

Accounting for model error due to unresolved scales within ensemble Kalman filtering

Lewis Mitchell^{a,b,*} and Alberto Carrassi^{c,d}

^aDepartment of Mathematics & Statistics, University of Vermont, Burlington, USA

^bSchool of Mathematical Sciences, University of Adelaide, SA, Australia

^cNansen Environmental and Remote Sensing Center, Bergen, Norway

^dInstitut Catala de Ciències del Clima, Barcelona, Spain

*Correspondence to: L. Mitchell, Ingkarni Wardli Building, North Terrace Campus, University of Adelaide, SA 5005, Australia.
E-mail: lewis.mitchell@adelaide.edu.au

We propose a method to account for model error due to unresolved scales in the context of the ensemble transform Kalman filter (ETKF). The approach extends to this class of algorithms the deterministic model error formulation recently explored for variational schemes and extended Kalman filter. The model error statistic required in the analysis update is estimated using historical reanalysis increments and a suitable model error evolution law. Two different versions of the method are described; a time-constant model error treatment where the same model error statistical description is time-invariant, and a time-varying treatment where the assumed model error statistics is randomly sampled at each analysis step. We compare both methods with the standard method of dealing with model error through inflation and localization, and illustrate our results with numerical simulations on a low-order nonlinear system exhibiting chaotic dynamics. The results show that the filter skill is significantly improved through the proposed model error treatments, and that both methods require far less parameter tuning than the standard approach. Furthermore, the proposed approach is simple to implement within a pre-existing ensemble-based scheme. The general implications for the use of the proposed approach in the framework of square-root filters such as the ETKF are also discussed.

Key Words: model error; data assimilation; ensemble Kalman filter; ensemble prediction; bias correction

Received 9 April 2014; Revised 18 August 2014; Accepted 1 September 2014; Published online in Wiley Online Library

1. Introduction

A fundamental concern within all data assimilation schemes is the treatment of model error, which can arise from, amongst other sources, the impossibility of characterizing phenomena at all scales of motion. It is inevitable that there will be some processes which remain unresolved in any numerical environmental model, making the problem of model error of fundamental importance.

Representing model error within data assimilation is an important consideration within modern-day operational data assimilation algorithms, and there has appeared in recent years a significant body of work dedicated to this problem. The state augmentation method (Jazwinski, 1970) was primarily introduced in the context of Kalman filtering. In this method, the state estimation problem is formulated in terms of an augmented state vector which includes, along with the state estimate, a set of parameters used for the model error representation. This approach has been applied successfully in both Kalman filter-like and variational assimilation methods (Zupanski, 1997; Nichols, 2003; Zupanski and Zupanski, 2006). Dee and Da Silva (1998) proposed an algorithm to estimate and remove the biases in the background field in a data assimilation system due to additive

systematic model errors. This method was implemented for the bias correction of the humidity analysis component of the Goddard Earth Observing System (GEOS) assimilation system (Dee and Todling, 2000) and more recently in the context of the European Centre for Medium-range Weather Forecasts (ECMWF) ocean data assimilation (Balmaseda *et al.*, 2007). A key ingredient of the state augmentation technique is the definition of a 'model' for the model error (e.g. Nichols, 2003). Recently Zupanski and Zupanski (2006) have described the model error evolution using a first-order Markov process. A similar assumption was already used in Daley (1992) to investigate the impact of time-correlated model errors in Kalman filtering. In Dee (2005), the model bias was modelled based on unknown bias parameters to be determined and on the use of analysis increments; these latter are also used here as explained in Section 4.

In ensemble-based data assimilation procedures, Houtekamer *et al.* (2009) examine both isotropic and stochastic model error representations within an operational numerical weather prediction (NWP) system, while Charron *et al.* (2010) study stochastic methods within the same system. One of the most direct methods to account for model error in ensemble filtering is

through tuning the inflation and localization parameters. These parameters, which were originally introduced to ameliorate the covariance underestimation (Anderson and Anderson, 1999) and spurious long-range correlations (Houtekamer and Mitchell, 1998; Hamill, 2001) associated with finite ensemble size effects, nonetheless allow the user some freedom to compensate for other sources of error including model error. Covariance inflation in particular can be useful for achieving this aim, as it has the effect of decreasing the amount by which the analysis ‘trusts’ the model forecast relative to the observations. Of course this parameter requires empirical tuning, which may be computationally expensive. Regardless, inflation has been shown to be an effective approach for dealing with model error in operational models (Deng *et al.*, 2011; Raynaud *et al.*, 2012).

We propose a deterministic approach to the model error problem in ensemble data assimilation, inspired by previous works on model error treatment for both variational (Carrassi and Vannitsem, 2010) and sequential (Carrassi and Vannitsem, 2011) schemes. This method assumes that the model error is a time-correlated process in order to derive expressions for the model error bias and covariance, along with an approximated dynamical law suitable for practical applications. The bias and covariance are then approximated from an assumed database of historical analysis increments.

Recognizing that there are many ways in which this information about model error may be incorporated into the data assimilation process, we investigate two ways in which we might implement our model error treatment within an ensemble Kalman filter. The first method treats model error as deterministic and constant in time similar to Carrassi and Vannitsem (2011), while the second allows for the model error correction to vary in time in a random manner, but where the model noise is sampled from the model error statistics derived using the deterministic hypothesis. The latter approach leads to a randomly perturbed forecast model somewhat analogous to stochastic climate models, which have been shown to be beneficial in ensemble data assimilation schemes (Harlim and Majda, 2008; Mitchell and Gottwald, 2012). The proposed approaches are studied and compared in the framework of the ensemble transform Kalman filter (ETKF; Bishop *et al.*, 2001), a prototype square-root filter (Tippett *et al.*, 2003). The implications of this choice for the performance of the model error treatment are also discussed. The advantages and drawbacks of the method in relation with standard inflation/localization procedures for treating model error are highlighted with regards to their implementation in a more realistic model and observational scenario.

An outline of the remainder of the article is as follows: in section 2 we formulate the problem of model error due to unresolved scales as a time-correlated process, and derive expressions for the model error bias and covariance in the short-time regime. In section 3 we describe the ETKF, and discuss strategies for dealing with model error in section 4. We present numerical results applying these methods to a low-dimensional nonlinear model with chaotic dynamics in section 5, and conclude with a discussion of our results in section 6.

2. Problem formulation

Let the model at our disposal be represented as:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}, \boldsymbol{\lambda}), \quad (1)$$

where \mathbf{f} is typically a nonlinear function, defined in \mathbb{R}^N and $\boldsymbol{\lambda}$ is a P -dimensional vector of parameters.

Model (1) is used to describe the evolution of a (unknown) ‘true’ dynamics or *nature*, whose evolution is assumed to be given by the following coupled equations:

$$\frac{d\widehat{\mathbf{x}}(t)}{dt} = \widehat{\mathbf{f}}(\widehat{\mathbf{x}}, \widehat{\boldsymbol{\lambda}}), \quad \frac{d\widehat{\boldsymbol{\lambda}}(t)}{dt} = \widehat{\mathbf{g}}(\widehat{\mathbf{x}}, \widehat{\boldsymbol{\lambda}}), \quad (2)$$

where $\widehat{\mathbf{x}}$ is a vector in \mathbb{R}^N , and $\widehat{\boldsymbol{\lambda}}$ is defined in \mathbb{R}^L and may represent scales that are present in the real world, but are neglected in model (1); the unknown parameter $\widehat{\boldsymbol{\lambda}}$ has dimension P' different from P . The true state is thus a vector of dimension $N + L$. While the model state vector \mathbf{x} and the variable $\widehat{\mathbf{x}}$ of the true dynamics span the same phase space, the difference in the functions \mathbf{f} and $\widehat{\mathbf{f}}$ implies that they do not in general have the same attractor.

When using the model (1) to describe the evolution of $\widehat{\mathbf{x}}$, estimation error can arise from the uncertainty in the initial conditions in the resolved scale ($\mathbf{x}(t_0) \neq \widehat{\mathbf{x}}(t_0)$) and from the approximate description of the nature afforded by Eq. (1); we refer to this type of error as *model error*. A number of different sources of model error are present in environmental modelling. Typical examples are those arising from the inadequate description of some physical processes, numerical discretization and/or the presence of scales in the actual dynamics that are unresolved by the model. In this study, we focus on the latter and assume furthermore that the model parameter $\widehat{\boldsymbol{\lambda}}$ is perfectly known so that $\boldsymbol{\lambda} = \widehat{\boldsymbol{\lambda}}$.

Following the approach outlined by Nicolis (2004), we derive the evolution equations for the dominant moments, the mean and covariance, of the estimation error $\delta\mathbf{x} = \mathbf{x} - \widehat{\mathbf{x}}$ in the resolved scale (i.e. in \mathbb{R}^N). The formal solutions of Eqs (1) and (2) read respectively:

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t d\tau \mathbf{f}(\mathbf{x}(\tau), \boldsymbol{\lambda}), \quad (3)$$

$$\widehat{\mathbf{x}}(t) = \widehat{\mathbf{x}}_0 + \int_0^t d\tau \widehat{\mathbf{f}}(\widehat{\mathbf{x}}(\tau), \widehat{\boldsymbol{\lambda}}), \quad (4)$$

where $\mathbf{x}_0 = \mathbf{x}(t_0)$, and $\widehat{\mathbf{x}}_0 = \widehat{\mathbf{x}}(t_0)$. By taking the difference between Eqs (3) and (4), and averaging over an ensemble of perturbations around the reference trajectory, we obtain the evolution equation for the mean error, or bias:

$$\langle \delta\mathbf{x}(t) \rangle = \langle \delta\mathbf{x}_0 \rangle + \int_0^t d\tau \langle \delta\mathbf{f}(\tau) \rangle, \quad (5)$$

with $\delta\mathbf{x}_0 = \mathbf{x}_0 - \widehat{\mathbf{x}}_0$ and $\delta\mathbf{f}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\lambda}) - \widehat{\mathbf{f}}(\widehat{\mathbf{x}}(t), \widehat{\boldsymbol{\lambda}})$.

