

Filtres de Kalman singuliers évolutifs pour l'assimilation de données en océanographie

Singular evolutive Kalman filters for data assimilation in oceanography

Dinh Tuan PHAM^{1*}, Jacques VERRON² et Lionel GOURDEAU³

¹ Laboratoire de modélisation et calcul, BP 53, 38041 Grenoble cedex

² LEGI, UMR CNRS 5519, BP 53X, 38041 Grenoble cedex

³ Orstom, BP A5, 98848 Nouméa, Nouvelle Calédonie, France

ABSTRACT

We propose a Kalman type filter for assimilating oceanic data into a numerical model. It essentially consists in approximating the error covariance matrix by a singular low rank matrix which leads to making corrections only in those directions for which the error is not sufficiently attenuated by the system. These directions evolve with time according to the model evolution, yielding its adaptative nature to the filter. The filter has been applied in a realistic framework on the Pacific tropical ocean, yielding quite satisfactory results with simulated altimetric data, distributed according to the spatio-temporal sampling from the Topex/Poseidon satellite. © Académie des sciences / Elsevier, Paris.

Keywords: Altimetric data, Data assimilation, Extended Kalman filter, Tropic oceanography, Pacific Ocean

RÉSUMÉ

Nous proposons un filtre de type Kalman pour l'assimilation séquentielle de données océaniques dans un modèle numérique. Il consiste essentiellement à approcher la matrice de covariance des erreurs par une matrice singulière de rang faible, qui conduit à faire des corrections seulement dans les directions où les erreurs ne sont pas assez atténuées par le système. Ces directions *évoluent dans le temps* selon la dynamique du modèle, ce qui confère au filtre son caractère évolutif. Le filtre a été appliqué dans un cadre géographique réaliste à l'océan Pacifique tropical et produit des résultats très satisfaisants, avec des données altimétriques simulées, distribuées selon l'échantillonnage spatio-temporel du satellite Topex/Poseidon. © Académie des sciences / Elsevier, Paris.

Mots clés : Assimilation de données, Données altimétriques, Filtre de Kalman étendu, Océanographie tropicale, Océan Pacifique

ABRIDGED VERSION

1. The extended Kalman filter for data assimilation

Consider a physical system described by (1.1) where \mathbf{x}^t represents the true state of the system, $M(t_{i-1}, t_i)$ is the transition operator from time t_{i-1} to t_i and η_i is the dynamical (or model) error. At each time t_p one observes a vector $\mathbf{y}_i = H_i \mathbf{x}(t_i) + \epsilon_i$, H_i being the observation operator and ϵ_i the

observation error. The (sequential) data assimilation problem consists in estimating $\mathbf{x}^t(t_i)$, given the observations up to time t_i . The extended Kalman filter realizes this sequentially in two stages, starting from an estimate $\mathbf{x}^a(t_{i-1})$ at time t_{i-1} and its corresponding error covariance matrix $\mathbf{P}^a(t_{i-1})$.

1. *Forecast.* The model (1.1) is used to forecast the state at time t_i as $\mathbf{x}^f(t_i) = M(t_p, t_{i-1})\mathbf{x}^a(t_{i-1})$. The forecast error covari-

Note présentée par Jean-Claude André.

Note remise le 6 février 1997, acceptée après révision le 15 septembre 1997.

* E-mail : dinh-tuan.pham@imag.fr

C. R. Acad. Sci. Paris, Sciences de la terre et des planètes / Earth & Planetary Sciences
1998, 326, 255-260

255



Fonds Documentaire ORSTOM
Cote : Bx 17149 Ex : 1

ance matrix is given by (1.2), $M(t_p, t_{i-1})$ being the gradient of $M(t_p, t_{i-1})$ evaluated at $\mathbf{x}^a(t_{i-1})$ and \mathbf{Q}_i being the covariance matrix of the dynamical error.

2. *Correction.* The observation \mathbf{y}_i at time t_i is used to correct the forecast according to

$$\mathbf{x}^a(t_i) = \mathbf{x}^f(t_i) + \mathbf{K}_i[\mathbf{y}^o(t_i) - \mathbf{H}_i \mathbf{x}^f(t_i)];$$

where \mathbf{K}_i is given by (1.3) and \mathbf{H}_i is the gradient of H_i at $\mathbf{x}^f(t_i)$. The corresponding filter error covariance matrix is given by (1.4) and from it \mathbf{K}_i can also be computed by (1.5).

