Nonlin. Processes Geophys., 14, 59–71, 2007
www.nonlin-processes-geophys.net/14/59/2007/

Nonlinear Processes in Geophysics

Model error estimation in ensemble data assimilation

S. Gillijns and B. De Moor

SCD-SISTA-ESAT, Katholieke Universiteit Leuven, Leuven, Belgium

Received: 21 April 2006 – Revised: 8 January 2006 – Accepted: 8 January 2006 – Published: 31 January 2007

Abstract. A new methodology is proposed to estimate and account for systematic model error in linear filtering as well as in nonlinear ensemble based filtering. Our results extend the work of Dee and Todling (2000) on constant bias errors to time-varying model errors. In contrast to existing methodologies, the new filter can also deal with the case where no dynamical model for the systematic error is available. In the latter case, the applicability is limited by a matrix rank condition which has to be satisfied in order for the filter to exist.

The performance of the filter developed in this paper is limited by the availability and the accuracy of observations and by the variance of the stochastic model error component. The effect of these aspects on the estimation accuracy is investigated in several numerical experiments using the Lorenz (1996) model. Experimental results indicate that the availability of a dynamical model for the systematic error significantly reduces the variance of the model error estimates, but has only minor effect on the estimates of the system state. The filter is able to estimate additive model error of any type, provided that the rank condition is satisfied and that the stochastic errors and measurement errors are significantly smaller than the systematic errors. The results of this study are encouraging. However, it remains to be seen how the filter performs in more realistic applications.

1 Introduction

Error in environmental forecasting is mainly due to two causes: inaccurate initial conditions and deficiencies in the model. Much of attention has focused on reducing the effect of the first cause. Several suboptimal filters have been developed to assimilate measurements into large-scale models in order to come up with a more accurate estimate of the initial condition. The ensemble Kalman filter (EnKF), introduced by Evensen (1994), has gained particular popularity for environmental state estimation thanks to its ease of implementation and its robustness against filter divergence. Nowadays, the number of data assimilation applications involving the EnKF is numerous, see (Evensen, 1994; Houtekamer and Mitchell, 2001; Reichle et al., 2002; Evensen, 2003) and the references therein.

However, apart from stochastic model uncertainties, the EnKF is based on a perfect model assumptions. It is thus not able to deal with deficiencies in the model, which may play a major role in environmental forecasting (Orrell et al., 2001). A number of authors have addressed this lack of the EnKF. The effect of systematic model errors on the estimation accuracy is investigated in (Mitchell and Houtekamer, 2002) and (Reichle et al., 2002). In (Mitchell and Houtekamer, 2002; Heemink et al., 2001), an ad hoc method is used to account for systematic errors by treating the errors like random white noise with prescribed error covariance matrix. Another heuristic technique is covariance inflation (Anderson and Anderson, 1999), where the spread of the ensemble is artificially enlarged to make the filter more robust against model errors. Although both methods are successfully used in practice, they do not make use of the observations which contain information about the model error. Furthermore, none of both methods is able to yield estimates of the model error.

A commonly used method to estimate and deal with model error in Kalman filtering, is to augment the state vector with the model error vector and then design a Kalman filter for the augmented model. To reduce the computational load of the augmented state filter, Friedland (1969) proposed the two-stage filter, where the estimation of the state and the model error are separated. An efficient suboptimal variation of the two-stage filter was first applied in the data assimilation community by Dee and Da Silva (1998); Dee and Todling (2000) to estimate constant bias errors in numerical weather prediction. The state augmentation method has been successfully

Correspondence to: S. Gillijns
(steven.gillijns@esat.kuleuven.be)

used for estimating systematic model error in ensemble based data assimilation as well as in variational data assimilation (Zupanski, 1997; Griffith and Nichols, 2000; Martin et al., 2002; Zupanski and Zupanski, 2006). The method has the advantage of being very flexible and being able to incorporate different types of prior knowledge about the model error into the assimilation procedure. However, the fact that a model which describes the dynamical evolution of the error must be available, limits the applicability of the method.

There are types of model error of which the dynamics are not known, for example certain types of time-varying bias errors, errors due to unresolved scales, discretization errors, unmodeled dynamics and unknown disturbances. In these cases, the state augmentation method can not be used.

Like (Dee and Da Silva, 1998; Dee and Todling, 2000), this paper addresses the problem of additive model error estimation and correction in data assimilation. Based on the optimal linear filters of Kitanidis (1987); Gillijns and De Moor (2007), we develop a rigorous and efficient method to deal with systematic model error in linear filtering as well as in nonlinear ensemble based filtering. In case a dynamical model for the systematic error is available, our results extend the work of Dee and Todling (2000) to time-varying model error. More precisely, using the same approximation, we develop a suboptimal but efficient filter where the estimation of the time-varying model error and the state are interconnected. However, provided that a certain matrix rank condition is satisfied, our method can also deal with the case where no dynamical model for the systematic error is available.

The performance of the filter developed in this paper is limited by the availability and the accuracy of observations and by the variance of the stochastic model error component. The effect of these aspects on the estimation accuracy is investigated in several numerical experiments using the Lorenz (1996) model. Due to the limitations, the method can in practice not be used to correct the entire state vector for all types of errors described above. However, it can be used to obtain, possibly for a limited number of state variables, an idea about the additive effect of the model error affecting these state variables, which is especially useful if the dynamics of the error are unknown. These estimates might give insight into the dynamics of the error, which might lead to a refinement of the simulation model or to the development of a "model error model" which can then be incorporated into the assimilation procedure.

This paper is outlined as follows. In the next section, we formulate the problem considered in this paper in more detail. In Sect. 3, we develop two linear filters which can deal with systematic model error. The first filter is based on the results of Kitanidis (1987); Gillijns and De Moor (2007) and assumes that no dynamical model for the error is available. The second filter is obtained by incorporating prior knowledge about the model error in the first filter and has a close connection to the result of Dee and Todling (2000). These filters are extended to the framework of nonlinear ensemble based filtering in Sect. 4. In Sect. 5, we discuss the relation between our filters, the state augmentation method and the filter of Dee and Todling (2000). Finally, in Sect. 6, we consider several numerical examples using the Lorenz model.

2 Problem formulation

Consider the nonlinear discrete-time model

$$\boldsymbol{x}_{k+1} = \mathbf{F}_k(\boldsymbol{x}_k, \boldsymbol{u}_k), \tag{1}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the state vector, $\boldsymbol{u}_k \in \mathbb{R}^l$ is a known external forcing term and the operator $\mathbf{F}_k(\cdot)$ maps the state vector at time instant k to time instant $k+1$. Assume that the model operator $\mathbf{F}_k(\cdot)$ is subject to both additive stochastic model error and systematic model error. The stochastic component is denoted by $\boldsymbol{w}_k \in \mathbb{R}^n$ and is assumed zero-mean white with covariance matrix $\mathbf{Q}_k = \mathbb{E}[\boldsymbol{w}_k \boldsymbol{w}_k^{\mathsf{T}}]$. Furthermore, assume that the errorneous equations of $\mathbf{F}_k(\cdot)$ are known. This type of prior knowledge about the systematic model error may be represented by a matrix $\mathbf{G}_k \in \mathbb{R}^{n \times m}$, where m is the number of independent errors. For example, a binary matrix can be used, where the i-th row contains a 1 if the i-th equation of the operator $\mathbf{F}_k(\cdot)$ is errorneous. If the i-th and the j-th equation of the operator $\mathbf{F}_k(\cdot)$ are subject to the same error, then the i-th and the j-th row of $\mathbf{G}_k$ contain a 1 in the same column. Under these assumptions on the stochastic and the systematic model errors, there exists a vector $\boldsymbol{d}_k \in \mathbb{R}^m$ such that the state of the true system at time instant $k+1$ is given by

$$\boldsymbol{x}_{k+1} = \mathbf{F}_k(\boldsymbol{x}_k, \boldsymbol{u}_k) + \mathbf{G}_k \boldsymbol{d}_k + \boldsymbol{w}_k, \tag{2}$$

where $\boldsymbol{x}_k$ is the true system state at time instant k. The vector $\boldsymbol{d}_k$, which will be called the *model error vector* or simply *model error*, is in general a nonlinear function of $\boldsymbol{x}_{k-1}$ and $\boldsymbol{d}_{k-1}$, that is,

$$\boldsymbol{d}_{k+1} = \mathbf{H}_k(\boldsymbol{d}_k, \boldsymbol{x}_k). \tag{3}$$

In previous work on data assimilation in the presence of systematic model errors, it was always assumed that the operator $\mathbf{H}_k(\cdot)$ is known. In this paper, we will also consider the case where $\mathbf{H}_k(\cdot)$ is unknown.

We assume that noisy measurements $\boldsymbol{y}_k \in \mathbb{R}^p$ are available, related to the system state $\boldsymbol{x}_k$ by

$$\boldsymbol{y}_k = \mathbf{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k, \tag{4}$$

where $\boldsymbol{v}_k \in \mathbb{R}^p$, assumed to be uncorrelated to $\boldsymbol{w}_k$, is a zero-mean white random vector with covariance matrix $\mathbf{R}_k = \mathbb{E}[\boldsymbol{v}_k \boldsymbol{v}_k^{\mathsf{T}}]$. The measurements are assumed not to be subject to systematic errors.

The first objective of this paper is to develop linear recursive filters which estimate both the model error $\boldsymbol{d}_k$ and the system state $\boldsymbol{x}_k$ from the observations $\boldsymbol{y}_k$ in case the operator $\mathbf{F}_k(\cdot)$ is linear. We will consider the case $\mathbf{H}_k(\cdot)$ known as well as the case $\mathbf{H}_k(\cdot)$ unknown. This objective is addressed in Sect. 3. The second objective of the paper is to extend the linear filters to the framework of nonlinear ensemble based filtering. This objective is addressed in Sect. 4.

