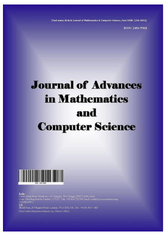

Journal of Advances in Mathematics and Computer Science

36(5): 84-97, 2021; Article no.JAMCS.69688

ISSN: 2456-9968

(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Extended Ensemble Filter for High-dimensional Nonlinear State Space Models

Oryiema Robert[1*], David Angwenyi[1] and Kevin Midenyo[1]

[1] *Masinde Muliro University of Science and Technology, P.O.Box 190-50100, Kakamega, Kenya.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMCS/2021/v36i530365

Editor(s):

(1) Dr. Metin Basarir, Sakarya University, Turkey.

(2) Dr. Doina Bein, California State University, USA.

Reviewers:

(1) Maria Beatriz Bernábe Loranca, Benemérita Universidad Autónoma de Puebla, Mexico.

(2) Anthony (Tony) Spiteri Staines, University of Malta, Malta.

Complete Peer review History: http://www.sdiarticle4.com/review-history/69688

Received: 12 April 2021

Accepted: 17 June 2021

Published: 22 June 2021

Original Research Article

Abstract

There are several functional forms for non-linear dynamical filters. Extended Kalman filters are algorithms that are used to estimate more accurate values of unknown quantities of internal dynamical systems from a sequence of noisy observation measured over a period of time. This filtering process becomes computationally expensive when subjected to high dimensional data which consequently has a negative impact on the filter performance in real time. This is because integration of the equation of evolution of covariances is extremely costly, especially when the dimension of the problem is huge which is the case in numerical weather prediction.

This study has developed a new filter, the First order Extended Ensemble Filter (FoEEF), with a new extended innovation process to improve on the measurement and be able to estimate the state value of high dimensional data. We propose to estimate the covariances empirically, which lends the filter amenable to large dimensional models. The new filter is derived from stochastic state-space models and its performance is tested using Lorenz 63 system of ordinary differential equations and Matlab software.

**Corresponding author: E-mail: oriema@gmail.com;*

The performance of the newly developed filter is then compared with the performances of three other filters, that is, Bootstrap particle Filter (BPF), First order Extended Kalman Bucy Filter (FoEKBF) and Second order Extended Kalman Bucy Filter (SoEKBF).
The performance of the FoEEF improves with the increase in ensemble size. Even with as low number of ensembles as 40, the FoEEF performs as good as the FoEKBF and SoEKBF. This shows, that the proposed filter can register a good performance when used in high-dimensional state-space models.

Keywords: Bayesian technique; filtering; estimation; state space dynamical system; extended ensemble Kalman filter.

2010 Mathematics Subject Classification: 62F15; 62F40; 62M05.

1 Introduction

Stochastic state space models incorporate random variables, or uncertainty, in their equation.

The dynamics of these models are given by the following stochastic differential equation, as developed by Kiyosi Itô [1]:

$$dw_t = f(w_t, t)dt + g(w_t, t)d\beta_t, \qquad t_0 \le t, \tag{1.1}$$

where w_t is a variable that represents an n by 1 dimensional state vector, $f(w_t, t)$ is an n by 1 dimensional drift function that enforces the evolution of the state vector w_t in any given change of time dt, $g(w_t, t)$ is an n by m dimensional diffusion function and β_t with $t > t_0$ is an m by 1 standard Brownian motion process that incorporates the uncertainties in the process.

1.1 Stochastic integrals

The solution to equation (1.1) can be presented as

$$w_t = w_{t_0} + \int_{t_0}^{t} f(w_\tau, \tau)d\tau + \int_{t_0}^{t} g(x_\tau, \tau)d\beta_\tau \tag{1.2}$$

The last integral, $\int_{t_0}^{t} g(w_\tau, \tau)d\beta_\tau$, is a stochastic integral, and can take any of the following formulations:

1. Itô
$$\int_a^b g(w_t, t)d\beta_\tau = l.i.m. \sum_{k=1}^{n} g(w_{t_k}, t_k)(\beta_{t_{k+1}} - \beta_{t_k}) \tag{1.3}$$
where $\delta t = t_{k+1} - t_k$, and,
2. Stratonovich
$$\int_a^b g(x_t, t)' d\beta_\tau = l.i.m \sum_{k=1}^{n} g(\frac{w_{t_k} + w_{t_{k+1}}}{2}, t_k)(\beta_{t_{k+1}} - \beta_{t_k}). \tag{1.4}$$

The relationship between Itô and Stratonovich integrals is given by [2]

$$\int_a^b g(w_t, t)' d\beta_\tau = \int_a^b g(w_t, t)d\beta_\tau + \frac{1}{2}\int_a^b \partial_w[g(w_\tau, \tau)]g(w_\tau, \tau)d_\tau \tag{1.5}$$

With $\partial_w[g]$ representing $\frac{\partial g}{\partial w}$ derivative.

