DELFT UNIVERSITY OF TECHNOLOGY

REPORT 00-03

VARIANCE REDUCED ENSEMBLE KALMAN FILTERING

A.W. HEEMINK, M. VERLAAN AND A.J. SEGERS

ISSN 1389-6520

Reports of the Department of Applied Mathematical Analysis

Delft 2000

Copyright © 2000 by Department of Applied Mathematical Analysis, Delft, The Netherlands.

No part of the Journal may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission from Department of Applied Mathematical Analysis, Delft University of Technology, The Netherlands.

VARIANCE REDUCED ENSEMBLE KALMAN FILTERING

A.W. Heemink, M. Verlaan and A.J. Segers

Abstract

A number of algorithms to solve large scale Kalman filtering problems has been introduced recently. The ensemble filter represents the probability density of the state estimate by a finite number of randomly generated system states. Other algorithms use a singular value decomposition to select the leading eigenvectors of the covariance matrix of the state estimate and to approximate the full covariance matrix by a reduced rank matrix. Both algorithms, however, still require a huge amount of computer resources. In this paper we propose to combine the two algorithms and to use a reduced rank approximation of the covariance matrix as a variance reductor for the ensemble filter. If the leading eigenvectors explain most of the variance, which is the case for most applications, the computational burden to solve the filtering problem can be reduced significantly (up to an order of magnitude).

1 Introduction

Kalman filtering is a powerful framework for solving data assimilation problems (Ghil and Malanotte-Rizzoli 1991). In order to use a Kalman filter for assimilating data into a numerical model, this model is embedded into a stochastic environment by introducing a system noise process. In this way it is possible to take into account the inaccuracies of the underlying deterministic system. By using a Kalman filter, the information provided by the resulting stochastic-dynamic model and the noisy measurements are combined to obtain an optimal estimate of the state of the system. The standard Kalman filter implementation however would impose an unacceptable computational burden. In order to obtain a computationally efficient filter simplifications have to be introduced.

The Ensemble Kalman filter (EnKF) was introduced by (Evensen 1994) and has been used successfully in many applications (Evensen and Van Leeuwen 1996), (Houtekamer and Mitchel 1998), (Canizares 1999). This Monte Carlo approach is based on a representation of the probability density of the state estimate by a finite number N of randomly generated system states. The algorithm does not require a tangent linear model and is very easy to

implement. The computational effort required for the EnKF is approximately N times as much as the effort required for the underlying model. The only serious disadvantage is that the statistical error of the approach decreases very slowly for larger sample size. This is a well known fundamental problem with all Monte Carlo methods. As a result for most practical problems the sample size has to be chosen rather large.

Another approach to solve large scale Kalman filtering problems is to approximate the full covariance matrix of the state estimate by a matrix with reduced rank. This approach was introduced by (Cohn and Todling 1995), (Cohn and Toding 1996) and (Verlaan and Heemink 1995), (Verlaan and Heemink 1997), where the latter used a more robust square root formulation for the filter implementation. Algorithms based on similar ideas have been proposed and applied by (Lermusiaux 1997) and (Pham, Verron, and Rouband 1998).

The reduced rank approaches can also be formulated as an Ensemble Kalman filter where the q ensemble members have not been chosen randomly, but in the directions of the q leading eigenvectors of the covariance matrix (Verlaan and Heemink 1997). As a result also these algorithms do not require a tangent linear model. The computational effort required is approximately $q + 1$ model simulations plus the computations required for the singular value decomposition ($O(q^3)$). In many practical problems the full covariance can be approximated accurately by a reduced rank matrix with relatively small value of q . However, reduced rank approaches often suffer from filter divergence problems for small values of q . This was observed first by (Cohn and Todling 1995) who tried several methods of compensating for the truncation error. The main reason for the occurrence of filter divergence is the fact that truncation of the eigenvectors of the covariance matrix implies that the covariance is allways approximated from below. It is well-known that underestimating the covariance may cause filter divergence. Filter divergence can be avoided by chosing q relatively large, but this of course reduces the computational efficiency of the method considerably.