2. The singular evolutive extended Kalman filter

The main idea is to make correction only along the directions for which the error will be amplified by the system. There is at least one such direction for chaotic systems, which are often the case of ocean models. Further, as the system also has an attractor, the errors along complementary directions will be attenuated and hence they can be ignored.

Practically, in the case where there is no dynamical error ($\mathbf{Q}_i = \mathbf{0}$), the extended Kalman filter is applied with an initial singular filter error covariance matrix $\mathbf{P}^a(t_0)$. It is decomposed into $\mathbf{L}_0 \mathbf{U}_0 \mathbf{L}_0^T$, where \mathbf{U}_0 is a positive definite matrix of (low) rank r , and it can be shown that the subsequent $\mathbf{P}^a(t_i)$ can also be decomposed as $\mathbf{L}_i \mathbf{U}_i \mathbf{L}_i^T$ with $\mathbf{L}_p \mathbf{U}_i$ given by (2.1) and (2.2). Thus, by (1.3) and (1.5), the 'directions of correction' are those generated by the columns of \mathbf{L}_p . In the case where dynamical error is present, we project this error vector onto the linear subspace spanned by the column of \mathbf{L}_p before applying the correction calculations of the extended Kalman filter. This results in a more complex updating equation for \mathbf{U}_p given by (2.2') instead of (2.2). A third version of our filter replaces this equation by (2.3''), where $\rho \in (0, 1)$ is a forgetting factor. This factor attenuates the effect of remote observations by giving more weight to the present data and less to the forecast. In this way, the filter may overcome the possible divergence due to the accumulated linearization error and the possible presence of a small dynamical error, which was not accounted for in the model.

3. Stability of the filter

In deriving the above filter, our objective is its stability, not optimality, since the last is too ambitious to attain. We note that for a filter with an arbitrary gain matrix \mathbf{K}_p , the error propagates according to (3.1). This propagation thus depends essentially on the behavior of the sequence of matrices $\mathbf{M}_i - \mathbf{K}_i \mathbf{H}_i \mathbf{M}_i$. To simplify the analysis, we shall regard them as constant during the time span under consideration (and thus drop the subscript i). Then the filter error would remain bounded or explode depending on whether the matrix $\mathbf{M} - \mathbf{KHM}$ is stable (that is having all its eigenvalues of moduli less than 1) or not. It has been shown that for our filter, this matrix is stable if and only if $\bar{r} \leq r$, \bar{r} denoting the number of eigenvalues of \mathbf{M} of modulus greater than or equal to 1.

4. The singular evolutive interpolated Kalman filter

This is a variant of the above filter in which interpolation is used instead of linearization (through the first order Taylor development). Its appeal is double: (i) it avoids the computation of the gradient which can be very complex numerically; and (ii) it would be more robust to model nonlinearities, since interpolation results in less error for large deviations than linearization. In this filter, there is a drawing stage before the forecast stage: one draws $r+1$ interpolating states $\mathbf{x}_1^i(t_{i-1}), \dots, \mathbf{x}_{r+1}^i(t_{i-1})$, randomly but such that its barycenter is $\mathbf{x}^a(t_{i-1})$ and (4.1) is satisfied. The state forecast $\mathbf{x}^f(t_i)$ at time t_i is now simply the barycenter of $\mathbf{x}_1^i(t_{i-1}), \dots, \mathbf{x}_{r+1}^i(t_{i-1})$, where $\mathbf{x}_j^i(t_{i-1}) = M(t_{i-1}, t_i) \mathbf{x}_j^i(t_{i-1})$, and its error covariance matrix is given by (4.2). The correction is similar as before, but with \mathbf{L}_i defined as $[\mathbf{x}_1^i(t_{i-1}) \dots \mathbf{x}_{r+1}^i(t_{i-1})] \mathbf{T}$ and $\mathbf{H}_i \mathbf{L}_i$ is replaced by (4.3); and \mathbf{T} denoting a matrix of r columns defining a basis of the space of vectors in \mathbb{R}^{r+1} with components summing to zero. Further, the matrix \mathbf{U}_{i-1}^{-1} in (2.2) or (2.2') or (2.2'') should be taken as $(r+1)\mathbf{T}^T \mathbf{T}$.

