4DVAR assimilation of subsurface and altimetry observations in the HOPE OGCM adjusting surface fluxes

Gerrian Appeldoorn[1]
Geert Jan van Oldenborgh
KNMI, De Bilt, Netherlands
March 10, 2003

Abstract
For seasonal forecasts a good ocean initial state is essential. This is obtained by combining observations of the ocean with the dynamics of an ocean model. The observations are generally sea surface temperature measurements, in-situ subsurface temperature measurements in some regions, and satellite altimetry measurements representing a measure of integrated heat content. To optimally combine the observations with an OGCM a 4DVAR data-assimilation scheme is implemented that varies the momentum, heat and freshwater forcings of the ocean model over 16-week periods.

To gain insight in how the method adjusts the ocean analysis, several identical twin experiments are carried out in which a model run serves as the 'truth'. Either the initial state or the forcing is perturbed with respect to this 'truth'. It is then approximated again by assimilation of sea surface temperature data, subsurface temperature data and/or sea surface height data generated by the truth run but with the same properties as real observations.

The method provides an ocean analysis in which the temperature structure of the ocean is improved. The assimilation of either altimetry data or subsurface temperature data gives comparable reductions in all cost functions. However, the analysis does depend on which dataset is assimilated. Subsurface temperature data leave relatively small residuals near the equator at the cost of larger residuals at higher latitudes. Altimetry data result in a more uniform adjustment, in which the ocean analysis at depth is also improved.

Different fluxes alter the state of the ocean in different regions. The effect of changes in the zonal wind stress is most obvious in a band up to 30° from the equator and at thermocline depth. Adjusting the heat flux also has effects at higher latitudes. It has impact through the ocean mixed layer. The effect of freshwater flux variations is negligible. To make improvements below the thermocline and in chaotic regions, the initial state will also have to be adjusted.

1. Correspondence to: KNMI, P. O. Box 201, NL-3730 AE De Bilt, the Netherlands

1 Introduction

Seasonal forecasts of the weather are possible whenever slowly varying boundary conditions perturb the chaotic weather in a predictable way. The most important of these boundary conditions is sea surface temperature (SST). SST influences the weather on a global scale, for example during an El Niño event, but also influences the weather on smaller, local scales, for example influencing coastal air temperatures. In order to make good SST forecasts, the state of the ocean has to be well-known at the beginning of the forecast. This ocean analysis is best determined by the combination of observations with model physics, through data assimilation. Most observations are sea surface temperature measurements, both from satellite and in-situ, in-situ subsurface temperature and satellite sea surface height observations.

Satellite and in-situ SST measurements are analysed at NCEP into a gridded global product with good coverage and accuracy. Subsurface observations are less homogeneous, in many regions they are sparse. However, some small areas are well sampled with temperature measurements. The observations are concentrated at buoys in coastal zones, the Tropical Atmosphere Ocean (TAO) array and the Pilot Research Moored Array in the Tropical Atlantic (PIRATA). Measurements from expendable bathythermographs (XBT) also contribute to a higher concentration along the main shipping routes. This dataset of irregular coverage directly describes the three-dimensional thermal structure of parts of the ocean.

Another source of ocean data is satellite altimetry. The TOPEX/POSEIDON (T/P) satellite measures SSH along tracks that are repeated every 10 days and are $\sim$ 300 km ($\approx$ 2.8 °) apart. The coverage of the SSH dataset is uniform and the resolution is relatively high. When simulating the thermal structure of the ocean, these properties are advantageous. However, a disadvantage of SSH measurements is that the 2D sea level information has to be projected into the vertical to obtain a 3D subsurface structure describing thermal and saline properties.

To describe both thermal and saline subsurface properties, Troccoli and Haines [1999] have applied a method in which only temperature profiles were assimilated. They suggest that using SSH data in addition to temperature data would be the most useful way forward in reconstructing the density field for ocean and coupled modeling experiments. In the context of a seasonal forecasting system, Alves et al. [2000] consider the projection of SSH data onto the temperature and salinity fields in an idealised identical twin setup. They also indicate that combining SSH data with subsurface in-situ data may provide an even greater benefit. Segschneider et al. [2001] then use the combination of datasets to initialize a forecast. Synthetic temperature observations were first derived from SSH data and then optimal interpolation was used to combine them with the directly-observed temperature profiles. Several measures of skill suggested that forecasts improved by initializing the model with an ocean analysis in which both temperature and SSH data were assimilated. Vossepoel et al. [2002] have applied salinity corrections in addition to temperature corrections while assimilating both temperature profiles and SSH data. The salinity correction makes the model follow both the temperature and the sea level more accurately.

In this study, we investigate how a 4DVAR data assimilation scheme uses temperature and SSH data to adjust the ocean state in the HOPE model. The system is tested on the same period and in the same way that was used by Alves et al. [2000], taking into account the findings of Segschneider et al. [2001]. The 4DVAR method used does not update the subsurface state directly, as in Weaver

et al. [2002] for example. Instead, it alters the ocean state by adjusting the fluxes into the ocean. When the ocean model is assumed to be perfect, the analysis errors are due to poorly known forcing fields and can be corrected by adjusting the fluxes. Some systematic model errors can also be compensated in this way. An advantage of adjusting the fluxes is that the ocean state is updated in a balanced way. An update in the density field does not result in deterioration of the velocity field [Burgers et al., 2002]. In addition, the ocean itself remains consistent during the assimilation period. A consistent analysis is a very useful tool when a theoretical study is carried out. A disadvantage is that some model errors, for example a too high diffusivity, cannot be corrected. Also the slow dynamics of the ocean, that occur below the thermocline, cannot be corrected by adjusting surface fluxes over a relatively short period.

Among others, Bonekamp et al. [2001] and Vossepoel [2002] have done earlier work in which they use fluxes to adjust the ocean state. They both focussed on the equatorial Pacific, since most improvements were expected here. However, it is also possible to improve the ocean state outside the equatorial Pacific. In this article, a global extension of Bonekamp et al. [2001] is described. The scheme now includes several improvements. First, the 4DVAR scheme is extended to use forcing updates of heat and freshwater in addition to wind forcing. By using all fluxes, it is expected that improvements extend out of the equatorial region. Second, the time period over which data is assimilated is doubled while the number of iterations was reduced to keep a constant computing time. By increasing the assimilation window to sixteen weeks, it is expected that the region of possible improvements increases and the amplitude of the errors decrease. The validity of the linearization of the adjoint model has been investigated in van Oldenborgh et al. [1999] and was found to be good when considering flux variations over 16 weeks. Finally, the scheme is extended with additional cost function terms of SST and SSH data.

This article illustrates how 4DVAR fluxes adjust the global thermal ocean state in our model. It gives a global impression of how the method adjusts the ocean state and where most of the improvements are achieved. Since initial state and forcing perturbations have different characteristics, we correct for them separately in our identical twin experiments. From a 'truth' run pseudo-observations are generated at the same place and time as real observations. The pseudo observations are then used to approximate the 'truth' again by assimilation into perturbed runs. Though the truth is exactly known, there are also disadvantages to using identical twin experiments. Because we have defined a 'truth' from the model, model errors are not taken into consideration. Furthermore, since we use measurements that directly originate from our model, in this configuration no measurement errors were taken into account. This includes representation errors due to sub-gridscale variability. In studying the adjustments of the 4DVAR fluxes, three questions were kept in mind. First, does the 4DVAR system improve the temperature and salinity structure of the ocean? Second, where are most of the improvements and how do they depend on the observations used? Third, which fluxes are responsible for these improvements?

The outline of the paper is as follows. In section 2 we briefly describe the HOPE ocean model and the observational data. Section 3 then deals with the assimilation method of the observations. An outline of the experiments is also given here. The initial state error and forcing error experiments are discussed in section 4.1 and 4.2 respectively. Section 5 finally summarizes the conclusions.

