

On an adaptive filter for altimetric data assimilation and its application to a primitive equation model, MICOM

By H. S. HOANG^{1*}, R. BARAILLE¹ and O. TALAGRAND², ¹*SHOM/CMO, 18 Avenue Edouard Belin, 31041 Toulouse Cédex, France;* ²*LMD, ENS, 24 Rue de Lhomond, 75231 Paris Cédex 05, France*

(Manuscript received 13 May 2003; in final form 20 August 2004)

ABSTRACT

In a reduced-order adaptive filtering approach, we demonstrate a possibility to overcome two major difficulties in estimating oceanic circulation: a very high dimension of the system state and uncertainties in specification of the model error statistics. This approach is based essentially on the assumption that a particular parametrized gain matrix has been selected and the tuning parameters are adjusted by minimizing the mean prediction error. In the present paper we apply a reduced-order adaptive filter for solving the problem of assimilating altimetric sea surface height into a primitive equation model: the Miami isopycnic coordinate ocean model (MICOM). A gain structure is described which is proven to be very efficient in the twin experiments. The assimilation algorithm to be employed in the identical twin experiment is a reduced-order filter whose reduced state consists of the layer thickness. The velocity update is calculated from the geostrophic hypothesis. The gain structure for the non-adaptive filter is obtained on the basis of three principal hypotheses: (H1) analysis error for the system output is cancelled in the case of noise-free observations (as is done naturally in a standard Kalman filter for noise-free observation); (H2) conservation of linear potential vorticity; (H3) no correction for the velocity at the bottom layer. The initial values of the parameters in the gain will be selected in such a way that the filter behaves exactly as the Cooper–Haines filter (CHF) at the first data update step. It is shown that the adaptive filter, which relaxes one or several of the above hypotheses, is capable of producing the better estimates for the ocean state (layer thickness and velocity) compared to that produced by the CHF in all layers, surface or subsurface. Numerical experiments demonstrate the excellent capacity of the adaptive filter to extract useful information from surface observations for inferring the oceanic circulation in the MICOM.

1. Introduction

Recently, there has been great interest in developing advanced assimilation methods for operational ocean monitoring and prediction systems, especially for assimilating sea surface height (SSH) into oceanic models. Theoretically, the minimum variance solution to the linear filtering problem, under the assumption on Gaussian noise processes (model and observation errors), can be obtained by the Kalman filter (KF; e.g. Jazwinski, 1970; Anderson and Moore, 1979). However, due to the non-linearity of the model equations, to the very high dimension of the ocean state (and on uncertainty in specification of the model error statistics), the KF is impossible to properly employ because it requires a linearization and a solution of the algebraic Riccati equation (ARE) for the forecast error covariance matrix (ECM; see Dee, 1991; Fukumori et al., 1993; Fukumori and Malanotte-Rizzoli,

1995). For a review of efforts to overcome these computational difficulties, see Todling and Cohn (1994). In fact, implementation of the Kalman-like filters for very high-dimensional systems requires inevitably introducing approximations to the solution of the ARE. In this regard it is of interest to cite Cohn and Todling (1996), who approximated the solution of the ARE using a leading part of the singular value decomposition (Golub and Van Loan 1993) of the linear tangent or a leading part of the eigendecomposition of the forecast (propagated analysis) ECM. Alternatively, Courtier (1993) proposed an approximation in retaining the leading modes of the analysis ECM. In the singular evolutive extended Kalman (SEEK) filter (see Pham et al. 1997) the propagated analysis ECM is approximated by a singular low-rank matrix and most often the reduced-rank subspace is initialized from an empirical orthogonal function (EOF) analysis of the free model variability. As for the ensemble Kalman filter (EnKF; see Evensen, 1994) the ECM is proposed to be calculated from an ensemble of model states obtained by integrating the numerical model. Thus, the rank of the approximated ECM is at most equal

*Corresponding author.
e-mail: hoang@thor.cst.cnes

to the number of simulated ocean states. It is seen that all the cited sequential approaches deal with the problem of reducing the dimension of the space representing a solution of the ARE.

Roughly speaking, at each assimilation step, the KF requires $(n \times p)$ elements of the gain matrix where n and p are the dimensions of the system state and the observation vector, respectively. From the computational point of view, a less expensive way is to apply the adjoint equation (AE) method in which the initial state (of dimension n) is chosen as a control vector, which reduces significantly the number of elements to be determined (3D-VAR, 4D-VAR; see Le Dimet and Talagrand, 1986; Talagrand and Courtier, 1987). Data assimilation then reduces to a problem of finding the best fit to the data, and for a noise-free linear system this algorithm is equivalent to the KF. The AE method, however, suffers from the assumption of a perfect numerical model. A possibility to take into account the model error in the variational method was recently elaborated in Bennett et al. (1997) and Amodei (1995). Note that, as in the KF approach, the AE has difficulty in the specification of the ECM of the background, which is equivalent to the forecast ECM in the KF. It should be emphasized that, due to the non-linearity of the model equations, both KF and AE methods are never truly optimal in the real case.

Alternatively, an adaptive filtering approach is proposed and developed in Hoang et al. (1997a; 1997b; 2000; 2001) to overcome the difficulties mentioned above. Two features distinguish this approach from others: (i) the optimality of the filter is understood in the minimum mean prediction error (MPE, see Ljung, 1987, page 171) sense (misfit between the observations and the forecast of the system output); (ii) the performance of the filter is optimized (in the MPE sense) by adjusting some parameters of the gain. In this way, the filter gain is a function of observations and is adaptive. As elements of the gain vary dependent on the values of the tuning parameters, the stability of the filter, even under the most favourable conditions, might not be ensured if we do not impose several constraints on the structure of the gain. It is well known (see Kailath 1980) that the stability of the filter is a most important issue in the design of any filter. This question is a subject of the study in Hoang et al. (2001; 2002). For example, in Hoang et al. (2001) for a time-invariant linear system, it is shown that under detectability of the input–output system, a stable filter can be obtained by projecting the innovation on to a subspace spanned at least by all unstable and neutral eigenmodes of the system dynamics (necessary and sufficient conditions for the existence of a stable filter). As for time-varying systems, it is demonstrated in Hoang et al. (2000) that one can construct a stable (L_2 norm) filter under s-detectability of the system (all the left and/or the right unstable and neutral singular vectors of the system dynamics are observable). By choosing a priori a particular stabilizing gain structure, no ARE that constitutes an essential computational burden in evaluating the KF is involved in the determination of the gain of the adaptive filter (AF). This major advantage allows us to implement the filter for

very high-dimensional systems. As there is no need to specify the statistics of the model and observation errors, all efforts are concentrated on enhancing filter performance by adjusting the tuning parameters in the gain to minimize the MPE. Flexibility of the parametrization of the gain allows us to drastically reduce the number of tuning parameters in the gain (or in other terms, control variables). However, the cost of the technique based on the computation of a set of leading singular modes/eigenmodes is probably still too high for its immediate application.

It is worth mentioning that historically the idea of selecting simplified gains to assimilate data into oceanic (and meteorological) numerical models is not new (see Ghil and Malanotte-Rizzoli, 1991 for a review of the development of assimilation methods, such as direct insertion, optimal interpolation (OI), etc., in meteorology and oceanography). The aim of our study is to demonstrate that it is possible to exploit the robust and simple techniques already existing in the field of data assimilation, to seek a structure of the gain and its parametrization which is inexpensive, efficient and easy to implement. With the SSH observations, for example, there exists in particular a statistical method (i.e. OI) in which the gain has a form similar to that of the Kalman gain, with the difference that the forecast ECM is prescribed or calculated approximately from the numerical models or from a sequence of historical observations such as correlations (e.g. De Mey and Robinson 1987; Mellor and Ezer 1991). Related to the Miami isopycnic coordinate ocean model (MICOM), Gavart and De Mey (1997) have presented a computation of isopycnal empirical modes for the North Atlantic in the MICOM. These modes allow us to represent in a simplified way the vertical variability of the ocean. These suboptimal schemes are proven to be efficient in many experiments (with generated or real data). However, all these schemes rely upon pre-derived correlation coefficients with the SSH for assimilation.

Another interesting approach to be taken into consideration as a starting point in the present paper was developed initially by Haines et al. (1993) and later by Cooper and Haines (1996), based on the conservation principle. They have shown that considerable success can be achieved when assimilating SSH data if the subsurface potential vorticity (PV) is conserved at assimilation instants. This technique has been tested with success in the twin experiment in Cooper and Haines (1996) for assimilating surface pressure data into a 21-level, eddy-resolving Cox model in a double-gyre configuration. This simple and efficient scheme, referred to as a Cooper–Haines filter (CHF), will be first formalized in this paper following a standard OI approach. Three principal hypotheses are formulated from which we obtain the equations serving after to seek a set of tuning parameters. The non-adaptive version (with appropriate guess values for tuning parameters) will coincide exactly with the CHF. As shown in this paper, the AF, which is designed to offer a tracking capacity (minimization of the prediction error), will improve significantly on the performance of the filter in producing both forecast and

analysis estimates of the ocean state, for all model variables and in all layers, in comparison with that produced by the CHF.

We will evaluate the performance of the proposed AF in the twin experiments for SSH data assimilation in the primitive equation model, MICOM, with four layers. It will become clear that the appropriate parametrization of the gain allows us to drastically reduce the number of tuning parameters in the gain. This approach for the AF is computationally realizable for solving data assimilation problems in complicated oceanic numerical models.

The paper is organized as follows. In Section 2, we outline a set of partial differential equations describing the oceanic circulation used in the twin experiment with the MICOM. The numerical model, MICOM, and the set of initial and boundary conditions for simulating the ‘true’ oceanic circulation, as well as that to be initialized in the AF, will also be given. In Section 3, we derive a non-adaptive version of the filter that is closely related to the OI scheme. The equations deduced from the constraints imposed by three hypotheses are obtained. In Section 4, we derive different sets of tuning parameters by relaxing one or several mentioned hypotheses. We show how the tuning parameters should be initialized to meet exactly the conditions required in the CHF. The numerical results of the twin experiment are presented in Section 5 for the set of noise-free observations, as well as the set of noisy observations. In Section 5.2, for the set of noise-free observations first, we simply integrate the MICOM to produce the ocean circulation, without assimilation. The results obtained by this simple integration are compared to those generated by the CHF. Here, we see the excellent capacity of the CHF to infer the subsurface circulation using the SSH observations. As for the AF, we apply here two adaptive versions, one of which relaxes only the hypothesis on no correction for the velocity at the bottom and the other preserves only the hypothesis that the filter would produce the analysis field for the system output identical to the observed SSH. A comparison between the performances of the AF and of the CHF is given also in this section. Assimilation with the set of noisy observations is presented briefly in Section 5.3. It can be seen that by adjusting tuning parameters, the AFs are capable of improving significantly on the forecast and analysis performances of the estimates for all variables of the model (layer thickness and velocity), in all layers and at all assimilation instants. Concluding remarks and future works are finally discussed in Section 6.

2. Primitive equation model: MICOM

The oceanic numerical model to be used in the twin assimilation experiments is a primitive equation model, MICOM (isopycnic), which has been developed at the University of Miami (Bleck, 1998). The MICOM has achieved great success in the USA as well as in Europe (e.g. Chassignet, 1992; Baraille and Filatoff, 1995). The model relies on one prognostic equation for each component of the horizontal velocity field and one equation for

mass conservation, per layer:

$$\begin{aligned} \frac{\partial u_i}{\partial t} + \frac{\partial}{\partial x} \frac{u_i^2 + v_i^2}{2} - (\phi_i + f)v_i + \frac{\partial P_i}{\rho_0 \partial x} &= \tau_x \delta_{i,1} \\ &+ \frac{1}{h_i} [\nabla \cdot (\mu h_i \nabla u_i)] \\ \frac{\partial v_i}{\partial t} + \frac{\partial}{\partial y} \frac{u_i^2 + v_i^2}{2} + (\phi_i + f)u_i + \frac{\partial P_i}{\rho_0 \partial y} &= \tau_y \delta_{i,1} \\ &+ \frac{1}{h_i} [\nabla \cdot (\mu h_i \nabla v_i)] \\ \frac{\partial h_i}{\partial t} + \frac{\partial(h_i u_i)}{\partial x} + \frac{\partial(h_i v_i)}{\partial y} &= 0. \end{aligned} \tag{1}$$

Here, index i refers to the i th layer, ρ_0 is an average in space and time of the density, $\delta_{i,j}$ is the Kronecker symbol, ∇ is the two-dimensional gradient operator, u and v are the two horizontal velocity components, h_i is the thickness of the i th layer, $\phi = \partial v / \partial x - \partial u / \partial y$ is the relative vorticity, f is the Coriolis parameter, τ_x and τ_y are the wind stress forcings, and μ is an eddy viscosity coefficient.

