Reducing Forecast Errors Due to Model Imperfections Using Ensemble Kalman Filtering

HIROSHI KOYAMA

Graduate School of Environmental Science, Hokkaido University, Sapporo, and Center for Climate System Research, University of Tokyo, Kashiwa, Japan

MASAHIRO WATANABE

Center for Climate System Research, University of Tokyo, Kashiwa, Japan

(Manuscript received 21 May 2009, in final form 18 February 2010)

ABSTRACT

A method is introduced for reducing forecast errors in an extended-range to one-month forecast based on an ensemble Kalman filter (EnKF). The prediction skill in such a forecast is typically affected not only by the accuracy of initial conditions but also by the model imperfections. Hence, to improve the forecast in imperfect models, the framework of EnKF is modified by using a state augmentation method. The method includes an adaptive parameter estimation that optimizes mismatched model parameters and a model ensemble initialized with the perturbed model parameter. The main features are the combined ensemble forecast of the initial condition and the parameter, and the assimilation for time-varying parameters with a theoretical basis.

First, the method is validated in the imperfect Lorenz '96 model constructed by parameterizing the small-scale variable of the perfect model. The results indicate a reduction in the ensemble-mean forecast error and the optimization of the ensemble spread. It is found that the time-dependent parameter estimation contributes to reduce the forecast error with a lead time shorter than one week, whereas the model ensemble is effective for improving a forecast with a longer lead time. Moreover, the parameter assimilation is useful when model imperfections have a longer time scale than the forecast lead time, and the model ensemble appears to be relevant in any time scale. Preliminary results using a low-resolution atmospheric general circulation model that implements this method support some of the above findings.

1. Introduction

It is widely known that forecast errors in numerical weather prediction arise from the chaotic nature of the atmosphere, in which small initial errors necessarily grow in the deterministic chaotic system and eventually result in the failure of the forecast for a longer lead time (Lorenz 1963). To overcome this problem, a probabilistic approach has been introduced into the forecasts on weekly and one-month time scales; that is, a single forecast is replaced by an ensemble forecast initiated from a set of initial perturbations added to the analysis field. The conventional ensemble forecast (called the initial condition ensemble throughout this paper) plays a significant role for more than a decade in the operational forecast.

The theory of initial error growth is based on a perfect model scenario that assumes no error in prediction models. In reality, however, atmospheric general circulation models (AGCMs) can never be perfect. They inevitably contain various errors arising from the discretization of fluid elements, parameterizations for unresolved subgrid-scale phenomena, and approximations in governing equations. To reduce the forecast error caused by these model imperfections, two main approaches can be used: improving the atmospheric model and introducing the model ensemble technique. First, the fundamental improvement of the model itself, which involves improving the parameterization schemes, narrowing the grid intervals, and using more accurate governing equations, is very important; therefore, continuous improvements must be made to the model. Indirect improvements to the model using the parameter estimation

Corresponding author address: Hiroshi Koyama, Center for Climate System Research, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8568, Japan.
E-mail: koyama@ccsr.u-tokyo.ac.jp

DOI: 10.1175/2010MWR3067.1

and by correcting the model bias have been attempted in many previous studies (Dee and da Silva 1998; Baek et al. 2006; Zupanski and Zupanski 2006). For example, Annan et al. (2005) applied a method based on the ensemble Kalman filter (EnKF; Evensen 1994) to estimate the optimal model parameters in the AGCM. However, few studies have estimated the time-varying parameters (Gillijns and De Moor 2007; Kondrashov et al. 2008). For example, Kondrashov et al. (2008) estimated the time-varying parameters of the intermediate coupled model by using an extended Kalman filter (EKF). They suggested in an experiment using actual sea surface temperature observations, that the parameters may switch between values that are associated with two distinct modes of the El Niño–Southern Oscillation (ENSO), namely the delayed oscillator and the westward-propagating modes. This supports the importance of estimating an adequate time-varying parameter.