Using equation (1.5), the Itô differential equation $dw_t = f(w_t, t)dt + g(w_t, t)d\beta_t$ leads to the Stratonovich differential equation

$$dw_t = f(w_t, t) - \frac{1}{2}g(w_t, t)\partial_w[g(w_t, t)]dt + g(w_t, t)'d\beta_t. \tag{1.6}$$

1.2 Fokker-Planck equation

Taking the density function as

$$\psi_t(w) = \psi(w, t) \tag{1.7}$$

and the transition probability density function as

$$\psi_{t|\tau}(w|Y) = \psi(w, t|Y, \tau) \tag{1.8}$$

The Fokker-Planck equation describes the evolution of the density function (1.7) and the conditional density function (1.8) and satisfies the Chapman-Kolmogorov equation [3, 2]

$$\psi_{t_3|t_1}(w|Y) = \int \psi_{t_3|t_2}(w|Z)\psi_{t_2|t_1}(y|Y)dy$$

$$t_1 < t_2 < t_3 \tag{1.9}$$

The equation for the evolution of transitional probability density function can be derived using Taylor series expansion [2]

$$\frac{\partial\psi_{t|\tau}(w|y)}{\partial t} = -\frac{\partial(\psi_{t|\tau}(w|y)f(w, t))}{\partial w} + \frac{1}{2}\frac{\partial^2(\psi_{t|\tau}(w|z)g^2(w, t))}{\partial w^2} \tag{1.10}$$

This equation is called the*Fokker-Planck* equation or the *Kolmogorov forward equation.*
Carrying out the expected value of the Kolmogorov forward equation yields the equation of evolution of probability density function

$$\frac{\partial\psi_t(w)}{\partial t} = -\frac{\partial(\psi_t(w)f(w, t))}{\partial w} + \frac{1}{2}\frac{\partial^2(\psi_t(w_t)g^2(w, t))}{\partial w^2}$$

given that

$$E_\tau[\psi_{t|\tau}(w|z)] = \int \psi_{t|\tau}(w|z)\psi_\tau(z)dz = \psi_t(w). \tag{1.11}$$

1.3 Weak form of Fokker-Planck equation (*equation of evolution of moments*)

The weak form of Fokker-Planck equation is obtained by multiplying the Fokker-Planck equation(1.11) with an infinitely differentiable function $\phi(w)$ and then integrating; that is,

$$\int \frac{\partial\psi_w}{\partial t}\phi(w)dw = -\int \frac{\partial(\psi_t(w)f(w, t)}{\partial w}\phi(w)dw + \frac{1}{2}\int \frac{\partial^2(\psi_t(w)g^2(x, t))}{\partial w^2}\phi(w)dw \tag{1.12}$$

Then integrating by parts and getting the expected value $E(.)$[3] yields

$$d\psi_t[\phi] = \psi_t[\ell^*\phi]dt \quad \text{where} \quad \ell^* = f\frac{\partial}{\partial w} + \frac{1}{2}g^2\frac{\partial^2}{\partial w^2}. \tag{1.13}$$

The idea of filtering being the incorporation of measurements into the model, it now remains to introduce the measurements equation, which is as follows:

$$dz_t = h(w_t, t)dt + R^{\frac{1}{2}}d\eta_t \tag{1.14}$$

whose dimensions and significations correspond to those of (1.1). We only need mention that η_t is white noise of appropriate dimension.

1.4 Kushner-Stratonovich equation

The Kushner-Stratonovich equation is a perturbation of Fokker-Planck equation (1.12), which is addition of knowledge from measurement using Bayesian approach. The equation estimates $\hat{w}_t$ at time t by combining noisy dynamics with noisy measurement, and is as follows: [4]

$$\psi_t(w/z_t) = \psi_{t_0}(w) + \int_{t_0}^{t} \ell(\psi_t(w/z_t))d\tau + \int_{t_0}^{t} \psi_\tau(w/z_\tau)(h - \hat{h}_\tau)^T R^{-1}(\tau)(dz_\tau - \hat{h}_\tau d\tau) \quad (1.15)$$

which can be written in its differential form as

$$d\psi_t(w/z_t) = \ell(\psi_t(w/z_t))d\tau + \psi_\tau(w/z_\tau)(h - \hat{h}_\tau)^T R^{-1}(\tau)(dz_\tau - \hat{h}_\tau d\tau) \quad (1.16)$$

where $\ell = f\frac{\partial}{\partial w} + \frac{1}{2}g^2\frac{\partial^2}{\partial w^2}$ and $\hat{h}_t = \int h(w,t)\psi_t(w/z_t)dw$.

The proof can be found in [3], but we here only give a sketch. The weak form of Kushner-Stratonovich equation is

$$d\psi_t[\phi] = \psi_t[\ell\phi]dt + (\pi_t[\phi(w)h] - \hat{\phi}\hat{h}_t)^T R^{-1}(t)(dz_t - \hat{h}_t dt)] \quad (1.17)$$

Deriving the innovation process

$$[d\psi_t]_{measurement} = [\pi_t(w/z_t + \delta z_t) - \psi_t(w/z_t)] \quad (1.18)$$

from

$$\delta z_t = h(w_t)\delta t + R^{\frac{1}{2}}(t)\delta\eta_t$$

and taking

$$E[\delta z_t \delta z_t^T] = R(t)\delta t + ...$$

we apply ***Bayesian equation***

$$\psi_t(w/z_t + \delta z_t) = \frac{\psi_t(\delta z/w_t)\psi_t(w/z_t)}{\int \psi_t(\delta z/w_t)\psi_t(w/z_t)dw}$$

where

$$\psi_t(\delta z/w_t) \sim N(h(w_t)\delta t, R(t)\delta t)$$

let

$$q(\delta t, \delta z_t) = \frac{\psi_t(w/z_t + \delta t)}{\psi_t(w/z_t)}$$

using Taylor expansion about (0,0) and taking the expectation

$$q(\delta t, \delta y_t) = 1 + (h - \hat{h})^T R^{-1}(\delta z_t - \hat{h}_t \delta t) + ...$$

hence applying equation

$$q(\delta t, \delta z_t)\psi_t(w/z_t) = \psi_t(w/z_t + \delta z_t)$$

we get

$$\psi_t(w/z_t + \delta z_t) = \psi_t(w/z_t) + (h - \hat{h}_t^T R^{-1}(\delta z_t - \hat{h}_t \delta t)\psi_t(w/z_t) + ...$$

