On estimation of model error by an adaptive filter

HONG SON HOANG
SHOM
HOM/REC
42 Av Gaspard Coriolis, Toulouse
FRANCE
hhoang@shom.fr

REMY BARAILLE
SHOM
HOM/REC
42 Av Gaspard Coriolis, Toulouse
FRANCE
remy.baraille@shom.fr

Abstract: This paper presents an optimal filtering approach to state and model error (ME) estimation problem, with a deterministic or stochastic ME. The approach is based on the adaptive filtering (AF) algorithm which is aimed at overcoming the difficulties in the filter design with very high dimensionality of the dynamic systems. The objective is to design a filtering algorithm offering potential for improvement of numerical accuracy and reduction of computational burden. A hypothesis on the structure of ME is introduced. The improvement of the AF performance is achieved by tuning some pertinent parameters of the filter gain as well as bias parameters to minimize the prediction error of the system output. Numerical experiments are presented to illustrate the performance of the proposed approach.

Key–Words: Dynamic system, Model error, Adaptive filter, Minimal mean prediction error, Filter stability.

1 Introduction

This work deals with the problem of state estimation in partially observed dynamic systems of very high dimensions, in presence of model error (ME). Given a set of measurements at each assimilation moment, the objective is to obtain system estimates that best exploit the measurements, while facing with ME. This issue happens frequently in data assimilation (DA) for geophysical systems (DA-GeoS) applications, where a numerical model is biased compared to the real physical system. The solution of such problem is important for designing a high quality forecasting system.

Consider the problem of state estimation in a high dimensional system (HdS)

$$\begin{aligned} x(t+1) &= \Phi x(t) + b + w(t), \\ z(t+1) &= Hx(t+1) + v(t+1), \end{aligned} \tag{1}$$

here $x(t)$ is the n-dimensional system state at the t assimilation instant, with the n being of order $10^7 - 10^8$, Φ is the (nxn) fundamental matrix, $z(t)$ is the p-dimensional observation vector, H is the (pxn) observation matrix, w, v are the model and observation noises. We assume $w(t)$ is a stochastic model error (SME), $w(t)$ and $v(t)$ are gaussian white noise processes of zero mean and time-invariant covariance Q and R respectively, and they are mutually uncorrelated.

In a more general form, the problem (1) is formulated as [3]

$$\begin{aligned} x(t+1) &= \Phi x(t) + Bb + w(t), \\ z(t+1) &= Hx(t+1) + Db + v(t+1), \end{aligned} \tag{2}$$

The approach to be developed in the sequel can be easily applied to the system (2). In (1) the function b represents an unknown perturbation, deterministic or stochastic. For simplicity and clarity of presentation, in this paper, we assume that b is a constant deterministic model error (DME). As to Q and R, in practice, if the covariance R is more or less known, the matrix Q is usually unknown. In [11] the adaptive filter (AF) is proposed to deal with HdS (1) under the condition $b(t) = 0$.

The necessity to introduce the AF [11] is dictated by the fact that the traditional Kalman filter (KF) is inappropriate for solving the filtering problems in high dimensional setting. Moreover, in practice, the statistics of the model error (ME) is not known (or poorly given). Using the KF in such situations can produce poor results, not to say on its possible instability.

There is a long history of ME estimation for filtering algorithms, in particular, with the bias and covariance estimation. One of the most original approaches, dealing with the treatment of bias in recursive filtering (known as bias-separated estimation - BSE), is carried out by Friedland in [3]. It has been shown in [3] that the least mean square state estimator for a linear dynamic system augmented with bias states can be decomposed into three parts: 1) a bias-free state estimator; 2) a bias estimator; and 3) a blender. Mention that this BSE has the advantage that it requires fewer numerical operations than the traditional augmented-state implementation. In addition, the BSE avoids numerical ill-conditioning compared to the case of state vectors of large dimension. However, as in the KF, the bias-estimation requires additional matrix equations to be solved which are impossible for HdSs.

For the time-varying bias, see [17]. Extension of Friedlands bias filtering technique to a class of nonlinear systems is done in [14], [16]. In [18], the extension of BSE is studied for randomly time-varying bias in nonlinear systems. For estimating parameters for a stochastic dynamic marine ecological system, see [2].

As to estimation of the model error covariance matrix (ECM), i.e. Q of $w(t)$ in (1), in the filtering algorithm for environmental HdS, it is worth of mentioning the work of Dee [1] on online estimation of model error ECM parameters in atmospheric DA. Recently a new algorithm is proposed in [8] for estimation of high-dimensional ECM of the state prediction error (PE). In this algorithm, the ECM is assumed to be of the structure of Schur product of two matrices, one is of horizontal coordinate, another - of vertical coordinates. The unknown parameters are estimated using a Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm [15]. The optimization problem is formulated as minimization of the mean squared error (MSE) of the difference between the data matrix (generated by the Prediction Error Sampling Procedure (PeSP) [6]) and estimated matrix.

In this paper, we examine in more detail the question on ME estimation in the context of AF approach. Two situations are investigated here, one concerns a DME $b(t)$, another - estimating the ECM of $w(t)$. It is assumed that the SME $w(t)$ is a sequence of zero mean. This assumption is not restricting the class of considered problems because, if $w(t)$ is of non-zero mean, its unknown mean can be considered as a DME $b(t)$.

It is important to stress that the ME estimation algorithms, developed in this paper, are based essentially on the hypothesis concerning the structure of ME and on minimization of the MSE of the innovation process in the AF. The tuning parameters are bias parameters and some pertinent parameters of the filter gain [12].

The simulation results, presented in this paper, show that the proposed ME estimation procedure has a considerable potential for improving a performance of the estimation procedure. Here the experiments are carried out for both small and HdSs.

