Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

A Low Cost Filter Design for State and Parameter Estimation in very High Dimensional Systems

Hong Son Hoang * **Remy Baraille** **

* *SHOM/HOM/REC, 42 av Gaspard Coriolis 31057 TOULOUSE FRANCE (Tel: 33 (0)5 67 76 68 23, e-mail: hhoang@shom.fr)*
** *SHOM/HOM/REC, 42 av Gaspard Coriolis 31057 TOULOUSE FRANCE (Tel: 33 (0)5 67 76 68 44, e-mail: remy.baraille@shom.fr)*

Abstract: In this paper we consider the filter design problem for very high dimensional systems. Assuming the hypothesis on separability of the vertical and horizontal structure for the error covariance, the number of unknown elements in the error covariance is reduced drastically and are estimated from generated error samples. A low-cost filtering algorithm is thus determined and parameterized up to some pertinent gain coefficients to be tuned to optimize the filter performance. Results from the experiment on assimilation of the sea surface height (SSH) into an oceanic numerical model demonstrate the high effectiveness of the proposed filtering approach.

Keywords: Filtering problems; Adaptation; Error estimation; Distributed-parameter systems; Environmental engineering.

1. INTRODUCTION

Partial Differential Equations (PDEs) are used practically in all scientific areas, from financial markets to mathematical biology ... not to say on the well-known domains like quantum mechanics, electrodynamics, oceanography and meteorology ... As a mathematical model is only a simplification or abstraction of a (complex) real world, the measurements (observations) constitute the most important source of information which should be used promptly in order to improve the model solution for practical problems. This task can be excellently accomplished by filtering algorithms.

This paper addresses the problem with the design of an efficient filtering algorithm for very high dimensional systems at a moderately low cost. High dimensional systems we mean here are that resulting from a set of PDEs. In such systems or numerical models, a typical dimension of the system state is of order $10^6 - 10^8$. As to observations, a typical dimension of the 2d image vector is $10^4 - 10^5$.

Various approaches have been pursued to overcome the difficulties associated with very high dimensionality of the dynamical systems. This includes the methods such as Optimal Interpolation (OI) Cooper et al. [1996] and statistical interpolation (Kriging; Stein [1999]); variational 3D-Var and 4D-Var Talagrand et al [1987]; Reduced-order adaptive filter Hoang et al [1997]; Kalman filter and its variants like reduced rank and ensemble Kalman filters (EnKF) Evensen [2003]; and computationally more demanding methodologies such as Sequential Monte Carlo (SMC), Particle Filters (PF) Doucet et al. [2000] ... Generally speaking there are two principal classes of methods used so far for solving data assimilation problems : sequential (filtering, or real-time assimilation systems) and non-sequential (or retrospective assimilation) where observation from the future can be used. For a review of data assimilation methods, see Ghil et al [1991]. The filtering algorithm developed in this paper is based essentially on the work Hoang et al [1997], with a more detailed assumption on the structure of the ECM (Section 2) and a new formulation of the optimization problem for seeking unknown parameters in the proposed structure (Section 3). The resulting structure of the filter gain is obtained in Section 4. The experiment on data assimilation in very high dimensional systems is presented in Section 5, principally with the ocean model MICOM. We want to stress that throughout the paper, the algorithms of simultaneous perturbation stochastic approximation (SPSA) Spall [2000] will be used to estimate the unknown parameters in the ECM as well as to optimize the filter performance (adaptive filter - AF). The efficiency of the SPSA for optimization problems in the filter design has been demonstrated in Hoang et al [2011b]. The conclusions are given in Section 6.

2. PROBLEM FORMULATION

Consider the standard filtering problem for the partially observed process $[x(k), z(k)]$,

$$x(k+1) = \Phi x(k) + w(k), k = 0, 1, 2, ... \quad (1)$$

$$z(k+1) = Hx(k+1) + \epsilon(k+1), k = 0, 1, 2, ... \quad (2)$$

here $x(k)$ is the n-dimensional system state at the time instant k, $z(k)$ is the p-dimensional observation vector. We assume $w(k), \epsilon(k)$ are uncorrelated sequences of zero mean and time-invariant covariance Q and R respectively. Mention that the system (1) is considered as a state-space representation of the numerical model derived from a set of PDEs discretized at some spatial grid. Let the state $x(k)$ be estimated by the filter

