ELSEVIER

Journal of Marine Systems 16 (1998) 323–340

JOURNAL OF MARINE SYSTEMS

A singular evolutive extended Kalman filter for data assimilation in oceanography

Dinh Tuan Pham [a,*], Jacques Verron [b], Marie Christine Roubaud [a]

[a] *Laboratoire de Modélisation et Calcul, Projet Idopt, INRIA / JNPG / UJF / CNRS, B.P. 53X, 38041 Grenoble Cédex, France*
[b] *Laboratoire des Écoulements Géophysiques et Industriels, UMR 5519 CNRS, UJF, INPG, B.P. 53X, 38041 Grenoble Cédex, France*

Received 23 September 1996; accepted 12 November 1997

Abstract

In this work, we propose a modified form of the extended Kalman filter (KF) for assimilating oceanic data into numerical models. Its development consists essentially of approximating the error covariance matrix by a singular low rank matrix, which amounts in practice to making no correction in those directions for which the error is the most attenuated by the system. This not only reduces the implementation cost but may also improve the filter stability as well. These 'directions of correction' evolve with time according to the model evolution, which constitutes the most original feature of this filter and distinguishes it from other sequential assimilation methods based on the projection onto a fixed basis of functions. A method for initializing the filter based on the empirical orthogonal functions (EOF) is also described. An example of assimilation based on the quasi-geostrophic (QG) model for a square ocean domain with a certain wind stress forcing pattern is given. Although this is only a simple test case designed to assess the feasibility of the method, the results are very encouraging.

Keywords: Kalman filter; direction of correction; assimilation; quasi-geostrophic model

1. Introduction

This work was motivated by the problem of data assimilation in oceanography (or in geophysical sciences generally). In general, data assimilation is seen as the means of obtaining a consistent picture of a geophysical system optimally blending all the information available on this system. This information may consist of data of all types, of varying accuracy and geographical distribution, and also of information derived from the theoretical knowledge already possessed on the system expressed in terms of physical laws, either deterministic or statistical, and therefore via various forms of models.

Several methods exist for assimilating observations into numerical models, most of them originally developed in meteorology. Two main approaches are usually seen in the assimilation techniques, following either the optimal control theory or the statistical estimation theory (see for example Ghil and Manalotte-Rizzoli,

* Corresponding author. Fax: +33-4-76-63-1263; e-mail: dinh-tuan.pham@imag.fr

0924-7963/98/$ - see front matter
PII: S0924-7963(97)00109-7

1991). In the present study, we are interested in the second approach in which the Kalman filtering theory is the primary framework. But the application of this theory encounters enormous difficulties due to the huge dimension of the state vector of the considered system. A further major difficulty is caused by its non-linear nature. To deal with this, one usually linearizes the ordinary Kalman filter (KF) leading to the so-called extended Kalman filter (EKF). We shall assume that the linearized system is close enough to the original one for the results in the linear case to be transferable to the non-linear case.

In this paper, we shall concentrate on the issue of dimension and propose a method to partially overcome this difficulty. To this end, a brief description of EKF together with some discussion of problems arising in its use in geophysical data assimilation will be first given in Section 2. From this, an algorithm based on the use of singular low rank error covariance matrix, will be introduced. This filter, called the singular evolutive extended Kalman (SEEK) filter, not only solves the practical problem of reducing the computational cost to an acceptable level, but in addition possesses a certain theoretical foundation. Indeed, as it is pointed out in Section 2, it would be utopic to look for an 'optimal' filter in such a high dimensional context, so our aim may simply be confined to finding a stable filter which reduces the propagation of error from one step to the next. Our SEEK filter is designed to achieve this. It operates by applying correction in certain directions only and thus it has some similarities to certain other recently proposed reduced order KF (see references in Section 3). But the novel feature of our filter is that these directions are constructed to be those for which the system least attenuates the error and therefore evolve in time as the state of the system changes (hence the name 'evolutive'). In this regard, our filter possesses a certain robustness with respect to initialization of possibly poor quality. However, we also provide a method for finding a good initial singular low rank error covariance matrix, based on the empirical orthogonal functions (EOFs) also known as principal components (PCs). The overall method is illustrated by an example in Section 4. In this simple academic test case performed with a simple quasi-geostrophic (QG) model, an assessment of the feasibility of the method is only sought. Real applications are currently being investigated in the North Atlantic and in the tropical Pacific with more complex models and these will be reported later.

2. The extended Kalman filter

For consistency with the existing literature on the subject, we adopt most notations proposed by Ide et al. (1997). Consider a physical system described by

$$\boldsymbol{x}^{\mathrm{t}}(t_i) = M(t_{i-1}, t_i)\,\boldsymbol{x}^{\mathrm{t}}(t_{i-1}) + \eta(t_i) \tag{2.1}$$

where $\boldsymbol{x}^{\mathrm{t}}$ is a vector representing its true state, $M(t_{i-1}, t_i)$ is an operator describing the system transition from time t_{i-1} to t_i and $\eta(t_i)$ a noise term. The transition operator $\mathbf{M}(s,t)$ is usually obtained from the integration of a certain partial differential system, typically a numerical model. At each t_i, one observes a vector

$$\boldsymbol{y}_i^{\mathrm{o}} = H_i\,\boldsymbol{x}^{\mathrm{t}}(t_i) + \boldsymbol{\epsilon}_i \tag{2.2}$$

where H_i is the observational operator and $\boldsymbol{\epsilon}_i$ represents the noise. The random vector $\eta(t_i)$ and $\boldsymbol{\epsilon}_i$ are assumed to be independent white noise processes with zero mean and covariance matrix $\mathbf{Q}(t_i)$ and $\mathbf{R}_i$, respectively.

The problem is to estimate $\boldsymbol{x}^{\mathrm{t}}(t_i)$ given the observations up to this time. Since the time of Kalman (1960), this problem has been the source for numerous developments in various fields of science and in particular that of atmospheric and oceanic flow (Ghil, 1986; Ghil and Manalotte-Rizzoli, 1991). The KF has been extended to non-linear system through linearization, yielding the so-called EKF (Ghil et al., 1982) which has been the subject of various investigations (Miller et al., 1994; Evensen, 1992, 1993, 1994). It may however result in instability in the covariance evolution equation (Evensen, 1992; Gauthier et al., 1993).

The EKF operates sequentially: from an analysis state vector $\boldsymbol{x}_a(t_{i-1})$ and its error covariance matrix $\mathbf{P}^{\mathrm{a}}(t_{i-1})$, it constructs the next $\boldsymbol{x}^{\mathrm{a}}(t_i)$ and $\mathbf{P}^{\mathrm{a}}(t_i)$ as follows.

(i) Forecasting. The model Eq. (2.1) is used to forecast the state at time t_i: $\boldsymbol{x}^{\mathrm{f}}(t_i) = \mathbf{M}(t_{i-1},t_i)\boldsymbol{x}^{\mathrm{a}}(t_{i-1})$. By linearizing Eq. (2.1) around $\boldsymbol{x}^{\mathrm{a}}(t_{i-1})$ as

$$\boldsymbol{x}^{\mathrm{t}}(t_i) \approx M(t_{i-1},t_i)\,\boldsymbol{x}^{\mathrm{a}}(t_{i-1}) + \mathbf{M}(t_{i-1},t_i)\left[\boldsymbol{x}^{\mathrm{t}}(t_{i-1}) - \boldsymbol{x}^{\mathrm{a}}(t_{i-1})\right] + \eta(t_i), \tag{2.1′}$$

where $\mathbf{M}(t_{i-1},t_i)$ denotes the gradient of $M(t_i,t_{i-1})$ *evaluated* at $\boldsymbol{x}^{\mathrm{a}}(t_{i-1})$, the forecast error covariance matrix can be seen to be approximately

$$\mathbf{P}^{\mathrm{f}}(t_i) = \mathbf{M}(t_{i-1},t_i)\mathbf{P}^{\mathrm{a}}(t_{i-1})\mathbf{M}(t_{i-1},t_i)^{\mathrm{T}} + \mathbf{Q}(t_i) \tag{2.3}$$

wherein T denotes the transpose.

(ii) Analysis (or correction). The new observation $\boldsymbol{y}_i^{\mathrm{o}}$ is used to correct the forecast state vector $\boldsymbol{x}^{\mathrm{f}}(t_i)$ according to $\boldsymbol{x}^{\mathrm{a}}(t_i) = \boldsymbol{x}^{\mathrm{f}}(t_i) + \mathbf{K}_i[\boldsymbol{y}^{\mathrm{o}}(t_i) - \mathbf{H}_i\boldsymbol{x}^{\mathrm{f}}(t_i)]$, where $\mathbf{K}_i$ is a certain matrix called the gain. The optimal gain can be shown to be

$$\mathbf{K}_i = \mathbf{P}^{\mathrm{f}}(t_i)\mathbf{H}_i^{\mathrm{T}}\left[\mathbf{H}_i\mathbf{P}^{\mathrm{f}}(t_i)\mathbf{H}_i^{\mathrm{T}} + \mathbf{R}_i\right]^{-1}, \tag{2.4}$$

where $\mathbf{H}_i$ is the gradient of $\mathbf{H}_i$ evaluated at $\boldsymbol{x}^{\mathrm{f}}(t_i)$. The corresponding analysis error covariance matrix is

$$\mathbf{P}^{\mathrm{a}}(t_i) = \mathbf{P}^{\mathrm{f}}(t_i) - \mathbf{P}^{\mathrm{f}}(t_i)\mathbf{H}_i^{\mathrm{T}}\left[\mathbf{H}_i\mathbf{P}^{\mathrm{f}}(t_i)\mathbf{H}_i^{\mathrm{T}} + \mathbf{R}_i\right]^{-1}\mathbf{H}_i\mathbf{P}^{\mathrm{f}}(t_i).$$

For the mathematical analysis, the whole algorithm can be described by two equations: (i) the Riccati equation which updates the error covariance matrix

$$\begin{aligned}\mathbf{P}_i^{\mathrm{a}} = {} & \mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}} + \mathbf{Q}_i - \left(\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}} + \mathbf{Q}_i\right)\mathbf{H}_i^{\mathrm{T}} \\ & \times\left[\mathbf{H}_i\left(\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}} + \mathbf{Q}_i\right)\mathbf{H}_i^{\mathrm{T}} + \mathbf{R}_i\right]^{-1}\mathbf{H}_i + \left(\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}} + \mathbf{Q}_i\right)\end{aligned} \tag{2.5}$$

putting $\mathbf{P}_i^{\mathrm{a}} = \mathbf{P}^{\mathrm{a}}(t_i)$, $\mathbf{Q}_i = \mathbf{Q}(t_i)$ and $\mathbf{M}_{i-1,i} = \mathbf{M}(t_{i-1},t_i)$ for short, and (ii) the filter equation

$$\boldsymbol{x}^{\mathrm{a}}(t_i) = M(t_{i-1},t_i)\,\boldsymbol{x}^{\mathrm{a}}(t_{i-1}) + \mathbf{K}_i\left[\boldsymbol{y}_i^{\mathrm{o}} - \mathbf{H}_i M(t_{i-1},t_i)\,\boldsymbol{x}^{\mathrm{a}}(t_{i-1})\right] \tag{2.6}$$

where $\mathbf{K}_i$ is related to $\mathbf{P}_{i-1}^{\mathrm{a}}$ through Eqs. (2.3) and (2.4). From Eq. (2.5), it can be shown that when $\mathbf{R}_i$ is invertible

$$\mathbf{K}_i = \mathbf{P}_i^{\mathrm{a}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}. \tag{2.7}$$

