Data Assimilation by Morphing Fast Fourier Transform Ensemble Kalman Filter for Precipitation Forecasts using Radar Images[1]

Jan Mandel[2], Jonathan D. Beezley[2], Kryštof Eben[3], Pavel Juruš[3],
Volodymyr Y. Kondratenko[2], and Jaroslav Resler[3]

April 2010

Abstract

The FFT EnKF takes advantage of the theory of random fields and the FFT to provide a good and cheap approximation of the state covariance with a very small ensemble. The method and predecessor components are explained and the method is extended to the case of observations given on a rectangular subdomain, suitable for assimilation of regional radar images into a weather model.

Contents

1 Introduction

Incorporating new data into computations in progress is a well-known problem in weather forecasting, and techniques to incorporate new data by sequential Bayesian estimation are known as data assimilation [12]. The basic framework is the discrete time state-space model in its most general form is an application of the Bayesian update problem. In the Gaussian case, direct application of

[1]This research was partially supported by the National Science Foundation under the grants CNS-0719641 and ATM-0835579.

[2]Center for Computational Mathematics and Department for Mathematical and Statistical Sciences, University of Colorado Denver, Denver CO 80217-3364, USA.

[3]Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodarenskou vezi 2, 182 07 Prague 8, Czech Republic.

the Bayes theorem with assumed state covariance becomes the classical optimal statistical interpolation. A more general method, the Kalman filter evolves the state covariance by the dynamics of the system, but it is not suitable for high-dimensional problems, due to the need to maintain the covariance matrix. The ensemble Kalman filter (EnKF) [7] is a Monte-Carlo implementation of the Kalman filter, which replaces the state covariance by sample covariance of an ensemble of simulations. The EnKF has become quite popular because it allows an implementation without any change to the model; the model only needs to be capable of exporting its state and restarting from the state modified by the EnKF. However, there are two main drawbacks to the EnKF.

First, a reasonable approximation of the state covariance by the sample covariance requires a large ensemble, easily many hundreds [7]. Techniques to improve the approximation of the state distribution by the ensemble include non-random selection of the initial ensemble such as Lyapunov vectors or bred vectors [12], and localization techniques such as covariance tapering [8], the Ensemble Adjustment Kalman Filter [2], and the Local Ensemble Transform Kalman Filter [10]. This paper builds on the Fast Fourier Transform (FFT) EnKF, introduced in [15, 17]. The FFT EnKF takes advantage of the fact that the state is approximately a stationary random field, that is, the covariance between two points is mainly a function of their distance vector. Then the multiplication of the covariance matrix and a vector is approximately a convolution, so the covariance matrix in the frequency domain can be well aproximated by its diagonal. This results in a good approximation of the covariance for very small ensembles with no need for tapering, as well as an efficient implementation of the EnKF formulas by the FFT. It should be noted that the related Fourier Domain Kalman filter (FDKF) [5] is something else, although it also uses a diagonal covariance matrix in the frequency domain. The FDKF is the Kalman filter used in each Fourier mode separately, and an ensemble in the FDKF context consists of independent relalizations for the purpose of post-processing only.

Second, although the EnKF can be and is widely used for quite general probability distributions, the EnKF still assumes that the probability distributions are Gaussian, otherwise the implementation of the Bayes theorem is no longer valid, and the EnKF performace deteriorates. In the Gaussian case, it can be proved that in the large ensemble limit, the ensemble generated by the EnKF converges to a sample from the filtering distribution computed by the Kalman filter [18]. Combinations of the EnKF and a particle filter for the non-Gaussian case also exist [14]. One particular case where strongly non-Gaussian distributions occur are nonlinear systems with sharp coherent moving features, such as weather fronts, firelines, and epidemic waves. In these cases, the distribution of the position of the coherent features may well be close to Gaussian but the distribution of the values of the physical variables at any given point is not. This paper uses techniques borrowed from image processing that can be used to transform the state to the so-called morphing representation, which contains both amplitude and position information and it has a distribution much closer to Gaussian. Then, EnKF techniques can be succesfully used on the morphing representation. [4, 15, 16, 17].

In this paper, we provide a more complete description, motivation, and investigation of the properties of the FFT EnKF than it was possible in the short conference papers [15, 17]. In addition, in [15, 17], the FFT EnKF was limited to the case when the observation consists of one state variable over the whole physical domain, such as the heat flux from a wildfire, or the number of infected individuals per unit area in an epidemic simulation. Here, we extend the FFT EnKF to the case when the values of the state variable are observed on a rectangle, which is the case of radar observations over an area of interest, while the weather simulation executes over a larger

domain. We also provide a new and more complete discussion of the crosscovariances, i.e., how the innovation in the observed variable affects other model variables, which was mentioned in [15, 17] only in passing.

The FFT EnKF method is combined with a new version of the Morphing EnKF [4] on a subdomain, resulting in a method suitable for assimilation of regional radar images into a weather model.

