RMetS
Royal Meteorological Society

Estimating model-error covariances in nonlinear state-space models using Kalman smoothing and the expectation–maximization algorithm

D. Dreano,[a] P. Tandeo,[b] M. Pulido,[c] B. Ait-El-Fquih,[a] T. Chonavel[b] and I. Hoteit[a]*

[a]*CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia*
[b]*Lab-STICC – Pôle CID, Telecom Bretagne, Brest, France*
[c]*Department of Physics, Universidad Nacional del Nordeste, Corrientes, Argentina*

*Correspondence to: I. Hoteit, CEMSE Division, King Abdullah University of Science and Technology (KAUST) 4700, 23955-6900 Thuwal, Saudi Arabia.
E-mail: ibrahim.hoteit@kaust.edu.sa

Specification and tuning of errors from dynamical models are important issues in data assimilation. In this work, we propose an iterative expectation–maximization (EM) algorithm to estimate the model-error covariances using classical extended and ensemble versions of the Kalman smoother. We show that, for additive model errors, the estimate of the error covariance converges. We also investigate other forms of model error, such as parametric or multiplicative errors. We show that additive Gaussian model error is able to compensate for non-additive sources of error in the algorithms we propose. We also demonstrate the limitations of the extended version of the algorithm and recommend the use of the more robust and flexible ensemble version. This article is a proof of concept of the methodology with the Lorenz-63 attractor. We developed an open-source Python library to enable future users to apply the algorithm to their own nonlinear dynamical models.

Key Words: expectation–maximization; EnKF; EnKS; extended Kalman filter; model error; state-space models

Received 8 September 2016; Revised 11 February 2017; Accepted 24 March 2017; Published online in Wiley Online Library

1. Introduction

One of the major difficulties in the practice of data assimilation is the specification of model errors (Daley, 1992; Dee, 1995; Hoteit *et al.*, 2007). Model errors can be due to different sources, including physical modelling errors, numerical errors, and misparametrization (Lahoz *et al.*, 2010). In assimilation schemes, model errors are conveniently represented by additive Gaussian random variables. This approach was shown to be very effective in many situations (Houtekamer *et al.*, 2009). Although inaccurate parametrization of the covariance matrix of the model error can have a negative impact on the performance of data assimilation schemes and may even lead to their failure (Dee, 1995; Mitchell and Houtekamer, 2000; Hoteit *et al.*, 2007), the development of accurate parametrization methodologies remains challenging.

Most research done on model error estimation for data assimilation have focused on innovation-based adaptive filters (Dee, 1995; Mitchell and Houtekamer, 2000; Berry and Sauer, 2013). In these algorithms, the model error is estimated sequentially along with the state by maximizing the likelihood of the innovations. This approach is particularly suited for the estimation of time-evolving model errors. These methods involve a parametric representation of the model-error covariance to reduce the number of degrees of freedom in the optimization problem (Dee, 1995; Mitchell and Houtekamer, 2000). More recently, Ueno *et al.* (2010) proposed an offline approach for estimating model errors in data assimilation schemes using an ensemble Kalman filter (EnKF), which relies on parameter space binning and optimization of the likelihood by exhaustive search.

Expectation–maximization (EM) is an iterative procedure to approximate the maximum likelihood estimate when hidden variables prevent a direct optimization of the log-likelihood (Dempster *et al.*, 1977). The EM algorithm has been implemented for estimating model-error covariance in linear Gaussian state-space models by Shumway and Stoffer (1982). In the expectation step, a Rauch–Tung–Striebel (RTS) Kalman smoother is used to compute the smoothing distributions (conditional on all past and future observations), which are used to derive a cost function. In the maximization step, this cost function is optimized to compute an update to the approximation of the maximum likelihood estimate. Unlike the approach by Ueno *et al.* (2010), the EM algorithm does not require a potentially expensive exhaustive search to optimize the log-likelihood of the state space model.

Recently, Tandeo *et al.* (2015) adapted the EM procedure in an ensemble framework to a case where the observations are nonlinearly related to the underlying states. They approximated the smoothing distribution using an ensemble version of the RTS Kalman smoother (Rauch *et al.*, 1965), the ensemble Kalman smoother (EnKS; Cosme *et al.*, 2012). EnKS is a boostrapped version of the Bayesian smoother, which, contrary to the particle filter, uses the best linear unbiased estimation (BLUE) to re-estimate the state when new observations are available (Smith

et al., 2013). EnKS is a two-pass smoother whose forward pass is given by the EnKF (Burgers *et al.*, 1998; Hoteit *et al.*, 2002, 2015).

The EM algorithm has also been used in conjunction with the EnKS and the Extended Kalman Smoother (EKS) to estimate physical parameters in chemical (Chitralekha *et al.*, 2010) and neuroscience (Kulkarni and Paninski, 2007) applications relying on nonlinear dynamical models. (Ueno and Nakamura, 2014) used the EM algorithm to sequentially compute an approximation of the maximum likelihood estimate of the observation-error covariance in a data assimilation scheme using the EnKF. More recently, (Ueno and Nakamura, 2016) applied the EM algorithm for online estimation of the observation-error covariance matrix in a Bayesian framework.

In this work, one of our goals is to evaluate the ability of the Gaussian additive errors to compensate for non-additive model errors. We extend the EM-EnKS approach proposed by Tandeo *et al.* (2015) for estimating the model- and observation-error covariances, as well as the parameters of the Gaussian initial background state in nonlinear dynamical systems. We then designate these statistical parameters as the hyperparameters of the state-space model. We also derive another approximation of the EM procedure based on the EKS, which we refer to as EM-EKS. EM-EKS and EM-EnKS are tested on trajectories simulated with the Lorenz 63 (L63) model (Lorenz, 1963), with additive and non-additive model error assumptions. Ultimately, we focus on estimating the appropriate model-error covariances.

In the following section, we derive the EM-EKS and the EM-EnKS schemes for estimating the hyperparameters of a nonlinear state-space model. In section 3, we present the experiments testing the EM-EnKS/EKS algorithms with different parametrization of model errors, and describe the results in section 4. In section 5, we conclude by discussing the implication of our results to the estimation of model error in general.

2. Methods

2.1. State-space models

Let $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_K\}$ denote the discrete-time state process, which is typically unknown to the observer. Let $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K\}$ represent the observation process, which describes the measurements of the system that are available to the observer. In many applications, such as geophysical fluid dynamics (Evensen, 2009), oil reservoir modelling (Aanonsen *et al.*, 2009; Luo *et al.*, 2012), or target tracking (Chui and Chen, 2009), these state and observation processes are related following a nonlinear state-space model:

$$\mathbf{x}_k = f_{k-1,k}(\mathbf{x}_{k-1}) + \boldsymbol{\eta}_k, \tag{1}$$

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\epsilon}_k. \tag{2}$$

Equation (1) describes how the system state $\mathbf{x}_k$ evolves through a nonlinear dynamical model $f_{k-1,k}$ between successive time steps t_{k-1} and t_k. $\boldsymbol{\eta} = \{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_K\}$ is the model noise process, which accounts for the imperfections of the model; it is assumed to be independent and identically distributed (i.i.d.). At each time step, $\boldsymbol{\eta}_k$ is assumed to be Gaussian with zero mean and covariance matrix $\mathbf{Q}$. Equation (2) models how an observation $\mathbf{y}_k$ at time t_k is obtained from the state $\mathbf{x}_k$ through a linear operator $\mathbf{H}$. $\boldsymbol{\epsilon} = \{\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_k\}$ is an i.i.d. process representing the observation errors; $\boldsymbol{\epsilon}_k$ is assumed Gaussian with zero mean and covariance matrix $\mathbf{R}$. Finally, the processes $\boldsymbol{\eta}$ and $\boldsymbol{\epsilon}$ are assumed to be jointly independent and independent of the initial (background) state $\mathbf{x}_0$, which is assumed to be Gaussian with mean $\mathbf{x}^{\mathrm{b}}$ and covariance $\mathbf{B}$.

In this work, we assume that the hyperparameters $\mathbf{x}^{\mathrm{b}}, \mathbf{B}, \mathbf{Q}$, and $\mathbf{R}$ of the model (1)–(2) are unknown and that $\mathbf{Q}$ and $\mathbf{R}$ are fixed in time. Our objective is to estimate these parameters using the EM algorithm. (Shumway and Stoffer, 1982) derived such an algorithm in the case where $f_{k-1,k}$ is linear. Tandeo *et al.* (2015) adapted this algorithm to the case where the observation operator $\mathbf{H}$ is nonlinear. Our approach, derived below, extends these algorithms to the case of nonlinear dynamical models.

2.2. Expectation–maximization algorithm

Let $\theta = \{\mathbf{x}^{\mathrm{b}}, \mathbf{B}, \mathbf{R}, \mathbf{Q}\}$. In the system (1)–(2), the probability density function (pdf) of all observations $\mathbf{y}_{1:K} = \{\mathbf{y}_1, \dots, \mathbf{y}_K\}$ depends on the hyperparameters θ and will be denoted $p_\theta(\mathbf{y}_{1:K})$. Similarly, $p_\theta(\mathbf{u})$ stands for the pdf of a random variable $\mathbf{u}$, parametrized as a function of θ. The goal is to compute an estimator of θ that maximizes the likelihood function $p_\theta(\mathbf{y}_{1:K})$:

$$\widehat{\theta}_{\mathrm{ML}} = \arg\max_\theta [p_\theta(\mathbf{y}_{1:K})] \tag{3}$$

$$= \arg\max_\theta [\ln p_\theta(\mathbf{y}_{1:K})]. \tag{4}$$

In state-space models, the passage from the complete data pdf $p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K})$ to the incomplete data pdf $p_\theta(\mathbf{y}_{1:K})$, which is difficult to express analytically and thus to optimize, is governed by the marginalization:

$$p_\theta(\mathbf{y}_{1:K}) = \int \dots \int p_\theta(\mathbf{y}_{1:K}, \mathbf{x}_{0:K})\, \mathrm{d}\mathbf{x}_0 \dots \mathrm{d}\mathbf{x}_K. \tag{5}$$

Because of the integral in (5), $p_\theta(\mathbf{y}_{1:K})$ can be a complex function of θ, often making the direct computation of θ_{ML} impossible.

