

Research Article

Modifying Covariance Localization to Mitigate Sampling Errors from the Ensemble Data Assimilation

Mingheng Chang , Hongchao Zuo , and Jikai Duan

College of Atmospheric Science, Lanzhou University, Lanzhou, China

Correspondence should be addressed to Hongchao Zuo; zuohch@lzu.edu.cn

Received 18 July 2022; Revised 5 September 2022; Accepted 12 October 2022; Published 26 October 2022

Academic Editor: Antonio Donateo

Copyright © 2022 Mingheng Chang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The ensemble-based Kalman filter requires at least a considerable ensemble (e.g., 10,000 members) to identify relevant error covariance at great distances for multidimensional geophysical systems. However, increasing numerous ensemble sizes will enlarge sampling errors. This study proposes a modified Cholesky decomposition based on the covariance localization (CL) scheme, namely a covariance localization scheme with modified Cholesky decomposition (CL-MC). Our main idea utilizes a modified Cholesky (MC) decomposition technique for estimating the background error covariance matrix; meanwhile, we employ the tunable singular value decomposition method on the background error covariance to improve the ensemble increment and avoid the imbalance of the system. To verify if the proposed method can effectively mitigate the sampling errors, numerical experiments are conducted on the Lorenz-96 model and large-scale model (SPEEDY model). The results show that the CL-MC method outperforms the CL method for different data assimilation parameters (ensemble sizes and localizations). Furthermore, by performing one year of assimilation experiments on the SPEEDY model, it is found that the 1-day forecast RMSEs obtained by the CL are approximately equal to the 5-day forecast RMSEs of CL-MC. So, the CL-MC method has potential advantages for long-term forecasting. Maybe the proposed CL-MC method achieves good prospects for widespread application in atmospheric general circulation models.

1. Introduction

Data assimilation (DA) is the combination of observations with “prior knowledge” (e.g., numerical models and model outputs) to achieve an estimate of the true state of a system (cf. Katzfuss et al. [1]). In the process of sequential data assimilation, the traditional Kalman filter (KF; Kalman and Bucy [2]) is used for the estimation and prediction of linear problems. For complex multidimensional systems, the classical ensemble DA filter (EnKF; Evensen [3]) derived from the merging of the linear estimation theory of Kalman filter and the nonlinear Monte Carlo estimation theory has been proposed, where is widely applied in the large-scale models (cf. Evensen and Van Leeuwen [4]). However, the ensemble members are always much smaller than the model components so sampling errors cannot be avoided. The sampling errors can introduce numerous other problems (e.g., inbreeding and filter divergence)

into the process of the ensemble filter. Inbreeding is a term used to describe the underestimate of the analysis error covariance (cf. Furrer and Bengtsson [5]), and filter divergence occurs when the analysis error distribution moves away from the truth and this can be caused by the underestimating of the background error covariance. Such problems can be addressed with the covariance inflation method, which operates on the background error covariance multiplied by an inflation factor and prevents the underestimation of the covariance (cf. Anderson and Anderson [6]). However, the size of the inflation factor will depend on many factors including the type of filter and the used dynamical system (cf. Hamill et al. [7]). Thus, covariance inflation does not address all problems, especially since the spurious long-range correlations between great distance observations and model components from the EnKF. It is common to use localization schemes to solve such spurious correlations.

Localization of the covariance matrix in the EnKF is necessary for large-scale models because the traditional ensembles are limited to only 10–100 states, which is considerably less than the degree of freedom of the model. Localization methods are typically implemented by taking the Schur product of the correlation matrix or the localization function, such that the localization methods contain covariance localization (CL; Hamill et al. [7]) and observation localization (OL; Hunt et al. [8]). Alternatively, the domain is decomposed as in domain localization (DL; Houtekamer and Mitchell [9]) and separate analysis is conducted for each domain. The CL method is commonly employed in ad hoc treatment of the spurious correlation problem. The OL and DL methods are frequently combined to use and achieve better observation weights according to their distance. Additionally, other localization techniques, such as local analysis (LA; Sakov and Bertino [10]), background error CL (BL; Greybush et al. [11]), and observation error CL (RL; Greybush et al. [11]) have been devised. Although such localization can be effective in avoiding spurious correlations and limiting the negative effects of sampling errors, it simultaneously introduces imbalances into the system (cf. Kepert [12]). Especially covariance localization is a major driver of potential imbalance in EnKF. Efforts are being made to find approaches to avoid the imbalance system. The imbalance is primarily produced by the computation of analysis increments. The incremental analysis update (IAU) procedure is used to prevent imbalance generation during the computation of the local increments (cf. Kleist and Ide [13]). In addition, it is possible to avoid imbalances by exploiting the shorter assimilation window (Houtekamer and Zhang [14]).

In the Localization scheme framework, the localization parameters (i.e., the localization length and the weighting function), a key indicator of the localization scheme, are often manually configured. These parameters are unknown a priori and can usually be found by trial and error. Kirchgessner et al. [15] investigated that the optimal localization radius of DL and OL are linearly related to the local observation dimension, and the OL effect was demonstrated to be better than the DL effect. With higher-resolution models, the use of a narrower localization radius reduces the sampling error, but it limits the use of observations. Miyoshi and Kondo [16] proposed a multiscale localization (500 km and 900 km) approach for EnKF to find analysis increments in independent subobservation spaces. However, it requires a large computational cost when the number of observations available is $O(10^7)$. In addition, to preserve the larger-scale structures from a limited ensemble size, Kondo et al. [17] employed a dual-localization approach to analyze small-scale and large-scale analysis increments separately using spatial smoothing. For complex geophysical models, although the impact of the observations will be decreased with longer localization ranges, some issues of localization require to be briefly discussed. First, distance-based localization methods employ unimodal segmentation functions (e.g., GC functions) to construct the corresponding matrices, which have difficulties in “localizing” nonlocal observations (Bai et al. [18]). These observations exhibit multimodal

correlations with model components (cf. Fertig et al. [19]; De La Chevrotière and Harlim [20]). Furthermore, in some cases, the model components or the corresponding observations may not have correlated physical locations, which may be against the principles of distance-based methods (cf. Bai et al. [21]). The fuzzy control functions are used for yielding more precise observation weights instead of localization functions (e.g., GC functions) (cf. Mingheng et al. [22]; Mingheng et al. [23]). For an ideal intermediate atmospheric general circulation model (AGCM), the negative effects of localization could be avoided with a sufficiently large ensemble size. Miyoshi et al. [24] investigated that using large ensemble sizes (up to 10,240 members) could capture long-range scale error correlations in the implementation of the local ensemble transform Kalman filter (LETKF; Miyoshi et al. [25]), however, it is essential for preventing memory overflow to employ relatively large ensemble members. Kondo and Miyoshi [26] modified the EnKF code without using the covariance localization prerequisite. It was found that using 10,240 ensemble members also reveals the long-term structural features of the background error covariance matrix. The aforementioned studies commonly considered the effect of horizontal scales on assimilation effects. Farchi and Bocquet [27] considered a realistic extension of the LEnSRF using covariance localization in the vertical direction. Wang et al. [28] introduced a multiscale local gain from ensemble transform Kalman filter (MLGETKF) to allow simultaneous update of multiple scales for the forecast ensemble mean and perturbations via assimilating all observations at once. Recently, The Cholesky decomposition method has been applied to the estimate of the forecast error covariance matrix (cf. Kang and Deng [29]).

In the abovementioned studies, the use of a multiscale localization scheme or large ensemble members (e.g., 10,000 members) has so far been necessary to mitigate the sampling errors. However, there are few investigations on the estimate of the forecast error covariance matrix. Here, a covariance localization scheme with modified Cholesky decomposition (CL-MC) for estimating the background error covariance is proposed. The effectiveness of the proposed CL-MC method is examined by the Lorenz-96 and the SPEEDY models. Thus, the article is organized as follows: Section 2 focuses on the background of the ensemble Kalman filter, covariance localization, modified Cholesky decomposition, and the proposed covariance localization with the modified Cholesky decomposition. Section 3 describes the configuration of the twin models and analyzes the experimental results. Section 4 provides the discussion and conclusion of the results.

2. Background

2.1. Ensemble Kalman Filter (EnKF). The EnKF is an ensemble-based Kalman filtering scheme, it was proposed by Evensen [3]; subsequently, it evolved many forms of EnKF methods (cf. Tippett et al. [30]; Houtekamer and Mitchell [9]; Evensen and Van Leeuwen [4]; Evensen [31]; Burgers et al. [32]. Anderson [33]). These methods can be roughly

classified into two categories: deterministic filters (cf. Burgers et al. [32]) and stochastic filters (cf. Whitaker and Hamill [34]). The main difference between them is whether or not perturbed observations are introduced to maintain the correct statistics of the filter.

The implementation of EnKF is to update the a priori estimate $x^f(k)$ of the atmosphere valid at atmosphere valid time k with the information in the new observation y^o to reach the latest estimate $x^a(k)$ of the atmosphere (equation (1)). For this purpose, the Kalman gain matrix K is used to dispatch appropriate weights to the observation error covariance R and the background error covariance P^f (equation (2)). In all equations, the superscripts a and f refer to the analysis and predictor variables, respectively. The observation operator H denotes the mapping from the model space to the observation space. The model operator M denotes the transfer of the estimate $x^a(k)$ to the next analysis time (equation (3)).

$$x_i^a(k) = x_i^f(k) + K(y_i^o - Hx_i^f(k)), i = 1, \dots, n, \quad (1)$$

$$K = P^f H^T (H P^f H^T + R)^{-1}, \quad (2)$$

$$x_i^f(k+1) = M[x_i^a(k)], i = 1, \dots, n, \quad (3)$$

where n denotes the ensemble size and k represents the atmosphere valid time. For the Kalman gain K in equation (2), the ensemble size is usually used to approximate the estimates $P^f H^T$ and $H P^f H^T$ (cf. Houtekamer and Mitchell [35]; Ghil and Malanotte-Rizzoli [36]). In the subsequent equations, we omit the superscript of the time index k for ease of notation.

For the stochastic EnKF, small spurious correlations between the ensemble background and observations lead the analysis to move away from the truth. Whitaker and Hamill [34] introduced the deterministic ensemble square root filter (EnSRF). However, using the EnSRF method to compute the analysis error covariance with a relatively small ensemble size, the Kalman gain is susceptible to contamination by sampling error, making it likely that the true analysis error is much larger than the background error. The ensemble may become too concentrated to the detriment of the assimilation results one would expect. It is common to use covariance inflation or covariance localization to compensate for this particular problem. Depending on the covariance inflation, it does not address all problems associated with undersampling. The covariance localization (CL) is essential for data assimilation.

There are two reasons for using covariance localization in EnKF, these are usually listed in the literature as ‘‘Spurious covariance’’ and insufficient ensemble rank (cf. Sakov and Bertino [10]). The ‘‘spurious covariance’’ means that the limitation of the ensemble size makes no connection between the covariances of the elements of the remote state vectors. It can lead to excessive reduction of the state error covariance in the analysis, resulting in ensemble collapse and filter divergence. The insufficient ensemble rank denotes that the ensemble size is less than the

model components. Therefore, the proper use of the ensemble size is one of the most basic rules to ensure the design of EnKF systems.

