AGU PUBLICATIONS

Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1002/2014MS000373

A modified ensemble Kalman particle filter for non-Gaussian systems with nonlinear measurement functions

Zheqi Shen[1] and Youmin Tang[1,2]

[1]State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic Administration, Hangzhou, China, [2]Environmental Science and Engineering, University of Northern British Columbia, Prince George, British Columbia, Canada

Key Points:
- We have extended the EnKPF method to systems with nonlinear measurement function
- The modified EnKPF is shown to perform better than the EnKF
- This method suggests a promising direction for non-Gaussian data assimilation

Correspondence to:
Y. Tang,
ytang@unbc.ca

Citation:
Shen, Z., and Y. Tang (2015), A modified ensemble Kalman particle filter for non-Gaussian systems with nonlinear measurement functions, *J. Adv. Model. Earth Syst.*, *7*, 50–66, doi:10.1002/2014MS000373.

Received 22 AUG 2014
Accepted 22 DEC 2014
Accepted article online 31 DEC 2014
Published online 19 JAN 2015

Abstract The ensemble Kalman particle filter (EnKPF) is a combination of two Bayesian-based algorithms, namely, the ensemble Kalman filter (EnKF) and the sequential importance resampling particle filter (SIR-PF). It was recently introduced to address non-Gaussian features in data assimilation for highly nonlinear systems, by providing a continuous interpolation between the EnKF and SIR-PF analysis schemes. In this paper, we first extend the EnKPF algorithm by modifying the formula for the computation of the covariance matrix, making it suitable for nonlinear measurement functions (we will call this extended algorithm nEnKPF). Further, a general form of the Kalman gain is introduced to the EnKPF to improve the performance of the nEnKPF when the measurement function is highly nonlinear (this improved algorithm is called mEnKPF). The Lorenz '63 model and Lorenz '96 model are used to test the two modified EnKPF algorithms. The experiments show that the mEnKPF and nEnKPF, given an affordable ensemble size, can perform better than the EnKF for the nonlinear systems with nonlinear observations. These results suggest a promising opportunity to develop a non-Gaussian scheme for realistic numerical models.

1. Introduction

The classical Kalman filter [*Kalman*, 1960] provides a complete and rigorous solution for the state estimation of linear systems under Gaussian noise. If the system is nonlinear, the extended Kalman filter (EKF) provides a suboptimal estimate by linearizing the nonlinear models. However, such linearization can cause large truncation errors for highly nonlinear systems, and in particular is very difficult, even intractable, for certain high-dimensional complex models such as general circulation models (GCMs). For this reason, the ensemble Kalman filter (EnKF) has been introduced to address the drawbacks of the EKF. The EnKF is a sophisticated sequential data assimilation method, initially proposed by *Evensen* [1994] and later on modified by *Burgers et al*. [1998]. The EnKF chooses a set of samples, referred to as the ensembles of states, to capture the statistical information of the forecast states. The ensembles are propagated with the full nonlinear model to approximate the covariance of the prediction error, and observations are incorporated into the model according to the Kalman filter formula. The square root Kalman filter (EnSRF), which avoids the perturbations of the observations induced by the EnKF, has been developed, further promoting the application of EnKF [*Anderson*, 2001; *Bishop et al*., 2001; *Whitaker and Hamill*, 2002; *Tippett et al*., 2003]. The EnKF and its variants, e.g., the Ensemble Adjustment Kalman Filter [*Anderson*, 2001, 2009], Ensemble Transform Kalman Filter [*Bishop et al*., 2001], and EnSRF [*Tippett et al*., 2003], have enjoyed great success in atmospheric and oceanic data assimilation. However, these methods contain an inherent assumption, namely, that the error statistics are Gaussian. This assumption often fails to hold for nonlinear systems. Even an initial Gaussian error often becomes non-Gaussian while propagating forward with nonlinear models [*Bocquet et al*., 2010].

Several ideas have been proposed to address the issue of non-Gaussian error statistics in the field of geophysical data assimilation [*Bocquet et al*., 2010]. Among these proposals, the particle filter (PF) is a highly promising technique because it does not invoke any Gaussian assumptions. It has been widely used and studied in many other fields. The PF estimates the full probability density function (pdf) of the forecasted state based on an ensemble of states with different weights. A member of this ensemble is also often termed a particle. Similar to the EnKF, the filter propagates all the particles forward in time according to the nonlinear model. When observation data become available, the weights of the particles are changed such that the information present in the data is incorporated into the swarm of particles. In contrast, all members

are equally weighted in the EnKF framework, and the analysis scheme only changes the value of the members rather than the weights. The PF suffers from the problem of filter degeneracy, i.e., the procedure collapses to a very small number of highly weighted particles among a horde of almost useless particles carrying a tiny proportion of the probability mass. Even if resampling techniques are used, the degeneracy cannot be completely avoided with limited ensemble size. The number of particles must grow substantially with the dimension of the system to avoid degeneracy [*Bengtsson et al*., 2003; *Snyder et al*., 2008], a requirement that is apparently too costly for large models such as GCMs. Various efforts have been made to resolve this issue, as documented in several excellent overviews [e.g., *Cappé et al*., 2007; *van Leeuwen*, 2009].

Several strategies are often employed to address the problem of filter degeneracy in applications of the particle filter. For example, Papadakis et al. proposed a weighted ensemble Kalman filter (WEnKF), which uses an ensemble-based Kalman filter as the proposal density, from which the particles are drawn [*Papadakis et al*., 2010]. van Leeuwen developed a fully nonlinear particle filter by exploiting the freedom of the proposal transition density, which ensures not only that all particles ultimately occupy high-probability regions of state space but also that most of the particles have similar weights [*van Leeuwen*, 2010, 2011; *Ades and van Leeuwen*, 2013]. The implicit particle filter [*Chorin et al*., 2010; *Morzfeld et al*., 2012] uses gradient descent minimization combined with random maps to find the region of high probability, avoiding the calculation of Hessians. Luo and Hoteit have proposed an efficient particle filter that uses residual nudging to prevent the residual norm of the state estimates from exceeding a prespecified threshold [*Luo and Hoteit*, 2014]. However, these particle filters still require a relatively large ensemble to maintain their advantage over the EnKF.

In this study, we focus on the ensemble Kalman particle filter (EnKPF), which is a blend of the EnKF and the PF [*Frei and Künsch*, 2013; *Rezaie and Eidsvik*, 2012]. The analysis step of the EnKPF combines both the EnKF and PF updating schemes with a continuous transition index $\gamma \in [0, 1]$. The index can be chosen automatically such that the analysis provides the particle filter with weighting appropriate for avoiding degeneracy, providing a trade-off scheme between effectiveness and feasibility with limited computational resources. The main advantage of the EnKPF over other filters is that the algorithm does not have to fit a Gaussian distribution to the forecast ensemble, nor does it requires a large ensemble to prevent filter degeneracy. Accordingly, the EnKPF is easier to implement than the other filters considered.

As noticed in the original paper [*Frei and Künsch*, 2013, hereinafter referred to as FK2013], the EnKPF was initially developed for state estimation with nonlinear dynamical models and linear measurement functions. However, nonlinear measurement functions often occur in many realistic systems, e.g., the estimation of atmospheric or oceanic states from satellite and remote sensing data. Thus, it is interesting to extend the EnKPF algorithm to systems with nonlinear measurement functions. To achieve this goal, this study attempts to modify the EnKPF algorithm, making it suitable for nonlinear measurement functions. Furthermore, the modified EnKPF is tested using the Lorenz '63 and '96 models, both of which are simplified atmospheric models with realistic features of the atmosphere. These models have been widely used as test beds in data assimilation due to their low dimensional but highly nonlinear nature.

Section 2 gives a brief review of the ensemble Kalman filter and the particle filter, in which a modified scheme of Kalman gain for nonlinear measurement functions is proposed in a rigorous statistical sense with detailed derivations. On this basis, the EnKPF is introduced for nonlinear dynamical systems with nonlinear measurement functions. The applications of the modified EnKPF to the Lorenz '63 and '96 models are presented in section 3. Discussions and conclusions are given in section 4.

2. Bayesian-Based Methods

Atmospheric and oceanic flow can be described by a system of stochastic partial differential equations (sPDE). Within this framework, the dynamic system can be stochastically forced, and the observations are also considered in the form of stochastic processes rather than single numerical values. The sPDE model can often be described in terms of a state-space model as follows:

$$x_k = f(x_{k-1}, \eta_{k-1}), \tag{1}$$

$$y_k = h(x_k, \zeta_k), \tag{2}$$

where x_k denotes the state variable at discrete time step t_k, and y_k denotes the measurement of the state variable. Functions f and h describe the evolution of the state vector over time and the relationship between measurements and states, respectively, which might be linear or nonlinear. At the same time, the dynamic model noise η and the measurement noise ζ are also incorporated into the models. These noise variables are assumed to be independent. The aim of state estimation is to construct the posterior probability density function (pdf) of the required state vector using all available information. The posterior pdf, written as $p(x_k|Y_k)$, is a complete description of the state of knowledge about the required vector. Y_k denotes all the observations up to time step t_k, and y_k is the observation at the current time. An equivalent probabilistic description of the evolution of the state by equation (1) is $p(x_k|x_{k-1})$, which is sometimes called the transition density. Additionally, an equivalent probabilistic model for equation (2) is the conditional pdf $p(y_k|x_k)$, which is the likelihood of y_k, given an estimate x_k. Thus, in summary, the probabilistic description of the state estimate by equations (1) and (2) is $p(x_0), p(x_k|x_{k-1})$ and $p(y_k|x_k)$, where $p(x_0)$ is the prior pdf of the state vector at time $k = 0$, before any measurements have been received.