The EM algorithm is an iterative procedure to approximate $\widehat{\theta}_{\mathrm{ML}}$ when the complete likelihood can be expressed analytically and easily optimized. The algorithm involves the log-likelihood of the complete data pdf $\ln p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K})$ rather than that of the incomplete data $\ln p_\theta(\mathbf{y}_{1:K})$. Since in $\ln p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K})$, $\mathbf{x}_{0:K}$ is not known, it is replaced by its expected value conditionally on $\mathbf{y}_{1:K}$, and given an estimator of the parameter θ, calculated at the previous iteration (say i) and denoted by $\theta^{(i)}$. At iteration $i+1$, $\theta^{(i+1)}$ is calculated by maximizing this expected value. Thus, the EM algorithm proceeds in two steps that can be summarized as follows.

2.2.1. Expectation (E) step

This step computes the auxiliary function:

$$G_{\theta^{(i)}}(\theta) = E_{p_{\theta^{(i)}}(\mathbf{x}_{0:K}|y_{1:K})}\left[\ln p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K})\right], \tag{6}$$

given by the expected value of $\ln[p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K})]$ with respect to the smoothing pdf of $\mathbf{x}_{0:K}$ given $\mathbf{y}_{1:K}$, parametrized by $\theta^{(i)}$.

2.2.2. Maximisation (M) step

This step provides an update $\theta^{(i+1)}$ of the hyper-parameters by maximizing $G_{\theta^{(i)}}(\theta)$; that is

$$G_{\theta^{(i)}}(\theta^{(i+1)}) \geq G_{\theta^{(i)}}(\theta),\ \forall\, \theta. \tag{7}$$

Dempster *et al.* (1977) have shown that the likelihood $p_{\theta^{(i)}}(\mathbf{y}_{1:K})$ does not decrease with the iterations (i); that is

$$p_{\theta^{(i+1)}}(\mathbf{y}_{1:K}) \geq p_{\theta^{(i)}}(\mathbf{y}_{1:K}),\ \forall i = 1, 2, \dots \tag{8}$$

An approximation of $\widehat{\theta}_{ML}$ is given by the last update of $\theta^{(i)}$ when some stopping criterion is achieved (for example when the difference between successive estimates of the hyperparameters $|\theta^{(i)} - \theta^{(i-1)}|$ is less than a predefined tolerance).

In the state-space system (1)–(2), the complete likelihood can be factorized using the independence properties of $\boldsymbol{\eta}$, $\boldsymbol{\epsilon}$, and $\mathbf{x}_0$:

$$p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K}) = p_\theta(\mathbf{x}_0) \prod_{k=1}^{K} p_\theta(\mathbf{x}_k|\mathbf{x}_{k-1}) \prod_{k=1}^{K} p(\mathbf{y}_k|\mathbf{x}_k), \tag{9}$$

which yields the log-likelihood:

$$L = \ln p_\theta(\mathbf{x}_{0:K}, \mathbf{y}_{1:K}) = \ln p_\theta(\mathbf{x}_0) + \sum_{k=1}^{K} \ln p_\theta(\mathbf{x}_k|\mathbf{x}_{k-1}) + \sum_{k=1}^{K} \ln p_\theta(\mathbf{y}_k|\mathbf{x}_k). \quad (10)$$

Since $\boldsymbol{\eta}$, $\boldsymbol{\epsilon}$ and $\mathbf{x}_0$ are Gaussian, one obtains:

$$L = -\frac{1}{2}\ln|\mathbf{B}| - \frac{1}{2}(\mathbf{x}_0-\mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_0-\mathbf{x}^{\mathrm{b}}) - \frac{K}{2}\ln|\mathbf{Q}| - \frac{1}{2}\sum_{k=1}^{K}\{\mathbf{x}_k - f_{k-1,k}(\mathbf{x}_{k-1})\}^{\mathrm{T}}\mathbf{Q}^{-1}\{\mathbf{x}_k - f_{k-1,k}(\mathbf{x}_{k-1})\} - \frac{K}{2}\ln|\mathbf{R}| - \frac{1}{2}\sum_{k=1}^{K}(\mathbf{y}_k-\mathbf{H}\mathbf{x}_k)^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}_k-\mathbf{H}\mathbf{x}_k) + C, \quad (11)$$

where C is a constant that does not depend on θ. Thus, the auxiliary function can be written as:

$$G_{\theta^{(i)}}(\theta) = -\frac{1}{2}\ln|\mathbf{B}| - \frac{1}{2}\mathrm{Tr}\left[\mathbf{B}^{-1}\Sigma_0^{(i)}\right] - \frac{K}{2}\ln|\mathbf{Q}| - \frac{1}{2}\mathrm{Tr}\left[\mathbf{Q}^{-1}\sum_{k=1}^{K}\Sigma_k^{(i)}\right] - \frac{K}{2}\ln|\mathbf{R}| - \frac{1}{2}\mathrm{Tr}\left[\mathbf{R}^{-1}\sum_{k=1}^{K}\Omega_k^{(i)}\right], \quad (12)$$

with

$$\Sigma_0^{(i)} = E_{p_{\theta^{(i)}}(\mathbf{x}_0|\mathbf{y}_{1:K})}\left[(\mathbf{x}_0-\mathbf{x}^{\mathrm{b}})(\mathbf{x}_0-\mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\right], \quad (13)$$

$$\Sigma_k^{(i)} = E_{p_{\theta^{(i)}}(\mathbf{x}_{k-1},\mathbf{x}_k|\mathbf{y}_{1:K})} \times\left[\{\mathbf{x}_k - f_{k-1,k}(\mathbf{x}_{k-1})\}\{\mathbf{x}_k - f_{k-1,k}(\mathbf{x}_{k-1})\}^{\mathrm{T}}\right], \quad (14)$$

$$\Omega_k^{(i)} = E_{p_{\theta^{(i)}}(\mathbf{x}_k|\mathbf{y}_{1:K})}\left[(\mathbf{y}_k-\mathbf{H}\mathbf{x}_k)(\mathbf{y}_k-\mathbf{H}\mathbf{x}_k)^{\mathrm{T}}\right]. \quad (15)$$

The computation of (13)–(15) requires the knowledge of the smoothing pdfs $p_{\theta^{(i)}}(\mathbf{x}_0|\mathbf{y}_{1:K}), p_{\theta^{(i)}}(\mathbf{x}_{k-1},\mathbf{x}_k|\mathbf{y}_{1:K})$ and $p_{\theta^{(i)}}(\mathbf{x}_k|\mathbf{y}_{1:K})$. We propose here to compute them using the EKS and EnKS algorithms, which will be presented in the following sections.

The maximization step consists of analytically optimizing $G_{\theta^{(i)}}(\theta)$. By taking the partial derivatives with respect to the parameters as zero, we obtain the following parameter updates:

$$\widehat{\mathbf{x}}^{b,(i+1)} = E_{p_{\theta^{(i)}}(\mathbf{x}_0|\mathbf{y}_{1:K})}[\mathbf{x}_0], \quad (16)$$

$$\widehat{\mathbf{B}}^{(i+1)} = \Sigma_0^{(i)}, \quad (17)$$

$$\widehat{\mathbf{Q}}^{(i+1)} = \frac{1}{K}\sum_{k=1}^{K}\Sigma_k^{(i)}, \quad (18)$$

$$\widehat{\mathbf{R}}^{(i+1)} = \frac{1}{K}\sum_{k=1}^{K}\Omega_k^{(i)}. \quad (19)$$

In the particular cases where $\mathbf{Q}$ is diagonal or of the form $\alpha\mathbf{Q}_0$, with $\alpha > 0$, and $\mathbf{Q}_0$ is positive definite, one can show that the updates are respectively:

$$\widehat{\mathbf{Q}}^{(i+1)} = \frac{1}{K}\sum_{k=1}^{K}\mathrm{diag}[\Sigma_k^{(i)}], \quad (20)$$

and

$$\widehat{\mathbf{Q}}^{(i+1)} = \alpha^{(i+1)}\mathbf{Q}_0, \quad (21)$$

with $\alpha^{(i+1)} = \mathrm{Tr}[\mathbf{Q}_0^{-1}\sum_{k=1}^{K}\Sigma_k^{(i)}]/Kn$, where n is the number of state variables.

2.3. EM equations for EKS

Here, we compute the smoothing pdfs by adapting the RTS smoother that was originally introduced in linear-Gaussian state-space models (Shumway and Stoffer, 1982; Ait-El-Fquih and Desbouvries, 2008; Cosme *et al.*, 2012) to the smoothing version of the extended Kalman filter (EKF; Anderson and Moore, 1979). The RTS smoother is a two-pass algorithm. First, a forward pass with the EKF to recursively compute the mean and covariance $(\mathbf{x}_k^{\mathrm{f}}, \mathbf{P}_k^{\mathrm{f}})$ of the Gaussian forecast pdf, $p(\mathbf{x}_k|\mathbf{y}_{1:k-1})$, and those $(\mathbf{x}_k^{\mathrm{a}}, \mathbf{P}_k^{\mathrm{a}})$ of the analysis pdf $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ for $k = 0, ..., K$. These quantities are then used in the backward pass, starting from the final time step t_K, to recursively compute backward in time the mean and covariance $(\mathbf{x}_k^{\mathrm{s}}, \mathbf{P}_k^{\mathrm{s}})$ of the Gaussian smoothing pdfs $p(\mathbf{x}_k|\mathbf{y}_{1:K})$. Since the dynamical model is nonlinear in (1), we linearize it, following the principle of the EKF, before we use the RTS smoother (Rauch *et al.*, 1965). The resulting EKS algorithm is summarized as follows.

