.A. Balmaseda and A. Tompkins, 2006: Monthly forecast of the Madden-Julian Oscillation using a coupled GCM. *Mon. Wea. Rev.*, In press.

Wolff, J., E. Maier-Reimer and S. Legutke, 1997. The Hamburg Ocean Primitive Equation Model. *Deutsches Klimarechenzentrum, Hamburg*, Technical Report No. 13.

Woolnough, S. J., F. Vitart and M. A, Balmaseda, 2006: The role of the ocean in the Madden-Julian Oscillation: Sensitivity of an MJO forecast to ocean coupling. *Q. J. R. Meteorol. Soc.*, in press.