Original article

A semi-evolutive filter with partially local correction basis for data assimilation in oceanography

Un filtre semi-évolutif avec une base de correction partiellement locale pour l'assimilation de données en océanographie

Ibrahim Hoteit [a,*], Dinh-Tuan Pham [a], Jacques Blum [b]

[a] *Laboratoire de Modélisation et Calcul, Tour IRMA BP 53, 38041 Grenoble cedex 9, France*
[b] *J.A. Dieudonné Laboratory, 06108 Nice, France*

Abstract

A new data assimilation scheme derived from the singular evolutive extended Kalman (Seek) filter is introduced. The novel feature of the new filter is its correction basis which is partially local in the sense that it consists of "global" and "local" vectors, the later obtained from a local empirical orthogonal functions (Eof) analysis. Such an analysis was introduced in order to better represent the local variability of the ocean. This not only significantly reduces the implementation cost but may also improve the representativeness of the correction basis of the filter. The performance of this scheme is evaluated through twin experiments conducted in a realistic setting of the OPA model over the tropical Pacific zone. The results are compared against those of the Seek filter. The new filter is shown to perform better while it is up to six times faster. Adaptive tuning of the forgetting factor was also used, which enhances performance and improves the stability of the filter during model unstable periods.

© 2003 Éditions scientifiques et médicales Elsevier SAS and Ifremer/CNRS/IRD. All rights reserved.

Résumé

Un nouveau schéma d'assimilation de données dérivé du filtre de Kalman singulier étendu et évolutif (Seek) est présenté. La nouveauté de ce filtre est sa base de correction partiellement localisée car elle consiste en une partie "globale" et une partie "locale". Ce dernier provient d'une analyse locale empirique de fonctions orthogonales. Une telle analyse permet une meilleure représentation de la variabilité locale de l'océan. Elle réduit non seulement le coût d'implémentation mais elle accentue également la représentativité de la base de correction du filtre. Sa performance est évaluée par deux expériences jumelles conduites avec une configuration réaliste du modèle OPA sur le Pacifique tropical. Les résultats sont comparés à ceux qui sont issus du filtre Seek. Le nouveau filtre est plus performant tout en étant six fois plus rapide. Un réglage adapté du facteur oublié est également utilisé, ce qui rehausse la performance et accentue la stabilité du filtre durant les périodes d'instabilité du modèle.

© 2003 Éditions scientifiques et médicales Elsevier SAS and Ifremer/CNRS/IRD. All rights reserved.

1. Introduction

One of the major challenges in the coming years in oceanography is the design of an operational data assimilation system. There is increasing interest in these systems for the purposes of improving short-term and mid-term weather prediction, developing climate predictions or for the specific objectives of the navies in various countries.

Among the various methodological approaches (see for example Ghil and Malanotte-Rizzoli, 1991, for a review), we

---

* Corresponding author.
E-mail address: ihoteit@fjord.ucsd.edu (I. Hoteit).

have chosen the statistical sequential technique based on the Kalman filter. For a linear system, this filter provides the best estimate, in the least squares sense, of the state vector at each observation time using all observations up to this time (Kalman, 1960). For non-linear models, a so-called extended Kalman (EK) filter has been derived (Jazwinski, 1970). However, brute force implementation of the EK filter in realistic ocean models is not possible because of the huge dimension of the system ($\sim 10^7$). Several approaches, which essentially consist of projecting the system state onto a low-dimensional sub-space, have been proposed to overcome this problem (Cane et al., 1996; Evensen, 1994; Fukumori and Malanotte-Rizzoli, 1995; Hoang et al., 1995; Verlaan and Heemink, 1995).

With the same aim in view, Pham et al. (1997) recently introduced the singular evolutive extended Kalman filter (Seek). This filter essentially approximates the error covariance matrix by a singular low rank matrix, which allows for ignoring corrections in the directions where the errors are naturally strongly attenuated by the system. The novel feature of this filter is that its "directions of correction", usually initialized through an empirical orthogonal function (Eof) analysis, evolve in time to follow the model dynamics. The Seek filter has been applied in different realistic ocean frameworks, yielding quite satisfactorily results (Pham et al., 1997; Verron et al., 1998; Ballabrera-Poy et al., 2001; Hoteit et al., 2002).

Nevertheless, the Seek filter remains an expensive choice for operational oceanographic applications since a model integration is needed for the evolution of each correction direction. To reduce the cost of this filter, Hoteit et al. (2002) have proposed different methods to simplify the evolution of these directions. But this does not address a built-in weakness of the Seek filter (in fact, it may even exacerbate it) in that the correction basis may not contain enough vectors to capture the short-range variability of the model. Increasing the number of basis vectors would of course increase the computational burden and, as shown in our experiments, only marginally increases the representativeness of the basis (Cane et al., 1996; Verron et al., 1998). Furthermore, in real applications, the model can be subject to errors and hence an Eof analysis based on a run of the model would be subjected to errors too. We believe that such errors mostly affect the last Eofs because they are the most numerically unstable.

For the above reasons, we shall introduce a variety of "local Eof analysis" which would provide better representativeness of the ocean state. The "local Eofs" will support a small region of the ocean and vanish elsewhere. Therefore, the computational cost is reduced allowing the use of more basis vectors for a given cost. This local analysis limits the spatial correlation length of the ocean variables, which is consistent with the idea that such correlation generally vanishes for large distances. Further, it allows for adapting the number of Eofs according to the variability of each sub-domain.

Improving the estimation of small-scale correlations with the Kalman filters has received more attention recently. Houtekamer and Mitchell (1997) simply used a cutoff radius to exclude the effects of observations far away from the analysis location. However, the introduction of an "artificial" radius may be inconsistent with the correlations presented in the estimation error covariance matrix, which would lead to numerical noise. Fukumori (2002) and Testut (2000) partitioned the model domain into several overlapping sub-domains and then applied the correction of the Kalman filter independent of each sub-domain. Additional state reductions were also applied to separate sub-domains using coarser grid (Fukumori, 2002) and Eof analysis (Testut, 2000). Although the idea proposed in this paper is similar to the one used by Testut (2000), our approach is different since the correction of the Kalman filter remains unchanged in our filter and is always applied globally (only the correction basis contains local functions). The use of a "mixed" correction basis of global and local functions is also introduced in this paper to improve the representativeness of the ocean long-range variability, which can be degraded, as shown in our experiments, if only local functions are used. Moreover, the global functions of such a basis can be let to evolve as in the Seek filter, since the local Eofs cannot be let to evolve without destroying their locality. The resulting semi-evolutive partially local Kalman (Seplek) filter is much less costly than the Seek filter and yet can yield better results. Our work was presented at the The Fourth International Marine Environmental Modeling Seminar 2000, for which some results appeared in the conference proceedings (Hoteit et al., 2001). However, this paper provides a more complete description and full results of the work.

The paper is organized as follows. Section 2 introduces different approaches to improve the representativeness of the Eof analysis. The Seplek filter is described in Section 3. Section 4 describes the numerical results of the assimilation experiments. Finally, concluding remarks are given in Section 5.

2. Global–local Eof analysis

The goal of the "classical" (or global) Eof analysis is to construct a representation as accurately as possible of a set of state vectors X_1, K, X_N in $\mathcal{R}^p$ in a low-dimension (denoted by r) linear sub-space. Let $\{\phi_k\}_{k=1, K, r}$ be a M-orthogonal basis of such a space (here M is some metric introduced to account for the different nature of the state variables as they represent different physical variables such as velocity, salinity, etc.), the best representation relative to this space of a vector X is given by

$$\tilde{X} = \bar{X} + \sum_{k=1}^{r} c_k \phi_k = \bar{X} + LL^T M(X - \bar{X}) \qquad (1)$$

where $\bar{X}$ is the average of X_1, K, X_N, $c_k = <X - \bar{X}, \phi_k>_M = \phi_k^T M(X - \bar{X})$ and L is the matrix having $\{\phi_k\}_{k=1, K, r}$ as columns. It has been shown that the M-norm of the represen-

tation error is smallest when $\{\phi_k\}_{k=1, \text{K}, r}$ are the first eigenvectors of the sample covariance matrix of X_1, K, X_N, the eigenvectors being ranked in decreasing order of their eigenvalues λ_1, K, λ_N. These eigenvectors are called Eofs. Furthermore, the fraction of variance (or inertia) explained by the first r Eofs is known to be equal to $\sum_{k=1}^{r} \lambda_k / \sum_{k=1}^{p} \lambda_k$ and thus can be used as a guide for choosing r (this fraction should be close to 1). The reader is referred to Preisendorfer (1988) for more details.

