Particle filter and EnKF as data assimilation methods for the Kuramoto-Sivashinsky Equation

M. Jardak [1] & I.M. Navon [2]

School of Computational Science
Florida State University
Tallahassee, FL 32306-4120, USA

Abstract

The Kuramoto-Sivashinsky equation plays an important role as a low-dimensional prototype for complicated fluid dynamics systems having been studied due to its chaotic pattern forming behavior. Up to now, efforts to carry out data assimilation with this 1-d model were quasi totally restricted to variational adjoint methods domain and only Chorin and Krause [26] tested it using a sequential Bayesian filter approach. In this work we compare the usual ensemble Kalman filter (EnKF) approach versus versions of the sequential Monte-Carlo particle filter approach and compare in detail their relative performance for both linear and nonlinear observation operators. Results of this sequential data assimilation tests are discussed using several versions of the particle filter addressing the important issue of resampling to avoid filter degeneracy.

Key words: Sequential data assimilation, ensemble Kalman filter, particle filter,resampling, Kuramoto-Sivashinsky equation.

[1] Email: mjardak@scs.fsu.edu
[2] Email: navon@scs.fsu.edu

1 Introduction

Sequential data assimilation is a relatively novel and versatile multidisciplinary methodology. It combines observations of the current (and possibly, past) state of a system with results from a mathematical model (the forecast) to produce an analysis, providing "the best" estimate of the current state of the system. Central to the concept of sequential estimation data assimilation is the propagation of flow dependent covariance errors.

In sequential estimation, the analysis and forecasts can be viewed of as probability distributions. The analysis step is an application of the Bayes theorem. Advancing the probability distribution in time, in the general case is done by the Chapman-Kolmogorov equation, but since it is unrealistically expensive, various approximations operating on representations of the probability distributions are used instead. If the probability distributions are normal, they can be represented by their mean and covariance, which gives rise to the Kalman filter (KF). However it is not feasible to store the covariance due to the large number of degrees of freedom in the state, so various approximations based on Monte-Carlo ensemble calculations are used instead.

In the present work we aim comparing standard ensemble Kalman filter (EnKF) to the particle filter (PF) using various resampling strategies. In addition, we compare the performance of each method in the presence of linear and nonlinear observation operators which leads us to draw conclusion regarding superiority of the particle filter vs EnKF.

While work on 4-D Var data assimilation using as model the Kuramoto-Sivashinsky equation has been carried out by very few research workers (see Protas et al. [19]), to the best of our knowledge only the work of Chorin and Krause [26] has used Bayesian filters.

We further discuss issues of computational efficiency of the two methods taking into account the additional effort required by PF due to various resampling strategies.

The structure of the present work is as follows. After the introduction, we present in section 2 the K-S equation and its numerical solution. In section 3 we present the formulation of the PF along with its various sampling strategies. The formulation of the standard EnKF is presented as well. Section 4 is providing discussion of the numerical experiments followed by a discussion of their significance. Finally, section 5 contains summary and conclusions along with directions for further research.

2 The Kuramoto-Sivashinsky Equation

The Kuramoto-Sivashinsky (K-S) equation is an evolution equation in one space dimension, with a Burgers nonlinearity, a fourth order dissipation term and a second anti-dissipative term. It assumes the form

$$u_t + u_{xxxx} + u_{xx} + uu_x = 0,\ (x,t) \in \mathbb{R} \times \mathbb{R}^+ \quad u(x,0) = u_0(x), x \in \mathbb{R} \quad (1)$$

The K-S equation models pattern formations in different physical contexts and is a paradigm of low-dimensional behavior in solutions to partial differential equations. It arises as a model amplitude equation for inter-facial instabilities in many physical contexts. It was originally derived by Kuramoto and Tsuzuki ([5], [6]) to model small thermal diffusive instabilities in laminar flame fronts in two space dimensions. It has also been derived in the context of angular-phase turbulence for a system of reaction-diffusion modeling the Belouzov-Zabotinskii reaction in three space dimensions. Sivashinsky([7],[8])derived it independently to model small thermal diffusive instabilities in laminar flame fronts. The equation also arises in modeling small perturbations from a reference Poiseuille flow of a film layer on an inclined plane [9], while Babchin et al.[10] derived (1) as a general mechanism modeling the nonlinear saturation of instabilities in flow films as in the Rayleigh-Taylor-type instability.

The K-S equation is non-integrable, and no explicit solutions exist. It is characterized by a second-order unstable diffusion term, responsible for an instability at large scales,a fourth-order stabilizing viscosity term, which provides damping at small scales; and a quadratic nonlinear coupling term which stabilizes by transferring energy between large and small scales. This is readily apparent in Fourier space, where one may write(1)with periodic boundary condition as

$$\frac{d\hat{u}_k}{dt} = (k^2 - k^4)\hat{u}_k + \frac{i}{2}\sum_{k'} k'\hat{u}_{k'}\hat{u}_{k-k'} \quad (2)$$

where

$$u(x,t) = \sum_k \hat{u}_k(t)\exp(ikx),\ k = n\frac{2\pi}{L},\ k' = m\frac{2\pi}{L},\ m,n \in \mathbb{Z},\ i = \sqrt{-1},$$

The zero solution is linearly unstable to modes with $|k| < 1$; whose number is proportional to the bifurcation parameter L, are coupled to each other and to damped modes at $|k| > 1$ through the non-linear term. In an effort to characterize, understand and predict the spatially and temporally nontrivial dynamical behavior including chaos of K-S, numerical simulation has been

conducted by Hyman and Nicolaenko [11] and by Kevrekidis et al.[12] to cite but few.

Foias et al. [13], Nicolaenko et al.[14] and Temam [16] established the existence of a unique compact inertial manifold for the K-S equation. They also demonstrated that the K-S equation is strictly equivalent to a low-dimensional dynamical system.That is, all orbits are attracted exponentially to a finite-dimensional, bounded, compact, smooth manifold and the dynamics take a place in this "inertial" manifold. This implies that the transition to chaos of the K-S equation can be analyzed using the tools developed for low-dimensional dynamics systems.

Desertion and Kazantzis [17] used the K-S equation to examine the performance improvement of a class of nonlinear transport processes subject to spatio-temporally varying disturbances through the employment of a comprehensive and systematic actuator activation policy. Lee and Tran [18] obtained a reduced-order system that can accurately describe the dynamics of the K-S equation by employing an approximate inertial manifold and a proper orthogonal decomposition. From this resulting reduced-order system, they designed and synthesized the feedback controller for the K-S equation. Recently, Protas et al.[19] came up with a comprehensive framework for the regularization of adjoint analysis in multiscale PDE systems. They examined the regularization opportunities available in the adjoint analysis and optimization of multiscale, and applied the proposed regularization strategies to the K-S equation

2.1 Mathematical Formulation

We consider the solutions of

$$u_t + u_{xxxx} + u_{xx} + uu_x = 0,\ (x,t) \in \mathbb{R} \times \mathbb{R}^+ \quad u(x,0) = u_0(x), x \in \mathbb{R} \quad (3)$$

which are space periodic of period L, $u(x,t) = u(x+L,t),\ L > 0$.

Let $\Omega \subset \mathbb{R}$, we denote by

$$\mathbb{L}^2(\Omega) = \left\{ u | u : \Omega \longrightarrow \mathbb{R}, u \text{ measurable and } \int_\Omega |u(x)|^2 dx < \infty \right\},$$

the Hilbert space of square integrable function over Ω endowed with the norm

$$\|u\|_2 = \left\{ \int_\Omega |u(x)|^2 dx \right\}^{1/2},$$

and by

$$\mathbb{H}^4(\Omega) = \left\{ u | u : \Omega \longrightarrow \mathbb{R}, u \text{ measurable and } \frac{d^i u}{dx^i} \in \mathbb{L}^2(\Omega), \text{ for } i = 0, \cdots, 4 \right\},$$

the fourth-order Sobolev space with the norm, see Adams [1] and also Mazya [2],

$$\|u\|_{4,2} = \left\{ \|u\|_2^2 + \sum_{i=1}^{4} \left\| \frac{d^i u}{dx^i} \right\|_2^2 \right\}^{1/2}.$$

Finally, we denote by $\mathbb{H}^4_{per}(\Omega)$ the closure of $\mathscr{C}^\infty_{per}(\Omega)$ for the $\mathbb{H}^4$-norm. $\mathscr{C}^\infty_{per}(\Omega)$ is the subset of $\mathscr{C}^\infty(\mathbb{R})$ of Ω-periodic functions.
We set

$$\mathscr{A} = \frac{d^4}{dx^4}, \quad \mathscr{H} = \left\{ u \in \mathbb{L}^2(\frac{-L}{2}, \frac{L}{2}) \right\}$$

and

$$D(\mathscr{A}) = \mathbb{H}^4_{per}(\frac{-L}{2}, \frac{L}{2}) \cap \mathscr{H}, \quad \mathscr{V} = D(\mathscr{A}^{1/2}).$$

$\mathscr{H}$ is a Hilbert space, the dissipative operator $\mathscr{A}$ is a linear self-adjoint unbounded operator in $\mathscr{H}$ with domain $D(\mathscr{A})$ and dense in $\mathscr{H}$. Assuming $\mathscr{A}$ positive closed and that $\mathscr{A}^{-1}$ is compact, $\mathscr{V}$ is a Hilbert space endowed with the norm $|\mathscr{A}^{1/2} \cdot |$. Using Leray's method (see Lions[3], Temam[4]), the K-S equation (1) with initial condition $u_0 \in \mathscr{H}$ has a unique solution defined for all $t > 0$ and such that

$$u \in \mathscr{C}(\mathbb{R}^+; \mathscr{H}) \cap \mathbb{L}^2(0, T; \mathscr{V}) \quad \forall T > 0.$$

Moreover, if $u_0 \in \mathscr{V}$, then

$$u \in \mathscr{C}(\mathbb{R}^+; \mathscr{V}) \cap \mathbb{L}^2(0, T; D(\mathscr{A})) \quad \forall T > 0.$$

Furthermore, it has been proved in Nicoleanko et al. ([14],[15]) that only the odd solutions of (1) are stable for large t. Consequently, the subspace $\mathscr{H}$ is restricted to

$$\mathscr{H} = \left\{ u \in \mathbb{L}^2(\frac{-L}{2}, \frac{L}{2}),\ u \text{ is odd} \right\}$$

2.2 Numerical Solution of the K-S Equation

We consider the one-dimensional PDE with initial data as used in [20] and [21]

$$\begin{cases} u_t = -u_{xxxx} - u_{xx} - uu_x, \quad x \in [0, 32\pi] \\ u(x+L,t) = u(x,t), \quad L = 32\pi,\ \forall t > 0 \\ \qquad u(x,0) = \cos(\frac{x}{16}) \left(1 + \sin(\frac{x}{16})\right) \end{cases} \tag{4}$$

The system (4) is known to be stiff. In fact, the stiffness is due to rapid exponential decay of some modes (the dissipative part), the stiffness is also due to rapid oscillations of some modes (the dispersive part).