With the hypothesis that the initial condition is unbiased, $\langle \delta\mathbf{x}_0 \rangle = 0$, Eq. (5) gives the evolution equation of the bias due to the model error, usually referred to as drift in the context of climate prediction (Doblas-Reyes *et al.*, 2013). The important factor driving the drift is the difference between the true and modelled velocity fields, $\langle \delta\mathbf{f}(\tau) \rangle$. We treat this difference as being correlated in time, and expand Eq. (5) in a Taylor time series around $t_0 = 0$ up to the first non-trivial order. By assuming an unbiased initial condition this expansion reads:

$$\mathbf{b}_m = \langle \delta\mathbf{x}(t) \rangle \approx \langle \delta\mathbf{f}(\tau) \rangle t. \quad (6)$$

Equation (6) gives the evolution of the bias, \mathbf{b}_m , the drift, in the linear and short-time approximation and the subscript ‘m’ stands for model error-related bias.

Similarly, by taking the expectation of the external product of the error anomalies $\delta\mathbf{x} - \langle \delta\mathbf{x} \rangle$ with themselves, we have:

$$\begin{aligned} \mathbf{P}(t) &= \langle \{\delta\mathbf{x}(t)\} \{\delta\mathbf{x}(t)\}^T \rangle \\ &= \langle (\delta\mathbf{x}_0) (\delta\mathbf{x}_0)^T \rangle + \left\langle (\delta\mathbf{x}_0) \left(\int_0^t d\tau \delta\mathbf{f}(\tau) \right)^T \right\rangle \\ &\quad + \left\langle \left(\int_0^t d\tau \delta\mathbf{f}(\tau) \right) (\delta\mathbf{x}_0)^T \right\rangle \\ &\quad + \left\langle \int_0^t d\tau \int_0^t d\tau' \{\delta\mathbf{f}(\tau)\} \{\delta\mathbf{f}(\tau')\}^T \right\rangle. \end{aligned} \quad (7)$$

Equation (7) describes the time evolution of the estimation error covariance in the resolved scale. The first term on the

right-hand side, which does not depend on time, represents the covariance of the initial error. The two following terms account for the correlation between the error in the initial condition and the model error, while the last term combines their effect on the evolution of the estimation error covariance. Following a standard hypothesis in most data assimilation procedures (Jazwinski, 1970), we assume that model and initial condition error do not correlate with each other, so that the second and third terms on the RHS of Eq. (7) are set to zero. By further assuming that the resolved-scale initial condition error is zero, Eq. (7) becomes

$$\mathbf{P}(t) = \left\langle \int_0^t d\tau \int_0^{\tau} d\tau' \{ \delta \mathbf{f}(\tau) \} \{ \delta \mathbf{f}(\tau') \}^T \right\rangle. \quad (8)$$

Equation (8) represents the forecast-error covariance in a deterministic process starting from perfect initial conditions but in presence of error due to the unresolved scale $\widehat{\mathbf{y}}$. The amplitude and structure of this covariance depend on the dynamical properties of the difference between the nature and model velocity fields. Note that the hypothesis of perfect resolved-scale initial conditions is only functional to our discussion in order to isolate the contribution to forecast error related to the unresolved scale. In realistic applications, $\delta \mathbf{x}_0 \neq 0$ and the use of Eq. (8) requires some tuning; in sections 4 and 5, we will discuss how to address this issue by means of correctly weighting the corrections due to model error. Furthermore, our assumption that initial condition and model errors do not correlate is only valid for short times. Over longer times the model error feeds into the state-estimation error and its evolution begins to be dominated by the Lyapunov modes so that model and initial-condition error become naturally correlated. We investigate over what forecast windows the short-time approximation is numerically reasonable in section 5.

Assuming as above that these differences are correlated in time, we can expand Eq. (8) in a Taylor series up to the first non-trivial order around the arbitrary initial time $t_0 = 0$, and obtain

$$\mathbf{P}_m(t) \approx \langle \{ \delta \mathbf{f}(0) \} \{ \delta \mathbf{f}(0) \}^T \rangle t^2. \quad (9)$$

Equation (9) describes the short-time evolution of the estimation error covariance when the initial condition error is uncorrelated with the model error. Note that, if the terms $\mathbf{f} - \widehat{\mathbf{f}}$ are delta-correlated, as in the case of uncorrelated model error, the short-time evolution of $\mathbf{P}(t)$ is bound to be linear instead of quadratic. This distinctive feature is relevant in data assimilation applications where model error is often assumed to be uncorrelated in time, a choice that can also reduce the computational cost, particularly in the case of variational assimilation (Tremolet, 2006; Carrassi and Vannitsem, 2010).

3. The ensemble transform Kalman filter (ETKF)

Let us assume that a set of $M \leq N$ noisy observations of the resolved scale, $\widehat{\mathbf{x}}$, is available at the regularly spaced discrete times $t_j = t_0 + j\tau$, $j = 1, 2, \dots$, with τ being the assimilation interval. The observations are stored as components of an M -dimensional observation vector \mathbf{y}^o , so that:

$$\mathbf{y}^o(t_j) = \mathbf{y}_j^o = \mathcal{H}(\widehat{\mathbf{x}}_j) + \boldsymbol{\epsilon}^o, \quad (10)$$

where $\widehat{\mathbf{x}}_j = \widehat{\mathbf{x}}(t_j)$ and $\boldsymbol{\epsilon}^o$ is the observation error, assumed to be Gaussian, uncorrelated in time, and with known covariance matrix \mathbf{R} . \mathcal{H} is the (possibly nonlinear) observation operator which maps the model solution to the observation space, and may involve spatial interpolations (or spectral to physical space transformation in the case of spectral models) as well as transformations based on physical laws for indirect measurements (Kalnay, 2002). To simplify notation, the time dependency is removed hereafter from all vectors and matrices, which are then assumed to be evaluated at the same arbitrary time unless specified differently.

In ensemble Kalman filters (EnKFs), an ensemble with k members \mathbf{x}_k

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k] \in \mathbb{R}^{N \times k}$$

is propagated by the full nonlinear model dynamics (Eq. (1)) according to

$$\frac{d\mathbf{X}}{dt} = \mathbf{f}(\mathbf{X}), \quad \mathbf{f}(\mathbf{X}) = [\mathbf{f}(\mathbf{x}_1), \mathbf{f}(\mathbf{x}_2), \dots, \mathbf{f}(\mathbf{x}_k)] \in \mathbb{R}^{N \times k} \quad (11)$$

to produce a forecast ensemble \mathbf{X}^f ; all members evolve using the same $\boldsymbol{\lambda}$. This ensemble is split into its mean

$$\mathbf{x}^f = \frac{1}{k} \sum_{i=1}^k \mathbf{x}_i^f = \mathbf{X}^f \mathbf{w} \quad \text{with} \quad \mathbf{w} = \frac{1}{k} \mathbf{e} \in \mathbb{R}^k,$$

where $\mathbf{e} = [1, \dots, 1]^T \in \mathbb{R}^k$, and ensemble deviation matrix

$$\mathbf{X}^{f'} = \mathbf{X}^f - \mathbf{x}^f \mathbf{e}^T = \mathbf{X}^f \mathbf{U}, \quad (12)$$

with

$$\mathbf{U} = \mathbf{I}_k - \mathbf{w} \mathbf{e}^T \in \mathbb{R}^{k \times k}$$

being the constant projection matrix and $\mathbf{I}_k \in \mathbb{R}^{k \times k}$ the identity matrix. The defining characteristic of the EnKF (Evensen, 2009) is that the sample covariance of the ensemble deviation matrix $\mathbf{X}^{f'}$ is used to approximate the forecast-error covariance matrix:

$$\mathbf{P}^f = \frac{1}{k-1} \mathbf{X}^{f'} [\mathbf{X}^{f'}]^T \in \mathbb{R}^{N \times N}. \quad (13)$$

This shows the computational advantage of using the EnKF; instead of storing the $N \times N$ covariance matrix \mathbf{P}^f , one can instead store the $N \times k$ ensemble perturbation matrix $\mathbf{X}^{f'}$, resulting in a significant computational saving for $k \ll N$. Note however that \mathbf{P}^f is rank-deficient for $k < N$; this is the typical situation in NWP, where N is of the order of 10^9 and k is of the order of 10^2 . This often leads to the problem of covariance underestimation, as the small number of ensemble members insufficiently capture the large number of degrees of freedom of the actual error.

The analysis mean \mathbf{x}^a which minimizes the cost function

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x}^f - \mathbf{x})^T (\mathbf{P}^f)^{-1} (\mathbf{x}^f - \mathbf{x}) + \frac{1}{2} \{ \mathbf{y}^o - \mathcal{H}(\mathbf{x}) \}^T \mathbf{R}^{-1} \{ \mathbf{y}^o - \mathcal{H}(\mathbf{x}) \}, \quad (14)$$

is

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K} (\mathbf{y}^o - \mathbf{H} \mathbf{x}^f), \quad (15)$$

where

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}, \quad (16)$$

is the *Kalman gain matrix* or optimal (in the sense of a minimum variance estimate) weight matrix for the observations. The analysis error covariance matrix can be found as

$$\mathbf{P}^a = (\mathbf{I}_N - \mathbf{K} \mathbf{H}) \mathbf{P}^f, \quad (17)$$

with $\mathbf{H} \in \mathbb{R}^{M \times N}$ being the linearized observation operator, and $\mathbf{I}_N \in \mathbb{R}^{N \times N}$ the identity matrix. To determine an ensemble \mathbf{X}^a which is consistent with the analysis-error covariance matrix \mathbf{P}^a as defined in Eq. (17) and which satisfies

$$\mathbf{P}^a = \frac{1}{k-1} \mathbf{X}^{a'} (\mathbf{X}^{a'})^T, \quad (18)$$

where the prime denotes the deviations from the analysis mean, one can use the method of randomly perturbed observations (Burgers *et al.*, 1998; Houtekamer and Mitchell, 1998; Evensen, 2009) or a deterministic method such as an ensemble square-root filter (Simon, 2006). We will use the ETKF as proposed in Bishop *et al.* (2001), Tippett *et al.* (2003) and Wang *et al.* (2004), which seeks a transformation $\mathbf{T} \in \mathbb{R}^{k \times k}$ to be applied to $\mathbf{X}^{f'}$ so that

$$\mathbf{X}^{a'} = \mathbf{X}^{f'} \mathbf{T}. \quad (19)$$

This method has the advantage of transforming the problem into ensemble space, greatly reducing the number of operations required when $k \ll N$, as is typically the case in operational data assimilation for environmental science. Note that the matrix \mathbf{T} is not uniquely determined for $k < N$, and the specific choice characterizes the type of square-root filter adopted (Tippett *et al.*, 2003). According to Wang *et al.* (2004), the ETKF transformation matrix \mathbf{T} can be obtained by

$$\mathbf{T} = \bar{\mathbf{V}} (\mathbf{I}_{k-1} + \bar{\mathbf{S}})^{-\frac{1}{2}} \bar{\mathbf{V}}^T, \quad (20)$$

where \mathbf{USV}^T is the singular value decomposition of

$$\mathbf{W} = \frac{1}{k-1} \mathbf{X}^{f'T} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{X}^{f'}. \quad (21)$$

The matrix $\bar{\mathbf{V}} \in \mathbb{R}^{k \times (k-1)}$ is obtained by erasing the last zero column from $\mathbf{V} \in \mathbb{R}^{k \times k}$, and $\bar{\mathbf{S}} \in \mathbb{R}^{(k-1) \times (k-1)}$ is the upper left $(k-1) \times (k-1)$ block of the diagonal matrix $\mathbf{S} \in \mathbb{R}^{k \times k}$. The presence of the null eigenvalue is due to the fact that the k ensemble deviation are not independent in view of Eq. (12). The transformation (Eq. (20)) is mean preserving, so that $\mathbf{T}\mathbf{e} = \mathbf{0}$, a condition which is not necessarily true for general square-root filters (Wang *et al.*, 2004).