2.1. Local Eof analysis

The classical Eof analysis provides an approximate representation of the correlations between all ocean variables. However, the resulting correlations often have long range if only a few Eofs are retained, while one can expect ocean variables to have a limited spatial correlation length (two variables evaluated at two distant ocean locations tend to be uncorrelated, or weakly correlated). Further, as the Eof analysis is based on extracting the global variability of the ocean, a local turbulent zone might not have enough energy to appear in the analysis while it may play an important role in the ocean dynamics.

To enhance the effectiveness of the Eof analysis, we propose to force the Eofs to be local. Our basic idea consists of constructing a set of Eofs, each supporting a small region of the ocean. This can be achieved by splitting the model domain into sub-domains and then applying separate Eof analyses on each of these sub-domains. To retain correlations between spatial points in spatially nearby sub-domains, we use overlapping sub-domains. Our approach is very flexible as we have at our disposal a wide range of options to choose from: (i) the number of sub-domains, (ii) the extent of their overlaps and possibly the size and shape of these sub-domains as well, and (iii) the number of local basis elements in each sub-domain. Such flexibility would enable us to construct a basis best adapted to the problem at hand. Another point is that calculations, which mostly involve scalar products, can be less costly with the local functions since these functions have local supports. Therefore, one can afford more (local) basis elements without increasing (with a possibility of even decreasing) the computation cost.

2.1.1. Construction of the local Eofs

The idea is to define a set of ocean sub-domains, on each of which a separate Eof analysis will be applied. As said before, these sub-domains should overlap. To do this properly, we consider a partition of unity of the ocean domain, i.e. a set of non-negative functions $\{\pi^{(j)}, j = 1, \text{K}, J\}$ defined on the ocean domain, whose sum equals 1 identically. Therefore, any state vector X can be written as:

$$X(x, y, z) = \sum_{j=1}^{J} X(x, y, z)\, \pi^{(j)}(x, y, z) = \sum_{j=1}^{J} X^{(j)}(x, y, z) \quad (2)$$

where x, y and z denote the spatial coordinates and $X^{(j)} = X\pi^{(j)}$.

A separate Eof analysis can be then carried out on each local field $X^{(j)}$, $1 \leq j \leq J$, to compute a basis for each ocean sub-domain. The representation formula for each $X^{(j)}$ is

$$X^{(j)}(x, y, z) = \bar{X}^{(j)}(x, y, z) + \sum_{l=1}^{r^{(j)}} c_l^{(j)}\, \phi_l^{(j)}(x, y, z) \\ + e^{(j)}(x, y, z) \quad (3)$$

where $\phi_l^{(j)}$ denote the Eofs resulting from the jth local Eof analysis, $\bar{X}^{(j)}$ is the mean vector, $r^{(j)}$ is the number of retained Eofs and $c_l^{(j)}$ are constant coefficients. By summing the last equation over the index j, one obtains a representation formula for the full state vector

$$X(x, y, z) = \bar{X}(x, y, z) + \sum_{j=1}^{J} \sum_{l=1}^{r1(j)} c_l^{(j)}\, \phi_l^{(j)}(x, y, z) \\ + e(x, y, z) \quad (4)$$

In the above representation, the error e is equal to the sum of all local representation errors $e^{(j)}$. However, this representation is not optimal as one can (and will) readjust the coefficients $c_l^{(j)}$ in Eq. (4) in order to obtain the smallest mean squared errors. This results in the representation formula (in matrix form)

$$X = \bar{X} + BC + e, \quad (5)$$

where B is the matrix whose $r = \sum_{j=1}^{J} r^{(j)}$ columns are the local Eofs $\phi_l^{(j)}$ and C is a column vector such that BC represents the M-orthogonal projection of $X - \bar{X}$ onto the sub-space spanned by the columns of B. These columns thus constitute a representation basis, called the local Eof basis. The number $r^{(j)}$ of Eofs in each sub-domain can be chosen as in the classical Eof analysis. Thus, the value of $r^{(j)}$ varies according to the variability of each sub-domain. While this number may serve as a first guess, one can still readjust it, for example by increasing it in the sub-domains with strong variability.

2.1.2. Choice of the ocean sub-domains

The size of the sub-domains is obviously characterized by the support of the partition of unity functions $\pi^{(j)}$. In practice such functions should support a small region of the ocean and vanish elsewhere. For example, in the case of a rectangular oceanic domain, one can take these functions in tensorial form as

$$\pi^{(j)}(x, y, z) = \pi_X^{(j_1)}(x)\, \pi_Y^{(j_2)}(y), \quad (6)$$

with j_1 varying from 1 to J_1, j_2 varying from 1 to J_2, $j = j_1 + J_1(j_2 - 1)$ varying from 1 to $J = J_1 J_2$ and the functions $\pi_X^{(j_1)}$ and $\pi_Y^{(j_2)}$ having local support and sum 1. Moreover, since the ocean vector field is manifestly continuous, it is a good idea to take these functions to be continuous, which would require them to have overlapping support (as $\pi^{(j)}$ must decrease continuously to 0 as one approaches the boundary of its support, it must be strictly less than 1 near this boundary, hence, since $\sum_{k=1}^{J} \pi^{(j)} = 1$, there must be some other

function $\pi^{(k)}$ which is non-zero there, meaning that the functions $\pi^{(j)}$ have overlapping support). In any case, one should not limit the correlation length in the vertical direction to let the surface information propagate to the ocean bottom (because altimetry data only provide information about the ocean surface).

The main difficulty with the local Eof basis is that it cannot be allowed to evolve with the model (see Section 3) without losing its locality. Although we can decide to use such a basis to initialize the Seek filter only, and abandon its locality afterward by allowing it to evolve with the model dynamics, this would be very costly because of the large dimension of the basis. We have also noticed in our experiments that the long-range variability was not well represented by the local Eofs. For these reasons, we propose to augment the local basis by some global elements, which results in a so-called mixed global–local basis.

2.2. Mixed Eof analysis

Local ocean variability is in general not well represented by the first few global Eofs. The residuals of the state vectors in the sub-space generated by the global Eof analysis thus mostly contain information about the short-range motions. One can therefore apply a local Eof analysis to extract the local variability from these vectors. By combining the local and the global bases one obtains a representation basis capable of representing the long range as well as the short-range oceanic phenomena. Moreover, as previously stated, the global part of this basis can be made to evolve as in the Seek filter to follow the model dynamics.

To construct such a basis, one first computes a global Eof basis L from a sample of ocean states $X_1, ..., X_N$. The representation of a state vector X by a vector in the sub-space generated by L is then given by the formula

$$X = \bar{X} + LL^{\mathrm{T}} \mathrm{M}(X - \bar{X}) + e, \qquad (7)$$

where e is the representation error (or residual). Next, by applying a local Eof analysis (not centered) on the above representation residuals $e_1, ..., e_N$ of the state vectors $X_1, ..., X_N$, one obtains a reconstruction formula for the residuals according to

$$e = BC + e', \qquad (8)$$

where the matrix B contains the retained local Eofs, C contains the coefficients of the reconstruction and e' is the representation error. Note that we do not center the vector e, as it represents a residual and thus should already be centered. By combining Eqs. (8) and (9), one gets

$$X = \bar{X} + LL^{\mathrm{T}} \mathrm{M}(X - \bar{X}) + BC + e', \qquad (9)$$

showing that $X - \bar{X}$ can be represented, with an error e', in terms of mixed Eof basis, defined by the matrix

$$L^B = [L : B]. \qquad (10)$$

But this representation is not optimal as one still can (as in Section 2.1) adjust the vector C^B so that:

$$X = \bar{X} + L^B C^B + e'' \qquad (11)$$

with e'' having the smallest M-norm.

3. The semi-evolutive partially local filter

The sequential data assimilation scheme consists of the estimation of the state of the system at each observation time, using only observations up to this time. In the linear case, this problem has been entirely solved by the well known Kalman filter. In the non-linear case, one often linearize the model around the current estimated state vector, which yields the extended Kalman (EK) filter (see for example Ghil and Malanotte-Rizzoli, 1991, for details). However, direct application of the EK filter to data assimilation in meteorology and oceanography is not possible due to the huge dimension ($n \approx 10^7$) of the system. Pham et al. (1997) proposed a sub-optimal EK filter very close to the original EK, called Seek filter, in which the error covariance matrix was assumed to be singular with a low rank $r << n$. This leads to a filter in which the error correction is made only along certain directions parallel to a linear sub-space of dimension r. They are the directions for which errors are not sufficiently attenuated by the system dynamics.

