

EAKF-Based Parameter Optimization Using a Hybrid Adaptive Method

LIGE CAO,^a XINRONG WU,^b GUIJUN HAN,^a WEI LI,^a XIAOBO WU,^a HAOWEN WU,^a CHAOLIANG LI,^a
YUNDONG LI,^a AND GONGFU ZHOU^a

^a School of Marine Science and Technology, Tianjin University, Tianjin, China

^b Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, Ministry of Natural Resources, Tianjin, China

(Manuscript received 7 April 2022, in final form 2 August 2022)

ABSTRACT: To effectively reduce model bias and improve assimilation quality, we adopt a hybrid adaptive approach of ensemble adjustment Kalman filter (EAKF) and multigrid analysis (MGA), called EAKF-MGA, to implement parameter optimization as follows. For each assimilation cycle, observations are used to adjust the prior ensembles of both state variables and parameters using the EAKF without inflation. Then, the MGA is adaptively triggered to extract multiscale information from the observational residual to innovate the ensemble mean of the state once again. Results of biased twin experiments consisting of a barotropic spectral model and idealized observation systems show that the proposed EAKF-MGA is insensitive to state variance inflation and localization during the parameter optimization process, compared with the EAKF with adaptive inflation. We also find that computational efficiency is another important advantage of the EAKF-MGA for both state estimation and parameter estimation since extremely small ensemble size is allowed, while the EAKF with adaptive inflation does not work anymore. In essence, the EAKF-MGA is designed to estimate and correct systematic errors jointly with model's state variables. Through alleviating biases, including the model bias caused by the biased parameter and the analysis bias resulting from the sampling noise given the limited ensemble size, it can be guaranteed that the analysis in the EAKF-MGA will be proceeded onward with the standard assumption of the unbiased model background field in modern data assimilation theory to be met.

KEYWORDS: Data assimilation; Model errors; Optimization

1. Introduction

The so-called model bias occurs when systematic errors are introduced by model deficiencies including, as pointed out by Danforth et al. (2007), approximations in numerical differentiation and integration, inaccurate forcing, and parameterizations used to represent subgrid-scale physical processes. Parameter values involved in the above practices yet to be determined empirically, however, can be estimated or optimized in tandem with the model state through assimilation of observations to reduce bias as much as possible in order to improve the model. This frees modelers, to certain extent, from a subjective parameter tuning process through a trial-and-error approach, which is computationally expensive and labor intensive. Such parameter estimation or parameter optimization (PO) can be realized by ensemble filtering and adjoint approaches (Zhang et al. 2012; Han et al. 2015; Zhang et al. 2020). With the former approach, taking the ensemble Kalman filter (EnKF) for example, which is able to effectively simulate the probability distribution of temporally varying background covariance, parameters can be directly estimated based on the error covariance between the observed model state and parameter, which is easy to implement (e.g., Evensen 1994; Zhang et al. 2005; Anderson 2007a; Evensen 2007; Han et al. 2013; Ruiz et al. 2013; Houtekamer and Zhang 2016; Zupanski 2016; Zhang et al. 2020). Based on the EnKF, Zhang et al. (2012) proposed a data assimilation scheme with enhance parameter correction (DAEPC), which performs state and parameter

estimation after the former reaches a “quasi-equilibrium” where the uncertainty of the model state is sufficiently constrained by observations. The DAEPC has been applied successfully in many studies related to PO (e.g., Zhang 2011a,b; Wu et al. 2012; Han et al. 2014; Zhang et al. 2015; Wu et al. 2016; Zhao et al. 2019).

However, the EnKF itself is sensitive to covariance inflation and localization (Anderson and Anderson 1999; Mitchell et al. 2002; Anderson 2007a; Bishop and Hodyss 2007, 2011; Wu et al. 2014, 2015; Flowerdew 2015). To tackle these challenges effectively, there are a number of widely used adaptive methods in the literature. In regard to covariance inflation, Anderson (2007a; hereafter A07) presented a temporally adaptive algorithm that uses observations to automatically calculate appropriate values for covariance inflation, which detects and compensates for model error significantly. Li et al. (2009) proposed to estimate the inflation factor and observational errors simultaneously within the EnKF, and examined the feasibility of the simultaneous method in the Lorenz-96 model and in the simplified parameterization, primitive equation dynamics (SPEEDY) model. Miyoshi (2011) used an advanced Gaussian approach to adaptive covariance inflation, and successfully applied it in a low-resolution atmospheric general circulation model (AGCM). Gharamti (2018) discussed an enhanced adaptive inflation algorithm that assumes an inverse gamma distribution for the prior inflation factor, which is capable of generating accurate and consistent state estimates. Regarding covariance localization, Anderson (2007b) utilized the hierarchical Monte Carlo method to limit the impacts of ensemble sampling error and spurious correlations by adaptively computing the regression confidence factor. Bishop and Hodyss (2009a,b) introduced an adaptive localization approach, in which the ensemble information is

Corresponding authors: Xinrong Wu, xrw_nmdis@163.com; Guijun Han, guijun_han@tju.edu.cn

used to apply flow-dependent localization functions by raising smoothed versions of ensemble correlations raised to a power (ECO-RAP). Anderson and Lei (2013) developed an empirical localization algorithm that computes localization from the output of an ensemble observing system simulation experiment (OSSE); Lei and Anderson (2014) successfully applied it in an AGCM. Flowerdew (2015) examined the optimal localization method and a new form of hybrid data assimilation in which localization damps the sample correlation toward its climatological mean rather than zero. Moosavi et al. (2019) successfully used machine learning algorithms to automatically tune the impact radius of the covariance localization; but the computational consumption of training the learning model is relatively high.

Is there an adaptive method that can concurrently address the issues that arise in the context of both covariance inflation and localization? The answer is yes. As Wu et al. (2015, hereafter W15) demonstrated in their adaptively compensatory method for the ensemble adjustment Kalman filter (EAKF) of Anderson (2001), at each analysis cycle, after the EAKF without inflation and with a prescribed impact radius for covariance localization is accomplished, a compensatory step is adaptively triggered to adjust the ensemble mean by employing the multi-grid analysis (MGA) to extract multiscale information from observational residual. It was found that such hybrid adaptive method is insensitive to the impact radius of localization, and has a superior performance over the EAKF with optimal inflation. This was confirmed by applying the hybrid adaptive method to an El Niño–Southern Oscillation (ENSO) prediction experiment using an intermediate coupled model (Wu 2016). Given its capacity to deal with problems related to both covariance inflation and localization, the efforts of the present paper, as a follow-up study of W15, focus on the application of the hybrid adaptive method, or EAKF-MGA for short, in PO experiment in order to reduce model bias. Its performance will be investigated within a biased twin assimilation experiment framework consisting of a barotropic spectral model.

The remainder of this paper is structured as follows. In section 2, we briefly describe the barotropic spectral model, and the algorithms for the implementation of PO using the hybrid adaptive method. In section 3, we introduce sensitivity tests and biased twin experiment configuration. Analysis of results is presented in section 4. Summary and discussion are given in section 5.

2. Methodology

a. The barotropic spectral model

We employ a global barotropic spectral model based on the equation of potential vorticity conservation (Haltiner and Williams 1980) to conduct experiments of state estimation only (SEO) and PO in this study:

$$\frac{d}{dt} \left(\frac{f + \zeta}{H} \right) = 0, \quad (1)$$

where ζ and f represent relative vorticity and planetary vorticity (i.e., Coriolis parameter), respectively; and H is the average atmospheric “equivalent depth.”

Equation (1) can be rewritten in terms of geostrophic streamfunction ψ :

$$\frac{\partial}{\partial t} (\nabla^2 - \lambda^2) \psi + J(\psi, \nabla^2 \psi) + \beta \frac{\partial \psi}{\partial x} + J(\psi, h') = 0, \quad (2)$$

where $h' = f_0 h_{\text{terrain}}/H$ represents the effect of topography; $\beta = df/dy$; y represents the northward meridional distance from the equator; $J(\cdot)$ denotes the Jacobian operator; and $\lambda^2 = f_0^2/(gH)$ is the Cressman parameter, which is the inverse square of the Rossby deformation radius.

The state variable in the context of data assimilation is the atmospheric streamfunction at 64 (longitude) \times 54 (latitude) Gaussian grid points. The leapfrog time step is used to integrate the model; and the Robert–Asselin time filter (Robert 1969; Asselin 1972) with the dimensionless filter coefficient γ is applied to damp spurious computational modes. The integration time step is 30 min.

b. The EAKF method

A deterministic EnKF, namely, the EAKF, is used in this paper. The EAKF can sequentially assimilate observations one by one under the assumption that observational errors are uncorrelated. The implementation of the EAKF for a single observation y^o can be summarized by the following two steps.

First, we compute the observational increment as follows:

$$\Delta y_i = \left(\sqrt{\frac{r^2}{r^2 + (\sigma_y^p)^2}} - 1 \right) (y_i^p - \bar{y}^p) + \frac{(\sigma_y^p)^2}{r^2 + (\sigma_y^p)^2} (y^o - \bar{y}^p), \quad (3)$$

where i indexes the ensemble member; y_i^p denotes the i th prior ensemble member of y^o , which is usually computed by interpolating the i th prior ensemble member of a state variable to the location of y^o ; \bar{y}^p and σ_y^p are computed from y_i^p , representing the prior ensemble mean and standard deviation of y_i^p , respectively; and r denotes the standard deviation of observational errors.

Second, we project the observational increment to model grids:

$$\Delta x_{i,j} = \rho_{j,y} \frac{\text{cov}_{j,y}^p}{(\sigma_y^p)^2} \Delta y_i, \quad (4)$$

where Δx represents the state increment, and j indexes the state variable. The $\text{cov}_{j,y}^p$ term is the prior error covariance between x_j and y^o , formulated as follows:

$$\text{cov}_{j,y}^p = \frac{\sum_{i=1}^N (x_{i,j} - \bar{x}_j)(y_i^p - \bar{y}^p)}{N - 1}, \quad (5)$$

where N is the ensemble size; $x_{i,j}$ and \bar{x}_j indicate the i th ensemble member and the ensemble mean of x_j , respectively.

In Eq. (4), $\rho_{j,y}$ is the covariance localization factor between x_j and y^o . The $\rho_{j,y}$ takes a widely used distance-based Gaspari–Cohn (GC) function (Gaspari and Cohn 1999) in this study, that is,

$$\rho_{j,y} = \Omega(a, b) = \begin{cases} -\frac{1}{4}\left(\frac{b}{a}\right)^5 + \frac{1}{2}\left(\frac{b}{a}\right)^4 + \frac{5}{8}\left(\frac{b}{a}\right)^3 - \frac{5}{3}\left(\frac{b}{a}\right)^2 + 1, & 0 \leq b \leq a; \\ \frac{1}{12}\left(\frac{b}{a}\right)^5 - \frac{1}{2}\left(\frac{b}{a}\right)^4 + \frac{5}{8}\left(\frac{b}{a}\right)^3 + \frac{5}{3}\left(\frac{b}{a}\right)^2 - 5\left(\frac{b}{a}\right) + 4 - \frac{2}{3}\left(\frac{b}{a}\right)^{-1}, & a < b \leq 2a; \\ 0, & b > 2a. \end{cases} \quad (6)$$

In Eq. (6), b denotes the distance between x_j and y^o , and a is the half-width of the GC function, which represents the impact radius of the observation involved in the localization of the EAKF in this study.

c. The EAKF-MGA method

As described in W15, implementation of the EAKF-MGA method at each analysis cycle for the state variables can be simply expressed in the following steps.

