JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. C1, PAGES 1393-1404, JANUARY 15, 1999

The time-mean circulation in the Agulhas region determined with the ensemble smoother

Peter Jan Van Leeuwen

Institute for Marine and Atmospheric Research Utrecht, Utrecht, Netherlands

Abstract. The time-mean circulation in the Agulhas Retroflection area is determined by combining TOPEX/POSEIDON data and a two-layer quasi-geostrophic model using the ensemble smoother. By taking the time-mean circulation as the unknown in the data assimilation procedure, the time-varying altimeter signal is used to constrain the time-mean field. The quasi-geostrophic model is applied as a strong constraint, with only the time-mean circulation containing errors. Inspection of the posterior penalty function showed that the inversion was successful. The errors in the time-mean sea surface topography reduced from about 10 to about 3 cm. A cyclonic recirculation cell over the Agulhas Plateau was found, related to the northward meander of the Agulhas Return Current. Another cyclonic recirculation cell was found west of Africa, probably related to the passage of anticyclonic Agulhas Rings south of it. The new field is compared with advanced very high resolution radiometer infrared satellite data, confirming the northward meander of the Agulhas Return Current.

1. Introduction

Satellite observations are becoming more and more important for our understanding and monitoring of the world oceans. A serious drawback of this kind of observations is the fact that they measure only at the surface of the ocean. Satellite infrared images measure the skin temperature, while optical sensors can reach depths of about 50 m in clear-water conditions. So these measurements contain only information of the mixed layer and the air- sea interaction. Satellite altimeters and also synthetic apperture radar (SAR) images contain information on the shape of the water surface, so they measure an integrated quantity. This is the reason why altimetry, in particular, has been used to study mesoscale ocean dynamics.

A serious problem with altimetry is that the shape of the ocean surface as dictated by gravity is not known with high accuracy. So although the measurements themselves have an accuracy of 2 to 5 cm, it is impossible to distinguish the geoid from the time-mean circulation signal [*Chelton,* 1988; *Nerem et al.,* 1990;*Fu et al.,* 1996]. Of course, the geoid is to a large extent time invariant, so the time-varying part of the altimeter signal can be used to observe the mesoscale activity of the ocean very accurately [e.g. *Cheney et al.,* 1993; *Wakker et al.,* 1990; *Tai and White,* 1988; *Shum et al.,* 1990;*Gordon and Haxby,* 1990; *Feron et al.,* 1992].

Paper number 1998JC900012.
0148-0227/99/1998JC900012$09.00

The time-mean circulation cannot be directly determined with accuracy. Clearly, this is an important quantity, to study quantitatively, for instance, transports such as interocean exchanges or ring-shedding processes (meanders cannot be separated from eddies). To solve these problems, different methods can be used.

One way to improve interpretation is application of statistical techniques that analyze decorrelation in time and space. For instance, *Feron et al.,* [1992] were able to determine the instances of ring shedding from the time-varying signal only. Another approach is to combine the altimeter signal with other independent observations of the circulation. *Vazquez et al.,* [1990] combined Geosat data with observations of sea surface temperature in the Gulf Stream area. Other studies [e.g. *Tai and White,* 1988; *Willebrand et al.,* 1990;*Gordon et al.,* 1990; *Ichikawa and Imawaki,* 1994] combined hydrographic data and/or surface drifter buoys with Geosat data to study rings detached from the mean flow.

Two methods are presently in use to estimate the long-wavelength part of the time-mean circulation from altimeter observations. The first method simultaneously adjusts the gravity field and this circulation. This so-called integrated least squares approach is successful for wavelengths larger than roughly 1400 km by using a priori information on the time-mean circulation from other measurements [e.g. *Marsh et al.,* 1990; *Engelis and Knudsen,* 1989;*Nerem et al.,* 1990; *Visser,* 1992]. In the second method the geoid is determined more accurately from the orbit of the satellite. The geoid reflects the shape of the gravity field, which influences the orbit of the satellite. The very high precision of the height measurements together with the very accurately known satellite orbit allow the geoid to be found [e.g., *Minster et al.*, 1993; *Naeije et al.*, 1993; *Fu et al.*, 1996]. Both methods still lack accuracy for wavelengths shorter than about 1400 km.

One can also try to determine the mean flow directly from the altimeter signal by assuming a certain shape of the current. *Kelly and Gille*, [1990] and *Qiu et al.* [1991] investigated the mean flow in the Gulf Stream and Kuroshio Extension respectively by fitting a synthetic current's height profile to Geosat residual height data along individual tracks. Their results were found to be in remarkable agreement with hydrographic and acoustic Doppler current profiles [e.g., *Kelly et al.*, 1991; *Teague et al.*, 1990]. More recently, *Gille*, [1994] used the same technique for mapping the Antarctic Circumpolar Current. Both the Sub-Antarctic front and the Polar front could be determined and analyzed from Geosat altimeter data.

All of the methods given above lack a very valuable piece of information: the dynamics of the water. Obviously, the time-mean and the time-varying parts of the ocean circulation are dynamically coupled. Recently, *Feron et al.*, [1998] used this dynamical coupling. Time averaging of the potential vorticity equation for the upper ocean layer, in which stretching was shown to be negligible, results in a differential equation for the averaged relative vorticity in which the mean divergence of the eddy vorticity fluxes acts as a source or sink. The essential part is that these eddy fluxes can be determined from the altimeter observations. Consequently, no parameterizations appear in the averaged vorticity equation. From the average vorticity field surface geostrophic velocities and related mean dynamic sea surface topography can then be simply derived. The utility of the method is established using 'perfect' data, namely numerical output from the U.K. Fine- Resolution Antarctic Model. The method appears applicable to areas of the ocean with strong enough mesoscale variability such as the major western boundary currents and their extensions and frontal regions of the Antarctic Circumpolar Current. Quite realistic results were reported for active regions of the world ocean, using comparisons with hydrographic observations. Also, an error estimate could be attached to the new time-mean circulation, so it can be used for quantitative studies. *Feron et al.*, [1998] showed that the repeated shedding of large rings can now be synoptically reconstructed as a continuous process, by combining the new time-mean signal with the time-varying altimeter signal.

In the present paper this inversion technique is developed further by using the time-mean circulation as the unknown in a data assimilation experiment. By using a so-called smoother, the time-mean circulation can be also found in data sparse regions or in regions with low variability, which is not possible in the averaging technique of *Feron et al.*, [1998].

The data assimilation method chosen is the ensemble smoother [see *Van Leeuwen and Evensen*, 1996]]. It has been applied successfully in the Agulhas area to study the shedding of large Agulhas Rings that transport heat, salt, vorticity, and energy from the Indian to the south Atlantic Ocean as part of the Indian South Atlantic supergyre [*Van Ballegooijen et al.*, 1994; *Byrne et al.*, 1995]. In the study by *Van Leeuwen*, [1998] the time-mean circulation of the Gordon Southern Ocean Atlas [*Gordon*, 1982] was added to TOPEX/-POSEIDON gridded altimeter fields, and this combination was assimilated in a two-layer quasi-geostrophic model.

In the ensemble smoother a Bayesian view is taken in which the prior probability density of the model and the probability density of the data are combined to form a posterior density. The mean and the covariance of this density give the optimal model evolution and its errors. The advantage of this smoother over all others is that no adjoint equations need to be integrated and error estimates are easily obtained.

This paper is organized as follows. First, the data assimilation method is described in some detail. Then the data treatment is explained and the model is outlined. In section 4, results from the inversion are given and a comparison is made with infrared satellite images. A summary with conclusions closes the paper.

