ADAPTIVE FILTERING AND ORDER-REDUCTION IN THE OCEAN

P De Mey, Hoang H S, R Baraille, M Gavart
CNRS/UMR5566
Toulouse, France

Summary: Ocean data assimilation schemes must accomodate large quantities of observations, and take into account the fact that the error statistics are not well known, in particular as far as the models are concerned. An adaptive filter scheme is considered in an attempt to overcome these difficulties. A steady-state gain matrix is identified by a variational method applied recursively. The essence of the algorithm is presented from a user perspective, as well as comments on cost, stability and extension to nonlinear and nonstationary systems. The rest of the paper is concerned with order reduction in the ocean. The reduced-order adaptive filter is shown, and classic choices of order reduction based on physical considerations are presented. Isopycnal empirical orthogonal functions are shown to be a promising tool for order-reduction in the vertical in the ocean.

1. CONTEXT: DATA ASSIMILATION FOR OCEAN CIRCULATION PROBLEMS

Data assimilation is a recent subject of investigation in the ocean. It did not really catch up before the eighties, except for tidal problems (which will not be mentioned here). Two trends in the eighties are responsible for the fast development of ocean data assimilation:

- The advent of sophisticated numerical models with high resolution and the ability to simulate in a realistic manner
- New global, routine, satellite observational systems (altimetry, SST), but limited to surface observations.

Two more trends in favour of a continued development are appearing in the nineties:

- The development of commercial ocean services and operational oceanography (e.g. GOOS)
- The need for ocean monitoring and prediction to improve climate prediction.

It is clear that our basic observations will remain (for some time) surface observations, measured with a complex space-time sampling. This leads to the definition of an inverse problem, with the following objectives:

- Estimate the state of the ocean
- Estimate uncertainties on physics and improve models through match with observations
- Prediction (not the original objective).

The largest problems remaining to be solved in the field of oceanic data assimilation include:

- The propagation of surface information on the vertical

- The huge dimension of the oceanic problem. Oceanic scales are much smaller than in the atmosphere (O(100km) for the mesoscale). For instance a low-resolution 1/4° model of the Mediterranean is of dimension 10^6. Overall the problem is at least one order of magnitude larger than in the atmosphere. Therefore we need efficient methods, possibly specially designed for use in the ocean.
- Numerical models and satellite observing systems are recent. Models have not been run for a long time in assimilation mode. Therefore model errors are still not well known. Also the redundancy in the observational network not well known. Only qualitative statements are available.

At this time, data assimilation in the ocean is carried out for scientific reasons, i.e. to improve our knowledge. Therefore most assimilation methods are hindcasting methods, and use off-line "best-quality" data. In the near future we will need forecasting schemes as well. Popular assimilation methods in the ocean include:

- **Nudging** — Simple and efficient, but not optimal, no error theory, simplistic. The only method applied to global models at the moment.
- **Kalman filter** — Mostly applied in tropical areas, where linear models do a good job. Very costly, not applicable to medium- to high-resolution problems at basin- to global-scale.
- **Adjoint variational method (4DVAR)** — Works well, costly, used regionally (tropics, limited-area models, marginal seas)
- **Optimal interpolation** — Works well, same drawbacks as in meteorology (forecast errors, oversimplifications), same applications as nudging but has a simplified error theory.

However we still need a flexible, powerful assimilation method which could accomodate large quantities of data, and which would be robust enough with regard to uncertainties in the error statistics.

2. AN ADAPTIVE FILTER ALGORITHM

2.1. Adaptive filter and Kalman filter

We consider a physical system whose state $\mathbf{x}$ of dimension n evolves according to the stochastic equation:

$$\mathbf{x}_{t+1} = \Phi \mathbf{x}_t + \mathbf{w}_t \qquad (1)$$

with observations $\mathbf{z}$ of dimension p given by:

$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{v}_t \tag{2}$$

$\mathbf{w}$ and $\mathbf{v}$ are uncorrelated Gaussian white noise sequences with zero mean and covariances $\mathbf{Q}$ and $\mathbf{R}$ (most of the time $\mathbf{Q}$ and $\mathbf{R}$ are unknown). The transition matrix Φ (representing the model) and the observation operator $\mathbf{H}$ are assumed to be constant for the moment.

