Acta Oceanol. Sin., 2022, Vol. 41, No. 2, P. 80–90
https://doi.org/10.1007/s13131-021-1876-1
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

The statistical observation localized equivalent-weights particle filter in a simple nonlinear model

Yuxin Zhao[1], Shuo Yang[1], Renfeng Jia[2], Di Zhou[3], Xiong Deng[1]*, Chang Liu[1], Xinrong Wu[4]

[1] College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China

[2] Harbin Marine Boiler & Turbine Research Institute, Harbin 150078, China

[3] China Ship Development and Design Center, Wuhan 430064, China

[4] Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, State Oceanic Administration, Tianjin 300171, China

Received 20 March 2021; accepted 5 June 2021

Abstract

This paper presents an improved approach based on the equivalent-weights particle filter (EWPF) that uses the proposal density to effectively improve the traditional particle filter. The proposed approach uses historical data to calculate statistical observations instead of the future observations used in the EWPF's proposal density and draws on the localization scheme used in the localized PF (LPF) to construct the localized EWPF. The new approach is called the statistical observation localized EWPF (LEWPF-Sobs); it uses statistical observations that are better adapted to the requirements of real-time assimilation and the localization function is used to calculate weights to reduce the effect of missing observations on the weights. This approach not only retains the advantages of the EWPF, but also improves the assimilation quality when using sparse observations. Numerical experiments performed with the Lorenz 96 model show that the statistical observation EWPF is better than the EWPF and EAKF when the model uses standard distribution observations. Comparisons of the statistical observation localized EWPF and LPF reveal the advantages of the new method, with fewer particles giving better results. In particular, the new improved filter performs better than the traditional algorithms when the observation network contains densely spaced measurements associated with model state nonlinearities.

Key words: data assimilation, particle filter, equivalent weights particle filter, localization methods

Citation: Zhao Yuxin, Yang Shuo, Jia Renfeng, Zhou Di, Deng Xiong, Liu Chang, Wu Xinrong. 2022. The statistical observation localized equivalent-weights particle filter in a simple nonlinear model. Acta Oceanologica Sinica, 41(2): 80–90, doi: 10.1007/s13131-021-1876-1

1 Introduction

Data assimilation algorithms are frequently used in a wide range of atmospheric and oceanic fields. The most extensive solution of data assimilation is to obtain the probability density function (PDF) of the state based on the model. Most current assimilation algorithms use Gaussian approximate error distributions for observations and model prediction, which have limitations for use in nonlinear models. There are two main schemes used in current assimilation algorithms (Chorin and Tu, 2009). One method uses the Gaussian and linear assumptions, such as the ensemble Kalman filter (EnKF) (Evensen, 1994). However, the linearization assumption causes truncation errors for high-dimensional nonlinear models. The other is the particle filter (PF) based on Bayes's theorem, which is used in statistics and mathematics (De Freitas et al., 2001). Similar to EnKF, the PF integrates all particles forward in time; then, when the observations are available, each particle is assigned a weight, which is proportional to the probability of this particle in the analysis stage (Shen et al., 2016); this weight also determines which particles will be removed in the PF algorithms, and particles with high probability are guaranteed to be retained by resampling methods (Van Leeuwen and Evensen, 1996). Because of computational limitations, only a few particles can represent the posterior PDF. As the dimensionality of the model states increases, a finite number of particles will not perfectly represent the PDF, but simply adding particles will waste computational resources (Zhu et al., 2016).

As the PF can solve fully nonlinear systems and calculate the posterior PDF, many improvements have been proposed to enable the PF to work for high-dimensional systems (Lei and Bickel, 2011). For example, Nakano et al. (2007) exploited a hybrid algorithm that uses the ensemble transform Kalman filter (ETKF) to obtain the proposal distribution. Van Leeuwen (2010) and Van Leeuwen et al. (2019) also presented the equivalent weights particle filter (EWPF) to address the problem. It uses a well-designed proposal density to move particles towards the future observation and improve the PF assimilation quality. Hybridisation is considered as a way to mitigate filter degeneracy about the PF, Stordal et al. (2011) develop the ensemble Kalman particle filter (EnKPF) which use two stages to adopt the EnKF and PF. The local EnKPF which are introduced localization is applied to a convective-scale numerical weather prediction (NWP) with the real observaions (Robert et al., 2018).

The localized PF (LPF), which was first introduced by Poterjoy (2016), uses a similar approach for localization to the localized ensemble adjustment Kalman filter (EAKF). The success of

Foundation item: The National Basic Research Program of China under contract Nos 2017YFC1404100, 2017YFC1404103 and 2017YFC1404104; the National Natural Science Foundation of China under contract No. 41676088.
*Corresponding author, E-mail: xiongdeng407@hrbeu.edu.cn

the EAKF in high-dimensional systems depends on the use of covariance localization. The LPF has a clear advantage in applying positioning in background covariance in a large spatial dimension (Houtekamer and Mitchell, 1998). This algorithm uses a Bayesian update of the particles near the location of the observation while preserving the a priori particles.

This paper proposes an improved EWPF that uses the reanalysis data to calculate the statistical observations (EWPF-Sobs). This statistical observations are used instead of the future observations that limit the development of the EWPF in real-time data assimilation. The statistical observations are used to collect and calculate historical reanalysis data during the assimilation cycle to improve the proposal density accuracy. This paper also refer to the localization function used in the LPF to propose a localized equivalent weights particle filter with statistical observations (LEWPF-Sobs). Specifically, the traditional EWPF algorithm lacks a localization scheme for models with irregularly distributed grid observations. The localization used in the LPF determines the weights using the distance of the model grid point to the observation. Therefore, the LEWPF-Sobs uses this approach to confirm the weights of the model grid points that are not observed around the given location at the time of assimilation. This manuscript describes the performance of the EWPF-Sobs and LEWPF-Sobs, and both are used in twin experiments with the 40-variable Lorenz 96 model (Lorenz, 1995).

This paper is organized as follows. Section 2 describes the EWPF in its most basic form and the localization method used in the LPF. Section 3 describes the proposal density in the EWPF, and obtaining a new EWPF with statistical observations (EWPF-Sobs). The localization strategy is used in the EWPF-Sobs to derive the LEWPF-Sobs. In Section 4 numerical experiments are used to compare the LEWPF with the LPF and the LEAKF. A summary and discussion are provided in Section 5.

2 Methodology

2.1 *The relaxation proposal density*

The PF can solve a fully nonlinear system, but resampling is not sufficient to avoid filter collapse. The EWPF uses the proposal density to ensure that all integrations end up close to the new observation and contribute to the posterior PDF. The assimilation process is divided into two parts, the relaxation proposal density and the equivalent weights proposal density. This section describes the relaxation proposal density. Bayes's theorem provides a way to propose the posterior PDF based on the prior PDF and the likelihood. The PF is based on the Bayes's theorem using the conditional posterior PDF to express the assimilation solution (Law et al., 2015). Given an observation y, the conditional PDF $p(x|y)$ is (Ades and Van Leeuwen, 2013)

$$p(x|y) = \frac{1}{p(y)} p(y|x)\, p(x), \tag{1}$$

where $p(x)$ is the prior PDF, $p(y|x)$ is the likelihood and $p(x|y)$ is the posterior PDF. The posterior PDF gives the probability of the state variable x^n at time step n and observations $y^{1:n}$ at times $1, 2, \cdots, n$. So the proposal transition density $q\left(x^n|x^{n-1}, y^n\right)$, can be shown as

$$q\left(x^n|x^{n-1}, y^n\right) = \frac{p\left(y^n|x^n\right)}{p\left(y^n|x^{n-1}\right)} p\left(x^n|x^{n-1}\right), \tag{2}$$

and the posterior PDF may be written as

$$p\left(x^n|y^{1:n}\right) = \frac{p\left(y^n|x^n\right)}{p\left(y^n\right)} \int \frac{p\left(x^n|x^{n-1}\right)}{q\left(x^n|x^{n-1}, y^n\right)} \times q\left(x^n|x^{n-1}, y^n\right) p\left(x^{n-1}|y^{1:n-1}\right) \mathrm{d}x^{n-1}. \tag{3}$$