The pressure P_i is a linear combination of the layer depths (hydrostatic hypothesis). Denoting by N the number of layers, we have

$$P_i = g \left(\sum_{k=1}^{i-1} \rho_k h_k + \rho_i \sum_{k=i}^N h_k + \rho_i \eta_b \right) \tag{2}$$

where g is the gravitational acceleration, ρ_i is the density of the i th layer, η_b is the bottom slope, and $\sum_{i=1}^0 (\cdot) = 0$ by definition.

The MICOM does not filter the external barotropic gravity waves as in the rigid lid approximation (Bryan, 1969). It splits the original equations between a barotropic part integrated with a small time-step and a set of baroclinic equations numerically constrained by a Courant–Friedrichs–Lewy (CFL) criterion based on the internal gravity wave velocity.

Introducing the new variables

$$\begin{aligned} \bar{u} &= \frac{\sum_{k=1}^N \rho_k h_k u_k}{\sum_{k=1}^N \rho_k h_k} & \bar{v} &= \frac{\sum_{k=1}^N \rho_k h_k v_k}{\sum_{k=1}^N \rho_k h_k} \\ u'_i &= u_i - \bar{u} & v'_i &= v_i - \bar{v} & h_i &= (1 + \eta) h'_i \end{aligned} \tag{3}$$

where η is the dimensionless quantity

$$\eta = \frac{\sum_{k=1}^N (\rho_k - \rho_0) h_k + \rho_0 \xi_1}{\rho_0 \sum_{k=1}^N h_k - \rho_0 \xi_1} \tag{4}$$

and ξ_1 is the SSH, we define the product $h'_b \eta$ by

$$h'_b \eta = \sum_{k=1}^N (\rho_k - \rho_0) h_k + \rho_0 \xi_1. \tag{5}$$

Finally, we obtain a new set of equations for the variables \bar{u} , \bar{v} , $h'_b \eta$, u'_i , v'_i and h'_i .

Approximating $(1 + \eta)$ by 1, we obtain the barotropic system

$$\begin{aligned} \frac{\partial \bar{u}}{\partial t} - f\bar{v} + g \frac{\partial h'_b \eta}{\partial x} &= \frac{\partial \bar{u}^*}{\partial t} \frac{\partial \bar{v}}{\partial t} + f\bar{u} + g \frac{\partial h'_b \eta}{\partial y} \\ &= \frac{\partial \bar{v}^*}{\partial t} \frac{\partial h'_b \eta}{\partial t} + \frac{\partial h'_b \bar{u}}{\partial x} + \frac{\partial h'_b \bar{v}}{\partial y} = 0 \end{aligned} \quad (6)$$

and the baroclinic equations

$$\begin{aligned} \frac{\partial u'_i}{\partial t} + \frac{\partial}{\partial x} \frac{u_i^2 + v_i^2}{2} - (\phi_i + f)v'_i - \phi_i \bar{v} + \frac{\partial P_i - gh'_b \eta}{\partial x} &= -\frac{\partial \bar{u}^*}{\partial t} \\ \frac{\partial v'_i}{\partial t} + \frac{\partial}{\partial y} \frac{u_i^2 + v_i^2}{2} + (\phi_i + f)u'_i + \phi_i \bar{u} + \frac{\partial P_i - gh'_b \eta}{\partial y} &= -\frac{\partial \bar{v}^*}{\partial t} \\ \frac{\partial h'_i}{\partial t} + \frac{\partial h'_i u_i}{\partial x} + \frac{\partial h'_i v_i}{\partial y} &= \frac{h'_i}{h'_b} \left(\frac{\partial h'_b \bar{u}}{\partial x} + \frac{\partial h'_b \bar{v}}{\partial y} \right). \end{aligned} \quad (7)$$

The term $\partial \bar{u}^* / \partial t$ ensures that $\bar{u}'_i = 0$. Note that this term is never calculated in the MICOM. For the definitions of \bar{u}^* and \bar{v}^* , see Hoang et al. (1997b). After some basic algebra, it can be shown that the sea surface elevation ξ_1 is related to the prognostic variables of system (7) by (a detailed derivation is given in Baraille and Filatoff, 1995)

$$\xi_1 = h'_b \eta / \rho_0 - \left(1 + \frac{h'_b \eta}{h'_b} \right) \sum_{k=1}^N \left(\frac{\rho_k}{\rho_0} - 1 \right) h'_k. \quad (8)$$

Relation (8) involves the depth-independent quantity $h'_b \eta$ and all the layer thicknesses h'_k .

3. Description of the adaptive filter

3.1. Adaptive filter formalism

It is seen from the description of the MICOM in the previous section that the ocean state $X(t)$ at each time instant t is composed of three variables distributed in space: the layer thickness, h , and two velocity components, u and v . In the discretized form, we have $[h(x_i, y_j, z_k; t)]_{ijk}$, $[u(x_i, y_j, z_k; t)]_{ijk}$ and $[v(x_i, y_j, z_k; t)]_{ijk}$ where $i \in II$, $j \in JJ$ and $k \in KK$. We assume that a set of observations modelled by sea surface elevation ξ_1 is available every 10 d at all horizontal grid points. As usual, the purpose of the assimilation experiment is to estimate as precisely as possible the system state $X(t)$ using these altimeter observations. We will implement here two algorithms: one is the well-known CHF and the other is an AF, which is identical to the CHF if no adaptation procedure is applied.

To be more precise, let us describe here a structure of the filter. The filter, in fact, is of a reduced-order type whose reduced state is composed only of layer thickness h . The velocity correction $(\delta u, \delta v)$ is computed from the geostrophy hypothesis (GH) on the basis of the layer thickness correction δh . Symbolically, the

filter is of the form

$$\begin{aligned} h_a(t + 1) &= h_f(t + 1) + \delta \hat{h}(t + 1), \\ \delta \hat{h}(t + 1) &= K^h \zeta(t + 1), \zeta(t + 1) \\ &= z(t + 1) - z_f(t + 1), \\ [u_a(t + 1), v_a(t + 1)] &= [u_f(t + 1), v_f(t + 1)] \\ &\quad + GH[\delta \hat{h}(t + 1)]. \end{aligned} \quad (9)$$

$z(t + 1)$ is the SSH observed at the time instant $(t + 1)$. The operator GH symbolizes the operation of computing $(\delta u, \delta v)$ from δh on the basis of the GH, and $\zeta(t + 1)$ is an innovation vector. The subscripts ‘a’ and ‘f’ denote the analysis and forecast estimate for corresponding variables.

As seen from structure (9), the most important and difficult task in the design of a filter is concerned with determining the gain K^h . In what follows, let the horizontal grid points (x_i, y_j) , $i = 1, \dots, II$; $j = 1, \dots, JJ$ of the mesh be ordered in some way. For a function $f(x_i, y_j)$ defined in the horizontal mesh, the notation f also means its vector representation (if it does not lead to misunderstanding) whose components are the values of $f(x_i, y_j)$ ranged in the order mentioned above.

The AF to be applied in this paper takes the CHF as a departure point. As the CHF is constructed on the basis of hypotheses such as conservation of the linear PV (for more details, see Section 3.2), it will be clearer to begin the description of the filter in the space of vertical displacement interface (VDI) variables. Let $\xi := (\xi_1, \dots, \xi_4)$ (in the experiments with the MICOM we have $KK = 4$) be a vector composed from VDI variables in all layers (see Fig. 1 for schematic representation of VDI variables). We have then the following relationships

$$h_k = \bar{h}_k + \xi_k - \xi_{k+1}, \quad k = 1, \dots, 4 \quad (10)$$

where \bar{h}_k is the (temporal) average of h_k ($\sum_k \bar{h}_k$ is the mean of the sea level). In practice, we can take \bar{h}_k as a climatology. This definition gives the observed SSH $z = \xi_1$ and for the flat bottom

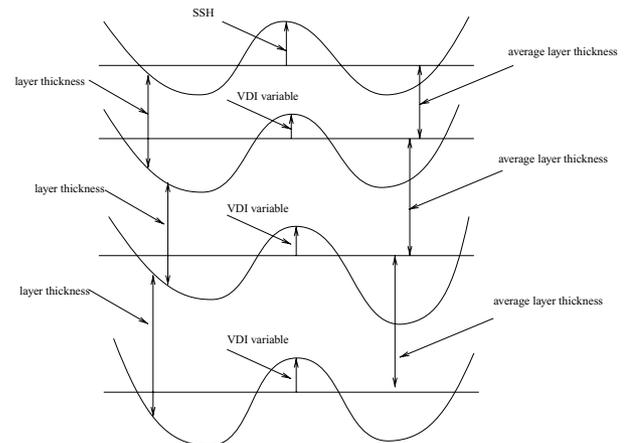


Fig 1. Schematic representation of VDI variables.

(which is the case in the experiment), $\xi_3 = 0$. Hence, we can write

$$\begin{aligned} \tilde{h} &= T\xi, & \tilde{h} &= h - \bar{h}, & h &= (h_1, \dots, h_4)^T, \\ T &= \begin{bmatrix} I_N & -I_N & 0 & 0 \\ 0 & I_N & -I_N & 0 \\ 0 & 0 & I_N & -I_N \\ 0 & 0 & 0 & I_N \end{bmatrix}. \end{aligned} \quad (11)$$

I_N is a unit matrix of dimension $N = II \times JJ$ – the total number of horizontal grid points.

Because $\sum_k \tilde{h}_k = \xi_1 = z$ we have

$$z = H_\xi \xi, \quad H_\xi = [I_N, 0, 0, 0] \quad (12)$$

where H_ξ denotes the observation operator in the space of VDI variables. Formally, the filter for VDI variables takes the form

$$\xi_a = \xi_f + K^\xi \zeta, \quad \zeta = z - H_\xi \xi_f \quad (13)$$

where ξ_a and ξ_f are the analysis and forecast estimates for ξ , and ζ is an innovation vector. Let B_ξ be the ECM for ξ_f . Then the optimal gain is given by $K^\xi = B_\xi H_\xi^T K_\zeta^+$ where $K_\zeta = H_\xi B_\xi H_\xi^T + R$ is the ECM for the innovation vector, and K_ζ^+ is the pseudo-inversion of K_ζ (Albert, 1972). We have then

$$K^\xi = \begin{bmatrix} K^\xi(1) \\ K^\xi(2) \\ K^\xi(3) \\ K^\xi(4) \end{bmatrix}, \quad K^\xi(k) = B_\xi(k, 1)[B_\xi(1, 1) + R]^+ \quad (14)$$

where $B_\xi(k, l)$ is the (k, l) block of B_ξ .