Although the cost reduction of the Seek filter is very important, this filter remains expensive for operational oceanographic applications since the evolution of its correction basis requires $r + 1$ times the model run. To further reduce the cost, one needs to simplify the evolution of its correction basis. Motivated by the fact that most of the estimation error in the numerical experiments of Pham et al. (1997) was reduced immediately after the first correction, while the evolution of the Eof basis L_0 was not yet effective, Brasseur et al. (1999) proposed to keep the initial correction basis of the Seek filter fixed in time. This approach leads to a very low-cost filter, called singular fixed extended Kalman (Sfek) filter, which operates in two stages exactly as for the Seek but without the evolution equation of the correction basis (Eq. (21)). It is thus almost $r + 1$ times faster than the Seek filter. It also provides a cheap way to test the representativeness of the Eof basis (Brasseur et al., 1999; Hoteit et al., 2002).

With the same goal, Hoteit et al. (2002) proposed different degraded forms of the Seek filter, which are less costly and yet perform reasonably well. Among them, the singular semi-evolutive extended Kalman (Sseek) filter can be most easily adapted to the present (global–local) approach. This filter lets to evolve a few correction directions only, while keeping the others fixed. But since the basis vectors which do not evolve are those which contribute the least to the error representation, it cannot be used directly in the present context because this will destroy the separation of the mixed Eof basis into global and local parts. Here, we simply propose to let the global part of the mixed Eof basis to evolve as in the Seek

filter, while keeping the local part fixed in time. This results in a new filter called semi-evolutive partially local Kalman (Seplek) filter. Of course, small-scale features are expected to evolve much faster than basin-scale features. In fact, the former could change so fast that the Seek filter would be unable to follow, as the design of this filter is aimed toward following slow changes only (Pham et al., 1997). This can result in filter instability as the correction basis vectors keep changing continually but are still unable to capture sufficiently the dynamic of the ocean. In this situation, it may be preferable to keep part of the correction basis (which corresponds to small scale) fixed but by using a much larger number of basis vectors. We are in effect not aiming to correct all those small-scale variation errors, but only a large number of them on average.

3.1. The semi-evolutive partially local EK filter

The Seplek filter proceeds in two stages after an initialization stage just as in the Seek filter. The notation proposed by Ide et al. (1995) will be adopted for the description of the filter algorithm. Let the physical system be described by

$$X^t(t_k) = M(t_k, t_{k-1}) X^t(t_{k-1}) + \eta_k, \qquad (12)$$

where $X^t(t)$ denotes the vector representing the true state at time t, $M(t,s)$ is an operator describing the system transition from time s to time t and η_k is the system noise vector. At each time t_k, one observes

$$Y_k^o = H_k X^t(t_k) + \varepsilon_k, \qquad (13)$$

where H_k is the observational operator and ϵ_k is the observational noise. The noises η_k and ϵ_k are assumed to be independent random vectors with mean zero and covariance matrices Q_k and R_k, respectively.

For the initialization, a long sequence of state vectors is first generated using the numerical model. A mixed B global–local Eof analysis is then performed on these vectors, which yields a mixed Eof basis L_0^B partitioned as

$$L_0^B = [L \,\vdots\, B] \qquad (14)$$

with $r_g + r_1$ columns, r_g and r_1 denoting the dimensions of the global and the local basis, respectively. However, such an analysis does not readily provide a rank $r_g + r_1$ error covariance matrix for the state vector. Therefore, based on the first observation Y_0^o, an objective analysis is used to make the first correction according to

$$X^a(t_0) = \bar{X} + L_0^B [(L_0^B)^{\mathrm{T}} H_0^{\mathrm{T}} R_0^{-1} H_0 L_0^B]^{-1}$$
$$(L_0^B)^{\mathrm{T}} H_0^{\mathrm{T}} R_0^{-1} [Y_0^o - H_0 \bar{X}] \quad , (15)$$

where $\bar{X}$ is the average of the state vectors (from the historical run) and H_0 is the gradient of H_0 evaluated at $\bar{X}$. The initial analysis error covariance matrix is then given by

$$P^a(t_0) = L_0^B U_0 (L_0^B)^{\mathrm{T}}$$
$$\text{where}$$
$$U_0 = [(L_0^B)^{\mathrm{T}} H_0^{\mathrm{T}} R_0^{-1} H_0 L_0^B]^{-1}. \qquad (16)$$

3.1.1. Forecast stage

At time t_{k-1}, $X^a(t_{k-1})$ and $P^a(t_{k-1})$ in the factorized form

$$P^a(t_{k-1}) = L_{k-1}^B U_{k-1} (L_{k-1}^B)^{\mathrm{T}} \qquad (17)$$

are available, where the mixed correction basis $L_{k-1}^B = [L_{k-1} \,\vdots\, B]$ and the matrix U_{k-1} are of dimension $n \times (r_g + r_1)$ and $(r_g + r_1) \times (r_g + r_1)$, respectively. The Eq. (13) is used to compute the forecast. The corresponding forecast error covariance matrix can then be approximated by:

$$P^f(t_k) = L_k^B U_{k-1} (L_k^B)^{\mathrm{T}} + Q_k \pi, \qquad (18)$$

where the new mixed correction basis is taken as

$$L_k^B = [L_k \,\vdots\, B]. \qquad (19)$$

The global part of the correction basis L_k evolves as in the Seek filter according to

$$L_k = M(t_k, t_{k-1}) L_{k-1}, \qquad (20)$$

where $M(t_k, t_{k-1})$ is the gradient of $M(t_k, t_{k-1})$ evaluated at $X^a(t_{k-1})$.

It may be, however, desirable for numerical stability reasons, to normalize the evolutive basis vectors at each filtering step so that their norm is still equal to a constant c. This amounts to replacing the jth column L_k^j of L_k by the vector $(c/\|L_k^j\|_M) L_k^j$ and then dividing the jth column and the jth row of U_{k-1} by the constant $c/\|L_k^j\|_M$. For the normalizing constant c, we choose the average of the squared norms of the Eofs (which are equal to the corresponding eigenvalues).

3.1.2. Correction stage

The new observation Y_k^o at time t_k is used to correct the forecast according to

$$X^a(t_k) = X^f(t_k) + G_k [Y_k^o - H_k X^f(t_k)], \qquad (21)$$

where G_k is the gain matrix given by

$$G_k = L_k^B U_k (L_k^B)^T H_k^T R_k^{-1} \qquad (22)$$

with H_k the gradient of H_k evaluated at $X^f(t_k)$ and U_k computed from

$$U_k^{-1} = [U_{k-1} + [(L_k^B)^{\mathrm{T}} L_k^B]^{-1} (L_k^B)^{\mathrm{T}} Q_k L_k^B$$
$$[(L_k^B)^{\mathrm{T}} L_k^B]^{-1}]^{-1} + (L_k^B)^{\mathrm{T}} H_k^{\mathrm{T}} R_k^{-1} H_k L_k^B \quad .(23)$$

The corresponding filter error covariance matrix is then equal to

$$P^a(t_k) = L_k^B U_k (L_k^B). \qquad (24)$$

The structure of the forecast and analysis error covariance matrices of this filter is, therefore, very similar to that of the Seek filter. Basically, these matrices are taken to be singular of low rank (or at least approximated to be so). Their image space is determined by the "correction space" spanned by the correction basis, partly evolving in time with the model dynamics. The magnitude of the error covariance matrices is controlled by the matrix U_k, which is updated in the correction stage.

Concerning the cost of the Seplek filter, it is almost entirely due to the evolution of the global part of its correction basis. More precisely, it is slightly $r_g + 1$ times larger than the cost of the model. The Seplek filter can, therefore, be much less costly than the Seek filter since r_g can be chosen to be much smaller than the dimension of the (global) correction basis of the Seek filter.

3.2. Algorithm with adaptive forgetting factor

The main reason behind the use of the forgetting factor in the Seek filter is to limit the effective filter memory length by discarding old data. This enables the filter to follow system changes and also to attenuate the propagation of the errors in time. By errors, we specifically mean the error included by linearization and the part of the dynamical and filter errors which lies outside the correction space and is ignored by the filter (see Eq. (24)), the normal error being correctly handled by the filter. The gain matrix is thus modified to avoid the "blow up" phenomena due to the above errors (Astrom and Wittenmark, 1989). Furthermore, the use of a forgetting factor does not require any extra cost for its implementation: the filter equations remain unchanged except for the emergence of the forgetting factor ρ in the time propagation error covariance equation (Pham et al., 1997). Specifically, the updating Eq. (24) for U_k now changes to:

$$U_k^{-1} = \left[\frac{1}{\rho} U_{k-1} + [(L_k^B)^T L_k^B]^{-1} (L_k^B)^T Q_k L_k^B \right.$$
$$\left. [(L_k^B)^T \quad L_k^B]^{-1} \right]^{-1} + (L_k^B)^T \quad H_k^T R_k^{-1} H_k L_k^B \quad .(25)$$

With $\rho = 1$, all data have the same weight, but with $\rho < 1$, recent data have exponentially larger weights than old data.