This estimation problem can classically be solved with a Kalman filter. However for reasons given above a Kalman filter is seldom applicable. An adaptive filter algorithm has been proposed by Hoang et al. (1994) to overcome some of the difficulties which accompany the implementation of a Kalman filter. It is summarised hereafter from a user perspective. The interested reader can find a thorough mathematical analysis and proofs in the Hoang et al. article.

We consider the following class of linear steady-state filters to produce an estimate $\hat{\mathbf{x}}$ of the state as a function of time:

$$\begin{aligned} \hat{\mathbf{x}}_{t+1} &= \Phi\hat{\mathbf{x}}_t + \mathbf{K}\zeta_{t+1} \\ \zeta_{t+1} &= \mathbf{z}_{t+1} - \mathbf{H}\Phi\hat{\mathbf{x}}_t \end{aligned} \tag{3}$$

The notations are the classic Kalman filter notations. The Kalman filter belongs to this class, but it is not the only member. In the Kalman filter solution, the gain would take the form:

$$\begin{aligned} \mathbf{K} &= \mathbf{M}\mathbf{H}^T\left(\mathbf{H}\mathbf{M}\mathbf{H}^T + \mathbf{R}\right)^{-1} \\ \mathbf{M} &= \Phi\mathbf{P}\Phi^T + \mathbf{Q} \end{aligned} \tag{4}$$

In the case of a Kalman filter, a steady-state gain is often reached asymptotically (under conditions such as "nice" properties on the unstable modes of the system, see e.g. Anderson and Moore, 1979). The "doubling algorithm" (e.g. Fukumori et al., 1993) timesteps in powers of two and allows to calculate the steady-state gain in an economical way. However, if a steady-state gain exists, even the doubling algorithm is out of reach if the system is of large dimension.

The (constant) gain matrix $\mathbf{K}$ is to be determined by minimising the cost function J, a weighted norm of the innovation vector and an implicit function of $\mathbf{K}$:

$$\begin{aligned} J &= E[\Psi_t] \\ \Psi_t &= \zeta_t^T \Sigma \zeta_t \end{aligned} \tag{5}$$

where Σ is a given weighting matrix. The minimisation problem is solved recursively over a running window of width T_0, producing a sequence of estimates of $\mathbf{K}$:

$$\hat{\mathbf{K}}_t = \hat{\mathbf{K}}_{t-1} - \gamma_t \nabla_{\hat{\mathbf{K}}} \Psi_t \tag{6}$$

Efficient choices for the γ function are discussed in Hoang et al. (1994). This function must tend to zero after some time for the algorithm to converge. This provides a saturation mechanism: after some time, the observations stop influencing the estimates of $\mathbf{K}$. Therefore, it may be necessary to restart the filter from time to time if the environment changes.

The calculation of the gradient of the cost function with respect to the gain, required in (6) to calculate the new estimate of the gain, involves the integration of the adjoint of the filtering equations (3) between the current time t and time t-T_0, forced by the model-data misfits:

$$\begin{aligned} \varphi_\tau^* &= \mathbf{L}^T \varphi_{\tau+1}^* - (\mathbf{H}\Phi)^T \Sigma \zeta_\tau \\ \mathbf{L} &= \left(\mathbf{I} - \hat{\mathbf{K}}_{t-1}\mathbf{H}\right)\Phi \end{aligned} \tag{7}$$

The gradient in (6) is a function of the adjoint variables:

$$\nabla_{\hat{\mathbf{K}}} \Psi_t = \sum_{t-T_0}^{t} \varphi_\tau^* \zeta_{\tau-1} \tag{8}$$

The algorithm is summarised in Table 1. At each restart, an initial estimate of $\mathbf{K}$ must be provided. For instance, a simplified form of the Kalman filter gain in (4) can be used for initialisation:

$$\hat{\mathbf{K}}_0 = \mathbf{H}^T \left(\mathbf{H}\mathbf{H}^T + \mathbf{I}\right)^{-1} \tag{9}$$

although a smoother form with extrapolation structure in $\mathbf{M}$ would probably achieve a better job. When restarting the filter, the previous estimate of $\mathbf{K}$ can be used if it is not too far from the new regime.