2.2. Covariance Localization (CL). For the CL method, it is used to modify and update equation (2) by multiplying the distance correlation coefficient matrix ρ with each element of the background error covariance (Schur). It is one of the localization methods that are applied to large-scale systems. The core idea behind this operation is to modify each element of the background error covariance to increase the rank of the covariance matrix.

$$P^f \longrightarrow \rho^o P^f. \quad (4)$$

However, the localized background error covariance matrix has hardly been explicitly computed except for small-scale systems. Instead, if we have enough observations, we can approximately compute the terms $(\rho^o P^f) H^T$ and $H(\rho^o P^f) H^T$ in equation (2) by the ensemble size. Algorithm 1 summarizes the analysis step of the resulting generic CL method in detail.

Where $X_{i_o}^f$ refers to the ensemble anomalies of each member’s position i_o around o observations, $\begin{Bmatrix} 0 \\ \end{Bmatrix}$ represents the observation operator with all rows set to zero except for the o th, \tilde{H} implies to the observation position matrix, such that $\tilde{H}_{oi} = 1, i = i_o; 0, i \neq i_o$, and p is the observation member. Although the introduction of covariance inflation and localization techniques can overcome the problems in the EnKF. There will be a restriction to apply these techniques only to the background error covariance matrix ($B \approx P^f = (n-1)^{-1} \cdot [X^f \cdot (X^f)^T]$). There still exists finite ensemble sizes. Therefore, the local sample covariance matrix may be rank-deficient. To avoid this problem, the background error covariance matrix B is estimated directly by an improved Cholesky decomposition. A sparse approximation of \tilde{B} is replaced by B . The aim is to avoid limiting the ensemble size as well as to prevent memory overflow.

2.3. Modifying Cholesky Decomposition. The accuracy of background error covariance estimation has been the core problem for the covariance localization schemes when using limited ensemble sizes. Recently, some researchers employ Cholesky decomposition to achieve the accuracy of background error covariance estimates (cf. Ueno and Tsuchiya [37]; Nino-Ruiz et al. [38]; Kang and Deng [29]; Boyles and Katzfuss [39]). In this regard, we achieve the estimation of the inverse of the sparse background error covariance matrix by using a modified Cholesky decomposition. The Cholesky decomposition is also known as the square root decomposition method. Its core principle is to represent a symmetric positive definite matrix as a decomposition of the product of a lower triangular matrix and its transpose. Consider (approximately) N samples of Gaussian random vectors $X = [X_1, \dots, X_j, \dots, X_n] \in \mathcal{Y}^{m \times n}$ with statistical moments $X_j \sim N(0_n, B)$ for $1 \leq j \leq n$. $x_i \in \mathcal{Y}^{m \times 1}$ indicates the i th model component for all samples. The modified Cholesky decomposition is derived from the previous component regression (Nino-Ruiz [40]).

- (i) **Require:** A statistical sample of N state estimates $x^i, i = 1, \dots, n$. E is the ensemble matrix, that is, $m \times n$, where column are the ensemble number. m is the model component. $X^f = [x_1^f - \bar{x}^f, x_2^f - \bar{x}^f, \dots, x_n^f - \bar{x}^f]$. Observation vector $y \in R^{N_y}$ (input). λ is a multiplicative inflation factor (parameter). 1 represents the vector with all elements being 1. \mathbf{I} is the identity matrix. ρ is the localizing function ($0 \leq \rho_{ij} \leq 1$).
- (1) **for** $step = 1, \dots, step_0$ **do**
 - (2) $x = n^{-1} \cdot (E \cdot 1)$ and $P^f = (n-1)^{-1} \cdot [X^f \cdot (X^f)^T]$
 - (3) $X = E \cdot [I - n^{-1} \cdot (11^T)]$
 - (4) $K = P^f \cdot H^T \cdot (H \cdot P^f \cdot H^T + R)^{-1}$
 - (5) $(\rho \circ P^f) \cdot H^T \equiv (n-1)^{-1} \cdot \sum_{o=1}^p X_o^f \cdot [(\overset{(o)}{H} \cdot X^f)]^T$
 - (6) $H \cdot (\rho \circ P^f) \cdot H^T \equiv (n-1)^{-1} \cdot (\bar{H} \cdot \rho \cdot \bar{H}^T) \circ (H \cdot X^f)(H \cdot X^f)^T$
 - (7) $x^a = x^f + K \cdot (y - H \cdot x^f)$
 - (8) $\bar{x}^a = n^{-1} \cdot (\sum_{i=1}^n x_i^a)$
 - (9) $X^a = x^a - \bar{x}^a$
 - (10) $E \leftarrow x^a + \lambda \cdot \sqrt{n-1} \cdot X^a$
 - (11) **return** $m \times n$ analysis ensemble E
 - (12) **end for**

ALGORITHM 1: Analysis step for CL.

Then, the estimate of B can be computed as follows:

$$(\rho \circ \hat{B}) = \hat{U} \cdot \hat{\Sigma} \cdot \hat{U}^T \in \Upsilon^{m \times m}, \quad (5)$$

$$(\rho \circ \hat{B})^{-1} = \hat{U}^{-T} \cdot \hat{\Sigma}^{-1} \cdot \hat{U}^{-1} \in \Upsilon^{m \times m}, \quad (6)$$

for equations (5) and (6), the eigenstructure of $(\rho \circ \hat{B})$ depends on the structure of \hat{U} . Based on the structural features of \hat{U} , the sparse estimates of $(\rho \circ \hat{B})$ are obtained by imposing the elements of \hat{U} to be zero. The main zero components in \hat{U} are the following: if the two model components are conditionally independent for a given localization radius r , the corresponding elements in $(\rho \circ \hat{B})$ are zero. Furthermore, based on the CL approach, the conditional independence of background errors between different model components can be achieved by using local neighborhoods (cf. Ueno and Tsuchiya [37]). For a given model component, its neighborhood is formed by some components within a localization radius r . If a model component exceeds the r distance, the correlation between the corresponding background errors is set to zero.

2.4. Covariance Localization with the Modified Cholesky Decomposition. This section investigates the estimation of the covariance matrix with modified Cholesky decomposition of the covariance localization (CL-MC). First, according to equation (5), the structure of \hat{B} depends on the structure of \hat{U} , considering the estimation of the background error covariance. If we assume that the correlation between the model components is local and there is no correlation outside the localization radius r , we obtain a lower triangular sparse estimate of \hat{U} , and \hat{U} is sparse and localized. According to literature (cf. Nino-Ruiz et al. [38]; Boyles and Katzfuss [39]), the regression equation can only be performed on the former for each model component, so the model components must be sorted (labeled) before

computing \hat{U} . For the grid model, we mainly considered the column order.

The estimation procedure for $(\rho \circ \hat{B})^{-1}$ is as follows:

1. For the i th model component, we use the truncated singular value decomposition method to estimate the regression coefficients $\eta_{i,j}$ for $1 \leq i \leq m$ and $j \in P(i, r)$, the model component x_i as follows:

$$x_i = \sum_{j \in P(i, r)} \eta_{i,j} \cdot x^j + \varepsilon_i \in \Upsilon^{m \times 1}, \quad (7)$$

where $P(i, r)$ represents the predecessor indexes of the model component i when the localization radius is r . x^j denotes the j th model component which precedes x_i under the condition of $1 \leq j \leq i \leq m$. Figure 1(a) shows the local predecessors and neighborhood of model element 28 when column order is employed. ε_i represents the error of the i th model component regression. η denotes the coefficients can be calculated by the solving the optimization problem, that is $\eta_i = \operatorname{argmin} \|x_i - Z_i \cdot \eta\|_2^2$, where $Z_i = [x_1, x_2, \dots, x_i]^T \in \Upsilon^{i \times n}$ ($1 \leq i \leq n$), $\eta_i = [\eta_{i,1}, \eta_{i,2}, \dots, \eta_{i,i}]^T \in \Upsilon^{i \times 1}$. The simplest solution is $\eta_i = (Z_i \cdot (Z_i)^T)^{-1} \cdot Z_i \cdot x_i$.

2. Create the sparse matrix

$$\{\hat{U}\}_{i,j} = \begin{cases} -\eta_{i,j} & j \in P(i, r), \\ 1 & j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

The residuals in equation (8) will never be zero, it implicitly estimates the covariance inflation. The application of covariance inflation may not be required for the CL-MC assimilation scheme.

Once $(\rho \circ \hat{B})$ is estimated, in the next section we present how to avoid the imbalance effect and handle it with $(\rho \circ \hat{B})$. However, here, the analysis formulation is as follows:

$$x^a = x^f + [(\rho \circ \hat{B})^{-1} + H^T \cdot R^{-1} \cdot H]^{-1} \cdot H^T \cdot R^{-1} \cdot \delta y_o \in \Upsilon^{m \times n}, \quad (9)$$

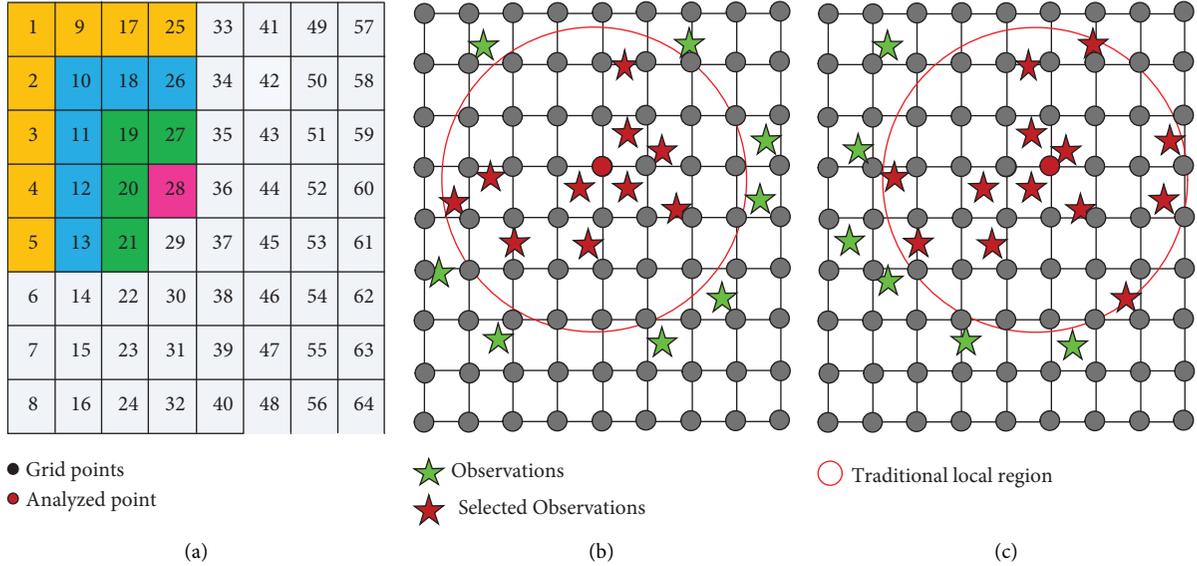


FIGURE 1: (a) shows the predecessors of the model component 28, the green region, blue region, and yellow region within the different localization radii. (b) and (c) show schematic diagrams of the observations used in the localization method from the i -th grid point to the $(i+1)$ th grid point. For each grid point (gray dots, the analyzed point is red). The localization region is drawn (red circle) and only those observation points (red pentacles) are included in the analysis.

