A Local Particle Filter Using Gamma Test Theory for High-Dimensional State Spaces

Wang, Zhenwu; Hut, Rolf; Van de Giesen, Nick

DOI
10.1029/2020MS002130

Publication date
2020

Document Version
Final published version

Published in
Journal of Advances in Modeling Earth Systems

Citation (APA)
Wang, Z., Hut, R., & Van de Giesen, N. (2020). A Local Particle Filter Using Gamma Test Theory for High-Dimensional State Spaces. *Journal of Advances in Modeling Earth Systems*, *12*(11), 1-16. Article e2020MS002130. https://doi.org/10.1029/2020MS002130

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002130

A Local Particle Filter Using Gamma Test Theory for High-Dimensional State Spaces

Zhenwu Wang[1], **Rolf Hut**[1], **and Nick Van de Giesen**[1]

[1]Water Resources Engineering, Delft University of Technology (TU Delft), Delft, Netherlands

Key Points:

- The new local particle filters based on the Gamma test theory is proposed
- The Gamma test theory is used to account for uncertainty in the process of data assimilation
- This algorithm has a better performance in the nonlinear case

Correspondence to:
Z. Wang,
zhenwu.wang@tudelft.nl

Citation:
Wang, Z., Hut, R., & Van de Giesen, N. (2020). A local particle filter using Gamma test theory for high-dimensional state spaces. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002130. https://doi.org/10.1029/2020MS002130

Received 17 JUN 2020
Accepted 6 OCT 2020
Accepted article online 14 OCT 2020

Abstract Particle filters are non-Gaussian filters, which means that the assumption that the error distribution of the ensemble should be Gaussian is unnecessary. Like the ensemble Kalman filter, particle filters are based on the Monte Carlo approximation to represent the distribution of model states. It requires a substantial number of particles to approximate the probability density function of states in high-dimensional models, which is prohibitive for real applications. In order to overcome problems with high dimensionality, localization was applied in an Ensemble-type data assimilation system. This study combines the localization in LETKF (Local Ensemble Transformation Kalman Filter) with particle filters and proposes a new local particle filter with the model state space correction using Gamma test theory for high-dimensional models. A series of tests with various parameter settings, including different the numbers of particles, observation intervals, localization scale, inflation factors, and observation operators, were used to evaluate the performance of this new method using a Lorenz model with 40 variables. Besides, the proposed filter was applied in the Lorenz model with 1,000 variables to evaluate its performance in the model with higher dimensions. The results show that this approach can deal with the issue of dimensionality, which otherwise leads to the collapse of the particle filters in high-dimensional systems. The local particle filter is stable and has considerable potential for complex higher-dimensional models.

Plain Language Summary Particle filters (PFs) are non-Gaussian filters, which means that the error distribution does not have to be Gaussian. This is an important advantage over ensemble Kalman filters(EnKF), which does assume a Gaussian distribution. Similar to EnKF, PFs need some particles to represent the distribution of model states. However, in high-dimensional systems, the need for particles is exceptionally large for standard PFs, which limits PF applications for real-world problems. To reduce the number of particles need, localization is a possible solution. We propose a new PF combining localization and the Gamma test theory, which can be used for high-dimensional systems. The performance of the proposed algorithm was tested using a Lorenz model with 40 variables, and a set of experiments was conducted. One more test was used to explore the potential of the proposed PFs in the Lorenz model with 1,000 variables. Results show that this new algorithm is stable in high-dimensional systems, which makes it potentially useful for more complex geophysical models in the real world.

1. Introduction

Numerical models are used to forecast and estimate model states in many fields including meteorology, ocean science, and hydrology. Accounting for different sources of uncertainty to improve the accuracy of such models has drawn considerable attentions in recent decades (Liu et al., 2012). Data assimilation (DA) provides a solution to combine information from model estimations and observations to achieve better prediction performance and quantify uncertainty (Bannister, 2017; Hut et al., 2015; van Leeuwen, 2015). DA has been applied widely in geoscience (Abolafia-Rosenzweig et al., 2019; Choi et al., 2017; Dong et al., 2015; Fang & Li, 2019; Fox et al., 2018; Irrgang et al., 2017; Jin et al., 2018, 2019; Zhang et al., 2017).

The ensemble Kalman filter (EnKF) and its variants have been applied frequently (Chen et al., 2013; Sun et al., 2016; Xie & Zhang, 2010). EnKF relies on the assumption that the error distribution is Gaussian (Katzfuss et al., 2016). Particle filters (PFs) are another ensemble-type data assimilation algorithm, which is proposed by Gordon et al. (1993) and has been used in many low-dimensional models (van Leeuwen, 2009). PFs, just like the EnKF, use a Monte Carlo approximation in which a certain number of particles are used to represent the distribution of model states. The distribution reflects the mean and spread of model states

and is updated by the prior weights, which can be calculated using the likelihood given to each particle. However, unlike EnKF and its variants, PFs do not rely on the assumption that the error distribution is Gaussian. It should be noted that the Gaussian assumption can lead to suboptimal results when a model system is nonlinear and errors are non-Gaussian, which is the main limitation of ensemble-Gaussian-type data assimilation strategies.

For PFs, the increasing number of dimensions of a model requires an exponentially growing number of particles to avoid filter divergence, as shown by Snyder et al. (2008), which is called the curse of dimensionality (Bengtsson et al., 2008). Bengtsson et al. (2008) also showed that for accurate representation of high-dimensional distributions, the number of particles needs to be increased exponentially with the number of model states. Therefore, the use of PFs in high-dimensional models is limited due to the curse of dimensionality and associated high demand on computational resources.

Currently, there are several strategies to deal with the dimensionality of PFs in high-dimensional systems (Pinheiro et al., 2019; Potthast et al., 2018). An equivalent weights PF has been proposed, which prevents filter degeneracy using the proposal transition density (Ades & van Leeuwen, 2013; van Leeuwen, 2010). The proposal transition density keeps particles close to observations. Consequently, it leads to a better statistic representation of the posterior distribution because none of particles is ignored. Ades and van Leeuwen (2014) explored the effect of the equivalent weights PF on the dynamical balance in a primitive equation model. Their results showed that this method had potential for large-scale geophysical applications. Ades and van Leeuwen (2015) used this method in the barotropic vorticity model with 65,500 states and the results illustrated its powerful abilities to make the filter stable and avoid filter divergence.

Localization is an useful method for preventing the collapse of PFs in high-dimension models, which is used commonly to solve the high-dimensional issues in EnKF and its variants, for example, the LETKF (Hunt et al., 2007). Researchers have attempted to apply localization in PFs to reduce the impact of dimensionality constraints. van Leeuwen (2003) and Bengtsson et al. (2003) tried to combine localization with PFs. Farchi and Bocquet (2018) reviewed several local PFs from both theoretical and practical aspects. In this review, the localization solutions were divided into two strategies, which were state-domain localization and sequential-observation localization (Farchi & Bocquet, 2018).

State-domain localization is used to update each model state independently only using the observations within localization scales, which has been applied in PFs (Chustagulprom et al., 2016; Lee & Majda, 2016; Penny & Miyoshi, 2015; Rebeschini & van Handel, 2015). Penny and Miyoshi (2015) introduced a localized PF (LPF) with state-domain localization. The particles of this algorithm were updated first by a transformation matrix and the final updated particles were a linear combination of particles around each point. This LPF could outperform the LETKF under some conditions (Penny & Miyoshi, 2015). However, this method still needs an inflation method, which is a common method in data assimilation to maintain the spread of the model states to prevent collapse of the algorithm and to keep the filter stable.

In sequential-observation localization, all observations are assimilated one by one. Only nearby model states at each observation site are updated and influenced and distant model states stay unchanged (Farchi & Bocquet, 2018). Poterjoy (2016) proposed a LPF with sequential-observation localization, which can be applied in high-dimensional systems using a small number of particles. It extended the weights in traditional PFs to a vector to remove the impact of distant observations. Kernel density distribution mapping (KDDM) was applied to the updated particles to obtain a desired posterior distribution, which was selected by prior particles and their weights. However, this method still suffered imbalance issues caused by localization. Poterjoy and Anderson (2016) used an idealized atmospheric general circulation model (GCM) to test physical consistency of posterior particles and scalability. The results of these two different tests showed the LPF proposed by Poterjoy (2016) is highly likely to work in more complex real models.

This study provides a new solution to prevent the divergence of PFs in high-dimensional models. We attempt to consider the impact of the uncertainty between prior and posterior particles. Our new method is based on the Bayesian theorem and combines the standard PFs with the so-called state-domain localization. Besides, the Gamma test is used to correct uncertainty in the posterior distribution. Localization divides all model states into local patch vectors and each model state has its own local patch vector (Penny & Miyoshi, 2015). In this way, PFs only assimilate observations within the localization scale. The Gamma test is a technique, which can be used to estimate uncertainty between prior and posterior model state space

(Evans & Jones, 2002), which can lessen impact of localization issues. Additionally, the proposed algorithm applies the Gamma test to modify the variance of the posterior uncertainty to stabilize the filter and avoid filter collapse. The main motivation for designing this local filter is to make PFs available and effective in high-dimensional systems with a comparatively small number of particles.

The manuscript is organized as follows. Section 2 describes the standard PFs and then the solutions to the curse of dimensionality, including localization methods and the Gamma test. Section 3 evaluates the performance of the proposed method using the Lorenz model with 40 and 1,000 variables and compares its results with the LETKF. In the final section, the limitations of the new filter and its possible applications for high-dimensional geophysical models are discussed.