2.3.1. Forward pass

Let $\mathcal{N}(\mathbf{x}_k^{\mathrm{f}}, \mathbf{P}_k^{\mathrm{f}})$ be the EKF approximation of the forecast pdf $p_\theta(\mathbf{x}_k|\mathbf{y}_{1:k-1})$, and let $\mathcal{N}(\mathbf{x}_k^{\mathrm{a}}|\mathbf{P}_k^{\mathrm{a}})$ be the EKF approximation of the filtering pdf $p_\theta(\mathbf{x}_k|\mathbf{y}_{1:k})$. Let the gradient at $\mathbf{x}_{k-1}^{\mathrm{a}}$ of the dynamical model be $\mathbf{M}_{k-1,k} = \partial f(\mathbf{x})/\partial\mathbf{x}|_{\mathbf{x}_{k-1}^{\mathrm{a}}}$. The filter is initialized with $(\mathbf{x}_0^{\mathrm{a}}, \mathbf{P}_0^{\mathrm{a}}) = (\mathbf{x}^{\mathrm{b}}, \mathbf{B})$. The EKF recursively computes the values $\mathbf{x}_k^{\mathrm{f}}, \mathbf{P}_k^{\mathrm{f}}, \mathbf{x}_k^{\mathrm{a}}, \mathbf{P}_k^{\mathrm{a}}$ through the following two steps.

- **EKF Forecast Step:**

$$\mathbf{x}_k^{\mathrm{f}} = f_{k-1,k}(\mathbf{x}_{k-1}^{\mathrm{a}}), \quad (22)$$

$$\mathbf{P}_k^{\mathrm{f}} = \mathbf{M}_{k-1,k}\mathbf{P}_{k-1}^{\mathrm{a}}\mathbf{M}_{k-1,k}^{\mathrm{T}} + \mathbf{Q}. \quad (23)$$

- **EKF Update Step:**

$$\mathbf{K}_k^{\mathrm{f}} = \mathbf{P}_k^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_k^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \quad (24)$$

$$\mathbf{x}_k^{\mathrm{a}} = \mathbf{x}_k^{\mathrm{f}} + \mathbf{K}_k^{\mathrm{f}}(\mathbf{y}_k - \mathbf{H}\mathbf{x}_k^{\mathrm{f}}), \quad (25)$$

$$\mathbf{P}_k^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}_k^{\mathrm{f}}\mathbf{H})\mathbf{P}_k^{\mathrm{f}}. \quad (26)$$

2.3.2. EKS smoothing step

Let $\mathcal{N}(\mathbf{x}_k^{\mathrm{s}}, \mathbf{P}_k^{\mathrm{s}})$ be the EKS approximation of the smoothing pdf $p_\theta(\mathbf{x}_k|\mathbf{y}_{1:K})$. For $k = K, K-1, ..., 0$, the forecast and analysis means and covariances $\mathbf{x}_k^{\mathrm{f}}, \mathbf{P}_k^{\mathrm{f}}, \mathbf{x}_k^{\mathrm{a}}, \mathbf{P}_k^{\mathrm{a}}$ are used to compute the smoothing means and covariances $\mathbf{x}_k^{\mathrm{s}}, \mathbf{P}_k^{\mathrm{s}}$ as:

$$\mathbf{K}_k^{\mathrm{s}} = \mathbf{P}_k^{\mathrm{a}}\mathbf{M}_{k,k+1}^{\mathrm{T}}(\mathbf{P}_{k+1}^{\mathrm{f}})^{-1}, \quad (27)$$

$$\mathbf{x}_k^{\mathrm{s}} = \mathbf{x}_k^{\mathrm{a}} + \mathbf{K}_k^{\mathrm{s}}(\mathbf{x}_{k+1}^{\mathrm{s}} - \mathbf{x}_{k+1}^{\mathrm{f}}), \quad (28)$$

$$\mathbf{P}_k^{\mathrm{s}} = \mathbf{P}_k^{\mathrm{a}} - \mathbf{K}_k^{\mathrm{s}}(\mathbf{P}_{k+1}^{\mathrm{f}} - \mathbf{P}_{k+1}^{\mathrm{s}})\mathbf{K}_k^{\mathrm{s}}, \quad (29)$$

with $\mathbf{K}_k^{\mathrm{s}}$ the smoothing gain. This pass is initialized at time step t_K with $(\mathbf{x}_K^{\mathrm{s}}, \mathbf{P}_K^{\mathrm{s}}) = (\mathbf{x}_K^{\mathrm{a}}, \mathbf{P}_K^{\mathrm{a}})$.

The two passes of EKS compute the smoothing means and covariances $(\mathbf{x}_k^{\mathrm{s}}, \mathbf{P}_k^{\mathrm{s}})$. They can be used to directly update the quantities Σ_0 and Ω_k in (13) and (15), respectively. However, Σ_k in (14) requires the knowledge of the cross-covariance

$$\mathbf{P}_{k,k-1}^{\mathrm{s}} \equiv E_{p(\mathbf{x}_k,\mathbf{x}_{k-1}|\mathbf{y}_{0:K})}[(\mathbf{x}_k-\mathbf{x}_k^{\mathrm{s}})(\mathbf{x}_{k-1}-\mathbf{x}_{k-1}^{\mathrm{s}})^{\mathrm{T}}],$$

which is not directly computed by the EKS algorithm above. We therefore take from Shumway and Stoffer (1982) (their Eqs A11 and A12) the following (backward) recursion:

$$\mathbf{P}_{k,k-1}^{\mathrm{s}} = \mathbf{P}_k^{\mathrm{a}}(\mathbf{K}_{k-1}^{\mathrm{s}})^{\mathrm{T}} + \mathbf{K}_k^{\mathrm{s}}(\mathbf{P}_{k+1,k}^{\mathrm{s}} - \mathbf{M}_{k,k+1}\mathbf{P}_k^{\mathrm{a}})(\mathbf{K}_{k-1}^{\mathrm{s}})^{\mathrm{T}}. \quad (30)$$

The recursion is initialized at the final time K by

$$\mathbf{P}^{\mathrm{s}}_{K,K-1} = (\mathbf{I} - \mathbf{K}^{\mathrm{f}}_K \mathbf{H})\mathbf{M}_{K-1,K}\mathbf{P}^{\mathrm{a}}_{k-1}.$$

In the maximization step, we use the results of the EKS and (30) to compute the quantities (13)–(15) at each iteration (i). For the background covariance (13):

$$\Sigma_0^{(i)} = \mathbf{P}^{\mathrm{s}}_0 + (E_{p_{\theta^{(i)}}(\mathbf{x}_0|\mathbf{y}_{1:K})}[\mathbf{x}_0] - \mathbf{x}^{\mathrm{b}})(E_{p_{\theta^{(i)}}(\mathbf{x}_0|\mathbf{y}_{1:K})}[\mathbf{x}_0] - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}, \quad (31)$$

$$= \mathbf{P}^{\mathrm{s}}_0. \quad (32)$$

For the cross-covariance (14) between successive states, $k = 1, \ldots, K$, after expansion and linearization, one obtains:

$$\Sigma_k^{(i)} = \mathbf{P}^{\mathrm{s}}_k + \left\{\mathbf{x}^{\mathrm{s}}_k - f_{k-1,k}(\mathbf{x}^{\mathrm{s}}_{k+1})\right\}\left\{\mathbf{x}^{\mathrm{s}}_k - f_{k-1,k}(\mathbf{x}^{\mathrm{s}}_{k+1})\right\}^{\mathrm{T}} + \mathbf{M}_{k-1,k}\mathbf{P}^{\mathrm{s}}_{k-1}\mathbf{M}^{\mathrm{T}}_{k-1,k} - \mathbf{P}^{\mathrm{s}}_{k-1,k}\mathbf{M}^{\mathrm{T}}_{k-1,k} - \mathbf{M}_{k-1,k}\mathbf{P}^{\mathrm{s}}_{k-1,k}. \quad (33)$$

For the covariance between states and observations (15):

$$\Omega_k^{(i)} = (\mathbf{y}_k - \mathbf{H}\mathbf{x}^{\mathrm{s}}_k)(\mathbf{y}_k - \mathbf{H}\mathbf{x}^{\mathrm{s}}_k)^{\mathrm{T}} + \mathbf{H}\mathbf{P}^{\mathrm{s}}_k\mathbf{H}^{\mathrm{T}}. \quad (34)$$

Finally, for the mean of the initial state, we get: $\widehat{\mathbf{x}}^{\mathrm{b},(i)} = \mathbf{x}^{\mathrm{s}}_0$.

2.4. EM equations for EnKS

EM-EnKS structure is similar to that of EM-EKS, with a forward filtering pass, a backward smoothing pass, and the computation of the covariances in (13)–(15).

2.4.1. Forward pass

We use the stochastic EnKF as proposed in Burgers *et al.* (1998), which approximates the forecast and analysis distributions from discrete samples $\mathbf{x}^{\mathrm{f}}_{k,j}$ and $\mathbf{x}^{\mathrm{a}}_{k,j}$ respectively, for $k = 1, \ldots, K$, $j = 1, \ldots, N_{\mathrm{e}}$, where N_{e} is the number of ensemble members. The samples are initialised with $\mathbf{x}^{\mathrm{a}}_{0,j} \sim \mathcal{N}(\mathbf{x}^{\mathrm{b}}, \mathbf{B})$. We then alternate forecast and analysis steps for $k = 1, \ldots, K$.