Table 1: Adaptive filter algorithm
Recurrent operations

Step	Operations at time t	Equations involved
1	Timestep filter from (t-1) to (t), calculate new state	$\hat{\mathbf{x}}_t = \Phi \hat{\mathbf{x}}_{t-1} + \hat{\mathbf{K}}_{t-1} \zeta_t$
2	Timestep adjoint equations from (t) to (t-T_0)	$\varphi^*_\tau = \mathbf{L}^T \varphi^*_{\tau+1} - (\mathbf{H}\Phi)^T \Sigma \zeta_\tau$
3	Calculate gradient of cost function	$\nabla_{\hat{\mathbf{K}}} \Psi_t = \sum_{t-T_0}^{t} \varphi^*_\tau \zeta_{\tau-1}$
4	Calculate new gain matrix to be used from (t) to (t+1)	$\hat{\mathbf{K}}_t = \hat{\mathbf{K}}_{t-1} - \gamma_t \nabla_{\hat{\mathbf{K}}} \Psi_t$
5	Return to step 1	

2.2. Discussion

Physical constraints — It may be useful in some instances to add weak physical constraints to the problem, because some of the solutions of the model and of its adjoint may not be acceptable. For instance in the ocean one might want to constrain the solution to leave the water masses undisturbed along isopycnal surfaces, to be close to isobath flow, or to a predetermined solution at depth where few observations are available. These constraints can (at least theoretically) be easily added to the problem: provided that they have "nice" mathematical properties, their derivatives can be incorporated as additional source terms in the adjoint equation (7). In practice, these ideas remain to be tested.

Model errors — This algorithm does not require the specification of the error covariance matrices **Q** and **R**, to the contrary of the Kalman filter. Therefore, to the contrary of the Kalman filter, it is not sensitive to the frequent and gross errors in those quantities, especially in the model error covariances (the observational error covariances are usually better known). Examples of damages caused by erroneous statistics are ubiquitous in the meteorological literature. In the ocean, the damage could be even worse, because models have not been compared to independent data over sufficiently long times to have a good idea of **Q**.

2.3. Cost

Tables 2 and 3 show that for a steady-state filter the adaptive filter algorithm is much cheaper than the Kalman filter. The number of additions and multiplications grows in n^2, not in n^3, because there are no matrix-matrix products to the contrary of the Kalman filter. The difference increases as the size of the state increases.

Table 2: Number of multiplications and additions for the Kalman Filter and Adaptive Filter

	Kalman Filter	Adaptive Filter
MULT	$n^3 + 2(1+p)n^2 + p^2n + p^3$	$(1+T_0)n^2 + (T_0+3)pn + p^2$
ADD	$2n^3 + 2pn^2 + (p^2-p-1)n + p^2$	$(T_0+1)n^2 + (T_0+2)(p-1)n + p^2 - p$

Table 3: Examples

n	p	T_0	MULT RATIO	ADD RATIO
1000	100	100	10	19
10^5	10^3	100	999	1980

The choice of T_0, the number of timesteps over which the adjoint equations are integrated backwards, is another cost- issue. Since the misfits only force the adjoint system at τ=t, contributions to the sum (8) start being negligible beyond some limit for a stable filter. For instance an exit criterion could be:

$$\left\|\left(\mathbf{I}-\hat{\mathbf{K}}\mathbf{H}\right)\Phi\right\|^{T_0} \leq \varepsilon \tag{10}$$

In practice, 3-5 timesteps are sufficient in most tests carried out in Toulouse.