The filter written for displacement variable ξ now takes the form

$$\xi_a = \xi_f + \delta\hat{\xi}, \quad \delta\hat{\xi} = K^\xi \zeta. \quad (15)$$

In the space of layer thickness variables h , we have

$$T\xi_a = T\xi_f + T\delta\hat{\xi} \quad \text{or} \quad \tilde{h}_a = \tilde{h}_f + K^h \zeta,$$

or

$$h_a = h_f + K^h \zeta, \quad K^h = TK^\xi, \quad K^h = \begin{bmatrix} K^h(1) \\ \dots \\ K^h(4) \end{bmatrix}, \quad (16)$$

$$K^h(k) = K^\xi(k) - K^\xi(k+1), \quad k = 1, 2, 3; \quad K^h(4) = K^\xi(4).$$

3.2. Simplifications

To see how the gain of the CHF is derived, let us introduce some simplifications. First, assume that R and $B_\xi(k, l)$ are spatially homogeneous, $R = \sigma_r^2 I_N$, $\sigma_r^2 > 0$, $B_\xi(k, l) = b_{k,l} I_N$. Then

we have

$$K^\xi = \begin{bmatrix} \gamma_1^\xi I_N \\ \gamma_2^\xi I_N \\ \gamma_3^\xi I_N \\ \gamma_4^\xi I_N \end{bmatrix}, \quad \gamma_k^\xi = \frac{b_{k1}}{b_{11} + \sigma_r^2}, \quad k = 1, \dots, 4 \quad (17)$$

$$K^h(k) = \gamma_k^h I_N, \quad \gamma_k^h = \gamma_k^\xi - \gamma_{k+1}^\xi, \quad 1 \leq k \leq 3,$$

$$\gamma_4^h = \gamma_4^\xi. \quad (18)$$

3.2.1. Geostrophy and no motion hypothesis at bottom. As seen in the previous section, the filter for the reduced state allows us to project the SSH anomaly to recover the subsurface correction for layer thickness. Having obtained these corrections, the current updates are derived from the GH as follows. For a given h_i , $i = 1, \dots, 4$,

$$u_k = \sum_{i=1}^4 \beta_{k,i} \partial h_i / \partial x, \quad v_k = - \sum_{i=1}^4 \beta_{k,i} \partial h_i / \partial y, \quad (19)$$

$$\beta_{k,i} = g\rho_i / f_0\rho_k, \quad i \leq k-1; \quad \beta_{k,i} = g/f_0, \quad i \geq k-1$$

where f_0 is the Coriolis parameter and $g = 980.6 \text{ cm s}^{-2}$ is the gravity constant. If we assume in addition that the pressure at the bottom is not altered (level of no motion), then the velocity update at the bottom satisfies $\delta u_4 = \delta v_4 = 0$. We have

$$\delta u_4 = \sum_{k=1}^4 \beta_{4,k} \partial(\delta h_k) / \partial x$$

and for $\delta h_k = \delta\hat{h}_k = K^h(k)\zeta$ it follows from $\delta u_4 = 0$ that

$$\delta u_4 = \sum_{k=1}^4 \beta_{4,k} \gamma_k^h (\partial\zeta / \partial x) = (\partial\zeta / \partial x) \sum_{k=1}^4 \beta_{4,k} \gamma_k^h = 0$$

because $(\partial\zeta / \partial x)$ does not depend on k . As this relation must be satisfied for any function $(\partial\zeta / \partial x)$ [and for any function $(\partial\zeta / \partial y)$] this constraint leads to the equation

$$\sum_{k=1}^4 \beta_{4,k} \gamma_k^h = 0. \quad (20)$$

3.2.2. Cooper–Haines filter. Let us first return to eqs. (17) and (18). Following the principle of water mass rearrangement, if we choose the gain so that all water parcels in a particular water column are displaced vertically by the same amount, then the linear PV is conserved (Cooper and Haines 1996). This technique corresponds to the choice of subsurface parameters

$$\gamma_2^\xi = \gamma_3^\xi = \gamma_4^\xi = \gamma = \text{const.} \quad (21)$$

Substituting eqs. (21) and (18) into eq. (20) yields the equation $[\gamma(\beta_{4,4} - \beta_{4,1}) + \gamma_1^\xi \beta_{4,1}] = 0$,

from which the value of the parameter γ , which ensures that the bottom pressure is not altered, is

$$\gamma = -\frac{\beta_{4,1}\gamma_1^\xi}{\beta_{4,4} - \beta_{4,1}}. \quad (22)$$

For noise-free observations, $\sigma_r^2 = 0$, hence $\gamma_1^\xi = 1$, the computation yields

$$\gamma_{\text{ch}} = -\frac{\beta_{4,1}}{\beta_{4,4} - \beta_{4,1}} = -184.964\,966 \quad (23)$$

for the ocean configuration in the MICOM to be implemented in the twin experiment. The CHF takes thus a simple form

$$h_a = h_f + \delta h, \quad \delta h = K_{\text{ch}}^h \zeta, \quad \zeta = z - z_f,$$

$$K_{\text{ch}}^h(1) = (1 - \gamma_{\text{ch}})I_N, \quad K_{\text{ch}}^h(2) = 0, \quad K_{\text{ch}}^h(3) = 0,$$

$$K_{\text{ch}}^h(4) = \gamma_{\text{ch}}I_N, \quad \gamma_{\text{ch}} = -184.964\,966,$$

$$(u_a, v_a) = (u_f, v_f) + QH(\delta h). \quad (24)$$

This CHF will be applied in the following to assimilate the SSH data in the MICOM, which serves as a reference to be compared with the results generated by the AF.

Comment 3.1. For a practical application of the CHF (24), it is necessary to introduce the constraints of the type of threshold for displacement to ensure a non-negativity of layer thickness (see Gavart, 2001). In the twin experiments presented in this paper, the positiveness of layer thickness is checked during the assimilation process and never this condition is violated.

Comment 3.2. In practice, the realistic ocean models have up to 30–40 layers and, as the bottom layers are the most stable layers, only several first layers are updated by assimilation. For the DYNAMO configuration (North Atlantic) of the MICOM (see Gavart, 2001), for example, the displacement modes are updated only at the first 15 layers (up to about 3000–4000 m). The thickness of the sixteenth layer is of the order of 2000 m and there are practically no changes in the velocity calculated by geostrophy from this layer.

Subject to the assumption $\sigma_r^2 = 0$, the CHF provides an analysis SSH, which is the same as the observed SSH. Really due to definition (12), $H_h = TH_\xi = [I_N, \dots, I_N]$, hence

$$H_h(h_a - \bar{h}) = H_h(h_f - \bar{h}) + H_h K_{\text{ch}}^h \zeta,$$

$$H_h K_{\text{ch}}^h \zeta = \sum_{k=1}^4 K_{\text{ch}}^h(k) \zeta = \gamma_1^\xi I_N \zeta = \zeta$$

for $\gamma_1^\xi = 1$. Thus, $\xi_{1,a} = \xi_{1,f} + \zeta = z$.

In summary, the gain in the CHF (24) is obtained on the basis of three hypotheses, as follows:

H1 – the filter produces an SSH analysis estimate, which is the same as the observed SSH;

H2 – conservation of linear PV;

H3 – no motion at the bottom level.

4. Parametrization of the gain in the adaptive filter

In this section we concentrate our attention on the derivation of the different possible parametrized gain structures needed for the adjustment procedure in the AF. As the gain in the CHF is constrained by three hypotheses (H1, H2 and H3), the main idea to be followed here is to examine the possibility of removing one (or two, or all) of these constraints to release free parameters to be adjusted. To attain this goal, we reformulate the constraints (H1, H2 and H3) in the form of equations from which we can identify the set of free parameters appropriate for adaptation.

4.1. Relaxing hypothesis H1

First, let us begin by examining hypothesis H1. We denote by h_k^* the true layer thickness at the k th layer.

As seen from eq. (10) $\sum_{k=1}^4 (h_k^* - \bar{h}_k) = \sum_{k=1}^4 \tilde{h}_k^* = \xi_1^*$ where ξ_1^* denotes a true SSH. In general, the observed SSH $z = \xi_1^o$ is different from ξ_1^* (in the noisy observation case) and $\xi_1^o = \xi_1^* + v_\xi$. For $\xi_{f,1}$ (SSH forecast), we have $\xi_{f,1} = \sum_{k=1}^4 \tilde{h}_{f,k}$. Substituting from both sides of the filter eq. (9), $h_{a,k} = h_{f,k} + \delta h_k$, and the value \bar{h}_k yields $\tilde{h}_{a,k} = \tilde{h}_{f,k} + \delta h_k$. For $\delta h_k = K^h(k)\zeta$, $K^h(k) = \gamma_k^h I_N$ we have

$$\begin{aligned} \xi_{a,1} &= \sum_{k=1}^4 \tilde{h}_{a,k} = \sum_{k=1}^4 (\tilde{h}_{f,k} + \delta h_k) \\ &= \sum_{k=1}^4 \tilde{h}_{f,k} + \zeta \sum_{k=1}^4 \gamma_k^h = \xi_{f,1} + \zeta \sum_{k=1}^4 \gamma_k^h. \end{aligned}$$

Hypothesis H1 requires $\xi_{a,1} = \xi_1^o$ and, from the last equation, $\xi_1^o - \xi_{f,1} = \zeta = \zeta \sum_{k=1}^4 \gamma_k^h$ from which follows

$$\sum_{k=1}^{k=4} \gamma_k^h = 1 \quad \text{or} \quad \gamma_1^h = 1 - \sum_{k=2}^{k=4} \gamma_k^h. \quad (25)$$

Constraint (25) thus ensures that the SSH analysis errors are equal to zero, i.e. the SSH analysis is the same as the observed SSH, independent of whether the observations are noisy or not. Evidently this constraint is reasonable only if observations are noise-free. For noisy observations, a weaker constraint should be taken into consideration. As seen from eq. (22) we have $K^h = \gamma_1^\xi K_{\text{ch}}^h$ and we can choose $\lambda = \gamma_1^\xi$ as a tuning parameter. Let us look at $\lambda = (b_{11}/b_{11} + \sigma_r^2)$. As $\sigma_r^2 \geq 0$, the value of λ lies in the interval $\lambda \in [0, 1]$. For satellite SSH observations, with small positive σ_r^2 the value of λ must be close to 1 (if b_{11} is relatively large in comparison with σ , i.e. usually the case in the beginning of the assimilation period). The difficulty we have here is that we do not know about b_{11} . However, as the value $\sigma_{\zeta(t)} = b_{11}(t) + \sigma_r^2$ is estimated well by the filter (variance of innovation vector), and for satellite SSH observations we are given with high precision

σ_r^2 , the parameter λ can be estimated adaptively by the formula

$$\lambda(t) = \frac{[\sigma_{\xi(t)}^2 - \sigma_r^2]}{\sigma_{\xi(t)}^2}. \quad (26)$$

Equation (26) will be used in the experiment with noisy observations in Section 5.3.

4.2. Relaxing hypothesis H2

As seen from Section 3.2.2, hypothesis H2 is the most important in the construction of CHF, which is based on the principle of water mass rearrangement. This simple rearrangement allows us to greatly simplify the application in real applications and it has proven to be efficient in many works concerned with SSH data assimilation. On the other hand, as shown in Gavart and De Mey (1997), the computation of isopycnal empirical modes in the Azores current region shows that there is a slight departure of the dominant empirical mode from that of the CHF. In fact, it is argued that even if the PV is not uniform, it is still a very useful constraint to apply during altimetric assimilation (Haines et al., 1993).

It should be mentioned that, in practice, due to the lack of space–time measurements on the ocean state, we are never able to extract exactly the vertical dependence between SSH variation and that of subsurface interface displacements. It is therefore of interest to search for a possibility to identify, when the hypothesis H2 is no longer valid, the dependence between SSH variation and variations of interface displacements. We describe here one possibility to relax hypothesis H2. Let us return to eq. (20). We have

$$\beta_{4,1} \left(1 - \sum_{k=2}^{k=4} \gamma_k^h \right) + \sum_{k=2}^{k=4} \beta_{4,k} \gamma_k^h = 0 \quad \text{or} \quad (27)$$

$$\sum_{k=2}^{k=4} a_{k-1} \gamma_k^h = -\beta_{4,1}, \quad a_{k-1} = \beta_{4,k} - \beta_{4,1}.$$

To express this equation in the space of displacement variables γ_k^ξ , it is enough to remember the relationship (10) between γ_k^h and γ_k^ξ . From eqs. (10) and (27) after some algebraic manipulations we come to

$$c^T \gamma^\xi = -\beta_{4,1}, \quad c^T = (c_1, c_2, c_3),$$

$$c_1 = a_1, \quad c_2 = a_2 - a_1, \quad c_3 = a_3 - a_2 \quad \text{or}$$

$$c_k = \beta_{4,k+1} - \beta_{4,k}, \quad k = 1, 2, 3. \quad (28)$$

Because the number of unknowns in eq. (27) is three (for each horizontal grid point), there exists an infinite number of solutions and the set of all solutions of this equation can be represented in the form (Albert, 1972)

$$\gamma^\xi(y) = -c^{T,+} \beta_{4,1} + Ay, \quad A = [I - c^{T,+} c^T], \quad (29)$$

where $c^{T,+} = c/||c||^2$ is the Moore–Penrose pseudo-inverse of c^T , and y is any vector of dimension 3. The solution $\gamma_0^\xi = \gamma^\xi(y = 0)$ is known as a solution of minimal norm. Thus, for whatever

$y \in R^3$, the filter with the gain specified by $\gamma^\xi = \gamma^\xi(y)$ will produce the analysis layer thickness estimate satisfying H1 and H3. By varying the control variable y , we have the possibility to relax the hypothesis on conservation of the linear PV. Let y_{ch} be such that $-c^{T,+} \beta_{4,1} + Ay_{\text{ch}} = \gamma_{\text{ch}}^\xi$ (the element y_{ch} exists as seen in the next section). Then, representing $y = y_{\text{ch}} + \delta y$ gives $\gamma^\xi(y) = \gamma_{\text{ch}}^\xi + A\delta y = \gamma_{\text{ch}}^\xi + \delta\gamma^\xi$. Thus, correction to the Cooper–Haines displacement modes takes the form $\delta\gamma^\xi = A\delta y$, i.e. it belongs to the null space of c^T . The vector δy can now be considered as a control vector to be adjusted in the filter, which should be initialized as $\delta y = 0$. The adaptation is aimed at improving the performance of the CHF by altering only the hypothesis on conservation of the PV.