Let us assume the particle filter selects the particles that have similar weights at the previous time step $n-1$ and drawing samples from the proposal density instead of the $p\left(x^n|x^{n-1}\right)$. So the posterior PDF becomes

$$p\left(x^n|y^{1:n}\right) = \sum_{i=1}^{N} \frac{1}{N} \frac{p\left(y^n|x_i^n\right)}{p\left(y^n\right)} \frac{p\left(x_i^n|x_i^{n-1}\right)}{q\left(x_i^n|x_i^{n-1}, y^n\right)} \delta\left(x^n - x_i^n\right), \tag{4}$$

where A is a normalization factor and the particle weights are

$$\omega_i = \frac{1}{A} p\left(y^n|x_i^n\right) \frac{p\left(x_i^n|x_i^{n-1}\right)}{q\left(x_i^n|x_i^{n-1}, y^n\right)}. \tag{5}$$

However, the EWPF aims to ensure that the majority of particles are close to the observations. It is reasonable to calculate the proposal density based on future observation information at future time n and determine the weights for model time steps at the time m. So particles need to calculate the equal weights and move toward the future observation at time $n-m$. So Eq. (5) can be written as

$$\omega_i = \frac{1}{A} p\left(y^n|x_i^n\right) \prod_{j=n-m+1}^{n} \frac{p\left(x_i^j|x_i^{j-1}\right)}{q\left(x_i^j|x_i^{j-1}, y^n\right)}. \tag{6}$$

In the EWPF the particles are updated using the model equation (Ades and Van Leeuwen, 2015):

$$x_i^j = f\left(x_i^{j-1}\right) + B(\tau)\left(y^n - h\left(x_i^{j-1}\right)\right) + d\widehat{\beta}_i^j, \tag{7}$$

where $h\left(x_i^{j-1}\right)$ is the observation operator and the random error $d\widehat{\beta}_i^j$ is distributed as a Gaussian with mean zero and covariance Q. The relaxation term $B(\tau)\left(y^n - h\left(x_i^{j-1}\right)\right)$ moves the particles toward the future observation, where the relaxation factor $B(\tau)$ is

$$B(\tau) = b\tau \boldsymbol{Q}\boldsymbol{H}^{\mathrm{T}} R^{-1}, \tag{8}$$

with $\tau = \left(t^j - t^0\right) / \left(t^n - t^0\right)$, t^0 is the previous observation time and t^n is the target observation time. The observation error covariance is $\boldsymbol{R}$ and $\boldsymbol{Q}$ is the model error covariance; $\boldsymbol{H}$ is the linearization of the observation operator; b represents a scaling factor that controls the degree of relaxation towards the observation (Ades and Van Leeuwen, 2015).

2.2 *The equivalent weights particle filter*

The EWPF has been developed by Van Leeuwen (2010, 2011). It ensures that most particles are guided towards the future observations and uses particles that have significant information on the posterior PDF. The EWPF uses two separate steps to avoid filter degeneracy. The first is the relaxation proposal density introduced in Section 2.1, and it uses a simple nudging scheme to move the particles towards the future observation based on Eq. (7).

The second is the equivalent-weights proposal density; the relaxation proposal density is used to adjust the particles so that they have similar weights in the final step. Based on the results of the proposed density, the equivalent-weights proposal density is briefly described as follows:

(1) Separation of the accumulated weight of each particle ω_i until the last step before an observation, the equivalent-weight may be written as (Van Leeuwen, 2011)

$$\omega_i^n = \omega_i^{\text{rest}} \frac{p\left(y^n \left| x_i^n \right.\right) p\left(x_i^n \left| x_i^{n-1} \right.\right)}{q\left(x^n \left| x_i^{n-1} \right., y^n\right)}, \tag{9}$$

where ω_i^{rest} represents the proposal density weights ω_i accumulated over the assimilation interval.

(2) Determine the target weight ω^{target} that is chosen based on the maximum weight each particle can achieve, and makes most of particles can have the equivalent-weight. so the target weight is given by

$$\begin{aligned} -\ln\omega_i^{\text{target}} = &-\ln\omega_i^{\text{rest}} + \frac{1}{2}\left(y^n - \boldsymbol{H}f\left(x_i^{n-1}\right)\right)^{\mathrm{T}} \times \\ &\left(\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}\right)^{-1}\left(y^n - \boldsymbol{H}f\left(x_i^{n-1}\right)\right). \end{aligned} \tag{10}$$

This choice would lead to all weights becoming equal to the minimum weights. However, the method can reach a compromise between the retained particles and the weight value, and those discarded particles are still returned by resampling.

(3) Determine the new particles x_i^n that have equal weights, the model state at time n may be written as

$$x_i^n = f\left(x_i^{n-1}\right) + \alpha_i K\left(y^n - \boldsymbol{H}f\left(x_i^{n-1}\right)\right), \tag{11}$$

where $K = \boldsymbol{Q}\boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}\right)^{-1}$, $\alpha_i = 1 + \sqrt{1 - b_i/a_i}$, $a_i = 0.5x_i^{\mathrm{T}}X \times \boldsymbol{R}^{-1}\boldsymbol{H}Kx$, $b_i = 0.5x_i^{\mathrm{T}}\boldsymbol{R}^{-1}x_i - \ln\left(\omega_i^{\text{target}}\right) - \ln\left(\omega_i^{\text{rest}}\right)$ and $x = y^n - \boldsymbol{H}f \times \left(x_i^{n-1}\right)$ (Van Leeuwen, 2015).

2.3 *The localization weights of the particle filter*

Poterjoy (2016) proposed the LPF algorithm using a framework similar to the EAKF that relies on the localization function to describe the relationship between weights and a single observation. The weights are confirmed using the likelihood of the observation in the PF framework. The LPF achieves localization by extending the original weights from scalars to vectors and the localization function $l\left[y, x_j; r\right]$ (Gaspari and Cohn, 1999) is used to conform the localization about the jth element of the ensemble local weight ω_i and their normalization vector $\boldsymbol{\Omega}$:

$$\omega_{i,j} = \left(p\left(y|x_{i,j}\right) - 1\right) l\left[y, x_j; r\right] \alpha + 1, \tag{12}$$

$$\boldsymbol{\Omega}_j = \sum_{i=1}^{N_e} \omega_{i,j}. \tag{13}$$

The localization function that controls the impact of observations which are from region surrounding the localization analysis on the weights. (Whitaker and Hamill, 2012). This function has a maximum value of 1 when y and x_j coincide and asymptotes to 0 if the y and x_j are far apart; the parameter r controls the rate of decay. The localization scheme $l\left[y, x_j; r\right] = \boldsymbol{\Omega}\left(d_{y,x_i}; c\right)$ which has been detailed calculation formula in the Shen et al. (2017). It uses a continuous function whose value is inversely proportional to the distance from the observation point to generate a multiplication factor for the state. Where d_{y,x_i} represents the distance between the observation y and the model grid value x_j. The parameter c related to the decorrelation length. The LPF parameter $0 < \alpha < 1$ used to make the weights more uniform in regions when the observation has a large impact to the update. This weights represent the posterior probability distribution for the model state variable. The PF uses the resampling steps after the update when the weights are determined. It eliminates particles that have small weights and duplicates those with large weights (Gordon et al., 1993). So the scalar weights are determined by

$$\omega_j^{t_k} \propto \frac{1}{\sqrt{2\pi\left|\boldsymbol{R}\right|}} \exp\left[-\frac{1}{2}\left(y^{t_k} - h\left(x_j^{t_k}\right)\right)^{\mathrm{T}} \boldsymbol{R}^{-1}\left(y^{t_k} - h\left(x_j^{t_k}\right)\right)\right]. \tag{14}$$