(Canizares 1999) compared the ensemble filter and the reduced rank approach for a number of practical large scale shallow water flow problems. His conclusion is that the computational efficiency of both approaches is comparable. The rank q of the reduced rank approximation could be chosen considerable smaller (3-5 times) than the ensemble size N of the ensemble filter. However the singular value decomposition becomes computationally the most expensive part of the reduced rank algorithm for larger values of q .

In this paper we propose to combined the Ensemble Kalman filter with the reduced rank approach by using the reduced rank approximation as a variance reductor for the ensemble filter (Hammersley and Handscomb 1964). The ensemble of the new filter algorithm now consists of two parts: q ensembles in the direction of the q leading eigenvalues of the covariance matrix and N randomly chosen ensembles. In the algorithm only the projection of the random ensemble members orthogonal to the first ensemble members is used to obtain the filter gain and the state estimate. This Partially Orthogonal Ensemble Kalman filter (POEnKF) does not suffer from divergence problems because the reduced rank approx-

imation is embedded in an ensemble filter. The ensemble filter acts as a compensating mechanism for the truncation error. At the same time POEnKF is much more accurate than the ensemble filter with ensemble size $N + q$ because the leading eigenvectors of the covariance matrix are computed accurately using the full (extended) Kalman filter equations.

In Section 2 of this paper we summarize the Ensemble Kalman filter and the Reduced Rank Square Root filter and introduce the Partially Orthogonal Ensemble Kalman filter algorithm and a few variants of this algorithm. We illustrate the performance of the various algorithms with an advection diffusion model application in Section 3. Here, in order to compare the results of the various suboptimal filter algorithms with the exact Kalman filter we concentrate our attention to linear problems. For a detailed discussion on the performance of a number of Kalman filter algorithms in highly nonlinear systems the reader is referred to (Verlaan and Heemink 1999).

2 Kalman filtering

2.1 State space model

Suppose that modelling techniques have provided us with a non-linear stochastic state space model of the form:

$$x(k+1) = f(x(k)) + \Gamma(k)w(k) \quad (1)$$

and with a linear measurement relation:

$$y(k) = M(k)x(k) + v(k) \quad (2)$$

where $x(k)$ is the state, $y(k)$ is the measurement vector, $w(k)$ and $v(k)$ are independent Gaussian white noise processes with covariance $\Sigma_w(k)$ and $\Sigma_v(k)$ respectively. The initial condition $x(0)$ is assumed to be Gaussian with mean $\hat{x}_0$ and covariance P_0 .

2.2 Ensemble Kalman filter

The Ensemble Kalman filter algorithm is based on a representation of the probability density of the state estimate by a finite number N of randomly generated system states.

The Ensemble Kalman filter (EnKF) for the model (1) – (2) can be summarized as follows (Burgers, Van Leeuwen, and Evensen 1998):

initialisation

An ensemble of N initial states $\xi_i(0|0)$ are generated to represent the uncertainty in $x(0)$.

time update

$$\xi_i(k+1|k) = f(\xi_i(k|k)) + \Gamma(k)w_i(k) \quad (3)$$

$$\hat{x}(k+1|k) = \frac{1}{N} \sum_{i=1}^N \xi_i(k+1|k) \quad (4)$$

$$E(k+1|k) = [\xi_1(k+1|k) - \hat{x}(k+1|k) \dots \xi_N(k+1|k) - \hat{x}(k+1|k)] \quad (5)$$

measurement update:

$$P(k+1|k) = \frac{1}{N-1} E(k+1|k) E(k+1|k)^T \quad (6)$$

$$K(k+1) = P(k+1|k) M(k+1)^T [M(k+1) P(k+1|k) M(k+1)^T + \Sigma_v(k+1)]^{-1} \quad (7)$$

$$\xi_i(k+1|k+1) = \xi_i(k+1|k) + K(k+1)(y(k+1) - M(k+1)\xi_i(k+1|k) + v_i(k+1)) \quad (8)$$

where $\xi_i(k|k)$ is an ensemble of state vectors generated with the realizations $w_i(k)$ and $v_i(k)$ of the noise processes $w(k)$ and $v(k)$ respectively. Note that in the final implementation of the algorithm P need not actually be computed (Evensen 1994).