State estimation problems are addressed mainly within Bayesian estimation theory. The Bayesian recursive filter consists of two steps, namely a prediction and a correction operation. The prediction operation propagates the posterior pdf of the state vector from time step $k - 1$ forward to step k with the dynamic model. Using the probabilistic description, this operation yields

$$p(x_k|Y_{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1})dx_{k-1}. \tag{3}$$

The correction (or analysis) operation incorporates the information in the measurements and updates the state vectors to give the posterior pdf at time step k. By Bayes' rule, it can be written as

$$p(x_k|Y_k) = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{A}, \tag{4}$$

where $A = p(y_k|y_{k-1}) = \int p(y_k|x_k)p(x_k|Y_{k-1})dx_k$ is a normalizing factor. In most cases, the likelihood is assumed to be Gaussian, i.e., $p(y_k|x_k) = \phi(y_k; h(x_k), R)$. Here and subsequently, $\phi(x; \mu, \Sigma)$ denotes the multivariate normal density with mean value μ and covariance matrix Σ at x.

For both the EnKF and the PF, the prediction steps are the same and use the nonlinear dynamic model in (1). The two methods differ in terms of their specific analysis schemes and inherent assumptions, as discussed below.

2.1. The Ensemble Kalman Filter

The EnKF is a suboptimal method for state estimation in which the error statistics are analyzed by numerically solving the Fokker-Planck equation using the Monte Carlo method [*Evensen*, 2003], namely, using an ensemble of states to approximate the covariance matrices. Initially, the term EnKF refers, in particular, to the stochastic ensemble Kalman filter [*Evensen*, 1994], which requires perturbing the observations. Subsequently, several deterministic EnKFs that avoid the use of "perturbed observations" were developed, e.g., the ensemble transform Kalman filter [*Bishop et al*., 2001], the ensemble adjustment Kalman filter [*Anderson*, 2001], and the ensemble square root filter [*Tippett et al*., 2003]. These filter designs are labeled as variants of the EnKF in the literature [e.g., *Papadakis et al*., 2010], because they are also based on the Kalman filtering formula and ensemble representations. A common feature of all of these EnKF-based filters is the inherent Gaussian assumption, which is the key issue targeted in the present study. For simplicity and convenience, we only consider the stochastic EnKF as follows.

For a dynamic system with L variables, we denote the prior ensemble as $X^f \in \mathbf{R}^{L\times N}$,

$$X^f = (x_1^f, \ldots, x_i^f, \ldots, x_N^f),$$

where the subscript i and superscript f refers to the ith forecast ensemble member, and the ensemble size is N. Because we are considering a single analysis cycle, the subscript k for the time step is omitted. The empirical mean and empirical covariance are defined as follows:

$$\overline{x^f} = \frac{1}{N}\sum_{i=1}^{N} x_i^f, \tag{5}$$

$$\overline{P^f} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i^f - \overline{x^f})(x_i^f - \overline{x^f})^T. \tag{6}$$

As N tends to infinity, these variables converge to the true mean and covariance of the forecast state vector.

The EnKF performs the analysis step by applying the Kalman filter formula to each ensemble member,

$$x_i^a = x_i^f + K[y_i - h(x_i^f)],$$

for $i=1,\ldots,N$. The perturbed observations y_i are given by

$$y_i = y + v_i,$$

where v_i is a random variable having a normal distribution with zero mean and covariance R. The error covariance matrix computed from v_i converges to R as $N \to \infty$.

If the measurement function is linear and the noise is additive, namely,

$$y_k = Hx_k + \zeta, \tag{7}$$

the Kalman gain is defined by

$$K = \overline{P^f}H^T(H\overline{P^f}H^T + R)^{-1}. \tag{8}$$

However, if the measurement function is nonlinear, as in equation (2), the two terms $\overline{P^f}H^T$ and $H\overline{P^f}H^T$, which occur in equation (8), are written as follows [*Houtekamer and Mitchell*, 2001]

$$\overline{P^f}H^T \equiv \frac{1}{N-1}\sum_{i=1}^{N}[x_i^f - \overline{x^f}][h(x_i^f) - \overline{h(x^f)}]^T, \tag{9}$$

$$H\overline{P^f}H^T \equiv \frac{1}{N-1}\sum_{i=1}^{N}[h(x_i^f) - \overline{h(x^f)}][h(x_i^f) - \overline{h(x^f)}]^T, \tag{10}$$

where $\overline{h(x^f)} = \frac{1}{N}\sum_{i=1}^{N} h(x_i^f)$. Equations (9) and (10) allow for a direct evaluation of the nonlinear measurement function h in calculating the Kalman gain. However, these equations were only presented intuitionally and equivalent signs were used. Tang and Ambadan have argued that equations (9) and (10) approximately hold only if the following is true [*Ambadan and Tang*, 2009]:

$$\overline{h(x^f)} = h(\overline{x^f}), \tag{11}$$

$$\mathrm{Norm}(x_i^f - \overline{x^f}) \text{ is small for } i=1,2,\ldots,N. \tag{12}$$

Under the conditions of equations (11) and (12), equations (9) and (10) actually linearize the nonlinear measurement functions h to H. Therefore, the direct application of the nonlinear measurement function in equations (9) and (10) imposes an implicit linearization process using ensemble members.

A general algorithm for the Kalman gain for the nonlinear model and the nonlinear measurement function can be written as follows [*Julier and Uhlmann*, 1997; *Tang et al*., 2014]:

$$K = P_{xy}P_{yy}^{-1}, \tag{13}$$

where the P_{xy} is the cross covariance between the state and observation errors, and the P_{yy} is the error covariance of the difference between the observation and its prediction of $h(x^f)$. If the estimate is unbiased and the ensemble size is infinite, we can use the ensemble mean to represent the true value, i.e.,

$$x^{tr} = E[x_i^f] + \eta = \overline{x^f} + \eta,$$

$$y^{tr} = h(\overline{x^f}) + \zeta.$$

in which the overbar represents the mean over all the ensemble members. The terms η and ζ were added due to the random nature of the true state and are mutually independent. The error covariance matrices in equation (13) can be computed as follows:

$$P_{xy}=\frac{1}{N-1}\sum_{i=1}^{N}[x_i^f-\overline{x^f}][h(x_i^f)-h(\overline{x^f})]^T, \tag{14}$$

$$P_{yy}=\frac{1}{N-1}\sum_{i=1}^{N}[h(x_i^f)-h(\overline{x^f})][h(x_i^f)-h(\overline{x^f})]^T+R. \tag{15}$$

A summary of the main equations of the modified EnKF is given below:

$$x_i^a=x_i^f+K[y_i-h(x_i^f)],\quad i=1,2,\ldots,N, \tag{16}$$

$$K=\left\{\frac{1}{N-1}\sum_{i=1}^{N}[x_i^f-\overline{x^f}][h(x_i^f)-h(\overline{x^f})]^T\right\}*\left\{\frac{1}{N-1}\sum_{i=1}^{N}[h(x_i^f)-h(\overline{x^f})][h(x_i^f)-h(\overline{x^f})]^T+R\right\}^{-1}. \tag{17}$$

The modified Kalman gain, equation (17), does not add additional computations compared with the traditional formula, equations (8)–(10). It has been found that the modified Kalman gain can result in better assimilation analysis for nonlinear measurement functions [*Tang et al.*, 2014].

It should be mentioned that an iterative scheme-based EnKF has recently been proposed as a tool for better analysis. The idea behind this iterative EnKF is to repeatedly adjust the model states by multiple assimilations of the same observations. It has been argued that the iterative EnKF can address the nonlinear models and observations [e.g., *Sakov et al.*, 2012]. However, the process of multiple iterations is computationally costly.

2.2. The Particle Filter

The main difference between the particle filter and the EnKF is that in the particle filter, the ensemble members (or the particles) are not modified, but are combined with different weights. We use the likelihood for an ensemble member $x_k^{(i)}$ given the observation y_k to update its weight. More specifically, the pdf of the analysis at time step $k-1$ is assumed to be a linear combination of Dirac-Delta functions

$$p(x_{k-1}|Y_{k-1})=\sum_{i=1}^{N}w_{k-1}^{(i)}\delta(x_{k-1}-x_{k-1}^{(i)}),$$

which is not necessarily Gaussian. Here $x_{k-1}^{(i)}$ is a particle at time step $k-1$, with corresponding weight $w_{k-1}^{(i)}$.