- **EnKF forecast step.** The forecast step propagates the analysis ensemble from time step t_{k-1} to time step t_k using the state equation (1) and computes samples of forecasted observation using the observation (2):

$$\mathbf{x}^{\mathrm{f}}_{k,j} = f_{k-1,k}(\mathbf{x}^{\mathrm{a}}_{k-1,j}) + \boldsymbol{\eta}_{k,j}, \qquad j = 1, \ldots, N_{\mathrm{e}}, \quad (35)$$

$$\mathbf{y}^{\mathrm{f}}_{k,j} = \mathbf{H}\mathbf{x}^{\mathrm{f}}_{k,j} + \boldsymbol{\epsilon}_{k,j}, \qquad j = 1, \ldots, N_{\mathrm{e}}, \quad (36)$$

  with $\boldsymbol{\eta}_{k,j} \sim \mathcal{N}(0, \mathbf{Q})$, $\boldsymbol{\epsilon}_{k,j} \sim \mathcal{N}(0, \mathbf{R})$, and $\mathbf{y}^{\mathrm{f}}_{k,j}$ the forecasted observation for sample j.
- **EnKF analysis step.** The forecasted ensemble $\mathbf{x}^{\mathrm{f}}_{k,j}$ is corrected with the new observation $\mathbf{y}_k$ to obtain the analysis ensemble $\mathbf{x}^{\mathrm{a}}_{k,j}$:

$$\mathbf{K}^{\mathrm{f}}_k = \bar{\mathbf{P}}^{\mathrm{f}}_k\mathbf{H}^{\mathrm{T}}[\mathbf{H}\bar{\mathbf{P}}^{\mathrm{f}}_k\mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-1}, \quad (37)$$

$$\mathbf{x}^{\mathrm{a}}_{k,j} = \mathbf{x}^{\mathrm{f}}_{k,j} + \mathbf{K}^{\mathrm{f}}_k(\mathbf{y}_k - \mathbf{y}^{\mathrm{f}}_{k,j}),\ j = 1, \ldots, N_{\mathrm{e}}, \quad (38)$$

  with $\bar{\mathbf{P}}^{\mathrm{f}}_k$ the sample covariance of the sample $\mathbf{x}^{\mathrm{f}}_{k,j}, j = 1, \ldots, N_e$.

2.4.2. Backward pass

We follow the RTS algorithm adapted by Cosme *et al.* (2012) to the EnKF. It is initialized with $\mathbf{x}^{\mathrm{s}}_{K,j} = \mathbf{x}^{\mathrm{a}}_{K,j}, j = 1, \ldots, N_{\mathrm{e}}$. For $k = K-1, \ldots, 0$:

$$\mathbf{K}^{\mathrm{s}}_k = \bar{\mathbf{P}}^{\mathrm{a,f}}_{k|k+1}(\bar{\mathbf{P}}^{\mathrm{f}}_{k+1})^{-1}, \quad (39)$$

$$\mathbf{x}^{\mathrm{s}}_{k,j} = \mathbf{x}^{\mathrm{a}}_{k,j} + \mathbf{K}^{\mathrm{s}}_k(\mathbf{x}^{\mathrm{s}}_{k+1,j} - \mathbf{x}^{\mathrm{f}}_{k+1,j}), \quad (40)$$

where $\bar{\mathbf{P}}^{\mathrm{a,f}}_{k|k+1}$ is the sample covariance matrix of $\mathbf{x}^{\mathrm{a}}_{k,j}$ and $\mathbf{x}^{\mathrm{f}}_{k+1,j}$, $j = 1, \ldots, N_{\mathrm{e}}$, and $\bar{\mathbf{P}}^{\mathrm{f}}_{k+1}$ is the sample covariance matrix of $\mathbf{x}^{\mathrm{f}}_{k+1,j}, j = 1, \ldots, N_{\mathrm{e}}$. The backward smoothing pass does not require dynamical model integrations, but uses only analyses and forecasts from the forward pass.

The inversion in (39) is implemented with the singular value decomposition (SVD) as describes in Cosme *et al.* (2012) (their Eq. 30). However, in the case where the state dimension is larger that the ensemble size, $\mathbf{P}^{\mathrm{f}}_k$ is not invertible. In that case, one can compute the smoothing gain using the formula proposed in Cosme *et al.* (2012) (their Eq. 31):

$$\mathbf{K}^{\mathrm{s}}_k = \mathbf{S}^{\mathrm{a}}_k[(\mathbf{S}^{\mathrm{f}}_{k+1})^{\mathrm{T}}\mathbf{S}^{\mathrm{f}}_{k+1}]^{-1}(\mathbf{S}^{\mathrm{f}}_{k+1})^{\mathrm{T}}, \quad (41)$$

where the columns of $\mathbf{S}^{\mathrm{a}}_k$ and $\mathbf{S}^{\mathrm{f}}_{k+1}$ are the centred ensembles members $\mathbf{x}^{\mathrm{a}}_{k,j}$ and $\mathbf{x}^{\mathrm{f}}_{k+1,j}$ $j = 1, \ldots, N_{\mathrm{e}}$, respectively. This formula is based on the fact that $\mathbf{K}^{\mathrm{s}}_k\bar{\mathbf{P}}^{\mathrm{f}}_{k+1} = \bar{\mathbf{P}}^{\mathrm{a,f}}_{k|k+1}$ for both (39) and (41).

Using the samples $\mathbf{x}^{\mathrm{s}}_{k,j}, k = 0, \ldots, K$, $j = 1, \ldots, N_e$, one can compute the sample estimates of equations (13)–(15) using:

$$\widehat{\mathbf{x}}^{\mathrm{b},(i)} = \frac{1}{N_{\mathrm{e}}}\sum_{j=1}^{N_{\mathrm{e}}}\mathbf{x}^{\mathrm{s}}_{0,j}, \quad (42)$$

$$\Sigma_0^{(i)} = \frac{1}{N_{\mathrm{e}}}\sum_{j=1}^{N_{\mathrm{e}}}(\mathbf{x}^{\mathrm{s}}_{0,j} - \widehat{\mathbf{x}}^{\mathrm{b},(i)})(\mathbf{x}^{\mathrm{s}}_{0,j} - \widehat{\mathbf{x}}^{\mathrm{b},(i)})^{\mathrm{T}}, \quad (43)$$

$$\Sigma_k^{(i)} = \frac{1}{N_{\mathrm{e}}}\sum_{j=1}^{N_{\mathrm{e}}}\{\mathbf{x}^{\mathrm{s}}_{k,j} - f(\mathbf{x}^{\mathrm{s}}_{k-1,j})\}(\{\mathbf{x}^{\mathrm{s}}_{k,j} - f(\mathbf{x}^{\mathrm{s}}_{k-1,j})\}^{\mathrm{T}}, \quad (44)$$

$$\Omega_k^{(i)} = \frac{1}{N_{\mathrm{e}}}\sum_{j=1}^{N_{\mathrm{e}}}(\mathbf{y}_k - \mathbf{H}\mathbf{x}^{\mathrm{s}}_{k,j})(\mathbf{y}_k - \mathbf{H}\mathbf{x}^{\mathrm{s}}_{k,j})^{\mathrm{T}}. \quad (45)$$

In the case where the state dimension is larger than the number of ensemble members, these sample covariance matrices are singular and therefore $\widehat{\mathbf{Q}}$ and $\widehat{\mathbf{R}}$ estimated in (18)–(19) can become singular as well (if the state dimension is larger than KN_{e}). In this case, $\widehat{\mathbf{Q}}$ and $\widehat{\mathbf{R}}$ can be parametrized as in (20)–(21) to ensure that they remain invertible.

2.5. Algorithms summary

Both EM-EKS and EM-EnKS are initialized with an arbitrary value of the parameter $\widehat{\theta}^{(0)} = \{\widehat{\mathbf{x}}^{b,(0)}, \widehat{\mathbf{B}}^{(0)}, \widehat{\mathbf{R}}^{(0)}, \widehat{\mathbf{Q}}^{(0)}\}$. For iterations $i = 1, \ldots, I$, the parameter is updated with the following steps:

(1) Filter the state with the observations using (22)–(26) for EM-EKS and (35)–(38) for EM-EnKS, assuming the value of the state-space parameter θ is its current estimate $\widehat{\theta}^{(i-1)}$. That is, the model and observation covariances are $\widehat{\mathbf{Q}}^{(i-1)}$ and $\widehat{\mathbf{R}}^{(i-1)}$, respectively, and the initial *a priori* state is specified by $\widehat{\mathbf{x}}^{\mathrm{b},(i-1)}$ and $\widehat{\mathbf{B}}^{(i-1)}$.

(2) Smooth the filtered states using the backward recursions (27)–(29) for EM-EKS and (39)–(40) for EM-EnKS. Additionally, EM-EKS requires running the backward recursion (30) to compute the cross-covariance matrix of smoothed states $\mathbf{P}^{\mathrm{s}}_{k,k-1}$.

(3) Compute the quantities $\Sigma_k^{(i)}$ and $\Omega_k^{(i)}$ using the smoothed states and covariances $\mathbf{x}^{\mathrm{s}}_k, \mathbf{P}^{\mathrm{s}}_k, \mathbf{P}^{\mathrm{s}}_{k,k-1}$ and (31)–(34) for EM-EKS, and use the smoothed ensemble states $\mathbf{x}^{\mathrm{s}}_{k,j}$ and (43)–(45) for EM-EnKS.

(4) Update the estimate of the hyperparameters $\widehat{\mathbf{Q}}^{(i)}, \widehat{\mathbf{R}}^{(i)}, \widehat{\mathbf{B}}^{(i)}, \widehat{\mathbf{x}}^{\mathrm{b},(i)}$ using (17)–(21) for parameters $\mathbf{Q}, \mathbf{R}$ and $\mathbf{B}$, respectively. The estimate of the mean initial value parameters $\mathbf{x}^{\mathrm{b}}$ is updated as $\mathbf{x}^{\mathrm{s}}_0$ for EM-EKS and using (42) for EM-EnKS.

These steps are repeated until a given stopping condition is achieved (e.g. on the log-likelihood or the estimated parameter values). The estimate of the parameter $\widehat{\theta}_{\mathrm{EM}}$ is given by the estimate at the last iteration I: $\widehat{\theta}^{(I)} = \{\widehat{\mathbf{Q}}^{(I)}, \widehat{\mathbf{R}}^{(I)}, \widehat{\mathbf{B}}^{(I)}, \widehat{\mathbf{x}}^{\mathrm{b},(I)}\}$.