The model ensemble technique can potentially have an effect similar to that of the initial condition ensemble. Different methods can be used for the model ensemble: an ensemble using different models, an ensemble using different parameterizations (physics ensemble), and a parameter ensemble using different parameters. Fujita et al. (2007) conducted experiments using an AGCM and indicated that the forecast obtained by using a combination of initial condition and physics ensembles was superior to that obtained by using only the initial condition ensemble. Rodwell and Palmer (2007) demonstrated the effectiveness of the parameter ensemble determined using a weight based on the initial tendency of state variables. However, unlike the initial condition ensemble, the model ensemble does not always work well. This is due to the different features among initial condition errors and model errors—that is, whether or not errors arise at only the initial time, and whether or not errors exist in the same phase space as the model variables. Because of these differences, the model ensemble approach is theoretically difficult to implement. In addition, it is essential that the cost effectiveness of reducing the forecast error by introducing the model ensemble in addition to the initial ensemble is superior to that by introducing only the initial ensemble.

In this study, we introduce a method for reducing forecast errors in an extended-range to one-month forecast based on EnKF with the state augmentation method. The state augmentation method, in which the parameter is regarded as a part of the state variable, is often used to estimate unknown model parameters (Cohn 1997; Annan and Hargreaves 2004; Zupanski and Zupanski 2006; Kondrashov et al. 2008; Fertig et al. 2009). Our modifications from the conventional method include the separation of EnKF for the state variable and that for the parameter, and the averaging process of state variables sampled in each EnKF. The main features of our method are that a combined ensemble forecast is performed with respect to the initial condition and parameter, and that the time-varying parameters for the parameter ensemble are objectively determined. In comparison with other model ensembles, the parameter ensemble has advantages—its introduction is feasible, and the difference among ensemble members is caused only by the different parameters. Previous studies focused on the ability of the state-augmented EnKF to estimate parameters as well as to assimilate state variables, whereas there are quite few studies that examined the effectiveness of the state-augmented EnKF in the extended-range to one-month forecast. Therefore, we focus not only on the accuracy of analysis but also on the forecast scores.

In the next section, we describe the theoretical framework of our method. In section 3, we describe the model and the experimental setup. In section 4, using a low-order atmospheric model, we investigate the effectiveness of our method. In section 5, we present a brief description of results obtained using a low-resolution AGCM that implements our method. In section 6, we present the summary and discussion.

2. Methodology

a. Conventional EnKF

In this subsection, we briefly describe the conventional EnKF on which our method is based: the serial ensemble square root filter (serial EnSRF) proposed by Whitaker and Hamill (2002). The time propagation from $t - \Delta t$ to t of the state variable $\mathbf{x}$ (i.e., the ensemble forecast) is defined as

$$\mathbf{x}_n^f(t) = M[\mathbf{x}_n^a(t - \Delta t)], \tag{1}$$

where M is the (nonlinear) model operator and $\mathbf{x}_n^f$ and $\mathbf{x}_n^a (n = 1, \ldots, N)$ represent the forecast and analysis of $\mathbf{x}$ at the nth member, respectively. Hereafter, we ignore the time t for convenience, except in the formulas of the ensemble forecast. By arranging the ensemble perturbation columnwise, we obtain the perturbation matrix for the forecast, $\mathbf{E}^f$:

$$\mathbf{E}^f = \frac{1}{\sqrt{N-1}}(\mathbf{x}_1^f - \overline{\mathbf{x}}^f, \ldots, \mathbf{x}_N^f - \overline{\mathbf{x}}^f), \tag{2}$$

where the ensemble mean $\overline{\mathbf{x}}^f$ is given by

$$\overline{\mathbf{x}}^f = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_n^f. \tag{3}$$