2. The Ensemble Smoother

The determination of the generalized inverse can be considered as the estimation of the unknown true model variables ψ, given the data and the model estimates, with information about their prior error statistics. Intuitively, it is clear that the probability density of the model and the probability density of the data contain all information needed to calculate the inverse estimate. Using Bayesian statisticsi, one can consider the probability density of the model forecast as prior information, which is 'updated' by the data. This results in a new probability density of the model, given the data. The procedure is described in detail by *Van Leeuwen*, [1998]. In the context of time-independent problems this idea has been used by *Tarantola*, [1987] (but slightly modified) and *Lorenc*, [1988].

The unknown state vector ψ is viewed as the value of a random variable $\underline{\psi}$, and the probability density of the data $\boldsymbol{d}$ is interpreted as the conditional probability density $f(\boldsymbol{d}|\psi)$ of $\underline{\boldsymbol{d}}$ assuming $\underline{\psi} = \psi$. The model with its error estimates is regarded as a priori information, and it is used to assign a density $f(\psi)$ to the random variable $\underline{\psi}$. The new probability density of ψ, given the data, is given by

$$f(\psi|\boldsymbol{d}) = \frac{f(\boldsymbol{d}|\psi)f(\psi)}{\int f(\boldsymbol{d}|\psi)f(\psi)d\psi} \tag{1}$$

To use this concept, the probability density of the data $f(\boldsymbol{d}|\psi)$ has to be determined. Usually, it is assumed to be known, for instance, a Gaussian.

More problematic is the prior probability density $f(\psi)$ of the model evolution. In general, the probability density for the model state has a huge number of variables, so it is not feasible computationally to determine its evolution for real oceanographic or meteorological applications. One alternative is to determine it from an ensemble calculation, but to construct the density from the ensemble members is again infeasible. However, one is not generally interested in $f(\psi|\boldsymbol{d})$ as a whole, only in its first few moments e.g., a best estimator of the truth and its error variance. In that case, ensemble or Monte Carlo experiments can be extremely useful.

The estimates can be obtained in a number of ways. The optimal estimator is the maximum-likelihood estimator. To make use of it, the exact shape of the joint probability density of model evolution and data has to be known. In simulated annealing and related methods this probability density is approximately generated. However, these methods require an unreasonable amount of storage and many iterations due to the random nature of the probing.

The most frequently used estimator is the minimum-variance estimator. It is well known that for Gaussian-distributed variables, the minimum-variance estimate is equal to the maximum-likelihood estimate. The ensemble smoother is a minimum-variance estimator, which is given by

$$\hat{\psi} = \int \psi f(\psi|\boldsymbol{d}) d\psi \tag{2}$$

Now use the expression for $f(\psi|\boldsymbol{d})$ from (1) to find

$$\psi = \frac{\int \psi f(\boldsymbol{d}|\psi) f(\psi) d\psi}{\int f(\boldsymbol{d}|\psi) f(\psi) d\psi}. \tag{3}$$

The probability density of the a priori model integration is partly determined by its mean and its covariance. For linear dynamics, assuming Gaussian statistics, these two completely determine the density, and the variance minimizing solution is easily found. For nonlinear dynamics the probability density may be separated in a Gaussian part ('G'), with the mean and covariance of the whole density, and a non-Gaussian part ('N'), describing the deviation from a Gaussian density. Thus the prior model density distribution becomes

$$f(\psi) = f_G(\psi) f_N(\psi). \tag{4}$$

Van Leeuwen, [1998] (see also *Van Leeuwen and Evcensen,* [1996]) showns that this leads to the following expression for the optimal model state:

$$\begin{aligned} \hat{\psi} &= \psi_F + \boldsymbol{r}^T \boldsymbol{b} \\ &\quad -A \left[\boldsymbol{r}^T (\boldsymbol{R} + \boldsymbol{w}^{-1})^{-1} \boldsymbol{r} - Q_{\psi\psi}\right] \\ &\quad \cdot \int \frac{\delta f_N(\psi)}{\delta \psi} f(\boldsymbol{d}|\psi) f_G(\psi) d\psi, \end{aligned} \tag{5}$$

in which $Q_{\psi\psi}$ is the model error covariance. The $\cdot$ denotes integration over space and time. The representer matrix $\boldsymbol{R}$ is the measurement-measurement covariance,

$$\boldsymbol{R} = E\left[\boldsymbol{\mathcal{L}}\left[\psi - \psi_F\right] \boldsymbol{\mathcal{L}}\left[\psi - \psi_F\right]\right] = \boldsymbol{\mathcal{L}}^T[\boldsymbol{r}]. \tag{6}$$

In (6) $E[..]$ is the expectation operator and $\boldsymbol{\mathcal{L}}$ is the measurement operator; $\boldsymbol{b}$ are the well-known representer coefficients [*Bennett,* 1992], that can be determined from the data $\boldsymbol{d}$, with error covariance $\boldsymbol{w}^{-1}$, as

$$(\boldsymbol{R} + \boldsymbol{w}^{-1}) \boldsymbol{b} = \boldsymbol{d} - \boldsymbol{\mathcal{L}}[\psi_F], \tag{7}$$

The representers $\boldsymbol{r}$, which are the model field-measurement covariances, are given by

$$\boldsymbol{r} = E\left[(\psi - \psi_F) \boldsymbol{\mathcal{L}}\left[\psi - \psi_F\right]\right] = \boldsymbol{\mathcal{L}}[Q_{\psi\psi}] \tag{8}$$

The field ψ_F is the mean of the a priori distribution, and so the first-guess field. Note that for nonlinear dynamics this is not the central forecast, which has no statistical meaning in that case.

If the model was linear and the first-guess field was initially distributed in a Gaussian manner, the first-guess evolution would stay Gaussian distributed. In that case the term with $f_N(\psi)$ would vanish and the well-known representer expression would appear.

In the ensemble smoother the non-Gaussian term is neglected, resulting in the first two terms in (5) being calculated using (7) and (8). This technique is similar to the methodology used in the Kalman filter. In the Kalman filter, only the first two moments are used to obtain a variance-minimizing estimate based on the current ensemble statistics. This estimate is calculated only at measurement times. In the ensemble smoother, however, the estimate is calculated over the entire spatial and temporal domain. Note that the ensemble smoother is as simple to calculate as the ensemble Kalman Filter [see *Evensen,* 1994] and does not require the storage of any fields as a function of both space and time. The computational load is either the same or twice that of the ensemble Kalman filter, depending on the fact that either the ensemble members are stored at diagnostic output times or the ensemble is recalculated. Details for construction of the smoother solution using ensemble statistics are given the appendix.

3. Data Assimilation Experiment

3.1. Observations

The standard corrections were applied to the TOPEX/-POSEIDON data (solid Earth and ocean tides, dry and wet troposphere, inverse barometer effect, and sea state bias as 2% of the significant wave height). Because of the excellent satellite orbits, I decided to apply no orbit error corrections. The geoid is not known very precisely, so only the time-varying part of the altimeter signal is used. The altimeter data were interpolated to the $1^o x 1^o$ grid in Figure 1, using the Gauss-Markov method, every 10 days for the period October to December 1992. These data fields are assimilated into the model every 10 days. *Tsaoussi and Koblinsky,* [1994] estimate the error in the time-varying part of the raw altimeter data

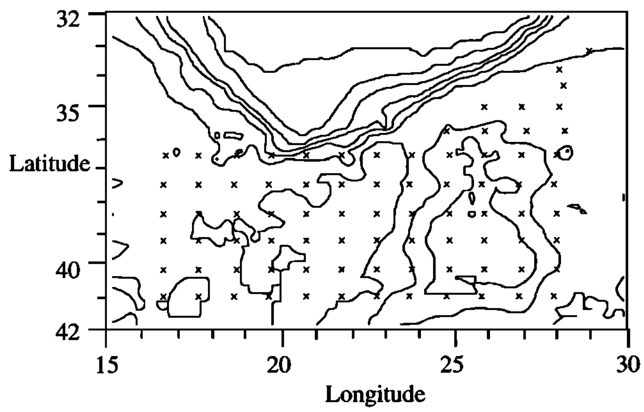

Figure 1. Bottom topography. Contour interval is 1000 m. The 100m contour is also drawn. Crosses denote 1^o interpolated values of TOPEX/POSEIDON.

to be 3 cm rms in the Agulhas area. The spatial and temporal smoothing will reduce this error even further. I estimated the error in each measurement at a rather low value of 1 cm. Error covariances between different measurements are neglected, mainly due to the large distance between the measurement points. Section 4 shows that the errors were probably correctly estimated.

