A modified iterative ensemble Kalman filter data assimilation method

Baoxiong Xu, Yulong Bai*, Yizhao Wang, Zhe Li, Boyang Ma

College of Physics and Electrical Engineering, Northwest Normal University, Lanzhou 730070, China
*: Corresponding author

Abstract. High nonlinearity is a typical characteristic associated with data assimilation systems. Additionally, iterative ensemble based methods have attracted a large amount of research attention, which has been focused on dealing with nonlinearity problems. To solve the local convergence problem of the iterative ensemble Kalman filter, a modified iterative ensemble Kalman filter algorithm was put forward, which was based on a global convergence strategy from the perspective of a Gauss-Newton iteration. Through self-adaption, the step factor was adjusted to enable every iteration to approach expected values during the process of the data assimilation. A sensitivity experiment was carried out in a low dimensional Lorenz-63 chaotic system, as well as a Lorenz-96 model. The new method was tested via ensemble size, observation variance, and inflation factor changes, along with other aspects. Meanwhile, comparative research was conducted with both a traditional ensemble Kalman filter and an iterative ensemble Kalman filter. The results showed that the modified iterative ensemble Kalman filter algorithm was a data assimilation method that was able to effectively estimate a strongly nonlinear system state.

1 Introduction

As in highly nonlinear dynamical and observational models, introducing iterations to a Kalman filter has proven to be beneficial for estimation problems [7]. To improve the performance of the traditional ensemble Kalman filter (EnKF) [5], the iterative ensemble Kalman filter (IEnKF) was recently proposed [13]. At the cost of additional iterations and propagations of the ensemble, this IEnKF was shown to significantly outperform the EnKF, especially in cases with large time intervals between updates [3]. The IEnKF should be considered for use in cases where there is a well-constrained system with nonlinear growth during the assimilation cycle (the propagation was initially nonlinear but became linear through repeated iteration). Later on, Bocquet and Sakov (2012) proved that the IEnKF can be taken as a lag-one smoother and be further extended to a fixed-lag smoother: the iterative ensemble Kalman smoother (IEnKS). The IEnKS is an ensemble variational method that does not require the use of the tangent linear of the evolution or the adjoint of observation models. By using a nonlinear least square formulation, Gu and Oliver [6] were able to introduce an ensemble randomized maximum likelihood (EnRML) method. In an EnRML, the sensitivities are calculated using an ensemble based method. However, for high-dimensional problems, the ensemble approximation of the sensitivity matrix is often poor [4]. An iterated EnKF [9] was also proposed with a running in place scheme, which was referenced by Sakov [13] and applied by Penny [12] in a realistic Ocean General Circulation Model (OGCM) reanalysis. In further study of all iteration-based Ensemble Kalman filters, we found that one downside was that only when certain local conditions are satisfied can the

convergence reach an ideal value, and there exists the possibility of failure under conditions with large initial errors.

Because the Newton method, which was implicitly adopted [13], is only one of many minimization schemes to minimize the cost function of data assimilation systems, there is still a large degree of freedom in choosing an iterative scheme, such as the Levenberg-Marquardt algorithm [3]. In this study, an IEnKF with a quick local convergence is combined with a global convergence strategy for the purpose of achieving a modified iterative ensemble Kalman filter algorithm (modified IEnKF: MIEnKF) that leads to every iteration in the process of assimilation to tend towards the optimal value. Thus, even if the initial value was not accurately estimated, the higher-accuracy solution of the iterative convergence can still be obtained. Experimental tests were conducted on this algorithm with both ensemble Kalman and iterative ensemble Kalman filters in some typical, corresponding chaotic systems to confirm the algorithm.

2 Gauss-Newton derivation of the IEnKF

A Gauss-Newton iteration algorithm is an approximate application of a Newton iteration algorithm, which is an iterative method type for nonlinear equations. In this study, the updated problem of the EnKF observation was transformed to solve the minimum cost function, and the observation iterative equation was deduced from the angle of a Gauss-Newton iteration algorithm, which thereby proved that an IEnKF is a type of Gauss-Newton iteration process.

