



Article An Investigation of Adaptive Radius for the Covariance Localization in Ensemble Data Assimilation

Xiang Xing ^{1,2}, Bainian Liu ^{1,*}, Weimin Zhang ^{1,3}, Jianping Wu ¹, Xiaoqun Cao ¹, and Qunbo Huang ⁴

- College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China; xingxiang14@nudt.edu.cn (X.X.); wmzhang104@139.com (W.Z.); wjp@nudt.edu.cn (J.W.); caoxiaoqun@nudt.edu.cn (X.C.)
- ² College of Computer, National University of Defense Technology, Changsha 410073, China
- ³ Laboratory of Software Engineering for Complex Systems, Changsha 410073, China
- ⁴ 93110 Troops of PLA, Beijing 100843, China; hqb09@163.com
- Correspondence: bnliu@nudt.edu.cn

Abstract: The covariance matrix estimated from the ensemble data assimilation always suffers from filter collapse because of the spurious correlations induced by the finite ensemble size. The localization technique is applied to ameliorate this issue, which has been suggested to be effective. In this paper, an adaptive scheme for Schur product covariance localization is proposed, which is easy and efficient to implement in the ensemble data assimilation frameworks. A Gaussian-shaped taper function is selected as the localization taper function for the Schur product in the adaptive localization scheme, and the localization radius is obtained adaptively through a certain criterion of correlations with the background ensembles. An idealized Lorenz96 model with an ensemble Kalman filter is firstly examined, showing that the adaptive localization scheme helps to significantly reduce the spurious correlations in the small ensemble with low computational cost and provides accurate covariances that are similar to those derived from a much larger ensemble. The investigations of adaptive localization radius reveal that the optimal radius is model-parameter-dependent, verticallevel-dependent and nearly flow-dependent with weather scenarios in a realistic model; for example, the radius of model parameter zonal wind is generally larger than that of temperature. The adaptivity of the localization scheme is also illustrated in the ensemble framework and shows that the adaptive scheme has a positive effect on the assimilated analysis as the well-tuned localization.

Keywords: ensemble data assimilation; spurious correlations; adaptive covariance localization

1. Introduction

Background error covariances play an important role in determining the influence of assimilated observations on the multivariate analysis of model states [1]. However, the background error covariances are generally specified under the assumptions of homogeneity and isotropy at the beginning of the DA window in the four-dimensional variation data assimilation system (4D-Var), and the characteristics of background error introduced by 4D-Var are still limited by the linearity assumption and the usage of a static covariance model at the beginning of each 4D-Var cycle in the mainstream numerical weather prediction (NWP) centers [2–4].

Ensemble data assimilation has been widely employed for estimating the background error covariances in numerical oceanographic and meteorological dynamic systems [5–7]. However, the ensemble data assimilation scheme suffers from a few problems in which the covariance matrix is essentially replaced with the sample covariance matrix. For the practical implementation in oceanic or atmospheric systems, the ensemble size is typically many orders of magnitude smaller than the dimension of the model states, considering the affordable computational load in the operational system. This phenomenon means that the sample covariance matrix is rank-deficient due to the finite ensemble size, which can lead



Citation: Xing, X.; Liu, B.; Zhang, W.; Wu, J.; Cao, X.; Huang, Q. An Investigation of Adaptive Radius for the Covariance Localization in Ensemble Data Assimilation. *J. Mar. Sci. Eng.* 2021, *9*, 1156. https:// doi.org/10.3390/jmse9111156

Academic Editor: Terence O'Kane

Received: 27 September 2021 Accepted: 18 October 2021 Published: 21 October 2021 Corrected: 21 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). to the sampling error and the long-range spurious correlations in the sample covariance matrix [8–10].

The consequence of the spurious correlations is that a state variable may be incorrectly impacted by an observation that is physically remote, and filter divergence occurs when the analysis adjusted by observation information is unable to more accurately represent the true state. For this reason, the localization technique has been developed to overcome this problem [11–14]. The most common localization scheme is Schur product covariance localization, which attempts to mitigate spurious correlations by applying a weighting function to cut off the spurious correlations at a long distance. While covariance localization has been useful for ameliorating spurious correlations over long distances, they require exhaustive tuning of the localization radius for determining at which point the spurious correlations should be cut off [15,16]. More recently, to avoid the empirical tuning of the localization parameters, several schemes of adaptive localization have been proposed and validated for ensemble data assimilation depending on their specific implementations [17–22], e.g., using an ensemble of ensembles to mitigate spurious correlations such as sampling error correction and reducing correlation sampling for adaptive localization [23–26]; assuming the localization function can be taken as a power of the background error correlation function [27–29]; proposing the theory of optimal linear filtering for adaptive localization [30]; proposing a correlation-based adaptive localization method [31-33]; and considering machine learning algorithms to capture the adaptive localization parameters [34–36]. The adaptive localization technique has been developed in detail, but some of these methods make strong assumptions or do not always work in ensemble data assimilation (EDA)-like frameworks.

In this paper, an adaptive localization scheme is proposed for Schur product covariance localization. We focus on the implementation of the adaptive localization for the ensemblebased background error covariances with two difficulties: which taper function is used for determining the localization weighting coefficient and how the localization radius can be updated with the weather scenario.