2. Methodology

2.1. Standard PFs

This subsection briefly introduces standard PFs (Doucet et al., 2001). Particles are used to represent the distribution of model states, which can capture the mean and uncertainty of the model states. The particles are updated by a resampling algorithm based on weights that are calculated by the likelihood, given observations. The resampling method is crucial to PFs (Hol et al., 2006), and its basic idea is to modify prior particles to posterior ones by eliminating particles with smaller weights and by duplicating particles having larger weights. The residual resampling algorithm (Lui & Chen, 1998) was applied in this study, as being one of the most frequently used resampling algorithm in PF data assimilation (Hol et al., 2006; Hong et al., 2004; Zhang et al., 2013).

Let us assume that $\boldsymbol{x}$ represents a model state of a model which can be propagated over time and $\boldsymbol{y}$ represents a vector of observations. The model captures our incomplete knowledge of the physical system under consideration. Uncertainty in model states is inevitable due to imperfection of the model. Therefore, the model states, $\boldsymbol{x}$, can only approximate the truth and cannot reach the truth. For similar reasons, observations $\boldsymbol{y}$ are only approximations of the truth because of observation uncertainties. The relation between observations and true model states can be expressed as follows:

$$\boldsymbol{y} = \mathrm{H}\left(\boldsymbol{x}_{\mathbf{true}}\right) + \epsilon \tag{1}$$

where H is the observation operator that transforms model states into observation space and ϵ is the observation error. PFs are used to estimate $\boldsymbol{x}$ given observations $\boldsymbol{y}$ by a Bayes's theorem expansion using a Monte Carlo estimation:

$$p(\boldsymbol{x}\,|\boldsymbol{y}) = \frac{p(\boldsymbol{y}\mid \boldsymbol{x})p(\boldsymbol{x})}{\int p(\boldsymbol{y}\mid \boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}} \tag{2}$$

where $p\left(\boldsymbol{x}_{|}\boldsymbol{y}\right)$ is the probability of model states $\boldsymbol{x}$ given all observations $\boldsymbol{y}$. $p(\boldsymbol{x}|\boldsymbol{y})$ can be obtained by a Monte Carlo approach. By drawing N_p particles, denoted $\boldsymbol{x}_n(n = 1,2,3, \ldots, N_p)$, $p(\boldsymbol{x})$ can be constructed as a discrete set of delta functions centered on every individual particle:

$$p(\boldsymbol{x}) \approx \frac{1}{N_p}\sum_{n=1}^{N_p}\delta\left(\boldsymbol{x}-\boldsymbol{x}_n\right) \tag{3}$$

Similarly, $p(\boldsymbol{x}|\boldsymbol{y})$ can be approximated by using

$$p(\boldsymbol{x}|\boldsymbol{y}) \approx \sum_{n=1}^{N_p}\frac{\boldsymbol{w_n}}{W}\delta\left(\boldsymbol{x}-\boldsymbol{x}_n\right) \tag{4}$$

where w_n/W are the normalized weights, provided by

$$\boldsymbol{w_n} = p\left(\boldsymbol{y}|\boldsymbol{x}_n\right) \tag{5}$$

$$W = \sum_{n=1}^{N_p}\boldsymbol{w_n} \tag{6}$$

Generally, a resampling algorithm is used to generate the posterior distribution by discarding particles with low weight and duplicating high weight particles (Lui & Chen, 1998). The resampling process guarantees that sufficient particles remain.

2.2. Degeneracy of Filters

As introduced in section 2.1, PFs do not rely on estimating the error distributions and useful particles are chosen based on weights using a resampling method. The degeneracy of PFs refers to the situation in which the particle with the largest weight is the only particle chosen by the resampling algorithm (Poterjoy, 2016). When an increase in the number of model states is not matched with an increase in the number of particles (Penny & Miyoshi, 2015; Snyder et al., 2008), it becomes difficult for PFs to find enough particles with sufficiently high probabilities. Increasing the number of particles can reduce the degeneracy of PFs. Snyder et al. (2008) analyzed the relationship between the number of particles and the number of model states theoretically. They indicated that the number of particles must increase exponentially with the dimensions of a model to obtain a posterior mean, which has a smaller error than the prior particles or observations. LETKF has a different fundamental reason why the filter collapses. The main cause of the collapse of PFs is that the weights of each model state are close to unity. For LETKF, just like other EnKF variants, the classic divergence of a filter is indicated by the decreasing or increasing spread of ensemble (Houtekamer & Zhang, 2016). The reason for this can be model errors, sample errors, nonlinearity in the system, and other uncertainties.

2.3. Localization Method

The localization method, which was proposed by Hunt et al. (2007), is commonly used to remove spurious error covariance outside the local scale due to sampling errors caused by too small an ensemble. It only assimilates observations within a given scale for each particle, which is efficient for high-dimensional systems. The local PFs in this study use a localization scheme inspired by the LETKF. Therefore, the localization in LETKF is applied.

Using a localization method can stabilize the filter and avoid filter degeneracy in EnKF and its variants when only a small number of particles is used in the data assimilation of high-dimensional models (Hunt et al., 2007). For the issues in PFs caused by dimensionality (Snyder et al., 2008), as mentioned and explained in section 2.2, localization has been applied to solve the dimensionality of PFs (Poterjoy, 2016; Penny & Miyoshi, 2015). However, the localization method can deal with the issue partially and additional methods are still needed to stabilize the posterior distribution. Poterjoy (2016) corrected updated ensemble by using kernel density distribution mapping, and Penny and Miyoshi (2015) still used an inflation method to prevent the filter collapse. As in LETKF, every model state is assimilated one by one in PFs and particles and observations are localized by a localization function to form local particles $\boldsymbol{x}_{loc}$ and local observations $\boldsymbol{y}_{loc}$. Then the weights of each state are calculated based on $\boldsymbol{y}_{loc} - \mathrm{H}\boldsymbol{x}_{loc}$. In this study, the localization method in LETKF is used for PFs to remove certain observations that fall outside the localization scale. The specific localization function used here for local PFs was proposed by Gaspari and Cohn (1999).

2.4. Gamma Test

The Gamma test is normally used to estimate the variance of the noise in a given data set, which is used to build the best smooth model, without knowledge of the specific model form. In this work, we will use the Gamma test to estimate the variance of the uncertainty between prior and posterior model states to correct posterior model states. In doing so, we assume that the uncertainty in particles is maintained. Only a brief introduction to the Gamma test theory is given in this subsection and further details can be found in corresponding papers (Evans & Jones, 2002; Jones et al., 2007). Let us assume we have prior particles $\boldsymbol{x}^{prior}$ and updated posterior particles $\boldsymbol{x}^{post}$:

$$\left\{ \left(\boldsymbol{x}_i^{prior}, \boldsymbol{x}_i^{post} \right) | 1 \le i \le N_p \right\} \tag{7}$$

where N_p is the number of particles.

In data assimilation, the prior particles are updated by observations and this process can be interpreted as a "data assimilation model" to generate the output-posterior particles according to the input-prior particles. Because of various uncertainty sources in data assimilation, such as the uncertainties caused by assumptions about observation error, forward operator error, and observation bias (Houtekamer & Zhang, 2016), there always is an uncertainty in this "data assimilation model" that cannot be estimated. In this research, the variance of the uncertainty is estimated by the Gamma test. To fit the Gamma test, the relationship between $\boldsymbol{x}^{prior}$ and $\boldsymbol{x}^{post}$ is expressed as follows:

$$\boldsymbol{x}^{post} = F_{DA}\left(\boldsymbol{x}^{prior} \right) + \boldsymbol{r}_{DA} \tag{8}$$

where F_{DA} represents a data assimilation process and $\boldsymbol{r}_{DA}$ are the errors with expectation 0. PFs are a non-Gaussian type of filter because when calculating weights of particles, the error distribution does not have to be Gaussian, but we still need to know the specific probability density function of the error distribution. However, in a nonlinear case, we would not be able to determine the error distribution. The Gamma test can estimate the variance of the noise $\mathrm{var}\left(\boldsymbol{r}_{DA}\right)$ regardless of the specific data assimilation algorithm used and the underlying error distribution because of the existence of nonlinearity and non-Gaussianity.

The Gamma test statistic is calculated by the following procedure. First, the Gamma test uses a kd-tree to find the kth ($1 \le k \le p$) nearest neighbors $\boldsymbol{x}_k^{prior}, \boldsymbol{x}_k^{post}$ of $\boldsymbol{x}^{prior}, \boldsymbol{x}^{post}$ for each particle member. Here p is set to 10 typically (Jones et al., 2007). Next, the algorithm computes

$$\delta(k) = \frac{1}{N_m}\sum_{i=1}^{N_m}\left|x_k^{prior} - x_i^{prior}\right|^2 \qquad (1 \le k \le p) \tag{9}$$

$$\gamma(k) = \frac{1}{2N_m}\sum_{i=1}^{N_m}\left|x_k^{post} - x_i^{post}\right|^2 \qquad (1 \le k \le p) \tag{10}$$

where $|\ldots|$ denotes Euclidean distance, $\delta(k)$ is the mean square of the k nearest neighbors of the prior distribution, and $\gamma(k)$ is derived from the k nearest neighbors of the posterior distribution, which is defined as the outcome of the Gamma test. Based on the points $(\delta(k), \gamma(k))$, the linear regression $\gamma(k) = A\delta(k) + \Gamma$ is computed and the intercept Γ is the estimation of $\mathrm{var}(r_{DA})$, as can be shown $\gamma(k) \to \mathrm{var}\left(r_{DA}\right)$ in probability as $\delta(k) \to 0$.