The IEnKF was introduced [13], in which the prior probability density function is $\mathbf{p}(\mathbf{x}\,|\,\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$:

$$x=\frac{1}{N}\sum_{l=1}^{N}\mathbf{x}_l \qquad P=\frac{1}{N-1}\sum_{l=1}^{N}(\mathbf{x}_l-\bar{\mathbf{x}})(\mathbf{x}_l-\bar{\mathbf{x}})^T=\frac{1}{N-1}[(\mathbf{E}-\mathbf{x}\mathbf{l}^T)/\varepsilon][(\mathbf{E}-\mathbf{x}\mathbf{l}^T)/\varepsilon]^T \tag{1}$$

where $\mathbf{x}_l$ indicates the l-th member forecast ensemble. It need not match the mean x_b, and the error covariance matrix B of the prior probability density function, $\mathbf{p}(\mathbf{x}\,|\,\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)$, E, is the ensemble of model states and $\varepsilon=const<<1$.

Suppose that we have the following dynamical system, then, the prediction and observation equations of the system are shown as follows:

$$\text{Prediction equation: } \mathbf{x}^f=M(\mathbf{x}) \tag{2}$$

$$\text{Observation equation: } y=h(\mathbf{x},\mathbf{R}) \tag{3}$$

where M indicates the nonlinear model operator of the system; R is the observation error covariance; the superscript f indicates the prediction value; and x indicates the model state vector; the effective prior probability density function satisfies:

$$\mathbf{p}(\mathbf{x}\,|\,\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)\propto N(\mathbf{x};\mathbf{x}^f,\mathbf{P}^-) \tag{4}$$

According to the Bayes theorem, the posteriori probability of the system state can be shown as follows:

$$\mathbf{p}(\mathbf{x}\,|\,\mathbf{y},\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)\propto\alpha\mathbf{p}(\mathbf{y}\,|\,\mathbf{x})\mathbf{p}(\mathbf{x}\,|\,\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) \tag{5}$$

where α is the normalized constant and y indicates the observation vector at a certain update step. $\mathbf{p}(\mathbf{y}\,|\,\mathbf{x})$ is the likelihood probability and satisfies:

$$\mathbf{p}(\mathbf{y}\,|\,\mathbf{x})\propto N(\mathbf{y};h(\mathbf{M}(\mathbf{x}^f)),\mathbf{R}) \tag{6}$$

where h() is the observation operator and R is the observation error covariance matrix.

Therefore, while solving the maximum posteriori estimation of x, α can be ignored, and the cost function of the state variable can be shown as follows:

$$J(\mathbf{x})=\frac{1}{2}((y-h(M(\mathbf{x}^f))^T\mathbf{R}^{-1}(y-h(M(\mathbf{x}^f)))+(\mathbf{x}-(\mathbf{x}^f)^T\mathbf{P}^{-1}(\mathbf{x}-\mathbf{x}^f))) \tag{7}$$

When $\mathbf{Y}=[\mathbf{y}^T\ \ (\mathbf{x}^f)^T]^T$, then it obeys a Gaussian distribution, the mean is $u(\mathbf{x})=[(h(M(\mathbf{x}^f)))^T\ \ \mathbf{x}^T]^T$, and with the variance being $\boldsymbol{\sigma}=diag([\mathbf{P},\mathbf{R}])$, the cost function is equivalent to the following:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{Y} - \mu(\mathbf{x}))^T \sigma^{-1}(\mathbf{Y} - \mu(\mathbf{x})) \tag{8}$$