At the bth observation point, the analysis for the state variable is defined as

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{K}_b[\mathbf{y}_b - (\mathbf{H}\overline{\mathbf{x}}^f)_b], \tag{4}$$

where $\mathbf{y}_b$ is the b component of the observation $\mathbf{y}$, and $\mathbf{H}$ is the matrix mapping the state variable to the observation point, which is assumed to be linear. A Kalman gain matrix $\mathbf{K}_b$ is given by

$$\mathbf{K}_b = \mathbf{E}^f(\mathbf{H}\mathbf{E}^f)_b^{\mathrm{T}}[|(\mathbf{H}\mathbf{E}^f)_b|^2 + \mathbf{R}_{b,b}]^{-1}, \tag{5}$$

where $\mathbf{R}_{b,b}$ is the (b, b) component of the covariance matrix of observation error $\mathbf{R}$. Note that $\mathbf{R}$ is diagonal because of the assumption of independent observations introduced in Whitaker and Hamill (2002). The perturbation matrix for the analysis, $\mathbf{E}^a$, is given as follows:

$$\mathbf{E}^a = \mathbf{E}^f - \alpha\mathbf{K}_b(\mathbf{H}\mathbf{E}^f)_b, \tag{6}$$

$$\alpha = \left[1 + \sqrt{\frac{\mathbf{R}_{b,b}}{|(\mathbf{H}\mathbf{E}^f)_b|^2 + \mathbf{R}_{b,b}}}\right]^{-1}. \tag{7}$$

In serial EnSRF, steps (4) to (7) are serially repeated for each observation. Finally, the analysis ensemble $\mathbf{x}_n^a$ for the next forecast is given as

$$(\mathbf{x}_1^a, \ldots, \mathbf{x}_N^a) = (\overline{\mathbf{x}}^a, \ldots, \overline{\mathbf{x}}^a) + \sqrt{N-1}\,\mathbf{E}^a, \tag{8}$$

and then step (8) loops back to (1) with an increment of Δt.

b. State augmentation method

As mentioned in the introduction, our method is based on EnKF with the state augmentation method. The conventional method is briefly described in this subsection. This method estimates unknown model parameters by regarding the model parameter $\mathbf{p}$ as a part of the state variable $\mathbf{x}$. Hence, the augmented state variable $\mathbf{x}^*$ is defined as

$$\mathbf{x}^* = \begin{pmatrix} \mathbf{x} \\ \mathbf{p} \end{pmatrix}. \tag{9}$$

The ensemble forecast for the state variable is modified as

$$\mathbf{x}_n^f(t) = M_n[\mathbf{x}_n^a(t - \Delta t)], \tag{10}$$

where M_n is the model operator with the analysis of the model parameter at the nth member, $\mathbf{p}_n^a(t - \Delta t)$. In previous studies (Zupanski and Zupanski 2006; Baek et al. 2006), the forecast for the parameter $\mathbf{P}_n^f$ is commonly given as

$$\mathbf{p}_n^f(t) = \mathbf{p}_n^a(t - \Delta t). \tag{11}$$

Other equations are given by introducing the augmented state variable for the analysis and forecast ($\mathbf{x}_n^{*f}$ and $\mathbf{x}_n^{*a}$) to (2)–(8). Note that the observation and the covariance matrix of the observation error are not used for the parameter because the parameter is not directly assimilated but instead is estimated by using the observation of the state variable.

c. Our method (pEnKF)

1) OUTLINE

We aim at performing both the model parameter estimation and the parameter ensemble forecast, as mentioned in the introduction. For this purpose, we construct a modified state-augmented EnKF, called pEnKF for convenience. The scheme and its difference from the conventional method reviewed in section 2b are described in this subsection. In short, the pEnKF consists of two EnKFs—one for the state variable, which is used to assimilate the state variable and generate the initial condition ensemble, and one for the parameter, which is used to estimate the parameter and generate the parameter ensemble. Figure 1 shows a schematic diagram of pEnKF.

2) ENKF FOR THE STATE VARIABLE

This EnKF is the same as the conventional EnKF described in section 2a except for the ensemble forecast and the definition of the state variables ($\mathbf{x}_n^a$ and $\mathbf{x}_n^f$). First, the ensemble forecast (1) is modified as

$$\mathbf{x}_{n,m}^f(t) = M_m[\mathbf{x}_{n,m}^a(t - \Delta t)], \tag{12}$$

where M_m is the model operator, with the model parameter $\mathbf{p}_m^a$ determined by the after mentioned EnKF for the parameter. The variables $\mathbf{x}_{n,m}^f$ and $\mathbf{x}_{n,m}^a$ ($n = 1, \ldots, N$; $m = 1, \ldots, M$) represent the forecast and analysis of $\mathbf{x}$ at the nth initial condition member and the mth parameter member, respectively. Second, $\mathbf{x}_n^f$ in (2) is replaced by