2 Background

2.1 Sequential data assimilation

In sequential statistical estimation, the modeled quantity is the probability distribution of the simulation state. The model is advanced in time until an *analysis* time. The distribution of the system state before, now called the *prior* or the *forecast* distribution, and the *data likelihood* are now combined to give the new system state distribution, called the *posterior* or the *analysis* distribution. This step is called Bayesian update or analysis step. This completes one *analysis cycle*, and the model is then advanced until the next analysis time.

In the Bayesian update, the probability density $p(u)$ of the system state u before the update (the prior) and the probability density $p(d|u)$ of the the data d given the value of the system state u (the data likelihood) are combined to give the new probability density $p^a(u)$ of the system state u with the data incorporated (the posterior) from the Bayes theorem,

$$p^a(u) \propto p(d|u)\, p(u). \tag{1}$$

where $\propto$ means proportional. Equation (1) determines the posterior density $p(u)$ completely, because $\int p^a(u)\, du = 1$.

Here, we consider the case of linear *observation operator* H: given system state u, the data value, d, would be Hu if the model and the data were without any errors. Of course, in general, data are given such that $d \neq Hu$, so the discrepancies are modeled by the data likelihood $p(d|u)$. Assume that the prior has normal distribution with mean μ and covariance Q, and the data likelihood is Gaussian with mean Hu and covariance R, that is,

$$p(u) \propto \exp\left(-\frac{1}{2}(u-\mu)^{\mathrm{T}} Q^{-1}(u-\mu)\right), \quad p(d|u) \propto \exp\left(-\frac{1}{2}(d-Hu)^{\mathrm{T}} R^{-1}(d-Hu)\right).$$

It can be shown by algebraic manipulations ([1], see also [3, p. 10]) that the posterior distribution is also Gaussian,

$$p^a(u) \propto \exp\left(-\frac{1}{2}(u-\mu^a)^{\mathrm{T}} (Q^a)^{-1}(u-\mu^a)\right),$$

where the posterior mean μ^a and covariance Q^a are given by the update formulas

$$\mu^a = \mu + K(d - Hu), \quad Q^a = (I - KH)\, Q, \tag{2}$$

and

$$K = QH^{\mathrm{T}}\left(HQH^{\mathrm{T}} + R\right)^{-1} \tag{3}$$

is called Kalman gain.

When the prior covariance Q is assumed to be known (that is, it is determined by expert judgement of the modeler), the Bayesian update (2) is known as *optimal statistical interpolation*. The Kalman filter [11] advances the state covariance by the model explicitly, assuming that the model is linear. Techniques such as the extended Kalman filter, which advances the state covariance by a linearization of the model, were developed to treat nonlinear systems. Although Kalman filter and its variants were succesfully used in many applications, they need to store and manipulate the covariance matrix of the state, which makes them usuitable for high-dimensional systems.

Other, different data assimilation methods were developed for high-dimensional and nonlinear system, such as variational data data assimilation (3DVAR and 4DVAR) [12]. These methods proceed by adjusting the initial conditions (3DVAR) as well as the state at various times in the past (4DVAR) and they require an additional code, consistent with the model, to implement the so-called adjoint model, which goes back in time to make the adjustment possible.

2.2 The EnKF

The EnKF is a Monte-Carlo implementation of the KF. The EnKF advances an ensemble of independent simulations $[u_k] = [u_1, \ldots, u_N]$ in time. The ensemble $[u_k]$ approximates the probability distribution of the model state u, which is a column vector in $\mathbb{R}^n$. In the analysis step, the ensemble $[u_k]$, now called the *forecast ensemble*, is combined with the data by the EnKF formulas [7, p. 41]

$$u_k^a = u_k + Q_N H^{\mathrm{T}} \left(H Q_N H^{\mathrm{T}} + R\right)^{-1} \left(d + e_k - H u_k^f\right), \quad k = 1, \ldots, N, \tag{4}$$

to yield the *analysis ensemble* $[u_k^a]$. Here, Q_N is an approximation of the covariance Q of the model state u, and e_k is sampled from $N(0, R)$. This completes the analysis cycle, and the ensemble is then advanced by the simulations in time again.

When Q_N is the ensemble covariance, the EnKF formulation (4) does not take advantage of any special structure of the model. Efficient evaluation of the analysis formula (4) is then made possible by representing the n-by-n sample covariance matrix by an n-by-N matrix E,

$$\begin{aligned} Q_N &= \frac{1}{N-1} \sum_{k=1}^{N} \left(X_k - \overline{X}\right) \left(X_k - \overline{X}\right)^{\mathrm{T}} = \frac{1}{N-1} E E^{\mathrm{T}}, \\ E &= \left[X_1 - \overline{X}, \ldots, X_N - \overline{X}\right], \\ \overline{X} &= \frac{1}{N} \sum_{k=1}^{N} \overline{X}. \end{aligned} \tag{5}$$

Efficient schemes for computing with the matrix E only are then possible, and the matrix Q_N is never formed explicitly. The method in [16, eq. (15)] is suitable in the case when R has a sparse inverse and it is based on the QR decomposition of E. The procedure in [7, Ch. 14] requires approximating R by the sample covariance of the data perturbations e_k. This method works for a general R, but it requires a more expensive SVD instead of QR decomposition, and the inversion o f the singular matrix in (4) needs special care.