3. Numerical experiments

The numerical experiments use the EM-EKS and the EM-EnKS algorithms to estimate the model-error covariance matrix $\mathbf{Q}$ ($\mathbf{R}$ is assumed known and fixed). Different model error assumptions are used to simulate the true trajectories, $\mathbf{x}_k^{\mathrm{t}}, k = 0, \ldots, K$. The observations are computed from $\mathbf{x}^{\mathrm{t}}$ according to (2). These experiments will assess the ability of EM-EKS and EM-EnKS to perform in realistic settings where the type of model error is unknown *a priori*. These experiments will also reveal to what extent the additive Gaussian model error assumption is reasonable in real systems. We will consider the following assumptions for model error: additive, parametric and multiplicative.

3.1. Lorenz-63 (L63) model and experimental design

We apply the EM algorithms described in the previous section to the L63 model (Lorenz, 1963). This model is low-dimensional and has a chaotic behaviour, which makes it an ideal reference model for testing data assimilation schemes. The L63 differential equations are:

$$\frac{\mathrm{d}x_1(t)}{\mathrm{d}t} = \sigma\{x_2(t) - x_1(t)\}, \tag{46}$$

$$\frac{\mathrm{d}x_2(t)}{\mathrm{d}t} = x_1(t)\{\rho - x_3(t)\} - x_2(t), \tag{47}$$

$$\frac{\mathrm{d}x_3(t)}{\mathrm{d}t} = x_1(t)x_2(t) - \beta x_3(t), \tag{48}$$

where $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. The model was integrated over a total duration of $T = 100$ (assimilation window), and the states were recorded every model step $\mathrm{d}t = 0.01$ (10 000 time steps). The ordinary differential equation (ODE) system is integrated using an explicit Runge–Kutta method of order 5 (Hairer *et al.*, 2009).

The background state of the EM-EKS and EM-EnKS algorithms is initialized with climatological estimates calculated based on the free integration of the model over 5000 model steps ($\mathrm{d}t = 0.01$). The initial state $\mathbf{x}^{\mathrm{b}}$ is taken as the average of the states over that period and the initial covariance $\mathbf{B}$ as the empirical covariance matrix of the simulated states. Unless mentioned otherwise, EM-EKS and EM-EnKS were iterated 200 times (EM iterations). To understand the impact of the ensemble size on EM-EnKS, we conducted experiments with 20, 100, and 500 ensemble members. For each experimental set-up, EM-EnKS was repeated 20 times to compute prediction intervals of the errors and log-likelihoods of the estimated state, and the model-error covariances.

For each of the model error assumptions mentioned above, we run the EM algorithms with the following types of observations:

(a) The full state is observed every model step ($\Delta t = 0.01$).
(b) The full state is observed every 10 model steps ($\Delta t = 0.1$).
(c) The first state variable is observed every model step ($\Delta t = 0.01$).
(d) The first state variable is observed every 10 model steps ($\Delta t = 0.1$).

The true trajectories are generated in different ways described below, depending on the type of simulated model errors. Then, we generate the observations using the true covariance error matrix (assumed known and not estimated) $\mathbf{R}^{\mathrm{t}} = \sigma_{\mathrm{obs}}^2\mathbf{I}$, with $\sigma_{\mathrm{obs}}^2 = 2$. In experiments (a) and (b), we set $\mathbf{H} = \mathbf{I}$ in (2), while for (c) and (d), $\mathbf{H} = [1, 0, 0]$. In EM-EKS and EM-EnKS, $\mathbf{Q}$ is initialized to $\mathbf{I}$, a value larger than the true model errors simulated below.

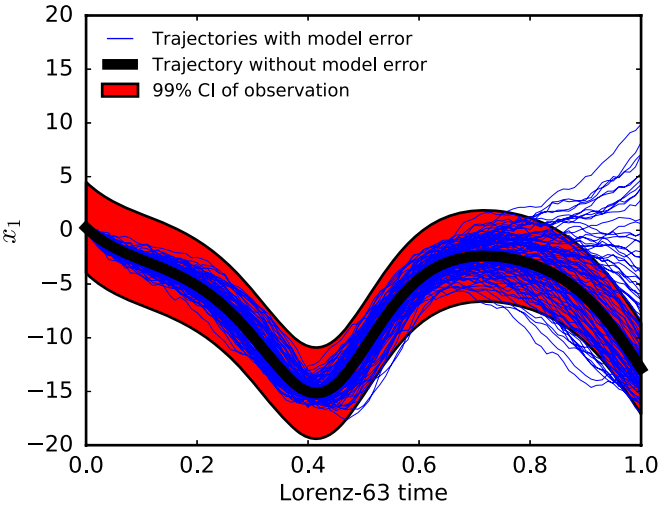

Figure 1. First variable of the L63 model for trajectories simulated with and without additive Gaussian model error, added every $\Delta t = 0.01$, with covariance $0.05\mathbf{I}$. The 99% prediction interval is of an observation of the first state variable of the error-free trajectory with observation error of covariance $2\mathbf{I}$. [Colour figure can be viewed at wileyonlinelibrary.com].

3.2. Experiments with additive model error

In the first experiments, the trajectories are simulated with Gaussian additive model errors. That is, a random Gaussian distributed error of covariance matrix $\mathbf{Q}^{\mathrm{t}}$ (true model covariance matrix) is added to the state at every model step $k = 1, \ldots, K$ of length $\mathrm{d}t = 0.01$.

$$\mathbf{x}_k^{\mathrm{t}} = f(\mathbf{x}_{k-1}^{\mathrm{t}}) + \boldsymbol{\eta}_k, \qquad \boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{\mathrm{t}}). \tag{49}$$

This model error assumption is the same as the one used to derive the EM algorithms in section 3. Thus, this experiment evaluates the ability of these algorithms to estimate $\mathbf{Q}$ in an ideal setting.

We compare the convergence of $\mathbf{Q}$ for the EM-EKS and the EM-EnKS algorithms as well as the log-likelihood and root mean squared error (RMSE, calculated over all states, sampled every $\mathrm{d}t = 0.01$). In the first part of this experiment, the true model-error covariance matrix is set to $\mathbf{Q}^{\mathrm{t}} = 0.05\mathbf{I}$. This value of $\mathbf{Q}^{\mathrm{t}}$ is small in comparison to $\mathbf{R}$ (40 times smaller). However, in a chaotic system like L63, small errors are quickly amplified.

To evaluate the sensitivity of the L63 system to this model error, 100 independent trajectories following (49) are simulated over 100 model steps ($\mathrm{d}t = 0.01$). The evolution of the first variable of L63 for each of these trajectories is shown in Figure 1 as well as the first variable of the reference trajectory (without error). A 99% prediction interval for an observation (with $\mathbf{R} = 2\mathbf{I}$) of the reference trajectory is shown. Initially, the trajectories are well bounded by the prediction intervals. However, they quickly diverge after $t = 0.8$ (80 model steps) and split between the two attractors.

To evaluate the convergence of the estimated $\mathbf{Q}$ with respect to the true value $\mathbf{Q}^{\mathrm{t}}$, we conducted another simulation where $\mathbf{Q}^{\mathrm{t}}$ is of the form $\sigma^2\mathbf{I}$, with $\sigma^2 \in \{0, 0.02, 0.04, \ldots, 0.2\}$. $\mathbf{Q}$ is estimated with assimilation cycles of different lengths (1, 4, 7, and 10 model steps), and the results are compared to the true value of $\mathbf{Q}^{\mathrm{t}} = 0.05\mathbf{I}$.

In real systems, we expect the model errors of different state variables to be correlated. Therefore, assuming $\mathbf{Q}^{\mathrm{t}}$ to be non-diagonal could be more realistic. However the inversion of full model-error covariance matrices when filtering in systems with high-dimensional states is prohibitive. In this case, low-dimensional parametrization of $\mathbf{Q}$ is needed. Therefore, we conducted an experiment where the true trajectory is generated with a non-diagonal $\mathbf{Q}^{\mathrm{t}}$ of the form:

$$\mathbf{Q}^{\mathrm{t}} = 0.05 \begin{pmatrix} 1.0 & 0.7 & 0.0 \\ 0.7 & 1.0 & 0.7 \\ 0.0 & 0.7 & 1.0 \end{pmatrix}. \tag{50}$$

We then estimate $\mathbf{Q}$ with the EM-EKS and the EM-EnKS algorithms, assuming the estimate is either full, diagonal or of the forms $\alpha\mathbf{I}$ or $\alpha\mathbf{B}$ (in which case the scalar α is estimated). We then compare the results of EM-EKS and EM-EnKS with these parametrizations of $\mathbf{Q}$ with runs of EKS and EnKS, assuming the true values $\mathbf{Q}^{\mathrm{t}}$.

3.3. *Parametric and multiplicative model error*

3.3.1. *Parametric error*

In this experiment, we consider a common source of model error: parameter error, that is, the error introduced in the model by misparametrization and/or parameter evolution. Here, we examine the effects of an inaccurate parametrization of parameter β from (48) of the L63 ODE system. The true trajectory is generated with β changing at every model step $\mathrm{d}t = 0.01$ and sampled from a uniform distribution $\beta \sim \mathcal{U}([\beta_{\min}, \beta_{\max}])$. If $f(\mathbf{x}; \beta)$ is the result of the state integration $\mathbf{x}$ over a $\mathrm{d}t$ time step with parameter β, then the true trajectory $\mathbf{x}_k^{\mathrm{t}}$ is generated in the following way:

$$\mathbf{x}_k^{\mathrm{t}} = f(\mathbf{x}_{k-1}^{\mathrm{t}}; \beta_k), \ \beta_k \sim \mathcal{U}([\beta_{\min}, \beta_{\max}]). \tag{51}$$

We set $\beta_{\min} = 8/3 - 2$ and $\beta_{\max} = 8/3 + 2$.