2 Model error and state estimation

2.1 Bias-separated estimation by filtering technique

It is common to treat the bias b as part of the system state and then estimate the bias as well as the system state. For simplicity of presentation, consider the situation when $b(t)$ is a constant unknown vector. For a more general class of ME, a suitable ME equation can be introduced, for example, by various types of deterministic or stochastic differential equations. In this situation we say on $b(t)$ as a structural ME (which may be periodic, sinusoidal, Markovian ...). If the second order statistics of $b(t)$ is given, a suitable dynamical model for $b(t)$ can be constructed as shown in [7]. In this situation $b(t)$ is a structural stochastic ME.

Under the assumption made above, we have

$$b(t+1) = b(t), t = 0, 1, 2, ... \tag{3}$$

To introduce a subspace for the values of b, one can write $b = Gd$ where $G \in R^{n \times n_e}$, $n \geq n_e$. Generally speaking, G is unknown, and finding a reasonable hypothesis for G is desirable but very difficult. This question will be addressed in the next section. It what follows, for simplicity of presentation, we assume $G = I$- identity matrix.

Introduce $x_g = col(x, b)$ - column vector consisting of two column vectors x and b - as an augmented state. The new input-output system is written for x_g on the basis of (1)

$$x_g(t+1) = \Phi_g x_g(t) + w_g(t),$$
$$\Phi_g := \begin{vmatrix} \Phi & \Phi_2 \\ 0 & \Phi_3 \end{vmatrix}, w_g(t) := col(w(t), 0), \tag{4}$$

where $\Phi_2 = I, \Phi_3 = I$. The observation system (8) now reads as

$$z(t+1) = H_g x_g(t+1) + v(t),$$
$$H_g := [H, 0], \tag{5}$$

and the filtering problem is solved using the observations $z(t)$. This leads to an augmented state Kalman filter (ASKF). whose implementation can be computationally intensive. The main idea in [3] is to separate the problem of estimation of x in the presence of systematic bias b into two estimation problems. First, form a modified bias-free filter by ignoring the bias term and by adding an external bias-compensating input. Second, take the bias into account and derive a bias filter to compensate the modified bias free filter in order to reconstruct the original filter. More concretely, for the filtering problem (1), instead of writing an ASKF for the augmented state $x_g = col(x, b)$ one has

$$\begin{aligned} \hat{x}(t+1) &= \tilde{x}(t+1) + V_x \hat{b}(t+1), \\ \tilde{x}(t+1) &= \Phi\tilde{x}(t) + K_x\zeta_x(t+1), \\ \hat{b}(t+1) &= \hat{b}(t) + K_b\zeta_b(t+1), \\ \zeta_x(t+1) &= z(t+1) - H\tilde{x}(t+1), \\ \zeta_b(t+1) &= \zeta_x(t+1) - H\hat{b}(t), t = 0, 1, 2, ... \end{aligned} \quad (6)$$

The estimate $\hat{x}_g(t) := col(\tilde{x}(t), \hat{b}(t))$, defined by (6), is an unbiased with minimal variance as such obtained by applying an ASKF for the augmented state $x_g(t)$ if V_x is obtained in an optimal way as shown in [3]. For the optimal gain matrices K_x, K_b, see [3].

2.2 Variational method (VM)

In the VM [13], for a set of observations $z(t), t = 1, ..., N$, one seeks the initial state for the augmented state x_g which minimizes the cost function

$$\begin{aligned} J[\theta] &\to \min_\theta, \theta := col(x(0), b(0)), \\ J[\theta] &:= (1/2)\sum_{k=1}^{N} ||z(k) - Hx_g(k)||^2_{M^{-1}}, \end{aligned} \quad (7)$$

Mention the objective function in (7) is not written explicitly as a function of θ. This can be done by expressing $x_g(k)$ through $x_g(0)$ using (4). Thus, in the VM, the bias is considered also as a part of the initial (augmented) state to be estimated along with $x(0)$ to minimize the objective function (7).

3 Adaptive filter and model error

3.1 Adaptive filter

Under the condition that $b(t) = 0$, the AF in [12] is designed as an optimal solution the class of filters of a given (stabilizing) structure. The objective of the AF is to minimize the MSE of the innovation. As seen in section 2.2, if in the VM, the initial state is chosen as a control vector for the optimization problem, in the AF the control vector consists of some pertinent parameters of the filter gain. These parameters are turned during the assimilation process to minimize the MSE of innovation [12]. The choice of tuning parameters is dictated by the wish to ensure a stability of the filter : we select only the parameters, whose values belong to prescribed intervals, so that the filter remains stable during the assimilation process.

The main goal in the development of an AF is to overcome the difficulties encountered in dealing with the systems of high state dimension. For example, it is impossible to solve the Algebraic Riccati Equation (ARE) in the KF. Construction of high performance DA-GeoS represents a great challenge for both theoreticians and practitioners in the field of filtering and estimation.

Consider a non-adaptive filter (NAF)

$$\begin{aligned} \hat{x}(t+1) &= \hat{x}(t+1/t) + K\zeta(t+1), \\ \hat{x}(t+1/t) &= \Phi\hat{x}(t), \end{aligned} \quad (8)$$

where $\zeta(t+1) = z(t+1) - H\hat{x}(t+1/t)$ is the innovation vector, $\hat{x}(t+1)$ is the filtered estimate, $\hat{x}(t+1/t)$ is the prediction for $x(t+1)$, where the structure of the gain $K := K(\theta)$, in difference with that in the KF, is assumed to be given a priori and parametrized by some vector of unknown parameters θ. The AF is obtained by tuning θ to minimize the PE of the system output,

$$J[\theta] = E[\Psi(\zeta(t)] \to \min \theta, \Psi(\zeta(t)) := ||\zeta(t)||^2, \quad (9)$$

where $E(.)$ denotes the mathematical expectation.