A practical difficulty with the EKF is that it assumes everything known. While one can reasonably assume that H_i and $M(t_{i-1},t_i)$ are known, it would be too optimistic to assume so for the covariance matrices $\mathbf{R}_i$, $\mathbf{Q}_i$ and $\mathbf{P}_0^{\mathrm{a}}$. For $\mathbf{R}_i$, one may have good knowledge of the instrumental error variances in situations such as altimetric observations from the satellite Topex/Poseidon over the ocean, for which the error estimates have become fairly solidly established, being based on now extensive verification and validation phases (see for example the two special issues of the Journal of Geophysical Research, Vol. 99, C12, Dec. 1994, and Vol. 100, C12, Dec. 1995). But it is not clear how the correlations of these errors can be obtained. One often assumes that they vanish but the possibility that errors in neighboring observations be correlated cannot be ruled out (in fact this occurs frequently as a result of geophysical correction). More difficult to obtain are the covariances matrices $\mathbf{Q}_i$ and $\mathbf{P}_0^{\mathrm{a}}$. The first would require the accurate knowledge of the statistical behavior of the state process $\boldsymbol{x}^{\mathrm{t}}(t_i)$, but how can this be acquired without actually observing it? As for the second, quite often very little is known concerning the initial state of the system and so even less about the error when some $\boldsymbol{x}^{\mathrm{a}}(t_0)$ is chosen. One should not forget that the systems encountered in meteorology and oceanography are of very large dimension so that the matrices $\mathbf{Q}_i$ and $\mathbf{P}_0^{\mathrm{a}}$ have a vastly larger number of independent elements (namely $N(N-1)/2$ where N is the dimension of the system, typically 10^5–10^6 or even higher). It is questionable whether one can have even a crude knowledge on such a huge number of parameters. This point is discussed at length in the work of Cane et al. (1996). Here we note, only briefly, that a statistical estimate based on a sample of length T has a convergence rate of $1/\sqrt{T}$ and in practice T would be only a tiny fraction of N. Therefore, if one tries to

estimate such a huge number of parameters, there will be inevitably large errors in many. As for the choice of $\mathbf{P}_0^a$, it might be of somewhat lesser importance in view of the property of the Riccati Eq. (2.5). The theory for such equation states that for linear autonomous systems and under appropriate conditions 'any solution will converge to the same limit regardless of the starting point $\mathbf{P}_0^a$' (Hager and Horowitz, 1976). Therefore, even if $\mathbf{P}_0^a$ is poorly specified, one may hopefully still have a good approximation to $\mathbf{P}_i^a$, in the long term. But whether this is true in practice is debatable. Our system is not linear autonomous, at most it is approximately so. More precisely, $\mathbf{M}_{i-1,i}$ may be regarded as constant in time only for a limited time span. As our system is huge, convergence may be so slow that $\mathbf{P}_i^a$ never approaches its limit before $\mathbf{M}_{i-1,i}$ has changed significantly. A further point is that for the above result to hold, one of the conditions needed is controllability. This condition, taken from the control theory with the dynamic noise $\eta(t_i)$ playing the role of the control, requires that the state of the system can be brought to any point in a neighborhood of its starting point by some realization of $\eta(t_i)$. Such condition is not verified if there is no dynamic noise or if such noise is highly directional. But on physical grounds and also from observations, it appears that the covariance matrix of the state vector can be represented to a good approximation by a low rank matrix, which implies that the dynamic noise, if it exists, should be highly directional.

From the above discussion, we may conclude that there is a real possibility that the matrices $\mathbf{R}_i$ and especially $\mathbf{Q}_i$ are imperfectly specified (the discussion of Fukumori (1995) with regards to this question of the model and data error estimates is particularly valuable). The KF is optimal, but only if the parameters $\mathbf{R}_i$ and $\mathbf{Q}_i$ are correctly specified. Thus, any KF is in practice somewhat suboptimal. As far as the EKF is concerned, there is also the crucial additional issue of the linearization error.

3. Singular evolutive extended Kalman filter

The above considerations suggest that it may be too ambitious to demand optimality, even to an approximate degree. Therefore, we will concentrate on a more modest aim: stability. By this we mean that the filtering error should remain bounded. Consider a general filter algorithm defined by the Eq. (2.6) and certain gain matrix $\mathbf{K}_i$, but not necessarily the one given by Eqs. (2.5) and (2.7). Then from Eq. (2.1′) the filter error is propagated according to

$$\boldsymbol{x}^a(t_i) - \boldsymbol{x}^t(t_i) \approx (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{M}_{i-1,i}\left[\boldsymbol{x}^a(t_{i-1}) - x^t(t_{i-1})\right] - \mathbf{K}_i\boldsymbol{\epsilon}_i - (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\,\eta(t_i). \tag{2.8}$$

This formula shows that the behavior of the error depends essentially on the sequence of matrices $(\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{M}_{i-1,i}$. To simplify things, consider the 'nearly linear autonomous' case where $\mathbf{M}_{i-1,i}$ may be regarded as constant (over a certain time span) and assume $\mathbf{H}_i$ as the constant in time, so that the subscript i will be dropped. If $\mathbf{K}_i$ is chosen constant $(=\mathbf{K})$, then what matters are the eigenvalues of $\mathbf{M} - \mathbf{KHM}$. The error will diverge if one eigenvalue has modulus greater than 1 and remain bounded if all eigenvalues have modulus less than 1. In the last case the matrix $\mathbf{M} - \mathbf{KHM}$ is said to be stable. Thus, our aim would be to make it is so. Referring to Eq. (2.8), this means that the aim is to reduce the error growth of the filter, so that it is attenuated at each step. It is clear from Eq. (2.8) that this depends on the choice of the filter gain $\mathbf{K}_i$ together with the system dynamics and the observation operator, but not on the dynamical and observational noise covariance matrices. Thus, by focusing on this property, we need not worry about these matrices being badly specified.

Practical constraints on actual implementation of the KF or EKF leave no other choice but to reduce the size of the system in some way. Schematically several strategies are possible. Firstly, one can simply reduce the dimension of the model state vector to make it compatible with the computational constraint (Miller and Cane, 1989; Gaspar and Wunsch, 1989; Gourdeau et al., 1992; Fukumori et al., 1993). Quite often, one also assumes a certain asymptotic property of constancy for the error covariance matrix (Gourdeau et al., 1992; Fu et al., 1993; Fukumori and Malanotte-Rizzoli, 1995; Fukumori, 1995). Secondly, one might consider reducing the KF working space until the computation becomes feasible (without necessarily reducing the size of the model!).

Fukumori (1995), for example, undersampled the computational grid used for the filter description. Several studies are based on some ways of projecting the filter evolution on a reduced dimension basis (Evensen, 1994; Cane et al., 1996, Hoang et al., 1995). Another approach consists of relying on physical considerations which will naturally decrease the degree of freedom of the system: long wave approximation in the tropics (Cane and Patton, 1984) and geostrophy (Dee, 1991) as examples, among others.

Physical considerations support the idea of reducing the 'working dimension' of the filter in some way. Indeed, the ocean is basically a forced and dissipated dynamical system possessing an attractor, meaning that asymptotically the trajectories of the state vector occupy only a small part of the phase space, the so-called 'attractor'. Clearly, the description of this attractor implies a much lower degree of freedom than the description of the whole system. However, such description does not come easily since the attractor is basically 'curved', so that methods using linear projections would produce some error. Here, we adopt a more flexible approach by retaining the state space but restricting the corrections to a linear space 'tangent' to the attractor in some sense. This way, the correction does not make the state vector leave the attractor once it is nearly on it. Note that in the case of the ocean (and of the atmosphere) the attractor is 'strange' (or chaotic): a slight perturbation of the state vector can develop into a marked divergence in the phase space. Our filter is constructed in such a way that it corrects in priority the errors in the directions for which they are most amplified. These directions are necessarily tangent to the attractor since when a state vector is 'pushed' outside the attractor, it will be pulled back (though not to the same place) to the attractor.

One may view our filter as a variant of the EKF with a singular low rank error covariance matrix $\mathbf{P}_i^{\mathrm{a}}$. In this respect, it is similar to the reduced rank Kalman filtering introduced by Cohn and Todling (1995, 1996) and Verlaan and Hemmink (1995). However, there are differences, in Cohn and Todling (1995, 1996), a singular value decomposition (SVD) of $\mathbf{M}_{i-1,i}$ or an eigenvalue decomposition (EVD) of $\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}}$ is performed in order to obtain a low rank approximation, before entering the filter calculation. We feel that this procedure may be too costly if such SVD or EVD has to be done step by step. In Verlaan and Hemmink (1995), a linear time invariant system (or a steady state filter) is considered so that such decomposition is done only once. No such decomposition is required in our filter: the algorithm is constructed in such a way that it always produces a low rank error covariance matrix. Further, it is meant to work in a non-linear time variant system (at least when the non-linearity is not too strong and the time evolution not too fast). There is also another subtle difference: the image space of our error covariance matrix does not correspond (asymptotically) to the eigenspace of $\mathbf{M}_{i-1,i}\mathbf{M}_{i-1,i}^{\mathrm{T}}$ or of $\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}}$ (as implied in the above authors's works), but to that of $\mathbf{M}_{i-1,i}$.

We shall present three versions of our algorithm: a no dynamic noise version, a general version and a version with forgetting factor. The reason for presenting the first version separately is that it is simpler and easier to understood, being a particular case of the EKF in the absence of dynamic noise ($\mathbf{Q}_i = 0$). Despite its name, it can be used in the case where dynamic noise do exists, since the feature which we are focused in, namely its stability, is not affected by the presence of such noise. Of course, there would be a loss of performance, so it would be more appropriate to use the general version when the dynamic noise is strong. This version is no longer a particular case of the EKF but it could have a good performance since it makes correction in the directions for which the error is likely to be high. More importantly still, it is designed to possess the stability property. But it is more complex and costly in terms of computation and requires the specification of the dynamic noise covariance matrix $\mathbf{Q}_i$ or at least its restriction to a low dimensional linear space. As discussed before, this specification is a big difficulty in the EKF. Here it is less severe because one does not have to specify the whole matrix $\mathbf{Q}_i$ but only its restriction. The version with forgetting factor has been developed as a compromise, since it can handle the presence of dynamical noise to some extent. It is the same as the first version but with a 'forgetting' factor which downweigh past observation. Its appeal is its simplicity and low cost.