5. Application to the tropical Pacific ocean

After testing our filter on a academic quasigeostrophic model, we have implemented it in the reduced-gravity primitive equations ocean model with σ coordinate in the vertical direction, of Gent and Cane (1989). The model has been ameliorated by the NASA/GSFC team (Murtugudde et al., 1996) and is applied to the tropical Pacific ocean. In the presented test cases, the model has been configured to cover the longitudinal extent of the tropical Pacific ocean between 120°E and 80°W and an meridional extension between -30° and +30° latitude. The resolution is 1° in longitude and variable in latitude, from 0.3° at the equator to 1° at 20° latitude. The area comprises between 20° and 30° latitude in the north and in the south is fact used as a buffer since the ocean is 'open'. There, a progressive relaxation to the climatology according to Levitus (1982) is imposed. On the vertical, the resolution is variable, having nine reference levels (at 50, 10, 15, 25, 50, 50, 50, 50 and 100 m).

The assimilation scheme of our filter has been implemented in its interpolated version. At this preliminary stage of our investigations, we are more interested in the evaluation of the method rather than a real application to reproduce a long sequence of assimilation covering, for example, many known significance cycles of ENSO. The model is forced by the FSU monthly average wind (Goldenberg and O'Brien, 1981), and a parametrisation of the type Kraus-Turner (1987) is used for the surface layer.

The results obtained from the twin experiments (which permits the evaluation of the performance of the assimilation since the true state of the ocean is known in the simulation) are excellent, especially those concerning vertical control of the subsurface temperature fields. They are displayed in figures 1-4.

1. Filtre de Kalman étendu pour l'assimilation de données

Dans les sciences de l'environnement planétaire, c'est surtout l'océanographie qui, après la météorologie, a vu un développement, extrêmement fructueux et prometteur ces dernières années, de l'assimilation de données dans les modèles numériques (voir par exemple Ghil et Manalotte-Rizzoli, 1991). Rappelons brièvement les éléments du problème et l'outil filtre de Kalman (en version étendue à des fins d'application à des modèles non linéaires). Nous adopterons les notations proposées dans (Ide et al., 1997). Le modèle physique est décrit par :

$$\mathbf{x}^t(t_i) = M(t_{i-1}, t_i) \mathbf{x}^t(t_{i-1}) + \eta_i \quad (1.1)$$

où \mathbf{x}^t est le vecteur représentant le vrai état du système, $M(t_{i-1}, t_i)$ est un opérateur décrivant sa transition du temps t_{i-1} au temps t_i et η_i représente l'erreur engendré par l'opérateur (erreur modèle). À chaque instant t_i , on dispose d'un vecteur d'observation $\mathbf{y}_i^o = H_i \mathbf{x}^t(t_i) + \epsilon_i$, où H_i est l'opérateur d'observation et ϵ_i est l'erreur d'observation. Les variables $\eta(t_i)$ et ϵ_i sont supposées centrées, indépendantes, de matrices de covariance \mathbf{Q}_i et \mathbf{R}_i .

L'assimilation (séquentielle) de données consiste à estimer $\mathbf{x}^t(t_i)$ en chaque instant t_i à l'aide des observations jusqu'à cet instant. Le filtre de Kalman étendu réalise cela séquentiellement en deux étapes, en partant d'une estimation courante $\mathbf{x}^a(t_{i-1})$ et d'une matrice de covariance des erreurs $\mathbf{P}^a(t_{i-1})$ à l'instant t_{i-1} .

1. *Prévision* : On se sert de (1.1) pour prédire l'état au temps t_i par $\mathbf{x}^f(t_i) = M(t_i, t_{i-1}) \mathbf{x}^a(t_{i-1})$. En se servant d'une linéarisation du modèle autour de $\mathbf{x}^a(t_{i-1})$, la matrice de covariance de l'erreur de prévision est

$$\mathbf{P}^f(t_i) = M(t_{i-1}, t_i) \mathbf{P}^a(t_{i-1}) M(t_{i-1}, t_i)^T + \mathbf{Q}_i \quad (1.2)$$

où $M(t_{i-1}, t_i)$ est le gradient de $M(t_i, t_{i-1})$ en $\mathbf{x}^a(t_{i-1})$ et T désigne la transposition.