The remainder of this paper is organized as follows: In Section 2, a brief introduction of the ensemble Kalman filter and traditional covariance localization approach is given. Section 3 presents the new adaptive localization scheme and its implementation. In Section 4, the adaptivity of this localization scheme is evaluated in the Lorenz96 model [37], and we also demonstrate the numerical results of the adaptive localization scheme in the Data Assimilation Research Testbed (DART) [38] and Weather Research and Forecasting (WRF) system [39]. Finally, conclusions and remaining issues are discussed in Section 5.

2. Background Material

In this section, we recall the basics of the ensemble Kalman filter (EnKF) and the localization technique. Table 1 describes some notations used in this paper.

Notation	Description
Ν	The ensemble size
m_{χ}	Model state dimension
m_{y}	Observation dimension
$\mathbb{E}[\cdot]$	Expectation
х	Model state variables
у	Observational data
r	Euclidean distance between two points
c	Cut-off radius

Table 1. Glossary of each notation.

2.1. Ensemble Kalman Filter

The EnKF is one of the variations developed from the Kalman filter equations by replacing the single forecast trajectory with an ensemble. The initial ensemble is a collection of perturbed estimates of the background state, and these ensemble members approximate a

random sample from the probability distribution of the background states and observations by a Monte Carlo procedure. There are two steps involved in using the Kalman filter to assimilate observations in the system: a forecast step and a Kalman-analysis step.

In the forecast step, the EnKF produces an ensemble of forecasts by propagating with the full nonlinear numerical model to the time when observations become available. This step can be performed over a long time period, and the background error covariance matrix could be flow-dependent. The approximation of the evolved background error covariance can be estimated in the ensemble data assimilation:

$$\mathbf{P}^{f} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\mathbf{x}^{fi} - \mathbb{E} \left[\mathbf{x}^{f} \right] \right) \left(\mathbf{x}^{fi} - \mathbb{E} \left[\mathbf{x}^{f} \right] \right)^{\mathrm{T}} = \frac{1}{N-1} \mathbf{S} \mathbf{S}^{\mathrm{T}}$$
(1)

In this expression, **S** indicates the matrix containing each forecast ensemble anomalies as a column, $\mathbf{S} = (\mathbf{x}^{f_1} - \mathbb{E}[\mathbf{x}^f], \mathbf{x}^{f_2} - \mathbb{E}[\mathbf{x}^f], \dots, \mathbf{x}^{f_N} - \mathbb{E}[\mathbf{x}^f])$; the superscript T stands for the matrix transpose, the superscript *f* represents the forecast.

In the Kalman-analysis step, a weighting value between the observation and background errors is calculated in the Kalman filter equations. This weighting determines the degree to which the forecast trajectory can be adjusted towards the observational data, and the update equations are commonly written as follows:

$$\mathbf{x}^{a} = \mathbf{x}^{f} + \mathbf{K} \Big[\mathbf{y} - \mathbf{H} \Big(\mathbf{x}^{f} \Big) \Big]$$
(2)

$$\mathbf{K} = \mathbf{P}^{f} \mathbf{H}^{\mathrm{T}} \left(\mathbf{H} \mathbf{P}^{f} \mathbf{H}^{\mathrm{T}} + \mathbf{R} \right)^{-1}, \tag{3}$$

where \mathbf{x}^{t} and \mathbf{x}^{a} denote the forecast and analysis variables, respectively; \mathbf{H} is the forward operator that maps the model state to the observational space ($m_{y} \times m_{x}$ matrix); \mathbf{R} is the observational error covariance ($m_{y} \times m_{y}$ matrix).

2.2. Traditional Covariance Localization

However, current NWP models have state spaces containing elements on the order of 10^7 , and a resource-limited ensemble size is far smaller than the dimension of the state spaces. The undersampling may produce a reduced rank representation of the background error covariance matrix and the development of spurious long-range correlations. Covariance localization is the most commonly method used to mitigate these issues by modifying the ensemble-based covariance (Equation (1)) with its element-wise product [40–42]:

$$\mathbf{P}^{loc} = \boldsymbol{\rho} \circ \mathbf{P}^{f},\tag{4}$$

where \circ indicates the Schur product. It is ordinarily achieved by a localization matrix ρ , whose structure is a band of nonzero elements around the main diagonal, with ones on the diagonal and falling to zero at the localization radius distance from the diagonal. The covariance localization attempts to cut-off the spurious correlations in the error covariances at a long distance, and effectively improves the estimates of the background error covariance. Thus, the update equations with covariance localization can be written as:

$$\mathbf{x}^{a} = \mathbf{x}^{f} + \left(\boldsymbol{\rho} \circ \mathbf{P}^{f}\right) \mathbf{H}^{\mathrm{T}} \left(\mathbf{H}\left(\boldsymbol{\rho} \circ \mathbf{P}^{f}\right) \mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \left[\mathbf{y} - \mathbf{H}\left(\mathbf{x}^{f}\right)\right]$$
(5)

For computational efficiency, the localization operations are implemented using the **H** in the observational space, $\rho \circ (\mathbf{P}^{f}\mathbf{H}^{T})$ and $\rho \circ (\mathbf{HP}^{f}\mathbf{H}^{T})$, respectively.