2.4. Reduced-order adaptive filter

In some cases, it may still be necessary to further reduce the number of unknowns because the system dimension is still too large. For instance, this can occur when the tangent linear system and the adjoint of the model are not available, and when it is not possible to calculate and store the adjoint transition matrix in (7). Alternatively, one may want to enforce physical constraints by projecting the stochastic forcing **w** in (1) onto some subspace or attractor.

In practice, it is possible to parametrise the gain matrix **K** in the form $\mathbf{PrK}_e$, where **Pr** is a given projection operator, and solve an identification problem for a smaller set of control variables in $\mathbf{K}_e$. For instance, if the influence of observations can be considered to be limited in space because of the underlying physics or because of coasts, one may use a block or band approximation for **K**. Another choice is to use the optimal structure of the gain matrix in the Kalman filter form (4) and identify elements of the error covariance matrices. However these choices have consequences on the stability of the filter.

A more detailed discussion on the stability and physically-relevant choices for **Pr** will be presented later.

In a follow-up paper, Hoang et al. (1995) study a reduced-order adaptive filter algorithm, and compare it with a reduced-order Kalman filter. The reduced-order adaptive filter algorithm is very similar to the full-order filter presented above. In the new algorithm, the filtering equation (solved in full space) becomes:

$$\begin{aligned} \hat{\mathbf{x}}_{t+1} &= \Phi\hat{\mathbf{x}}_t + \mathbf{PrK}_e\zeta_{t+1} \\ \zeta_{t+1} &= \mathbf{z}_{t+1} - \mathbf{H}\Phi\hat{\mathbf{x}}_t \end{aligned} \qquad (11)$$

and the adjoint equations (solved in reduced space) become:

$$\begin{aligned} \varphi^*_\tau &= \mathbf{L}_e^T\varphi^*_{\tau+1} - (\mathbf{H}_e\Phi_e)^T\Sigma\zeta_\tau \\ \mathbf{L}_e &= \left(\mathbf{I} - \hat{\mathbf{K}}_{e,t-1}\mathbf{H}\right)\Phi_e \\ \Phi_e &= \mathbf{Pr}^+\Phi\mathbf{Pr} \end{aligned} \qquad (12)$$

The calculation of the gradient of the cost function is still calculated as in (8), and the reduced gain is updated as the full gain was in (6). The new algorithm still minimises the cost function in (5), a measure of the model-data distance, but this time we look for solutions with structure given by the projection operator **Pr**. It can be noted that there is no such optimality criterion for the equivalent reduced-order Kalman filter.

Of course, the cost and performance of the reduced-order filter, as well as the physical sense of the results, will critically depend on the choice of the projection. One could envisage to adaptively identify **Pr** for a given model and class of processes, but this is still to be done.

An example of use of reduced-order adaptive filtering in a quasi-geostrophic model of an unstable oceanic jet is presented in Hoang et al. (1995). Similar work with the Miami isopycnal model (MICOM) is being carried out in Toulouse and Brest with very promising results.

2.5. Stability

This is a "research-in-progress" item.

The identifiability of the Kalman gain by the adaptive filter is not guaranteed for assimilation problems arising in meteorology and oceanography. Therefore the filter may be ill-behaved. It is important to satisfy stability criteria for the filter, i.e. to make sure that all eigenvalues of the filter transition matrix **L** in equation (7) are less than 1:

$$\left\| \left(\mathbf{I} - \hat{\mathbf{K}}\mathbf{H} \right) \Phi \right\| \leq 1 \tag{13}$$

For the steady-state gain to exist, the unstable modes must be observable and controllable (detectability and stabilisability). However the filter (in particular in reduced space) can still be unstable. If in addition the model is stable (in Lyapunov sense), then the adaptive filter is stable. However ocean models usually have neutral and unstable modes, and not all of them are well known!

