An Ensemble Smoother with Error Estimates

PETER JAN VAN LEEUWEN

IMAU, Utrecht University, Utrecht, Netherlands

(Manuscript received 14 July 1999, in final form 19 June 2000)

ABSTRACT

A smoother introduced earlier by van Leeuwen and Evensen is applied to a problem in which real observations are used in an area with strongly nonlinear dynamics. The derivation is new, but it resembles an earlier derivation by van Leeuwen and Evensen. Again a Bayesian view is taken in which the prior probability density of the model and the probability density of the observations are combined to form a posterior density. The mean and the covariance of this density give the variance-minimizing model evolution and its errors. The assumption is made that the prior probability density is a Gaussian, leading to a linear update equation. Critical evaluation shows when the assumption is justified. This also sheds light on why Kalman filters, in which the same approximation is made, work for nonlinear models. By reference to the derivation, the impact of model and observational biases on the equations is discussed, and it is shown that Bayes's formulation can still be used. A practical advantage of the ensemble smoother is that no adjoint equations have to be integrated and that error estimates are easily obtained. The present application shows that for process studies a smoother will give superior results compared to a filter, not only owing to the smooth transitions at observation points, but also because the origin of features can be followed back in time. Also its preference over a strong-constraint method is highlighted. Furthermore, it is argued that the proposed smoother is more efficient than gradient descent methods or than the representer method when error estimates are taken into account.

1. Introduction

When one wants to study the dynamics in an area where numerical models have serious shortcomings, a possible step forward is to use data assimilation. Since the emphasis has mostly been on forecasting model fields, much more effort has been devoted to filters than to smoothers. Indeed, it is a simple exercise to show that an optimal filter will give identical results to a smoother at the end of the assimilation interval for linear model dynamics. Recently, Evensen and van Leeuwen (2000) showed that this is even true for nonlinear models, by using a probabilistic description of the data assimilation problem.

Deriving filters for nonlinear problems is not trivial. Extensions of the Kalman filter, which is optimal for linear problems, to nonlinear dynamics has led to a wide variety of suboptimal methods. For weakly nonlinear problems the extended Kalman filter can be used, but this filter often fails for strongly nonlinear systems (see, e.g., Evensen 1992; Miller et al. 1994a,b; 1999). Another problem is the fact that the central forecast or control (defined here as a pure forward model integration of the initial optimal state) is taken as the optimal solution.

To understand the importance of the central forecast problem a more general viewpoint on data assimilation is useful. The formulation of the nonlinear inverse problem is done most naturally in terms of probability densities (see, e.g., Jazwinski 1970; van Leeuwen and Evensen 1996). Intuitively it is clear that the probability density of the model and the probability density of the observations contain all information needed to calculate the inverse estimate. Using Bayesian statistics one can consider the probability density of the model forecast as prior information, which is ''updated'' by the observations. This results in a new probability density of the model, given the observations.

For the problem with the central forecast one has to realize that we always lack knowledge on certain model parameters and/or forcing fields. Clearly, different parameter settings or forcing fields give rise to a different model evolution. Because the model is nonlinear the most probable (or optimal) choices for parameters and forcing will not lead to the most probable model evolution. To get a grip on what a nonlinear model does we have to have some information of the probability density of the model evolution. Of course, one cannot calculate the probability density function of the model for realistic oceanographic or atmospheric problems. But that is not what one can handle either, so the use

Corresponding author address: Peter Jan van Leeuwen, IMAU, Utrecht University, P.O. Box 80005, 3508 TA Utrecht, Netherlands. E-mail: leeuwen@phys.uu.nl

of suboptimal methods seems to be justified. A variety of methods is presently in use, all having their own strong and weak points [see, e.g., Todling and Cohn (1994) and van Leeuwen and Evensen (1996) for a discussion].

A method that uses the probability density idea is the ensemble Kalman filter (Evensen 1994b; see also Burgers et al. 1998). In that method, an ensemble of model states is generated from a randomly perturbed initial state. This might not be too efficient, but Evensen (1994b) and Evensen and van Leeuwen (1996) have shown that that ensemble sizes of 100–500 are enough to describe the model evolution accurately for their two-layer quasigeostrophic models. The method has no problems with nonlinear dynamics and the mean of the ensemble is used as the best model evolution instead of the central forecast. Other methods, like the SEEK filter (Pham et al. 1998; Brasseur et al. 1999) and the RRSQRT filter (Verlaan and Heemink, 1995; see also Lermusiaux and Robinson 1999) use more sophisticated methods to derive an ensemble. Indeed, they report ensembles of much smaller size, but the price that has to be paid is that the optimal solution is given by the central forecast. It can be expected that this is in error for strongly nonlinear systems, as these authors mention. Another problem is that by concentrating only on the singular vectors of the covariance matrix, certain potentially important parts of state space are not probed. This means that the error covariance is always approximated from below, which potentially leads to filter divergence (Todling and Cohn 1995). At least partially random probing seems to be necessary to solve this problem (A. W. Heemink 1999, personal communication). Of interest is the paper by Houtekamer and Mitchell (1998), in which a double ensemble is used to avoid inbreeding during the analysis step. This inbreeding is due to the fact that the same ensemble is used for the covariances as for the Kalman gain in an ensemble Kalman filter. In a comment on this paper van Leeuwen (1999b) showed that inbreeding can also be due to the nonlinearity of the gain itself but, more importantly, that this inbreeding effect is only apparent for relative small ensemble sizes (typically $\sim$100) (see also the reply by Houtekamer and Mitchell 1999). It must be mentioned that Lermusiaux and Robinson (1999) and Brasseur et al. (1999) also discuss variants of their method in which a random component is included.

A serious drawback of filters in general is that they lack propagation of information back in time. It not only causes discontinuities at measurement times, but it can also hamper process studies, especially in data-void areas. This is the main reason why a smoother is considered in this paper.

For a linear model it has been shown by Bennett (1992; see also Jazwinski 1970; Gelb 1974) that the optimal solution can be written as a linear combination of a first-guess field and a weighted sum of correction fields that represent the influence of the observations. Each observation gives rise to one correction field, or representer, as function of space and time. He showed that each representer can be obtained by performing the corresponding measurement on the model covariance. So, although the representer contains information on the characteristics of the corresponding measurement, it is independent of the specific value of that measurement. That information is contained in weighting factors, called the representer coefficients. These are the unknowns in the estimation problem. The search for the optimal solution does not have to be performed in the model space, but in the space with the dimension of the number of observations. Bennett (1992) shows that any correction to the first guess that is orthogonal to the space spanned by the representers is unobservable by the observation array. Because these corrections increase the total error of the optimal state, they should be discarded.

Several ways of solving the weak-constraint linear problem have been provoked, among which are fast implementations of the representer method mentioned above (Bennett 1992; Egbert et al. 1994), the gradient descent method, and the Kalman smoother. Not all methods are able to give proper error estimates. The representer method does not provide error estimates on the model fields themselves, although it does provide them for the initial condition, the model dynamics, and the observations. To obtain errors for the model fields one has to turn to ensemble calculations (see, e.g., Bennett 1992). For a linear problem, errors can be obtained in the gradient descent method from the inverse of the Hessian, but the Hessian is a huge matrix in general. The Kalman smoother propagates the error covariance matrix forward and backward in time.

For nonlinear problems the situation is not that clear. In general, the optimal solution cannot be written as a first guess and a sum of correction fields that are directly related to the observations. In the representer method this problem is addressed iteratively by solving a set of linear inverse problems. Again no error estimate for the optimal state is available. Gradient descent methods can easily be extended to nonlinear problems. Used most frequently are so-called strong-constraint methods, which are, misleadingly, also called adjoint methods (both the representer method and the Kalman smoother use adjoint equations). The Hessian is not a good measure of the errors of the optimal solution because it describes the local curvature of the cost function at the optimal solution, which is not equal to the inverse of the error covariance. So these methods are also lacking an error estimate. One can define an iterative Kalman smoother for nonlinear problems, but that is prohibitively expensive for the size of the problems at hand in meteorology and oceanography.

Recently, van Leeuwen and Evensen (1996) introduced the ensemble smoother, based on a probabilistic description. A practical advantage of the method is that no adjoint equations have to be formulated, because all

knowledge on how information is propagating forward and backward in time is present in the ensemble.

The main points of this paper are as follows. It is shown that the ensemble smoother is able to handle these strongly nonlinear processes, that information is transported effectively backward in time without the need for adjoint equations, and that neither filters nor strong-constraint methods are able to describe the ring shedding process this accurately. Another point is that in deriving the ensemble smoother equations, it is shown that data assimilation does not require either model or observations to be unbiased. In the present paper the ensemble smoother is applied to a strongly nonlinear region in the world ocean, the Agulhas retroflection region. This region couples the Indian to the Atlantic Ocean. The exchange between the two oceans is mainly brought about by large Agulhas rings, which are shed from the retroflecting Agulhas Current. De Ruijter et al. (1999a) give a review of our physical knowledge of the system.

The next section gives a short introduction to the ensemble smoother, discusses the assumptions made in connection with the Kalman filter, and discusses the bias problem. Section 3 describes the data assimilation experiment, and section 4 deals with the results. Finally, conclusions and a discussion of the results are given in section 5.

2. The ensemble smoother

In this section the ensemble smoother is derived using Bayes's theorem. For details the reader is deferred to van Leeuwen and Evensen (1996), which the present derivation greatly resembles. It is shown here explicitly to illustrate that no reference has to be made to an unknown truth, so a bias will not prevent us from using the results.

a. Bayesian statistics

At the heart of nonlinear data assimilation lies the notion of combining probability densities of model and observations. By expressing the problem in terms of probability density functions a Bayesian estimation problem can be formulated. In Bayesian statistics the unknown model evolution ψ is viewed as the value of a random variable ψ. The density $f(\psi)$ of ψ is obtained from the model somehow and is called the prior probability density. Using the definition of a conditional probability density we can derive the new, or posterior, probability density of ψ given the observations $\mathbf{d}$:

$$f(\psi|\,\mathbf{d}) = \frac{f(\mathbf{d}\,|\,\psi)f(\psi)}{f(\mathbf{d})}. \tag{1}$$

The first factor in the numerator, the density $f(\mathbf{d}\,|\,\psi)$, is the probability of the observations given that the model random variable $\psi = \psi$. The second factor is the a priori model density $f(\psi)$. The denominator is the probability density of the observations. Unfortunately, we have no way to determine it because we only have one realization $\mathbf{d} = \mathbf{d}$. A way to circumvent this problem is to realize that $f(\mathbf{d})$ can be considered as the marginal density of the joint probability density of model and observations:

$$f(\mathbf{d}) = \int f(\mathbf{d}, \psi)\, d\psi = \int f(\mathbf{d}\,|\,\psi) f(\psi)\, d\psi. \tag{2}$$

The term on the right-hand side is well understood, as mentioned above. To use this concept in practice, the probability density of the observations $f(\mathbf{d}\,|\,\psi)$ has to be determined. Usually it is assumed to be known, for instance, a Gaussian. Because the model variable ψ is given, the mean of the Gaussian density will be the measurement of the optimal model state, while its variance is the measurement error.

The prior probability density $f(\psi)$ of the model evolution is more difficult to obtain. In principle the Kolmogorov equation describes its evolution. However, because the probability density for the model state has a huge amount of variables, it is computationally not feasible for real oceanographic or meteorological applications to determine its evolution. The evolution of the density could be determined from ensemble integrations. In simulated annealing and related methods this probability density is generated approximately. However, these methods need a huge amount of storage and iterations due to the random nature of the probing. On the other hand, knowledge of the complete density is too much information. One is interested only in its first few moments, for example, a best estimator of the truth and its error variance. In that case ensemble or Monte Carlo experiments can be extremely useful. The most-used estimator is the minimum-variance estimator, in which only integrated properties of the density are needed. The ensemble smoother is an example of an approximate variance-minimizing estimator.