One may argue that the data have to be assimilated as purely as possible, i.e., the along-track data as soon as they are available. The disadvantage of this procedure may be that data errors propagate in the model as Rossby waves, which would be difficult to remove with only along-track data. This effect can be more severe in a primitive equation model, where spurious gravity waves will disturb the optimal solution. So I therefore used the gridded fields, but arguments can be made for either approach.

3.2. Model

The ocean model used is the quasi-geostrophic model of *Holland et al.,* [1991] and consists of two layers in the vertical of 1 and 4 km deep and 25 km resolution in the horizontal. The density difference between the layers is 2 kg m^{-3}. The time step is 1 hour. Small-scale noise is removed with a Shapiro filter of order 8. This model describes the ocean circulation around the southern tip of Africa from $32^oS - 42^oS$ and $15^oE - 30^oE$. It has open boundaries to account for the inflow of the Agulhas Current, the inflow of the Antarctic Circumpolar Current, the outflow of the Agulhas Return Current, and the westward migration of Agulhas Rings.

The African continent is extended to the 100 m depth contour to lessen the influence of coastal shelf dynamics, which are not described well by quasi-geostrophic dynamics. The derivative of the stream function along the boundary is kept zero (free slip) by applying the modified capacitance matrix method as described by *Milniff,* [1990]. The influence of the bottom topography is reduced by a factor 0.2 to compensate for the fact that the influence of bottom topography is overestimated in the dynamics of a quasi-geostrophic model.

The open boundaries are, of course, problematic. A radiation condition is chosen, so baroclinic waves are propagated out of the domain. This means that no small-scale disturbances enter the domain. Experiments showed that mesoscale features leave the domain without velocity changes and without spurious wave reflection (see, for instance, *Evensen,* [1996]. In this study I assume no errors in the model dynamics. This is, of course, not very realistic; it is well known that a quasi-geostrophic model cannot describe the shedding of Agulhas Rings correctly. However, the error in the unknown time-mean field is assumed to be much larger, so the main correction will be in the mean field. (Note that incorporating these errors in the model dynamics is just as expensive.)

3.3. Initial Conditions

As stated in the previous section, the state vector consists of the time-mean circulation and therefore a two-layer stream function field. Our first-guess time-mean field is calculated from the *Gordon,* [1982] data set. To generate the ensemble, an ensemble of random fields is produced and added to each time-mean field. The random fields have a covariance given by

$$C(x,y) = A \exp\left[-\frac{(x-y)^2}{2d^2}\right] \tag{9}$$

in which d is a characteristic length scale for the mean field. I choose this scale to be 4 times the Rossby deformation radius, which is about 45 km for the chosen density difference. So $d = 180$ km.

The variances in the mean field are $A = 10^8 \mathrm{m}^2\,\mathrm{s}^{-1}$ and for the lower layer $A = 10^6 \mathrm{m}^2\,\mathrm{s}^{-1}$. These values are about 10% of the initial values and correspond to errors in initial sea level of about 10 cm. This may seem large, but a small shift in the currents will give rise to this order of magnitude. Clearly, this is not true for relatively quiet areas, so a position-dependent initial error will be a better description of the error field. However, there is no way to determine beforehand how the major current and the eddies will move. Hence I assumed position-independent errors.

It is difficult to generate an ensemble with a known probability density of time-mean states. This is due to the fact that each ensemble member will evolve differently over time, leading to a different mean (i.e., time averaged) state that cannot be anticipated beforehand. So starting off with a Gaussian distribution of the initial field does not mean that the distribution in mean fields will stay Gaussian during the ensemble run. Note that data do not play a role yet; we are just trying to obtain a Gaussian ensemble for the mean fields. I dealt with this problem by using the different scales of eddy field and mean field. In each ensemble member the Gordon time-mean field plus the random field is taken as the time-mean field of that member. It turned out to be a very good guess; the time-averaged field after 100 days was not really different from the time-mean field we wanted to put in. The rms differences were less then 1% of the prior error covariances. Again, I think this is due to the difference in spatial scales between the eddy fields and the mean field. To test the degree to which the prior distribution was Gaussian in shape, the skewness was calculated as 10^{-4}, which is indeed a low value.

To each upper layer field an interpolated TOPEX/-POSEIDON altimeter field was added to account for the initial eddy field. That same TOPEX/POSEIDON field, now multiplied by 0.4, was added to the lower layer. The factor 0.4 is the correlation coefficient between upper and lower layers, determined from a 2-year model run. That factor is chosen also for the TOPEX/-POSEIDON data because the assumption is that the model is unbiased. The validity of that assumption is questionable but is outside the scope of this paper. Note that (nearly) all data assimilation methods start from this assumption. I stress that this factor is only used in the initial condition; later on, there is no need to use this factor.

3.4. Assimilation Statistics

The size of the ensemble was 500. Increasing this size to 600 did not change the model results significantly (rms changes less than 1% of the variance) , so it was concluded that 500 members was enough in this case.

The penalty function reduces from about 1,000,000 to 957. The number of measurements was 738, but the number of independent measurements determined from conditioning of the representer matrix was 163. For a linear model these numbers should be about the same (the penalty function is a χ^2 variable with the number of independent measurements as the number of degrees of freedom). For a nonlinear model I cannot expect this to be the case. The fact that they are of the same order of magnitude gives confidence in the a priori error estimates in model and data. It is interesting to note that in the experiment described by *Van Leeuwen*, [1998], in which the optimal evolution of the stream function fields was determined by the same method, the number of independent measurements was 209. The fact that I obtain a lower number in the present case is caused by the larger characteristic length scale in the error covariance used here.

The representer matrix, as defined in section 2, can be used to locate areas that have the greatest influence on the new time-mean field. These areas are prime candidates for observational efforts [see e.g. *Bennett*, 1992]. To this end, the eigenvectors with the largest eigenvalues are determined. The largest elements in these eigenvectors indicate the locations at which measurements would have the most impact. In Figure 2 the eigenvector elements are contoured for the first four eigenvectors, which explain more than 90% of the variance.

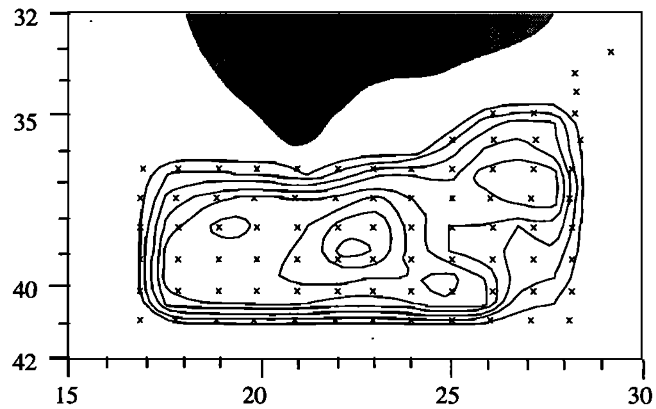

Figure 2. Eigenvector elements of the first four eigenvectors that explain more than 94% of the variance, showing the measurement positions that had the greatest influence on the final solution. These are the positions at which future measurements must be done.

A maximum can be distinguished in the middle of the domain, in the area just west of the Agulhas Plateau. This is the position where measurements of sea-surface height will most constrain the solution. The significance of this position is probably related to the northward meander in the Agulhas Return Current, leading to a time invariant recirculation cell over the Agulhas Plateau. I elaborate on this in the next section. Of course, altimeters have a much larger coverage area so the exercise is not that impressive for this experiment. However, the same could be done with in-situ measurements by calculating the representer matrix for a hypothetical cruise, and the economical impact will be tremendous!