Our filter possesses the interesting property of being able to correct against bad initialization: the sequence $\mathbf{P}_i^{\mathrm{a}}$ is in the long run little affected by the initial $\mathbf{P}_0^{\mathrm{a}}$. Nevertheless, in high dimensional systems, the choice of $\mathbf{P}_0^{\mathrm{a}}$

remains important: the time taken by the filter to correct the effect of bad initialization could prove long, which would then hamper its capacity to track the evolution of the system dynamics. Taking this into consideration, we also provide an initialization method based on the EOFs.

For quick reference, we provide here a summary of our filter.

(i) Initialization. Choose an initial analysis state $\boldsymbol{x}^{\mathrm{a}}(t_0)$ and a low rank error covariance matrix of the form $\mathbf{L}_0\mathbf{U}_0\mathbf{L}_0^{\mathrm{T}}$.

(ii) Forecasting. $\boldsymbol{x}^{\mathrm{f}}(t_i)=\mathbf{M}(t_{i-1},t_i)\boldsymbol{x}^{\mathrm{a}}(t_{i-1})$ and $\mathbf{L}_i=\mathbf{M}_{i-1,i}\mathbf{L}_{i-1}$.

(iii) Correction. Compute $\mathbf{U}_i$ either by:

$$\boldsymbol{U}_i^{-1}=\rho\mathbf{U}_{i-1}^{-1}+\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i$$

with $\rho=1$ in the no dynamic noise version and $0\le\rho<1$ in the version with forgetting factor, or by

$$\mathbf{U}_i^{-1}=\left[\mathbf{U}_{i-1}+\left(\mathbf{L}_i^{\mathrm{T}}\mathbf{L}_i\right)^{-1}\mathbf{L}_i^{\mathrm{T}}\mathbf{Q}_i\mathbf{L}_i\left(\mathbf{L}_i^{\mathrm{T}}\mathbf{L}_i\right)^{-1}\right]^{-1}+\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i$$

in the general version. Then compute $\boldsymbol{x}^{\mathrm{a}}(t_i)=\boldsymbol{x}^{\mathrm{f}}(t_i)+\mathbf{L}_i\mathbf{U}_i\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}[\boldsymbol{y}_i^{\mathrm{o}}-\mathbf{H}_i\boldsymbol{x}^{\mathrm{f}}(t_i)]$.

(iv) Renormalization. One can change $\mathbf{L}_i$, $\mathbf{U}_i$ to $\mathbf{T}\mathbf{L}_i$, $(\mathbf{T}^{\mathrm{T}})^{-1}\mathbf{U}_i\mathbf{T}^{-1}$ without changing the algorithm. This should be done periodically to avoid the column of $\mathbf{L}_i$ from becoming large and nearly parallel each to the others and $\mathbf{U}_i$ becoming ill conditioned as a result. Practically, one can take $\mathbf{T}$ to be the Cholesly factor of $\mathbf{U}_i^{-1}$ so as to change $\mathbf{U}_i$ to the identity matrix.

3.1. No dynamic noise version

3.1.1. Filter description

This version is a particular case of the EKF with $\mathbf{Q}_i=0$ and a starting error covariance matrix $\mathbf{P}_0^{\mathrm{a}}$ of low rank. In this case, the Riccati Eq. (2.5) simplifies to:

$$\mathbf{P}_i^{\mathrm{a}}=\mathbf{M}_{i-1,i}\left[\mathbf{P}_{i-1}^{\mathrm{a}}-\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\left(\mathbf{H}_i\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\mathbf{M}_{i-1,i}^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}+\mathbf{R}_i\right)^{-1}\mathbf{H}_i\mathbf{M}_{i-1,i}\mathbf{P}_{i-1}^{\mathrm{a}}\right]\mathbf{M}_{i-1,i}.$$

Let us factorize $\mathbf{P}_0^{\mathrm{a}}$ into $\mathbf{L}_0\mathbf{U}_0\mathbf{L}_0^{\mathrm{T}}$ where $\mathbf{U}_0$ is a positive definite matrix with dimension equal to the rank of $\mathbf{P}_0^{\mathrm{a}}$ (this being low in practical applications). Then:

$$\mathbf{P}_1^{\mathrm{a}}=\mathbf{L}_1\left[\mathbf{U}_0-\mathbf{U}_0\mathbf{L}_1^{\mathrm{T}}\mathbf{H}_1^{\mathrm{T}}\left(\mathbf{H}_1\mathbf{L}_1\mathbf{U}_0\mathbf{L}_1^{\mathrm{T}}\mathbf{H}_1^{\mathrm{T}}+\mathbf{R}_1\right)^{-1}\mathbf{H}_1\mathbf{L}_1\mathbf{U}_0\right]\mathbf{L}_1^{\mathrm{T}},$$

where $\mathbf{L}_1=\mathbf{M}_{0,1}\mathbf{L}_0$. Denoting the expression inside the above bracket by $\mathbf{U}_1$ and repeating the same computation, one gets

$$\mathbf{L}_i=\mathbf{M}_{i-1,i}\mathbf{L}_{i-1},$$

$$\mathbf{P}_i^{\mathrm{a}}=\mathbf{L}_i\mathbf{U}_i\mathbf{L}_i^{\mathrm{T}},$$

$$\mathbf{U}_i=\mathbf{U}_{i-1}-\mathbf{U}_{i-1}\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\left(\mathbf{H}_i\mathbf{L}_i\mathbf{U}_{i-1}\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}+\mathbf{R}_i\right)^{-1}\mathbf{H}_i\mathbf{L}_i\mathbf{U}_{i-1}. \tag{3.1}$$

In the case where $\mathbf{R}_i$ is invertible, we can derive a more interesting updating equation for $\mathbf{U}_i$, which shows that the $\mathbf{U}_i$ are invertible and hence $\mathbf{P}_i^{\mathrm{a}}$ all have the same rank. By Eq. (3.1)

$$\mathbf{U}_i\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i=\mathbf{U}_{i-1}\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\left(\mathbf{H}_i\mathbf{L}_i\mathbf{U}_{i-1}\mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}+\mathbf{R}_i\right)^{-1}\mathbf{H}_i\mathbf{L}_i.$$

If $\mathbf{U}_{i-1}$ invertible, the last expression, by Eq. (3.1) is no other that $\mathbf{I} - \mathbf{U}_i\mathbf{U}_{i-1}^{-1}$. Hence $\mathbf{U}_i$ is also invertible with inverse given by

$$\mathbf{U}_i^{-1} = \mathbf{U}_{i-1}^{-1} + \mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i. \tag{3.2}$$

Since $\mathbf{U}_0$ is invertible, we deduce that $\mathbf{U}_i$ is and Eq. (3.2) holds, for all i.

The filter gain $\mathbf{K}_i$ is computed from Eqs. (2.7), (3.1) and (3.2) and the filter equation is then given by Eq. (2.6). From these equations, one sees that corrections are made parallel to the space spanned by the columns of $\mathbf{L}_i$.

3.1.2. Stability consideration and interpretation

We consider here the near linear autonomous case where $\mathbf{M}_{i-1,i}$, $\mathbf{H}_i$, $\mathbf{Q}_i$, $\mathbf{R}_i$ can all be considered constant [1] (over a certain time span) and we thus drop the subscript i where appropriate. One can then recognize the iteration $\mathbf{L}_i = \mathbf{M}\mathbf{L}_{i-1}$ as being the power method for computing the eigenvectors of $\mathbf{M}$. Indeed, in the case where $\mathbf{L}_i$ is a column vector, it is well known that it tends to align with the first eigenvector of $\mathbf{M}$, the eigenvectors being ordered according to decreasing moduli of their eigenvalues. In general when $\mathbf{L}_i$ is a matrix with r columns, the linear space spanned by them converges to the one spanned by the first r eigenvectors of $\mathbf{M}$. Indeed, let $\mathbf{V}$ and $\mathbf{V}_\ominus$ be matrices whose columns constitute a basis of the invariant subspaces associated with the first r eigenvalues of $\mathbf{M}$ and with those remaining. In order that they are well defined, we assume that the $(r+1)$-th eigenvalue is (strictly) less than the r-th in modulus. By definition, $\mathbf{M}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$, $\mathbf{M}\mathbf{V}_\ominus = \mathbf{V}_\ominus\mathbf{\Lambda}_\ominus$ for some matrices $\mathbf{\Lambda}$ and $\mathbf{\Lambda}_\ominus$ having as eigenvalues the first r eigenvalues of $\mathbf{M}$ and the remaining ones. Now decompose $\mathbf{L}_0$ into $\mathbf{V}\mathbf{N} + \mathbf{V}_\ominus\mathbf{N}_\ominus$, the column of $\mathbf{N}$ and $\mathbf{N}_\ominus$ representing the coordinate of the columns of $\mathbf{L}_0$ with respect to the above bases. Assuming that $\mathbf{N}$ (a square matrix) is invertible, one can absorbed it into $\mathbf{V}$ by changing $\mathbf{V}\mathbf{N}$ to $\mathbf{V}$ and $\mathbf{N}^{-1}\mathbf{\Lambda}\,\mathbf{N}$ to $\mathbf{\Lambda}$. Then,

$$\mathbf{L}_i\mathbf{\Lambda}^{-i} = \mathbf{V} + \mathbf{V}_\ominus\mathbf{\Lambda}_\ominus^{i}\mathbf{N}_\ominus\mathbf{\Lambda}^{-i}.$$

But $\|\mathbf{A}^i\| \le C\rho(\mathbf{A})^i$ where $\|\ \|$ is a matrix norm, C a constant and $\rho(\mathbf{A})$ is the maximum modulus of the eigenvalues of $\mathbf{A}$. Further, any matrix norm satisfies the inequality $\|\mathbf{A}\mathbf{B}\| \le \|\mathbf{A}\|\ \|\mathbf{B}\|$. Therefore $\mathbf{L}_i\mathbf{\Lambda}^{-i} \to \mathbf{V}$ as $i \to \infty$, implying the above-mentioned convergence.