3. An Adaptive Localization Scheme

In this section, the adaptive localization scheme will be introduced in detail with the localization taper function and adaptive radius. The implementation of the adaptive scheme in the ensemble framework is also illustrated.

4 of 15

3.1. Localization Taper Function

The covariance localization scheme cuts off the long-range spurious correlations with its element-wise product like a weighting function. There are different ways of choosing the localization weighting function in localization schemes, such as taper functions [43], the fifth-order polynomial correlation function [44], and the distance-based localization functions for multiphase flows [45].

In the previous section, it was shown that the flow-dependent background error covariance estimated by the ensemble members is an $m_x \times m_x$ real symmetric matrix. For any nonzero column vector u of m_x real numbers, the scalar $u^{T}P^{f}u$ is strictly positive:

$$P^{f} = \mathbb{E}\left(\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)^{\mathrm{T}}\right)$$
(6)

$$u^{\mathrm{T}}P^{f}u = u^{\mathrm{T}}\mathrm{E}\left(\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)^{\mathrm{T}}\right)u$$
$$= \mathrm{E}\left(u^{\mathrm{T}}\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)^{\mathrm{T}}u\right)$$
$$= \mathrm{E}\left(\left\|u\left(\mathbf{x}^{fi} - \mathbb{E}\left[\mathbf{x}^{f}\right]\right)\right\|^{2}\right) \ge 0$$
(7)

Therefore, the sample covariance matrix estimated from the ensembles is a positive semi-definite matrix, and the localized covariance $\rho \circ \mathbf{P}^f$ is also positive semi-definite since it is the Schur product of two positive semi-definite matrices [46,47]. Another advantage of the Schur product is that it increases the rank of the covariance matrix by cutting off all long-range spurious covariances to zero, and the increasing rank would introduce the extra degrees of freedom to the dimension of the model states and the observations.

The Gaspari–Cohn function is very popularly used as localization taper function in ensemble frameworks because it is a compactly supported 5th-order piecewise rational function and its Gaussian-shaped is similar to the probability density function of a normal distribution [44,48]; thus the Gaspari–Cohn function is selected as the taper function for the adaptive localization scheme. The Gaspari–Cohn function has a value decreasing from one to zero in an approximation of a Gaussian distribution between zero and a $2 \times$ cutoff distance, shown in Figure 1, that will be transformed and constructed into a localization matrix with the same dimensions as the covariance matrix.

$$f = \begin{cases} -\frac{1}{4} \left(\frac{|r|}{c}\right)^5 + \frac{1}{2} \left(\frac{|r|}{c}\right)^4 + \frac{5}{8} \left(\frac{|r|}{c}\right)^3 - \frac{5}{3} \left(\frac{|r|}{c}\right)^2 + 1, & 0 \le |r| \le c \\ \frac{1}{12} \left(\frac{|r|}{c}\right)^5 - \frac{1}{2} \left(\frac{|r|}{c}\right)^4 + \frac{5}{8} \left(\frac{|r|}{c}\right)^3 + \frac{5}{3} \left(\frac{|r|}{c}\right)^2 - 5 \left(\frac{|r|}{c}\right) + 4 - \frac{2}{3} \left(\frac{c}{|r|}\right), & c \le |r| \le 2c \\ 0, & |r| \ge 2c \end{cases}$$
(8)



Figure 1. The Gaspari–Cohn function.

3.2. The Threshold Value of the Localization Radius

At present, defining the localization radius is a heuristic process (i.e., it is based on experience), which is computationally expensive and often needs to be redone whenever

the weather situation changes or the ensemble system is modified. If the threshold value is larger, then there is a larger localization radius such that the long-range spurious correlations will not be eliminated and more observational data are used for the model state update in Equation (5). Similarly, If the threshold value is small, then the physical dynamic correlations are excessively damped, and more observational data will be excluded. The root mean square error (RMSE) indicates how closely the estimation matches the reference field and is defined as

$$RMSE = \sqrt{\frac{\Sigma \left(\bar{x} - x_{ref}\right)^2}{N_{total}}}$$
(9)

where \bar{x} corresponds to the mean of the updated fields in the ensemble, x_{ref} represents the reference field, and N_{total} is the number of model parameters.

The Lorenz96 model is a dynamic system formulated by Edward Lorenz in 1996 [37]. It is defined as follows. For $i = 1, ..., m_x$,

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2}) \times x_{i-1} - x_i + F,$$
(10)

where it is assumed that $x_{-1} = x_{m_x-1}$, $x_0 = x_{m_x}$ and $x_{m_x+1} = x_1$. Here, x_i is the state of the system, $m_x = 40$, F is a forcing constant, and F = 8 is a common value known to cause chaotic behavior.



Figure 2. The RMSE of the ensemble mean analysis averaged over the points for different localization radii. Loc1, loc2, loc3, and loc4 represent 0.05, 2, 4, and 6 localization scales, respectively.

Figure 2 shows the RMSE corresponding to different radii in the Lorenz96 model with a 20-member ensemble. It can be noticed that different localization radii have different RMSE in the ensemble assimilation. A well-tuned radius will appropriately remove the unphysical spurious correlations and make the data match well, such as loc2 and loc3 in Figure 2. However, a larger radius, such as the label "loc4", will weaken the effect of removing the spurious correlation and have a poor data match with larger RMSE. And when the localization radius is adjusted to an excessively small value, the specified dynamic correlations will be deteriorated and a smaller radius will exclude many useful neighboring observations.