$$\hat{\mathbf{x}}_n^f = \frac{1}{M}\sum_{m=1}^{M}\mathbf{x}_{n,m}^f. \tag{13}$$

Equations (2)–(8) in which $\mathbf{x}_n^f$ is replaced with $\hat{\mathbf{x}}_n^f$ result in the analysis ensemble, $\hat{\mathbf{x}}_n^a$. Then the analysis ensemble for the next forecast, $\mathbf{x}_{n,m}^a$, is simply given as $\mathbf{x}_{n,m}^a = \hat{\mathbf{x}}_n^a$; that is, individual members differ among initial condition members but are the same among parameter members. Note that the ensemble forecasts can spread on multiple attractors unlike the conventional EnKF because of different M_m for each m, and (13) means the average over different attractors; this is discussed in section 6.

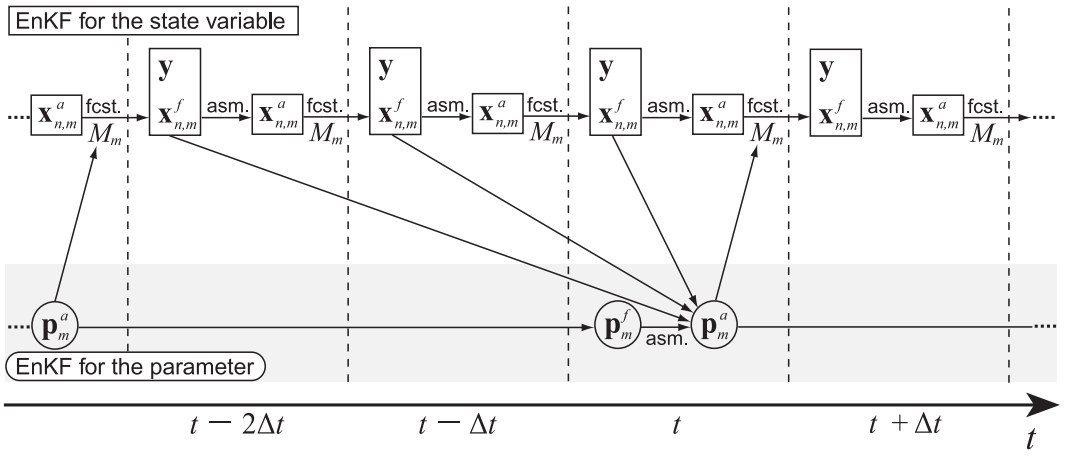

FIG. 1. A schematic diagram describing the procedure of pEnKF ($K = 3$). For the notation, see text.

3) ENKF FOR THE PARAMETER

This EnKF is performed every K cycles (the *time window* is defined as $\tau \equiv K\Delta t$) of the EnKF for the state variable (lower part of Fig. 1). The ensemble forecasts that have previously been calculated by EnKF for the state variable are used. This EnKF is also based on the augmented EnKF described in section 2b. Hence, we describe differences from the conventional augmented EnKF.

First, the augmented state variable for the forecast in section 2b, $\mathbf{x}_n^{*f}$, is replaced by

$$\check{\mathbf{x}}_m^{*f} = \left\{ \begin{array}{c} \dfrac{1}{K}\displaystyle\sum_{k=0}^{K-1}\left[\dfrac{1}{N}\sum_{n=1}^{N}\mathbf{x}_{n,m}^{f}(t-k\Delta t)\right] \\ \mathbf{p}_m^f \end{array} \right\}. \tag{14}$$

The ensemble mean in (14) is used to reduce the forecast error due to inaccurate initial conditions and the time average is used to reduce or remove the high-frequency variability of the state variable. Note that n and N in section 2b are replaced by m and M because the initial condition member is replaced by the parameter member.