4. Results

The upper stream function field as deduced from *Gordon*, [1982] is given in Figure 3. From now on this field is referred to as the prior field. The Agulhas Current can be detected along the east coast of South Africa. This is where the highest velocities are present, up to 3 m/s. Following this current southward, I find that it retroflects, i.e., turns back to the Indian Ocean as the Agulhas Return Current that leaves the domain at the eastward side. One can clearly identify a 'ring corridor', the preferred path of newly formed Agulhas Rings on the western side of the domain [see, e.g., *Garzoli and Gordon*, 1996; *Byrne et al.*, 1995]. Being anticyclonic, they form a 'ridge' of positive sea-surface elevations (negative stream function values). At the southern end of the domain the eastward flowing part of the Indian-South Atlantic supergyre is visible.

In Figure 3 the smoother solution, i.e. the new time-mean circulation, is also depicted. Note that this new field cannot be compared directly with the original field because the mean is determined over only 3 months time, while the original field is a multiyear average. (Strictly speaking, this is not true. It is based on hy-

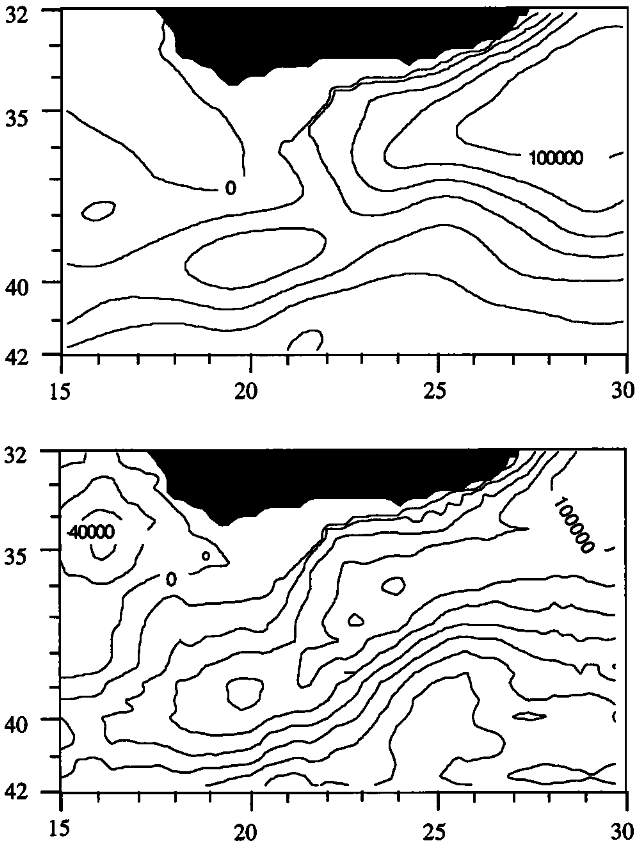

Figure 3. (top) Prior and (bottom) posterior time-mean upper layer stream function fields in square meters per second.

drographic measurements, which are biased toward non-winter conditions on the southern hemisphere. For reference, the altimeter data used here cover the period October to December 1992.) Many small-scale features appear, which are independent of ensemble size. The influence of the Agulhas Plateau is more pronounced than in the prior field. The retroflection is more westward than in the original field and seems to be split in two, one retroflection at 22°E and another at 18° E. Also an early retroflection more eastward at 28 E is present, with a recirculation cell centered at 24 E.

It is difficult to compare the three retroflections with other sources because this is only a 3-month average of October-December 1992. The analysis of Quartly and Srokosz, [1993] of advanced very high resolution radiometry data on the position of fronts in the retroflection area shows a mean position for September-November of 18.5° E and for December-January of 19° E. The time period of the analysis was March 1985 to February 1988, so 3 years. The new main retroflection at 18° E is in agreement with this, but it must be stressed that the mean position is variable from year to year. An early retroflection has been observed earlier [see *Lutjeharms and Van Ballegooyen,* 1988] and indeed seems to take place during this period [see *Van Leeuwen,* 1998].

The assimilated field differs from the prior field in three other features, as can be seen more clearly in Figure 4, showing the difference between the prior and the posterior fields. They are cyclonic recirculation cells southwest of Africa and at the outflow of the Agulhas Return Current and an anticyclonic cell in the middle of the area. Their formation can be explained as follows. Starting with the latter, this feature is due to the farther westward penetration of the Agulhas Current before retroflection. This results in a more pronounced northward meander of the Agulhas Return Current when it flows over the Agulhas Plateau. This meander can be understood by considering the balance between vortex stretching and the β effect: when the current feels a sloping bottom, vortex squeezing is balanced by a northward shift. On the other side of the plateau the reverse takes place. Result is a steady meander northward [see also *Belkin and Gordon, 1996*]. This meander will induce a cyclonic circulation in the loop, that is clearly evident in Figure 3.

The cyclonic circulation in the difference field (Figure 4) at the eastern boundary of the domain is probably due to two effects: the more northward directed return current at about 25° E due to the effect described above and a tighter recirculation cell in the northeast corner of the domain. Such recirculation cells are found close to the separation points of all western boundary currents, strongly increasing the transport of these currents above the Sverdrup values.

The cyclonic feature west of Africa is probably related to the passage of anticyclonic eddies south of it. The eddies will induce a westward flow on their northern side. The water will return eastward farther north, resulting in a recirculation cell. It is important to realize that the features are significant because the error in the stream function field is, at most, 10% of the maximum values of the features (see Figure 5). On the other hand, the feature in the prior field west of Africa is hardly significant.

We have to be careful with the interpretation of these results. In the 3 months investigated here, only one

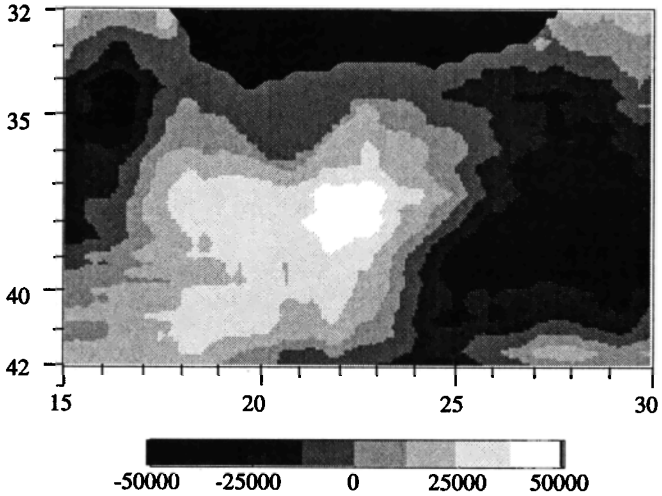

Figure 4. Difference between prior and posterior time-mean upper layer stream function fields in square meters per second.