If new particles at step k are generated by the transition density $p(x_k|x_{k-1})$, i.e., $x_k^{(i)}=f(x_{k-1}^{(i)})$, the prior pdf at step k is

$$p(x_k|Y_{k-1})=\sum_{i=1}^{N}w_{k-1}^{(i)}\delta(x_k-x_k^{(i)}).$$

Based on Bayes' rule by equation (4), the posterior pdf can be computed as follows:

$$p(x_k|Y_{k-1})=\frac{1}{A}\sum_{i=1}^{N}p(y_k|x_k^{(i)})w_{k-1}^{(i)}\delta(x_k-x_k^{(i)})=\sum_{i=1}^{N}w_k^{(i)}\delta(x_k-x_k^{(i)}), \tag{18}$$

where

$$w_k^{(i)}\propto p(y_k|x_k^{(i)})w_{k-1}^{(i)}. \tag{19}$$

With the estimated pdf, each statistical moment of the updated state vectors x_k can be derived. For example, the mean value of x_k is

$$\overline{x_k}=\int x_k p(x_k|Y_k)dx_k=\sum_{i=1}^{N}w_k^{(i)}x_k^{(i)}.$$

The particle filter allows us to infer high-order moments of the posterior state. Suppose that $f(x)$ is a function of the state vector x_k, the expectation of $f(x)$ is then

$$\overline{f(x_k)} = \int f(x_k) p(x_k|Y_k) dx_k = \sum_{i=1}^{N} w_k^{(i)} f(x_k^{(i)}).$$

High-order moments of the posterior state can be computed using $f(x) = x^n$. The likelihood function $p(y_k|x_k^{(i)}) = \phi(y_k; h(x_k^{(i)}), R)$ usually takes the form of a Gaussian, although other forms of likelihood functions are also applicable. To simplify the discussion, we assume that the observation noise is Gaussian, and the likelihood function is as follows [*van Leeuwen*, 2009]:

$$p(y|x) = \phi(y; h(x), R) = A\exp\left\{-\frac{1}{2}[y-h(x)]^T R^{-1}[y-h(x)]\right\}. \tag{20}$$

The bootstrap particle filter is commonly considered the first example of modern particle filters [*Cappé et al*., 2007]. Its basic idea is to use the transition density $p(y_k|x_k^{(i)})$ as the proposal density, producing new particles in the next time step. The implementation of the bootstrap PF is easy, but its main drawback is that the particles are not modified. As a result, when all particles move away from the observations, they are not pulled back. Experiments have shown that most of the weights of the particles vanish after a few steps, leading to a poor analysis. Therefore, the weights to collapse has been reported previously in the geophysical literature [e.g., *Bengtsson et al*., 2003], where it is often referred to as "degeneracy" or "impoverishment." To overcome degeneracy, particle filters invariably employ certain forms of resampling or selection steps after the updated weights are calculated, i.e., removing particles with very small weights and duplicating those with large weights. The Sequential Importance Resampling (SIR) algorithms, including the Systematic Resampling (SR) and Residual Resampling (RR) algorithms [*Doucet et al*., 2001], essentially import the information contained in the weights into the swarm of posterior particles, making the weights become uniform again.

Certain advanced particle filters are based on choosing more sophisticated resampling schemes to prevent filter degeneracy. These filters include the merging particle filter (MPF) [*Nakano et al*., 2007], and the PF with Gaussian resampling (PFGR) [*Xiong et al*., 2006]. They are good attempts but still require a large number of particles to produce superior performance compared with the EnKF.

To explore the problem of filter degeneracy, one can define the effective ensemble size N_{eff} as follows [*Arulampalam et al*., 2002]:

$$N_{eff} = 1/\sum_{i=1}^{N} (w^{(i)})^2. \tag{21}$$

The effective ensemble size is a measurement of degeneracy and is related to the distance between the weights $w^{(i)}$ and uniform weights. Because the weights sum to 1, the value of N_{eff} varies between 1 and N. A value close to 1 indicates that there is only one useful sample in the set, i.e., severe degeneracy. Conversely, if the weights are spread uniformly among the particles, the effective ensemble size approaches N. In practical applications of the SIR-PF, the resampling step is required only if the effective ensemble size is below a prespecified threshold; typically, $N_{eff} = N/2$. If N_{eff} is below the threshold, the assimilation system is most likely degenerate. This threshold relationship is often termed the effective sample size criterion [*Liu*, 2008].

2.3. The Ensemble Kalman Particle Filter for Nonlinear Measurement Functions

Several comparisons between the performance of the EnKF and the SIR-PF have been reported [e.g., *Bocquet et al*., 2010; *Nakano et al*., 2007]. These comparisons have yielded the conclusion that each method has its own advantages and drawbacks. On the one hand, the EnKF provides good estimates with a very small ensemble. This property makes it suitable for large models. On the other hand, as the ensemble size increases, the root-mean-squared error of the EnKF analyses rapidly converges to a lower bound, and the performance of the EnKF cannot be further improved by increasing the ensemble size. The limitation of the EnKF is due to its inherent Gaussian assumption. The SIR-PF is able to outperform the EnKF if the ensemble size is sufficiently large to prevent filter degeneracy. However, even for the Lorenz '96 model with $N = 10$

variables, the particle filter still requires 10^4 members to match the performance of the EnKF [*Bocquet et al.*, 2010].

As stated above, both EnKF and PF methods are based on the Bayesian estimation theory, but they approximate the probability density function of the state in different ways. The EnKF only approximates the mean and covariance of the state through a series of equally weighted ensemble members. The analysis of EnKF, which is a weighted combination of the prediction and observation through Kalman gain, updates each ensemble member based on its distance from the observation in the state space. In contrast, the particle filter only updates the weight of each particle in the analysis step without updating the particle itself. Because most particles may have small weights, a large number of particles are required to prevent filter degeneracy, making the particle filter impractical for high-dimensional models.

The EnKPF takes advantage of both methods by combining the analysis schemes of the EnKF and the SIR-PF using a controllable index (i.e., tuning) parameter. In contrast with both the EnKF and the SIR-PF, the analysis scheme of the EnKPF not only updates the particles but also considers the weights.

For convenience, we first present the analysis scheme of EnKPF for the linear measurement function, as in FK2013. Here we assume that the forecast ensemble x_i^f, $i=1, 2, \ldots, N$ and the observation data y are available and that the forecast covariance P^f can be calculated using equation (6).

1. Choose $\gamma \in [0, 1]$ and apply the EnKF, which is based on the inflated observation error covariance R/γ as follows:

$$K_1(\gamma)=P^f H^T (HP^f H^T+R/\gamma)^{-1}=\gamma P^f H^T (\gamma HP^f H^T+R)^{-1}, \tag{22}$$

$$v_i=x_i^f+K_1(\gamma)(y-Hx_i^f), \tag{23}$$

$$Q=\frac{1}{\gamma}K_1(\gamma)RK_1(\gamma)^T. \tag{24}$$

2. Compute the weights w_i for each updated member v_i as follows:

$$w_i=\phi(y; Hv_i, \frac{R}{1-\gamma}+HQH^T), \tag{25}$$

and normalize the weights by $\hat{w}_i=w_i/\sum_{i=1}^N w_i$.

3. Calculate the resampling index $s(i)$ for each member v_i according to $\hat{w}_i$ with the residual resampling schemes [*Arulampalam et al.*, 2002], then set

$$x_i^u=v_{s(i)}+K_1(\gamma)\frac{\epsilon_{1,i}}{\sqrt{\gamma}}, \tag{26}$$

where $\epsilon_{1,i}$ is a random observation error drawn from the Gaussian $\mathbf{N}(0, R)$.

4. Compute $K_2(1-\gamma)=(1-\gamma)QH^T[(1-\gamma)HQH^T+R]^{-1}$, generate $\epsilon_{2,i}$ from $\mathbf{N}(0, R)$ and apply the EnKF with the inflated observation error again as follows:

$$x_i^a=x_i^u+K_2(1-\gamma)[y+\frac{\epsilon_{2,i}}{\sqrt{1-\gamma}}-Hx_i^u]. \tag{27}$$

The analysis scheme of the EnKPF consists of two parts, including updating the particles with the EnKF (step 1 and step 4) and redistributing the particles with the SIR-PF (step 2 and step 3). However, the EnKF procedure is split into two parts in practice, as given by equations (23) and (27), processed using tempered likelihood functions $\phi(y; h(x), R/\gamma)$ and $\phi(y; h(x), R/(1-\gamma))$, respectively. The covariance of the observations is divided by γ or $(1-\gamma)$ to prevent overfitting. Equation (25) indicates that the SIR-PF procedure is also based on the tempered likelihood, in which the covariance consists of R divided by (1 – γ) and the covariance of the estimate from the EnKF. The observation y in equation (23) is supposed to be perturbed in the regular EnKF. However, the perturbation is delayed until the resampling step is completed via equation (26). By doing this, the non-Gaussian information in the ensemble can be assimilated.