3 Linear filtering in the presence of model error

In case the model operator $\mathbf{F}_k(\cdot)$ is linear, the dynamics of the true system (2) can be written as

$$\boldsymbol{x}_{k+1} = \mathbf{A}_k\boldsymbol{x}_k + \mathbf{B}_k\boldsymbol{u}_k + \mathbf{G}_k\boldsymbol{d}_k + \boldsymbol{w}_k. \tag{5}$$

In Sect. 3.1, we investigate what happens if $\boldsymbol{d}_k$ is neglected and the Kalman filter is used to estimate the state vector $\boldsymbol{x}_k$. Next, in Sect. 3.2, we discuss the filters of Kitanidis (1987); Gillijns and De Moor (2007) which take the model error into account and yield optimal estimates of $\boldsymbol{x}_k$ under the assumption that $\mathbf{H}_k(\cdot)$ is unknown. Finally, in Sect. 3.3, we show how the knowledge of the operator $\mathbf{H}_k(\cdot)$ can be incorporated in the filter of Gillijns and De Moor (2007).

3.1 The flaws of the Kalman filter

Assume that we neglect the model error $\boldsymbol{d}_k$ and apply the Kalman filter to estimate the state of system (5). The resulting filter equations are then given by,

$$\hat{\boldsymbol{x}}_k^{\mathrm{f}} = \mathbf{A}_{k-1}\hat{\boldsymbol{x}}_{k-1}^{\mathrm{a}} + \mathbf{B}_{k-1}\boldsymbol{u}_{k-1}, \tag{6}$$

$$\hat{\boldsymbol{x}}_k^{\mathrm{a}} = \hat{\boldsymbol{x}}_k^{\mathrm{f}} + \mathbf{K}_k(\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}}), \tag{7}$$

where $\hat{\boldsymbol{x}}_k^{\mathrm{f}}$ denotes the estimate of $\boldsymbol{x}_k$ given measurements up to time instant $k-1$ and $\hat{\boldsymbol{x}}_k^{\mathrm{a}}$ denotes the estimate of $\boldsymbol{x}_k$ given measurements up to time instant k. The Kalman gain $\mathbf{K}_k$ is given by

$$\mathbf{K}_k = \mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}(\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k)^{-1}, \tag{8}$$

where $\mathbf{P}_k^{\mathrm{f}}$ is updated by

$$\mathbf{P}_k^{\mathrm{f}} = \mathbf{A}_{k-1}\mathbf{P}_{k-1}^{\mathrm{a}}\mathbf{A}_{k-1}^{\mathsf{T}} + \mathbf{Q}_{k-1}, \tag{9}$$

$$\mathbf{P}_k^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}_k\mathbf{C}_k)\mathbf{P}_k^{\mathrm{f}}. \tag{10}$$

Let $\hat{\boldsymbol{x}}_{k-1}^{\mathrm{a}}$ be unbiased, then it follows from (6) that $\hat{\boldsymbol{x}}_k^{\mathrm{f}}$ is biased because the model error is neglected. Furthermore, it follows from (7) that for the choice of $\mathbf{K}_k$ given by (8), also the updated state estimate $\hat{\boldsymbol{x}}_k^{\mathrm{a}}$ is biased. The optimal linear analysis is thus not given by the Kalman filter update.

3.2 An extension of the Kalman filter

Kitanidis (1987) developed a filter for the system (5) which can deal with $\mathbf{H}_k(\cdot)$ unknown and actually is optimal only if $\mathbf{H}_k(\cdot)$ is unknown. His filter takes the form (6)–(7) of the Kalman filter. However, the optimal gain matrix is not given by (8) but is obtained by minimizing the variance of $\hat{\boldsymbol{x}}_k^{\mathrm{a}}$ under an unbiasedness condition. The result of Kitanidis was extended in (Gillijns and De Moor, 2007), where a new design method for the filter was given and where it was shown that optimal estimates of $\boldsymbol{d}_{k-1}$ can be obtained from the innovation $\boldsymbol{y}_k-\mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}}$.

In this section, we summarize the equations of the filter developed in (Gillijns and De Moor, 2007). The filter takes the recursive from

$$\hat{\boldsymbol{x}}_k^{\mathrm{f}} = \mathbf{A}_{k-1}\hat{\boldsymbol{x}}_{k-1}^{\mathrm{a}} + \mathbf{B}_{k-1}\boldsymbol{u}_{k-1}, \tag{11}$$

$$\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}} = \mathbf{M}_k(\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}}), \tag{12}$$

$$\hat{\boldsymbol{x}}_k^{\mathrm{a}*} = \hat{\boldsymbol{x}}_k^{\mathrm{f}} + \mathbf{G}_{k-1}\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}}, \tag{13}$$

$$\hat{\boldsymbol{x}}_k^{\mathrm{a}} = \hat{\boldsymbol{x}}_k^{\mathrm{a}*} + \mathbf{K}_k(\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{a}*}), \tag{14}$$

where the estimation of the state vector and the model error vector are interconnected. As discussed in the previous section, (11) yields a biased estimate of the system state $\boldsymbol{x}_k$. Therefore, in the second step, $\mathbf{M}_k$ is determined such that (12) yields a minimum-variance unbiased estimate of $\boldsymbol{d}_{k-1}$ based on the innovation $\boldsymbol{y}_k-\mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}}$. This estimate is used for compensation in (13), such that $\hat{\boldsymbol{x}}_k^{\mathrm{a}*}$ is unbiased. In the final step, $\mathbf{K}_k$ is determined such that (14) yields a minimum-variance unbiased estimate of the system state $\boldsymbol{x}_k$. Note that (14) takes the form of the analysis step of the Kalman filter. Furthermore, note that (13)–(14) can be rewritten as

$$\hat{\boldsymbol{x}}_k^{\mathrm{a}} = \hat{\boldsymbol{x}}_k^{\mathrm{f}} + \mathbf{L}_k(\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}}), \tag{15}$$

where $\mathbf{L}_k$ is given by

$$\mathbf{L}_k = \mathbf{K}_k + (\mathbf{I} - \mathbf{K}_k\mathbf{C}_k)\mathbf{G}_{k-1}\mathbf{M}_k. \tag{16}$$

As shown in (Gillijns and De Moor, 2007), the gain matrix $\mathbf{K}_k$ minimizing the variance of $\hat{\boldsymbol{x}}_k^{\mathrm{a}}$ is not unique. One of the optimal values for $\mathbf{K}_k$ takes the form of the Kalman gain,

$$\mathbf{K}_k = \mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}(\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k)^{-1}, \tag{17}$$

where the covariance matrix $\mathbf{P}_k^{\mathrm{f}}$ is defined by

$$\mathbf{P}_k^{\mathrm{f}} = \mathbb{E}[\tilde{\boldsymbol{x}}_k^{\mathrm{f}}\tilde{\boldsymbol{x}}_k^{\mathrm{f}\mathsf{T}}], \tag{18}$$

$$= \mathbf{A}_{k-1}\mathbf{P}_{k-1}^{\mathrm{a}}\mathbf{A}_{k-1}^{\mathsf{T}} + \mathbf{Q}_{k-1}, \tag{19}$$

with $\tilde{\boldsymbol{x}}_k^{\mathrm{f}}=\boldsymbol{x}_k-\mathbf{G}_{k-1}\boldsymbol{d}_{k-1}-\hat{\boldsymbol{x}}_k^{\mathrm{f}}$, and with $\mathbf{P}_k^{\mathrm{a}}$ the covariance matrix of $\hat{\boldsymbol{x}}_k^{\mathrm{a}}$,

$$\mathbf{P}_k^{\mathrm{a}} = \mathbb{E}[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{a}})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{a}})^{\mathsf{T}}]. \tag{20}$$

It follows from (11) and (4)–(5) that there is a linear relation between the innovation $\boldsymbol{y}_k-\mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}}$ and the model error $\boldsymbol{d}_{k-1}$, given by

$$\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}} = \mathbf{E}_k\boldsymbol{d}_{k-1} + \boldsymbol{e}_k, \tag{21}$$

where $\mathbf{E}_k=\mathbf{C}_k\mathbf{G}_{k-1}$ and where $\boldsymbol{e}_k$ is given by

$$\boldsymbol{e}_k = \mathbf{C}_k\tilde{\boldsymbol{x}}_k^{\mathrm{f}} + \boldsymbol{v}_k. \tag{22}$$

Since $\mathbb{E}[\tilde{\boldsymbol{x}}_k^{\mathrm{f}}]=0$, $\boldsymbol{e}_k$ is a zero-mean random variable with covariance matrix

$$\tilde{\mathbf{R}}_k = \mathbb{E}[\boldsymbol{e}_k\boldsymbol{e}_k^{\mathsf{T}}] = \mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k. \tag{23}$$

It follows from (21) that a minimum-variance unbiased estimate of $\boldsymbol{d}_{k-1}$ can be obtained from the innovation by weighted least-squares estimation with weighting matrix $\tilde{\mathbf{R}}_k^{-1}$. The optimal value for $\mathbf{M}_k$ is thus given by

$$\mathbf{M}_k = \left(\mathbf{E}_k^{\mathsf{T}}\tilde{\mathbf{R}}_k^{-1}\mathbf{E}_k\right)^{-1}\mathbf{E}_k^{\mathsf{T}}\tilde{\mathbf{R}}_k^{-1}, \tag{24}$$

and the variance of the corresponding model error estimate $\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}}$ by

$$\mathbf{P}_{k-1}^{\mathrm{d}} = \mathbb{E}[(\boldsymbol{d}_{k-1} - \hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}})(\boldsymbol{d}_{k-1} - \hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}})^{\mathsf{T}}], \tag{25}$$

$$= (\mathbf{E}_k^{\mathsf{T}}\tilde{\mathbf{R}}_k^{-1}\mathbf{E}_k)^{-1}. \tag{26}$$

Note that the inverses in (24) and (26) exist under the condition that

$$\operatorname{rank} \mathbf{C}_k\mathbf{G}_{k-1} = \operatorname{rank} \mathbf{G}_{k-1} = m. \tag{27}$$

Equation (27) gives the condition under which the model error can be uniquely determined from the innovation. Note that this condition implies $n \geq m$ and $p \geq m$.