Moreover, when $\mathbf{S}^T\mathbf{S} = \boldsymbol{\sigma}^{-1}$, $l(x) = \mathbf{S}(\mathbf{Y} - \mu(x))$, then the following can be obtained:

$$J(\mathbf{x}) = \frac{1}{2}\left\| l(\mathbf{x}) \right\|^2 \tag{9}$$

According to Eq. (9), the IEnKF update can be transferred to the solution of a nonlinear least squares method, and the Gauss-Newton iterative algorithm can be used to obtain the solution of Eq. (9) as follows:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - (l'(\mathbf{x}_i)^T l'(\mathbf{x}_i))^{-1} l'(\mathbf{x}_i)^T l(\mathbf{x}_i) \tag{10}$$

where $l'(\mathbf{x}_i)$ is the Jacobian matrix of $l(\mathbf{x}_i)$; i is the iteration quantity; $l'(\mathbf{x}_i) = -\mathbf{S}[(\mathbf{HM})^T \mathbf{I}]^T$; $\mathbf{H}$ is the Jacobian matrix of $h(x)$; and $\mathbf{I}$ is the unit matrix. Then, when substituted into Eq. (10), the following can be obtained:

$$\mathbf{x}_{i+1} = \mathbf{x}^f + \mathbf{P}^{i+1}(\mathbf{HM})^T(\mathbf{R})^{-1}\{\mathbf{y} - h[M(\mathbf{x}_i)] + \mathbf{HM}(\mathbf{x}_i - \mathbf{x}^f)\} \tag{11}$$

When the gain is set as $\mathbf{K} = \mathbf{P}^{i+1}(\mathbf{HM})^T(\mathbf{R})^{-1}$, then:

$$\mathbf{x}_{i+1} = \mathbf{x}^f + \mathbf{K}\{\mathbf{y} - h[M(\mathbf{x}^i)] + \mathbf{HM}(\mathbf{x}_i - \mathbf{x}^f)\} \tag{12}$$

This equation is the updated iterative formula of the iterative ensemble Kalman filter.

It was determined that the observed iteration of a traditional iterative ensemble Kalman filter algorithm was a Gauss-Newton iterative process.

Due to the large spread of initial value deviations, as well as other factors during the process of data assimilation, the most recent iteration value was chosen as the final value in every iteration of the original IEnKF, and the new iteration's result was considered to be closer to the expected value when compared to the former iteration. However, this method may reduce the convergent performance in certain applications [13].

3 A Modified IEnKF scheme

The Gauss-Newton iteration algorithm was found to have stronger local convergence. When the initial value was close enough to the optimal value, the convergence of the iterative sequences was ensured. During this process, the Newton step is always used for examining whether it is closer to the expected value by iterations. However, when the initial estimate is quite different from the expected value, the iteration of the Newton step may lead to a greater deviation from the expected value. In this study, through adding a step factor, λ, the value of the step factor will adaptively adjust every iteration to ensure that each step of the iteration tends towards the expected. It has been proven that a Gauss-Newton method with a step factor λ presented a general convergence [14].

To determine the step factor, the cost function J (x) was minimized with reference to search step factor λ_i along the Newton direction, enabling J (x) to be minimized along the Newton direction, and so on, as follows:

$$J(\mathbf{x}_i + \lambda_i \delta_i) = \min_{i \geq 0} J(\mathbf{x}_i + \lambda \delta_i) \tag{13}$$

where $\delta_i = -J(\mathbf{x}_i)/J'(\mathbf{x}_i)$. Through solving Eq. (10), it was found that the multiple step factor λ met the conditions, so it could be used to calculate $J(\mathbf{x}_i + \lambda\delta_i)$, which corresponded to each λ many times, and the λ value was then taken to enable the $J(\mathbf{x}_i + \lambda\delta_i)$ minimum through comparison. Due to the multiple calculations for the status value, a large amount of calculating was consumed by this process in each iteration. When the data assimilation system showed higher requirements for real time, it would determine the step factor λ according to the J (x) gradient descent direction.