First, we use the EAKF without inflation to assimilate all available observations with a fixed impact radius, and obtain the posterior ensemble mean $\bar{\mathbf{x}}_{\text{EAKF}}$ and deviations from the ensemble mean.

Second, we compute the observational residual \mathbf{y}^{res} , i.e., the difference between the observation vector and its model counterpart projected to the observation space as follows:

$$\mathbf{y}^{\text{res}} = \mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}_{\text{EAKF}}, \quad (7)$$

where \mathbf{y}^o and \mathbf{H} represent the observation vector and the bilinear interpolation operator from the state space to the observation space, respectively.

Third, we compute the root-mean-square error (RMSE) between \mathbf{y}^o and $\mathbf{H}\bar{\mathbf{x}}_{\text{EAKF}}$ as follows:

$$\text{RMSE}^{\text{res}} = \sqrt{\frac{1}{K} \sum_{k=1}^K (y_k^o - \mathbf{H}_k \bar{\mathbf{x}}_{\text{EAKF}})^2}, \quad (8)$$

where K indicates the number of observations; y_k^o and \mathbf{H}_k indicate the k th observation and the bilinear projection operator from the model grids to the observation location corresponding to y_k^o , respectively.

Fourth, we compute the critical value of RMSE^{res} , i.e., $\theta = r\sqrt{\chi_{1-\alpha}^2(K)/K}$, where $\chi_{1-\alpha}^2(K)$ represents the $(1 - \alpha)$ upper fractile of the χ^2 distribution with K degrees of freedom, with α being the significance level. If RMSE^{res} is greater than θ , the MGA will be triggered to extract multiscale information from \mathbf{y}^{res} through refining the analysis grid. The cost function for the l th analysis grid is formulated as

$$J(\delta\mathbf{x}^l) = \frac{1}{2}(\mathbf{H}^l \delta\mathbf{x}^l - \mathbf{d}^l)^T (\mathbf{R}^l)^{-1} (\mathbf{H}^l \delta\mathbf{x}^l - \mathbf{d}^l) + \frac{1}{2}(\delta\mathbf{x}^l)^T (\mathbf{S}^l)^{-1} \delta\mathbf{x}^l, \quad l = 1, \dots, L_{\text{MGA}}, \quad (9)$$

where L_{MGA} is the number of levels of the analysis grid, and \mathbf{H}^l is the bilinear projection operator from the l th grid to the observation space. The resolution of the l th grid is set to $(2^{l-1} + 1) \times (2^{l-1} + 1)$, which is nearly twice that of

the $(l - 1)$ th grid. The expressions $\delta\mathbf{x}^l$, \mathbf{d}^l , \mathbf{R}^l , and \mathbf{S}^l are, respectively, the increment of state vector, the observational innovation vector, the observational error covariance matrix, and the smoothing matrix for the l th grid. For each grid level, \mathbf{d}^l is calculated as follows:

$$\mathbf{d}^l = \begin{cases} \mathbf{y}^{\text{res}}, & l = 1 \\ \mathbf{d}^{l-1} - \mathbf{H}^{l-1}(\delta\mathbf{x}^{l-1})^a, & l = 2, \dots, L_{\text{MGA}}, \end{cases} \quad (10)$$

where $(\delta\mathbf{x}^{l-1})^a$ represents the analysis result of $\delta\mathbf{x}^{l-1}$. The matrix \mathbf{R}^l is set to the identity matrix. The smoothing term in Eq. (9) can be expressed as follows:

$$\begin{aligned} \frac{1}{2}(\delta\mathbf{x}^l)^T (\mathbf{S}^l)^{-1} \delta\mathbf{x}^l &= \frac{1}{2} \sum_{i=2}^{2^{l-1}-1} \sum_{j=1}^{2^{l-1}+1} (\delta x'_{i-1j} - 2\delta x'_{ij} + \delta x'_{i+1j})^2 \\ &+ \frac{1}{2} \sum_{i=1}^{2^{l-1}-1} \sum_{j=2}^{1+2^{l-1}} (\delta x'_{ij-1} - 2\delta x'_{ij} + \delta x'_{i+1j})^2, \end{aligned} \quad (11)$$

where $\delta x'_{ij}$ represents the (i, j) th element of $\delta\mathbf{x}^l$.

With Eq. (9) and the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm (Liu and Nocedal 1989), the numerical approximation of $(\delta\mathbf{x}^l)^a$ can be obtained. For each grid level, $(\delta\mathbf{x}^l)^a$ is linearly interpolated to the L_{MGA} th grid with the operator \mathbf{F}^l . Then, the analysis result of the MGA is calculated by

$$\delta\mathbf{x}_{\text{MGA}} = \mathbf{E} \sum_{l=1}^{L_{\text{MGA}}} \mathbf{F}^l (\delta\mathbf{x}^l)^a, \quad (12)$$

where \mathbf{E} represents the projection operator from the L_{MGA} th grid to the model grid.

Fifth, we add $\delta\mathbf{x}_{\text{MGA}}$ to the posterior ensemble mean produced by the EAKF without inflation in the first step to obtain the final ensemble mean analysis.

Last, we add the final ensemble mean analysis to the posterior deviations produced in the first step to form the final analyses of the ensemble members for the state variables.

As noted in W15, the EAKF-MGA is different from the running in place and quasi-outer-loop (RIP/QOL) method presented in Yang et al. (2012). In the RIP/QOL method, the same observation is assimilated more than once if the second loop is activated according to a criterion. In the EAKF-MGA, however, an observation at each data assimilation cycle is divided into two parts: $\mathbf{H}\bar{\mathbf{x}}_{\text{EAKF}}$ and \mathbf{y}^{res} as formulated by Eq. (7). The first term is the observational information extracted by the EAKF; and the second term is left behind by

the EAKF, which may contain multiscale information. The MGA can be adaptively triggered, according to the critical value θ , to retrieve the multiscale signals from the second term to compensate for the loss (contamination) of longwave information caused by an overly small (large) impact radius. As a result, an observation is actually used only once in the implementation of the EAKF-MGA. Note that if RMSE^{res} is less than or equal to the critical value θ , the MGA would not be activated; thus, the EAKF-MGA will reduce to the EAKF without inflation.

d. Implementation of PO

Under the framework of the EAKF, it is easy to realize PO. At each cycle of data assimilation, for each single observation the observational increment computed by Eq. (3) is also mapped to the increment of parameter ensemble as follows:

$$\Delta\beta_i = \frac{\text{cov}_{\beta,y}^p}{(\sigma_y^p)^2} \Delta y_i, \quad (13)$$

where $\text{cov}_{\beta,y}^p$ is the prior error covariance between parameter β and y^p , which can be calculated by substituting the parameter ensemble member in place of the state ensemble member in Eq. (5).

PO is thus embedded into the EAKF and the first step of the EAKF-MGA by keeping other steps involved in the estimation of the state variables unchanged.

Different from the state variables, parameters have no dynamically supported internal variability, the ensemble spread of a parameter is always decreased by PO. It is therefore indispensable to introduce the inflation of the prior ensemble of the parameter in PO. Following Zhang et al. (2012), we present the parameter inflation scheme:

$$\tilde{\beta}_i = \max\left(1, \frac{1}{\kappa} \frac{\sigma_{\beta,t_0}}{\sigma_{\beta,t}}\right) (\beta_i - \bar{\beta}) + \bar{\beta}, \quad (14)$$

where β_i and $\tilde{\beta}_i$ represent the i th prior and inflated ensembles of parameter β , respectively; $\bar{\beta}$ represents the ensemble mean; and κ is defined as the dimensionless sensitivity of the model state with regard to β . Calculation of κ will be described in section 3c. The terms σ_{β,t_0} and $\sigma_{\beta,t}$ denote the prior spread of β at the initial time t_0 and assimilation time t , respectively.

3. Setup of twin experiments

a. Model configuration

A biased twin experiment framework is designed to explore the performance of the EAKF with adaptive inflation (A07) and the EAKF-MGA. We set the Cressman parameter λ^2 in Eq. (2) with a biased value in the assimilation model to simulate the model bias. In the truth model, λ^2 is set to 10^{-12} m^{-2} . In the assimilation model, the initial ensemble mean of λ^2 is set to $1.2 \times 10^{-12} \text{ m}^{-2}$. A Gaussian white noise with the standard deviation being 1% of the biased value is added to the biased value to form the initial ensemble member of λ^2 . Following previous studies (Wu et al. 2012, 2014; W15), the time

filter coefficient γ is set to 0.01 in both models. Starting from the streamfunction at 1200 UTC 1 January 1991 derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis u and v fields at 500 hPa, both the truth model and assimilation model are spun up for 20 days each, to obtain their initial states.

b. Observation systems

The truth model is integrated for 200 days to provide “observations.” The observational interval is set to 6 h (12 time steps). A Gaussian noise with the standard deviation of $10^6 \text{ m}^2 \text{ s}^{-1}$ is superimposed on the “truth” to simulate the observational error. Two observation systems are designed in the following assimilation experiments. The first is a globally uniform network in which observations are available at all model grid points (hereafter referred to as “complete observations”). In this case, the value of K (number of observations) in Eq. (8) at each data assimilation cycle is 3456, which is equal to the dimension of the state vector. The second is the network used in W15 as shown in Fig. 1, in which 1872 observations are randomly scattered in 0° – 180° , 0° – 90°N (area A), 180° – 360° , 0° – 90°N (area B), and the Southern Hemisphere (area C) (hereinafter referred to as “random observations”). The spatial sampling density of the observations in area A is twice of that in area B and 3 times of that in area C. Then, the value of K is 1872 in the case of random observations. To determine the threshold θ for these two observation networks, we set the significance level to the typical value of 0.1; and the value of θ is $1.015 \times 10^6 \text{ m}^2 \text{ s}^{-1}$ for the complete observations and $1.021 \times 10^6 \text{ m}^2 \text{ s}^{-1}$ for the random observations, respectively.