The filter described in this section can thus deal with the case where $\mathbf{H}_k(\cdot)$ is unknown. Note that it can estimate model errors of any type. However, its applicability is limited by the matrix rank condition (27). Furthermore, as will be discussed further in the paper, the variance of the model error estimate (12) can be rather high.

3.3 Incorporating prior knowledge about the model error

If prior information about the model error is available, the variance of the model error estimate (12) can be reduced. Consider the case where an unbiased estimate $\hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}}$ with covariance matrix $\mathbf{P}_{k-1}^{\mathrm{f,d}}$ is available. The least-squares problem obtained by combining the information in the innovation and in $\hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}}$, is given by

$$\begin{bmatrix} \boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}} \\ \hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}} \end{bmatrix} = \begin{bmatrix} \mathbf{E}_k \\ \mathbf{I} \end{bmatrix} \boldsymbol{d}_{k-1} + \begin{bmatrix} \boldsymbol{e}_k \\ \tilde{\boldsymbol{d}}_{k-1}^{\mathrm{f}} \end{bmatrix}, \tag{28}$$

where $\tilde{\boldsymbol{d}}_{k-1}^{\mathrm{f}} = \hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}} - \boldsymbol{d}_{k-1}$ is a zero-mean random vector with covariance matrix $\mathbf{P}_{k-1}^{\mathrm{f,d}}$. Under the assumption that

$$\mathbb{E}[\tilde{\boldsymbol{d}}_{k-1}^{\mathrm{f}}\boldsymbol{v}_k^{\mathsf{T}}] = 0, \tag{29}$$

$$\mathbb{E}[\tilde{\boldsymbol{d}}_{k-1}^{\mathrm{f}}(\tilde{\boldsymbol{x}}_k^{\mathrm{f}})^{\mathsf{T}}] = 0, \tag{30}$$

the least-squares solution $\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}}$ of (28) which coincides with the linear minimum-variance unbiased estimate of $\boldsymbol{d}_{k-1}$, can be written as

$$\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}} = \hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}} + \mathbf{P}_{k-1}^{\mathrm{f,d}}\mathbf{E}_k^{\mathsf{T}}\left(\mathbf{E}_k\mathbf{P}_{k-1}^{\mathrm{f,d}}\mathbf{E}_k^{\mathsf{T}} + \tilde{\mathbf{R}}_k\right)^{-1} (\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}} - \mathbf{E}_k\hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}}), \tag{31}$$

see (Kailath et al., 2000). Note that (31) has a structure similar to the analysis step of the Kalman filter. Furthermore, note that the inverse in (31) also exists if $\mathbf{E}_k$ does not have full column rank. If prior information about the model error is available, the existence condition (27) does not have to be necessarily satisfied in order for the filter to exist.

Substituting (12) by (31), we obtain the following filter,

$$\hat{\boldsymbol{x}}_k^{\mathrm{f}} = \mathbf{A}_{k-1}\hat{\boldsymbol{x}}_{k-1}^{\mathrm{a}} + \mathbf{B}_{k-1}\boldsymbol{u}_{k-1}, \tag{32}$$

$$\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}} = \hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}} + \mathbf{K}_k^{\mathrm{d}}(\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{f}} - \mathbf{E}_k\hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}}), \tag{33}$$

$$\mathbf{K}_k^{\mathrm{d}} = \mathbf{P}_{k-1}^{\mathrm{f,d}}\mathbf{E}_k^{\mathsf{T}}(\mathbf{E}_k\mathbf{P}_{k-1}^{\mathrm{f,d}}\mathbf{E}_k^{\mathsf{T}} + \mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k)^{-1}, \tag{34}$$

$$\hat{\boldsymbol{x}}_k^{\mathrm{a*}} = \hat{\boldsymbol{x}}_k^{\mathrm{f}} + \mathbf{G}_{k-1}\hat{\boldsymbol{d}}_{k-1}^{\mathrm{a}}, \tag{35}$$

$$\hat{\boldsymbol{x}}_k^{\mathrm{a}} = \hat{\boldsymbol{x}}_k^{\mathrm{a*}} + \mathbf{K}_k^{\mathrm{x}}(\boldsymbol{y}_k - \mathbf{C}_k\hat{\boldsymbol{x}}_k^{\mathrm{a*}}), \tag{36}$$

$$\mathbf{K}_k^{\mathrm{x}} = \mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}(\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k)^{-1}. \tag{37}$$

If conditions (29)–(30) hold, this filter is optimal in the minimum-variance unbiased sense. Indeed, under these conditions the gain matrix (37) minimizes the variance of (36), see Appendix A for an outline of the proof.

Now, assume that $\mathbf{H}_{k-2}(\cdot)$ is known and linear. Then the optimal estimate $\hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}}$ is given by

$$\hat{\boldsymbol{d}}_{k-1}^{\mathrm{f}} = \mathbf{H}_{k-2}(\hat{\boldsymbol{d}}_{k-2}^{\mathrm{a}}, \hat{\boldsymbol{x}}_{k-2}^{\mathrm{a}}). \tag{38}$$

Consider the filter consisting of (32)–(38). Note that for this filter the optimality condition (29) obtains. However, it is straightforward to verify that the optimality condition (30) is not satisfied, so that the filter is suboptimal. As will be shown in Sect. 5, this suboptimal filter has a strong connection to the efficient filter developed by Dee and Da Silva (1998); Dee and Todling (2000).

4 Nonlinear filtering in the presence of model error

In this section, we extend the filters discussed in the previous section to the framework of large-scale nonlinear ensemble based filtering. In Sect. 4.1, we show that the EnKF suffers from the same flaws as the Kalman filter. Next, in Sect. 4.2, we develop an ensemble based version of the Kitanidis filter which can deal with additive model error of any type. In Sect. 4.3, we show how prior information can be incorporated into the latter filter. Finally, in Sects. 4.4 and 4.5, we discuss computational aspects and limitations with respect to applicability.

4.1 The flaws of the ensemble Kalman filter

The EnKF can be seen as an ad hoc extension of the Kalman filter to large-scale nonlinear systems. It propagates an ensemble of q $(q \ll n)$ members, $\{\boldsymbol{\xi}^i, i=1 \ldots q\}$, which capture the mean and the covariance of the current state estimate. Covariance information is thus propagated implicitly in the ensemble. The EnKF is widely used in data assimilation applications due to the ease of implementation, the low computational cost and the low storage requirements.

First, consider the model (2) with $\boldsymbol{d}_k=0$. The algorithm of the EnKF consists of two steps which are repeated recursively.

The first step of the algorithm, the forecast step, projects the q ensemble members ahead in time, from time instant $k-1$ to k. This step is given by

$$\boldsymbol{\xi}_k^{\mathrm{f},i} = \mathbf{F}_{k-1}(\boldsymbol{\xi}_{k-1}^{\mathrm{a},i}, \boldsymbol{u}_{k-1}) + \boldsymbol{w}_{k-1}^i, \qquad i = 1\ldots q, \tag{39}$$

$$\bar{\boldsymbol{\xi}}_k^{\mathrm{f}} = \frac{1}{q}\sum_{i=1}^{q} \boldsymbol{\xi}_k^{\mathrm{f},i}, \tag{40}$$

where $\bar{\boldsymbol{\xi}}_k^{\mathrm{f}}$ denotes the estimate of the system state at time instant k given measurements up to time $k-1$. The forecast step thus comprises q runs of the numerical model, one run for each of the q ensemble members $\boldsymbol{\xi}_{k-1}^{\mathrm{a},i}$. To account for the stochastic model error, q random realizations $\boldsymbol{w}_{k-1}^i$, sampled from a distribution with mean zero and variance $\mathbf{Q}_{k-1}$, are added to the forecasted ensemble members in (39).

In the second step, the analysis step, the q ensemble members are updated with the observation $\boldsymbol{y}_k$ through a procedure which emulates the Kalman filter measurement update. Defining the error covariance matrix $\breve{\mathbf{P}}_k^{\mathrm{f}}$ by

$$\breve{\mathbf{P}}_k^{\mathrm{f}} = \mathbb{E}[(\boldsymbol{x}_k - \bar{\boldsymbol{\xi}}_k^{\mathrm{f}})(\boldsymbol{x}_k - \bar{\boldsymbol{\xi}}_k^{\mathrm{f}})^{\mathsf{T}}], \tag{41}$$

this step starts by approximating $\breve{\mathbf{P}}_k^{\mathrm{f}}\mathbf{C}_k$ and $\mathbf{C}_k\breve{\mathbf{P}}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}$ using the q ensemble members,

$$\overline{\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}} = \frac{1}{q-1}\sum_{i=1}^{q}\left(\tilde{\boldsymbol{\xi}}_k^{\mathrm{f},i}(\mathbf{C}_k\tilde{\boldsymbol{\xi}}_k^{\mathrm{f},i})^{\mathsf{T}}\right), \tag{42}$$

$$\overline{\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}} = \frac{1}{q-1}\sum_{i=1}^{q}\left((\mathbf{C}_k\tilde{\boldsymbol{\xi}}_k^{\mathrm{f},i})(\mathbf{C}_k\tilde{\boldsymbol{\xi}}_k^{\mathrm{f},i})^{\mathsf{T}}\right), \tag{43}$$

where $\tilde{\boldsymbol{\xi}}_k^{\mathrm{f},i}=\bar{\boldsymbol{\xi}}_k^{\mathrm{f}}-\boldsymbol{\xi}_k^{\mathrm{f},i}$. Next, the gain matrix $\bar{\mathbf{K}}_k$ is computed using the formula for the Kalman gain,

$$\bar{\mathbf{R}}_k = \overline{\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}} + \mathbf{R}_k, \tag{44}$$

$$\bar{\mathbf{K}}_k = \overline{\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}}\bar{\mathbf{R}}_k^{-1}, \tag{45}$$

and the ensemble members are updated with the measurements,

$$\boldsymbol{\xi}_k^{\mathrm{a},i} = \boldsymbol{\xi}_k^{\mathrm{f},i} + \bar{\mathbf{K}}_k\left(\boldsymbol{y}_k - \mathbf{C}_k\boldsymbol{\xi}_k^{\mathrm{f},i} + \boldsymbol{v}_k^i\right), \quad i = 1\ldots q \tag{46}$$

$$\bar{\boldsymbol{\xi}}_k^{\mathrm{a}} = \frac{1}{q}\sum_{i=1}^{q}\boldsymbol{\xi}_k^{\mathrm{a},i}, \tag{47}$$

where random realizations $\boldsymbol{v}_k^i$, sampled from a distribution with mean zero and variance $\mathbf{R}_k$, have to be added to the observations to account for the measurement noise (Burgers et al., 1998).