In regards to the use of the Newton step, once it approached the expected value, a super-linear convergence rate was achieved. In each iteration, it was necessary to test whether the current step reduced the error or not. If the answer was negative, back-tracking was conducted along the Newton

direction, until a step meeting the conditions was located. This step did not necessarily support the J(x) minimum in such a direction, but was the needed step to be able to make it decline. Because the Newton direction was downward for the J(x) direction, an acceptable step needed to be found in each of the back-tracking processes. Then, the following was obtained:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \lambda\delta_i,\ 0 < \lambda < 1, \tag{14}$$

The comparison between $J(\mathbf{x}_i)$ and $J(\mathbf{x}_{i+1})$ was used to decide whether to choose the step factor obtained by the back-tracking. If inadequately met, the $J(\mathbf{x_{i+1}}) < J(\mathbf{x_i})$ criterion could potentially lead to a failure in the convergence of the $J(\mathbf{x})$ to the minimum in certain cases. As demonstrated by the experiments of this study, a simple method was determined to enable the average decline rate of the $J(\mathbf{x}_{i+1})$ to at least reach α of $\nabla J^T \bullet \delta_i$ (Press et al., 1988) as follows:

$$J(\mathbf{x}_{i+1}) \le J(\mathbf{x}_i) + \alpha\nabla J^T \bullet \delta_i \qquad 0 < \alpha < 1 \tag{15}$$

In an actual situation, $\alpha = 10^{-4}$ should be realized. Meanwhile, according to the specific situation, the step factor should be as large as possible. Through the back-tracking step, the easiest way was to halve the step factor in each of the back-tracking processes. Only when λ was approaching 1, could the super-linear convergence rate of the Newton iterative method be obtained.

Based on the above global convergence strategy, it was determined that the Newton iterative algorithm may cause a larger norm of the system output value at the point of the new iteration value, under the local condition of not satisfying Newton's lemma. For example, the new iteration value deviated more from the optimal value, which sometimes resulted in such divergence as $|J(\mathbf{x}_{i+1})| > |J(\mathbf{x}_i)|$. Therefore, a Newton iterative algorithm with a global convergence strategy was used to solve this problem. If $|J(\mathbf{x}_{i+1})| > |J(\mathbf{x}_i)|$, $(\mathbf{x}_{i+1} + \mathbf{x}_i)/2$ is taken as a new iteration value $\mathbf{x}'_{i+1}$. If $|J(\mathbf{x}'_{i+1})| > |J(\mathbf{x}_i)|$, and $(\mathbf{x}'_{i+1} + \mathbf{x}_i)/2$ is taken as the new iteration value $\mathbf{x}''_{i+1}$. Until the assimilation norm of the new iteration value was less than $|J(\mathbf{x}_i)|$, the value satisfying the conditions was assigned to $\mathbf{x}_{i+1}$ to satisfy the local conditions of Newton's lemma in the next assimilation. However, multiple calculations were required for this method, which aggravated the burden of the calculation and affected the convergence speed so that the weighted average method could be used to improve the speed of the convergence. Table 1 illustrates the algorithm: a modified iterative ensemble Kalman filter.

Table 1: Algorithm: a modified iterative ensemble Kalman filter

Require	The forecast ensemble $\mathbf{x}_k^f$, k $=1,2,3.....N$, the observations $\mathbf{y}$ and error covariance matrix $\mathbf{R}$
1:	Compute $\mathbf{Y} = [\mathbf{y}^T \ \ (\mathbf{x}^f)^T]^T$ ~($u(\mathbf{x}) = \left[(h(M(\mathbf{x}^f)))^T \ \ \mathbf{x}^T\right]^T$ $\boldsymbol{\sigma} = diag([\mathbf{P},\mathbf{R}])$)
2:	find the minimum $J(\mathbf{x}) = \frac{1}{2}((\mathbf{y} - h(M(\mathbf{x}^f))^T \mathbf{R}^{-1}$ $(\mathbf{y} - h(M(\mathbf{x}^f))) + (\mathbf{x} - \mathbf{x}^f)^T \mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}^f))$
3:	compute the Newton step $\delta_i = -J(x_i)/J'(x_i)$
4:	compute $J(x_{i+1}) = J(x_i + \lambda_i\delta_i) = \min_{i \ge 0} J(x_i + \lambda\delta_i)$, to choose a proper step factor λ within $0 < \lambda < 1$
5:	compute $J(x_{i+1}) < J(x_i)$ and $J(x_{i+1}) \le J(x_i) + \alpha\nabla J^T \bullet \delta_i$, $0 < \alpha < 1$
6:	if $\lvert J(x_{i+1})\rvert > \lvert J(x_i)\rvert$, $(x_{i+1} + x_i)/2 \to x'_{i+1}$ and if $\lvert J(x'_{i+1})\rvert > \lvert J(x_i)\rvert$, $(x'_{i+1} + x_i)/2 \to x''_{i+1}$, ..., until $\lvert J(x^n_{i+1})\rvert < \lvert J(x_i)\rvert$, $x^n_{i+1} \to x_{i+1}$