Second, the observation and covariance matrix of the observation error are modified as

$$\check{\mathbf{y}}_b = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{y}_b(t-k\Delta t), \tag{15}$$

$$\check{\mathbf{R}}_{b,b} = \frac{1}{K}\mathbf{R}_{b,b}. \tag{16}$$

This modification accompanies the time averaging of the state variable; here, (16) is defined under the assumption that $\mathbf{R}_{b,b}$ behaves like a temporally Gaussian noise.

Third, the parameter is estimated by the parameter part of (8) replaced by the augmented state variables:

$$(\mathbf{p}_1^a, \ldots, \mathbf{p}_M^a) = (\overline{\mathbf{p}}^a, \ldots, \overline{\mathbf{p}}^a) + \sqrt{M-1}\,\mathbf{E}_{\mathbf{p}}^{*a}, \tag{17}$$

where $\mathbf{E}_{\mathbf{p}}^{*a}$ represents the parameter part of the augmented perturbation matrix for the analysis, $\mathbf{E}^{*a}$. Note that in this EnKF the analysis and perturbation of the state variable are not stored because they are not used in the next cycle.

We summarize the procedure of pEnKF. In EnKF for the state variable (upper part of Fig. 1), the time evolution M_m of the analysis state variable $\mathbf{x}_{n,m}^a$ yields the forecast $\mathbf{x}_{n,m}^f$, and then the forecast averaged over the parameter ensemble $\hat{\mathbf{x}}_n^f$ is assimilated to the observation $\mathbf{y}$ to obtain the analysis for the next forecast. This cycle of analysis and forecast is repeated as shown in Fig. 1. By contrast, in EnKF for the parameter (lower part of Fig. 1), which is performed every K cycles of EnKF for the state variable, the parameter is estimated by using the forecast averaged over the initial condition ensemble $\check{\mathbf{x}}_m^{*f}$ and observation of state variables $\check{\mathbf{y}}$. The mth parameter associated with M_m is then dynamically updated by using the estimated parameter $\mathbf{p}_m^a$ (upward arrow in Fig. 1). The essential difference of pEnKF from the conventional augmented EnKF in section 2b is the two-level nested EnKF, which necessarily requires the averaging process in (13) and (14). In other words, the assimilation processes for the state variable and for the parameter are only partially coupled in pEnKF. The ensemble size with respect to the initial perturbation of the state is reduced in pEnKF if the total ensemble size is prescribed; however, this shortcoming is overcome by advantages arising from the two-level nesting, as demonstrated in section 4c.

4) COVARIANCE INFLATION AND LOCALIZATION

We use a method called *multiplicative covariance inflation*, in which the perturbation matrix is multiplied by a factor to stably operate EnKF (Anderson and Anderson 1999). Hence, we multiply the perturbation matrix for state variables by the factor $\sqrt{1+c_x}$, and the perturbation matrix for parameters by $\sqrt{1+c_p}$. The Kalman gain matrix for the state variable is localized by a fifth-order piecewise rational function of Gaspari and Cohn (1999). This function is tuned by the localization length scale l_c as in Hamill et al. (2001) (e.g., if the distance between the state variable and observation is more than $l_c \times 2\sqrt{10/3}$, the function is set to 0). The component between the state variable and the parameter of the Kalman gain matrix is simply multiplied by a constant factor L_p that adjusts the increment of the parameter to some extent.

3. Model

a. Lorenz '96 model

To demonstrate that pEnKF is effective for improving the forecast, we use the two-level Lorenz '96 model (Lorenz 1996, hereafter L96); this model has been widely used for validating data assimilation and forecast schemes. In the present test, the original model is regarded as a perfect model without model errors, whereas a modified model is also used as an imperfect model with model errors. The original L96 model consists of the slow large-scale variables x_i ($i = 1, \ldots, I$) and the fast small-scale variables $y_{i,j}$ ($i = 1, \ldots, I$; $j = 1, \ldots, J$):