For most practical problems equation (3) is computationally dominant. As a result the computational effort required for the EnKF is approximately N model simulations. The errors in the state estimate are of statistical nature and converge very slowly with the sample size ($\approx \frac{1}{\sqrt{N}}$). This is one of the very few drawbacks of this Monte Carlo approach.

2.3 Reduced Rank Square Root filter

The Reduced Rank Square Root (RRSQRT) filter algorithm (Verlaan and Heemink 1997) is based on a factorization of the covariance matrix P of the state estimate according to $P = LL^T$, where L is a matrix with the q leading eigenvectors l_i , $i = 1, \dots, q$, of P as columns. The algorithm can be summarized as follows:

initialisation

$$\begin{aligned} x(0|0) &= x_0 \\ L(0|0) &= [l_1(0|0) \dots l_q(0|0)] \end{aligned} \quad (9)$$

time update

$$\hat{x}(k+1|k) = f(\hat{x}(k|k)) \quad (10)$$

$$l_i(k+1|k) = \frac{1}{\varepsilon} (f(\hat{x}(k|k) + \varepsilon l_i(k|k)) - f(\hat{x}(k|k))) \quad (11)$$

$$\tilde{L}(k+1|k) = [l_1(k+1|k) \dots l_q(k+1|k) \quad \Gamma(k)\Sigma_w(k)^{1/2}] \quad (12)$$

$$L(k+1|k) = \Pi(k+1|k) \tilde{L}(k+1|k) \quad (13)$$

where $\Pi(k+1|k)$ is a projection onto the q leading eigenvectors of the matrix $\tilde{L}(k+1|k)\tilde{L}(k+1|k)^T$ and where ε is chosen close to 1.

measurement update

$$P(k+1|k) = L(k+1|k)L(k+1|k)^T \quad (14)$$

$$K(k+1) = P(k+1|k)M(k+1)^T[M(k+1)P(k+1|k)M(k+1)^T + \Sigma_v(k+1)]^{-1} \quad (15)$$

$$\hat{x}(k+1|k) = \hat{x}(k+1|k) + K(k+1)(y(k+1) - M(k+1)\hat{x}(k+1|k)) \quad (16)$$

$$\tilde{L}(k+1|k+1) = [I - K(k+1)M(k+1) \quad K(k+1)\Sigma_v(k+1)] \quad (17)$$

$$L(k+1|k+1) = \Pi(k+1|k+1)\tilde{L}(k+1|k+1) \quad (18)$$

where $\Pi(k+1|k+1)$ is a projection onto the q leading eigenvectors of the matrix $\tilde{L}(k+1|k+1)\tilde{L}(k+1|k+1)^T$. Note that the full covariance need not to be computed (Verlaan and Heemink 1997). This measurement update is not the most efficient procedure. Equations (17)–(18) are, however, more general than measurement update proposed by (Verlaan and Heemink 1997). Equations (17)–(18) holds for arbitrary filter gains $K(k)$ and not only for gain matrices satisfying equations (15). This becomes important if the RRSQRT algorithm is used as part of the Partially Orthogonal Ensemble Kalman filter described in the next Subsection.

For smaller values of q equation (11) is computationally dominant resulting in a computational effort of $q+1$ model simulations. The projection (13) requires $O(q^3)$ computations. As a result for very large q this part of the algorithm becomes time consuming too (see (Canizares 1999)). The errors of the algorithm are caused by the linearized dynamics in equation (11) and the representation of the covariance matrix by only the q leading eigenvectors. Because of the truncation the covariance matrix is approximated from below. As a result the algorithm is sensitive to filter divergence problems. This problem can be avoided by choosing q relatively large, but this obviously reduces the computational efficiency.