c. Parameter sensitivity test

To determine which parameter to be estimated, we first perform parameter sensitivity test with 20-member ensemble following Zhang et al. (2012). For the barotropic model, we focus on the two parameters of λ^2 and γ as noted before. As described in Wu et al. (2012) and Zhang et al. (2012), the ensemble spread of model state resulted from parameter perturbation is used to represent the sensitivity of the model state, which is denoted as α_β . However, α_β is unsuitable to be directly applied to parameter variance inflation unless we convert it to a dimensionless value. Therefore, we define the dimensionless sensitivity as the ratio of α_β and a constant α_0 , that is, $\kappa = \alpha_\beta/\alpha_0$. In this study, α_0 is set to $10^6 \text{ m}^2 \text{ s}^{-1}$, which is exactly the standard deviation of the observational error because of its same unit and similar magnitude as α_β . Zhang et al. (2012) indicated that the inflation amplitude of a parameter ensemble is inversely proportional to the model’s sensitivity with respect to the parameter. Thus, we apply the inverse of κ to parameter variance inflation as formulated in Eq. (14). For each parameter, 20 Gaussian random numbers with the standard deviation being 5% of the truth value are generated to represent parameter perturbations. Starting from the initial condition of the truth model described in section 3a, the barotropic model, with one parameter perturbed and the other set to the default value, is integrated for 20 days. By spatially and

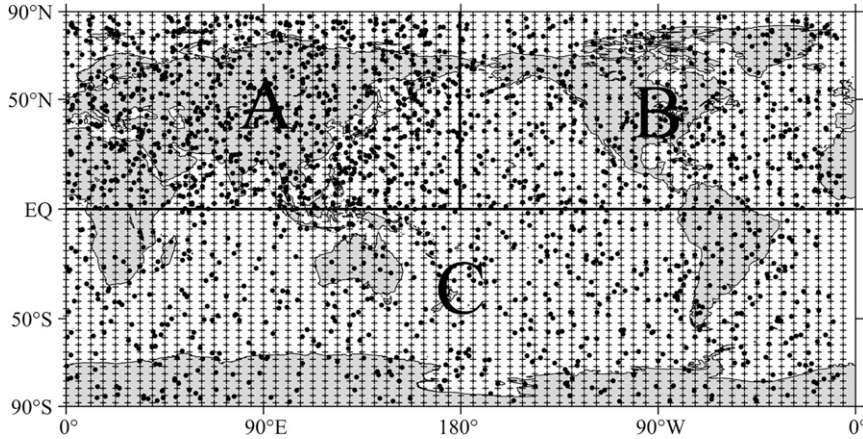


FIG. 1. Model grids (pluses) and randomly distributed observations (dots). The bold lines divide the globe into three areas. Labels A, B, and C represent the three areas with different sampling densities of observations.

temporally averaging values of α_β during the last integration day, means of α_β with regard to λ^2 and γ are 1.29×10^7 and $1.19 \times 10^7 \text{ m}^2 \text{ s}^{-1}$, respectively, indicating that the model state responds to changes in the value of λ^2 slightly more. It means that the streamfunction is slightly more sensitive to changes in λ^2 than to changes in γ .

We design the following three experiments to further examine parameter sensitivity. One is the SEO experiment that only uses observations to adjust the model state variables. The other two experiments are the single parameter optimization (SPO) for estimating λ^2 and γ , respectively. All three experiments use the same assimilation method (i.e., the EAKF with adaptive inflation), initial conditions, and complete observations. The assimilation period is 200 days; and the impact radius is set to 1500 km. The two SPO experiments use observations to adjust the model state for the first 100 days, while adjusting both the model state and one of the two parameters for the last 100 days. We use the space–time-averaged RMSE of the prior ensemble mean of the streamfunction during the last 10 days to measure the assimilation quality, which is calculated by

$$\overline{\text{RMSE}} = \frac{1}{S} \sum_{s=1}^S \text{RMSE}_s = \frac{1}{S} \sum_{s=1}^S \sqrt{\frac{1}{M} \sum_{j=1}^M (\bar{\psi}_{j,s}^{\text{prior}} - \psi_{j,s}^{\text{true}})^2}, \quad (15)$$

where $\bar{\psi}$ represents the ensemble mean of the streamfunction; superscripts “prior” and “true” represent the prior and true values, respectively; M and S are the numbers of model grid points and analysis steps, respectively; and RMSE_s represents the space-averaged RMSE at the s th analysis step.

The RMSEs of the SEO experiment and the two SPO experiments for λ^2 and γ are 0.86×10^6 , 0.42×10^6 , and $0.81 \times 10^6 \text{ m}^2 \text{ s}^{-1}$, respectively, suggesting that the parameter optimization experiment for λ^2 can obtain better assimilation quality than the other two experiments. Thus, λ^2 is chosen to perform PO in the following experiments.

d. Assimilation experiment setting

A Gaussian white noise with the standard deviation of $10^6 \text{ m}^2 \text{ s}^{-1}$ is added to the initial field of the assimilation model to form the initial ensemble of the streamfunction. Based on the observation networks described above, initial conditions, assimilation model, covariance localization, and inflation schemes, we design four assimilation experiments (see Table 1): EAKF_SEO, which only performs state estimation utilizing the EAKF with adaptive inflation for 200 days; EAKF_PO, which uses the EAKF with adaptive inflation to perform state estimation (simultaneous state and parameter estimation) for the first (last) 100 days; and MGA_SEO and MGA_PO, which are similar to EAKF_SEO and EAKF_PO, respectively, except for using the EAKF-MGA method instead. For each experiment, we consider 15 impact radii and 6 ensemble sizes. Note that PO is activated after 100 days of state estimation to ensure a “quasi-equilibrium” to be reached according to Zhang et al. (2012).

4. Results

a. Dependence on impact radius and ensemble size

Performance of the EnKF depends on covariance localization and ensemble size, as well as observation density (Houtekamer and Mitchell 1998; Mitchell et al. 2002; Ying et al. 2018). To get an overall sense of the effect of ensemble size combined with impact radius (i.e., a) on the performances of the EAKF with adaptive inflation and the EAKF-MGA, we initially employ complete observations to implement the four assimilation experiments illustrated in Table 1. To facilitate the description in the following text, we use the EAKF to simply represent the EAKF with adaptive inflation unless otherwise stated. Figure 2 presents the space–time-averaged RMSE of the prior ensemble mean for the streamfunction calculated using the results of the last 10 days produced by EAKF_SEO, EAKF_PO, MGA_SEO, and MGA_PO in terms of a and ensemble size. The first thing to be noticed is the blank area found in the upper left corner of

TABLE 1. Settings of assimilation experiments.

Experiment	Assimilation method	State estimation	Parameter optimization	State variance inflation	Impact radius (km)	Ensemble size
EAKF_SEO	EAKF with adaptive inflation	Yes	No	Yes	125, 250, 500, 750, 1000, 1250, 1500, 1750,	5, 10, 15, 20, 30, and 40
EAKF_PO	EAKF with adaptive inflation	Yes	Yes	Yes	2000, 2500, 3000, 3500, 4000, 4500, and 5000	
MGA_SEO	EAKF-MGA	Yes	No	No		
MGA_PO	EAKF-MGA	Yes	Yes	No		

both Figs. 2a and 2b, where the values of RMSE are not available because the model blows up due to filter divergence as the ensemble size decreases and a increases in the two experiments of EAKF_SEO and EAKF_PO. This means that the filter divergence caused by long-range spurious correlations in the model covariance cannot be resolved by means of the adaptive variance inflation. This does not happen for either MGA_SEO (Fig. 2c) or MGA_PO (Fig. 2d) even for extremely large a or extremely small ensemble size. In the two experiments of the EAKF, for a given ensemble size there is an optimal value of a ; and the optimal a slightly increases as the ensemble size increases. This is a standard result from

statistical theory, as explained in Houtekamer and Mitchell (1998). Compared to the EAKF, for a given ensemble size the EAKF-MGA helps alleviate the dependence of the assimilation quality on a , especially in the experiment of PO. That is to say, thanks to the function of the adaptive MGA in the EAKF-MGA method, the RMSE of MGA_SEO or MGA_PO first decreases with a gradually and then remains almost unchanged after the minimum is reached. In other words, when a is fixed, the RMSE of EAKF_SEO or EAKF_PO increases as the ensemble size decreases, while reducing ensemble size does not have much impact on MGA_SEO or MGA_PO in comparison. Particularly for extremely small ensemble size, say 5, the filter

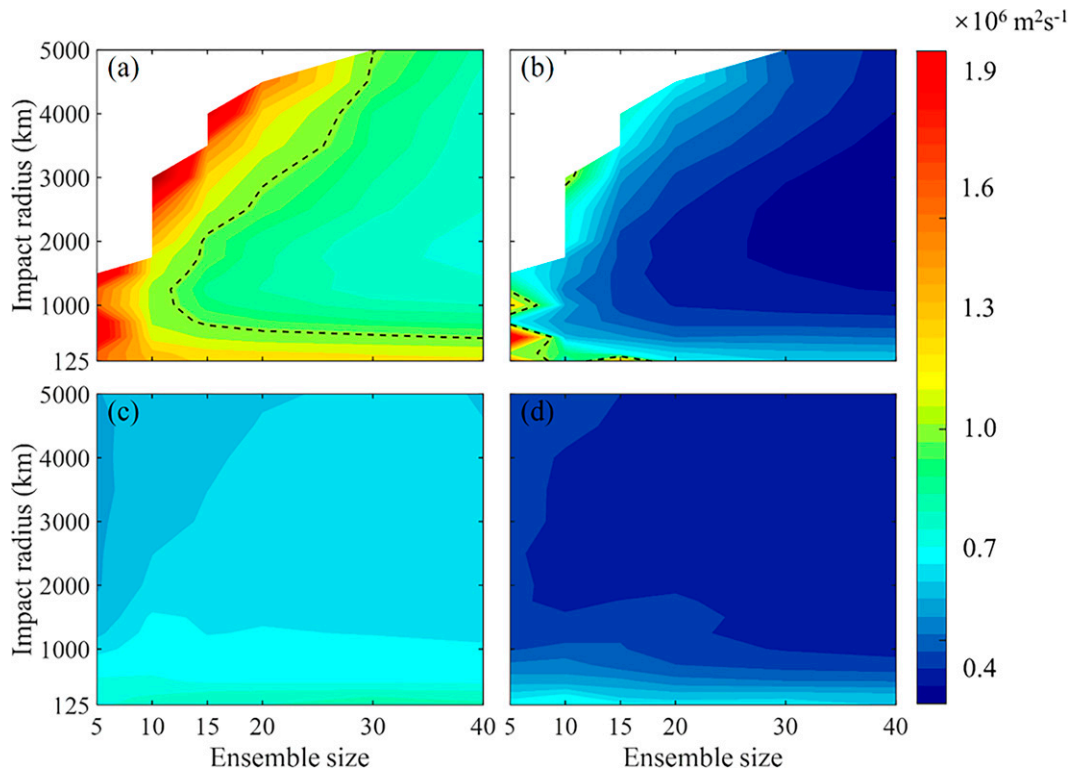


FIG. 2. RMSE ($10^6 \text{ m}^2 \text{ s}^{-1}$) of the prior ensemble mean for the streamfunction in (a) EAKF_SEO, (b) EAKF_PO, (c) MGA_SEO, and (d) MGA_PO with different a (125, 250, 500, 750, 1000, 1250, 1500, 1750, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 km) and ensemble sizes (5, 10, 15, 20, 30, and 40) using complete observations. The dashed curve indicates the standard deviation of observation error (i.e., $10^6 \text{ m}^2 \text{ s}^{-1}$). The blank areas in (a) and (b) denote no available outputs, because the model blows up.