As the equation is periodic, a Fourier spectral method is used for spatial discretization. Despite the remarkable success of the spectral and pseudo-spectral methods for a wide range of applications [22] and [23], the set of ODEs for the mode amplitudes is stiff, due to the time scale associated with the nth mode scales as $O(n^{-m})$ for large n, where m is the order of the highest spatial derivative, so that the highest modes evolve on short time scales.

In order to carry out numerical solution of K-S, a modification of the exponential time-differencing fourth-order Runge-Kutta method (ETDRK4) has been used. This method has been proposed by Cox and Matthews [25] and further modified by Kassam and Trefethen [20]. A short review of the ETDRK4 is as follows:

First we transform (4) to Fourier space

$$\hat{u}_t = -\frac{ik}{2}\hat{u^2} + (k^2 - k^4)\hat{u}, \tag{5}$$

set

$$\mathscr{L}\hat{u}(k) = (k^2 - k^4)\hat{u}(k), \qquad \mathscr{N}(\hat{u}, t) = \mathscr{N}(\hat{u}) = -\frac{ik}{2}(\mathcal{F}((\mathcal{F}^{-1}(\hat{u}))^2)), \tag{6}$$

$\mathscr{L}$ and $\mathscr{N}$ stand for linear and nonlinear operators, respectively. $\mathcal{F}$ denotes the discrete Fourier transform. Write (5) in an operational form

$$\hat{u}_t = \mathscr{L}\hat{u} + \mathscr{N}(\hat{u}, t). \tag{7}$$

Define $v = e^{-\mathscr{L}t}u$ where $e^{-\mathscr{L}t}$ the integrating factor to obtain

$$v_t = e^{-\mathscr{L}t}\mathscr{N}(e^{\mathscr{L}t}v). \tag{8}$$

Let h denote the time step length, then integrating (8) we obtain

$$u_{n+1} = e^{\mathscr{L}h}u_n + e^{\mathscr{L}h}\int_0^h e^{-\mathscr{L}\tau}\mathscr{N}(u(t_n+\tau), t_n+\tau)d\tau, \tag{9}$$

where u_n is the solution at the time $t = nh$ and $0 < \tau < h$.
The equation (9) is exact, and the various order EDT schemes differ only on the way one approximates the integral in (9). Cox and Matthews [25] proposed the generating formula

$$u_{n+1} = e^{\mathscr{L}h}u_n + h\sum_{m=0}^{s-1} g_m \sum_{k=0}^{m}(-1)^k \begin{pmatrix} m \\ k \end{pmatrix} \mathscr{N}_{n-k} \tag{10}$$

where s is the order of the scheme. The coefficients g_m are provided by the recurrence relation

$$\begin{cases} \mathscr{L}(hg_0) = e^{\mathscr{L}h} - I, \\ \mathscr{L}(hg_{m+1}) + \mathbf{I} = g_m + \frac{1}{2}g_{m-1} + \frac{1}{3}g_{m-2} + \cdots + \frac{1}{m+1}g_0, \quad m \geq 0. \end{cases} \tag{11}$$

We solve the K-S equation employing 64 Fourier spectral modes and integrate from $t = 0$ to $t = 250$ (nondimensional time units) using the EDTRK4 time stepping. The time evolution for the K-S equation is depicted in (Fig. 1), while the time evolution for the K-S at four different locations is presented

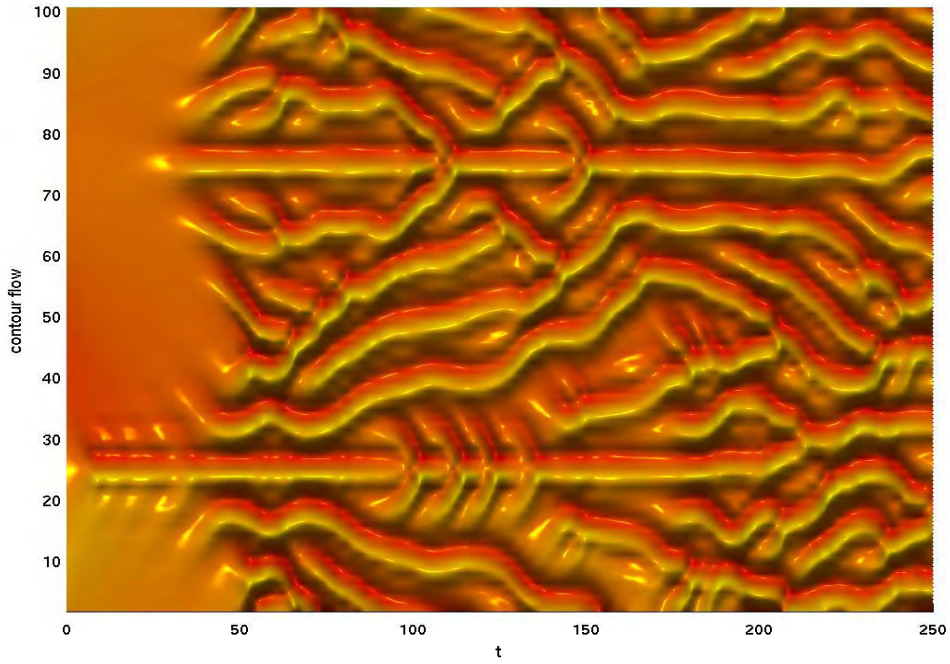

Fig. 1. Contours for the time evolution for the K-S equation

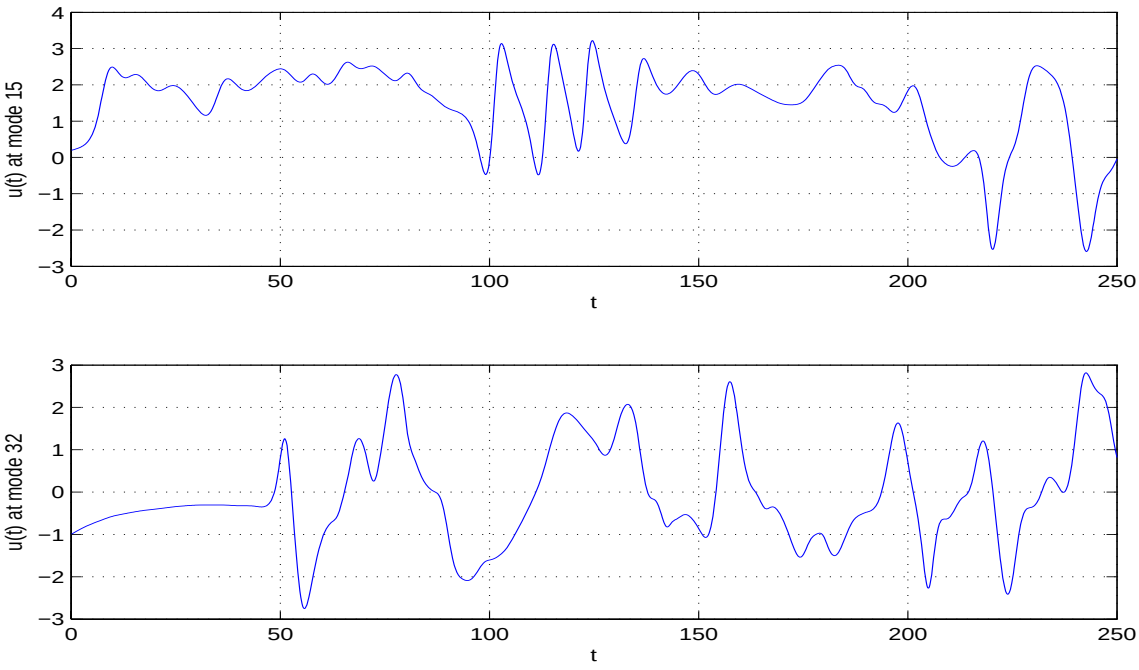

Fig. 2. K-S solution at different spatial positions

3 Data Assimilation for the K-S Equation

Data assimilation is the process by which observational data distributed in space and time are fused with mathematical model forecast information. The probabilistic state space formulation and the requirement for the updating of information when new observations are encountered are ideally suited for the Bayesian approach, and thus constitute an appropriate framework for data assimilation. The Bayesian approach and in particular ensemble or particle filtering methods are a set of efficient and flexible Monte-Carlo methods to solve the optimal filtering problem. Here one attempts to construct the posterior probability density function (PD) of the state based on all available information, including the set of received observations. Since this PD embodies all available statistical information, it may be considered to be a complete solution to the estimation problem.

In the field of data assimilation, there are only few contributions in sequential estimation (EnKF or PF filters) using the K-S equation. Chorin and Kruse [26] used particle filters and proposed a strategy for reducing the size of the system of equations to be solved for evolving the particles by adaptively finding subsets of variables that do not need to be recomputed and solving only in directions in which the dynamic system is expanding. In the work of Hu and Temam [27], a robust boundary control method for the K-S equation has been proposed, and a data assimilation problem corresponding to the K-S equation has been considered.