4. Strategies for dealing with model error

The formulation of the ETKF described in section 3 does not explicitly include a treatment of the model error in the state estimation procedure. This is the topic of the present section. We first review a common approach used to deal with model error in most ensemble-based schemes including the ETKF, and then introduce two novel methods based on the deterministic formulation outlined in section 2.

4.1. ETKF with inflation and localization

The standard way to account for model error in most ensemble-based data assimilation schemes is through artificial inflation and localization of the error covariances. These solutions are often adopted with the general purpose of mitigating the misrepresentation of the actual error covariance due to the use of a finite size ensemble, even in the absence of model error.

Two main types of inflation procedures are known, referred to as multiplicative (Anderson and Anderson, 1999) and additive (Hamill and Whitaker, 2005). In the former case, the matrix \mathbf{P}^f is multiplied by a scalar coefficient, $(1 + \delta)$, with δ slightly larger than zero in most relevant applications. By doing so, the variance explained by \mathbf{P}^f is increased while its range and rank are left unchanged. The increase in the explained variance is taken to account for the portion of forecast error due to model error. Optimization of δ is usually done with numerical tuning, but some advanced solutions allowing for an adaptive estimation have recently appeared (Sacher and Bartello, 2008; Whitaker and Hamill, 2012), as well as a new ensemble-based scheme without the intrinsic need for inflation (Bocquet, 2011). That the span of \mathbf{P}^f is preserved makes

multiplicative inflation attractive in applications with highly chaotic systems, where the ensemble is successful in tracking instabilities that grow along the data assimilation cycle and the range of \mathbf{P}^f has a significant projection on the unstable subspace (Palatella *et al.*, 2013). In additive inflation, random noise is added along the diagonal of \mathbf{P}^f or \mathbf{P}^a at each analysis step. This procedure refreshes the covariance matrix, preventing the collapse of the ensemble along few dominant modes, and maintaining diversity amongst the ensemble members. However, this gain in ensemble spread is obtained at the price of breaking the dynamical consistency of the ensemble and may cause imbalances. Finally, a key issue is the prescription of the statistical properties of the noise, which are difficult to determine and whose knowledge reflect assumptions made on the properties of the model error. Note that the deterministic approach described in the following implicitly addresses this issue and proposes a way to describe the model error mean and covariance.

The localization procedure is designed to remove long-range spatial correlations which, due to the finite size of the ensemble, are often poorly estimated and may introduce spurious correlations into the assimilation cycle (Houtekamer and Mitchell, 1998; Hamill *et al.*, 2001). The standard method for implementing localization is through the Schur product of \mathbf{P}^f by a localization matrix whose entries are obtained on the basis of a predefined length-correlation function $\Omega(r)$, r being the distance between model grid points. Recently Bishop and Hodyss (2011) have proposed a flow-dependent localization method.

In the numerical experiments that follow, the ETKF is implemented with the simultaneous use of multiplicative inflation and localization. Before being used in the analysis update (Eqs (14)–(16)), the forecast-error covariance matrix is transformed according to:

$$\mathbf{P}^f \implies (1 + \delta) \mathbf{P}^f \circ \Omega(r). \quad (22)$$

The update step is then completed with the transformation of the ensemble of forecast deviations into the analysis ensemble using Eq. (19).

4.2. ETKF with deterministic model error treatment

This approach is inspired by previous studies on deterministic model error treatment in the context of both variational and sequential schemes (Carrassi and Vannitsem, 2010, 2011). Model error is regarded as a time-correlated process and the short-time evolution laws (Eqs (6) and (9)) are used to estimate the bias, \mathbf{b}_m , and the model error covariance matrix, \mathbf{P}_m , respectively. The sequential nature of the ETKF justifies adoption of the short-time approximation, however an important practical concern is the ratio between the duration of the short-time regime and the length of the assimilation interval τ over which the approximation is used (Nicolis, 2004).

A key issue is the estimation of the first two statistical moments of the advection mismatch, $\mathbf{f} - \hat{\mathbf{f}}$, required in Eqs (6) and (9) respectively. This problem is addressed here assuming that a reanalysis dataset of relevant geophysical fields is available and is used as a proxy of the true nature evolution. Reanalysis programs constitute the best-possible estimate of the Earth system over an extended period of time, using an homogeneous model and data assimilation procedure, and are of paramount importance in climate diagnosis (Dee *et al.*, 2011).

Let us suppose to have access to such a reanalysis which includes the analysis, \mathbf{x}_r^a , and the forecast field, \mathbf{x}_r^f , so that $\mathbf{x}_r^f(t_j + \tau_r) = \mathcal{M}(\mathbf{x}_r^a(t_j))$, and τ_r is the assimilation interval of the data assimilation scheme used to produce the reanalysis. The operator \mathcal{M} represents the forward model propagator, relative to the model (1), over the reanalysis assimilation interval τ_r ; the subscript r stands for reanalysis. Under this assumption the

following approximation is made:

$$\begin{aligned}
 \mathbf{f}(\mathbf{x}, \lambda) - \widehat{\mathbf{f}}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \lambda) &= \frac{d\mathbf{x}}{dt} - \frac{d\widehat{\mathbf{x}}}{dt} \\
 &\approx \frac{\mathbf{x}_t^f(t + \tau_r) - \mathbf{x}_t^a(t)}{\tau_r} - \frac{\mathbf{x}_t^a(t + \tau_r) - \mathbf{x}_t^a(t)}{\tau_r} \\
 &= \frac{\mathbf{x}_t^f(t + \tau_r) - \mathbf{x}_t^a(t + \tau_r)}{\tau_r} \\
 &= -\frac{\delta \mathbf{x}_t^a}{\tau_r}. \tag{23}
 \end{aligned}$$

The difference between the analysis and the forecast, $\delta \mathbf{x}_t^a$, is usually referred to in the data assimilation literature as an *analysis increment*. From Eq. (23) we see that the vector of analysis increments can be used to estimate the difference between the model and the true nature. A similar approach was originally introduced by Leith (1978), and has been used recently to account for model error in data assimilation (Li *et al.*, 2009). Rodwell and Palmer (2007) also used a similar approach of averaging analysis increments to represent the ‘systematic forecast tendencies’ of a model and diagnosed differences between its climate and the true climate.

Note that the estimate (Eq. (23)) neglects the analysis error, so that its accuracy is connected to that of the data assimilation algorithm used to produce the reanalysis. This in turn is related to characteristics of the observational network such as the number, distribution and frequency of the observations. However, this analysis error is present and acts as an initial condition error, a contribution which is already partially accounted for in the ETKF update by the ensemble-based forecast-error covariance, \mathbf{P}^f . As a consequence, when Eq. (9) is used to estimate only the model error component, an overestimation is expected. This effect, which was also found by Carrassi and Vannitsem (2011) in the context of the extended Kalman filter, can be overcome by an optimal tuning of the amplitude of $\bar{\mathbf{b}}_m$ and $\bar{\mathbf{P}}_m$.

Substituting Eq. (23) into Eq. (6), we obtain an approximation of the bias at the analysis time (i.e. the drift between t and $t + \tau$) as

$$\bar{\mathbf{b}}_m = \langle \delta \mathbf{x}_t^a \rangle \frac{\tau}{\tau_r}. \tag{24}$$

Inserting Eq. (23) into Eq. (9), the model error contribution to the forecast-error covariance can be estimated taking the external product of Eq. (23) after removing the mean and reads

$$\bar{\mathbf{P}}_m = \left\langle \left(\delta \mathbf{x}_t^a - \langle \delta \mathbf{x}_t^a \rangle \right) \left(\delta \mathbf{x}_t^a - \langle \delta \mathbf{x}_t^a \rangle \right)^T \right\rangle \frac{\tau^2}{\tau_r^2}. \tag{25}$$

4.2.1. ETKF with time-constant model error treatment: ETKF-TC

In this approach, $\bar{\mathbf{b}}_m$ and $\bar{\mathbf{P}}_m$ are estimated once using Eqs (24) and (25) and then kept constant along the entire assimilation cycle.

In the ETKF-TC the ensemble forecast mean and covariance are transformed according to

$$\mathbf{x}^f \implies \mathbf{x}^f - \alpha \bar{\mathbf{b}}_m, \tag{26}$$

$$\mathbf{P}^f \implies (1 + \delta) \mathbf{P}^f \circ \Omega(r) + \alpha^2 \bar{\mathbf{P}}_m. \tag{27}$$

The inflation and localization apply to the ensemble-based forecast-error covariance matrix \mathbf{P}^f alone. The scalar term $\alpha > 0$ is a tunable coefficient aimed at optimizing the bias size to account for the expected overestimation connected with the use of Eq. (23). The new first guess and forecast-error covariance, Eqs (26) and (27), are used in the ETKF analysis formulae (Eqs (14)–(16)) to obtain the analysis. The update of the ensemble members is

done as in the standard ETKF transforming the analysis deviation ensemble into the forecast ones according to Eq. (19). In the ETKF-TC, model error is repeatedly corrected in the subspace spanned by the range of $\bar{\mathbf{P}}_m$ where model error is supposed to be confined. This choice reflects the assumption that the impact of the unresolved scales on the forecast error does not fluctuate greatly along the analysis cycle.