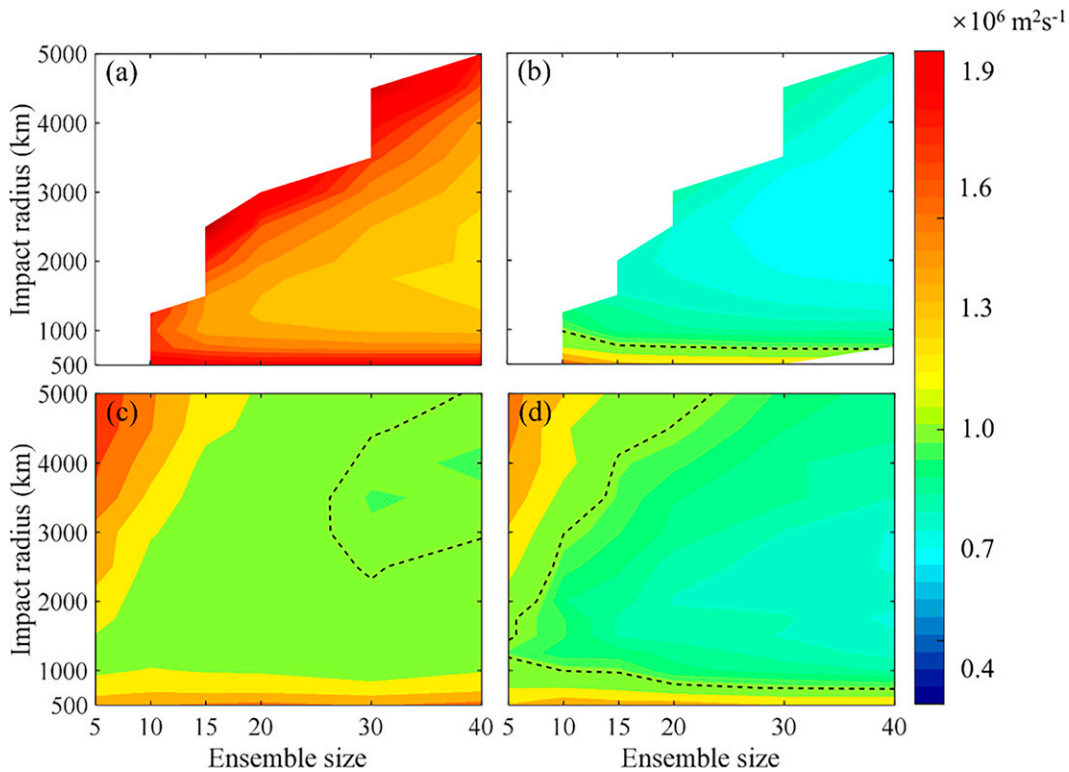


FIG. 3. As in Fig. 2, but using random observations.

divergence occurs in the EAKF for a larger than 1500 km, while SEO and PO with the EAKF-MGA still work effectively for all a . For a specific pair of $(a, \text{ensemble size})$, EAKF_PO (MGA_PO) can further reduce the state error relative to EAKF_SEO (MGA_SEO), and the RMSE of MGA_SEO is always smaller than that of EAKF_SEO. In terms of all pairs of $(a, \text{ensemble size})$, the RMSE of MGA_SEO or MGA_PO is less than the uncertainty of the observation (dashed curve shown in Figs. 2a,b), and MGA_PO achieves the lowest RMSE among these four experiments even for overly small ensemble size. The RMSE is much smaller than the standard deviation of observation error (i.e., $10^6 \text{ m}^2 \text{ s}^{-1}$) in MGA_PO for the given ensemble size of 5, which means that the accuracy of predicted model state is very high.

By employing random observations, we then explore the performances of the EAKF and EAKF-MGA in the PO experiments based on SEO. Because the longest distance from an observation to the nearest model grid point is 368 km when random observations are considered, the filter divergence of the EAKF undoubtedly occurs for a smaller than 500 km. Hence, as shown in Fig. 3, there are no results for $a = 125$ and 250 km. The distributions of the RMSE in Fig. 3 confirm many of the main features noted in Fig. 2, except that more blank areas appear in Figs. 3a and 3b, indicating that the EAKF using random observations diverges for more pairs of $(a, \text{ensemble size})$ compared with that using complete observations; and the RMSE tends to become increasingly larger for all four experiments, showing that the minimum RMSE in

EAKF_SEO (Fig. 3a) and part of the RMSE in EAKF_PO (Fig. 3b), EAKF_SEO (Fig. 3c) and EAKF_PO (Fig. 3d) is higher than the uncertainty of the observation (i.e., $10^6 \text{ m}^2 \text{ s}^{-1}$; dashed curve shown in Figs. 3b–d). Note that, due to the success in reducing the model bias through correcting the biased parameter, EAKF_PO (Fig. 3b) significantly reduces the state error with respect to EAKF_SEO and acquires the minimum RMSE even smaller than that of MGA_PO (Fig. 3d) with a given ensemble size larger than 20, while MGA_PO achieves high quality assimilation results for a broader range of a and ensemble size (see the area surrounded by the dashed curve in Fig. 3d). The fact that MGA_PO outperforms EAKF_PO for smaller ensembles, particularly the five-member ensemble in the case of random observations brings us conviction of the feasibility and potential advantage of the EAKF-MGA for practical applications.

b. Performance skill with five-member ensemble

Considering the advantage of the EAKF-MGA with regard to extremely small ensemble size, let us fully examine and compare the performances of the EAKF and EAKF-MGA applied in SEO and PO with ensemble size of 5. Figure 4 shows the space–time-averaged RMSE of the prior ensemble mean of the streamfunction as a function of a computed using the results of the last 10 days produced by the four experiments. Same as W15, the superimposed bar in Fig. 4 denotes the $\pm\zeta$ bound of the space–time-averaged RMSE, where ζ represents the standard deviation of the RMSE:

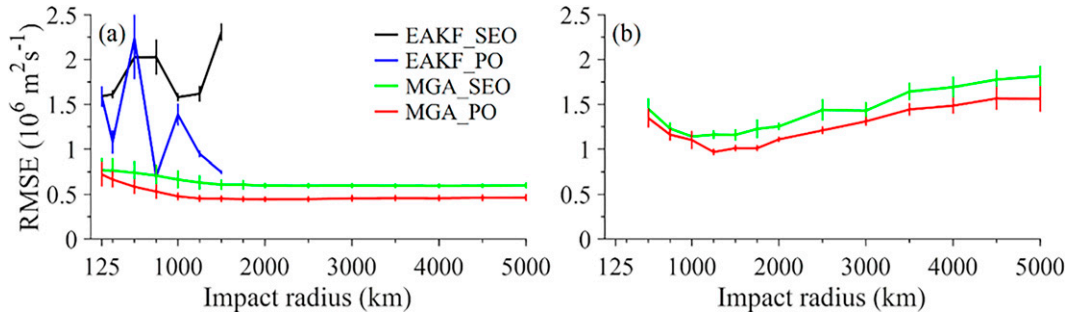


FIG. 4. Space-time-averaged RMSE ($10^6 \text{ m}^2 \text{ s}^{-1}$) of the prior ensemble mean of the streamfunction produced by EAKF_SEO (black curve), EAKF_PO (blue curve), MGA_SEO (green curve), and MGA_PO (red curve) with five-member ensemble in terms of different a (km) using (a) complete and (b) random observations, respectively. The superimposed bar denotes the $\pm\zeta$ bound of the space-time-averaged RMSE, where ζ represents the standard deviation of the RMSE.

$$\zeta = \sqrt{\frac{1}{S} \sum_{s=1}^S (\text{RMSE}_s - \overline{\text{RMSE}})^2}. \quad (16)$$

If the error produced by experiment A falls out of the $\pm\zeta$ bound of the error produced by experiment B, we define that the result of experiment A is significantly different from experiment B.

When complete observations are considered, as demonstrated in Fig. 4a, the irregularly oscillated RMSE suggests that EAKF_SEO (black curve) and EAKF_PO (blue curve) fail to work effectively even though a is not overly large, which corresponds to the large error area in the bottom left corner of Fig. 2a. The EAKF-MGA, however, works well in implementing both SEO (green curve) and PO (red curve) for the whole range of a compared with the EAKF. The RMSE of MGA_SEO or MGA_PO initially decreases gradually, and then remains at a fairly low level with increasing a . At the same time, through mitigating the model bias, MGA_PO significantly reduces the RMSE for all a with respect to MGA_SEO. When the observing system with random observations is employed, as illustrated in Fig. 4b, the EAKF no longer works, and the EAKF-MGA in both SEO (green curve) and PO (red curve) still works normally without filter divergence. The RMSE in SEO or PO decreases at first to a level less than the uncertainty of the observation (i.e., $10^6 \text{ m}^2 \text{ s}^{-1}$), and then grows slowly to a level greater than the uncertainty of the observation. In the meantime, when a is greater than 1000 km, the overall decrease in error that occurs with MGA_PO with respect to MGA_SEO can be seen by examining the two curves. Compared to the case of complete observations, while the error levels in SEO and PO are slightly larger compared to their counterparts in Fig. 4a, the results are acceptable especially when an appropriate a is chosen (e.g., 1250 km) for PO.