Consider now the sequence of matrices $\mathbf{U}_i$, which by Eq. (3.2), can be written as

$$\mathbf{U}_i = \left[\mathbf{U}_0^{-1} + \sum_{j=1}^{i}\mathbf{L}_j^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{L}_j\right]^{-1}. \tag{3.3}$$

Therefore

$$(\mathbf{\Lambda}^{\mathrm{T}})^{-i}\mathbf{U}_i^{-1}\mathbf{\Lambda}^{-i} = (\mathbf{\Lambda}^{\mathrm{T}})^{-i}\mathbf{U}_0^{-1}\mathbf{\Lambda}^{-i} + \sum_{k=0}^{i-1}(\mathbf{\Lambda}^{\mathrm{T}})^{-k}\left(\mathbf{L}_{i-k}\mathbf{\Lambda}^{k-i}\right)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\left(\mathbf{L}_{i-k}\mathbf{\Lambda}^{k-i}\right)\mathbf{\Lambda}^{-k}.$$

Since $\lim_{i\to\infty}\mathbf{L}_{i-k}\mathbf{\Lambda}^{k-i} = \mathbf{V}$ for each k, each term in the above sum converges and can be bounded in matrix norm by a constant times $\rho(\mathbf{\Lambda}^{-1})^{2k}$. Hence, by the Lebesgue-dominated convergence theorem, if $\rho(\mathbf{\Lambda}^{-1}) < 1$, then

$$(\mathbf{\Lambda}^{\mathrm{T}})^{-i}\mathbf{U}_i^{-1}\mathbf{\Lambda}^{-i} \to \sum_{k=0}^{\infty}(\mathbf{\Lambda}^{\mathrm{T}})^{-k}\mathbf{V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{V}\mathbf{\Lambda}^{-k} = \mathbf{\Pi}^{-1}, \text{ as } i \to \infty. \tag{3.4}$$

Obviously, the notation $\mathbf{\Pi}^{-1}$ makes sense only if this matrix is invertible. This will be assumed throughout and can be seen to correspond precisely to the observability condition. It means roughly that the observation operator must provide sufficient information to determine the system state uniquely.

[1] In applications, $\mathbf{H}_i$ may vary in time in a cyclical manner, an example is in the case of oceanographic satellite observations. But a similar analysis can be done based on the consideration of a complete cycle.

The above result depends on the assumption that $\rho(\mathbf{\Lambda}^{-1}) < 1$, which is the same as that the number r^* of eigenvalues of $\mathbf{M}$ with modulus greater than 1, is at least r. In the case where $r^* < r$, we need to change our definitions of $\mathbf{V}$, $\mathbf{V}_\ominus$ and $\mathbf{\Lambda}$ somewhat. We now take $\mathbf{V}$ and $\mathbf{V}_\ominus$ to be matrices whose columns constitute a basis of the invariant spaces associated with the r^* first eigenvectors and with those remaining. We define $\mathbf{\Lambda}$ and $\mathbf{\Lambda}_\ominus$ as before. Assume that $\mathbf{M}$ admits no eigenvalue of unit modulus so that $\rho(\mathbf{\Lambda}_\ominus) < 1$. As before, we decompose $\mathbf{L}_0$ into $\mathbf{VN} + \mathbf{V}_\ominus\mathbf{N}_\ominus$. The matrix $\mathbf{N}$ is no longer square but we assume that it is of full rank. Then there is a $r \times r$ invertible matrix $\mathbf{T}$ such that the first r^* columns of $\mathbf{NT}$ form an invertible matrix while the remaining ones are zero. As before, we absorb this invertible matrix into $\mathbf{V}$, so that $\mathbf{L}_0\mathbf{T} = [\mathbf{V} + \mathbf{V}_\ominus\tilde{\mathbf{N}}\ \mathbf{V}_\ominus\tilde{\mathbf{N}}_\ominus]$ for some matrices $\tilde{\mathbf{N}}$ and $\mathbf{N}_\ominus$. Then, in a similar way as before

$$\mathbf{L}_i\mathbf{T}\begin{bmatrix}\mathbf{\Lambda}^{-i} & 0\\ 0 & \mathbf{I}\end{bmatrix} = \left[\mathbf{V} + \mathbf{\Lambda}^i_\ominus\mathbf{V}_\ominus\tilde{\mathbf{N}}_\ominus\mathbf{\Lambda}^{-i} \mid \mathbf{\Lambda}^i_\ominus\mathbf{N}_\ominus\mathbf{\Lambda}^{-i}\right],$$

$$\begin{bmatrix}\mathbf{\Lambda}^{-i} & 0\\ 0 & \mathbf{I}\end{bmatrix}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{U}_i^{-1}\mathbf{T}\begin{bmatrix}\mathbf{\Lambda}^{-i} & 0\\ 0 & \mathbf{I}\end{bmatrix} = \sum_{k=0}^{i-1}\begin{bmatrix}\mathbf{\Lambda}^{-i} & 0\\ 0 & \mathbf{I}\end{bmatrix}^{\mathrm{T}}\left(\mathbf{L}_{i-k}\mathbf{T}\begin{bmatrix}\mathbf{\Lambda}^{k-i} & 0\\ 0 & \mathbf{I}\end{bmatrix}\right)^{\mathrm{T}}$$
$$\times\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\left(\mathbf{L}_{i-k}\mathbf{T}\begin{bmatrix}\mathbf{\Lambda}^{k-i} & 0\\ 0 & \mathbf{I}\end{bmatrix}\right)\begin{bmatrix}\mathbf{\Lambda}^{k-i} & 0\\ 0 & \mathbf{I}\end{bmatrix}.$$

By a similar argument as before based on the inequality $\|\mathbf{A}^k\| \le C\rho(\mathbf{A})^k$, it can be shown that the above right hand sides converge as $i \to \infty$ to $[\mathbf{V}\ 0]$ and $\begin{bmatrix}\mathbf{\Pi}^{-i} & 0\\ 0 & *\end{bmatrix}$, respectively, $\mathbf{\Pi}$ being as before and $*$ denoting an unspecified positive definite matrix.

Thus, in all cases, $\mathbf{P}_i^{\mathrm{a}} \to \mathbf{V\Pi V}^{\mathrm{T}}$ as $i \to \infty$. This limit is of rank $\min(r, r*)$ and is independent of the initial $\mathbf{P}_0^{\mathrm{a}}$ as long as it is of the rank of at least $r*$ and certain generic assumptions (such as $\mathbf{N}$ of full rank…) are met. Hence by Eq. (2.7), $\mathbf{K}_i \to \mathbf{K} = \mathbf{V\Pi V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$.

We now show that the matrix $\mathbf{M} - \mathbf{KHM}$ is stable if and only if $r \ge r*$. Indeed

$$(\mathbf{M} - \mathbf{KHM})\mathbf{V} = \mathbf{V\Lambda} - \mathbf{V\Pi V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HV\Lambda} = \mathbf{V\Pi}(\mathbf{\Pi}^{-1} - \mathbf{V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HV})\,\mathbf{\Lambda}.$$

But the matrix $\mathbf{\Pi}^{-1}$, from its definition, Eq. (3.4), can be seen to satisfy the (Lyapounov) equation: $(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1}\mathbf{\Lambda}^{-1} + \mathbf{V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HV} = \mathbf{\Pi}^{-1}$. It follows that $(\mathbf{M} - \mathbf{KHM})\mathbf{V} = \mathbf{V\Pi}(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1}$. This yields

$$(\mathbf{M} - \mathbf{KHM})[\mathbf{V}\ \mathbf{V}_\ominus] = [\mathbf{V}\ \mathbf{V}_\ominus]\begin{bmatrix}\mathbf{\Pi}(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1} & \mathbf{\Pi V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{HV}_\ominus\mathbf{\Lambda}_\ominus\\ 0 & \mathbf{\Lambda}_\ominus\end{bmatrix}. \tag{3.5}$$

As the last matrix in Eq. (3.5) is bloc triangular, it has the same eigenvalues as those of $\mathbf{\Pi}(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1}$ and $\mathbf{\Lambda}_\ominus$. But the first matrix has the same eigenvalues as those of $\mathbf{\Lambda}^{-1}$ and by construction, $\mathbf{\Lambda}$ and $\mathbf{\Lambda}_\ominus$ have as eigenvalues the first $\min(r, r^*)$ eigenvalues of $\mathbf{M}$ and the remaining ones. This yielded the announced result.

An interpretation of how our filter works can now be provided. Suppose that we have committed an error $\mathbf{e}$ in the analysis state vector $\boldsymbol{x}^{\mathrm{a}}(t_i)$. If no further correction is made, then at time t_{i+k}, it will become $\mathbf{M}^k\mathbf{e}$ (since no dynamic noise enters the system). Thus if the error lies on the space spanned by the columns of $\mathbf{V}_\ominus$, that is $\mathbf{e} = \mathbf{V}_\ominus\mathbf{f}$ for a certain $\mathbf{f}$, then it dies out automatically, since $\mathbf{M}^k\mathbf{V}_\ominus = \mathbf{V}_\ominus\mathbf{\Lambda}^k_\ominus$ and all the eigenvalues of $\mathbf{\Lambda}_\ominus$ have modulus less than 1. On the other hand, if the error lies on the space spanned by the columns of $\mathbf{V}$, then it will diverge since $\mathbf{M}^k\mathbf{V} = \mathbf{V\Lambda}^k$ and all the eigenvalues of $\mathbf{\Lambda}$ have modulus greater than 1. It is therefore intuitively clear that all efforts should be concentrated on correcting the latter type of error and there is no need to correct the former type. Our filter operates precisely on this principle. To see this, consider (for simplicity) the limiting case where $\mathbf{K}_i$ has converged to $\mathbf{K}$. By Eq. (2.8), a filter error will be transmitted through the multiplication with $\mathbf{M} - \mathbf{KHM}$. If the error is of the form $\mathbf{Vf}$, it is transformed into $\mathbf{V\Pi}(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1}\mathbf{f}$ by Eq. (3.5) and is thus attenuated since the matrix $\mathbf{\Pi}(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1}$ have the same eigenvalues as those of $\mathbf{\Lambda}^{-1}$.

Moreover, the error remains of the same form. If the error is of the form $\mathbf{V}_\ominus \mathbf{f}$, it is simply attenuated by the system dynamics, the correction even creating an additional error lying on the space spanned by the columns of $\mathbf{V}$ (because $\mathbf{K}$ contains the factor $\mathbf{V}$). But this error will be attenuated next time. Finally, the observational errors $\boldsymbol{\epsilon}_i$ enter the system only through $\mathbf{K}$, hence lie on the space spanned by the columns of $\mathbf{V}$ and will thus be also attenuated.

3.2. General version

3.2.1. Filter description

In the case where dynamic noise is present, it can be seen from Eq. (2.5) that $\mathbf{P}_i^{\mathrm{a}}$ will not, in general, have low rank even if one starts the algorithm with a low rank one. Therefore, we shall abandon the EKF and content ourselves with a stable filter. In order that it is so, it is intuitively clear that one must make corrections along all directions for which the noise is not attenuated by the system. However, for non-linear system, instability may arise from the linearization error, so it is prudent to make corrections along all directions for which the noise is not strongly attenuated by the system. This principle could yield a filter with good performance (though not optimal) since the directions for which the error is amplified are likely those where it is large.