Under certain local conditions, a Gauss-Newton iterative algorithm descends along the Newton direction. In regards to an IEnKF, if the norm of the new iteration assimilation result is greater than that of the previous assimilation result, it has been determined that both assimilation results become distributed on both sides of the optimal value absolutely. For example, when $\mathbf{x}_i < \mathbf{x}_{true} < \mathbf{x}_{i+1}$ and $|J(\mathbf{x}_{i+1})| > |J(\mathbf{x}_i)|$, the weighted average algorithm can be used. The new iteration value is made to return along the old iteration value direction, where the optimal solution is closer to the true value and can be obtained as the new iteration value. The comparison with IEnKF algorithm is as shown in Figure 1.

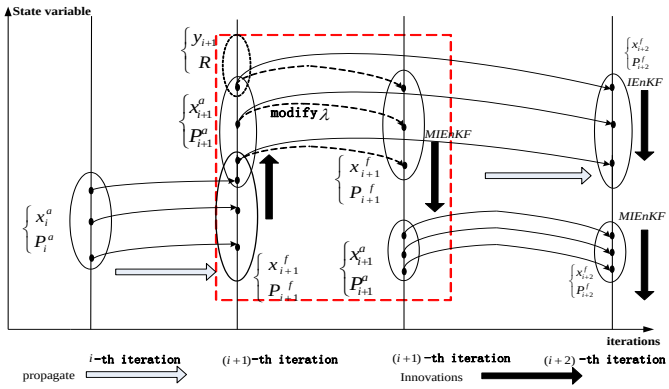

Figure 1: Schematic diagram of IEnKF and MIEnKF algorithms.

In Figure 1, the x-axis shows the number of iterations, while the y-axis indicates the state changes. The red dotted box indicates the specific process of using the MIEnKF for reducing the errors at the i+1-th iteration, when the i+1-th iteration value greatly deviates from the i-th iteration value. According to the gradient descent direction of the cost function, the step λ is adaptively adjusted to make assimilation results eventually converge, and so on. So, the MIEnKF can better adapt to a strongly nonlinear system.

4 Numerical experiments

With different parameterizations and physical representations determining the sensitivity of the procedure of different models, observing system simulation experiments (OSSE) are designed to enable examination of the performance of data assimilation procedures [12]. In this section, the MIEnKF is compared with a standard IEnKF method, as well as a traditional EnKF method for research purposes.

4.1 Low-dimensional chaotic system Lorenz-63

A Lorenz-63 model is a three-variable chaotic dynamical system, of which the control equation of a state evolution can be given by the following ordinary differential equation:

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = rx - y - xz \\ \dot{z} = xy - bz \end{cases} \tag{16}$$

In this study, the parameter values were set as follows: $\sigma = 10$, $b = 8/3$ and $r = 28$. In the numerical discrete method, a fourth-order Runge-Kutta method was used with the integral step of $dt = 0.01$, the total integration time S=1000, and the initial state of $X = [1.508870; 21.531271; 25.46091]$, which was a point in the state of the chaotic system.