3.1 Sequential Bayesian Filter

The sequential Bayesian filter employs a large number N of random samples or "particles" advanced in time by a stochastic evolution equation, to approximate the probability densities. In order to analyze and make inference about the dynamic system at least a model equation along with an observation operator are required. First, a model describing the evolution of the state with time, and an observation operator for noisy observations of the state. Generically, stochastic filtering problem is a dynamic system that assumes the form

$$\dot{\mathbf{x}}_t = f(t, \mathbf{x}_t, u_t, \mathbf{v}_t) \tag{12}$$

$$\mathbf{z}_t = h(t, \mathbf{x}_t, u_t, \mathbf{n}_t) \tag{13}$$

The equation (12) is state equation or the system model, (13) is the observation operator equation, $\mathbf{x}_t$ is the state vector, $\mathbf{z}_t$ the observation vector and $\mathbf{u}_t$ is the system input vector serving as the driving force. $\mathbf{v}_t$ and $\mathbf{n}_t$ are the state and

observation noises, respectively. In practical application, however, we are more concerned about the discrete-time filtering, and we consider the evolution of the state sequence $\{\mathbf{x}_k, k \in \mathbb{N}\}$, given by

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}), \tag{14}$$

where the deterministic mapping $f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_d} \longrightarrow \mathbb{R}^{n_x}$ is a possibly non-linear function of the state $\mathbf{x}_{k-1}$, $\{\mathbf{v}_{k-1}, k \in \mathbb{N}\}$ is an independent identically distributed (i.i.d) process noise sequence, n_x, n_d are dimensions of the state and process noise vectors, respectively, and $\mathbb{N}$ is the set of the natural numbers. The objective is to recursively estimate $\mathbf{x}_k$ from observations

$$\mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{n}_k), \tag{15}$$

where $h_k : \mathbb{R}^{n_\mathbf{x}} \times \mathbb{R}^{n_n} \longrightarrow \mathbb{R}^{n_z}$ is a possibly non-linear function, $\{\mathbf{n}_k, k \in \mathbb{N}\}$ is an i.i.d. observation noise sequence, and n_x, n_n are dimensions of the state and observation noise vectors, respectively.

We denote by $\mathbf{z}_{1:k}$ the set of all available observations $\mathbf{z}_i$ up to time $t = k$, $\mathbf{z}_{1:k} = \{\mathbf{z}_i | i = 1, \cdots, k\}$. From a Bayesian point of view, the problem is to recursively calculate some degree of belief in the state $\mathbf{x}_k$ at time $t = k$, taking different values, given the data $\mathbf{z}_{1:k}$ up to the time $t = k$.Then the Bayesian solution would be to calculate the PDF $p(\mathbf{x}_k|\mathbf{z}_{1:k})$. This density will encapsulate all the information about the state vector $\mathbf{x}_k$ that is contained in the observations $\mathbf{z}_{1:k}$ and the prior distribution for $\mathbf{x}_k$.
Suppose that the required PDF $p(\mathbf{x}|\mathbf{z}_{1:k-1})$ at time $k-1$ is available. The prediction stage uses the state equation (14) to obtain the prior PDF of the state at time k via the Chapman-Kolmogorov equation

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}. \tag{16}$$

The probabilistic model of the state evolution, $p(\mathbf{x}_k|\mathbf{x}_{k-1})$, is defined by the state equation (14) and the known statistics of $\mathbf{v}_{k-1}$.

At time $t = k$, a measurement $\mathbf{z}_k$ becomes available, and it may be used to update the prior via the Bayes rule

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{1:k-1})}, \tag{17}$$

where the normalizing constant

$$p(\mathbf{z}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) d\mathbf{x}_k. \tag{18}$$

depends on the likelihood function $p(\mathbf{z}_k|\mathbf{x}_k)$, defined by the measurement equation (15) and the known statistics of $\mathbf{n}_k$.

The relations (16) and (17) form the basis for the optimal Bayesian solution. This recursive propagation of the posterior density is only a conceptual solution. It cannot be determined analytically. Solutions exist only in a very restrictive set of cases like that of the Kalman filters for instance. (namely, if f_k and h_k are linear and both v_k and n_k are Gaussian)

3.2 Particle Filters

Particle filters (see [30], [31], [32] and [28]) approximate the posterior densities by population of states. These states are called "particles". Each of the particles has an assigned weight, and the posterior distribution can then be approximated by a discrete distribution which has support on each of the particles. The probability assigned to each particle is proportional to its weight. The different (PF) algorithms differ in the way that the population of particles evolves and assimilates the incoming observations.

We use the Sampling Importance Resampling (SIR) or the Bayesian bootstrap filter of Gordon et al. [33] see also Berliner and Wikel([28],[29]). The SIR algorithm generates a population of equally weighted particles to approximate the posterior at some time k. This population of particles is assumed to be an approximate sample from the true posterior at that time instant.

The PF algorithm proceeds as follows:

- **Initialization:** The filter is initialized by drawing a sample of size N from the prior at the initial time. The algorithm is then started with the filtering step.
- **Preliminaries:** Assume that $\{\mathbf{x}_{k-1}^i\}_{i=1,\cdots,N}$ is a population of N particles, approximately distributed as in an independent sample from $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$
- **Prediction** Sample N values, $\{w_k^1, \cdots, w_k^N\}$, from the distribution of $\mathbf{v}_k$. Use these to generate a new population of particles, $\{\mathbf{x}_{k|k-1}^1, \mathbf{x}_{k|k-1}^2, \cdots, \mathbf{x}_{k|k-1}^N\}$ via the equation

$$\mathbf{x}_{k|k-1}^i = f_k(\mathbf{x}_{k-1}^i, \mathbf{v}_k^i) \tag{19}$$

- **Filtering:** Assign each $\mathbf{x}_{k|k-1}^i$, a weight q_k^i. This weight is calculated by

$$q_k^i = \frac{p(\mathbf{z}_k|\mathbf{x}_{k|k-1}^i)}{\sum_{j=1}^M p(\mathbf{z}_k|\mathbf{x}_{k|k-1}^i)} \tag{20}$$

This defines a discrete distribution which, for $i \in \{1, 2, \cdots, N\}$, assigns

probability mass q_k^i to element $\mathbf{x}_{k|k-1}^i$

- **Resampling:** Resample independently N times, with replacement, from the distribution obtained in the filtering stage. The resulting particles, $\{\mathbf{x}_k^i\}_{i=1,\cdots,N}$, form an approximate sample from $p(\mathbf{x}_k|\mathbf{z}_{1:k})$.

The method outlined above can be justified as follows. If the particles at time $t = k-1$ were an i.i.d sample from the posterior at time $t = k-1$, then the predictive stage just produces an i.i.d. sample from the prior at time $t = k$. The filtering stage can be viewed as an importance sampling approach to generate an empirical distribution which approximates the posterior.

The proposal density is just the prior $p(\mathbf{x}_k|\mathbf{z}_{1:k-1})$, and as a result of Bayes formula, we obtain

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}, \mathbf{z}_k) \propto p(\mathbf{x}_k|\mathbf{z}_{1:k-1})p(\mathbf{z}_k|\mathbf{x}_k), \tag{21}$$

the weights are proportional to the likelihood $p(\mathbf{z}_k|\mathbf{x}_k)$. As N tends to infinity, the discrete distribution which has probability mass q_i at point $\mathbf{x}_{k|k-1}^i$, converges weakly to the true posterior. The resampling step is a crucial and computationally expensive part in a particle filter. It is used to generate equally weighted particles aimed at avoiding the problem of degeneracy of the algorithm, that is, avoiding the situation that all but one of the weights are close to zero. The performance of the algorithm is affected by the proper choice of the resampling method. Generically, it is implemented as follows:

- Draw N particles $\{\tilde{\mathbf{x}}_k^i\}_{i=1,\cdots,N}$ from the uniform distribution.
- Assign the resampled particles $\{\tilde{\mathbf{x}}_k^i\}_{i=1,\cdots,N}$ to $\{\mathbf{x}_k^i\}_{i=1,\cdots,N}$ and assign equal weights $\frac{1}{N}$ to each particle.

In addition to the systematic resampling proposed by Kitagawa [34] and also by Liu and Chen [35] which is preferred by many authors ([32],[31]), both in terms of resampling quality and computational complexity, we also implemented a stratified resampling to emphasize the importance of resampling for the particle filter.

3.3 The Ensemble Kalman Filter

The ensemble Kalman filter (EnKF) was first proposed by Evensen [37] and further developed by Burgers et al. [38] and Evensen ([39],[40]). It is related to particle filters in the context that a particle is identical to an ensemble member. EnKF is a sequential filter method, which means that the model is integrated forward in time and, whenever observations are available, these are used to reinitialize the model before the integration continues. The EnKF originated as

a version of the Extended Kalman Filter (EKF) ([43],[44]) for large problems. The classical KF [36] method is optimal in the sense of minimizing the variance only for linear systems and Gaussian statistics. Similar to the particle filter method, the EnKF stems from a Monte Carlo integration of the Fokker-Planck equation governing the evolution of the PDF that describes the prior, forecast, and error statistics. In the analysis step, each ensemble member is updated according to the KF scheme and replaces the covariance matrix by the sample covariance computed from the ensemble. However, the EnKF presents two potential problems namely:

1) Even though the EnKF uses full non-linear dynamics to propagate the forecast error statistics, the EnKF assumes that all probability distributions involved are Gaussian.
2) The updated ensemble preserves only the first two moments of the posterior. Let $p(\mathbf{x})$ denote the Gaussian prior probability density distribution of the state vector $\mathbf{x}$ with mean μ and covariance $\mathcal{Q}$

$$p(\mathbf{x}) \propto exp\left(\frac{-1}{2}(\mathbf{x}-\mu)^T\mathcal{Q}^{-1}(\mathbf{x}-\mu)\right)$$

We assume the data $\mathbf{z}$ to have a Gaussian PDF with covariance $\mathcal{R}$ and mean $\mathcal{H}\mathbf{x}$, where $\mathcal{H}$ is the so-called the observation matrix, is related to h of equation (13), and where the value $\mathcal{H}\mathbf{x}$ assumes the value of the data $\mathbf{z}$ would be for the state $\mathbf{x}$ in absence of observation errors. Then the conditional probability or likelihood $p(\mathbf{z}|\mathbf{x})$ assumes the form

$$p(\mathbf{z}|\mathbf{x}) \propto exp\left(\frac{-1}{2}(\mathbf{z}-\mathcal{H}\mathbf{x})^T\mathcal{R}^{-1}(\mathbf{z}-\mathcal{H}\mathbf{x})\right).$$

According to the Bayes theorem the posterior probability density follows from the relation

$$p(\mathbf{x}|\mathbf{z}) \propto p(\mathbf{z}|\mathbf{x})p(\mathbf{x}). \tag{22}$$

The basic formulation of the EnKF ([37], [38], [42] [41], [48]) may be divided into three steps, as follows:

- **Setting and matching**

  ■ Define the ensemble

  $$\mathcal{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N] \tag{23}$$

  be an $n_x \times N$ matrix whose columns are a sample from the prior distribution. N being the number of the ensemble members.

■ Form the ensemble mean

$$\bar{\mathcal{X}} = \mathcal{X} \cdot \mathbf{1}_N, \tag{24}$$

where $\mathbf{1}_N \in \mathbb{R}^{N\times N}$ is the matrix where each element is equal to 1.