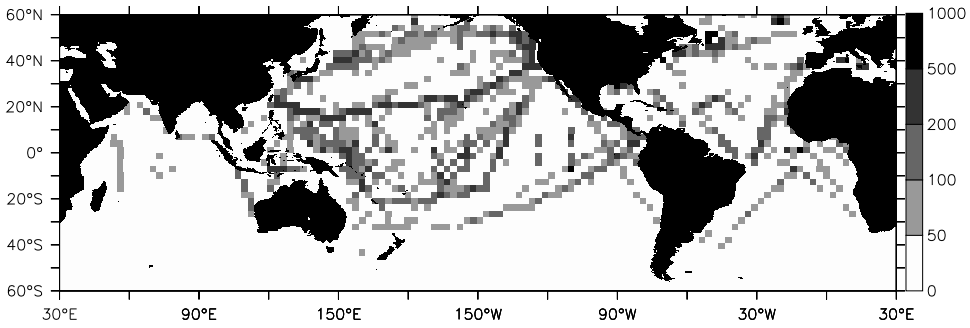

Figure 1: Numer of observations in the subsurface temperature dataset during a period of sixteen weeks (last assimilation cycle), summed over depth.

2 Model and Data

2.1 The HOPE Model

For this study, the Hamburg Primitive Equation (HOPE) OGCM is used. The version is identical to the one used in the ECMWF system-1 seasonal forecasts. A detailed description of the model is given by Wolff et al. [1997]. The prognostic variables are temperature, salinity, horizontal velocities and sea surface height. The availability of the latter variable makes inclusion of SSH observations in a cost function straightforward.

The equations in HOPE are spatially discretized on an Arakawa E-grid. The model has a zonal resolution of 2.8° at all latitudes. The meridional resolution varies from 0.5° near the equator to 2.8° at higher latitudes. In the vertical, 10 levels cover the first 300 meters below the sea surface and another 10 levels cover the deep ocean. The time step is 2 hours. The vertical mixing in HOPE is based on a Richardson-number dependent formulation and a simple mixed-layer scheme to represent the effects of wind stirring.

The model is forced with daily surface fluxes of momentum, heat and freshwater. As will be discussed later, these fluxes are adjusted by the 4DVAR scheme to improve the analysis.

2.2 Data

Three different kinds of observations are used in this study: subsurface temperature, SSH anomalies and SST. From the observational datasets, only time, place and error estimates are used in the identical twin experiments. The subsurface temperature observations originate from buoys or ships and are obtained from the ECMWF [Alves et al., 2000]. The measurements are concentrated near the western boundary currents, the PIRATA and TAO/TRITON arrays and some ship routes in the Pacific Ocean, see Figure 1.

The SSH dataset consists of T/P measurements and is based on Schrama et al. [2000]. The experiments were situated in early 1992 to allow for comparisons with Alves et al. [2000]. However, T/P data are only available from the end of 1992 onward. As identical twin experiments require only the properties of the observations, not the values themselves, these were taken from the next year (1993). The data have been aggregated to measurements every second (7 km). They are quality controlled and further reduced to the HOPE grid. The averaging procedure also

yields an error estimate (see appendix A).

The last dataset used is the SST dataset from Reynolds and Smith [1994]. It consists of daily data on the atmospheric grid of HOPE (2.8°by 2.8°). The observations in this dataset are dependent in space and time. Assimilation of SST data can be interpreted as an alternative to the relaxation to SST with a mixed-layer depth dependent relaxation time.

2.3 The 4DVAR scheme

The model and data described above are tied together by a 4DVAR assimilation scheme that is based on the system of Bonekamp et al. [2001]. The current version comprises cost function terms for SST and SSH in addition to subsurface temperature measurements. Originally, the scheme only adjusted the fluxes of momentum, but in the current version all surface fluxes are adjusted.

The 4DVAR data assimilation method attempts to minimize a cost function J that quantifies the misfit between model values and observations. It is made up of a background term, J_{bg}, and the observational terms J_{SST}, J_{SUB}, and J_{SSH}.

$$J = J_{\mathrm{bg}} + w_{\mathrm{SST}} J_{\mathrm{SST}} + w_{\mathrm{SUB}} J_{\mathrm{SUB}} + w_{\mathrm{SSH}} J_{\mathrm{SSH}} \; . \tag{1}$$

The subscript of the observational terms indicates which dataset is used. The weight terms w are used to compensate for dependencies in the data that are not explicitly resolved.

The first term J_{bg} is the background term. This term penalizes for too strong deviations from the prescribed background forcing,

$$J_{\mathrm{bg}} = \mathbf{c}^T \mathsf{B}^{-1} \mathbf{c} \tag{2}$$

The vector $\mathbf{c}$ contains the control parameters, two-weekly grid point values of momentum, heat and freshwater fluxes. B is the background covariance matrix. The covariances of all fluxes are estimated from the variability of the ERA dataset in the same manner as described in Bonekamp et al. [2001], but following their recommendations the error decorrelation lengths have been estimated to be half the variability decorrelaton lengths.

In principle, the observational cost function terms have the standard form:

$$J_i' = (\mathsf{H}_i \mathbf{x}(\mathbf{c}, \mathbf{x}_0) - \mathbf{z})^T \mathsf{E}_i^{-1} (\mathsf{H}_i \mathbf{x}(\mathbf{c}, \mathbf{x}_0) - \mathbf{z}), \quad i = \mathrm{SUB}, \mathrm{SSH}, \mathrm{SST} \; . \tag{3}$$

H is the observation operator which projects the model output on the model counterpart of the observational data. The model output depends on the initial state $\mathbf{x}_0$ and the control vector $\mathbf{c}$. The vector $\mathbf{x}$ contains the HOPE model trajectory, $\mathbf{z}$ the observations. E is the observational error covariance matrix which accounts both for the errors in the measurements and in the projection H.

When using SSH anomalies, a model and observational climatology are needed. Although the model climatology should be derived by applying the assimilation procedure iteratively, we use a model climatology of the ECMWF (Alves et al. [2002]) instead. The observational climatology is computed from the T/P data in the 1992–2000 period. Due to model errors and uncertainties in the geoid, the model climatology is not equal to the observational climatology. In the context of identical twin experiments, the climatology is irrelevant however.

The subsurface measurements (daily in the case of buoys) are far enough apart to allow the approximation of E being diagonal. The SSH along-track measurements

are reduced to effective measurements in diagonal form following the method described in Appendix A. The dependencies of the gridded SST field are compensated for by a weight value $w_{\mathrm{SST}} < 1$ rather than an explicit estimate of the error covariance matrix. The weight of the SST cost function term in equation 1, w_{SST}, was taken as 0.5. The background term J_{bg} is relatively small. Since the dependency of the altimetry measurements in a single grid box is already taken care of, w_{SSH} only has to take into account the dependence between the gridboxes and is set to 1. w_{SUB} is set to 2.

Some measurement errors or model deficiencies can have a large effect on the ocean analysis during assimilation [Bonekamp et al., 2001]. Large, localized measurement errors for example can lead to unwanted deformations in the ocean analysis. Also in places where the ocean is not represented correctly in the coarse-resolution model, for example in western boundary currents, the large deviations between model and observations should not affect the analysis too much. Consequently, the following non-quadratic form was used to reduce the weight of the observations that deviate very strongly from the model:

$$J_i = \sum_n \frac{j_n}{j_n + 2^2}, \quad j_n = \frac{(H_n x_n - z_n)^2}{(\delta z_n)^2} \tag{4}$$

The error on observation n is denoted by δz_n and the sum is taken of all n observations.

3 Set-up of the Numerical Experiments

3.1 The assimilation procedure

The data assimilation experiments are set up in such a way that they simulate an operational forecasting cycle, see figure 2. First, a model trajectory is calculated by integrating the model forward in time for a period of sixteen weeks. The model trajectory and the (pseudo-) observations are used to make a first evaluation of the cost function J. This value reflects the misfit between the first-guess model trajectory and the (pseudo-) observations, since the additional momentum, heat and freshwater fluxes are still zero. From here, we calculate a best-guess trajectory by constructing nine two-weekly sets of new fluxes that minimize J.

The minimization is done with the quasi-Newton routine M1QN3 [Gilbert and C. Lemarechal, 1989]. The required evaluations of the gradient of the cost function are obtained from backward integration with adHOPE, described in van Oldenborgh et al. [1999]. Because the integrations with adHOPE are computationally expensive, the allowable number of iterations with M1QN3 is restricted. In all the experiments, the maximum number of iterations has been set to two.

After the first period of sixteen weeks is dealt with, we repeat the assimilation cycle with a time window shifted by two weeks. As initial state the best-guess estimate (analysis) at $t = t_1$ of the first cycle is taken. For the first eight two-weekly sets, the best-guess flux updates of the previous cycle are reused, whereas the new flux updates at the end are set to zero. Each assimilation experiment consists of seven shifted assimilation cycles.