Before we continue it is important to realize what these probability densities exactly mean. The idea is that the true system (the world) evolves according to a set of stochastic differential equations. (One could argue that the true system evolves deterministically, at least at the scales we are looking at. In that case the stochastic terms describe the unknown physics in a stochastic way. The idea is that the statistics of the stochastic process are known.) Because we do not know the random forcing that describes the evolution of the true system, we describe our knowledge in terms of a probability density function. Also, we are unaware of the exact initial condition of the true system, leading again to a probability density function describing our knowledge. If we could know the stochastic differential equation exactly we could determine the evolution of the probability density function in time by means of the Kolmogorov equation. An approximate way of solving this equation is by means of ensemble integrations. If the probability den-

sity of the observations is known too, we can increase our knowledge of the true system evolution by using Bayes's formula, leading of a narrowing of the probability density of the model evolution (at least in general). Because both the probability density of the model evolution and that of the data is unbiased, the conditional probability density of the model given the observations will be unbiased too (e.g., Jazwinski 1970).

The fact is, however, that we do not know the exact stochastic differential equations, so the probability density that we derive from the Kolmogorov equation is biased. (No doubt our probabilistic description of the initial condition is biased too.) This means that not only the mean, or the mode, or other statistical estimates of the true evolution of the system, are biased, but also the covariances and the higher-order moments. If the probability density of the observations is not biased, we could estimate the biases in all moments as wanted. Dee and da Silva (1998) concentrate on the bias in an optimal estimate of the true evolution and neglect the other biases.

If one wants to reduce the bias in the whole probability density (the ultimate goal), the Bayes formula gives us the key to do that. What it tells us is how we should combine two sources of information on the system. Clearly, the best estimate of the information that comes from the model is given by our best estimate of the probability density of the model evolution. It might (or in fact will) be biased, but it is the best estimate we have. The same is true for the observations. Our probabilistic description of the errors in them is very often biased too, especially in oceanography. Bayes just tells us that the optimal way to handle the (all) information we have is by combining the probability densities according to his formula. If the observations are much less biased than the model, we end up with exactly what we want: a less biased estimate of the probability density. When we have that we can calculate less biased estimates of whatever moment we want. This is the reason why I argue that the Kalman filter equations are still giving us a optimal estimate of the true system even when the model and/or the observations are biased. Note that if the observations are much less biased than the model, the covariances will be less biased too. This is only true indirectly (i.e., via a better mean) in the method proposed by Dee and da Silva (1998).

b. Minimum-variance estimator

With the above discussion in mind, we derive the smoother equations. The variance-minimizing estimator is the variable $\hat{\psi}$ such that the variance

$$e(\hat{\psi}) = E[(\psi - \hat{\psi})^2] = \int (\psi - \hat{\psi})^2 f(\psi|\mathbf{d})\, d\psi \quad (3)$$

is minimal. Note that the integration variable is ψ, so it is a complete model evolution in space and time. The minimum can be obtained from the calculus of variations and the notion that $\hat{\psi}$ is independent of ψ as

$$\hat{\psi} = \int \psi f(\psi|\mathbf{d})\, d\psi, \quad (4)$$

which can be rewritten, using (1) and (2), as

$$\hat{\psi} = \frac{\int \psi f(\mathbf{d}|\psi) f(\psi)\, d\psi}{\int f(\mathbf{d}|\psi) f(\psi)\, d\psi}. \quad (5)$$

This equation shows that to find the first moment of the posterior density, all moments of the prior density are needed. As explained above, these cannot be determined for the problem size we have in mind. An approximation used in the Kalman filter is to assume that the prior density is Gaussian distributed in state space at the time of the observation. (This assumption does not have to be made in the case of a linear model, but it is the probabilistic interpretation for a nonlinear model.) In that case, only two moments describe the complete density. In the ensemble smoother we do the same, with this difference: we make this assumption for the prior probability density function at all times. In the following an expression for the ensemble smoother estimator is given, together with an expression for the error that is made this way.

For nonlinear dynamics the probability density can 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 becomes

$$f(\psi) = f_G(\psi) f_N(\psi), \quad (6)$$

in which the Gaussian part is given by

$$f_G(\psi) = A_G \exp\left[-\frac{1}{2}(\psi - \psi_F) \bullet W_{\psi\psi} \bullet (\psi - \psi_F)\right]. \quad (7)$$

In this equation A_G is a normalization factor, ψ_F is the mean of the distribution, and $W_{\psi\psi}$ is the inverse of the model covariance $Q_{\psi\psi}$, found to fulfill

$$Q_{\psi\psi}(\mathbf{x}_1, t_1, \mathbf{x}_3, t_3) \bullet W_{\psi\psi}(\mathbf{x}_3, t_3, \mathbf{x}_2, t_2) = \delta(\mathbf{x}_1 - \mathbf{x}_2)\delta(t_1 - t_2). \quad (8)$$

The $\bullet$ denotes integration over space and time. Also the boundary and initial errors are included in (7). For further reference it is noted that this density might be biased. The above equation just expresses all knowledge we have of the system.

Assume the probability density of the observations to be Gaussian too with error covariance matrix $\mathbf{W}^{-1}$. The posterior density then becomes

$$f(\psi|\mathbf{d}) = Af_N(\psi)\exp\left\{-\frac{1}{2}(\psi-\psi_F)\bullet W_{\psi\psi}\bullet(\psi-\psi_F) - \frac{1}{2}[\mathbf{d}-\mathbf{L}(\psi)]^{\mathrm{T}}\mathbf{W}[\mathbf{d}-\mathbf{L}(\psi)]\right\}. \tag{9}$$

in which $\mathbf{L}(\cdot)$ is the measurement operator, assumed to be linear, and A a normalization factor. Now take the variational derivative of this equation and multiply with $Q_{\psi\psi}$ to find

$$-Q_{\psi\psi}\bullet\frac{\delta f(\psi)}{\delta\psi} = \{\psi-\psi_F-\mathbf{L}^{\mathrm{T}}(Q_{\psi\psi})\mathbf{W}[\mathbf{d}-\mathbf{L}(\psi)]\}f(\psi|\mathbf{d}) - AQ_{\psi\psi}\bullet\frac{\delta f_N(\psi)}{\delta\psi}f_G(\psi)f(\mathbf{d}|\psi). \tag{10}$$

Integrate this equation over the ψ domain, integrate over space and time, and use (4) to find

$$\hat{\psi} = \psi_F + \mathbf{L}^{\mathrm{T}}(Q_{\psi\psi})\mathbf{W}[\mathbf{d}-\mathbf{L}(\hat{\psi})] + Q_{\psi\psi}\bullet\int A\frac{\delta f_N(\psi)}{\delta\psi}f_G(\psi)f(\mathbf{d}|\psi)\,d\psi. \tag{11}$$

To proceed, this equation is measured and subtracted from the measurements $\mathbf{d}$ to find

$$\mathbf{d}-\mathbf{L}(\hat{\psi}) = \mathbf{W}^{-1}(\mathbf{R}+\mathbf{W}^{-1})^{-1}[\mathbf{d}-\mathbf{L}(\psi_F)] - \mathbf{W}^{-1}(\mathbf{R}+\mathbf{W}^{-1})^{-1}\mathbf{L}(Q_{\psi\psi})\bullet\int A\frac{\delta f_N(\psi)}{\delta\psi}f_G(\psi)f(\mathbf{d}|\psi)\,d\psi, \tag{12}$$

in which the representer matrix $\mathbf{R}$ is the measurement operator covariance,

$$\mathbf{R} = E[\mathbf{L}(\psi-\psi_F)\mathbf{L}^{\mathrm{T}}(\psi-\psi_F)] = \mathbf{L}[\mathbf{L}^{\mathrm{T}}(Q_{\psi\psi})]. \tag{13}$$

Here $E(\cdot)$ is the expectation operator. Using this result in (11) we find

$$\hat{\psi} = \psi_F + \mathbf{r}^{\mathrm{T}}\mathbf{b} + [Q_{\psi\psi}-\mathbf{r}^{\mathrm{T}}(\mathbf{R}+\mathbf{W}^{-1})^{-1}\mathbf{r}]\bullet\int A\frac{\delta f_N(\psi)}{\delta\psi}f(\mathbf{d}|\psi)f_G(\psi)\,d\psi, \tag{14}$$

in which $\mathbf{b}$ are the well-known representer coefficients, which can be determined from

$$(\mathbf{R}+\mathbf{W}^{-1})\mathbf{b} = \mathbf{d}-\mathbf{L}[\psi_F], \tag{15}$$

and we introduced the representers $\mathbf{r}$, which are the model field measurement operator covariances, as

$$\mathbf{r} = E[(\psi-\psi_F)\mathbf{L}(\psi-\psi_F)] = \mathbf{L}(Q_{\psi\psi}) \tag{16}$$

The representers are crucial in the data-assimilation problem. A representer describes how the information of the corresponding measurement influences the solution at all space–time points. Bennett (1992) gives an excellent treatment of the meaning of representers. The representer coefficients determine how strong each representer should be counted in the final solution. They depend on the model–observation misfit and on their error covariances. Clearly, an important ingredient in the success of the inversion procedure is the conditioning of the sum of the representer matrix and the error covariance of the observations, as given in (15).

The error covariance of the optimal estimate is given by

$$Q_{\hat{\psi}\hat{\psi}} = Q_{\psi\psi} - \mathbf{r}^{\mathrm{T}}(\mathbf{R}+\mathbf{W}^{-1})^{-1}\mathbf{r} + \beta, \tag{17}$$

in which β a complicated term related to the non-Gaussian part of the prior density, as shown in the appendix.

If the model were linear and the first-guess field was distributed Gaussian initially, 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 appears in (14).

To find the nonlinear smoother one has to specify f_N, which is nontrivial. Furthermore, even for simple forms of f_N complicated expressions for $\hat{\psi}$ arise. In the ensemble smoother the nonlinear contribution is neglected altogether. This is legitimate in a few cases. First, if the non-Gaussian density contains only even moments centered on $\hat{\psi}$, the contribution of this term vanishes. A second possibility is found by close inspection of (14). It shows that the effect of the nonlinear term is proportional to $Q_{\psi\psi}-\mathbf{r}^{\mathrm{T}}(\mathbf{R}+\mathbf{W}^{-1})^{-1}\mathbf{r}$. For the ensemble smoother this is the error covariance of the optimal state after assimilation, $Q_{\hat{\psi}\hat{\psi}}$. So the smaller this error, the better the ensemble smoother performs. Note that this will be one of the reasons that data assimilation methods that use the Kalman filter update equation work at all in (strongly) nonlinear situations: as long as the filter does not diverge the optimal estimate will be relatively close to the minimum-variance estimate. Finally, one can simply look at higher moments of the prior density; if the third moment stays small compared to the second we are relatively safe. (Note the word ''relatively''; it is extremely difficult to quantify these matters for a real-sized problem.) Details for efficient construction of the smoother solution using ensemble statistics can be found in van Leeuwen and Evensen (1996).