2.5. Local PFs With the Gamma Test

The local PF with the Gamma test (LPF-GT) is explained in this subsection. The foundation of PFs is still Bayes's theorem and the Monte Carlo method. With the localization procedure, each model state can be updated independently in DA. The Gamma test provides an estimation of potential uncertainty for DA. Under the assumption that the observation errors are independent, weights can be calculated by

$$\boldsymbol{w}_n = p\left(\boldsymbol{y}|\boldsymbol{x}_n\right) = \prod_{i=1}^{N_{obs}} p\left(y_i|\boldsymbol{x}_n\right) \tag{11}$$

where N_{obs} is the number of observations. When calculating weights in PFs, and when only a few particles carry almost all the weight, filter collapse is inevitable. Therefore, in order to avoid filter collapse, the probability $p\left(y_i|\boldsymbol{x}_n\right)$ in Equation 11 is calculated in the following way:

$$p_\beta\left(y_i|\boldsymbol{x}_n\right) = \left[p\left(y_i|\boldsymbol{x}_n\right) * loc\left(y_i, \boldsymbol{x}_n, r\right) + \beta_a\right]\beta_m \tag{12}$$

where $loc\left(y_i, \boldsymbol{x}_n, r\right)$ are the localization coefficients and * represents elementwise product. In current research, we use (4.10) of Gaspari and Cohn (1999) for $loc\left(y_i, \boldsymbol{x}_n, r\right)$, which has a Gaussian-type structure with a width r. The parameters β_a and β_m are used to control weights. Hence, after replacing $p\left(y_i|\boldsymbol{x}_n\right)$ with $p_\beta\left(y_i|\boldsymbol{x}_n\right)$, the localized weights in LPF-GT are given by

$$\begin{aligned}\boldsymbol{w}_n &= \prod_{i=1}^{N_{obs}^{loc}} p_\beta\left(y_i|\boldsymbol{x}_n\right) \\ &= \prod_{i=1}^{N_{obs}^{loc}} \left[p\left(y_i|\boldsymbol{x}_n\right) * loc\left(y_i, \boldsymbol{x}_n, r\right) + \beta_a\right]\beta_m\end{aligned} \tag{13}$$

where N_{obs}^{loc} indicates the number of observations within the localization scale. The β_a is an additive factor, and β_m is a multiplicative factor.

The mean effective number of particles N_{eff} (shown in Equation 14) is used to evaluate the quantity of the particles (Penny & Miyoshi, 2015) and the factor β_a and β_m in Equation 13 can tune the value of N_{eff}.

$$N_{eff} = \left[\sum_{i=1}^{N_p}\left(\frac{w_i}{\sum_{i=1}^{N_p} w_i}\right)^2\right]^{-1} \tag{14}$$

The role of β_a and β_m is mainly to avoid filter degeneracy. Because they influence the value of N_{eff}. Tuning N_{eff} can change the percentage of removed particles in the resampling procedure, which was discussed in the next section. We found that keeping N_{eff} close to a certain value by tuning β_a and β_m can avoid filter collapse. To fix N_{eff}, values of β_a and β_m are changed dynamically by finding appropriate values in the parameter space of β_a and β_m. There are definitely more than more than one pair of β_a and β_m that fix N_{eff}. We simply stopped the search when the first pair of these two parameters had been found.

Equation 13 attempts to avoid the divergence of PFs by rescaling the weights of each particle twice. The weights of distant observations are reduced gradually by introducing localization coefficients to the probabilities of local observation errors. The factors β_a and β_m are used to maintain the stability of the weights, which can control the proportion of particles removed by residual resampling algorithm.

Next, the modified particles $\boldsymbol{x}_n^a$ and the factor α corrected by the Gamma test are given in Equations 15 and 19, respectively.

$$\boldsymbol{x}_n^a = \overline{\boldsymbol{x}_n^{a'}} + \eta\left(\boldsymbol{x}_n^{a'} - \overline{\boldsymbol{x}_n^{a'}}\right) + (1-\eta)\,\alpha\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right) \tag{15}$$

where $\boldsymbol{x}_n^{a'}$ indicates the updated particles generated by a resampling method. The prime indicates a posterior, or updated, particle. And $\overline{\boldsymbol{x}_n^{a'}}$ is its mean. $\boldsymbol{x}_n^b$ denotes the prior particles. In Equation 15, $\alpha\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right)$ represents the uncertainty $\boldsymbol{r}_{DA}$ brought by data assimilation, which is shown in Equation 8. We assume that $\boldsymbol{r}_{DA}$ follows a Gaussian distribution with mean 0 and its variance is estimated by the Gamma test. Using $\boldsymbol{x}_n^{a'}$ to minus $\boldsymbol{x}_n^b$ directly, we obtain a sample, which is not the error we expect. To achieve our goal, the sample $\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right)$ needs to be normalized, which is given as $\boldsymbol{X}$. Next, we set the variance of $\boldsymbol{X}$ equal to Γ estimated by the Gamma test. According to the definition of the variance, we define Γ and normalized var $\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right)$ as follows.

$$\Gamma = \frac{\sum(\alpha\boldsymbol{X} - \mu)^2}{N} \tag{16}$$

$$\text{var}\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right) = \frac{\sum(\boldsymbol{X} - \mu)^2}{N} \tag{17}$$

where Γ is the desired variance and α refers to the corrected factor of the normalized sample $\boldsymbol{X}$. The mean of $\boldsymbol{X}$ is 0. N is the number of samples. We can obtain α in Equation 15 based on Equations 18 and 19.

$$\frac{\Gamma}{\text{var}\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right)} = \frac{\sum(\alpha\boldsymbol{X} - \mu)^2}{\sum(\boldsymbol{X} - \mu)^2} = \frac{\alpha^2\sum(\boldsymbol{X})^2}{\sum(\boldsymbol{X})^2} = \alpha^2 \tag{18}$$

$$\alpha = \sqrt{\frac{\Gamma}{\text{var}\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right)}} \tag{19}$$

In Equation 15, the uncertainty of posterior particles is a linear combination of the particles updated by PFs $\left(\boldsymbol{x}_n^{a'} - \overline{\boldsymbol{x}_n^{a'}}\right)$ and particles modified by the Gamma test $\alpha\left(\boldsymbol{x}_n^{a'} - \boldsymbol{x}_n^b\right)$. The mean of $\boldsymbol{x}_n^a$ is still obtained using the resampling algorithm. In this research, we assume that the uncertainty of $\boldsymbol{x}_n^a$ consists of two parts. One part is comes from the resampling method. The other part comes from the data assimilation framework, which can be estimated by the Gamma test. These two parts are combined by η and the value of η is between 0 and 1. The parameter η can be tuned to change the impact on the uncertainty of $\boldsymbol{x}_n^a$.

3. Numerical Experiments and Results

In this research, the Lorenz model (Lorenz, 1996; L96 model) was used to evaluate our proposed local PFs (LPF-GT) and to compare its performance to LETKF. More detailed introduction to the L96 model can be found in Lorenz (1996). As is common in testing new data assimilation schemes for high dimensions models, in this study, 40 variables are chosen and F remains fixed at 8.0 to maintain the chaotic behavior in L96. Furthermore, the differential equations of the L96 model are integrated by a fourth-order RungeKutta method with a 0.05 time step (Lorenz, 1996 which is defined as 6h).