However, the use of too small a forgetting factor when the system evolution is stable would degrade the filter performance, especially when there is little information in the measurements. To maximize the benefit of the forgetting factor, Hoteit et al. (2002) used an adaptive one: such factor is set close to 1 when the model is stable and much less than 1 when the model is unstable. Indeed, old data should have less impact in the latter case, in order to adapt the filter to the new model's mode. Model unstable periods are detected by tracking the filter's state by computing short-term and long-term averages of the prediction error variance. The reader is referred to Hoteit et al. (2002) for more details.

A different approach proposed by Dee and Da Silva (1999) is to treat the forgetting factor as an unknown parameter in the forecast error covariance matrix, then to estimate it adaptively (on-line) by the maximum likelihood method. Mitchell and Houtekamer (2000) follow a similar approach but parameterize directly the model error covariance matrix (which is more natural) in the context of the ensemble Kalman filter. However, these methods would be much more costly than our simple adjustment of the forgetting factor.

4. Application to altimetric data assimilation in the tropical Pacific

The performances and the capabilities of the Seplek filter are assessed using a twin experiment approach with a realistic setting of the OPA model in the tropical Pacific Ocean under a "perfect model" assumption ($Q_k = 0$), whereby the same model is used to generate a true solution, and to generate forecasts. The assimilation is based on pseudo-observations extracted from the twin experiments.

4.1. OPA model in the tropical Pacific

The OPA model (OPA for océan parallélisé) is a primitive equation ocean general circulation model, which was developed at Lodyc (Laboratoire d'Océanographie Dynamique et de Climatologie) to study large-scale ocean circulation. It solves the Navier–Stokes equations which express the momentum balance, the heat and salt balances and a non-linear equation of state under the assumptions of hydrostatic equilibrium and incompressibility, plus the rigid lid assumption and some hypotheses made from scale considerations. The system equations are written in curvilinear z–coordinates and discretized using the centered second-order finite difference approximation on a three-dimensional generalized C–Arakawa grid (Arakawa, 1972). Time stepping is achieved by two time differencing schemes: a basic leap-frog scheme associated to an Asselin filter for the non-diffusive processes and a forward scheme for diffusive terms. The sub-grid scale physics are parameterized by tracer diffusive operators of second-order on the vertical, the eddy coefficients being computed from a turbulent closure model (Blanke and Delecluse, 1993). The lateral diffusive and viscous operators can either be of second or of fourth order. The reader is referred to the OPA reference manual (Madec et al., 1997) for more details.

The model domain covers the entire tropical Pacific basin extending from 120° E to 70° W and from 33° S to 33° N, and the level depth varies from 0 at the sea surface to 4000 m. Two buffer zones are included between 20° and 33° in the north and south of the domain, for the connection with the sub-tropical gyres. The number of horizontal grid points is 171 × 59 on 25 vertical levels. The model equations are solved on a horizontal grid with a 1° zonal resolution and a maximum meridional resolution of 0.5° at the equator and decreasing to 2° at the northern and southern boundaries. The vertical resolution is approximately 10 m between the sea surface and 120 m depth, and then decreases to 1000 m near the floor. The time step is 1 h.

The bathymetry is relatively coarse. It was obtained from The Levitus data mask (Levitus, 1982). The forcing fields are interpolated every month from the ECMWF data. They are composed of find stress, heat and fresh water fluxes. Zero fluxes of heat and salt and non-slip conditions are applied at solid boundaries. A second-order horizontal friction and diffusion scheme for momentum and tracers is chosen with a coefficient of 2000 m²/s in the strip 10° N–10° S, increasing

up to 10,000 m²/s at the northern and southern boundaries. Static instabilities are resolved in a turbulent closure scheme. The model starts from rest (i.e. with zero velocity fields). The salinity and the temperature stem from seasonal Levitus (1982) climatology.

4.2. Experiments design

4.2.1. The state vector

The state vector X^T is the set of prognostic variables, which, in the OPA model, consists of the zonal and meridional velocity U and V, the salinity S and temperature T, thus

$$X^T = (U, V, S, T)^T. \tag{26}$$

However, the observation, which is the sea surface height (SSH), is directly related to the surface pressure P_S but the latter is only indirectly related to the state variables through a set of partial differential equations. More precisely, P_S is a diagnostic variable which can be computed through the system of equations (Pinardi et al., 1995)

$$\nabla(h\nabla P_S) = -\nabla(fh\mathbf{k} \wedge \mathbf{u}) - \nabla \int_{-h}^{o} (z+h) \nabla\Omega \, dz + \delta\nabla\mathrm{B}, \tag{27}$$

where ∇ denotes the horizontal gradient, h is the ocean depth, f is the Coriolis factor, $\mathbf{k}$ is the vertical unit vector, $\mathbf{u}$ is the horizontal vector velocity field, Ω is the density, δ is the Rossby number and $\mathbf{B}$ is a term describing the non-linear advection and dissipation effects. The term $\mathbf{B}$ is rather complex but is also needed for solving the system equations concerning X^t. In fact, the numerical code for integrating the OPA model also computes, as a by-product, the diagnostic variable P_S. Therefore, for purely technical reasons, it is advantageous to augment the state vector by including the variable P_S, that is we now take

$$X^T = (U, V, T, P_S)^T. \tag{28}$$

Of course, the model equations must now include the extra equation (28) and the derivation of our filter algorithm must be based on this extended model, not on the original model (the Eof analyses are also carried out on the augmented state vectors). The overhead is insignificant in any case since P_S is only defined on the surface of the ocean. The number of state variables increases from $4 \times 171 \times 59 \times 25 = 1,008,900$ to $4 \times 171 \times 59 \times 25 + 171 \times 59 = 1,018,989$. The numerical cost of the algorithm would increase by the same proportion since this cost is roughly proportional to the dimension of the state vector. This increase would be offset by the fact that the calculation of the observation operator is now straightforward. If we work with the original state vector, we would have to pull out the portions of the OPA code for computing directly P_S from (U, V, S, T), with some unavoidable redundancy with the integration of the model equations. But the main point is that this would considerably increase the complexity of the programming work without incurring real differences both in terms of cost and in terms of methodology.

4.2.2. Eof analysis and initial correction basis

As explained in Section 3, the choice of the initial state estimator and the corresponding correction basis is made through a simulation of the model itself. Thus, in the first experiment, the model was spun up for 7 years from 1980 to 1986 with the aim of reaching a statistically steady state of mesoscale turbulence. Then, another integration of 4 years was carried out from 1987 to 1990 to generate a historical sequence H_S of model realization. A sequence of 480 state vectors was retained by storing one state vector every 3 d to reduce the calculation costs since successive states are quite similar. Since the state variables in Eq. (29) are not of the same nature, we shall apply a multivariate Eof analysis (local and global). The matrix M (see Section 2) is chosen to be diagonal with diagonal elements being the inverse of the average variances of the state variables at the grid points.

4.2.2.1. Local Eof analysis. We performed two subdivisions of the tropical Pacific domain: (i) in three subdomains in the zonal direction according to the natural subdivision of the tropical Pacific into a western warm pool, the central pool and the cold eastern pool and (ii) in nine subdomains in order to further limit the spatial correlation length in the meridional direction. The choice of the partition of unity functions is shown in Fig. 1. After applying separate
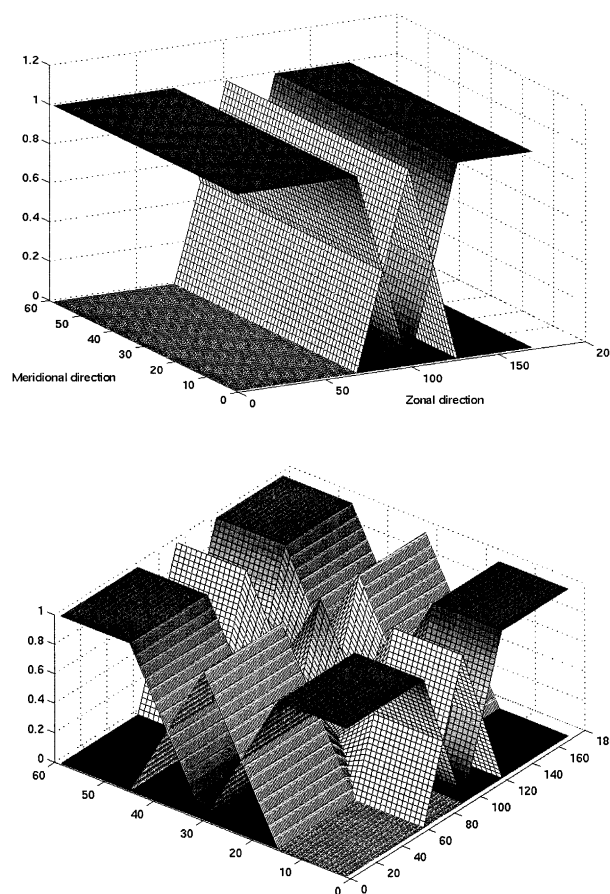

Fig. 1. Two-dimensional partition of unity functions in three and nine sub-domains.