4.2.2. ETKF with time-varying model error treatment: ETKF-TV

This is a modification of the ETKF-TC in which the model error correction is made time-varying by incorporating a random sampling from the assumed model error statistics into the forecast step. A model error vector is added to each ensemble forecast member, before being used in the analysis update. In the ETKF-TV the forward propagation of the ensemble members, which is done in the ETKF using the model according to Eq. (11), is substituted by:

$$\mathbf{x}_{i,j}^f = \mathcal{M}(\mathbf{x}_{i,j}^a) - \alpha \boldsymbol{\eta}_{i,j} \frac{\tau}{\tau_r}, \tag{28}$$

$$\text{with } \boldsymbol{\eta}_{i,j} \in \mathcal{N}(\bar{\mathbf{b}}_m, \bar{\mathbf{P}}_m), \quad i = 1, \dots, k,$$

where j refers to the arbitrary analysis time t_j . The model error vectors, $\boldsymbol{\eta}_{i,j} \in \mathbb{R}^N$, are sampled from a distribution of the model error statistics, with mean and covariance equal to $\bar{\mathbf{b}}_m$ and $\bar{\mathbf{P}}_m$ calculated using Eqs (24) and (25). The ETKF-TV preserves the assumed model error statistics while exploring a different realization of the model error at each analysis time. We remark that adding a random perturbation ameliorates problems due to rank deficiency of the finite-size ensemble, but can upset balance and introduce spurious variability. This formulation bears a formal resemblance to stochastic climate models (Majda *et al.*, 2001), which have recently been found to be beneficial in the context of data assimilation (Harlim and Majda, 2008; Mitchell and Gottwald, 2012). Whereas these stochastic climate models rely on the mathematical technique of homogenization which may be difficult or impossible to apply to large models, the ETKF-TV only requires access to an historical reanalysis dataset to generate $\bar{\mathbf{P}}_m$. While therefore less mathematically rigorous than stochastic climate model methods, the ETKF-TV may nonetheless provide a tractable computational approach for dealing with model error in data assimilation systems with large geophysical models.

The forecast ensemble members (Eq. (28)) are then used to build the forecast error covariance matrix \mathbf{P}^f according to Eq. (13) as in the standard ETKF and, if desired, inflation and localization are applied using Eq. (22). The analysis update (Eqs (14)–(16)), including the forecast–analysis deviations transformation, Eq. (19), is done as in the ETKF.

4.3. Comments on the forecast–analysis ensemble transformation

The ETKF forecast–analysis transformation, \mathbf{T} , in Eq. (19), is obtained under the requirements (Wang *et al.*, 2004) that the ensemble analysis-error covariance \mathbf{P}^a given by Eq. (18) is equal to the one predicted by the least-square update (Eq. (17)), and that the analysis ensemble is centred around the mean analysis field, \mathbf{x}^a , obtained via Eqs (14)–(16). As mentioned in section 3, when $k < N$ this transformation, which accounts for the search of a square root of the matrix \mathbf{P}^a , is not unique and its choice characterizes the specific formulation of the square-root implementation (Tippett *et al.*, 2003). A distinguishing feature of the ETKF is that the analysis ensemble members are orthonormal in the observation space (Wang and Bishop, 2003). The search for a square root of \mathbf{P}^a is solved for the transformation matrix, which is then applied to the forecast ensemble. The ensemble transformation implies only the rescaling and rotation of the forecast members, without introducing additional noise. The

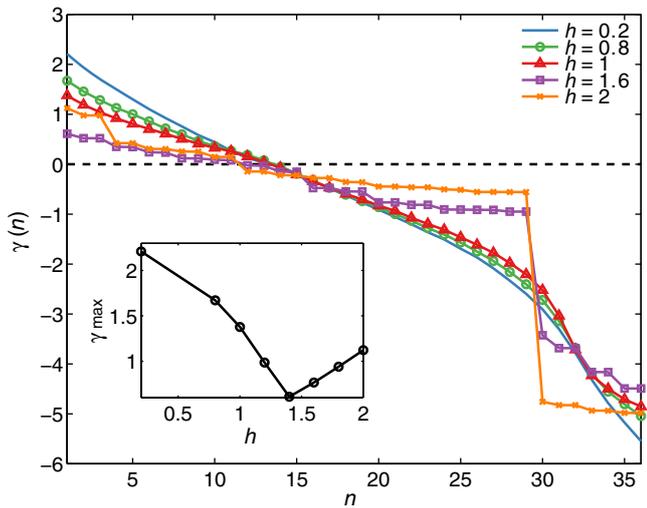


Figure 1. Lyapunov spectra for a range of values of h . The inset shows the maximal Lyapunov exponent γ_{\max} , as a function of h .

latter feature of the deterministic filters with respect to their stochastic counterpart causes major differences in their respective performance when applied to chaotic systems, particularly with relation to the ensemble size (Lawson and Hansen, 2004; Sakov and Oke, 2008; Carrassi *et al.*, 2009).

Let us consider now the standard ETKF as given in section 3, with inflation and localization as described in section 4.1, so that the ensemble-based forecast-error covariance, \mathbf{P}^f , entering the analysis update, Eqs (14)–(17), is now given by Eq. (22): the analysis field and the predicted error covariance are obtained assuming the inflated/localized forecast-error covariance. Nevertheless the transformation to obtain analysis deviations consistent with this predicted \mathbf{P}^a is applied to the forecast perturbations, $\mathbf{X}^{f'}$, evolved over the previous analysis interval. The fact that the inflated/localized matrix, $(1 + \delta)\mathbf{P}^f \circ \Omega(r)$, is used in the analysis update of the mean state, while the unaltered forecast deviations are used in the analysis-error covariance update, causes an inconsistency: the matrix $\mathbf{X}^{f'}$ is not a square root of $(1 + \delta)\mathbf{P}^f \circ \Omega(r)$. Note that we have restricted our discussion to the standard ETKF with inflation/localization only for the sake of clarity; the same arguments holds for the ETKF-TC, to an extent related to the properties of $\bar{\mathbf{P}}_m$. This issue arises in all circumstances when the ensemble-based forecast-error covariance of a square-root filter is adjusted to account for errors misrepresented by the ensemble alone: the new matrix \mathbf{P}^f , used to update the mean, is no longer the square of $\mathbf{X}^{f'}$, and $\mathbf{X}^{a'} = \mathbf{X}^{f'}\mathbf{T}$ is not the square root of the predicted \mathbf{P}^a . This problem is not present in the ETKF-TV, unless localization and/or inflation is applied. In fact the model error treatment in this case acts directly on the ensemble members used to build the forecast-error covariance matrix whose deviations give the corresponding square-root matrix, and represents a potential advantage of the ETKF-TV. The problem of obtaining a consistent mean and ensemble deviations update was originally recognized by Whitaker and Hamill (2002), while Tippett *et al.* (2003) outlined a general method for incorporating model error statistics into an ensemble square-root filter.

5. Results

To compare how well the ETKF-TC and ETKF-TV perform when compared with a standard ETKF, we perform twin data assimilation experiments where a model trajectory taken to be the truth has noise added to create synthetic observations. As a testbed for our experiments we take as our truth the slow-fast Lorenz-96 model (Lorenz, 1996), a prototype model for atmospheric dynamics consisting of a ring of N large-scale, slow oscillators \hat{x}_i

coupled to a ring of $L = N \times J$ small-scale, fast oscillators $\hat{y}_{j,i}$:

$$\frac{d\hat{x}_i}{dt} = (\hat{x}_{i+1} - \hat{x}_{i-2})\hat{x}_{i-1} - \hat{x}_i + F - \frac{hc}{b} \sum_{j=1}^J \hat{y}_{j,i}, \quad (29)$$

$$\frac{d\hat{y}_{j,i}}{dt} = cb(\hat{y}_{j-1,i} - \hat{y}_{j+2,i})\hat{y}_{j+1,i} - c\hat{y}_{j,i} + \frac{hc}{b}\hat{x}_i, \quad (30)$$

with $\hat{x}_{i\pm N} = \hat{x}_i$, $\hat{y}_{j\pm J,i} = \hat{y}_{j,k\pm 1}$ and $\hat{y}_{j,i\pm N} = \hat{y}_{j,i}$.

Our model consists of the first four terms of Eq. (29) for the slow variables only:

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \quad (31)$$

with $x_{i\pm N} = x_i$, which define the resolved scale and the model has no knowledge of the fast subsystem dynamics. The variables \hat{x}_i , $\hat{y}_{j,i}$ and x_i refer to grid point values with specified spacing on a periodic domain, enabling the use of localization in the following numerical experiments. In all experiments we fix $N = 36$, $J = 10$, $F = 10$, and $c = b = 10$ so that the fast modes oscillate ten times faster than the slow modes. We integrate the model using a fourth-order Runge–Kutta scheme with time step $dt = 0.005$. For these parameters and with $h = 1$, the climatological standard deviation is $\sigma_{\text{clim}} = 3.54$ and maximal Lyapunov exponent is $\gamma_{\max} = 1.3775$. In Figure 1 we show the Lyapunov spectra of Eq. (29) for a few different values of h . These spectra were calculated using a QR decomposition (Parker and Chua, 1989) over 10^6 time steps. The inset shows the maximal Lyapunov exponent, γ_{\max} , as a function of h . The maximal Lyapunov exponent decreases with increasing h up to $h = 1.4$, whereafter it increases again. As h increases past 2, the system eventually becomes non-chaotic, with all Lyapunov exponents $\gamma < 0$. A more in-depth exploration of the dynamical properties of the Lorenz-96 system can be found in Frank *et al.* (2014).

In all experiments we take 12 equally spaced observations with observational error covariance $\mathbf{R} = 0.5\mathbf{I}$ or approximately 4% of the climatological variance. It is assumed that the observation-error covariance matrix \mathbf{R} is perfectly known and is used to sample the simulated observational noise. The use of a diagonal matrix \mathbf{R} reflects the assumption of mutually uncorrelated observations. Except where otherwise stated, we use $k = 2N = 72$ ensemble members for the ETKF analysis.