The tuned radius cannot reflect the real-time characteristics of the ensemble system. The adaptive radius needs to correspond adaptively to the correlation coefficient of the ensemble system since the localization technique is used for ameliorating the spurious correlations.

We define the x_i are the ensemble members, $\bar{x}(\bar{x} = \mathbb{E}[x])$ is the estimate of the mean μ_x . Since ensemble members x are the random sampling using the Monte Carlo method in the EnKF, independent and uncorrelated. According to the Strong Law of Large Numbers, when the sample is infinite $(N \to \infty)$ that we can get the estimate from the ensemble is exact:

$$\mathsf{P}\left\{\lim_{N\to\infty}\langle \bar{x}\rangle = \mu_x\right\} = 1\tag{11}$$

The estimate error is $\varepsilon = x - \mu_x$, we can get the estimation is unbiased with:

$$\mathbb{E}[\varepsilon] = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x) = \frac{1}{N-1} \left(\sum_{i=1}^{N} x_i - N \mu_x \right) = 0$$
(12)

The variance $\operatorname{Var}[x](\operatorname{Var}[x] = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2)$ is the estimate of the true variance σ_x^2 . According to the Central Limit Theorem, the estimate values are normally distributed around the true, and the samples are independent, thus:

٦

$$\operatorname{Var}\left[\frac{1}{N-1}\sum_{i=1}^{N}x_{i}\right] = \frac{1}{(N-1)^{2}}\sum_{i=1}^{N}\operatorname{Var}[x_{i}] = \frac{N}{(N-1)^{2}}\sigma_{x}^{2}$$
(13)

The standard error of the ensemble estimates approximates the square root of the ensemble size $\frac{1}{\sqrt{N-1}}(N \to \infty)$. The role of covariance localization is to cut-off spurious correlation between state variables that are spatially distant or not physically correlated. Since the estimation error of the ensemble members based on Monte Carlo sampling is $\frac{1}{\sqrt{N-1}}$, the spurious correlation in the correlation coefficient has seriously affected the accuracy of the estimated value when the correlation coefficient is less than $\frac{1}{\sqrt{N-1}}$. Therefore, it is considered that the corresponding distance at which the correlation coefficient is less than $\frac{1}{\sqrt{N-1}}$. Therefore, it is correlation can be cut-off with such radius. To avoid the effect of negative correlation coefficients, we use the square of the correlation coefficient for the determination of the localization radius $\mathbb{E}[C(r)^2] \geq \frac{1}{N-1}$.

Some studies have shown that there are strong links between localization and correlation, when the ensemble distribution is gaussian with the assumption of the sampling error is unbiased [49,50]:

$$\widetilde{\rho}_{ij} \approx \frac{(N-1)}{(N-2)(N+1)} \left((N-1) - \frac{1}{\mathbb{E}\left[\widetilde{C}_{ij}^2\right]} \right)$$
(14)

where $\tilde{\rho}_{ij}$ is the first-order approximation of the localization coefficient ρ_{ij} and C_{ij} is the element of sample correlation. The threshold value of the localization radius for the Gaspari-Cohn function should make $\rho \ge 0$ in the adaptive localization scheme, so that spurious correlations can be accurately eliminated. Thus, the Equation (14) can be written as $\mathbb{E}\left[\tilde{C}_{ij}^2\right] \ge \frac{1}{N-1}$.

The taper coefficient of the localization weighting function should be zero when $\mathbb{E}[C(r)^2]$ is less than $\frac{1}{N-1}$, in which the corresponding distance *r* is the localization radius threshold value. The localization radius can be computed adaptively with known statistical properties of sample covariances and updated with the real-time characteristics in the ensemble system. Figure 3a–d shows the square of the correlation coefficient in the Lorenz96 model with different ensemble members, and the dashed line corresponds to the position of the localization radius calculated by the adaptive localization scheme. It can be noticed that the spurious correlation weakens significantly with the increase of ensemble members (Figure 3e–h), and the adaptive radius reflects the range of spurious correlation well and gives a reasonable radius threshold. The larger ensemble has a larger radius than that of the small ensemble size, which can be explained that the accuracy of the estimated correlations of the small ensemble have been severely contaminated, but it is still accurate at the same distance in a larger ensemble.



Figure 3. The square of the correlation coefficient in the Lorenz96 model with (**a**) 50 members, (**b**) 100 members, (**c**) 200 members and (**d**) 400 members. (**e**–**h**) represent the spurious correlation squares, respectively.

3.3. Implementation

The adaptive covariance localization scheme can be implemented directly as an independent module in the ensemble framework, and the flowchart is shown in Figure 4. The practical implementation in the ensemble framework can be summarized by the following steps:

- 1. Integrate all ensemble members forward to the assimilation time step;
- 2. Estimate the sample covariance for the forecast ensemble members;
- 3. Calculate the correlation coefficient with the background ensembles;
- 4. Update the adaptive localization radius and transform the localization matrix;
- 5. Update the analysis field with the observation data, operator and error information;
- 6. Integrate the ensemble members forward to the next time step;
- 7. Repeat steps 2–6 until the assimilation process is completed.



Figure 4. The flowchart of the adaptive covariance localization in the ensemble framework.