Comment 4.1. From Hoang et al. (2001) there is a constraint for δy . To see, let us look at the constraint for the elements of the matrix Λ for ensuring a stability of the filter. In fact, $K^\xi(k; \theta) = \theta_k \gamma_{\text{ch}}^\xi I_N$, and the constraint is $\theta_k \in [b, 2 - b]$, $b \in (0, 1)$. Writing $\gamma_k^\xi = \theta_k \gamma_{\text{ch}}^\xi = \gamma_{\text{ch}}^\xi + \delta\gamma_k^\xi$, $\delta\gamma_k^\xi := (1 - \theta_k) \gamma_{\text{ch}}^\xi$, we see that $\delta\gamma_k^\xi \in [1 - b, 2 - b - 1]$. Thus, the component of $A\delta y$ has the order of $\delta\gamma_k^\xi$. For example, for $b = 0.5$ we have $||a_k \delta y|| \in [-0.5|\gamma_{\text{ch}}^\xi|, 0.5|\gamma_{\text{ch}}^\xi|]$, where a_k is the k th row of A . These relations can serve as a constraint to check whether the estimate \hat{y} is ‘reasonable’. Of course, the more important index is the achievement of the MPE minimization process, which can be observed during the assimilation process.

4.3. Relaxing hypothesis H3

Consider a possibility to relax hypothesis H3 but conserving H1 and H2.

As shown above (and in Cooper and Haines, 1996), one simple way to conserve the PV is to displace all water parcels vertically by the same amount. This requirement, in terms of γ^ξ , is equivalent to saying that the displacement modes γ^ξ must belong to the linear subspace $R[1_3]$ generated by the vector $1_3 = (1, 1, 1)^T$. From a mathematical point of view, this physical constraint imposed for the members of eq. (29) is justified only if we can show that there exists an element of the subspace $R[1_3]$, which is also a member of the set of solutions (29). Let $\gamma^\xi \in R[1_3]$ have the form $\gamma^\xi = \gamma 1_3$ where γ is some constant. Substituting $\gamma^\xi = \gamma 1_3$ into eq. (28), it is seen then that this equation now is led to

$$\gamma c^T 1_3 = \gamma \sum_{k=2}^{k=4} (\beta_{4,k} - \beta_{4,k-1}) = -\beta_{4,1} \quad \text{or}$$

$$\gamma(\beta_{4,4} - \beta_{4,1}) = -\beta_{4,1}. \quad (30)$$

For $\gamma = -\beta_{4,1}/(\beta_{4,4} - \beta_{4,1})$ the vector γ^ξ is now a solution of eq. (28) and the requirement on displacing all water parcels vertically by the same amounts is acceptable. In fact, this is the solution we have already assigned to the gain of the CHF (see eq. 23). Let us write $\gamma = \theta \gamma_{\text{ch}}$. We see that to relax hypothesis H3 from the CHF, it is sufficient to allow θ to vary from

$\theta = 1$. Following the approach in Hoang et al. (1997b), the AF then seeks an optimal element θ^* in $R[1_3]$, which minimizes the prediction error

$$J[\theta] = E[\Psi(\zeta)] \rightarrow \min_{\theta}, \quad \Psi(\zeta) = \|\zeta\|^2. \quad (31)$$

As $J[\theta]$ is unknown (we do not know, for example, the distribution functions of the model and observation errors or their statistics), the optimization problem in eq. (31) can be solved only by a stochastic optimization (SA) technique (e.g. Tsypkin 1971). This technique will use a gradient of the sample objective function $\Psi(\zeta)$ to search optimal parameters. In the present case, the AF is very simple to implement because there is only one tuning parameter θ and the gradient can be easily estimated, for example, by two integrations of the direct model, without the need to have an adjoint code associated with the linear tangent of the numerical model. This filter (denoted AF1) will be applied in the twin experiment in the next section.

4.4. Other classes of adaptive filters

Each of the three classes of AFs presented in the previous sections corresponds to weakening one of the three hypotheses H1, H2 or H3. Evidently other possibilities can be considered by moving more than one hypothesis. For example, one interesting class of AF is to remove two hypotheses H2 and H3. In this situation only constraint (25) must be hold. We write eq. (25) in the form

$$a^T \gamma^h = 1, \quad a^T = (1, 1, 1, 1). \quad (32)$$

The set of all solutions of eq. (32) is expressed as

$$\gamma^h(y) = a^{T,+} + Ay, \quad \forall y, \quad A = [I - a^{T,+} a^T],$$

$$a^+ = a^T / 4. \quad (33)$$

Thus, whatever y is, the filter will produce the analysis layer thickness estimate satisfying H1.

Similar to what was done in Section 4.2, we first represent the vector $y = y_0 + \delta y$, where y_0 is a vector such that $a^{T,+} + Ay_0 = \gamma_{ch}^h$, $\gamma_{ch}^h = (1 - \alpha_{ch}, 0, 0, \alpha_{ch})^T$. By varying the control variable $\theta = \delta y$ we have a possibility to relax simultaneously two hypotheses H2 and H3. The correction to the Cooper–Haines gain matrix belongs to the null space of a^T . The vector δy should be initialized as $\theta = 0$. This filter (denoted AF2) will be used in the experiment with the MICOM in next section.

Finally, we want to stress that the choice of a particular structure for the AF is dictated by a specific problem we have at hand, but not by the intention to increase as much as possible the number of free parameters to be adjusted.

5. Twin experiments with the MICOM: numerical results and discussion

5.1. Parameters

In this section we report on the results of a set of experiments on altimetric data assimilation using the primitive equation model, MICOM. The model configuration used in the experiment is described in Chassignet (1992) and it is an idealized description of the Gulf Stream circulation on a beta plane. It is configured in a flat bottom rectangular oceanic basin ($1860 \times 2380 \times 5 \text{ km}^3$) driven by a zonal symmetric wind forcing: $\vec{\tau} = [-\tau_m \cos(2\pi y/L), 0]$ where $\tau_m = 10^{-4} \text{ m}^2 \text{ s}^{-2}$. The wind energy input is compensated by both lateral biharmonic viscosity ($A_4 = 8 \times 10^{10} \text{ m}^4 \text{ s}^{-1}$) and a linear bottom drag ($\mu = 2.65 \times 10^{-7} \text{ s}^{-1}$). Free-slip boundary conditions are imposed on each side of the basin. The mesh grid spacing is constant in both horizontal directions, and equal to around 13 km. In the vertical, the spacing (four layers) is non-uniform and the average layer thicknesses (from top to bottom) are 440, 608, 978 and 2974 m, respectively, with corresponding density values of 1000, 1001.70, 1002.91 and 1003.61 kg m^{-3} . The first baroclinic deformation radius is 44.7 km. The Coriolis parameter f linearly varies with latitude according to $f = f_0 + \beta y$ with $f_0 = 0.93 \times 10^{-4} \text{ s}^{-1}$ and $\beta = 2.0 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$ (40° mean north latitude).

The standard model configuration used here is a domain situated in the North Atlantic from 30°N to 60°N and 80°W to 44°W ; the exact model domain is shown in Fig. 2a. Grid spacing is about 0.2 in longitude and in latitude, requiring 25 200 (140×180) grid points \times four vertical modes. The model time-step is 18 min. The problem is to produce the estimate for the ocean state at each assimilation instant. Note that the system state consists of three variables (h, u, v), which are two-dimensional fields defined at four vertical modes. Thus, the dimension of the state is equal to 302 400.

The model, initialized at a state of rest, has been spun up for 10 yr. The resulting circulation essentially consists of two gyres (a counterclockwise one in the north and a clockwise one in the south) separated by a mid-latitude eastward jet whose amplitude decays from the western boundary into the interior of the domain. This jet forms unstable meanders and creates rings which propagate westward to the boundaries under the beta effect. For a complete study of ring formation process and ring statistics in this model, see Chassignet (1992).

Next, we integrate the model over 2 yr and time average the states produced during these 2 yr, giving a so-called climatology. The ‘true’ ocean state, used as a reference to be compared with estimates produced by different filters, is obtained by integration of the model from the climatology during the next 350 d. Every 10 d, the state, composed of (h, u, v), is stored and the SSH is calculated as a linear combination of displacements of four modes. Thus, the computed set of SSHs available at all surface grid points will be used as a set of observations,

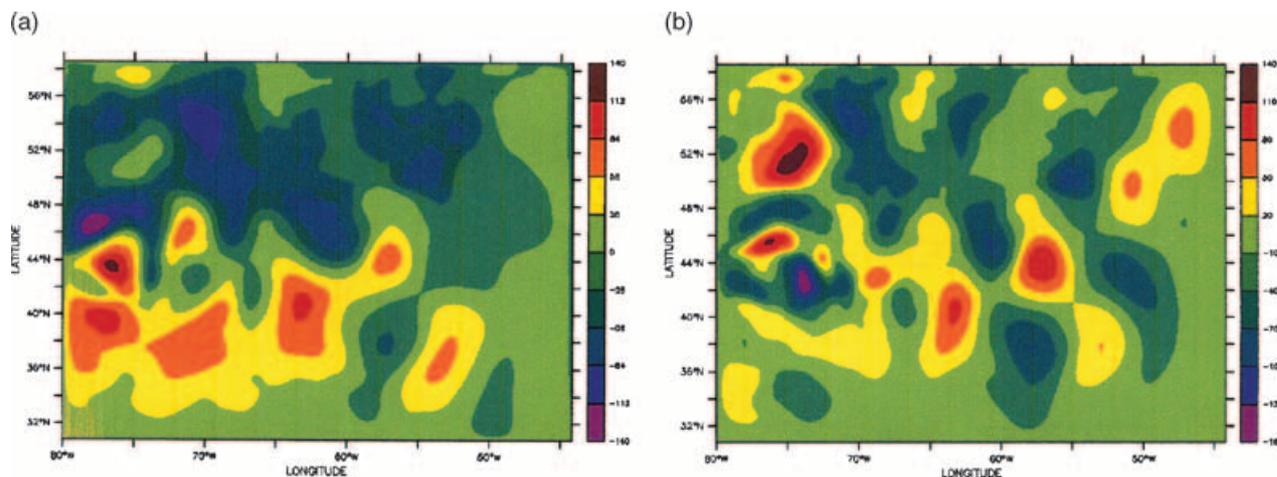


Fig 2. (a) SSH resulting from the true ocean state at May 31, 1993, which is used as an initial condition in the MODEL (no observations being assimilated), CHF and AF. (b) The difference between the SSH in (a) and that resulting from the true ocean state at January 1, 1993.

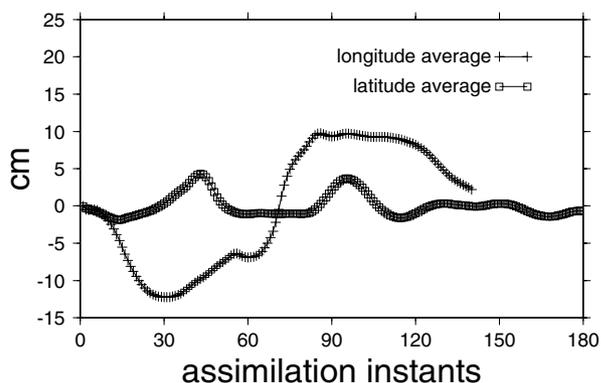


Fig 3. Longitude and latitude averages of the mean SSH obtained by averaging 35 SSH observations.

denoted SO(1), for assimilation experiments in Section 5.2. In total, we have a time sequence of 35 observations (350 d). In all experiments, the filters are initialized by the true state at day 150. The SSH calculated from this true state is shown in Fig. 2a. The difference between the initialized SSH and the true SSH at the beginning of the assimilation procedure is shown in Fig. 2b. In Fig. 2a we see a double-gyre structure with a positive SSH level to the south and a negative SSH area to the north. To examine whether such a structure is typical for the observed true SSH, let us first calculate the mean SSH from the sequence of observed SSHs (over 35 observations) and, afterwards, average the obtained SSH over longitude. The procedure of latitude average is also performed for this mean SSH. In Fig. 3 we display the two curves derived from these averaging procedures. One sees that the transition from a negative level to positive level occurs in the mid-latitude. The latitude average curve has a mildly neutral characteristic, with two positive picks at nearly 70°W and 60°W. Note that the mean layer thickness (computed from the sequence of true states) at the first and fourth layers has a ten-

dency to grow from the north to the south, whereas the inverse tendency is observed at the second and third layers. Computation of the mean velocity reveals that the u -velocity component is more homogeneous than the v -component (Figs. 4a and b).