Our experiments consist of two stages: a *reanalysis* stage where the model error biases (Eq. (24)) and covariance matrices (Eq. (25)) are estimated from the analysis increments $\delta\mathbf{x}_T^a$ of a suitably optimized ETKF; and an *experiment* stage where this $\bar{\mathbf{P}}_m$ is used in an Observing System Simulation Experiment (OSSE), to compare analyses made by the ETKF-TC and ETKF-TV with the standard ETKF.

We show estimates of the model error distribution in Figure 2 for different values of the forecast interval τ , and with $h = 1$. Note that we have taken $\tau_f = \tau$ in all experiments, meaning that the reanalysis interval is always the same as the forecast horizon of our verification. The blue curve in Figure 2 shows histograms of the ‘true’ model error distribution obtained by making forecasts using the model (31) initialized with x -values from one 10-year long nature trajectory of Eqs (29) and (30) and considering the differences. The green and red curves show approximations to this model error distribution arising from the analysis increments (Eq. (23)) obtained from a standard ETKF without inflation or localization for an arbitrary selected observed and unobserved variable, respectively. The inset J -values indicate the intersection or joint probability between the true model error distribution and unobserved analysis increments.

Reflecting the progressive degradation of the information for increasing time intervals, both the true and approximate distributions become wider as forecasts are made longer, and the distributions move further apart from one another. This gives rise to the appearance of a bias in the distributions, a

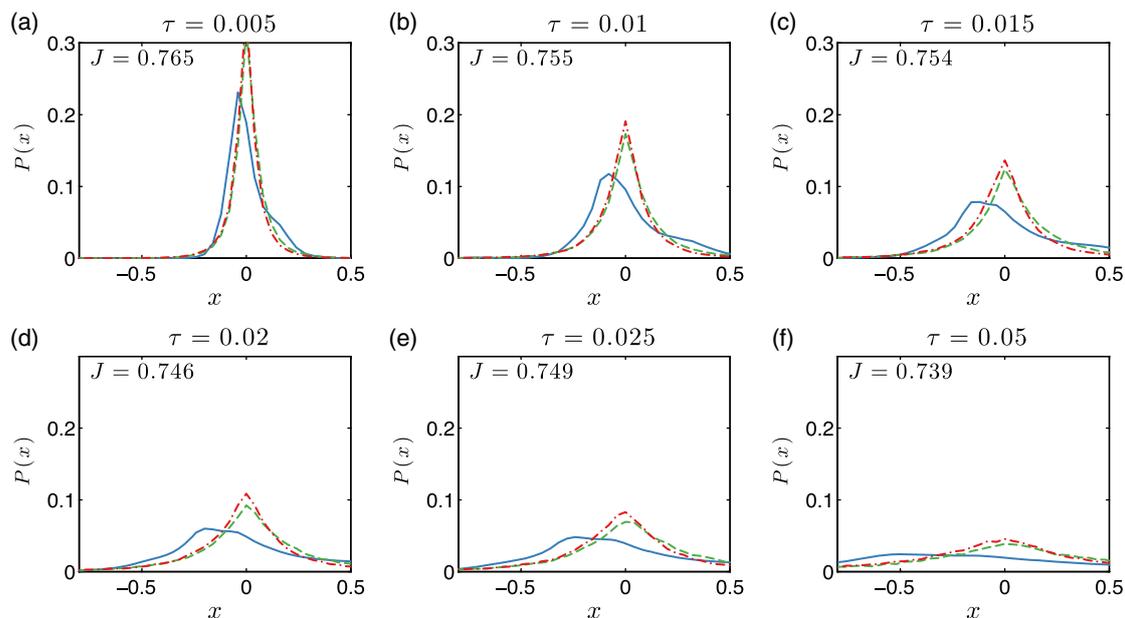


Figure 2. Distribution of analysis increments for arbitrary selected observed (dashed line, green in online) and unobserved (dashed dotted line, red in online) components of the Lorenz-96 system with $h = 1$, as well as the ‘true’ model error distribution (solid line, blue in online), for six different forecast lengths τ (a) 0.005, (b), 0.01, (c) 0.015, (d) 0.02, (e) 0.025, and (f) 0.05. The inset J -values indicate the intersection or joint probability between the true model error distribution and unobserved analysis increments. Statistics were collated over 10 years of analyses, and we have taken $\tau_r = \tau$ in all experiments.

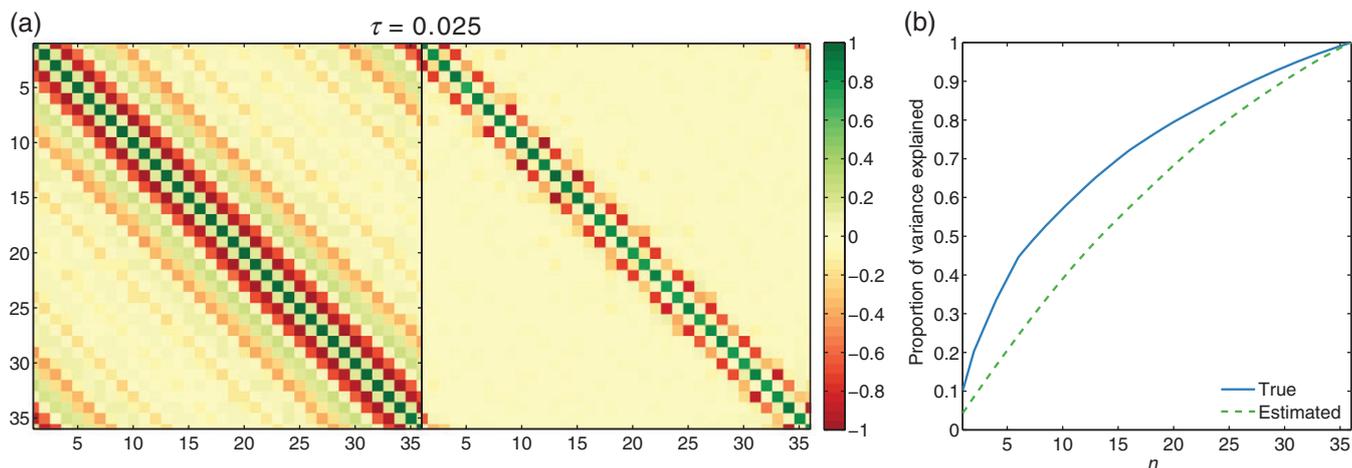


Figure 3. (a) True (left) and estimated (right) model error covariance matrices $\bar{\mathbf{P}}_m$ for $\tau = 0.025$. Both matrices are normalized such that all values lie between -1 and 1 . (b) Proportion of the variance explained as a function of eigenvalue number for the true and estimated covariances.

direct consequence of neglecting a scale of motion in our model. Note that this drift would not have appeared if model error acted as a white noise. Figure 2 shows also that the bias is not captured by our approximation based on the reanalysis increment. This is because the ETKF, similarly to most Kalman filter-like methods, assumes the forecast and the observations are unbiased. Such schemes are unable to correct this error unless a bias correction procedure is explicitly incorporated (Dee and Da Silva, 1998).

In Figure 3(a), we show the true model error covariances and analysis error-estimated model error covariances $\bar{\mathbf{P}}_m$ for the same 10-year experiments as in Figure 2. Each heat map is normalized such that values range between -1 and 1 , with large-magnitude values indicating strong positive or negative correlation between sites. The estimated model error covariances $\bar{\mathbf{P}}_m$ capture some of the negative correlations that exist between neighboring sites, but represent little of the long-range correlation structures present in the true covariance matrix. Correlations between pairs of observed variables are also better preserved than those between pairs involving unobserved variables.

Based on the above, we set the analysis interval $\tau = 0.025$, approximately equivalent to 3 h, for all subsequent experiments. This is as a compromise to make our short-time approximation

work adequately and for the observation interval to be similar to that of a realistic observational network.

We first calibrate the inflation factor δ and localization radius r in order to produce the best possible reanalysis dataset. Figure 4 shows a contour plot of RMS analysis errors for a standard ETKF, averaged over 200 realizations of 1 month in length. The RMS errors are normalized by the climatological standard deviation σ_{clim} . The best results are obtained with a strong inflation, $\delta = 0.9$, and a moderate localization radius, $r = 3$. This suggests that the ETKF tends to underestimate the actual error variance and misrepresent the error spatial correlations at a length-scale larger than three grid points. Based on the results in Figure 4 we set $r = 3$, $\delta = 0.9$ in all subsequent experiments to ensure both accuracy and stability in the ETKF.

Using a 10-year reanalysis with this choice of r and δ , we then perform twin data assimilation experiments with the ETKF, ETKF-TC and ETKF-TV. One realization of this experiment is shown in Figure 5, where we show 300 analysis cycles of analyses made by the ETKF, ETKF-TC and ETKF-TV for (a) an observed variable and (b) an unobserved variable, with the black curves showing the true time series. (c) shows the RMS analysis error for each method, as averaged over all variables. For the observed variables, the ETKF-TC and ETKF-TV both outperform the

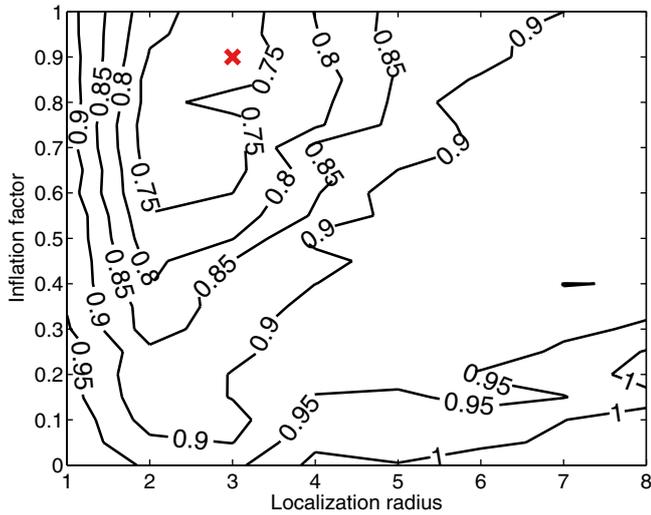


Figure 4. Contour plot of RMS errors for the ETKF as a function of inflation factor δ and localization radius r . The cross (red in online) shows the location of the minimum. RMS errors are normalized by the climatological standard deviation σ_{clim} , and are averaged over 200 realizations of 1 month in duration.