here, the matrix of perturbed observations $y_s = [y^o + \varepsilon^1, y^o + \varepsilon^2, \dots, y^o + \varepsilon^n] \in Y^{m \times n}$, $\varepsilon^i \sim N(0, R)$ for $1 \leq i \leq n$. For equation (9), we obtain it by using the basic decomposition matrix identities.

$$x^a = x^f + \hat{U}^{-1} \cdot \hat{\Sigma}^{1/2} \cdot W_B^T \cdot \left[R + W_B \cdot W_B^T \right]^{-1} \cdot \delta y_o, \quad (10)$$

where $W_B = H \cdot \hat{U}^{-1} \hat{\Sigma}^{1/2} \in Y^{m \times p}$ and p is the observation members. The $\delta y_o = [y_s - Y^f]$ term in equation (10) represents the perturbation observation. Equation (10) is well known as the EnKF dual formulation and equation (9) is the incremental form of the primal formulation. Using equation (10), we conclude that the computational effort of the method is bounded by the following equation:

$$O(p \cdot m + p \cdot m \cdot n). \quad (11)$$

2.5. Balance and Truncated Singular Value Decomposition (svd). In the above section, we introduced the CL method with the modified Cholesky decomposition. However, covariance localization is a major source of significant potential imbalances in the EnKF, especially for short localization distances. Thus, how to effectively reduce the source of imbalance has become one of the key issues for data assimilation. There are many main ways to the reduction of sources of imbalance (such as modifying the initial spin-up procedure and incremental analysis update (IAU)) (cf. Bloom et al. [41]; Kleist and Ide [13]). In practical operations, modifying the initial spin-up procedure or applying IAU has so far been essential to mitigate the imbalance in the system. In our localization scheme, we use the sparse background error covariance ($\rho^o \hat{B}$) as an alternative to the \hat{B} , the distribution of the model components, and the analyzed point in the spatial

dimensions as shown in Figures 1(b) and 1(c). Schematic diagrams of the observations used in the localization scheme from the i -th grid point to the $(i+1)$ -th grid point. For efficiency, we compute the distances for only observations with the coordinates $limits_x$ and $limits_y$. Valid local observations reside within the circle of the local range radius l , which is determined according to this distance. The physical distance between the analyzed point and the observation position is calculated before the state estimate is updated. The closer the distance to the observation position is, the greater the weight of the observation point, and vice versa.

Although we use a modified Cholesky decomposition to estimate the $(\rho^o \hat{B})$, $(\rho^o \hat{B})$ and \hat{B} exhibit a strong correlation. The correlation is retained within a defined distance but has been eliminated beyond the defined distance. However, with high-dimensional systems, the ensemble covariance becomes full rank, and we have to deal with a great covariance matrix (by exploiting symmetry, we can represent it by using $(m \cdot (m-1))/2$ elements). One does not want to represent such a matrix in memory, and it is preferable to formulate the filter in some matrix-free manner. Most of the matrix-free formulations rely on one of two things: domain decomposition and Schur product. Here, we employ the truncated svd and represent such a matrix in memory.

Suppose that we have a truncated svd of \hat{B} given by the following equation:

$$\hat{B} = \rho^o \hat{B} = \rho^o \left[\hat{X}^{f'} \cdot \left(\hat{X}^{f'} \right)^T \right] = \hat{Q} \cdot \hat{D} \cdot \hat{Q}^T \in Y^{m \times m}, \quad (12)$$

where $\hat{Q} \in Y^{m \times n}$ orthogonal matrix, $\hat{D} \in Y^{m \times n}$ diagonal matrix, $\hat{X}^{f'} = (n-1)^{-1/2} \cdot X^f$, and $X^f = [x_1^f - \bar{x}^f, x_2^f - \bar{x}^f, \dots, x_n^f - \bar{x}^f]$. Because \hat{B} is symmetric positive

definite, equation (12) is a truncated singular value decomposition. In truncated svd scheme, if the ensemble size is smaller than the number of local observations, then the local ensemble perturbation is written as $\widehat{X}^a = (I - 0.5 \cdot K \cdot H) \cdot \widehat{X}^{f'}$; in the opposite case, it becomes $\widehat{X}^a = (I - 0.5 \cdot (\rho^\circ \widehat{B})H^T R^{-1}H) \widehat{X}^{f'}$. Let $\widehat{X}^{f'}$ be an $m \times (n+1)$ matrix such that

$$\widehat{X}^{f'} \cdot \left(\widehat{X}^{f'}\right)^T = \left(\widehat{Q} \cdot \widehat{D}^{1/2}\right) \cdot \left(\widehat{Q} \cdot \widehat{D}^{1/2}\right)^T, \quad (13)$$

$$\widehat{X}^{f'} \cdot \mathbf{1} = 0. \quad (14)$$

Note that $\widehat{X}^{f'} \in \Upsilon^{m \times \widehat{n}}$ represents \widehat{n} ensemble perturbations, which call that ‘‘augmented ensemble’’ for simplicity; and \widehat{n} denotes the ensemble increments. Then, $\widehat{X}^{f'}$ is a solution to equation (12) with $\widehat{n} = n + 1$ perturbations. For equation (13), how can we efficiently construct $\widehat{X}^{f'}$ from $\widehat{Q} \cdot \widehat{D}^{1/2}$. Let $X^{f'} \in \Upsilon^{m \times (n-1)}$ matrix. We want to construct a matrix that has the same empirical covariance matrix and which is centered. Since $X^{f'}$ has rank at most $n - 1$, we need to find a matrix $Z \in \Upsilon^{m \times n}$ such that

$$Z \cdot Z^T = X^{f'} \cdot \left(X^{f'}\right)^T, \quad (15)$$

$$Z \cdot \mathbf{1} = 0. \quad (16)$$

For $\forall \xi \in \{-1, 1\}$, we define $\lambda = \sqrt{n}(\sqrt{n} - \xi)^{-1}$ (multiplicative inflation factor) and Q_ξ as the $n \times n$ matrix whose coefficients are as follows:

$$[Q_\xi]_{i,j} = \begin{cases} n^{-1/2} \cdot \xi & \text{if } i = 1 \text{ or } j = 1, \\ 1 - (n^{-1} \cdot \lambda) & \text{if } i = j \geq 2, \\ n^{-1} \cdot \lambda & \text{else.} \end{cases} \quad (17)$$

It can be easily checked that $Q_\xi \cdot Q_\xi^T = \mathbf{I}$ (i.e., Q_ξ is an orthogonal matrix) and $Q_\xi \cdot \mathbf{1} = e_1$, the first basis vector.

Let $W \in \Upsilon^{m \times n}$ matrix whose first column is zero and whose other columns are those of $X^{f'}$, that is,

$$W = \left[0, X^{f'}\right]. \quad (18)$$

By construction, $Z = W \cdot Q_\xi$ is a solution of equations (15) and (16).

For the study of covariance localization, it is mainly limited to the effect of each observation on the analyzed distance within a fixed horizontal distance L and vertical distance V . In addition, each model component within the background covariance matrix B is correlated with each other. Therefore, in the context of the sequential assimilation scheme, the elements of the Kalman gain matrix K are zero beyond these distances (cf. Houtekamer and Zhang [14]; equation (20)).

The observations in the grid point are applied to different weights based on their physical distance from the grid points. These weights are mainly imposed on the inverse of the Schur product of the observation covariance matrix R and the correlation coefficient matrix to implement (cf.

Hunt et al. [8]). The traditional localization function is an adjustable function with a distribution function similar to the probability density function of a normal distribution. In addition, efforts are being made to find the Gaussian function that can be fitted to the widely used Gaspari-Cohn function with compact support (cf. Gaspari and Cohn [42]), where the observations in this localization region are applied different observation weights and the Gaspari-Cohn function outside this localization region is zero. The step function is as follows:

$$w(z, l) := \begin{cases} 1 & \text{if } |z| \leq l, \\ 0 & \text{else,} \end{cases} \quad (19)$$

where l denotes the adjustable cutting radius. For observation weighting, there are mainly weights inside the observation region and weights outside the observation region (usually set to zero). Therefore, equation (19) is also considered as domain localization (DL) (cf. Kirchgessner et al. [15]).

3. Twin Models and Experimental Settings

3.1. Experimental Settings for the Lorenz-96 Model. For the simple 40-dimensional Lorenz-96 model (cf. Lorenz and Emanuel [43]), the model parameters are chosen as follows: to ensure sufficient integration frequency for each assimilated variable in the model, all 40 grid points are observed. We set the 0.05-time step (i.e., corresponds to 6 h atmosphere valid time), and this 0.05-time step is labeled as a DA step. To introduce more observation errors in a more challenging way, the observation error covariance is $3.05 \cdot \mathbf{I} \in \Upsilon^{p \times p}$. Here, selecting the observation error covariance value of 3.05 is used for validating the assimilation performance of CL and CL-MC, this is mainly because introducing larger observation error covariance values whether the CL-MC method still achieves better DA performance. The Lorenz-96 model state $x_j(t)$, $j = 1, \dots, n$ is based on the following equation:

$$\bullet x_j = (x_{j+1} - x_{j-2}) \cdot x_{j-1} - x_j + F, \quad (20)$$

Here, $n = 40$, and the observation members are 20 (i.e., $p = 20$). The input parameters (e.g., inflation factor, localization radius, and ensemble size) are also considered. The initial conditions are generated from the filter’s ensemble members. The model was first integrated forward over several years of real-world time without assimilating any variables. The initial forecast ensemble members are similar except for the imposition of additional Gaussian noise. The true values are produced randomly by small perturbations of the instability. In this case, the initial ensemble members move away from the model attractor. There is no clear relationship between the spread and the error. The spin-up period of one and a half years is sufficient to eliminate any unfavorable effects (cf. Hoteit et al. [44]). Constant multiplicative inflation factor (cf. Anderson and Anderson [6]) is used to enlarge the ensemble spread and prevent filter divergence. The default strength forcing is 8. A set of data

TABLE 1: Manually tuned optimal localization parameters and default DA parameters for the filters with the Lorenz-96 model.

Method	Localization radius	Inflation factor	Ensemble size (m)	Observation error variance (R)	Observation (p)
CL	9	1.051	25	3.05	20
CL-MC	12	1.035	25	3.05	20

assimilation experiments for each data assimilation method is performed for 365 days, using the default settings and different assimilation parameters (inflation and localization). The manually tuned optimal parameters are shown in Table 1.

3.2. Experimental Settings for the SPEEDY Model. The moderately complex SPEEDY model (cf. Ritchie [45]; Miyoshi [46]) is similar to weather dynamics. Its assimilation variables contain temperature, dimensional wind component, meridional wind component, specific humidity, surface pressure, and precipitation (cf. Bourke [47]). It has eight vertical levels of 96×48 resolution in the sigma coordinate system. The SPEEDY model creates an output file every 6 hours for weather prediction (cf. Greybush et al. [11]). A more detailed SPEEDY model is available in Molteni [48]. The initial ensemble of states is generated randomly by the model with a spin-up period of 1 year (cf. Kondo and Miyoshi [26]). The natural runtime starts from a stationary standard atmosphere (i.e., no wind) at 0000 UTC 1 January. The total model component is $m = 133632$. For all filter cases, the model state space exceeds the ensemble size ($m \gg n$).