It is stressed again that the above derivation does not refer to an unknown ''true'' evolution. So, contrary to what large groups of meteorologists and oceanographers think, the equations for the optimal model evolution require neither the model nor the observations to be unbiased. Clearly, an unbiased model and unbiased observations will lead to a better estimate of the true evolution of the system, but that is not a prerequisite for data assimilation. This is fortunate, because it is extremely difficult to find model biases. Sometimes it is even impossible, because the observations are likely to be biased too. An example are observations of the Southern Ocean, which are strongly biased to austral summer situations. More generally, a lack of observa-

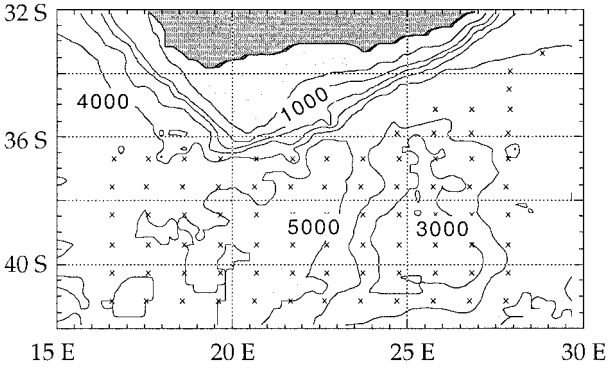

FIG. 1. Bottom topography. Contour interval is 1000 m, with the continental shelf at 100-m depth shaded. Crosses denote 1° interpolated values of TOPEX/Poseidon.

tions will prevent us from determining model biases, but, as stated above, data assimilation is still well defined.

3. Data assimilation experiment

TOPEX/Poseidon data are assimilated in a nonlinear quasigeostrophic ocean model with the ensemble smoother. Observations and model are presented and much effort is put in the determination of proper error statistics.

a. Observations

The standard corrections were applied to the TOPEX/Poseidon data (solid Earth and ocean tides, the dry and wet troposphere, the inverse barometer effect, and the sea state bias as 2% of the significant wave height). The excellent satellite orbits needed no orbit error corrections. Because the geoid is not known very precisely we only used the time-varying part of the altimeter signal. The time mean part is obtained from the Gordon (1982) *Southern Ocean Atlas* dataset. The altimeter data where Gauss–Markov interpolated to a one-by-one degree grid is given in Fig. 1, every 10 days. The decorrelation length was 100 km in space and 10 days in time. These observation fields are assimilated every 10 days in the model.

Tsaoussi and Koblinsky (1994) estimate the error in the time-varying part of the raw altimeter data to be 3 cm rms. The space and time smoothing will reduce this error even further. The error in the time mean field is incorporated in the error of the initial condition. Clearly, incorporating this error at each measurement time will lead to highly correlated error estimates of initial conditions and measurements, which is not realistic. Furthermore, the error in the mean field will have larger spatial scales than those present in the altimeter fields. We finally estimated the error in each measurement to be 1 cm, with the above discussion in mind. Error covariances between different measurements are neglected, mainly due to the large distance between the measurement points. The next section gives some statistical evidence that the errors were probably chosen correctly.

One may argue that the observations have to be assimilated as purely as possible, that is, including the alongtrack observations as soon as they are available. A disadvantage of this procedure is that the observations contain subgrid-scale information alongtrack, while they are sparse acrosstrack. By interpolating to a grid the subsampling problem is avoided. So gridded fields are chosen, but it is recognized that the choice is subjective. However, the exact choice is not of relevance to the main part of the paper.

Figure 2 shows an interpolated set of eight consecutive fields that are assimilated into the model. These fields are interpolated with a decorrelation scale of 100 km and a decorrelation time of 10 days. It should be realized that the interpolation results in rather smooth fields, which sometimes even cross the continental shelf (see below). A large anticyclonic Agulhas ring is in the process of being shed in the first image. At day 30 it seems to leave the domain around 37°S. A second ring, or the remaining part of the first ring, leaves the domain on day 60, at 38°S. The large cyclonic ring, centered at 36°S, 18°E seems to be at least partly responsible for the shedding of the rings. We will come back to this later. The Agulhas Current, as it flows along the continent, is rather smooth, as can be expected from the interpolation. This will also be commented on later. It should be noted that we assimilate only the observations at the individual points given in Fig. 1.

b. Model and initial state

The model describes the ocean in a box from 32°S, 15°E to 42°S, 30°E and describes the circulation of the ocean around the southern tip of the African continent. Figure 1 gives the bottom topography in this area. Characteristic here is the large continental shelf, given by the shaded area and the steep slope seaward of it. The relatively shallow Agulhas Plateau centered at 38°S, 25°E is important for the current system. The Agulhas Current flows along the east coast of the continent southward with velocities of up to 3 m s^{-1}. The main portion of the current retroflects at about 38°S, 20°E and flows back as the Agulhas Return Current into the Indian Ocean (see Fig. 3). The Agulhas plateau forces the return current to make a large northward meander. Sometimes a direct connection seems to exist between the Agulhas Current and the Return Current close to the plateau. At the retroflection loop large anticyclonic rings are formed, the Agulhas rings, which travel west into the South Atlantic, thus carrying large amounts of relatively warm and salty water from one ocean to the other. Farther south the Antarctic Circumpolar Current runs from west to east.

To model these features the numerical model has open boundaries to account for the inflow of the Agulhas

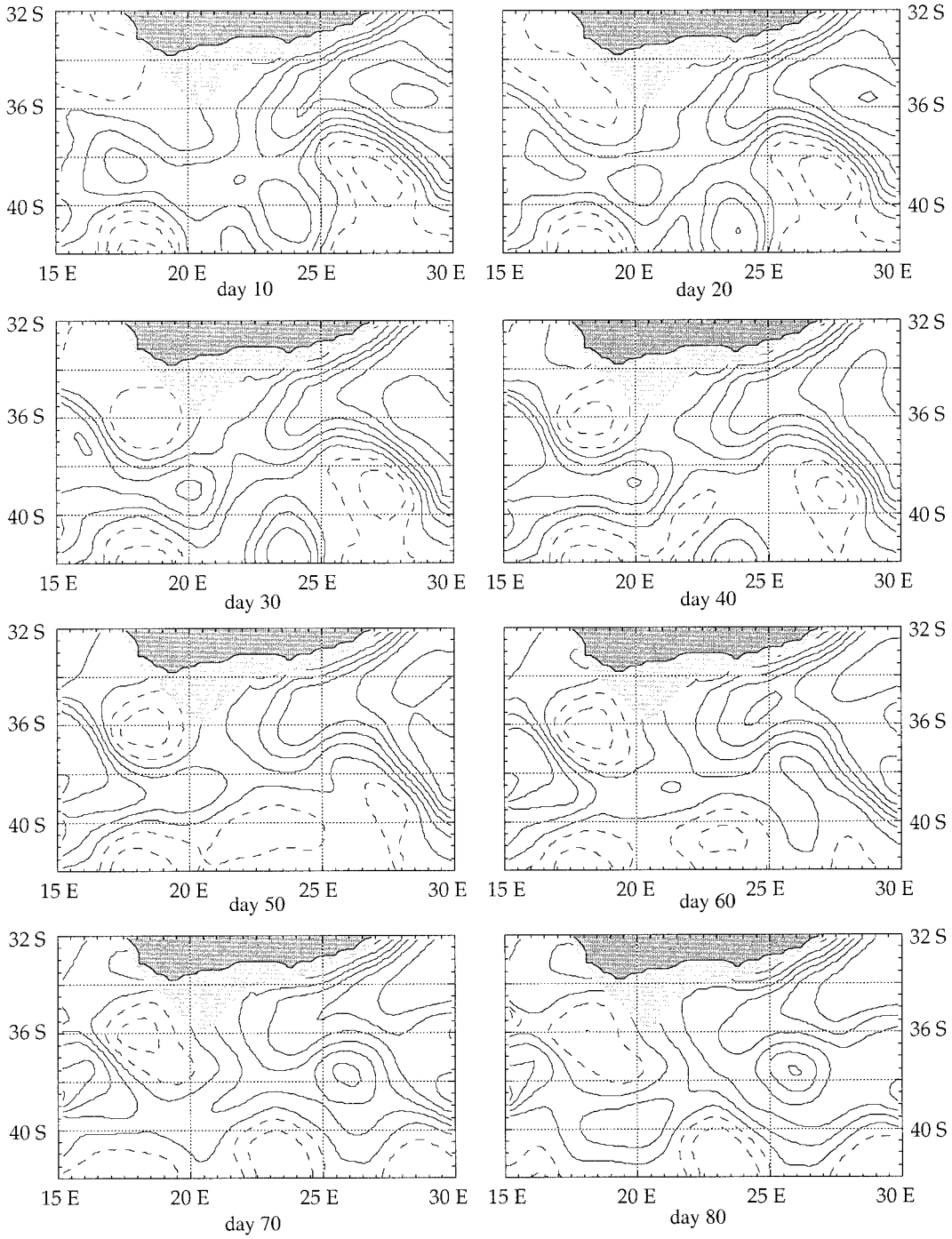

FIG. 2. TOPEX/Poseidon altimeter fields interpolated to the model grid. Contour interval 10 cm, zero line not indicated, dashed lines denote negative values.

Current, the in- and outflow of the Antarctic Circumpolar Current, the outflow of the Agulhas Return Current, and the migration of Agulhas rings westward. The quasigeostrophic approximation is made and the African continent is extended to the 100-m depth contour to prevent too much influence of coastal shelf dynamics, which is not described well by quasigeostrophic dynamics. The most energetic processes have a larger scale and it is assumed that the influence of the shelf sea can be neglected. The model consists of two layers in the vertical of 1- and 4-km depth, and 25-km resolution in the horizontal. The density difference between the layers is 2.5 kg m^{-3}. The time step is 1 h. Small-scale features are removed with a Shapiro filter of order 8 once every

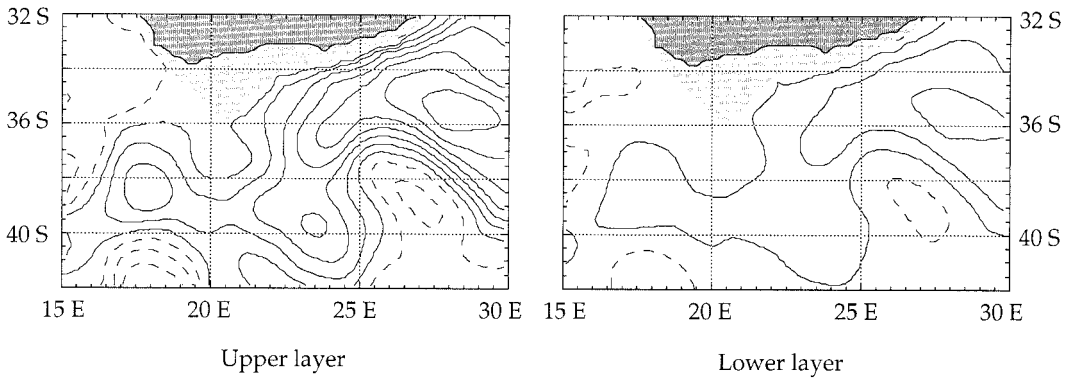

FIG. 3. Initial condition for upper- and lower-layer streamfunction. The water flows approximately along the contours. East of the continent flows the Agulhas Current, which turns back to the Indian Ocean below the tip of the continent. A large eddy is about to be pinched off. Contour interval 20 and 10 Sv, respectively. Dashed contours denote positive values.

day. The along-boundary derivative of the streamfunction is kept zero on the land boundary (free slip) by applying the modified capacitance matrix method as described by Millif (1990). The influence of the bottom topography is reduced by a factor of 0.2 to compensate for the overestimation of bottom topography in a quasigeostrophic model. At the open boundaries baroclinic waves are propagated out of the domain, perpendicular to the boundary, as in Ikeda and Apel (1981) and Ikeda et al. (1989). This means that no small-scale disturbances enter the domain. Experiments showed that mesoscale features leave the domain with a speed comparable to that in observations, that is, about 5 cm s^{-1} (see Olson and Evans 1986), without shape distortions. It should be noted, however, that the precise boundary condition is not that important, as long as it is not totally wrong. The data-assimilation scheme allows the boundaries to contain errors and the assimilated fields have boundary values consistent with the observations and the model dynamics. [The same model configuration is used in van Leeuwen (1999a), but in that paper an attempt was to reconstruct the time mean field using the time-varying part of the altimeter signal and the model dynamics.] In comparison, Holland et al. (1991) used sponge layers at the horizontal boundaries. A disadvantage of that approach is extra computations in the sponge layers. As mentioned above, the boundary conditions used here are satisfactory. They also used a five-layer model in an extended domain. The main reason to keep to two layers here is that this is meant as a demonstration of the capabilities of the smoother method used here, not for an optimal description of the ocean circulation around South Africa.