EM-EKS and EM-EnKS are used to estimate the covariance matrix $\mathbf{Q}$ of an additive Gaussian model error, that compensates for the parametric error. The algorithms are evaluated in the experiments (a)–(d) presented at the beginning of this section.

3.3.2. *Multiplicative error*

In this last experiment, we consider a multiplicative model error. At every model step, the current state is obtained by multiplying each state variable by the corresponding component of a multivariate lognormal error. That is, for $k = 1, \ldots, K$,

$$\begin{aligned} &[\mathbf{x}_k^{\mathrm{t}}]_l = [\boldsymbol{\nu}_k]_l . [f(\mathbf{x}_{k-1}^{\mathrm{t}})]_l, \\ &\text{with } l = 1, 2, 3, \quad \boldsymbol{\nu}_k \sim \log\mathcal{N}(\mathbf{0}, \mathbf{Q}^{\mathrm{mult}}), \end{aligned} \tag{52}$$

with $\mathbf{Q}^{\mathrm{mult}} = 0.001\mathbf{I}$, and $[\mathbf{x}]_l$ denotes the ith component of vector $\mathbf{x}$.

Multiplicative noise is useful to model uncertainties that depend on the state (Sura *et al.*, 2005), and have been used for modelling errors in precipitation data (Tian *et al.*, 2013). As with the parametric model error, we evaluate the ability of EM-EnKS and EM-EKS to estimate an appropriate covariance matrix of an additive Gaussian error.

4. Results

4.1. *Additive model error*

Figure 2 shows the RMSE and log-likelihood of the smoothed state after 100 iterations of the EM-EKS and EM-EnKS using 20, 100, and 500 ensemble members with the full state observed at every model step (experiment (a)). After around 80 EM iterations, both EM-EKS and EM-EnKS converge to a solution close to those of EnKS (RMSE 0.39, with 100 ensemble members) and EKS (RMSE 0.38) with the true values of $\mathbf{Q}$. EM-EKS and EM-EnKS with 500 ensemble members provide the best performance in terms of log-likelihood and RMSE (RMSE 0.38). As the ensemble size decreases, the EM-EnKS RMSE and log-likelihood degrade and their variability increases.

The estimated $\mathbf{Q}$ converges towards a value close to $\mathbf{Q}^{\mathrm{t}} = 0.05\mathbf{I}$. Figure 3 displays the convergence of the average of on-diagonal terms of the estimated $\mathbf{Q}$ with EM-EKS and EM-EnKS with the number of iterations. In all cases, the diagonal terms converge towards the true value of 0.05. The off-diagonal terms stay very close to zeros ($< 10^{-2}$). We note the faster convergence of EM-EnKS for smaller ensemble size and the small variability in the value of $\mathbf{Q}$ in each algorithm.

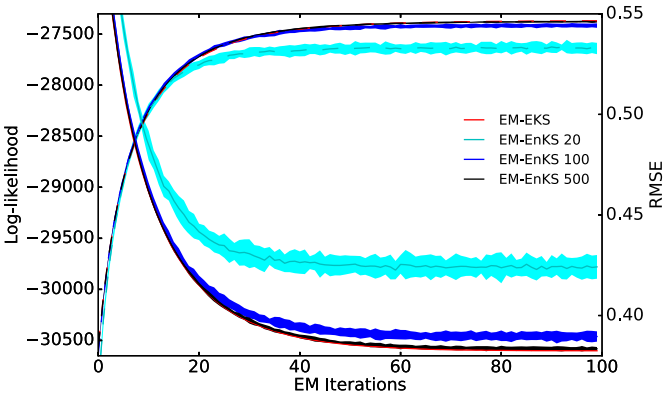

Figure 2. Evolution of the RMSE (solid line) and log-likelihood (dashed line) with the iterations of EM-EKS and EM-EnKS when the full state is observed every $\Delta t = 0.01$ with observation-error covariance $\mathbf{R} = 2\mathbf{I}$ and additive Gaussian model error of covariance $\mathbf{Q}^{\mathrm{t}} = 0.05$. EM-EnKS was repeated 20 times, and the solid and dashed lines represent the mean RMSE and log-likelihood respectively, while the light ribbon represents the 95% prediction interval. [Colour figure can be viewed at wileyonlinelibrary.com].

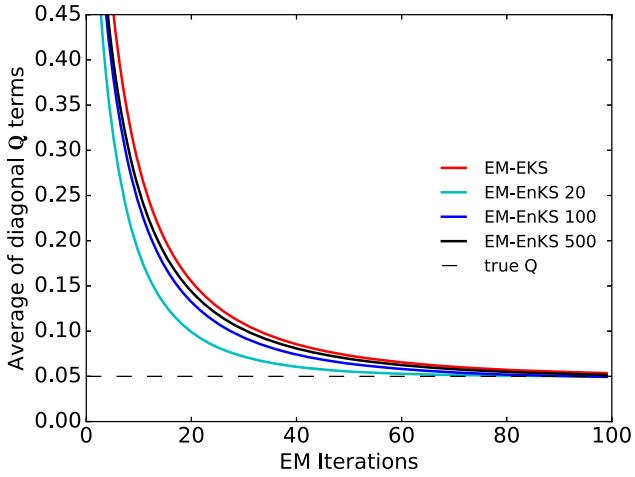

Figure 3. Evolution of the estimated model-error covariance $\mathbf{Q}$ with the iterations of EM-EKS and EM-EnKS when the full state is observed every $\Delta t = 0.01$ with observation-error covariance $\mathbf{R} = 2\mathbf{I}$ and additive Gaussian model error of covariance $\mathbf{Q}^{\mathrm{t}} = 0.05$. EM-EnKS was repeated 20 times, and the solid line represents the averaged $\mathbf{Q}$, while the light ribbon represents the 95% prediction interval (not visible if the prediction interval is too thin). [Colour figure can be viewed at wileyonlinelibrary.com].

When the full state is observed every 10 model steps (experiment (b)), we find that EM-EnKS log-likelihood and RMSE converge faster for smaller ensembles sizes. Overall, the RMSE of EM-EnKS (with 100 ensemble members) converges to 0.63 after 400 EM iterations, a value close to that of EKS and EnKS with the true values of $\mathbf{Q}$ (0.62 and 0.63 respectively). We also observe that EM-EKS either converges very slowly or converges towards a suboptimal local minimum solution (Figure 4) corresponding to a larger value of the estimated $\mathbf{Q}$ (Figure 5). We conducted further experiments (results not shown) in which we changed the size of the assimilation window and found that the performance of EM-EKS degrades with shorter assimilation windows. This seems to indicate that, with less frequent observations, the smoothing error increases because of the nonlinearity of the dynamical model. This may lead to an overestimation of $\mathbf{Q}$ and ultimately to slower convergence with the EM iterations, and increased log-likelihood and RMSE. On the other hand, larger assimilation windows smooth these errors in the estimation of $\mathbf{Q}$ (18) and lead to better performance for EM-EKS.

Experiments (c) and (d), with assimilation of observations of only the first state component, yield similar results, with a

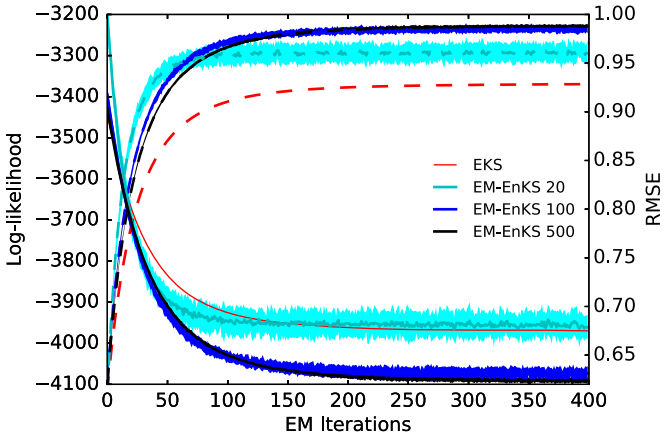

Figure 4. As Figure 2, but with the full state observed every $\Delta t = 0.1$. [Colour figure can be viewed at wileyonlinelibrary.com].

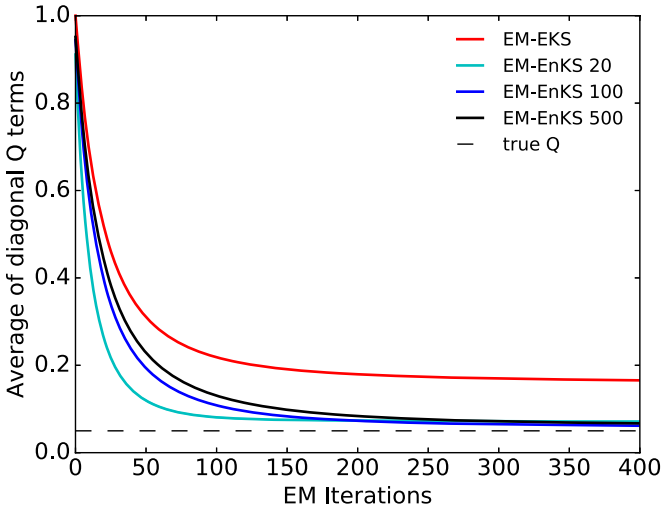

Figure 5. As Figure 3, but with the full state observed every $\Delta t = 0.1$ and with EM-EnKS repeated 20 times. [Colour figure can be viewed at wileyonlinelibrary.com].

faster convergence of EM-EnKS (in terms of log-likelihood and estimated **Q**) as the ensemble size gets smaller. In addition, EM-EKS provides better performance than EM-EnKS for short assimilation time steps, but diverges when the time between successive observations increases.

Figure 6 shows the diagonal average and spread of the estimated **Q** for EM-EnKS with 100 ensemble members for increasing values of $\mathbf{Q}^t$ and duration between observations. There is a clear linear relationship between $\mathbf{Q}^t$ and **Q** estimated with a slope close to 1. There is a tendency of EM-EnKS to slightly overestimate **Q**, which is more pronounced as the interval between observations increases. This suggests that the EM algorithm compensates for nonlinearities in the dynamical model by increasing the model-error covariance.