■ Define the ensemble perturbation matrix $\mathcal{X}'$ and set the $\mathbb{R}^{n_x\times n_x}$ ensemble covariance matrix $\mathcal{C}$

$$\mathcal{X}' = \mathcal{X} - \frac{1}{N}\bar{\mathcal{X}}, \tag{25}$$

$$\mathcal{C} = \frac{\mathcal{X}'\mathcal{X}'^T}{N-1}, \tag{26}$$

- **Resampling**

■ Generate

$$\mathcal{Z} = [\mathbf{z}_1, \cdots, \mathbf{z}_N] \tag{27}$$

be an $n_z \times N$ matrix whose columns are a replicate of the measurement vector $\mathbf{z}$ plus a random vector from the normal distribution $\mathcal{N}(0, \mathcal{R})$.

■ Form the $\mathbb{R}^{n_z\times n_z}$ measurement error covariance

$$\mathcal{R} = \frac{\mathcal{Z}\mathcal{Z}^t}{N-1}, \tag{28}$$

- **Updating** Obtain the posterior $\mathcal{X}^p$ by the linear combinations of members of the prior ensemble

$$\mathcal{X}^p = \mathcal{X} + \mathcal{C}\mathcal{H}^T(\mathcal{H}\mathcal{C}\mathcal{H}^T + \mathcal{R})^{-1}(\mathcal{Z} - \mathcal{H}\mathcal{X}) \tag{29}$$

The matrix

$$\mathcal{K} = \mathcal{C}\mathcal{H}^T(\mathcal{H}\mathcal{C}\mathcal{H}^T + \mathcal{R})^{-1} \tag{30}$$

is the Kalman gain matrix. Since $\mathcal{R}$ is always positive definite(i.e. covariance matrix), the inverse $(\mathcal{H}\mathcal{C}\mathcal{H}^T + \mathcal{R})^{-1}$ exists. An easy computation shows that the mean and covariance of the posterior or updated ensemble are given by

$$\bar{\mathcal{X}^p} = \mathcal{X}^p + \mathcal{K}\left[\mathbf{z} - (\mathcal{H}\mathcal{X}^p + \mathbf{d})\right], \tag{31}$$

and

$$\mathcal{C}^p = \mathcal{C} - \mathcal{K}\left[\mathcal{H}\mathcal{C}\mathcal{H}^T + \mathcal{R}\right]\mathcal{K}^T, \tag{32}$$

the vector $\mathbf{d}$ which appears in (31) stems from the affine measurement relation

$$h(\mathbf{x}) = \mathcal{H}\mathbf{x} + \mathbf{d}. \tag{33}$$

In the case of nonlinear observation operators, a modification to the above algorithm is advised. As presented in Evensen [39], let $\hat{\mathbf{x}}$ the augmented state vector made of the state vector and the predicted observation vector (nonlinear in this case).

$$\hat{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ \mathcal{H}(\mathbf{x}) \end{pmatrix}. \tag{34}$$

Define the linear observation operator $\hat{\mathcal{H}}$ by

$$\hat{H} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{y} \tag{35}$$

and carry out the steps of the EnKF formulation in augmented state space $\hat{\mathbf{x}}$ and $\hat{\mathcal{H}}$ instead of $\mathbf{x}$ and $\mathcal{H}$. Superficially, this technique appears to reduce the nonlinear problem to the previous linear observation operator case. However, whilst the augmented problem, involving linear observation problem, is a reasonable way of formulating the EnKF, it is not as well-founded as the linear case, which can be justified as an approximation to the exact and optimal KF.

4 Numerical Experiments

4.1 Filtering with a Linear Operator

In order to compare PF with various sampling strategies to the standard EnKF the following numerical experiments were carried out. We will start with description of the numerical experiments with a linear observation operator.

For the PF, 2000 particles were employed for each of the experiments. The observations were provided at a frequency consisting of one observation every 300 time steps. Each run was carried out for 250 non-dimensional time units with a time step of $\Delta t = 1/10$. The number of Fourier modes in (2) was set to 64 modes. The observation error standard deviation was taken to be $\sqrt{2}$.

In Fig.(3) we present the contour flow of the mean forecast solution of the K-S equation. On the right hand side of the same figure, we have the equivalent mean PF filtered solution and as one can observe, there is a strong similarity between the two.

Fig.(4) illustrates two different (modes) locations for the above mentioned PF experiment. The trends show an almost perfect matching between the forecast and the PF filter results and this fact is further confirmed by the small value of the variances between forecast and PF as depicted in Fig.(5). We also present in Fig.(6) the root-mean-square (rms) error. The rms error is calculated according to the relation

$$rms(x,t) = \sqrt{\frac{1}{Ns}\sum_{i=1}^{Ns}(u_i(x,t) - u_{true}(x,t))^2}, \tag{36}$$

it is assumed in the rms equation 36 that the true analysis solution denoted by $u_{true}(x,t)$, is given by the forecast solution used to produce the observations. This is assumed acceptable as pointed out by Gleeson [46] and also by Anderson and Anderson [45].

An identical setting of the time step and the observation operator has been employed for the standard EnKF. The number of ensemble members employed in EnKF has been set to 40. The observations are available every 70 model time steps. In view of further experimentation the aim to assess the impact of number of ensemble members

on EnKF, performance with a double number of ensemble members, i.e. 80 which was found to yield almost identical result. Hence we decided that using EnKF with 40 ensemble members was sufficient for the purpose of this

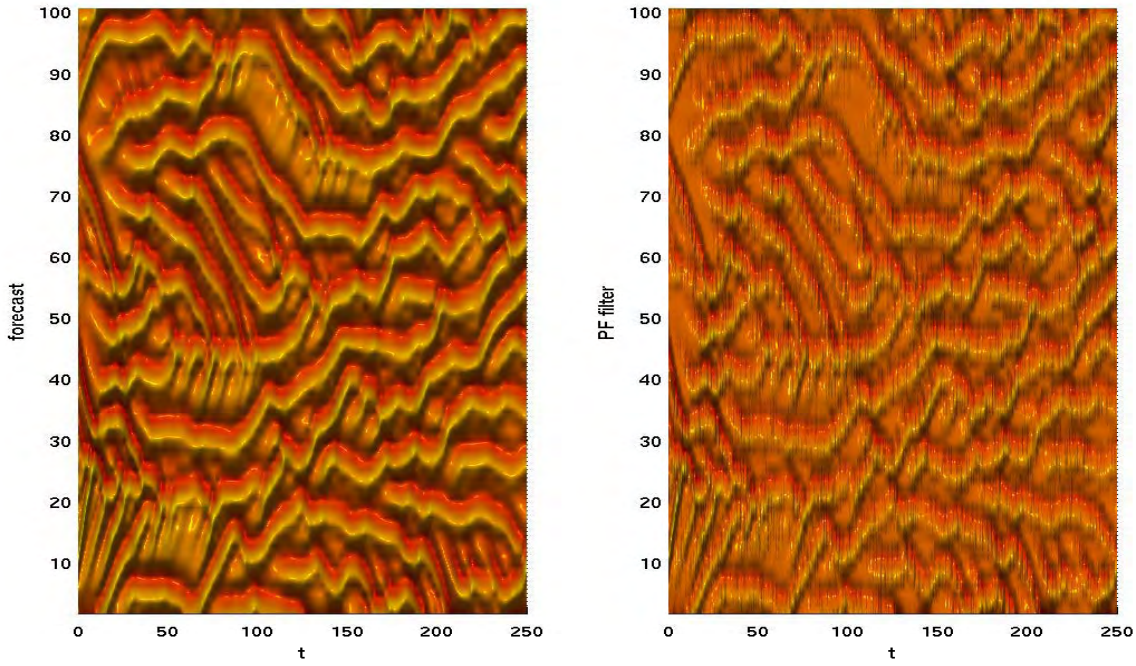

Fig. 3. Contours of the forecast & PF filter K-S solution

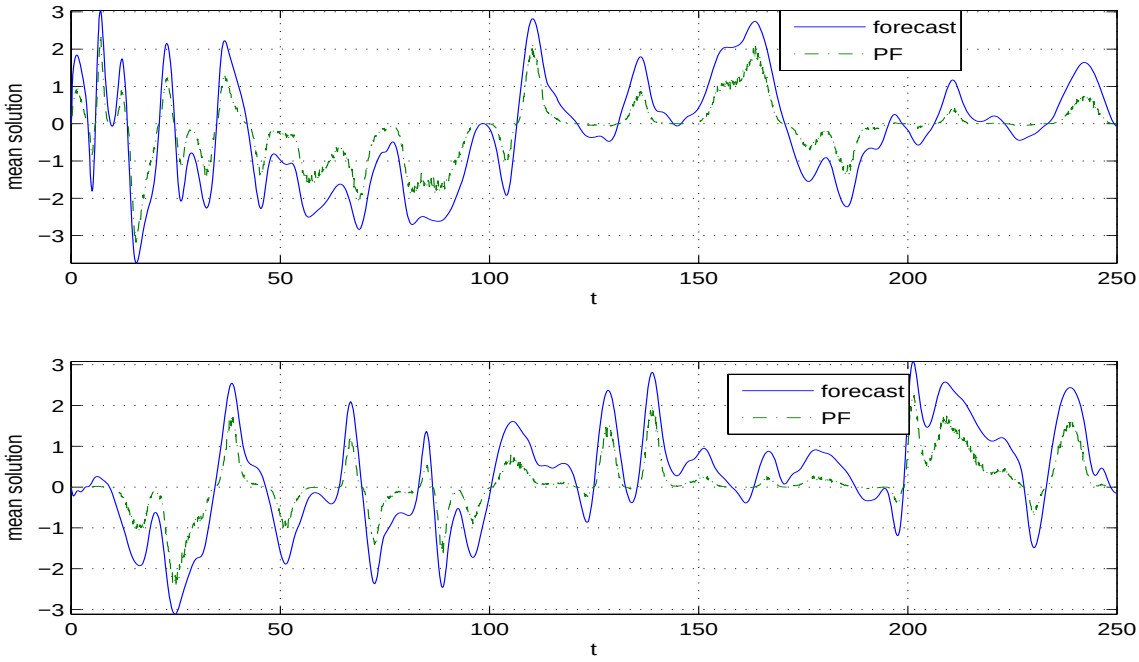

Fig. 4. Forecast and PF filter mean K-S solution at two different locations

experiment.