For the SPEEDY model, identical observations are generated by adding Gaussian-distributed random numbers with the natural run every 6 h at radiosonde-like locations for 8 levels (100 hPa, 200 hPa, 360 hPa, 520 hPa, 800 hPa, 850 hPa, 880 hPa, 925 hPa). The standard deviations of the observation errors for the zonal wind, temperature, specific humidity, and surface pressure are fixed at 1.0 ms^{-1} , 1.0 K , 1.0 gkg^{-1} , and 1.0 hPa .

3.3. Validation Statistics. If there are m -dimensional model components, the root mean square errors (RMSEs cf. equation (18) in Bai et al. [21]) and the ensemble spread (cf. equation (22) in Hoteit et al. [44]) are applied to evaluate the assimilation performance error factors, the equation is as follows:

$$RMSEs = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_i^a - x_i^f)^2}, \quad (21)$$

$$Spread = \sqrt{\frac{1}{n} \sum_{i=1}^n \sigma_{x_i}^2}, \quad (22)$$

where the subscript i denotes the index of the state components and the subscript j refers to the index of the ensemble members; x_i^a and x_i^f denote the i -th analyzed value and the true value, respectively; σ_{x_i} denotes the ensemble variance of x_i^f , the analysis quality of the assimilation algorithm can be measured more comprehensively by

averaging the computations in space. Usually, better assimilation results correspond to smaller RMSEs.

3.4. Experimental Results for the Lorenz-96 Model. This section analyzes the CL and CL-MC algorithms assimilation results using the Lorenz-96 model. Sensitivity experiments are conducted on the CL and CL-MC algorithms by configuring different localization scales with the default ensemble size, the fixed observation members, the strength forcing, and the manually tuned optimal inflation value (Table 1). To improve the accuracy of the experiments, each experiment was repeated twice, applying their mean value as the evaluation criterion for each algorithm.

The power spectral density (PSD; cf. Sakov and Bertino [10]) evaluation criterium is used to test the performance of the CL and CL-MC algorithms. With the different localization scales, the CL and CL-MC algorithms of PSD change correspondingly as shown in Figure 2. The blue curve ‘‘Forecast’’ in the graph represents the Fourier spectrum of the ensemble singular value of the forecast values, and the green curve ‘‘CL’’ and the red curve ‘‘CL-MC’’ represent the Fourier spectrum of the ensemble singular values of the analysis values. The PSD of the vertical coordinate indicates the PSD of the ensemble mean and the area under the panel represents the variance.

Figure 2 shows the influence of the CL and CL-MC algorithms on the PSD with different localization radius. Here, the default ensemble size, observation fraction, observation error variance, strength forcing, and the optimization of the inflation factor are configured. The result shows that the areas under the curves gradually become small while the PSD value is less than the forecast value for the CL and CL-MC algorithms as the spectral component increases. In addition, as the spectral component increases, the PSD values for the CL and CL-MC algorithms exhibit different results. (1) The area under the curve for the CL-MC algorithm is always smaller than the CL algorithm; however, as the spectral component increases, the PSD values for the CL and CL-MC methods are far less than forecast values. (2) Based on the modified Cholesky decomposition leads to the magnitude of the CL-MC algorithm between 0.1 and 0.2, in the CL scheme, the magnitudes are between 0.10 and 0.30 when the localization radius is 10 (Figure 2(a)). When the localization radius is increased to 15, the modified Cholesky decomposition leads to magnitudes in the PSD of the CL-MC algorithm between 0.05 and 0.15 and in the CL algorithm between 0.13 and 0.25 (Figure 2(b)). It is worth noting that the PSD value of the CL and CL-MC methods become small when the localization radius is 20. However, the PSD value of CL-MC is still smaller than the CL. Overall, the CL-MC algorithm exhibits better performance than the CL algorithm.

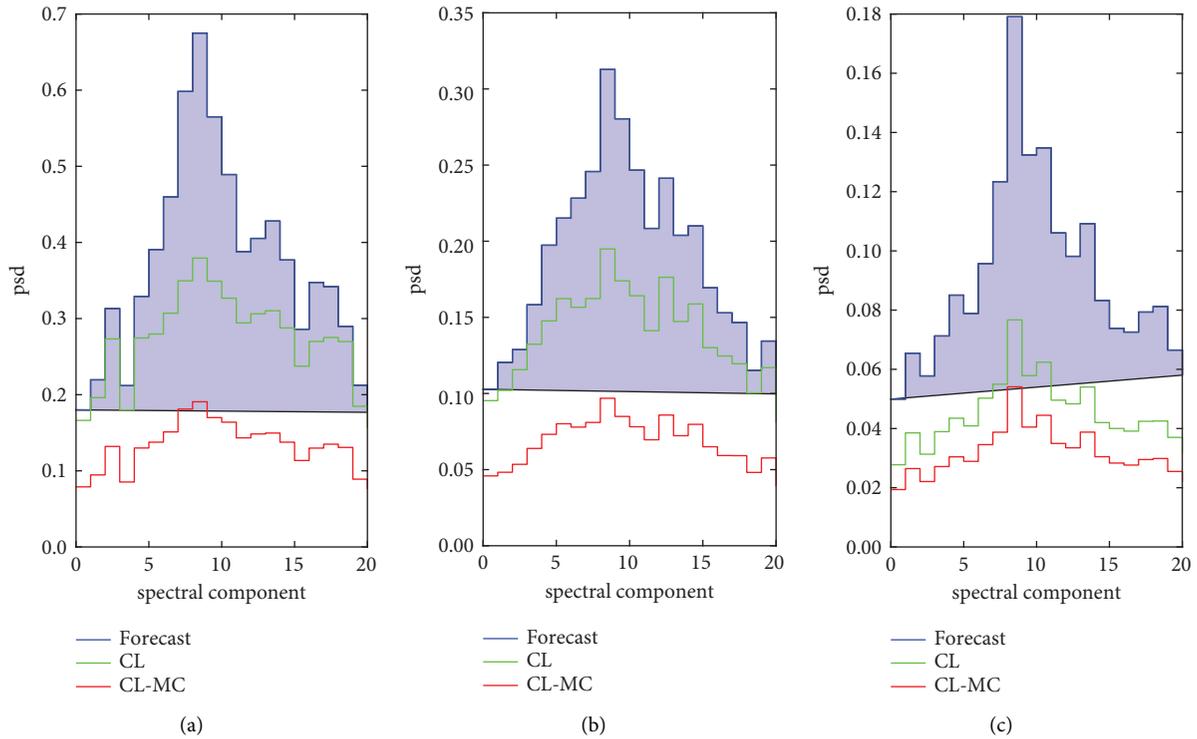


FIGURE 2: Influence of the CL and CL-MC algorithms on the PSD with the different localization radii. (a) Localization radius = 10. (b) Localization radius = 15. (c) Localization radius = 20.

In the local analysis process, it is known from equation (1) that the update of Kalman gain plays an essential role in the analysis value, which affects the final analysis accuracy. Here, the default ensemble size, observation fraction, observation error variance, strength forcing, and the optimization of the inflation factor and localization radius are used (Table 1). Figure 3 shows the impact of CL and CL-MC methods on Kalman gain for different localization radii. In this case, the Kalman gain is a 40×20 matrix. The horizontal axes in Figure 3 indicate the index numbers of the corresponding observations, and the vertical axes indicate the index numbers of the state variables. The Kalman gain values were obtained by the CL and CL-MC methods, as well as without using the localization scheme. The figure is similar to literature Sakov and Bertino [10] in Figure 3. The results show that the Kalman gain without the localization scheme retains the spurious correlation from the long-range observations, whereas the CL and CL-MC methods eliminate the spurious correlation of the long-range in the gain matrix. Furthermore, we observed that the Kalman gain matrix exhibits significant diagonality and symmetry due to the use of localization schemes (including CL and CL-MC). However, the superior performance of the CL-MC method may be more significant compared to the CL scheme (as seen from the colorbar scale). In terms of the gain obtained for each observation (Kii), increasing the localization radius is favorable for the CL and CL-MC methods to obtain the small Kalman gain, whereas the CL, CL-MC, and without using the localization scheme have larger Kalman gain values

when the observation index approaches 20 (fourth column of Figure 3).

Time series of the RMSEs for zonal wind (ms^{-1}) at the fourth model level (~ 510 hPa) with different ensemble sizes: (a) ensemble size = 75, (b) ensemble size = 100, (c) ensemble size = 125, and (d) ensemble size = 200. Here, 14600 assimilation steps correspond to 10 years from 0000 UTC 1 January 1982 to 0000 UTC 1 January 1992. The localization method is beneficial for the ensemble Kalman filter to eliminate spurious long-range correlations. Especially, the choice of localization parameters is unknown a priori, and most researchers obtain the optimal localization and inflation factor parameters sufficient to handle the corresponding assimilation problem with continuous iterative experiments. In this study, the inflation factor and localization are manually tuned with optimal parameters. Here, the default ensemble size, observation members, observation error variance, strength forcing, and the optimization of the inflation factor are used (Table 1). Figure 4 shows the time series of the RMSEs for the CL and CL-MC algorithms with different localization radii. In the operational model, the assimilation cycle is 1460 steps. One assimilation cycle is approximately equivalent to 6 hours of “weather”. The 1460 assimilation cycles are one year (365 days). The results show that the CL and CL-MC methods achieve better filter performance with increasing localization radius. However, the CL and CL-MC methods perform large RMSEs since the instability of the system state at the initial moment. The RMSEs of the CL and CL-MC methods present convergence