Now, consider the case $\boldsymbol{d}_k\neq 0$ and assume that we apply the EnKF to estimate the system state. Like in the Kalman filter, the forecasted state estimate $\bar{\boldsymbol{\xi}}_k^{\mathrm{f}}$ is then biased, even for $q\rightarrow\infty$. Consequently, it follows from (46) and (47) that the updated state estimate $\bar{\boldsymbol{\xi}}_k^{\mathrm{a}}$ is also biased.

4.2 The ensemble Kitanidis filter

An ensemble based filter which can deal with $\mathbf{H}_k(\cdot)$ unknown is obtained by extending the Kitanidis to the framework of ensemble based filtering. The resulting filter is called the ensemble Kitanidis filter (EnKiF) and consists of three steps.

In the first step, the ensemble members $\boldsymbol{\xi}_{k-1}^{\mathrm{a},i}$ are projected ahead in time. Like in the EnKF, this step comprises q runs of the numerical model and is given by (39)–(40). Due to the model error, this step introduces a bias error in the forecasted ensemble members $\boldsymbol{\xi}_k^{\mathrm{f},i}$.

In the second step, this bias error is accounted for by estimating the model error from the innovations and by using the resulting estimates for compensation. More precisely, an ensemble of model error estimates $\{\boldsymbol{\delta}_{k-1}^i, i=1\ldots q\}$ is computed from the measurement $\boldsymbol{y}_k$ and the forecasted ensemble $\{\boldsymbol{\xi}_k^{\mathrm{f},i}, i=1\ldots q\}$ by using an ensemble version of (12). To this aim, the matrix $\tilde{\mathbf{R}}_k^{-1}$ in (24) is replaced by its approximation (44),

$$\bar{\mathbf{M}}_k = \left(\mathbf{E}_k^{\mathsf{T}}\bar{\mathbf{R}}_k^{-1}\mathbf{E}_k\right)^{-1}\mathbf{E}_k^{\mathsf{T}}\bar{\mathbf{R}}_k^{-1}. \tag{48}$$

The ensemble members $\boldsymbol{\delta}_{k-1}^i$ are then computed by

$$\boldsymbol{\delta}_{k-1}^i = \bar{\mathbf{M}}_k(\boldsymbol{y}_k - \mathbf{C}_k\boldsymbol{\xi}_k^{\mathrm{f},i} + \boldsymbol{v}_k^i), \qquad i = 1\ldots q, \tag{49}$$

and the estimate of the model error is given by

$$\bar{\boldsymbol{\delta}}_{k-1} = \frac{1}{q}\sum_{i=1}^{q}\boldsymbol{\delta}_{k-1}^i. \tag{50}$$

As will be shown further in the paper, random vectors $\boldsymbol{v}_k^i$ with mean zero and variance $\mathbf{R}_k$ have to be added to the observation $\boldsymbol{y}_k$ in (49) in order that the sample variance of the ensemble of model error estimates converges to (26) for $q\rightarrow\infty$. This is similar to the analysis step of the EnKF where perturbed observations have to be used in order that the variance of the updated ensemble members converges to the correct value (Burgers et al., 1998). Finally, the forecasted ensemble members $\boldsymbol{\xi}_k^{\mathrm{f},i}$ are updated with $\boldsymbol{\delta}_{k-1}^i$ using an ensemble version of (13),

$$\boldsymbol{\xi}_k^{\mathrm{a},i*} = \boldsymbol{\xi}_k^{\mathrm{f},i} + \mathbf{G}_{k-1}\boldsymbol{\delta}_{k-1}^i, \qquad i = 1\ldots q. \tag{51}$$

In the third step, the variance of the ensemble $\{\boldsymbol{\xi}_k^{\mathrm{a},i*}, i=1\ldots q\}$ is reduced by emulating (14) in the same way as in the analysis step of the EnKF,

$$\boldsymbol{\xi}_k^{\mathrm{a},i} = \boldsymbol{\xi}_k^{\mathrm{a},i*} + \bar{\mathbf{K}}_k\left(\boldsymbol{y}_k - \mathbf{C}_k\boldsymbol{\xi}_k^{\mathrm{a},i*} + \boldsymbol{v}_k^i\right), \tag{52}$$

where $\bar{\mathbf{K}}_k$ is given by (45). Finally, the updated estimate of the system state is given by (47).

The random vectors $\boldsymbol{v}_k^i$ in (52) may be the same as in (49). Furthermore, if the same random vectors are used, (49), (51) and (52) can be combined to

$$\boldsymbol{\xi}_k^{\mathrm{a},i} = \boldsymbol{\xi}_k^{\mathrm{f},i} + \bar{\mathbf{L}}_k \left(\boldsymbol{y}_k - \mathbf{C}_k \boldsymbol{\xi}_k^{\mathrm{f},i} + \boldsymbol{v}_k^i \right), \quad (53)$$

where $\bar{\mathbf{L}}_k$ is given by

$$\bar{\mathbf{L}}_k = \bar{\mathbf{K}}_k + (\mathbf{I} - \bar{\mathbf{K}}_k \mathbf{C}_k) \mathbf{G}_{k-1} \bar{\mathbf{M}}_k. \quad (54)$$

In case of a linear model operator $\mathbf{F}_k(\cdot)$, this filter converges for $q \to \infty$ to the filter of Gillijns and De Moor (2007), see Appendix B for an outline of the proof.

4.3 Incorporating prior knowledge in the EnKiF

If a prior estimate of the model error is available, e.g. in the form of an operator $\mathbf{H}_k(\cdot)$, equations (32)-(38) can be extended to the framework of ensemble based filtering by making use of the analogy of (33)–(34) to the analysis step of the Kalman filter. As will be discussed in Sect. 5, the resulting filter has a close connection to the filter developed by Dee and Da Silva (1998). Therefore, it will be called the DDS-EnKiF.

It follows from (32)–(38) that the DDS-EnKiF needs a prior estimate $\hat{\boldsymbol{d}}_{-1}^{\mathrm{f}}$ with known variance to be initialized. However, if no prior estimate is available, but rank $\mathbf{C}_0\mathbf{G}_{-1}{=}m$, the DDS-EnKiF can be initialized by running the EnKiF for one or a few steps.

4.4 Computational aspects

Under the assumption that $\mathbf{C}_k$ and $\mathbf{G}_{k-1}$ are sparse, the matrix $\mathbf{E}_k^{\mathsf{T}} \bar{\mathbf{R}}_k^{-1} \mathbf{E}_k \in \mathbb{R}^{m\times m}$ in (48) can be efficiently computed by applying the matrix inversion lemma to (44) (Tippett et al., 2003), even if the number of measurements is very high. However, the calculation of the model error vector requires the inverse of $\mathbf{E}_k^{\mathsf{T}} \bar{\mathbf{R}}_k^{-1} \mathbf{E}_k$ to be computed, which is computationally very demanding if m is large. Consequently, the number of errors which can be accounted for by the EnKiF is limited by the available computational power.

It is well known that the use of a limited number of ensemble members ($q \ll n$) introduces sampling errors in the forecasted ensemble of the EnKF due to spuriously large correlation estimates between greatly separated grid points. Houtekamer and Mitchell (2001); Hamill et al. (2001) showed that the analysis can be improved by using *covariance localization*, a technique where the covariance estimates obtained from the ensemble are multiplied by a distance-dependent correlation function. In the *local ensemble Kalman filter* (Ott et al., 2002), a method where the analysis at each grid point is based on the forecasted ensemble members within a local cube of a few grid points, spurious correlations are avoided by assuming the correlation zero beyond the local cube. Similar techniques may be used to reduce the effect of spurious correlations in the EnKiF, where not only the forecasted state ensemble, but also the ensemble of model error estimates is affected by sampling errors.

The use of perturbed observations also introduces sampling errors in the EnKF and thus also in the EnKiF. Since the third step of the EnKiF is equivalent to the analysis step of the EnKF, a *square root filter* (Whitaker and Hamill, 2002; Bishop et al., 2001; Anderson, 2001; Tippett et al., 2003) can be employed to avoid the perturbed observations in (52). Note that the ensemble of model error estimates also suffers from sampling errors due to perturbed observations. A technique similar to square root filtering, where the mean and the variance of the model error estimate are computed separately, might reduce the effect of sampling errors due to perturbed observations.

4.5 Limitations with respect to applicability

The applicability of the EnKiF is hampered by the existence condition (27). For a constant bias error affecting all state variables in the same way, one measurement is in theory sufficient to estimate and account for the error. If all state variables are affected by independent errors, the method can not be used to correct the entire state vector because this would require that values of all state variables are incorporated into the measurement. In this case, the EnKiF can be used to obtain, possibly for a limited number of state variables, an idea about the additive effect of the model error affecting these state variables, which is especially useful if the dynamics of the error are unknown. The estimates of the model error might give insight into the dynamics of the errors, which might lead to a refinement of the simulation model or might lead to the development of a "model error model" which can then be incorporated into the assimilation procedure.

The EnKiF and DDS-EnKiF are based on the assumption that observational errors are zero-mean white with known covariance. If measurements with systematic errors are assimilated without preprocessing, the model error estimates and state estimates will be biased because the filter can not distinguish between systematic errors in the forecast model and in the observations. Therefore, if possible, observational bias must be removed. Also, a limited subset of unbiased observations may be used for the purpose of model error estimation (Dee and Da Silva, 1998).