$$\hat{x}(k+1) = \hat{x}(k+1/k) + K(k+1)\zeta(k+1),$$
$$\hat{x}(k+1/k) = \Phi\hat{x}(k), \quad (3)$$

where $\zeta(k+1) = z(k+1) - H\hat{x}(k+1/k)$ is the innovation vector, $\hat{x}(k+1)$ is the filtered (or analysis) estimate, $\hat{x}(k+1/k)$ is the prediction for $x(k+1)$. From the Kalman filtering theory Kalman [1960], the optimal gain $K(k+1)$ is given in the form

$$K(k+1) = M(k+1)H^T[HM(k+1)H^T + R]^{-1} \quad (4)$$

where $M(k+1)$ is the ECM for the prediction error (PE) $e_p(k+1) := \hat{x}(k+1/k) - x(k+1)$ (will be denoted as PECM). The algorithm for computing $M(k+1)$ is given by the Kalman filter (KF). For the system with the state of dimension of order $10^6 - 10^7$, it is impossible to apply the KF since the matrix $M(k)$ is composed from $10^{12} - 10^{14}$ elements. The sub-optimal filtering problem is stated as follows: given the class $E_M(\theta)$ of matrices $M(\theta)$ known up to the vector of parameters θ, find an optimal θ^o which minimizes

$$J(\theta) = E[||M(\theta) - M_d||_F^2] \to \min_\theta \quad (5)$$

where $E(.)$ denotes mathematical expectation, $||.||_F$ is the Frobenius matrix norm, M_d is a given "data" matrix (to be specified later).

2.1 On the matrix M_d

If the true $M(k)$ is known, the best way is to put $M_d := M(k)$. However, as $M(k)$ is impossible to compute, we propose to obtain its samples $M_k[n_u]$ from the set of PE samples $S_k[n_u]$ which consists of generated approximations of n_u leading real Schur vectors of the system dynamics Φ. These samples are obtained by applying the sampling procedure (SP) in Hoang et al [2011a] : At each time instant k we integrate $n_u + 1$ times the dynamical system, one from the state $x(k)$ and n_u from the perturbed states $x(k) + \delta x^l(k), l = 1, ..., n_u$. This produces a set of n_u perturbations,

$$S_{k+1}[n_u] := [s_1(k+1), ..., s_{n_u}(k+1)] \quad (6)$$

which are orthogonalized and used next to generate a new set of perturbations ...

2.2 Structure for $M(\theta)$

Suppose the solution of numerical model is a space-time stochastic process $x(k, \mathbf{s})$ where k represents a time instant t_k, $s \in R^d$, $d \geq 1$. For $d = 3$, $s = (i, j, l)$, (i, j) denotes a horizontal grid point and l - vertical coordinate. Let $M \in R^{n \times n}$ be the ECM, i.e. $M = M(\mathbf{s}, \mathbf{s}')$. We will assume that the ECM M has a separable vertical-horizontal structure (SeVHS),

$$M(s, s') = M_v(s_v, s_v') \otimes M_h(s_h, s_h'), s_v := l, s_h := (i, j), \quad (7)$$

where $\otimes$ denotes the Kronecker product on two matrices,

$$M_v(s_v, s_v') \otimes M_h(s_h, s_h') = M(i, j, l; i', j', l') =$$
$$\begin{pmatrix} m_v(1,1)M_h & m_v(1,2)M_h & ... & m_v(1,n_v)M_h \\ m_v(2,1)M_h & m_v(2,2)M_h & ... & m_v(2,n_v)M_h \\ ... & ... & ... & ... \\ m_v(n_v,1)M_h & m_v(n_v,2)M_h & ... & m_v(n_v,n_v)M_h \end{pmatrix} \quad (8)$$

3. ESTIMATION OF PARAMETERS IN M_V AND M_H

3.1 Parametrized ECM

For the given M_d, our task is to determine two matrices M_v and M_h from (7) to minimize (5), where $M(s, s') := M_\theta(s, s')$ with $M_v(\theta)$ and $M_h(\theta)$ depending on the vector of unknown parameters θ. We have thus to find the *best separable* (Kronecker product) approximation for M_d.