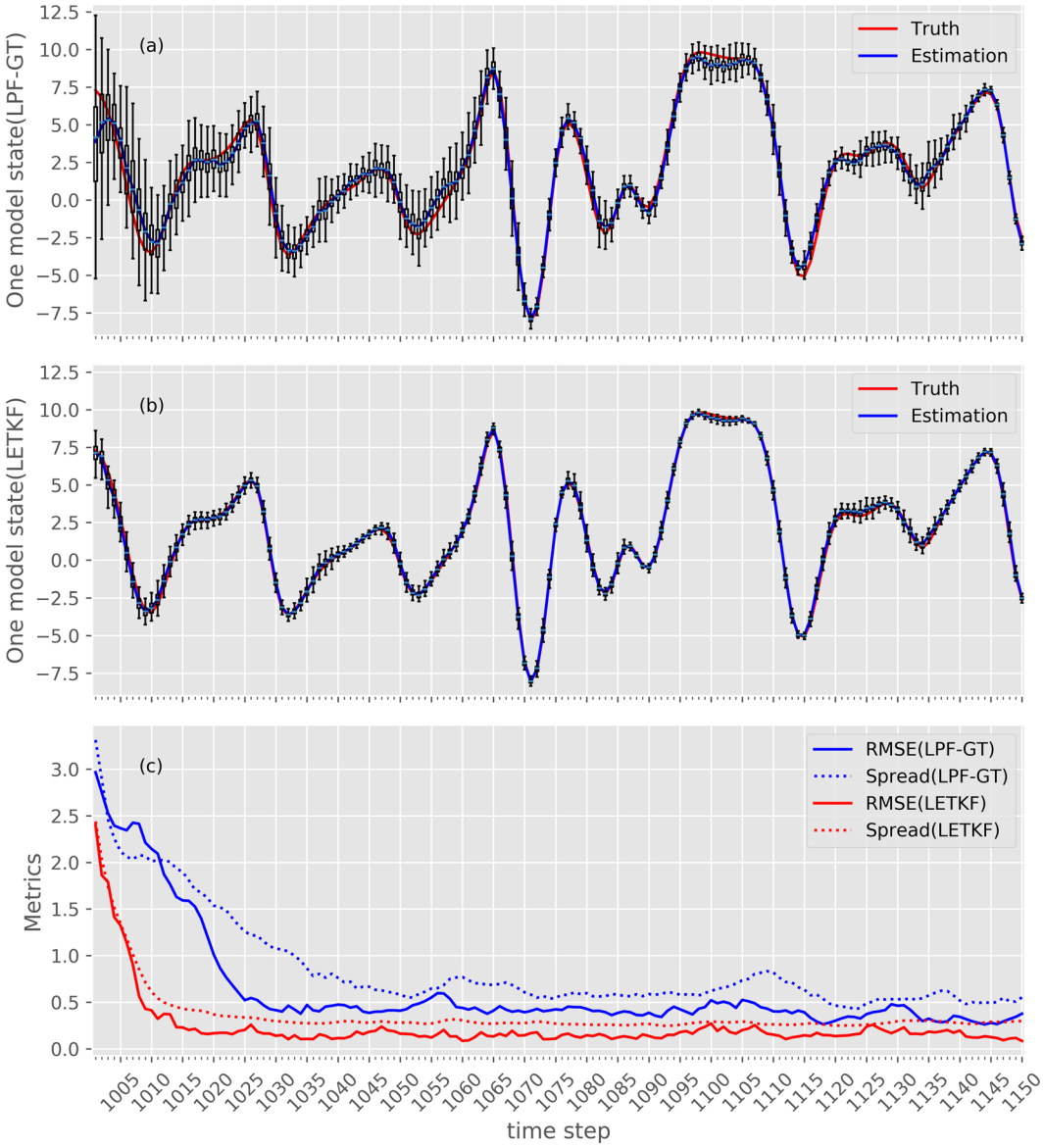

Figure 1. Trajectories of (a) the first model state using the LPF-GT, (b) the first model state using the LETKF, and (c) RMSE and spread for both methods. In the first two figures, the blue curve is the truth and the red curve is the estimation of the model state. In the third one, the solid lines indicate the RMSE and the dash lines is the spread. The assimilation starts at time step 1,000.

3.1. Experimental Setup

A set of experiments were conducted to test the validity of LPF-GT for various parameter configurations that mimic real applications. In these experiments, LPF-GT and LETKF were examined by a variety of parameter settings to test the effectiveness and disadvantages of LPF-GT. The parameters in these experiments include various observation intervals, the number of particles N_p, the number of observations N_{obs}, inflation factors, localization scales, and two different observation operators H.

The default configuration consists of the linear H operator. Only half of the model states are observed, 20 observations, were assimilated in every experiment which were chosen evenly and were fixed spatially over time. An inflation methods was used for LETKF, and the inflation factor had been tuned to a fixed value 1.05 as the default. The parameters β_a and β_m in LPF-GT have a similar role as in the inflation method. The influence to N_{eff} by tuning β_a and β_m was investigated through numerical experiments. The default number of particles was 100. Observations were derived from the truth with an uncertainty

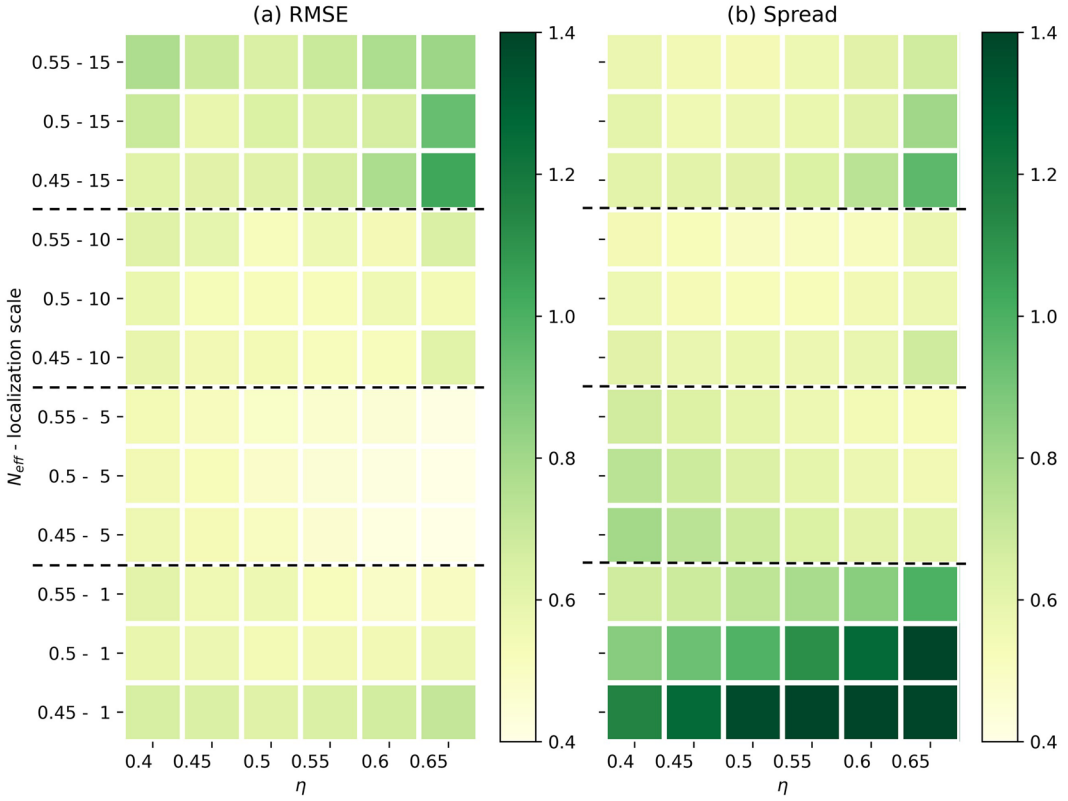

Figure 2. Prior mean RMSE(a) and spread (b) as a function of parameter N_{eff}, η, and localization scale v_{local}. The value in brackets indicates the localization scale and each cell represents one experiment.

$\epsilon \sim N(0,0.5)$. Other configurations, which are different from the default, will be given at the beginning of each experiment. A spinup time with 1,000 time steps was added to each test and the following 10,000 steps were used to summarize and analyze results. All experiments were executed on the DAS-4 supercomputer (Bal et al., 2016).

In all experiments, the root-mean-square error (RMSE) was used to evaluate the performance of the new filter. The ensemble spread, defined as the square root of the variance of the ensemble averaged over all model states, is another metric for the evaluation. These two metrics are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N_m}\sum_{k=1}^{N_m}\left(x^t_{\text{truth},k} - \bar{x}^t_k\right)^2} \tag{20}$$

$$\text{Spread} = \sqrt{\frac{1}{N_m}\sum_{k=1}^{N_m}\sum_{i=1}^{N_p}\left(x^t_{i,k} - \bar{x}^t_k\right)^2} \tag{21}$$

in which $\bar{x}^t_k$ is the ensemble mean for filters and $x^t_{\text{truth},k}$ is the corresponding truth. $x^t_{i,k}$ represents each particle. In this research, we use time-averaged values for both metrics.

3.2. Results

The first experiment was to examine the behavior of one model state, and the performance of the local filter when a linear H operator was used in the simulation. The results of prior error statistics from LETKF and LPF-GT were compared by calculating domain averages of RMSE and ensemble spreads for all model states over time. Time series of the first model state and the corresponding truth have been plotted together in Figure 1. The constants η and N_{eff} in LPF-GT can be tuned, and using different combinations can impact the final performance. In this experiment, we used $\eta = 0.55$ and $N_{eff} = 0.65$, and the behavior of the system for these two parameters was investigated later. For LETKF, after tuning the inflation factor, the fixed

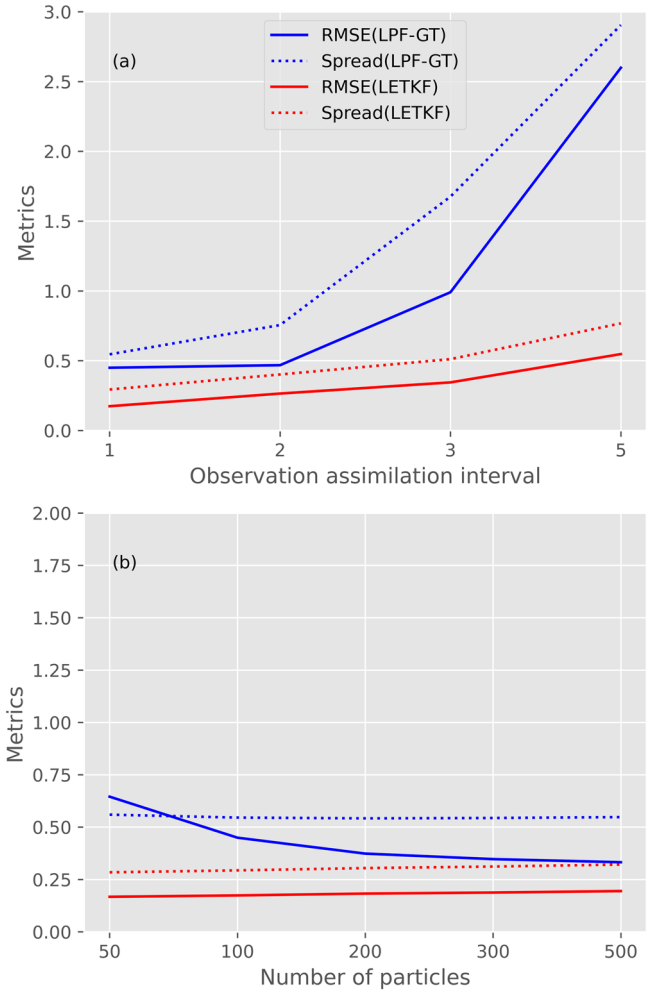

Figure 3. Prior mean RMSE and spread as a function of (a) ensemble size and (b) assimilation time interval for experiments using the linear observation operator.

value for inflation 1.05 was used. In this linear-Gaussian case, at the beginning of data assimilation, LPF-GT took more time than LETKF to reach a stable status. It is probably because LPF-GT is more sensitive to sampling errors. When both of them become stable, they produce low RMSEs and have similar performance. For the entire simulation time, results show that LETKF outperforms LPF-GT. The time-averaged RMSE and ensemble spread for LPF-GT are 0.38 and 0.55, respectively, compared to 0.17 and 0.29 for LETKF. This experiment demonstrates that LPF-GT can work stably for high-dimensional models using the linear observation operator and confirms that the application of localization in PFs prevents the filter collapse using fewer particles.