Presented in Fig.(7) is the contour flow of the mean forecast solution of the K-S equation. On the right hand side of the same figure, we have the equivalent mean EnKF solution. As in the PF case, a strong similarity between the two

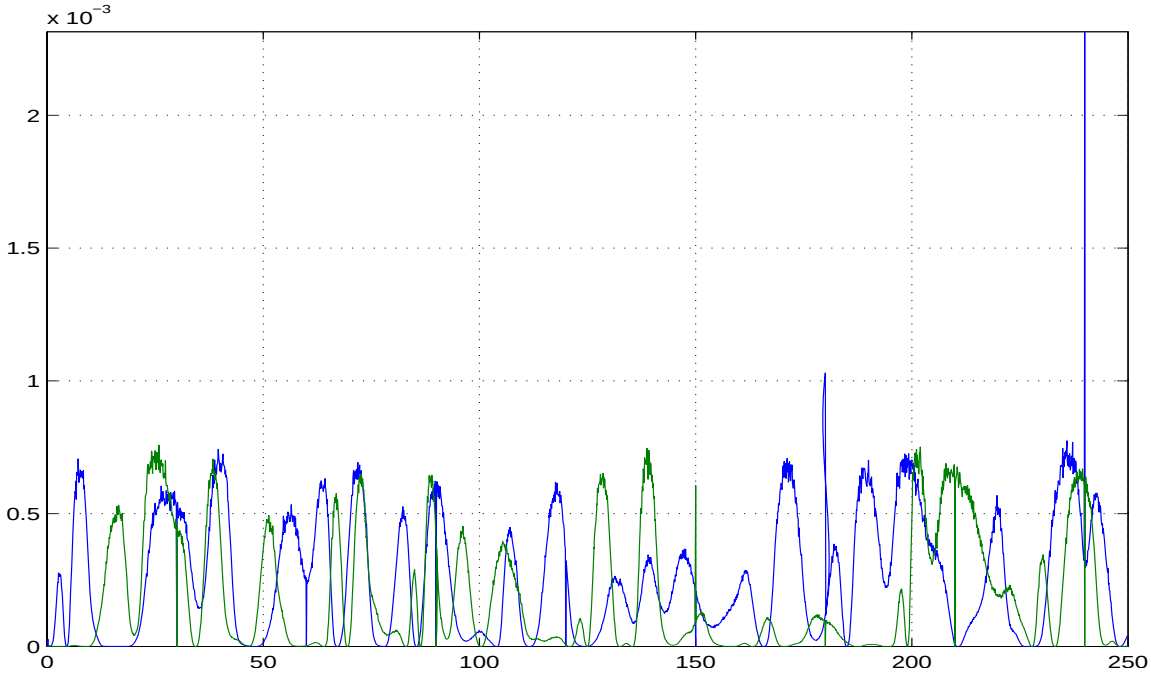

Fig. 5. Variances between the forecast and filtered mean K-S solution at two different locations

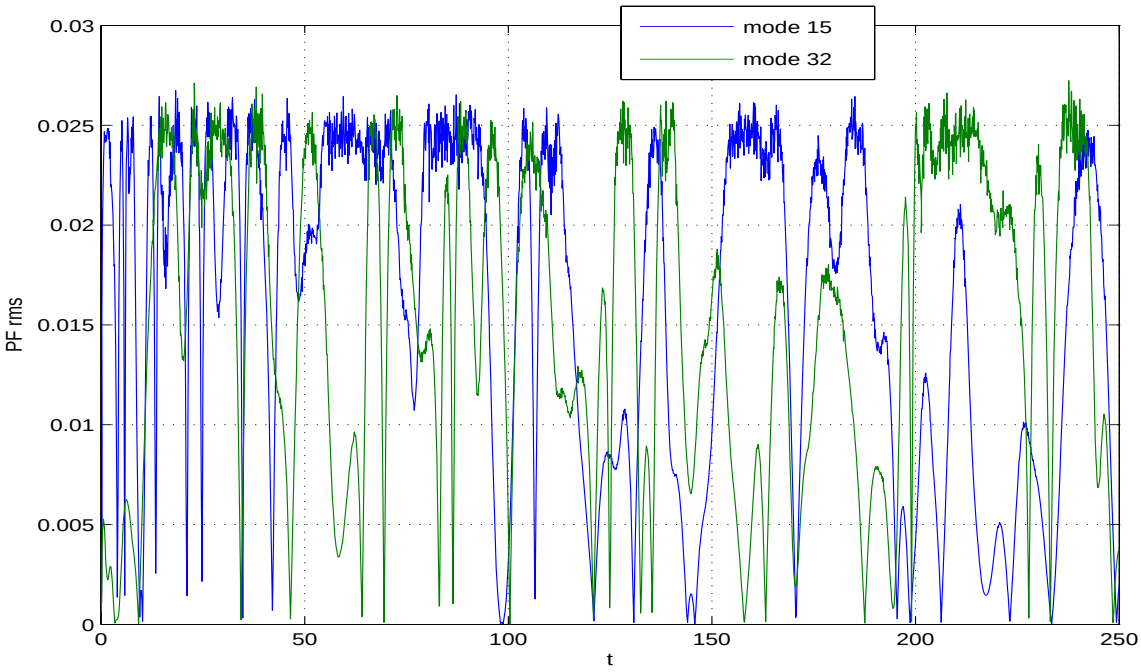

Fig. 6. rms at two different locations for the PF filter

was observed. As for the PF experiment, we added Fig.(8) to emphasize the perfect matching between the forecast and the EnKF filtered solution at two different locations. In Fig.(9) we combine the forecast, PF, and EnKF mean solutions results on the same plot. These results point to the conclusion that

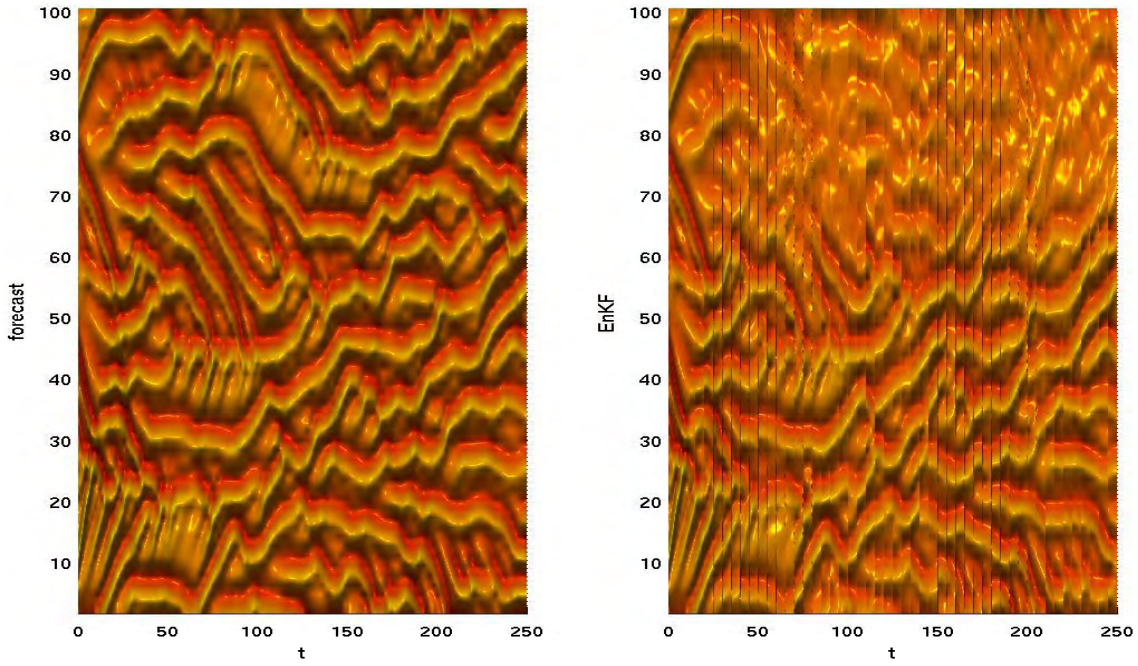

Fig. 7. Contours of the forecast & EnKF filter K-S solution

for the case of linear observation operator the EnKF and PF results for the K-S equation are in good agreement.

To conclude this subsection, it is of importance to emphasize that the two resampling procedures used for the PF method, namely the systematic and stratified resampling performed equally well in quality and computational complexity. Similar conclusion has been drawn by Hol and al. [47]. In future work, we shall further investigate the use of the merging particle filter (MPF) see Nakano et al. [50], as a way to reduce the computational cost even more.

4.2 Filtering with Nonlinear Observation Operator

A similar setting to the linear observation case in terms of number of ensemble member and particles has been used to carry out numerical experiments. The EnKF version tailored to nonlinear observation operator as described in section 3 has been implemented and $\mathcal{H}(x) = x^2$ has been considered. From Fig. (11), which represents the contours of the forecast and EnKF K-S solution, one can easily see that EnKF does not provide good estimates in the case of nonlinear observation operator applied to the K-S model. The discrepancy between the forecast and EnKF is better seen in Fig.(12).

Similarly, we considered the nonlinear observation operator $\mathcal{H}(x) = x^2$ and applied the PF to the K-S model. As expected, the PF method performed