The EnKiF is also based on the assumption that measurements are available at every assimilation time. If this is not the case, the EnKiF can still be used to estimate the model error which is build up during the consecutive time instants at which no measurements are available.

It follows from (26) that the variance of the model error estimate is determined by the variance of the measurement noise and by the variance of the forecasted state ensemble. In case the measurements are very noisy or the spread of the forecasted ensemble is very large (e.g. due to stochastic model error with high variance), the model error estimates obtained with the EnKiF will be very noisy too. Consequently, the model error estimates obtained with the EnKiF will be appropriate and accurate only if the stochastic model error and the measurement error are significantly smaller than the systematic model error.

The effect of these limitations on the accuracy of the state estimates and the model error estimates obtained with the EnKiF and DDS-EnKiF is investigated in several numerical studies in Sect. 6.

5 Comparison to existing methods

A standard approach to deal with systematic model error in Kalman filtering and data assimilation, is to augment the state vector with a vector of model error variables (Zupanski, 1997; Griffith and Nichols, 2000; Martin et al., 2002; Zupanski and Zupanski, 2006). This so-called method of state augmentation is very flexible and can incorporate different types of prior information into the problem. In its most general form, the method can estimate model error which nonlinearly interacts with the state vector. Let the model be given by (1), then the method can deal with the case where the true system is given by $\boldsymbol{x}_{k+1}{=}\bar{\mathbf{F}}_k(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{d}_k)$, provided that the interaction between the model error $\boldsymbol{d}_k$ and state vector $\boldsymbol{x}_k$ is known and provided that a model for the dynamical evolution of $\boldsymbol{d}_k$ is available, which is in its most general form given by (3). Note that the filters presented in Sect. 4 are not able to estimate the model error $\boldsymbol{d}_k$ in this general setting. However, they can be used to compensate and estimate the additive effect of these types of errors on the state vector, provided that the errorneous model equations are known.

In case of constant bias errors, the method of Dee and Da Silva (1998); Dee and Todling (2000) is usually applied. This method is based on the two-stage Kalman filter introduced by Friedland (1969), which can be seen as an augmented state filter where the estimation of the state and the model error have been separated. Dee and Todling (2000) developed a suboptimal, but efficient variation of the two-stage filter where, in contrast to the two-stage filter itself, information between the bias estimator and the state estimator is exchanged in two directions. The latter filter has a strong connection to the suboptimal filters developed in Sects. 3.3 and 4.3. More precisely, our results extend the work of Dee and Todling (2000) to time-varying model error. Indeed, in Sect. 3.3 we used the same approximation as Dee and Todling (2000) to develop an efficient filter which has a structure very similar to that of Dee and Todling (2000). The main difference is that our method estimates the model error with one step delay.

6 Numerical examples

In this section, we consider three numerical examples. The first example deals with bias errors, the second example with non-smooth disturbances and the third example with errors due to unresolved scales.

6.1 Bias errors

In a first experiment, we consider the example which was also used in (Anderson, 2001) for state estimation under constant bias errors. Consider the nonlinear one-level Lorenz (1996) model with $N{=}40$ and $F{=}8$ (the equations are given in Appendix C). This model is discretized using a fourth order Runge-Kutta scheme with time step $\Delta t{=}0.005$. The “true” states of the system are taken as the trajectories obtained with the Runge-Kutta scheme, where Gaussian white process noise is added to the discretized state variables. It is assumed that the exact value of F is unknown. The model is thus subject to a constant bias error. Noisy measurements of all state variables are available.

We compare the assimilation results obtained with an augmented EnKF based on the error model $\boldsymbol{d}_{k+1}{=}\boldsymbol{d}_k$ to the results of the DDS-EnKiF based on the same error model and to the results of the EnKiF. In the (DDS)-EnKiF, the matrix $\mathbf{G}$ is chosen as $\mathbf{G}{=}\mathbf{1}_m^{\mathsf{T}}$, which reflects that all state variables are affected by the same error. The initial bias estimate in the augmented EnKF was $F_0{=}10$. The DDS-EnKiF is initialized by first running one step of the EnKiF so that no initial estimate of the bias is needed.

Figure 1 compares the estimation results for 20 ensemble members and $\mathbf{Q}{=}10^{-5}\mathbf{I}$, $\mathbf{R}{=}10^{-3}\mathbf{I}$. Part (a) of the figure shows the estimated values of F. The variance of the estimates obtained with the EnKiF is clearly much higher than for the other two methods. Incorporating prior knowledge thus significantly reduces the variance of the bias estimate. Note the rather slow convergence of the augmented EnKF compared to the DDS-EnKiF where convergence is almost immediate. Part (b) of the figure shows the estimated values of the system state. Note that the high variance of the bias estimates obtained with the EnKiF has no detrimental effect on the estimated state trajectory.

Table 1 compares the mean square error (MSE) of the estimated F-values as function of the measurement noise variance and the variance of the stochastic model error. The values shown in the table were obtained by averaging the MSE over 1000 consecutive steps, after a converging time of 1000 steps. Results are shown for 20 ensemble members. The model error estimates are more accurate when $\mathbf{R}$ decreases. The MSE of the model error estimates also decreases with $\mathbf{Q}$. However, if $\mathbf{Q}$ is very small, the estimates degrade due to the fact that the spread of the ensemble is very small. This may lead to filter divergence because the filter gives very low confidence to the observations. We note filter divergence for values of $\mathbf{Q}$ smaller than $10^{-8}\mathbf{I}$. Again, it should be noted that

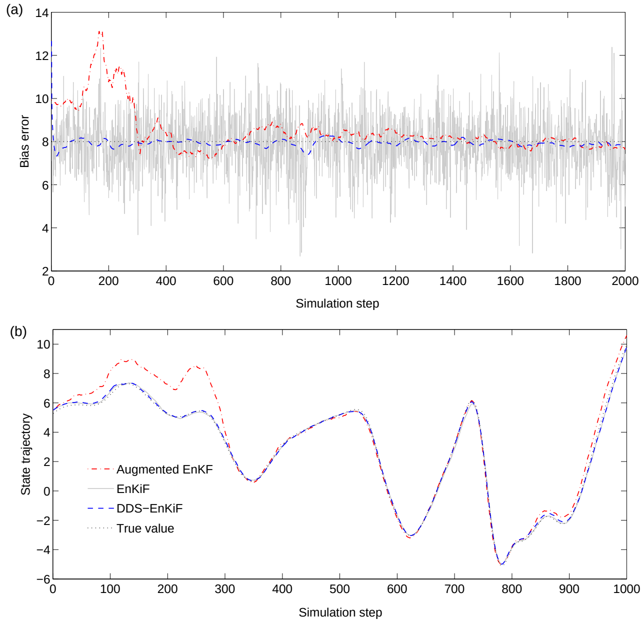

Fig. 1. Comparison between the assimilation results of an augmented EnKF, the EnKiF and the DDS-EnKiF for the example dealing with constant bias errors. **(a)** The variance of the bias estimates obtained with the EnKiF is much higher than for the other two methods. **(b)** However, this has no detrimental effect on the estimated state trajectory. Results are shown for 20 ensemble members and $\mathbf{Q}$=$10^{-5}\mathbf{I}$, $\mathbf{R}$=$10^{-3}\mathbf{I}$.

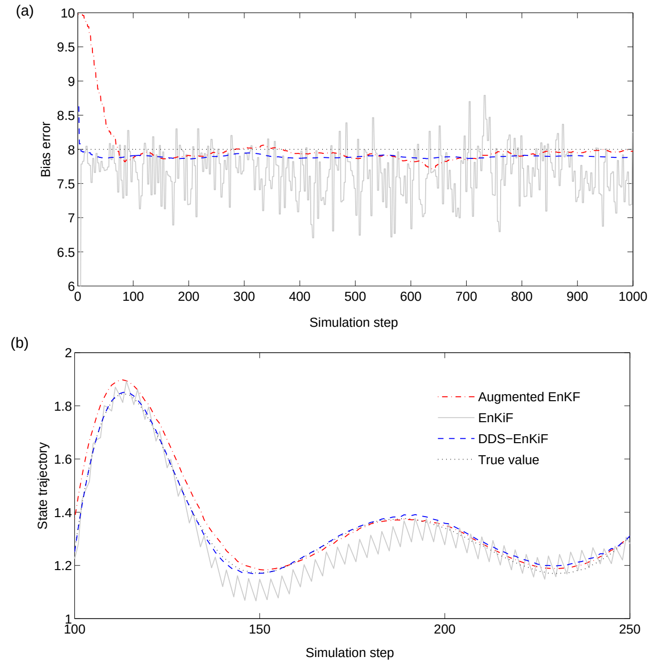

Fig. 2. Comparison between the model error estimates **(a)** and the state estimates **(b)** of the augmented EnKF, the EnKiF and the DDS-EnKiF when the time between measurements equals $3\Delta t$. Results are shown for $\mathbf{Q}=10^{-6}\mathbf{I}$, $\mathbf{R}=10^{-4}\mathbf{I}$ and 20 ensemble members.

Table 1. Comparison between the mean square error of the estimated $F-$values obtained with the EnKiF and the DDS-EnKiF as function of the measurement noise variance $\mathbf{R}$ and the variance of the stochastic model error $\mathbf{Q}$. Results are shown for 20 ensemble members.

Q		**R**			
		$10^{-2}\mathbf{I}$	$10^{-4}\mathbf{I}$	$10^{-6}\mathbf{I}$	$10^{-8}\mathbf{I}$
$10^{-2}\mathbf{I}$	DDS-EnKiF	2.10^{-2}	7.10^{-3}	9.10^{-3}	7.10^{-3}
	EnKiF	45	30	31	31
$10^{-4}\mathbf{I}$	DDS-EnKiF	4.10^{-3}	8.10^{-4}	1.10^{-3}	7.10^{-4}
	EnKiF	26	4.10^{-1}	3.10^{-1}	3.10^{-1}
$10^{-6}\mathbf{I}$	DDS-EnKiF	1.10^{-2}	4.10^{-3}	7.10^{-4}	7.10^{-4}
	EnKiF	23	3.10^{-2}	7.10^{-3}	6.10^{-3}
$10^{-8}\mathbf{I}$	DDS-EnKiF	2.10^{-1}	6.10^{-3}	3.10^{-3}	7.10^{-4}
	EnKiF	45	7.10^{-1}	3.10^{-2}	3.10^{-3}

the high variance of the model error estimates obtained with the EnKiF has no detrimental effect on the state estimates.