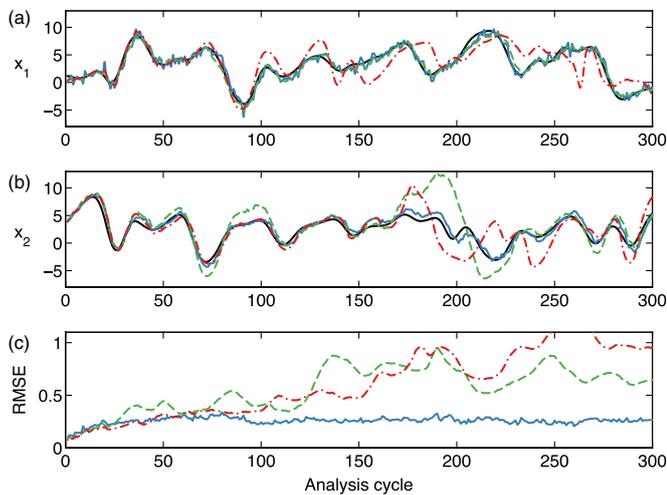


Figure 5. Time series of analyses made by the ETKF (dashed dotted line, red in online), ETKF-TC (dashed line, green in online) and ETKF-TV (solid line, blue in online) for (a) an observed variable, (b) an unobserved variable, with the black curves showing the true time series, and (c) RMS errors.

standard ETKF, which diverges after approximately 100 analysis cycles. The ETKF-TC and ETKF-TV provide an efficient track of the truth throughout the entire duration of the experiment. However, while the ETKF-TV tracks the unobserved variable similarly well to the observed variable, the ETKF-TC diverges from the unobserved reference trajectory between analyses 175 and 225. The relative performance of the three algorithms is summarized in the time series of the RMSE; errors are normalized using the system's natural variability σ_{clim} as above. The RMSE for the ETKF reaches values as large as 100% of the climatological standard deviation, meaning the assimilation process is unable to carry the information coming from the observations into the state estimate. However, the ETKF-TV is very successful at this task, with the RMSE converging to 25% of the climatological standard deviation σ_{clim} . The performance of the ETKF-TC is between those of the ETKF and ETKF-TV, with the RMSE converging to about 60% of σ_{clim} . Note that the RMSE here is averaged over all variables; in the observed subdomain the RMSE of both ETKF-TV and ETKF-TC goes well below the observational accuracy level, a sign that both filters are functioning well. The ETKF RMSE in the observed subdomain does not attain a similarly low level. Such a marked difference in the skill between observed and unobserved areas for the ETKF-TC is possibly due to the use of a constant $\bar{\mathbf{P}}_m$. In the unobserved

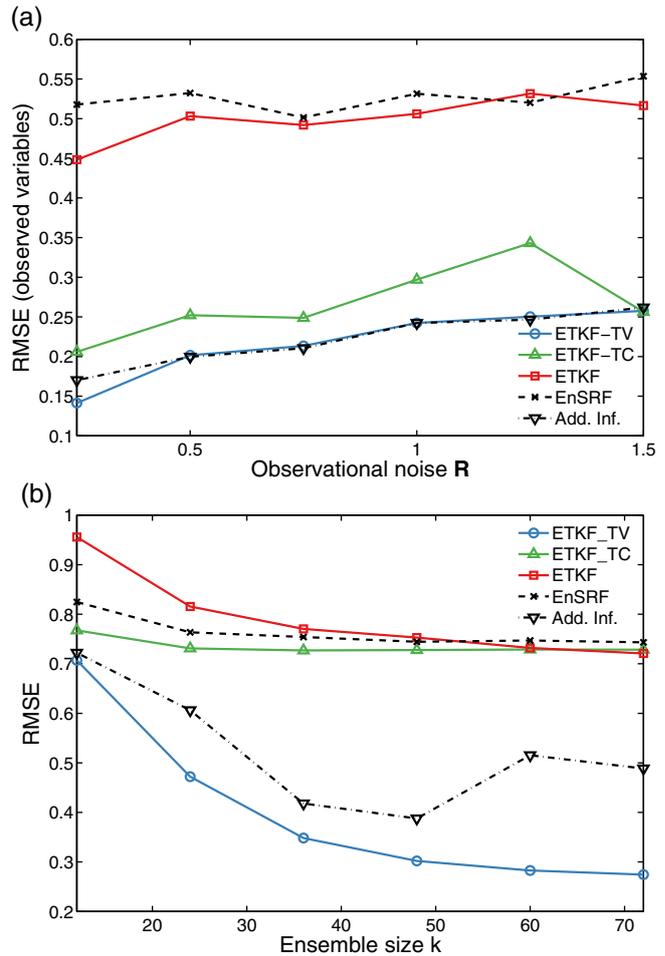


Figure 6. Dependence of (a) RMSE averaged over the observed variables upon the magnitude of the observational noise \mathbf{R} and (b) RMSE averaged over all variables upon the number of ensemble members for different methods for ETKF with inflation and localization (red); ETKF-TC (green); ETKF-TV (blue); serial EnSRF (black dashed); and ETKF with additive inflation (black dash-dotted). Results are averaged over 200 realizations, and all errors are normalized by the climatological standard deviation σ_{clim} .

areas the analysis correction projects mostly on the span of the forecast error covariance \mathbf{P}^f , making its correct specification crucial in those regions. Note from Figure 2 that the estimated model error bias $\bar{\mathbf{b}}_m$ is approximately equal to zero; this means that skill improvements made by the ETKF-TC and ETKF-TV come largely from second-moment corrections in the analysis update.

In Figure 6 we show how the methods perform as a function of the magnitude of (a) the observational error \mathbf{R} and (b) the number of ensemble members. We again set $\delta = 0.9$, $r = 3$ in all experiments and average our results over 200 realizations, each of 1 month in length. The RMS errors as a function of ensemble size (b) are shown as averaged over all variables, while for the magnitude of the observational error (a) we focus on the RMSE averaged over the observed variables only. The general trends observed in Figure 5 are preserved in both cases, with the ETKF-TV showing a large improvement in analysis quality over the ETKF, with the ETKF-TC showing a slightly more modest improvement. The skill of each method decreases roughly monotonically as the observational error is increased, and as the ensemble size is decreased. Interestingly, small ensemble sizes impact the ETKF-TV more adversely than they do the ETKF-TC, with the latter method almost becoming preferable for $k = 12$. The semi-independence of the ETKF-TC upon k is due to the fact that the model error treatment is independent of k , and the range of the matrix $\bar{\mathbf{P}}_m$ provides an additional subspace for error correction that is not subject to ensemble collapse. This is an attractive property in the common situation where the ensemble

size cannot be increased due to computational limitations for both the model and the ensemble-based scheme. However, the overall error level for the ETKF-TC remains high. This is an indication that the model-error covariance matrix is dominating the ensemble-based forecast-error covariance, and may require further tuning through the model-error inflation parameter α (Figure 8 below).

As a further comparison of the proposed model error treatments to standard techniques, we also include in Figure 6 curves corresponding to the serial ensemble square-root filter (EnSRF) of Whitaker and Hamill (2002), as well as an ETKF with additive inflation to account for model error as described by Hamill and Whitaker (2005). We took the EnSRF as an example of an ensemble Kalman filter which avoids the inconsistency between analysis mean and perturbations described in section 4.3, and additive inflation as an example of a standard model error treatment with some similarity to the ETKF-TV. While we refer the reader to the above articles for details of those methods, we remark that we set $\delta = 0.9$, $r = 3$ for the EnSRF as we had for the ETKF and used our reanalysis as a proxy for a high-resolution model run in the additive inflation experiment. The EnSRF performs comparably to the ETKF in both experiments, while additive inflation performs as well as the ETKF-TV for the observed variables but less well for the unobserved variables, leading to a higher overall RMS error in Figure 6(b). It is not surprising that additive inflation as implemented here should perform comparably to the ETKF-TV, as the only difference between the two methods is in the sampling of the dataset of reanalysis increments to produce $\eta_{i,j}$ in Eq. (28). Following Hamill and Whitaker (2005), in the additive inflation experiment the sampling is done directly from the record of reanalysis increments, while in the ETKF-TV this is done from the multivariate covariance matrix \mathbf{P}_m . While this is expected to lead to significant differences when the model error distribution is far from Gaussian, from Figure 3 we know that this is not the case here and the skill of additive inflation and ETKF-TV are comparable.

We now investigate whether our model error treatments can make further improvements over the standard ETKF by re-tuning the inflation and localization parameters δ and r in the experiment stage. To do this we use the same reanalysis dataset with $\delta = 0.9$ and $r = 3$ to build the model error biases (Eq. (24)) and covariance matrices (Eq. (25)), but vary the δ and r used in the ETKF-TC and ETKF-TV experiments. The contour plots in Figure 7 show the results of these experiments for (a) the ETKF-TC and (b) the ETKF-TV, displaying the RMS errors over all variables and as averaged over 20 realizations each. Firstly, the ETKF-TV produces the best analyses for all values of δ and r , showing RMS errors approximately half those of the ETKF-TC for $r > 1$. Secondly, the two methods are optimized by slightly different values of δ and r . The ETKF-TC obtains its best performance when $\delta = 0.05$ and $r = 8$, while the ETKF-TV is optimized when $\delta = 0$ and $r = 5$ (approximately). This implies that the ETKF-TC underestimates the actual forecast error variance by only approximately 5%, and it is essentially correctly estimated by the ETKF-TV (compared with approximately 90% underestimation for the ETKF). Moreover, the large values found for the optimal localization radius r indicate that both model error treatments have essentially mitigated the misrepresentation of the long-range spatial correlations by spurious ensemble effects, compared to the standard ETKF. Both methods outperform the optimized standard ETKF, with RMSE equal to 0.256 for the ETKF-TV with optimal tuning and 0.690 for the ETKF-TC with optimal tuning compared with 0.709 for the standard ETKF with optimal tuning. Interestingly, the skill dependence of ETKF-TC and ETKF-TV on δ and r is very similar and clearly different from that of the standard ETKF. From Figure 7 we observe that for $r \leq 3$ the performance of the two methods depends only on the localization factor. For larger r the skill of either method almost converges, with only minor dependence on δ and r . Both