The sample covariance Q_N, however, is generally a poor approximation on the true state covariance. It is expected that the state covariance has small or zero entries between variables that correspond to points distant in space, since the small variation in one place should affect places far

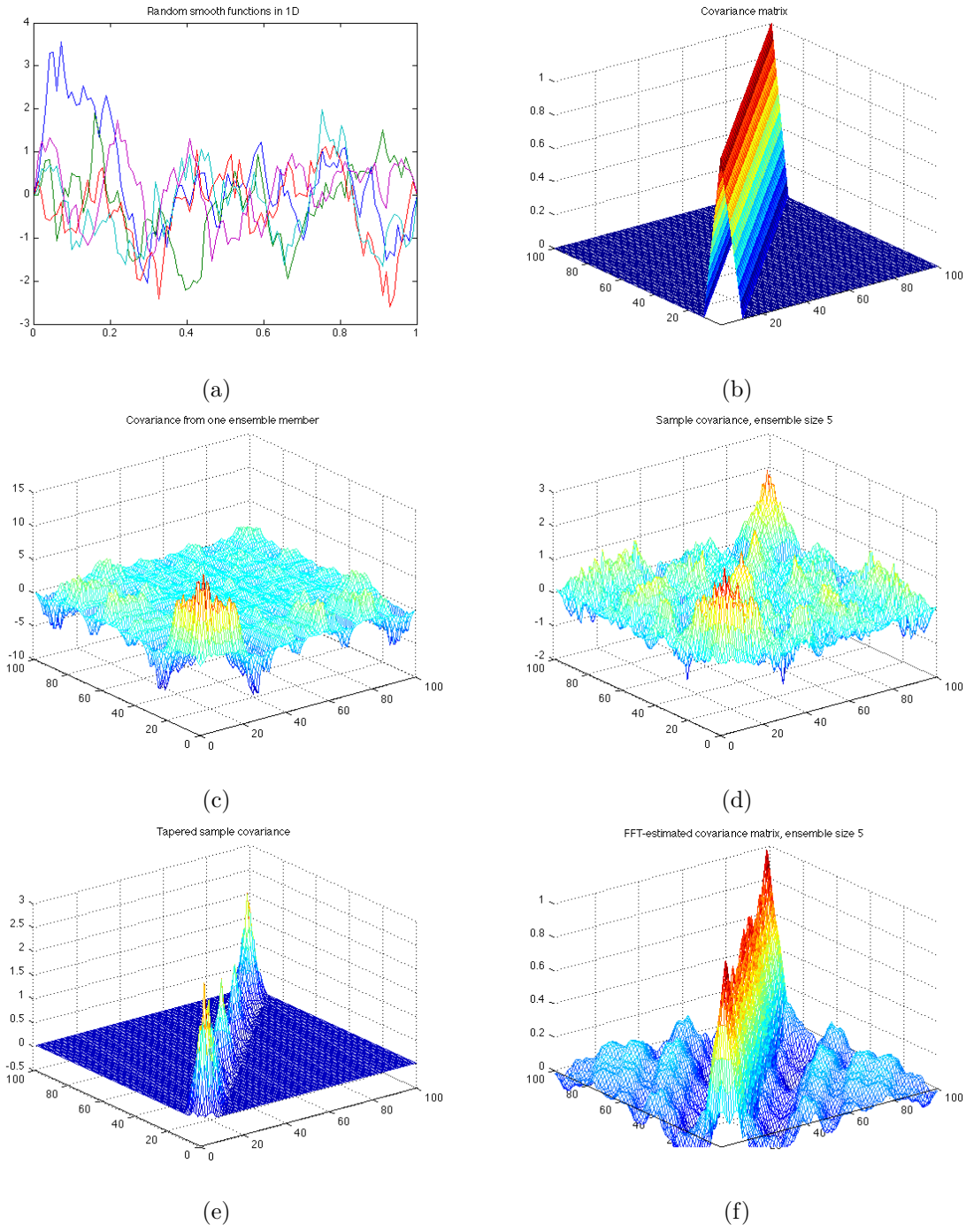

Figure 1: Covariance estimation from an ensemble of 5 random smooth functions in 1D. Random smooth functions (a) are generated by sampling a 1D random field with the given covariance (b). Covariance of one of the functions (c) show non-local corellations. The sample covariance (d) is a poor approximation due to the small sample size. Tapering the sample covariance (e) forces the non-local correlations to zero. The covariance obtained from the FFT estimation (d) is much more accurate and does not require any artificial modifications.

away. This is also a common assumption in the statistical theory of random fields [6]; the covariance of random smooth functions drops off with the distance from the diagonal. However, Q_N is the sum of N rank one matrices, with $N \ll n$. In a rank-one matrix, all columns are proportional to each other and all rows are proportional to each other, thus having larger entries only close to the diagonal is not possible (Figure 1c). The sample covariance is the sum of a small number of rank-one matrices and the same conclusion applies (Figure 1d). Therefore, it was proposed to modify the sample covariance by multiplication entry-by-entry with a given tapering matrix T to force the long-range covariances to zero (Figure 1e). That is, the tapered covariance matrix is the Schur product

$$\widetilde{Q}_N = Q_N \cdot T, \quad (\widetilde{Q}_N)_{ij} = (Q_N)_{ij}\, T_{ij}. \tag{6}$$

It is known that if T is positive definite then the tapered covariance matrix is at least positive semidefinite.