3.2 Description of the experiments

A list of the experiments is given in table 1. The reference run (runT) defines the 'truth' and generates the pseudo-observations for the assimilation experiments. The

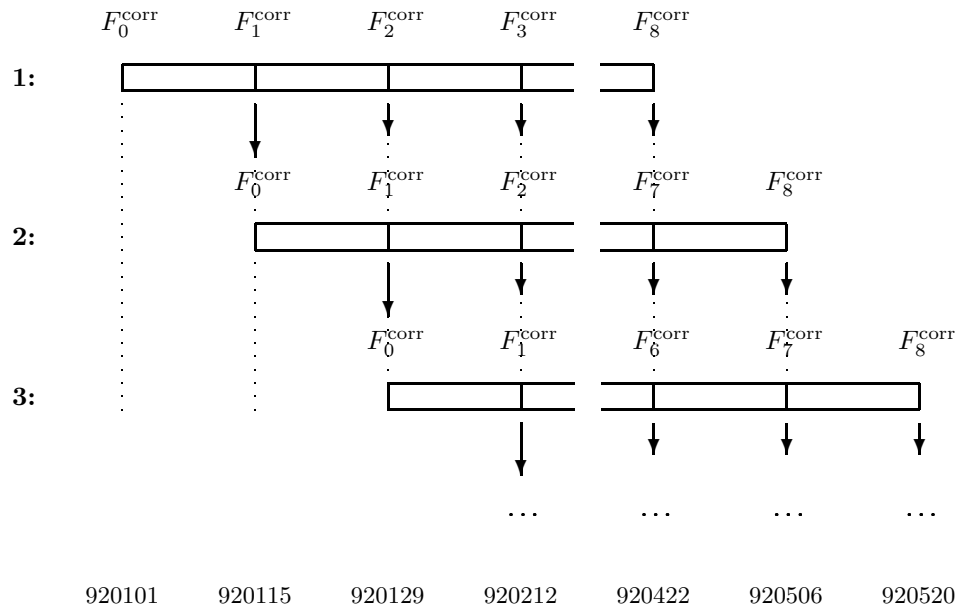

Figure 2: Schematic diagram of three out of seven assimilation cycles. Time is from left to right. The re-use of the two-weekly flux corrections from the previous assimilation window is indicated by the arrows. In addition, the long arrows mark the re-use of the model state as the initial state of the next assimilation window.

initial state of runT is the same as the one used in Alves et al. [2000]. It is obtained by integrating the model from Levitus initial conditions at 1985 to January 1st, 1992. During this integration ECMWF reanalysis fluxes (ERA-15) were applied. SST was strongly relaxed to Reynolds SST on a three day time scale and sea surface salinity (SSS) was weakly relaxed to Levitus SSS on a one month time scale. The reference run starts at 1 January 1992 and is integrated forward in time by again applying the ECMWF reanalysis fluxes. During the reference run, a weaker SST relaxation was applied of 40 $\mathrm{Wm}^{-2}\mathrm{K}^{-1}$.

The experiments ending with the number 1 are done to test the capability of the 4DVAR scheme to correct for errors in the initial state. These errors can also be partly corrected by adjusting the fluxes. The experiments start from an initial state that differs from the truth and approximate it again. The initial state used, is the perturbed initial state as is described in Alves et al. [2000]. It is obtained from the same spin up integration as described earlier, but now it is forced from 1990 onward by ECMWF operational fluxes instead of reanalysis fluxes. So in the final two years, different fluxes are applied.

The experiments ending with the number 2 are done to test to what extent the 4DVAR scheme is able to correct for errors in the surface forcing. These experiments start from the same initial state as the truth, while surface fluxes are applied from the ECMWF operational archives (OPS).

The first-guess runs, indicated by 'FG', set a baseline to which the data assimilation experiments can be compared. In the initial state error experiments, FG1 shows to what extent the model converges from the perturbed initial state towards the truth by only applying the correct surface fluxes. In the forcing error experiments, FG2 illustrates to what extent the model diverges from the truth because of different forcing fields. In these experiments, no data are assimilated yet.

The performance in reducing errors with respect to the truth is examined in the assimilation experiments, named SST, SSH, SUB and BOTH. These experiments clarify the effect of assimilating data with varying properties. In the SST experiments, SST data is assimilated only. In the SSH experiments, SSH observations are assimilated additionally. In the SUB experiments, subsurface temperature observations are assimilated additionally. Experiment BOTH assimilates both SSH and subsurface temperature data in addition to the SST data. It should be noted that in all assimilation experiments the SST relaxation is omitted. Instead, the heat flux resulting from the SST relaxation in runT is imposed additionally.

4 Identical Twin Experiments

In identical twin experiments the truth is generated in a model run and pseudo observations are generated from that model run. The pseudo observations are then assimilated to approximate the truth again. The next subsections examine how the 4DVAR scheme corrects for errors from the truth. Paragraph 4.1 first discusses the performance in reducing the initial state errors. Paragraph 4.2 then discusses the correction for forcing errors.

4.1 Correction for Initial State Errors

Fig. 3 first shows the decrease of the cost function terms J_{SST}, J_{SUB} and J_{SSH} as a function of iteration and assimilation cycle. Since each assimilation cycle starts with an evaluation of the cost function, this value is displayed first. After that, values

Experiment	Initial state	Forcing	Data assimilated
runT	ERA forced	ERA	none
FG1	OPS forced	ERA	none
SST1	OPS forced	ERA	SST
SSH1	OPS forced	ERA	SST, altimetry
SUB1	OPS forced	ERA	SST, subsurface
BOTH1	OPS forced	ERA	SST, altimetry and subsurface
FG2	ERA forced	OPS	none
SST2	ERA forced	OPS	SST
SSH2	ERA forced	OPS	SST, altimetry
SUB2	ERA forced	OPS	SST, subsurface
BOTH2	ERA forced	OPS	SST, altimetry and subsurface

Table 1: Overview of the experiments

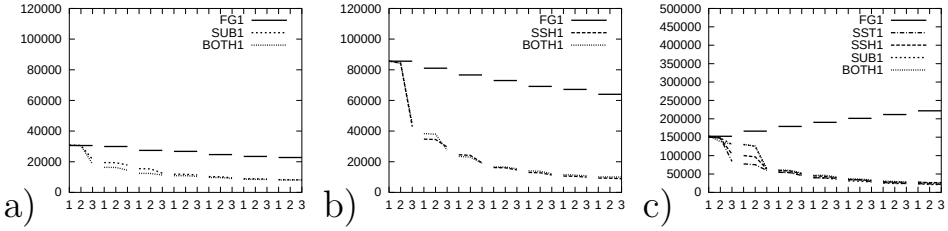

Figure 3: Decrease of the cost functions J_{SUB} (a), J_{SSH} (b) and J_{SST} (c) as a function of iteration and assimilation cycle.

are depicted from the two iterations, in which the derivatives are computed again and the value of the cost function decreases. Two iterations in each cycle already give a large decrease of the cost function. Increasing the number of iterations any further does not lead to large reductions of the cost function any more [Bonekamp et al., 2001].

In general, the first-guess terms J_{SUB} and J_{SSH} decrease when the assimilation window is shifted. Starting from an incorrect initial state, by applying the correct fluxes the ocean is forced towards the correct state again. The first-guess term J_{SST} does not decrease significantly. The explanation for this is two-sided. First, the SST of the perturbed initial state starts off close to the SST of the 'true' initial state. Both initial states are generated with the same strong relaxation to SST. So J_{SST} starts off small. Thereafter, the relaxation to SST in the 'truth' is replaced by a slightly different forcing term in the analysis. So it is possible for J_{SST} to even increase.

In all assimilation experiments the cost function terms are reduced compared to their first-guess value. In the first cycles, the decrease of the cost function depends on the dataset that is assimilated. After a few cycles, the cost function terms seem to converge to a constant value. Note that the value of the cost function at the start of a new cycle is not as high as the first-guess value. The new cycle uses the updated initial state and most fluxes from the preceding cycle.

By using the cost function to measure the quality of the analysis, global information is projected into a single value. In addition, the cost function only takes into account the deviation from reality at the observation points. A more

detailed view can be obtained by looking at the differences between the analyzed and the 'true' ocean state, which is exactly known in identical twin experiments. The integral over depth and longitude of the absolute value of these residuals is shown for the Pacific, Indian and Atlantic Ocean basins in Figure 4. Because we are mainly interested in the subsurface adjustments, the two uppermost layers (at 10m and 30m) are omitted. The residuals are given for the first week of the last assimilation cycle. Here the analysis is best, because the adjusted fluxes have updated the ocean state most frequently.