2.4 Partially Orthogonal Ensemble Kalman filter

The Ensemble Kalman filter and the Reduced Rank Square Root filter are both of the more robust square root type since both algorithms are formulated in terms of the square root of the covariance matrix (respectively $E(k+1|k)$ and $L(k+1|k)$). Also, both algorithms are of the ensemble type. The ensemble filter is based on N randomly chosen ensemble members, while in the reduce rank filter the ensemble members are chosen deterministically in the direction of the q leading eigenvectors. Because the two algorithms both have a very similar algorithmic structure, they can be integrated relatively easy.

The ensemble of the Partially Orthogonal Ensemble Kalman filter (POEnKF) consists of two parts. The q leading eigenvectors l_i of the covariance matrix plus N randomly chosen ensemble members ξ_i . The ensemble members l_i are updated using a RRSQRT algorithm and ξ_i by using the EnKF algorithm. The two algorithms interact with each other at the measurement update. Here only the information of the random ensemble members orthogonal to the q leading eigenvectors is used.

The POEnKF algorithm can be summarized as follows:

initialisation

$$[L(0|0) \quad E(0|0)] = [l_1(0|0) \dots l_q(0|0) \quad \xi_{q+1}(0|0) \dots \xi_{q+N}(0|0)] \quad (19)$$

where l_i are the leading eigenvector of P_0 and ξ_i are generated randomly to represent the uncertainty in $x(0)$.

time update

time update equations (3)–(5) of the EnKF algorithm

time update equations (9)–(13) of the RRSQRT algorithm

measurement update

$$E^\perp(k+1|k) = \Pi^\perp(k+1|k)E(k+1|k) . \quad (20)$$

where $\Pi^\perp(k+1|k)$ is a projection of the random ensemble members orthogonal to the first q ensemble members l_i .

$$P(k+1|k) = L(k+1|k)L(k+1|k)^T + \frac{1}{N-1}E^\perp(k+1|k)E^\perp(k+1|k) \quad (21)$$

$$K(k+1) = P(k+1|k)M(k+1)^T[M(k+1)P(k+1|k)M(k+1)^T + \Sigma_v(k+1)]^{-1} \quad (22)$$

measurement update equation (8) of the EnKF algorithm for the ensemble $E(k+1|k)$

measurement update equations (16)–(18) of the RRSQRT algorithm

For small values of q the time propagation equations (3) and (11) for the ensemble is computationally dominating. As a result for most practical problems the computational effort for the POEnKF is approximately $N + q$ times the effort required for one model simulation. By integrating the Ensemble Kalman filter and the Reduced Rank Square Rootfilter the best of both are combined. The reduced rank part acts as a variance reductor for the ensemble filter reducing the statistical errors of this Monte Carlo approach significantly. At the other hand by embedding the reduced rank filter in an Ensemble Kalman filter the covariance is not underestimated, eliminating the filter divergence problems of the reduced rank approach (also for very small numbers of q). As a result q can be chosen on the basis of efficiency arguments and not for stabilizing the filter algorithm.

Many generalizations and modifications of the POEnKF exists. One important modification to make the POEnFK more robust for strongly nonlinear problems is to take into account the information of the random ensemble members in the direction of the q leading eigenvectors of the covariante matrix. If

$$E''(k+1|k) = \Pi(k+1|k)E(k+1|k) \quad (23)$$

is the projection of the random ensemble members onto the first q ensemble members l_i , then

$$P(k+1|k) = \frac{1}{N-1}E''(k+1|k)E''(k+1|k)^T \quad (24)$$

is the Ensemble Kalman filter approximation of the reduced rank covariance matrix with rank q . Instead of equation (21), where this information is completely ignored, it is also possible to use the equation

$$\begin{aligned} P(k+1|k) = & \gamma L(k+1|k)L(k+1|k) + \frac{1-\gamma}{N-1}E''(k+1|k)E''(k+1|k) \\ & + \frac{1}{N-1}E^\perp(k+1|k)E^\perp(k+1|k)^T \end{aligned} \quad (25)$$

where γ is a weighting coefficient. The first term of the right hand side suffers from errors caused by the approximation of the nonlinear dynamic of the system in the time update, the second term is dominated by the statistical noise of the Monte Carlo method. For most practical problems the errors of the first term are relatively small and $\gamma = 1$ is a good choice. However for very nonlinear data assimilation problems it may be better to choose γ smaller.