Eof analysis on the state vectors of the sub-domains, we have retained for the sub-division (i) 18, 30 and 23 Eofs to attain 85% of the inertia in the first, second and third sub-domains, respectively. Thus, the dimension of the local Eof basis is equal to $18 + 30 + 23 = 71$. For the sub-division (ii) we have retained 7, 10, 7, 10, 20, 13, 5, 17, 12 Eofs to attain 85% of the inertia in the nine sub-domains, respectively. The dimension of this local Eof basis is then equal to 101.

4.2.2.2. Mixed Eof analysis. Since better results were obtained with the sub-division (i) (see Section 4), we have applied this sub-division only on the residuals of the ocean states in the sub-space spanned by the first five global Eofs which almost explain 50% of the total global inertia. Results of the corresponding local analysis show that 19, 19 and 20 Eofs have to be retained in order to explain 65% of the inertia of the residuals in the first, second and the third sub-domains, respectively. The dimension of the mixed Eof basis is, therefore, equal to $5 + 19 + 19 + 20 = 63$.

4.2.3. Data description and filters validation

The data for the assimilation experiments were again simulated but in a way unrelated to the above run. A reference experiment was performed and the reference state X^t retained to be compared later with the fields produced during the assimilation experiments. More precisely, a sequence of 250 state vectors was retained every day during the period from 1 March 1991 to 10 November 1991. The assimilation experiments were performed using the pseudo-measurements which were extracted from the reference experiment. The SSH was assumed to be observed at every grid point of the model surface with a nominal accuracy of 3 cm. The observation error was simulated by adding randomly generated Gaussian noise to the synthetic observations of SSH. Note that in the assimilation interval, a period of strong model instability occurs between July and September (Hoteit et al., 2002).

The performance of all the filters is evaluated by comparing the relative root mean square (RRMS) error for each state
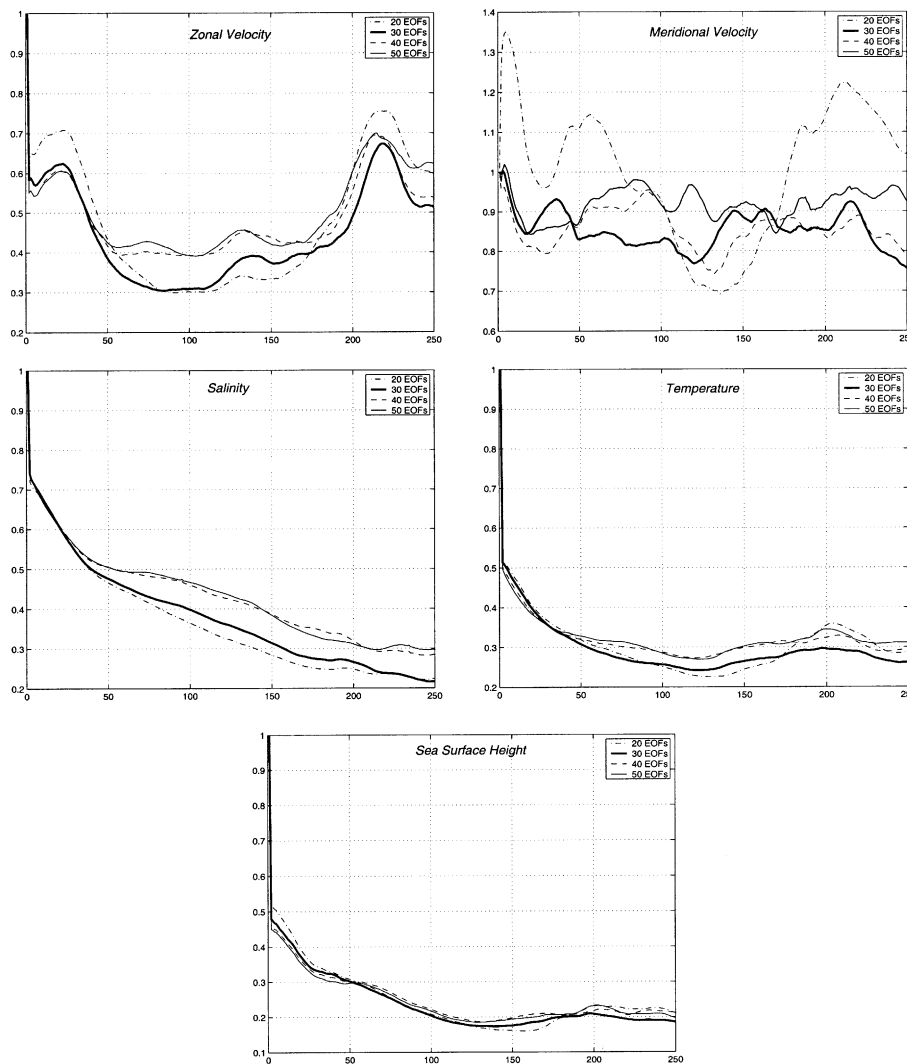

Fig. 2. Evolution in time of the RRMS on the whole model domain for the Sfek filter with 20, 30, 40 and 50 Eofs. The forgetting factor is set to 0.8.

variable, in each layer or in the whole domain of the ocean model. The RRMS is defined as

$$\text{RRMS} = \frac{\| X^t(t_k) - X^a(t_k) \|}{\| X^t(t_k) - \bar{X} \|}, \qquad (29)$$

where $\bar{X}$ is the mean state of the sample H_S and $\|\cdot\|$ denotes the Euclidean norm. Thus the error is relative to the "free-run" error, since the denominator represents the error when there is no observation and the analysis vector is simply the mean state vector.

4.3. Results of assimilation experiments

Several experiments were carried out to study the sensitivity of the filters to the different Eofs basis and to demonstrate the relevance of the Seplek filter. The results of these experiments are presented hereafter.

4.3.1. Sensitivity of the Sfek filter to the number of global Eofs

Fig. 2 plots the assimilation results of the Sfek filter with four different numbers of retained (global) Eofs: 20, 30, 40 and 50, which explain 79%, 85%, 88% and 90% of the inertia of the sample H_S, respectively. One can see that the best results were obtained using only 30 Eofs. This suggests that adding more Eofs in the correction basis does not guarantee better performance of the filter and over a certain point it can even make it worse. Such an observation was also made by Cane et al. (1996) and Verron et al. (1998). We believe that the first Eofs are good for representing long-range ocean phenomena but not suitable for short-range fluctuations. Other Eofs corresponding to low eigenvalues are often unreliable; they are quite sensible to statistical errors, even computational round-off errors and of course model error. One has also to keep in mind that Eofs are optimal (at best) in a time-mean sense only.
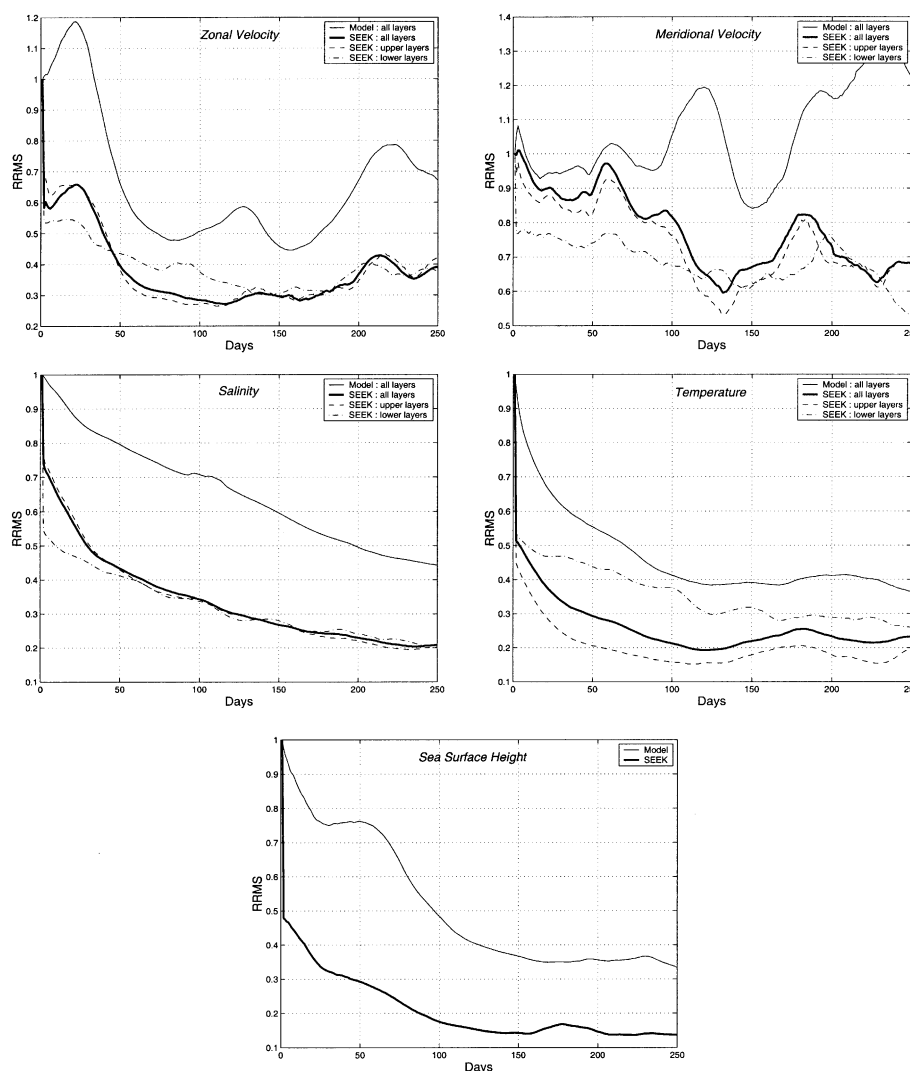