In the first-guess the residuals peak near 10–15° from the equator, where the trade winds and their variability reach a maximum. Over the period of two years in which the perturbed initial state is generated, the different fluxes have disturbed a region around the equator. In the three months of the analysis, the correct fluxes can only partly compensate for these disturbances, in a smaller region. The assimilation experiments manage to reduce the residuals of the first-guess even further. Note the difference from the similar figures in Bonekamp et al. [2001] showing the much larger reduction in the cost function, which only includes points that are used in the minimization.

The residuals differ when different observational datasets are used. In the equatorial region, large first-guess residuals can be found in the Pacific. By assimilating SST data these residuals are only slightly reduced. When another dataset is assimilated additionally, the residuals are further reduced. The subsurface dataset has many measurements concentrated near the equator and reduces the largest of the residuals quite well in this region. At higher latitudes, however, we observe larger residuals. The cost function is undefined in the regions without data, so fluxes can be defined here arbitrarily in favour of a reduction of the cost function in data-rich regions. As a result, the analysis is hardly improved or even deteriorated at higher latitudes (near 25°). Since the coverage of altimetry data is more regular, assimilating altimetry data results in a more uniform adjustment. The residuals are somewhat larger near the equator, but they are considerably reduced at higher latitudes. Another example can be found in the Indian Ocean. Here subsurface data cannot bring the analysis as close to the truth as the altimetry data, simply because there are fewer subsurface observations.

At high latitudes, residuals are large where the ocean model behaves chaotically, even on the coarse HOPE grid. The deviation of the first-guess from the truth is large in regions of the western boundary currents and the Antarctic Circumpolar Current, which are visible in the Atlantic near 50°N and 30°S. The residuals of the first-guess are hardly reduced in the assimilation experiments, even when data are present. At some places, data assimilation leads to slightly increased residuals (near 50°N).

Looking at the vertical distribution of the residuals, large differences are found between the tropics and extratropics. Figure 5 shows the residuals as a function of depth, distinguishing between low latitudes (20° S to 20° N) and high latitudes (30° to 60°). In the equatorial region, relatively small differences between the first-guess and the truth are found near the surface and large differences are found near the thermocline, around 100m. At higher latitudes, the differences between the truth and the first-guess are larger near the surface than the relatively small differences at depth. In the analyses the largest differences are reduced. Zonal wind updates change the depth of the thermocline in the equatorial region when subsurface or altimetry data are assimilated, which leads to smaller residuals at depth. When only SST data are assimilated, temperatures at depth barely change. The thermo-

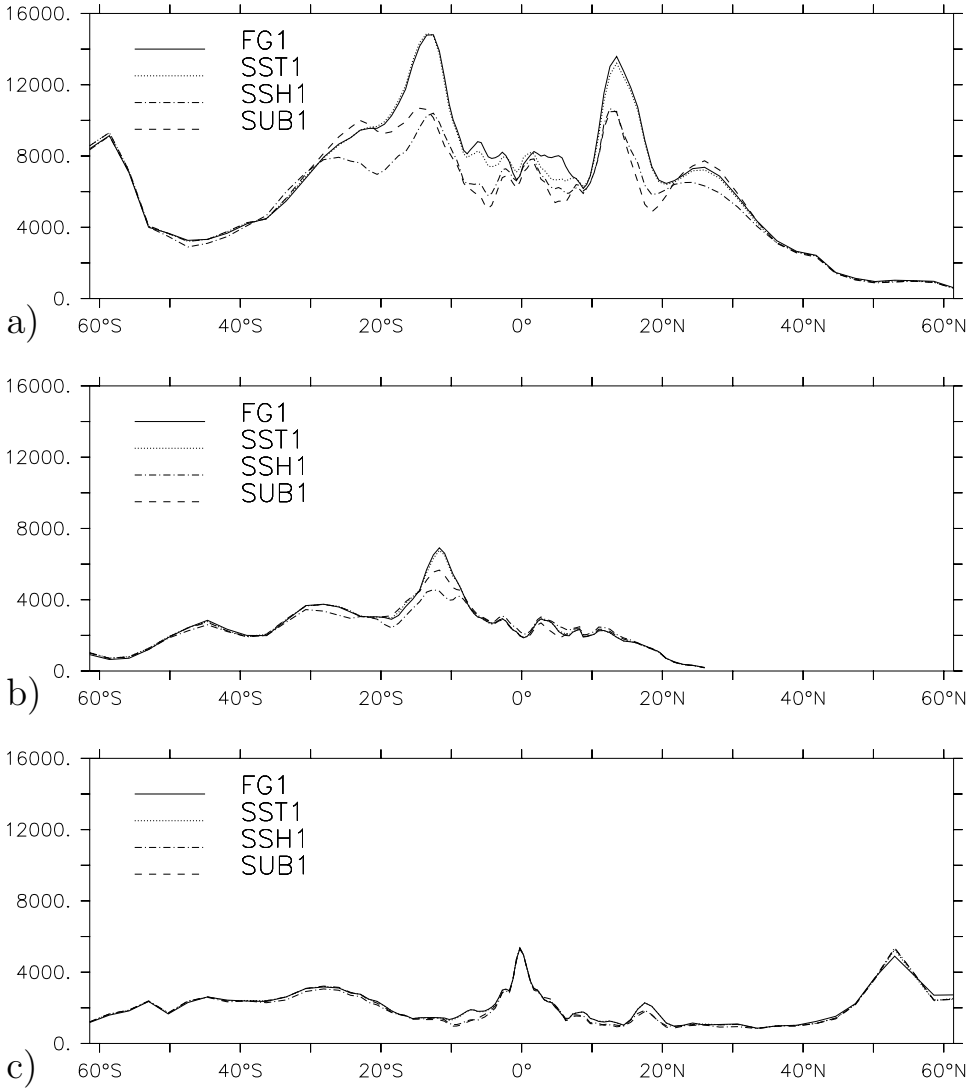

Figure 4: Residuals ($\int_{z>40\mathrm{m}} |T_{\mathrm{truth}} - T_{\mathrm{analysis}}| dxdz$) as a function of latitude of FG1, SST1, SUB1 and SSH1 in the Pacific (a), Indian (b) and Atlantic (c) Ocean.

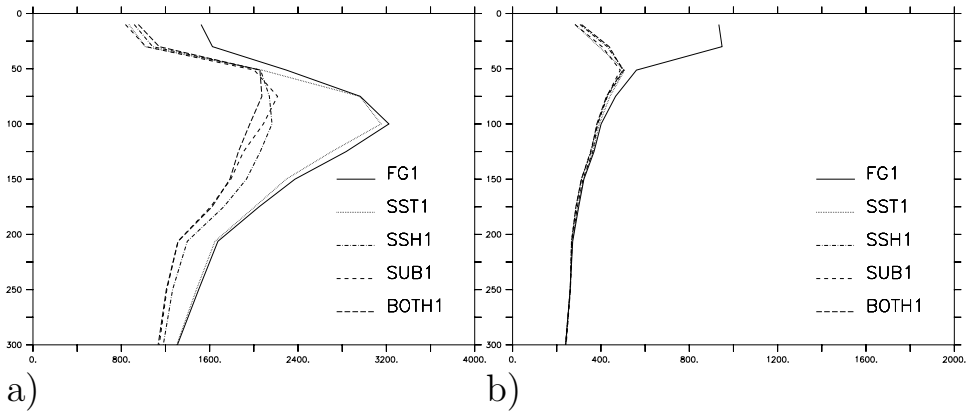

Figure 5: Residuals ($\int |T_{\text{truth}} - T_{\text{analysis}}| dxdy$) as a function of depth of FG1, SST1, SUB1 and SSH1 from 20°S to 20°N (a) and from 30° to 60° (b).

cline can reach depths up to 1000m in this latitudinal band, so that improvements are possible up to that depth. As will be discussed later, zonal wind stress cannot effectively change the state of the ocean at higher latitudes due to the limited assimilation window. At higher latitudes, no improvements are visible near thermocline depth. Here improvements from heat fluxes are most pronounced. Instead of displacing the thermocline by zonal wind stress, heat fluxes gradually reduce the residuals through the surface. These heat fluxes can be determined by assimilation of SST only; assimilating subsurface or altimetry data barely improves the analysis at depth. Differences between the analysis and the truth continue to exist in the period of assimilation.