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F - \frac{hc}{b}\sum_{j=1}^{J} y_{i,j}, \quad (18)$$

$$\frac{dy_{i,j}}{dt} = cby_{i,j+1}(y_{i,j-1} - y_{i,j+2}) - cy_{i,j} + \frac{hc}{b}x_i, \quad (19)$$

where h, c, and b are scaling parameters and F is a forcing. This model has terms that mimic advection, damping, forcing, and multiscale coupling. In addition, both x_i and $y_{i,j}$ require a cyclic boundary (i.e., $x_{i-I} = x_i$, $x_{i+I} = x_i$, and $y_{i+I,j} = y_{i,j}$, $y_{i,j-J} = y_{i-1,j}$). Therefore, we may consider this model to be a prototype of an atmospheric model having an extremely simple structure.

The imperfect L96 model is constructed as follows. On the rhs of (18), the last term containing $y_{i,j}$ is approximated by the linear regression of x_i (Smith 2000; Orrell 2003; Roulston and Smith 2003):

$$-\frac{hc}{b}\sum_{j=1}^{J} y_{i,j} \simeq \alpha + \beta x_i. \quad (20)$$

The negative correlation visible in the scatterplot of the lhs of (20) against x_i supports the validity of the above approximation (not shown). The substitution of (20) into (18) yields a single equation:

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F + \alpha + \beta x_i. \quad (21)$$

This approach implies the parameterization of y with respect to x, which is analogous to the parameterization encountered with realistic models such as AGCM; in other words, x is a resolvable variable, whereas y is an unresolved subgrid-scale variable.

The experimental setup for the L96 model is described. The number of variables is $I = 8$ and $J = 4$, for which the perfect and imperfect models have 40 and 8 dimensions, respectively. We set $F = 10$ and $h = 1$, which are kept constant, as in other studies (Smith 2000; Orrell 2003). The parameters c and b are time dependent in this study:

$$c = b = 2.5 \sin\left(\frac{2\pi t}{T_p}\right) + 7.5, \quad (22)$$

where T_p is the period of c and b. For constant parameters such as $c = b = 7.5$, the lhs of (20) behaves like white noise, and thus the model imperfections do not have a preferred time scale. This is not favorable for the purpose of this study. On the other hand, using (22), the lhs of (20) has an explicit time scale characterized by T_p. Therefore, we can control the model imperfection such that it has a time scale that is either shorter or longer than that of the extended-range to one-month forecast.

To relate the time scale in the real atmosphere, henceforth we multiply the nondimensional time unit of the above models by a factor of 5. This is because on the basis of the evaluation of the doubling time of the forecast error, one time unit in the L96 model is known to correspond to roughly 5 days in a global weather forecast model. This definition was given with the forcing $F = 8$ by Lorenz and Emanuel (1998). Using the same definition with $F = 10$ in this study causes a more difficult situation for assimilation and forecast because of a faster error growth. This difference, however, does not seriously affect our conclusions. Here, we set $T_p = 90$ days in (22) unless otherwise stated so that the autocorrelation time of the sinusoidal function in (22) is 17 days—that is, the time scale akin to the extended-range to one-month forecast. We also confirm that the solutions behave chaotically with the above conditions. The models are integrated for 30 years using a fourth-order Runge–Kutta scheme using a time step of 0.025 days for x and 0.0025 days for y; the latter is contained only in the perfect model.

The 30-year-long ''true'' states, denoted as x_i^{true}, are obtained by the single time integration of the perfect model. Using them, ''observations'' are then generated by adding a Gaussian random noise with a variance of 0.4 for x and 0.04 for y (only in the perfect model). The diagonal element of the covariance matrix of the observation error is $\mathbf{R}_{b,b} = (0.4)^2$ for x and $\mathbf{R}_{b,b} = (0.04)^2$ for y. These values are approximately 10% of the standard deviation of the natural variability, which is 4.1 for x and 0.37 for y. Observations are assimilated at all the ith points, and in a time interval of 6 h. For each experiment, the covariance inflation factor c_x is tuned to minimize the analysis error; we find that this tuning leads to a nearly minimal forecast error. The localization length scale is $l_c = 1.8$. In contrast, the covariance inflation factor for the parameter c_p is tuned such that it results in the best estimation for the parameter. The localization factor for the parameter is $L_p = 1.0$ (i.e., nonlocalization) because this value leads to good estimation for the parameter.