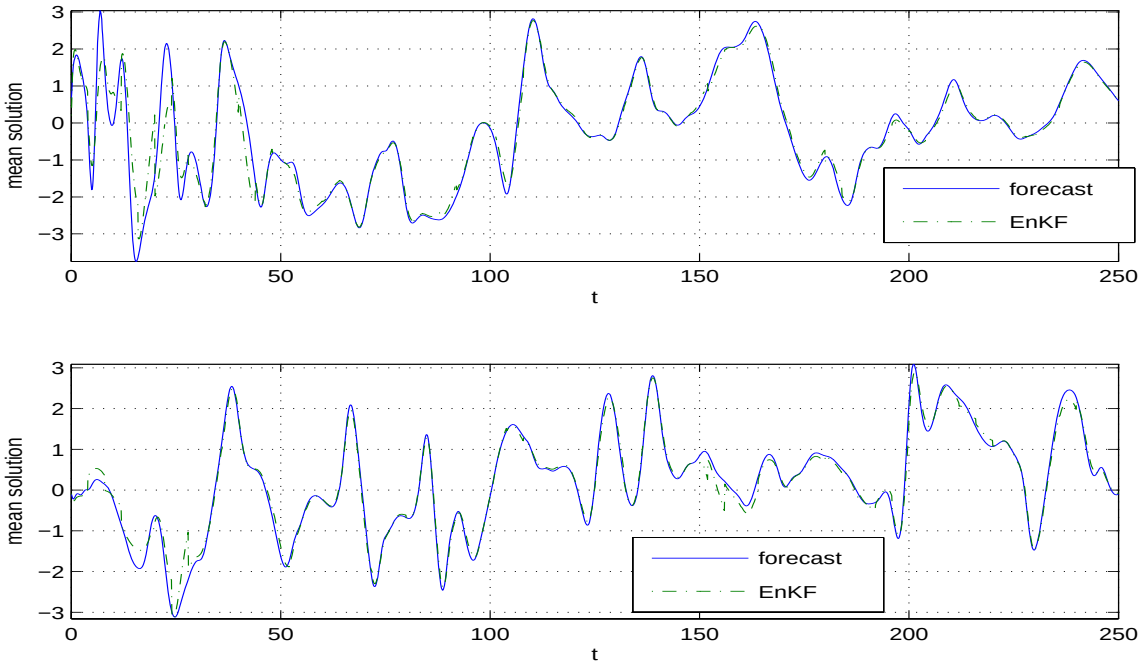

Fig. 8. Forecast and EnKF filter mean K-S solution at two different locations

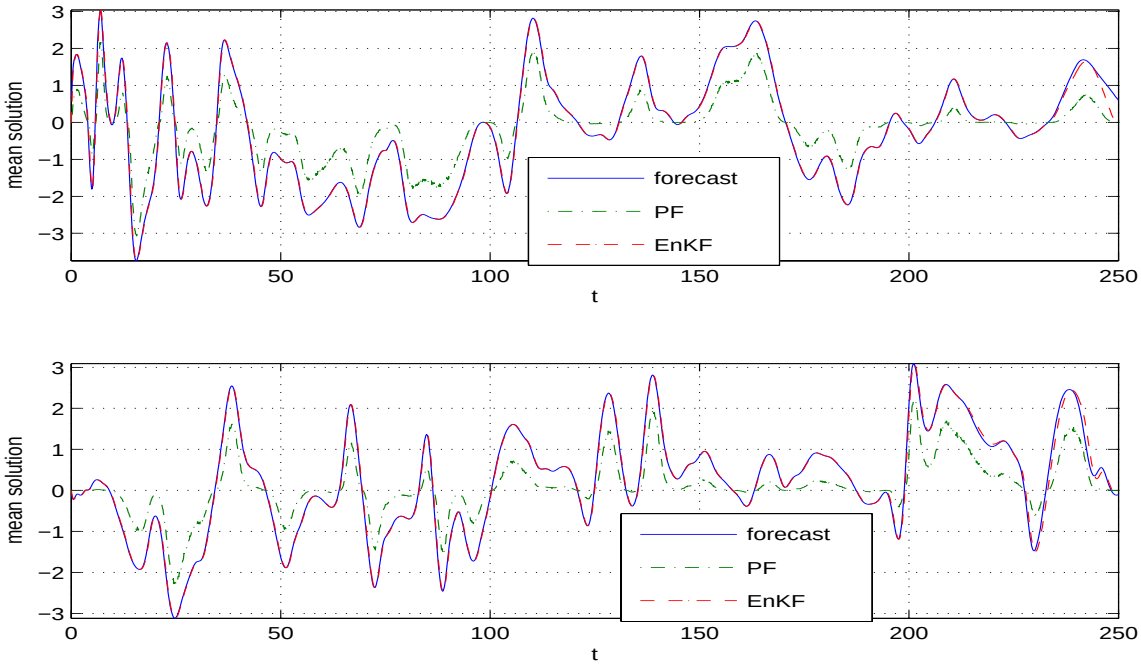

Fig. 9. Means of the forecast, EnKF and PF filtered K-S solution at different spatial locations

well as shown in Fig.(13) and Fig.(12). We went further in experimenting with the observation operator and considered the highly nonlinear observation operator $\mathcal{H}(x) = \exp x$, here again PF did provide a good estimate. Fig.(13) shows good agreement between the contours of the forecast and the PF filtered

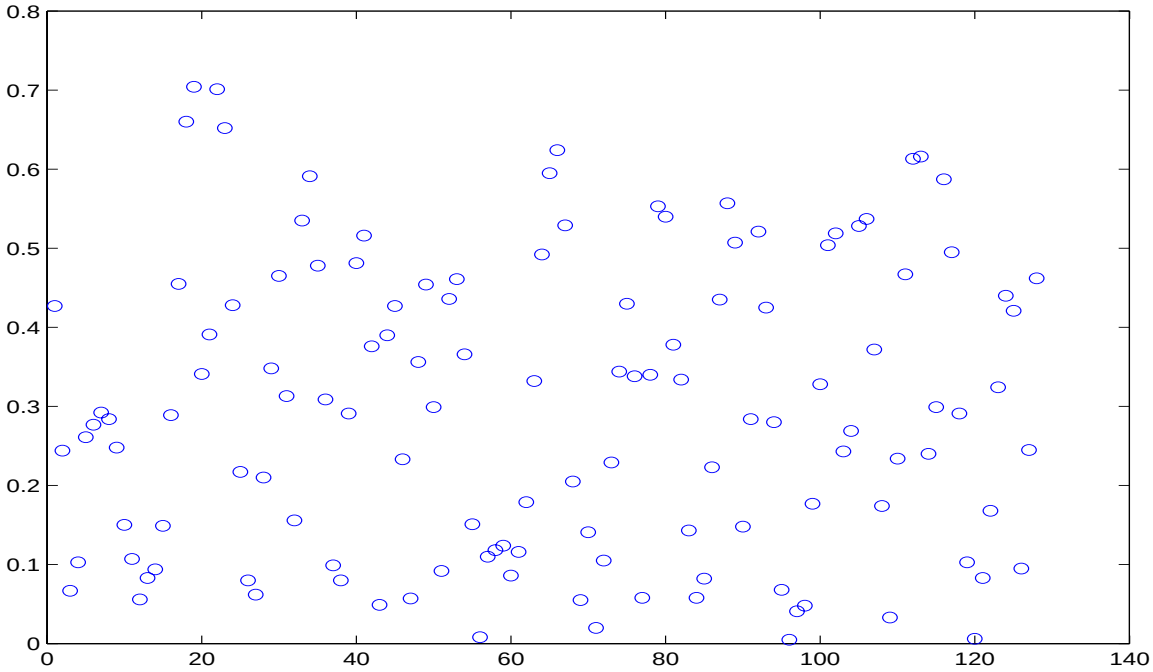

Fig. 10. Importance sampling weights

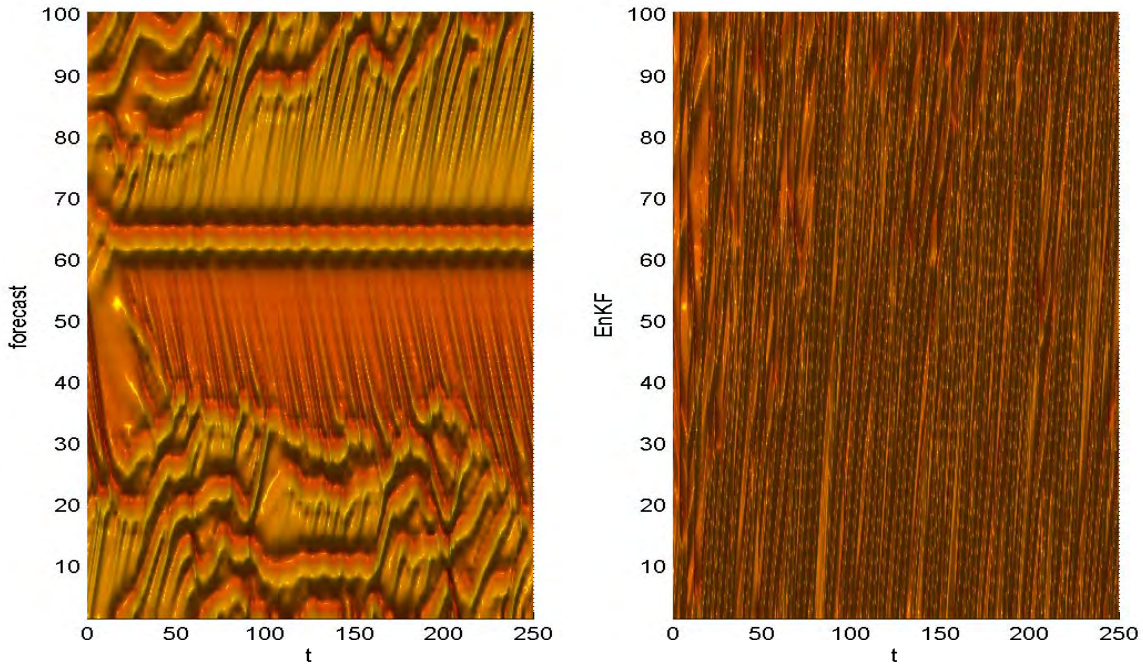

Fig. 11. Contours of K-S solution using EnKF filter in presence of nonlinear observation operator, the square observation operator

K-S solution. Fig.(16) which depicts the forecast and filtered means of the K-S solution using SIR particle filter at two different spatial positions reinforces the conclusion that the PF method is better suited to handle nonlinear observation operator in the case of the K-S model.