The difference between the tropics and extratropics can be seen directly by considering the effect of momentum and heat flux adjustments on the cost function. Eq. 5 divides the change of the cost function into the contributions of the separate fluxes:

$$\Delta J \approx \sum_i \left\{ \frac{\partial J}{\partial \tau_{x,i}} \Delta \tau_{x,i} + \frac{\partial J}{\partial \tau_{y,i}} \Delta \tau_{y,i} + \frac{\partial J}{\partial Q_i} \Delta Q_i + \frac{\partial J}{\partial F_i} \Delta F_i \right\} \tag{5}$$

Here, J is the cost function. τ_x, τ_y, Q and F are the zonal momentum flux, the meridional momentum flux, the heat flux and the freshwater flux respectively. The sum refers to all surface grid points. Contributions of the meridional momentum and the freshwater flux proved to be minimal and are not shown.

Figure 6 maps the contribution according to equation 5 of zonal wind stress and heat flux updates at the beginning of the last BOTH1 cycle. The figure shows that many improvements in the Pacific region from 20°N to 20°S can be ascribed to the zonal momentum flux updates. In the Indian and Atlantic ocean basins the impact is less pronounced. An estimate of the highest latitude where improvements from zonal wind updates are visible can be calculated from the planetary wave speed. An approximation of the phase speed of this wave is $v = -c^2 \cos\phi / 2\Omega R \sin^2\phi$ (see Gill [1982]). Here c is the speed of a Kelvin wave at the equator, Ω and R are the angular velocity and radius of the earth respectively, and ϕ represents latitude. The planetary wave speed permits information to be propagated between two successive tracks of altimetry data (2.8°) during the time of the assimilation window (16 weeks) up to 25-30°latitude. As the cost function is only defined on the tracks, at higher

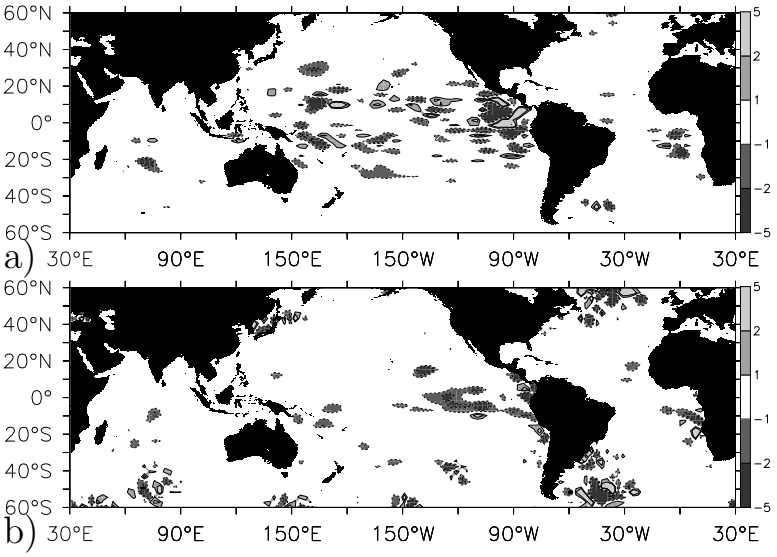

Figure 6: The contributions $\partial J/\partial \tau_x \, \Delta\tau_x$ (a) and $\partial J/\partial Q \, \Delta Q$ (b) to the decrease of the cost function.

latitudes its derivative to the wind stress is zero on part of the gap between tracks and the computed (large-scale) updates will be too small. Improvements from zonal wind updates are indeed visible in the latitudinal band from about 30°S to 30°N. At higher latitudes, the wave speed is such that improvements during the assimilation window are not significant anymore.

Improvements from heat updates are both visible in the equatorial region and at higher latitudes. In the equatorial region, e.g. near 120°W, adjusting the heat flux is an efficient way in counteracting the errors near the surface. At higher latitudes some large-scale subsurface errors are also corrected for. An example is shown in figure 7, where the deviations south-east of Australia are almost cleared in the analysis. However, there are also small-scale errors at higher latitudes that are not corrected, for example near western boundaries. Here, adjacent small areas increase and decrease the cost function. The deviation from the truth changes sign at small spatial scales. The term $\partial J/\partial Q$ therefore also changes sign. However, large correlation scales in J_{bg} restrict the heat update ΔQ to large-scale patterns of the same sign, giving rise to the small-scale structures. Smaller-scale updates are not permitted as it is not natural to use heat fluxes to cancel the chaotic behaviour in these regions.

Figure 7 compares the performance of FG1 and BOTH1 in reproducing the ocean thermal state at the end of the experiment (July 15th), the starting point for a seasonal forecast. The sections show that the scheme can correct for initial state errors both near the equator and at higher latitudes. Improvements are due to momentum and heat fluxes, which do not affect the ocean below the thermocline in the 16-week assimilation window. In the Pacific, it is obvious that temperature differences are reduced at many places. In the equatorial section however, some large but small-scale deviations (e.g. at 130°W) are reduced at the expense of small, large-scale deviations, such as an upward shift of the thermocline. The fluxes have less impact on the ocean state in the Indian and Atlantic basins.

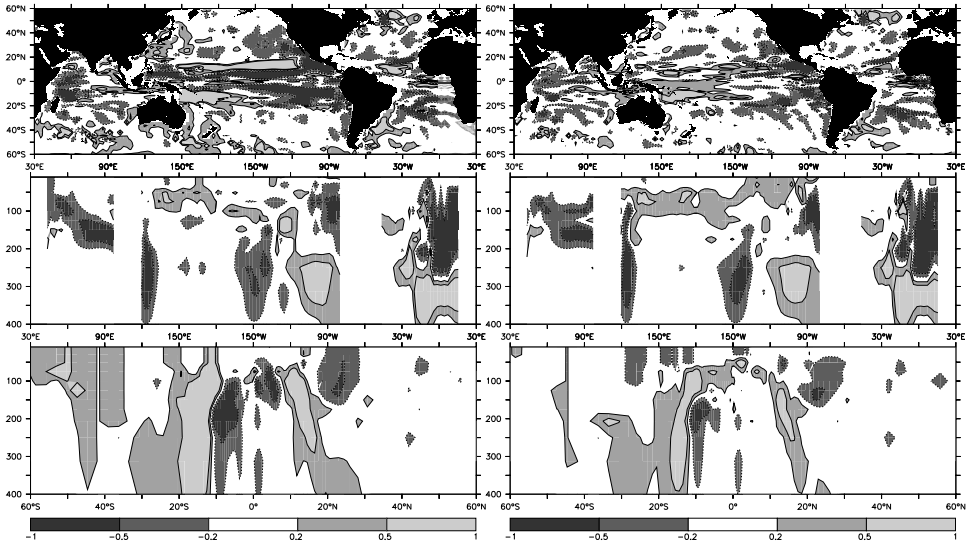

Figure 7: Final temperature differences between the truth and FG1 (L), and between the truth and BOTH1 (R) at 100m (top), at the equator (middle) and at the dateline (bottom).

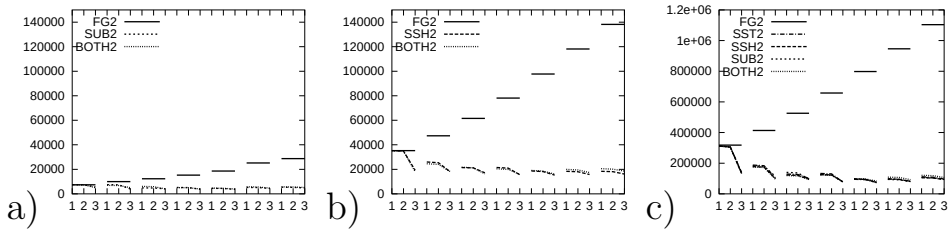

Figure 8: Decrease of the cost functions J_{SUB} (a), J_{SSH} (b) and J_{SST} (c) as a function of iteration and assimilation cycle.

Salinity observations are not assimilated in the experiments. However, the simultaneous use of subsurface temperature observations and altimetry can in principle be used to adjust the salinity [Vossepoel et al., 2002]. An example of this was observed in the western equatorial Pacific, where the mismatch is translated in a precipitation update that improves the salinity analysis.