Next, we explore the impact of η and N_{eff}, when different localization scales v_{local} are used. Therefore, a set of experiments using the linear H operator for different combinations of these parameters were conducted, and the RMSE and spread of prior particles averaged over the entire domain are shown in Figure 2. LPF-GT is tuned optimally in this case, and the optimal configuration comes from the experiment, which yields the lowest prior RMSEs. Four localization scales v_{local} used in this test are 1.0, 5.0, 10.0, and 15.0, respectively. The tested values for the parameter η were 0.45, 0.5, and 0.55 and parameter N_{eff} varies between 0.4 and 0.65 with 0.05 steps. Each pixel is the result of running the data assimilation algorithm with different settings of these three parameters. For the current settings $\eta = 0.5, N_{eff} = 0.65$, and $v_{local} = 5$ yielded the best result.

These results in Figure 2 clearly show that changes in these two parameters, N_{eff}, and η can impact the performance, and both of their roles are significant. When $v_{local} = 5$ is used, the RMSE becomes lower with the growth of N_{eff}. In cases with larger localization scales, in general, the performance of the LPF-GT becomes worse with increasing N_{eff}. As for the impact of the localization, increasing localization scales degrade the accuracy of the LPF-GT generally, which is consistent with the results in Penny & Miyoshi's (2015) research. The η parameter is used to adjust the analysis errors, which are derived from analysis errors given by the PF and the other one corrected by the Gamma test, but there is no apparent monotonic relationship between the two error sources. Therefore, appropriate parameter values of η should be tuned to obtain the desired performance. The parameter N_{eff} influences the performance of LPF-GT by reassigning values of weights. Using a proper value of N_{eff} can draw model states to the truth, and it turns out that changing the variance of particles by adjusting their weights is a potential strategy to avoid the collapse for PFs.

Next, we check the sensitivity of the proposed LPF-GT to the number of particles and the impact of the observation assimilation intervals in both linear and nonlinear cases. In the nonlinear case, $y = \ln|\mathbf{x}|$ is used as the nonlinear observation operator. In this research, the number of particles varies as 50, 100, 200, 300, and 500. The standard deviation of the error in observation is set to 0.5 in the linear case, and both 0.1 and 0.5 are chosen as observational errors in the nonlinear case.

Figure 3 shows the RMSE, and the corresponding ensemble spread as a function of the number of particles and update interval in the linear case. For the linear/Gaussian experiment, both filters produce low RMSEs and the LETKF has lower RMSEs than the LPF-GT but from resulting prior statistics the performance of the LPF-GT is still acceptable. LPF-GT is more sensitive to the number of particles and, as expected, the RMSEs decrease when the number of particles is increased. For both filters, using 50 particles or ensemble members can provide satisfactory results. The LETKF is optimal for the linear observation type, and because of the application of localization, it does not need a large number of ensemble members to maintain the Gaussianity of the ensemble. Performance is limited by sampling errors in prior ensembles. In Figure 3, the spread given by LPF-GT is larger than results from LETKF. It is mainly because we use several methods to avoid the collapse of the filter including the localization procedure, tuning the effective number of particles

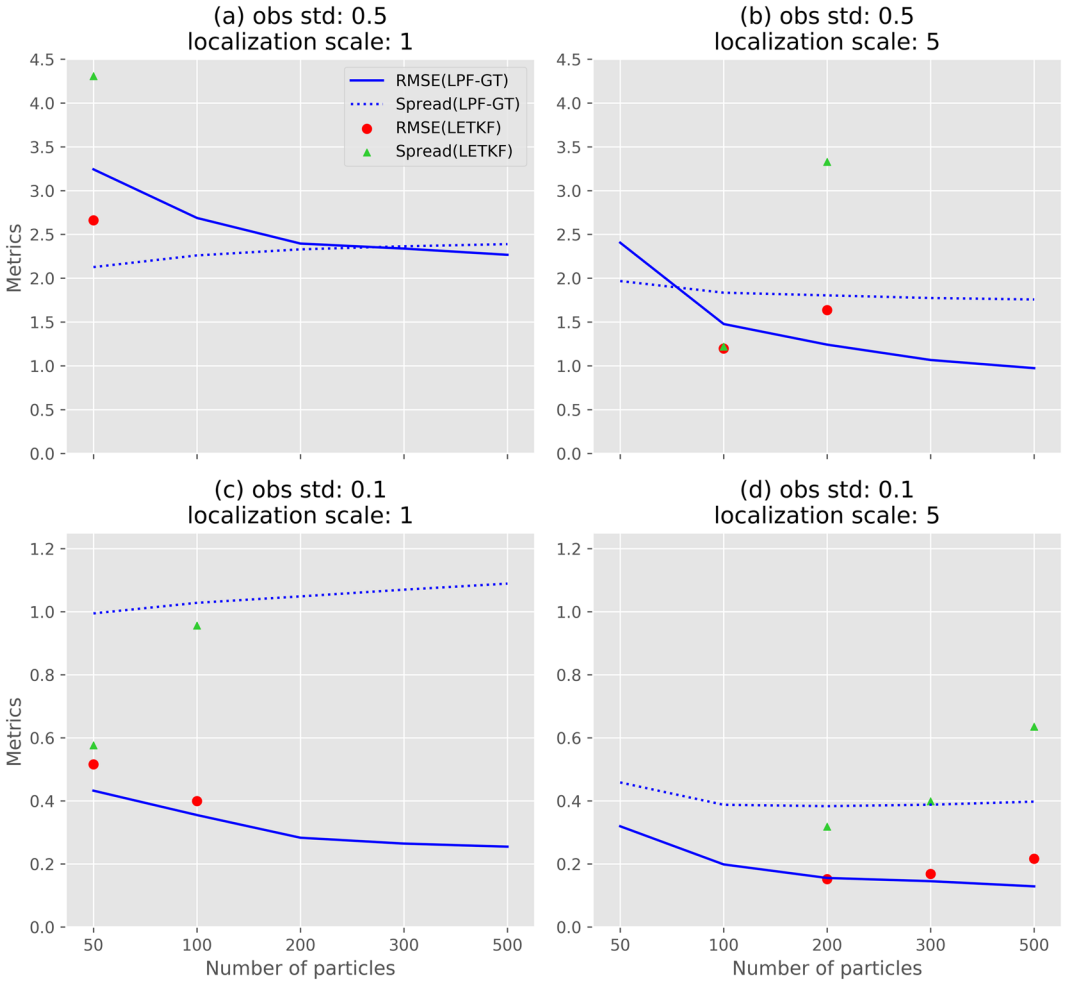

Figure 4. Prior mean RMSE and spread as a function of ensemble size for (a) obs std (the standard deviation of observational error) 0.5 and Localization Scale 1, (b) obs std 0.5 and Localization Scale 5, (c) obs std 0.1 and Localization Scale 1, and (d) obs std 0.1 and Localization scale 5, using the nonlinear observation operator $y = \ln |\mathbf{x}|$.

and considering potential uncertainty provided by the Gamma test. All these methods keep the filter stable, but the variance of particles is inflated inevitably.

To examine the performance of the LPF-GT when nonlinear observations are assimilated, a more comprehensive assessment of its performance is given. The experimental configurations for the nonlinear operator include two localization scales and two observation errors. Figure 4 shows the prior mean RMSE and the corresponding spread as a function of the number of particles in the nonlinear case, when using various localization radius and observational errors. The nonlinear H operator was used for the nonlinear experiments. However, because the LETKF is suboptimal for the nonlinear operator: $y = \ln |\mathbf{x}|$, its performance is not as stable and predictable as LPF-GT and the filter degeneracy occurs in LETKF in many experiments. Thus, in Figure 4, we only demonstrate results from the LETKF when the collapse did not happen.