In fact, the u -component is significant only along the western boundary (near the longitude 80°W). The v -component has a strip structure along the longitude, but only up to the mid-latitude. The strip width is about 280 km and the v -component has opposite signs in neighbourhood strips. The deeper the layer is, the smaller the velocity is. We have also calculated the rms distribution resulting from the sequence of true states (around the mean state). The layer thickness rms is most strong at the first layer (see Figs. 4c and d), especially in the area of (80°W, 70°W) and (40°N, 48°N). Layer thickness variability is nearly uniform at the second and third layers, but it remains still significant at the fourth layer. Perhaps this is the reason why the AF1 based on relaxing hypothesis H3 is capable of considerably improving its performance compared to that of the CHF (Section 5.2.2). For the velocity, for example, the rms is much bigger in the area (80°W, 70°W) and (40°N, 48°N), especially at the two first layers (and for the v -component). In the fourth layer, it becomes more or less uniform.

In general, the rms becomes smaller as the layer becomes deeper. Note that the rms of the v -component is higher than that of the u -component. All these statistics suggest that in comparison with the u -component, the v -component will be more difficult to estimate precisely.

To test the optimality of the tuning parameters found by the AF from assimilating the observations from SO(1), as well as to see how the AF works in the more realistic situation of noisy observations, we simulate two other sets of observations: SO(2) (noise-free observations) and NSO(2) (noisy observations). The set SO(2) of SSH observations is calculated from the model states obtained in the last 2 yr of the sequence of the 20-yr integration of the model from a state of rest. As for NSO(2), it is obtained from

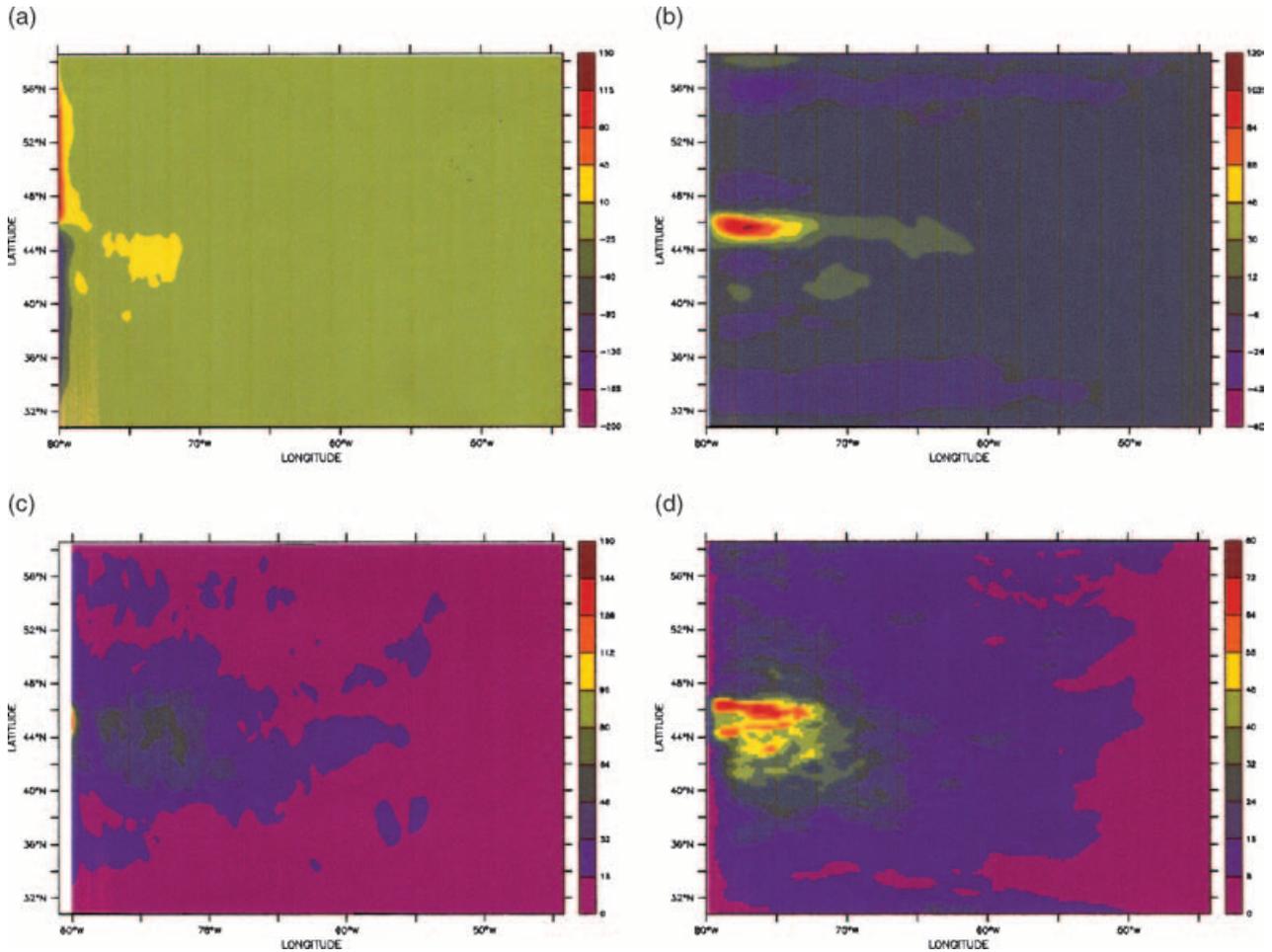


Fig 4. Mean velocity components u (a) and v (b) at the first layer. Note that the u -component is more homogeneous than the v -component. The corresponding rms (around the mean) is shown for (c) the u -component and (d) the v -component.

SO(2) by adding to the observations from SO(2) the uncorrelated (temporally and spatially) Gaussian noise with zero mean and variance 16 cm^2 . For more details, see Section 5.3.

5.2. Numerical results from assimilation of SO(1)

In this section we present the experiments based on the set SO(1) of noise-free observations.

5.2.1. Cooper–Haines and adaptive filters. Figure 2b shows the difference between the true SSH (the SSH resulting from the true initial system state) and that to be initialized in the filters (Fig. 2a). It is seen that a significant difference is observed in the domain 70°W – 80°W and 40°W – 54°W (the maximum values are about 160 and 130 cm). On average, the rms of this difference is about 32 cm.

First, we present in more detail the results of the AF1 described in Section 4.3, which relaxes hypothesis H3 on no motion at the bottom level. We have chosen this filter because it is very simple to implement (with one tuning parameter) and hence is

very promising for future applications. The performance of AF2 will be summarized only briefly in Table 3 to compare with that generated by the AF1.

Thus, the filter has the form

$$\begin{aligned}
 h_{a,k} &= h_{f,k} + K^h(k)\zeta, \\
 K^h(1) &= (1 - \theta\gamma_{\text{ch}})I_N, \quad K^h(2) = 0, \quad K^h(3) = 0, \\
 K^h(4) &= \theta\gamma_{\text{ch}}I_N, \quad \gamma_{\text{ch}} = -184.964966
 \end{aligned} \tag{34}$$

where θ is a scalar parameter to be adjusted to minimize the prediction error (see eq. 31). As the CHF corresponds to $\theta = 1$, in the AF the parameter θ is initialized at the beginning of the assimilation process as $\theta = \theta(0) = 1$. The value of $\theta = \theta(t)$ at the assimilation time instant t will change according to a multistage SA algorithm (See Polyak, 1990)

$$\begin{aligned}
 \theta'(t+1) &= \theta'(t) - v(t)\nabla_{\theta(t)}\{\Psi[\zeta(t+1)]\}, \\
 \theta(t) &= \frac{1}{t} \sum_{\tau=1}^{\tau=t} \theta'(\tau).
 \end{aligned} \tag{35}$$

Here, $v(t)$ is a parameter providing a convergence of the algorithm, $v(t) = c/(t^\alpha + p)$, where c and p are positive constant, $\alpha \in (0.5, 1)$, $\nabla_{\theta(t)}[\Psi]$ is the gradient vector of the sample objective function Ψ evaluated at the point $\theta(t)$, $\Psi(\zeta) = \|\zeta\|^2$, and $\|\zeta\|$ is the L_2 -norm. Note that algorithm (35) possesses a convergence rate equivalent to the second-order SA procedure (Polyak, 1990).

5.2.2. *Numerical results: Cooper–Haines filter and AF1.* In the following formulae, the bar over a character denotes the time average. To evaluate the performances of the filters, we introduce the following quantities to express estimation errors.

(1) For the objective function

$$e_\zeta(t) = \frac{1}{N_h} \sum_{i,j} \zeta(t, i, j)^2, \quad (36)$$

$N_h = IIXJJ$ is the number of horizontal grid points

$$\bar{e}_\zeta(t) = \frac{1}{t} \sum_{\tau=1}^t e_\zeta(\tau).$$

Thus, $e_\zeta(t)$ is the horizontal average of the innovation vector and $\bar{e}_\zeta(t)$ is its temporal average.

(2) For the layer thickness forecast error (for some k th component of h and for h),

$$\begin{aligned} e_f^{hk}(t) &= \frac{1}{N_h} \sum_{i,j} |e_f^{hk}(t; i, j)|, \\ e_f^{hk}(t; i, j) &= h_{r,k}(t; i, j) - h_k^*(t; i, j), \\ \bar{e}_f^{hk}(t) &= \frac{1}{t} \sum_{\tau=1}^t m_f^{hk}(\tau), \quad e_f^h(t) = \frac{1}{4} \sum_{k=1}^4 e_f^{hk}(t), \\ \bar{e}_f^h(t) &= \frac{1}{t} \sum_{\tau=1}^t e_f^h(\tau). \end{aligned} \quad (37)$$

Note that e_f^h and \bar{e}_f^h are vertical averages. Similar quantities are introduced for the analysis errors.

(3) Analogously for the velocity component u , we introduce

$$\begin{aligned} e_f^{uk}(t) &= \frac{1}{N_h} \sum_{i,j} [e_f^{uk}(t; i, j)]^2, \\ e_f^{uk}(t; i, j) &= u_{r,k}(t; i, j) - u_k^*(t; i, j), \\ \bar{e}_f^{uk}(t) &= \frac{1}{t} \sum_{\tau=1}^t e_f^{uk}(\tau), \quad e_f^u(t) = \frac{1}{4} \sum_{k=1}^4 e_f^{uk}(t), \\ \bar{e}_f^u(t) &= \frac{1}{t} \sum_{\tau=1}^t e_f^u(\tau). \end{aligned} \quad (38)$$

Similar quantities are defined for the analysis error for u . It is evident how to write the same error indices for forecast and analysis errors for the v component.

Thus, for example, $e_\zeta(t)$ represents a sample objective function normalized by the number of horizontal grid points (horizontally averaged), and $\bar{e}_\zeta(t)$ is the estimated objective function (the ensemble average is estimated by a time average over the interval $[1, t]$).

Before applying the filters, we first examine the model simulation without use of the observations (denoted MODEL). Fig. 5a shows the errors in reproducing the forecast SSH by the CHF and simply by running the MICOM (initialized by the same state as in the CHF). It is seen that the MODEL itself cannot reduce the errors for SSH (the errors remain nearly stable during the whole assimilation period). The CHF, as expected, improves significantly on what the MODEL alone could generate. In Fig. 5b, the objective functions resulting from the CHF and AF1 are plotted. It is seen that in the AF1 the objective function decreases quickly from the beginning up to day 170, and it is stabilized thereafter. The time series of the tuning coefficient in the AF1 is plotted in Fig. 5c. Let us look at estimation errors for the system state (h, u, v) . We remark from Table 1 that although the MODEL itself reduces the forecast estimation error for layer thickness, with rms from about 31.1 m at the beginning to 16.7 m at the end of the assimilation period (this happens since the initial condition used in the MODEL is too far from the true initial ocean state. For better initial estimates (climatology, for example), this forecast error may increase but it converges to the same error level at the end of the assimilation period as reported in Table 1), nothing is similar to the case of estimating the forecast SSH by MODEL as seen in Fig. 5a (and also of estimating the velocity (u, v) plotted in Fig. 6a).