We compare the results in detail in Fig. 5 by illustrating the case of $a = 1250$ km. Results shown in the left panels of Fig. 5 are obtained using complete observations, and those in the right panels are the counterparts employing random observations. Let us look first at the left panels of Fig. 5 when

complete observations are considered. Figure 5a shows the time series of the daily space-averaged RMSE of the prior ensemble mean of the streamfunction produced by the four experiments. On account of the increased sampling error from the overly small ensemble size, the RMSE of EAKF_SEO (black curve) gradually increases within the assimilation window with the error level being larger than the standard deviation of observation error (i.e., $10^6 \text{ m}^2 \text{ s}^{-1}$), as the dashed line shown in Fig. 5a. The RMSE of EAKF_PO (blue curve) shows an irregular oscillation within the assimilation window. Note that by examining the results of EAKF_SEO and EAKF_PO beyond the provided assimilation window (not shown), it is found that the filter divergence occurs in EAKF_SEO after 377 days and in EAKF_PO after 286 days. By contrast, MGA_SEO (green curve) has a low-level error of $0.72 \times 10^6 \text{ m}^2 \text{ s}^{-1}$ after the spinup of the assimilation, demonstrating that through adaptively triggering the MGA after the EAKF, the EAKF-MGA can retrieve the multiscale information from the observational residual to reduce the sampling noise and thus further decrease the state error. MGA_PO (red curve) significantly reduces the state error with regard to MGA_SEO. The RMSE of MGA_PO stays at the level of $0.49 \times 10^6 \text{ m}^2 \text{ s}^{-1}$ after the spinup of the assimilation. To further investigate the performance of PO, we show the time series of the ensemble mean of λ^2 produced by EAKF_PO (blue curve) and MGA_PO (red curve) for $a = 1250$ km in Fig. 5b. We can see that as PO is activated on the 100th day in MGA_PO, the ensemble mean of λ^2 is quickly reduced from the biased value (i.e., $1.2 \times 10^{-12} \text{ m}^{-2}$) to the true value (i.e., 10^{-12} m^{-2}), as the dashed line shown in Fig. 5b; it then oscillates around the truth. The EAKF apparently does not succeed in estimating the true value of λ^2 . To figure out how the MGA works in the EAKF-MGA, we show the time series of RMSE^{EES} defined in Eq. (8) obtained from MGA_SEO (green curve) and MGA_PO (red curve) in Fig. 5c. In MGA_SEO, the MGA is always triggered because of the low assimilation quality from the first step of the EAKF without inflation. When the MGA is activated, it makes the estimated ensemble means get closer to observations, and the assimilation quality is thus improved. Otherwise, the adaptive

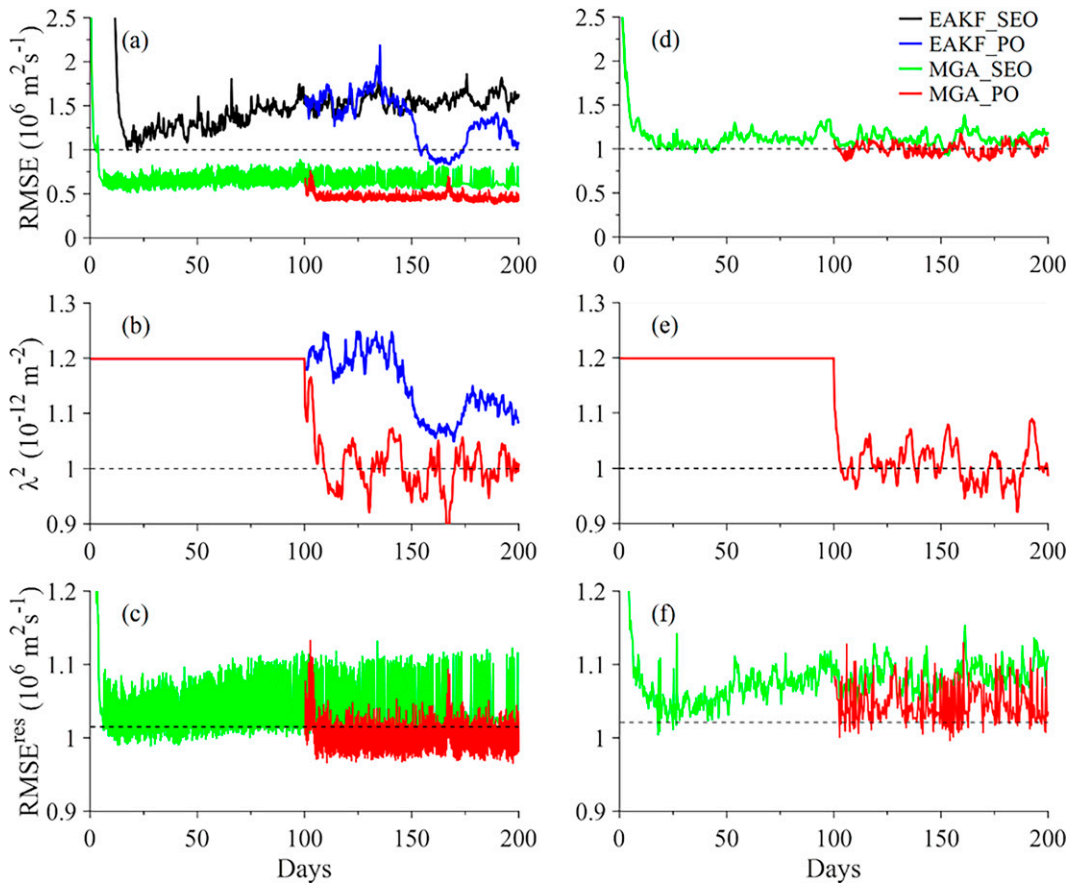


FIG. 5. (a) Time series of daily space-mean RMSE ($10^6 \text{ m}^2 \text{ s}^{-1}$) of the prior ensemble mean of the streamfunction produced by EAKF_SEO (black curve), EAKF_PO (blue curve), MGA_SEO (green curve), and MGA_PO (red curve) for $a = 1250 \text{ km}$, with the dashed line indicating the standard deviation of observation error (i.e., $10^6 \text{ m}^2 \text{ s}^{-1}$). (b) Time series of the ensemble mean of λ^2 produced by EAKF_PO (blue curve) and MGA_PO (red curve) for $a = 1250 \text{ km}$, with the dashed line representing the truth value of λ^2 . (c) Time series of RMSE^{res} produced by MGA_SEO (green curve) and MGA_PO (red curve) for $a = 1250 \text{ km}$, with the dashed line representing the threshold when deciding whether to trigger the MGA. (a)–(c) Obtained by using complete observations; (d)–(f) as in (a)–(c), but for using random observations. Note that no results of EAKF_SEO and EAKF_PO are available in (d) and (e) since the model blows up.

method reduces to the EAKF without inflation, and the RMSE increases as expected. As a result, the state error produced by the EAKF-MGA presents relatively large-amplitude oscillations as shown in Fig. 5a (green curve). Moreover, since the model bias is reduced after PO is activated in MGA_PO, the MGA is not called as frequently as in MGA_SEO, but still plays a role in improving assimilation accuracy by compensating the impact of the localization and sampling error.

Let us examine the right panels of Fig. 5 when random observations are employed. Note that no results are available for the EAKF since it no longer works in this case. Figure 5d shows that the RMSE of MGA_PO (red curve) is lower than MGA_SEO (green curve) and oscillates around its equilibrium position that coincides with the dashed line of $10^6 \text{ m}^2 \text{ s}^{-1}$, indicating that jointly estimating state and parameter with the EAKF-MGA can provide high accuracy of predicted model state. As shown in Fig. 5e, MGA_PO-estimated λ^2 oscillates around the truth, demonstrating that the biased model parameter has been

optimized successfully. In Fig. 5f, no matter PO is activated or not, the MGA is always called in the EAKF-MGA, which is slightly different from that shown in Fig. 5c. The reason lies in the fact that under the circumstance of sparse observations and very limited ensemble size, although the analysis bias is reduced, assimilation quality produced by the EAKF without inflation is not satisfactory. The MGA thereby needs to be called more frequently than that with complete observations to ensure assimilation quality and prevent filter divergence.

Additionally, taking $a = 1250 \text{ km}$ for example, we examine the spatial distribution of the time-averaged RMSE of the prior ensemble mean of the streamfunction over the last 10 days produced by the two experiments of MGA_SEO and MGA_PO using complete observations (Figs. 6a,b) and random observations (Figs. 6c,d), respectively. We can see that the substantial overall decrease in error occurs with complete observations in the top panels of Fig. 6 compared to those using random observations in the bottom panels, especially

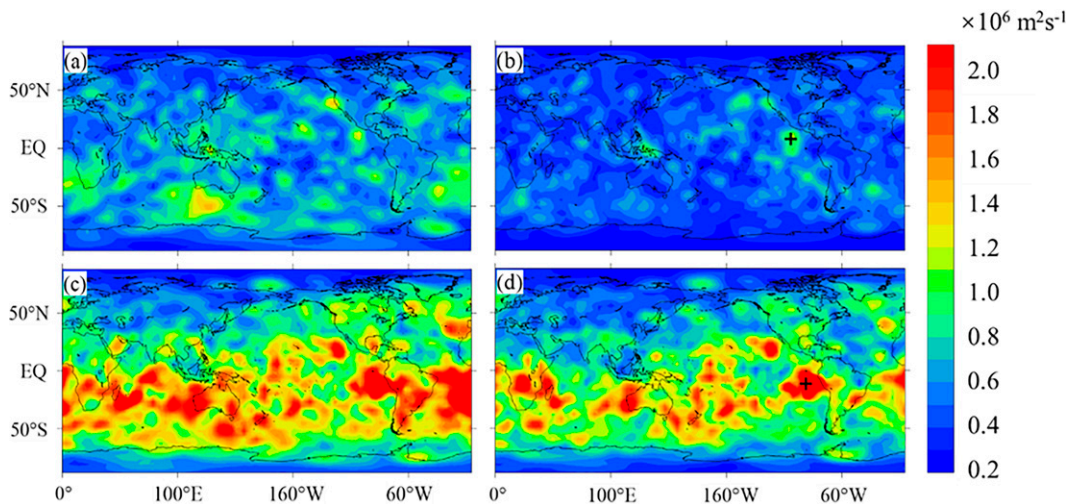


FIG. 6. Spatial distributions of time-averaged RMSE ($10^6 \text{ m}^2 \text{ s}^{-1}$) of the prior ensemble mean of the streamfunction over the last 10 days produced by (a),(c) MGA_SEO and (b),(d) MGA_PO with five-member ensemble, for $a = 1250$ km: (top) complete observations and (bottom) random observations. The black plus signs in (b) and (d) indicate the (46, 29)th model grid at (5.80317°N , 106.87500°W) and the (48, 27)th model grid at (2.52815°S , 95.62500°W), where the RMSE of MGA_PO is relatively large.

in the Southern Hemisphere where observations are sparser and more uneven. For each observation system, the RMSE of MGA_PO (right panels of Fig. 6) is globally decreased with respect to MGA_SEO (left panels of Fig. 6), demonstrating that PO implemented through the EAKF-MGA method with only five-member ensemble succeeds in reducing the model bias and improving the accuracy of state estimation.

As noted previously, there is no state variance inflation in the implementation of the EAKF-MGA. One may wonder about the impact of the uninflated ensemble of state variables on the performance of the EAKF-MGA. Figure 7 shows the averaged spread of the prior ensemble of the streamfunction over the last 10 days produced by EAKF_SEO (black curve), EAKF_PO (blue curve), MGA_SEO (green curve), and MGA_PO (red curve) as a function of a , with the ensemble size of 5. With complete observations (Fig. 7a), rigorous oscillation found in the ensemble spread produced by EAKF_SEO

(black curve) or EAKF_PO (blue curve) for small a is consistent with the RMSE presented in Fig. 4a. The EAKF fails to work for a larger than 1500 km, as noted before. Due to the absence of state variance inflation in the EAKF-MGA, the ensemble spreads produced by MGA_SEO (green curve) and MGA_PO (red curve) are almost the same; in general, they decrease gradually as a increases with both observation systems. Overall, larger ensemble spread is obtained with random observations (Fig. 7b) than with complete observations (Fig. 7a); this is because for the same a , fewer observations are assimilated to adjust the state ensembles by using random observations than by using complete observations.