Tapering improves the accuracy of the approximate covariance for small ensembles [8], but it makes the implementation of (4) much more expensive: the sample covariance matrix can no longer be efficiently represented as the product of two small dense matrices as in (5), but it needs to be manipulated as a large, albeit sparse, matrix, requiring much more sophisticated and expensive sparse linear algebra.

In general, the state has more than one variable. The state u, the state covariance Q, and the observation matrix H then have the block form

$$u = \begin{bmatrix} u^{(1)} \\ \vdots \\ u^{(n)} \end{bmatrix}, \quad Q = \begin{bmatrix} Q^{(11)} & \cdots & Q^{(1M)} \\ \vdots & \ddots & \vdots \\ Q^{(M1)} & \cdots & Q^{(MM)} \end{bmatrix}, \quad H = \begin{bmatrix} H^{(1)} & \cdots & H^{(M)} \end{bmatrix}, \tag{7}$$

with one block per variable. The EnKF analysis formulas (4) with the sample covariance (5) are completely oblivious to this structure of the state. However, the choice of the initial ensemble needs to respect it.

2.3 Smooth random functions by FFT

Suppose $\{v_k\}$ is an orthonormal sequence in a Hilbert space H. Then one can construct a Gaussian random element in H,

$$U = \textstyle\sum_{k=1}^{\infty} u_k d_k^{1/2} \theta_k$$

where the constant coefficients d_k satisfy

$$d_k \geq 0, \quad \textstyle\sum_{k=1}^{\infty} d_k < \infty,$$

and $\theta_k \sim N(0,1)$ are independent. Then $E(U) = 0$ and the covariance C of U is given by

$$\begin{aligned} \langle u, Cv \rangle &= E\left(\langle U, u \rangle \langle U, v \rangle \right) \\ &= E\left(\left\langle \textstyle\sum_{k=1}^{\infty} u_k d_k^{1/2} \theta_k, u \right\rangle \left\langle \textstyle\sum_{k=1}^{\infty} u_k d_k^{1/2} \theta_k, v \right\rangle \right) \\ &= \textstyle\sum_{k=1}^{\infty} d_k \langle u_k, u \rangle \langle u_k, v \rangle, \end{aligned}$$

so

$$C = \textstyle\sum_{k=1}^{\infty} d_k u_k u_k^{\mathrm{T}}$$

and d_k are the eigenvalues of C.

When H is a space of functions $u(w)$, we get the covariance function

$$C\left(w, w'\right) = \sum_{k=1}^{\infty} d_k u_k\left(w\right) u_k\left(w'\right).$$

When u_k are smooth functions, such as trigonometric functions, U is a smooth random function.The faster the coefficients d_k decay, the smoother the random function is. For the sin functions and a rectangle, we have

$$U\left(x, y\right) = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} d_{k\ell}^{1/2} \sin\left(\frac{x}{a}k\pi\right) \sin\left(\frac{y}{b}\ell\pi\right) \theta_{k\ell},$$

with the covariance function

$$C\left(\left(x, y\right), \left(x', y'\right)\right) = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} d_{k\ell} \sin\left(\frac{x}{a}k\pi\right) \sin\left(\frac{y}{b}\ell\pi\right) \sin\left(\frac{x'}{a}k\pi\right) \sin\left(\frac{y'}{b}\ell\pi\right).$$

Choose $d_{k\ell} = \lambda_{k\ell}^{-\alpha}$, where $\lambda_{k\ell}$ are eigenvalues of the Laplace operator. Then

$$U\left(x, y\right) = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \left(\left(\frac{k\pi}{a}\right)^2 + \left(\frac{\ell\pi}{b}\right)^2\right)^{-\alpha/2} \sin\left(\frac{x}{a}k\pi\right) \sin\left(\frac{y}{b}\ell\pi\right) \theta_{k\ell}$$

and the covariance of the random field U is

$$\begin{aligned} C\left(\left(x, y\right), \left(x', y'\right)\right) &= E\left(U\left(x, y\right) U\left(x', u'\right)\right) \\ &= \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \left(\left(\frac{k\pi}{a}\right)^2 + \left(\frac{\ell\pi}{b}\right)^2\right)^{-\alpha/2} \\ &\quad \sin\left(\frac{x}{a}k\pi\right) \sin\left(\frac{y}{b}\ell\pi\right) \sin\left(\frac{x'}{a}k\pi\right) \sin\left(\frac{y'}{b}\ell\pi\right). \end{aligned}$$

The use of the Green's function (that is, the inverse) of the Laplace equation as covariance was suggested earlier in [13].