As initial condition we choose the time mean part from the Gordon (1982) dataset. The time mean circulation derived by van Leeuwen (1999a) could not be used because the same observations are used in that paper as are used here. This means that the model and the observations would become correlated, while the assumption made in the derivation above, and in fact in most (if not all) data assimilation methods, is that they are uncorrelated. To the upper layer we added a stream function field obtained from the first one-by-one degree TOPEX/Poseidon altimeter grid, interpolated to the model grid.

The difference with Evensen and van Leeuwen (1996) is that an ensemble Kalman filter is used there. Furthermore, the model is slightly different, the observations are from a different satellite for a different period of time, and the error statistics of model and observations are different.

c. Initial, boundary, and dynamical model errors

In the ensemble smoother a number of ensemble members are integrated forward in time. The members are found by adding random fields to the initial state. These random fields are Gaussian distributed with covariances given by

$$\mathrm{cov}(x, y) = A \exp\left[-\frac{(x - y)^2}{d^2}\right], \qquad (18)$$

in which $A = 10^8$ m^2 s^{-2} for the first layer and $A = 10^6$ m^2 s^{-2} for the second layer. Here $d = R_d$ is the Rossby deformation radius, which is about 50 km for the chosen density difference. The correlation between upper and lower layer error fields is 0.4. This is the correlation of the two layers found from a multiyear integration of the model. The errors do not have to be distributed in the same way, but it is probably the best one can do: these vertical correlations of the model error fields are extremely difficult to determine.

The initial model errors are about 10% of the initial values. This value has been chosen to incorporate the errors in the Gordon dataset and the errors in the TOPEX/Poseidon data, both due to real observation errors and due to representation errors. It corresponds to an error in initial sea level of about 10 cm. Compared to characteristic sea level variations of the order of 1 m

this may seem large, but a small shift in the main currents will give rise to this order of magnitude. For relatively quiet areas this is not a good approximation, but there the errors in the interpolated observations will be smaller too. So the weighted mean given by the ensemble smoother will not be too sensitive to this choice. Another point is that the initial field might contain a quiet area that will be more variable during the assimilation run: we do not know beforehand which way the major current and the eddies will move. With this in mind, position-independent errors are adopted.

The errors in model dynamics are more difficult to access. We want to describe the combined effect of missing model physics, like poorly represented stretching terms and bottom topography, representation errors due to the two-layer approach, and forcing. In this paper we focus on the internal dynamics and excluded the wind forcing.

As a first estimate one can use the fact that the errors in the momentum equation are of the order of the Rossby number, which is about 0.05 in this case. However, the error will be dominated by the assumption that the interface displacement is much smaller than the undisturbed layer depth, thus causing errors in the continuity equation. That missing physics can partly be described by assuming conservation of potential vorticity II in each layer, defined as

$$\Pi_1 = \frac{\zeta_1 + f}{H_1 + \eta + h} \quad \text{and} \tag{19}$$

$$\Pi_2 = \frac{\zeta_2 + f}{H_2 - b - h}, \tag{20}$$

in which ζ is the relative vorticity, f the planetary vorticity, H_i the layer depth in absence of motion, η the sea surface height, h the interface height measured positive downward, and b the bottom topography. The conservation of the potential vorticity in each layer can be split in a quasigeostrophic part and a part that is not modeled:

$$H_1\frac{d\Pi_1}{dt} = 0$$

$$= \frac{d(\zeta_1 + f)}{dt} - \frac{f}{H_1}\frac{dh}{dt} + \left(\frac{f}{H_1} - \frac{\zeta_1 + f}{H_1 + h}\right)\frac{dh}{dt}, \tag{21}$$

in which η is neglected compared to h and H_1, and

$$H_2\frac{d\Pi_2}{dt} = 0 = \frac{d(\zeta_2 + f)}{dt} + \frac{f}{H_2}\frac{d(h + \epsilon b)}{dt} - \left(\frac{f}{H_2} - \frac{\zeta_2 + f}{H_1 - h}\right)\frac{dh}{dt} - \left(\frac{\epsilon f}{H_2} - \frac{\zeta_2 + f}{H_2 - b - h}\right)\frac{db}{dt}, \tag{22}$$

in which $\epsilon = 0.2$ to reduce the overestimation of bottom topography changes in quasigeostrophic context, as mentioned above. In both equations the first two terms are modeled in the quasigeostrophic model, while the rest is not. An order of magnitude for these terms can be found as follows. A characteristic velocity is 0.5 m s^{-1} and a characteristic horizontal scale is 100 km, leading to a relative vorticity of 5×10^{-6} s^{-1}. This is small compared to the planetary vorticity, which is -9.3×10^{-5} at this latitude. Interface depressions are about 100 m, from quasigeostrophic theory. If we use these values in the nonmodeled part of the evolution of the upper layer conservation of potential vorticity, we find an order of magnitude of 10^{-11} s^{-2}. This error estimate is also used for the boundaries.

For the lower layer a characteristic velocity is 0.05 m s^{-1} and a characteristic horizontal scale is 100 km, leading to a relative vorticity of 5×10^{-7} s^{-1}. The bottom topography has an amplitude of 1000 m and a horizontal scale of 500 km. This leads to a nonmodeled part of about 10^{-12} s^{-2}.

These error estimates have to be combined with errors due to the bad representation of stratification. It is extremely difficult to say anything sensible on these errors. It is assumed that they are at most of the same order of magnitude as the errors due to the quasigeostrophic assumption. Hence we choose the total errors to be 10^{-11} s^{-2} and 10^{-12} s^{-2} for the upper and lower layers, respectively. To account for the fact that the model errors will not be white in time, a decorrelation time of 0.5 days is used for these errors. It turns out that the model is not very sensitive to this choice; white noise gives differences in the model fields that are not detectable in the figures given below.

The spatial distribution is again an issue here. For the active areas the above-defined error estimates are probably reasonable, but quiet areas west of Africa may need smaller values. Because this area is probably not of great importance for the shedding of the Agulhas rings, the errors are kept position independent.

Finally, it is worthwhile to mention that the errors were chosen differently in Evensen and van Leeuwen (1996). As is mentioned in the next section, the errors used in this paper are consistent with the statistics of the inversion. That test was not done in Evensen and van Leeuwen (1996).

4. Results

Initial conditions for the upper and lower layers are given Fig. 3, showing the streamfunction in each layer. The Agulhas Current can be detected along the east coast of South Africa, with velocities of up to 3 m s^{-1}. At the southernmost point of the continent it retroflects and flows back into the Indian Ocean as the Agulhas Return Current, which leaves the domain at the eastward side. In the retroflection area a large Agulhas eddy is in the process of being pinched off. After pinching off it will travel westward into the Atlantic Ocean and leave

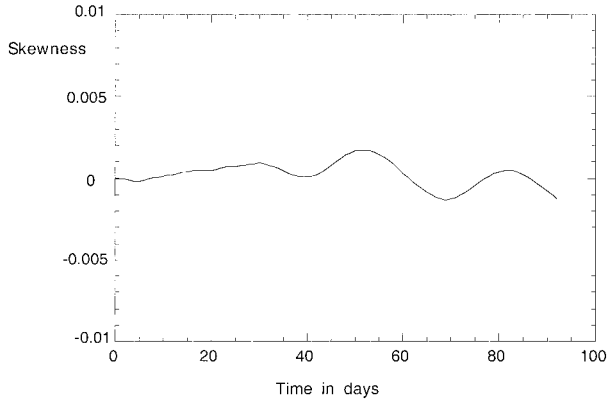

FIG. 4. Skewness of the prior probability density with time.

the model domain at the westward side. Note that these eddies are mainly responsible for the transport of, for instance, heat, salt, and energy from the Indian to the South Atlantic Ocean and, as such, may be of importance for the global climate (Veronis 1973; Gordon 1986; Lutjeharms and Roberts 1988; De Ruijter et al., 1999b).

a. Assimilation statistics

The size of the ensemble was 500. Increasing this size to 600 did not change the model results more than 1%, so we conclude that 500 members was enough in this case. Even an ensemble of 100 members did show the features described below; however, increasing the ensemble size still altered details of the resulting fields. To test the assumption of Gaussianity of the prior distribution the skewness was calculated every time step, see Fig. 4. It stayed close to its initial value of 10^{-4}.

A time evolution of the upper-layer streamfunction field is shown in Fig. 5. It shows the mean of an ensemble of 500 members without data assimilation. These fields show the best guess the model can give of the evolution of the circulation, given the model errors. Note that due to the nonlinearity of the model the central forecast, defined here as the pure model run from the optimal initial condition, is not the optimal estimate of the model evolution. In fact, its statistical meaning is that it is a first-order estimate of the statistical mean (see Jazwinski 1970). One might argue that the central forecast has more dynamical meaning than the mean of an ensemble. However, since we allowed for errors in the model dynamics this is not true. We know what the dynamical evolution of the central forecast is, but we also know that it is wrong.

The spatial scales in the evolution of the upper layer are increased somewhat compared to the initial field due to two effects. First, a quasigeostrophic model has the tendency to increase the scales due to the ''inverse energy cascade'' (see, e.g., Pedlosky 1987). This leads to an increase in scale for each member of the ensemble individually. Second, the errors in the model simulation grow rather large (see Fig. 7), so the ensemble mean will be smoother than each individual member.

An impression of the errors involved in the solution can be obtained from Fig. 6, which gives the errors at day 90 before and after the assimilation run for the upper layer. (Note that, once the ensemble is created, the errors are obtained easily in this data-assimilation method.) The errors have been reduced by a factor of about 10 in some places. Prior to the assimilation, the errors at day 90 were up to 70% of the streamfunction values in the area of largest errors. After the assimilation this number reduces to below 10% of the assimilated streamfunction fields. Another way to visualize the error evolution is given in Fig. 7. The variances at all model area points from both layers are added and normalized by the initial total error before assimilation. The variance of the pure ensemble integration grows rapidly during the integration. Close to day 80, however, the growth seems to level off. Further integration of the ensemble shows that the total variance keeps on growing, but at a much smaller pace (not shown). This is what one would expect in a nonlinear model owing to the nonlinear interactions. A linear model would most likely show a continuation of this fast growth.

The variance of the posterior ensemble shows the characteristic depressions around the measurement times. Clearly, the solution is pulled toward the measurements before the actual measurement times. A striking example of this can be seen at $t = 0$. The posterior variance is a factor of 2 lower than the prior variance while observations are only available 10 days later. Here we see the difference with a filter like the Kalman filter, in which discontinuities arise at measurement times. The overall error over both layers, depicted in Fig. 7, is reduced by a factor of 30 at final time.