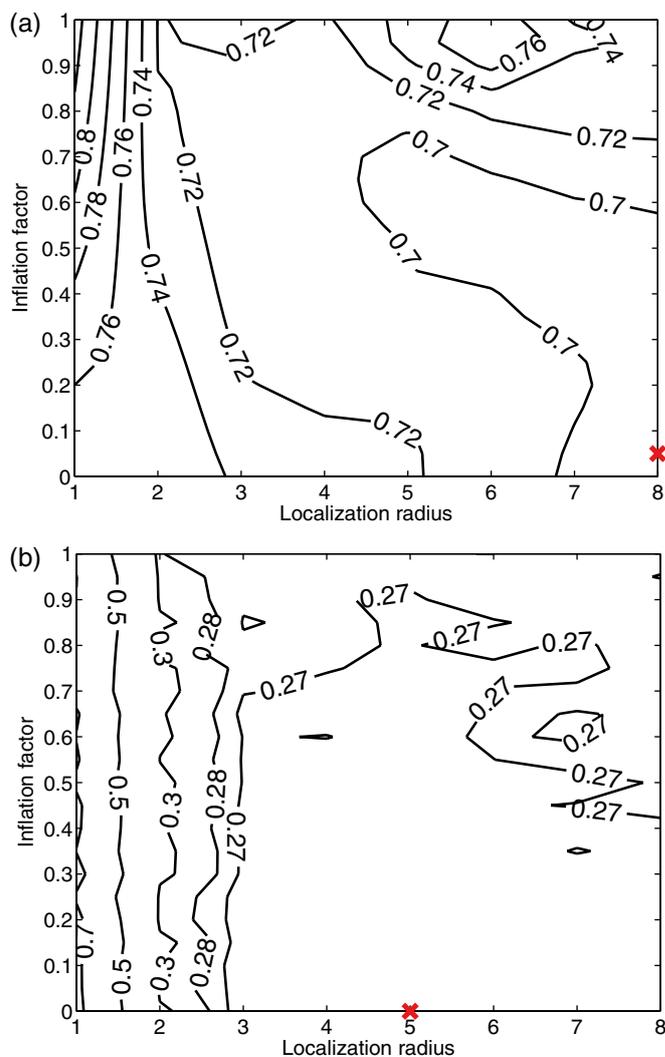


Figure 7. Contour plot of RMS errors for (a) the ETKF-TC and (b) ETKF-TV as a function of inflation factor δ and localization radius r . The crosses (red in online) show the location of the minima. RMS errors are normalized by the climatology standard deviation σ_{dim} , and are averaged over 200 realizations of 1 month in duration.

the ETKF-TC and ETKF-TV therefore appear very stable with respect to the tunable parameters δ and r , an attractive property when considering their implementation with large numerical systems.

Following Carrassi and Vannitsem (2011), we investigate the effect of varying the model error inflation factor α in Eqs (26)–(28). As was discovered in Carrassi and Vannitsem (2011), we again find here that some deflation of the model error covariance matrix $\bar{\mathbf{P}}_m$ used in both the ETKF-TC and ETKF-TV can actually be beneficial to those methods. In Figure 8 we show how the RMS analysis errors in ETKF-TC depend upon the model error inflation/deflation factor α . Each point is averaged over 200 realizations, and shows that a deflation of the model error term through setting $\alpha = 0.2$ improves the performance of the ETKF-TC by approximately 40%, while setting $\alpha = 0.5$ improves the ETKF-TV very slightly (however this improvement is not statistically significant). That the best results are obtained for $\alpha < 1$ and that the RMS errors worsen for $\alpha > 1$ implies that the approximations Eqs (24) and (25) are actually overestimations of the actual bias and covariance. This is because Eqs (24) and (25) incorporate some initial condition error into the estimation of $\bar{\mathbf{P}}_m$, so optimal results are obtained when α is tuned accordingly. However the response to α for the two algorithms is very different, with the ETKF-TC showing the largest dependence. This is not surprising given the time-constant model error treatment of the ETKF-TC, and is also consistent with the discussion on Figure 6(b). The ETKF-TV depends less on α because different

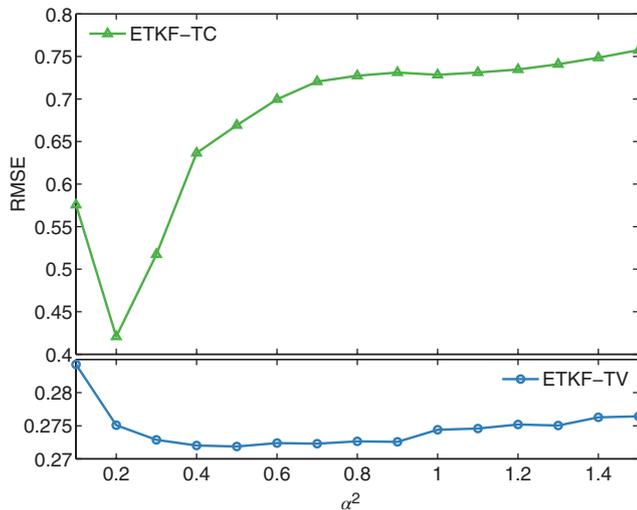


Figure 8. Dependence of RMS errors for the ETKF-TC (green) and ETKF-TV (blue) upon the magnitude of the model error inflation factor α .

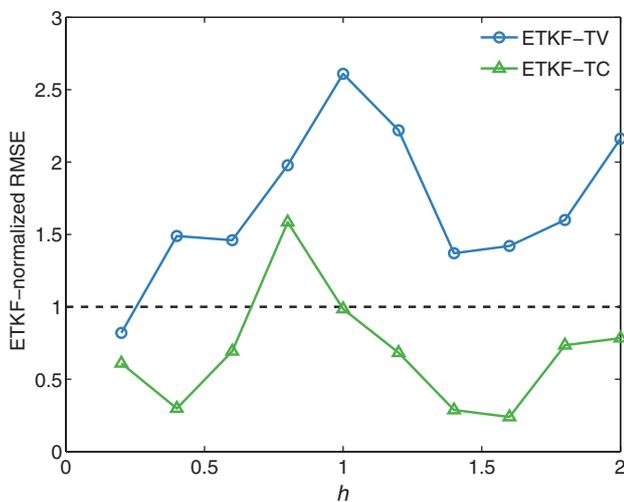


Figure 9. Dependence of the skill of ETKF-TC (green) and ETKF-TV (blue) upon the magnitude of the size of the model error h , shown as the ratio of the ETKF errors divided by each method, so values > 1 represent a skill improvement. All values are averaged over 200 realizations.

optimal model error sizes may be required at each analysis step.

It is worth studying the relative performance of the algorithms for different levels of model error. This is controlled in the Lorenz-96 model by modifying the parameter h . However this factor has also a strong interplay with other dynamical aspects of the model, such the spectrum of the Lyapunov exponents and the amplitude of the dominant Lyapunov exponent γ_{\max} . As such, it is difficult to draw a clear relationship between h and the performance of our two model error treatments and so we cannot expect the relative performance of ETKF-TC/TV over the ETKF to improve monotonically with h . In Figure 9 we show the ratio of the ETKF error to that of the ETKF-TC and ETKF-TV, so that values greater than 1 represent an improvement in analysis skill. Figure 9 shows that a monotonic trend is almost found for $h \leq 1$, as the model error decreases. For larger h the curves resemble the dependence of γ_{\max} over h (the inset of Figure 1), suggesting that in a stabler model configuration the gain in skill of the ETKF-TC/TV is smaller than in more unstable cases.

6. Discussion

Following the formulation of Nicolis (2004), we have derived a method for the treatment of time-correlated model error in the

short-time limit for an ensemble Kalman filter. In the absence of complete knowledge of the true atmospheric state, we used a database of historical reanalysis increments to approximate the model error biases and covariance matrices required in the least-square-based analysis update. One clear drawback of this choice is that its accuracy depends on the properties of the observational network, so that in unobserved or poorly observed areas, the analysis increments will be almost zero even in the presence of model error. This dependency is however less dramatic than in the closely related approach of using the innovations (differences between observation and forecast) to estimate or tune background or observational-error covariance matrix (e.g. Desroziers *et al.*, 2005). In this latter case in fact only the projection of the error covariance matrix on the observational space can be estimated. The statistics of the analysis increments have been also used by Cullen (2010b) to estimate the so-called ‘regularization matrix’ (also Cullen, 2010a), a newly introduced term used to improve the fit to the observations. In this view, the regularization term is intended to account for all the system error (also connected to the assimilation procedure itself) and not just the error originating from the underlying model dynamics.

We studied two potential implementations of this method for an ETKF, namely the ETKF-TC which assumed the model error correction term to be constant in time and the ETKF-TV where the model error correction was allowed to vary randomly after each forecast. In the latter case, the noise is sampled from the assumed model error statistics and its contribution to the forecast error is propagated forward using a short-time approximation. These methods were compared with the standard approach for dealing with model error in the ETKF of tuning the inflation factor and localization radius empirically to minimize RMS errors through numerical experiments on the Lorenz-96 model.

Both the ETKF-TC and ETKF-TV showed improvements upon the ETKF with optimized inflation and localization parameters, with the ETKF-TV showing the lowest RMS errors of all three methods. The two methods with deterministic model error treatment showed further improvements once inflation and localization were retuned, but importantly were found to be less sensitive to the precise tuning of these parameters than the ETKF was. This is a desirable property when considering the potential implementation of ETKF-TC or ETKF-TV in an operational setting. Room for additional improvements in the ETKF-TC and -TV are based on the recognition that the adopted model error treatment is prone to suffer from an overestimation of the actual model error, due to the presence of initial condition error. Results confirmed this conjecture and showed that the skill of both ETKF-TC and -TV can be further improved by deflating the associated model error contribution to the state-estimation error. This is particularly relevant for the ETKF-TC in view of its description of the model error being constant in time, which has both advantages and disadvantages. A stationary model error treatment suggests that the ETKF-TC should be expected to perform well only on average, and indeed we did find that the ETKF-TV was better able to capture intermittent events such as are expected in forecasting on a meteorological time-scale. However, the independence over the ensemble size of the model error contribution in the ETKF-TC makes this algorithm less sensitive to errors arising from small ensemble size effects than the time-varying ETKF-TV was. The overall performance of the ETKF-TV is markedly better than the ETKF-TC, however the aforementioned features make the ETKF-TC attractive in situations where a pre-existing ensemble scheme is already implemented and the computational recourses are already fully exploited.