Here, y^{t_k} is a vector containing all the observations of all model variables at time steps of t_k, $\boldsymbol{R}$ is the observation error covariance, and h is the measurement function that maps the model states onto the observational site corresponding to y^{t_k}. Consequently, the LPF localizes the roles of each observation by extending the scalar weights to vector weights (Shen et al., 2017). So this method use x_i to denote the ith particle and the observation y has m components that represent the observational dataset around the model grid point x_i, so the distances between particles and observational site values of y_j are given as $l\left[y_j, x_i; r\right]$, for $j = 1, 2, \cdots, m$. So from Eq. (12), Eq. (14) can be written as

$$\begin{aligned} \omega_i &\propto \exp\left[-\frac{1}{2}\sum_{j=1}^{m} l\left[y_j, x_i, r\right]\left(y_j - h_j\left(x_i\right)\right)^2/\sigma_j^2\right] \\ &= \prod_{j=1}^{m} \exp\left[-\frac{1}{2} l\left[y_j, x_i, r\right]\left(y_j - h_j\left(x_i\right)\right)^2/\sigma_j^2\right], \end{aligned} \tag{15}$$

where σ_j^2 is the observation variance and h_j is the function that maps the model state to the observations. The weight before reaching the observation y_{j-1} are denoted as $\left(\omega_j^n\right)_{y_{j-1}}$, and then multiply each term to get $\left(\widehat{\omega}_j^n\right)_{y_j}$ from the Eq. (16):

$$\left(\widehat{\omega}_j^n\right)_{y_j} = \exp\left[-\frac{1}{2} l\left(y_j - h_j\left(x_i\right)\right)^2/\sigma_j^2\right]. \tag{16}$$

Therefore, the vector weights corresponding to the observation y_j can be obtained recursively by assimilating the weights based on the individual observations

$$\left(\omega_i\right)_{y_j} = \left(\omega_i\right)_{y_{j-1}}\left(\widehat{\omega}_{y_j}\right), \quad j = 1, 2, \cdots, m, \tag{17}$$

and $\omega_i = \left(\omega_i\right)_{y_m}$. For applications where observations are assimilated over a large area most of the weights will be equal to 1 when $l\left[y, x_i; r\right] = 0$ in Eq. (12) (Shen et al., 2017). So the useful weights depend on the observations in the localization region defined by $l\left[y, x_i; r\right]$. Likewise, the normalization step ensures that the sum of weights is equal to 1. The most important features for the LPF are the sampling and merging. The observations y_j are assimilated one by one to calculate the local weights and combine the priori particles with sampled particles and calculate local samples from the distributions that will be introduced in Section 3.2 (Poterjoy, 2016).

3 The localized equivalent weights particle filter

3.1 *The statistical observation proposal density*

The EWPF ensures that the majority of particles are close to the observations and avoids filter degeneracy. However, it is difficult to obtain future observations in practical applications. Nevertheless, the reanalysis data can be used in practical studies to cover as much information as possible for the whole region and time period. Although these reanalysis data are still inaccurate compared to the actual observations and it is difficult to accurately represent the real-time observations. Therefore, the new algorithm calculate statistical observations by averaging the historical reanalysis data over time steps in order to reduce errors while adjusting particles. So this research use statistical observations to calculate the proposal density used in the EWPF (EWPF-Sobs).

The proposal density uses τ to determine the time when particles are adjusted toward the future observation in the Section 2.1. As shown in the Fig 1, where t^n is the next observation time, t^j is the current time and $t^n - t^0$ is the assimilation interval.

The reanalysis data are the most comprehensive in both space and time, so this paper use historical reanalysis data to calculate statistical observations and it adopts the proposal density parameter τ determine the time to start the cumulative observations. To avoid confusion, using $y_{r,}^{t^0} \cdots y_{r,}^{t^a} \cdots, y_{r,}^{t^n}$ to denote the historical reanalysis data of the time-series distribution of assimilated observations over the assimilation interval. The t^a determined from τ is the time to start accumulating the reanalysis observations, so the statistical observation at the time t^b is given by

$$y_s^{t^b} = \frac{\sum_{d=t^a}^{t^b} y_r^d}{(t^b - t^a) + 1} \quad \left(t^a \leqslant t^b \leqslant t^n\right). \tag{18}$$

Based on the τ , the time $t^b \left(t^a \leqslant t^b \leqslant t^n\right)$ is used to calculate the proposal density and determine the time of the statistical observations which use reanalysis data in the assimilation interval. The particles and weights at the time t^b are written as $(x_i)^{t^b}$ and $(\omega_i)^{t^b}$.

There are two steps to make the particles move toward the statistical observation that use the proposal density before the time of assimilation:

(1) Calculate and normalize the proposal density weights based on the statistical observation $y_s^{t^b}$ and Eq. (6).

$$(\omega_i)^{t^b} = \frac{1}{A} p\left(y_s^{t^b} | x_i^{t^b}\right) \frac{p\left(x_i^{t^b} | x_i^{t^b-1}\right)}{q\left(x_i^{t^b} | x_i^{t^b-1}, y_s^{t^b}\right)} \quad \left(t^a \leqslant t^b \leqslant t^n\right). \tag{19}$$

(2) The particles are moved towards the statistical observation $y_s^{t^b}$ and updated the model Eq. (7) at time t^b.

$$(x_i)^{t^b} = f\left(x_i^{t^b-1}\right) + B(\tau)\left(y_s^{t^b} - h\left(x_i^{t^b-1}\right)\right) + \widehat{d\beta_i^{t^b}}. \tag{20}$$

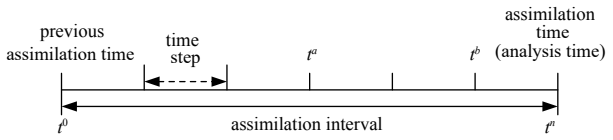

Fig. 1. Schematic showing the assimilation interval in terms of the model integration step. Here t^a is the statistical observation start time and t^n is the analysis time.

At the assimilation moment t^n the proposal density also uses the reanalysis data to update the new statistical observation $y_s^{t^n}$ to update the states and weights. The equivalent-weights proposal density retains the formulas of the traditional algorithm. Using Eq. (10) for the proposal density weights to determine the target weights and use Eq. (11) to update the final particles.

3.2 *The local equivalent weights particle filter*

The core of the LPF which proposed by the Poterjoy is to solve the filter degeneracy using a localization scheme. The localization scheme of LPF is mainly to suppress the long-range correlation due to insufficient particles. However, the EWPF method uses the proposed density adjust particles to improve the problems of traditional particle filter and has reliable assimilation results with complete observations. But when there are few available observations, it is difficult to improve the overall assimilation results by model tuning alone (Chen et al., 2020), so the localization scheme with reference to the LPF method is designed to improve the state of the particles at the location where observations are missing.

The new method still choose the EWPF method to calculate the adjustment of the particles at the corresponding positions of the observations, and adjust the particles at other positions according to the localization scheme to ensure the best possible application for the observations. The EWPF-Sobs uses Eq. (9) to obtain the scalar weights ω_i^n for the ith state vector and generates local samples based on the local distribution when y_j is used to update the local weights.

The states and weights until the observation point y_{j-1} are written as $x_m^{(y_j-1)}$ and $\omega_{y_{j-1}}^m$ for $m = 1, 2, \cdots, N_e$. The EWPF-Sobs with localization function (LEWPF-Sobs) is briefly described as follows:

(1) Use the statistical observation $y_s^{t^n}$ and Eq. (6) to determine the proposal density weights ω_i.

(2) Based on the final proposal density weights, use Eqs (9)–(11) to calculate the final model states $\left(x_i^{t^n}\right)_{\text{eq}}$ that correspond to the observation position $y_j^{t^n}$.

(3) Use the $\left(x_i^{t^n}\right)_{\text{eq}}$ and Eq. (12) to obtain the weights and use standard resampling to obtain the resampling indices $k_1, k_2, \cdots, k_{N_e}$ and these resampled particles are shown as x_{k_m}.

(4) Update the vector weights $(\omega_i^m)_{y_j}$ around the observed position according to y_j and the normalization factor for each j using $S_j = \sum_{m=1}^{M} (\omega_i^m)_{y_j}$.