Remember that the closely related nearest Kronecker product problem (NKP) is formulated in Golub [1996] (p. 712) as follows: Suppose that $A \in R^{m \times n}, m = m_1 m_2, n = n_1 n_2$. The KNP involves minimizing

$$\phi(B, C) = ||A - B \otimes C||_F$$

where $B \in R^{m_1 \times n_1}, C \in R^{m_2 \times n_2}$. The solution to this problem has been presented from a linear algebraic point of view in Golub [1996] using singular value decomposition of a permuted version of A and requires to rearranging A into another matrix $A_R \in R^{m_1 n_1 \times m_2 n_2}$ such that the sum of squares that arise in $||A - B \otimes C||_F$ is exactly the same as the sum of squares that arise in $||A_R - vec(B)vec(C))||_F$ where $vec(X)$ is the vector representation for the matrix X, i.e. $vec(X)$ is obtained by stacking the columns of X on top of each other.

In what follows we present a new simple algorithm for estimating the parameters in M_v and M_h. This new algorithm minimizes $E(||.||_F^2)$ - the mean of the square of $||.||_F$ (see (5)) subject to unknown parameters of M_v and M_h. This results in linear equations for the unknown elements of M_v and the formulas for computing these unknowns follow immediately. As to the matrix M_h, a parametrized structure shall be introduced. Mention that separable covariance models are a common way to model spatial covariances : the joint vertical-horizontal covariance is factored into the product of covariance functions each of which depends only on vertical or horizontal coordinates (Daley [1991], Section 4.3). In many works (for example, Simonovski et al [2004]), the unknown parameters in M_h are estimated directly from the set of 2d PE samples.

Return to Eq. (8). Today, in meteorology and oceanography usually the number of vertical layers $n_v < 100$. It is therefore possible to estimate all the elements $c_{km} := m_v(k, m)$ of the matrix M_v without assuming hypotheses like homogeneity or isotropy for the vertical error covariances. As to M_h, we will assume that it is analytically well determined up to some vector of unknown parameters. For example, the ECM M_h can be assumed to have the structure like a Gaussian, first-order (second-order) auto-regressive models (FOAR, SOAR)... In what follows, for illustration purpose, let M_h be represented in the form $M_h = DC_hD$ with C_h - correlation matrix, $D = diag(\sigma_1, ..., \sigma_{n_v})$, $\sigma_{s_h}^2$ is the error variance at the point s_h. We shall assume two following FOAR structures for C_h,

$$C_h(s_h, s_h') = exp[-d/L_d], d = d(s_h, s_h') \quad (9)$$

$$C_h(s_h, s_h') = exp[-(d_x/L_x + d_y/L_y)], \quad (10)$$

where $d = d(i,j;i'j') = \sqrt{(i-i')^2 + (j-j')^2}, d_x = |i - i'|, d_y = |j - j'|$, L_d has the meaning of correlation length (L_x and L_y are correlation lengths in x- and y-direction). Thus for the model (8),(9) the vector of parameters θ have $\frac{(n_v+1)n_v}{2} + 1$ parameters to be estimated. As to the model (8),(10), this number is equal to $\frac{(n_v+1)n_v}{2} + 2$.

3.2 Parameter estimation

Suppose we are given the ensemble of PE samples $S_k[n_u]$ (see (6)) which are obtained by applying the Sampling Procedure (Samp-Proc) in Hoang et al [2011a]. Let the ECM $M(k)$ in (4) be estimated from $S_k[n_u]$ as $\hat{M}(k)$,

$$\hat{M}(k) = \frac{1}{k'} \sum_{k'=1}^{k} M_{k'}, M_{k'} := \frac{1}{n_u} S_{k'}[n_u] S_{k'}^T[n_u] \quad (11)$$

For the problem (8)-(9), define the vector of unknown parameters as

$$\theta := (c_{11}, ..., c_{1n_v}, c_{21}, ..., c_{2n_v}, c_{n_v 1}, ..., c_{n_v n_v}, L_d)^T. \quad (12)$$

where $c_{lm} := m_v(l,m)$ (see (8)).