2. *Correction (ou analyse)* : On se sert de \mathbf{y}_i^o pour corriger $\mathbf{x}^f(t_i)$ selon la formule :

$$\mathbf{x}^a(t_i) = \mathbf{x}^f(t_i) + \mathbf{K}_i [\mathbf{y}_i^o(t_i) - H_i \mathbf{x}^f(t_i)]$$

la matrice \mathbf{K}_i , appelée gain, étant choisie pour minimiser l'erreur commise. Elle est donnée par :

$$\mathbf{K}_i = \mathbf{P}^f(t_i) \mathbf{H}_i^T [\mathbf{H}_i \mathbf{P}^f(t_i) \mathbf{H}_i^T + \mathbf{R}_i]^{-1} \quad (1.3)$$

où \mathbf{H}_i est le gradient de H_i évalué en $\mathbf{x}^f(t_i)$. La matrice de covariance de l'erreur correspondante est :

$$\mathbf{P}^a(t_i) = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{P}^f(t_i) \quad (1.4)$$

En tenant compte de cette égalité, on a aussi :

$$\mathbf{K}_i = \mathbf{P}^a(t_i) \mathbf{H}_i^T \mathbf{R}_i^{-1} \quad (1.5)$$

2. Le filtre de Kalman singulier évolutif étendu

La grande difficulté pour appliquer le filtre de Kalman en assimilation de données océanographiques provient de la très grande taille du système. Cela conduit non seulement à un coût de calcul prohibitif, mais aussi à des difficultés pour spécifier adéquatement les matrices \mathbf{Q}_i et \mathbf{R}_i . Si ces dernières sont mal spécifiées, comme cela se produit nécessairement en pratique, ne serait-ce qu'en raison de la grande taille du système, le filtre de Kalman n'est plus optimal. Par suite, nous allons abandonner le pur objectif d'optimalité et nous concentrer sur la stabilité du filtre. Notre but est donc simplement de contenir les erreurs à un niveau raisonnable. Pour cela, on note que l'océan est fondamentalement un système dynamique forcé dissipatif admettant un attracteur. Ce qui signifie que les trajectoires du système sont « attirées » vers une variété (de faible dimension en général) de l'espace des phases, appelé attracteur. Au voisinage de ce dernier, les perturbations « transversales » à l'attracteur seront naturellement atténuées. Mais au vu du caractère chaotique (étrange) de l'attracteur, les perturbations « tangentes » à celui-ci seront fortement amplifiées. L'idée de base est donc de corriger d'abord les erreurs dans les dernières directions, tandis que celles dans les premières directions peuvent être ignorées. On verra ci-dessous que cette idée de correction « sélective » se réalise simplement par l'utilisation d'une matrice de covariance des erreurs singulières de rang faible.

Pour simplifier l'exposé, commençons par le cas où l'erreur modèle est absente ($\mathbf{Q}_i = \mathbf{0}$). On montre alors qu'en initialisant le filtre de Kalman par une matrice $\mathbf{P}^a(t_0)$ de rang r , toutes les matrices $\mathbf{P}^a(t_i)$ auront le même rang : en factorisant $\mathbf{P}^a(t_0)$ en $\mathbf{L}_0 \mathbf{U}_0 \mathbf{L}_0^T$ où \mathbf{U}_0 est une matrice définie positive d'ordre r , on a $\mathbf{P}^a(t_i) = \mathbf{L}_i \mathbf{U}_i \mathbf{L}_i^T$ où

$$\mathbf{L}_i = M(t_{i-1}, t_i) \mathbf{L}_{i-1} \quad (2.1)$$

$$\mathbf{U}_i = (\mathbf{U}_{i-1}^{-1} + \mathbf{L}_i^T \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{L}_i)^{-1} \quad (2.2)$$

La matrice du gain est toujours donnée par (1.5). On voit clairement que les « directions de corrections » sont celles engendrées par les colonnes de \mathbf{L}_i .

Dans le cas où l'erreur modèle est présente, la matrice $\mathbf{P}^a(t_i)$ dans le filtre de Kalman ne serait pas de rang faible, même si elle est initialisée de telle sorte. Toutefois, une matrice de rang faible pourrait en constituer une bonne approximation. La difficulté est que cette dernière est difficile à construire et dépendrait de \mathbf{Q}_i , que l'on connaît mal. Notre idée est de remplacer simplement le vecteur erreur η_i par sa projection sur le sous-espace vectoriel engendré par les colonnes de \mathbf{L}_i . On montre alors que les directions de corrections restent inchangées. La matrice de gain est encore donnée par (1.5) avec $\mathbf{P}^a(t_i) = \mathbf{L}_i \mathbf{U}_i \mathbf{L}_i^T$ et les \mathbf{L}_i calculées par (2.1). Toutefois, la relation (2.2) est remplacée par :

$$U_i^{-1} = [U_{i-1} + (L_i^T L_i)^{-1} L_i^T Q_i L_i (L_i^T L_i)^{-1}]^{-1} (2.2') \\ + L_i^T H_i^T R_i^{-1} H_i L_i$$