Different parameterized structures of K are proposed in [12] which are obtained from the point of view of filter stability: K must be chosen in such a way to ensure a stability of the filter. One of stabilizing gain structures is

$$K = P_r Ke, K_e = M_e H_e^T [H_e M_e H_e^T + R]^{-1}, Me > 0 \quad (10)$$

or

$$K = MH^T[HMH^T + R]^{-1}, M = P_r M_e P_r^T \quad (11)$$

where M_e is a symmetric positive definite (SPD) matrix which represents the ECM for the reduced state (in the reduced space). The matrix $P_r \in R^{n \times n_e}$ is a projection from the reduced space R^{n_e} into the full spaceR^n. The ECM of the full-order state PE $e_p(t+1) = \hat{x}(t+1/t) - x(t+1)$ is given by M. Under the detectability condition, a stability of the filter is guaranteed if the columns of P_r are the unstable and stable eigenvectors (EiVecs) (or singular vectors, or real Schur vectors) of the system dynamics Φ. In [6] the PeSP has been proposed to generate an ensemble of PE samples ($EnPE$) which are samples of the leading real Schur vectors. These samples used as columns of P_r in order to ensure a stability of the filter.

3.2 Model errors

We are interested in two types of the ME : (i) DME $b(t)$; (ii) SME $w(t)$.

Mention in practice of DA-GeoS, only a very limited number of observations (in time) are given hence our aim is to use the observations in the best way possible to estimate $b(t)$. Moreover, as the AF is optimal only in a class of admissible stable filters, the problem of identifiability of $b(t)$ is not of our interest in this paper. As to the SME, the attention is drawn on how one can account for the ECM Q in the structure of the filter gain ?

For both the DME and SME, we will assume in this paper a hypothesis on the subspace for the ME. This hypothesis is based on the observation that in practice of data assimilation, the model time step δt (required for advancing in time the numerical solution) is much less than the assimilation window ΔT (time interval separating two observations). Mention that δt is chosen in order to ensure a stability of numerical scheme and to provide an accuracy of the numerical solution.

4 Estimation of deterministic model error

4.1 Non-adaptive filter : augmented state approach

The filter for solving the problem (4)(5), is proposed of the form

$$\begin{aligned}\hat{x}_g(t+1) &= \hat{x}_g(t+1/t) + K_g\zeta(t+1),\\ \zeta(t+1) &= z(t+1) - H_g\hat{x}_g(t+1/t),\\ \hat{x}_g(t+1/t) &= \Phi_g\hat{x}_g(t). \qquad (12)\end{aligned}$$

The gain K_g can be constructed as proposed in [12] (analogous to (10)(11) for the input-output system (1)(8)).

$$\begin{aligned}K_g &= P_{g,r}K_{g,e},\\ K_{g,e} &= M_{g,e}H_{g,e}^T[H_{g,e}M_{g,e}H_{g,e}^T + R]^{-1},\\ H_{g,e} &= H_g P_{g,r}, \qquad (13)\end{aligned}$$

where $M_{g,e}$ is SPD matrix.

Introduce the Jordan decomposition for Φ [5],

$$\begin{aligned}\Phi = UJU^{-1}, J &= \text{diag}[\sigma_1, \sigma_2, ..., \sigma_n]\\ &|\lambda_1| \geq |\lambda_2| ... \geq |\lambda_n|,\\ U = [U_1, U_2], D &= \text{block diag } [D_1, D_2],\\ \tilde{U} &= U^{-1} = [\tilde{U}_1^T, \tilde{U}_2^T]^T, \qquad (14)\end{aligned}$$

where $U_1, \tilde{U}_1 \in R^{n \times n_s}$, $D_{n_s} \in R^{n_s \times n_s}$, n_s is the number of all unstable and neutral EiVs of Φ. In the future, for simplicity, unless otherwise stated, we say on the set of all unstable EiVs as that including all unstable and neutral EiVs. In the sequel, for simplicity we assume that D_1, D_2 are diagonal and the presentation is done in term of eigenvalue (EiV) decomposition (21). However, all the results presented below are valid for the Schur decomposition of Φ and in practice it is more efficient to work with the real Schur decomposition. In the Schur decomposition, U is an orthogonal matrix.

Looking at Φ_g one sees that according to the AF approach in [12], in order to ensure a stability of the filter, the projection operator $P_{g,r}$ can be chosen in the form

$$P_{g,r} = [P^{(1)}, P^{(2)}] = \left| \begin{array}{cc} U_1 & P_2 \\ 0 & P_3 \end{array} \right|, \qquad (15)$$

where the columns of U_1 are the unstable EiVs (or Schur vectors) of Φ. As to P_2, P_3, they are the matrices of dimensions $R^{n \times n}$. The difficulty we have here is related to computation of $P^{(2)} := [P_2^T, P_3^T]^T$. From the point of view of the stability of the AF [12], as $\Phi_3 = I$ - the transition matrix for the system $b(t+1) = b(t)$, all the EiVs of Φ_3 are neutrally stable, all the EiVs, associated with them, must be taken into account in the construction of $P_{g,r}$. The following Lemma shows that there are two simple ways for the choice of P_2, P_3.

Lemma 4.1. The projection operator $P_{g,r}$ can be chosen in the form (15) subject either to

$$P_2 = U, P_3 = U\Sigma, \Sigma = \text{diag}[1-\lambda_1, ..., 1-\lambda_n], \quad U = [U_1, U_2]. (16)$$

where $\lambda_i, i = 1, ..., n_s$ are all the unstable and neutral EiVs of Φ, or in the form (15) s.t.

$$P_2 = I = [x_1^{(1)}, ..., x_1^{(n)}], x_1^{(i)} = e_i,$$
$$P_3 = I - \Phi = [x_2^{(1)}, ..., x_2^{(n)}], x_2^{(i)} = e_i - \phi^{(i)}, \quad (17)$$

where $e_i := (0, ..., 1, ..., 0)^T, i = 1, ..., n$ with all components of e_i equal to 0 except the i^{th} equal to 1 and $\phi^{(i)}$ is the i^{th} column of Φ.

Comment 4.1. The projection operator based on the formula (16) is difficult to realize because it requires computation of all EiVs of Φ. As to (17), if one wants to have Φ, this requires to make n integrations of the model by $e_i, i = 1, ..., n$ to compute all the columns of Φ (in a linear case). One way to overcome this difficulty is to employ the numerical method proposed in [9] to estimate Φ by perturbing Φ and solve the associated optimization problem using the SPSA algorithm.