Proceeding on this principle, we consider as before the near linear autonomous case so that $\mathbf{M}_{i-1,i}$, $\mathbf{Q}_i$, $\mathbf{H}_i$ and $\mathbf{R}_i$ can be regarded as independent of i (over a certain time span) and we shall drop this subscript where appropriate. Let $\mathbf{V}$ be a matrix whose columns constitute a basis of the invariant space associated with the first r eigenvectors of $\mathbf{M}$, where r is chosen such that the remaining eigenvalues have small modulus, less than 1 in any case. We consider the change of basis

$$\boldsymbol{x} = [\mathbf{V}\ \mathbf{V}_\ominus]\begin{bmatrix}\mathbf{z}\\ \mathbf{z}_\ominus\end{bmatrix}, \qquad \begin{bmatrix}\mathbf{z}\\ \mathbf{z}_\ominus\end{bmatrix} = \begin{bmatrix}\mathbf{W}\\ \mathbf{W}_\ominus\end{bmatrix}\boldsymbol{x},$$

where $\mathbf{W} = (\mathbf{V}^{\mathrm{T}}\mathbf{V})^{-1}\mathbf{V}^{\mathrm{T}}$ and $\mathbf{V}_\ominus$ and $\mathbf{W}_\ominus$ are chosen such that above matrices (of basis change) are inverse to each other. (This is always possible since one already has $\mathbf{WV} = \mathbf{I}$.) This change of basis will be applied simultaneously to $\boldsymbol{x}^{\mathrm{a}}$, $\boldsymbol{x}^{\mathrm{f}}$ and $\boldsymbol{x}^{\mathrm{t}}$. Thus, $\boldsymbol{x}^{\mathrm{f}}(t_i) = \mathbf{V}\mathbf{z}^{\mathrm{f}}(t_i) + \mathbf{V}_\ominus \mathbf{z}_\ominus^{\mathrm{f}}(t_i)$ where $\mathbf{z}^{\mathrm{f}}(t_i) = \mathbf{W}\boldsymbol{x}^{\mathrm{f}}(t_i)$ and $\mathbf{z}_\ominus^{\mathrm{f}}(t_i) = \mathbf{W}_\ominus \boldsymbol{x}^{\mathrm{f}}(t_i)$. The idea is to perform correction only on $\mathbf{z}^{\mathrm{f}}(t_i)$, that is, to take as correction equations:

$$\mathbf{z}^{\mathrm{a}}(t_i) = \mathbf{z}^{\mathrm{f}}(t_i) + \tilde{\mathbf{K}}_i\left[\boldsymbol{y}_i^{\mathrm{o}} - \mathbf{H}_i \boldsymbol{x}^{\mathrm{f}}(t_i)\right], \qquad \mathbf{z}_\ominus^{\mathrm{a}}(t_i) = \mathbf{z}_\ominus^{\mathrm{f}}(t_i),$$

where $\tilde{\mathbf{K}}_i$ is a gain matrix to be defined. Since $\boldsymbol{x}^{\mathrm{a}}(t_i) = \mathbf{V}\mathbf{z}^{\mathrm{a}}(t_i) + \mathbf{V}_\ominus \mathbf{z}_\ominus^{\mathrm{a}}(t_i)$, the above equations reduce to Eq. (2.6) with $\mathbf{K}_i = \mathbf{V}_i$.

Note that

$$\begin{bmatrix}\mathbf{W}\\ \mathbf{W}_\ominus\end{bmatrix}\mathbf{M}[\mathbf{V}\ \mathbf{V}_\ominus] = \begin{bmatrix}\boldsymbol{\Lambda} & \mathbf{WMV}_\ominus\\ 0 & \boldsymbol{\Lambda}_\ominus\end{bmatrix}, \qquad \boldsymbol{\Lambda}_\ominus = \mathbf{W}_\ominus \mathbf{M}\mathbf{V}_\ominus,$$

which shows that the first r eigenvalues of $\mathbf{M}$ are those of $\boldsymbol{\Lambda}$ and the remaining ones are those of $\boldsymbol{\Lambda}_\ominus$. Also from this equality, Eq. (2.1′) can be written as

$$\mathbf{z}^{\mathrm{t}}(t_i) \approx \mathbf{z}^{\mathrm{f}}(t_i) + \boldsymbol{\Lambda}\left[\mathbf{z}^{\mathrm{t}}(t_{i-1}) - \mathbf{z}^{\mathrm{a}}(t_{i-1})\right] + \mathbf{WMV}_\ominus\left[\mathbf{z}_\ominus^{\mathrm{t}}(t_{i-1}) - \mathbf{z}_\ominus^{\mathrm{a}}(t_{i-1})\right] + \mathbf{W}\eta(t_i) \tag{3.6}$$

$$\mathbf{z}_\ominus^{\mathrm{t}}(t_i) \approx \mathbf{z}_\ominus^{\mathrm{f}}(t_i) + \boldsymbol{\Lambda}_\ominus\left[\mathbf{z}_\ominus^{\mathrm{t}}(t_{i-1}) - \mathbf{z}_\ominus^{\mathrm{a}}(t_{i-1})\right] + \mathbf{W}_\ominus \eta(t_i).$$

The last equation, together with $\mathbf{z}_\ominus^{\mathrm{a}}(t_i) = \mathbf{z}_\ominus^{\mathrm{f}}(t_i)$, yield the error propagation equation for $\mathbf{z}_\ominus^{\mathrm{t}}$: $\mathbf{z}_\ominus^{\mathrm{a}}(t_i) - \mathbf{z}_\ominus^{\mathrm{t}}(t_i) = \boldsymbol{\Lambda}_\ominus[\mathbf{z}_\ominus^{\mathrm{a}}(t_{i-1}) - \mathbf{z}_\ominus^{\mathrm{t}}(t_{i-1})] - \mathbf{W}_\ominus \eta(t_i)$. Because of the attenuation effect of $\boldsymbol{\Lambda}_\ominus$, the error $\mathbf{z}_\ominus^{\mathrm{a}} - \mathbf{z}_\ominus^{\mathrm{t}}$ will be small (but will not die out, unless $\mathbf{W}_\ominus \eta(t_i) \equiv 0$). Ignoring this error, Eq. (3.6) has the same form as Eq. (2.1) in

the EKF, with $\mathbf{z}$ playing the role of $\boldsymbol{x}$, $\boldsymbol{\Lambda}$ the role of $\mathbf{M}$ and $\mathbf{W}\boldsymbol{\eta}$ the role of $\boldsymbol{\eta}$. Note that the linearized observation Eq. (2.2) can be written as

$$\mathbf{y}_i^o \approx \mathbf{H}_i \boldsymbol{x}^f(t_i) + \mathbf{HV}\left[\mathbf{z}^t(t_i) - \mathbf{z}^f(t_i)\right] + \mathbf{HV}_\ominus\left[\mathbf{z}_\ominus^t(t_i) - \mathbf{z}_\ominus^f(t_i)\right] + \boldsymbol{\epsilon}_i,$$

which can be considered to be the 'observation equation for $\mathbf{z}^t$' if one regards $\mathbf{z}_\ominus^t$ as equal to $\mathbf{z}_\ominus^f$ (that is ignoring the error $\mathbf{z}_\ominus^t - \mathbf{z}_\ominus^f$). These considerations suggest our taking $\mathbf{K}_i$ as the Kalman gain in the filtering problem for $\mathbf{z}$, that is $\tilde{\mathbf{K}}_i$ is given by Eqs. (2.5) and (2.7) with $\mathbf{M}_{i-1,i}$ replaced by $\boldsymbol{\Lambda}$, $\mathbf{H}_i$ by $\mathbf{HV}$, $\mathbf{R}_i$ by $\mathbf{R}$ and $\mathbf{Q}_i$ by $\mathbf{WQW}^T$. Explicitly, according to Eqs. (2.5) and (2.7), $\tilde{\mathbf{K}}_i = \boldsymbol{\Pi}_i \mathbf{V}^T \mathbf{H}^T \mathbf{R}^{-1}$, where

$$\boldsymbol{\Pi}_i = \boldsymbol{\Lambda}\boldsymbol{\Pi}_{i-1}\boldsymbol{\Lambda}^T + \mathbf{WQW}^T - \left(\boldsymbol{\Lambda}\boldsymbol{\Pi}_{i-1}\boldsymbol{\Lambda}^T + \mathbf{WQW}^T\right)\mathbf{H}^T\mathbf{V}^T$$
$$\times\left[\mathbf{HV}\left(\boldsymbol{\Lambda}\boldsymbol{\Pi}_{i-1}\boldsymbol{\Lambda}^T + \mathbf{WQW}^T\right)\mathbf{V}^T\mathbf{H}^T + \mathbf{R}\right]^{-1}\mathbf{HV}\left(\boldsymbol{\Lambda}\boldsymbol{\Pi}_{i-1}\boldsymbol{\Lambda}^T + \mathbf{WQW}^T\right)$$

To compute $\boldsymbol{\Pi}_i$, one needs $\mathbf{V}$ and $\boldsymbol{\Lambda}$, which can be obtained from $\mathbf{M}$. But our assumption $\mathbf{M}_{i-1,i} = \mathbf{M}$ was only postulated in order to simplify the arguments. In fact $\mathbf{M}_{i-1,i}$ evolves with time, albeit slowly. To take this into account, we draw our inspiration from the computations in Section 3.1.2. Define the sequences

$$\mathbf{V}_i = \mathbf{M}_{i-1,1}\mathbf{V}_{i-1}\boldsymbol{\Lambda}_i^{-1}, \qquad \boldsymbol{\Lambda}_i = \mathbf{W}_{i-1}\mathbf{M}_{i-1,i}\mathbf{V}_{i-1}, \qquad \mathbf{W}_{i-1} = \left(\mathbf{V}_{i-1}^T\mathbf{V}_{i-1}\right)^{-1}\mathbf{V}_{i-1}^T \tag{3.7}$$

starting from some $\mathbf{V}_0$. Then, if all the $\mathbf{M}_{i-1,i}$ are equal to $\mathbf{M}$, it can be shown that $\mathbf{V}_i$ converges generally to a limit $\mathbf{V}$ with columns spanning the invariant space associated with the first r eigenvectors of $\mathbf{M}$. The $\boldsymbol{\Lambda}_i$ also converge to $\boldsymbol{\Lambda}$, satisfying $\mathbf{MV} = \mathbf{V}\boldsymbol{\Lambda}$. Thus, we shall take as our filter gain the one given by Eq. (2.7) with $\mathbf{P}_i^a = \mathbf{V}_i\boldsymbol{\Pi}_i\mathbf{V}_i^T$ and $\boldsymbol{\Pi}_i$ defined as above but replacing $\mathbf{H}$, $\mathbf{V}$, $\boldsymbol{\Lambda}$, $\mathbf{W}$, $\mathbf{R}$, $\mathbf{Q}$ by $\mathbf{H}_i$, $\mathbf{V}_i$, $\boldsymbol{\Lambda}_i$, $\mathbf{W}_i$, $\mathbf{R}_i$, $\mathbf{Q}_i$. The resulting equation for $\boldsymbol{\Pi}_i$ can alternatively be written as (using the same computations in Section 3.2.1)

$$\boldsymbol{\Pi}_i^{-1} = \left(\boldsymbol{\Lambda}_i\boldsymbol{\Pi}_{i-1}\boldsymbol{\Lambda}_i^T + \mathbf{W}_i\mathbf{Q}_i\mathbf{W}_i^T\right)^{-1} + \mathbf{V}_i^T\mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{V}_i. \tag{3.8}$$