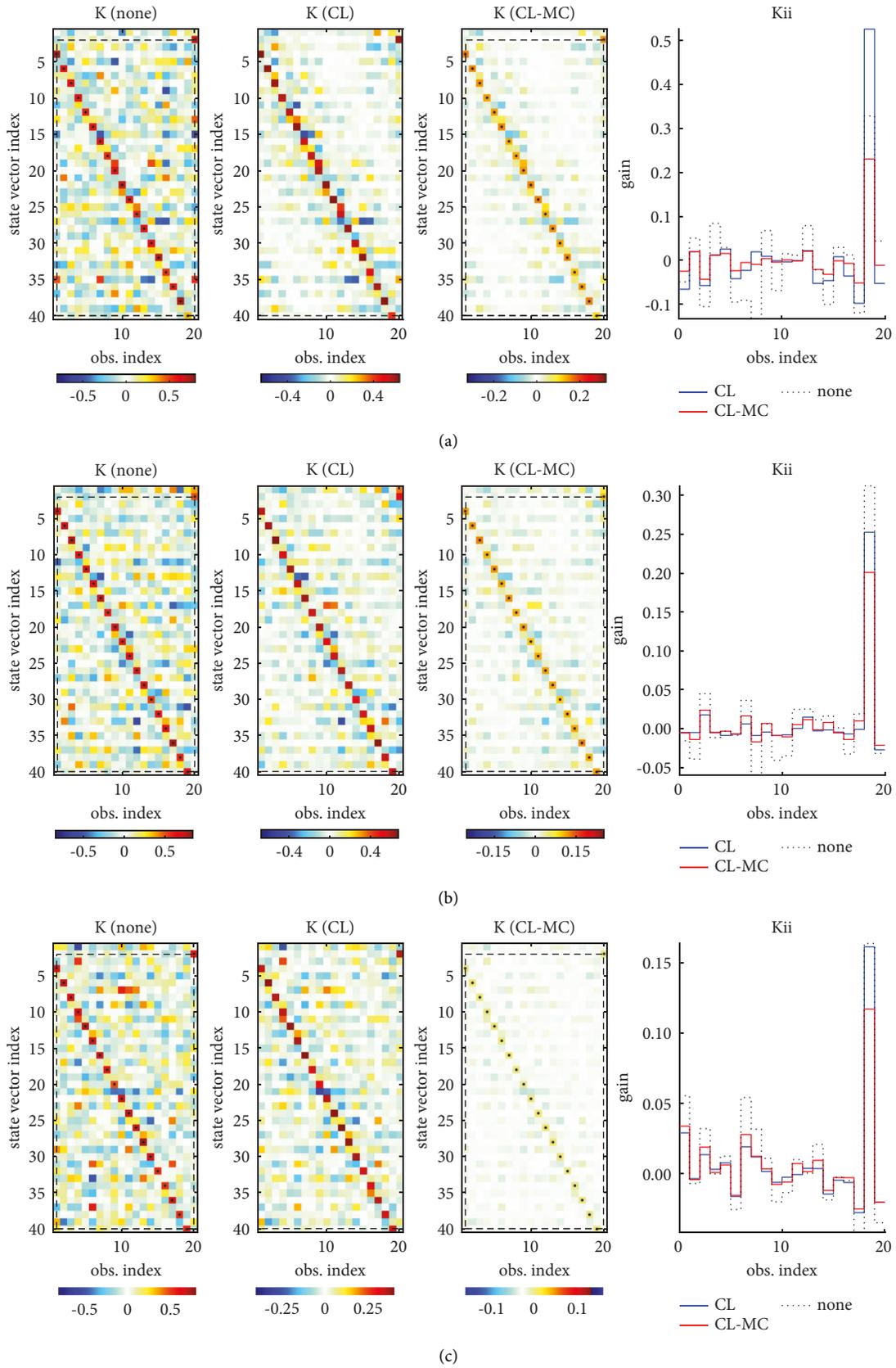


FIGURE 3: Influence of the CL and CL-MC on the Kalman gain matrix for model component $i = 40$. (a) Localization radius = 10. (b) Localization radius = 15. (c) Localization radius = 20. “K(none)” represents the without using localization scheme; “K(CL)” and “K(CL-MC)” represent using and without using modified Cholesky decomposition, respectively. “Kii” denotes the Kalman gain from each observation for CL, CL-MC, and without using the localization scheme.

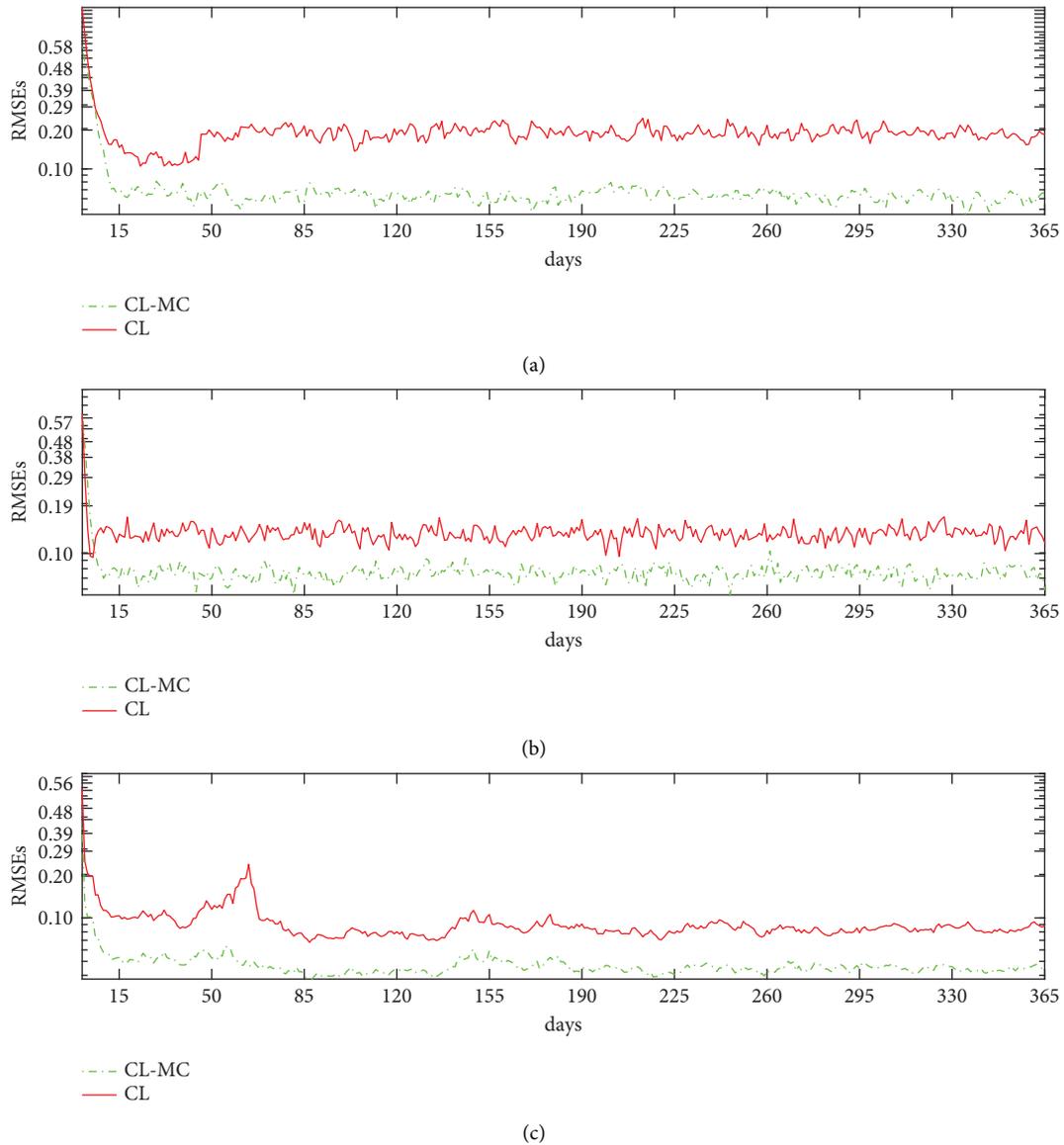


FIGURE 4: Experiment with the Lorenz-96 model. The time series of RMSEs for the CL and CL-MC methods with the localization radii are (a) localization radius = 10, (b) localization radius = 15, and (c) localization radius = 20, respectively.

results as the assimilation step increases. After a period of assimilation steps, the RMSEs do not reduce and achieve a convergent trend, this may be due to the optimization of the inflation factor and the default ensemble members of 25, since small inflation and a large localization radius are desired. In addition, it is worth noting that after about 50 days of assimilation steps, the CL exhibits inferior balance (Figure 4(c)), this may be due to the missing part of the analysis increment in updating the local state variables, resulting in poor assimilation results. However, the CL-MC method shows better assimilation results and achieves well-balanced during the whole assimilation process.

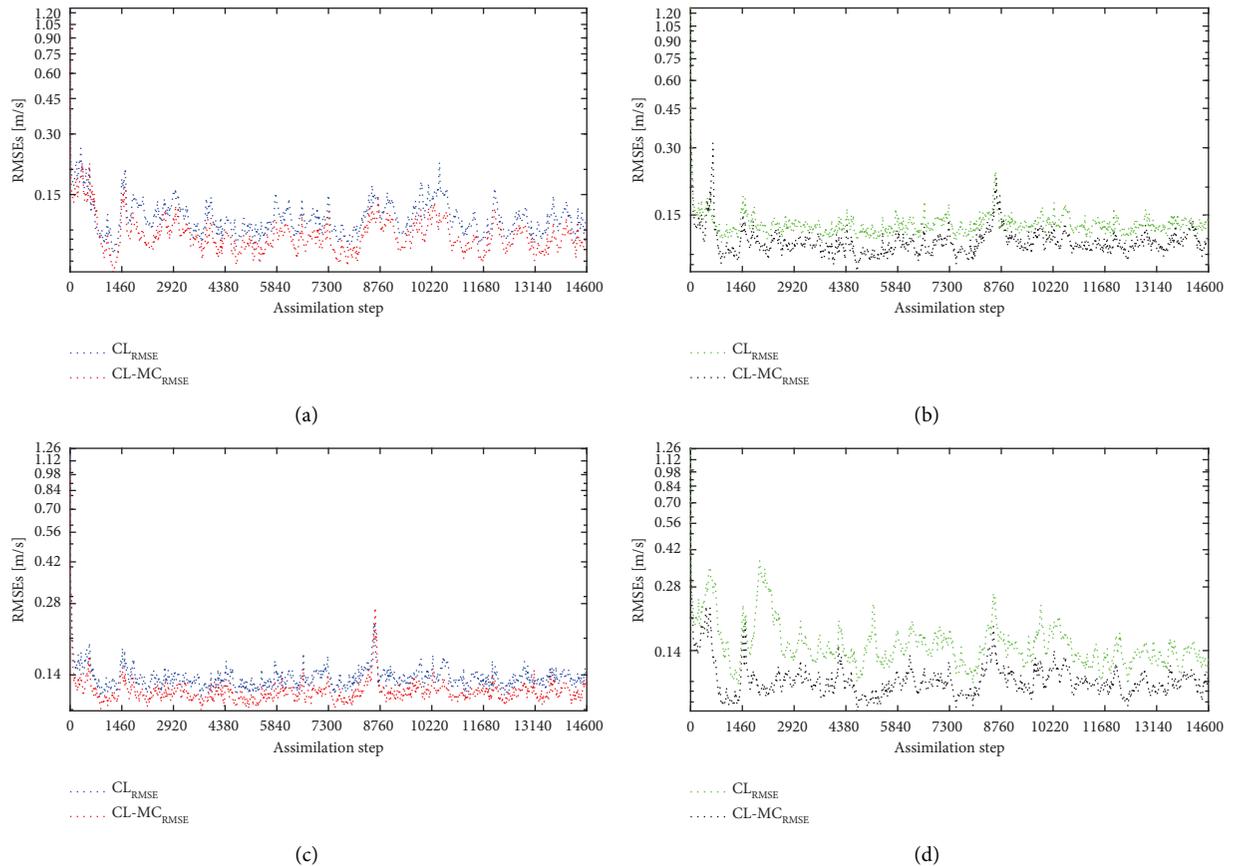
3.5. Experimental Results for the SPEEDY Model. In the abovementioned experiments, the effectiveness of CL and CL-MC methods was checked by the toy Lorenz-96 model

with different assimilation parameters. To further investigate the application of CL and CL-MC methods in large-scale models, in this section, the performance of the CL and CL-MC are further examined on the SPEEDY model. All experiments are initialized by the initial ensemble randomly chosen from the nature run at 0000 UTC 1 January (cf. Miyoshi [46] and Kang et al. [49]). Here, to avoid manual optimization of the inflation parameters, the adaptive covariance inflation method was applied (cf. Miyoshi [50]).