This gives equation the innovation process as $\delta t \to 0$

1.5 Evolution of the mean and co-variance

The mean is the first moment and the variance is the second moment [3]. For a function of w given by $\phi(w)$ and conditional density $\psi_t(w/z_t)$

The first moment is

$$\psi_t[\phi(w)] = \int \phi(w)\pi_t(w/z_t)dw \quad (1.19)$$

Using the weak form of Kushner-Stratonovich equation (1.17) and taking $\psi_t[\phi(w)] = \hat{\phi}_t(w)$ we get

$$d\hat{\phi}_t(w) = \psi_t[f\frac{\partial \phi}{\partial w}]dt + \frac{1}{2}\psi_t[g^2\frac{\partial^2 w}{\partial w^2}]dt + (\psi_t[\phi h] - \hat{\phi}_t\hat{h}_t)(R)^{-1}(t)(dz_t - \hat{h}_t dt) \tag{1.20}$$

When we substitute $\phi(w)$ with w, our equation becomes

$$d\hat{w}_t = \hat{f}_t dt + (\psi_t[wh] - \hat{w}_t\hat{w}_t)R^{-1}(t)(dz_t - \hat{h}_t dt). \tag{1.21}$$

The conditional variance is derived from (see [2] for finer details)

$$\begin{aligned} p_t &= E[(w-\hat{w}_t)^2|z_t] \\ &= \psi_t[w^2] - \hat{w}_t^2 \\ \Rightarrow dp_t &= d\psi_t[w^2] - d\hat{w}_t^2. \end{aligned} \tag{1.22}$$

Substituting $\phi(w)$ for w^2 we get

$$d\psi_t[w^2] = 2\psi_t[wf]dt + \psi_t[g^2]dt + (\psi_t[w^2h] - \psi_t[w^2]\hat{h}_t)R^{-1}(t)(dz_t - \hat{h}_t dt). \tag{1.23}$$

Applying Itô formula [2] of measurements

$$dh = \partial_t[h]dt + \nabla[h]^T dw_t + \frac{1}{2}tr[g(w_t,t)g^T(w_t,t)]\Delta[h]dt,$$

the equation becomes

$$d\hat{w}_t^2 = 2\hat{w}_t d\hat{w}_t + (\psi_t[wh] - \hat{w}_t\hat{h}_t)^2 R^{-1}(t)dt. \tag{1.24}$$

Substituting from (1.21), we get

$$d\hat{w}_t^2 = 2\hat{w}_t\hat{f}dt + (2\hat{w}_t\pi_t[wh] - 2\hat{w}^2\hat{h}_t)R^{-1}(t)(dy_t - \hat{h}_t dt) + (\pi_t[wh] - \hat{w}_t\hat{h}_t)^2 R^{-1}(t)dt. \tag{1.25}$$

The evolution of co-variance is developed from (1.5), (1.24) and (1.25) to obtain

$$\begin{aligned} dp_t = (2\psi_t[wf] - 2\hat{w}_t\hat{f}_t)dt + \psi_t[g^2dt - (\psi_t[xh] - \hat{w}_t\hat{h}_t)^2R^{-1}(t)dt + \\ (\psi_t[w^2h] - 2\hat{w}_t[wh] - \psi_t[w^2]\hat{h}_t + 2\hat{w}_t^2\hat{h}_t)R^{-1}(t)(dy_t - \hat{h}_t dt). \end{aligned} \tag{1.26}$$

2 Literature Review

Equations (1.25) and (1.5), are exact solutions to the filtering problem. However, they are infinite dimensional, especially when expected values of nonlinear functions are involved. For the linear case, the equations naturally yield the Kalman Bucy filter [5], which is optimal. In the nonlinear setting, the Kalman Bucy filter is not optimal. Extensions of the Kalman Bucy filter are, however, made by way of approximating the expected values of nonlinear functions using Taylor series expansion about the mean of the state variable. These filter approximations are derived in the sequel. But first particle filters.

A Particle filter [6] or sequential Monte Carlo method applies numerical methods in approximating the state based on importance sampling [7]. The idea is to draw, say, M samples, $\{w^i_{t_0:t_N}, w^{*i}_{t_N}\}_{i=1}^M$, where w^* represents corresponding weights, from a posterior density

$$\pi_{t_0:t_N}\left(w_{t_0:t_N} \mid \delta z_{t_1:t_N}\right),$$

where $w_{t_0:t_N} = (w_{t_0}^{\mathrm{T}},\, w_{t_1}^{\mathrm{T}},\, \ldots w_{t_N}^{\mathrm{T}})^{\mathrm{T}}$ and $\delta z_{t_0:t_N} = (\delta z_{t_0}^{\mathrm{T}},\, \delta z_{t_1}^{\mathrm{T}},\, \ldots \delta z_{t_N}^{\mathrm{T}})^{\mathrm{T}}$. In this case, $t_0 : t_N$ is the partition of time. However, as in most cases, drawing from the full posterior can be cumbersome. in stead, we may choose a simpler density which resembles the posterior, from which we can draw samples. Suppose such a proposal be the same as the transition density. Then the weights can be represented by the likelihood. This leads to the so-called Bootstrap particle filter (BPF). Resampling is done–in a bid to curb particle degeneracy–when the effective sample size (ESS) is below a certain threshold, α.

$$\text{ESS}_{t_n} = \frac{1}{\sum_{i=1}^{M} (\tilde{w}_{t_n}^{*i})^2}, \tag{2.1}$$

where $\{w_{t_n}^{*i}\}_{i=1}^{M}$ are normalized weights.