Nous avons aussi introduit une autre version du filtre, où on remplace la matrice $(L_i^T L_i)^{-1} L_i^T Q_i L_i (L_i^T L_i)^{-1}$ dans (2.2') par un multiple (positif) de U_{i-1} . Cela résulte en une nouvelle équation pour U_i :

$$U_i^{-1} = \rho U_{i-1}^{-1} + L_i^T H_i^T R_i^{-1} H_i L_i \quad (2.2)$$

où $0 < \rho < 1$. Nous appelons ρ le facteur d'oubli, car celui-ci s'interprète effectivement comme le facteur d'oubli couramment employé dans les algorithmes adaptatifs. En fait, on montre que cette version avec facteur d'oubli peut être considérée comme une extension de la version sans erreur modèle, où l'on a décompté les observations passées d'un facteur ρ à chaque pas de temps. L'introduction du facteur d'oubli atténue l'effet des observations lointaines en donnant plus de poids aux données présentes et moins de poids à la prévision. De cette façon, il permet au filtre de mieux résister à des divergences possibles, causées par l'effet cumulé des erreurs de linéarisation et des faibles erreurs éventuelles de dynamique, ignorées dans le modèle.

3. Stabilité du filtre

Considérons un filtre défini par une suite quelconque de matrices de gain K_i , pas nécessairement donnée par (1.2), (1.4) et (1.3) ou (1.5). Alors, en linéarisant l'opérateur $M(t_{i-1}, t_i)$ autour de $x^a(t_{i-1})$, on voit que l'erreur d'analyse se propage comme :

$$x^a(t_i) - x^f(t_i) \approx (I - K_i H_i) M(t_{i-1}, t_i) \\ [x^a(t_{i-1}) - x^f(t_{i-1})] + K_i \epsilon_i - (I - K_i H_i) \eta_i \quad (3.1)$$

Le comportement de cette erreur dépend donc essentiellement de la suite de matrices $(I - K_i H_i) M(t_{i-1}, t_i)$. En pratique, la dynamique du système étant relativement lente et pas trop non linéaire, on peut supposer que $M(t_{i-1}, t_i)$ est constant (dans une fenêtre de temps) et on l'écrira M . On suppose de même pour R_i et H_i et on les écrit R, H (ces hypothèses ne sont en fait pas contraignantes et pourrait être facilement relâchées). Il est alors clair que, si on prend K_i constant et égal à K , les erreurs de filtrage restent bornées, si et seulement si la matrice $M - KHM$ a toutes ses valeurs propres de module strictement inférieur à 1 (on dira que cette matrice est stable). De façon plus générale, le filtre est stable si la suite des gains K_i converge vers une matrice K , telle que $M - KHM$ soit stable.

Pour la version du filtre sans erreur modèle, on montre, sous ces conditions et une condition d'observabilité, que la suite des matrices $P^a(t_i)$ converge quand $i \rightarrow \infty$ vers VIV^T , où V est telle que ses colonnes engendrent le même sous-espace vectoriel que celui engendré par les premiers

vecteurs propres de M (rangés dans l'ordre décroissant de module de leurs valeurs propres associées) et Π^{-1} est la solution d'une équation de Lyapounov (l'inversibilité de cette matrice est assurée par la condition d'observabilité). De plus, le nombre de colonnes de V est le minimum de r , le rang de $P^a(t_i)$ et r^* le nombre de valeurs propres de M de module supérieur ou égal à 1. On déduit de ce résultat que la suite K_i converge vers $K = VIV^T H^T R^{-1}$ et la matrice $M - KHM$ est stable, si et seulement si $r^* \leq r$. Pour la version avec erreur modèle, on obtient un résultat analogue : $P^a(t_i)$ converge encore quand $i \rightarrow \infty$ vers VIV^T , où V est définie comme précédemment et Π est la solution d'une équation de Riccati. On déduit encore que la limite de K_i est telle que $M - KHM$ soit stable, si et seulement si $r^* \leq r$. Enfin, les arguments précédents peuvent être généralisés pour prouver la stabilité de la version avec facteur d'oubli (voir Pham et al., 1998 pour les détails).

4. Le filtre de Kalman singulier évolutif interpolé

C'est une variante du filtre de Kalman singulier précédent, où l'on a utilisé l'interpolation (linéaire) à la place de la linéarisation (par le développement de Taylor du 1^{er} ordre). Son intérêt est double : (i) il évite le calcul du gradient, qui peut être très complexe numériquement, et (ii) il semble être plus robuste vis-à-vis de la non-linéarité du modèle, car l'interpolation résulte en une moindre erreur pour les grandes déviations que la linéarisation. Ce filtre s'opère en trois étapes, en partant d'une estimation courant $x^a(t_{i-1})$ et d'une matrice de covariance des erreurs $P^a(t_{i-1})$, à l'instant t_{i-1} .