From results in Figure 4, it is clear that LPF-GT can provide more accurate solutions than the LETKF under these four conditions and the non-Gaussianity introduced by the nonlinear measurement operators makes LETKF less effective. When observational errors with standard deviation 0.1 are used, LPF-GT only needs 50 particles to work stably and continuously, and LETKF only outperforms LPF-GT slightly with 200 particles. In the case with standard deviation 0.5, LPF-GT requires 100 particles to achieve a relatively acceptable performance. With limited ensemble members, LETKF needs to concentrate on the first two moments of the posterior distribution and these two moments will not be unbiased in the nonlinear case. The additional nonlinearity makes LETKF more sensitive to changes of parameters like ensemble sizes and localization scales, which accounts for the collapse of LETKF in these cases partially. The LPF-GT can maintain its ability

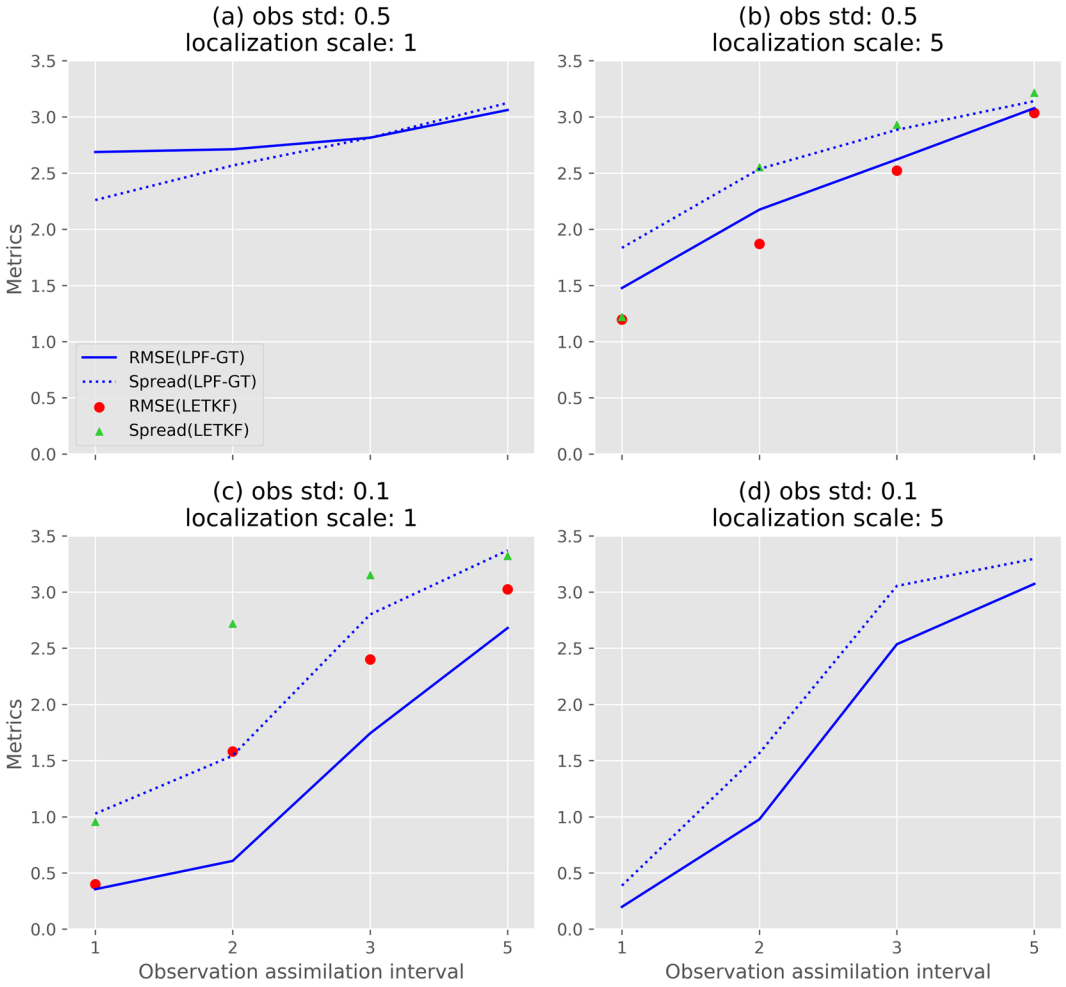

Figure 5. Prior mean RMSE and spread as a function of assimilation time interval for (a) obs std (the standard deviation of observational error) 0.5 and Localization Scale 1, (b) obs std 0.5 and Localization Scale 5, (c) obs std 0.1 and Localization Scale 1, and (d) obs std 0.1 and Localization Scale 5, using the nonlinear observation operator $y = \ln |\mathbf{x}|$.

and stability with relatively high accuracy even with different localization radius, and observational errors, which produces significant improvements over LETKF. Meanwhile, these experiments also provide useful information about the impact of localization scales. For LPF-GT, its performance becomes better when the localization scale is increased to 5.

The performance of LPF-GT is also investigated for cases in which observations are assimilated at different frequencies. The timescale $dt = 0.05$ of the Lorenz model was applied in this research, and it is comparable to the error doubling happening over 6 hr in the operational forecasting systems (Lorenz, 1996). Values of 1, 2, 3, and 5 are used for different time intervals, which represent 6, 12, 36, and 72 hr update frequencies, respectively.

Similar to experiment settings for the number of particles, we ran simulations in both linear and nonlinear cases using different time intervals. The results for the linear operator are shown in Figure 5. From simulation results, the LPF-GT has no practical benefit over LETKF when observations are assimilated less frequently. Using a longer update time interval can accumulate model errors, which is the main reason why the performance of both filters becomes worse when increasing the time interval. When using the nonlinear operator, LPF-GT offers advantages over the LETKF, which is shown in Figure 5.

When using localization radius five and observational errors with standard deviation 0.5, LETKF achieves a somewhat better performance than LPF-GT. However, in other cases, LPF-GT offers substantial benefits over LETKF, which shows that LETKF is more easily influenced by the localization radius. Probably,

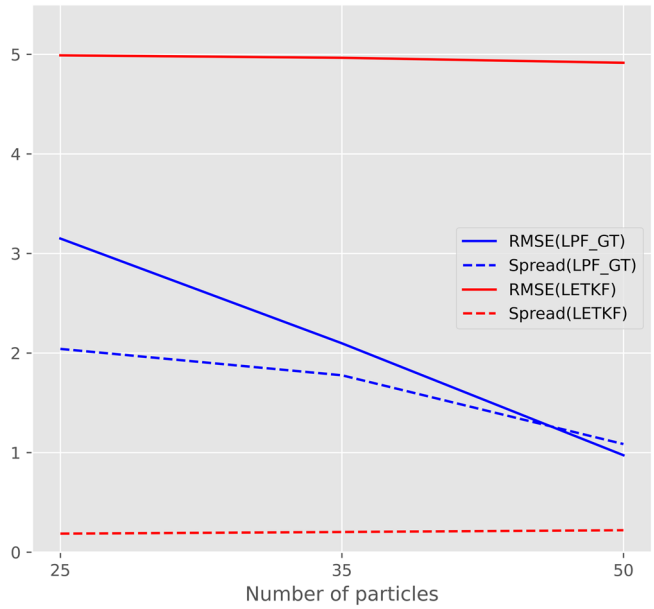

Figure 6. Prior mean RMSE and spread as a function of ensemble size for the standard deviation of observational error 0.1 and Localization Scale 5 using the nonlinear observation operator $y = \ln|\mathbf{x}|$ and 1,000 particles.

searching more of the parameter space for LETKF can produce better results, but it increases the cost. Besides, similar to results in experiments of the number of particles, LPF-GT yields better results when the localization scale five used.

Since the performance of LPF-GT in high-dimensional models is of high interest, we investigate the behavior of LPF-GT in a nonlinear case using the Lorenz model with 1,000 variables. We used fewer particles to explore the performance limit of LPF-GT. Half of all model states were observed, and nonlinear observations generated from the truth and the operator $y = \ln|\mathbf{x}|$ with the observation error $\epsilon \sim N(0, 0.1)$ were assimilated by using 25, 35, and 50 particles in the experiments. The rest of the experimental settings remained the same.

Figure 6 shows average prior mean RMSEs and spread for the different numbers of particles. For the case without data assimilation in this research, when a number of particles are propagated in the Lorenz model over time, the average mean RMSE of particles is about 3.5. It means that when the value of RMSE is smaller than 3.5, data assimilation improves model estimations. Otherwise, data assimilation is likely to reduce the accuracy of the model. From the results in Figure 6, it becomes clear that applying LETKF to the nonlinear case, causes filter collapse, and its domain-averaged prior RMSEs in all experiments are close to 5.0. For LPF-GT, increasing the number of particles improves the performance of data assimilation as expected. In the case with 25 or 35 particles, we can conclude that LPF-GT helps improve the accuracy of the model. But to achieve a satisfactory result, LPF-GT needs at least 50 particles.

To investigate the stability of filters, time series of the second model state, which is unobserved, and the development of the spread and the RMSE of particles in the experiment using 50 particles, are shown in Figure 7. The time series starts from the beginning of data assimilation, and it needs some time to stabilize. The first two rows show the comparison of the truth and the estimation of the second model state for LPF-GT and LETKF. LPF-GT gives a more accurate result and shows more stability than LETKF. The estimation of LETKF follows the truth in a short time, but after that, it deviates from the truth. The performance of LPF-GT is much more stable, and its estimations always stay close to the truth. From the development of the RMSE and the spread over time, it is clear that LETKF collapses at the beginning. By contrast, the spread and the RMSE of LPF-GT gradually decrease until the filter becomes stable. In this case, LPF-GT completely outperforms LETKF.

4. Conclusions

This research proposes a local PF for nonlinear high-dimensional applications. Similar to the localization method used in LETKF, LPF-GT assimilates observations within the localization scale, and the influence of distant observations is decreasing gradually in large-scale geophysical systems. Because of the use of the localization method, in this method, each model state needs much fewer observations than what is needed by typical PFs. The LPF-GT updates particles sequentially when observations are available. Posterior weights for each particle are obtained based on the localized likelihood of observations for each state in a model. Particles with lower weights are removed by the resampling method. The mean of new particles is based on the resampled particles, and its uncertainty is a linear combination of sampled particles and the uncertainty estimated by the Gamma test. The proposed filter can prevent filter collapse, even when the number of particles is relatively small. Another advantage of the use of the localization method is to make computation of the new filter affordable for large applications. For each state, all calculations made are within the localization scales and can be parallelized easily. Correction by the Gamma test needs more computing time. However, for even larger systems, the computational cost of the Gamma test does not increase because the estimate of variance approaches a constant with an increase in the number of samples. It is unnecessary to use all samples for the Gamma test, and only a small number of samples are enough for the estimation of the variance.