(5) Combine the prior particles and resampled particles using the formula:

$$x_n^{(y_j)} = \bar{x}^{(y_j)} + r_{1,i} \circ \left(x_{k_m}^{(y_{j-1})} - \bar{x}^{(y_j)}\right) + r_{2,i} \circ \left(x_m^{(y_j-1)} - \bar{x}^{(y_j)}\right). \tag{21}$$

The parameters $r_{1,i}$ and $r_{2,i}$ that decide which of the priori and resampled particles are retained and calculated as follows (Poterjoy and Anderson, 2016):

$$r_{1,i} = \sqrt{\frac{\sigma_j^{(y_j)^2}}{\frac{1}{N_e}\sum_{m=1}^{N_e}\left[x_{k_{m,i}}^{(y_{j-1})} - \bar{x}^{(y_j)} + c_j\left(x_{m,i}^{(y_{j-1})} - \bar{x}_i^{(y_j)}\right)\right]^2}}, \tag{22}$$

$$r_{2,i} = c_i r_{1,i}, \tag{23}$$

$$c_j = \frac{N_e\left(1 - l\left[x_i, y_j, r\right]\right)}{l\left[x_i, y_j, r\right]\ \alpha W}, \tag{24}$$

where $(\sigma_i)^2_{y_j} = \sum_{m=1}^{N_e} \frac{(\omega_i^m)_{y_j}}{S_i}\left(x_j^m - \left(\bar{x}_i\right)_{y_j}\right)^2$ is the weighted variance of the prior particles and $\tilde{W} = \sum_{n=1}^{Ne} p\left(y_j | x_n^{y_{j-1}}\right)$. The derivations of Eqs (21)–(24) may be found in Poterjoy (2016). The formula for locally adjusted particles is derived to make the mean and variance of the particles the same as the weighted mean and weighted variance of the prior particles. These adjustments are based on Eq. (21), which uses the sampling and merging to update the weights and generates new particles when particles have a single observation.

4 Numerical experiments

4.1 *The Lorenz 96 model*

Due to the complexity of model physical processes, this research first investigate the quality and characteristics of the EWPF, EWPF-Sobs and EAKF in the Lorenz 96 with the standard observation distribution. Secondly, using Lorenz 96, this paper further verify the localization characteristics of the LPF, LEAKF and LEWPF-Sobs using the sparse observations. This paper also research the improvement of localization for EWPF-Sobs comparing LEWPF-Sobs and EWPF-Sobs. The following is a brief introduction to the simple model used in the paper.

The Lorenz 96 model contains N_x equally spaced variables x_j that evolve according to the equations:

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = \left(x_{j+1} - x_{j-2}\right) x_{j-1} - x_j + F, \tag{25}$$

where $j = 1, 2, \cdots, N_x$ represents the spatial coordinates. The cyclical boundaries are $x_{j+N_x} = x_j$ and $x_{j-N_x} = x_j$. Standard parameters F and N_x determine the system dynamics and are fixed at 8 and 40, respectively. The model is integrated using the Runge–Kutta scheme with a time step of 0.05 time units.

4.2 *Experimental set up*

In this study, perfect and biased experiments are designed in the twin framework. In both experiments observations are provided by the truth model that is integrated from initial conditions that are set as (Shen et al., 2017):

$$x_j = \begin{cases} 1 & \text{for } j \neq 1 \\ 1.001 & \text{for } j = 1 \end{cases}. \tag{26}$$

This is integrated through 100 000 iterations for sufficient spin-up and then for another 50 000 iterations to obtain the truth. In the twin experiments, a "truth" model using the standard parameter to generate the truth model states (x^t) and the nonlinear observations that are related to the truth state by the function $y = H(x^t) + \varepsilon$, where ε is white noise whose mean value is 0 and variance is 2. The model state at the new time step is given by:

$$x_i^j = f\left(x_i^{j-1}\right) + d\beta_i^j, \tag{27}$$

the distribution of the $d\beta_i^j$ can be chosen appropriately for the model, it is assumed to be Gaussian with mean zero and covariance Q whose value is 1 (Ades and Van Leeuwen, 2013).

Three separate sets of observations are used in the twin experiments to verify the effect of nonlinear properties (Zhang, 2011). They differ in the specification of H, the function that maps the model space to observation space: in the first, $H(x) = x$, in the second $H(x) = |x|$ and in the third $H(x) = \ln(|x|)$ to interpolate the values. Each assimilation experiment is integrated for 50 000 iterations and only those data obtained in the final 10 000 iterations are used for the error statistics. Due to the simplicity of this experimental model, the historical reanalysis data used in the experiments are constructed with reference to the observations, and the reanalysis observations are constructed by adding white noise to the truth values based on the 40 000 iterations. In this experiment, the assimilation interval is chosen to be 13 time steps (Poterjoy, 2016) and the assimilation observations are obtained for assimilation every 13 steps. The reanalysis data are used in the time series of the assimilation interval, and this study assume that there is available historical reanalysis data at each moment in the assimilation interval for the computation of statistical observations. The biased experiment is similar to the perfect framework and starts from the same initial conditions used in the perfect framework. The biased experiments is designed, which assumes that the imperfect of the model is caused by inaccurate parameter, so the parameter $F = 5$ is used in this biased experiment.

Validation of the localization scheme is performed in the framework of the second sparse observation twin experiments. This study assume that observations are only available to x_i when i is an odd number. Then each assimilation experiment is repeated 20 times starting from 20 different initial particles and their means and characteristics are analyzed.

The root-mean-square-error (RMSE) is an important criterion for evaluating assimilation quality:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_x}\sum_{i=1}^{N_x}(x_i - x_i^t)^2}. \tag{28}$$

The RMSE is used to evaluate the assimilation results and both experiments are performed using the same initial particles.

4.3 *Experimental results for standard observations*

The value of τ used to establish the time to start calculating the statistical observations and proposal density is explained in Section 3.1. Determining a suitable τ can effectively improve the assimilation quality of the EWPF-Sobs with different observation operators.

Based on Section 3.1, Fig. 2a shows results for the RMSE in the perfect experiments: using the observation operator $H(x) = x$ (green point), $\tau = 0.4$ (red point) has the lowest RMSE; when the $H(x) = |x|$ (dark blue point), $\tau = 0.6$ has the lower result and using the observation operator $H(x) = \ln|x|$ (light blue point), $\tau = 0.6$ is also the threshold value. In the biased experiments in Fig. 2b, the RMSE is lowest when $\tau = 0.6$ is used with observation operators $H(x) = |x|$ and $H(x) = \ln|x|$, and when the operator $H(x) = x$, the RMSE is lowest for $\tau = 0.7$.

The threshold value is therefore mostly close to $\tau = \left(t^j - t^0\right) / \left(t^n - t^0\right) = 8/13 \approx 0.6$. If approximate value of τ is too small, more observations are used in the statistics and more observational errors are included in the statistical observations and the proposal density, thus affecting the final assimilation result. So this research use the best thresholds chosen from Fig. 2 for the subsequent experiments.

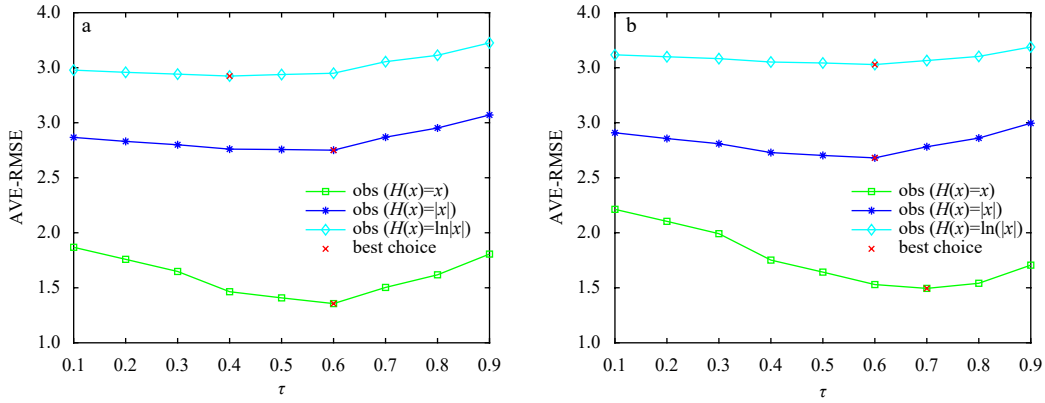

Fig. 2. Average RMSE with different threshold values τ for the three different observation operators in the perfect experiments (a). Green: observation operator $H(x) = x$; dark blue: observation operator $H(x) = |x|$ and light blue: $H(x) = \ln|x|$. Red dots represent the optimal solutions for the threshold. Biased model experiments using the same observation operators are shown in (b).