We then examine the evolution of the ensemble spread produced by the two PO experiments with complete observations, by taking the (46, 29)th model grid at (5.80317°N , 106.87500°W) as an example, where the RMSE of MGA_PO is relatively large marked with a black plus sign as shown in Fig. 6b. The time series of the spread of the prior ensembles

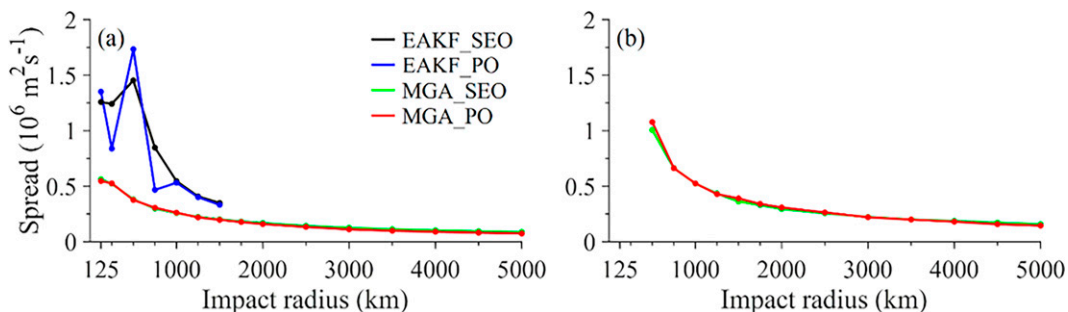


FIG. 7. Space-time-averaged ensemble spread ($10^6 \text{ m}^2 \text{ s}^{-1}$) of the prior ensemble mean of the streamfunction produced by EAKF_SEO (black curve), EAKF_PO (blue curve), MGA_SEO (green curve), and MGA_PO (red curve) with five-member ensemble in terms of different a (km): (a) complete observations and (b) random observations.

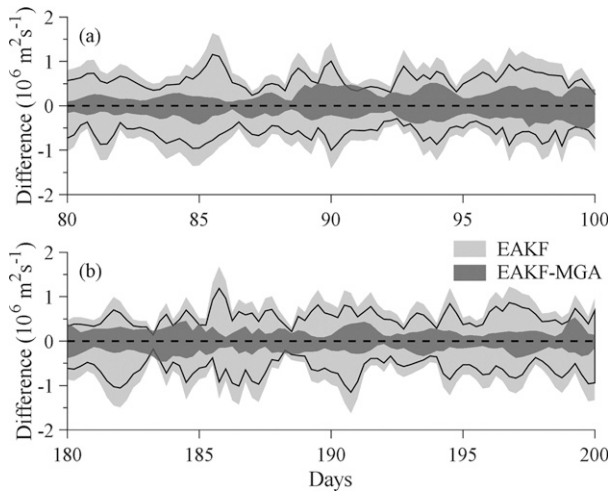


FIG. 8. Time series of the ensemble spread produced by the EAKF (light gray shading) and EAKF-MGA (dark gray shading), respectively, with five-member ensemble, for $a = 1250$ km, by using complete observations (a) from the 80th to 100th day corresponding to state estimation and (b) from the 180th to 200th day corresponding to state and parameter estimation at the (46, 29)th model grid. Note that the superimposed solid black curves indicate the bounds of the ensemble spread before state variance inflation is applied in EAKF_PO.

produced by the EAKF (light gray shading) and EAKF-MGA (dark gray shading) with complete observations for $a = 1250$ km are shown in Figs. 8a and 8b, respectively, corresponding to state estimation from the 80th to 100th day, and to state and parameter estimation from the 180th to 200th day, respectively. Note that the superimposed solid black curves indicate the bounds of the ensemble spread before state variance inflation is applied in EAKF_PO. Before PO is activated, the state ensemble produced by the EAKF is greatly inflated so as to improve the assimilation quality (Fig. 8a). Because of the absence of state variance inflation, the ensemble spread produced by the EAKF-MGA is much smaller. With PO activated (Fig. 8b), there is no significant change in the ensemble spread produced by the EAKF. The reason is that, due to the overly small ensemble size and the consequent increased sampling error, a sufficient variance inflation is indispensable even though the model bias is corrected by PO simultaneously. Actually, EAKF_PO with five-member ensemble does not really succeed in reducing the model bias (see the oscillated blue curve in Fig. 4a), and the slightly high quality state estimation is achieved for $a = 1250$ km only by chance as shown in Fig. 4a. In comparison, thanks to the function of the MGA on the basis of the EAKF without inflation, even with smaller ensemble spread, MGA_PO can always work well to retrieve the multiscale information from observations and provide high analysis accuracy.

The ensemble spread at the (48, 27)th model grid, marked with a black plus sign in Fig. 6d, produced by MGA_SEO and MGA_PO with random observations is illustrated in Fig. 9. Note that no data are available for the EAKF since it blows up. We can see that without state variance inflation, the EAKF-MGA

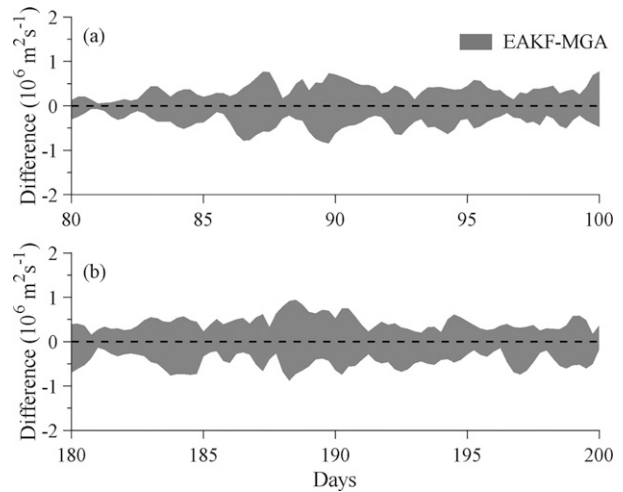


FIG. 9. As in Fig. 8, but at the (48, 27)th model grid for the EAKF-MGA using random observations. Note that no data are available for the EAKF since it blows up.

produces similar ensemble spread both before and after PO takes effect. Moreover, the ensemble spread of MGA_PO with random observations (Fig. 9b) is slightly larger than its counterpart using complete observations (dark gray shading in Fig. 8b) for the same reason noted for Fig. 7b.

To further evaluate the EAKF-MGA through comprehensively considering the model error and ensemble spread, we focus on the ensemble spread score (ESS) presented by Palmer et al. (2006) and by Keller and Hense (2011), which is defined as the ratio of the average ensemble variance and the mean squared error of the state ensemble mean in the verification interval. If the ESS is smaller (greater) than unity, the state ensemble is considered to be under-dispersive (over-dispersive), and the prediction system has higher (lower) potential than actual prediction skill (Glowienka-Hense et al. 2018, manuscript submitted to *Geophys. Model Dev. Discuss.*). Kadow et al. (2016) used the log-transform of the ESS for the retrospective hindcast verification of “Mittelfristige Klimaprognose” (MiKlip) system. Polkova et al. (2019) used the ESS as one of the metrics to assess the impact of several methods of initialization and ensemble generation on the prediction skill for decadal climate prediction. Figure 10 shows the space-time-averaged values of ESS over the last 10 days produced by EAKF_SEO (black curve), EAKF_PO (blue curve), MGA_SEO (green curve), and MGA_PO (red curve), for different a by using complete observations (Fig. 10a) and random observations (Fig. 10b). In Fig. 10a, values of ESS produced by EAKF_SEO (black curve) and EAKF_PO (blue curve) are smaller than unity although irregular oscillations occur for a less than 1500 km. Accordingly, the state ensembles produced by the two experiments of the EAKF are underdispersive when the model does not blow up. Given complete observations (Fig. 10a) or random observations (Fig. 10b), values of ESS produced by MGA_SEO (green curve) and MGA_PO (red curve) decrease gradually as a increases, and are overall smaller than unity, indicating that the state ensembles obtained from SEO and PO implemented with the EAKF-MGA are

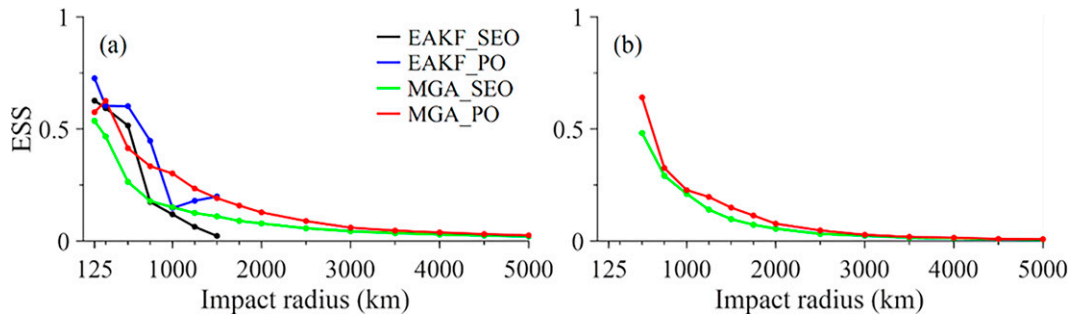


FIG. 10. Space-time-averaged values of ESS over the last 10 days produced by EAKF_SEO (black curve), EAKF_PO (blue curve), MGA_SEO (green curve), and MGA_PO (red curve) with five-member ensemble in terms of different a (km) by using (a) complete observations and (b) random observations.

underdispersive. Note that values of ESS of MGA_SEO are slightly lower than those of MGA_PO; and the reason is that the state error produced by MGA_SEO is larger than that by MGA_PO as shown in Fig. 4, while the ensemble spreads of the two experiments are almost the same as shown in Fig. 7. To sum up, although there is no state variance inflation applied, even with five-member ensemble, the EAKF-MGA shows a good potential for building efficient prediction system.

Based on what have been discussed above, we can conclude that, by adaptively introducing the compensation to the EAKF to target the issues that arise in the context of both covariance inflation and localization, the EAKF-MGA performs better than the EAKF in SEO and PO experiments in general. The reason why the EAKF-MGA successfully achieves performance improvement over the EAKF lies in the following facts. One of the essential characteristics of modern data assimilation theory is the standard assumption that the background field used in the statistical-analysis equation is unbiased. If the short-term forecast, used as the background field, is in fact biased, so will the resulting analysis be biased (Dee and Da Silva 1998). There is no doubt that this is precisely the case in our study under the biased twin assimilation experiment framework. The biased parameter in the barotropic spectral model will certainly lead to biased analysis. Therefore, the EAKF-MGA is actually designed to estimate and correct systematic errors jointly with the model state variables. Once biases, including the model bias caused by the biased parameter and the analysis bias resulted from the sampling noise given the limited ensemble size, are corrected

by the MGA through adjusting the ensemble mean obtained from the EAKF, it can guarantee that the analysis will be proceeded onward with the assumption of the unbiased model background field to be met. In particular, due to the key role played by the MGA, the EAKF-MGA can work normally even with extremely small ensemble size and overly large impact radius given random observations.

c. Computational cost

As pointed out in Zhang and Anderson (2003), the 20-member ensemble is applicable for studying the capability of the EAKF in the global barotropic spectral model. Actually, both EAKF and EAKF-MGA are able to provide satisfactory performance in PO with ensemble size of 20 by using both complete and random observations, as shown in Figs. 2 and 3, respectively. Thus, before discussing the computational cost, we examine the results of both methods with 20-member ensemble first.