Considering $M_{k'}$ as a sample for the ECM $M(k)$, introduce the optimization problem for determining the vector θ,

$$J[\theta] = E[\Psi(M(k), \theta)] \to \min_\theta,$$
$$\Psi(M(k), \theta)] := ||M(k) - M_v(s_v, s_v') \otimes M_h(s_h, s_h')||_F^2 \quad (13)$$

Comment 3.1. Compared to the original NKP described above, the optimization problem (13) is different in the sense that it is aimed at minimizing the mean squared of the Frobenius norm of the difference between the sample M_k and Kronecker product of two matrices $M_v(s_v, s_v')$ and $M_h(s_h, s_h')$.

1) Estimation of the vertical covariance matrix M_v

Using a SPSA algorithm Spall [2000], the optimal parameters c_{lm} can be approached asymptotically by

$$c_{lm}(k+1) = c_{lm}(k) - \gamma(k+1)\nabla_{c_{lm}(k)}\Psi[M_k, \theta(k)],$$
$$\nabla_{c_{lm}(k)}\Psi[M_k, \theta(k)] =$$
$$\frac{1}{n_u}\sum_{l=1}^{n_u}\sum_{i,j;i',j'}[\delta x_p^l(i,j,l)\delta x_p^l(i',j',m) -$$
$$c_{lm}(k)exp[-\frac{d}{L_d(k)}]]exp[-\frac{d}{L_d(k)}] \quad (14)$$

where $\{\gamma(k)\}$ is a scalar sequence ensuring a convergence of $\{c_{lm}(k)\}$ Spall [2000]. A more quick convergence for $\{c_{lm}(k)\}$ can be obtained by writing out the cost function (13) in terms of the time average of PE samples

$$J[\theta] = E[\Psi(M_k, \theta)] \approx J_k[\theta] = \frac{1}{k}\sum_{k'=1}^{k}\Psi(M_{k'}, \theta) \quad (15)$$

Without loss of generality, let $n_u = 1$. Taking a derivative of J_t with respect to c_{lm} implies the system of equations

$$\nabla_{c_{lm}} J_k = 0, l = 1, ..., n_v, m = 1, ..., n_v \quad (16)$$

and the solution c_{lm} is given by

Theorem 3.1. Consider the optimization problem (13) under the conditions (8)-(9). Then for a given $L_d = L_d(k)$, the elements $c_{lm}(k)$ of the vertical covariance matrix M_v which solve the system of equations (15), are defined uniquely by

$$c_{lm}(k) = \frac{\bar{a}_{lm}(k)}{\bar{b}(k)}, \bar{a}_{lm}(k) = \frac{1}{k}\sum_{\tau=1}^{k} a_{lm}(\tau) =$$
$$\bar{a}_{lm}(k-1) + \frac{1}{k}[a_{lm}(k) - \bar{a}_{lm}(k-1)],$$
$$\bar{b}(k) = \frac{1}{k}\sum_{\tau=1}^{t} b(\tau) = \bar{b}(k-1) + \frac{1}{k}[b(k) - \bar{b}(k-1)],$$
$$a_{lm}(\tau) := \sum_{i,j;i'j'} \delta x_p^1(i,j,k;\tau)\delta x_p^1(i',j',k';\tau)exp[-\frac{d}{L_d}],$$
$$b(\tau) := \sum_{i,j;i'j'} exp[-2\frac{d}{L_d}]$$
$$(17)$$

Comment 3.2. The elements $c_{lm}(k)$, determined by Theorem 3.1, can be thus computed recursively as that given by (14).

Comment 3.3. For simplicity Theorem 3.1 is proved subject to the horizontal covariance function (9). The proof remains identical, with minor modifications, subject to other correlation functions.

2) Estimation of correlation length L_d.

Taking a derivative of J in (13) with respect to the correlation length L_d leads to $\nabla_{L_d} J[\theta] = 0$ from which, similarly to (14), the recursive equation for estimating the correlation length can be obtained.

4. STRUCTURE OF FILTER GAIN

It is interesting to see how looks the filter gain when the ECM has a SeVHS. Introduce the notations: at the instant k, let $x(i,j,l)$ be the value of the system state defined at the grid points (i,j,l). Let $vec(x) = (vec(x)_1^T, vec(x)_2^T, ..., vec(x)_{n_v}^T)^T$ be a vector representation for x where $vec(x)_l$ is a vector whose components are the values of x at all the horizontal grid points (ordered in some way) at the l^{th}- vertical layer.