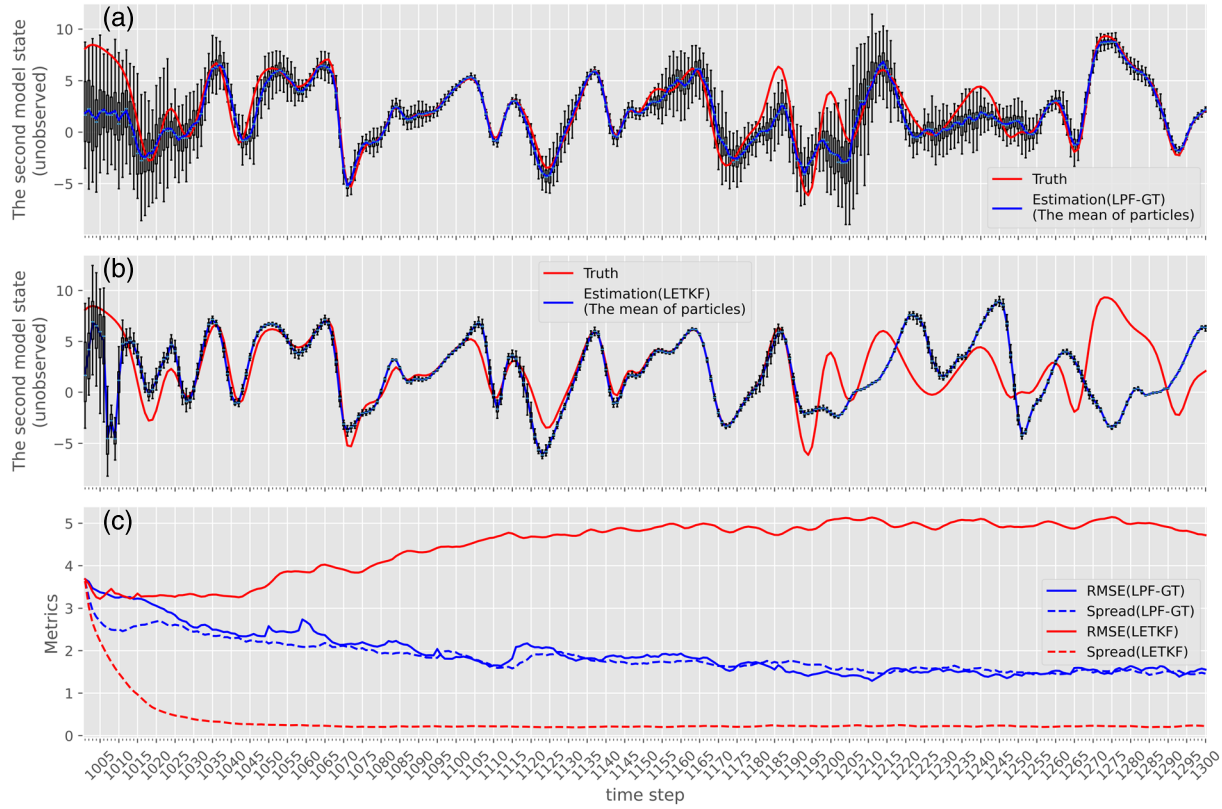

Figure 7. Time series of (a) the second model state using LPF-GT, (b) the second model state using LETKF, and (c) RMSE and spread for both methods. In this experiment, the standard deviation of observational error and localization scale are set to 0.1 and 5, respectively, and the nonlinear observation operator $y = \ln|\mathbf{x}|$ and 1,000 particles are used. In the first two figures, the blue curve is the truth and the red curve is the estimation of the model state. In the third one, the solid lines indicate the RMSE and the dash lines is the spread. The assimilation starts at time step 1,000.

The LPF-GT algorithm proposed in this research can be problematic for some applications in geoscience because of the imbalance caused by localization in model states, which is a disadvantage of the localization method. Similar to other data assimilation methods with localization, like LETKF and localized EnKF, LPF-GT can also break the physical consistency and cause an imbalance in posterior states. Poterjoy (2016) found the imbalance issue when updated using their local PFs. The local PFs proposed by Penny and Miyoshi (2015) showed a certain level of imbalance when they decreased the localization scale. The imbalance issue is common in most data assimilation methods with localization, and a proper localization scale is always needed for a specific application. One possible solution is to tune localization scales, but the cost of it may be expensive for some larger models. Thus, developing an adaptive localization method deserves more attention.

The LPF-GT has been tested by using the Lorenz system with 40 and 1,000 variables. All results from a set of experiments show that the new filter is stable and avoids filter degeneracy successfully. In the ideal case with the linear observation operator, LETKF outperforms LPF-GT slightly. For nonlinear cases, LPF-GT provides a significant benefit over LETKF. The successful application of LPF-GT with a Lorenz model does not guarantee its success in other models. Therefore, its application in real-world models with high dimensions will be the main topic of future studies to explore the limitations and advantages of this new filter.

Appendix A: Pseudocode Descriptions of LPF-GT

To be consistent with the notations in previous sections and for convenience, the notation used in pseudocodes is listed as follows:

N_p the number of particles
N_m the dimension of model states

N_{obs} the number of observations
y observations
x model states

Algorithm 1 Pseudocode description of LPF-GT

for $n = 1 \rightarrow N_m$ **do**
call LOCALIZATION to obtain local index id_{loc} and local coefficients c_{loc} and then get the local observations $\left(y_{i,local}\right)$ and local particles $x_{\text{local}}^{(y_{i-1})}$
for $j = 1 \rightarrow N_p$ **do**
$w_{n,j} \leftarrow p\left(x_{j,local}^{(y_{i-1})} \mid y_{i,local}\right)$
end for
call a bisection function to find factors β_a and β_m, which brings N_{eff} close to a certain value
$w_n \leftarrow \prod \left(w_{n,j} + \beta_a\right) * \beta_m$
$w_{n,\text{nor}} = \frac{w_n}{\sum w_n}$
Obtain resampled particles $x_n^{(y_i)}$ based on $w_{n,\text{nor}}$
end for
call GAMMATEST$(x^{(y_i)}, x^{(y_{i-1})})$ to obtain Γ
$\alpha = \sqrt{\frac{\Gamma}{variance\left(x^{(y_i)}, x^{(y_{i-1})}\right)}}$
$\bar{x} = \text{mean}\left(x^{(y_i)}\right)$
$x^{(y_i)} = \bar{x} + \eta\left(\bar{x} - x^{(y_i)}\right) + (1-\eta)\,\alpha\left(x^{(y_i)} - x^{(y_{i-1})}\right)$, where $\eta \in [0, 1]$

function GAMMATEST$(x^{(y_i)}, x^{(y_{i-1})})$
calculate the Gamma test statistic Γ
return Γ
end function

function LOCALIZATION
calculate local index id_{loc} and local coefficients c_{loc}
return $id_{\text{loc}}, c_{\text{loc}}$
end function

Acknowledgments
This research is part of the eWaterCycle project funded through the Netherlands eSciencecenter with project number 027.017.F01 and is supported by China Scholarship Council with the project reference number of 201606320249.

References

Abolafia-Rosenzweig, R., Livneh, B., Small, E., & Kumar, S. (2019). Soil moisture data assimilation to estimate irrigation water use. *Journal of Advances in Modeling Earth Systems*, *11*, 3670–3690. https://doi.org/10.1029/2019MS001797

Ades, M., & van Leeuwen, P. J. (2013). An exploration of the equivalent weights particle filter. *Quarterly Journal of the Royal Meteorological Society*, *139*(672), 820–840. https://doi.org/10.1002/qj.1995

Ades, M., & van Leeuwen, P. J. (2014). The effect of the equivalent-weights particle filter on dynamical balance in a primitive equation model. *Monthly Weather Review*, *143*(2), 581–596. https://doi.org/10.1175/MWR-D-14-00050.1

Ades, M., & van Leeuwen, P. J. (2015). The equivalent-weights particle filter in a high-dimensional system. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 484–503. https://doi.org/10.1002/qj.2370

Bal, H., Epema, D., de Laat, C., van Nieuwpoort, R., Romein, J., Seinstra, F., et al. (2016). A medium-scale distributed system for computer science research: Infrastructure for the long term. *Computer*, *49*(5), 54–63.