2.5 Complementary Orthogonal subspace Filter For Efficient Ensembles

The POEnKF described in the previous section uses a RRSQRT time-update for a small number of directions that correspond to the largest eigenvalues of the forecast error covariance. A separate ensemble update is used for the remaining space. However, the covariance for this orthogonal subspace is obtained by first generating an ensemble for the whole state space and subsequently removing all contributions in the directions with the largest variance. Since the contribution of the ensemble is ignored in the subspace spanned by modes of the RRSQRT-part of the POEnKF it seems more efficient to remove those contributions from the ensemble altogether. This can be achieved with a simple modification of the POEnKF, which will be referred to in the sequel as the Complementary Orthogonal subspace Filter For Efficient Ensembles (COFFEE) algorithm.

The COFFEE algorithm can be summarized as follows:

initialisation

$$[L(0|0) \ E(0|0)] = [l_1(0|0) \dots l_q(0|0) \ \xi_{q+1}(0|0) \dots \xi_{q+N}(0|0)] \quad (26)$$

where l_i again denotes the leading eigenvalues of P_0 . However, the random ensemble members are now generated to represent the covariance $P_0 - L(0|0)L(0|0)^T$.

time update

L is updated using the RRSQRT update (10)-(13) and the ensemble ξ_i is updated using (3)-(5). However, an additional ensemble is generated with a covariance equal to the covariance of the truncated part in the RRSQRT half, i.e. (5) changes to:

$$E(k+1|k) = [\xi_1(k+1|k) - \hat{x}(k+1|k) + \omega_1(k+1) \dots \xi_N(k+1|k) - \hat{x}(k+1|k) + \omega_N(k+1)] \quad (27)$$

with $E(\omega_i(k+1)\omega_i(k+1)') = (I - \Pi(k+1|k))\tilde{L}(k+1|k)\tilde{L}(k+1|k)^T(I - \Pi(k+1|k))^T$.

measurement update

The measurement update is almost identical to that of the POEnKF, except that no projection of the ensemble is needed, i.e. (20) can be omitted.

3 Application

3.1 Model and experimental setup

Consider the 2-D advection diffusion equation

$$\frac{\partial c}{\partial t} + u \frac{\partial c}{\partial x} + v \frac{\partial c}{\partial y} = \nu \frac{\partial^2 c}{\partial x^2} + \nu \frac{\partial^2 c}{\partial y^2} \quad (28)$$

with a square domain $[0, L] \times [0, L]$ and zero initial conditions. The concentration at the boundary is zero for inflow. A backward Lagrangian scheme is used to discretize these equations on a 30 by 30 grid. Preliminary experiments indicated that the numerical diffusion was still so large that the physical diffusion terms could be omitted. For all experiments in the sequel the velocity field was assumed to be known and constant in time. Figure 1 shows the velocity field which is similar to that of the well-known Molenkamp test.

In order to compare the algorithms proposed to other existing algorithms some twin experiments were carried out. A reference solution was generated by inserting constant emissions at grid cells $\{(5, 8), (7, 8), (15, 5), (3, 18), (20, 15)\}$. The increase of concentration per timestep for these location was $\{0.2, 0.5, 0.4, 0.2, 0.25\}$ respectively.

The measurements were generated from simulated true concentrations, which were computed by adding fluctuations to the mean emissions. For this purpose an AR(1) process was used, i.e. the fluctuations to the emissions per timestep were computed according to:

$$\tilde{w}_j(k+1) = \alpha_j \tilde{w}_j(k) + w_j(k) \quad (29)$$

with w_j independent Gaussian white noise processes with $E\{w_j(k)\} = 0$ and $var\{w_j(k)\} = 1$. The index j refers to measurement location $\{(m, n)_1 \dots (m, n)_p\} = \{(4, 11), (13, 3), (26, 15), (15, 10), (23, 2), (11, 9), (15, 20), (23, 10), (5, 25)\}$. The decays per timestep are $\{\alpha_1 \dots \alpha_p\} = \{0.9, 0.8, 0.8, 0.8, 0.9, 0.9\}$.