Tables 1 and 2 show the results of EM-EKS and EM-EnKS (100 ensemble members) on trajectories with a full $\mathbf{Q}^t$ and the form of **Q** to be estimated is assumed to be either full, diagonal, or of the form $\alpha\mathbf{I}$ or $\alpha\mathbf{B}$, where the scalar α is estimated. Table 1 shows the RMSE and log-likelihood obtained with these different parametrizations as well as those obtained with the true value of **Q**, with the full state observed every model step (experiment (a)). EM-EKS performs better than EM-EnKS, and their performances are very close to those with the true **Q** value. The parametrization of **Q** only slightly degrades the performances of the EM-EKS and EM-EnKS algorithms.

Table 2 outlines the results of the same experiment with the full state observed every 10 model steps (experiment (b)). As in the previous experiment, no significant degradation is found when **Q** is parametrized. EM-EnKS is slightly better than

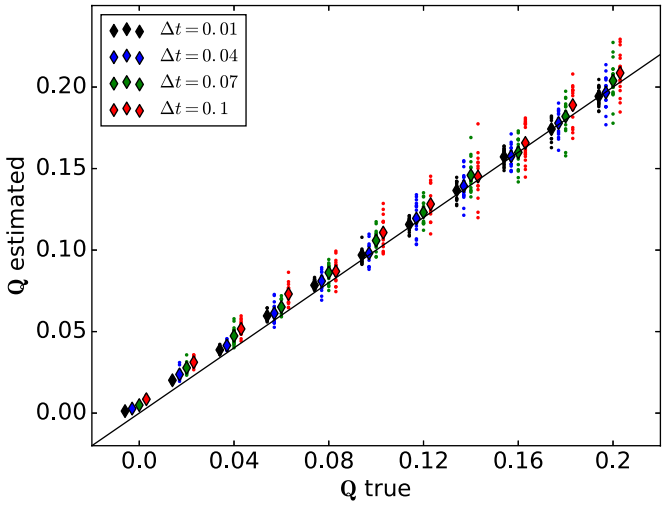

Figure 6. Average diagonal of **Q** estimated after 500 iterations of EM-EnKS (100 ensemble members) for different $\mathbf{Q}^t$ and assimilation times Δt between observations. The dots show the RMSE obtained with single runs of EM-EnKS, and the diamonds show the RMSE averaged over 20 runs of RM-EnKS. [Colour figure can be viewed at wileyonlinelibrary.com].

Table 1. Performance of EKS, EnKS, EM-EKS and EM-EnKS with the full state observed every $\Delta t = 0.01$ (10 000 observations) with $\mathbf{R} = 2\mathbf{I}$.

		EKS[a]	EnKS[a,b,c]
	RMSE	0.36	0.37 (±0.02)
	LL*	−27 270	−27 321 (±16)
Q type		EM-EKS[d]	EM-EnKS[b,c,d]
Full	RMSE	0.36	0.37 (±0.0)
	LL	−27 290	−27 328 (±16)
Diagonal	RMSE	0.38	0.39 (±0.0)
	LL	−27 428	−27 491 (±11)
$\alpha\mathbf{I}$	RMSE	0.38	0.39 (±0.0)
	LL	−27 429	−27 498 (±18)
$\alpha\mathbf{B}$	RMSE	0.37	0.38 (±0.0)
	LL	−27 362	−27 411 (±24)

* LL = Log-likelihood.
[a] Assuming the model-error covariance is $\mathbf{Q}^t$.
[b] With 100 ensemble members.
[c] Mean and 95% prediction interval based on 20 runs.
[d] With 500 EM iterations.

Table 2. As Table 1, but for $\Delta t = 0.1$ (1000 observations).

		EKS[a]	EnKS[a,b,c]
	RMSE	0.62	0.70 (±0.14)
	LL	−3258	−3271 (±8)
Q type		EM-EKS[d]	EM-EnKS[b,c,d]
Full	RMSE	0.67	0.64 (±0.01)
	LL	−3359	−3282 (±11)
Diagonal	RMSE	0.68	0.65 (±0.01)
	LL	−3360	−3289 (±11)
$\alpha\mathbf{I}$	RMSE	0.68	0.65 (±0.0)
	LL	−3358	−3290 (±9)
$\alpha\mathbf{B}$	RMSE	0.68	0.65 (±0.01)
	LL	−3405	−3308 (±9)

EM-EKS and its RMSE is very close to that of the EKS and EnKS with the true values of **Q**. However, when increasing the length of the assimilation window to $T = 200$, the difference between EM-EnKS and EM-EKS becomes very small, indicating again that EM-EKS converges faster when more data are available for estimating **Q** (results not shown).

When only the first state component is observed (experiments (c) and (d)), the convergence of the EM algorithm is degraded. In particular, EM-EKS diverges when the duration

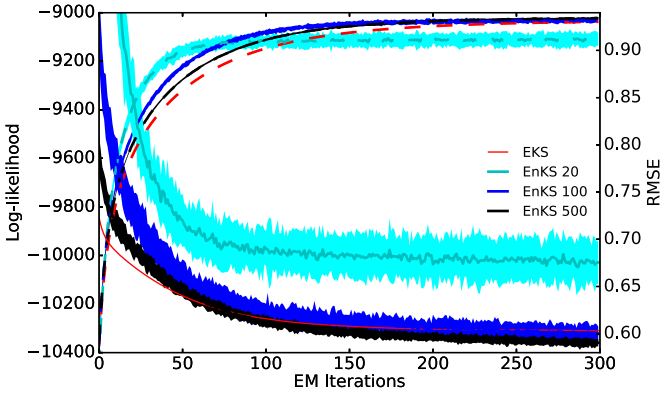

Figure 7. Evolution of the RMSE (solid line) and log-likelihood (dashed line) with the iterations of EM-EKS and EM-EnKS when the first state variable is observed every $\Delta t = 0.01$ with observation-error covariance $\mathbf{R} = 2\mathbf{I}$ and parametric model EM-EnKS was repeated 20 times. Line styles are as Figure 2. [Colour figure can be viewed at wileyonlinelibrary.com].

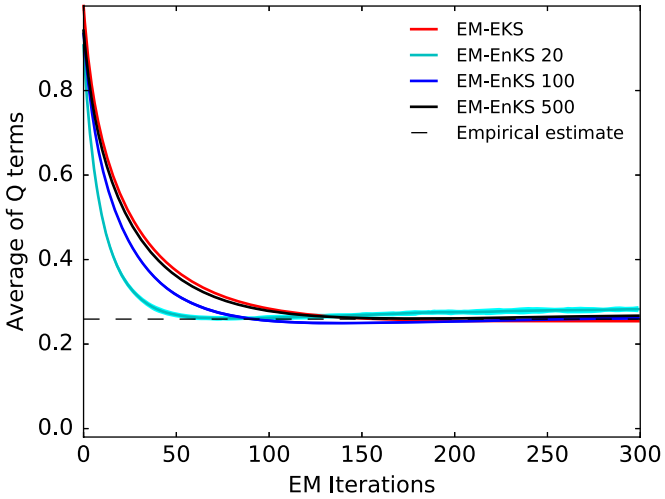

Figure 8. Evolution of the estimated model-error covariance $\mathbf{Q}$ with the iterations of EM-EKS and EM-EnKS when the state variable is observed every $\Delta t = 0.01$ with observation-error covariance $\mathbf{R} = 2\mathbf{I}$ and multiplicative model with $\mathbf{Q}^{\text{mult}} = 0.001\mathbf{I}$. EM-EnKS was repeated 20 times, and line styles are as Figure 2. [Colour figure can be viewed at wileyonlinelibrary.com].

between successive observations increases. We also notice that parametrizing $\mathbf{Q}$ improves the results, particularly when fewer observations are available.

4.2. *Parametric and multiplicative model error*

The same experiments as those described in Figures 2 and 4 were performed, assimilating observations generated from true trajectories with parametric and multiplicative error assumptions. In these cases, there is no $\mathbf{Q}^{\text{t}}$ to which the estimated $\mathbf{Q}$ should converge, as the true model errors are non-additive. However, we estimate an appropriate $\mathbf{Q}$ that compensates for them. In general, the results are very similar to the additive model error assumptions for the convergence of log-likelihood, RMSE, and estimated $\mathbf{Q}$. In the case of parametric noise, we can see in Figure 7 that when the first state variable is observed at every $\Delta t = 0.01$, the log-likelihood of EM-EnKS and EM-EKS converge. The convergence of EM-EnKS is better with a large ensemble, whereas with small ensemble sizes, the algorithm converges towards suboptimal values of log-likelihood and RMSE. Figure 8 shows the convergence of the estimated $\mathbf{Q}$ when the first state variable is observed at every $\Delta t = 0.01$, assuming a multiplicative model error. The estimated $\mathbf{Q}$ rapidly converges towards a value close to that of the empirically estimated $\mathbf{Q}$. The empirically estimated $\mathbf{Q}$ is computed by taking the sample covariance of the difference between the true trajectory and the trajectory estimated by the smoothing algorithm.

5. Conclusion

In this work, we focus on estimating the model-error covariance matrix $\mathbf{Q}$ for a state equation with an additive Gaussian model error (as stated in the classical Kalman framework). A proper tuning of this parameter is expected to improve the estimation of the hidden state and avoid filter divergence. We explore the possibility of this term to take into account classical model error structures including additive, parametric and multiplicative error terms. To do this, we extended the EM algorithm to the case of nonlinear dynamics, using the EnKS as proposed by (Tandeo *et al.*, 2015). The main objective of this procedure is to approximate the maximum likelihood estimators of the parameters of the state-space model, that is the error covariance matrices $\mathbf{Q}$ and $\mathbf{R}$ along with the state initial background conditions $\mathbf{x}^{\text{b}}$ and $\mathbf{B}$. The expectation step is implemented using classical Kalman-like recursion approximations (both for filtering and smoothing): the extended (EKF/EKS) and the ensemble (EnKF/EnKS) versions. For the maximization step, under the Gaussian assumption, analytical expressions for the optimal parameters are obtained. These expressions are directly evaluated in the EM-EKS, whereas sample estimates are computed in the EM-EnKS.