The difficulty in data assimilation lies in the determination of the model error covariances. In fact, as Bennett and Thorburn (1992) note, data assimilation is hypothesis testing in which the null hypothesis is given by the chosen error covariances. It is here that the science comes in, in this otherwise purely technical exercise. The test statistic is the value of the penalty function. For a linear model Bennett and Thorburn (1992) shows that the penalty function is a χ^2 variable with the number of independent measurements as the number of degrees of freedom. (This shows that the weights in a least square procedure are not just weighting factors, they have a physical meaning: the inverses of the error covariances.) In our case the penalty function reduces from about 1 000 000 to 875. The number of measurements was 738, but the number of independent measurements determined from the conditioning of the representer matrix was 209. This is not within the error bounds from the value of the penalty function, which is given by $\sqrt{875} = 30$, pointing to a bad choice of the initial error covariances. However, our model is nonlinear, so we cannot reject the null hypothesis on this

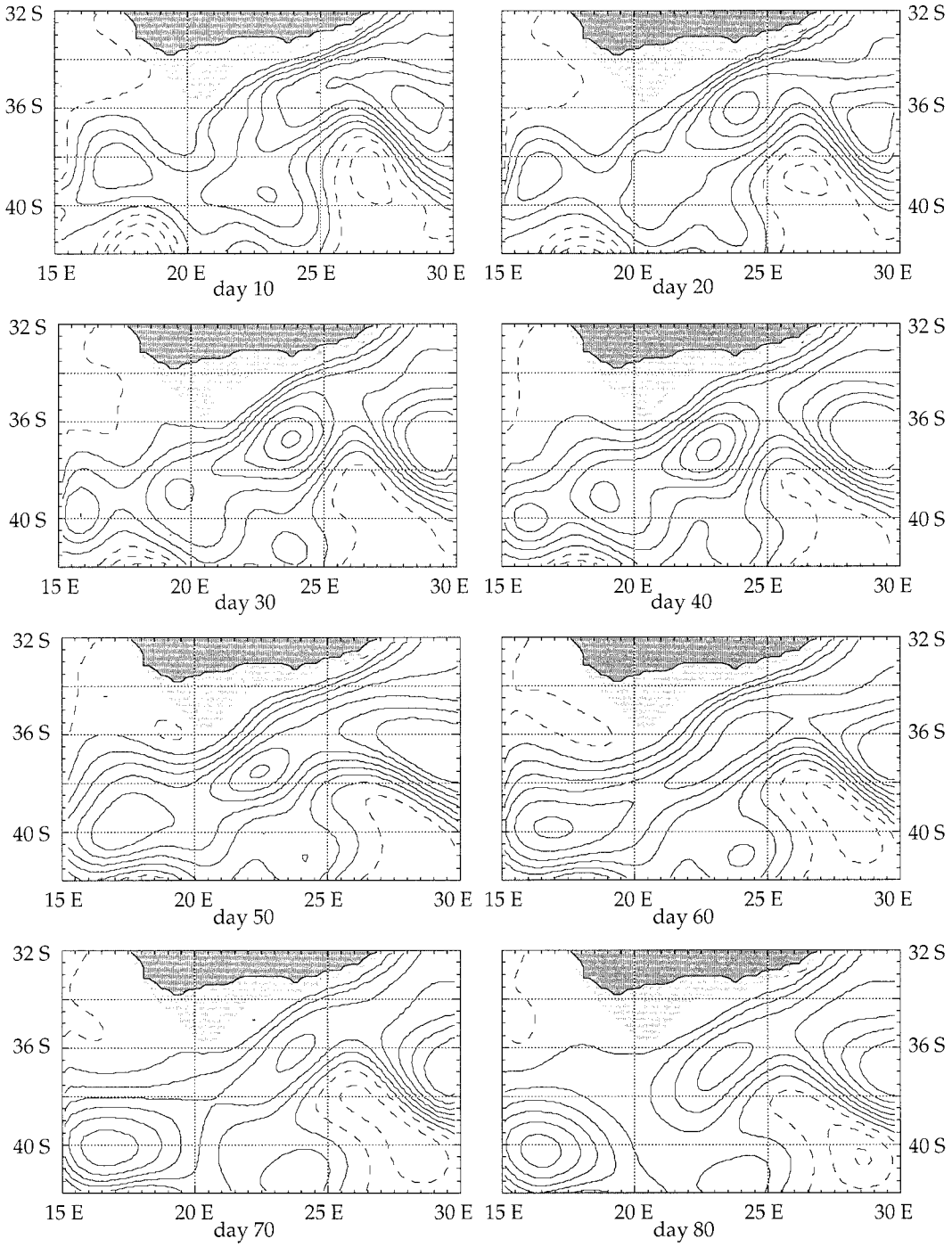

FIG. 5. Prior streamfunction of upper layer, contour interval = 20 Sv, dashed lines denote positive values.

basis. The fact that the value of the penalty function was reduced by a factor of 1000 to end up close to the number of independent measurements gives us confidence in the a priori estimated errors in model and observations.

The representer matrix can be used to obtain some idea where to do measurements that will have the greatest influence on the prior model evolution. To this end the eigenvectors with largest eigenvalues are determined. The largest eigenvector elements in these eigenvectors give the positions of the measurements that have most impact, because they point to positions where the model measurement operator covariances are largest. In Fig. 8 the eigenvector elements are contoured for

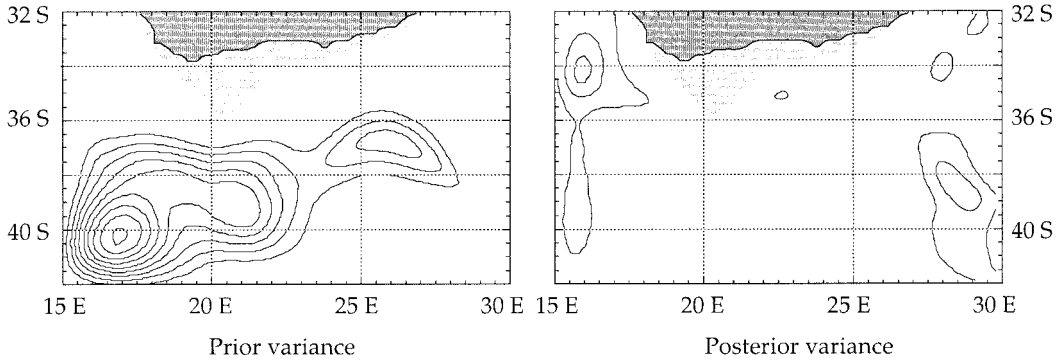

FIG. 6. Prior and posterior error estimates at day 90, contour interval = 5 Sv. The latter are multiplied by 10 to be visible, so a huge error reduction has been achieved.

the first 10 eigenvectors, which explain 90% of the variance. Three maxima can be distinguished, more or less in a southwest to northeast band. These are the positions where future measurements of sea surface height should be done. Of course, altimeters cover the whole area, so the exercise does not teach us that much here. However, if we have chosen a model with its error structure, we could do the same for hypothetical in situ measurements. This will tell us where to go on a cruise to have maximum impact on the final solution. Given the average cost of about $25 000 per day for a research vessel, the economical impact can be quite large.

b. Physical results

We have seen in Fig. 5 that the mean of the prior probability density was rather smooth. As a result of the larger scales, the Agulhas eddy that finally pinched of at day 80 is also too large compared to the measurements (e.g., van Ballegooyen et al. 1994). The nearly formed eddy at day 60 seems to be unable to leave the model domain, which could direct to a wrong boundary condition. However, this is only the effect of the averaging over the members. In some members it leaves the domain and in others it does not even come this far.

It is interesting to see that the eddy that nearly pinches of in the beginning of the run is recaptured by the Agulhas Current to be shed much later again. This also happens in reality (B. D. Olson and W. P. M. De Ruijter 1999, personal communication), but in this case it seems to be an artifact of the model, because it does not seem to happen in the observations (see Fig. 2). There, instead, the ring leaves the domain followed by a second ring 30 days later. The model also has the tendency to shed the ring too far south. Several possibilities for this misbehavior of the pure model can be identified. First, in a quasigeostrophic model instabilities grow too fast compared to reality but the final shedding of rings occurs too slow. This is due to the fact that just before the rings are shed and a small filament connects the eddy to the mean flow, small-scale processes are important, beyond the scale of quasigeostrophic dynamics (see, e.g., Drijfhout 1990). On average the ageostrophic terms are of moderate importance; the Rossby number is $R = U/Lf \approx 0.05$ for a characteristic velocity of 0.5 m s^{-1} and a lengthscale $L \approx 100$ km ($f \approx 10^{-1}$ s^{-1}). However, in the neck region of a ring that is being shed the velocity gradients increase to 4×10^{-5} s^{-1}, leading to $R \approx 0.4$.

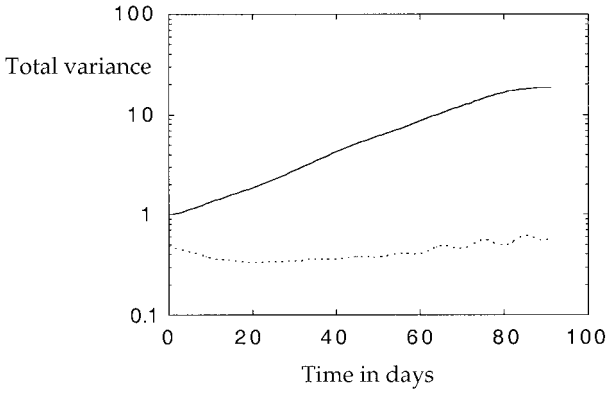

FIG. 7. Total errors of prior (drawn) and posterior (dashed) streamfunction fields, normalized by the initial error of the prior fields. A reduction of a factor of 30 is visible at day 90. Note that the errors are also reduced at the initial time, due to the use of a smoother instead of a filter.

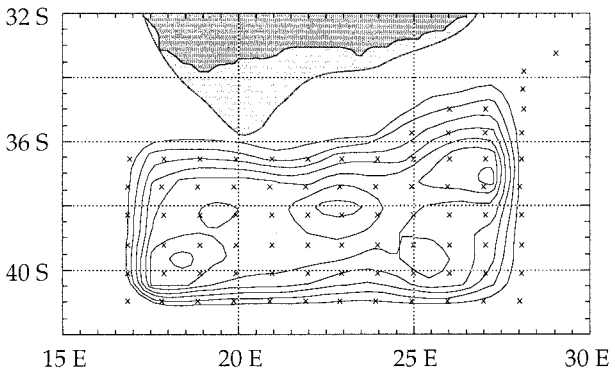

FIG. 8. Eigenvector elements of the first 10 eigenvectors, showing the measurement positions that had greatest influence on the final solution.

Another problem with quasigeostrophy is that the interface displacements are assumed to be of the order of 100 m, while van Ballegooyen et al. (1994) show from hydrography that the displacements can easily be 500 m in a ring. So, without a doubt, ageostrophic terms are of importance in the shedding area. Second, the mean field added to the observations, and hence to the model initial field, is not correct. The mean field used is probably the best available, but because it is based on hydrography it has the problem of the unknown reference level. Finally, it might be a resolution problem. Instabilities with a wave length of the order of the Rossby deformation radius are just on the edge of being resolved. Because the Shapiro filter is applied once every day, the instabilities do have the possibility to grow, but their evolution will be effected.

Figure 9 shows the streamfunction fields of the assimilated run. The first feature that strikes the eye is the much smaller spatial scales. So indeed, the TOPEX/Poseidon observations have their influence on the solution. The scales in this figure compare well with those obtained from infrared images, considering the 25-km grid of the model (npcitelut88). Furthermore, a ring leaves the domain on day 30, and a larger one pinches off on day 60, about 20 days earlier than in the prior run. This in agreement with the observations, but does not follow them exactly. The reason is that the model dynamics has a voice here too. It is difficult to say which is correct, the assimilated run or the observations. A possible source of error for the observations, which is of importance for the shedding, is the time mean field, as mentioned above. This points to the van Leeuwen (1999a) paper in which the time mean field was the main unknown in the data assimilation procedure. We come back to this issue later on.