It is possible to build stable filters even if there are unstable modes. We need special structures for the gain matrix to be identified. For instance, the following structure is being explored in Toulouse:

$$\hat{\mathbf{K}} = \Theta \mathbf{W} \mathbf{H}^T \mathbf{V} \tag{14}$$

where the matrices $\mathbf{V}$ and $\mathbf{W}$ are symmetric positive definite. This structure resembles the form of the Kalman gain, with an additional coefficient matrix Θ in front. It can be proven (Hoang, pers. comm.) that the following choices lead to a stable filter:

$$\begin{aligned} \Theta &= \mathbf{I} \\ \mathbf{V} &= \left(\mathbf{HWH}^T + \mathbf{R}\right)^{-1} \end{aligned} \tag{15}$$

where $\mathbf{R}$ is a symmetric positive definite matrix.

In the case of a reduced-order adaptive filter, the stability criterion must first be satisfied in reduced space. However this does not guarantee the stability in full space, unless all unstable modes are in the reduced space. It is simple good sense to try to maximise the observability in the reduced space (hence observing most of the unstable modes), and to somehow constrain the remaining unstable modes (bogus observations, relaxation to climatology, etc.).

2.6. Other issues: nonlinearity, nonstationarity

Nonlinearity — The adaptive filter algorithm can be extended to nonlinear cases (e.g. Hoang et al., 1995). The linearised equations are only needed to build the adjoint model. In the Kalman filter equivalent, the tangent linear equations are needed to propagate the analysis errors, and optimality is not guaranteed.

Nonstationarity — Most physical systems are nonstationary. For instance the regime (and hence the gain) exhibit seasonal and interannual variations. Also, in the ocean, there is no permanent observational network yet. A solution is the Kalman filter, which does not require stationarity (except in the steady-state version). However there are several possible solutions with the adaptive filter. For instance, if the observational network is time-dependent, one can adaptively identify $\mathbf{M}$ and $\mathbf{R}$ (or the diagonal elements thereof) in:

$$\mathbf{K}(t) = \mathbf{MH}(t)^T\left[\mathbf{H}(t)\mathbf{MH}(t)^T + \mathbf{R}\right]^{-1} \tag{16}$$

If in addition the background regime changes with time, one can identify R, P, and Q (or the diagonal elements thereof) in (16) above and:

$$\mathbf{M}(t) = \Phi(t)\mathbf{P}\Phi(t)^T + \mathbf{Q} \tag{17}$$

Note: most of this still have to be tested in realistic cases, but it gives guidelines for doing so. Another simple possibility in the case of a nonstationary system is to restart the adaptive filter from time to time, i.e. restart the γ function in (6) and reinitialise the gain. However the interdependence between γ, T_0, the flow's advective scales, and the scales over which the regime changes are not yet fully understood.

3. CANDIDATES FOR ORDER REDUCTION IN THE OCEAN

3.1. Classic choices in the ocean

Generally speaking, the first thing we should think of to reduce the order of the problem is a priori physical knowledge. For instance, in a primitive-equation model, the error in the state equation is negligible, and the background error covariances are generally chosen as geostrophic (which means that the ageostrophic part is only corrected in the advective phase following assimilation): the result is that we can remove the density and the velocities from the assimilation problem, and hence from the state. More generally, as suggested by Fukumori (1993), a general class of possibilities is offered by the choice of a transformation, or projection operator, which will yield a reduced-order state of much smaller size. In the case of the projection operator **Pr** appearing in (11,12) above, the reduced state $\mathbf{x}_e$ is such that

$$\mathbf{x} = \mathbf{Pr}\,\mathbf{x}_e + \mathbf{N} \tag{20}$$

where **N** spans the null space of **Pr**. **N** is the part of the state which is not described by the reduced form. One must be very cautious with the unstable modes appearing in the null space of **Pr**.