Fig. 3. Evolution in time of the RRMS for the model free run (without assimilation) on the whole model domain and for the Seek filter on the whole model domain, on the (mean of the 5) upper and the (mean of the 5) lower layers.

Following the results of this experiment, the number of retained Eofs for the global basis was set to 30 in all the above experiments.

4.3.2. The Seek filter

Fig. 3 plots the RRMS of the Seek filter with a fixed forgetting factor set to 0.8. As the model is driven by an external forcing, it is of interest to see if the use of correct forcing alone suffices to produce adequate results. We have, therefore, compared the RRMS of the Seek filter to that of the free model run (without assimilation and starting from the initial state of the filter). It can be seen that the Seek filter highly improves the model behavior. The performance of the Seek is also very satisfactorily both in the upper layers and in the lower layers. Although it appears to degrade somewhat in the presence of instabilities, it still behaves fairly well during this period. One may think that the meridional velocity is not sufficiently well-estimated because the assimilation error is only reduced by less than a half. But it is worthwhile to point

out that, since the velocity field of the tropical Pacific Ocean is primarily zonal, the meridional velocity fields are generally, and especially the referenced field in our experiment on 1 March 1991, well-approximated by the average of the meridional velocity. Since this average serves as our initial analysis, the initial error is already low and, therefore, it would be hard to reduce it much further.

The assimilation results of the Seek filter have been presented in both the upper and lower layers for completeness. But we noticed that the difference between the RRMS of the other filters and that of the Seek filter computed in all the layers are quite similar to that computed on each individual layer. Therefore, in the following, we will only present results in all domains, to save space.

4.3.3. Sensitivity of the Sfek filter to the choice of the local domains

Fig. 4 plots the evolution of the filtering error of the Sfek filter with the local Eofs computed from a domain sub-
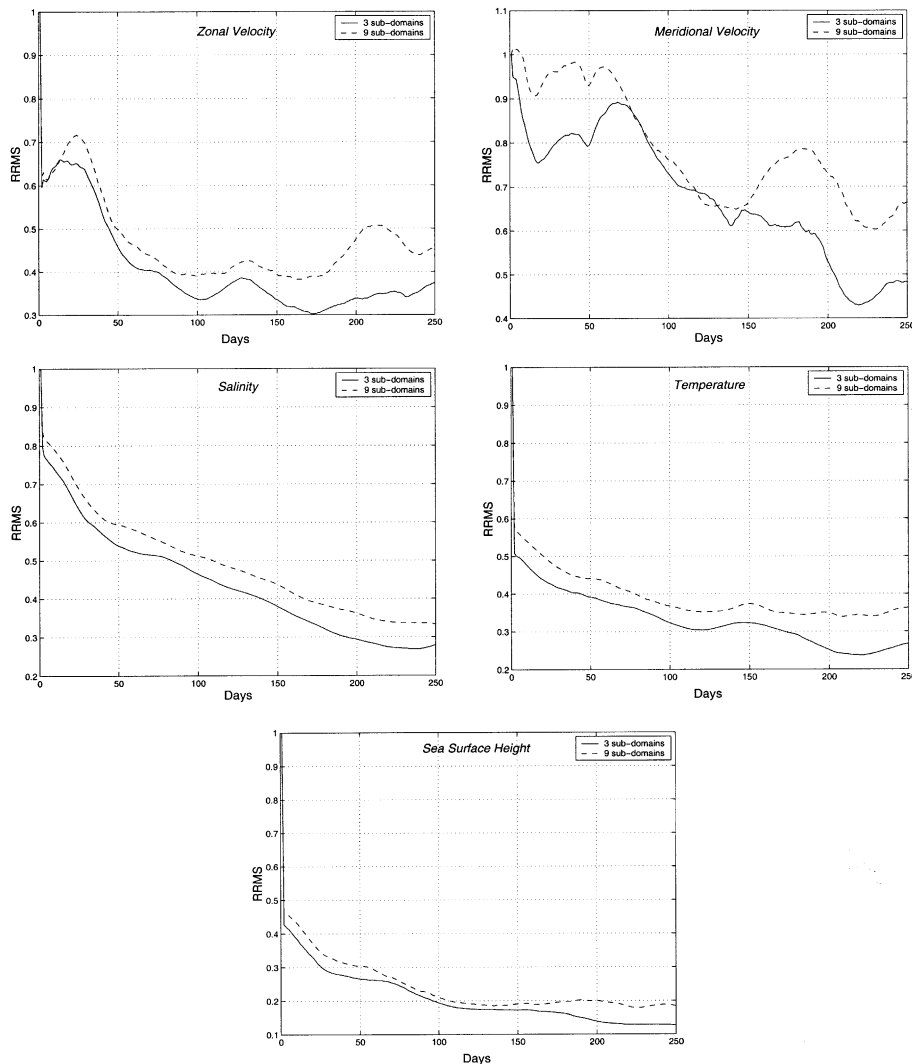

Fig. 4. Evolution in time of the RRMS for the Sfek with the local Eof basis obtained from a sub-division of the domain in three and nine sub-domains.

division into three and nine sub-domains. It can be clearly seen that better results are obtained from the "natural" sub-division into three zonal zones. This suggests that there is no need to limit the spatial correlation length of the tropical Pacific variables in the meridional direction, since the length of the tropical Pacific domain is much larger than its width. Furthermore, one should discount the north and south buffer zones which do not participate actively in the model dynamics. Another reason for this result may also be related to the divergence of the horizontal velocity on the vertical which should be equal to zero in the OPA model. Indeed, the local Eofs, however, do not fulfill this condition on the border of the sub-domains. The analysis state $X^a(t_k)$ has been, therefore, corrected at every filtering step (before using the model to forecast the state) by projecting it onto a zero-divergence linear sub-space. This entails a loss of precision on the estimation of $X^a(t_k)$ which is likely to increase with the number of sub-domains.