In order to compare the assimilation specialties of EWPF-Sobs, EWPF and EAKF under the same conditions, the three algorithms use the same 20 particles to investigate the average RMSE in the L96 model. To mitigate the effect of random initial system synthesis, 10 experiments were conducted for different initial particles. Figure 3 shows the average RMSE in the final 5 000 model steps in the twin experiment as histograms. The EAKF (dark blue column) has the lowest average RMSE for the $H(x) = x$ observation operators (Figs 3a and b) in the twin framework experiment. Therefore, the EWPF-Sobs (light blue column) has lower average RMSE than the EWPF (yellow) and EAKF (dark blue column) with both the observation operators $H(x) = |x|$ and $H(x) = \ln|x|$ (Fig. 3a), and there are similar results for the same operators in the biased framework (Fig. 3b). From the experimental results, it can be seen that the EAKF method is more suitable for the observation of Gaussian distribution while EWPF is more suitable for the observation of non-Gaussian distribution. However, the EWPF-Sobs reliable assimilation results with different observation operators.

To further demonstrate the advantage of the EWPF-Sobs, this study compute the average RMSE and the average Spread, which is determined by the ensemble spread over the final 2 000 steps. Figure 4 shows the corresponding RMSE/Spread for the three algorithms with the observation operator $H(x) = x$. In Fig. 4a the ratio for EWPF-Sobs (blue line) is closer to one than that for the EWPF (green dot–dash line) and EAKF (pink dash line). In the biased experiments, the results of EAKF are more stable than the EWPF-Sobs, but both of the two algorithms have similar results which is little better than the traditional EWPF. The two figures demonstrate that EWPF-Sobs is generally better than the EWPF.

4.4 *Results of the localization experiments*

In the LEWPF-Sobs, the localization function is used with spare observation to calculate the likelihood of the observations given the particles that are near this observation point. According to the experimental program presented in Section 4.2 to study the localization scheme, the sparse observation distribution is constructed assuming that odd variables in the Lorenz 96 have observation for the LEWPF-Sobs. The value of τ determined the time to start calculating the statistical observations before the subsequent experiments. Calculate the RMSE of all odd variables in the experiment according to equation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{20}\sum_{a=1}^{39}(x_a - x_a^t)^2}, \tag{29}$$

where a is the odd number and it choose the most appropriate τ for subsequent experiments based on the final results which are shown in Fig. 5. It shows the relationship between τ and RMSE in

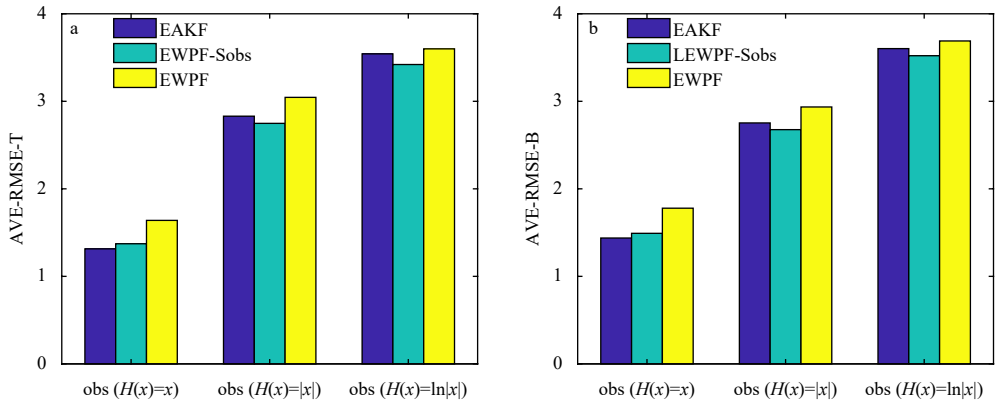

Fig. 3. Average RMSEs of the EWPF (yellow), EWPF-Sobs (light blue) and EAKF (dark blue) expressed as histograms based on the model states during the final 5 000 steps. a. Perfect experiments that use the observation functions $H(x) = x$, $H(x) = |x|$ and $H(x) = \ln|x|$. b. Same as a but for the biased experiments.

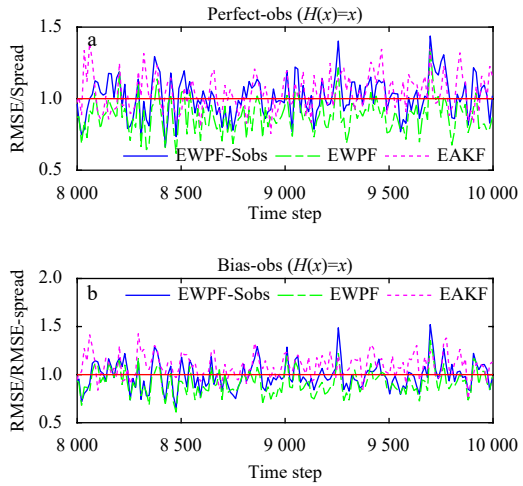

Fig. 4. Time series of the RMSE/Spread values of the EWPF-Sobs (blue line), EWPF (green dot–dash line) and EAKF (pink dash line) in the perfect experiment during the final 2 000 time steps (a). b. As a but for the biased experiments.

twin experiments using three different observation operators.

In the perfect experiments Fig. 5a, the RMSE is lowest when $\tau = 0.4$ is used with observation operators $H(x) = |x|$ and $\tau = 1/13 \approx 0.1$ for the $H(x) = \ln|x|$, and when the operator $H(x) = x$, the RMSE is lowest for $\tau = 0.5$. Figure 5b shows the $H(x) = x$ use the $\tau = 0.3$, $H(x) = |x|$ use the $\tau = 0.2$ and $\tau = 0.1$ is used with $H(x) = \ln|x|$ in the bias experiments.

Before further experiments, this study also should determine the effect of the radius (c) which has been described in the Section 2.3. It is related to the decorrelation, the distance between the observational site and the grid point (Shen et al., 2017) on the final average RMSE with three different observation operators. Figure 6 shows the average RMSE with the LEWPF-Sobs using 20 particles in the twin experiments for the observation radius $c \in \{1, 2, 4, 8\}$ with different observation operators $H(x) = x$ (green line), $H(x) = |x|$ (blue line) and $H(x) = \ln(|x|)$ (red line). The average RMSE increases with the parameter c: when $c = 1$ (black point) the LEWPF-Sobs has the lowest RMSE with different operators in the perfect experiment (Fig. 6a). The localization parameter $c = 1$ (black point) is also the best choice in the biased experiments (Fig. 6b). The main reason is that using a smaller localization parameter c introduces less observation error. So for this ensemble size $c = 1$ is preferable.

In the subsequent experiments use the localization parameter $c = 1$ to compare the LEWPF-Sobs and LPF with the LEAKF. In order to further analyze the improvement of the localization method on the EWPF, the EWPF-Sobs was added in this twin experiment for the same initial fields and the same 20 particles conditions. The experiments are duplicated 10 times with different initial particles and the average RMSEs of the last 5 000 model steps are shown as histograms in Fig. 7. The LEAKF (dark blue) has the lowest average RMSE when the observation operator is $H(x) = x$ in the twin experiments. However, the LEWPF-Sobs (light blue) has the lowest average RMSE for non-Gaussian distribution observations in the perfect experiments (Fig. 7a). In the biased experiments (Fig. 7b), the LEWPF-Sobs still maintains better results than the LPF (green column) and LEAKF for non-Gaussian distribution observations, while the LEAKF also keeps lower RMSE than the PF when use the Gaussian distribution observations. Since the experiment uses a sparse observation distribution, the EWPF-Sobs (yellow column), which lacks the localization scheme, maintains a high RMSE in these experiments.

In order to further analyze the necessity of using the localization scheme, this experiments choose the variables that lacks observations to study the adjustment of the localization scheme on the particles during the assimilation process. The construction of the sparse observation distribution is described in Section 4.2. In the experiments it is assumed that the odd number of variables in the 40 variables of the Lorenz-96 model are short of observations. In order to analyze the effect of the localization scheme for the lack of observed variables, the RMSE is calculated for all even variables in the model which use the observation operator $H(x) = x$ with three different algorithm. This part of results are given in Figs 8a and b. It clearly shows that the LEWPF-Sobs (blue circle line) and LPF (green rhombus line) have similar results and both of them better than EWPF-Sobs (brown cross line). The main reason is there is no localization scheme for EWPF-Sobs and it is difficult to adjust the particles at the missing locations, so the RMSE is significantly higher than other methods.