The CHF, in turn, reduces the initial error for h to about 11.7 m, or with the gain of about 30% compared to the no-assimilation procedure. Finally, the AF1 yields at the end of the assimilation period the rms error which is equal to 9.8 m, or with the gain of 16% compared to the CHF (see Table 1) and 41% compared to the MODEL (not shown in Table 1). In Table 1 we present also the time-average rms errors for different variable estimates. For example, in time average the AF1 reduces 22% of the rms of the innovation vector compared to that of the CHF, 8.8% (rms) for the layer thickness estimate and 21% for the velocity estimates.

As seen here, the MODEL has difficulty reducing the error in estimating the velocity. This means that it is very important to make a correction to the layer thickness forecast to reduce the error for velocity estimate, because without assimilation the MODEL is incapable of reducing this error. Fig. 6b displays the rms for estimation errors for the analysis estimates obtained in the CHF and the AF1. In general, we see here that the differences between the CHF and the AF1 (in terms of estimation errors) become more and more significant as more observations are assimilated; this proves the optimality of the tuning coefficient found in the AF1.

In Table 2 we summarize the performances of the CHF and the AF1 in terms of analysis estimation errors at the end of the assimilation period. We see that, as in the case of forecast estimates, the AF1 produces better analysis estimates in comparison with those generated by the CHF. In time average, at the end of the assimilation period the performance gain obtained by the AF1 in comparison with the CHF is of the order of 9.2%, 19% and 16.6% for h , u and v estimates, respectively. Looking

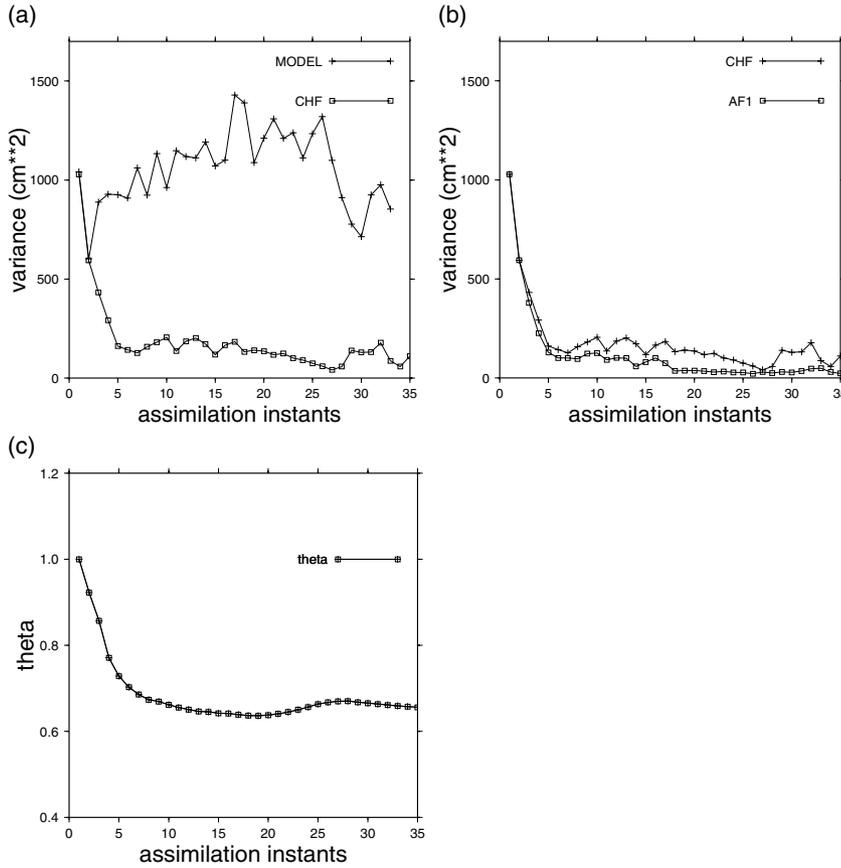


Fig 5. Time series of sample objective functions $m_{\zeta}(t)$ (in cm^2 ; see eq. 34) resulting from a simple model run (initialized by the same state as in the CHF and AF) and that of the CHF. We see here that the MODEL cannot reduce the error in forecasting the SSH (a). The CHF reduces considerably the forecast SSH error. It is curious to note that the forecast SSH error in CHF began to (slightly) increase after day 260 (the same picture is observed for the MODEL case). The same time series of sample objective functions, but for the CHF and AF1, are shown in (b). The tuning parameter θ , initialized as $\theta(0) = 1$, is adjusted in time to optimize the filter performance, and its evolution is shown in (c).

Table 1. Estimation errors (differences between forecast estimates and true ocean states) produced by the CHF and the AF1 in the experiment with SO(1). The AF1 was run subject to the initial value of $\theta = 1$, which corresponds to applying the CHF at the first assimilation instant. The MODEL, CHF and AF1 were initialized by the same state, which is the ‘true’ state at day 150. Column 5 presents a comparison between the CHF and AF1

	MODEL	CHF	AF1	Improvement by AF1
$\sqrt{e_{\zeta}^f(35)}$	29.2 (cm)	10.6 (cm)	5.1 (cm)	52%
$\sqrt{\bar{e}_{\zeta}^f(35)}$	32.5 (cm)	13.6 (cm)	10.6 (cm)	22%
$e_{\zeta}^a(35)$	16.7 (m)	11.7 (m)	9.8 (m)	16.1%
$\bar{e}_{\zeta}^a(35)$	21.5 (m)	16 (m)	14.5 (m)	8.8%
$\sqrt{e_{\zeta}^u(35)}$	18.7 (cm s^{-1})	8.52 (cm s^{-1})	5.5 (cm s^{-1})	35 %
$\sqrt{\bar{e}_{\zeta}^u(35)}$	19.4 (cm s^{-1})	9.8 (cm s^{-1})	7.5 (cm s^{-1})	21.2 %
$\sqrt{e_{\zeta}^v(35)}$	16.4 (cm s^{-1})	8.1 (cm s^{-1})	5.2 (cm s^{-1})	35.7%
$\sqrt{\bar{e}_{\zeta}^v(35)}$	17.9 (cm s^{-1})	9.2 (cm s^{-1})	7.2 (cm s^{-1})	21.2%

at the forecast and analysis rms errors for velocities (u, v) resulting from the AF1, it is seen that important performance improvements are observed at the beginning of the assimilation process, which correspond to the period when there are large corrections in the adjustment of the tuning parameter. This fact is also a consequence of the large difference between the initial-

ized state and the true system state, which pushes the filter to make a large correction for the velocity forecast. To see how the corrections to forecast made by the CHF and the AF1 can reduce forecast rms errors, we calculate averaged forecast and analysis rms errors for two velocity components (u, v) in two filters. It is found that, in time average, the gain matrices in the CHF and

Fig 6. The rms for forecast velocity (u, v) in the CHF and MODEL (a). For the velocity, the MODEL itself is incapable of reducing the forecast error, as it can do for the layer thickness case. We observe here a slight growth of this error for the CHF estimate after day 260. This growth of velocity corresponds to the increase of variance of SSH forecast error, as seen in Fig. 5a. The same rms for forecast velocity, but for CHF and AF1, is shown in (b).

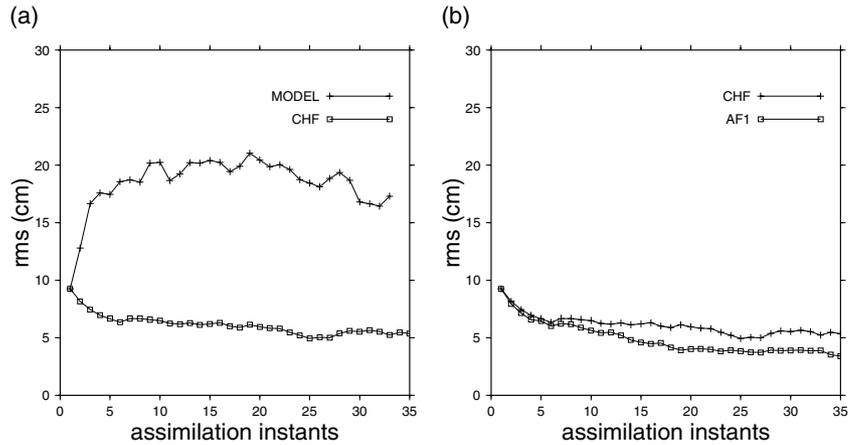


Table 2. The same as in Table 1, but for the analysis estimates

	CHF	AF1	Improvement by AF1
e_a^h	11.49 (m)	9.68 (m)	15.7%
\bar{e}_a^h	15.27 (m)	13.86 (m)	9.2%
$\sqrt{e_a^u(35)}$	5.27 (cm s ⁻¹)	3.57 (cm s ⁻¹)	32.3 %
$\sqrt{\bar{e}_a^u(35)}$	6.12 (cm s ⁻¹)	4.96 (cm s ⁻¹)	19 %
$\sqrt{e_a^v(35)}$	4.46 (cm s ⁻¹)	3.76 (cm s ⁻¹)	31.2%
$\sqrt{\bar{e}_a^v(35)}$	6.07 (cm s ⁻¹)	5.06 (cm s ⁻¹)	16.6%

Table 3. Forecast and analysis performance of the AF2 (at the end of the assimilation period)

Estimation error	Forecast	Analysis
$b_\zeta(35)$	15.76 (cm ²)	–
$\bar{b}_\zeta(35)$	109.4 (cm ²)	–
$b_{f,a}^h(35)$	9.17 (m)	9.2 (m)
$\bar{b}_{f,a}^h(35)$	14.49 (m)	13.86 (m)
$\sqrt{b_f^u(35)}$	4.54 (cm s ⁻¹)	3.12 (cm s ⁻¹)
$\sqrt{\bar{b}_{f,a}^u(35)}$	7.26 (cm s ⁻¹)	4.78 (cm s ⁻¹)
$\sqrt{b_{f,a}^v(35)}$	4.51 (cm s ⁻¹)	3.39 (cm s ⁻¹)
$\sqrt{\bar{b}_{f,a}^v(35)}$	7.08 (cm s ⁻¹)	4.88 (cm s ⁻¹)

the AF1 allow us to reduce 29% and 32% rms forecast errors, respectively.

5.2.3. *Numerical results: AF1 and AF2.* To examine whether it would be useful in relaxing simultaneously two hypotheses H2 and H3 to enhance a performance of the AF1, we apply here the AF2 described in Section 4.4. The assimilation results of this experiment are summarized in Table 3. Comparing the results in Table 3 with those in Tables 1 and 2 shows that the AF2 is

effectively able to generate better layer thickness and velocity estimates (for both forecast and analysis) than the AF1. The difference is more obvious for forecast estimates. At the end of the assimilation period, the value of the sample objective function produced by the AF2 is 15.8 cm², which is much lower than that of the AF1 (26.5 cm²). This fact shows that there is really an interest in relaxing simultaneously two hypotheses H2 and H3, although it will be (much) more expensive than the AF1 if the adjoint code is not available.

During the experiment we observe that the filter is very sensitive to departure from hypothesis H2, and hence a large departure from that hypothesis can lead to very poor state estimates or even to filter divergence (see also Comment 4.1). A careful choice of the gain parametrization (see eqs. 32 and 33, for example) as well as that of the amplitude of initial correction are important for guaranteeing a high filter performance. As the AF2 makes corrections to the layer thickness forecast at the second and third layers, it is interesting to examine what happens at these layers during assimilation. Figure 7a shows two curves representing the corrections to layer thickness analysis at the second and third layers, at the center of the domain, at the grid point (70, 90). Note that these corrections are equal to zero in the CHF. We found here large corrections at the beginning of the assimilation period. As seen in Fig. 7b, the gain components of the AF2 at the second and third layers are of opposite signs, positive for the third layer and negative for the second layer. The amplitudes of correction are larger at the third layer (about three to four times, after day 60) and converge more slowly compared to those observed in the second layer. To explain this, let us first look at the velocity errors introduced at the beginning of the assimilation period in these two layers. As seen in Table 4, the rms for velocity errors is bigger at the third layer compared to that at the second layer; they are 11 and 6.6 cm s⁻¹ for the u -component and 7.9 and 5.3 cm s⁻¹ for the v -component.