4.3.4. Representativeness of the different Eof bases

A simple and low-cost way to test the relevance of the different Eof bases is to use them as a fixed correction basis for the Sfek filter and then to examine the performance of the filter. Therefore, three experiments were conducted using the Sfek filter with the different correction bases obtained from the global, local and the mixed Eof analyses (with three sub-domains). The forgetting factor was chosen, fixed to be equal to 0.8. Results plotted in Fig. 5 show that the local Eof basis is much more representative of the variables *U*, *V* and SSH, than the global Eof basis. However, the assimilation results with the variables *S* and *T*, which may be thought of as being controlled by phenomena of long-range variability, are not as good as those obtained from the global basis. Using the mixed Eof basis, the Sfek filter performs very well with the variables *U*, *V* and SSH, much like the local Eof basis but considerably improving the assimilation of *S* and *T* with respect to that obtained from the latter basis. This suggests
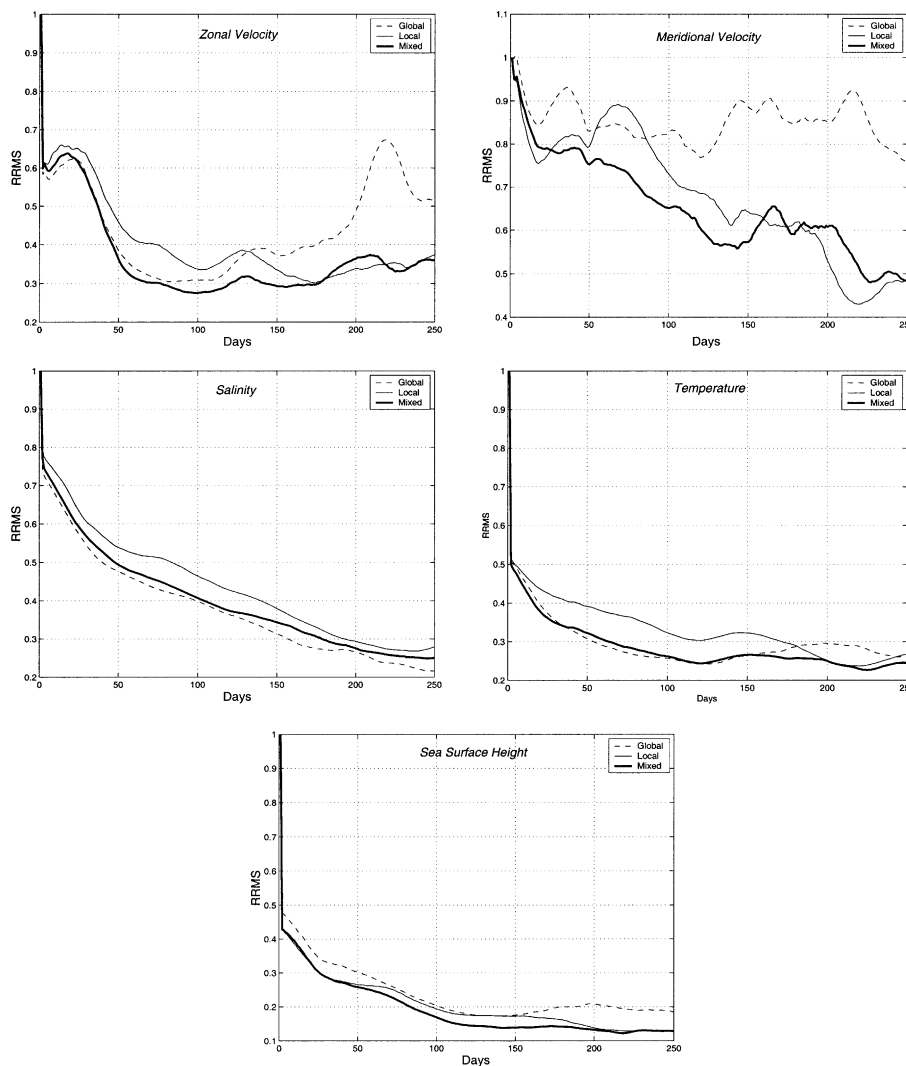

Fig. 5. Evolution in time of the RRMS for the Sfek with the global, local and the mixed Eof basis.

that the mixed Eof analysis is able to represent both the short and the long-range ocean variability.

4.3.5. The Seplek filter

The Seplek filter was implemented with a fixed forgetting factor, $\rho = 0.8$. The global part of the mixed basis was made to evolve as in the Seek filter while the local part was kept fixed. Since the dimensions of the (global) Eof basis and the global part of the mixed basis were taken equal to 30 and 5, respectively, the Sfek the Seplek were almost 30 and 6 times faster, respectively, than the Seek filter. It can be seen from Fig. 6 that the Seplek performs very well. Its assimilation results of the velocity components U and V and the SSH are shown to be much better than those obtained with the Seek filter during the unstable period. As far as the salinity S and the temperature T are concerned, the Seplek filter performs almost as well as the Seek filter. One can also notice the positive impact

of the evolution of the global basis part of the mixed correction basis by comparing the results of the Seplek filter with those obtained with the Sfek filter when the mixed Eof basis is used as a fixed correction basis.

4.3.6. Adaptive forgetting factor with the Sfek and Seplek filters

The use of an adaptive forgetting factor was shown to enhance the performance of the Seek filter and its variants (Hoteit et al., 2002). Here, we have used an adaptive forgetting factor with the Sfek (with the mixed Eof basis) and Seplek filters. Such forgetting factor takes one of the two values 1 or 0.8 according to the relative magnitudes of the short-term and long-term prediction errors. Fig. 7 compares the RRMS error for these experiments with those of the Seplek filter with a fixed forgetting factor equal to 0.8. These results show the benefit of the adaptive tuning scheme of the
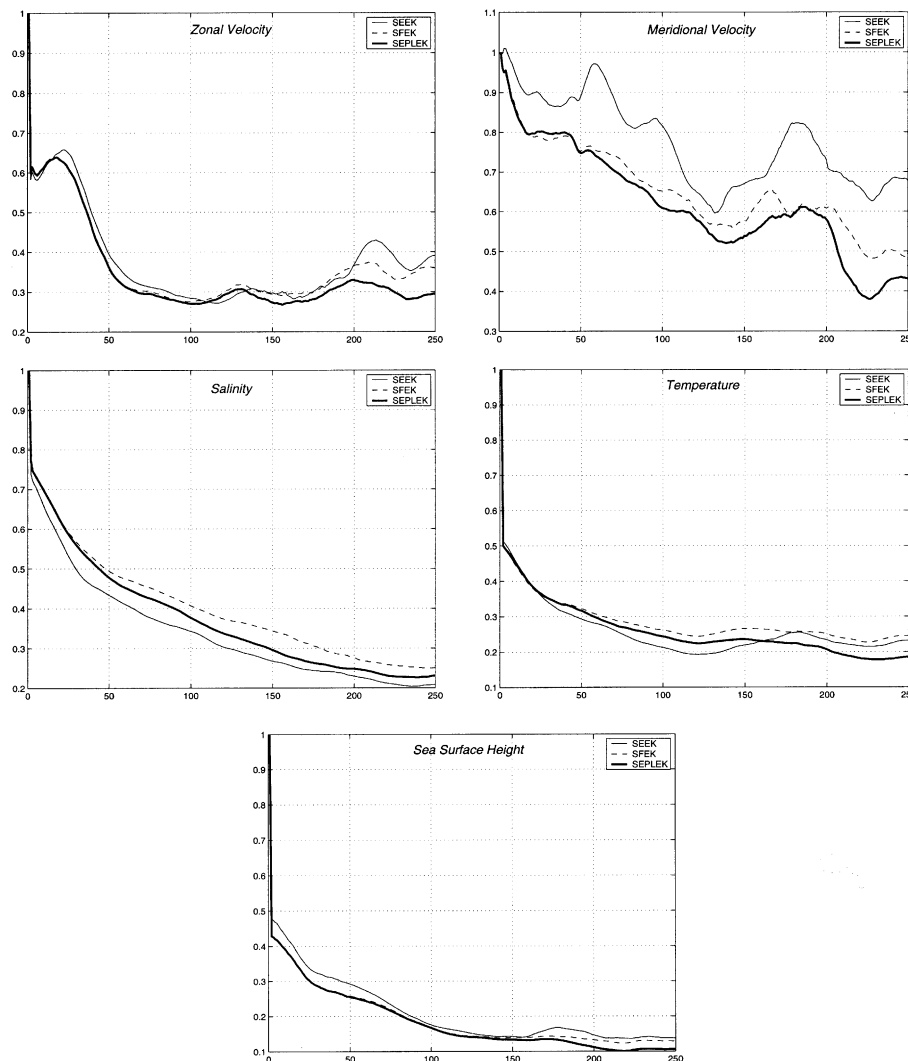