The case of simulated additive and Gaussian model errors corresponds to a twin experiment that allows us to evaluate the performance of the EM algorithms. We demonstrate that, even when very few data are available (only one model variable observed at scattered times) and for different forms of $\mathbf{Q}$ (diagonal or full, parametric or not), the solutions of the proposed algorithms converge to the true covariances. Moreover, experiments examining the estimated model-error covariance matrix for increasing amplitudes of $\mathbf{Q}$ and for different frequencies of observations indicated that the estimate is unbiased. When comparing EM-EKS to EM-EnKS, we find that the former achieves the best performances when the frequency of the observations is high. This suggests that when the linearization assumption holds, EKS is a better approximation than EnKS. However, when the time interval between observations increases or when the length of the assimilation window decreases, EM-EKS either converges to a local minimum or diverges. On the other hand, EM-EnKS is more robust and its log-likelihood converged in all the experiments we performed, given a large enough ensemble. EM-EnKS also tends to overestimate $\mathbf{Q}$ (and consistently so for $\mathbf{Q} = 0$ and very small values), and this likely prevents the ensemble from collapsing.

Concerning the experiments with non-additive model error (i.e. parametric and multiplicative), results are similar to those of the additive error model. This shows the ability of the additive model error assumption to efficiently deal with diverse sources of model uncertainties – confirming the findings of (Houtekamer *et al.*, 2009) – and that the EM algorithm is an efficient strategy to estimate an accurate error covariance matrix. This property may be of paramount importance for the application of EM to real systems, where the sources of model error are difficult to characterize.

In our experiments, we addressed different concerns for the use of the EM algorithm for estimating $\mathbf{Q}$ in large-scale systems. One of these concerns is the computational burden of estimating and using the full covariance matrix. In practice, $\mathbf{Q}$ is often chosen low-rank, and we tested the use of simple parametrizations to reduce the number of coefficients to estimate. We found that the parametrization only incurred a small degradation of performance. The number of iterations required for the convergence of the EM algorithm is also a concern in high-dimensional applications. However twin experiments with large models are needed to explore this issue in detail. One of the issues with our current implementation, which is based on the stochastic EnKF, is the need of a large ensemble (100–500) for best performances. For real large-scale data assimilation systems, alternative filtering schemes could be considered to reduce the ensemble size, such as the deterministic ensemble Kalman filter. Additionally, the RTS smoother is memory-inefficient in

large-scale systems as it requires saving the ensemble forecast and analysis states over the smoothing window (Raanes, 2016). More efficient smoothing algorithms could be considered, such as the Ensemble Smoothing (Cosme *et al.*, 2012), which has been shown to be equivalent to the ensemble formulation of the RTS smoother (Raanes, 2016) and consists of smoothing the whole state sequence by defining an augmented (full) state vector and performing a BLUE analysis using the complete sequence of observations.

In this work, we make an important assumption on the form of the model-error covariances matrices: we assume that they are constant in time. In practice, this is a very convenient hypothesis allowing us to retrieve robust maximum likelihood estimators. However, this assumption is not completely realistic because model errors generally evolve in space and time. A possible adaptive solution to tackle this problem could be to apply the EM procedure on different assimilation windows to track possible changes in covariance coefficients.

This work proposes general algorithms (EM-EnKS and EM-EKS) to estimate the model-error covariance using a state-space model with a nonlinear dynamical model. Although the focus of this work is on the model error, these algorithms can also estimate the observation-error covariance and the initial state parameters. An advantage of the ensemble-based version of the EM algorithm is its independence from the dynamical model. The open-source Python library is available (http://github.com/ptandeo/CEDA; accessed 23 April 2017) so that interested users can easily test the algorithm on their own dynamical models.

Acknowledgement

We are thankful to the two reviewers whose constructive comments helped significantly to improve this work.

References

Aanonsen SI, Nævdal G, Oliver DS, Reynolds AC, Vallès B. 2009. The ensemble kalman filter in reservoir engineering–a review. *SPE J.* **14**: 393–412.

Ait-El-Fquih B, Desbouvries F. 2008. On Bayesian fixed-interval smoothing algorithms. *IEEE Trans. Autom. Control Inst. Electr. Electron. Eng.* **53**: 2437–2442.

Anderson BDO, Moore JB. 1979. *Optimal Filtering*. Dover Publications: Mineola, NY.

Berry T, Sauer T. 2013. Adaptive ensemble Kalman filtering of non-linear systems. *Tellus Ser. A–Dyn. Meteorol. Oceanogr.* **65**: 20331. https://doi.org/10.3402/tellusa.v65i0.20331.

Burgers G, van Leeuwen PJ, Evensen G. 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.* **126**: 1719–1724.

Chitralekha SB, Prakash J, Raghavan H, Gopaluni RB, Shah SL. 2010. A comparison of simultaneous state and parameter estimation schemes for a continuous fermentor reactor. *J. Process Control* **20**: 934–943.

Chui CK, Chen G. 2009. *Kalman Filtering, with Real-Time Applications*. Springer: Berlin.

Cosme E, Verron J, Brasseur P, Blum J, Auroux D. 2012. Smoothing problems in a Bayesian framework and their linear Gaussian solutions. *Mon. Weather Rev.* **140**: 683–695.

Daley R. 1992. Estimating model-error covariances for application to atmospheric data assimilation. *Mon. Weather Rev.* **120**: 1735–1746.

Dee DP. 1995. Online estimation of error covariance parameters for atmospheric data assimilation. *Mon. Weather Rev.* **123**: 1128–1145.

Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B. Methodol.* **39**: 1–38.

Evensen G. 2009. *Data Assimilation: The Ensemble Kalman Filter*. Springer: Berlin.

Hairer E, Nørsett SP, Wanner G. 2009. *Solving Ordinary Differential Equations I. Non-stiff Problems* (2nd edn). Springer: Berlin.

Hoteit I, Pham DT, Blum J. 2002. A simplified reduced order Kalman filtering and application to altimetric data assimilation in tropical pacific. *J. Mar. Syst.* **36**: 101–127.

Hoteit I, Triantafyllou G, Korres G. 2007. Using low-rank ensemble Kalman filters for data assimilation with high-dimensional imperfect models. *J. Numer. Anal. Ind. Appl. Math.* **2**: 67–78.

Hoteit I, Pham DT, Gharamti ME, Luo X. 2015. Mitigating observation perturbation sampling errors in the stochastic ENKF. *Mon. Weather Rev.* **143**: 2918–2936.

Houtekamer PL, Mitchell HL, Deng X. 2009. Model error representation in an operational ensemble Kalman filter. *Mon. Weather Rev.* **137**: 2126–2143.

Kulkarni JE, Paninski L. 2007. Common-input models for multiple neural spike-train data. *Network-Comput. Neural Syst.* **18**: 375–407.

Lahoz WA, Khattatov B, Ménard R. 2010. Data Assimilation and Information, Chapter 1 in *Data Assimilation: Making Sense of Observations*, 3–12. Springer: Berlin.

Lorenz EN. 1963. Deterministic non-periodic flow. *J. Atmos. Sci.* **20**: 130–141.

Luo X, Hoteit I, Duan L, Wang W. 2012. Review of nonlinear Kalman, ensemble and particle filtering with application to the history matching problem. Chapter 7 in *Nonlinear Estimation and Applications to Industrial Systems Control*. Rigatos G. (ed.): 197–224. Nova Science Publishers: Hauppauge, NY.

Mitchell HL, Houtekamer PL. 2000. An adaptive ensemble Kalman filter. *Mon. Weather Rev.* **128**: 416–433.

Raanes PN. 2016. On the ensemble Rauch–Tung–Striebel smoother and its equivalence to the ensemble Kalman smoother. *Q. J. R. Meteorol. Soc.* **142**: 1259–1264. https://doi.org/10.1002/qj.2728.

Rauch HE, Striebel C, Tung F. 1965. Maximum likelihood estimates of linear dynamic systems. *AIAA J.* **3**: 1445–1450.

Shumway BH, Stoffer DS. 1982. An approach to time series smoothing and forecasting using the EM algorithm. *J. Time Ser. Anal.* **3**: 253–264. https://doi.org/10.1111/j.1467-9892.1982.tb00349.x.

Smith A, Doucet A, de Freitas N, Gordon N. 2013. *Sequential Monte Carlo Methods in Practice*. Springer: Berlin.

Sura P, Newman M, Penland C, Sardeshmukh P. 2005. Multiplicative noise and non-Gaussianity: A paradigm for atmospheric regimes? *J. Atmos. Sci.* **62**: 1391–1409.

Tandeo P, Pulido M, Lott F. 2015. Offline parameter estimation using EnKF and maximum likelihood error covariance estimates: Application to a subgrid-scale orography parametrization. *Q. J. R. Meteorol. Soc.* **141**: 383–395. https://doi.org/10.1002/qj.2357.

Ueno G, Nakamura N. 2014. Iterative algorithm for maximum-likelihood estimation of the observation-error covariance matrix for ensemble-based filters. *Q. J. R. Meteorol. Soc.* **140**: 295–315.

Ueno G, Nakamura N. 2016. Bayesian estimation of the observation-error covariance matrix in ensemble-based filters. *Q. J. R. Meteorol. Soc.* **142**: 2055–2080. https://doi.org/10.1002/qj.2803.

Ueno G, Higuchi T, Kagimoto T, Hirose N. 2010. Maximum likelihood estimation of error covariances in ensemble-based filters and its application to a coupled atmosphere–ocean model. *Q. J. R. Meteorol. Soc.* **136**: 1316–1343. https://doi.org/10.1002/qj.654.