4. Results and Analysis

In this section, experimental studies of the adaptive scheme application on ensemble framework are examined in Lorenz96 model and WRF model.

4.1. Preliminary Evaluation in Lorenz96 Model

The effects of the adaptive scheme on eliminating the spurious correlations is firstly examined with the Lorenz96 model in the EnKF. A perfect model scenario is simulated by a fourth-order Runge–Kutta scheme(RK4) for model integration with time step t = 1/40, which is used to guarantee stability with RK4. We perform a free run of the model from an initial condition, and this integration is considered to be the true state. The observations are created from the true state by adding uncorrelated random noise, and only one observation is available at two model grid points.



Figure 5. Representations of the covariance matrix estimated by (**a**) 20 members, (**b**) 100 members, (**c**) 20 members with localization. And (**d**) the localization matrix.

The representations' covariance matrices of the Lorenz96 model calculated from background ensembles using Equation (1) are shown in Figure 5. Each pixel in the grid represents the covariance between two state variables, and the pixels are colored according to the covariance value. Figure 5a shows the sample background error covariance estimated from the 20-member ensemble, and there are still large undesirable correlations in the covariance matrix far from the leading diagonal, and most of them are contributed by the spurious correlations. Accordingly, observations at one location may have an improper impact on state variables that are far away from the observation location, and the weighting placed on the forecasted state and the observations is inferior in the Kalman gain (Equation (3)), which will greatly damage the quality of the analysis.

Figure 5c shows the representation background error covariance matrix in Figure 5a after it has been localized. A striking feature is that the localization technique is like a weighting function to cut off long-range spurious correlations and the correlations close to the diagonal are maintained. The accuracy of the localized covariance matrix from the 20-member ensemble is almost the same as that from the 100-member ensemble (Figure 5b). Figure 6 shows the rank histogram of Lorenz96 model variables with different localization scheme 20 ensemble members. The statistical frequencies of each interval are more uniformly distributed with the adaptive localization scheme, reflecting a more reasonable distribution of ensemble members in the assimilation process, which can well suppress the filtering divergence.



Figure 6. Rank histogram with adaptive localization (line 1) and no localization (line 2), the three columns represent the 1st, 20th and 40th model variables, respectively.

The signal-to-noise ratios (SNR) of the raw and localized covariance matrix are 1.282 and 6.102, respectively, and a larger SNR of localized covariance strongly supports the fact that the adaptive localization effectively ameliorates most of the spurious correlations. Another benefit of the localization technique is that the covariance matrix becomes sparse after applying the Schur product (Equation (4)), which can lead to important computational savings in the calculations.



Table 2. Adaptive localization radii during the last 10 steps of the assimilation cycle.

Figure 7. RMSE and spread with the (a) adaptive scheme and (b) fixed scheme tuned manually.

The spread is a direct measurement that qualifies the uncertainty contained in the ensemble members. If the spread of the ensemble members is small, then all the members are in a similar state, and there is a high confidence in the background. In contrast, if the spread of the ensemble members is large, then the ensemble members are widely spaced. After the assimilation of observational data, the RMSE and spread of the ensemble members are reduced, and the spread should ideally match the RMSE if the ensembles estimate the uncertainty of the model states correctly.

Figure 7 shows the RMSE and spread of the analysis field with adaptive scheme and the localization tuned manually. The RMSE and spreads of the well-tuned localization scheme and the adaptive scheme descended stably and continuously. At the stable stage, the RMSE and spread of the two experiments are close to each other. The RMSE and the spread of the adaptive scheme reach 0.3604 and 0.4267, respectively, which implies that the ensembles work well on data assimilation. The localization radii during the last 10 steps of the assimilation cycle of the adaptive scheme are represented in Table 2. The localization radius is not constant in the cycle, and the adaptive scheme works well in updating the radius values. These results indicate that the adaptive localization radius is efficient in

suppressing the inferior effect of spurious correlations and has positive effects in the data assimilation.

4.2. Application in an Atmospheric Model

The performance of the adaptive localization scheme will be investigated in the Weather Research and Forecasting (WRF) model [39] and the Data Assimilation Research Testbed (DART) framework [38]. We select Typhoon Dujuan (2015) as a numerical example, which was generated in the Northwest Pacific on 20 September 2015 and weakened to a cyclone on 30 September. The center's maximum wind speed was about 79 m/s, and its pressure was about 925 hPa. The center of the simulation area is 23.0° N, 130.0° E at 1800 UTC 26 September 2015, which is shown in Figure 8; the grid size of the assimilation region is 500×400 ; the horizontal resolution is 10 km. The initial and lateral boundary conditions for the WRF simulation were derived from the National Centers for Environmental Prediction (NCEP) operational Global Forecast System analysis data. A pair of 40-member ensembles is generated by the random perturbations based on the initial conditions. For a more realistic environment, the observations (PREPBUFR files) from the NCEP are assimilated using the ensemble adjustment Kalman filter in DART [51].



Figure 8. Weather situation on 1800 UTC 26 September 2015.

The expectations of correlation coefficients at different distances are calculated from the ensemble states by random sampling in the simulation area, which is shown in Figure 9 that the point pair connected by the lime solid line is used to calculate the correlation coefficients at distances of 200 km. The adaptive scheme is easy to implement online in the frameworks of the traditional covariance localization and has a low computational cost that the realizations of random sampling for expectations is roughly in the order of 10².