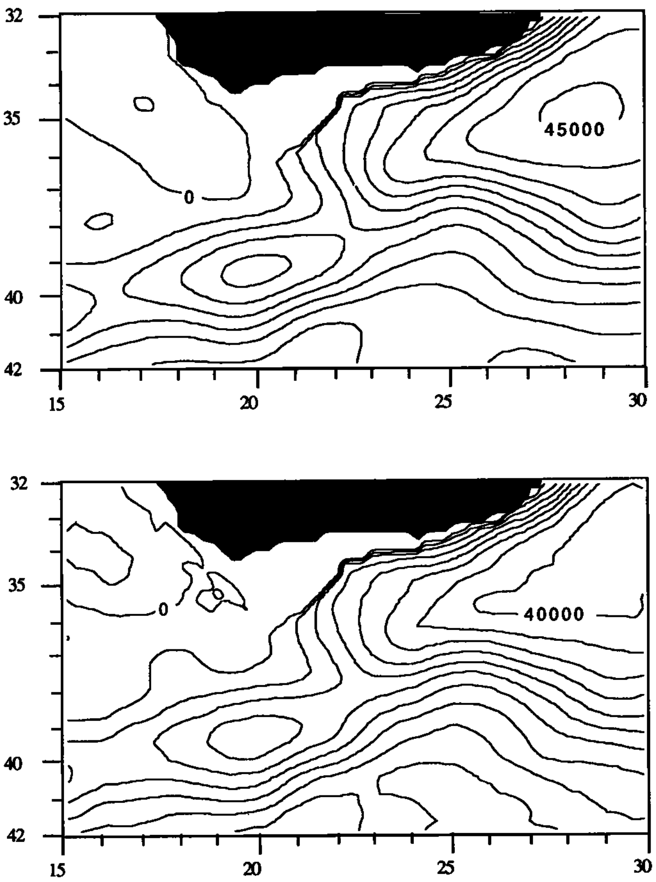

Figure 5. (top) Prior and (bottom) posterior time-mean lower layer stream function fields in square meters per second.

per layer. Figure 8 shows the same for the lower layer. In both figures the finiteness of the ensemble can be observed. I specified a space-independent variance for both the upper and lower layer prior variance fields, but both fields show structure. Increasing the ensemble size leads to a more homogeneous picture, owing to the fact that a larger ensemble is needed to describe a variance field than a mean field. Interestingly enough, the posterior error fields are not seriously affected by a larger ensemble.

Clearly, a large reduction in the uncertainty of the stream function values has been obtained. Over most of the area the variance has been reduced by a factor of about 10, so the errors are reduced by about a factor 3.5. Converting to sea level height errors, the original 10 cm uncertainty has been reduced to about 3 cm. So we have obtained a time-mean sea-surface elevation that has the same accuracy, on the large scale, as individual altimeter heights! It is stressed again that the errors are obtained with negligible effort in the ensemble smoother, because they can easily be calculated from the updated ensemble. This is contrary to all other data assimilation methods. Obviously, a time-mean field without a proper error estimate is useless.

In Figure 8 we see that although the time-mean circulation in the lower field has not changed much, the associated error has. A reduction in variance of a factor 5, so, there is an error reduction of about a factor 2. This may seem strange, but it is not. It just means that the first guess of the model compares well with the guess associated with the data. However, the data constrain the probability density and force it to become much narrower.

It is interesting to compare the present results with an independent data set. In Figure 9 the mean temperature for October to December 1992 is given. The temperature data are obtained from the monthly mean temperatures from infrared images of the *Vazquez et*

large ring was shed, which could, for, instance lead to a westward bias of the mean flow. On the other hand, five to six rings are being shed each year [see e.g. *Feron et al.*, 1992], so perhaps the 3 months form one complete cycle of ring growth and shedding. However, the highly variable altimeter signal in the retroflection area suggests that no two ringshedding events are the same.

In Figure 5 the lower layer is shown, both the original and the assimilated field. The difference, given in Figure 6, is less pronounced than in the upper layer fields (note the scales), but still some interesting features appear. Again, the cyclonic circulation cell southwest of Africa can be found. Also, a westward shift of the Agulhas Current at about 18° E is visible. Finally, the whole Agulhas Return Current is shifted northward over and northeast of the Agulhas Plateau. It is interesting to see that the lower layer seems to be less influenced by the Agulhas Plateau than the upper layer. This is probably related to an artifact of quasi-geostrophic dynamics. In order to reduce the unrealistically strong influence of bottom topography in the two-layer quasi-geostrophic model, the topography was scaled down with a factor 0.2. So the plateau, which reaches up to a depth of 2.5 km, was scaled down to a mountain of only 1 km.

Figure 7 shows what we have gained, the variance estimates before and after the assimilation for the up-

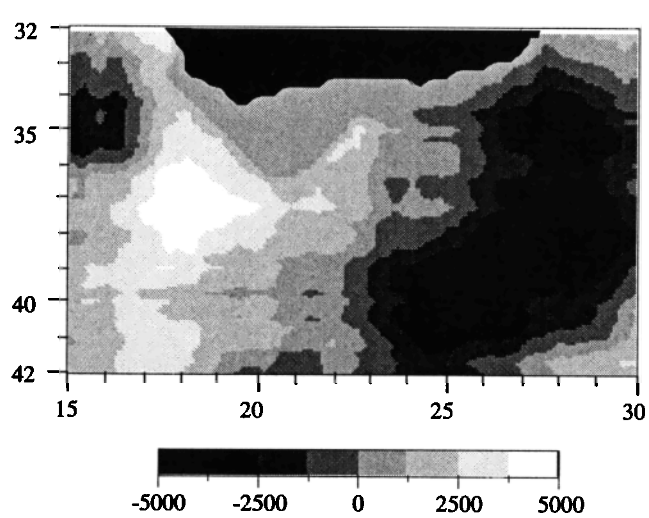

Figure 6. Difference between prior and posterior time-mean lower layer stream function fields in square meters per second.

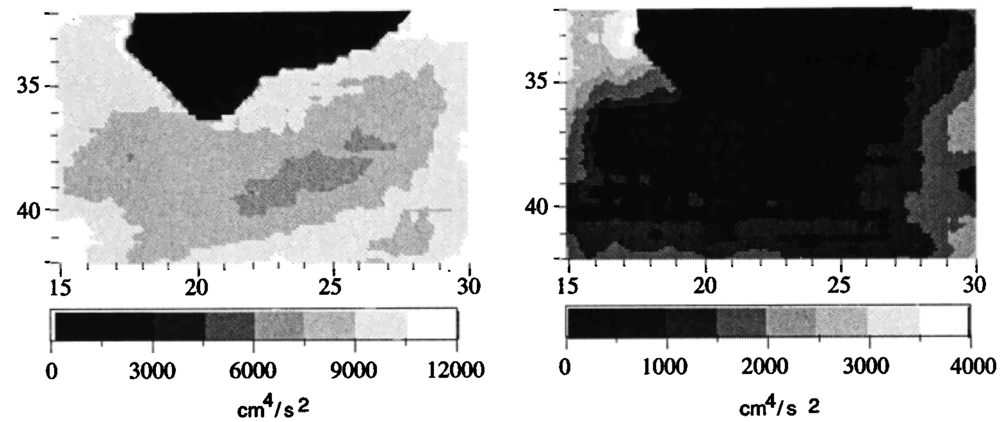

Figure 7. (left) Prior and (right) posterior error variances in the upper layer in cm^4/s^2. Note the difference in scale between the plots.

al., [1994]. The Figure 9 also shows the time-mean sea-surface elevation over the 3 month period of the new upper layer sea-surface height. This has been determined from the upper layer stream function using geostrophic balance. I imposed the 100-depth contour on the temperature image to make the comparison with the sea-surface height easier. The dark ribbon of the Agulhas indeed closely follows this contour, which gives us confidence in the chosen model domain.

The first thing that strikes the eye is the rapid loss of the temperature signal as the Agulhas flows southwestward, free from the continent. The temperature field does show the large northward meander of the Agulhas Return Current over the Agulhas Plateau, confirming this feature in the assimilated run. The recirculation cell cannot be viewed in sea-surface temperature because the upper layer loses its temperature signal extremely fast in this area. Note that the path of the return flow can best be obtained by the temperature gradient, which is a better measure than a temperature contour due to the excess heat transfer to the atmosphere. The longer a patch of water is exposed to the strong winds, the colder it will get (apart from upwelling and downwelling). Therefore it is expected that temperature gradients will last longer.