2.4 Morphing EnKF

Given an initial state u as in (7), the initial ensemble in the morphing EnKF [4, 16] is given by

$$u_k^{(i)} = \left(u_{N+1}^{(i)} + r_k^{(i)}\right) \circ \left(I + T_k\right), \quad k = 1, \ldots, N, \quad i = 1, \ldots, M, \tag{8}$$

where $\circ$ denotes composition of mappings, $u_{N+1} = u$ is considered an additional member of the ensemble, called the *reference member*, $r_k^{(i)}$ are random smooth functions on Ω, and T_k are random smooth mappings $T_k : \Omega \to \Omega$. Thus, the initial ensemble varies both in amplitude and in position, and the change in position is the same for all variables.

The case considered in EnKF is when the data d is a complete observation of one of the variables, $u^{(1)}$. The data d and the first blocks of all members $u_1, \ldots, u_N$ are then registered against the first block of u_{N+1} by

$$\begin{aligned} d &\approx u_{N+1}^{(1)} \circ \left(I + T_0\right), \quad T_0 \approx 0, \quad \nabla T_0 \approx 0. \\ u_k^{(1)} &\approx u_{N+1}^{(1)} \circ \left(I + T_k\right), \quad T_k \approx 0, \quad \nabla T_k \approx 0, \quad k = 1, \ldots, N, \end{aligned} \tag{9}$$

where $T_k : \Omega \to \Omega$, $k = 0, \ldots, N$ are called *registration mappings*. The registration mappings T_k are found by multilevel optimization [3, 4, 9]. The *morphing transform* maps each ensemble member u_k into the extended state vector

$$u_k \mapsto \widetilde{u}_k = M_{u_{N+1}}(u_k) = \left(T_k, r_k^{(1)}, \ldots, r_k^{(M)}\right), \tag{10}$$

called *morphing representation*, where

$$r_k^{(j)} = u_k^{(j)} \circ (I + T_k)^{-1} - u_{N+1}^{(j)}, \quad k = 0, \ldots, N, \tag{11}$$

are called *registration residuals*. Note that when $u_k^{(1)} = u_{N+1}^{(1)} \circ (I + T_k)$ in (9), then $r_k^{(j)} = 0$; this happens when all differences between $u_k^{(1)}$ and $u_{N+1}^{(1)}$ are resolved by position changes. See [4] for an explantion and motivation of the particular form (11) of the registration residual. Likewise, the data is mapped into the *extended data vector*, given by

$$d \mapsto \widetilde{d} = \left(T_0, r_0^{(1)}\right).$$

and the observation matrix becomes

$$\left(T, r^{(1)}, \ldots, r^{(M)}\right) \mapsto \left(T, r^{(1)}\right).$$

The EnKF is then applied to the transformed ensemble $[\widetilde{u}_1, \ldots, \widetilde{u}_N]$, giving the transformed analysis ensemble $[\widetilde{u}_1^a, \ldots, \widetilde{u}_N^a]$ and the the new transformed reference member is given by

$$\widetilde{u}_{N+1}^a = \frac{1}{N} \sum_{k=1}^{N} \widetilde{u}_k^a. \tag{12}$$

The analysis ensemble $u_1^a, \ldots, u_{N+1}^a$ including the new reference member is then obtained by the *inverse morphing transform*, defined by

$$u_k^{a,(i)} = M_{u_{N+1}}^{-1}\left(\widetilde{u}_k^a\right) = \left(u_{N+1}^{(i)} + r_k^{a,(i)}\right) \circ \left(I + T_k^a\right), \quad k = 1, \ldots, N+1, \quad i = 1, \ldots, M, \tag{13}$$

similarly as in (8). The analysis ensemble is then advanced by $N + 1$ independent simulations and the analysis cycle repeats.

3 FFT EnKF

We explain the FFT EnKF in the 1D case; higher-dimensional cases are exactly the same.

3.1 The basic algorithm

Consider first the case when the model state consists of one variable only. Denote by $u(x_i)$, $i = 1, \ldots, n$ the entry of vector u corresponding to node x_i. If the random field u is stationary, then the covariance matrix satisfies $Q(x_i, x_j) = c(x_i - x_j)$ for some covariance function c, and multiplication by Q is the convolution

$$v(x_i) = \sum_{j=1}^{n} Q(x_i, x_j)\, u(x_j) = \sum_{j=1}^{n} u(x_j)\, c(x_i - x_j), \quad i = 1, \ldots, n.$$

In the frequency domain, convolution becomes entry-by-entry multiplication of vectors, that is, the multiplication by a diagonal matrix. This is easy to verify for the complex FFT and circular convolution; however, here we use the real discrete Fourier transform (DFT), either sine or cosine, and convolution with the vectors extended by zero. Then the relation between convolution and entry-by-entry multiplication in the frequency domain is more complicated [19].