To further analyze the adjustment effect of the localization scheme for all unobservable even variables with different observation operator. This part of experiments use four different localization assimilation algorithms to calculate the RMSE for all even variables under different observation operators according to the Eq. (29) and the final results are shown in Figs 9a and b. It shows

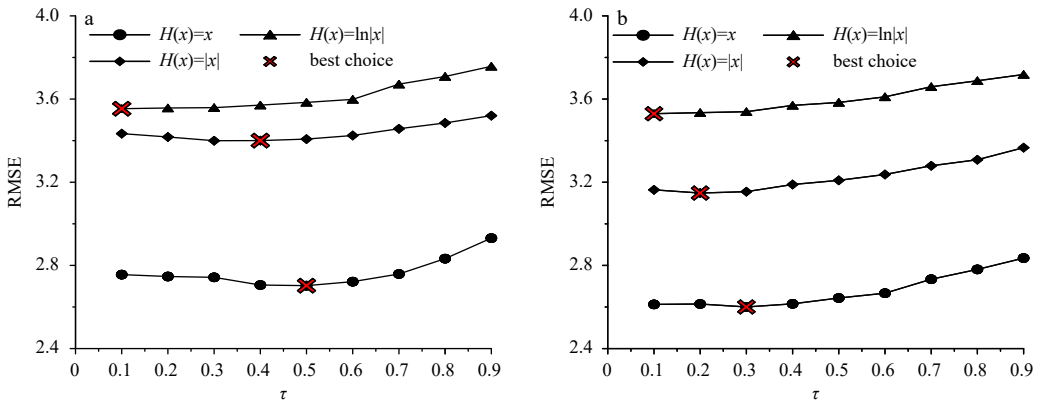

Fig. 5. Average RMSE with different threshold values τ for the three different observation operators in the perfect experiments (a). Circle: observation operator $H(x) = x$; diamond: observation operator $H(x) = |x|$ and triangle: $H(x) = \ln|x|$. Red cross represents the optimal solutions for the threshold. Biased model experiments using the same observation operators are shown in b.

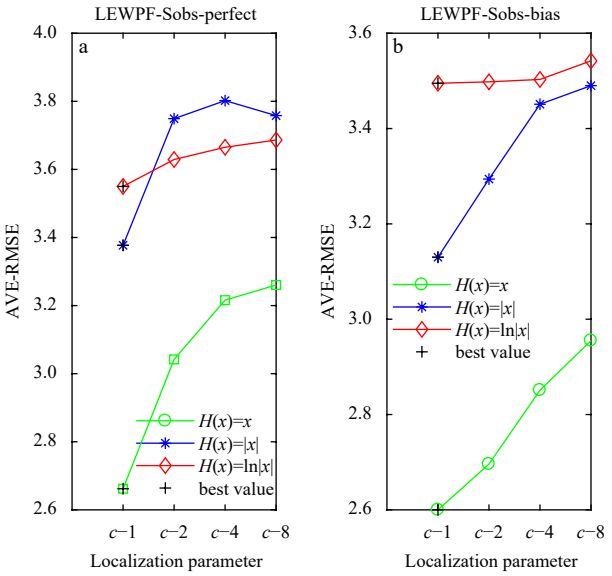

Fig. 6. Average RMSE with different localization parameters. a. Perfect experiment using the observation functions $H(x) = x$ (green line), $H(x) = |x|$ (blue line) and $H(x) = \ln|x|$ (red line). The black point represents the best results. b. Same as (a) but for the biased experiment.

the LEWPF-Sobs (pink column) has the lower RMSE than the LEAKF (grey column) and LPF (light blue column) which both of them have similar results. Because of the lack of adjustment of the localization scheme, the assimilation results of EWPF-Sobs (green column) are significantly higher than the other methods with different observational operators. It can be seen that the localization scheme used in EWPF-Sobs has a moderating effect on the unobserved particles.

Based on previous studies, the LPF method is sensitive to the number of particles, but the main advantage of EWPF is that it uses a small number of particles to obtain better results. To further demonstrate the characteristics of the LEWPF-Sobs, Fig. 10 illustrates the average RMSE obtained by the LPF (red line) with different ensemble sizes and compares it with those obtained with the LEAKF (blue line) and LEWPF-Sobs (pink line) assimilation results using 20 particles. These experiments are also repeated 10 times to reduce the random error. The average RMSEs and standard deviations are expressed as solid lines and error bars in this figure. Figure 10a shows that, with the observation function $H(x) = \ln|x|$ in the perfect experiments, the RMSEs of the LPF and LEAKF seem similar to the LEWPF-Sobs results when the number of particles is smaller than 80, but when 160 particles are used, the LPF is better than the LEWPF-Sobs. The results for the biased experiments are shown in Fig. 10b; LEWPF-Sobs has lower RMSE than the LPF using 160 particles and the LEWPF-Sobs is clearly better than the LEAKF when using the same number (20) of particles. When the LPF use 80 particles, the

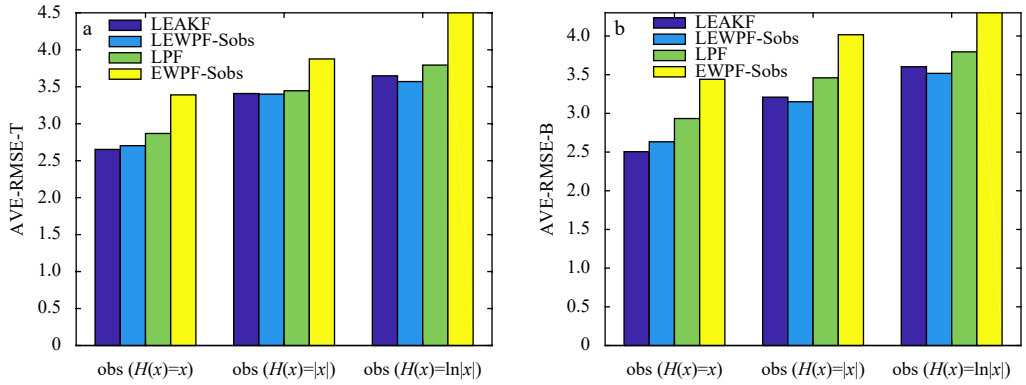

Fig. 7. Average RMSEs over the final 5 000 model steps of the LEAKF (dark blue), LEWPF-Sobs (light blue), LPF (green), and EWPF-Sobs (yellow) expressed as histograms. a. Perfect experiments that use the observation functions $H(x) = x$, $H(x) = |x|$ and $H(x) = \ln|x|$. b. Same as a but for the biased experiments.

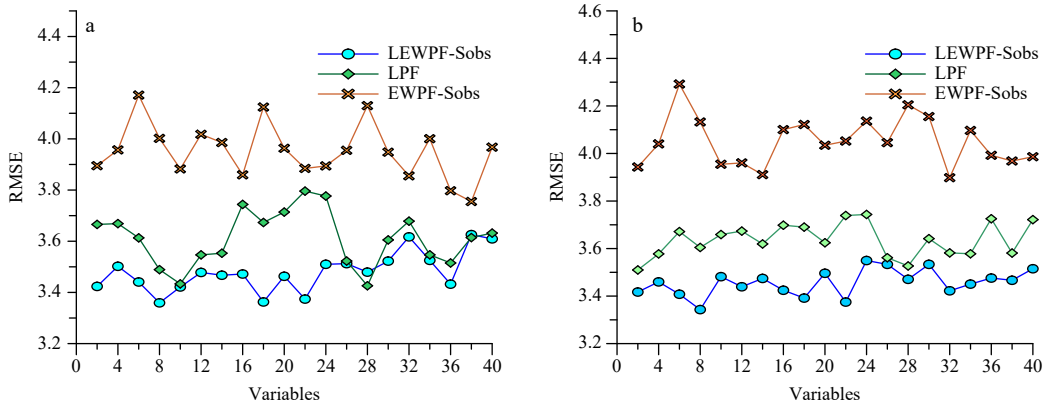

Fig. 8. The RMSE values with all even variables of LEWPF-Sobs (blue circle line), LPF (green rhombus line), and EWPF-Sobs (brown cross line). a. Perfect experiments that use the observation operator $H(x) = x$. b. Same as a but for the biased experiments.