Bannister, R. N. (2017). A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 607–633. https://doi.org/10.1002/qj.2982

Bengtsson, T., Bickel, P., & Li, B. (2008). Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems. In D. Nolan & T. Speed (Eds.), *IMS Collections: Probability and Statistics: Essays in Honor of David A. Freedman* (Vol. 2, pp. 316–334). Beachwood, Ohio, USA: Institute of Mathematical Statistics. https://doi.org/10.1214/193940307000000518

Bengtsson, T., Snyder, C., & Nychka, D. (2003). Toward a nonlinear ensemble filter for high-dimensional systems. *Journal of Geophysical Research*, *108*(D24), 8775. https://doi.org/10.1029/2002JD002900@10.1002/(ISSN)2169-8996.SPCTME1

Chen, H., Yang, D., Hong, Y., Gourley, J. J., & Zhang, Y. (2013). Hydrological data assimilation with the ensemble square-root-filter: Use of streamflow observations to update model states for real-time flash flood forecasting. *Advances in Water Resources*, *59*, 209–220. https://doi.org/10.1016/j.advwatres.2013.06.010

Choi, Y., Cha, D.-H., Lee, M.-I., Kim, J., Jin, C.-S., Park, S.-H., & Joh, M.-S. (2017). Satellite radiance data assimilation for binary tropical cyclone cases over the western North Pacific. *Journal of Advances in Modeling Earth Systems*, *9*, 832–853. https://doi.org/10.1002/2016MS000826

Chustagulprom, N., Reich, S., & Reinhardt, M. (2016). A hybrid ensemble transform particle filter for nonlinear and spatially extended dynamical systems. *SIAM/ASA Journal on Uncertainty Quantification*, *4*(1), 592–608. https://doi.org/10.1137/15M1040967

Dong, J., Steele-Dunne, S. C., Judge, J., & van de Giesen, N. (2015). A particle batch smoother for soil moisture estimation using soil temperature observations. *Advances in Water Resources*, *83*, 111–122. https://doi.org/10.1016/j.advwatres.2015.05.017

Doucet, A., de Freitas, N., & Gordon, N. (2001). An introduction to sequential Monte Carlo methods, *Sequential Monte Carlo methods in practice* (pp. 3–14). New York, NY: Springer. https://doi.org/10.1007/978-1-4757-3437-91

Evans, D., & Jones, A. J. (2002). A proof of the Gamma test. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *458*(2027), 2759–2799. https://doi.org/10.1098/rspa.2002.1010

Fang, M., & Li, X. (2019). An artificial neural networks-based tree ring width proxy system model for paleoclimate data assimilation. *Journal of Advances in Modeling Earth Systems*, *11*, 892–904. https://doi.org/10.1029/2018MS001525

Farchi, A., & Bocquet, M. (2018). Review article: Comparison of local particle filters and new implementations. *Nonlinear Processes in Geophysics*, *25*(4), 765–807. https://doi.org/10.5194/npg-25-765-2018

Fox, A. M., Hoar, T. J., Anderson, J. L., Arellano, A. F., Smith, W. K., Litvak, M. E., et al. (2018). Evaluation of a data assimilation system for land surface models using CLM4.5. *Journal of Advances in Modeling Earth Systems*, *10*, 2471–2494. https://doi.org/10.1029/2018MS001362

Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, *125*(554), 723–757. https://doi.org/10.1002/qj.49712555417

Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F (Radar and Signal Processing)*, *140*(2), 107–113. https://doi.org/10.1049/ip-f-2.1993.0015

Hol, J. D., Schon, T. B., & Gustafsson, F. (2006). On resampling algorithms for particle filters. Paper presented at Nonlinear Statistical Signal Processing IEEE Workshop, Linkoping, Sweden.

Hong, S., Bolic, M., & Djuric, P. M. (2004). An efficient fixed-point implementation of residual resampling scheme for high-speed particle filters. *IEEE Signal Processing Letters*, *11*(5), 482–485. https://doi.org/10.1109/LSP.2004.826634

Houtekamer, P. L., & Zhang, F. (2016). Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, *144*(12), 4489–4532. https://doi.org/10.1175/MWR-D-15-0440.1

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, *230*(1–2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008

Hut, R., Amisigo, B. A., Steele-Dunne, S., & van de Giesen, N. (2015). Reduction of used memory ensemble Kalman filtering (RumEnKF): A data assimilation scheme for memory intensive, high performance computing. *Advances in Water Resources*, *86, Part B*, 273–283. https://doi.org/10.1016/j.advwatres.2015.09.007

Irrgang, C., Saynisch, J., & Thomas, M. (2017). Utilizing oceanic electromagnetic induction to constrain an ocean general circulation model: A data assimilation twin experiment. *Journal of Advances in Modeling Earth Systems*, *9*, 1703–1720. https://doi.org/10.1002/2017MS000951

Jin, J., Lin, H. X., Heemink, A., & Segers, A. (2018). Spatially varying parameter estimation for dust emissions using reduced-tangent-linearization 4DVar. *Atmospheric Environment*, *187*, 358–373. https://doi.org/10.1016/j.atmosenv.2018.05.060

Jin, J., Lin, H. X., Segers, A., Xie, Y., & Heemink, A. (2019). Machine learning for observation bias correction with application to dust storm data assimilation. *Atmospheric Chemistry and Physics*, *19*(15), 10,009–10,026. https://doi.org/10.5194/acp-19-10009-2019

Jones, A. J., Evans, D., & Kemp, S. E. (2007). A note on the Gamma test analysis of noisy input/output data and noisy time series. *Physica D: Nonlinear Phenomena*, *229*(1), 1–8. https://doi.org/10.1016/j.physd.2006.12.013

Katzfuss, M., Stroud, J. R., & Wikle, C. K. (2016). Understanding the ensemble Kalman filter. *The American Statistician*, *70*(4), 350–357. https://doi.org/10.1080/00031305.2016.1141709

Lee, Y., & Majda, A. J. (2016). State estimation and prediction using clustered particle filters. *Proceedings of the National Academy of Sciences*, *113*(51), 14,609–14,614. https://doi.org/10.1073/pnas.1617398113

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., et al. (2012). Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, *16*(10), 3863–3887. https://doi.org/10.5194/hess-16-3863-2012

Lorenz, E. N. (1996). Predictability: A problem partly solved. In *Proceedings of a Seminar held at ECMWF on Predictability* (pp. 1–18). Reading, UK. https://scholar.google.com/scholar?hl=en&q=Lorenz%2C+E.+N.+%281996%29.+Predictability%3A+A+problem+partly+solved.+In+Proceedings+of+a+Seminar+held+at+ECMWF+on+Predictability+%28pp.+1%E2%80%9318%29.+Reading%2C+UK

Lui, J. S., & Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association; Alexandria*, *93*(443), 1032–1044.

Penny, S. G., & Miyoshi, T. (2015). A local particle filter for high dimensional geophysical systems. *Nonlinear Processes in Geophysics Discussions*, *2*, 1631–1658. https://doi.org/10.5194/npgd-2-1631-2015

Pinheiro, F. R., van Leeuwen, P. J., & Geppert, G. (2019). Efficient nonlinear data assimilation using synchronization in a particle filter. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2510–2523. https://doi.org/10.1002/qj.3576

Poterjoy, J. (2016). A localized particle filter for high-dimensional nonlinear systems. *Monthly Weather Review*, *144*(1), 59–76. https://doi.org/10.1175/MWR-D-15-0163.1

Poterjoy, J., & Anderson, J. L. (2016). Efficient assimilation of simulated observations in a high-dimensional geophysical system using a localized particle filter. *Monthly Weather Review*, *144*(5), 2007–2020. https://doi.org/10.1175/MWR-D-15-0322.1

Potthast, R., Walter, A., & Rhodin, A. (2018). A localized adaptive particle filter within an operational NWP framework. *Monthly Weather Review*, *147*(1), 345–362. https://doi.org/10.1175/MWR-D-18-0028.1

Rebeschini, P., & van Handel, R. (2015). Can local particle filters beat the curse of dimensionality? *The Annals of Applied Probability*, *25*(5), 2809–2866. https://doi.org/10.1214/14-AAP1061

Snyder, C., Bengtsson, T., Bickel, P., & Anderson, J. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, *136*(12), 4629–4640. https://doi.org/10.1175/2008MWR2529.1

Sun, L., Seidou, O., Nistor, I., & Liu, K. (2016). Review of the Kalman-type hydrological data assimilation. *Hydrological Sciences Journal*, *61*(13), 2348–2366. https://doi.org/10.1080/02626667.2015.1127376

van Leeuwen, P. J. (2003). Nonlinear ensemble data assimilation for the ocean. In *Seminar on recent developments in data assimilation for atmosphere and ocean. 8–12 September 2003, ECMWF, Reading, UK*. https://scholar.google.com/scholar?hl=en&q=van+Leeuwen%2C+P.J.+%282003%29.+Nonlinear+ensemble+data+assimilation+for+the+ocean.+In+seminar+Recent+Developments+in+Data+Assimilation+for+Atmosphere+and+Ocean.+8%E2%80%9312+September+2003%2C+ECMWF%2C+Reading%2C+UK

van Leeuwen, P. J. (2009). Particle filtering in geophysical systems. *Monthly Weather Review*, *137*(12), 4089–4114. https://doi.org/10.1175/2009MWR2835.1

van Leeuwen, P. J. (2010). Nonlinear data assimilation in geosciences: An extremely efficient particle filter. *Quarterly Journal of the Royal Meteorological Society*, *136*(653), 1991–1999. https://doi.org/10.1002/qj.699

van Leeuwen, P. J. (2015). Nonlinear data assimilation for high-dimensional systems. In P. J. Van Leeuwen, Y. Cheng & S. Reich (Eds.), *Nonlinear data assimilation* (pp. 1–73). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-18347-31

Xie, X., & Zhang, D. (2010). Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. *Advances in Water Resources*, *33*(6), 678–690. https://doi.org/10.1016/j.advwatres.2010.03.012

Zhang, Y., Hou, J., Gu, J., Huang, C., & Li, X. (2017). SWAT-based Hydrological Data Assimilation System (SWAT-HDAS): Description and case application to river basin-scale hydrological predictions. *Journal of Advances in Modeling Earth Systems*, *9*, 2863–2882. https://doi.org/10.1002/2017MS001144

Zhang, H., Qin, S., Ma, J., & You, H. (2013). Using residual resampling and sensitivity analysis to improve particle filter data assimilation accuracy. *IEEE Geoscience and Remote Sensing Letters*, *10*(6), 1404–1408. https://doi.org/10.1109/LGRS.2013.2258888