The EnKPF was initially designed for data assimilation for nonlinear dynamical systems with Gaussian linear observations. If the measurement function is nonlinear, i.e., $y_k = h(x_k, \zeta_k)$, equations (22) and (25) cannot be directly used. One solution is to directly linearize the nonlinear function and obtain H, but this approach is often technically difficult. The other solution is to implicitly linearize the nonlinear function without obtaining H, as shown in equations (9) and (10) in section 2.1.

For the nonlinear measurement function h, the H used in the above algorithm can be obtained by linearization in a way similar to that used in the EKF. However, the linearization is most likely difficult, even intractable, for certain complicated measurement functions, e.g., if, the measurement function is a nonlinear dynamical model. Here for the EnKPF, we propose to adopt the strategy used in the EnKF by *Houtekamer and Mitchell* [2001]. Specifically, the EnKPF algorithm can be modified for nonlinear measurement functions (this modified algorithm will be termed the nEnKPF hereafter) as follows:

1. Choose γ and perform the analysis with an EnKF that has an inflated observation error covariance R/γ, namely,

$$K_1(\gamma) = \left\{ \frac{1}{N-1} \sum_{i=1}^{N} [x_i^f - \overline{x^f}][h(x_i^f) - \overline{h(x^f)}]^T \right\} * \left\{ \frac{1}{N-1} \sum_{i=1}^{N} [h(x_i^f) - \overline{h(x^f)}][h(x_i^f) - \overline{h(x^f)}]^T + R/\gamma \right\}^{-1}, \quad (28)$$

$$v_i = x_i^f + K_1(\gamma)[y - h(x_i^f)], \quad (29)$$

$$Q = \frac{1}{N-1} \sum_{i=1}^{N} \omega_i \omega_i^T, \quad (30)$$

   where $\omega_i = \frac{K_1(\gamma)\epsilon_{1,i}}{\sqrt{\gamma}}$, and $\epsilon_{1,i}$ is the random observation error drawn from the Gaussian $\mathbf{N}(0, R)$.

2. Compute the weights w_i as follows:

$$w_i = \phi(y; h(v_i), \frac{R}{1-\gamma} + HQH^T), \quad (31)$$

   where the term HQH^T is defined by

$$HQH^T \equiv \frac{1}{N-1} \sum_{i=1}^{N} [h(\omega_i) - \overline{h(\omega)}][h(\omega_i) - \overline{h(\omega)}]^T, \quad (32)$$

   and normalize the weights by $\hat{w}_i = w_i / \sum_{i=1}^{N} w_i$.

3. Calculate the resampling index $s(i)$ or each member v_i according to $\hat{w}_i$ and set $x_i^u = v_{s(i)} + \omega_i$.

4. Generate $\epsilon_{2,i}$ from $\mathbf{N}(0, R)$ and apply the EnKF as follows

$$K_2(1-\gamma) = \left\{ \frac{1}{N-1} \sum_{i=1}^{N} (\omega_i - \overline{\omega_i})[h(\omega_i) - \overline{h(\omega)}]^T \right\} * \left\{ \frac{1}{N-1} \sum_{i=1}^{N} [h(\omega_i) - \overline{h(\omega)}][h(\omega_i) - \overline{h(\omega)}]^T + \frac{R}{1-\gamma} \right\}^{-1}, \quad (33)$$

$$x_i^a = x_i^u + K_2(1-\gamma)[y + \frac{\epsilon_{2,i}}{\sqrt{1-\gamma}} - h(x_i^u)]. \quad (34)$$

As discussed in section 2.1, the direct use of nonlinear measurement h in equation (28) actually contains an implicit linearization, which can cause truncation errors. Thus, a preferable solution is to process the nonlinear measurement function based on equation (17). Specifically, equation (28) can be replaced by

$$K_1(\gamma) = \left\{ \frac{1}{N-1} \sum_{i=1}^{N} [x_i^f - \overline{x^f}][h(x_i^f) - h(\overline{x^f})]^T \right\} * \left\{ \frac{1}{N-1} \sum_{i=1}^{N} [h(x_i^f) - h(\overline{x^f})][h(x_i^f) - h(\overline{x^f})]^T + R/\gamma \right\}^{-1}. \quad (35)$$

This approach generates a slightly different scheme from nEnKPF. The modified scheme will be termed the mEnKPF. In fact, equations (32) and (33) can also be replaced by the modified algorithm for the Kalman gain, but this approach would make very little difference because the mean value of ω_i is zero.

The parameter γ plays an important role in adjusting the proportions of analysis represented by the EnKF and by the PF. If $\gamma = 0$, it is easy to check that $K_1(\gamma) = 0$, and $K_2(1-\gamma) = 0$, corresponding to the SIR-PF. In

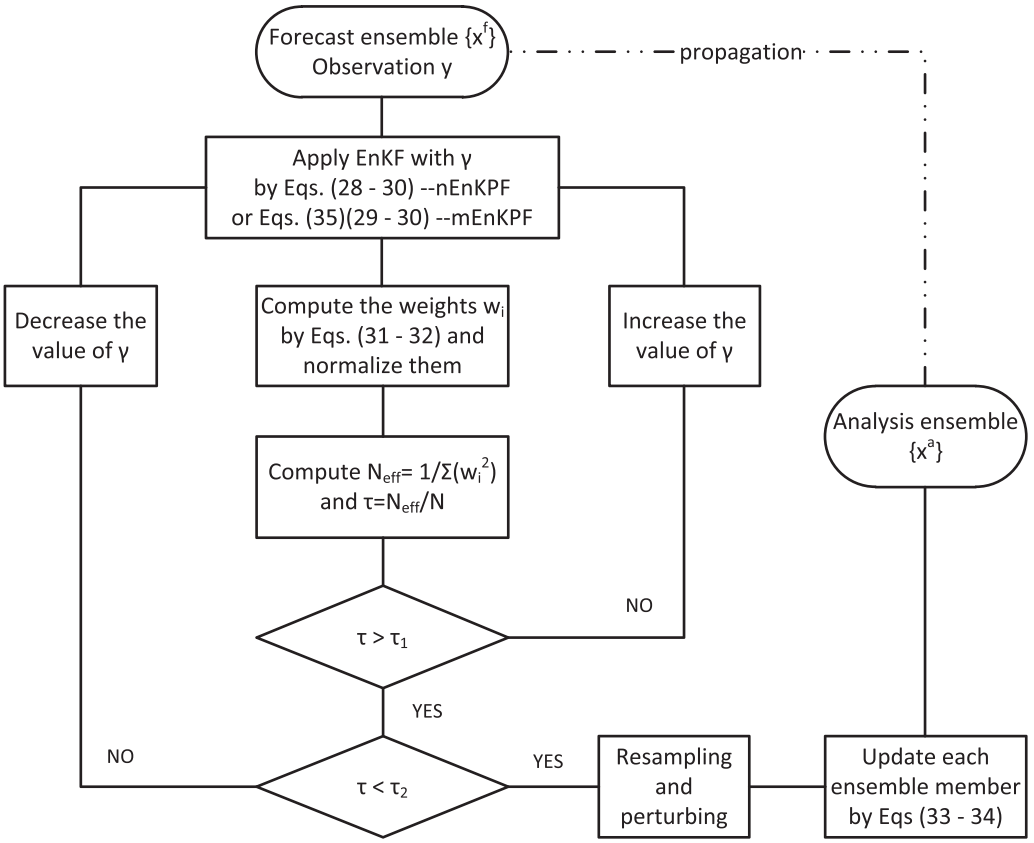

Figure 1. Flow diagram of the EnKPF algorithms with adaptive γ to limit the ensemble diversity τ to the prespecified interval $[\tau_1, \tau_2]$.

contrast, $\gamma = 1$ implies all $w_i = 1/N$ and $K_2(1-\gamma) = 0$, corresponding to the EnKF. The EnKPF algorithm essentially provides a continuous interpolation between the particle filter and the ensemble Kalman filter. A larger γ means that we place more emphasis on moving the particles toward the observation, reducing the variance in the weights, which results in a large N_{eff}, and makes degeneracy less likely. Meanwhile, a large γ also implies that less non-Gaussian information is incorporated into the ensemble by the PF, weakening the performance of the EnKPF.

The criterion for choosing γ follows the principle described in FK2013, namely that, γ should be neither too large nor too small, to achieve a suitable balance between the EnKF and the PF. As stated in section 2.2, the effective ensemble size N_{eff} is a measurement of degeneracy, and it is also a monotonic function of γ. The parameter γ is chosen adaptively in each data assimilation step to ensure that N_{eff} is neither too large nor too small. Thus, if we define the diversity of the ensemble by $\tau = N_{eff}/N$, then $0 < \tau < 1$, the EnKPF can always find a value of γ such that the corresponding $\tau \in [\tau_1, \tau_2]$ in each step.