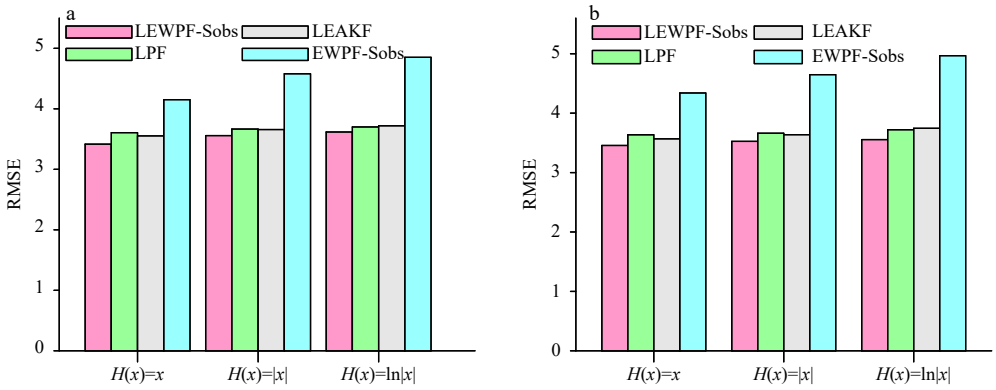

Fig. 9. The average RMSE with missing observations variables of LEWPF-Sobs (pink), LPF (green), LEAKF (grey) and EWPF-Sobs (light blue) expressed as histograms when use the observational function $H(x) = x$, $H(x) = |x|$ and $H(x) = \ln|x|$ (a). b. Same as a but for the biased experiments.

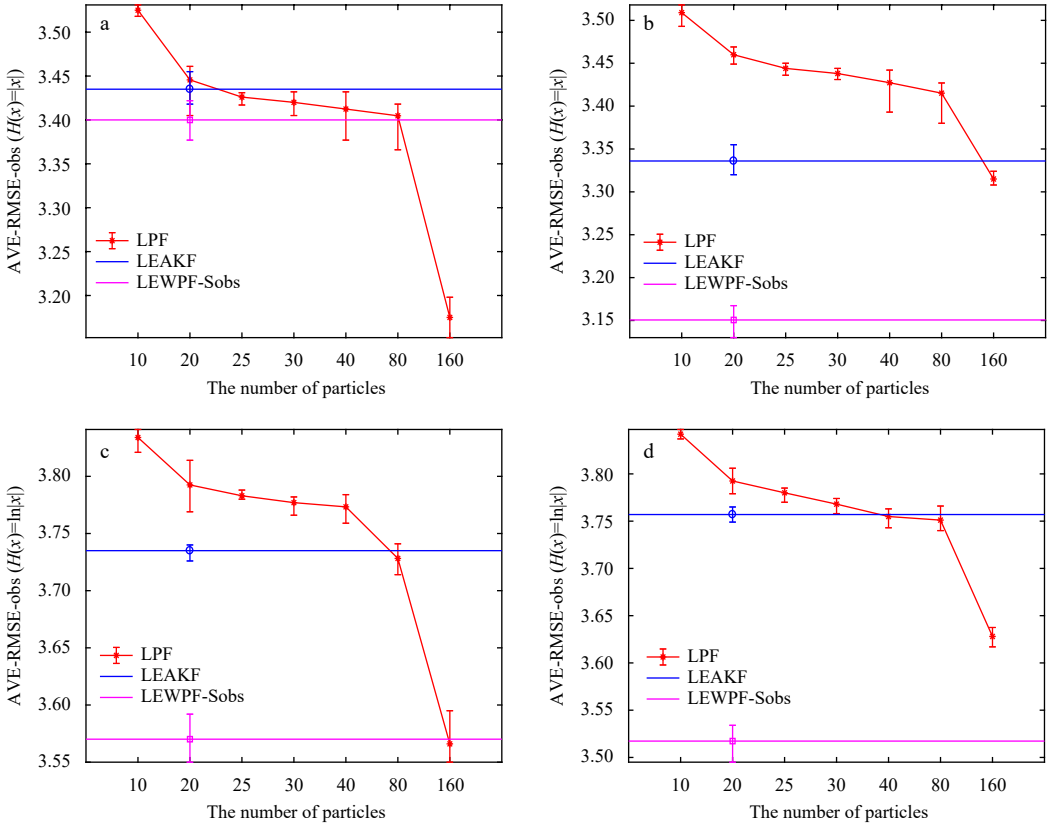

Fig. 10. Average RMSEs of the data assimilation with LEAKF (blue line) and LEWPF-Sobs (pink line) that only use 20 particles, and for different ensemble sizes of the LPF (red line). a, c. Results in the perfect experiments for the different algorithms using the observation function $H(x) = |x|$ and $H(x) = \ln|x|$, respectively. b, d. Same as a and c but for the biased experiments.

assimilation result is similar to the LEAKF in the twin experiments. Similar results are obtained with the observation operator $H(x) = \ln|x|$ and are shown in Fig. 10c for the perfect experiments and Fig. 10d for the biased experiments.

The research also use the relationship between Spread and RMSE to evaluate the LPF, LEAKF and LEWPF-Sobs. The scatter plots of the mean Spread relative to the RMSE for the 20 variables used in the sparse observation experiments are shown in Fig. 11 for the three algorithms in the twin experiments. The average Spread is acquired by the ensemble spread over the final 5 000 steps. Figure 11 is consistent with Fig. 4 and shows the LEWPF-Sobs and the LEAKF have a ratio closer to 1 in the perfect experiments and the LPF that uses 160 particles has the largest RMSE and Spread (Figs 11a–c). Figures 11d and e shows the same relationship in the bias experiments. The LEWPF-Sobs and LEAKF have a similar relationship, but both of them are better than the LPF. Comparison of these results reveals that the LEWPF-Sobs generates good ensemble particles and provides an estimate of uncertainty.

5 Summary and discussion

It is generally acknowledged that the EWPF is suitable for ap-

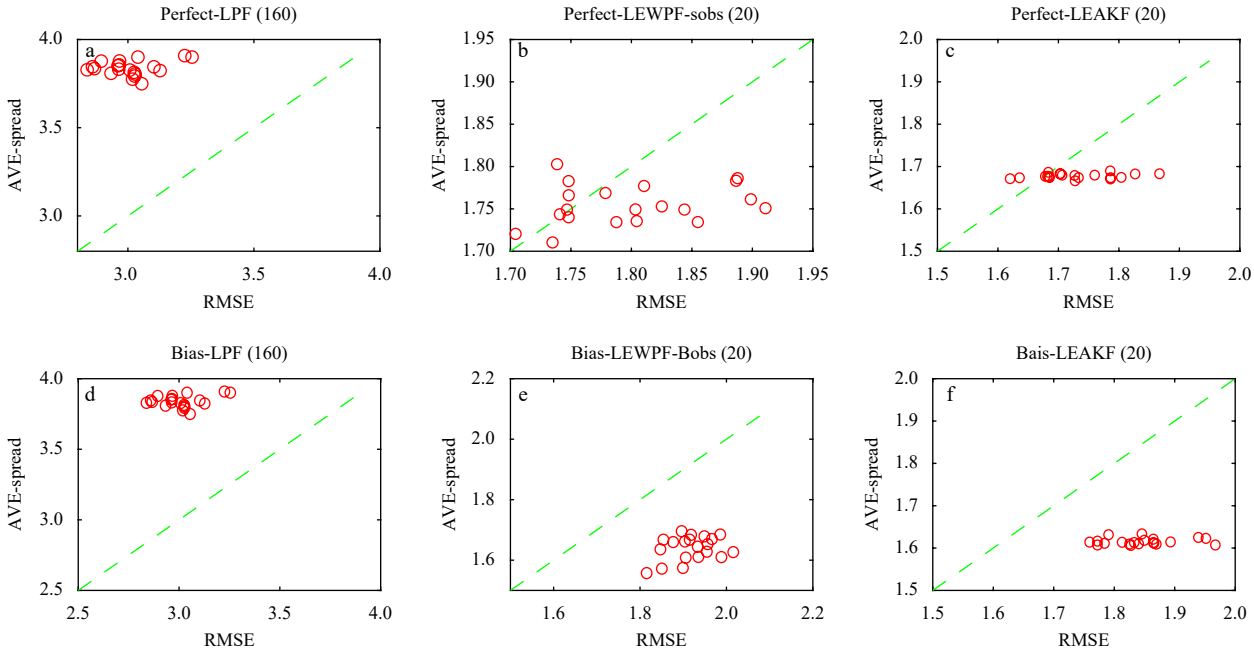

Fig. 11. Scatter plots of average Spread against the RMSE for the 20 variables. Results for the perfect experiments for the LPF (a), the LEWPF-Sobs (b) and the LEAKF (c). Dashed lines indicate when their ratio is 1. d–f. Same as a–c but for the biased experiments.