The temperature field suggests that Agulhas eddies leave the domain at a more northward position, but again, care has to be taken in the interpretation of the temperature field owing to the large air-sea exchange in this area (e.g. *Olson et al.,* [1992]). On the other hand, the assimilated field does show a more northward extension of the retroflection/eddy field southwest of Africa than the prior field, but this extension does not reach to the western boundary. To investigate this further, the TOPEX/POSEIDON gridded fields are added to the time-mean field in Figure 8. To obtain a complete picture, the original TOPEX/POSEIDON measurements are interpolated to the model grid on the complete model domain.

From Figure 10 it is clear that the Agulhas Rings do leave the domain more northward than the time-mean field suggests; but how can this be if the time-mean field is more or less the mean of the images of Figure 10? The answer to this question is twofold. First, the ring corridor of the time-mean field is more northward in the area where the data is assimilated; see Figure

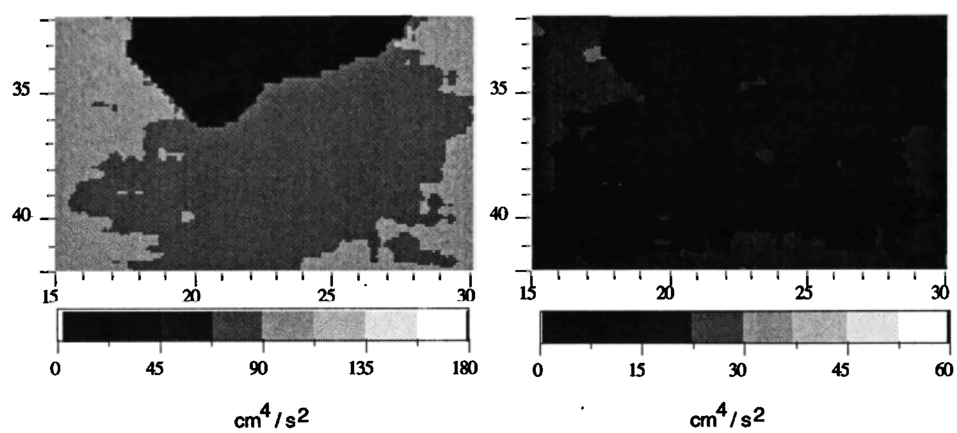

Figure 8. (left) Prior and (right) posterior error variances in the lower layer in cm^4/s^2. Note the difference in scale between the plots.

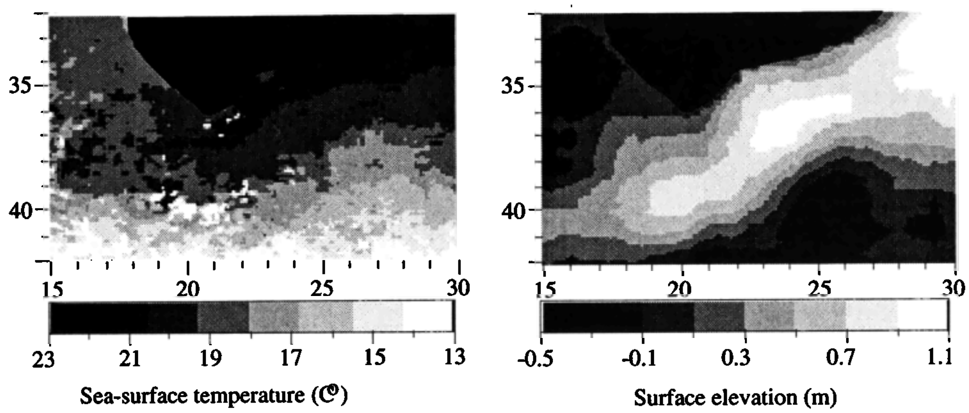

Figure 9. (left) Sea-surface temperature field, time average for October-December 1992, and (right) time-mean sea-surface topography determined from the posterior upper layer stream function.

1. West of this area the data have no direct influence, although one would expect a sea-level elevation there too. The reason that this does not happen is the initial condition. The prior field shows a southward displaced corridor, so a balance is achieved between data and this field, and the result is the assimilated field.

The second part of the answer lies in the model dynamics. The quasi-geostrophic model tends to form rings that move more southward than the data show [see *Van Leeuwen*, 1998]. This points to a bias in the model, while the data assimilation scheme assumes that no bias is present. This is an ingredient of all data assimilation methods, and it clearly needs more attention.

5. Summary and Discussion

The new data-assimilation method proposed by *Van Leeuwen and Evensen*, [1996], the ensemble smoother, is used to determine the time-mean circulation in the Agulhas Retroflection area. The time-varying part of TOPEX/POSEIDON measurements was combined with a two-layer quasi-geostrophic model, imposing the time-mean circulation as the unknown. So no time-mean field was added to the time-varying altimeter data before assimilation. The idea underlying the choice of assimilating only the time-varying part of the altimeter data is that the time-mean field and the time-varying fields are dynamically coupled through the (model) dynamics. (Note that this can only be done in this way by using a smoother.)

The basic idea in the ensemble smoother is to combine the probability densities of the model and the data in a Bayesian way to obtain the probability density of the model, given the data. The probability density of the time-mean model state, the prior distribution, is described by the time-averaged fields of an ensemble of possible model evolutions. Only 500 members are needed in this experiment. The density is assumed to be Gaussian, and inspection of higher-order moments showed that this was indeed the case. The probability density of the data is taken to be Gaussian, too. In the particular experiment described here the error reduction is about a factor of 3 overall. The penalty function reduced a factor of 1000 to a value close to the number of measurements, indicating that the a priori errors in model and data have been chosen consistently.

By studying the representer matrix, the locations that most constrain the solution were determined. This was found to be the northern edge of the Agulhas Plateau. Indeed, for the area covered by the data, this is where the largest changes were obtained.

The new time-mean circulation showed some interesting features. A cyclonic recirculation cell is found over the Agulhas Plateau. This recirculation cell is probably related to the steady northward meander of the Agulhas Return Current. Another recirculation cell appeared in the area west of Africa, probably due to the passage of Agulhas Rings south of it. The error estimates show that these features are indeed significant. One would expect a similar recirculation cell south of the eddy corridor. The fact that this cell is not found can have two reasons. First, the new time-mean field is probably not that reliable in that part due to the southward bias of the eddy corridor in the prior field and in the model. Second, the southern edge of the Indian-South Atlantic supergyre will blur the simple picture. The northward meander of the Agulhas Return Current over the Agulhas Plateau becomes more pronounced in the assimilated run compared to the prior. This is consistent with the 3-month mean sea-surface temperature obtained from the Pathfinder project.

A serious question is why the recirculation cells found in this study have not been found previously with great confidence. The first thing that comes to mind is that they are not visible with the conventional measurements. In infrared satellite images they are invisible because their temperature signal is the same as that of the surrounding water. In satellite altimetry they cannot be detected because only the time-varying signal has enough accuracy. Finally, in hydrographic measure-