The fact that the pure ensemble run and the assimilated run are so different points to the importance of the ageostrophic terms, in combination with the other two potentially problematic artifacts of the model mentioned above. To illustrate this, Fig. 10 shows the evolution of a strong constraint run. This run was obtained by putting the errors in the model dynamics to zero and repeating the assimilation experiment. The assimilated field at day 0 is then the initial field for a pure model run forward in time, and the result of that run is depicted in Fig. 8. The boundary conditions for this run are found from the model run itself, as explained above. One could also determine the optimal boundary conditions for a strong constraint inversion, but that is not done here. A ring is being shed on day 40, close to the observations, but the next ring, centered at 39°S, 19°E, is too large and moves southward instead of westward. So, although the run is close to the observations initially, after about 40 days the model (as it is) is unable to show a realistic evolution. This is probably related to the sensitivity of nonlinear systems to initial conditions. (Optimal boundary conditions might have led to a better model evolution, but given the fact that the ring moves too far south even when it is still far from the boundary, this is unlikely.)

Another promising feature in the weak constraint inversion (Fig. 9) is the appearance of cyclonic depressions northward of the retroflection area, at 36°S, 18°E. These features seem to play an important role in the final shedding, as can be seen in Fig. 9. In Fig. 11 a space–time diagram is depicted showing the sea level elevation along the maximum mean velocity in the Agulhas Current as it runs along the continent. Clearly visible is a depression that starts at day 0 and moves downstream with a velocity of about 25 km day^{-1}. Meanders of this kind have been found by Lutjeharms (1989) and Lutjeharms and Roberts (1988) in infrared satellite images and named Natal pulses, after their place of initiation in the upstream Agulhas, in the Natal Bight area. They move along with the Agulhas Current with a speed of $\approx$20 km day^{-1}. The fact that the meanders move faster in our simulation is probably related the highly simplified vertical structure in our model. De Ruijter et al. (1999b) found from Geosat, *ERS-1*, TOPEX/Poseidon, and infrared satellite observations that five to six pulses are formed every year. They grow in size from 30 km at initiation up to a few 100 km in the ring-shedding area. Their formation in probably related to barotropic instability of the Agulhas Current in the Natal Bight (see van Leeuwen et al. 2000). Van Leeuwen et al. (2000) have shown that these features have a strong relation to the shedding of Agulhas rings: each ring is preceded by a Natal pulse. So these pulses might play an important role in the global transport of heat and salt in the ocean.

These findings show one of the advantages of using a smoother over a filter. The cyclonic depression is not in the model dynamics (see Fig. 4) but arises from the observations. However, no observations are incorporated along the east coast of South Africa. What happened is that the information of the observations is transported back in time by the model evolution measurement operator covariance, the representers. We thus see that with a smoother it is possible to trace features back in time that are not present in a pure model run due to incorrect initial and boundary conditions and of which only the end product is seen in the observations. By neglecting errors in model dynamics, a strong-constraint inversion, Natal pulses do arise (see Fig. 10), but they do not have the correct influence on the shedding of the Agulhas rings. The fact that Natal pulses do arise is due to the possibility of the data assimilation method to propagate information back in time, thus to alter the initial condition. However, in this simulation it is essential to include errors in the model dynamics to obtain a correct evolution of the fields.

It is interesting to compare these results with the ensemble Kalman filter calculations of Evensen and van Leeuwen (1996). In their Fig. 7 a large meander develops at the position where the Agulhas Current makes a right turn onto the continental shelf south of the con-

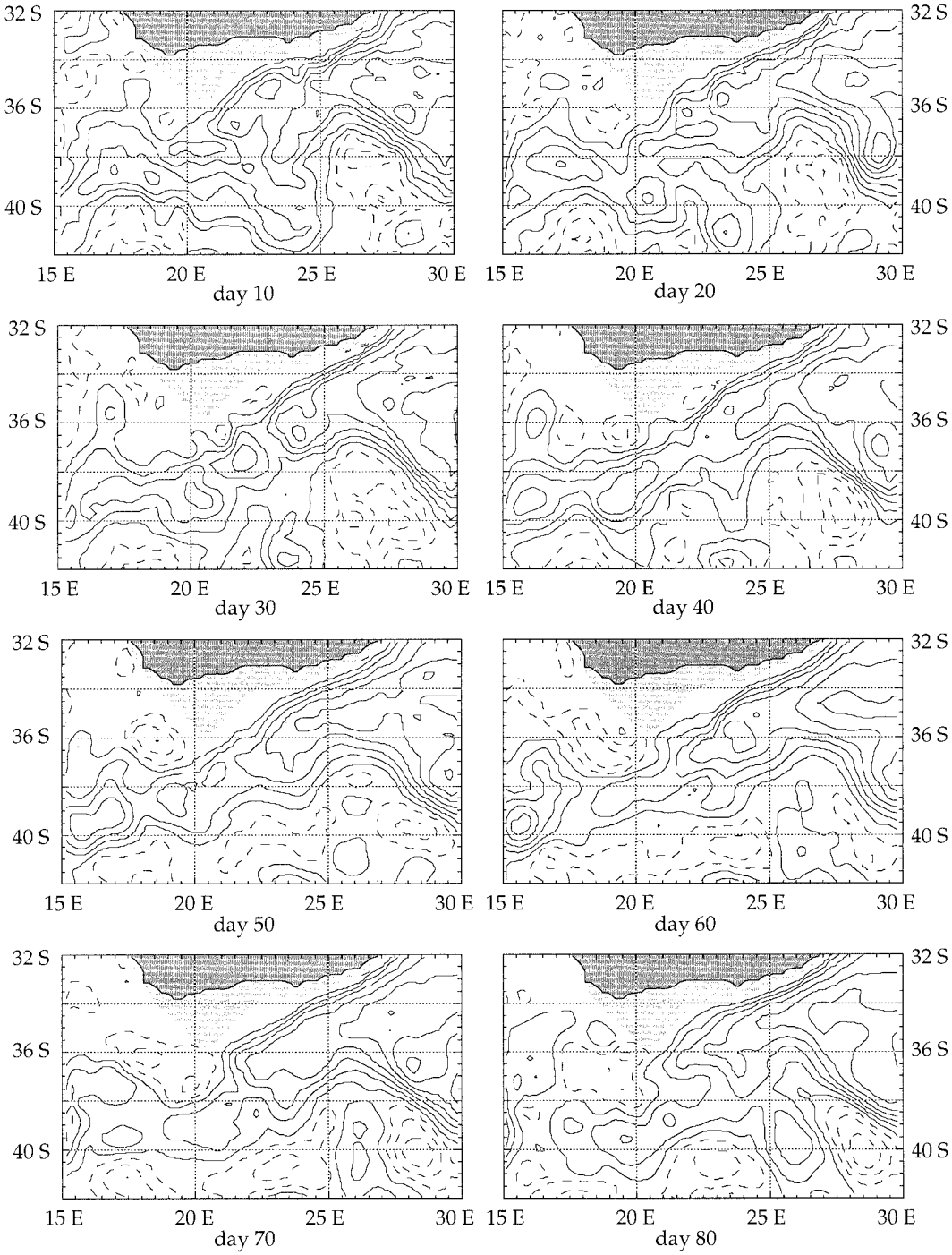

FIG. 9. Posterior streamfunction of upper layer, contour interval = 20 Sv, dashed lines denote positive values.

tinent. [They did not exclude the continental shelf from their calculations and the assimilated fields are unrealistic in this respect, e.g., see Lutjeharms (1989).] This meander cannot be followed upstream but originates at that location, probably related to inertial overshoot. Also, the meander does not propagate downstream. We can thus conclude that no Natal pulse is present, although it should be there (see van Leeuwen et al. 2000). Still, the situation is not quite comparable. The interpolated Geosat altimeter fields used in that paper do not contain Natal pulses as they travel along the east coast of South Africa, because of their small scale. However, observations are assimilated at those locations in that paper, making it hard for a pulse to develop. But, even

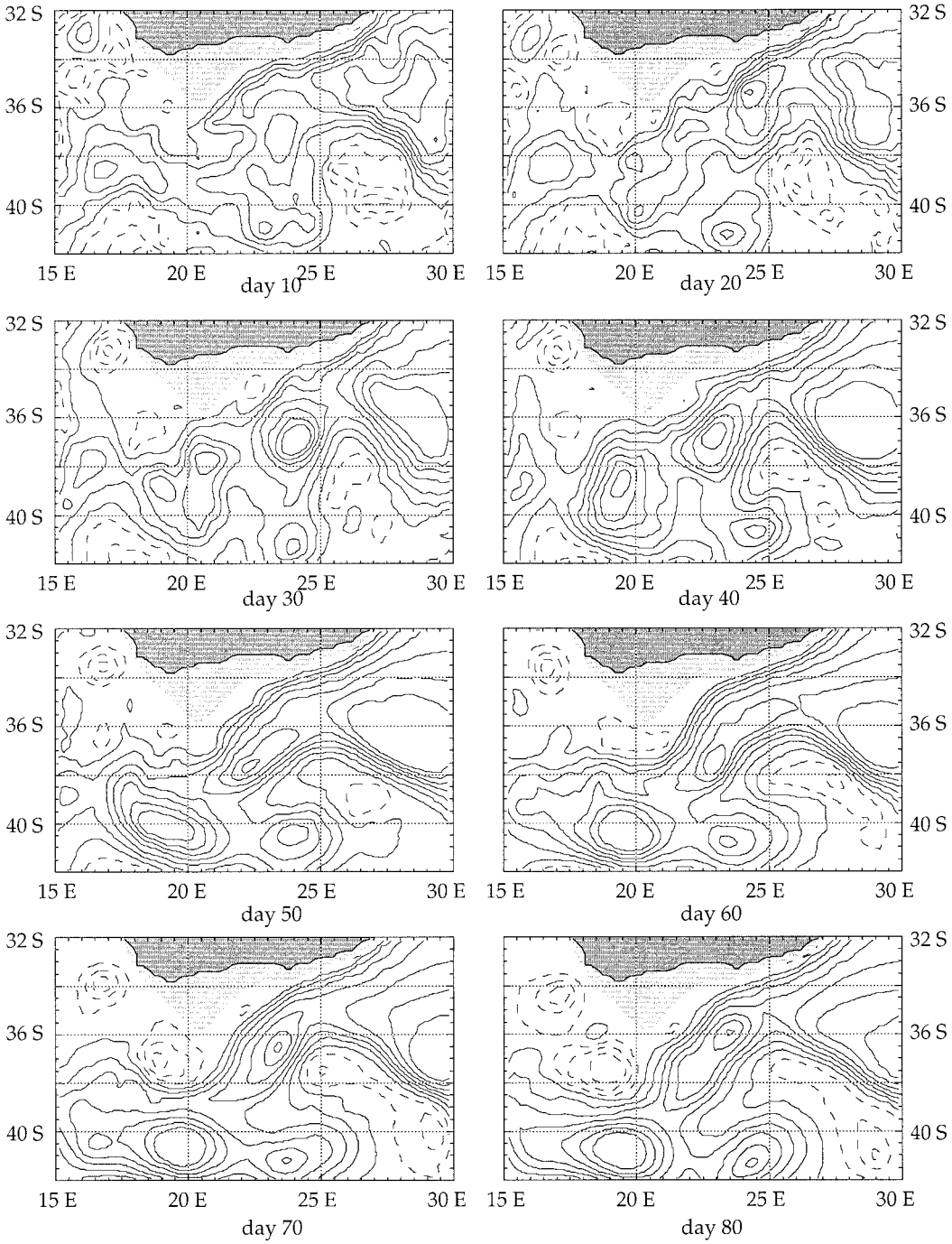

FIG. 10. Strong constraint inversion of upper layer, contour interval = 20 Sv, dashed lines denote positive values.

regarding these matters, it is clear that if one would use the same observation set as in this paper, a filter cannot create Natal pulses. Furthermore, due to sudden changes to the model fields at observation times the process of ring shedding is difficult to study.