plication to nonlinear and non-Gaussian modes. However, the traditional EWPF relies on future observations to determine the proposal density and lacks a suitable localization function in high-dimensional models.

This paper analyzed the proposal density that relies on future observations as used in the EWPF algorithm and proposed the EWPF-Sobs that uses statistical observations which are determined by the historical reanalysis data instead of future observations to improve the EWPF's ability to perform real-time observational assimilation. Using the Lorenz 96 model to compare EWPF, EAKF and EWPF-Sobs assimilation quality and analyzed their performance using three different non-Gaussian observation operators in the twin experiments. The EWPF-Sobs have similar results to EAKF when use the Gaussian observations and the EWPF-Sobs gave better results than the traditional EWPF in the non-linear model. This shows that statistical observations can improve the EWPF's assimilation quality.

This paper also analyzed the localization function used in the LPF and used a similar localization function with the EWPF-Sobs to construct the LEWPF-Sobs. This part of research again used the Lorenz 96 model with sparse observations and conducted twin experiments using LEWPF-Sobs, LEAKF and LPF under the same experimental nonlinear conditions and non-Gaussian observation operator. Compare the EWPF-Sobs and LEWPF-Sobs in this experiment to further investigate the extent of improvement of the localization scheme for this method. The experiments show that the LEWPF-Sobs is better than the LEAKF and requires only 20 particles to give better results than the LPF, which uses 160 particles in the biased experiments. The LEWPF-Sobs using the localization scheme under sparse observation conditions is significantly better than the EWPF-Sobs in this twin experiments. The results prove the great potential of the LEWPF-Sobs for application in complex coupled models.

The present experiments use only the simple Lorenz 96 model. Further study of LEWPF-Sobs assimilation systems suitable for coupled cycle models will require the use of more complex models to demonstrate the application of the algorithms. The localization scheme may inevitably introduce discontinuity and imbalance to the model. Therefore, this localization scheme will be refined in the subsequent study to ensure the balance of the model. Nevertheless, the results obtained here with the LEWPF-Sobs will encourage further research and application of the corresponding PF methods.

Acknowledgements

Thanks go to Mengbin Zhu for his comments and suggestions at the early version of this manuscript. Conversations with Ting Zhao, Rixu Hao, Dequan Yang and Hengde Zhao led to modifications of many for the first version of the manuscript. And Harbin Engineering University and China Scholar Council (awarded to Deng Xiong for two and half years' study abroad at GFDL/UW-Madison/Ohio State University Visiting Program).

References

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, doi: 10.1002/qj.1995

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, doi: 10.1002/qj.2370

Chen Yan, Zhang Weimin, Wang Pingqiang. 2020. An application of the localized weighted ensemble Kalman filter for ocean data assimilation. Quarterly Journal of the Royal Meteorological Society, 146(732): 3029–3047, doi: 10.1002/qj.3824

Chorin A J, Tu Xuemin. 2009. Implicit sampling for particle filters. Proceedings of the National Academy of Sciences of the United States of America, 106(41): 17249–17254, doi: 10.1073/pnas.0909196106

De Freitas N, Andrieu C, Højen-Sørensen P, et al. 2001. Sequential monte Carlo methods for neural networks. In: Doucet A, De Freitas N, Gordon N, eds. Sequential Monte Carlo Methods in Practice. New York, NY, USA: Springer, 359–379

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans,

99(C5): 10143–10162, doi: 10.1029/94JC00572

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, doi: 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, doi: 10.1049/ip-f-2.1993.0015

Houtekamer P L, Mitchell H L. 1998. Data assimilation using an ensemble Kalman filter technique. Monthly Weather Review, 126(3): 796–811, doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2

Law K, Stuart A, Zygalakis K. 2015. Data Assimilation: A Mathematical Introduction. Cham, Switzerland: Springer

Lei Jing, Bickel P. 2011. A moment matching ensemble filter for nonlinear non-Gaussian data assimilation. Monthly Weather Review, 139(12): 3964–3973, doi: 10.1175/2011MWR3553.1

Lorenz E N. 1995. Predictability: a problem partly solved. In: Proceedings Seminar on Predictability. Reading, UK: ECMWF

Nakano S, Ueno G, Higuchi T. 2007. Merging particle filter for sequential data assimilation. Nonlinear Processes in Geophysics, 14(4): 395–408, doi: 10.5194/npg-14-395-2007

Poterjoy J. 2016. A localized particle filter for high-dimensional nonlinear systems. Monthly Weather Review, 144(1): 59–76, doi: 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, doi: 10.1175/MWR-D-15-0322.1

Robert S, Leuenberger D, Künsch H R. 2018. A local ensemble transform Kalman particle filter for convective-scale data assimilation. Quarterly Journal of the Royal Meteorological Society, 144(713): 1279–1296, doi: 10.1002/qj.3116

Shen Zheqi, Tang Youmin, Li Xiaojing. 2017. A new formulation of vector weights in localized particle filter. Quarterly Journal of the Royal Meteorological Society, 143(709): 3269–3278, doi: 10.1002/qj.3180

Shen Zheqi, Zhang Xiangming, Tang Youmin. 2016. Comparison and combination of EAKF and SIR-PF in the Bayesian filter framework. Acta Oceanologica Sinica, 35(3): 69–78, doi: 10.1007/s13131-015-0757-x

Stordal A S, Karlsen H A, Nævdal G, et al. 2011. Bridging the ensemble Kalman filter and particle filters: the adaptive Gaussian mixture filter. Computational Geosciences, 15(2): 293–305, doi: 10.1007/s10596-010-9207-1

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, doi: 10.1002/qj.699

Van Leeuwen P J. 2011. Efficient nonlinear data-assimilation in geophysical fluid dynamics. Computers & Fluids, 46(1): 52–58

Van Leeuwen P J. 2015. Aspects of particle filtering in high-dimensional spaces. In: First International Conference on Dynamic Data-Driven Environmental Systems Science. Cambridge, UK: Springer, 251–262

Van Leeuwen P J, Evensen G. 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. Monthly Weather Review, 124(12): 2898–2913, doi: 10.1175/1520-0493(1996)124<2898:DAAIMI>2.0.CO;2

Van Leeuwen P J, Künsch H R, Nerger L, et al. 2019. Particle filters for high-dimensional geoscience applications: a review. Quarterly Journal of the Royal Meteorological Society, 145(723): 2335–2365, doi: 10.1002/qj.3551

Whitaker J S, Hamill T M. 2012. Evaluating methods to account for system errors in ensemble data assimilation. Monthly Weather Review, 140(9): 3078–3089, doi: 10.1175/MWR-D-11-00276.1

Zhang S. 2011. A study of impacts of coupled model initial shocks and state–parameter optimization on climate predictions using a simple pycnocline prediction model. Journal of Climate, 24(23): 6210–6226, doi: 10.1175/JCLI-D-10-05003.1

Zhu Mengbin, Van Leeuwen P J, Amezcua J. 2016. Implicit equal - weights particle filter. Quarterly Journal of the Royal Meteorological Society, 142(698): 1904–1919, doi: 10.1002/qj.2784