In the next section we will show that in fact the NAF can be implemented at the cost of making one more additional integration of the numerical model without the need to store Φ. By this way it is possible to apply the AF with the augmented state approach for very HdSs.

4.2 Non-adaptive filter based (17)

Consider the filtering problem (4)-(5). Let us calculate the filter gain K_g in (20). According to Lemma 4.1, using the projection operator $P_{g,r}$ given in the form (17) and under the assumption $M_{g,e} = I$ one has

$$H_{g,e} = H_g P_{g,e} = [HU_1, H],$$
$$K_{g,e} = [K_e(1), K_e(2)] = H_{g,e}^T \Sigma^{-1},$$
$$K_e(1) = (HU_1)^T \Sigma^{-1}, K_e(2) = H^T \Sigma^{-1},$$
$$\Sigma := [H_{g,e} H_{g,e}^T + R].$$

Using the formula for $P_{g,r}$ we have now

$$K_g = [K_g^T(1), K_g^T(2)]^T,$$
$$K_g(1) = (I + U_1 U_1^T) H^T \Sigma^{-1},$$
$$K_g(2) = (I - \Phi) H^T \Sigma^{-1}. \quad (18)$$

It is seen that Φ participates only in $K_g(2)$. The filter (12) is rewritten as

$$\hat{x}(t+1) = \hat{x}(t+1/t) + c(1),$$
$$\hat{b}(t+1) = \hat{b}(t) + c(2),$$
$$\hat{x}(t+1/t) = \Phi\hat{x}(t) + \hat{b}(t),$$
$$c(1) := (I + U_1 U_1^T) H^T \eta,$$
$$c(2) := (I - \Phi) H^T \eta, \eta = \Sigma^{-1}\zeta(t+1),$$
$$\Sigma = [HUU_1 H^T + HH^T + R]$$
$$\zeta(t+1) = z(t+1) - H\hat{x}(t+1/t), \quad (19)$$

Mention that for HdS, η is found by iteratively solving the equation $\Sigma\eta = \zeta(t+1)$.

In comparison with the case of no bias estimation, the implementation of (19) requires, in addition, only one integration of Φ to perform $c(2) = K_g(2)\zeta(t+1)$. It means that there exists no principal difficulty to implement the filter (19) for very HdS. The obtained result is formulated in the following

Theorem 4.1 Consider the filtering problem (1). Under the *observability* condition of (1), the filter (19) is stable.

The proof of Theorem 4.1 is given in Appendix A.

Mention that including the equation for b makes the filtering problem more complex and using only the information from the subspace U_1 is insufficient for ensuring a stability of the filter for the augmented state. That is why the condition *detectability* of (1) in the case with $b = 0$ is now replaced by the stronger condition *observability* of the system (1). Note that in fact observability of the system (1) is equivalent to detectability of the augmented state system (4)-(5).

4.3 Adaptive filter

It is well known [4] that if the system (1) is stationary, applying the KF filter to (1) s.t. $b := b^*$, where b^* is the true value of b, the Kalman gain $K_{kf}(t)$ will tend to an optimal constant gain

$K_{kf}(\infty) = K^*$ as $t \to \infty$ and $\zeta(t)$ is of minimal variance. This idea has been used in [11] for the design of the AF.

In this paper, we will follow the same idea to construct the AF. Let the gain matrix K_g be chosen as a stabilizing gain for the filter (12), with an appropriate parameterization as done in [12]. We remark that in the ideal situation, if α - the vector of unknown parameters of the gain K_g - consists of all elements of K_g, under the suitable condition (detectability of (4)-(5)), by solving the optimization problem (9), in asymptotic, we obtain a stationary (constant) optimal estimate K_g^* for K_g, with the corresponding innovation $\zeta(t)$ being of minimal variance.

Based on this idea, consider the problem of minimizing (9) s.t. the control vector $\alpha := (\theta^T, \gamma^T)^T$ where θ is the vector of tuning parameters in the gain $K_g := K_g(\theta)$, γ - the vector of unknown parameters in $b := b(\gamma)$. As before, when θ consists of all elements of K, γ - comprises all elements of b, under suitable condition, the innovation $\zeta(t)$ will be of minimal variance and in asymptotic one obtains the optimal gain K_g^*.

Based on the NAF, developed in the previous section, we can write out now a structure for the AF. According to [12], one of the AFs can be of the form (12) where the gain K_g now takes the form

$$\begin{aligned} K_g &= P_{g,r}\Xi_g K_{g,e}, \\ K_{g,e} &= M_{g,e}H_{g,e}^T[H_{g,e}M_{g,e}H_{g,e}^T + R]^{-1}, \\ H_{g,e} &= H_g P_{g,r}, \end{aligned} \quad (20)$$

where

$$\begin{aligned} \Xi_g &= \text{block diag}[\Theta, \Gamma], \\ \Theta &\in R^{n_e \times n_e}, \Gamma \in R^{n \times n}. \end{aligned} \quad (21)$$

The matrices Θ, Γ in (21) may be diagonal with positive diagonal elements. For more details on the constraints for the diagonal elements, see [12]. The diagonal elements of Θ, Γ (denoted by the two vectors (θ, γ)) are the tuning parameters to be adjusted to minimize the objective function (9). Substituting $P_{g,r}$ from (17), Ξ_g from (21), into the gain K_g in (20) yields

$$\begin{aligned} K_g &= [K^T(1), K^T(2)]^T, \\ K(1) &= \text{block diag}[\Theta, \Gamma], \\ \Theta &\in R^{n_s \times n_s}, \Gamma \in R^{n \times n}. \end{aligned} \quad (22)$$

The AF is obtained by tuning $\alpha = col(\theta, \gamma)$ to mimimize the objective function (9). For HdSs, the SPSA algorithm (c.f., [15]) is very useful for solving such optimization problem : This follows from the fact that the SPSA algorithms require to evaluate only the gradient of the sample cost function $\Psi := ||\zeta(t)||^2$, at each assimilation instant and $||\zeta(t)||^2$ is quadratic with respect to (w.r.t.) the vector α. It makes the optimization process much easier and accelerates a convergence of estimated gain parameters. According to the SPSA approach,

$$\alpha(t+1) = \alpha(t) - \mu(t+1)\nabla_{\alpha(t)}\Psi[\zeta(t+1)]. \quad (23)$$

where $\nabla_{\alpha(t)}\Psi[\zeta(t+1)]$ is the gradient of $\Psi[\zeta(t+1)]$ with respect to (wrt) $\alpha(t)$; $\mu(t+1)$ is a scalar or matrix sequence chosen for ensuring a convergence of the algorithm (23).