In this study, to verify the performance of the MIEnKF filter algorithm, a sensitivity experiment was used with different ensemble sizes, covariance factors, and observation frequency and analysis

cycle length. The following evaluation indices, named the root mean square error (rmse), were adopted:

$$rmse(\mathbf{x},\mathbf{x}^{true})=\left[\frac{1}{n}\left(\mathbf{x}-\mathbf{x}^{true}\right)^{T}\left(\mathbf{x}-\mathbf{x}^{true}\right)\right]^{1/2} \tag{17}$$

where $\mathbf{x}$ and $\mathbf{x}^{true}$ are the estimate and true value of the model state, respectively, and n is the dimension of the model state vector.

4.2 Experimental results and analyses

In this section, the assimilation performances of the IEnKF and MIEnKF are examined. For the same setup, the modified effects were found to be different.

4.2.1 Influences of ensemble size. In the experiment, we changed ensemble sizes from N=5 to N=80; the model step T = 25 was selected, the inflation factor was set as infl = 1.05, and the observation error obeyed the Gaussian distribution with a mean of 0 and a variance of $\sigma_{obs}^{2}=1$. Figure 2 shows the experimental results of the three methods. Due to the different domain of the values, we present the forecast-rmse and analysis-rmse in separate diagrams.

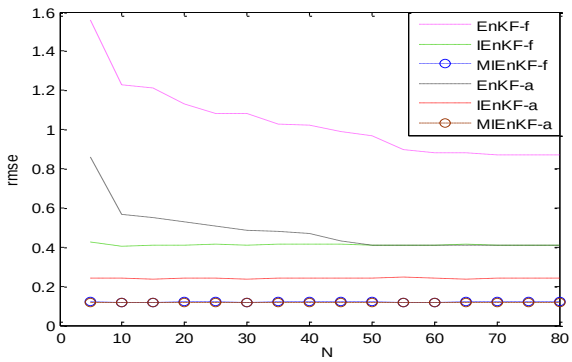

Figure 2: Influence of ensemble size on comparison of the three algorithms with forecast-rmse and analysis-rmse

The research showed that: (1) with an increase of ensemble size, the performance of the three algorithms improved in agreement with typical results from the literature [1]. While the EnKF seems to improve with increased ensemble size, the IEnKF and MIEnKF show very little improvement. The larger the ensemble size, the better the assimilation effect. However, the assimilation time was correspondingly longer. (2) Among the three types of algorithms, the performance of the EnKF was found to be the least effective. In cases with a small ensemble size, the rmse was larger. The IEnKF was determined to have a slightly better performance than the EnKF. However, the modified MIEnKF was able to meet the demand of the actual assimilation more effectively, and the rmse was minimal.

4.2.2 Influences of the inflation factor. An inflation factor is a type of error processing method for improving assimilation effects. If the inflation factor is changed, the performance of the different algorithms will change greatly.

In this study, the results showed the following:

(1) When the inflation factor was increased, three types of algorithms had diverse results, wherein the performance of the EnKF improved the most, the improvement of the IEnKF was not obvious, and no influence was made on the MIEnKF.

(2) When the step of the inflation factor was equally reduced, it was found that the EnKF had the worst performance, and the rmse of the IEnKF decreased from 0.237 to 0.236, while no changes were made on the MIEnKF. Therefore, when combined with the performance curve diagram of the three algorithms, as shown in Figure 3, it can be confirmed that no matter what type of situation, the performance of the EnKF was the worst. For example, the rmse value of the EnKF presented multiple

peak-trough phenomena as the equal interval of the inflation factor increased or decreased. Therefore, it was particularly important to select the inflation factor. To focus on this problem, a genetic algorithm was used to select the optimal inflation factor from previous studies [1] and to obtain the optimal data assimilation results under limited computing resources.

(3) When compared with the IEnKF, this algorithm was able to adapt well to the data assimilation of a strongly nonlinear system. In addition, when combining the algorithm performance variations caused by ensemble size changes, it could be seen that the rmse of the MIEnKF was smaller, and the same was found for the inflation factor. Therefore, the MIEnKF had a more robust performance than the IEnKF.