4.2 Correction for Forcing Errors

Figure 8 shows the evolution of the cost function terms J_{SST}, J_{SUB} and J_{SSH} in the experiments that were started from the correct initial state, but where the forcing was taken from the ECMWF operational fluxes rather than the reanalysis. In this case fluxes are changed to compensate for incorrect forcing fields. In the experiments, all three first-guess terms increase when the assimilation window is shifted. The increase was expected since incorrect surface fluxes were applied. The model obviously diverges from the 'true' subsurface state from which it was started.

At the end of the assimilation experiments all observational cost function terms are reduced with respect to their first-guess value. The term J_{bg}, which en-

sures large-scale updates with a reasonable amplitude, is the main reason why the data assimilation procedure cannot decrease the cost function any further. The cost functions remain sligthly higher in BOTH2, when assimilation of subsurface data slightly increases the misfit with altimetry data and vice versa.

In the different ocean basins (figure 9), the first-guess shows large residuals near the equator from approximately 25°S to 25°N. Here the ocean response to wind forcing is fastest and forcing perturbations can easily bring about disturbances of the ocean state. The first-guess residuals, which are caused by the perturbed fluxes, are of course smaller than the residuals from the initial state error experiments, which are generated by perturbing the fluxes for two years before the start date of the experiments.

When only SST data is assimilated, some of the residuals at depth increase and some of the residuals decrease with respect to the first-guess residuals. As will be shown later, here the fluxes are only defined to minimize the residuals at the surface. The deviations from the first guess are larger than in the previous set of experiments, because SSTs were close to the truth in the perturbed initial state. The heat fluxes perturbations in this set of experiments are not small.

In the other assimilation experiments the residuals of the first-guess are further reduced. The reduction is not as large as the reduction in the cost function shown in Bonekamp et al. [2001], since all grid points were taken into account again rather than only the observational points. Different residuals were obtained when different observational datasets were used. In the equatorial Pacific, from 8°S to 8°N, many subsurface temperature data are generated by the TAO array. In this region, the subsurface data lead to smaller residuals than the altimetry data do. Since more altimetry data are available at higher latitudes, here the opposite is true. At these latitudes residuals from altimetry are smaller. In the Indian Ocean the benefit of more datapoints being available is also illustrated. Compared to altimetry data, the residuals are only slightly reduced when a few subsurface temperature data are assimilated.

In the Atlantic, chaotic model behaviour results in residuals that are hardly reduced compared to their first-guess value by varying the fluxes into the ocean. The Gulf Stream region near 55°N illustrates that data assimilation in chaotic regions can even lead to an enlargement of the residuals.

The residuals as a function of depth are shown in Fig, 10 , again distinguishing between the low and high latitudes. When perturbed fluxes are applied, the largest residuals are found near the surface. These residuals are effectively reduced due to the good SST observations and the easy way SST deviations can be corrected for by adjusting the heat flux. At depth in the tropics, residuals are also reduced when subsurface or sea level data are assimilated in addition to SST data. When SST data are only assimated however, equatorial fluxes are defined which minimise the residuals at the surface at the cost of residuals at depth. The residuals at depth are even larger than the first-guess residuals. At higher latitudes the residuals are mainly reduced by fluxes through the surface again. At depth, the residuals remain close to zero.

Figure 11 shows the contribution to the reduction of the cost according to variations of zonal wind stress and heat flux at the beginning (cf. equation 5) of the last BOTH2 assimilation cycle. Again, zonal wind updates dominate the equatorial region and the heat flux updates also appear at higher latitudes. Wind stresses are more explicit in the equatorial region now than in the first set of experiments and are even clear in the Indian and Atlantic basins. Adjusting the fluxes is a

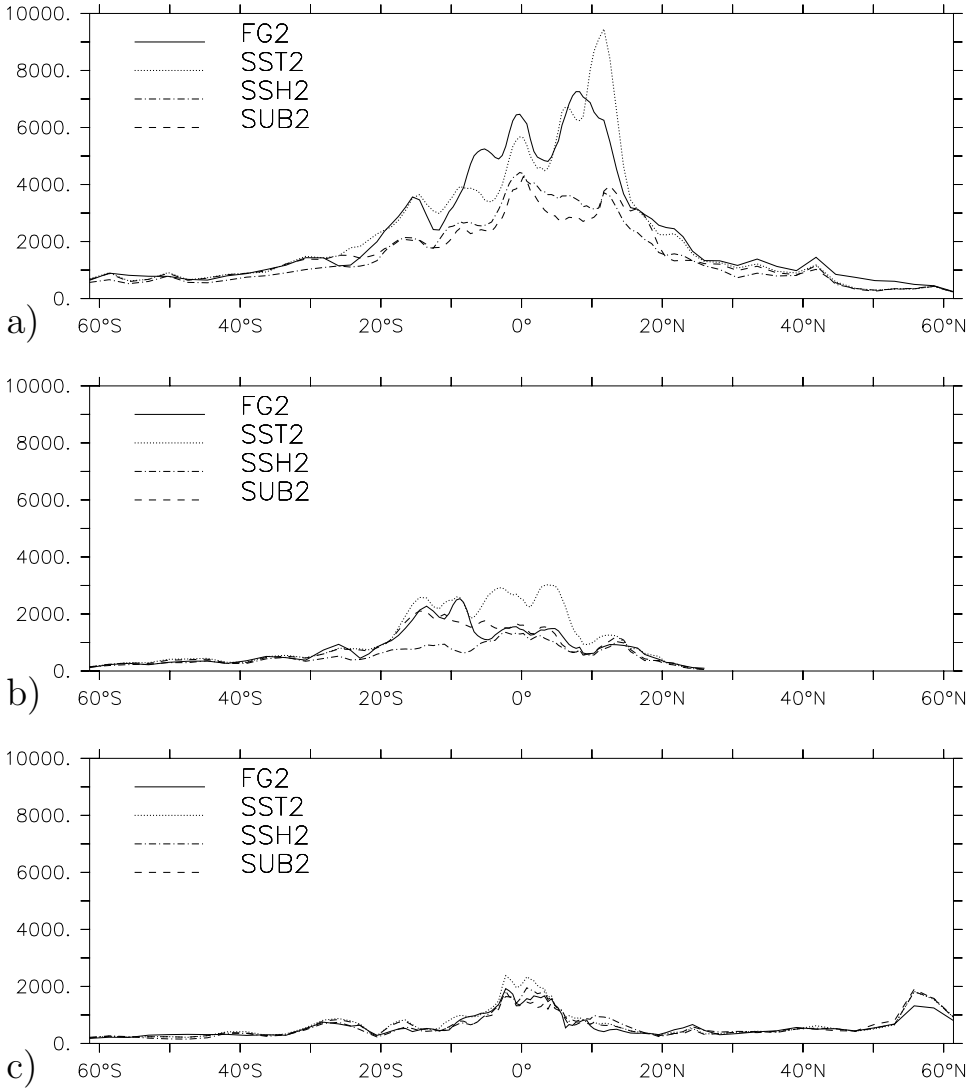

Figure 9: Residuals ($\int_{z>40\mathrm{m}} |T_{\mathrm{truth}} - T_{\mathrm{analysis}}| dxdz$) of FG2, SST2, SUB2 and SSH2 in the Pacific(a), Indian(b) and Atlantic(c) Ocean.

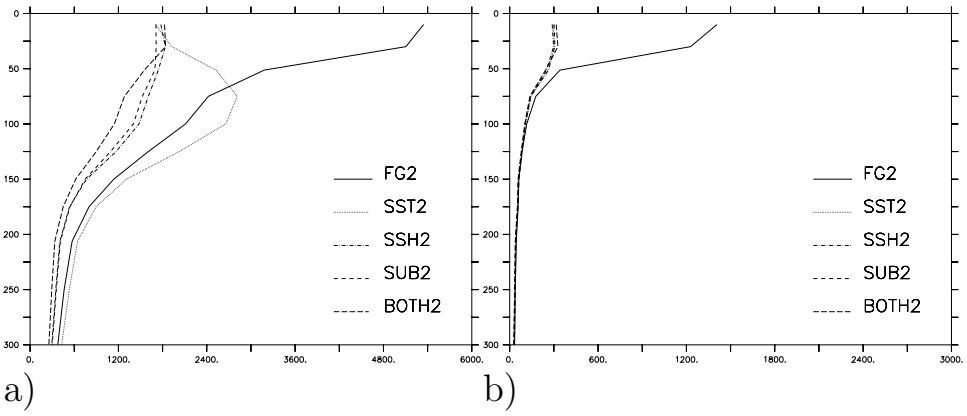

Figure 10: Residuals ($\int |T_{\text{truth}} - T_{\text{analysis}}| dxdy$) as a function of depth of FG2, SST2, SUB2 and SSH2 from 20°S to 20°N (a) and from 30°to 60°(b).

straightforward method to counteract forcing errors and the larger impact from the heat flux in these experiments is not really surprising. At higher latitudes, the heat fluxes are again not efficient in compensating for chaotic behaviour.