To investigate the assimilation results of the CL and CL-MC methods in the large-scale models, an experiment was conducted using different ensemble sizes, since this choice plays an essential part in the DA process. In practical applications, the feasible ensemble size is usually smaller than the degrees of freedom of the system (or dimension of the unstable subspace or error subspace). The other assimilation parameters, such as localization scale (containing horizontal

TABLE 2: Experiment settings using the SPEEDY model to verify the sensitivity of the ensemble size for the CL and CL-MC filters.

Expt	E75	E100	E125	E200
Ensemble size	75	100	125	200
Horizontal localization radius		2000 km		
Vertical localization radius (SD)		$0.12 \ln p$		
Initial time and date		0000 UTC 1 Jan 1982		
Final time and date		0000 UTC 1 Jan 1992		
Inflation	Adaptive (1.045/ 1.022)	Adaptive (1.112/ 1.018)	Adaptive (1.028/ 1.012)	Adaptive (1.052/ 1.002)

FIGURE 5: Time series of the RMSEs for zonal wind (ms^{-1}) at the fourth model level (~ 510 hPa) with different ensemble sizes. (a) Ensemble size = 75. (b) Ensemble size = 100. (c) Ensemble size = 125. (d) Ensemble size = 200. Here, 14600 assimilation steps correspond to 10 years from 0000 UTC 1 January 1982 to 0000 UTC 1 January 1992.

and vertical scales), observation member, and the observation error variance (1.0 ms^{-1}) are also applied as the default values (Table 2), and an adaptive covariance inflation method is also employed to enlarge the ensemble spread and prevent filter divergence. Figure 5 shows the time series of RMSEs of the CL-MC and CL methods are examined for the zonal wind at the fourth level mode (~ 510 hPa). The results show that increasing the ensemble size improves the assimilation results for the CL-MC and CL methods. When the ensemble size is 75 and 125, the RMSEs profiles of the CL-MC and CL methods show a similar assimilation behavior (Figures 5(b) and 5(c)). However, as the assimilation step increases, in particular when the ensemble size is 100, 125, and 200, the RMSEs of CL-MC and CL methods present

peaks (at the same assimilation step) (Figures 5(b) to 5(d)). This may be the adaptive inflation for larger ensemble sizes, since small inflation and broad localization scales are what we expect as the ensemble size increases. The RMSEs of the CL-MC method are much smaller than CL when the ensemble size is 200, this implies that a larger ensemble size is beneficial for the CL-MC method (Figure 5(d)). Overall, the CL-MC method outperforms CL, especially when the ensemble size is larger.

To better understand the mechanisms behind the CL and CL-MC methods and to test their applicability to real atmospheric systems. The zonal wind of the spatial distribution of the time-averaged analysis RMSEs for the CL-MC and CL methods are examined from 0000UTC 1 March 1982

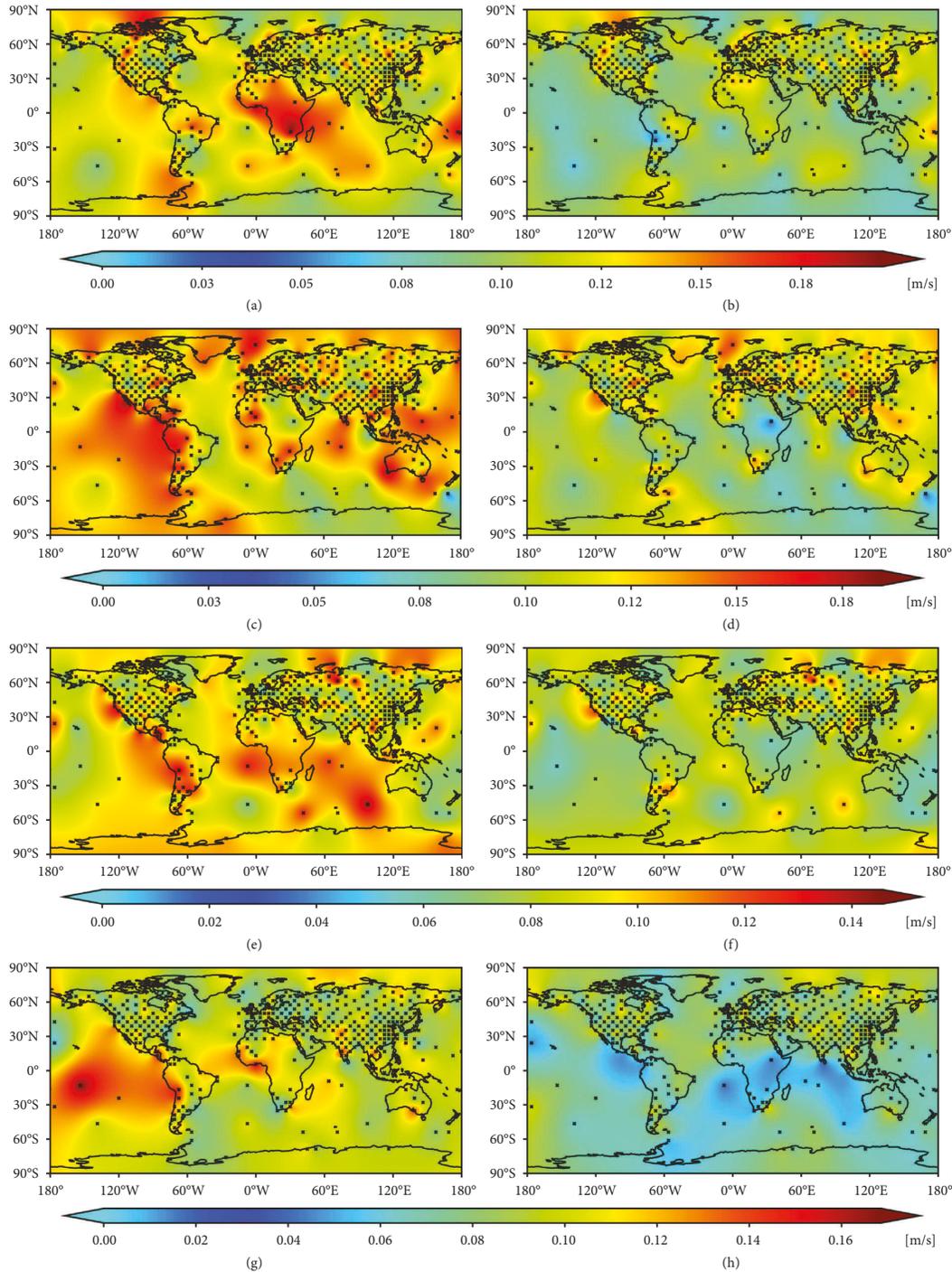


FIGURE 6: The spatial distributions of the time-mean analysis RMSEs for zonal wind from 0000UTC 1 March 1982 to 0000 UTC 1 April 1982 with different ensemble sizes, and (a), (b) ensemble size = 75; (c), (d) ensemble size = 100; (e), (f) ensemble size = 125; (g), (h) ensemble size = 200. (a), (c), (e), and (g) refer to the analysis result from the CL, and (b), (d), (f), and (h) refer to the analysis result from the CL-MC. The \times symbols indicate the observation stations.

to 0000UTC 1 April 1982 with different ensemble sizes. The other assimilation parameters, localization scale, observation member, and the observation error variance (1.0 ms^{-1}) are used as their default values. Figure 6 displays the spatial distribution of the time-averaged analysis RMSEs for the CL-MC and CL methods from 0000UTC 1 March 1982 to 0000UTC 1 April 1982 with different ensemble sizes. The

results show that the CL and CL-MC methods exhibit inconsistent assimilation results when the ensemble sizes are 75 and 100. (1) The time-averaged analysis RMSEs of the CL method are particularly pronounced, especially when the ensemble size is 100, in the longitude region $0^\circ \text{ W} - 150^\circ \text{ W}$ (observations are sparse), where the average analysis RMSEs of the zonal wind reaches about 0.152 ms^{-1} . The time-

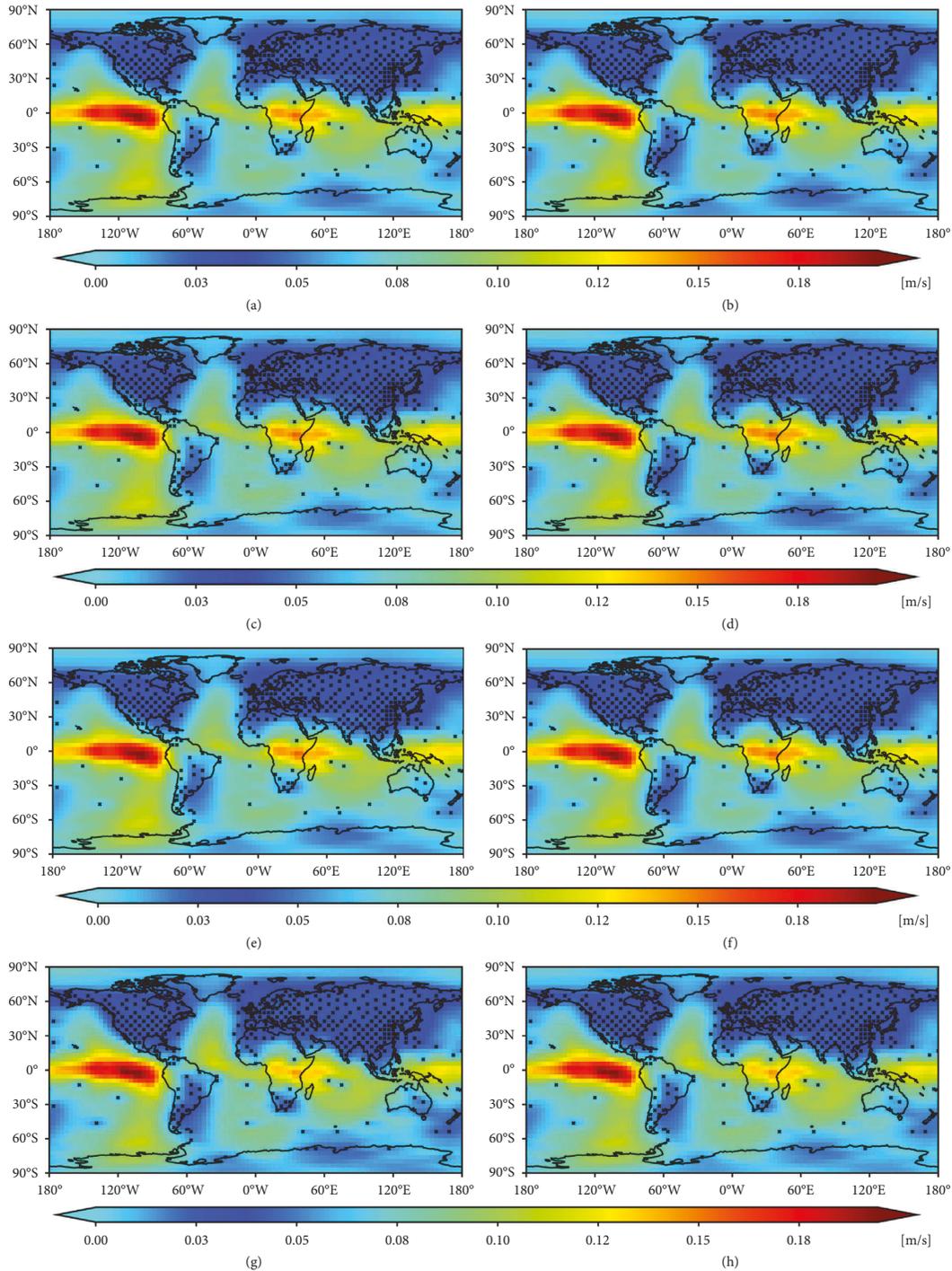


FIGURE 7: The spatial distributions of the time-mean analysis spread instead of analysis RMSEs for zonal wind from 0000UTC 1 March 1982 to 0000 UTC 1 April 1982 with different ensemble sizes, and (a), (b) ensemble size = 75; (c), (d) ensemble size = 100; (e), (f) ensemble size = 125; (g), (h) ensemble size = 200. (a), (c), (e), and (g) refer to the analysis result from the CL, and (b), (d), (f), and (h) refer to the analysis result from the CL-MC.