Tirage des états interpolants : On choisit $r + 1$ vecteurs $x_j^i(t_{i-1})$, ..., $x_{r+1}^i(t_{i-1})$ au hasard, mais qui satisfont les contraintes suivantes :

- que leur barycentre soit $x^a(t_{i-1})$ et
- que $P^a(t_{i-1}) = \frac{1}{r+1} \sum_{j=1}^{r+1} [x_j^i(t_{i-1}) - x^a(t_{i-1})] \cdot [x_j^i(t_{i-1}) - x^a(t_{i-1})]^T$ (4.1)

Prévision : On applique le modèle pour amener chaque $x_j^i(t_{i-1})$ en $x_j^f(t_i) = M(t_{i-1}, t_i) x_j^i(t_{i-1})$, $i = 1, \dots, r + 1$. La prévision $x^f(t_i)$ de l'état en t_i sera le barycentre des $x_j^f(t_i)$ et a pour matrice de covariance des erreurs :

$$P^f(t_i) = Q_i + \frac{1}{r+1} \sum_{j=1}^{r+1} [x_j^f(t_i) - x^f(t_i)] \cdot [x_j^f(t_i) - x^f(t_i)]^T \quad (4.2)$$

Correction : On calcule

$$L_i = [x_1^f(t_i) \dots x_{r+1}^f(t_i)]^T,$$

où T est une matrice de r colonnes, définissant une base de l'espace des vecteurs de \mathbb{R}^{r+1} dont la somme des compo-

santes est nulle. Comme précédemment, on projette le vecteur erreur modèle η_i sur le sous-espace vectoriel engendré par les colonnes de L_i . On est alors conduit à la formule de correction, avec comme gain $K_i = L_i U_i (HL)_i^T R_i^{-1}$ et à la matrice de covariance des erreurs $P^a(t_i) = L_i U_i L_i^T$, où :

$$(HL)_i = [H_i x_1^i(t_i^-) \dots H_i x_{r+1}^i(t_i^-)]^T$$

$$U_i^{-1} = \{ [(r+1)T^T T]^{-1} + (L_i^T L_i)^{-1} L_i^T Q_i L_i (L_i^T L_i)^{-1} \}^{-1} + (HL)_i^T R_i^{-1} (HL)_i$$

La première matrice du membre de droite précédente se réduit à $(r+1)T^T T$ dans la version sans erreur modèle et à $\rho(r+1)T^T T$ dans la version avec facteur oublié.

5. Application à l'océan Pacifique tropical

Après de premiers tests académiques dans un modèle quasigéostrophique simple (Pham et al., 1998), notre filtre a été implanté dans le modèle aux équations primitives, à coordonnée σ sur la verticale, de Gent et Cane (1989), amélioré par l'équipe de la NASA/GSFC (Murtugudde et al., 1996), sur la zone géographique de l'océan Pacifique tropical. L'avantage de ce modèle est son « poids » informatique relativement modeste, malgré la dimension du vecteur d'état (supérieure à 400 000). La raison ne tient pas à la nature des équations (qui sont essentiellement les mêmes que celles des autres modèles beaucoup plus coûteux aux équations primitives), mais à l'hypothèse de gravité réduite qui ramène la couche active de l'océan à quelques centaines de mètres seulement. Les effets des couches océaniques profondes et de la topographie, par exemple, ne sont donc pas pris en compte. C'est évidemment une approximation importante. Mais on sait depuis longtemps que l'approximation en « deux couches » de l'océan tropical est solide : les résultats récents obtenus avec ce modèle sont d'ailleurs tout à fait convaincants à cet égard et semblent montrer, en particulier, des évolutions très réalistes de la Warm Pool (énorme étendue d'eau chaude délimitée par l'isotherme 29 °C).