Consider the ECM (8) and the observation equation (2). Represent the observation matrix H in a block-matrix form

$$H = [H_1, ..., H_{n_v}] \quad (18)$$

which corresponds to the vector representation $vec(x)$, i.e.

$$Hvec(x) = \sum_{\nu=1}^{n_v} H_\nu vec(x)_\nu$$

Compute the gain according to (4). We have

$$M(k)H^T = M_v \otimes M_h H^T = [\Sigma_1^T, ..., \Sigma_{n_v}^T]^T,$$

$$\Sigma_l = M_h \sum_{k=1}^{n_v} c_{lk} H_k^T = M_h G_{v,l},$$
$$M(k)H^T = M_d G_v, M_d = \text{block diag}[M_h, ..., M_h],$$
$$G_v = [G_{v,1}^T, ..., G_{v,n_v}^T]^T, G_{v,l} := \sum_{k=1}^{n_v} c_{lk} H_k^T,$$
$$\Sigma := HM(k)H^T + R = \sum_{k=1,l=1}^{n_v} c_{lk} H_l M_h H_k^T + R, \tag{19}$$

This result is formulated in

Theorem 4.1. Suppose that the ECM has a SeVHS (7)-(8). Then the filter gain (4) has the following form

$$K = M_d K_v, K_v = [K_{v,1}^T, ..., K_{v,n_v}^T]^T = G_v \Sigma^{-1},$$
$$K_{v,l} = G_{v,l} \Sigma^{-1}, l = 1, ..., n_v, \tag{20}$$

where $G_{v,l}, \Sigma$ are defined by (19).

In particular when H_l are identical for all l we have

Corollary 4.1. Consider the situation in Theorem 4.1 and suppose that $H_l = H_c, \forall l = 1, ..., n_v$, H_c is some constant matrix. Then the filter gain (4) has the structure (20) subject to

$$K = K_v \otimes K_h,$$
$$K_v := [K_{v,1}, ..., K_{v,n_v}]^T, K_{v,l} = \gamma_l,$$
$$K_h = M_h H_c^T \Sigma_h^{-1}, \Sigma_h := H_c M_h H_c^T + R_c,$$
$$\gamma_l := \frac{\sum_{k=1}^{n_v} c_{lk}}{\sigma_c}, R_c = \frac{R}{\sigma_c}, \sigma_c = \sum_{k=1,l=1}^{n_v} c_{lk}. \tag{21}$$

Comment 4.1. As σ_c is independent on vertical layers, from Corollary 4.1 it follows that the gain K has a SeVHS too. The gain (19) is a particular form of the reduced-order gain postulated in Hoang et al [1997]

$$K = P_r K_e \tag{22}$$

subject to $P_r := K_v \otimes I_p, K_e := K_h$.

Comment 4.2. The gain (21) is obtained in Hoang et al [2001] directly from the assumption (21). In general the assumption (21) can be considered as an induced approximation for the gain structure.

5. EXPERIMENT WITH OCEANIC MICOM MODEL

5.1 MICOM model

The MICOM model used in this experiment is exactly as that presented in Hoang et al [2011a]. We recall only that the model configuration is a domain situated in the North Atlantic from 30^0 N to 60^0 N and 80^0 W to 44^0 W. The grid spacing is about 0.2^0 in longitude and in latitude, requiring the horizontal mesh $i = 1, ..., 140; j = 1, ..., 180$. The distance between two points $\Delta x = x_{i+1} - x_i \approx 20km$, $\Delta y = y_{j+1} - y_j \approx 20km$. The number of layers in the model $n_v = 4$. We note that the state of the model $x := (h, u, v)$ where $h = h(i, j, l)$ is the thickness of l^{th} layer, $u = u(i, j, l), v = v(i, j, l)$ are two velocity components. The "true" ocean is simulated by running the model from "climatology" during two years. Each ten days the sea surface height (SSH) are stored at the grid points $i_o = 1, 10, 20, ..., 140; j_o = 1, 10, 20, ..., 180$ which are considered as observations in the assimilation experiment. The sequence of true states will be available and allows us to compute the estimation errors. Thus the observation operator H is constant at all assimilation instants.

The assimilation experiment consists of using the SSH to correct the model solution, which is initialized by some arbitrarily chosen state resulting from the control run.