averaged analysis RMSEs of the CL-MC method are relatively small (about 0.116 ms^{-1}) (Figures 6(c) and 6(d)). (2) In the high latitudes, the time-averaged analysis RMSEs reach 0.148 ms^{-1} for the CL method (Figure 6(c)) and about 0.135 ms^{-1} for the CL-MC method (Figure 6(d)). In sparse observations, it is worth noting that EnKF-based CL methods are susceptible to large adaptive covariance

inflation factors, with a fixed localization scale, smaller covariance inflation factors are what one would expect as the ensemble size increases. When the ensemble size is 125, the spatial distribution of the time-averaged analysis RMSEs of the CL and CL-MC methods reaches inconsistent results. (1) The CL scheme achieves larger RMSEs, mainly in the low latitudes ($120^\circ \text{ W}-120^\circ \text{ E}$), where the averaged analysis

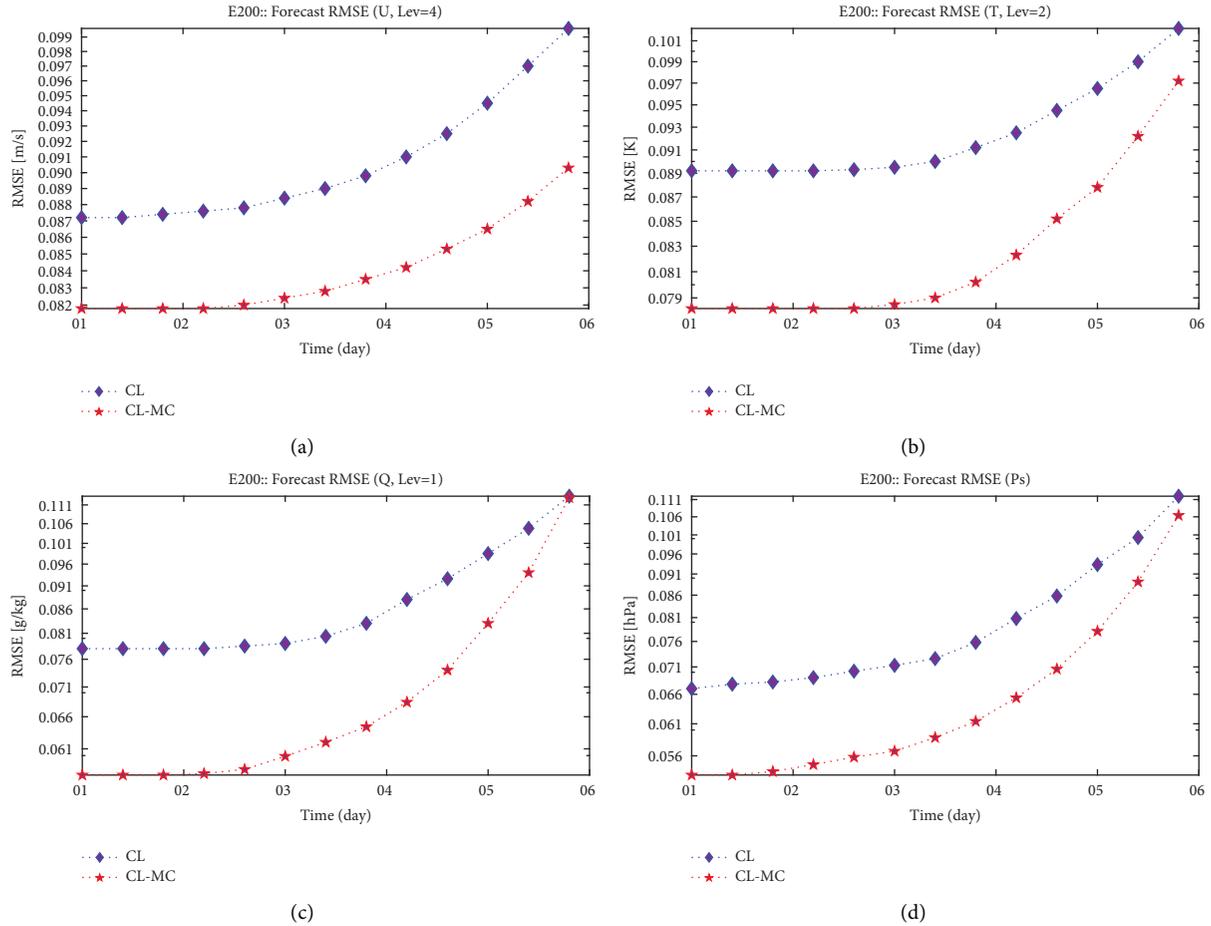


FIGURE 8: Experiment results with the SPEEDY model. Global mean 6-day forecast RMSEs averaged from 0000 UTC 1 March to 0000 UTC 1 May for (a) zonal wind (ms^{-1}) at the fourth model level (~ 510 hPa), (b) temperature (K) at the second model level (~ 835 hPa), (c) specific humidity (gkg^{-1}) at the lowest level (~ 925 hPa), and (d) surface pressure (hPa).

RMSEs are approximately 0.120 ms^{-1} (Figure 6(e)). (2) In most of the regions, the CL-MC method obtains the smaller analysis RMSEs (the maximum value reaches 0.112 ms^{-1}) (Figure 6(f)). This is probably due to the CL-MC scheme retaining the main characteristics of background error covariance for larger localization radius. CL scheme neglects some short to medium-range correlations. When the ensemble size is 200, over the Arctic, especially near 60°W – 60°E where observations are relatively sparse, the time-averaged analysis RMSEs for the CL-MC method reach almost 0.04 ms^{-1} (Figures 6(h)), the CL scheme reaches about 0.10 ms^{-1} (Figure 6(g)), as a reference, observations are available only at the position shown at the intersection with an error standard deviation of 1.0 ms^{-1} . In general, these results demonstrate that the CL and CL-MC techniques tested here resulted in a well-constrained analysis, and the error is significantly reduced with increasing ensemble sizes, but when the ensemble size increases from 75 to 125, the error of the CL scheme does not present a significant reduction in the few areas where observations are sparsely distributed. This may be due to the broader localization scale and the larger inflation factor used for the CL scheme since the less inflation and the wider localization scale is what one

would expect when the ensemble size increases (Figures 6(a), 6(c) and 6(e)). Therefore, the superior behavior of the CL-MC method here may be more significant when considering practical application aspects.

Finally, the stability of the CL and CL-MC methods for adaptive covariance inflation factor is explored by using time-averaged analysis ensemble spread instead of time-averaged analysis RMSEs (cf. Kondo and Miyoshi [26]). The results show that in regions with few observations, the effect of the CL and CL-MC methods on the time-averaged analysis ensemble spread is larger with increasing ensemble size. In contrast, in regions with a larger number of observations, the time-averaged analysis ensemble spread values are surprisingly small (Figures 7(a) to 7(h)). Thus, the adaptive covariance inflation factors exhibit strong robustness for different ensemble sizes. Table 1 lists the global averages of the adaptive covariance inflation parameters for CL (left value in the last row) and CL-MC (right bold value in the last row) at the fourth model level.

The previous experimental cases are based on analysis RMSEs to examine the effectiveness of CL and CL-MC methods, where the CL-MC method outperforms CL, and the CL and CL-MC are susceptible to local instability since

relatively large adaptive covariance inflation factors are used. To prevent such a situation, the short-term forecast RMSEs of the CL and CL-MC methods are examined with the default observations, observation error covariance, localization scale, and adaptive covariance inflation factors. Here, the ensemble size is 200 members and the analysis ensemble mean is used as the initial condition ensemble members. For all variables (including zonal wind, temperature, specific humidity, and surface pressure), the average 6-day forecast RMSEs for 245 forecast cases from 0000UTC 1 March to 0000UTC 1 May are investigated (Figure 8). The results show that for all variables, except for the specific humidity, the profiles of the average 6-day forecast RMSEs are never crossed. The average forecast RMSEs of the CL-MC method is always smaller than CL for the entire 6-day forecast from the initial time. By calculating the average 6-day forecast RMSEs for 245 forecast cases, we find that the 1-day forecast RMSEs of CL is approximately equal to the 5-day forecast RMSEs of CL-MC. This suggests that the CL-MC method has potential advantages for relatively large ensemble sizes and for considering the correlation of short-term forecast errors. However, during the experimental settings, the results may be overly optimistic since assumptions such as fully known observation error statistics and a simplified model with low resolution are used. In general, the CL-MC clearly outperforms the CL method in terms of short-term forecasting.

4. Conclusion

In ensemble DA, since using a large ensemble size leads to sampling error and imbalance in the DA system, this paper proposes a covariance localization method (CL-MC) with Cholesky decomposition to eliminate sampling error and achieve the estimation of the inverse of the background covariance matrix. Several advantages of using such a new approach are listed below. (1) Typically, the ensemble covariance is a full-rank matrix, and most rank-deficient matrices depend on the Schur-product. Since the matrix has large dimensionality, it is not favorable for matrix storage. The predefined sparse matrix structure is constructed as the inverse of the covariance coefficient by Cholesky decomposition. If two distant model components are uncorrelated, the corresponding entries in the inverse covariance matrix are zero, thus avoiding the use of excessive memory; (2) Since the only nonzero entries in the Cholesky factors correspond to model components close to each other, the use of sparse structure on the inverse background covariance matrix is a form of covariance localization. (3) To facilitate the storage of the background error covariance matrix, an adjustable singular value decomposition method is applied to the background error matrix.