Algorithm 2.1 Bootstrap particle filter

Input: δt, α, M, N, π_{t_0}, π_{t_n} and w_{t_0}.
Output: $\{w_{t_n}^{i}\}_{n,i=1}^{N,M}$.

1: Draw $w_{t_0}^{i} \sim \pi_{t_0}\left(w_{t_0}^{i} \mid w_{t_0}\right)$
2: Compute initial weights $w_{t_0}^{*i} = \dfrac{1}{M}$
3: **for** $n = 1$ **to** N, $\delta t > 0$ **do**
4: Draw $w_{t_n}^{i} \sim \pi_{t_n}\left(w_{t_n}^{i} \mid w_{t_{n-1}}^{i}\right)$
5: Compute weights $w_{t_n}^{*i} \sim \pi_{t_n}\left(\delta z_{t_n} \mid w_{t_n}^{i}\right) w_{t_{n-1}}^{*i}$
6: Normalise the weights
7: Compute $\text{ESS}_{t_n} = \dfrac{1}{\sum_{i=1}^{M} (\tilde{w}_{t_n}^{*i})^2}$
8: **if** $\text{ESS}_{t_n} \leq \alpha$ **then**
9: Resample the particles
10: Set the weights to $w_{t_n}^{*i} = \dfrac{1}{M}$
11: **end if**
12: **end for**

The need for resampling disallows the use of parallel computers, especially when the dimensions are high. Furthermore, the sample size required to achieve accuracy of particle filters scales according to the exponent of the dimension of the filter, a phenomenon dubbed the *curse of dimensionality* [8]. All these challenges work against the particle filters.

Kushner-Stratonovic equation gives solution to filtering problems by propagating the conditional density in time. The equation of conditional density is known to be infinite[2] hence it suffers from the curse of dimensionality [3]. Non-linear filtering problems have high order moments in solving equations of the mean which, that is, as the dimension of the state increases, the space becomes large and the data becomes sparse. This problem leads to the needs of large amount of data. Alternatively, approximation of the equations for both the mean and the covariance are done of the first order and the second order by using Taylor expansion in order to find a numerical solution to the filtering problem.

3 Extended Filters

Consider the following Itô expression of continuous state space models:

$$\text{Signal:} \quad dw_t = f(w_t, \theta)dt + G(w_t)Q^{1/2}(t)d\beta_t \tag{3.1}$$
$$w_{t_0} = w(0), t_0 \leq t$$

$$\text{Measurement:} \quad dz_t = h(w_t)dt + R^{1/2}(t)d\eta_t \quad t_0 \le t \tag{3.2}$$

where w_t is state vector with dimensions n by *1*, $f(w_t, \theta)$ is the drift term with dimensions n by *1*, θ is the vector of parameters with dimension d by *1*, $G(w_t)$ is a matrix function with dimensions n by m $Q(t)$ is a time function matrix with dimensions m by m, $\beta_t, t > t_0$ is the Brownian motion with dimensions m by *1*, z_t is an output vector with dimensions r by *1*, $h(x_t)$ is the measurement function with dimensions n by *1*, $R(t)$ is a time function matrix with dimensions r by r, $\eta_t, t > t_0$ is the Brownian motion process with dimensions r by *1*,

An estimate of the state w_t can be obtained from the historical measurements $z_t, t > t_0$ by using conditional density function.

For equation (3.1) and (3.2), the mean, $\hat{w}_t$, and co-variance p_t satisfies, respectively,

$$d\hat{w}_t = \hat{f}dt + (\widehat{w_t h^T} - \hat{w}_t \hat{h}^T)R^{-1}(t)(dy_t - \hat{h}dt) \text{ where } \quad w_{t_0} = w(0), \tag{3.3}$$

and

$$(dp)_{ij} = (\widehat{w_i f_j} - \widehat{w_i \hat{f}_j})dt + (\widehat{f_i w_j})dt + (\widehat{GQG^T})_{IJ}dt - (\widehat{w_i h} - \hat{w}_i \hat{h})^T R^{-1}(\widehat{h w_j} - \hat{h}\hat{w}_j)dt$$
$$+(\widehat{w_i w_j h} - \widehat{w_i w_j}\hat{h}) - (\hat{w}_i \widehat{w_j h} - \hat{w}_j \widehat{w_j h} + 2\hat{w}_i \hat{w}_j \hat{h})^T R^{-1}(t)(dy_t - \hat{h}dt) \tag{3.4}$$

where $p_{t_0} = p(0)$.