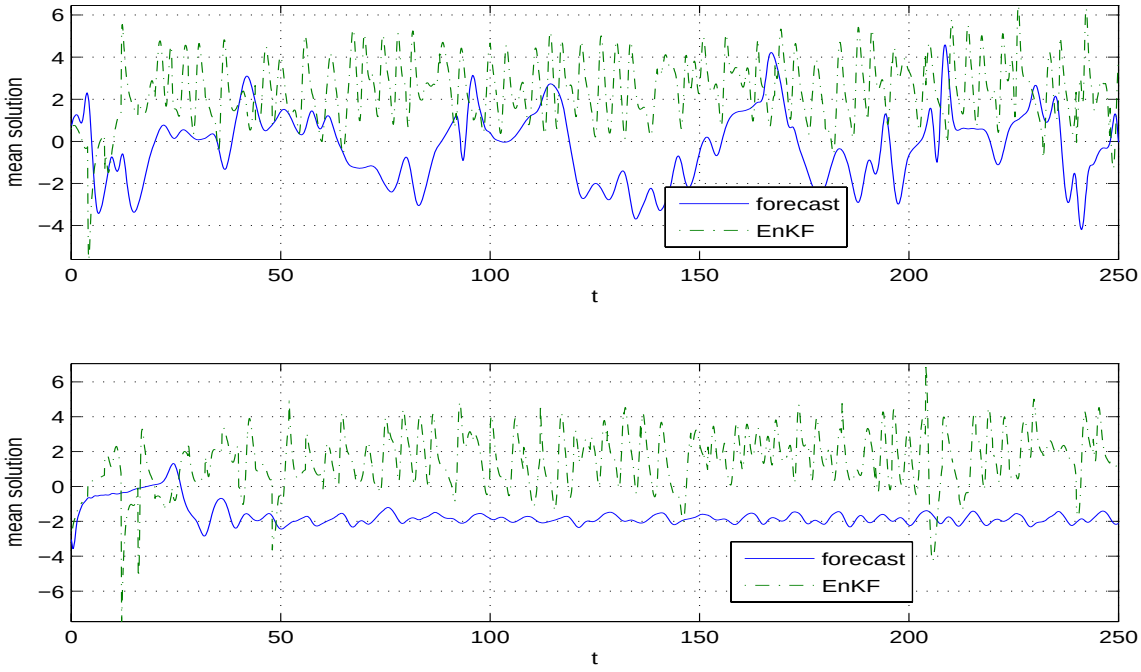

Fig. 12. Means of the forecast and the EnKF filtered K-S solution in the case of a nonlinear observation operator

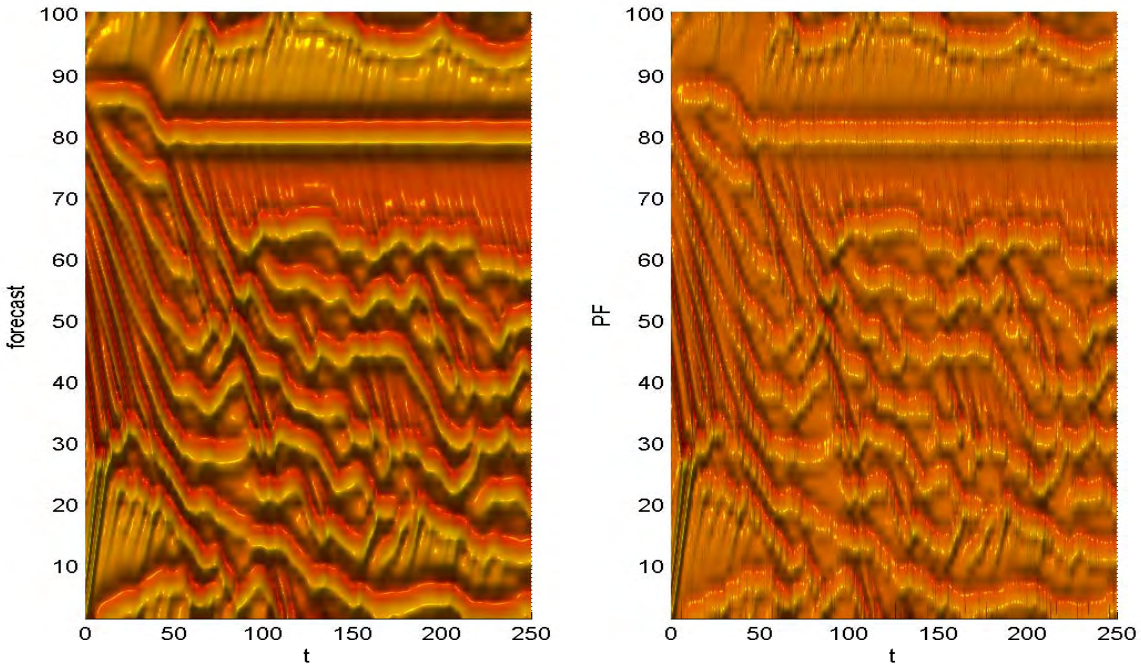

Fig. 13. Contours of the K-S solution using SIR particle filter for the square observation operator

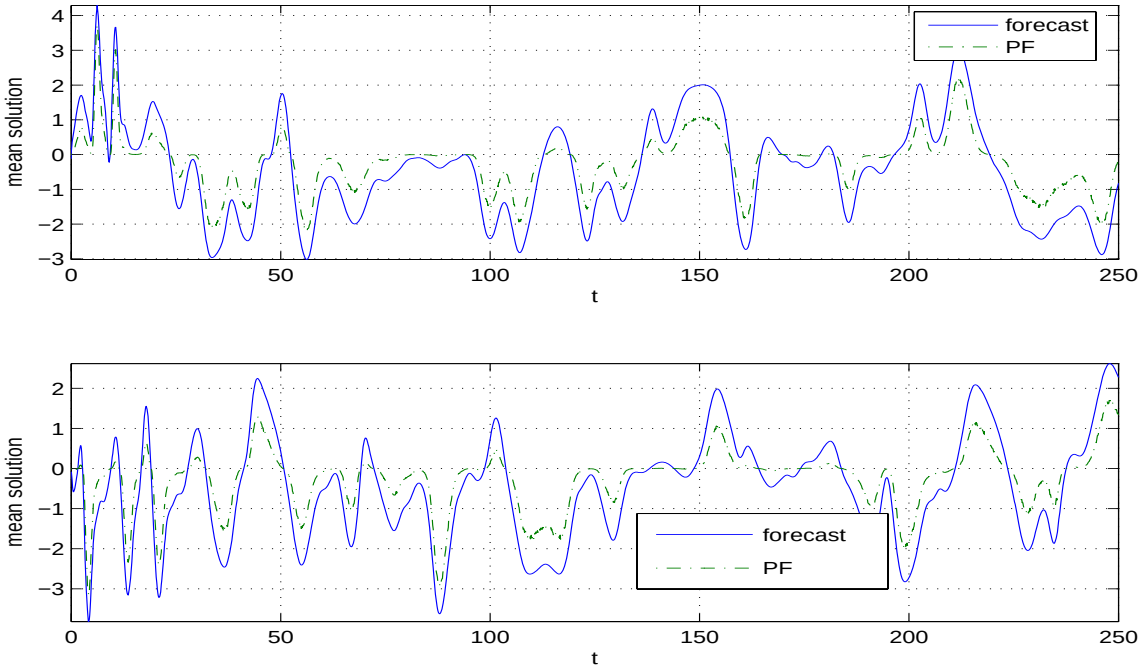

Fig. 14. Means of the K-S solution using SIR particle filter in the presence of non-linear observation operator

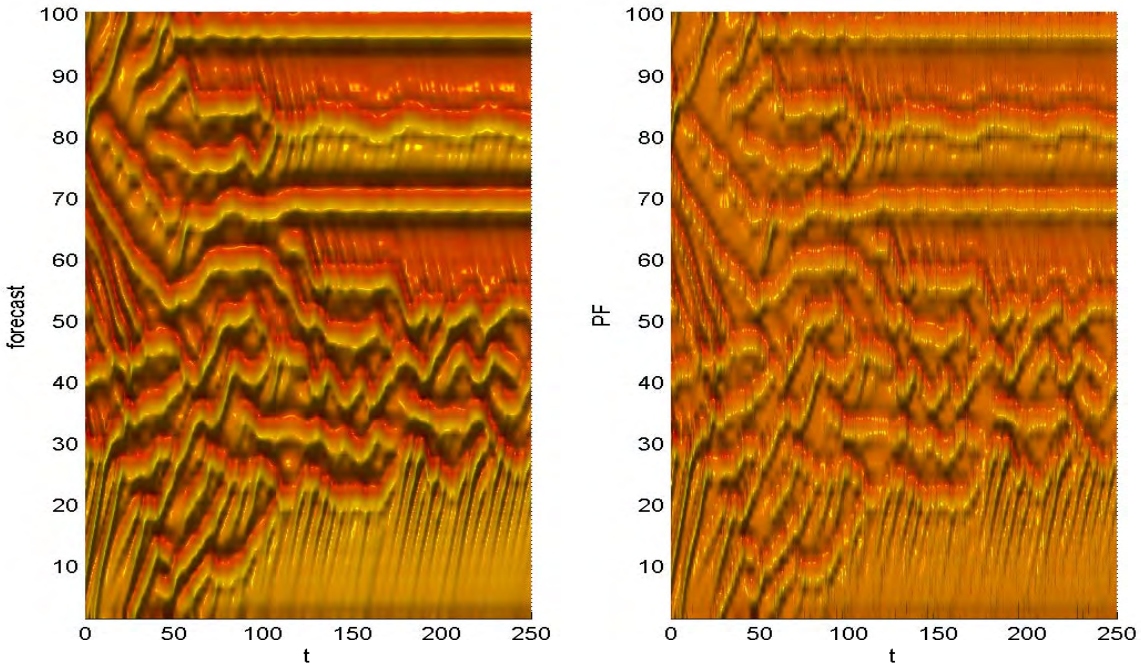

Fig. 15. Contours of the K-S solution using SIR particle filter for the exponential observation operator

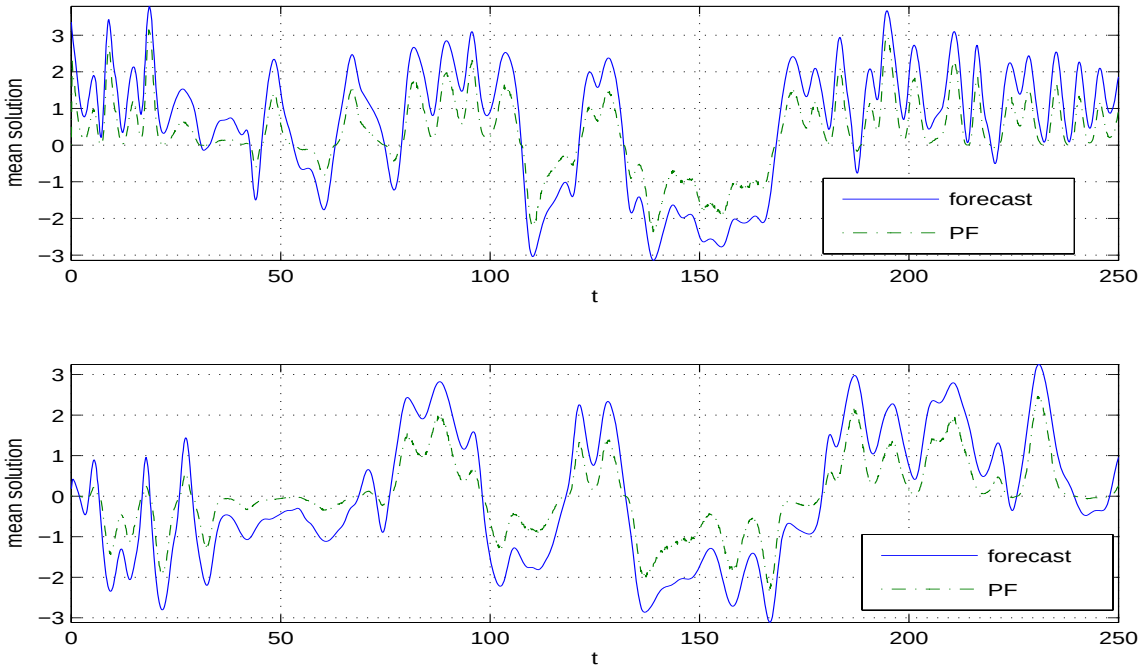

Fig. 16. Means of the K-S solution using SIR particle filter in the presence of non-linear observation operator

5 Summary & conclusions

We have presented a data assimilation study for the Kuramoto-Sivashinsky equation. We have used the EnKF and PF methods as a sequential-based methods for Bayesian filtering. For each of the methods, linear and nonlinear observation operators have been employed and numerical experiments have been carried out. For the PF filtering, two resampling technics served to avoid the degeneration problem, which is inevitable in the particle filter (PF).