To verify the effectiveness of the proposed method, a series of experiments are conducted in the simple Lorenz-96 model and the SPEEDY model. The results of the Lorenz-96 model show that the RMSEs obtained by CL-MC are smaller than the CL method with sparse observation results ($p = 20$). In addition, increasing the localization radius is beneficial for using more observation information and improving the

assimilation result of the proposed algorithm. The results of the SPEEDY model show that increasing the ensemble number is favorable for the CL and CL-MC methods to obtain better assimilation results. However, the differences in the spatial distribution of the analysis ensemble spread and the analysis RMSEs between the CL and CL-MC methods are more obvious. This may be due to the fact that the use of adaptive covariance localization is susceptible to large inflation factors in the filtering process and reflects the changing characteristics of the background field of the chaotic model, which is not a good choice for the DA system. Thus, investigating adaptive covariance localization methods is a future topic. In addition, calculating the average 6-day forecast RMSEs for 245 forecast cases, we find that the 1-day forecast RMSEs of CL are approximately equal to the 5-day forecast RMSEs of CL-MC. This suggests that the CL-MC method has potential advantages for relatively large ensemble sizes considering the correlation of short-term forecast errors.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Mingheng Chang performed conceptualization, methodology, software, validation, original draft preparation, and reviewing. Hongchao Zuo supervised conceptualization and methodology. Jikai Duan performed investigation and Editing.

Acknowledgments

This work was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) under grant no. 2019QZKK0103, and the National Natural Science Foundation of China under grant no. 422750699. This work was also supported by Lanzhou University, Fundamental Research Funds for the Central Universities under grant number lzujbky-2021-sp61. Many thanks to Pavel Sakov for providing the publicly available EnKF-Matlab software.

References

- [1] M. Katzfuss, J. R. Stroud, and C. K. Wikle, "Understanding the ensemble kalman filter," *The American Statistician*, vol. 70, no. 4, pp. 350–357, 2016.
- [2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.
- [3] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research: Oceans*, vol. 99, no. C5, pp. 10143–210162, 1994.

- [4] G. Evensen and P. J. van Leeuwen, "An ensemble kalman smoother for nonlinear dynamics," *Monthly Weather Review*, vol. 128, no. 6, pp. 1852–1867, 2000.
- [5] R. Furrer and T. Bengtsson, "Estimation of high-dimensional prior and posterior covariance matrices in kalman filter variants," *Journal of Multivariate Analysis*, vol. 98, no. 2, pp. 227–255, 2007.
- [6] J. L. Anderson and S. L. Anderson, "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Monthly Weather Review*, vol. 127, no. 12, pp. 2741–2758, 1999.
- [7] T. M. Hamill, J. S. Whitaker, and C. Snyder, "Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter," *Monthly Weather Review*, vol. 129, no. 11, pp. 2776–2790, 2001.
- [8] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, "Efficient data assimilation for spatiotemporal chaos: a local ensemble transform kalman filter," *Physica D: Nonlinear Phenomena*, vol. 230, no. 1-2, pp. 112–126, 2007.
- [9] P. L. Houtekamer and H. L. Mitchell, "Data assimilation using an ensemble kalman filter technique," *Monthly Weather Review*, vol. 126, no. 3, pp. 796–811, 1998.
- [10] P. Sakov and L. Bertino, "Relation between two common localisation methods for the enfk," *Computational Geosciences*, vol. 15, no. 2, pp. 225–237, 2011.
- [11] S. J. Greybush, E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, "Balance and ensemble kalman filter localization techniques," *Monthly Weather Review*, vol. 139, no. 2, pp. 511–522, 2011.
- [12] J. D. Kepert, "Covariance localisation and balance in an ensemble kalman filter," *Quarterly Journal of the Royal Meteorological Society*, vol. 135, no. 642, pp. 1157–1176, 2009.
- [13] D. T. Kleist and K. Ide, "An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants," *Monthly Weather Review*, vol. 143, no. 2, pp. 452–470, 2015.
- [14] P. L. Houtekamer and F. Zhang, "Review of the ensemble kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 144, no. 12, pp. 4489–4532, 2016.
- [15] P. Kirchgeßner, L. Nerger, and A. Bunse-Gerstner, "On the choice of an optimal localization radius in ensemble kalman filter methods," *Monthly Weather Review*, vol. 142, no. 6, pp. 2165–2175, 2014.
- [16] T. Miyoshi and K. Kondo, "A multi-scale localization approach to an ensemble kalman filter," *Sola*, vol. 9, pp. 170–173, 2013.
- [17] K. Kondo, T. Miyoshi, and T. Tanaka, "Parameter sensitivities of the dual-localization approach in the local ensemble transform kalman filter," *Sola*, vol. 9, pp. 174–178, 2013.
- [18] Y.-L. Bai, M.-h. Chang, B.-x. Xu, and Y.-j. Ma, "Application of fuzzy control to observational error covariance matrices for data assimilation," *Fuzzy Systems and Data Mining*, vol. IV, pp. 553–562, 2018.
- [19] E. J. Fertig, B. R. Hunt, E. Ott, and I. Szunyogh, "Assimilating non-local observations with a local ensemble kalman filter," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 59, no. 5, pp. 719–730, 2007.
- [20] M. De La Chevrotière and J. Harlim, "A data-driven method for improving the correlation estimation in serial ensemble kalman filters," *Monthly Weather Review*, vol. 145, no. 3, pp. 985–1001, 2017.
- [21] Y. Bai, X. Ma, and L. Ding, "A fuzzy-logic-based covariance localization method in data assimilation," *Atmosphere*, vol. 11, no. 10, p. 1055, 2020.
- [22] C. Mingheng, B. Yulong, M. Xiaoyan, M. Ruoyu, and W. Lili, "Localization analysis of data assimilation methods coupled with fuzzy control algorithms," *Advances in Earth Science*, vol. 33, no. 8, pp. 874–883, 2018.
- [23] C. Mingheng, Z. Hongchao, B. Yulong, and D. Jikai, "The two localization methods based on fuzzy control," *Advances in Earth Science*, vol. 36, no. 2, pp. 874–883, 2021.
- [24] T. Miyoshi, K. Kondo, and T. Imamura, "The 10,240-member ensemble kalman filtering with an intermediate agcm," *Geophysical Research Letters*, vol. 41, no. 14, pp. 5264–5271, 2014.
- [25] T. Miyoshi, S. Yamane, and T. Enomoto, "Localizing the error covariance by physical distances within a local ensemble transform kalman filter (letkf)," *Sola*, vol. 3, pp. 89–92, 2007.
- [26] K. Kondo and T. Miyoshi, "Impact of removing covariance localization in an ensemble kalman filter: experiments with 10 240 members using an intermediate agcm," *Monthly Weather Review*, vol. 144, no. 12, pp. 4849–4865, 2016.
- [27] A. Farchi and M. Bocquet, "On the efficiency of covariance localisation of the ensemble kalman filter using augmented ensembles," *Frontiers in Applied Mathematics and Statistics*, vol. 5, p. 3, 2019.
- [28] X. Wang, H. G. Chipilski, C. H. Bishop, E. Satterfield, N. Baker, and J. S. Whitaker, "A multiscale local gain form ensemble transform kalman filter (mlgetkf)," *Monthly Weather Review*, vol. 149, no. 3, pp. 605–622, 2021.
- [29] X. Kang and X. Deng, "An improved modified cholesky decomposition approach for precision matrix estimation," *Journal of Statistical Computation and Simulation*, vol. 90, no. 3, pp. 443–464, 2020.
- [30] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, "Ensemble square root filters," *Monthly Weather Review*, vol. 131, no. 7, pp. 1485–1490, 2003.
- [31] G. Evensen, "The ensemble kalman filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, no. 4, pp. 343–367, 2003.
- [32] G. Burgers, P. Jan van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble kalman filter," *Monthly Weather Review*, vol. 126, no. 6, pp. 1719–1724, 1998.
- [33] J. L. Anderson, "An ensemble adjustment kalman filter for data assimilation," *Monthly Weather Review*, vol. 129, no. 12, pp. 2884–2903, 2001.
- [34] J. S. Whitaker and T. M. Hamill, "Ensemble data assimilation without perturbed observations," *Monthly Weather Review*, vol. 130, no. 7, pp. 1913–1924, 2002.
- [35] P. L. Houtekamer and H. L. Mitchell, "A sequential ensemble kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 129, no. 1, pp. 123–137, 2001.
- [36] M. Ghil and P. Malanotte-Rizzoli, "Data assimilation in meteorology and oceanography," *Advances in Geophysics*, vol. 33, pp. 141–266, 1991.
- [37] G. Ueno and T. Tsuchiya, "Covariance regularization in inverse space," *Quarterly Journal of the Royal Meteorological Society*, vol. 135, no. 642, pp. 1133–1156, 2009.
- [38] E. D. Nino-Ruiz, A. Sandu, and X. Deng, "An ensemble kalman filter implementation based on modified cholesky decomposition for inverse covariance matrix estimation," *SIAM Journal on Scientific Computing*, vol. 40, no. 2, pp. A867–A886, 2018.
- [39] W. Boyles and M. Katzfuss, "Ensemble kalman filter updates based on regularized sparse inverse cholesky factors," *Monthly Weather Review*, vol. 149, no. 7, pp. 2231–2238, 2021.

- [40] E. D. Nino-Ruiz, "A matrix-free posterior ensemble kalman filter implementation based on a modified cholesky decomposition," *Atmosphere*, vol. 8, no. 7, p. 125, 2017.
- [41] S. Bloom, L. Takacs, A. da Silva, and D. Ledvina, "Data assimilation using incremental analysis updates," *Monthly Weather Review*, vol. 124, no. 6, pp. 1256–1271, 1996.
- [42] G. Gaspari and S. E. Cohn, "Construction of correlation functions in two and three dimensions," *Quarterly Journal of the Royal Meteorological Society*, vol. 125, no. 554, pp. 723–757, 1999.
- [43] E. N. Lorenz and K. A. Emanuel, "Optimal sites for supplementary weather observations: simulation with a small model," *Journal of the Atmospheric Sciences*, vol. 55, no. 3, pp. 399–414, 1998.
- [44] I. Hoteit, D.-T. Pham, M. Gharamti, and X. Luo, "Mitigating observation perturbation sampling errors in the stochastic enkf," *Monthly Weather Review*, vol. 143, no. 7, pp. 2918–2936, 2015.
- [45] H. Ritchie, "Application of the semi-Lagrangian method to a multilevel spectral primitive-equations model," *Quarterly Journal of the Royal Meteorological Society*, vol. 117, no. 497, pp. 91–106, 1991.
- [46] T. Miyoshi, *Ensemble Kalman Filter Experiments with a Primitive-Equation Global Model*, University of Maryland, College Park, MD, USA, 2005.
- [47] W. Bourke, "A multi-level spectral model. i. formulation and hemispheric integrations," *Monthly Weather Review*, vol. 102, no. 10, pp. 687–701, 1974.
- [48] F. Molteni, "Atmospheric simulations using a gcm with simplified physical parametrizations. i: model climatology and variability in multi-decadal experiments," *Climate Dynamics*, vol. 20, no. 2-3, pp. 175–191, 2003.
- [49] J.-S. Kang, E. Kalnay, T. Miyoshi, J. Liu, and I. Fung, "Estimation of surface carbon fluxes with an advanced data assimilation methodology," *Journal of Geophysical Research: Atmospheres*, vol. 117, no. D24, 2012.
- [50] T. Miyoshi, "The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform kalman filter," *Monthly Weather Review*, vol. 139, no. 5, pp. 1519–1535, 2011.