We assume that the random field is approximately stationary, so we can neglect the off-diagonal terms of the covariance matrix in the frequency domain, which leads to the the following FFT EnKF method. First apply FFT to each member to obtain

$$\widehat{u}_k = F u_k.$$

Similarly, we denote by $\widehat{\ }$ the DFT of other quantities. For an n by n matrix M, the corresponding matrix in the frequency domain is

$$\widehat{M} = F M F^{-1},$$

so that

$$u = Mv \Leftrightarrow \widehat{u} = \widehat{M}\widehat{v}. \tag{14}$$

Note that, with a suitable scaling, the DFT is unitary, $F^*F = I$ and

$$F^* = F^{-1}, \tag{15}$$

where * denotes conjugate transpose. This is true for the real DFT (sine and cosine) as well.

Let $\widehat{u}_{ik}$ be the entries of the column vector $\widehat{u}_k$. Using (15), we have the sample covariance in the frequency domain

$$\widehat{Q}_N = F Q_N F^{-1} = \frac{1}{N-1}\sum_{k=1}^{N} F\left(u_k - \overline{u}\right)\left(\widehat{u}_k - \overline{u}\right)^{\mathrm{T}} F^* \tag{16}$$

$$= \frac{1}{N-1}\sum_{k=1}^{N}\left(\widehat{u}_k - \overline{\widehat{u}}\right)\left(\widehat{u}_k - \overline{\widehat{u}}\right),$$

where

$$\overline{\widehat{u}} = \frac{1}{N}\sum_{k=1}^{N}\widehat{u}_k,$$

so the entries of $\widehat{Q}_N$ are

$$\widehat{Q}_{Nij} = \frac{1}{N-1}\sum_{k=1}^{N}\left(\widehat{u}_{ik} - \overline{\widehat{u}}_i\right)\left(\widehat{u}_{jk} - \overline{\widehat{u}}_j\right), \text{ where } \overline{\widehat{u}}_i = \frac{1}{N}\sum_{k=1}^{N}\widehat{u}_{ik}. \tag{17}$$

We approximate the forecast covariance matrix in the frequency domain by the diagonal matrix $\widehat{C}$ with the diagonal entries given by

$$\widehat{c}_i = \widehat{Q}_{Nii} = \frac{1}{N-1}\sum_{k=1}^{N}\left|\widehat{u}_{ik} - \overline{\widehat{u}}_i\right|^2. \tag{18}$$

The resulting approximation $C = F^{-1}\widehat{C}F$ of the covariance of u tends to be a much better approximation than the sample covariance for a small number of ensemble members N (Figure 1f).

Multiplication of a vector $\widehat{u}$ by the diagonal matrix $\widehat{C}$ is the same as entry-by-entry multiplication by the vector $\widehat{c} = [\widehat{c}_i]$:

$$\widehat{C}\widehat{u} = \widehat{c} \bullet \widehat{u}, \quad (\widehat{c} \bullet \widehat{u})_i = \widehat{c}_i \widehat{u}_i. \tag{19}$$

Now assume that the observation function $H = I$, that is, the whole state u is observed. By (14), the evaluation of the EnKF formula (4) in the frequency domain, with the spectral approximation $\widehat{C}$ in the place of $\widehat{Q}_N$, becomes

$$\widehat{u}_k^a = Fu_k + FCF^{-1}\left(FCF^{-1} + FRF^{-1}\right)^{-1}\left(Fd + Fe_k - Fu_k^f\right), \tag{20}$$

which is simply

$$\widehat{u}_k^a = \widehat{u}_k + \widehat{C}\left(\widehat{C} + \widehat{R}\right)^{-1}\left(\widehat{d} + \widehat{e}_k - \widehat{u}_k^f\right), \tag{21}$$

Further assume that the data error covariance R is such $\widehat{R}$, and denote by $\widehat{r}$ the diagonal of $\widehat{R}$. For example, when the data error is identical and independent at the mesh points (i.e., white noise), then R is the multiple of identity. A more general diagonal $\widehat{R}$ can be used to model smooth data error (Section 2.3). Then, (21) becomes

$$\widehat{u}_k^a = \widehat{u}_k + \frac{\widehat{c}}{\widehat{c} + \widehat{r}} \bullet \left(\widehat{d} + \widehat{e}_k - \widehat{u}_k^f\right). \tag{22}$$

where the operations on vectors are performed entry by entry. The analysis ensemble is obtained by inverse FFT at the end,

$$u_k^a = F^{-1}\widehat{u}_k^a.$$

The FFT EnKF as described here used the sine DFT, which results in zero changes to the state on the boundary.

In the rest of this section, we describe several variations and extensions of this basic algorithm.