When exploring the potential of a projection operator for simplifying an assimilation problem, various difficulties can arise. For instance, the choice of **Pr** may make physical sense while being difficult to put in mathematical form. Also, if the null space of **Pr** contains unstable modes, the resulting system will not be stable even if the reduced-order filter is. Another potential problem is that the errors in the null space, even when bounded, can contaminate the reduced state through nonlinear processes or climatic drift. Therefore we look for projections whose null space is not dynamically coupled to the reduced space, i.e. such that null space vector will in the null space after integration by the model. This will almost never happen exactly, but we can choose the transformation in order to minimise that coupling. We also want to maximise observability of the system in reduced space.

A few examples of relevant transformations for oceanic problems follow:

I. **The fit of feature models.** This can work well in specific cases, when the initialization fields can be described with a few parameters, for instance the Gulf Stream and its rings (Robinson et al., 1988). However the null space can be very large in most cases where typical feature models do not exist.

II. **3-D to 2-D,** also called **vertical extrapolation.** Clearly, the vertical dimension is a special dimension in the ocean. Also, the only routine observations are collected at the surface. We can distinguish two classes here:

A. The dynamical methods, such as the potential vorticity method (Haines, 1991), the water property conservation method (Cooper and Haines, 1996), and the neutral-surface method (Bindoff and Mc Dougall, 1994). These methods, which actually are very close to each other, are justified mostly on the large scales because they give preference to the conservation of water properties over the advective effects. One negative side is the fact that the null space of those transformations contain synoptic-scale dynamics such as vortex stretching, which participates in the mesoscale potential vorticity balance and is coupled to the reduced space for eddy-resolving models.

B. The statistical methods, such as the vertical empirical orthogonal function (VEOF) method (De Mey and Robinson, 1987; Dombrowsky and De Mey, 1992; Fischer and Latif, 1994; Rienecker and Adamec, 1995; Benkiran and De Mey, 1996) and its isopycnal variant (Gavart and De Mey, 1996), the regression method (Mellor and Ezer, 1991), and the water-mass method (largely statistical - Oschlies, 1994). These are flexible methods which are well suited to quasi-homogeneous media such as synoptic-scale geophysical fluids, but the isopycnal EOF method works on the large scale too. The VEOF method explicitly minimises the norm of the null-space vectors, and the null space is statistically uncoupled to the stochastic forcing space. In addition, the isopycnal EOF method minimises the observability in null space. For the mesoscale, one or two dominant modes yield a robust representation of the three-dimensional synoptic fields. The representativity of EOFs at larger scales and in multivariate cases has also been explored (e.g. Fukumori and Wunsch, 1991).

III. The local inversion (Dombrowsky and De Mey, 1992; Rienecker and Adamec, 1995; Benkiran and De Mey, 1996) consists in selecting misfits only within a preselected space/time horizon around the interpolation point. This is close to the "data volume method" put forward by Lorenc (1981) at the European Centre for Medium Range Weather Forecasting (ECMWF) and used by Cummings (1995) in an ocean model. It is an order-reduction method because it leads to a simpler structure of the gain matrix. This approach is particularly suited to mesoscale problems, and enhances the portability to parallel machines. However, the null space information, defined as the remote data forcing, may be advected inside the horizon by strong currents for instance, which would in effect couple both subspaces.

IV. The use of a coarse grid for the filter (Fukumori et al., 1993). This method is not very useful at the synoptic scale, because we want the full resolution, but at the larger scales it is very well conditioned. It is being used in Toulouse to assimilate simulated observations in MICOM with the adaptive filter method.

3.2. Isopycnal EOFs: a tool for order-reduction in the ocean

The isopycnal EOF method (Gavart and De Mey, 1996) aims at significantly reducing the dimension of the assimilation problem while maximising the observability of the reduced system (in particular at depth). We also expect the contribution of mean thermohaline gradients to the statistics to be much smaller when calculating EOFs in isopycnal coordinates than in depth coordinates. The isopycnal representation of the ocean is not only fundamental for large scale dynamical studies, but also because of the growing interest in isopycnal circulation models such as MICOM implied in data assimilation studies. The so-called dynamical method of altimetric assimilation proposed by Cooper and Haines (1996) is also based on isopycnal considerations with a local vertical adjustment of the water column.