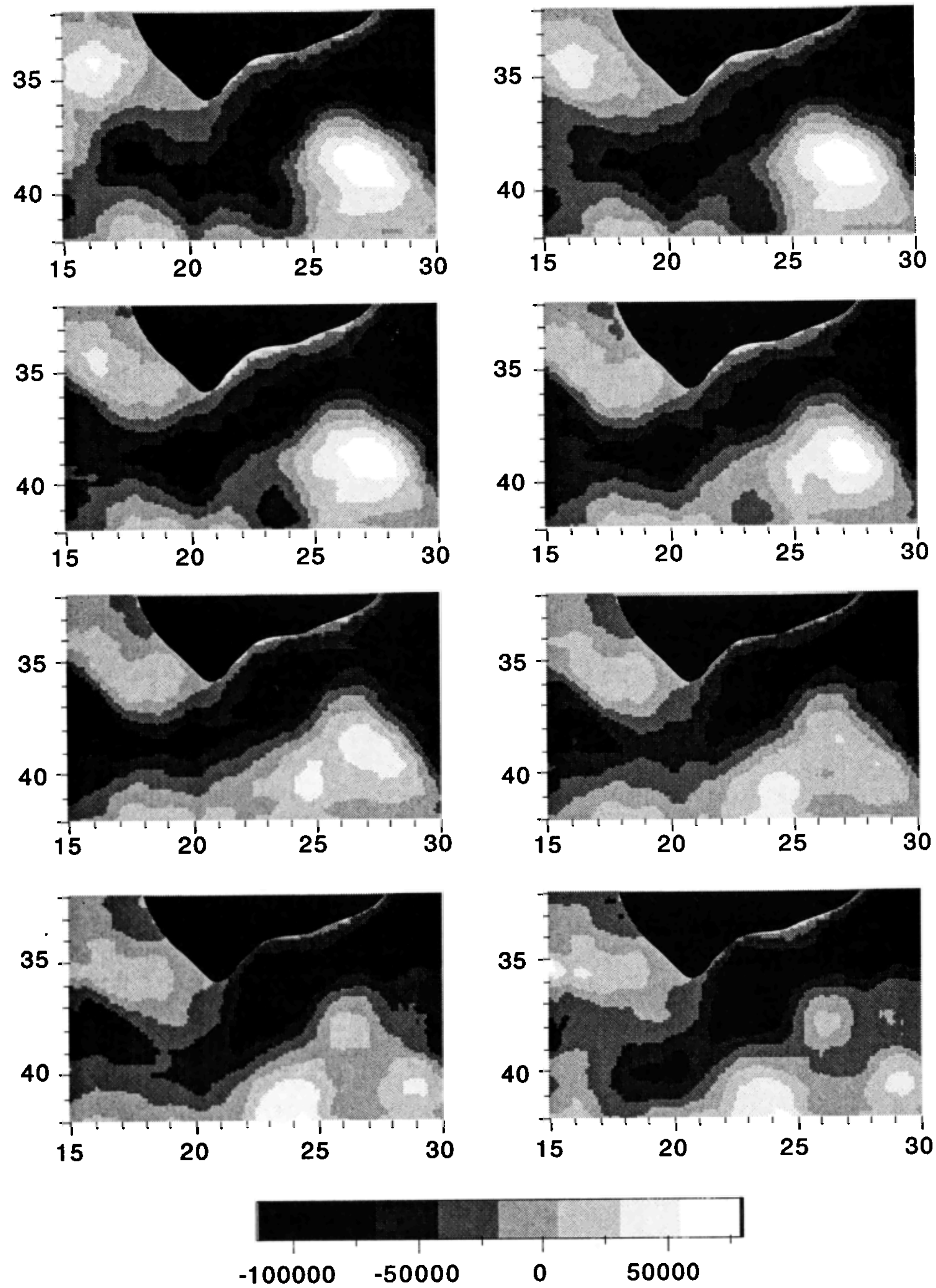

Figure 10. Sequence of TOPEX/POSEIDON data superimposed on the time-mean upper layer stream function field for days 10 to 80 with a 10 day interval.

ments they are difficult to find because they are strongly barotropic. The reason that they are more barotropic than the flow that induces them is probably the following. Imagine a two-layer flow in which the initial flow is mainly confined to the upper layer. This means that the interface displacement is such that it counteracts the sea-surface elevation, resulting in a negligible deep flow. Water parcels close to this flow will be dragged with the flow in the upper layer. Geostrophy will lead to a sea-surface elevation that supports this flow. There is no reason why the interface will move such as to compensate for the sea-surface elevation, and hence a deep flow will arise. Note that numerous numerical experiments exist that do show induced flows that are more barotropic than the initial flow; see, for instance, Drijfhout, [1990].

Another serious question is what the present results actually mean. Altimeter data suggest that no two ring-shedding event are similar, so a time-mean field based on only one shedding cannot be very accurate. Indeed, the present results will be biased to this particular event. This is one of the reasons why the error covariance of the prior field is chosen to have a large horizontal decorrelation length. The horizontal extension of the new mean field covers more or less the region of high variability found from altimetry, so it may not be that bad after all. Furthermore, seasonal variability is present in this area [e.g. Quartly and Srokosz, 1993], so maybe the new time-mean field is quite representative for the October to December period. This deserves more study.

Appendix: Implementation of the Ensemble Smoother

The ensemble of model states is integrated over the complete time interval of interest, but the prior error covariance is not calculated explicitly. This would be a huge computation, and the equations in section 3 show that only measurements of $Q_{\psi\psi}$ are needed. The actual algorithm can be illustrated as follows. (1) Integrate each ensemble member ψ in time and store only its measurements $\mathcal{L}(\psi)$. (2) Determine the mean measurement vector $\mathcal{L}(\psi_F)$ and calculate the representer matrix from

$$\boldsymbol{R} = \frac{\boldsymbol{S}\boldsymbol{S}^T}{n-1}. \qquad \text{(A1)}$$

with

$$\boldsymbol{S} = \mathcal{L}(\psi - \psi_F) \qquad \text{(A2)}$$

(3) Find the representer coefficients from

$$b_{ij} = \left(\boldsymbol{R}_{ik} + \boldsymbol{w}_{ik}^{-1}\right)^{-1} \left(d_{kj} - \mathcal{L}_k[\psi_j]\right) \qquad \text{(A3)}$$

in which d_kj is measurement K for ensemble J. Note that to each measurement a random error from a distribution with variance $\boldsymbol{w}$ has to be added to avoid a too strong correlation between the ensemble members after the update [see *Burgers et al.*, 1998]. (4) Recalculate the ensemble until the first wanted output time and store each member at that time. (5) Each member could be updated as

$$\psi_j^{\text{new}} = \psi_j^{\text{old}} + \sum_{i=1}^{M} r_i b_{ij} \qquad \text{(A4)}$$

but that would require the storage of the representers. Better is to update the moments of the new density (e.g. mean and variance) separately, so that only sums of the old members are needed. (6) Integrate the ensemble to the next time level of interest and repeat (5). (7) Repeat stage (6) until the end of the integration interval.

Note that one only has to store the ensemble at one time level, so the storage requirements are relatively mild. Clearly, this computation is embarrassingly parallel!

Acknowledgments. The author was sponsored by the Space Research Organization Netherlands (SRON) under grant EO-002 and by the National Research Program (NOP II), grant 013001237.10. All calculations were performed on the CRAY-C98 under contract SC-399, which was sponsored by the Stichting Nationale Computerfaciliteiten (National Computing Facilities Foundation (NCF)) for the use of supercomputer facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research, NWO). Marc Naeije from the Technical University Delft supplied the fully corrected TOPEX/POSEIDON data. The Pathfinder Ocean project team is thanked for the excellent AVHRR data set. Three anonymous reviewers are thanked for their constructive remarks.

References

Belkin, I. M. and A. L. Gordon, Southern ocean fronts from the Greenwich meridian to Tasmania, *J. Geophys. Res.*, *101*, 3675-3696, 1996.

Bennett, A. F., *Inverse Methods in Physical Oceanography*, Cambridge Univ. Press, New York, 1992.

Burgers, G., P. J. van Leeuwen, and G. Evensen, On the analysis scheme in the ensemble Kalman Filter, *Mon. Weather Rev.*, *126*, 1719-1724, 1998.

Byrne, D. A., A. L. Gordon, and W. F. Haxby, Agulhas eddies: A synoptic view using GEOSAT ERM data, *J. Phys. Oceanogr.*, *25*, 902-917, 1995.

Chelton, D. B., WOCE/NASA altimewter algoritm workshop, *U.S. WOCE tech. rep.*, *2*, U.S. Planning Office for the World Ocean Circulation Experiment, College Station, Tex, 1988.

Cheney, R. E., J. G. March, and B. D. Beckley, Global mesoscale variability from collinear tracks of SEASAT altimeter data, *J. Geophys. Res.*, *88*, 4343-4354, 1983.

Drijfhout, S. S., Ring genesis and the related transports of heat, momentum and vorticity: A parameter study, *J. Phys. Oceanogr.*, *20*, 1645-1665, 1990.