The gradient $\nabla_{\alpha(t)}\Psi[\zeta(t+1)]$ can be computed by simultaneously perturbing (stochastically) all the components of $\alpha(t)$. This method requires integrating only two or three times the numerical model Φ to generate the forecast and its perturbations for estimating a sample gradient vector. For more detail, see [9]. According to Theorem 5.4 [12], the constraint for the i^{th} component $\alpha_i(t)$ of $\alpha(t)$ is expressed as following

$$\begin{aligned} &\alpha_i(t) \in [1-\epsilon_i, 1+\epsilon_i], \epsilon_i \in (0,1), \\ &\epsilon_i \to 1 \text{ as } |\lambda_i| \to 1, \epsilon_i \to 0 \text{ as } |\lambda_i| \to \infty \end{aligned} \quad (24)$$

From (24) it is seen that the interval of admissible values for $\alpha_i(t)$ approaches to [0,2] for a neutrally stable EiV and it shrinks to consist of one point 1 for a large (by modulus) EiV. It means that there is a relatively large freedom to vary $\alpha_i(t)$ when the corresponding Eiv is close to be neutrally stable. In contrast, for large unstable EiV, no margin left to vary $\alpha_i(t)$. As all $\{\alpha_i(t)\}$ do not have any physical sense, no normalization procedure is imposed for $\{\alpha_i(t)\}$ in the systems with different physical variables. This simplifies the implementation procedure for optimizing the filter performance.

5 Structure of the model error

5.1 Hypothesis on the structure of model error

In this section we will concentrate our attention on the question on whether it is possible to describe a subspace for the values of the ME in the framework of the data assimilation problem ? In the context of the MEs $b(t)$ and $w(t)$, this question is equivalent to saying on whether one can find the relationships

$$b = G_b d, w = G_w \psi, \\ G_b \in R^{n \times n_d}, G_w R^{n \times n_w}, n_d < n, n_w < n. \quad (25)$$

where G_b, G_w are known ? The information, given by (25), allows to better estimate the DME as well as the filter gain, especially for $n_d << n, n_w << n$ in a HdS setting. For example, for the SME, if in Eqs (10)(11) the ECM M is assumed to be constructed on the basis of only the PE samples (from application of the PeSP in [6]), with the assumption $w = G_w \psi$, one can obtain a richer structure for M in the form

$$M_Q = M + Q_w, Q_w = G_w Q_e G_w^T. \quad (26)$$

As to the DME b, there is now a need to estimate only $d \in R^{n_b}$ with a much lower dimension $n_b << n$. In vue of very large n and a small number of observations (compare to n), this structure allows to accelerate a convergence of estimation procedure and to produce a more precise estimate for the vector b.

The difficulty, encountered in the practice of operational forecasting systems (OFS), is that (practically) nothing is given a priori on the ME. This concerns, for example, a subspace spanned by the values of the ME. To overcome this difficulty, a simple hypothesis on this subspace will be introduced. This hypothesis is possible to be postulated if one takes into consideration the fact that for a large number of data assimilation problems, the model time steps δt (chosen for ensuring a stability of numerical scheme and for guaranteeing a high precision of the discrete solution) is much smaller than ΔT - the assimilation window (time interval between two observation arrivals). More precisely, suppose $\Delta t = n_a \delta t$ and $\mu(\tau)$ represents an ME between two model time steps. Symbolically we have (for simplicity, $x'(0) := x(t), x'(n_a \delta t) = x'(\Delta t) = x(t+1), x(t)$ is the system state in (1)). Hence

$$x'(\tau+1) = \phi x'(\tau) + \mu(\tau), \tau = 0, 1, ..., n_a, \\ x'(n_a) = \phi(n_a, 0)x'(0) + \\ \textstyle\sum_{\tau=1}^{n_a-1} \phi(n_a-1, \tau)\mu(\tau),\ \phi(\tau+m, \tau) := \phi^m, \\ \phi(\tau, \tau) = I,\ \phi(\tau, \tau+m) := 0, m > 0,$$

hence

$$x(t+1) = \Phi x(t) + w(t), \Phi = \phi(n_a, 0), \\ w(t) = \sum_{\tau=1}^{n_a-1} \nu(\tau), \nu(\tau) := \phi^{n_a-1-\tau}\mu_(\tau). \quad (27)$$

It is not hard to see that the dominant part of $w(t)$ consists of the members $\nu(\tau)$ for which $n_a - 1 - \tau$ is large, if ϕ is unstable. It means that approximately one can consider $w(t)$ belongs to the subspace of unstable Schur vectors of Φ if n_a is relatively large.

Hypothesis. Under the condition that n_a is relatively large, the ME belongs to the subspace spanned by the unstable and neutral EiVecs (or Schur vectors) of the system dynamics Φ.

In the sequel we will refer to the formulated hypothesis as the Hypothesis on ME (HME).