It is instructive to note the similarity of this filter and the one in Section 3.1.1. Indeed, putting

$$\mathbf{L}_i = \mathbf{V}_i\boldsymbol{\Lambda}_i\ldots\boldsymbol{\Lambda}_1, \qquad (\mathbf{L}_0 = \mathbf{V}_0)$$

and

$$\mathbf{U}_i = \boldsymbol{\Lambda}_i^{-1}\ldots\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Pi}_i\left(\boldsymbol{\Lambda}_i^T\right)^{-1}\ldots\left(\boldsymbol{\Lambda}_i^T\right)^{-1}, \qquad \mathbf{U}_0 = \boldsymbol{\Pi}_0.$$

Then clearly $\mathbf{P}_i^a = \mathbf{L}_i\mathbf{U}_i\mathbf{L}_i^T$ and Eqs. (3.7) and (3.8) can be rewritten in terms of $\mathbf{L}_i$, $\mathbf{U}_i$ as

$$\mathbf{L}_i = \mathbf{M}_{i-1,i}\mathbf{L}_{i-1},$$

$$\mathbf{U}_i^{-1} = \left[\mathbf{U}_{i-1} + \left(\mathbf{L}_i^T\mathbf{L}_i\right)^{-1}\mathbf{L}_i^T\mathbf{Q}_i\mathbf{L}_i\left(\mathbf{L}_i^T\mathbf{L}_i\right)^{-1}\right]^{-1} + \mathbf{L}_i^T\mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i. \tag{3.9}$$

The last equation is the same as Eq. (3.2) except for an extra term involving $\mathbf{Q}_i$. The first equation is also the same as the first equation of Eq. (3.1). Thus this algorithm is reduced to the one in Section 3.1.1 in the case where $\mathbf{Q}_i = 0$. However, the algorithm in term of $\mathbf{L}_i$ and $\mathbf{U}_i$ may be numerically unstable, since $\mathbf{L}_i$ and $\mathbf{U}_i^{-1}$ tend to diverge. (As is shown in Section 3.1.2, $\mathbf{L}_i\boldsymbol{\Lambda}^{-i}$ and $(\boldsymbol{\Lambda}^T)^{-i}\mathbf{U}_i^{-1}\boldsymbol{\Lambda}^{-i}$ converge, but since $\boldsymbol{\Lambda}$ has eigenvalue of modulus greater than 1, $\mathbf{L}_i$ and $\mathbf{U}_i^{-1}$ themselves diverge.) The algorithms Eqs. (3.7) and (3.8) are numerically stable since $\mathbf{V}_i$ is 'renormalized' at each step by the division by $\boldsymbol{\Lambda}_i$. If the first algorithm is used, $\mathbf{L}_i$ needs to be renormalized periodically to avoid numerical instability.

3.2.2. Stability consideration

A remarkable property of the Riccati Eq. (2.5) in the linear autonomous case (that is the case where $\mathbf{M}_{i-1,i}$, $\mathbf{H}_i$, $\mathbf{Q}_i$, $\mathbf{R}_i$ do not depend on i) is that under appropriate conditions, its solution converges to a limit such that $\mathbf{M} - \mathbf{KHM}$ is stable, $\mathbf{K}$ denoting the corresponding limit of the $\mathbf{K}_i$ as given by Eq. (2.7). Thus, consider again

the nearly linear autonomous case so that we may apply the above result to the $\mathbf{\Pi}_i$ sequence, which satisfies a similar Riccati equation. This yields that $\mathbf{\Pi}_i$ converges to $\mathbf{\Pi}$ such that the matrix

$$\mathbf{\Lambda} - \mathbf{\Pi}\mathbf{V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{V}\mathbf{\Lambda} \tag{3.10}$$

is stable. We must emphasize that such an argument does not apply to the case $\mathbf{Q}_i = 0$ considered in Section 3.1.2, since the required controllability condition is violated. Here the matrices $\mathbf{\Pi}_i$ are of a low dimension and the controllability condition holds quite generally, in fact as soon as $\mathbf{W}\mathbf{Q}\mathbf{W}^{\mathrm{T}}$ is non-singular. Therefore, the gain matrices $\mathbf{K}_i$ of our filter converge to $\mathbf{K} = \mathbf{V}\mathbf{\Pi}\mathbf{V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$. Then, a direct computation yields

$$(\mathbf{M} - \mathrm{KHM})\mathbf{V} = \mathbf{V}(\mathbf{\Lambda} - \mathbf{\Pi}\mathbf{V}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{V}\mathbf{A})$$

which leads to the same formula as Eq. (3.5) but with $\mathbf{\Pi}(\mathbf{\Lambda}^{\mathrm{T}})^{-1}\mathbf{\Pi}^{-1}$ replaced by the matrix Eq. (3.10). Since the latter matrix is stable, using the same argument as that in the end of Section 3.1.2, we obtain that our filter is indeed stable.

3.3. Version with forgetting factor

In the above derivations of the filter, we have ignored the linearization error in Eq. (2.1′). This error would add to the dynamic noise $\eta(t_i)$ in Eq. (2.1′) and thus changes the forecast error covariance matrix $\mathbf{P}^{\mathrm{f}}(t_i)$ as given by Eq. (2.3). A simple and (very rough) way to take into account this error is to increase the dynamic noise covariance matrix $\mathbf{Q}(t_i)$ which amounts to increasing $\mathbf{P}^{\mathrm{f}}(t_i)$. Note that the linearization error does not necessarily increases this matrix but may decrease it, but we argue that it is safer to err on the conservative side. Indeed, increasing $\mathbf{P}^{\mathrm{f}}(t_i)$ would lead to the filter relying more on the observed data than on the model, which is what one would naturally do when there is doubt on the accuracy of the model. The difficulty, of course, is that we do not know how and how much to increase the matrix $\mathbf{Q}(t_i)$. But as pointed out in Section 2, even the specification of $\mathbf{Q}(t_i)$ is not obvious. So the choice of $\mathbf{Q}(t_i)$ contains inevitably some part of arbitrariness based on convenience. By the same reasoning, even when there is no dynamic error, it is still prudent to add an artificial dynamic noise covariance matrix $\mathbf{Q}(t_i)$. For convenience, we propose taking $\mathbf{Q}_i$ to be such that $(\mathbf{L}_i^{\mathrm{T}}\mathbf{L}_i)^{-1}\mathbf{L}_i^{\mathrm{T}}\mathbf{Q}_i\mathbf{L}_i(\mathbf{L}_i^{\mathrm{T}}\mathbf{L}_i)^{-1} = \alpha\mathbf{U}_{i-1}$, where α is a positive constant. This choice is not standard, since $\mathbf{Q}_i$ is not explicit and dependent on i through $\mathbf{U}_{i-1}$, but since the linearization error should increase with the filter error $\boldsymbol{x}^{\mathrm{a}}(t_{i-1}) - \boldsymbol{x}^{\mathrm{t}}(_{i-1})$, this choice seems logical. The main reason for it is that it simplifies the algorithm: Eq. (3.9) becomes

$$\mathbf{U}_i^{-1} = \rho\mathbf{U}_{i-1}^{-1} + \mathbf{L}_i^{\mathrm{T}}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i. \tag{3.11}$$

where $\rho = 1/(1+\alpha)$. This results in an essentially the same filter as the no dynamic noise version in Section 3.1.1, except that the matrix $\mathbf{U}_{i-1}$ is amplified by a factor of $1/\rho$ before entering the updating equation.

The factor ρ may be interpreted as a forgetting factor. In fact, as is shown in the work of Pham et al. (1995), the filter in Section 3.1.1 can be viewed as a recursive implementation of the minimization of the least squares criterion with constraint and penalty on the initial state. If, instead of the usual sum of the squares of the errors, one takes the weighted sum of the squares with weight decreasing with the number of steps k (going back in time) as ρ^k, then the above filter is obtained. This kind of weighting is precisely what a forgetting factor produces in adaptive algorithms. The primary purpose is to down-weight the earlier observations with respect to the recent ones, thus enabling the algorithm to follow changes in the underlying process. As a beneficial side effect, it also attenuates the error propagation. This is a type of error compensation technique (Jazwinski, 1970). It actually enhances the stability of the filter. Although the arguments in Section 3.2.2 are no longer valid here ($\mathbf{Q}_i$ being not constant), those in Section 3.1.1 may be adapted to show this stability.

3.4. Initialization of the filter: the EOFs technique

To initialize the Kalman algorithm, one needs an initial analysis state vector $\boldsymbol{x}^{\mathrm{a}}(t_0)$ and its error covariance matrix $\mathbf{P}_0^{\mathrm{a}}$. A choice thus needs to be made for $\boldsymbol{x}^{\mathrm{a}}(t_0)$ and $\mathbf{P}_0^{\mathrm{a}}$. To this end, we propose to use the EOFs technique applied to observed, or possibly simulated, state sequences from the system. Simulation might prove necessary if no sufficiently long sequence of observed state vectors is available. In the no dynamic noise case, it is also quite easy to generate long sequence of state vectors from the model Eq. (2.1). The initial state may be set arbitrarily if one has taken care to wait until the model has been settled into a stable regime (from the statistical point of view) thus discarding the spin-up phase irrelevant for our purpose.

It is quite natural to take as $\boldsymbol{x}^{\mathrm{a}}(t_0)$ the average of the simulated (or observed) state vectors and as $\mathbf{P}_0^{\mathrm{a}}$ the low rank approximation of the sample covariance matrix $\mathbf{P}_0$ of these vectors. [2] To obtain such an approximation, we propose to use the EOFs approach, which provides, in a certain sense explained below, the best approximation. Let $\mathbf{V}_1, \mathbf{V}_2, \ldots$, be the eigenvectors of $\mathbf{P}_0$ of unit L^2-norm, ordered according to their eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$. These vectors are called EOFs. In the EOFs method, $\mathbf{P}_0$ approximates to $\mathbf{L}_0\mathbf{U}_0\mathbf{L}_0^{\mathrm{T}}$ where

$$\mathbf{L}_0 = [\mathbf{V}_1 \ldots \mathbf{V}_r], \qquad \mathbf{U}_0 = \mathrm{diag}(\lambda_1, \ldots, \lambda_r).$$

This is justified by the following results which state that, ''if $\boldsymbol{x}$ is a random vector of mean zero and covariance matrix $\mathbf{P}_0$, then among all its projections onto a linear subspace of dimension r, the one for which the error vector has smallest expected squared norm is the projection onto the linear space spanned by $\mathbf{V}_1, \ldots \mathbf{V}_r$'' (Seber, 1984, p. 179). It can be seen further that this projection have covariance matrix $\mathbf{L}_0\mathbf{U}_0\mathbf{L}_0^{\mathrm{T}}$ and the error vector has expected squared norm $\sum_{j>r}\lambda_j$. Since $\boldsymbol{x}$ has expected squared norm $\mathrm{Tr}(\mathbf{P}_0)$ (the trace of $\mathbf{P}_0$), the ratio $\sum_{j>r}\lambda_j/\mathrm{Tr}(\mathbf{P}_0)$ represents the relative error in square L^2-norm and can thus be used to assess the accuracy of the approximation for choosing the appropriate value for r.