An adaptive search procedure to find a suitable γ value can be based on including $\tau = N_{eff}/N$ into a prespecified interval $[\tau_1, \tau_2]$. In this procedure, an arbitrary initial guess (e.g., 1/2) is used for the γ, and τ can then be computed from the weights using equation (21). If $\tau < \tau_1$, we should use a larger γ, whereas if $\tau > \tau_2$, a smaller γ would be preferable. To avoid excessive computations, a binary search tree [*Knuth*, 2013] is used to find a suitable γ. For convenience, we assume that γ only takes the values of multiples of 1/16 as its value. In this situation, at most four steps are required to determine the smallest γ such that $\tau \in [\tau_1, \tau_2]$ if possible.

In short, we have proposed two modified EnKPF algorithms for nonlinear measurement functions, nEnKPF and mEnKPF. These algorithms differ in the approach used to formulate the Kalman gain. The details of the procedures used in these algorithms can be summarized in a schematic figure (Figure 1).

3. Numerical Experiments

In this section, we will test the nEnKPF and mEnKPF algorithms using assimilation experiments. For comparative purposes, the same experiment will also be performed using the EnKF. For both the nEnKPF and the mEnKPF, the parameter γ will be chosen adaptively during the assimilation to control the relative diversity τ in a prespecified interval, as discussed above. Two numerical models that have been widely used in the assimilation community will be used for the purpose of validation.

3.1. Lorenz '63 Model

The Lorenz '63 model [*Lorenz*, 1963] is the benchmark in the field of chaos, which was derived from the simplified equations of convection rolls arising at the Equator. It shares many common features with the atmospheric dynamical system, and can be used to simulate nearly regular oscillations or highly nonlinear fluctuations occurring in reality by adjusting the model parameters. In data assimilation, this model has served as the test bed for examining the properties of various assimilation methods.

The Lorenz '63 model consists of a system of three coupled and nonlinear ordinary differential equations,

$$\frac{dx}{dt}=\sigma(y-x)+q^x, \tag{36}$$

$$\frac{dy}{dt}=\rho x-y-xz+q^y, \tag{37}$$

$$\frac{dz}{dt}=xy-\beta z+q^z, \tag{38}$$

where $x(t)$, $y(t)$, and $z(t)$ are the dependent variables, and related to the intensity of convective motion, and the temperature gradients in the horizontal and vertical directions, respectively. q^x, q^y, and q^z represent the unknown model errors, assumed to be uncorrelated in time and unbiased. The parameters of σ, ρ, and β are set to 10, 28, and 8/3, respectively, as in *Miller et al*. [1994]. This model is integrated with the a fourth-order Runge-Kutta scheme, using a time step of $\Delta t=0.01$. One integration step is often called a model step in the following discussions.

For validation, we create the true state vector by integrating the model without model noises (i.e., perfect model), where the initial conditions are set to 1.508870, −1.531271, and 25.46091. The observations are made every T model steps with a randomly designed nonlinear measurement function, namely,

$$Y_k=10\tanh(X_k)+\zeta_k, \tag{39}$$

where the state vector $X=[x;y;z]$ contains all the model states; the observation vector Y is related to the model state space by a randomly designed nonlinear measurement function. Here we use the hyperbolic tangent function as an example. The observation noises are assumed to be Gaussian and additive. For a nonadditive noise system (e.g., multiplicative noise), Gaussian-based assimilation methods, such as the EnKF, are often invalid. All through the experiment, the observation noises are randomly drawn from the Gaussian distribution $\mathbf{N}(0,\sqrt{2})$.

The assimilation experiment is designed to estimate model states from an inaccurate initial condition using the imperfect model. The initial conditions are perturbed by a random noise drawn from $\mathbf{N}(0,1)$, and the model noises $[q^x;q^y;q^z]$ are randomly drawn from $\mathbf{N}(0,0.04)$. The system is run for 5500 assimilation cycles, and the first 500 cycles are discarded, where one assimilation cycle is defined as an assimilation process.

Figure 2 shows the estimates of the variable x for the last 100 assimilation cycles, as a function of the model step. The EnKF, nEnKPF, and mEnKPF are performed with the same ensemble size $N=64$, and observations are assimilated every $T=25$ model steps. Both the nEnKPF and the mEnKPF use the same adaptive strategy to limit the diversity τ to a value in $[0.1,0.3]$. The variable x is randomly chosen but an examination of other two variables reveals similar estimation accuracy (not shown). Figure 3 shows the RMSE of all state variables, which is calculated at all model steps during the 100 assimilation cycles, namely,

$$\mathrm{RMSE}(X,X^{true})=\sqrt{\frac{1}{L}(X-X^{true})^T(X-X^{true})},$$

where X and X^{true} are the assimilated state vector and true vector, respectively, and $L=3$ is the length of the state vector.

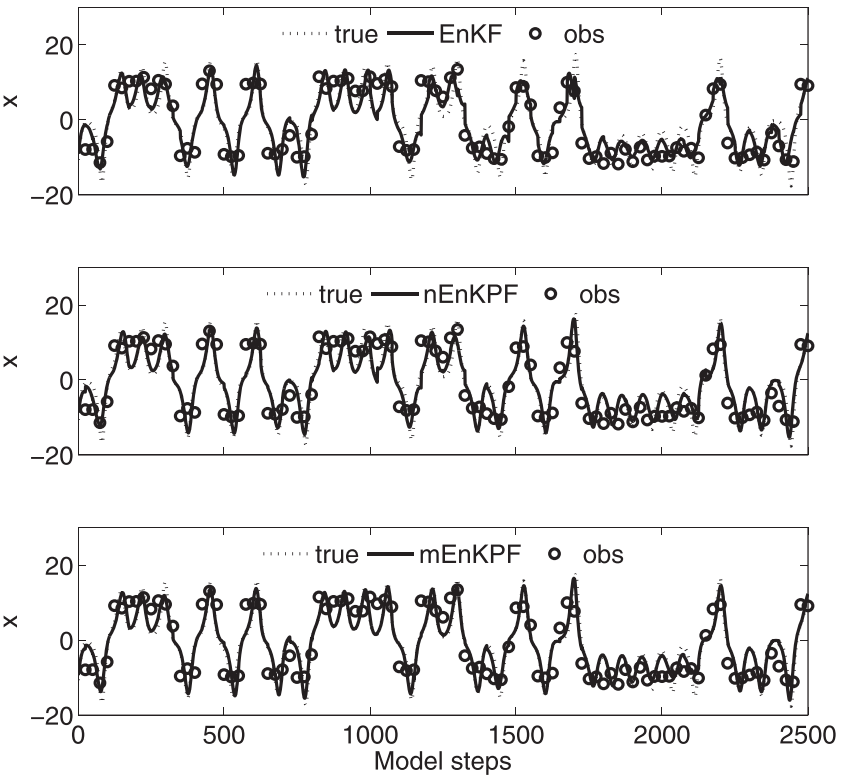

Figure 2. State estimates of variable *x* of the Lorenz '63 model for the last 100 data assimilation cycles.

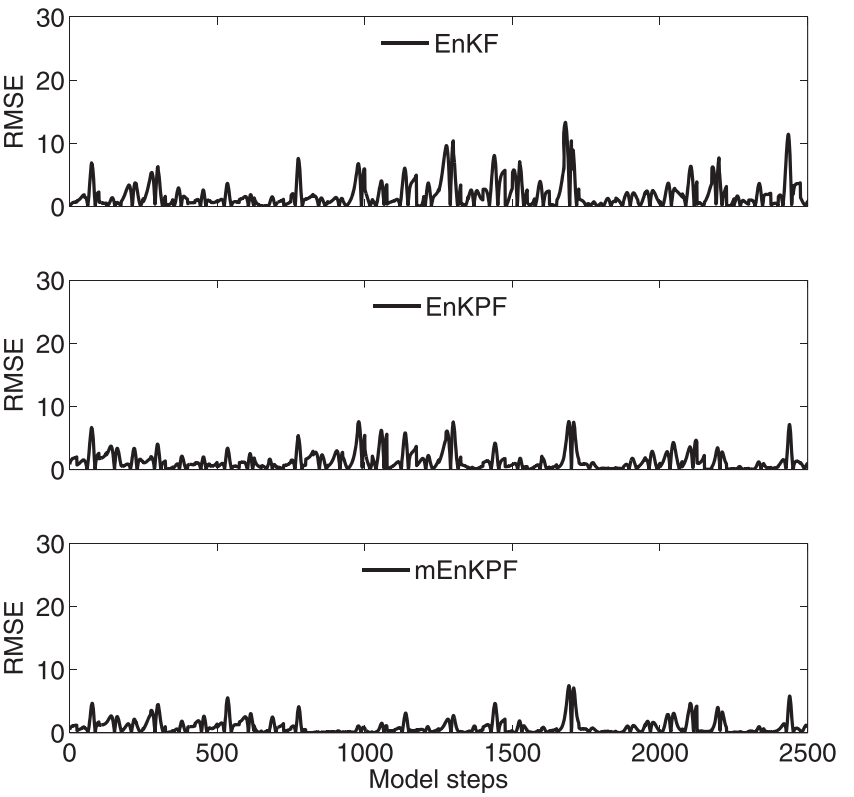

Figure 3. RMSE of all variables of the state estimates of Lorenz '63 model for the last 100 data assimilation cycles.