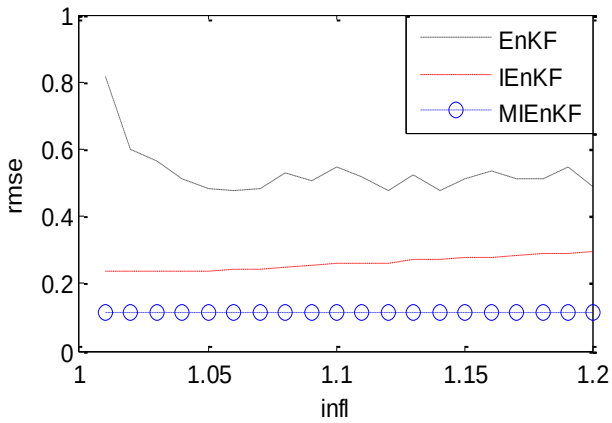

Figure 3: Influence of inflation factor on data assimilation.

4.3 High-dimensional chaotic system Lorenz-96

In this section, to further detect the performance of the modified iterative ensemble Kalman filter, a Lorenz-96 model that had 40 state variables [11] was chosen. A Lorenz-96 system is a more complex nonlinear dynamic system, and it is used to simulate the time evolution of atmospheric variables. The equation is as follows:

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F \tag{18}$$

The discrete numerical experiments adopted a fourth-order Runge-Kutta method, where $j = 1, 2, \ldots, M$ represent the loop label, forcing parameters of $F = 8$, and variables of $M = 40$. The time step dt was taken as a 0.05 dimensionless unit fixed value.

The performances were compared among the three methods (EnKF, IEnKF, and MIEnKF) by changing the observation error and the inflation factors. The ensemble size was set as N = 30, the observation variance as obs_var = 1, and the model step as T = 25. The inflation factor infl changed from 1.01 to 1.20 and was set at intervals of 0.01. Its performance in data assimilation was then observed. Figure 4 illustrates the observation error influence on the performance of the three algorithms' assimilation.

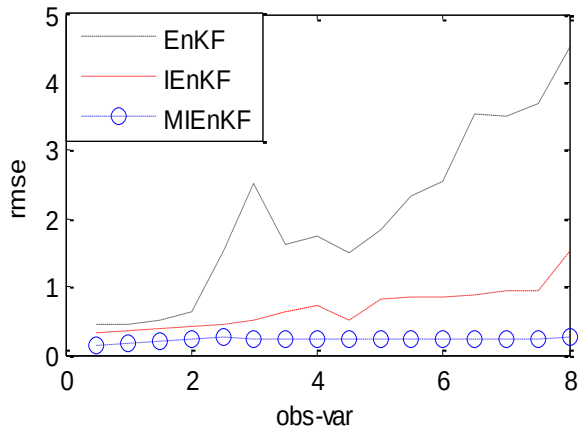

Figure 4: Influence of observation error on data assimilation.

In the experiment, the ensemble size was set as N = 30, the model step as T = 25, and inflation factor as infl = 1.05. The observation error, obs_var, ranged between 0.5 and 8, and the step size was 0.5. The results confirmed that the modified algorithm could still finish with an accurate estimation, even when the interval between the observation errors was larger by contrast. In the early stages, the EnKF could basically meet the true state. However, great deviations appeared in later stages. The floating range of the IEnKF was relatively small, and therefore the assimilation performance was relatively poor because the analogue system error was larger. When combined with the previous analysis, the new method was found to have a greatly improved performance compared to the first two algorithms.

In addition, in terms of computational costs, the average time for the EnKF was 18.25 s, while the average time of the IEnKF and MIEnKF were approximately 35.25 s and 44.5 s, respectively. We found that the MIEnKF needed a larger amount of calculation, which took longer; therefore, reducing the computation time while maintaining the premise of satisfying the estimation precision is still a subject for further study.