Figure 12 finally compares the performance of FG2 and BOTH2 at the end of the experiment, July 15th. Differences between the first-guess and the 'truth' show that most improvements are found near the surface and in the equatorial region. The sections show that the analyses are able to correct for forcing perturbations even at higher latitudes. Most pronounced are the improvements in the Pacific Ocean basin. In the Indian and Atlantic basins, they are visible but less pronounced again. The salinity structure of the ocean was again slightly improved as well (not shown).

5 Discussion and conclusions

In this study, the 4DVAR scheme of Bonekamp et al. [2001] has been extended to a global domain. Apart from subsurface temperature data, the scheme is now also able to assimilate SSH and SST data. To enlarge the impact in off-equatorial regions, the scheme has been adapted at three other points. The first point is that the scheme uses heat and freshwater fluxes in addition to the wind stress to adjust the ocean state. Now also heat fluxes are able to improve the ocean analysis, for example in regions in which cloud cover is not represented correctly. The second point is that the length of the assimilation window is doubled. Wind-generated planetary waves can now adjust the ocean state at higher latitudes during the assimilation time. The performance in the equatorial region is further improved by a reduction of the error decorrelation length of the fluxes.

Various identical twin experiments have been carried out to quantify the impact of the 4DVAR data assimilation on the ocean analysis. In the experiments SST data were assimilated in combination with SSH data, subsurface temperature data or both datasets to correct for either initial state errors or surface forcing errors. Both by assimilating SSH data and by assimilating subsurface temperature data, the analysis approximates the ocean state of the truth again.

From different points of view, the performance of the 4DVAR method was illustrated. Within several cycles, the 4DVAR scheme was able to reduce all cost

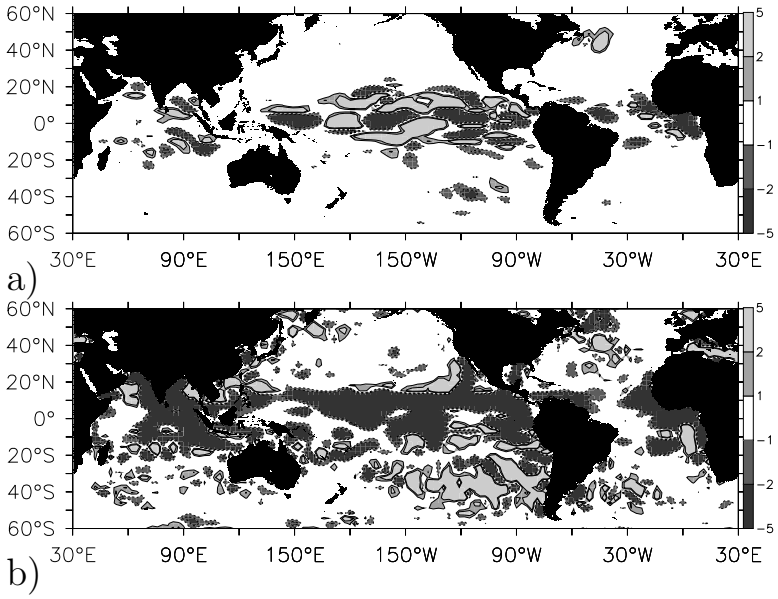

Figure 11: The contributions $\partial J/\partial \tau_x \; \Delta\tau_x$ (a) and $\partial J/\partial Q \; \Delta Q$ (b) to the decrease of the cost function.

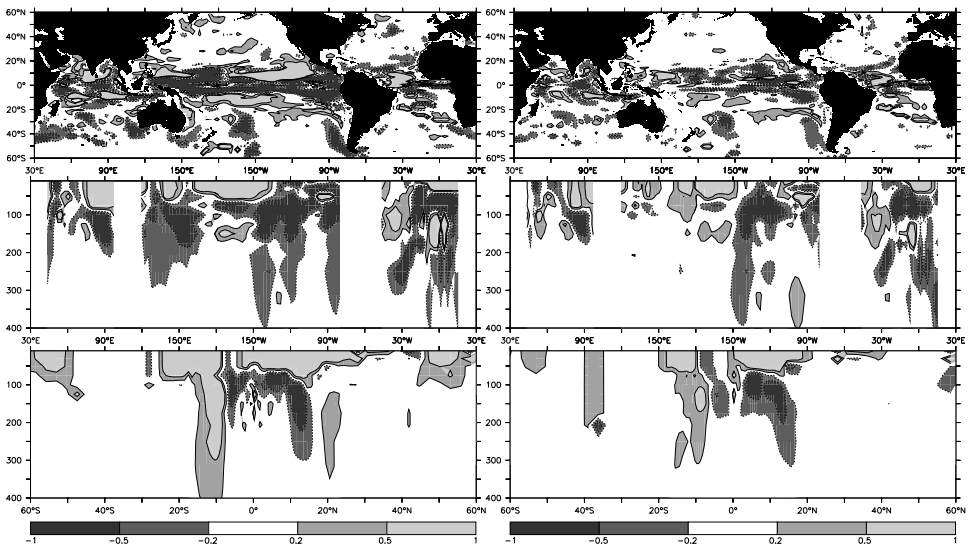

Figure 12: Final temperature differences between the truth and FG2 (L), and between the truth and BOTH2 (R) at 100m (top), at the equator (middle) and at the dateline (bottom).

function terms to an almost constant value. The final value of the specific cost function was barely dependent on which dataset was assimilated: assimilating altimetry data reduces the subsurface cost function almost as effectively as assimilating the subsurface data and vice versa. This equivalence was also found by Segschneider et al. [2001].

Dependent on which dataset is assimilated, differences also arise between the resulting analyses. By assimilating SST data only, the improvements are confined to some surface layers. Subsurface and altimetry data reach somewhat deeper. Subsurface temperature data are highly concentrated in the equatorial region and give information in depth. They leave relatively small residuals near the equator, at the cost of larger residuals at higher latitudes. Altimetry data on the other side are global with uniform coverage and resolution, but only give information at the surface of an integrated heat content. Assimilating altimetry data results in a more uniform adjustment, which also alters the thermal ocean structure at depth.

In the tropics, improvements by adjusting the wind stress are most obvious. Planetary (Rossby) waves can bridge the gap between observations within the 16-week assimilation window up to a latitude of 25–30° latitude, giving a rough limit on the area where the zonal wind stress can improve the ocean state in this configuration. The area is extended with respect to that of the shorter assimilation window of Bonekamp et al. [2001]. The momentum induced improvement principally implies a shift of the thermocline via Ekman pumping, which is often equivalent to a lifting or lowering of a temperature and salinity profile. On the equator advection over a temperature gradient is also used by the 4DVAR system.

Improvements from heat fluxes occur both in the tropics and in the extratropics. Especially large scale errors are nicely reduced by the scheme. The heat fluxes can not reduce small-scale errors, for example near western boundary currents. This is partly due to the high decorrelation lengths. Instead of shifting the thermocline, heat fluxes mainly act through the surface.

The analysis at the end of the last cycle can be used as initial state for seasonal forecasts. The subsurface structure of this analysis is investigated in some more detail. In both experiments, the final thermal subsurface ocean state is improved in the tropics as well as at higher latitudes. Most of the improvements are visible in the Pacific and in the equatorial region. Contrary to the 4DVAR scheme of Weaver et al. [2002], the flux adjustments can not correct the ocean state below the thermocline. Deviations due to small-scale chaotic ocean behaviour also have to be reduced by additionally adjusting the initial state of the analysis. In some areas the salinity analysis is also improved. This is due to freshwater fluxes, and wind updates affecting both temperature and salinity.

Acknowledgements We thank CNES/NASA and DEOS for the use of the TOPEX/POSEIDON data. Also acknowledgement is made for the use of ECMWF computing and archiving facilities in this research.