Table 1. The Average RMSE Over 5000 Cycles for the Lorenz '63 Model Using the EnKF and the Two Algorithms of EnKPF, i.e., Alg.1 = nEnKPF, Alg.2 = mEnKPF[a]

Data Frequency		25 Steps		30 Steps		35 Steps		40 Steps	
	$[\tau_1,\tau_1]$	Alg.1	Alg.2	Alg.1	Alg.2	Alg.1	Alg.2	Alg.1	Alg.2
EnKPF	[0.1,0.3]	1.23	1.07	1.41	1.23	1.84	1.63	1.94	1.75
	[0.3,0.5]	1.26	1.09	1.35	1.23	1.85	1.59	1.96	1.72
	[0.5,0.7]	1.3	1.08	1.41	1.23	1.97	1.65	2.05	1.79
	[0.7 0.9]	1.45	1.15	1.5	1.25	2.07	1.8	2.24	1.95
EnKF		1.83		2.15		2.33		2.42	

[a]Several different constrained diversity intervals ($[\tau_1, \tau_2]$) and different assimilation frequencies are considered (see text).

As shown in Figures 2 and 3, both the nEnKPF and the mEnKPF are better than the EnKF, indicating that the EnKPF that was originally proposed for a linear measurement function can also perform well for nonlinear measurement functions if its algorithm is properly modified. It is consistent with the finding in FK2013 that the EnKPF with linear measurement functions outperforms the EnKF. Figure 3 also shows that the mEnKPF seems slightly better than the nEnKPF, in particular, during the model step from 750 to 1500. The most likely explanation for this result is that the use of the Kalman gain equation (35) in the mEnKPF can better estimate the prediction error covariance matrix than the Kalman gain equation (28) of the nEnKPF, as discussed above.

The results from the SIR-PF are not presented here because filter degeneracy was observed to occur during its assimilation cycles. The degeneracy of PF due to finite particles has been a challenging issue, and considerable attention has been devoted to it in recent years. It has been reported that some significant progress has been made in avoiding filter degeneracy, although there is still a long way to go to completely solve the filter degeneracy problem with affordable numbers of particles for high-dimensional systems in practice [*Luo and Hoteit*, 2014; *van Leeuwen*, 2010]. The main objective of this paper is to extend the EnKPF, and further investigations on modern particle filters are beyond the scope of this study. As a result, the performance of the EnKPF is only compared with that of the EnKF in this study.

Table 1 shows the averaged RMSE over the last 5000 assimilation cycles for three assimilation schemes. To explore the impact of assimilation frequency on assimilation analysis, we conducted four sensitivity experiments, namely, assimilation performed every 25, 30, 35, and 40 model steps, respectively. The ensemble size was set to 64 in all experiments. The constrained diversity was tuned in different intervals $\tau \in [\tau_1, \tau_2]$ for the nEnKPF and mEnKPF. It can be observed from Table 1 that the averaged RMSE is smaller in both the EnKPF algorithms than in the EnKF for all experiments. This finding is consistent with the results shown in Figures 2 and 3. The optimal performance of the nEnKPF and the mEnKPF is obtained if $[\tau_1, \tau_2]$ is [0.1,0.3] or [0.3,0.5].

To compare our results with those obtained from a more advanced EnKF algorithm, we also applied the iterative ensemble Kalman filter (IEnKF) [*Sakov et al*., 2012]. The EnKF with an iterative part has attracted a broad attention in recent years, for its ability to address nonlinearities in data assimilation systems. Previous studies have shown that the IEnKF significantly outperforms the EnKF if the model is nonlinear and the measurement function is linear or weakly nonlinear. In our experiment, however, we found that the IEnKF had very poor performance, even worse than the EnKF. The reason for this result is that the nonlinearity in the measurement function was extremely strong, violating the IEnKF stipulation that the ensemble anomalies should propagate in a generally linear manner. This stipulation is adopted to make the iterative scheme converge [*Sakov et al*., 2012].

As mentioned in section 2.3, a control factor that affects the performance of both the nEnKPF and the mEnKPF is the prespecified interval of diversity given by $[\tau_1, \tau_2]$. We found from a series of sensitivity experiments that the optimal interval of τ values that yields the best assimilation analysis is dependent on the ensemble size. For example, the optimal value of the τ interval was [0.7,0.9] for ensemble size $N = 8$ and 16, [0.3,0.5] for $N = 32$, and [0.1,0.3] for larger N. Figure 4 shows the RMSE of the EnKF, nEnKPF, and mEnKPF with the optimal interval of τ as a function of ensemble size. One significant feature of this figure is that these three assimilation methods can all approach a minimum RMSE value (i.e., a "saturated RMSE") as the ensemble size increases. The assimilation performance of the methods cannot be further improved beyond

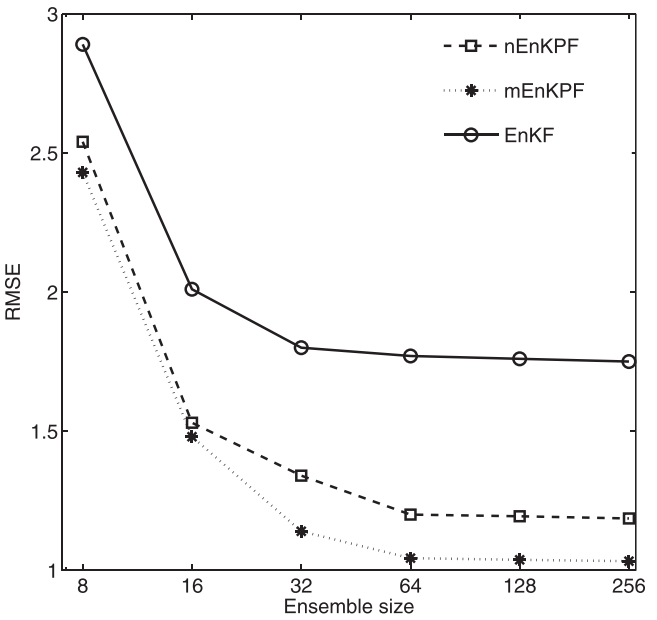

Figure 4. The averaged RMSE of the EnKF, nEnKPF, and mEnKPF as a function of ensemble size.

this value. A comparison of the methods shows that the EnKF has a much larger saturated RMSE than either the nEnKPF or the mEnKPF, indicating the advantage of the latter for non-Gaussian systems. However, the EnKF converges to the saturated value of the error more rapidly, requiring only half the ensemble size of the nEnKPF or the mEnKPF. This comparison illustrates the relative computational efficiency of the EnKF. In addition, the mEnKPF is superior to the nEnKPF for all ensemble sizes in Figure 4, and this result is consistent with the conclusion drawn from Table 1. As stated in section 2.3, the two methods differ only in the formulation of the Kalman gain, i.e., equation (28) in the nEnKPF and equation (35) in the mEnKPF. As discussed in section 2.1, the nEnKPF imposes an implicit linearization process, whereas the mEnKPF removes this process. Therefore, the mEnKPF performs better than the nEnKPF if the measurement function is nonlinear.

To explore the impact of the nonlinearity of the measurement function on the EnKPF, we slightly modified the measurement function of equation (39), as follows:

$$Y_k = 10 \tanh (X_k/a) + \zeta_k,$$

where a is a scale parameter. We set $a = 1, 5, 10$ to represent different degrees of nonlinearity in the measurement function. We repeated the above experiments with an assimilation frequency of 25 model steps and calculated the RMSE using the optimal interval for τ of [0.1,0.3]. The results are shown in Table 2, where two ensemble sizes, 64 and 256, are considered.

The value $a = 1$ implies a strongly nonlinear measurement function. In this case, both the nEnKPF and mEnKPF are better than the EnKF, and the mEnKPF outperforms the nEnKPF significantly for $N = 64$. For $N = 256$, the advantage is insignificant, most likely because both methods are approaching the saturated value of RMSE, as shown in Figure 2. In the other two cases, $a = 5$ and $a = 10$, the nonlinearity of the measurement functions decreases, with the result that the nEnKPF and the mEnKPF perform better than the EnKF only at a very large ensemble size. This finding is consistent with FK2013, in which the EnKPF is proposed to address linear or weakly nonlinear measurement functions and requires a large ensemble size of the order of 10^2 to outperform the EnKF. As discussed above, we found that if the measurement function is highly nonlinear, the performance of both the nEnKPF and the mEnKPF, in particular the mEnKPF, can be better than that of the EnKF, even with a relatively small ensemble size.

Table 2. The Average RMSE Over 5000 Cycles Using the EnKF, nEnKPF, and mEnKPF for Lorenz '63 Model With Assimilation Every 25 Model Steps ($T = 25$)[a]

	EnKF		nEnKPF		mEnKPF	
T = 25	N = 64	N = 256	N = 64	N = 256	N = 64	N = 256
a = 1	1.5	1.48	1.03	0.91	0.92	0.9
a = 5	0.91	0.91	1.15	0.74	1.11	0.74
a = 10	1.14	1.11	2.02	1.07	2.03	1.07

[a]Three types of nonlinear measurement functions are used to generate the observations.