5.2 Different filters

The different filters will be implemented to solve this assimilation problem. First the filter PEF will be constructed whose PECM is obtained on the basis of (11) and Theorem 3.1. Parallelly two other filters, one is a Cooper-Haines filter (CHF) Cooper et al. [1996] and another is an EnOI (Ensemble based Optimal Interpolation) filter Greenslade et al [2005] will be constructed too. These well known filters will serve as references to compare their performances with that of the PEF.

We assume that the covariance has a SeVHS structure. The horizontal covariance function (9) will be used in the designed filters. Here $d = \sqrt{(i - i')^2 + (j - j')^2}$, $s_h(i, j) := (x_i, y_j), s'_h(i', j') := (x'_i, y'_j)$ and (i, j) is denoted as the (x_i, y_j) grid point. Thus the correlation length is expressed in term of the number of grid points.

Mention that the CHF applies a vertical rearrangement of water parcels (see also Hoang et al [2011a]). The method conserves the water masses and maintains geostrophy. The main difference between PEF and EnOI is lying in the way to generate the ensembles of PE samples for simulating the PE realizations. In the PEF, the ensemble of PE samples is generated using the Samp-Proc (and it will be denoted as $En(PEF)$). As for the EnOI, the ensemble of background errors samples (the term used in Greenslade et al [2005] and will be denoted by $En(EnOI)$) will be used. The elements of $En(EnOI)$ are constructed according to the method in Greenslade et al [2005]. It consists of using 2-year mean of true states as the background field and the error samples are calculated as differences between individual 10-day true states during this period and the background. Based on these two ensembles of PE samples, the vertical gain coefficients $K_{v,l}, l = 1, ..., 4$ and length scales for two filters PEF and EnOI are computed as described in Section 3.2.

Fig. 1 shows the estimated coefficients $K_{v,l}^{pef}, l = 1, ..., 4$ obtained on the basis of $En(PEF)$. It is seen that the estimates converge quite quickly. The estimated gain coefficients based on two ensembles $En(PEF), En(EnOI)$ at the iteration $t = 72$ are

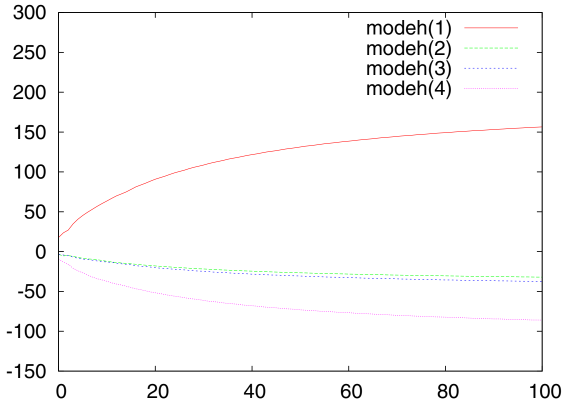

Fig. 1. Vertical gain coefficients obtained during application of the Samp-Proc for layer thickness correction

$$K^{pef}_{v,l} = [144.594, -29.532, -34.439, -80.120]^T$$
$$K^{enoi}_{v,l} = [34.047, -7.532, -3.305, -22.210]^T \quad (23)$$

The reason for the choice $k = 72$ is that in practice the ensemble $En(EnOI)$ has only a limited number of samples and for the comparison purpose we want to use two ensembles of the same number of samples. We remark that all the gain coeffcients in two filters are of identical sign but the elements of $K^{enoi}_{v,l}$ are of much less magnitudes than that of $K^{pef}_{v,l}$,

Two gains in (23) will be used in the two filters PEF and EnOI to assimilate the observations.

As to the correlation length, Fig. 2 displays its estimates obtained on the basis of two ensembles $En(PEF)$ and $En(EnOI)$. The values produced at the iteration $k = 72$ for each ensemble will be taken as correlation lengths in two filters PEF and EnOI. The correlation length in the CHF is assigned the value $L_d(CHF) = 20$ which is the guess value in the procedures for estimating $L_d(PEF), L_d(EnOI)$. The gain coefffficient for CHF is taken from Hoang et al [2011a] and is equal to

$$K^{chf}_{v,l} = [185.965, 0, 0, -184.964]^T \quad (24)$$

5.3 Numerical results

In Fig. 3 we show the instantaneous variances of the SSH innovation produced by three filters EnOI, CHF and PEF. It is seen that initialized by the same initial state, if the innovation variances in EnOI, CHF have a tendency to increase, this error remains stable for the PEF during all assimilation period. At the end of assimilation, the PE in the CHF is more than two times greater than that produced by the PEF. The EnOI has produced very poor estimates, with error about two times greater than the CHF has done. For the velocity estimates, the same tendency is observed as seen from Fig. 4 for the surface velocity PE errors. These results prove that statistically the members from $En(PEF)$ much better represent the PE compared to the samples taken from $En(EnOI)$.