Figure 11 shows the prior space-time-averaged RMSE produced by the four experiments for different a using complete observations (Fig. 11a) and random observations (Fig. 11b). From Fig. 11a, we can see that the RMSE of EAKF_SEO (black curve) decreases from $a = 125$ to 1250 km, but gradually increases as a gets larger. For extremely large a (e.g., 5000 km), the EAKF fails to work in SEO, which corresponds to filter divergence. Compared with EAKF_SEO, the RMSE of MGA_SEO (green curve) is significantly smaller especially for overly small and large a because of the effect of the adaptively triggered MGA. Note that the results of the SEO experiments in this study are different from those in W15 where the performance of the

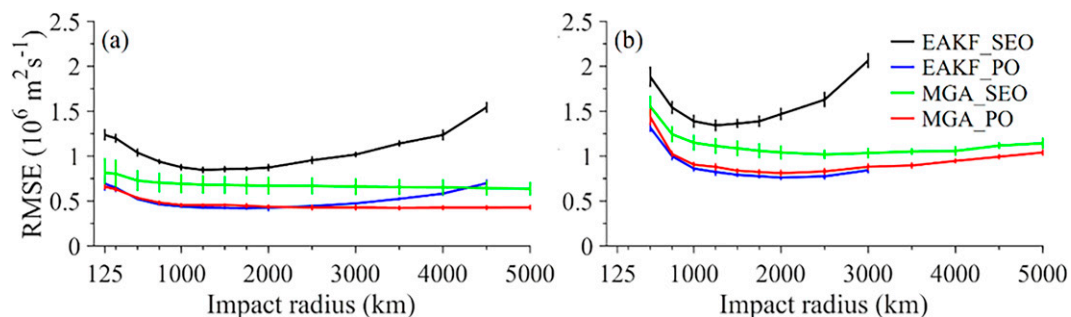


FIG. 11. As in Fig. 4, but for ensemble size of 20.

EAKF-MGA in SEO was slightly worse than the EAKF with optimal inflation for moderate a . As noted, to emphasize the influence of the model bias on assimilation quality, we set λ^2 , a high-sensitivity parameter comparing to γ , to be the biased value in this study instead of the biased γ in W15. When a strong bias is introduced, through correcting the ensemble mean produced by the EAKF step, the adaptive MGA in the EAKF-MGA ensures a better response of the state ensemble to the observations than the EAKF does. It turns out that the performance of the EAKF-MGA applied in SEO is better than that of the EAKF. PO with the EAKF-MGA (i.e., MGA_PO; red curve in Fig. 11a) significantly reduces the RMSE to reach a lower level with respect to MGA_SEO for all a . There is almost no difference among the results produced by EAKF_PO (blue curve) and MGA_PO (red curve) for a smaller than 2500 km. In comparison, as a gets larger, the RMSE produced by EAKF_PO gradually increases while the assimilation quality of MGA_PO remains almost unchanged. From the view of Zhang et al. (2012), the low accuracy of state estimation with the EAKF resulted from overly large a and the consequent low signal-to-noise ratio (SNR) of the error covariance between the observed model state and parameter leads to the bad performance of EAKF_PO. In contrast, the adaptive MGA ensures the high accuracy of state estimation, and consequently leads to the better performance of MGA_PO.

From Fig. 11b, we can see that with the reduction in the number and uneven distribution of the observations, the RMSEs of the four experiments are correspondingly larger than their counterparts by employing complete observations shown in Fig. 11a. The RMSE of EAKF_SEO (black curve) decreases for a smaller than 1500 km and then increases as a gets larger. EAKF_PO (blue curve) works well for a not larger than 3000 km, with the RMSE being much smaller with respect to EAKF_SEO. For a larger than 3000 km, however, the EAKF blows up, which indicates that, compared with the case of complete observations, it only works in a narrower range of a due to the sparse observations. The performance of MGA_SEO (green curve) is much better than EAKF_SEO; and MGA_PO (red curve) succeeds in reducing the RMSE with regard to MGA_SEO, which is very close to that of EAKF_PO for a not larger than 3000 km. Thanks to the adaptively triggered MGA, the EAKF-MGA works effectively in both SEO and PO for a broader range of a (as is the case with complete observations) than the EAKF.

Similar to the analysis for the ensemble spread in section 4a, we also checked the space–time average of the prior ensemble spread of the streamfunction obtained with given ensemble size of 20 (not shown). As expected, larger ensemble spread is obtained with random observations than those with complete observations. Regarding EAKF_PO, the state ensemble is obviously inflated before PO is activated. However, with PO activated, the ensemble spread decreases to the level close to that of MGA_PO, which is different from that shown in Fig. 8b. Because the model bias has been mitigated and the sampling error decreases with 20-member ensemble, EAKF_PO does not need a large state variance inflation factor to be introduced. Furthermore, values of the space–time-averaged ESS produced by MGA_SEO and MGA_PO with the given

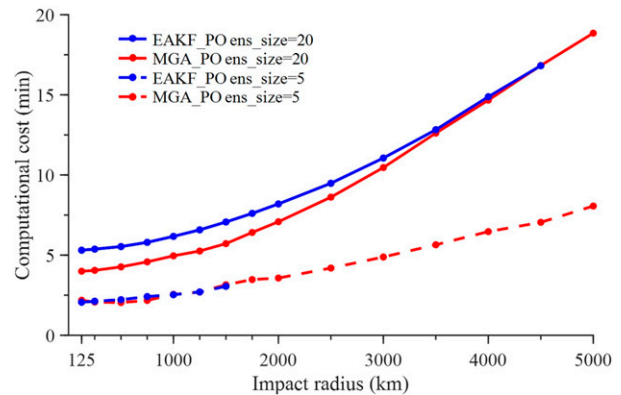


FIG. 12. Computational cost (min) with respect to a (km) produced by EAKF_PO (blue curve) and MGA_PO (red curve) with the ensemble size of 20 (solid curve) and 5 (dashed curve). No data are shown with EAKF_PO for a larger than 1500 km and given ensemble size of 5 because the EAKF no longer works.

ensemble size of 20 are smaller than unity (not shown). As a result, the prior ensemble obtained by SEO and PO with the EAKF-MGA is underdispersive, indicating that the model system initialized using the EAKF-MGA has good prediction skill.

As pointed out in Chang and Kalnay (2022), finding a balance between accuracy and computational cost is an inevitable challenge for modern EnKF applications; and this is particularly true in the context of operational data assimilation for large-scale systems. Therefore, although parameter optimization implemented with both EAKF and EAKF-MGA succeed in reducing the model bias and improving assimilation accuracy under appropriate conditions, with a view toward the practical application for large-scale systems the computational cost of these two methods is worthy of further investigation. Both the adaptive state variance inflation and the adaptive MGA increase the computational cost on basis of the EAKF. We compare the computational costs that are measured using the wall-clock time (in min) of the assimilation experiments carried out with complete observations. Figure 12 shows, when the observing system with complete observations is employed, the computational cost with respect to a required by implementing EAKF_PO and MGA_PO given the ensemble size of 20 and 5, respectively. We can see that the computational cost in each experiment increases gradually as a gets larger. In comparison, cost rising in the case of five-member ensemble is more slowly. Given ensemble size of 20, EAKF_PO costs more time than MGA_PO does for a smaller than 3500 km. The difference in time consumption between the two is the adaptive calculation of the state variance inflation factor and the adaptively activated MGA. It seems that, as a variational method, the latter is more computationally efficient than the former with moderate a . Although for the 20-member ensemble, Fig. 11 shows that the RMSE is not very different for parameter estimation with both EAKF and EAKF-MGA, and may even suggest that the EAKF is a little bit better than the EAKF-MGA within a certain range of a , the $\pm\zeta$ bound of the RMSE of MGA_PO is nearly overlapped by that of EAKF_PO, with ζ indicating the standard

deviation of the RMSE as described in Eq. (16), demonstrating that the difference between the two experiments is not significant. It means that PO with the EAKF-MGA can achieve good enough results with less computational cost, provided that one chooses an a within a reasonable range rather than spends much time tuning for an optimal a when applying the EAKF. Undoubtedly, the computational cost of the EAKF-MGA given ensemble size of 5 is much less. In summary, the EAKF-MGA is of promising prospects in operational data assimilation for large-scale systems from the viewpoint of low computational cost.

5. Summary and discussion

Model bias is an important factor affecting the quality of forecast. Parameter uncertainty is a source of model bias, which can be reduced by parameter optimization (PO) via data assimilation. Due to the advantages of the hybrid adaptive EAKF-MGA method proposed in W15 over the EAKF with optimal inflation, we adopt it to implement PO by performing series of assimilation experiments in this study. At each assimilation cycle, the state and parameter ensembles are first adjusted by the EAKF without inflation. Then, the observational residual is computed and compared with the threshold to see whether the MGA should be triggered to retrieve the multiscale information from the observational residual to correct the state ensemble mean.

The barotropic spectral model is used to study the feasibility of the EAKF-MGA in mitigating model bias. One of the two key parameters contained in the model, i.e., the topographical Cressman parameter λ^2 , is to be optimized through sensitivity study. Biased twin experiments are carried out: in one case, a globally uniform observing system with observations available at all model grid points; and in the other, an observing system with observations distributed randomly. For different impact radii and ensemble sizes, we analyze and compare the performances of the EAKF with adaptive inflation and the EAKF-MGA in the state estimation only (SEO) and PO experiments. Results show that because of the compensation of the adaptive MGA, the assimilation quality produced by SEO with the EAKF-MGA is much better than the EAKF with adaptive inflation. PO with the EAKF-MGA can effectively reduce the model bias and further improve the analysis quality compared to SEO. The hybrid adaptive method also works well in SEO and PO for extremely small ensemble size, even in the case of randomly distributed observations. The adaptively triggered MGA is actually a practice of bias correction and can effectively target the issues of both state variance inflation and localization. In conclusion, as stated in W15, when the method is implemented for SEO, PO implemented within the EAKF-MGA is labor saving since neither state variance inflation nor intensive trial-and-error tuning for covariance localization is needed. Computational efficiency is another important feature of the EAKF-MGA since small ensemble size is allowed. Given the advantages of the EAKF-MGA found in this study, which is actually designed to estimate and correct systematic errors jointly with the model state variables to make sure that the assumption of the unbiased model background field in modern data

assimilation theory holds, we anticipate that it will have a promising, broad prospect in operational data assimilation for large-scale systems.

Acknowledgments. This research is cosponsored by grants from the National Natural Science Foundation of China (41876014 and 41976019), and the Open Fund Project of Key Laboratory of Marine Environmental Information Technology, Ministry of Natural Resources, China. We sincerely appreciate the editor and reviewers for their valuable comments and suggestions that helped us improve the quality of the article.