3.2. Lorenz '96 Model

In this section, we use the Lorenz '96 model as the test bed to further examine the nEnKPF and mEnKPF. The Lorenz '96 model [*Lorenz*, 1996] represents an atmospheric variable X at J equally spaced points around a circle of the constant latitude. The jth component is propagated forward in time according to the differential equation

$$\frac{dX_j}{dt} = (X_{j+1} - X_{j-2})X_{j-1} - X_j + F, \tag{40}$$

where $j = 0, 1, \ldots, J-1$ represents a spatial coordinates. The boundary conditions are cyclic, i.e., $X_{J-1} = X_{-1}, X_J = X_0$, and $X_{J+1} = X_1$. This specification implies that the distance between two adjacent grid points roughly represents the midlatitude Rossby radius (approximately 800 km), assuming that the circumference of the midlatitude belt is approximately 30,000 km. F is a constant external forcing term that dominates the system dynamics. In this experiment, we set $F = 8$ to make the system highly chaotic.

This dynamic model is integrated with a time step of 0.05, corresponding to 6 hours in the realistic atmospheric physics. The initial condition is produced after a spin up integration of 10 years (i.e., 14,400 steps) with a rather arbitrary vector as in [*Lorenz*, 1996]. The true state vectors are obtained by integrating equation (40), and the observation data are generated by a nonlinear measurement function every T model steps, in which the observational noises are also drawn from $\mathbf{N}(0, \sqrt{2})$. The model noise q is set with a mean of zero and a standard deviation of 0.05. The system is run for 2500 assimilation cycles, and the first 500 cycles are discarded.

We first consider the Lorenz '96 model with $J = 10$ variables, and the observations are available for all variables in assimilation. Such a setting has been used as a test bed for the SIR-PF by *Bocquet et al*. [2010], who found that the SIR-PF performed better than the EnKF only at ensemble sizes greater than 10,000. In our experiment, four experiments were conducted to explore the sensitivity of the assimilation performance to the assimilation frequency and to the strength of nonlinearity of the measurement function. In detail, we performed the assimilation every $T = 8$ or 12 model steps, and we used the nonlinear measurement functions $Y_k = 5\tanh(X_k) + \zeta_k$ or $Y_k = 5\tanh(2\pi X_k) + \zeta_k$. Figure 5 shows the averaged RMSE over 2000 cycles, as a function of ensemble size for the four cases. The optimal diversity that has the smallest RMSE among several chosen diversity intervals was used for Figure 5. As shown in this figure, the performance of the EnKPF is comparable with that of the EnKF if the ensemble size is relatively small, especially for a strongly nonlinear measurement function. However, if the ensemble size is greater than 64, both the nEnKPF and the mEnKPF significantly outperformed the EnKF.

It should be noted that the difference between the nEnKPF and the mEnKPF appears insignificant in this experiment. Actually, the nEnKPF is slightly better than mEnKPF if the ensemble size is very small. The reason for this difference is most likely that the modified Kalman gain equation (35) in the mEnKPF assumes the ensemble mean $\overline{x^f}$ to be the true value, an assumption that holds well at large ensemble sizes.

Next, we consider the Lorenz '96 model with $J = 40$ variables and the observations are available for all variables in assimilation. Similar to the above experiments with $J = 10$, the assimilation frequency was set at 8 or 12 model steps, and the measurement function was $Y_k = 5\tanh(X_k) + \zeta_k$. Table 3 shows the averaged RMSE over 2000 cycles for the optimal diversity value τ as a function of the ensemble size N. As Table 3 shows, both the nEnKPF and the mEnKPF offer an advantage if the ensemble size exceeds 256. This finding is consistent with the results of FK2013 for linear measurement functions. Those results showed that an ensemble size of $N = 400$ was required to show the advantage of the EnKPF. Again, the performance of the nEnKPF and the mEnKPF was similar, but the mEnKPF was slightly better than the former for a large ensemble.

It is clearly recognized that if the model is high dimensional, it is not easy for the PF to outperform the EnKF due to the computational costs of the PF. For the basic PF, the weight for each member is a scalar, and members are duplicated or removed according to those scalars. The diversity among the ensembles will decrease very rapidly during the data assimilation cycles. Although many efforts have been made to maintain the diversity among ensembles, the weights are still in the form of a scalar, as in equation (20). A localization scheme is not available for particle filters due to the global algorithm of the PF [*van Leeuwen*, 2009] which further increases the difficulty of increasing its ensemble diversity. For a high-dimensional system such as a GCM model, a large population of ensembles is required to make the particle filter effective. The

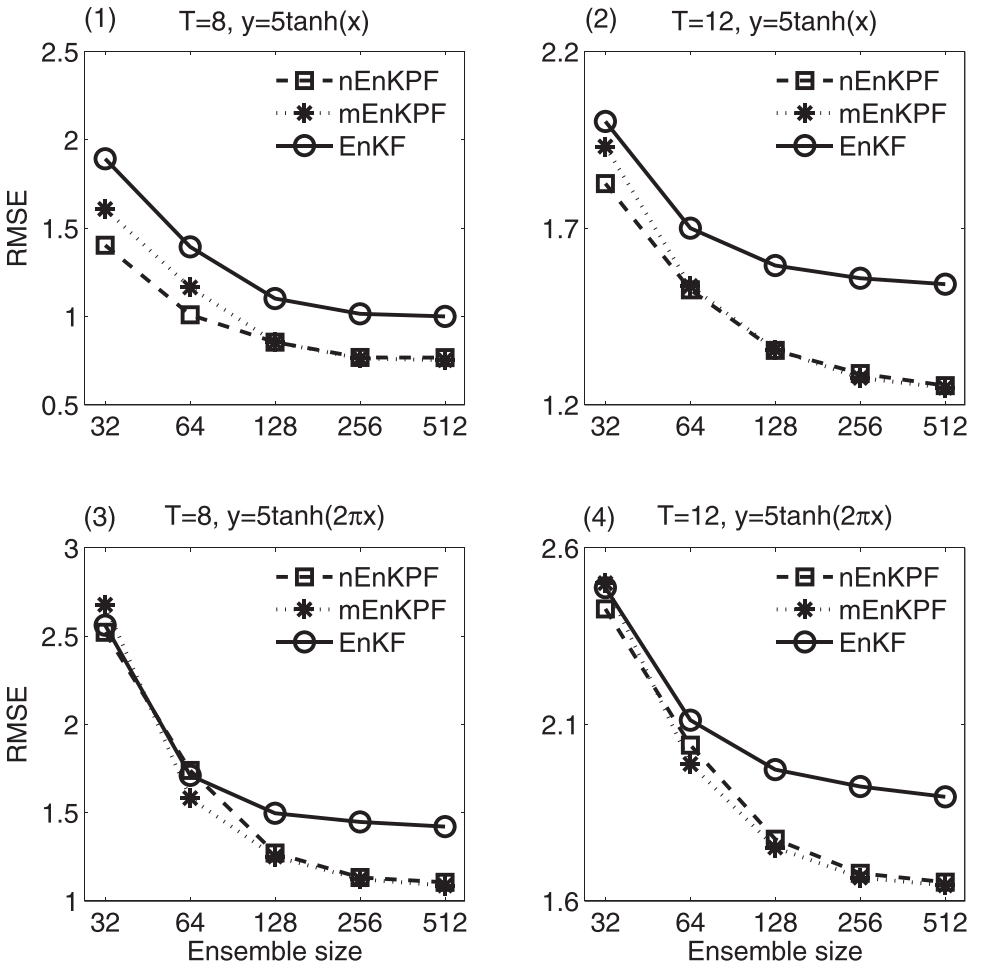

Figure 5. The averaged RMSE of all variables of the Lorenz '96 model with $J = 10$ over 2000 assimilation cycles for the nEnKPF, mEnKPF, and EnKF as a function of the ensemble size, with different data frequencies and measurement functions.

EnKPF, although its weight values can be substantial as γ tends to 1, still requires many ensembles to realize its advantage. Significant additional effort will be necessary before the EnKPF can be applied to a realistic GCM model.

Another concern regarding the mEnKPF is its associated computational burden. In the adaptive γ algorithm of the EnKPF, the Kalman gain and the weights of particles require recursive computations to search for the optimal value of γ. Equations (28), (29), and (30) or equations (35), (29), and (30) are often run several times in an assimilation cycle. This procedure differs from that employed by the EnKF, which only runs equations (16) and (17) once. Usually, the γ value is changed at the multiples of $1/2^{k+1}$ at the kth search step, and the binary search tree algorithm is used. Thus, we can set a reasonable value of the maximum iterative steps to determine the γ value in the adaptive algorithm.