An EOF (Empirical Orthogonal Function) analysis of the vertical space-time variability of the thermohaline structure of the ocean has been performed by Gavart and De Mey in summer and fall of 1993 in the Azores Current region. The analysis used data from CTD casts of the SEMAPHORE intensive experiment, referred to a seasonally-varying climatology. The relative merits of depth-coordinate and isopycnal-coordinate representations have been examined in regard to the quality of identification of physical processes, and to the effectiveness of the extrapolation from Sea-Level Anomaly (measured by satellite altimeters), in an attempt to reduce the number of degrees of freedom in oceanic assimilation problems.

The isopycnal EOFs consistently proved more efficient in capturing the vertical structure of both dominant processes in the area of investigation: the variability linked to the Mediterranean Water, and the coherent physical system made up of the the Azores Front and Azores Current. The isopycnal analysis also proved more robust when meddies (Mediterranean Water lenses) were included in the analysis. In addition, isopycnal EOFs appeared to be more observable from altimetry: residual sea-level anomaly variance after projection on the dominant mode (representing the Azores Front/Azores Current) was 8.4 cm^2 in depth-coordinate representation and 3.6 cm^2 in isopycnal-coordinate representation.

In addition, as an attempt to estimate the vertical structure of errors needed by assimilation schemes, differential isopycnal EOFs were calculated from pairs of casts close in space and time. Despite the fact that only 15 pairs were available for the chosen radii, the dominant process was a quasi-homogeneous vertical displacement of isopycnals with quasi-conservation of water masses and potential vorticity on the isopycnal grid.

REFERENCES

Anderson, B.D.O., and J.B. Moore, "Optimal Filtering", Prentice-Hall, Inc., Englewood Cliffs, 1979.

Benkiran, M., and P. De Mey, J. Phys. Oceanogr. , submitted, 1996.

Bindoff, N.L., and T.J., J. Phys. Oceanogr., 24, 1137-1152, 1994.

Cooper, A., and K. Haines, J. Geophys. Res., 101, 1059-1077, 1996.

Cummings, J.A., Proc. 2nd Int. Symposium on assimilation of observations in meteorology and oceanography, WMO/TD No. 651, Vol. 2, 509-512, 1995.

De Mey, P., and A. Robinson, J. Phys. Oceanogr., 17, 2280-2293, 1987.

Dombrowsky, E., and P. De Mey, J. Geophys. Res., 97, 9719-9731, 1992.

Fischer, M., and M. Latif, J. Marine Systems, 6, 31-46, 1995.

Fukumori, I., and C., Prog. Oceanogr., 27, 111-195, 1991.

Fukumori, I., J. Benveniste, C. Wunsch, and D. B. Haidvogel, J. Phys. Oceanogr., 23, 1831-1855, 1993.

Gavart, M., and P. De Mey, J. Phys.Oceanogr., submitted, 1996.

Haines, K., J. Phys. Oceanogr., 21, 843-868, 1991.

Hoang, H.S., P. De Mey, and O. Talagrand, Proc. 33rd IEEE Conf. on Decision and Control, 1994.

Hoang, H.S., O. Talagrand, N.T. Loan, P. De Mey, and R. Baraille, to appear in Kybernetika, 1995.

Hoang, H.S., P. De Mey, O. Talagrand, and R. Baraille, Proc. IFAC Int. Conf. on Adaptive Systems in Control and Signal Proc., 153-158, 1995. To appear in Automatica, 1996.

Lorenc, A.C., Mon. Wea. Rev., 109, 701-721, 1981.

Mellor, G. L., and T. Ezer, J. Geophys. Res., 96(C5), 8779-8795, 1991.

Oschlies, A., Ph.D. Thesis, IfM, D-24105 Kiel, Germany, 130pp, 1994.

Rienecker, M.M., and D. Adamec, J. Marine Systems, 6, 125-143, 1995.

Robinson, A.R., M.A. Spall, and N. Pinardi, J. Phys. Oceanogr., 18, 1811-1853, 1988.