5 Conclusions and further discussions

In terms of a linear Gaussian system, the EnKF was the optimal algorithm for data assimilation. However, as the operational-scale models increase in resolution and complexity, nonlinearities pose greater problems to today's most widely used data assimilation methods. These nonlinearities are particularly critical for determining the optimal assimilation method for a strongly nonlinear system. A MIEnKF was proposed in this study for the purpose of exploring narrating an IEnKF from the perspective of a Newton iteration and to determine the optimal iteration value by judging global convergence assimilation results. This iterative version of an EnKF is used to minimize the local (i.e., limited to on assimilation cycle) cost function and investigate the performance of such a scheme in a number of experiments withtwo common nonlinear chaotic models. The only difference between the MIEnKF method and the IEnKF method is the method of minimization: we used a Gauss-Newton derivation of the IEnKF. This is the reason the results of some experiments with the MIEnKF method are almost a straight line. Due to the OSSE setup for such synthetic experiments, "truth" is defined as when the model is integrated (run) for a set of initial conditions. "Observations" are generated from the "truth" by adding some type of noise. Assuming the error from the observations and forecast ensemble are ideal because of the applied minimization method, we can guarantee the assimilation performance is better than other methods. Meanwhile, when compared with the EnKF, the computing burden of the MIEnKF is increased with greater iteration numbers. Therefore, to satisfy the premise of estimation precision, a solution for reducing the calculation burden of the MIEnKF is an important direction for future study.

Acknowledgements
This work is supported by the NSFC (National Natural Science Foundation of China) project (grant number: 41461078 and 41061038).

References

[1] Bai, Y. L, and Li. X. "Evolutionary Algorithm Based Error Parametrization Methods for Data Assimilation", *Mon. Wea. Rev.*, **139**, pp. 2668-2685, (2011).
[2] Bell F, Cathey B. "The iterated Kalman filter update as a Gauss-Newon method", IEEE Trans On Automatic Control, **38**, pp. 294-297, (1993).
[3] Bocquet, M. and Sakov, P. "Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems", *Nonlin. Processes Geophys.*, **19**, pp. 383–399.
[4] Chen, Y and Oliver, D. S. "Ensemble randomized maximum likelihood method as an iterative smoother", *Math. Geosci*, **44**, pp. 1-26, (2012).
[5] Evensen G. "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal Geophys Res*, **99**, pp. 10143-10162, (1994).
[6] Gu, Y and Oliver, D. S. "An iterative ensemble Kalman filter for multiphase fluid flow data assimilation", *SPE Journal*, **12**, pp. 438-446, (2007).
[7] Jazwinski, A.H. "Stochastic Process and Filtering Theory", Academic Press: Waltham, MA, USA, pp.332–365, (1970).
[8] Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy."4-D-Var or ensemble Kalman filter", *Tellus*, **59A**, pp. 758-773, (2007).
[9] Kalnay, E., and S.-C. Yang. "Accelerating the spin-up of Ensemble Kalman Filtering", *Quart. J. Roy. Meteor. Soc.*, **136**, pp. 1644–1651, (2010).
[10] Li. X, and Bai, Y. L. "A Bayesian Filter framework for sequential data assimilation", *Advances in Earth Science*, **25**, pp. 515-522, (2010).
[11] Lorenz, E. N., and Emanuel, K. A. "Optimal Sites for Supplementary Weather Observations: Simulati-on with a Small Model", *Journal of American Meteorological Society*, **55**, pp. 399-414, (1998).
[12] Penny, S G, Kalnay, E, Carton, J A, Hunt, B R, Ide, K; Miyoshi, T, Chepurin, G A. "The local ensemble transform Kalman filter and the running-in-place algorithm applied to a global ocean general circulation model", *Nonlinear Processes in Geophysics*, **20**, pp. 1031, (2013).
[13] Sakov, Pavel, Dean S. Oliver, Laurent Bertino. "An Iterative EnKF for Strongly Nonlinear Systems," *Mon. Wea. Rev.*, **140**, pp. 1988-2004, (2012).
[14] Yuan Yaxiang, Sun Wenhan. "Optimization theory and method," Beijing: Science Press, pp. 108-330, (1997).