A An along-track altimetry data error estimate

Along-track altimetry data was assimilated rather than gridded values, as the error covariance structure is much easier to estimate. The raw 1Hz (7km) data were obtained from DEOS [Schrama et al., 2000] as a set of values per (nominal) obervation

point, one for each cycle (data up to cycle 234 were used). Bad tracks had been removed, but the following quality control cuts were necesary:

1. Observation points with less than 80% data were rejected. This eliminates regions that are sometimes covered with ice and coastal points that are dry at low tide, in all 9.75%.
2. Measurements at an observation point that deviate more than four times the standard deviation with respect to a 110-day running mean at that point were rejected. These 65923 (0.06%) observations are rather uniformly distributed, although the less stringent cut in areas of high variability means fewer are rejected there.
3. The standard deviation with respect to a 110-day running mean was required to be less than 52cm, the highest value found in the Agulhas retroflection area. This eliminates 74527 observations (0.07%), mainly in estuaria where the tide model is not accurate enough, but some in mid-ocean.

A visual inspection of the remaining anomalies show that the the largest remaining problem is the geoid gradient. The sea surface height is not always measured at the nominal observation position, but can deviate a few kilometers, mainly along-track, but also cross-track. In regions with large geoid gradients, such as the islands in the West Pacific and Caribean, these offsets give rise to sizable spurious anomalies. The along-track component shows up as large anti-correlated spurious signals on either side of an island or seamount, up to 30cm. These could be removed by estimation of the geoid from the data [Dorandeu et al., 2003], however this requires more information than we had available.

The data are used in Eq. 3 together with an estimate of the error covariance matrix $\mathsf{E}_{\mathrm{SSH}}$. This matrix is approximated by a block diagonal structure

$$\mathsf{E} = \left(\begin{array}{c|c|c|c} \ddots & \vdots & \vdots & \\ \hline \cdots & \mathsf{E_n} & 0 & \cdots \\ \hline \cdots & 0 & \mathsf{E_{n+1}} & \cdots \\ \hline & \vdots & \vdots & \ddots \end{array}\right) \tag{6}$$

The boundaries between the blocks are halfway between the latitudes of the HOPE grid points (around 20 points when treating the odd and even grid separately). All correlations between different blocks are neglected.

Within each block an effective sea surface height $S(t) = \sum_i \alpha_i s_i(t)$ is defined. The $s_i(t)$ are the measurements and $\sum_i \alpha_i = 1$. The weights α_i are determined by minimising the error

$$(\delta S)^2 = \sum_{i,j=0}^{N-1} (s_i(t) - S(t))\mathsf{C}_{ij}(s_j(t) - S(t)) \ , \tag{7}$$

with $(\mathsf{C}^{-1})_{ij} = \overline{s_i' s_j'}$ the covariance matrix of the measurements with a running mean subtracted, which is used as an approximation for the unknown error covariance matrix. This means that all sub-gridscale variability is considered noise.

The minimisation gives an effective sea surface height for each box

$$S(t) = \sum_{i,j=0}^{N-1} s_i(t)\mathsf{C}_{ij} / \sum_{i,j=0}^{N-1} \mathsf{C}_{ij} \ . \tag{8}$$

and an estimate of the error of this effective measurement:

$$(\delta S)^2 = \sum_{i,j=0}^{N-1} \alpha_i \alpha_j \overline{s'_i s'_j}. \tag{9}$$

The position of the effective measurement is $(X, Y)(t) = \sum_i \alpha_i (x, y)_i(t)$. The block diagonal assumption means that the effective measurements are assumed to be uncorrelated and can be used directly in Eq. 4.

A few practical consequences of this method are noted. In a first approximation, neglecting off-diagonal elements of C, it weighs the points according to their variance. Points with large errors are simply given less weight.

The covariance also solves the along-track geoid gradient problem. Two highly negatively correlated measurements are averaged to give a much more accurate estimate of SSH. Suppose $N = 2$ and

$$\overline{s'_i s'_j} = \begin{pmatrix} 1 & \delta - 1 \\ \delta - 1 & 1 \end{pmatrix}, \tag{10}$$

then $\alpha_1 = \alpha_2 = 1/2$ and $S(t) = (s_1(t) + s_2(t))/2$, $(\delta S)^2 = \delta/2 \ll (\delta s_1)^2$.

Finally, highly correlated measurements simply count as a single measurement. Again, for $N = 2$:

$$\overline{s'_i s'_j} = \begin{pmatrix} 1 & 1 - \delta \\ 1 - \delta & 1 \end{pmatrix}, \tag{11}$$

gives again $\alpha_1 = \alpha_2 = 1/2$ and hence $S(t) = (s_1(t) + s_2(t))/2$, but $(\delta S)^2 = 1 - \delta/2 \approx (\delta s_1)^2$.

A problem is immediately visible: measurements are often highly correlated, even after subtraction of a running mean. This leads to an inverse that is very badly conditioned, giving rise to large eigenvalues α_i that cancel each other. In order to regularize the solution one has to perform a singular value decomposition and reject eigenvalues until all α_i have a reasonable magnitude; we choose the condition $\alpha_i > 0 \, \forall i$.

References

Alves, J. O. S., M. A. Balmaseda, D. L. T. Anderson, and T. N. Stockdale: 2002, Sensitivity of dynamical seasonal forecasts to ocean initial conditions. Technical Report 369, ECMWF, Reading, U.K., 26 pp.

Alves, J. O. S., K. Haines, and D. L. T. Anderson: 2000, Sea level assimilation experiments in the tropical Pacific. *J. Phys. Oceanogr.*, **31**, 305–323.

Bonekamp, H., G. J. van Oldenborgh, and G. Burgers: 2001, Variational assimilation of TAO and XBT data in the HOPE OGCM, adjusting the surface fluxes in the tropical ocean. *J. Geophys. Res.*, **C106**, 16693–16709.

Burgers, G., M. A. Balmaseda, F. C. Vossepoel, G. J. van Oldenborgh, and P. J. van Leeuwen: 2002, Balanced ocean-data assimilation near the equator. *J. Phys. Oceanogr.*, **32**, 2509–2529.

Dorandeu, J., M. Ablain, and P. Le Traon: 2003, Reducing cross-track geoid gradient errors around topex/poseidon and jason-1 nominal tracks. application to calculation of sea level anomalies. *J. Atmos. Oceanic Technol.*, to appear.

Gilbert, J. and C. C. Lemarechal: 1989, Some numerical experiments with variable-storage quasi-newton algorithms. *J. Atmos. Oceanic Technol.*, **45**, 407–435.

Gill, A. E.: 1982, *Atmosphere–Ocean Dynamics*. Academic Press, Orlando, 662 pp.

Reynolds, R. W. and T. M. Smith: 1994, Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.

Schrama, E. J. O., R. Scharroo, and M. Naeije: 2000, Radar Altimeter Database System (RADS): Towards a generic multi-satellite altimeter database system. Technical Report 00-11, Remote Sensing Board (BCRS), Delft, Netherlands.

Segschneider, J., D. L. T. Anderson, J. Vialard, M. Balmaseda, T. N. Stockdale, A. Troccoli, and K. Haines: 2001, Initialization of seasonal forecasts assimilating sea level and temperature observations. *J. Climate*, **14**, 4292–4307.

Troccoli, A. and K. Haines: 1999, Use of the temperature-salinity relation in a data assimilation context. *J. Atmos. Oceanic Technol.*, **16**, 2011–2025.

van Oldenborgh, G. J., G. Burgers, S. Venzke, C. Eckert, and R. Giering: 1999, Tracking down the ENSO delayed oscillator with an adjoint OGCM. *Mon. Wea. Rev.*, **127**, 1477–1496.

Vossepoel, F. C.: 2002. Personal communication.

Vossepoel, F. C., G. Burgers, and P. van Leeuwen: 2002, Effects of correcting salinity with altimeter measurements in an equatorial pacific ocean model. *J. Geophys. Res.*, **107**, 10.1029/2001JC000816.

Weaver, A., J. Vialard, D. L. T. Anderson, and P. Delecluse: 2002, Three- and four-dimensional variational assimilation with a general circulation model of the tropical pacific ocean. Technical Report 365, ECMWF, 74 pp.

Wolff, J.-O., E. Maier-Reimer, and S. Legutke: 1997, The Hamburg Ocean Primitive Equation model HOPE. Technical Report No. 13, Deutsches Klimarechenzentrum, Bundesstr. 55, D-20146 Hamburg, Germany, Hamburg.