This study has been carried out using a simplified chaotic dynamical system which has allowed for an extensive and robust exploration of the methods under consideration. Furthermore, the use of this model has enabled us to use the ETKF in all numerical experiments, which is known to be impractical in

operational contexts due to the inability to store the forecast covariance matrix \mathbf{P}^f . An important step forward will be through applications of the same approach to model and observational set-ups of increasing complexity. In parallel to this application-oriented research direction, the authors are studying alternative formulations of the square-root filter in which the model error treatment does not give rise to the inconsistency described in section 4.3. It may also be possible to use methods such as the LETKF or EnSRF which allow for localization to be implemented to the forecast perturbations. How such filters perform with the proposed model error treatment is a direction for future research, and how such methods compare with standard procedures is of great relevance in the ongoing debate over model error and ensemble filtering.

A logical further extension of this work would be to explicitly incorporate the temporal evolution of model error. Model error temporal correlations are allowed in a weak-constraint variational framework (Tremolet, 2006), and the use of a deterministic method for their estimation has been introduced in Carrassi and Vannitsem (2010). An extension of the present study along this same line implies the use of ensemble-based smoother, a research direction that we leave for future work.

Acknowledgements

The authors thank Patrick Raanes for helpful discussions, as well as the three reviewers for helping us to improve this manuscript. LM was supported by NSF grant DMS-0940271, and is grateful for the computational resources provided by the Vermont Advanced Computing Core which is supported by NASA (NNX 08A096G), and the Vermont Complex Systems Center. AC was financed through the IEF Marie Curie Project INCLIDA of the FP7, as well as EU-FP7 project SANGOMA under grant agreement no. 283580.

References

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

Balmaseda MA, Dee D, Vidard A, Anderson DLT. 2007. A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Q. J. R. Meteorol. Soc.* **133**: 167–179, doi: 10.1002/qj.12.

Bishop CH, Hodys D. 2011. Adaptive ensemble covariance localization in ensemble 4D-VAR state estimation. *Mon. Weather Rev.* **139**: 1241–1255, doi: 10.1175/2010MWR3403.1.

Bishop CH, Etherton B, Majumdar S. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **129**: 420–436.

Bocquet M. 2011. Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlinear Processes Geophys.* **18**: 735–750, doi: 10.5194/npg-18-735-2011.

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

Carrassi A, Vannitsem S. 2010. Accounting for model error in variational data assimilation: A deterministic formulation. *Mon. Weather Rev.* **138**: 3369–3386, doi: 10.1175/2010MWR3192.1.

Carrassi A, Vannitsem S. 2011. Treatment of the error due to unresolved scales in sequential data assimilation. *Int. J. Bifurc. Chaos* **21**: 3619–3626, doi: 10.1142/S0218127411030775.

Carrassi A, Vannitsem S, Zupanski D, Zupanski M. 2009. The maximum likelihood ensemble filter performances in chaotic systems. *Tellus* **61**: 587–600.

Charron M, Pellerin G, Spacek L, Houtekamer PL, Gagnon N, Mitchell HL, Michelin L. 2010. Toward random sampling of model error in the Canadian ensemble prediction system. *Mon. Weather Rev.* **138**: 1877–1901, doi: 10.1175/2009MWR3187.1.

Cullen MJP. 2010a. A demonstration of 4D-Var using a time-distributed background term. *Q. J. R. Meteorol. Soc.* **136**: 1301–1315, doi: 10.1002/qj.645.

Cullen MJP. 2010b. A demonstration of cycled 4D-Var in the presence of model error. *Q. J. R. Meteorol. Soc.* **136**: 1379–1395, doi: 10.1002/qj.653.

Daley R. 1992. The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Weather Rev.* **120**: 178–196.

Dee DP. 2005. Bias and data assimilation. *Q. J. R. Meteorol. Soc.* **131**: 3323–3343, doi: 10.1256/qj.05.137.

Dee DP, Da Silva A. 1998. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* **117**: 269–295.

Dee DP, Todling R. 2000. Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Mon. Weather Rev.* **128**: 3268–3282.

Dee DP, Uppala S, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P, Bechtold P, Beljaars A, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer A, Haimberger L, Healy S, Hersbach H, Hólm EV, Isaksen I, Kallberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz B, Morcrette J-J, Park B, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F. 2011. The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597.

Deng Z, Tang Y, Freeland HJ. 2011. Evaluation of several model error schemes in the EnKF assimilation: Applied to Argo profiles in the Pacific Ocean. *J. Geophys. Res.* **116**: C09027, doi: 10.1029/2011JC006942.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.* **131**: 3385–3396, doi: 10.1256/qj.05.108.

Doblas-Reyes F, Garca-Serrano J, Lienert F, Pintò Biescas A, Rodrigues L. 2013. Seasonal climate predictability and forecasting: Status and prospects. *Interdiscip. Rev. Clim. Change* **4**: 245–268.

Evensen G. 2009. *Data Assimilation: The Ensemble Kalman Filter*. Springer: New York, NY.

Frank MR, Mitchell L, Danforth CM. 2014. Standing swells surveyed showing surprisingly stable solutions for the Lorenz '96 model. *Int. J. Bifurcation Fluids*. In press. Also arXiv preprint 1312.5965.

Hamill TM. 2001. Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Weather Rev.* **129**: 550–560.

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

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

Harlim J, Majda AJ. 2008. Filtering nonlinear dynamical systems with linear stochastic models. *Nonlinearity* **21**: 1281–1306, doi: 10.1088/0951-7715/21/6/008.

Houtekamer PL, Mitchell HL. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.* **126**: 796–811.

Houtekamer PL, Mitchell HL, Deng X. 2009. Model error representation in an operational ensemble Kalman filter. *Mon. Weather Rev.* **137**: 2126–2143, doi: 10.1175/2008MWR2737.1.

Jazwinski AH. 1970. *Stochastic Processes and Filtering Theory*. Academic Press: New York, NY.

Kalnay E. 2002. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press: Cambridge, UK.

Lawson W, Hansen J. 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Weather Rev.* **132**: 1966–1981.

Leith C. 1978. Predictability of climate. *Nature* **276**: 352–355.

Li H, Kalnay E, Miyoshi T. 2009. Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Q. J. R. Meteorol. Soc.* **135**: 523–533, doi: 10.1002/qj.371.

Lorenz EN. 1996. Predictability: A problem partly solved. In *Seminar on Predictability*, 1–18. ECMWF: Reading, UK.

Majda AJ, Timofeyev I, Vanden-Eijnden E. 2001. A mathematical framework for stochastic climate models. *Commun. Pure Appl. Math.* **54**: 891–974.

Mitchell L, Gottwald GA. 2012. Data assimilation in slow–fast systems using homogenized climate models. *J. Atmos. Sci.* **69**: 1359–1377, doi: 10.1175/JAS-D-11-0145.1.

Nichols N. 2003. Treating model error in 3D and 4D data assimilation. In *Data Assimilation for the Earth System*, Swinbank R, Shutyaev V, Lahoz WA. (eds.) NATO Science Series 26: 127–135. Kluwer: Amsterdam, Netherlands, doi: 10.1007/978-94-010-0029-1_11.

Nicolis C. 2004. Dynamics of model error: The role of unresolved scales revisited. *J. Atmos. Sci.* **61**: 1740–1753.

Palatella L, Carrassi A, Trevisan A. 2013. Lyapunov vectors and assimilation in the unstable subspace: Theory and applications. *J. Phys. A: Math. Theor.* **46**: 254020, doi: 10.1088/1751-8113/46/25/254020.

Parker TS, Chua LO. 1989. *Practical Numerical Algorithms for Chaotic Systems*. Springer: New York, NY.

Raynaud L, Berre L, Desroziers G. 2012. Accounting for model error in the Météo-France ensemble data assimilation system. *Q. J. R. Meteorol. Soc.* **138**: 249–262, doi: 10.1002/qj.906.

Rodwell M, Palmer TN. 2007. Using numerical weather prediction to assess climate models. *Q. J. R. Meteorol. Soc.* **133**: 129–146, doi: 10.1002/qj.23.

Sacher W, Bartello P. 2008. Sampling errors in ensemble Kalman filtering. Part I: Theory. *Mon. Weather Rev.* **136**: 3035–3049, doi: 10.1175/2007MWR2323.1.

Sakov P, Oke PR. 2008. Implications of the form of the ensemble transform in the ensemble square root filters. *Mon. Weather Rev.* **136**: 3035–3049.

Simon DJ. 2006. *Optimal State Estimation*. John Wiley & Sons, Inc.: New York, NY.

- Tippett MK, Anderson JL, Bishop CH, Hamill TM, Whitaker JS. 2003. Ensemble square root filters. *Mon. Weather Rev.* **131**: 1485–1490, doi: 10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.
- Tremolet Y. 2006. Accounting for an imperfect model in 4D-Var. *Q. J. R. Meteorol. Soc.* **132**: 2483–2504, doi: 10.1256/qj.05.224.
- Wang X, Bishop CH. 2003. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60**: 1140–1158.
- Wang X, Bishop CH, Julier S. 2004. Which is better, an ensemble of positive–negative pairs or a centered spherical simplex ensemble? *Mon. Weather Rev.* **132**: 1590–1605.
- Whitaker JS, Hamill TM. 2002. Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.* **130**: 1913–1924, doi: 10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.
- Whitaker JS, Hamill TM. 2012. Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Weather Rev.* **140**: 3078–3089, doi: 10.1175/MWR-D-11-00276.1.
- Zupanski D. 1997. A general weak constraint applicable to operational 4D-Var data assimilation systems. *Mon. Weather Rev.* **125**: 2274–2292.
- Zupanski D, Zupanski M. 2006. Model error estimation employing an ensemble data assimilation approach. *Mon. Weather Rev.* **134**: 1337–1354.