Engelis, T. and P. Knudsen, Orbit improvement and determination of the ocean geoid and topography from 17 days of SEASAT data, *Report Dep. of Geod. Sci. and Surv.*, Ohio State Univ., Columbus, 1989.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*, 10,143–10,162, 1994.

Evensen, G., and P. J. van Leeuwen, Assimilation of GEOSAT altimeter data for the Agulhas Current using the ensemble Kalman Filter with a quasi-geostrophic model, *Mon. Weather Rev.*, *124*, 87-96, 1996.

Feron, R. C. V., W. P. M. De Ruijter, and D. Oskam, Ring shedding in the Agulhas Current system, *J. Geophys. Res.*, *97*, 9467-9477, 1992.

Feron, R. C. V., W. P. M. De Ruijter, and P. J. Van Leeuwen, A new method to determine the mean sea surface topography from satellite altimeter observations, *J. Geophys. Res.*, *103*, 1343-1362, 1998.

Fu, L.-L., C. J. Koblinsky, J.-F. Minster, and J. Picaut, Reflecting on the first three years of TOPEX/POSEIDON, *Eos Trans. AGU*,*77*, 109, 111, 117, 1996.

Garzoli, S. L., and A. L Gordon, Origins and variability of the Benguela Current, *J. Geophys. Res.*, *101*, 897-906, 1996.

Gille, S. T., Mean sea-surface height of the Antarctic Circumpolar Current from GEOSAT data: Method and application, *J. Geophys. Res.*, *99*, 18,255-18,273, 1994.

Gordon, A. L., *Southern Ocean Atlas*, Columbia Univ. Press, New York, 1982.

Gordon, A. L., and W. F. Haxby, Agulhas eddies invade the South Atlantic: Evidence from GEOSAT altimeter and shipboard conductivity-temperature-depth survey, *J. Geophys. Res.*, *95*, 3117-3125, 1990.

Holland, W. R., V. Zlotnicki, and L.-L. Fu, Modelled time-dependent flow in the Agulhas Retroflection Region as deduced from altimeter data assimilation, *S. Afr. J. Mar. Sci.*, *10*, 407-427, 1991.

Ichikawa, K., and S. Imawaki, Life history of a cyclonic ring detached from the Kuroshio Extention as seen by the GEOSAT altimeter, *J. Geophys. Res.*, *99*, 15,953-15,966, 1994.

Kelly, K. A., and S. T. Gille, Gulf Stream surface transport and statistics at 69 W from the GEOSAT altimeter, *J. Geophys. Res.*, *95*, 3149-3161, 1990.

Kelly, K. A., T. M. Joyce, D. M. Schubert, and M. J. Caruso, The mean sea surface height and geoid along the GEOSAT subtrack from Bermuda to Cape Cod, *J. Geophys. Res.*, *96*, 12,699-12,709, 1991.

Lorenc, A. C., Optimal nonlinear objective analysis, *Q. J. R. Meteorol. Soc.*, *28*, 879–899, 1988.

Lutjeharms, J. R. E., and R. C. Van Ballegooyen, The retroflection of the Agulhas Current, *J. Phys. Oceanogr.*, *18*, 1570-1583, 1988.

Marsh, J. G., C. J. Koblinsky, F. Lerch, S. M. Klosko, J. W. Robbins, R. G. Williamson and G. B. Patel, Dynamic sea-surface topography, gravity and improved orbit accuracies from direct evaluation of SEASAT altimeter data, *J. Geophys. Res.*, *95*, 13,129-13,150, 1990.

Milnliff, R. F., Modified capacitance matrix method to implement coastal boundaries in the Harvard open ocean model, *Math. Comput. Simul.*, *31*, 541–564, 1990.

Minster, J. F., C. Brossier and P. Femenias, Global ocean dynamic topography and how it varies between October and December 1992, *AVISO altimetry newsl.*, vol. 2, pp. 22-23, AVISO, Toulouse, 1993.

Naeije, M. C., E. Wisse, R. Scharroo, and K. F. Wakker, Ocean dynamics from ERS-1 35-repeat mission, in *Proceedings of the second ERS-1 symposium*, ESA SP-361, 501-506, 1993.

Nerem, R. S., B. D. Tapley, and C. K. Shum, Determination of the ocean circulation using GEOSAT altimetry, *J. Geophys. Res.*, *95*, 3163-3179, 1990.

Olson, D. B., R. A. Fine, and A. L. Gordon, Convective modifications of water masses in the Agulhas *Deep-Sea Res., Part A*, *39S*, S163-S181, 1992.

Qiu, B., K. A. Kelly, and T. M. Joyce, Mean flow and variability in the Kuroshio Extension from GEOSAT altimetry data, *J. Geophys. Res.*, *96*, 18,491-18,507, 1991.

Quartly, G. D., and M. A. Srokosz, Seasonal variations in the region of the Agulhas Retroflection; studies with GEOSAT and FRAM, *J. Phys. Oceanogr.*, *23*, 2107-2124, 1993.

Shum, C. K., R. A. Werner, D. T. Sandwell, B. H. Zhang, R. S. Nerem, and B. D. Tapley, Variations of global mesoscale eddy energy observed from GEOSAT, *J. Geophys. Res.*, *95*, 17865-17876, 1990.

Tai, C. K., and W. B. White, Eddy variability in the Kuroshio Extension as revealed by GEASAT altimetry, *J. Geophys. Res.*, *93*, 10,621-10,629, 1988.

Tarantola, A., *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, New York, 1987.

Teague, W. J., M. J. Carron, and P. Hogan, A comparison between the Generalized Digital Environment Model and Levitus climatologies, *J. Geophys. Res.*, *95*, 7167-7183, 1990.

Tsaoussi, L., and C. Koblinsky, An error covariance model for the sea level surface topography and velocity derived from TOPEX/POSEIDON altimetry, *J. Geophys. Res.*, *99*, 24,669–24,683, 1994.

Van Ballegooyen, R. C., M. L. Grundlingh, and J. R. E. Lutjeharms, Eddy fluxes of heat and salt from the southwest Indian Ocean into the southeast Atlantic Ocean: A case study. *J. Geophys. Res.*, *99*, 14053-14070, 1994.

Van Leeuwen, P. J., and G. Evensen, Data assimilation and inverse problems in a probabilistic formulation, *Mon. Weather Rev.*, *124*, 2898–2913, 1996.

Vazquez, J., V. Zlotnicki, and L.-L. Fu, Sea-level variabilities in the Gulf Stream between Cape Hatteras and 50 W: A Geosat study, *J. Geophys. Res.*, *95*, 17,957-17,964, 1990.

Vazquez, J., E. Smith, M. Hamilton, A. Van Tran, and R. Sumagaysay, NOAA/NASA AVHRR Pathfinder Ocean Project, data management plan, Jet Propulsion Lab., Pasadena, 1994.

Visser, P. N. A. M., The use of satellites in gravity field determination and model adjustment, Ph.D. thesis, Tech. Univ. Delft, Delft, Netherlands, 1992.

Wakker, K. F., M. C. Naeije, R. Scharroo, and B. A. C. Ambrosius, Extraction of mesoscale ocean current information from GEOSAT altimeter data, *ESA SP-312*, 221-226, 1990.

Willebrand, J., R. H. Kase, D. Stammer, H.-H. Hinrichsen, and W. Kraus, Verification of GEOSAT sea-surface topography in the Gulf Stream Extension with surface drifting buoys and hydrographic measurements, *J. Geophys. Res.*, *95*, 3007-3014, 1990.

P.J. Van Leeuwen, Institute for Marine and Atmospheric Research Utrecht, P.O.Box 80005, 3508 TA Utrecht, Netherlands. (leeuwen@phys.uu.nl)

(Received May 19, 1997; revised August 25, 1998; accepted August 31, 1998.)