For a scalar case, equation (3.1) becomes

$$\text{Signal: } dw_t = f(w_t, \theta)dt + g(w_t)q^{1/2}(t)d\beta_t \text{ where } \quad w_{t_0} = w(0), t_0 \le t \tag{3.5}$$

and equation (3.2) becomes

$$\text{Measurement: } dz_t = h(w_t)dt + r^{1/2}(t)d\eta_t, \quad t_0 \le t. \tag{3.6}$$

The corresponding equations for evolution of the mean and covariance in the scalar case are:

$$d\hat{w}_t = \hat{f}dt + (\widehat{w_t h^T} - \hat{w}_t \hat{h}^T)r^{-1}(t)(dy_t - \hat{h}dt), \quad w_{t_0} = w(0) \tag{3.7}$$

and

$$(dp_t) = 2(\widehat{wf})dt + (qg^2)dt - (\widehat{wh} - \hat{w}\hat{h})^2 r^{-1} + (\widehat{w^2 h} - \hat{w^2}\hat{h} - 2\hat{w}\widehat{wh} + 2\hat{w}^2\hat{h})r^{-1}(t)(dz_t - \hat{h}dt) \quad p_{t_0} = p(0) \tag{3.8}$$

where $\hat{f} = \int f(w)p(w/z_t)dw$. Solutions for equations (3.7) and (3.8) give exact filter except it involves calculation of conditional expectation, an integration over non-linear functions which is not feasible. In this case we calculate an approximation of expected values by negating terms of the second and higher order terms in the Taylor series expansion of nonlinearities about the mean of the state, $\hat{w}_t$.

3.1 First order approximate filter

Proposition 3.1. *Suppose $f(w)$ and $h(w)$ are continuous functions whose first order derivatives f_w and h_w exist, then the first approximation equation for the exact filter equations* (3.7) *and* (3.8) *and omitting second and higher order moments are*

$$d\hat{w}_t = f(\hat{w})dt + ph_w(\hat{w})r^{-1}(t)(dy_t - h(\hat{w})dt) \tag{3.9}$$

$$dp_t = 2pf_w(\hat{w})dt + (q(t)g^2(\hat{w}) + pq(t)g_w^2(\hat{w})dt - (p_t h(\hat{w}))^2 r^{-1} \tag{3.10}$$

Proof. Proof by Taylor expansion about $\hat{w}$ and taking the expectations of the series $E(\bullet)$ upto the first order with $p = (\widehat{w-\hat{w}})^2 = w(\widehat{w-\hat{w}})$ and $(\widehat{w-\hat{w}}) = 0$

$$f(w) \approx f(\hat{w}) + (w-\hat{w})f_w(\hat{w}) + \frac{1}{2}(w-\hat{w})^2 f_w(\hat{w}) + ... \quad (3.11)$$

$$\hat{f} \approx f(\hat{w}) + \frac{1}{2}pf_{ww}(\hat{w})$$

$$\hat{f} \approx f\hat{w}$$

$$h(w) \approx h(\hat{w}) + (x-\hat{w})h_w(\hat{w}) + \frac{1}{2}(w-\hat{w})^2 h_{ww}(\hat{w}) + ... \quad (3.12)$$

$$\hat{h} \approx \hat{h(w)} + \frac{1}{2}ph_{ww}(w)$$

$$\hat{h} \approx \hat{h(w)}$$

$$q(t)(g^2(w) \approx q(t)(g(\hat{w}) + (w-\hat{w})g_w(\hat{w}) + \frac{1}{2}(w-\hat{w})^2 g_{ww}(\hat{w}))^2 + ...$$

$$\approx q(t)(g(\hat{w}) + \frac{1}{2}pg_{ww}(\hat{w}))^2$$

$$q(t)(g(\hat{w})^2 + 2(\frac{1}{2}pg_{ww}(\hat{w})g(\hat{w}) + (\frac{1}{2}pg_{ww}(\hat{w}))^2$$

$$\approx q(t)g(t\hat{w})^2 + pq(t)g_{ww}(\hat{w})g(\hat{w}) + \frac{1}{4}q(t)p^2(g_{ww}(\hat{w}))^2$$

$$q(t)g^2 \approx q(t)g^2(\hat{w}) + pq(t)g_x^2(\hat{w})$$

$$\widehat{q(t)g^2} \approx q(t)g^2\hat{w} + pq(t)g_w^2(\hat{w}) \quad (3.13)$$

$$wf(w) \approx wf(\hat{w}) + w(w-\hat{w})f_w(\hat{w}) + \frac{1}{2}w(w-\hat{w})^2 f_{ww}(\hat{w}) + ...$$

$$\hat{wf} - \hat{w}\hat{f} \approx \hat{w}f(\hat{w}) + pf_w(\hat{w}) + \frac{1}{2}\hat{w}pf_{ww}(\hat{w}) - \hat{w}f(\hat{w}) - \frac{1}{2}pf_{ww}(\hat{w})$$

$$= pf_w(\hat{w}) \quad (3.14)$$

We now seek the approximation of the term

$$\widehat{w^2h} - \widehat{w^2}\hat{h} - 2\hat{w}\widehat{wh} + 2\hat{w}^2\hat{h}.$$

By Taylor series expansion,

$$\widehat{w^2} \approx \hat{w}^2 + (w-\hat{w})2\hat{w} + (w-\hat{w})^2. \quad (3.15)$$

Using equations (3.12) and (3.15) we get

$$\widehat{w^2h} \approx E[(\hat{w}^2 + (w-\hat{w})2\hat{w} + (w-\hat{w})^2) * (h\hat{w}) + (w-\hat{w})h_w(\hat{w}) + \frac{1}{2}(w-\hat{w})^2 h_{ww}(\hat{w})]$$

$$= \hat{w}^2 h(\hat{w}) + (\hat{w}) + 2\hat{w}ph_w(\hat{w}) + ph(\hat{w}) \quad (3.16)$$

similarly

$$\hat{w^2}\hat{h} \approx E[(\hat{w}^2 + (w-\hat{w})2\hat{w} + (w-\hat{w})^2)] * E[(h(\hat{w}) + (w-\hat{w})h_w(\hat{w}) + \frac{1}{2}(w-\hat{w})^2 h_{ww}(\hat{w}))]$$