5.2 Structure of NAF under hypothesis

Under the HME, as $b = U_1 d$, the system (3) becomes

$$b(t+1) = U_1 d(t+1), d(t+1) = d(t), \quad (28)$$

As U_1 is known, there is a need to estimate only the vector d. Introduce $x_g = col(x, d)$. From (1)(28),

$$x_g(t+1) = \Phi_g x_g(t) + w_g(t), \\ \Phi_g := \left| \begin{array}{cc} \Phi & U_1 \\ 0 & I_{n_s} \end{array} \right|, w_g(t) := col(w(t), 0), \quad (29)$$

where $\Phi_2 = U_1, \Phi_3 = I_{n_s}$. The observation system (8) reads

$$z(t+1) = H_g x_g(t+1) + v(t), \\ H_g := [H, 0_{n \times n_s}], \quad (30)$$

Applying the filter, similar to (12)? with the gain K_g of the structure (20) to the filtering problem (29)-(30) yields

Lemma 5.1. The projection operator $P_{g,r}$ can be chosen in the form (15) s.t.

$$P_2 = U_1, P_3 = U_1 S_1, S_1 = \\ \text{diag}[1 - \lambda_1, ..., 1 - \lambda_{n_s}], \quad (31)$$

where $\lambda_i, i = 1, ..., n_s$ are all the unstable and neutral EiVs of Φ, i.e.

$$P_{g,r} = \begin{vmatrix} U_1 & P_2 \\ 0 & P_3 \end{vmatrix} = U_1 \begin{vmatrix} I_{n_s} & I_{n_s} \\ 0 & S_1 \end{vmatrix}.$$

Comment 5.1. The second choice, similar to that done in Lemma 4.1, is of no interest. In fact, we have now for a given $\mu_i, i = 1, ..., n_s$, the equations for the corresponding $x^{(i)} = col(x_1^{(i)}, x_2^{(i)})$ are $x_1^{(i)} = e_i$, $G_b x_2^{(i)} = e_i - \phi^{(i)}, i = 1, ..., n_s$. One sees there exists no interesting choice for $x_2^{(i)}$ as it does for $n_s = n$.

5.3 Structure of the AF under HME

According to [12], one class of stabilizing gains for the AF based on NAF under HME is

$$K_g = P_{g,r} \text{ blockdiag } [\Theta, \Gamma] H_{g,r}^T \Sigma \zeta(t+1),$$
$$\Sigma := [H_{g,r} H_{g,r}^T + R]^{-1}. \quad (32)$$

where $P_{g,r}$ is defined in Lemma 5.1. The diagonal matrices Θ, Γ are similar to those described in Section 4.3.

6 Experiments

6.1 Illustration of the hypothesis

To illustrate the postulated hypothesis, let us consider

$$\Phi = [\phi_{ij}]_{i,j=1}^2, \phi_{11} = 1.02,$$
$$\phi_{12} = 0.1, \phi_{21} = 0, \phi_{22} = 0.9. \quad (33)$$

Numerically one finds the first Schur vector of ϕ equal to $\hat{u}_1 = (-1.0, -7.0 \times 10^{-7})^T$.

Fig. 1 shows the results computed on the basis of (27). One sees that, for $n_a > 10$, the 2nd component is close to 0 whereas the 1st component becomes bigger and bigger (in absolute value) as n_a increases. Here $\mu(\tau)$ is a sequence of independent 2-dimensional Gaussian random vectors of zero mean and variance 1. This means that w becomes more and more close to the subspace spanned by $\hat{u}_1$ hence the hypothesis holds for $n_a > 10$ in this example. Mention that, as a rule, in ocean numerical models, n_a is of order $O(100)$ ($n_a = 800$ for the MICOM model in the experiment in section 6.4).

6.2 Two dimensional systems

Consider the system (1) with Φ given in (33) and $H = [1, 1]$ with the DME $b^*(1) = b^*(2) = 0.1$.

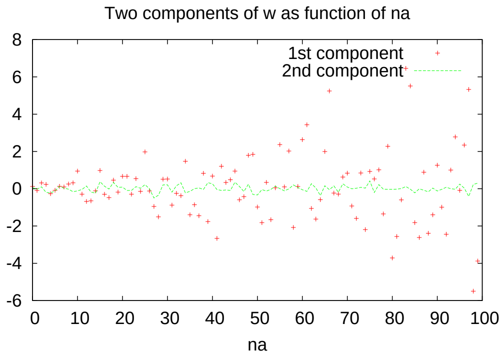

Figure 1: Two components of w as function of n_a

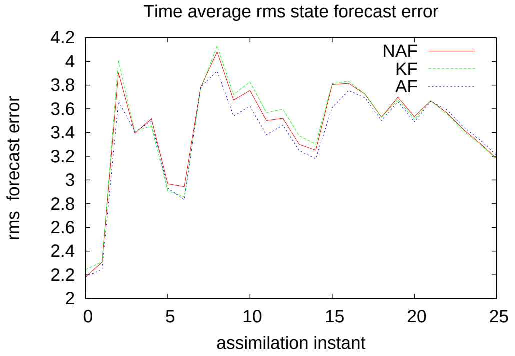

Figure 2: Sample time average RMS of the state forecast error produced by the NAF, KF and AF.

In the experiment, $t = 0, 1, ..., 390$ and the observations are given only at the assimilation moments $k = 1, 2, ..., 26$ which correspond to the model time instants $t = 15, 30, ..., 390$, i.e. $k = t/15$. As seen from Fig. 1, for $n_a = 15$ the separation of the two components of the first Schur vector is well established. The true system states $x^*(t)$ are generated by (1) s.t. $b^* = col(0.1, 0.1)$, $Q = I$, $R = 0.16$. For the KF and NAF, the forecast is obtained at each assimilation instant k as $\hat{x}(k + 1/k) = \Phi_k \hat{x}(k) + b'(k)$ where $\Phi_k = \Phi^{15}$. The vector $b' = col(b'(1), b'(2))$ is obtained by integrating the numerical model s.t. b^* over the assimilation window $t = 1, ..., 15$ and is equal to $b' = (0.2296, 2.0589E{-}02)$. Thus b' represents approximately a correct bias produced by the model over the assimilation window.

The initial gain of the KF is assigned to the gain in the NAF as well as the initial gain for the AF, which is equal to $K = col(0.584, 0.386)$. For the AF, α is updated during the assimilation process along with the gain parameters θ as shown in section 4.3.