In the paper by van Leeuwen (1999a) the same model and observations have been used in the same area to obtain a better description of the time mean circulation in the Agulhas region. In that paper the time mean circulation was taken as the unknown and errors in model dynamics were neglected. An ensemble of model runs that differed in the time mean circulation over a 100-day interval defined the prior probability density of the time mean circulation. Only the time-varying part of the altimeter observations was used to constrain the model evolution and hence the time mean circulation over

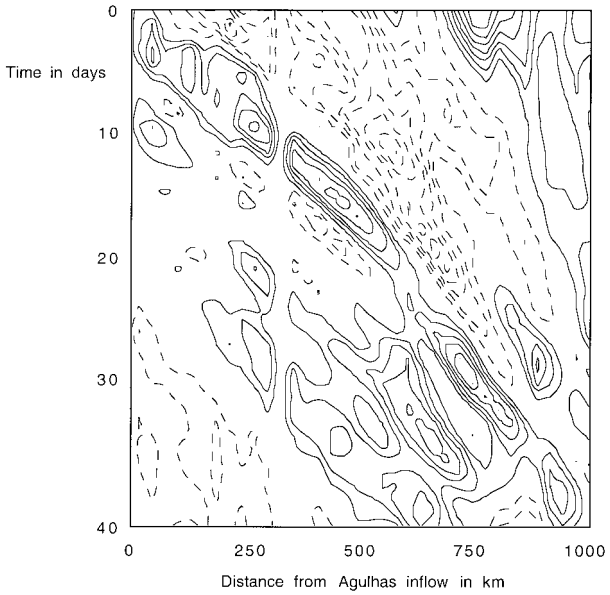

FIG. 11. Space–time diagram showing the propagation of a Natal pulse along the east coast of the African continent in the ensemble smoother inversion. Contour interval 5 Sv.

those 100 days. The representers contain the correlation between the measurements and the time mean field of the model. In Fig. 10 of van Leeuwen (1999a) the time-varying altimeter observations are added to the new time mean circulation. If we compare that figure with Fig. 9 of this paper the circulations have similarities, but also a few striking differences. The relatively strong cyclone west of Africa is found in that paper. The Agulhas rings leave the domain farther north, and Natal pulses seem to be less influential. Previously it was argued that the anticyclonic rings that leave the domain at the western side entrain water resulting in a cyclonic circulation on their northern side. In our case the cyclonic feature west of Africa not present. The reason why we do not find the cyclone might be the following. The cyclonic depression that becomes visible in the observations on day 40 at 37°S, 19°E was traced back in the present paper as a Natal pulse, coming from the Agulhas Current, from the eastern side of the continent. In van Leeuwen (1999a) at least part of this depression seems to originate from the strong cyclone west of the continent. So, in van Leeuwen (1999a) the time mean field was altered by creating a cyclone west of Africa, kept in place by the more northward travel path of Agulhas rings. Note that these areas of the domain were not restricted by observations. In our case the model chose the Natal pulse solution. It is interesting to see how the same model and observations and comparable data assimilations techniques can lead to completely different descriptions of a physical system. Now that data assimilation techniques are maturing, our knowledge of the system, that is, the places where we think the errors cannot be neglected, becomes more and more important.

5. Summary and discussion

In this paper the new data-assimilation method proposed by van Leeuwen and Evensen (1996), the ensemble smoother, is shown to work in a realistic oceanic situation. The generalized inverse of a two-layer quasigeostrophic model of the Agulhas Retroflection area and TOPEX/Poseidon altimeter data was determined. The idea underlying the new method is to combine the probability densities of the model and the observations in a Bayesian way to obtain the probability density of the model given the observations. The general solution to the minimum-variance smoother is presented. The probability density of the observations is assumed to be Gaussian. In addition, the probability density of the model, the prior density, is assumed to be Gaussian in the ensemble smoother. As noted, a similar procedure is adapted for a Kalman filter. The effect of the non-Gaussianity is discussed, and it was shown that the smaller the posterior covariances, the better the approximation. This result gives us a clue to why Kalman filters work for nonlinear models too. Inspection of the third moment of the prior density showed that this assumption of Gaussianity was acceptable for the present experiment.

The probability density of the model, the prior distribution, is described by an ensemble of possible model evolutions. We need 500 members in the present experiment. In an experiment with 600 members, changes in the model fields were below 1%. We even found that only 100 members gave an inverse estimate close to the 500-member case.

An important practical advantage is that no adjoint of the model has to be determined. There are still practical and fundamental problems with adjoint equations for models with discontinuous dynamics (''if statement'') (see Miller et al. 1994b). Furthermore, error estimates are obtained relatively easily, *so the accuracy of the results is known.* This is contrary to all other existing smoothers in which a large extra effort is involved. (Clearly, in analogy to the situation with observations, an inversion without a proper error estimate is of much less value.) In the particular experiment described here the error reduction is about a factor of 6 overall.

By studying the representer matrix an optimal measurement antenna for this model configuration was determined. This was found to be the area in which the Agulhas rings form and shed from the Agulhas Current. Clearly, quasigeostrophic dynamics is insufficient to describe this process accurately. For altimeter measurements this is an academic example because the data coverage is global. However, the same method can be used for in situ measurements too. A large gain in efficiency is expected here.

It has been shown that data assimilation in general requires neither the model nor the observations to be unbiased. This is an important point, because the model biases are generally extremely difficult to determine. If unbiased observations are present one may determine the model bias, as in Dee and da Silva (1998), but biases in observations are difficult to detect in general. Aside from this, it is unclear how to define a model bias precisely. Since the atmosphere or the ocean both vary on all timescales, with most energy in the slowest motions, a statistical steady state cannot be determined (if it exists at all). The point made here is that the present derivation of the smoother does not assume that either model or observations are unbiased. Rather, the more general viewpoint is taken that all the knowledge we have about the system is contained in two probability densities, one from the observations, and one from the dynamical equations, represented by the model. Data assimilation is just combining the two densities, whether they are biased or not.

The question remains why only 500 members are needed for the inversion. After all, the ensemble has to describe the probability density of a quasi-geostrophic model run over 100 days, leading to 12.5 million grid points! An answer can be found in the following way. At each time step we have 5600 unknowns and 500 independent guesses. The factor of about 10 between the two numbers can be explained by the spatial correlation of the error fields, thus by correlations among the 5600 unknowns. The characteristic horizontal length scale of about 50 km combined with a horizontal resolution of 25 km count for a factor of about 4. Another factor of 2 is due to the correlation of the error fields in the vertical, which is about 0.4. So an ensemble of 500 members should do a reasonable job, especially if we take into account that the noise is correlated in time in this experiment. In fact, Evensen and van Leeuwen (1996), van Leeuwen and Evensen (1996), van Leeuwen (1999a), and the present paper all find convergence by using 500 ensemble members. This is less strange than it seems because all papers use two-layer quasigeostrophic ocean models of comparable size, with comparable characteristic length scales. At this stage it is unclear how much ensemble members are needed for a primitive equation model. The number of degrees is much larger, pointing to more ensemble members, but the model errors will be much smaller, or neglible for some variables.

Another important factor lies in the model dynamics. In Fig. 7 we see that the model variance started to level off after about 80 days. This is not likely to happen with a linear model. The nonlinearity of the model keeps the ensemble relatively narrow, although it keeps slowly growing even after 125 days. Strange outliers are quickly removed by nonlinear processes, greatly restricting the number of possible model evolutions. This is probably also the reason why the skewness of the distribution stays relatively low, so the prior distribution remains Gaussian approximately. A more detailed discussion on these matters can be found in, for example, Houtekamer and Derome (1995) and Lermusiaux and Robinson (1999), who concentrate on the error subspace, thus on an optimal description of the model state assuming Gaussian statistics.

It is interesting to compare the smoother results with what can be expected from other methods. A nudging technique has been use by Holland et al. (1991), Cai (1994), Stutzer and Kraus (1998), Florenchie and Verron (1998), and Zhang and Marotzke (1999) in this area of the world ocean. Florenchie and Verron (1998) use an iterative procedure to obtain a time mean field that is consistent with the TOPEX/Poseidon altimeter data. The nudging technique can be viewed as a mixture of a filter and a smoother, in the sense that the model notices the observations ahead in time. However, this is just used to pull the model to the data, not to propagate information from the observation back in time. It was shown that the observations detected cyclonic structures that seemed to influence the ring shedding process. Thus, when the mechanism of ring shedding is studied, it is important to know where these cyclonic features came from. The smoother traced them back to so-called Natal pulses, meanders in the Agulhas Current when it flows along the east coast of South Africa, an area void of observations. With the nudging technique the smoother interval is relatively short, at least less than the time interval between observations at a certain point, so the technique will not be able to produce Natal pulses. Obviously, all filters have the same problem.

Another problem with nudging and related methods is that it is unclear how to judge the final results because no error information is available. It is clear that each estimate should contain an estimate of how good that particular estimate is. Because the model is forced to the observations, the shedding of Agulhas Rings will be according to the observations. However, the process of ring shedding cannot be studied very accurately because the model dynamics are severely violated around times when observations are present. This latter point also holds for filters in general.

A strong constraint assimilation procedure shows their origin, but the difference with the present results is the following. Both the ensemble smoother and the strong constraint method will transport the feature all the way back to the northward boundary consistent with quasigeostrophic dynamics. However, in a strong constraint approach the feature also moves downstream forward in time consistent with quasigeostrophic dynamics. As was shown, the strong constraint method is in agreement with the observations initially but deviates from it when rings are to be shed, some 40 days later. It seems that quasigeostrophic dynamics is not adequate for describing the ring-shedding process because of strong ageostrophic motions at the separation point and possible resolution problems. Implementations of the strong constraint method using a gradient descent algorithm

have the further disadvantage that error estimates are difficult to obtain for nonlinear model dynamics.

To conclude, neither a filter nor a strong constraint technique will be able to describe the mechanism responsible for ring shedding the way a weak constraint smoother can. The ensemble smoother seems to be an interesting method because it is relatively cheap to obtain both the inverse solution and a proper error estimate. A final remark is that it is extremely simple in ensemble methods to include model errors that are nonwhite in time. For methods based on the minimalization of a cost function this is much more problematic.

As a final point it is mentioned that none of the abovementioned methods has attacked the linear update equation. One of strong points in the ensemble Kalman filter is that each ensemble member can be updated independent from all others. This will not be the case for an update equation that takes the nonlinearities into account. So the question here is how to update an ensemble without destroying the nonlinear features (in probabilistic sense) that are represented by the ensemble. This is an area of active research.

Acknowledgments. P. J. van Leeuwen was sponsored by the Space Research Organization Netherlands (SRON) under Grant EO-002; by the National Research Program, Grant 013001237.10; and by the EC MAST-III project DIADEM under Contract MAS3-CT98-0167. All calculations were performed on the CRAY-C98 under Contracts SC-399 and SC-527, 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.