Using the second-order SPSA algorithm, we have adjusted the parameters in the filter gain to minimize the mean variance of innovation vector. The gain parameterization is done exactly as described in Hoang Hoang et al [2012] by changing variables from layer thickness to layer interface variables. For a four layer model, by desiring the filter to produce its output to be matched with the observation (the observations are noise-free), the first coefficient is kept unchanged $\alpha = 1$ and the other three parameters $\alpha_k, k = 2, 3, 4$ are adjusted. The variances of innovation in the CHF and its adaptive version (ACHF) are displayed in Fig. 5 (the curves CHF and ACHF). It is seen that adaptation allows to stabilize the ACHF and improve considerably its performance: the ACHF has produced almost the same error level as that obtained by the PEF. In the same manner, the adaptation allows to reduce sensibly the estimation error in the PEF and EnOI.

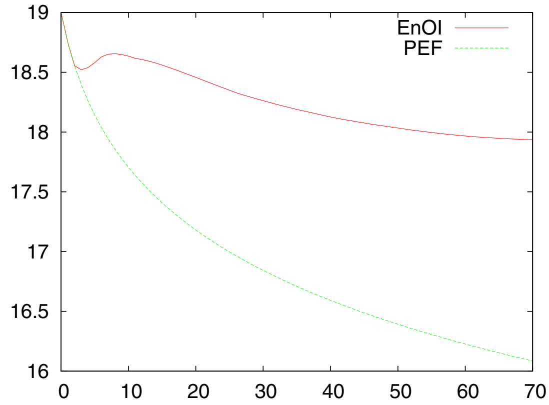

Fig. 2. Length scales estimated on the basis of two ensembles of simulated PE samples $En(EnOI)$ and $En(PEF)$.

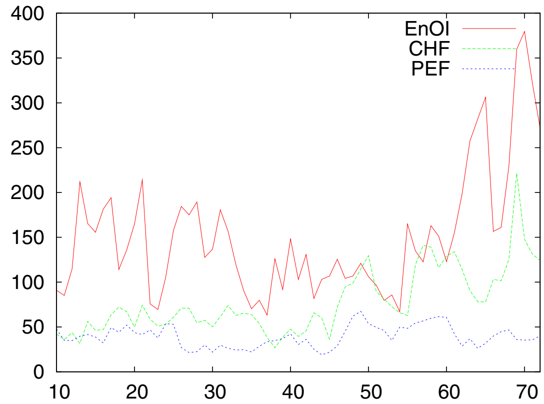

Fig. 3. Performance comparison of EnOI, CHF and PEF : Variance of SSH innovation resulting from the filters EnOI, CHF and PEF

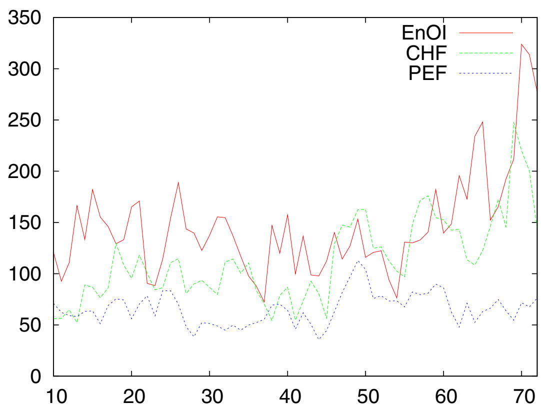

Fig. 4. The prediction error variance of the u velocity component at the surface (cm/s) resulting from the EnOI, CHF and PEF

.