In real-life data assimilation applications, measurements may not be available at every assimilation time. Figure 2 explores what happens when the time between measurements equals $3\Delta t$. The second and the third step of the EnKiF can then be applied at only one out of three assimilation times. The estimates $\hat{\boldsymbol{d}}^{\mathrm{a}}_{k-1}$ obtained with the EnKiF thus represent the build-up of the systematic model error over three steps. Part (a) of Fig. 2 compares the estimated values of the model error obtained with the augmented EnKF, the DDS-EnKiF and the EnKiF. The estimates of the EnKiF shown in the figure are obtained by dividing $\hat{\boldsymbol{d}}^{\mathrm{a}}_{k-1}$ by three. Part (b) of Fig. 2 shows the estimated values of the system state. Due to the bias error which is not accounted for in the EnKiF, the state estimate diverges from the true value during two consecutive steps and then re-converges when measurements are assimilated. This leads to the behavior seen in Fig. 2. The non-availability of measurements at all assimilation times has minor effect on the augmented EnKF and the DDS-EnKiF, but is detrimental for the accuracy of the EnKiF.

The effect of systematic measurement error and incomplete measurements is investigated in Fig. 3. This figure shows results for 10 ensemble members, $\mathbf{Q}$=$10^{-6}\mathbf{I}$ and $\mathbf{R}$=$10^{-4}\mathbf{I}$. The measurements are subject to systematic errors which have a maximal value of $2,5.10^{-1}$. In addition, one out of five state variables is not measured. Part (a) of the figure compares the bias estimates obtained with an augmented EnKF, the EnKiF and the DDS-EnKiF. Part (b) shows the estimated value of state variable x_{20}, which is not measured. The MSE of the estimated bias error obtained with the

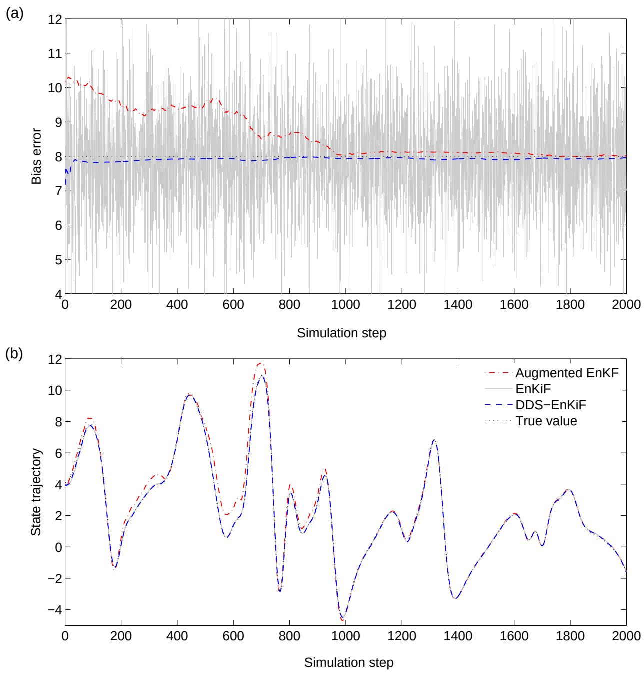

Fig. 3. Effect of systematic measurement error and incomplete measurements on the estimation accuracy. **(a)** Comparison between the bias estimates obtained with an augmented EnKF, the EnKiF and the DDS-EnKiF. **(b)** Estimated trajectory of state variable x_{20}, which is not measured. Results are shown for $\mathbf{Q}$=$10^{-6}\mathbf{I}$, $\mathbf{R}$=$10^{-4}\mathbf{I}$ and 10 ensemble members.

DDS-EnKiF increases from 8.10^{-3} in case of unbiased measurements to $1, 1.10^{-2}$ in case of systematic measurement error. The systematic measurement error has thus only small detrimental effect on the accuracy of the state estimates.

Now, consider the case where the model is subject to a time-varying bias error of which the dynamics are not known, such that the DDS-EnKiF and the augmented EnKF can not be used. Figure 4 shows the true value of the bias error and the estimate obtained with the EnKiF for $\mathbf{Q}$=$10^{-6}\mathbf{I}$, $\mathbf{R}$=$10^{-4}\mathbf{I}$ and 20 ensemble members. Like in the example dealing with constant bias errors, the estimates obtained with the EnKiF are rather noisy. However, the EnKiF is able to follow the fast variations in the bias error.

6.2 Non-smooth disturbances

In a second example, the true states of the system are taken as the trajectories of the one-level Lorenz model (with N=40 and F=8) obtained with the Runge-Kutta scheme, where Gaussian white process noise with variance $\mathbf{Q}$=$10^{-2}\mathbf{I}$ is added to the discretized state variables and where a non-smooth disturbance is added to state variable x_{21} at time instant $500\Delta t$. This disturbance has a peak value of 5 and a duration of $10\Delta t$. We compare the assimilation results obtained with the EnKF, where the disturbance is neglected, to

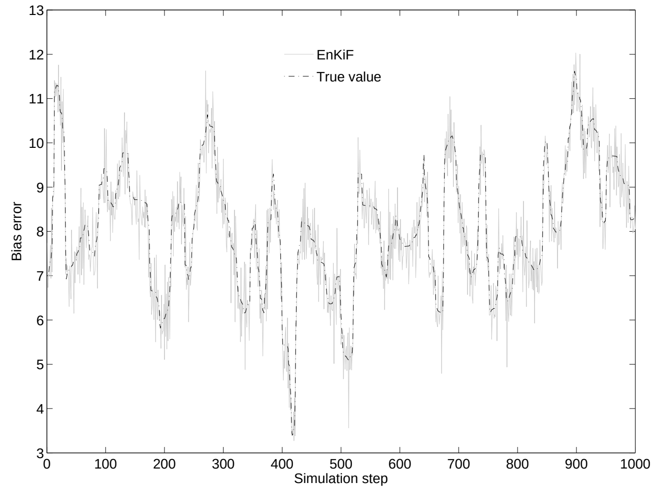

Fig. 4. Model error estimates obtained with the EnKiF for the example dealing with time-varying bias errors. Like in the example dealing with constant bias errors, the estimates obtained with the EnKiF are rather noisy. However, the EnKiF is able to follow the fast variations in the bias errors. Results are shown for $\mathbf{Q}$=$10^{-6}\mathbf{I}$, $\mathbf{R}$=$10^{-4}\mathbf{I}$ and 20 ensemble members.

the results of the EnKiF. Results are presented for 20 ensemble members and it is assumed that noisy measurements of all state variables, except x_{20}, are available. The measurement noise is Gaussian white with variance $\mathbf{R}$=$10^{-3}\mathbf{I}$. Figure 5a shows the true trajectory of state variable x_{21} and the trajectory that would be obtained if no disturbance would be present. The estimates of the EnKF and the EnKiF are also shown. The EnKF looses the true trajectory at the time the disturbances strikes, but quickly re-converges when the disturbance has disappeared. The performance of the EnKiF is better, it almost performs as if no disturbance is present. Figure 5b shows the trajectories for state variable x_{20} which is not affected by a disturbance, but not measured either. The same conclusions apply here.

6.3 Errors due to unresolved scales

Finally, in the third example, we emulate errors due to unresolved scales. The true system is taken to be the two-level Lorenz (1996) model (see Appendix C for the equations) with N=32, M=16 and F=10, consisting of 32 large-scale variables (the x-variables) and 512 fine-scale variables (the y-variables). The parameters c=10 and b=10 are chosen so that the fine-scale variables fluctuate ten times more rapidly, but with ten times smaller magnitude than the large-scale variables. The system is discretized using a fourth order Runge-Kutta scheme with time step Δt=0.005. The model is the one-level Lorenz model with $N = 32$ and F=10. As pointed out in (Orrell et al., 2001), this situation is analogous to that encountered in real weather models, where a

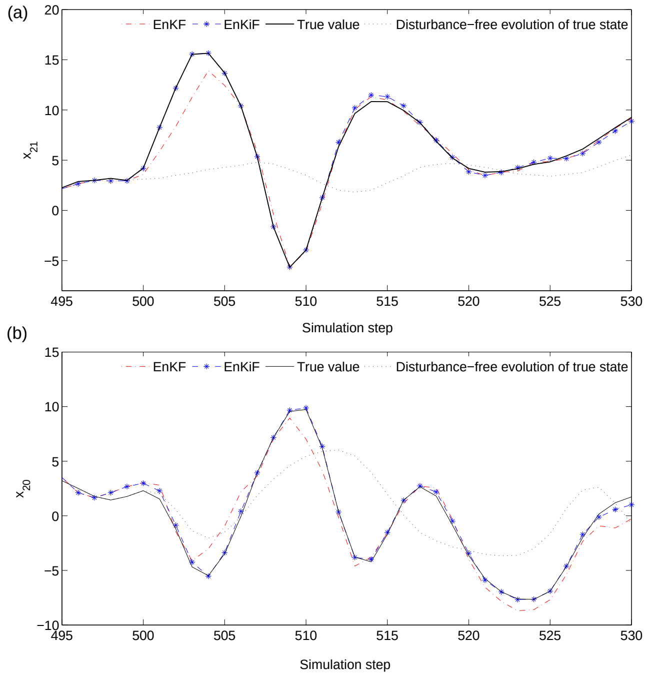

Fig. 5. Comparison between the assimilation results of the EnKF and the EnKiF for the Lorenz model subject to a high non-smooth disturbance. Results are shown for 20 ensemble members and $\mathbf{Q}=10^{-2}\mathbf{I}$, $\mathbf{R}=10^{-3}\mathbf{I}$. **(a)** Results for state variable x_{21}, which is measured, but affected by a disturbance. **(b)** Results for state variable x_{20}, which is not affected by a disturbance, but not measured either.