Negative emissions are truncated, which creates some non-linearity in the data-assimilation problem. Finally, white observational noise with variance 0.01 is added to the true concentrations.

3.2 Results of the suboptimal Kalman filter algorithms

In evaluating the filter algorithms we compute the RMS error

$$RMS := \sqrt{1/(M^2 K) \sum_{m,n,k} (c_{m,n}(k) - \hat{c}_{m,n}(k))^2} \quad (30)$$

where $c_{m,n}(k)$ are the exact generated concentrations and $\hat{c}_{m,n}(k)$ are the estimates computed, M is the number of gridpoints in one direction and K is the number of timesteps.

In Figures 2, 3 the concentration fields of the truth-run and the reference-run are shown after $K = 100$ timesteps. The +-signs indicate measurement locations and the diamond-signs the locations of the emissions. It can be seen clearly that the true fields are perturbed with time-varying fluctuations, while the reference solutions only contains a steady emission which is advected and spreading smoothly. The RMS difference between both runs is 0.665 without any data assimilation.

In the figures 5, 6 we show the results for the Ensemble Kalman filter and the Reduced

Rank Square Root filter at the final time $k = 100$ for 20 ensemble members and 20 modes respectively. These settings correspond to approximately the computational requirements for 20 model-runs. For these settings both filters do not perform well. Their respective RMS errors are 0.838 and 0.797, which is worse than without data-assimilation. The contours clearly show that both algorithms create spurious correlations if not fully converged. This result is consistent with the findings of (Houtekamer and Mitchel 1998) who introduced a data-selection procedure to reduce these effects. Although 20 eigenvalues explain almost 98% of the error covariance both the EnKF and the RRSQRT filter still show divergence problems. See Figure 4 for the variance explained by the leading eigenvectors of the analysis error covariance at the final time $k = 100$, as computed with the RRSQRT filter with 100 modes.

In Figures 7, 8 the results for the Partially Orthogonal Ensemble Kalman filter and the Complementary Orthogonal Filter For Efficient Ensembles are shown at the final time. In both computations 15 modes and an ensemble with $N = 5$ members was used. Although the computational requirements are approximately the same as in the previous experiments these filters do not diverge and actually improve on the reference solution. The RMS's are 0.642 and 0.480 respectively. This also shows that the reduced variance in the COFFEE algorithm as compared to the POEnKF does indeed reduce the errors even further in this case. In other experiments not shown here we also found that of the various variants that are possible for the COFFEE algorithm, it is most efficient to project the Kalman gain onto the subspace spanned by the modes of the RRSQRT part of the algorithm, i.e. to effectively ignore correlation computed from the EnKF part of the algorithm but to include their variance in the innovation error covariance.

Figure 9 shows the convergence of the four suboptimal schemes used in this paper. For the POEnKF and the COFFEE algorithm the number of ensemble members was kept fixed at $N = 5$. We observed during the experiments that for a small number of modes or ensemble members the RMS error are sensitive to the random numbers used in the computation. To reduce this random effect on the results the same set of random numbers was used for all the experiments, which made the results more comparable. However, the seemingly random behaviour for small ensembles (or modes) is still influenced by the set of random numbers used. When comparing the algorithms, it can be seen that although the EnKF is quite robust for small ensemble sizes the convergence is slow, whereas the RRSQRT algorithm seems to diverge up to a certain point and then converge rapidly. It should be noted here that the model used here is nearly linear, so that the better characteristics for nonlinear models of the EnKF are not visible. However, if one uses a very large ensemble, e.g. $N = 500$ the result, $RMS = 0.440$, is slightly better than the limit value for the RRSQRT filter $RMS = 0.464$.

Both the POEnKF and the COFFEE algorithm improve the convergence for small ensemble sizes, where the COFFEE variant is slightly more efficient. However, if the number of modes is increased the random part of these algorithms seems to hamper quick convergence

of the RRSQRT part of the algorithm. In this experiment accurate results are obtained with the COFFEE algorithm at only 40% of the computation time compared to the EnKF ($q = 20$ and $N = 50$).