Table 3. The Average RMSE Over 2000 Cycles for the Lorenz '96 Model With $J = 40$

N		32	64	128	256	512	1024
T = 8	nEnKPF	4.04	3.19	1.76	1.23	1.09	1.07
	mEnKPF	4.07	3.15	1.85	1.21	1.07	1.06
	EnKF	4.06	3.16	1.89	1.3	1.2	1.18
T = 12	nEnKPF	3.7	2.73	2.01	1.79	1.7	1.64
	mEnKPF	3.7	2.72	2.03	1.75	1.67	1.62
	EnKF	3.75	2.8	2.04	1.82	1.78	1.7

Note that the prediction step is often much more costly than the analysis step for high-dimensional systems because the prediction step requires the forward integration of a high-dimensional model. Thus, the computational cost of the EnKPF may be slightly greater than that of the EnKF but much less than that of the IEnKF.

4. Conclusions

Significant progress has been made in the field of data assimilation in recent years. However, it is still a great challenge to address nonlinear, non-Gaussian estimation problems. At present, the Kalman-based filters use a Gaussian pdf to approximate a non-Gaussian pdf. This approach is inefficient and can result in large estimation errors. The particle filter, which directly approximates the pdf using a combination of finite samples (particles) with the weights updated by the likelihoods, provides a complete image that serves to characterize all error statistics of the model states. However, a substantial difficulty associated with the particle filter is the high cost of the extensive computations that are required to prevent filter degeneracy. Under present conditions, these costs appear impractical for a high-dimensional model such as a GCM. The EnKPF was developed to combine the ensemble Kalman filter and particle filter, namely, assimilating the Gaussian information with the EnKF and the non-Gaussian information with the PF. This approach has the strengths of both the EnKF and the PF and produce a filter that reflects the trade-off between estimation accuracy and affordable cost. The results of the current study show that the EnKPF outperforms both the EnKF and the SIR-PF at the expense of additional computations in the analysis step.

Initially, the EnKPF was proposed to address assimilation systems with linear measurement functions. In this study, we extend the algorithm of EnKPF to nonlinear measurement functions in a rigorous statistical framework, and also introduce a new formula for the Kalman gain to further improve the EnKPF. Two numerical models, the Lorenz '63 and '96 models, are used to test the new algorithms. The assimilation experiments show that if the measurement function is nonlinear, both the nEnKPF and the mEnKPF can still provide more accurate estimates than the EnKF, as the original EnKPF did for linear cases. Thus, this modified algorithm further supports the application of the EnKPF.

In this study, we explored the modified EnKPF using a highly simplified nonlinear model with a designed nonlinear measurement function. One concern is the performance and efficiency of the mEnKPF when applied to a realistic GCM. Additional studies (e.g., localization in the EnKF, the formulation of weights, non-Gaussian likelihood functions) are needed for better implementation of these techniques applied to data assimilation problems in atmospheric or ocean GCMs. Nevertheless, the present study represents a step in the pursuit of advanced data assimilation algorithms based on a simple nonlinear model that shares several common features with complicated atmospheric and oceanic models.

Acknowledgments

This work was supported by the National Science Foundation of China (41276029, 41321004), the scientific research fund of the Second Institute of Oceanography of China (14202), and the State Key Laboratory of Satellite Ocean Environment Dynamics (14085G). The authors would also like to acknowledge and thank the Oregon Graduate Institute, Eric A. Wan, and Rudolph Van der Merwe for providing the ReBEL (Recursive Bayesian Estimation library, 2003, available at http://choosh.csee.ogi.edu/rebel/index.html) tool kit, part of which has been used in this research work.

References

Ades, M., and P. Van Leeuwen (2013), An exploration of the equivalent weights particle filter, *Q. J. R. Meteorol. Soc.*, *139*(672), 820–840.

Ambadan, J. T., and Y. Tang (2009), Sigma-point Kalman filter data assimilation methods for strongly nonlinear systems, *J. Atmos. Sci.*, *66*(2), 261–285.

Anderson, J. L. (2001), An ensemble adjustment Kalman filter for data assimilation, *Mon. Weather Rev.*, *129*(12), 2884–2903.

Anderson, J. L. (2009), Ensemble Kalman filters for large geophysical applications, *IEEE Control Syst.*, *29*(3), 66–82.

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Process.*, *50*(2), 174–188.

Bengtsson, T., C. Snyder, and D. Nychka (2003), Toward a nonlinear ensemble filter for high-dimensional systems, *J. Geophys. Res.*, *108*(D24), 8775, doi:10.1029/2002JD002900.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar (2001), Adaptive sampling with the ensemble transform Kalman filter. Part i: Theoretical aspects, *Mon. Weather Rev.*, *129*(3), 420–436.

Bocquet, M., C. A. Pires, and L. Wu (2010), Beyond gaussian statistical modeling in geophysical data assimilation, *Mon. Weather Rev.*, *138*(8), 2997–3023.

Burgers, G., P. J. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, *126*(6), 1719–1724.

Cappé, O., S. J. Godsill, and E. Moulines (2007), An overview of existing methods and recent advances in sequential Monte Carlo, *Proc. IEEE*, *95*(5), 899–924.

Chorin, A. J., M. Morzfeld, and X. Tu (2010), Implicit particle filters for data assimilation, *Commun. Appl. Math. Comput. Sci.*, *5*(2), 221–240.

Doucet, A., N. De Freitas, and N. Gordon (2001), *Sequential Monte Carlo Methods in Practice*, vol. 1, Springer, N. Y.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162.

Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, *53*(4), 343–367.

Frei, M., and H. R. Künsch (2013), Bridging the ensemble Kalman and particle filters, *Biometrika*, *100*(4), 781–800.
Houtekamer, P. L., and H. L. Mitchell (2001), A sequential ensemble Kalman filter for atmospheric data assimilation, *Mon. Weather Rev.*, *129*(1), 123–137.
Julier, S. J., and J. K. Uhlmann (1997), A new extension of the Kalman filter to nonlinear systems. Signal Processing, Sensor Fusion, and Target Recognition VI, edited by I. Kadar, International Society for Optical Engineering (SPIE Proceedings, vol. 3068), 182–193.
Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, *82*(1), 35–45.
Knuth, D. E. (2013), *Art of Computer Programming, vol. 4, Fascicle 4: Generating All Trees–History of Combinatorial Generation*, Addison-Wesley, Upper Saddle River, N. J.
Liu, J. S., (2008), *Monte Carlo Strategies in Scientific Computing*, Springer, N. Y.
Lorenz, E. N. (1963), Deterministic nonperiodic flow, *J. Atmos. Sci.*, *20*(2), 130–141.
Lorenz, E. N. (1996), Predictability: A problem partly solved, Proc. ECMWF Seminar on Predictability, vol. 1, ECMWF, 1–18, Reading, U. K.
Luo, X., and I. Hoteit (2014), Efficient particle filtering through residual nudging, *Q. J. R. Meteorol. Soc.*, *140*(679), 557–572.
Miller, R. N., M. Ghil, and F. Gauthiez (1994), Advanced data assimilation in strongly nonlinear dynamical systems, *J. Atmos. Sci.*, *51*(8), 1037–1056.
Morzfeld, M., X. Tu, E. Atkins, and A. J. Chorin (2012), A random map implementation of implicit filters, *J. Comput. Phys.*, *231*(4), 2049–2066.
Nakano, S., G. Ueno, and T. Higuchi (2007), Merging particle filter for sequential data assimilation, *Nonlinear Processes Geophys.*, *14*, 395–408, doi:10.5194/npg-14-395-2007.
Papadakis, N., E. Mémin, A. Cuzol, and N. Gengembre (2010), Data assimilation with the weighted ensemble Kalman filter, *Tellus, Ser. A*, *62*(5), 673–697.
Rezaie, J., and J. Eidsvik (2012), Shrinked (1- α) ensemble Kalman filter and α Gaussian mixture filter, *Comput. Geosci.*, *16*(3), 837–852.
Sakov, P., D. S. Oliver, and L. Bertino (2012), An iterative EnKF for strongly nonlinear systems, *Mon. Weather Rev.*, *140*(6), 1988–2004.
Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson (2008), Obstacles to high-dimensional particle filtering, *Mon. Weather Rev.*, *136*(12), 4629–4640.
Tang, Y., J. Ambandan, and D. Chen (2014), Nonlinear measurement function in the ensemble Kalman filter, *Adv. Atmos. Sci.*, *31*(3), 551–558.
Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker (2003), Ensemble square root filters*, *Mon. Weather Rev.*, *131*(7), 1485–1490.
van Leeuwen, P. J. (2009), Particle filtering in geophysical systems, *Mon. Weather Rev.*, *137*(12), 4089–4114.
van Leeuwen, P. J. (2010), Nonlinear data assimilation in geosciences: An extremely efficient particle filter, *Q. J. R. Meteorol. Soc.*, *136*(653), 1991–1999.
van Leeuwen, P. J. (2011), Efficient nonlinear data assimilation in geophysical fluid dynamics, *Comput. Fluids*, *46*(1), 52–58.
Whitaker, J. S., and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Mon. Weather Rev.*, *130*(7), 1913–1924.
Xiong, X., I. M. Navon, and B. Uzunoglu (2006), A note on the particle filter with posterior Gaussian resampling, *Tellus, Ser. A*, *58*(4), 456–460.