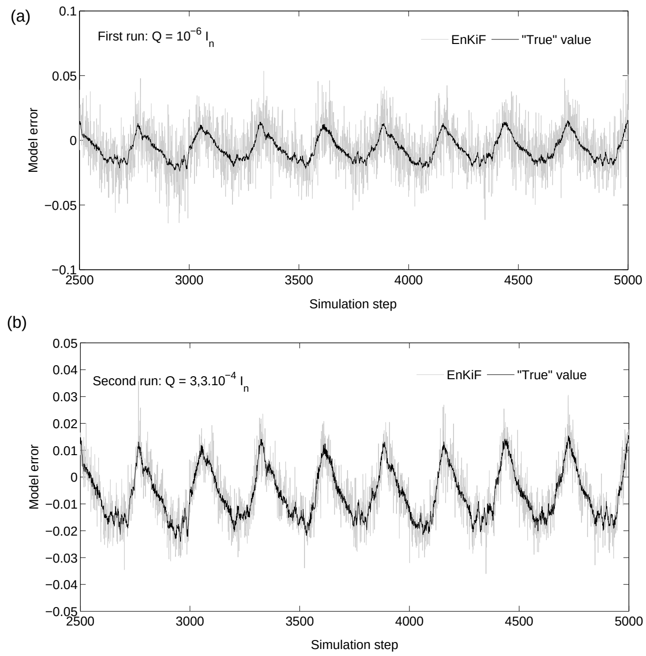

Fig. 6. True and estimated value of the error due to unresolved scales affecting state variable x_{16}. The model is taken to be the one-level Lorenz model, while measurements are generated using the two-level Lorenz model. **(a)** Results for $\mathbf{Q}=10^{-6}\mathbf{I}$. **(b)** Results for $\mathbf{Q}=3, 3.10^{-4}\mathbf{I}$.

constant forcing term is adopted to model the influence of unresolved fine-scale variables on the large-scale variables. The stochastic model error is assumed to be Gaussian white with variance $\mathbf{Q}_x=10^{-6}\mathbf{I}$ for the discretized large-scale variable and variance $\mathbf{Q}_y=10^{-8}\mathbf{I}$ for the discretized small-scale variable. It is assumed that noisy measurements of all large-scale variables are available. The measurement noise is Gaussian white with $\mathbf{R}=10^{-6}\mathbf{I}$. For these choices, the error in the measurements is approximately ten times smaller than the magnitude of the error due to unresolved scales.

The aim of this experiment is twofold. Firstly, we want to obtain an accurate estimate of the model error affecting state variables x_{15}, x_{16} and x_{17}. Secondly, we want to account for the model error affecting all other state variables by using an extension of the additive error approach developed by Mitchell and Houtekamer (2002) and used by Hamill and Whitaker (2005) to account for errors due to unresolved scales. In this approach, systematic model errors are accounted for by treating them like random white noise with artificially chosen variance. The aim of this experiment is to design a procedure in which this variance is computed from the estimates of the filter.

In order for the EnKiF to yield estimates of the model error affecting state variables x_{15} to x_{17}, we choose the $\mathbf{G}$-matrix as $\mathbf{G}=[\mathbf{0}_{3\times 14}\ \mathbf{I}_3\ \mathbf{0}_{3\times 15}]^{\mathsf{T}}$. The value of $\mathbf{Q}$ in the EnKiF is chosen to be $\mathbf{Q}=10^{-6}\mathbf{I}$, which is the variance of the stochastic model error affecting the large-scale variable in the true system. In the second step of the EnKiF algorithm, we apply covariance localization such that the model error affecting state variables x_{15} to x_{17} is estimated from innovations depending on estimates of state variables x_{15} to x_{17} only. All other innovations are inappropriate for estimating the model error affecting x_{15} to x_{17} due to the fact that these innovations depend on state estimates which are not accounted for model error. The true and estimated value of the model error affecting the state variable with index 16, are shown in Fig. 6a. The true value of the model error at time instant k, is computed by

$$\boldsymbol{d}_k = \mathbf{F}^{\text{tl}}_{k-1}(\boldsymbol{x}^{\text{tl}}_{k-1}) - \mathbf{F}^{\text{ol}}_{k-1}(\mathbf{T}(\boldsymbol{x}^{\text{tl}}_{k-1})), \tag{55}$$

where $\mathbf{F}^{\text{tl}}_{k-1}(\cdot)$ is the two-level Lorenz model operator, $\mathbf{F}^{\text{ol}}_{k-1}(\cdot)$ is the one-level Lorenz model operator and where $\mathbf{T}(\cdot)$ projects the state of the two-level model to the one-level model.

In a second step, we proceed as if all state variables are affected by independent zero-mean Gaussian errors with equal variance. In that case, the optimal value for $\mathbf{Q}$ equals $\sigma^2\mathbf{I}$, where σ^2 is the variance of the errors. We approximate σ by computing the standard deviation of the estimated model error affecting x_{16} over 5000 consecutive steps. The computed standard deviation equals $s_1=0,018$. Next, we apply the EnKiF with $\mathbf{Q}=10^{-6}\mathbf{I}$ for state variables x_{15} to x_{17}, but with $\mathbf{Q}=s_1^2\mathbf{I}$ for all other state variables. The true and estimated value of the model error affecting x_{16} are shown in Fig. 6b. Estimation accuracy has clearly increased. This improvement is also noticeable in the MSE of the state estimates, which has dropped from $1,3.10^{-3}$ in the first run to $3,6.10^{-4}$ in this run. The standard deviation of the estimated model error affecting x_{16} now equals $s_2=0,011$.

In a third step, we repeat the same procedure, with $\mathbf{Q}=10^{-6}\mathbf{I}$ for state variables x_{15} to x_{17} and with $\mathbf{Q}=s_2^2\mathbf{I}$ for all other state variables. The standard deviation of the model error affecting x_{16} now equals $s_3=0,012$. This values lies close to s_2, which indicates that s has almost converged to the optimal value which lies around $0,012$. Table 2 summarizes the results obtained in the three steps.

The method described above can be used to tune the variance of the random numbers in the additive error approach of Mitchell and Houtekamer (2002). In real-life applications, where the dimension of the measurement vector is much smaller than the dimension of the state vector, the matrix $\mathbf{G}_k$ can for example be chosen to estimate the errors affecting a limited number of state variables of which the value is incorporated into the measurements. For such a choice of $\mathbf{G}_k$, the rank condition (27) is always satisfied. The method described above can then be used to obtain an estimate of the errors affecting these state variables. Based on these estimates of the model error, the variance of the random numbers to be used in the approach of Mitchell and Houtekamer (2002) can be tuned.

Table 2. Results obtained in the three consecutive experiments dealing with errors due to unresolved scales. The matrix **Q** denotes the variance of the random vectors which are added to the forecasted ensemble members to account for the model error. The column "MSE" shows the mean square error of the state estimates. The last column shows the standard deviation of the estimates of the model error affecting x_{16}, which is used to compute the **Q**-matrix of the next step.

Step number	**Q**	MSE	s
1	$10^{-6}\mathbf{I}$	$1,3.10^{-3}$	0,018
2	$3,3.10^{-4}\mathbf{I}$	$3,6.10^{-4}$	0,011
3	$1,2.10^{-4}\mathbf{I}$	$3,3.10^{-4}$	0,012

7 Conclusion and discussion

A new methodology was developed to estimate and account for additive systematic model error in linear filtering as well as in nonlinear ensemble based data assimilation. In contrast to existing methodologies, the approach adopted in this paper can also deal with the case where no dynamical model for the error is available.

In case no model for the error is available, the filter is referred to as EnKiF. The applicability of the EnKiF is limited by the available computational power and by a matrix rank condition which has to be satisfied in order for the filter to exist. The EnKiF can therefore not be used to correct the entire state vector for all possible types of systematic errors. The intended use is therefore to obtain, possibly for a limited number of state variables, an idea about the additive effect of the model error affecting these state variables. This is especially useful if the dynamics of the error are unknown, e.g. in case of unknown time-varying bias errors or errors due to unresolved scales. The estimates of the model error might give insight into the dynamics of the error, which might lead to a refinement of the simulation model or might lead to the development of a "model error model" which can then be incorporated into the assimilation procedure.

In case a model for the error is available, the filter is referred to as DDS-EnKiF. It was shown that there is strong connection between the DDS-EnKiF and the efficient suboptimal filter developed by Dee and Todling (2000). More precisely, our results extend the latter work to time-varying bias errors.

Simulation results on the chaotic Lorenz (1996) model indicate that the model error estimates obtained with the EnKiF have a rather high variance. Estimation accuracy is mainly determined by the variances of the measurement error and the stochastic model error. It was shown that the availability of an accurate dynamical model for the error in the DDS-EnKiF strongly reduces the variance of the model error estimates. However, results also indicate that the high variance of the model error estimates obtained with the EnKiF has only minor detrimental effect on the state estimates.

Furthermore, simulation results indicate that the EnKiF and DDS-EnKiF are robust against systematic errors in the measurements. The non-availability of measurements at all assimilation times is detrimental for the accuracy of the EnKiF, but has only minor effect on the DDS-EnKiF because of the error model. The example dealing with constant bias errors indicates that both methods behave similarly as the number of measurements in space decreases.

This study indicates that the EnKiF might be preferable over the DDS-EnKiF when little or no prior information of the model error is available and when accurate measurements are available at every assimilation time. However, when relatively little information is available from measurements, additional information, e.g. in the form of a prior for the model error or an assumption on its evolution, will be necessary to account for systematic model error.

Appendix A

Calculation of optimal gain matrix

In this Appendix, we prove the optimality of the filter (32)–(38) for the case where conditions (29)–(30) are satisfied. We show that under the latter conditions the gain matrix (37) minimizes the variance of (36).