$$= (\hat{w}^2 + p)(h(\hat{w}) + (w-\hat{w}) + \frac{1}{2}(w-\hat{w})^2 h_{ww}(\hat{w}))$$

$$= \hat{w}^2 h(\hat{w}) + \frac{1}{2}\hat{w}^2 p h_{ww}(\hat{w}) + ph(\hat{w}) + \frac{1}{2}p^2 h_{ww}(\hat{w})$$

$$\widehat{w^2 h} = \hat{w}^2 h(\hat{w}) + ph(\hat{w}). \tag{3.17}$$

The remaining term

$$-2\hat{w}\widehat{wh} + 2\hat{w}^2\hat{h}$$

becomes

$$-2\hat{w}\widehat{wh} + 2\hat{w}^2\hat{h} = -2\hat{w}(\widehat{wh} - \hat{w}\hat{h}) = 2\hat{w}ph_w(\hat{w}) \tag{3.18}$$

The approximate filter for exact filter (3.7) and (3.8) can be obtained by writing the expected values of the terms with their approximations from above. Equations (3.9) and (3.10) jointly form the first order extended Kalman Bucy filter (FoEKBF).

The second order extended Kalman Bucy filter (SoEKBF) involves retaining of second order terms in the Taylor series expansion of the expected values of nonlinear terms in the exact filter [3]. The SoEKBF is given by (in vector form):

$$\begin{aligned} d\hat{w}_t &= f(\hat{w}_t)dt + \frac{1}{2} \triangle [f](\hat{w}_t) : P_t dt \\ &\quad + P_t[h]^{\mathrm{T}}(\hat{w}_t)R^{-1}(t)(dz_t - (h(\hat{w}_t) + \frac{1}{2} \triangle [h](\hat{w}_t) : P_t)dt), \end{aligned} \tag{3.19a}$$

$$\begin{aligned} dP_t &= P_t[f]^{\mathrm{T}}(\hat{w}_t)dt + [f](\hat{w}_t)P_t dt \\ &\quad + g(t)g^{\mathrm{T}}(t)dt + -P_t[h]^{\mathrm{T}}(\hat{w}_t)R^{-1}[h](\hat{w}_t)P_t dt \\ &\quad + \frac{1}{2}P_t : \triangle[h]^{\mathrm{T}}(\hat{w}_t)R^{-1}(t)\left(dz_t - (h(\hat{w}_t) + \frac{1}{2} \triangle [h](\hat{w}_t) : P_t)dt\right)P_t. \end{aligned} \tag{3.19b}$$

Now the first order–and second order–approximate filter suffers from one demerit; namely, the cost of integrating the equation of evolution of covariances. This is a major problem, especially where high dimensional models are involved. Moreover, the computation of Jacobian can be expensive when nonlinearities abound. These challenges led to the invention of ensemble filters, in which the covariances are empirically estimated [9].

3.2 First order extended ensemble Kalman filter

In the proposed filter, an M hypothesis of state, denoted as $X_t = w_i^M$ at given time t, the particles are propagated according to the first order approximate equation, (3.9); that is,

$$dw_t^i = f(w_t^i, \theta)dt + g(w_t^i)q^{\frac{1}{2}}(t)d\beta_t^i + ph_w(w_t^i)r^{-1}(t)(dz_t - 0.5(h(w_t^i) + h(\hat{w}_t))dt) \tag{3.20}$$

The mean w_t and co-variance p_t are estimated empirically as follows:

$$\hat{w}_t = \frac{1}{M}\sum_{i=1}^{M} w_t^i \tag{3.21}$$

$$t_0 \geq t$$

and

$$p_t = \frac{1}{M-1}\sum_{i=1}^{M}(w_t^i - \hat{w}_t)^2 \tag{3.22}$$

$$t_0 \geq t$$

This is the first order extended ensemble filter (FoEEF).

4 Filter Validation Using Multi-Dimensional Model

In this section, we test the performance of FoEEF using Lorenz63 equations and we develop the algorithm in Matlab software. The same experiment is done for BPF, FoEKBF and SoEKBF.

The new FoEEF filter is validated by comparing its graphical output with the output from three other filters, that is ,Bootstrap Particle Filter(BPF), First order Extended Kalman-Bucy Filter (FoEKBF) and Second order Extended Kalman-Bucy Filter(SoEKBF) by using lorenz63 systems of ordinary differential equations simulated on Matlab software. We plot the trajectories for the three variables x_1, x_2 and x_3, for the different filters, together with the root mean square errors (RMSEs).

But first we briefly introduce the Lorenz63 model.

4.1 Stochastic Lorenz63 equation

Lorenz63 equations is a system of coupled non-linear ordinary differential equations that best describes chaotic dynamics. The equations are known to simulate regular oscillations or highly non-linear fluctuations[10]. Lorenz63 equations oscillates around fixed point $p = (0, 0, 0)$ and a slight change in initial conditions leads to a significant change in oscillations. The coupled equations describe the rate of change of three quantities x, y, z with respect to time t.

Lorenz63 model is defined in three time dependant variables.

$$\frac{dx}{dt} = a(x_2 - x_1)$$

$$\frac{dy}{dt} = x_1(b - x_3) - x_2$$

$$\frac{dz}{dt} = x_1x_2 - cx_3 \tag{4.1}$$

Where $a,\ b,\ c$ are the parameters generally with values $a = 10$, $b = 28$ and $c = \frac{8}{3}$

Stochastic Lorenz63 contains an additive noise part in Lorenz63 model. The noise is assumed to be Gaussian in nature. Lorenz determined that a slight perturbation in the initial stage changes the outcome of a dynamical system drastically. This sensitivity to initial value condition has led to application of Lorenz63 equations to data assimilation and filtering problems. Stochastic Lorenz 63 equations is used in finding chaotic solutions for parameters with given initial conditions.