To apply the SEEK filter, one also needs to specify the matrices $\mathbf{R}_i$ and $\mathbf{Q}_i$, which are generally unknown. When the dynamic noise is thought to be weak or absent, the no noise version with forgetting factor in Section 3.3 has a special appeal, since it does not require the specification of $\mathbf{Q}_i$. If $\mathbf{R}_i$ is taken as σ^2 times a constant matrix $\tilde{\mathbf{R}}$ (often an identity matrix), this σ^2 need not be known either. Indeed, from Eqs. (2.7), (3.1) and (3.11), one obtains

$$\mathbf{K}_i = \mathbf{L}_i\left(\mathbf{U}_i/\sigma^2\right)\mathbf{L}_i^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{R}}^{-1}, \qquad \sigma^2\mathbf{U}_i^{-1} = \rho\sigma^2\mathbf{U}_{i-1}^{-1} + \mathbf{L}_i\mathbf{H}_i^{\mathrm{T}}\tilde{\mathbf{R}}^{-1}\mathbf{H}_i\mathbf{L}_i.$$

This means that only the $\mathbf{U}_i/\sigma^2$ enters the computation and hence only $\mathbf{U}_0/\sigma^2$ needs to be specified. Usually σ^2 is very small with respect to $\mathbf{U}_0$. In any case, since $\boldsymbol{x}^{\mathrm{a}}(t_0)$ is chosen somewhat arbitrarily, it is safe to err on the conservative side, by taking $\mathbf{U}_0$ to be very large. Then it makes little difference putting $\sigma^2\mathbf{U}_0^{-1} = 0$.

4. An example of assimilation

As an example illustrating the application of our filter, we have considered a simple application of a QG model of ocean circulation, commonly used to describe the dynamical evolution of geophysical fluids such as the atmosphere or the ocean (Pedlosky, 1987). In the present case, it is considered in its layered formulation (Holland, 1978). The ocean is thus subdivided into N layers of constant density. Vertical momentum is transferred vertically by the pressure coupling between layers. Horizontal circulation in each layer is represented

[2] The rank of $\mathbf{P}_0$ is the number of generated vectors minus 1 and thus can (and should) be rather high, although it is usually far less than the system dimension.

by N horizontal streamfunction fields $\psi_1, \ldots, \psi_N$. This model has been extensively used in our group for process studies or more realistic application problems. Although these models are less complete than the fully diabatic primitive equation ones, they are known to be nicely adapted to representing fairly faithfully the basic features of some oceanic processes such as mid-latitude ocean eddy interactions.

4.1. Simulation setup

In the mid-latitudes the ocean circulation is often characterized by very turbulent activity arising from strong interactions between the so-called mesoscale eddies. The prototype region of this mesoscale activity is the Gulf Stream system where the barotropic and baroclinic instability of the intense inertial currents produces strong mesoscale eddies. The energy present in these eddies is at least of the order of the mean current energy and often much greater. Other western boundary ocean current systems in the world also exhibit such behavior. Turbulent eddy activity is also encountered in the quieter areas of the ocean universe such as the interior regions of ocean basins, although with a lesser energy level.

In the present case, a very schematic box-model of this turbulent eddy activity was considered. The domain is a simple square ocean 4323–340 km wide. A schematic wind stress pattern forces a circulation pattern made up of a number of mesoscale eddies having in strong non-linear interactions. Dissipation occurs through bottom friction over the ocean floor and lateral friction through subgrid scale processes.

The numerical domain was divided into a 41×41 grid. The stratification was assumed to be three-layered. The QG equations were solved with an integration step of 22 500 s, with six such steps (the first being Euler forward, the others leap-frog) between each observation, yielding a time interval of 37.5 h between observations.

Observations are supposed to be of the altimetric type, i.e., observations of the instantaneous dynamical topography. This sea-surface height was directly proportional to the surface streamfunction which is a prognostic variable of the model. By definition, only the first layer is observed, but the whole state vector, all the three layers, was of course simulated. One challenge raised by the assimilation process, in addition to that of coping with the non-linearity, is to succeed in reconstructing flow evolution at great depth while disposing only of surface information.

The approach chosen is that of the so-called ‘twin experiments’. A reference experiment is performed and the reference fields will be compared to the fields produced during the assimilation experiment. The latter is performed using the ‘pseudo-observations’ which are extracted (at the surface only) from the reference experiment. The assimilation will be successful if the flow converges with time towards the reference situation especially in flow at depth.

This test case clearly has many drawbacks due to the relatively simple model formulation together with the use of synthetic data and the simplistic domain configuration. It should be therefore regarded as a simple test case which has to be proved satisfactory before more realistic problems are even contemplated.

4.2. Specification of initial state and error covariance matrix

Following the strategy explained in Section 3.4, the choice of the initial flow field and the corresponding error covariance matrix is made through the model simulation itself. Note that this was done only once and the results can be (and have been) filed for use in different assimilation experiments. In the present study, data for the assimilation experiments is again simulated (i.e., synthetic data) but in a way unrelated to that in the above simulation.

To generate the flow fields we always consider a statistically steady regime in which the flow has been integrated for a sufficiently long time in order to avoid the transitional spin-up phase. In the present experiment, a long sequence of 8000 state vectors was generated of which the first 4000 vectors were discarded to avoid

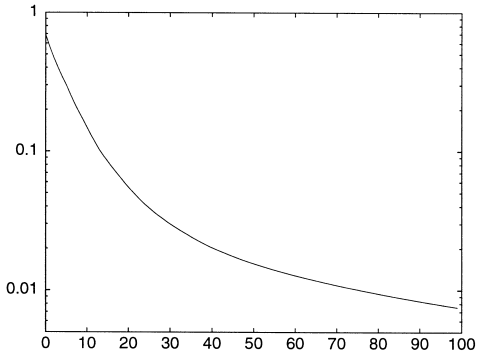

Fig. 1. Relative error in square L^2-norm vs. the number of retained EOFs.

such transient effects. The remaining sequence of 4000 state vectors (6250 days) was reduced into a sequence of 1000 by retaining only one vector out of four. This operation, known as decimation, was to reduce the calculation since successive states are quite similar. From this decimated sequence, we estimate the covariance matrix of the state vector and perform an EOFs analysis. Fig. 1 plots the relative error in square L^2 norm (as defined in Section 3.4) vs. the number of retained EOFs. It shows that the covariance matrix of $\boldsymbol{x}^t(t_0)$ is well approximated by a matrix of much lower rank, since the above relative error decreases rapidly with the number of retained EOFs. For 51 retained EOFs, it is already down to 1.55%. However, it decreases much slower in the range of higher numbers of retained EOFs (for 80 EOFs it is still 0.96%).

4.3. Results of assimilation experiments

We again generated a sequence of reference flow fields starting from an initial field totally unconnected to that used in the simulation of the above section. Practically, we take as initial field, a computed field at a random time far removed from the time segment which was used in this simulation. Sequences of 200 state vectors, corresponding to a period of 312.5 days, were thus obtained. The first layer was taken as the (pseudo)-observation and was inputted into our filter algorithm (in some experiments, an observational noise was also added). The other layers were used as reference fields to evaluate the performance of the algorithm.

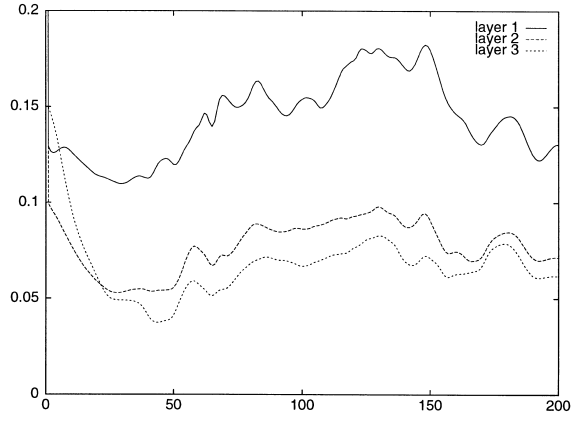

Fig. 2. Relative assimilation errors for rank 51 with no forgetting factor.

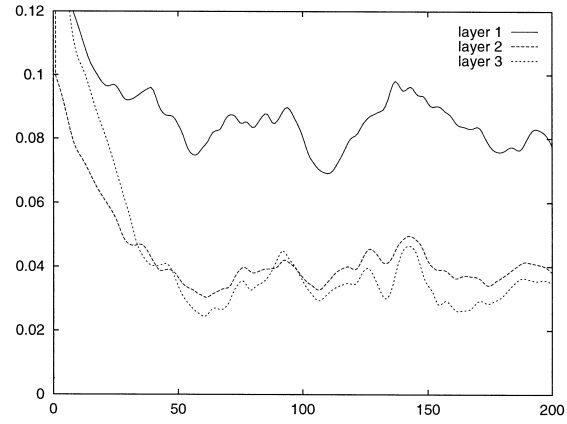

Fig. 3. Relative assimilation errors for rank 51 with forgetting factor 0.5.

Only the version with no dynamic noise was used (there is no such noise in our simulation either), but possibly with a forgetting factor.

Fig. 2 plots the assimilation results using the SEEK filter of rank 51 with no forgetting factor. It can be seen that the relative errors in L^2-norm (not squared), layer by layer, decrease very rapidly for the first few steps (the initial errors are 95, 99 and 101% which correspond to points far outside Fig. 2). However, after several tens of steps, the errors start to increase again. This seems to result from to the fact that the filter under-evaluates the error and does not therefore make sufficient correction. Also, using a SEEK filter of only rank 51 induces an intrinsic error limiting the filter's performance.

In a second assimilation experiment we used the same setup as before except that the filter now had a forgetting factor of 0.5. The results are plotted in Fig. 3. It can be seen that the introduction of the forgetting factor enhances performance by a factor which is roughly two.

The chosen rank of the initial error covariance matrix has a direct effect on the performance of the algorithm. This is illustrated in Fig. 4, where the setup is the same as in Fig. 3 but the rank is increased to 80. One can see that the algorithm performs much better. Improvement in performance can also be achieved by a better initial state estimate. In Fig. 5, we have deliberately started the algorithm at a point $\boldsymbol{x}^{\mathrm{a}}(t_0)$ much closer to the

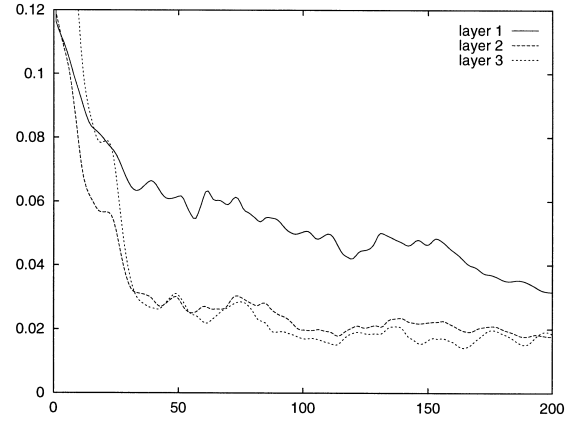

Fig. 4. Relative assimilation errors for rank 80 with forgetting factor 0.5.