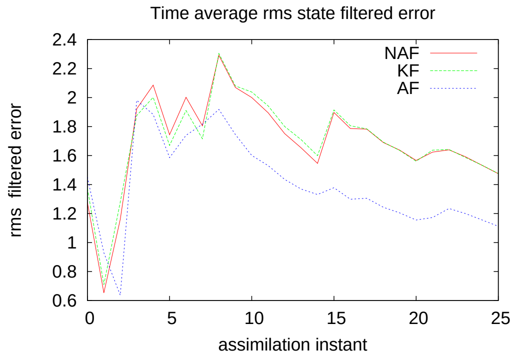

Figure 3: Sample time average RMS of the state filtered error produced by the NAF, KF and AF.

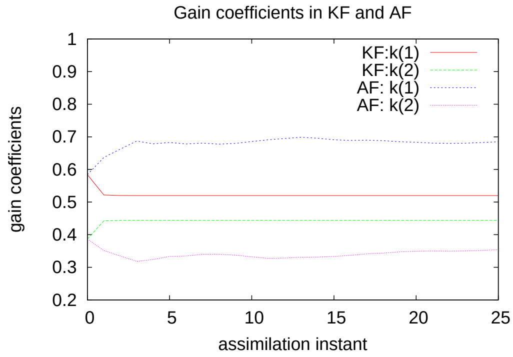

Figure 4: Gain coefficients in KF and AF : The gains in KF and AF are identical at the beginning of the assimilation process.

Fig. 2 displays the sample time average RMS of the state forecast error produced by the NAF, KF and AF. From Fig. 2 and Fig. 3, if there is no significant difference in the forecast performance produced by the three filters, the AF outperforms the NAF and KF in term of the filtered error (Fig. 3) performance. Thus tuning the three parameters $\alpha, \theta(1), \theta(2)$ during the assimilation process allows to improve considerably the performance of the AF. Fig. 4 shows a very quick convergence of the KF gain. This explains why two filters KF and NAF have produced nearly the same performance.

6.3 High dimensional system

For the HdS, the assimilation experiment on the MICOM (Miami Isopycnic Coordinate Ocean Model) has been carried out. For the real operational HdS, unfortunately, the information on the space of model error is usually unavailable. If there exists a model error and we do not take it into account, the error will be accumulated and grows mostly in directions of leading EiVecs (or Schur vectors) of the dynamical system matrix. This motivates us to postulate the hypothesis that $R[Q_w]$ is spanned by a few leading EiVecs of the ECM M, i.e. by assuming

$$Q_w = U_1 \Lambda U_1^T,$$
$$M = UDU^T = U_1 \Sigma_1 U_1^T + U_2 \Sigma_2 U_2^T. \quad (34)$$

The columns of $U = [U_1, U_2]$ are its EiVecs, Σ is diagonal with the EiVs $\sigma_1 \geq \sigma_2 .... \geq \sigma_n \geq 0$, $U_1 \in R^{\nu \times n}, \nu \leq n$.

The ocean model MICOM used in the experiment has 4 layers and the observations are the sea surface height (SSH). The ocean state has the dimension $n = 302400$. For the more detail, see [10]. The ECM Q_w is assumed to be of the form of Kronecker product of the vertical ECM M_v and the horizontal ECM M_h [8]. For simplicity, the model error ECM is accounted only in the vertical ECM hence

$$M_v = M_v^m + Q_v, \quad (35)$$

where $M_v^m \in R^{n_v \times n_v}$ is the ECM estimated from an ensemble of PE samples generated by the PeSP [6], Q_v represents the covariance of the associated model error. The structure of Q_v is supposed to be of the form

$$Q_v = U_v(1) D \Sigma_v(1) U_v^T(1), M_v^m = U_v \Sigma_v U_v^T,$$
$$U_v = [U_v(1), U_v(2)],$$
$$\Sigma_v = \text{block diag } [\Sigma_v(1), \Sigma_v(2)], \quad (36)$$

Computation reveals that the EiVs decomposition of M_v^m has the first mode with the explained variance 67%, the second and third modes - with the corresponding explained variances 17% and 15%. The explained variance of the fourth mode is only 0.7E-07 %.

In the experiment $U_v(1)$ consists of the three first eigenmodes, $D = 0.2I$. The corresponding AF is denoted ad AF3U.

Fig. 5 shows rms of filtered error for the u-component of surface velocity resulting from the AF0U - the AF whose non-adaptive version has the gain computed on the basis of an ensemble of PE samples (PeSP) hence $Q_v = 0$. It is seen that by introducing a simple hypothesis on the subspace for the model error ECM, it is possible to improve the AF performance compared to the case of no specification of the model error ECM.

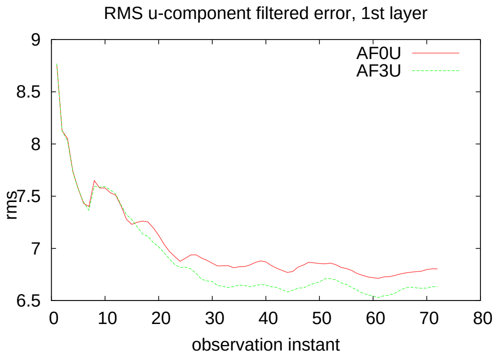

Figure 5: Improvement of the performance of the AF where the model error ECM has been taken into account (AF3U) compared to AF0U (without the use of model error ECM).

7 Conclusions

We have presented in this paper the first results on an AF approach to deal with the ME (deterministic and stochastic) present in the dynamics of a HdS. The advantage of the traditional BSE (or two-stage estimation) in separating the estimation of the bias from that of the dynamic state is that it allows to reduce the size of the matrices and vectors involved in the filtering computations. However, due to matrix equations involved in estimation algorithms, the BSE is inappropriate for HdSs as such encountered in practice of DA-GeoS. The main objective in this study is to develop the algorithms without recourse to solve the matrix equations for gain computation. This requires to introduce a new optimality criterion for the filter and the choice of appropriate tuning parameters. The AF approach, developed initially in [11], is completely adaptable for this purpose.