APPENDIX

The Posterior Error Covariance

In this appendix a derivation is given of the posterior covariance using the probabilistic description. It is noted beforehand that again no reference is made to an unknown true model evolution, so this derivation is valid even if biases in model and/or observations are present. We start with Eq. (14), which we write as

$$\hat{\psi} = \psi_F + \mathbf{r}^T\mathbf{b} + \alpha, \tag{A1}$$

in which α contains the effect of the non-Gaussian part of the prior probability density. The posterior covariance can be written as

$$\begin{aligned} Q_{\hat{\psi}\hat{\psi}} &= \int (\psi - \hat{\psi})^2 f(\psi|\mathbf{d})\, d\psi \\ &= A \int (\psi - \hat{\psi})^2 f_G(\psi) f_N(\psi) f(\mathbf{d}|\psi)\, d\psi. \end{aligned} \tag{A2}$$

If we use (A1) in this equation we obtain

$$Q_{\hat{\psi}\hat{\psi}} = A \int [(\psi - \psi_F)^2 - 2(\mathbf{r}^T\mathbf{b} + \alpha)(\psi - \psi_F)] f(\psi|\mathbf{d})\, d\psi + \mathbf{r}^T\mathbf{b}\mathbf{r}^T\mathbf{b} + 2\mathbf{r}^T\mathbf{b}\alpha + \alpha\alpha. \tag{A3}$$

Now use

$$\psi - \psi_F = -Q_{\psi\psi} \bullet \frac{\delta f_G}{\delta \psi} \tag{A4}$$

and perform the partial integrations to find

$$Q_{\hat{\psi}\hat{\psi}} = Q_{\psi\psi} + \int [(\psi - \psi_F) - 2(\mathbf{r}^T\mathbf{b} + \alpha)]\mathbf{r}^T\mathbf{W}[\mathbf{d} - \mathbf{L}(\psi)] f(\psi|\mathbf{d})\, d\psi + \mathbf{r}^T\mathbf{b}\mathbf{r}^T\mathbf{b} + \hat{\gamma}, \tag{A5}$$

in which $\hat{\gamma}$ is given by

$$\hat{\gamma} = A \int [(\psi - \psi_F) - 2(\mathbf{r}^T\mathbf{b} + \alpha)] Q_{\psi\psi} \bullet \frac{\delta f_N}{\delta \psi} f_G(\psi) f(\mathbf{d}|\psi)\, d\psi + 2\mathbf{r}^T\mathbf{b}\alpha + \alpha\alpha. \tag{A6}$$

The trick that is used now is to add $\hat{\psi}$ to $(\psi - \psi_F)$ in (A5) and subtract it again. Do the same with $\mathbf{L}(\hat{\psi})$ in $[\mathbf{d} - \mathbf{L}(\psi)]$ in the same integral. By performing the multiplication of the two terms and noting that from (12)

$$\begin{aligned} \mathbf{b} = \mathbf{W}[\mathbf{d} - \mathbf{L}(\hat{\psi})] + A(\mathbf{R} + \mathbf{W}^{-1})^{-1}\mathbf{L}(Q_{\psi\psi}) \\ \bullet \int \frac{\delta f_N(\psi)}{\delta \psi} f_G(\psi) f(\mathbf{d}|\psi)\, d\psi, \end{aligned} \tag{A7}$$

we arrive at

$$Q_{\hat{\psi}\hat{\psi}} = Q_{\psi\psi} - \mathbf{r}^{\mathrm{T}}\mathbf{w}\mathbf{L}(Q_{\hat{\psi}\hat{\psi}}) + \gamma, \tag{A8}$$

in which γ is given by

$$\gamma = \hat{\gamma} + \mathbf{r}^{\mathrm{T}}\mathbf{b}\alpha + (\mathbf{r}^{\mathrm{T}}\mathbf{b} + \alpha)A(\mathbf{R} + \mathbf{w})^{-1}\mathbf{L}(Q_{\psi\psi}) \cdot \int \frac{\delta f_N(\psi)}{\delta\psi} f_G(\psi) f(\mathbf{d}\,|\,\psi)\, d\psi. \tag{A9}$$

To proceed we measure this equation (a procedure followed before):

$$\mathbf{L}(Q_{\hat{\psi}\hat{\psi}}) = \mathbf{r} - \mathbf{R}\mathbf{w}\mathbf{L}(Q_{\hat{\psi}\hat{\psi}}) + \mathbf{L}(\gamma) \quad \text{or} \tag{A10}$$

$$\mathbf{L}(Q_{\hat{\psi}\hat{\psi}}) = (\mathbf{1} + \mathbf{R}\mathbf{w})^{-1}[\mathbf{r} + \mathbf{L}[\gamma]]. \tag{A11}$$

Use this result in (A8) to find finally

$$Q_{\hat{\psi}\hat{\psi}} = Q_{\psi\psi} - \mathbf{r}^{\mathrm{T}}(\mathbf{R} + \mathbf{w})^{-1}\mathbf{r} + \gamma - \mathbf{r}^{\mathrm{T}}(\mathbf{R} + \mathbf{w})^{-1}\mathbf{L}[\gamma]. \tag{A12}$$

The last two terms are called β in the main text. If the non-Gaussian term are neglected we end up with the linear update as can be found in Bennett (1992), among others, and as is used in the Kalman filter.

REFERENCES

Bennett, A. F., 1992: *Inverse Methods in Physical Oceanography*. Cambridge University Press, 346 pp.

——, and M. A. Thorburn, 1992: The generalized inverse of a nonlinear quasi-geostrophic ocean circulation model. *J. Phys. Oceanogr.,* **22,** 213–230.

Brasseur, P., J. Ballabrera-Poy, and J. Verron, 1999: Assimilation of altimetric data in the mid-latitude oceans using the Singular Evolutive Extended Kalman filter with an eddy-resolving, primitive equation model. *J. Mar. Syst.,* **22,** 269–294.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the analysis scheme of the Ensemble Kalman Filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

Cai, W. J., 1994: Circulation driven by observed surface thermohaline fields in a course resolution ocean general circulation model. *J. Geophys. Res.,* **99,** 10 163–10 181.

Dee, D. P., and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.,* **124,** 269–295.

De Ruijter, W. P. M., A. Biastoch, S. S. Drijfhout, J. R. E. Lutjeharms, R. P. Matano, T. Pichevin, P. J. van Leeuwen, and W. Weijer, 1999a: Indian-Atlantic interocean exchange: Dynamics, estimation and impact. *J. Geophys. Res.,* **104,** 20 885–20 910.

——, P. J. van Leeuwen, and J. R. E. Lutjeharms, 1999b: Triggering mechanisms for solitary meanders in the Agulhas Current. *J. Phys. Oceanogr.,* **29,** 3043–3055.

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

Egbert, G. D., A. F. Bennett, and M. G. G. Foreman, 1994: TOPEX/POSEIDON tides estimated using a global inverse model. *J. Geophys. Res.,* **99** (C12), 24 821–24 852.

Evensen, G., 1992: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *J. Geophys. Res.,* **97** (C11), 17 905–17 924.

——, 1994a: Inverse methods and data assimilation in nonlinear ocean models. *Physica D,* **77,** 108–129.

——, 1994b: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

——, and P. J. van Leeuwen, 1996: Assimilation of Geosat altimeter data for the Agulhas current using the ensemble kalman filter with a quasi-geostrophic model. *Mon. Wea. Rev.,* **124,** 85–96.

——, and ——, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.,* **128,** 1852–1867.

Florenchie, P., and J. Verron, 1998: South Atlantic ocean circulation: Simulation experiments with a quasi-geostrophic model and assimilation of TOPEX/Poseidon and ERS 1 altimeter data. *J. Geophys. Res.,* **103,** 24 737–24 758.

Gelb, A., Ed., 1974: *Applied Optimal Estimation.* The MIT Press, 212 pp.

Gordon, A. L., 1982: *Southern Ocean Atlas.* Columbia University Press, 234 pp.

——, 1986: Interocean exchange of thermocline water at the Agulhas Retroflection. *J. Geophys. Res.,* **91,** 5037–5046.

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

Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.,* **123,** 2181–2196.

——, and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

——, and ——, 1999: Reply. *Mon. Wea. Rev.,* **127,** 1378–1379.

Ikeda, M., and J. R. Apel, 1981: Mesoscale eddies detached from spatially growing meanders in an eastward-flowing oceanic jet using a two-layer quasi-geostrophic model. *J. Phys. Oceanogr.,* **11,** 1638–1661.

——, J. A. Johannessen, K. Lygre, and S. Sandven, 1989: A process study of mesoscale meanders and eddies in the Norwegian coastal current. *J. Phys. Oceanogr.,* **19,** 20–35.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376 pp.

Lermusiaux, P. F. J., and A. R. Robinson, 1999: Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Mon. Wea. Rev.,* **127,** 1385–1407.

Lutjeharms, J. R. E., 1989: The role of mesoscale turbulence in the Agulhas Current system. *Mesoscale/Synoptic Coherent Structures in Geophysical Turbulence,* J. C. J. Nihoul and B. M. Jamart, Eds., Elsevier, 357–372.

——, and H. R. Roberts, 1988: The natal pulse: An extreme transient on the Agulhas current. *J. Geophys. Res.,* **93,** 631–645.

Miller, R. N., M. Ghil, and F. Gauthiez, 1994a: Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.,* **51,** 1037–1056.

——, E. D. Zaron, and A. F. Bennett, 1994b: Data assimilation in models with convective adjustment. *Mon. Wea. Rev.,* **122,** 2607–2613.

——, E. F. Carter, and S. T. Blue, 1999: Data assimilation into nonlinear stochastic models. *Tellus,* **51A,** 167–194.

Milliff, R. F., 1990: Modified capacitance matrix method to implement coastal boundaries in the Harvard open ocean model. *Math. Comp. Simulation,* **31,** 541–564.

Olson, B. D., and R. E. Evans, 1986: Rings of the Agulhas Current. *Deep-Sea Res.,* **33,** 27–42.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics.* Springer-Verlag, 710 pp.

Pham, D. T., J. Verron, and M. C. Roubaud, 1998: Singular evolutive extended Kalman filter with EOF initialization for data assimilation in oceanography. *J. Mar. Syst.,* **16,** 323–340.

Stutzer, S., and W. Kraus, 1998: Mean circulation and transports in the South Atlantic Ocean: Combining model and drifters data. *J. Geophys. Res.,* **103,** 30 985–31 002.

Todling, R., and S. E. Cohn, 1994: Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Mon. Wea. Rev.,* **122,** 2530–2557.

Tsaoussi, L., and C. Koblinsky, 1994: 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.

van Ballegooyen, R. C., M. L. Gruündlingh, and J. R. E. Lutjeharms,

1994: Eddy fluxes of heat and salt from the southwest Indian Ocean into the southeast Atlantic Ocean: A case study. *J. Geophys. Res.,* **99,** 14 053–14 070.

van Leeuwen, P. J., 1999a: The time-mean circulation in the Agulhas region determined with the ensemble smoother. *J. Geophys. Res.,* **104,** 1393–1404.

——, 1999b: Comment on ''Data assimilation using an ensemble Kalman filter technique.'' *Mon. Wea. Rev.,* **127,** 1374–1377.

——, and G. Evensen, 1996: Data assimilation and inverse problems in a probabilistic formulation. *Mon. Wea. Rev.,* **124,** 2898–2913.

——, W. P. M. de Ruijter, and J. R. E. Lutjeharms, 2000: Natal pulses and the formation of Agulhas Rings. *J. Geophys. Res.,* **105,** 6425–6436.

Verlaan, M., and A. W. Heemink, 1995: Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrol. Hydraul.,* **11,** 349–368.

Veronis, G., 1973: Model of the world ocean circulation: I. Wind-driven two-layer. *J. Mar. Res.,* **31,** 228–288.

Zhang, K. Q., and J. Marotzke, 1999: The importance of open-boundary estimation for an Indian Ocean GCM-data synthesis. *J. Mar. Res.,* **57,** 305–334.