Dans les cas tests présentés, le modèle a été configuré sur toute l'étendue longitudinale du Pacifique tropical de 120°E à 80°W et une extension méridionale comprise entre -30° et +30° de latitude. La résolution est de 1° en longitude, et est variable en latitude de 0,3° à l'équateur à 1° à 20° de latitude. Les zones comprises entre 20° et 30° de latitude au nord et au sud sont utilisées en fait comme « zones tampons », puisque l'océan est « ouvert » à ces niveaux. Dans ces régions, un rappel progressif à la climatologie de Levitus (1982) est donc imposé. Ces régions sont de toute manière peu actives (et choisies d'ailleurs pour cette raison) par rapport aux régions proches de l'équateur. Sur la verticale, la résolution est variable et en 9 niveaux (respectivement 50, 10, 15, 25, 50, 50, 50, 50 et 100 m de la surface au fond) de référence. Le modèle est

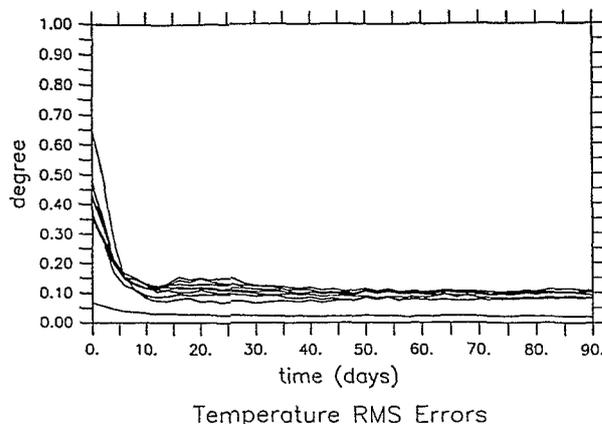
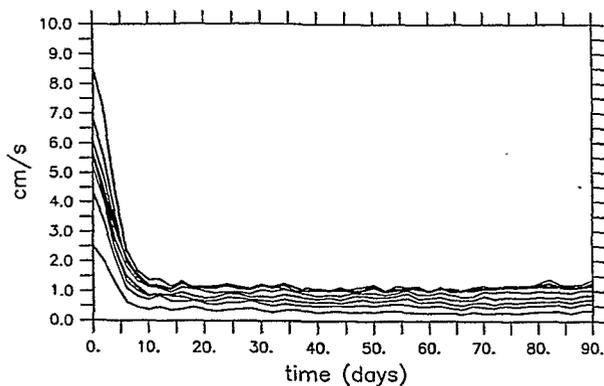


Figure 1. Convergence des erreurs rms sur les champs de température dans les neuf couches du modèle au cours des trois mois d'expérience d'assimilation.

Convergence of the rms errors on the temperature fields in the nine model layers during the three months of the assimilation experiment.

forcé par les vents moyens mensuels FSU (Goldenberg et O'Brien, 1981) et une paramétrisation de type Kraus-Turner (1987) est utilisée pour la couche de mélange superficielle.

Les conclusions de faisabilité de notre filtre sont positives dans une large mesure, bien que certains réglages (et donc la compréhension de certaines propriétés) restent à réaliser, en particulier le choix de r . L'analyse théorique montre que celui-ci doit être plus grand ou égal à un seuil r^* , mais on ne connaît pas ce dernier. Pratiquement, nous choisissons r de manière empirique, pour avoir un résultat satisfaisant sans trop alourdir le coût de calcul (dans cette expérience, r peut être aussi faible que 6). Les résultats obtenus avec des expériences jumelles (qui permettent de mesurer réellement la performance de l'assimilation, puisque le « monde réel » est supposé connu) sont excellents, notamment ceux portant sur le contrôle vertical des champs de température de subsurface : la seule donnée altimétrique de surface échantillonnée selon Topex/Poseidon permet un très bon contrôle de la structure thermique dans toutes les couches océaniques traitées (environ 400 m). La figure 1 montre, par exemple, la convergence de l'erreur sur la température dans les neuf couches du modèle à partir d'un état initial quelconque, par comparaison avec l'expérience de référence. La figure 2 montre la courbe équivalente pour le champ de vitesse zonale. D'une manière plus générale, on constate qu'après une séquence d'assimilation couvrant environ deux cycles de Topex/Poseidon, l'erreur sur la hauteur de la mer, les champs de température, de vitesse zonale et méridionale décroît à des niveaux inférieurs à 5 % de l'erreur initiale, et ceci dans toutes les couches du modèle. Pour la même expérience, les figures 3 et 4 montrent le champ de température superficiel et le champ d'erreur sur l'ensemble du bassin Pacifique tropical après le cycle de trois mois d'assimilation. On voit ainsi les régions géographiques où le contrôle d'erreur est le mieux réalisé.



Zonal Velocity RMS Errors

Figure 2. Convergence des erreurs rms sur les champs de vitesse zonale dans les neuf couches du modèle au cours des trois mois d'expérience d'assimilation.