3.2 Multiple variables

Consider the state with multiple variables and the covariance and the observation matrix in the block form (7). Assume that the first variable is observed, so

$$H = [I, 0, \ldots, 0].$$

Then

$$Q_N H^T = \begin{bmatrix} Q^{(11)} & \ldots & Q^{(1M)} \\ \vdots & \ddots & \vdots \\ Q^{(M1)} & \ldots & Q^{(MM)} \end{bmatrix} \begin{bmatrix} I \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} Q^{(11)} \\ \vdots \\ Q^{(M1)} \end{bmatrix},$$

$$HQ_N H^T = [I, 0, \ldots, 0] \begin{bmatrix} Q^{(11)} \\ \vdots \\ Q^{(M1)} \end{bmatrix} = Q^{(11)},$$

and (4) becomes

$$u_k^a = u_k + Q_N H^T \left(HQ_N H^T + R\right)^{-1} \left(d + e_k - Hu_k^f\right), \tag{23}$$

$$\begin{bmatrix} u_k^{(1),a} \\ \vdots \\ u_k^{(M),a} \end{bmatrix} = \begin{bmatrix} u_k^{(1)} \\ \vdots \\ u_k^{(M)} \end{bmatrix} + \begin{bmatrix} Q_N^{(11)} \\ \vdots \\ Q_N^{(M1)} \end{bmatrix} \left(Q_N^{(11)} + R\right)^{-1} \left(d + e_k - u_k^{(1)}\right). \tag{24}$$

In the case when all variables in the state u are based on the same nodes in space and so the blocks have the same dimension, one can proceed just as in Section 3.1 and approximate the covariances $Q^{(j1)}$ by their diagonals in the frequency domain. Then, (24) becomes

$$\widehat{u}_k^{(j),a} = \widehat{u}_k^{(j)} + \frac{\widehat{c}^{(j1)}}{\widehat{c}^{(11)} + \widehat{r}} \bullet \left(\widehat{d} + \widehat{e}_k - \widehat{u}_k\right), \tag{25}$$

where

$$\widehat{c}_i^{(j1)} = \frac{1}{N-1} \sum_{k=1}^{N} \left(\widehat{u}_{ik}^{(j)} - \overline{\widehat{u}}_i^{(j)}\right) \left(\widehat{u}_{ik}^{(1)} - \overline{\widehat{u}}_i^{(1)}\right), \quad \overline{\widehat{u}}_i^{(\ell)} = \frac{1}{N} \sum_{k=1}^{N} \widehat{u}_{ik}^{(\ell)}. \tag{26}$$

Note that we are using real DFT, so the complex conjugate in (26) can be dropped.

In the general case, however, the covariance matrices $Q_N^{(j1)}$ are rectangular. Fortunately, one can use FFT with approximation by the diagonal in the frequency space for the efficient evaluation the inverse in (24) only, while using some other approximation $\widetilde{Q}^{(j1)}$of the covariances $Q^{(j1)}$, which gives

$$\begin{bmatrix} u_k^{(1),a} \\ \vdots \\ u_k^{(M),a} \end{bmatrix} = \begin{bmatrix} u_k^{(1)} \\ \vdots \\ u_k^{(M)} \end{bmatrix} + \begin{bmatrix} \widetilde{Q}^{(11)} \\ \vdots \\ \widetilde{Q}^{(M1)} \end{bmatrix} F^{-1} \left(\frac{1}{\widehat{c}^{(11)} + \widehat{r}} \bullet \left(Fd + Fe_k - Fu_k^{(1)}\right)\right), \tag{27}$$

where the operations on vectors are again entry by entry.

Possible approximate covariance matrices include the sample covariance, a suitably tapered sample covariance, and matrices obtained by dropping selected entries of the sample covariance in the frequency domain. Because the approximate covariances $\widetilde{Q}^{(j1)}$ are not involved in a matrix inversion, they do not need to be diagonal in the frequency domain.

3.3 Observations on a subdomain

We now modify the method from Section 3.1 to the case when the observations are given as the values of one variable on a subdomain (an interval in 1D, a rectangle in 2D), rather than on the whole domain, and the observations are based on the same nodes as the state. Then,

$$H = [P, 0, \ldots, 0], \tag{28}$$

where P is a zero-one matrix such that multiplication Px selects the values of a vector x in the observed subdomain. Substituting this observation matrix into (23), (24) becomes

$$\begin{bmatrix} u_k^{(1),a} \\ \vdots \\ u_k^{(M),a} \end{bmatrix} = \begin{bmatrix} u_k^{(1)} \\ \vdots \\ u_k^{(M)} \end{bmatrix} + \begin{bmatrix} Q_N^{(11)} \\ \vdots \\ Q_N^{(M1)} \end{bmatrix} P^{\mathrm{T}} \left(PQ_N^{(11)} P^{\mathrm{T}} + R\right)^{-1} \left(d + e_k - u_k^{(1)}\right), \tag{29}$$

where $PQ_N^{(11)}P^{\mathrm{T}}$ is the sample covariance of the vectors Pu_1. Denote by F_P the DFT on the observed subdomain and replace the sample covariance in the frequency domain by its diagonal part. Denote the diagonal by $\widehat{c}_P^{(11)}$. Then, just as in (27), we have

$$\begin{bmatrix} u_k^{(1),a} \\ \vdots \\ u_k^{(M),a} \end{bmatrix} = \begin{bmatrix} u_k^{(1)} \\ \vdots \\ u_k^{(M)} \end{bmatrix} + \begin{bmatrix} \widetilde{Q}^{(11)} \\ \vdots \\ \widetilde{Q}^{(M1)} \end{bmatrix} P^{\mathrm{T}} F_P^{-1} \left(\frac{1}{\widehat{c}_P^{(11)} + \widehat{r}} \bullet \left(F_P d + F_P e_k - F_P u_k^{(1)} \right) \right).$$