Using (32)–(38), we find that

$$\mathbf{P}_k^{\mathrm{a}} = \mathbf{K}_k^{\mathrm{x}}\bar{\bar{\mathbf{R}}}_k\mathbf{K}_k^{\mathrm{x}\mathsf{T}} - \mathbf{K}_k^{\mathrm{x}}\bar{\bar{\mathbf{S}}}_k - \bar{\bar{\mathbf{S}}}_k^{\mathsf{T}}\mathbf{K}_k^{\mathrm{x}\mathsf{T}} + \mathbf{P}_k^{\mathrm{a}*}, \tag{A1}$$

where

$$\bar{\bar{\mathbf{R}}}_k = \mathbf{C}_k\mathbf{P}_k^{\mathrm{a}*}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k - \mathbf{E}_k\mathbf{K}_k^{\mathrm{d}}\mathbf{R}_k - \mathbf{R}_k\mathbf{K}_k^{\mathrm{d}\mathsf{T}}\mathbf{E}_k^{\mathsf{T}}, \tag{A2}$$

$$\bar{\bar{\mathbf{S}}}_k = \mathbf{C}_k\mathbf{P}_k^{\mathrm{a}*} - \mathbf{R}_k\mathbf{K}_k^{\mathrm{d}\mathsf{T}}\mathbf{G}_{k-1}^{\mathsf{T}}, \tag{A3}$$

$$\mathbf{P}_k^{\mathrm{a}*} = \mathbb{E}[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{a}*})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^{\mathrm{a}*})^{\mathsf{T}}], \tag{A4}$$

$$= (\mathbf{I} - \mathbf{G}_{k-1}\mathbf{K}_k^{\mathrm{d}}\mathbf{C}_k)(\mathbf{P}_k^{\mathrm{f}} + \mathbf{G}_{k-1}\mathbf{P}_{k-1}^{\mathrm{f,d}}\mathbf{G}_{k-1}^{\mathsf{T}})$$
$$(\mathbf{I} - \mathbf{G}_{k-1}\mathbf{K}_k^{\mathrm{d}}\mathbf{C}_k)^{\mathsf{T}} + \mathbf{G}_{k-1}\mathbf{K}_k^{\mathrm{d}}\mathbf{R}_k\mathbf{K}_k^{\mathrm{d}\mathsf{T}}\mathbf{G}_{k-1}^{\mathsf{T}}. \tag{A5}$$

Note that these equations are valid only if conditions (29)–(30) are satisfied. The gain matrix $\mathbf{K}_k^{\mathrm{x}}$ minimizing the trace of (A1), is given by

$$\mathbf{K}_k^{\mathrm{x}} = \bar{\bar{\mathbf{S}}}_k^{\mathsf{T}}\bar{\bar{\mathbf{R}}}_k^{-1}. \tag{A6}$$

Finally, substituting (A2) and (A3) in (A6), yields after a straightforward calculation

$$\mathbf{K}_k^{\mathrm{x}} = \mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}(\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}} + \mathbf{R}_k)^{-1}. \tag{A7}$$

Appendix B

Proof of convergence

In this Appendix, we give an outline of the proof that, in case of a linear model operator, the EnKiF converges to the filter developed by Gillijns and De Moor (2007) for $q \to \infty$. Using the fact that the EnKF converges to the Kalman filter in case of a linear model, we only need to show that Eqs. (48)–(50) converge to the corresponding equations in Sect. 3.2. This basically comes down to showing that the sample variance of $\boldsymbol{\delta}_{k-1}^i$ converges to (26). This sample variance is given by

$$\frac{1}{q-1}\sum_{i=1}^{q}\left(\bar{\boldsymbol{\delta}}_{k-1} - \boldsymbol{\delta}_{k-1}^i\right)\left(\bar{\boldsymbol{\delta}}_{k-1} - \boldsymbol{\delta}_{k-1}^i\right)^{\mathsf{T}} = \bar{\mathbf{M}}_k\bar{\mathbf{R}}_k\bar{\mathbf{M}}_k^{\mathsf{T}}. \tag{B1}$$

It follows from the convergence of the EnKF to the Kalman filter that $\bar{\mathbf{R}}_k$ converges to $\tilde{\mathbf{R}}_k$ for $q \to \infty$. Consequently (B1) converges to (26). If no perturbed observations are used in (49), the sample variance would converge to $\mathbf{M}_k\mathbf{C}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{C}_k^{\mathsf{T}}\mathbf{M}_k^{\mathsf{T}}$ and would thus underestimate (26).

Appendix C

The Lorenz (1996) model

The equations for the one-level Lorenz (1996) model are given by

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \tag{C1}$$

where the index $i=1,\ldots,N$ is cyclic so that $x_{i-N}=x_{i+N}=x_i$.

The equations for the two-level model are given by

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F - \frac{c}{b}\sum_{j=1}^{M} y_{i,j}, \tag{C2}$$

$$\frac{dy_{i,j}}{dt} = cb(y_{i,j-1} - y_{i,j+2})y_{i,j+1} - cy_{i,j} + \frac{c}{b}x_i, \tag{C3}$$

for $i=1,\ldots,N$ and $j=1,\ldots,M$. The indices are cyclic so that for example $y_{i,j+M}=y_{i+1,j}$ and $y_{i+N,j}=y_{i,j}$.

Acknowledgements. Our research is supported by Research Council KULeuven: GOA AMBioRICS, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), G.0452.04 (new quantum algorithms), G.0499.04 (Statistics), G.0211.05 (Nonlinear), research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, GBOU (McKnow); Belgian Federal Science Policy Office: IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling', 2002-2006); PODO-II (CP/40: TMS and Sustainability); EU: FP5-Quprodis; ERNSI; Contract Research/agreements: ISMC/IPCOS, Data4s, TML, Elia, LMS, Mastercard.

Edited by: O. Talagrand
Reviewed by: two referees

References

Anderson, J.: An ensemble adjustment Kalman filter for data assimilation, Mon. Wea. Rev., 129, 2884–2903, 2001.

Anderson, J. and Anderson, S.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, Mon. Wea. Rev., 127, 2741–2758, 1999.

Bishop, C., Etherton, B., and Majundar, S.: Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects, Mon. Wea. Rev., 129, 420–436, 2001.

Burgers, G., van Leeuwen, P., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, Mon. Wea. Rev., 126, 1719–1724, 1998.

Dee, D. and Da Silva, A.: Data assimilation in the presence of forecast bias, Quart. J. Roy. Meteorol. Soc., 117, 269–295, 1998.

Dee, D. and Todling, R.: Data assimilation in the presence of forecast bias: the GEOS moisture analysis, Mon. Wea. Rev., 128, 3268–3282, 2000.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10 143–10 162, 1994.

Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation, Ocean Dyn., 53, 343–367, 2003.

Friedland, B.: Treatment of bias in recursive filtering, IEEE Trans. Autom. Control, 14, 359–367, 1969.

Gillijns, S. and De Moor, B.: Unbiased minimum-variance input and state estimation for linear discrete-time systems, Automatica, 43, 111–116, 2007.

Griffith, A. and Nichols, N.: Adjoint methods in data assimilation for estimating model error, J. Flow Turbulence Combust., 65, 469–488, 2000.

Hamill, T. and Whitaker, J.: Accounting for the error due to unresolved scales in ensemble data assimilation: a comparison of different approaches, Mon. Wea. Rev., 133, 3132–3147, 2005.

Hamill, T., Whitaker, J., and Snyder, C.: Distance-dependent filtering of background-error covariance estimates in an ensemble Kalman filter, Mon. Wea. Rev., 129, 2776–2790, 2001.

Heemink, A., Verlaan, M., and Segers, A.: Variance reduced ensemble Kalman filtering, Mon. Wea. Rev., 129, 1718–1728, 2001.

Houtekamer, P. and Mitchell, H.: A sequential ensemble Kalman filter for atmospheric data assimilation, Mon. Wea. Rev., 129, 123–137, 2001.

Kailath, T., Sayed, A., and Hassibi, B.: Linear Estimation, Prentice Hall, Upper Saddle River, New Jersey, 2000.

Kitanidis, P.: Unbiased minimum-variance linear state estimation, Automatica, 23, 775–778, 1987.

Lorenz, E.: Predictability: A problem partly solved, in: Proc. ECMWF Seminar on Predictability, Vol. I, pp. 1–18, 1996.

Martin, M., Bell, M., and Nichols, K.: Estimation of systematic error in an equatorial ocean model using data assimilation, Int. J. Numer. Fluids, 40, 435–444, 2002.

Mitchell, H. and Houtekamer, P.: Ensemble size, balance, and model-error representation in an ensemble Kalman filter, Mon. Wea. Rev., 130, 2791–2808, 2002.

Orrell, D., Smith, L., Barkmeijer, J., and Palmer, T.: Model error in weather forecasting, Nonlin. Processes Geophys., 8, 357–371, 2001, http://www.nonlin-processes-geophys.net/8/357/2001/.

Ott, E., Hunt, B., Szunyogh, I., Zimin, A., Kostelich, E., Corazza, M., Kalnay, E., Patil, D., and Yorke, J.: A local ensemble Kalman filter for atmospheric data assimilation, Tellus, 56A, 415–428, 2002.

Reichle, R., McLaughlin, D., and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, Mon. Wea. Rev., 130, 103–114, 2002.

Tippett, M., Anderson, J., Bishop, C., Hamill, T., and Whitaker, J.: Ensemble square root filters, Mon. Wea. Rev., 131, 1485–1490, 2003.

Whitaker, J. and Hamill, T.: Ensemble data assimilation without perturbed observations, Mon. Wea. Rev., 130, 1913–1924, 2002.

Zupanski, D.: A general weak constraint applicable to operational 4DVAR data assimilation systems, Mon. Wea. Rev., 125, 2274–2292, 1997.

Zupanski, D. and Zupanski, M.: Model error estimation employing an ensemble data assimilation approach, Mon. Wea. Rev., 134, 1337–1354, 2006.