Figure 9. Shows the random sampling for calculating the correlation coefficients.

Figure 10a shows the expectations of the correlation coefficient calculated from random sampling. It is obvious that the characteristics of the correlation coefficient vary with the model parameter, which is illustrated by the temperature and zonal wind correlation

11 of 15

at the surface. The correlation coefficient gradually decreases with increasing distance, drops from one to approximately zero at first within a certain distance, and then starts to oscillate around zero with increasing distance. In addition, the correlation coefficients of different model parameters have different trends; for example, the correlation coefficient of temperature gradually decreases from one to approximately zero, while the correlation coefficient of coefficient of zonal wind first decreases below zero, then increases to approximately zero, and then oscillates around zero.

The localization radii of zonal wind and temperature on the vertical levels calculated by the adaptive scheme are illustrated in Figure 10b. There is a striking feature that the optimal localization radius is variable on the model vertical levels; for instance, the localization radius of temperature around the surface is 550 km, which is smaller than 930 km at 500 hPa. The trend of the radius varying with model levels is different with model parameters where the radius of zonal wind is generally larger than that of temperature at the same levels. The optimal radius of the localization function is not fixed for the model parameters, which is flow-dependent with the background ensemble weather scenario. A uniform localization radius tuned manually on the observational data of different variables may reduce the effect of localization on the spurious correlations issue and deteriorate the balance relationship between the model parameters [52].



Figure 10. (a) Temperature (solid) and zonal wind (dashed) correlations at the surface, (b) the corresponding localization radius on the model vertical levels.

Single observational data simulations with the adaptive scheme are implemented in the DART-WRF system. The spatiotemporal elution of the observational data increment of different model parameters and its link with the meteorological flow are investigated through Figure 11, from 18:00 UTC on 26 September to 00:00 UTC on 27 September 2015. An overview of Figure 11 outlines the general features of the adaptive localization scheme, where the adaptive localization based on the correlations is changed in accordance with the meteorological flow in the background without tuning the localization radius manually. The temporal evolution of the adaptive localization function is examined in Figure 12a for temperatures in which the localization radius at 00:00 UTC is shorter than that at 18:00 UTC.



Figure 11. The wind pressure field and increments of signal observations for zonal wind (**a**) and temperature (**c**) near the surface at 18:00 UTC on 26 September 2015 (left column). The right column show the same at 00:00 UTC on 27 September.



Figure 12. The adaptive localization function of temperature at different times, and of temperature and zonal wind at 18:00 UTC on 26 September 2015.

It can be observed that the range of the zonal wind increments is greater than the range of temperature at the same time (Figure 11) and the different increments of the different model parameters indicate that the optimal localization radius vary across the parameters, which is in agreement with the localization radius shown in Figure 12b that the localization radius of zonal wind obtained by the adaptive method is larger than that of temperature. A more detailed examination shows that the observation increment of temperature has almost no negative increment and zonal wind has a large amount of negative increment information, which corresponds to the correlations in Figure 10a that the correlation coefficient of temperature drops directly from one to near zero, while the correlation coefficient of zonal wind drops from one to below zero and then oscillates near zero. The characteristics of different variable observations with different radii are reflected reasonably by the adaptive localization scheme.

5. Discussion and Conclusions

The ensemble-based covariances estimated from the short-term forecast ensembles are a reduced rank representation of the background error covariance and always suffer

from spurious correlation problems. In order to ameliorate the spurious correlations in the ensemble-based covariances, an adaptive localization scheme is presented in this study.

The theoretical part of the ensemble-based covariance shows that the estimated matrix from the ensembles is a positive semi-definite matrix. In addition, this supports the claim that the positive semi-definite feature of the covariance matrix should be maintained after covariance localization. For that purpose, a Gaussian-shaped taper function is used for determining the localization weighting coefficient. The investigation of the threshold value of the localization radius reveals that the adaptive radius has a strong quantitative relationship with the correlation and ensemble size. The spurious correlations generated by the finite ensembles have the same scale as the true correlation coefficient at a long distance and become weaker in the larger ensemble. Note that the accuracy of the estimated correlations of the small ensemble have been severely contaminated, but it is still accurate at the same distance in a larger ensemble, which means that the localization radius of the large ensemble size is larger than that of the small ensemble size. In a real NWP context, the localization radius obtained by the background ensemble is model parameter- and vertical level-dependent, which is reflected in the variation of observational increments. The interesting properties are also closely consistent with the change characteristics of the corresponding correlation coefficient.

Experimental studies of the adaptive scheme application on the ensemble framework present that the adaptive scheme has a positive impact on the RMSE in the analysis compared with the well-tuned localization scheme. The spurious correlations in the ensemble-based covariance matrix have been significantly suppressed, and the assimilation effect is improved with the localization radius is updated adaptively. In addition, it is easy to implement the adaptive scheme in the traditional covariance localization procedure when running an ensemble assimilations.

However, the homogeneous and uniform type of the localization taper function may weaken the anisotropic characteristics of the increment. In the next step, combining the adaptive localization and a wavelet approach may be worth considering. Second, further adaptive applications are currently being explored for the multiple domain according to the weather situation.