The FFT EnKF with observations on a subdomain assumes the subdomain is well inside the simulation domain, and uses the cosine DFT.

4 The complete method

The overall method combines morphing (Section 2.4) and FFT EnKF with multiple variables (Section 3.2) and observations on a subdomain (Section 3.3). Morphing matches the data on the subdomain (radar observations), resulting in two variables representing the registration mapping T, defined on the whole domain, and the residual, defined on the subdomain only. The observation operator H then has three nonzero blocks: two identity blocks for the components of T, and a zero-one block P as in (28) that selects variables in the subdomain. The covariance $Q^{(11)}$ in (24) consists of three diagonal matrices and we neglect the off-diagonal blocks, so that the FFT approach (25), (29) can be used. As in [15, 17], the sine DFT is used in the blocks for the two components of the registration mapping T. This is compatible with the requirement that T needs to be zero on the boundary. On the observed subdomain, however, cosine DFT is used for the registration residual, because there is no reason that the change on the residual should be zero on the subdomain boundary.

References

[1] B. D. O. Anderson and J. B. Moore. *Optimal filtering.* Prentice-Hall, Englewood Cliffs, N.J., 1979.

[2] J. L. Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129:2884–2903, 1999.

[3] J. D. Beezley. *High-Dimensional Data Assimilation and Morphing Ensemble Kalman Filters with Applications in Wildfire Modeling.* PhD thesis, University of Colorado Denver, 2009.

[4] J. D. Beezley and J. Mandel. Morphing ensemble Kalman filters. *Tellus*, 60A:131–140, 2008.

[5] E. Castronovo, J. Harlim, and A. J. Majda. Mathematical test criteria for filtering complex systems: plentiful observations. *J. Comput. Phys.*, 227(7):3678–3714, 2008.

[6] N. A. C. Cressie. *Statistics for Spatial Data.* John Wiley & Sons Inc., New York, 1993.

[7] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter.* Springer Verlag, 2nd edition, 2009.

[8] R. Furrer and T. Bengtsson. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivariate Anal.*, 98(2):227–255, 2007.

[9] P. Gao and T. W. Sederberg. A work minimization approach to image morphing. *The Visual Computer*, 14(8-9):390–400, 1998.

[10] B. Hunt, E. Kostelich, and I. Szunyogh. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1-2):112–126, 2007.

[11] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering, Series D*, 82:35–45, 1960.

[12] E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

[13] P. K. Kitanidis. Generalized covariance functions associated with the Laplace equation and their use in interpolation and inverse problems. *Water Resources Research*, 35(5):1361–1367, 1999.

[14] J. Mandel and J. D. Beezley. An ensemble Kalman-particle predictor-corrector filter for non-Gaussian data assimilation. In G. Allen, J. Nabrzyski, E. Seidel, G. D. van Albada, J. Dongarra, and P. M. A. Sloot, editors, *ICCS (2)*, volume 5545 of *Lecture Notes in Computer Science*, pages 470–478. Springer, 2009. arXiv:0812.2290, 2008.

[15] J. Mandel, J. D. Beezley, L. Cobb, and A. Krishnamurthy. Data driven computing by the morphing fast Fourier transform ensemble Kalman filter in epidemic spread simulations. arXiv:1003.1771, 2010. International Conference on Computational Science, Amsterdam May 31–June 2, 2010 (ICCS 2010), Procedia Computer Science, Elsevier, to appear.

[16] J. Mandel, J. D. Beezley, J. L. Coen, and M. Kim. Data assimilation for wildland fires: Ensemble Kalman filters in coupled atmosphere-surface models. *IEEE Control Systems Magazine*, 29:47–65, June 2009.

[17] J. Mandel, J. D. Beezley, and V. Y. Kondratenko. Fast Fourier transform ensemble Kalman filter with application to a coupled atmosphere-wildland fire model. arXiv:1001.1588, 2010. International Conference on Modeling and Simulation, Barcelona July 15–17, 2010, (MS'2010), World Scientific, to appear.

[18] J. Mandel, L. Cobb, and J. D. Beezley. On the convergence of the ensemble Kalman filter. arXiv:0901.2951, January 2009. Applications of Mathematics, to appear.

[19] S. A. Martucci. Symmetric convolution and the discrete sine and cosine transforms. *IEEE Transactions on Signal Processing*, 42(5):1038–1051, 1994.