In the numerical part, based on the postulated hypothesis, the structure of the ME has been chosen in accordance with the postulated hypothesis and the optimization of the filter performance has been carried out. The numerical results clearly show a relevance and significance of the introduced hypothesis, its usefulness in the optimization of the filter performance.

In the future, we plan to extend this approach to the class of time-varying MEs which belong to the class of random processes with separable correlation matrices as studied in [7]. That would be an important step forward towards the goal of achieving a solution of GeoS-DA for HdSs for a wide class of MEs.

8 Appendix

Introduce the two classical definitions

Definition A.1. A linear system (1) is *observable* if all its EiVecs are observable.

Definition A.2 A linear system (1) is *detectable* if all unstable EiVecs (modes) are observable.

Mention that the criteria for checking an observability of (1) is that its observability matrix $\emptyset$ has the rank n where

$$\emptyset := [H^T, (H\Phi)^T, ..., (H\Phi^n)^T]^T \qquad (37)$$

Checking the rank of the observability matrix provides an algebraic test for observability, however this method is generally less interesting compared to the Popov-Belevich-Hautus (PBH) tests. The more useful is the Popov-Belevich-Hautus EiVec and rank tests for observability.

Theorem A1. (PBH Eigenvector Tests for Observability) The state equation (1) specified by (Φ, H) is observable if and only if there exists no right EiVec of Φ orthogonal to the rows of H.

Proof of Theorem 4.1. For $b = 0$, a stability of the filter under observability of (1) is proved in [12] (in fact, it requires only a detectability of (1), but observability implies detectability). The filter (19) is stable if we can show that observability of (1) implies observability of (4)-(5).

Return the the structure (16) in Lemma 4.1. One sees that the right EiVecs of $\tilde{\Phi}$ are of the form $u_{g,i} = (u_i^T, 0)^T, i = 1, ..., n$ and $u_{g,n+i} = (u_i^T, [(1 - \lambda_1)u_i]^T)^T, i = 1, ..., n$. We have $H_g u_{g,i} = Hu_i \neq 0, i = 1, ..., n$ since the system (1) is observable which implies $Hu_i \neq 0$ by the PBH Test for Observability. Moreover, as $H_g u_{g,n+i} = Hu_i, i = 1, ..., n$ it implies $H_g u_{g,n+i} \neq 0, i = 1, ..., n$ too. **(End of Proof)**

References:

[1] D.P. Dee, On-line estimation of error covariance parameters for atmospheric data assimilation, *Mon. Weather Rev.*, 123, pp. 1995, 1128 - 1145.

[2] M. Dowd, Estimating parameters for a stochastic dynamic marine ecological system, *Environmetrics*, 22, 2011, pp. 501 - 515.

[3] B. Friedland, Treatment of bias in recursive filtering, *IEEE Trans. Automat. Contr.*, AC-14, 1969,pp. 359 - 367.

[4] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, New York: Academic, 1970.

[5] G.H. Golub and C.F. Van Loan, 1993. *Matrix Computations*. 2 edn. Johns Hopkins.

[6] H.S. Hoang and B. Baraille, Prediction error sampling procedure based on dominant Schur decomposition: Application to state estimation in high dimensional oceanic model, *J. Appl. Math. Comput.* 218, 2011, pp. 3689–3709.

[7] H.S. Hoang and B. Baraille, Dynamical Systems-based Approach for Simulation and Estimation of Random Processes with Separable Covariance Functions, In *Mathematical models and methods in modern science*, Ed. A. J. Viamonte, 2012, pp. 36-41.

[8] H.S. Hoang and B. Baraille, On the efficient low cost procedure for estimation of high-dimensional prediction error covariance matrices *Automatica* 83, 2017, pp. 317 - 330.

[9] H.S. Hoang and B. Baraille, A simple numerical method based simultaneous stochastic perturbation for estimation of high dimensional matrices, *Multidimensional Systems and Signal Processing*, 2019, Issue 1, pp 195 - 217.

[10] H.S. Hoang, B. Baraille and O. Talagrand, On an adaptive filter for altimetric data data assimilation and its application to a primitive equation model MICOM, *Tellus* 57A, 2005, pp. 153-170.

[11] H.S. Hoang, P. DeMey, B. Baraille and O. Talagrand, A new reduced-order adaptive filter for state estimation in high dimensional systems, *Automatica* 33, 1997, pp. 1475-1498.

[12] H.S. Hoang, O. Talagrand and R. Baraille, On the design of a stable adaptive filter for high dimensional systems. *Automatica*, 37, 2001, pp. 341-359.

[13] F.-X. Le Dimet and V. Shutyaev. On deterministic error analysis in variational data assimilation. *Nonlinear Processes in Geophysics*, **12**, (4), (2005), p. 481–490. doi:10.5194/npg-12-481- 2005.

[14] J. M. Mendel, Extension of Friedlands bias filtering technique to a class of nonlinear systems, *IEEE Trans. Automat. Contr.*, AC-21, 1976, pp. 296 - 298.

[15] J.C. Spall, Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation", *IEEE Trans. Autom. Contr.*, vol. 37(3), pp. 332-341, 1992.

[16] E. L. Shreve and W. R. Hedrick, Separating bias and state estimation in a recursive second-order filter, *IEEE Trans. Automat. Contr.*, AC-19, 1974, pp. 585 - 586.

[17] E. C. Tacker and C. C. Lee, Linear filtering in the presence of time-varying bias, *IEEE Trans. Automat. Contr.*, AC-17, 1972, pp. 828 - 829.

[18] D. H. Zhou, Y. X. Sun, Y. G. Xi, and Z. J. Zhang, Extension of Friedlands separate-bias estimation to randomly time-varying bias of nonlinear systems, *IEEE Trans. Automat. Contr.*, vol. 38, 1993, pp. 1270 - 1273.