4 Conclusions

In this paper we, propose to use a Reduced Rank Square Root Kalman filter as a variance reductor for the Ensemble Kalman filter. The resulting Partially Orthogonal Ensemble filter and its variants combine the best of both. It is less sensitive to divergence problems and much faster than both the reduced rank approach and the Ensemble Kalman filter. Next, the POEnKF and COFFEE algorithms will be applied to real life, non-linear data assimilation problems in atmospheric-chemistry and shallow water flow forecasting.

References

- Burgers, G., P.J. Van Leeuwen and G. Evensen (1998). Analysis scheme in the ensemble Kalman filter. *Geophysical Res. Lett.* 26, 1027–1030.
- Canizares, R. (1999). *On the application of Data assimilation in Regional Coastal Models*. Ph. D. thesis, Delft University of Technology. A.A. Balkema, Rotterdam, 133pp.
- Cohn, S.E. and R. Toding (1996). Approximate data assimilation schemes for stable and unstable dynamics. *J. Met. Soc. Japan* 74(1), 63–75.
- Cohn, S.E. and R. Todling (1995). Approximate Kalman filters for unstable dynamics. In *Second Int. symposium on Assimilation of observations in Meteorology and Oceanography*, Tokyo, pp. 241–246. WMO.
- Evensen, E. (1994). Sequential data assimilation with a nonlinear QG model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99, 10143–10162.
- Evensen, G. and P.J. Van Leeuwen (1996). Assimilation of geosat altimeter data for the agulhas current using the Ensemble Kalman filter with a quasigeostrophic model. *Mon. Weather Rev.* 124(1), 85–96.
- Ghil, M. and P. Malanotte-Rizzoli (1991). Data assimilation in meteorology and oceanography. *Advances in Geophysics* 33, 141–266. Academic Press, San Diego, Calif.
- Hammersley, J. and D. Handscomb (1964). *Monte Carlo Methods*. London: Wiley and sons inc.
- Houtekamer, P.L. and A.L. Mitchel (1998). Data assimilation using ensemble Kalman filter technique. *Monthly Weather Reviews* 126, 796–811.
- Lermusiaux, P. (1997). *Error subspace data assimilation methods for ocean field estimation: theory, validation and applications*. Ph. D. thesis, Harvard University, Cambridge, Mass.

- Pham, D., J. Verron, and M. Rouband (1998). A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine systems* 16, 323–340.
- Verlaan, M. and A.W. Heemink (1995). Data assimilation schemes for non-linear shallow water flow models. In *Proceedings of the Second International Symposium on Assimilation of Observations*, Tokyo, Japan, pp. 247–252. WMO.
- Verlaan, M. and A.W. Heemink (1997). Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrology and Hydraulics* 11, 349–368.
- Verlaan, M. and A.W. Heemink (1999). Non linearity in Data assimilation applications, a practical method for analysis. *Monthly Weather Reviews*, (submitted).

FIGURES

Figure 1: Velocity field.

Figure 2: Reference concentration field after $k = 100$ timesteps. Countours at shown with increment 1 for reference concentrations.

Figure 3: True concentration field after $k = 100$ timesteps. Contours are shown with increment 1 for reference concentrations.

Figure 4: Eigenvalues of the error covariance matrix $P(100|100)$. Computed with the RRSQRT-filter with $q = 50$.

Figure 5: Concentrations computed with the EnKf with $N = 20$ at $k = 100$. Contours at shown with increment 1 for true concentrations (-) and filter solution (...).

Figure 6: Concentrations computed with the RRSQRT filter with $N = 20$ at $k = 100$. Contours at shown with increment 1 for true concentrations (-) and filter solution (...).

Figure 7: Concentrations computed with the POEnKF with $q + N = 15 + 5$ at $k = 100$. Contours are shown with increment 1 for true concentrations (-) and filter solution (...).

Figure 8: Concentrations computed with the COFFEE KF with $q + N = 15 + 5$ at $k = 100$. Contours are shown with increment 1 for true concentrations (-) and filter solution (...).

Figure 9: RMS-errors versus computation-times for various SOS's.