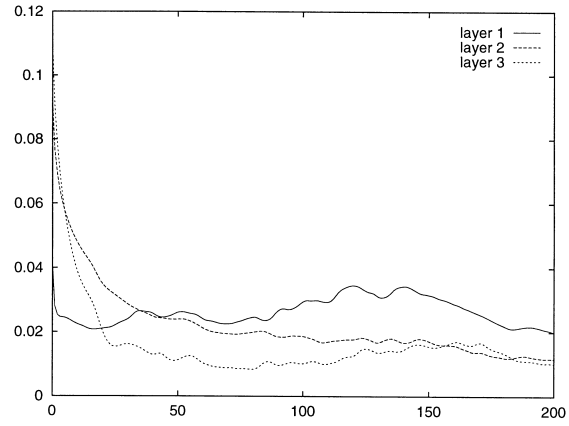

Fig. 5. Relative assimilation errors for rank 80 with forgetting factor 0.5, special initial estimate.

reference (the relative errors are 4.2, 9.2 and 10.9%). It can be seen from Fig. 5 that the error is significantly lower. However, this gain of performance is attenuated over time and is not quite as appreciable near the end of the assimilation period, thus suggesting the relatively weak impact of the choice of the initial state.

Note that in the above simulation experiments, observations were simply taken as being components of the state vector corresponding to the first layer, no noise being added (whereas in the KF model an observational noise is assumed). But it would seem that observational noise has no appreciable effect on the performance of the filter. This is illustrated in Fig. 6, which corresponds to the same situation as that of Fig. 4, except that the observation is corrupted by a (spatially) white noise of standard deviation 30 000 (this corresponds to roughly 10% of the average standard deviation of the state variables). It can be seen that the curves in Fig. 6 are very similar to those in Fig. 4. The latter are in fact slightly lower for most (but not all) of the times.

Looking at the above results, one might be surprised that the error in the first layer stays somewhat higher than that in the others, while only this layer has been observed (without error in some experiments!). This is likely due to the higher level of turbulence which is observed at the surface, due to the role of the stratification. Further, in our filter, correction was made for the whole ocean domain along a limited number of directions, hence some local accuracy on the first layer may be sacrificed to achieve a global accuracy.

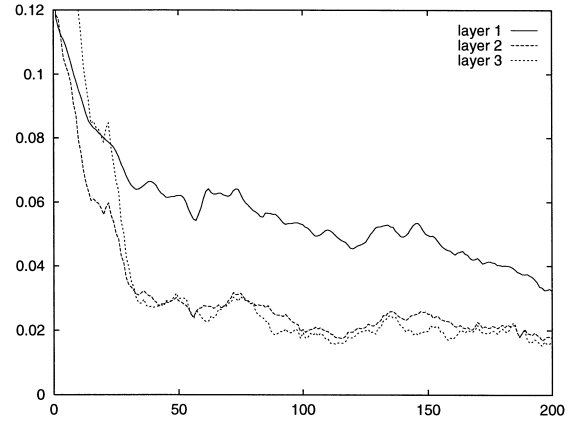

Fig. 6. Relative assimilation errors for rank 80 with forgetting factor 0.5, observational noise 30 000.

5. Conclusion

A modified EKF for the assimilation of oceanographic data into numerical models has been presented. A pilot implementation in a test case using altimeter (pseudo)-data from a strongly non-linear QG model is shown, which preliminary proves its feasibility. The basic feature for this SEEK filter is the use of a low rank approximation to the error covariance matrix. This results in the introduction of a reduced size basis of statistical functions on which one relies in order to describe the evolution of the error covariance matrix. The basis evolves in time following the model, as a consequence of its non-linearity. On the one hand, the gradient of the system transition operator depends on the system state so that the error propagation depends on the actual state of the ocean. On the other hand, the reduced statistical basis was designed specifically to concentrate on the correction on those directions for which the error is amplified. Although EOFs analysis has been found to be a convenient and fairly effective way of initializing the reduced basis, this technique, due to its statistical nature, could only provide a good basis 'on the average'. Evolution with time is therefore needed in order to adapt the basis to the particular state of the ocean under analysis at that time. We believe this evolutive property of our basis is crucial in keeping its dimension low. Fixed basis indeed has been considered in the literature, but since it remains the same for possibly quite different ocean states, it is likely that a much higher number of basis functions would be required. Finally, the introduction of a memory effect ('forgetting factor') was also found to be quite beneficial and it greatly improves the performance of the algorithm.

The chosen test situation was that of an ocean box-model within which several mesoscale eddies were interacting in an intrinsically active turbulent process. In such testing conditions, our SEEK filter was found to be fairly effective in monitoring the flow state and evolution disposing of surface-only pseudo-altimeter data. Further work will consider a more realistic situation, both from the model point of view (real coastal geometry, bathymetry, real wind forcing) and the observational one (using real data from the Topex/Poseidon satellite). However, this preliminary application was a necessary step before realistic applications and it provided us with encouraging results regarding that purpose. We are currently working to implement this filter in a model of primitive equations for the tropical Pacific, with very promising preliminary results (Pham et al., 1998)

Acknowledgements

This work was performed within the framework of the IDOPT project supported by INRIA/INPG/UJF/CNRS. The authors would like to thank the project leader J. Blum, with whom they have had many fruitful discussions, and all reviewers who have carefully read the manuscript and made many constructive comments.

References

Cane, M.A., Patton, R.J., 1984. A numerical model for low-frequency equatorial dynamics. J. Phys. Oceanogr. 14, 1853–1863.

Cane, M.A., Kaplan, A., Miller, R.N., Tang, B., Hackert, E.C., Busalacchi, A.J., 1996. Mapping tropical Pacific sea level: data assimilation via a reduced state Kalman filter. J. Geophys. Res. 101 (C10), 549–617.

Cohn, S.E., Todling, R., 1995. Approximate Kalman filters for unstable dynamics. Second International Symposium on Assimilation of Observations in Meterology and Oceanography, WMO, pp. 241–246.

Cohn, S.E., Todling, R., 1996. Approximate data assimilation scheme for stable and unstable dynamics. J. Meteorol. Soc. Jpn. 74 (1), 63–75.

Dee, D.P., 1991. Simplification of the Kalman filter for meteorological data assimilation. Q. J. R. Meteorol. Soc. 117, 365–384.

Evensen, G., 1992. Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. J. Geophys. Res. 97, 17905–17924.

Evensen, G., 1993. Open boundary conditions for the extended Kalman filter with a quasi-geostrophic ocean model. J. Geophys. Res. 98 (C9), 16529–16546.

Evensen, G., 1994. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99 (C5), 10143–10162.

Fu, L.L., Fukumori, I., Miller, R.N., 1993. Fitting dynamic models to the Geosat sea level observations in the tropical Pacific ocean: II. A linear, wind-driven model. J. Phys. Oceanogr. 23, 2162–2181.

Fukumori, I., 1995. Assimilation of Topex sea level measurements with a reduced-gravity, shallow water model of the tropical Pacific ocean. J. Geophys. Res. 100 (C12), 25027–25039.

Fukumori, I., Malanotte-Rizzoli, P., 1995. An approximate Kalman filter for ocean data assimilation: an example with an idealized Gulf Stream model. J. Geophys. Res. 100 (C12), 6777–6793.

Fukumori, I., Benveniste, J., Wunsch, C., Haidvogel, D.B., 1993. Assimilation of sea surface topography into an ocean circulation model using a steady state smoother. J. Phys. Oceanogr. 23, 1831–1855.

Gaspar, P., Wunsch, C., 1989. Estimates from altimeter data of barotropic Rossby waves in the northwestern Atlantic ocean. J. Phys. Oceanogr. 19, 1821–1844.

Gauthier, P., Courtier, P., Moll, P., 1993. Assimilation of simulated wind lidar data with a Kalman filter. Mon. Wea. Rev. 121, 1803–1820.

Ghil, M., 1986. Sequential estimation and satellite data assimilation in meteorology and oceanography. In: Sasaki, Y.K., Gat-Chen, T., Whie, L., Zaman, M.M., Ziegler, C., Chang, L.P., Rusk, D.J. (Eds.), Variational Methods in Geosciences. Elsevier, Amsterdam, pp. 91–100.

Ghil, M., Manalotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography. Adv. Geophys. 23, 141–265.

Ghil, M., Cohn, S.E., Dalcher, A., 1982. Sequential estimation, data assimilation and initialization. In: Williamson, D. (Ed.), The Interaction Between Objective Analysis and Initialization, Publ. Meteorol. 127, Proc. 14th Standstead Seminar, McGill University, Montreal, pp. 83–97.

Gourdeau, L., Arnault, S., Menard, Y., Merle, J., 1992. Geosat sea-level assimilation in a tropical Atlantic model using Kalman filter. Oceanol. Acta 15, 567–574.

Hager, W., Horowitz, L.L., 1976. Convergence and stability property of the discrete Riccati operator equation and the associated optimal control and filtering problems. Siam J. Contol Optimization 14 (2), 295–312.

Hoang, H.S., De Mey, P., Talagrand, O., Baraille, R. 1995. Assimilation of altimeter data in multilayer quasi-geostrophic model by simple nonlinear adaptive filter. Proc. Int. Symp. Assimilation Obser. Meteorol. Oceanogr., Tokyo, Japan, pp. 521–526.

Holland, W.R., 1978. The role of mesoscale eddies in the general circulation of the ocean: numerical experiments using a wind-driven quasi-geostrophic model. J. Phys. Oceanogr. 8, 363–392.

Ide, K., Bennett, A.F., Courtier, P., Ghil, M., Lorenc, A.C., 1997. Unified notation for data assimilation: operational, sequential and variational. J. Meteorol. Soc. Jpn. 75 (1B), 181–189.

Jazwinski, A.A., 1970. Stochastic and Filtering Theory. Mathematics in Sciences and Engineering Series, Series 64. Academic Press, New York.

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

Miller, R.N., Cane, M., 1989. A Kalman filter analysis of sea level height in the tropical Pacific. J. Phys. Oceanogr. 19, 773–790.

Miller, R.N., Ghil, M., Gauthiez, F., 1994. Advanced data assimilation in strongly non-linear dynamical systems. J. Atmos. Sci. 51, 1037–1056.

Pham, D.T., Verron, J., Roubaud, M.C., 1995. Assimilation of Oceanic altimeter data using a quasi-geostrophic model and the extended Kalman filter. Technical Report, IDOPT Project INRIA-CNRS, RT141, October 1995.

Pham, D.T., Verron, J., Gourdeau, L., 1998. Singular evolutive Kalman filters for data assimilation in oceanography. To appear in C. R. Acad. Sci.

Pedlosky, J., 1987. Geophysical Fluid Dynamics. Springer-Verlag, New York.

Seber, G.A.F., 1984. Multivariate Observations. Wiley, New York.

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