It was clear a priori that the EnKF is subject to two major drawbacks. One of them is common to all Kalman filtering schemes applied to non-Gaussian distributions. The other drawback is when a nonlinear relationship exists between a state and observed data. Here the EnKF is simply not effectual. The experiments presented here reveal the poor performance of the EnKF filtering in the case of the nonlinear observation operator. Even with the solution proposed by Evensen [39], the SIR particle filter presents a clear superiority. This issue is of high importance for atmospheric and oceanic data assimilation where the observation operators are mostly nonlinear.

6 Acknowledgments

The research of Prof. I.M. Navon and Dr. M. Jardak is supported by the National Science Foundation (NSF), grant ATM-03727818. The authors also acknowledge the support by NASA Modeling, Analysis, and Prediction Program under Award NNG06GC67G.

References

[1] Adams, R.A, *Sobolev Spaces*, Academic Press, 1975.

[2] Mazya, V.G, *Sobolev Spaces*, Springer-Verlag, 1985.

[3] Lions, J.L., *Quelques méthodes de résolution de problèmes aux limites non linéaires*, Dunod, Paris, (1969)

[4] Temam, R., *Navier-Stokes Equations*, Third edition, North-Holland, Amsterdam, New York (1984)

[5] Kuramoto, Y. and Tsuzuki, T., *On the formation of dissipative structures in reaction-diffusion systems*, Prog. Theor. Phys. **54** (1975), 687–699

[6] Kuramoto, Y. and Tsuzuki, T., *Persistent propagation of concentration waves in dissipative media far from thermal equilibrium*, Prog. Theor. Phys. **55** (1976), 356–369.

[7] Sivashinsky, G.I., *Nonlinear analysis of hydrodynamic instability in laminar flames I. Derivation of basic equations*, Acta Astrn. **4** (1977), 1177–1206

[8] Sivashinsky, G.I., *On irregular wavy flow of a liquid film down a vertical plane*, Progr. Theoret. Phys. **63** (1980), 2112–2114

[9] Benny, D.J., *Long waves in liquid film*, J. Math and Phys. **45** (1966), 150–155

[10] Babchin, A.J., Frenkel,A.L.,Levich, B.G.and Sivashinsky,G.I., *Nonlinear saturation of Rayleigh-Taylor instability in thin films*, Phys. Fluids. **26** (1983), 3159–3161

[11] Hyman, J.M. and Nicolaenko, B., *The Kuramoto-Sivashinsky Equation: A bridge between PDE'S and dynamical systems*, Physica D **18** (1986), 113–126

[12] Kevrekidis, I.G., Nicolaenko, B. and Scovel,J.C., *Back in the saddle again: A computer assisted study of the Kuramoto-Sivashinsky equation*, SIAM J. APPL. MATH. Vol.50 **3** (1990) 760–790

[13] Foias, C., Nicolaenko, B., Sell, G.R.and Temam, R., *Inertial manifolds for the Kuramoto-Sivashinsky equation and an estimate of their lowest dimension*, J.Math.Pures Appl. **67** (1988), 197–226

[14] Nicoleanko, B., Scheurer, B. and Temam, R., *Some global dynamical properties of the Kuramoto-Sivashinsky equations: Nonlinear stability and attractors*, Physica D. **16** (1985), 155–185

[15] Nicoleanko, B., Scheurer, B. and Temam, R., *Some global dynamical properties of a class of pattern formation equations* Comm. Partial Differential Equations, **14** (1989), 245–297

[16] Temam, R., *Infinite-dimensional dynamical systems in mechanics and physics* Applied Mat.Sci. **68**, (1997)

[17] Desertion, K.A. and Kazantzis, N., *Improved performance of the controlled Kuramoto-Sivashinsky equation via actuator and controller switching*, Proceeding of the 2004 American Control Conference 267–272

[18] Lee, C.H. and Tran, H.T., *Reduced-order-based feedback control of the Kuramoto-Sivashinsky equation* Journal of Computational and Applied Mathematics **173** (2005) 1–19

[19] Protas, B., Bewley, T. R and Hagen, G., *A computational framework for the regularization of adjoint analysis in multiscale PDE systems*, J. Comput. Phys. **195**,Issue 1 (2004), 49–89

[20] Kassam, A.K., and Trefethen, L.N., *Fourth-order time-stepping for stiff PDS's* SIAM J.SCI.COMPT. Vol.26 **4** (2005), 1214–1233

[21] Krogstad, S., *Generalized integrating factor methods for stiff PDE's* J. Comput. Phys. **203** (2005), 72–88

[22] Canuto, C., Hussaini, M. Y., Quarteroni, A. and Zang, T.A., *Spectral methods in fluids dynamics* Springer-Verlag (1988)

[23] Trefethen, N.L., *Spectral methods in Matlab*, Software Environ. Tools 10, SIAM, Philadelphia, 2000

[24] Boyd, J. P., *Eight definitions of the slow manifold: Seiches, pseudoseiches and exponential smallness* Dyn. Atmos. Oceans **22** (1995) 49

[25] Cox, S.M. and Matthews, P.C., *Exponential time differencing for stiff systems* J. Comput. Phys. **176** (2002) 430–455

[26] Chorin, A.J and Krause, P., *Dimensional reduction for a Bayesian filter*, PNAS, **101** no.42 (2004) 15013–15017

[27] Hu, C. and Temam, R., *Robust boundary control for the Kuramoto-Sivashinsky equation*, Dynamics of Continuous Discrete and Impulsive Systems-Series B-Applications & Algorithms. **8** (3) (2001) 315–338

[28] Berliner, M.L. and Wikle C.K., *A Bayesian tutorial for data assimilation*, Physica D **30** (2007) 1–16

[29] Berliner, M.L. and Wikle C.K., *Approximate importance sampling Monte Carlo for data assimilation*, Physica D **30** (2007) 37–49

[30] Doucet, A. , Godsill, S., and Andrieu C., *On Sequential Monte Carlo sampling methods for Bayesian filtering*, Statistics and Computing **10**(3) (2000) 197–208

[31] Doucet, A. , de Freitas, J. F. G.,and Gordon, N. J., *An introduction to sequential Monte Carlo methods,*, in Sequential Monte Carlo Methods in Practice, eds Doucet, A. , de Freitas, J.F.G,and Gordon, N.J (2001), Springer-Verlag, New York

[32] Arulampalam, M.S., Maskell, S., Gordon, N. and Clapp, T, *A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking*, IEEE transactions on signal processing vol.(150) no.2 (2002) 174–188

[33] Gordon, N. J., Salmond, D. J. and Smith, A.F.M., *Novel approach to nonlinear non-Gaussian Bayesian state estimate*, IEEE Proc. F, **140** (1993) 107–113

[34] Kitagawa, G., *Monte Carlo filter and smoother for non-Gaussian nonlinear state space models*, Journal of Computational and Graphical Statistics **5**,1 (1996) 1–25

[35] Liu, J. S. and Chen R., *Sequential Monte Carlo Methods for Dynamic Systems* Journal of the American Statistical Association, Vol.93, No.442 (1998) 1032–1044

[36] Kalman, R. E., *A new approach to linear filtering and prediction problems*, Transaction of the ASME - Journal of basic Engineering, Series D, 82 (1960), 35–45

[37] Evensen, G., *Sequential data assimilation in non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, J. Geophys Res **99** (C5):10 (1994) 108–129

[38] Burgers, G., van Leeuwen, P. J. and Evensen, G., *Analysis scheme in the ensemble Kalman Filter*, Mon Weather Rev **126** (1998) 1719–1724

[39] Evensen, G., *The Ensemble Kalman Filter: theoretical formulation and practical implementation* Ocean Dynamics **53** (2003) 343367

[40] Evensen, G., *Data Assimilation: The ensemble Kalman filter*, Springer, Berlin, 2007

[41] Mandel, J., *A brief Tutorial on the ensemble Kalman filter* manuscript, (2007).

[42] Mandel, J., *Efficient Implementation of the Ensemble Kalman Filter*, CCM Report 231, (2006)

[43] Jazwinski, A. H., *Stochastic processes and filtering theory*, Academic Press, New York, 1970.

[44] Bucy, R. S., *Recurrent sets*, Ann. Math. Statist. **36** (1965) 535–545.

[45] Anderson, J. L. and Anderson, S. L., *A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts* Mon. Wea. Rev.,**127** (1999) 2741–2785.

[46] Gleeson, T. A.,*Statistical-dynamical prediction.* J. Appl. Meteor.,**9**, (1970) 333–344.

[47] Hol J. D., Schön, T. B. and Gustafsson, F., *On Resampling Algorithms for Particle Filters.* Nonlinear Statistical Signal Processing Workshop, Cambridge, United Kingdom, September 2006.

[48] Lewis, J.L., Lakshivarahan, S. and Dhall, S. K., *Dynamic Data Assimilation: A least squares approach* Cambridge University Press (2006) **Volume 104** in the series on Encyclopedia of Mathematics and its Applications

[49] Xiong, X. , Navon, I. M. and Uzunoglu, B. *A Note on the Particle Filter with Posterior Gaussian Resampling* Tellus 58A, (2006) 456–460.

[50] Nakano, S., G. Ueno, G., and Higuchi, T. *Merging particle filter for sequential data assimilation* Nonlin. Processes Geophys., 14, (2007) 395–409.