At the end of the assimilation period, the AF2 reduces these initial errors in the third and second layers to 3.4 and 2.5 cm s⁻¹, respectively, for the u -component, or a reduction of 69% of

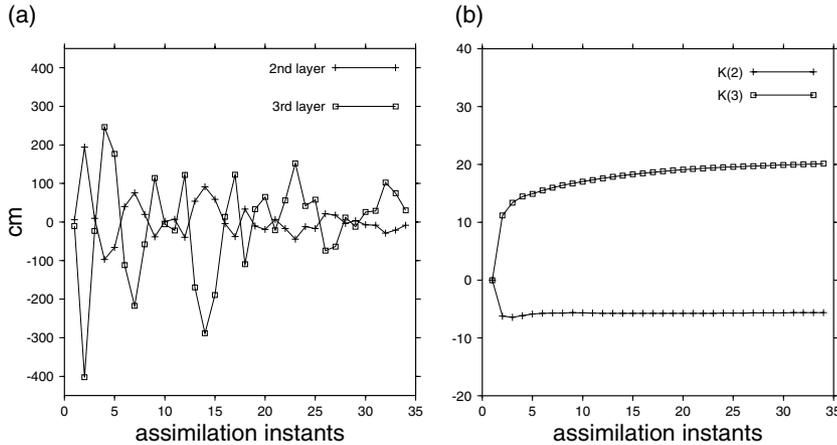


Fig 7. (a) The corrections, in cm, at the horizontal grid point (70,90), to layer thickness forecast estimates (at the second and third layers) made by AF2. Note that in the CHF these corrections are equal to zero. (b) The time series of gain components $K^h(2)$ and $K^h(3)$ at the second and third layers in the AF2.

Table 4. The rms for analysis velocity errors at the second and third layers generated by the AF2 and the CHF

Estimation error	$t = 1$ (cm s^{-1})	AF2 ($t = 35$) (cm s^{-1})	CHF ($t = 35$) (cm s^{-1})
$\sqrt{m_a^{u2}(t)}$	6.6	2.5	3.1
$\sqrt{m_a^{u3}(t)}$	11	3.4	5.3
$\sqrt{m_a^{v2}(t)}$	5.3	2.6	3.5
$\sqrt{m_a^{v3}(t)}$	7.9	3.7	5.7

the initial error in the third layer and 62% in the second layer. Similarly, for the v -component, we observe here a reduction of 53% and 51% of the initial errors in the third and second layers. Thus, by making larger corrections at the third layer, the AF2 is able to reduce more estimation errors at the third layer than in the second layer. As for the CHF, the rms for the u -component is equal to 5.3 and 3.1 cm s^{-1} at the third and second layers (and 5.7 and 3.5 cm s^{-1} for the v -component) at the end of the assimilation period. The reductions of the initial errors are hence 51% and 53% for the u -component, and 28% and 34% for the v -component, at the third and second layers, respectively. Thus, contrary to the AF2, the CHF reduces more velocity errors in the second layer than in the third layer. Note also that the amplitudes of corrections become smaller and smaller as the assimilation progresses. This does not mean that the gain components $K^h(2)$ and $K^h(3)$ in the AF2 tend to zero (see Fig. 7b), but it is a consequence of the fact that the forecast SSH error becomes smaller and smaller.

5.3. Noisy observation experiments

In this section we present briefly some results on the assimilation of noisy SSH observations. The model is first spun up during 20 yr from the state of rest, as done in Section 5.2. During the last 2 yr, years 19 and 20, the variables h , u and v are stored each 10 d, which will be used as the known ‘true’ ocean state

to evaluate filter performances. The set of 70 noise-free SSH observations, $SO(2)$, is first calculated from the sequence of true system states. The set of noisy observations, $NSO(2)$, is obtained by adding a Gaussian noise having zero mean and variance $\sigma_r^2 = 16 \text{ cm}^2$ (temporally and spatially uncorrelated) to the noise-free observations from $SO(2)$. The reason we have used a greater number of observations (in comparison with the experiments in Section 5.2) is that this enables us to check whether there really is a convergence of the adaptation algorithm for a long period of observations. In all the filters used in the experiments, the initial state is the mean of the sequence of true states.

5.3.1. *Experiments with the set of noise-free observation, $SO(2)$.* For the set of noise-free observations, $SO(2)$, the following filters are used: CHF, AF1 and CHF1. The CHF and the AF1 are of the forms described in Section 5.2. The filter CHF1 is of the form of the CHF with the difference that instead of $\gamma_{\text{ch}} = -184.964966$ (see eq. 23) now $\theta^* \gamma_{\text{ch}}$ where $\theta^* = 0.6567$ is the value of the parameter θ obtained by the AF1 in the experiment in Section 5.2 with $SO(1)$. Thus, the value θ^* has been obtained by the AF1 on the basis of the observations $SO(1)$, which is completely disconnected with the set of $SO(2)$. This experiment has the goal of checking whether the performance improvement of the AF1 over the CHF observed during the adaptation process, as done in the experiment in Section 5.2, has a transient character or not. These results will be used also to compare with the performances of the CHF and AF for the noisy observations case.

5.3.2. *Experiments with the set of noisy observations $NSO(2)$.* For the set of noisy observations, the following filters will be used.

(1) The CHF, which is the filter described in Section 4.1 for the case of noisy observations. Thus, we have the gain

$$K_{\text{rch}}^h = \lambda_{\text{ch}} K_{\text{ch}}^h. \quad (39)$$

We specify $\lambda_{\text{ch}} = 0.91$ (this is equivalent to assuming that the observation error is 10 times less than the SSH forecast error).

Fig 8. Data set SO(2). Time evolution of the ratios of rms of cost function and layer thickness (a), as well as rms of analysis errors for velocity estimates (b). We see here the superiority of the AF1 over the CHF. The smaller these ratios are, the better the accuracy of the AF estimates compared to those of the CHF.

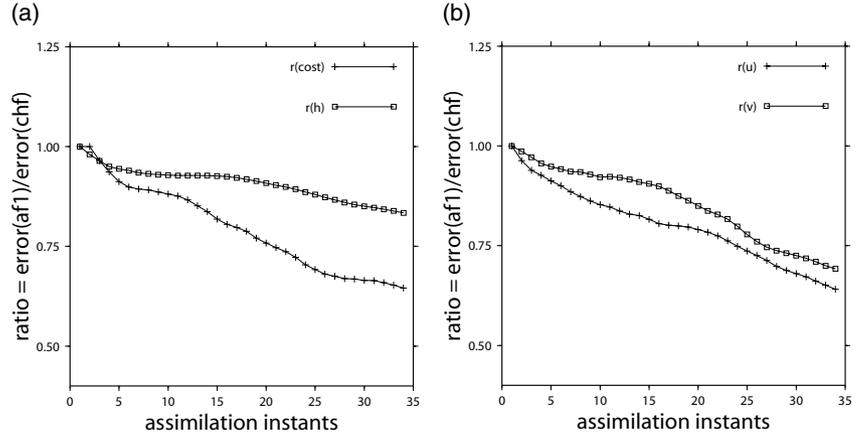
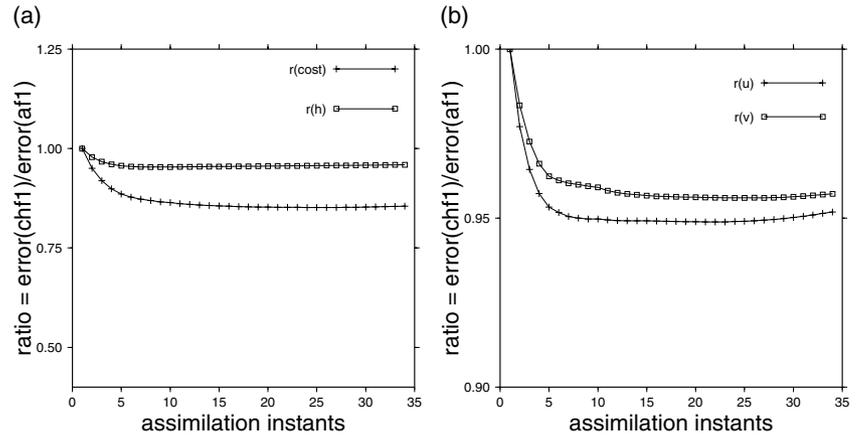


Fig 9. Comparison between the CHF1 and the AF1. (a) Ratio of rms for cost functions and layer thickness forecast errors. (b) The same ratios of rms but for forecast errors for velocity estimates.



(2) The AF. Let K_{af1}^h be the gain for AF1 (see eq. 34). Then the new AF has a gain

$$K_{af}^h = \lambda K_{af1}^h. \quad (40)$$

For this AF, two parameters are adjusted during assimilation process: λ and θ . The initial values for λ and θ are equal to 0.91 and 1, respectively (they correspond to the initialization of the AF as the CHF, eq. 39). The parameter λ is updated using eq. (26) in Section 4.1. The value $\sigma_{\zeta(t)}^2 = e_{\zeta(t)}$ is obtained by the filter and is defined by eq. (36). The value σ_r^2 is mis-specified and is equal to $\sigma_r^2 = 14 \text{ cm}^2$.

5.3.3. *Numerical results: noise-free observations.* First, we present the performances of AF1 and CHF for the set SO(2). We introduce for the cost function the ratio

$$r(\text{cost}) = \sqrt{\bar{e}_{\zeta(t)}(\text{af1})} / \sqrt{\bar{e}_{\zeta(t)}(\text{chf})}.$$

Similarly, for the forecast and analysis errors of the layer thickness h and velocity (u , v) estimates, we let

$$r(h) = \bar{e}_{h(t)}(\text{af1}) / \bar{e}_{h(t)}(\text{chf}),$$

$$r(u) = \sqrt{\bar{e}_{u(t)}(\text{af1})} / \sqrt{\bar{e}_{u(t)}(\text{chf})},$$

$$r(v) = \sqrt{\bar{e}_{v(t)}(\text{af1})} / \sqrt{\bar{e}_{v(t)}(\text{chf})}$$

(the subscripts ‘f’ or ‘a’ signifying ‘forecast’ or ‘analysis’ are omitted). The curves in Figs. 8a and b calculated for the analysis errors confirm, as in Section 5.2, the superiority of the AF1 over the CHF. The improvement is more significant for the velocity estimates shown in Fig. 8b. At the end of the assimilation period, in time average, in comparison with the CHF, the AF1 has reduced about 25% and 45% analysis errors for layer thickness and velocity, respectively. As for the forecast errors, these reductions are 25% and 35%. These results are much better than those observed in the experiment with SO(1) (see Tables 1 and 2), where the reduction of forecast errors is of the order of 9% for h and 21% for velocity (and 9% and 18% for analysis errors). Probably, this occurs because the ocean is more or less in a steady-state regime and the data sequence is longer. These conditions are ideal for the SA procedure to seek an optimal parameter(s).

Next, in Figs. 9a and b, we compare the results produced by the CHF1 and the AF1 for the forecast errors. It is seen that although the AF1 has improved considerably the quality of the forecast estimates in comparison with those of the CHF, the CHF1 is proven to be more efficient compared to the AF1, especially at the beginning of the assimilation process. Note that their relative performances become more and more identical at the end

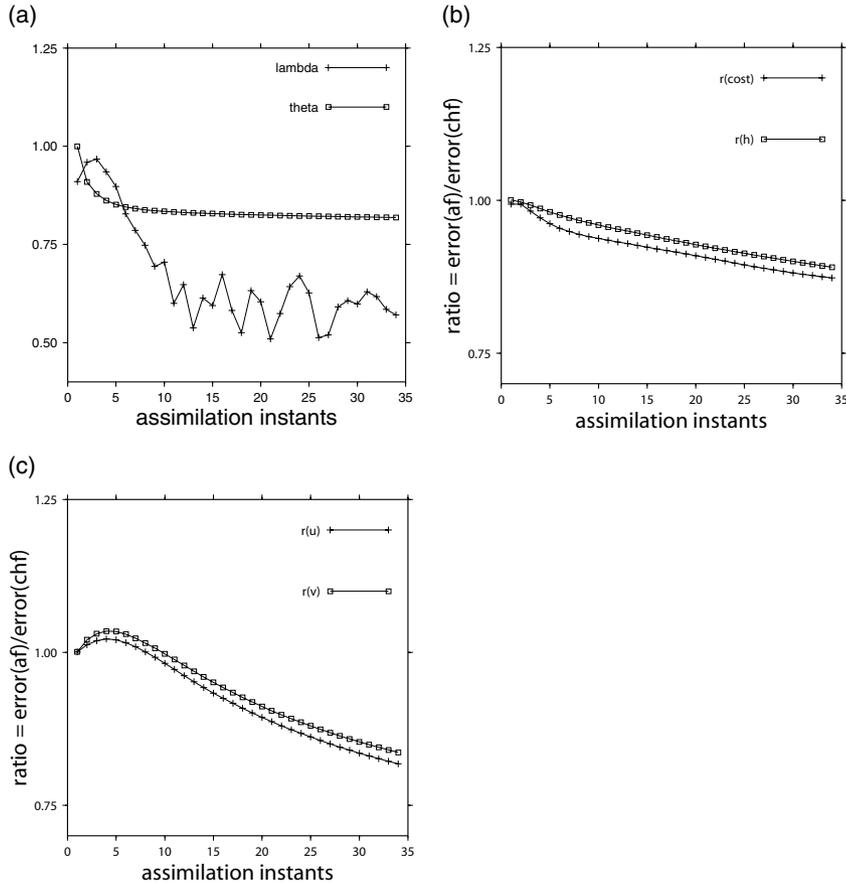


Fig 10. Noisy data set NSO(2). (a) Time evolution of two parameters λ and θ in the AF. (b), (c) The same ratios as shown in Figs. 8a and b, but for the set NSO(2).

of the assimilation process. This seems to be natural because in the AF1 the parameter θ is initialized at the beginning as done in the CHF as $\theta = 1$, whereas in the CHF1 it is assigned the value θ^* for the whole assimilation period. The adaptation allows the AF1 to approach the CHF1 at the end of the assimilation period. These results prove that the value θ^* is not so far from the optimal one. This fact shows also that the performance improvement of the AF1 over the CHF, observed in Section 5.2, is not simply due to the transient character of the adaptation process. The ‘optimality’ of the value θ^* thus has the sense of ‘statistical average’ over the physical processes under consideration, but not only has a ‘local’ character for the particular set SO(1).