5 Results

Fig. 1 shows a plot of RMSE against the reciprocal of ensemble size M for different filters. The trend indicates that the FoEEF improves in performance with the increase in ensemble size. The FoEEF performs in about the same way as the FoEKBF and SoEKBF.

From Fig. 1, output measurement of Root Mean Square Error against the number of the ensemble, it is observed that there is no significant deviation in performance from the new FoEEF away from the other three filters, BPF, FoEKBF and SoEKBF, especially at large ensemble sizes.

Fig. 2 has two graphs. The first graph compares the trajectories of evolution of new FoEEF (dark green) with the trajectories of SoEKBF (purple), FoEKBF (yellow), BPF (brown) and the trajectory of true state(blue) generated from Lorenz63 model.

The trajectories are conducted on filter estimates using nine ensemble number that are measured between times $T = 0$ and $T = 30$, with a time step of 0.001.

The sefcond graph is a measure of filter error against times $T = 0$ and $T = 30$ for FoEEnKF model(purple), SoEKBF model(yellow), FoEKBF model(orange), BPF model(blue) in the first variable.

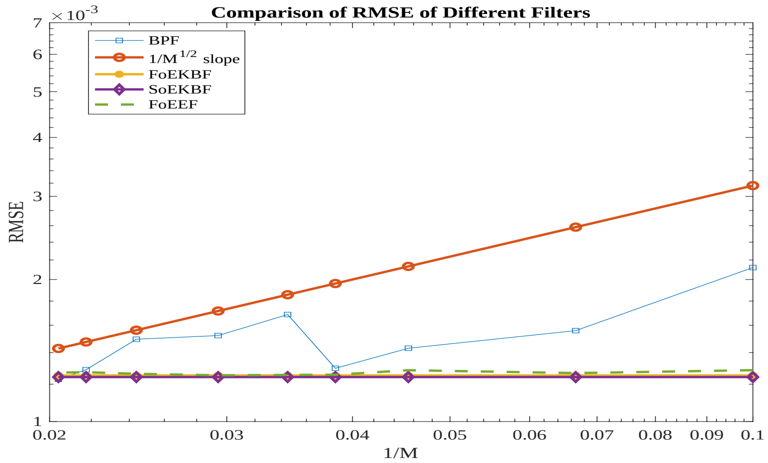

Fig. 1. Root Mean Square Error for the reciprocal of Ensemble. The sizes of the ensemble used are 10, 15, 22, 26, 29, 34, 41, 46, 49. Other settings are as follows: $dt = 0.001$, $R = 0.17$, **and** $g = 0.2$

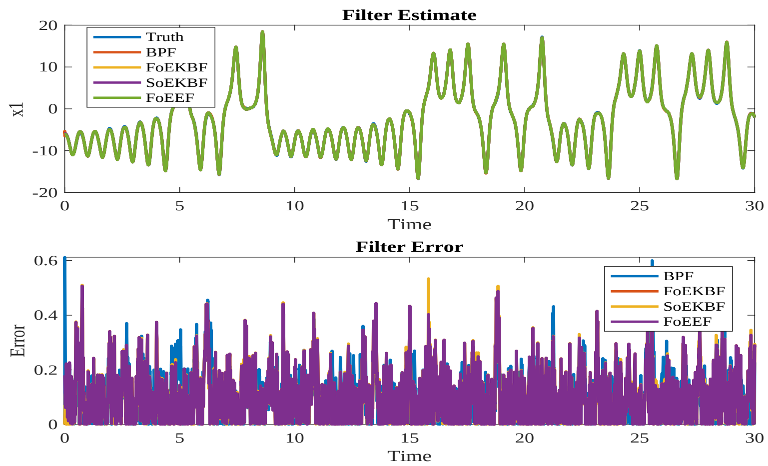

Fig. 2. Filter estimate and filter error for x_1 **in lorenz63 model**

The observations from filter estimate in Fig. 3 shows that there is no deviations in trajectory of evolution between FoEEF (green) and the trajectory of the true state generated from Lorenz63 model (blue). It can also be seen that the three other models, that is SoEKBF (purple), FoEKBF (orange) and BPF (red), also shows similar trajectories of evolution with both the true state trajectory and FoEKBF model trajectory.

From Fig. 3, it can be observed that there is no significant deviation that occurs between the output of FoEEF model and the output from other three models. The output of new FoEKBF model closely resembles the output from SoEKBF model, FoEKBF model and BPF model.

Fig. 4 has two graphical outputs with the first output comparing the trajectories of evolution of new FoEEnKF (dark green), SoEKBF (purple), FoEKBFn(yellow), BPF (brown) and the trajectory of true state(blue) for the third variable in Lorenz63 model. The trajectories are conducted for Filter Estimates using nine ensemble number which are measured between times $T = 0$ and $T = 30$.

The second graph measures the output from filter error against times $T = [0, 30]$ for FoEEnKF (purple), SoEKBF (yellow), FoEKBF (orange), BPF (blue).