Data availability statement. The output data of this study are obtained by the model numerical computation and available from the corresponding authors on request. The reanalysis data used in this study are publicly available from the fifth-generation ECMWF analysis (ERA5) via an online interface. A description of the dataset and data access instructions can be found at <https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5>.

REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2).
- , 2007a: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59A**, 210–224, <https://doi.org/10.1111/j.1600-0870.2006.00216.x>.
- , 2007b: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99–111, <https://doi.org/10.1016/j.physd.2006.02.011>.
- , and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, [https://doi.org/10.1175/1520-0493\(1999\)127<2741:AMCIOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2).
- , and L. Lei, 2013: Empirical localization of observation impact in ensemble Kalman filters. *Mon. Wea. Rev.*, **141**, 4140–4153, <https://doi.org/10.1175/MWR-D-12-00330.1>.
- Asselin, R., 1972: Frequency filter for time integrations. *Mon. Wea. Rev.*, **100**, 487–490, [https://doi.org/10.1175/1520-0493\(1972\)100<0487:FFFTI>2.3.CO;2](https://doi.org/10.1175/1520-0493(1972)100<0487:FFFTI>2.3.CO;2).
- Bishop, C. H., and D. Hodyss, 2007: Flow adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 2029–2044, <https://doi.org/10.1002/qj.169>.
- , and —, 2009a: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus*, **61A**, 84–96, <https://doi.org/10.1111/j.1600-0870.2008.00371.x>.
- , and —, 2009b: Ensemble covariances adaptively localized with ECO-RAP. Part 2: A strategy for the atmosphere. *Tellus*, **61A**, 97–111, <https://doi.org/10.1111/j.1600-0870.2008.00372.x>.
- , and —, 2011: Adaptive ensemble covariance localization in ensemble 4D-VAR state estimation. *Mon. Wea. Rev.*, **139**, 1241–1255, <https://doi.org/10.1175/2010MWR3403.1>.
- Chang, C.-C., and E. Kalnay, 2022: Applying prior correlations for ensemble-based spatial localization. *Nonlinear Processes Geophys.*, **29**, 317–327, <https://doi.org/10.5194/npg-29-317-2022>.

- Danforth, C. M., E. Kalnay, and T. Miyoshi, 2007: Estimating and correcting global weather model error. *Mon. Wea. Rev.*, **135**, 281–299, <https://doi.org/10.1175/MWR3289.1>.
- Dee, D. P., and A. M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295, <https://doi.org/10.1002/qj.49712454512>.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162, <https://doi.org/10.1029/94JC00572>.
- , 2007: *Data Assimilation: The Ensemble Kalman Filter*. Springer, 187 pp.
- Flowerdew, J., 2015: Towards a theory of optimal localisation. *Tellus*, **67A**, 25257, <https://doi.org/10.3402/tellusa.v67.25257>.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, <https://doi.org/10.1002/qj.4971255417>.
- Gharant, M. E., 2018: Enhanced adaptive inflation algorithm for ensemble filters. *Mon. Wea. Rev.*, **146**, 623–640, <https://doi.org/10.1175/MWR-D-17-0187.1>.
- Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology*. 2nd ed. Wiley, 477 pp.
- Han, G., X. Wu, S. Zhang, Z. Liu, and W. Li, 2013: Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and a simple pycnocline ocean model. *J. Climate*, **26**, 10218–10231, <https://doi.org/10.1175/JCLI-D-13-00236.1>.
- , X.-F. Zhang, S. Zhang, X.-R. Wu, and Z. Liu, 2014: Mitigation of coupled model biases induced by dynamical core misfitting through parameter optimization: Simulation with a simple pycnocline prediction model. *Nonlinear Processes Geophys.*, **21**, 357–366, <https://doi.org/10.5194/npg-21-357-2014>.
- , X. Wu, S. Zhang, Z. Liu, I. M. Navon, and W. Li, 2015: A study of coupling parameter estimation implemented by 4D-Var and EnKF with a simple coupled system. *Adv. Meteor.*, **2015**, 530764, <https://doi.org/10.1155/2015/530764>.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, [https://doi.org/10.1175/1520-0493\(1998\)126<0796:DAUAEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2).
- , and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, <https://doi.org/10.1175/MWR-D-15-0440.1>.
- Kadow, C., S. Illing, O. Kunst, H. W. Rust, H. Pohlmann, W. A. Müller, and U. Cubasch, 2016: Evaluation of forecasts by accuracy and spread in the MiKlip decadal climate prediction system. *Meteor. Z.*, **25**, 631–643, <https://doi.org/10.1127/metz/2015/0639>.
- Keller, J. D., and A. Hense, 2011: A new non-Gaussian evaluation method for ensemble forecasts based on analysis rank histograms. *Meteor. Z.*, **20**, 107–117, <https://doi.org/10.1127/0941-2948/2011/0217>.
- Lei, L., and J. L. Anderson, 2014: Empirical localization of observations for serial ensemble Kalman filter data assimilation in an atmospheric general circulation model. *Mon. Wea. Rev.*, **142**, 1835–1851, <https://doi.org/10.1175/MWR-D-13-00288.1>.
- Li, H., E. Kalnay, and T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 523–533, <https://doi.org/10.1002/qj.371>.
- Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization. *Math. Program.*, **45**, 503–528, <https://doi.org/10.1007/BF01589116>.
- Mitchell, H. L., P. L. Houtekamer, and G. Pellerin, 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, **130**, 2791–2808, [https://doi.org/10.1175/1520-0493\(2002\)130<2791:ESBAME>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2791:ESBAME>2.0.CO;2).
- Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535, <https://doi.org/10.1175/2010MWR3570.1>.
- Moosavi, A., A. Attia, and A. Sandu, 2019: Tuning covariance localization using machine learning. *Proc. 19th Int. Conf. on Computational Science*, Faro, Portugal, IEEE, 199–212.
- Palmer, T., R. Buizza, R. Hagedorn, A. Lawrence, M. Leutbecher, and L. Smith, 2006: Ensemble prediction: A pedagogical perspective. *ECMWF Newsletter*, No. 106, ECMWF, Reading, United Kingdom, 10–17.
- Polkova, I., and Coauthors, 2019: Initialization and ensemble generation for decadal climate predictions: A comparison of different methods. *J. Adv. Model. Earth Syst.*, **11**, 149–172, <https://doi.org/10.1029/2018MS001439>.
- Robert, A., 1969: The integration of a spectral model of the atmosphere by the implicit method. *Proc. WMO/IUGG Symp. on Numerical Weather Prediction*, Tokyo, Japan, Japan Meteorological Society, 19–24.
- Ruiz, J., M. Pulido, and T. Miyoshi, 2013: Estimating model parameters with ensemble-based data assimilation: A review. *J. Meteor. Soc. Japan*, **91**, 79–99, <https://doi.org/10.2151/jmsj.2013-201>.
- Wu, X., 2016: Improving EnKF-based initialization for ENSO prediction using a hybrid adaptive method. *J. Climate*, **29**, 7365–7381, <https://doi.org/10.1175/JCLI-D-16-0062.1>.
- , S. Zhang, Z. Liu, A. Rosati, T. L. Delworth, and Y. Liu, 2012: Impact of geographic-dependent parameter optimization on climate estimation and prediction: Simulation with an intermediate coupled model. *Mon. Wea. Rev.*, **140**, 3956–3971, <https://doi.org/10.1175/MWR-D-11-00298.1>.
- , W. Li, G. Han, S. Zhang, and X. Wang, 2014: A compensatory approach of the fixed localization in EnKF. *Mon. Wea. Rev.*, **142**, 3713–3733, <https://doi.org/10.1175/MWR-D-13-00369.1>.
- , —, —, L. Zhang, C. Shao, C. Sun, and L. Xuan, 2015: An adaptive compensatory approach of the fixed localization in the EnKF. *Mon. Wea. Rev.*, **143**, 4714–4735, <https://doi.org/10.1175/MWR-D-15-0060.1>.
- , G. Han, S. Zhang, and Z. Liu, 2016: A study of the impact of parameter optimization on ENSO predictability with an intermediate coupled model. *Climate Dyn.*, **46**, 711–727, <https://doi.org/10.1007/s00382-015-2608-z>.
- Yang, S.-C., E. Kalnay, and B. R. Hunt, 2012: Handling nonlinearity in an ensemble Kalman filter: Experiments with the three-variable Lorenz model. *Mon. Wea. Rev.*, **140**, 2628–2646, <https://doi.org/10.1175/MWR-D-11-00313.1>.
- Ying, Y., F. Zhang, and J. L. Anderson, 2018: On the selection of localization radius in ensemble filtering for multiscale quasi-geostrophic dynamics. *Mon. Wea. Rev.*, **146**, 543–560, <https://doi.org/10.1175/MWR-D-17-0336.1>.
- Zhang, S., 2011a: Impact of observation-optimized model parameters on decadal predictions: Simulation with a simple pycnocline prediction model. *Geophys. Res. Lett.*, **38**, L02702, <https://doi.org/10.1029/2010GL046133>.
- , 2011b: A study of impacts of coupled model initial shocks and state-parameter optimization on climate predictions using a simple pycnocline prediction model. *J. Climate*, **24**, 6210–6226, <https://doi.org/10.1175/JCLI-D-10-05003.1>.

- , and J. L. Anderson, 2003: Impact of spatially and temporally varying estimates of error covariance on assimilation in a simple atmospheric model. *Tellus*, **55A**, 126–147, <https://doi.org/10.3402/tellusa.v55i2.12087>.
- , M. Harrison, A. T. Wittenberg, A. Rosati, J. L. Anderson, and V. Balaji, 2005: Initialization of an ENSO forecast system using a parallelized ensemble filter. *Mon. Wea. Rev.*, **133**, 3176–3201, <https://doi.org/10.1175/MWR3024.1>.
- , Z. Liu, A. Rosati, and T. Delworth, 2012: A study of enhance parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model. *Tellus*, **64A**, 10963, <https://doi.org/10.3402/tellusa.v64i0.10963>.
- , and Coauthors, 2020: Coupled data assimilation and parameter estimation in coupled ocean-atmosphere models: A review. *Climate Dyn.*, **54**, 5127–5144, <https://doi.org/10.1007/s00382-020-05275-6>.
- Zhang, X., S. Zhang, Z. Liu, X. Wu, and G. Han, 2015: Parameter optimization in an intermediate coupled climate model with biased physics. *J. Climate*, **28**, 1227–1247, <https://doi.org/10.1175/JCLI-D-14-00348.1>.
- Zhao, Y., X. Deng, S. Zhang, Z. Liu, and C. Liu, 2019: Sensitivity determined simultaneous estimation of multiple parameters in coupled models. Part 1: Based on single model component sensitivities. *Climate Dyn.*, **53**, 5349–5373, <https://doi.org/10.1007/s00382-019-04865-3>.
- Zupanski, M., 2016: Data assimilation for coupled modeling systems. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, S. K. Park and L. Xu, Eds., Vol. 3, Springer, 55–70.