Author Contributions: Conceptualization, X.X. and B.L.; methodology, X.X., B.L. and W.Z.; formal analysis and investigation, J.W., X.C. and Q.H.; supervision, B.L. and W.Z.; writing original draft, X.X.; writing review and editing, B.L., W.Z., J.W., X.C. and Q.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant Nos. 42005003), the National Key Research and Development Program of China (Grant Nos. 2018YFC1506704) and the National Natural Science Foundation of China (Grant Nos. 41675097, 41375113, and 42005120).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are the short-term forecast by a global NWP system at the National University of Defense Technology that are not publicly available.

Conflicts of Interest: The authors declare that there is no conflict of interest.

References

- 1. Kalnay, E. Atmospheric Modeling, Data Assimilation and Predictability; Cambridge University Press: Cambridge, UK, 2002.
- 2. Lewis, J.M.; Derber, J.C. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A* **1985**, *37*, 309–322. [CrossRef]
- 3. Dimet, F.-X.L.; Talagrand, O. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus A* **1986**, *38*, 97–110. [CrossRef]
- 4. Rabier, F.; Jarvinen, H.; Klinker, E. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **2000**, *126*, 1143–1170. [CrossRef]

- Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 1994, 99, 10143–10162. [CrossRef]
- 6. Buehner, M. Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 1013–1043. [CrossRef]
- Houtekamer, P.L.; Mitchell, H.L. Data Assimilation Using an Ensemble Kalman Filter Technique. Mon. Weather Rev. 1998, 126, 796–811. [CrossRef]
- 8. Hamill, T.M.; Whitaker, J.S.; Snyder, C. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Weather Rev.* 2001, 129, 2776–2790. [CrossRef]
- 9. Lorenc, A.C. The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Q. J. R. Meteorol. Soc.* 2003, 129, 3183–3203. [CrossRef]
- 10. Houtekamer, P.L.; Zhang, F. Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Weather Rev.* 2016, 144, 4489–4532. [CrossRef]
- 11. Petrie, R. Localization in the Ensemble Kalman Filter. Master's Thesis, Atmosphere, Ocean and Climate University of Reading, London, UK, 2008.
- 12. Bergemann, K.; Reich, S. A localization technique for ensemble Kalman filters. Q. J. R. Meteorol. Soc. 2010, 136, 701–707. [CrossRef]
- 13. Zhu, J.; Zheng, F.; Li, X. A new localization implementation scheme for ensemble data assimilation of non-local observations. *Tellus A* **2011**, *63*, 244–255. [CrossRef]
- 14. Farchi, A.; Bocquet, M. On the efficiency of covariance localisation of the ensemble Kalman filter using augmented ensembles. *Front. Appl. Math. Stat.* **2019**, *5*, 3–5. [CrossRef]
- 15. Kirchgessner, P.; Nerger, L.; Bunse-Gerstner, A. On the choice of an optimal localization radius in ensemble Kalman filter methods. *Mon. Weather Rev.* 2014, 142, 2165–2175. [CrossRef]
- 16. Flowerdew, J. Towards a theory of optimal localisation. Tellus A 2015, 67, 25257. [CrossRef]
- 17. Devegowda, D.; Arroyo-Negrete, E.; Datta-Gupta, A. Flow relevant covariance localization during dynamic data assimilation using EnKF. *Adv. Water Resour.* **2010**, *33*, 129–145. [CrossRef]
- 18. Perianez, A.; Reich, H.; Potthast, R. Optimal localization for ensemble Kalman filter systems. J. Meteorol. Soc. Jpn. Ser. II 2014, 92, 585–597.
- 19. Wu, X.; Li, W.; Han, G.; Zhang, L.; Shao, C.; Sun, C.; Xuan, L. An adaptive compensatory approach of the fixed localization in the EnKF. *Mon. Weather Rev.* **2015**, *143*, 4714–4735. [CrossRef]
- 20. Zhang, H.-Q.; Tian, X.-J.; Zhang, C.-M. An Economical Approach to Flow-Adaptive Moderation of Spurious Ensemble Correlations and Its Application in the Proper Orthogonal Decomposition-Based Ensemble Four Dimensional Variational Assimilation Method. *Atmos. Ocean. Sci. Lett.* **2015**, *8*, 320–325.
- 21. Lei, L.; Anderson, J.L.; Whitaker, J.S. Localizing the impact of satellite radiance observations using a global group ensemble filter. *J. Adv. Model. Earth Syst.* **2016**, *8*, 719–734. [CrossRef]
- 22. Popov, A.A.; Sandu, A. A Bayesian approach to multivariate adaptive localization in ensemble-based data assimilation with time-dependent extensions. *Nonlinear Process. Geophys.* **2019**, *26*, 109–122. [CrossRef]
- 23. Anderson, J.L. Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Phys. D Nonlinear Phenom.* **2007**, 230, 99–111. [CrossRef]
- 24. Anderson, J.L. Localization and Sampling Error Correction in Ensemble Kalman Filter Data Assimilation. *Mon. Weather Rev.* 2012, 140, 2359–2371. [CrossRef]
- 25. Anderson, J.L. Reducing correlation sampling error in ensemble Kalman filter data assimilation. *Mon. Weather Rev.* 2016, 144, 913–925. [CrossRef]
- Ying, Y.; Zhang, F.; Anderson, J.L. On the selection of localization radius in ensemble filtering for multiscale quasigeostrophic dynamics. *Mon. Weather Rev.* 2018, 146, 543–560. [CrossRef]
- 27. Bishop, C.H.; Hodyss, D. Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Q. J. R. Meteorol. Soc.* 2007, 133, 2029–2044. [CrossRef]
- 28. Bishop, C.H.; Hodyss, D. Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus A* 2009, *61*, 84–96. [CrossRef]
- 29. Bishop, C.H.; Hodyss, D. Ensemble covariances adaptively localized with ECO-RAP. Part 2: A strategy for the atmosphere. *Tellus A* **2009**, *61*, 97–111. [CrossRef]
- Ménétrier, B.; Montmerle, T.; Michel, Y.; Berre, L. Linear filtering of sample covariances for ensemble-based data assimilation. Part I: Optimality criteria and application to variance filtering and covariance localization. *Mon. Weather Rev.* 2015, 143, 1622–1643. [CrossRef]
- Luo, X.; Bhakta, T.; Naevdal, G. Correlation-based adaptive localization with applications to ensemble-based 4D-seismic history matching. SPE J. 2018, 23, 396–427. [CrossRef]
- 32. Luo, X.; Lorentzen, R.J.; Valestrand, R.; Evensen, G. Correlation-based adaptive localization for ensemble-based history matching: Applied to the Norne field case study. *SPE Reserv. Eval. Eng.* **2019**, 22, 1084–1109. [CrossRef]
- 33. Luo, X.; Bhakta, T. Automatic and adaptive localization for ensemble-based history matching. J. Pet. Sci. Eng. 2020, 184, 106559. [CrossRef]
- 34. Moosavi, A.; Attia, A.; Sandu, A. A machine learning approach to adaptive covariance localization. arXiv 2018, arXiv:1801.00548.