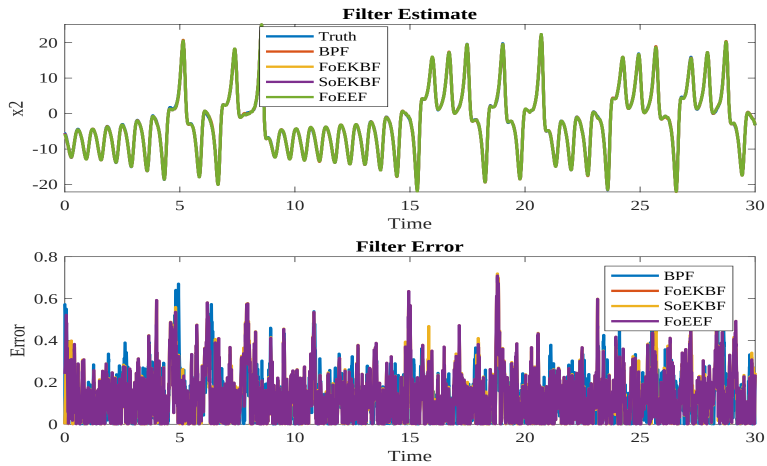

Fig. 3. Filter estimate and filter error for x_2 in Lorenz 63 model

Like in the previous figures above, it is observed that the trajectories of FoEEF (dark green), SoEKBF (purple), FoEKBF (yellow), BPF (brown) and true state (blue) that has been generated from Lorenz63 have similar trends by having no observable deviation between the five models.

The measurement on the filter error against time from the same Fig. 4 indicates no significant difference between the outputs from the four filter models(FoEEF, SoEKBF, FoEKBF, and BPF).

As with the other trajectories, it is observed that all the trajectories closely follow same pattern. There is no significant deviation of the FoEEF trajectories from those of BPF, FoEKBF and SoKBF.

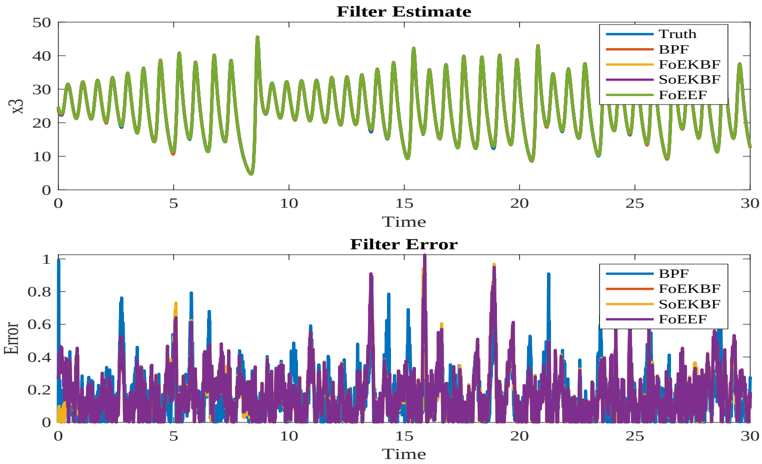

Fig. 4. Filter Estimate and Filter Error for x_3 in Lorenz63 model

6 Conclusion

The observation from the experiment of FoEEF, FoEKBF and SoEKBF conducted on a three dimensional Stochastic Lorenz63 has shown that the performance FoEEF, as also shown in RMSE plot, does not deviate in any significant way from the performances of FoEKBF and SoEKBF. The graphical trajectories of FoEEF in the same three dimensional Lorenz63 also indicates similarity with trajectories of FoEKBF and SoEKBF for the three variables, x_1, x_2 and x_3. It is well known that as the dimensions increases, solutions to filtering problems become computationally expensive since it requires integration of the equation of evolution of covariances. This cost is drastically reduced in the case of FoEEF, which registers a remarkable performance even at low ensemble sizes (refer to Fig. 1).

In FoEEF, the covariance matrix is estimated empirically which makes it computationally cheap as compared to FoKBF and SoKBF, where integration of an equation of evolution of covariances is involved. Further more, FoEEF does not involve re-sampling, which is common in particle filter as a way of circumventing the so called particle degeneracy. It is well known that re-sampling prohibit parallel computing. This makes FoEEF amenable to parallel computing hence it is more suitable for high dimensional data.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Kiyosi It. Stochastic Integral. Proc. Imperial Acad. Tokyo. 1944;20:51952.

[2] Jazwinski A. Stochastic Processes and Filtering Theory, Academic Press, Inc., New York, Dover Edition; 2007.

[3] Angwenyi D. Time-Continuous State and Parameter Estimation with application to SPDEs. PhD Thesis, Institute Fur Mathematik AG Numerische Mathematik, University of Potsdam; 2019.

[4] Kushner H. On the differential equations satisfied by conditional probability densities of markov processes, with applications. J. SIAM Control, Ser. A. 1962;2:1.

[5] Kalman R, Bucy R. New result in linear filtering and prediction theory. Journal of Basic Engineering; 1961.

[6] Arulampalam S, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on Signal Processing. 2002;50(2):174-188.

[7] Kantas N, Doucet A, Maciejowski J. An overview of sequential Monte Carlo methods for parameter estimation in general state-space models; 2017.

[8] Reich S, Cotter C. Probabilistic forecasting and Bayesian data assimilation. Cambridge University Press; 2015.

[9] Evensen G. Data Assimilation. The Ensemble Kalman Filter, Springer-Verlag, New York; 2006.

[10] Zheqi Shen, Yaoumin Tang. A modified ensemble Kalman particle filter for non-Gaussian systems with non-linear measurement functions. Journal of Advances in Modelling Earth Sciences; 2014.
Available:https://doi.org/10.1002/2014MS000373

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/69688