Fig. 6. Evolution in time concerning the RRMS for the Sfek (with the mixed Eof basis), Seplek and Seek filters with a forgetting factor 0.8.
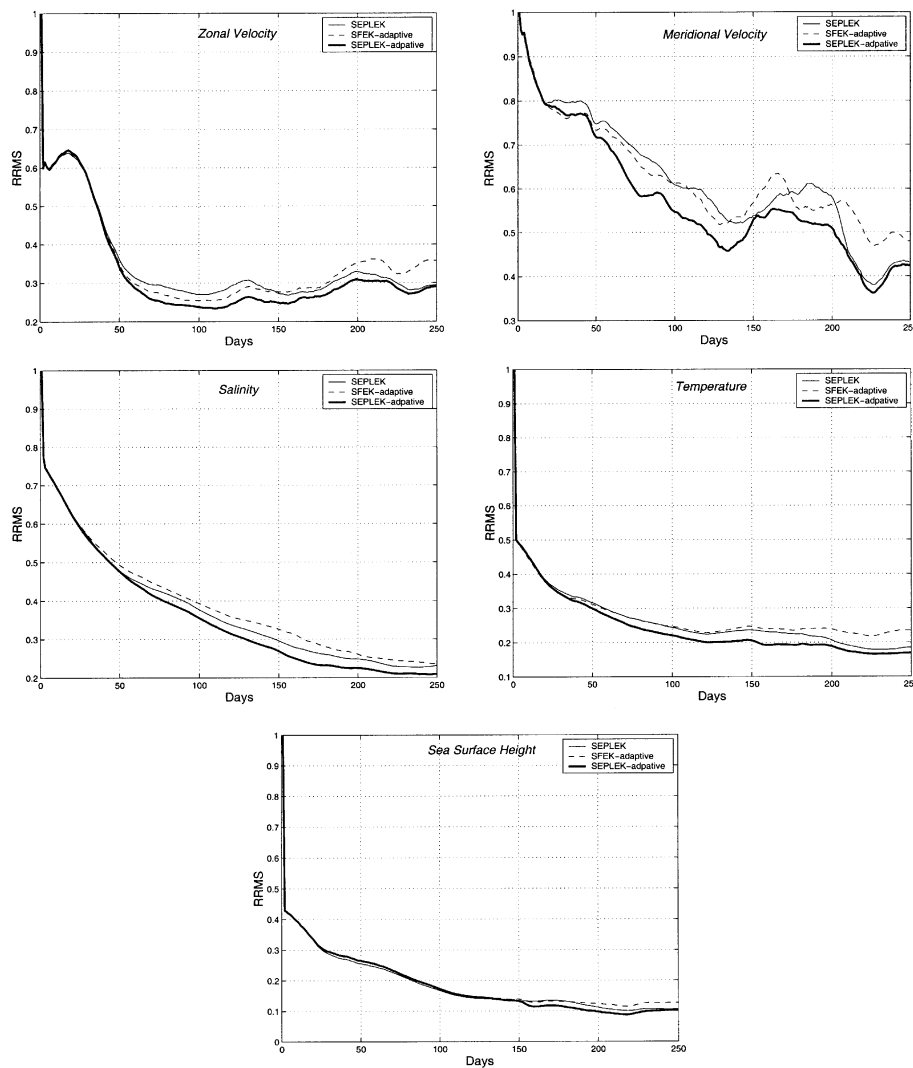
Fig. 7. Evolution in time of the RRMS for the Sfek and Seplek filters with a variable forgetting factor and for the Seplek filter with a constant forgetting factor.

forgetting factor and also its ability to significantly enhance the stability of the filters during model unstable periods.

5. Conclusions

A new data assimilation scheme derived from the Seek filter was introduced, namely, the semi-evolutive partially local extended Kalman (Seplek) filter. The motivation for developing such a filter was to reduce the cost of the Seek filter and possibly to obtain a better representativeness for its correction basis. Our approach was to construct a mixed Eof basis composed of global and local Eofs and to let only the global part evolve with model dynamics. To compute the local basis part, a "local" Eof analysis was applied on the residuals of the states in the linear sub-space generated by the global part. In such a local analysis, the Eofs support a small region of the ocean domain and vanish elsewhere. Such a procedure permits, on the one hand, to limit the spatial correlation length of the ocean variables if needed, and on the

other, to obtain a basis better adapted to the variability of each sub-domain according to its variability. The performance of the Seplek filter was assessed with a realistic setting of the OPA model in the tropical Pacific, and its results compared to those of the Seek filter as well as a fixed basis Seek filter. Based on these experiments, our main conclusions are as follows.

1. The local Eofs represent the variability of the velocity and the sea surface height much better than the classical Eofs but not that of the salinity and the temperature which seem to be essentially of a global scale. However, finding a way to let the local basis evolve with the model dynamics remains an open problem.

2. The mixed Eof basis is shown to perform very well even if we keep it fixed. It enhances the filter performance for the salinity and the temperature, with respect to the local Eof basis. This suggests that this basis is able to represent the local and the global ocean variability well. When its global part is made to evolve as in the Seek filter, it provides a dynamically evolutive filter which

performs even better than the Seek filter, while requiring much lower cost.

3. By tracking the forecast error, one can obtain information about the filter state and then adapt the filter parameters to the present situation. In particular, the adaptive tuning of the forgetting factor considerably enhances the performance of the Seek filter and its variants.

In twin experiments, the Seplek filter was found to be very effective in assimilating surface-only pseudo-altimeter data. Further work will consider more realistic situations, such as the addition of the model error or the use of more realistic observations (according to satellite tracks and real satellite data). However, these preliminary twin experiments were necessary steps before starting realistic applications. They provided us with encouraging results with regard to that goal. We are currently working to implement this filter in the OPA model of the North Atlantic Ocean.

References

Arakawa, A., 1972. Design of the Ucla general circulation model. Numerical integration of weather and climate, Rep. 7. Dept. of Meteorology, University of California 116 p.

Astrom, K.J., Wittenmark, B., 1989. Adaptive control. Addison-Wesley Publishing Company.

Ballabrera-Poy, J., Brasseur, P., Verron, J., 2001. Dynamical evolution of the error statistics with the Seek filter to assimilate altimetric data in eddy-resolving ocean models. Q.J.R. Met. Soc. 127, 233–253.

Blanke, B., Delecluse, P., 1993. Variability of the tropical Atlantic ocean simulated by a general circulation model with two different mixed layer physics. J. Phys. Oceanogr. 23, 1363–1388.

Brasseur, P., Ballabrera-Poy, J., Verron, J., 1999. Assimilation of altimetric observations in a primitive equation model of the gulf stream using a singular evolutive extended Kalman filter. J. Mar. Syst. 22, 269–294.

Cane, M.A., Kaplan, A., Miller, R.N., Tang, B., Hackert, E.C., Busalacchi, A.J., 1996. Mapping tropical Pacific sea level: data assimilation via a reduced state Kalman filter. J. Geophys. Res. 101 (No. C10), 599–617.

Dee, D.P., Da Silva, A.M., 1999. Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. Mon. Wea. Rev. 127, 1822–1834.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99, 10143–10162.

Fukumori, I., 2002. A partitioned Kalman filter and smoother. Mon. Wea. Rev. 130, 1370–1383.

Fukumori, I., Malanotte-Rizzoli, P., 1995. An approximate Kalman filter for ocean data assimilation: an example with an idealized gulf stream model. J. Geophys. Res. 100, 6777–6793.

Ghil, M., Malanotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography. Adv. Geophys. 33, 141–266.

Hoang, H.S., De Mey, P., Tallagrand, O., Baraille, R., 1995. Assimilation of altimeter data in multilayer quasi-geostrophic model by simple nonlinear adaptive filter. Proceedings of International Symposium on Assimilation Assimilation of Observations in Meteorology and Oceanography, Tokyo, Japan. pp. 521–526.

Hoteit, I., Pham, D.T., Blum, J., 2001. A semi-evolutive partially local filter for data assimilation. Mar. Pol. Bull. 43, 164–174.

Hoteit, I., Pham, D.T., Blum, J., 2002. A simplified reduced order Kalman filtering and application to altimetric data assimilation in tropical Pacific. J. Mar. Syst. 36, 101–127.

Houtekamer, P., Mitchell, H., 1997. Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev. 126, 796–811.

Ide, K., Bennett, A.F., Courtier, P., Ghil, M., Lorenc, A.C., 1995. Unified notation for data assimilation: operational, sequential and variational. J. Met. Soc. Japan 75, 181–189.

Jazwinski, A.H., 1970. Stochastic processes and filtering theory. Academic Press, New York.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Trans. ASME Ser. D. J. Basic Eng. 82D, 35–45.

Levitus, S., 1982. Climatological atlas of the world ocean. US Government Printing Office, Washington, DC NOAA Prof. Paper No. 13, 173 p.

Madec, G., Delecluse, P., Imbard, M., Levy, C., 1997. Ocean general circulation model reference manual. Technical Report. University Pierre and Marie Curie, Paris VI.

Mitchell, H.L., Houtekamer, P.L., 2000. An adaptive ensemble Kalman filter. Mon. Wea. Rev. 128, 416–433.

Pham, D.T., Verron, J., Roubaud, M.C., 1997. Singular evolutive Kalman filter with EOFs initialization for data assimilation in oceanography. J. Mar. Syst. 16, 323–340.

Preisendorfer, R., 1988. Principal component analysis in meteorology and oceanography. Elsevier Sci. Publ 425 p.

Pinardi, N., Rosati, A., Pacanowski, R., 1995. The sea surface pressure formulation of rigid lid models: implications for altimetric data assimilation studies. J. Mar. Syst. 6, 109–119.

Testut, C.M., 2000. Assimilation de données sattelitaires avec un filtre de Kalman de rang réduit dans un modèle aux équations primitives de l'océan Atlantique. Ph.D. thesis, Université Joseph Fourier, Grenoble, France.

Verlaan, M., Heemink, A.W., 1995. Reduced rank square root filters for large scale data assimilation problem. Second International Symposium on Assimilation of Observations in Meteorology and Oceanography. WMO, pp. 247–252.

Verron, J., Gourdeau, L., Pham, D.T., Murtugudde, R., Busalacchi, A.J., 1998. An extended Kalman filter to assimilate satellite altimeter data into a nonlinear numerical model of the tropical pacific: method and validation. J. Geophys. Res. 104, 5441–5458.