- 35. Attia, A.; Constantinescu, E. An optimal experimental design framework for adaptive inflation and covariance localization for ensemble filters. *arXiv* 2018, arXiv:1806.10655.
- 36. Moosavi, A.; Attia, A.; Sandu, A. Tuning Covariance Localization Using Machine Learning. *Int. Conf. Comput. Sci.* 2019, 3, 199–212.
- 37. Lorenz, E.N. Predictability: A problem partly solved. Proc. Semin. Predict. 1996, 1, 1.
- Anderson, J.; Hoar, T.; Raeder, K.; Liu, H.; Collins, N.; Torn, R.; Avellano, A. The data assimilation research testbed: A community facility. *Bull. Am. Meteorol. Soc.* 2009, 90, 1283–1296. [CrossRef]
- 39. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Wang, W.; Powers, J.G. A description of the Advanced Research WRF version 3. NCAR Technical note-475+ STR. *Atmos. Res.* **2008**. [CrossRef]
- 40. Chen, Y.; Oliver, D.S. Localization and regularization for iterative ensemble smoothers. Comput. Geosci. 2017, 21, 13–30.
- 41. Oke, P.R.; Sakov, P.; Corney, S.P. Impacts of localisation in the EnKF and EnOI: Experiments with a small model. *Ocean Dyn.* **2007**, 57, 32–45. [CrossRef]
- 42. Ehrendorfer, M. A review of issues in ensemble-based Kalman filtering. Meteorol. Z. 2007, 3, 795–818. [CrossRef]
- 43. Furrer, R.; Bengtsson, T. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivar. Anal.* 2007, *98*, 227–255. [CrossRef]
- 44. Gaspari, G.; Cohn, S.E. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **1999**, *125*, 723–757. [CrossRef]
- 45. Chen, Y.; Oliver, D.S. Cross-covariances and localization for EnKF in multiphase flow data assimilation. *Comput. Geosci.* 2010, 14, 579–601. [CrossRef]
- 46. Janjic, T.; Nerger, L.; Albertella, A.; Schröter, J.; Skachko, S. On domain localization in ensemble-based Kalman filter algorithms. *Mon. Weather Rev.* 2011, 139, 2046–2060. [CrossRef]
- 47. Horn, R. The hadamard product. Proc. Symp. Appl. Math. 1990, 40, 87–169.
- 48. Houtekamer, P.L.; Mitchell, H.L. Ensemble kalman filtering. Q. J. R. Meteorol. Soc. 2005, 131, 3269–3289. [CrossRef]
- 49. Ménétrier, B.; Montmerle, T.; Michel, Y.; Berre, L. Linear Filtering of Sample Covariances for Ensemble-Based Data Assimilation. Part II: Application to a Convective-Scale NWP Model. *Mon. Weather Rev.* **2015**, *143*, 1644–1664.
- 50. Sample Covariance Filtering. Available online: https://github.com/benjaminmenetrier/covariance_filtering/blob/master/ covariance_filtering.pdf (accessed on 10 September 2020). [CrossRef]
- 51. Anderson, J.L. An ensemble adjustment Kalman filter for data assimilation. Mon. Weather Rev. 2001, 129, 2884–2903. [CrossRef]
- 52. Kepert, J.D. Covariance localisation and balance in an ensemble Kalman filter. *Q. J. R. Meteorol. Soc.* **2009**, *135*, 1157–1176. [CrossRef]