Convergence of the rms errors on the zonal velocity in the nine model layers during the three months of the assimilation experiment.

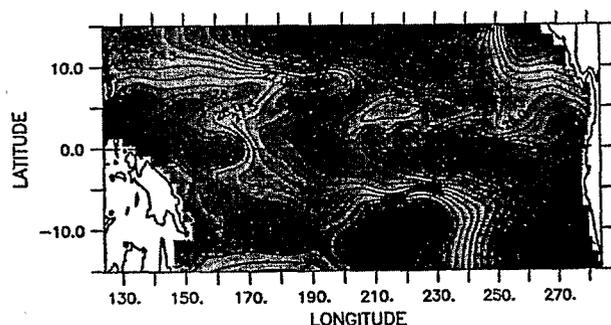


Figure 3. Champ de température de surface au 31 décembre 1992.

Sea surface temperature on 31 December 1992.

De premières expériences ont été réalisées sur les trois premiers mois de la mission Topex/Poseidon entre octobre et décembre 1992. Les premières validations sont positives, mais restent encore à être affinées, à la fois sur le plan du comportement des propriétés statistiques du filtre et sur

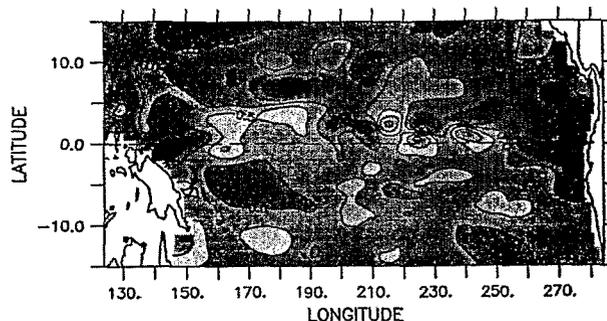


Figure 4. Erreur du champ de température de surface au 31 décembre 1992.

Error on the sea surface temperature on 31 December 1992.

celui de la validation, avec des données indépendantes. Il reste surtout à prolonger cette expérience préliminaire, sur des périodes de temps plus longues, pour lesquelles une réelle validation avec des observations in situ sera plus significative.

6. Conclusion

Au-delà de ces études en cours sur l'assimilation de longue durée des données Topex/Poseidon, les étapes suivantes porteront d'abord sur l'utilisation conjointe de ces dernières et de celles du réseau TAO (réseau international de mouillages fixes sur la ceinture équatoriale Pacifique). À terme, l'objectif est d'intégrer les autres données disponibles spatiales (ERS1, ERS2) et in situ (XBT et thermosalinographes des réseaux de navire marchands), etc.

Sur le plan méthodologique, notre filtre semble avoir démontré, au moins de manière préliminaire, sa faisabilité et son potentiel. Par ailleurs, la méthode présente un caractère relativement « portable » (sous réserve de la définition adéquate du vecteur d'état) et est donc susceptible de généralisation à d'autres modèles. Elle est d'ailleurs en cours d'implantation dans deux modèles d'océan aux équations primitives appliqués au bassin Atlantique.

Remerciements : Les auteurs remercient O. Talagrand pour ses commentaires judicieuses sur ce travail.

RÉFÉRENCES

Gent P.R. et Cane M.A. 1989. A reduced gravity primitive equation model of the upper equatorial ocean, *J. Comput. Phys.*, 81, 444-480
 Ghil M. et Manalotte-Rizzoli P. 1991. Data assimilation in meteorology and oceanography, *Adv. Geophys.*, 33, 141-266
 Goldenberg S.B. et O'Brien J.J. 1981. Time and space variability of tropical Pacific wind stress, *Mon. Weather Rev.*, 109, 1190-1207
 Ide K., Bennett A.F., Courtier P., Ghil M. et Lorenc A.C. 1997. Unified notation for data assimilation: operational, sequential and variational, *J. Met. Soc. Japan*, 75, 1B, 71-79

Kraus E.B. et Turner J.S. 1967. A one-dimensional model of the seasonal thermocline. Part II, *Tellus*, 19, 98-105

Levitus S. 1982. *Climatological Atlas of the World Oceans* - NOAA Prof. Pap. 13, US Govt. Print. Off., Washington, D.C., 173 p.

Murtugudde R., Seager R. et Busalacchi A.J. 1996. Simulation of the Tropical Oceans with an ocean GCM coupled to an atmospheric mixed layer model, *J. Climate*, 9, 1795-1815

Pham, D.T., Verron J. et Roubaud M. C. 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography, *J. Marine Syst.*, à paraître