5.3.4. Numerical results: noisy observations case. First, note that when the observations are noisy, the performance of the CHF with $\lambda = 1$ seriously degrades in comparison with the assignment $\lambda = 0.91$. This is the reason why we put $\lambda = 0.91$ in the CHF for assimilating the noisy observations. Secondly, the adjustment of the parameter λ plays a very important role in performance improvement of the AF for assimilating noisy observations. This fact is observed from the results of the experiment with noisy observations in which only one parameter θ is adjusted, as done in AF1 (these results are not shown here due to the space limi-

tations of the paper). We observe a much less significant performance improvement of the AF over the CHF compared to the case of simultaneous adjustment of λ and θ . The time evolution of the two tuning parameters λ and θ in the AF is presented in Fig. 10a. As seen here, the parameter λ contributes a large correction for the analysis estimate. Figures 10b and c show the comparison of analysis performances of the CHF and AF. As for the case of noise-free SO(2), the improvement is more significant for the velocity estimation. In time average, the AF is capable of reducing 25% errors for the analysis velocity estimate (and 15% errors for the layer thickness estimate). This reduction is less significant compared to the case of noise-free observations SO(2) reported in the previous section.

6. Concluding remarks

The AF presented in this paper for assimilating SSH observations is based on the adaptive approach developed in previous work (Hoang et al., 1997b, 2001), which is aimed at overcoming two major difficulties related to the very high dimension of the state of oceanic numerical models and to the uncertainty in specification of the model and observation error statistics. The main features of this approach are as follows. (i) The objective

function is to minimize the mean prediction error. (ii) The tuning parameters are chosen as some parameters of the gain matrix. This approach is simple to implement because it is numerically inexpensive and does not require much a priori information on relevant statistics of the model error. In fact, at each data update iteration only one forward integration of the direct model and one backward integration of the AE are to be performed for adjusting the parameters.

A most important and difficult step in applying this method is concerned with the selection of the structure of the gain and its parametrization. One of the natural criteria for this selection is related to the stability of the closed-loop filtering system (Hoang et al., 2000; 2001). Experiments with the adaptive filtering constructed on the basis of leading singular modes or eigenmodes of the linear tangent are under way to be implemented with the MICOM, and will be reported in forthcoming papers.

In the present paper we have developed a technique for construction of the filter gain for assimilating SSH observations, which satisfies some physical constraints required in modelling of the oceanic circulation. These constraints are proven to be reliable and to ensure a stability of the oceanic circulation, as proven in many research works on dynamical oceanography (Cooper and Haines, 1996) as well as in the present work. These constraints define in fact a subspace on to which the innovation is projected to find a correction to be added to the system forecast. The adaptation made by adjusting the tuning parameters is equivalent to seeking the correction lying in the neighbourhood of this subspace.

As seen from Section 2, first the structure of the gain is presented in a form of the optimal interpolation approach. Under the assumption of homogeneity of the forecast ECM and of the observation covariance matrix, the elements of the gain are specified in such a way that three hypotheses H1–H3 are satisfied. By relaxing one or several of these hypotheses, it becomes clear how one can select appropriate tuning parameters to optimize filter performance by adaptation. If it is relatively easy to choose tuning parameters by removing hypotheses H1 (noise-free observation) and H3 (no correction for the velocity at the bottom), relaxing only hypothesis H2 requires more careful examination of the set of all solutions satisfying H3 (or the set of all solutions satisfying H1 if we want to meet only H1).

Numerical experiments with noise-free and noisy observations presented in this paper illustrate the success and high efficiency of the proposed adaptive filtering in the academic context. By adjusting only one or several parameters, the adaptive filtering is capable of improving significantly on the forecast and analysis performance of the filter (for all system variables, layer thickness and velocity), in all layers. These experiments demonstrate also that the proposed technique is presumably useful for enhancing the performance of SSH data assimilation systems.

There remain many questions to be studied in more detail in order to make the adaptive filtering efficient for solving realistic data assimilation problems. For example, this is a question re-

lated to the hypothesis on horizontal spatial homogeneous ECM. One possibility to remove this hypothesis is to seek the best representation of the parameter function in some functional subspace. In this regard, we have tested the adaptive filtering where the parameter function is found in the subspace of polynomial functions of second and third orders. By representing the parameter function as a linear combination of these polynomials, the tuning coefficients are adjusted to minimize the prediction error. The obtained numerical results (not shown in this paper) justify this approach; really, in this way the adaptive filtering can produce a better performance in comparison with the case of homogeneous error covariance structure. Another possibility to choose a functional subspace is concerned with the widely used method of empirical orthogonal functions (see De Mey and Robinson, 1987; Pham et al., 1997) where several dominant eigenvectors of the forecast ECM, for the system state calculated from available observations as well as from model simulations, are used to approximate the true forecast ECM. The questions arising here are, to what extent and how better to combine the thus approximated ECM with the constraints required by H1–H3? Similar experiments are planned to be implemented in cooperation with Dr P. De Mey at LEGOS, Toulouse.

7. Acknowledgments

The authors express their gratitude to the reviewers for their constructive comments and suggestions, which were very helpful in making improvements to the paper. This work is partially supported by SHOM, Groupement de Recherche ‘Methodes Variationnelles en Meteorologie et Oceanographie’, CNRS, Météo-France and IFREMER, France.

References

- Albert, R. 1972. *Regression and the Moore-Penrose pseudoinverse*. Academic Press, New York.
- Amodei, L. 1995. Approximate solution to a problem of meteorological data assimilation with model error. *C. R. Acad. Sci., Paris*, **321**, Series IIa, 1087–1094 (in French).
- Anderson, B. D. O. and Moore, J. B. 1979. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ.
- Baraille, R. and Filatoff, N. 1995. Miami multilayer isopycnal shallow-water model. *Research Report CMO*, EPSHOM/CMO, Brest, France.
- Bennett, A. F., Chua, S. and Leslie, L. M. 1997. Generalized inversion of a Global Numerical Weather Prediction Model. *Meteorol. Atmos. Phys.* **60**, 165–178.
- Bleck, R. 1998. Ocean modeling in isopycnal coordinates. In: *Ocean Modeling and Parameterization* (eds E. P. Chassignet, and J. Verron). NATO Science Series, Kluwer, Dordrecht, 423–448.
- Bryan, K. 1969. A numerical method for the study of the circulation of the world ocean. *J. Comput. Phys.* **4**, 347–376.
- Chassignet, E. 1992. Rings in numerical models of ocean general circulation: a statistical study. *J. Geophys. Res.* **97**, 9479–9492.
- Cohn, S. and Todling, R. 1996. Approximate data assimilation schemes for stable and unstable dynamics. *J. Meteorol. Soc. Japan* **74**(1), 63–75.

- Cooper, M. and Haines, K. 1996. Altimetric assimilation with water property conservation. *J. Geophys. Res.* **101**, 1059–1077.
- Courtier, P. 1993. Introduction to numerical weather prediction data assimilation methods. In: Proceedings of ECMWF workshop on development in the use of satellite data in numerical weather prediction, Shinfield Park, Reading, Berkshire, 180–207.
- Dee, D. 1991. Simplification of the Kalman filter for meteorological data assimilation. *Q. J. R. Meteorol. Soc.* **117**, 365–384.
- De Mey, P. and Robinson, A. R. 1987. Assimilation of altimeter eddy fields in a limited-area quasi-geostrophic model. *J. Phys. Oceanogr.* **17**, 2280–2293.
- Evensen, E. 1994. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**(C5), 10 143–10 162.
- Fukumori, I. and Malanotte-Rizzoli, P. 1995. An approximate Kalman filter for ocean data assimilation: an example with an idealized Gulf Stream model. *J. Geophys. Res.* **100**(C4), 6777–6793.
- Fukumori, I., Benveniste, J., Wunsch, C. and Haivogel, D. 1993. Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother. *J. Phys. Oceanogr.* **23**, 1831–1855.
- Gavart, M. 2001. Assimilation in a basin model and regional modelisation. *Research Report CMO*, EPSHOM/CMO, Brest, France.
- Gavart, M. and De Mey, P. 1997. Isopycnal EOFs in the Azores Current region: a statistical tool for dynamical analysis and data assimilation. *J. Phys. Oceanogr.* **27**, 2146–2157.
- Ghil, M. and Malanotte-Rizzoli, P. 1991. Data assimilation in meteorology and oceanography. In: *Advances in Geophysics*, Vol. 33. Academic Press, New York, 141–266.
- Golub, G. H. and Van Loan, C. 1993. *Matrix Computation*. Johns Hopkins Univ. Press.
- Haines, K., Malanotte-Rizzoli, P., Young, R. E. and Holland, W. R. 1993. A comparison of two methods for the assimilation of altimeter data into a shallow water model. *Dyn. Atmos. Oceans*, **19**, 89–133.
- Hoang, H. S., Baraille, R., Talagrand, O., De Mey, P. and Carton, X. 1997a. Adaptive filtering: application to satellite data assimilation in oceanography. *J. Dyn. Atmos. Oceans* **27**, 257–281.
- Hoang, H. S., De Mey, P., Talagrand, O. and Baraille, R. 1997b. A new reduced-order adaptive filter for state estimation in high dimensional systems. *Automatica* **33**, 1475–1498.
- Hoang, H. S., Baraille, R., Cohn, S. E. and Talagrand, O. 2000. On stability of the partial singular value decomposition filter for stable and unstable dynamics. In: *Proceedings of the International Conference on Mathematical Theory Networks and Systems (MTNS)*, June, Perpignan, France (available at <http://www.univ-perp.fr/mtns2000/articles/B274.pdf>).
- Hoang, H. S., Baraille, R. and Talagrand, O. 2001. On the design of a stable adaptive filter for high dimensional systems. *Automatica* **37**, 341–359.
- Jazwinski, A. H. 1970. *Stochastic Processes and Filtering Theory*. Academic Press, New York.
- Kailath, T. 1980. *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ.
- Le Dimet, F.-X. and Talagrand, O. 1986. Variational algorithm for analysis and assimilation of meteorological observations. *Tellus* **38A**, 97–110.
- Ljung, L. A. 1987. *System Identification*. Prentice Hall, Englewood Cliffs, NJ.
- Mellor, G. L. and Ezer, T. 1991. A Gulf stream model and an altimetry assimilation scheme. *J. Geophys. Res.* **96**(C5), 8779–8795.
- Pham, D. T., Verron, J. and Roubaud, M. C. 1997. Singular evolutive Kalman filter with EOF initialization for data assimilation in oceanography. *J. Mar. Syst.* **16**, 323–340.
- Polyak, B. T. 1990. New method of stochastic approximation type. *Autom. Remote Contr.* **51**, 937–946.
- Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I. Theory. *Q. J. R. Meteorol. Soc.* **113**, 1311–1328.
- Todling, R. and Cohn, S. 1994. Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Mon. Wea. Rev.* **122**, 2530–2557.
- Tsympkin, Ya. 1971. *Adaptation and Learning in Automatic Systems*. Academic Press, New York.