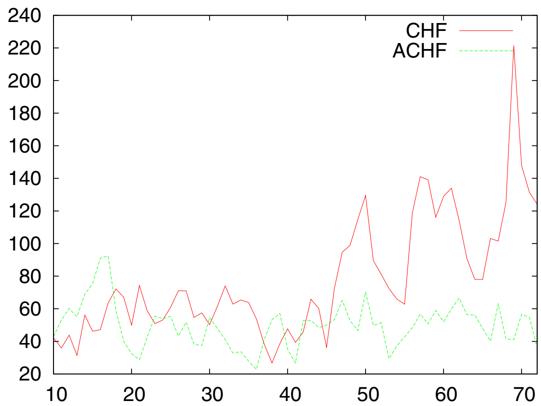

Fig. 5. Variance of SSH innovation resulting from CHF and its adaptive version (AOI)

6. CONCLUSIONS

We have presented a design procedure for a low-cost filter that produces the state estimate in very high dimensional systems. While the construction of a filter gain based on unstable eigenvectors (EVs) or Schur vectors (ScVs) system dynamics constitutes a solid theoretical background and allows to reduce enormous computational and memory requirements, its implementation is possible today only for systems with a moderate number of unstable eigenvalues (order about O(100)). Otherwise, seeking directly the correction in the subspace of all unstable EVs or ScVs is still practically unrealizable.

To work with systems independently on their dimensions, the SeVHS hypothesis on ECM structure has been introduced. The objective is twofold. Firstly it allows to ensure a positive definiteness of the ECM whereas involving an insufficient number of samples to approximate the ECM leads to its rank deficiency and to poor estimation results. Secondly, this hypothesis allows to parametrize the ECM by a very few number of unknown parameters to be estimated.

The effectiveness of the algorithm has been verified by numerical experiment on a high dimensional oceanic model MICOM.

REFERENCES

R. Daley. *Atmospheric data analysis*. Cambridge University Press, New York, 1991.

G. Evensen. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53, pages 343-367, 2003.

A. Doucet, S. Godsill, C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10 (3): pages 197-208. doi:10.1023/A:1008935410038, 2000.

M. Cooper and K. Haines. Altimetric assimilation with water property conservation. *J. Geophys. Res.*, 101(C1), pages 1059-1077, 1996.

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

G.H. Golub and C.F. van Loan. (1996). *Matrix Computations*, Cambridge University Press, 1996.

D.J.M. Greenslade and I.R. Young. The Impact of Altimeter Sampling Patterns on Estimates of Background Errors in a Global Wave Model. *J. of Atmos. and Ocean Tech.*, December, pages 1895-1917, 2005.

H.S. Hoang and R. Baraille. Prediction error sampling procedure based on dominant Schur decomposition. Application to state estimation in high dimensional oceanic model. *Appl. Math. Comp.*, v.218, no.7, pages 3689-3709, 2011 (a).

H.S. Hoang and R. Baraille. On Efficiency of Simultaneous Perturbation Stochastic Approximation Method for Implementation of an Adaptive Filter *Comput. Tech. Appl.*, 2, pages 948-962, 2011 (b).

H.S. Hoang H.S. and R. Baraille. On Gain Initialization and Optimization of Reduced-Order Adaptive Filter. *Journal IAENG IJAM*, 42:1, pages 19-33, 2012.

H.S. Hoang, R. Baraille and O. Talagrand. On the design of a stable adaptive filter for state estimation in high dimensional system. *Automatica* Vol 37, No 3, pages 341-359, 2001.

H.S. Hoang, P. De Mey, O. Talagrand and R. Baraille. A new reduced-order adaptive filter for state estimation in high dimensional systems. *Automatica*, 33, No 8, pages 1475-1489, 1997.

R.E. Kalman. A New Approach to Linear Filtering and Prediction Problems, *Transactions of the ASME - Journal of Basic Engineering*, Vol. 82: pages 35-45, 1960.

I. Simonovski, M. Kovac and L. Cizelj. Estimating the correlation length of inhomogeneities in a polycrystalline material, *Materials Science and Engineering A*, 381(1-2), pages 273-280, 2004.

J.C. Spall. Adaptive Stochastic Approximation by the Simultaneous Perturbation Method, *IEEE Transactions on Automatic Control*, vol. 45, pages 1839 -1853, 2000.

M.L. Stein. *Statistical Interpolation of Spatial Data: Some Theory for Kriging*, Springer, New York, 1999.

O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, 113, pages 1311-1328, 1987.