b. Forecast scores

We use the score given by averaging the root-mean-square of the ensemble-mean forecast error over all integration time T. This score is simply called RMSE:

$$\text{RMSE} = \sqrt{\frac{1}{T}\int_{t=1}^{T}\frac{1}{I}\sum_{i=1}^{I}[\overline{x}_i(t) - x_i^{\text{true}}(t)]^2\,dt}, \quad (23)$$

where $\overline{x}$ represents the ensemble mean. In addition, the improvement rate of a target RMSE R_t against a reference RMSE R_r is defined by

$$\lambda = \frac{R_r - R_t}{R_r} \times 100(\%). \quad (24)$$

We also use the ensemble spread averaged over the number of variables and all integration time T. This score is simply called spread:

$$\text{spread} = \sqrt{\frac{1}{T}\int_{t=1}^{T}\frac{1}{I}\sum_{i=1}^{I}\left\{\frac{1}{N}\sum_{n=1}^{N}[x_i^{(n)}(t) - \overline{x}_i(t)]^2\right\}dt}, \quad (25)$$

where $x_i^{(n)}$ is the forecast at the ith point and nth member. The ratio of the spread to the RMSE indicates the validity of a probability density distribution (Ziehmann 2000):

$$\gamma = \sqrt{\frac{N+1}{N-1}}\frac{\text{spread}}{\text{RMSE}}. \quad (26)$$

4. Application to the L96 model

a. Initial condition ensemble with conventional EnKF

Using the L96 model, we investigate the forecast skill roughly corresponding to the extended-range to one-month forecast, which implies a forecast within a lead time of six time units (=30 days), as described in section 3. Before introducing pEnKF, the prediction skill of the conventional ensemble forecast is examined by applying the conventional EnKF to the imperfect model. The model parameters are estimated by statistical linear regression—that is, by fitting the lhs of (20) to x during the period of τ days around time t (i.e., from $t - \tau/2$ to $t + \tau/2$). The interval of τ is the time window described in section 2c. The above parameters denoted as $\hat{\alpha}(t,\tau)$ and $\hat{\beta}(t,\tau)$, are thus time dependent. Note that $\hat{\alpha}$ and $\hat{\beta}$ can be obtained only when true states are known. It is evident that a sufficiently large τ leads to the climatology of the parameter. In fact, we estimate the climatological constant parameters $\hat{\alpha}_c$ and $\hat{\beta}_c$ with $\tau_c = 30$ years [i.e., $\hat{\alpha}(t,\tau_c) = \hat{\alpha}_c$ and $\hat{\beta}(t,\tau_c) = \hat{\beta}_c$]. The climatological parameters can be a good choice for improving the forecast when we consider only constant parameters. This is because they satisfy the condition for minimizing a short-term forecast error (Orrell 2003). For convenience, the experiments using the imperfect model with the parameters $(\hat{\alpha}, \hat{\beta})$ and $(\hat{\alpha}_c, \hat{\beta}_c)$ are referred to as IS[1] and IC,[2] respectively. Similarly, the experiment using the perfect model is denoted IP.[3] Table 1 summarizes the experiments.

Figure 2 shows the relationship between the lead time and the two scores of RMSE and λ in (23) and (24) using an ensemble size of 40. The reference RMSE R_r is fixed at the RMSE in IC ($N = 40$) afterward. The predictability limit when the RMSE exceeds the RMSE in the climatological forecast is approximately 21 days in IP and approximately 15 days in IC. The improvement rate λ in IP is approximately 30% for a lead time of up to 5 days, following which it gradually decreases (Fig. 2b). This improvement will come from two facts: the analysis of the state variables assimilated by EnKF is more accurate in the perfect model scenario, and the model having no model errors works to forecast more accurate trajectory.

Figure 3a shows λ in IS, which represents the rate of error reduction by using the statistically estimated set of parameters, $\hat{\alpha}(t,\tau)$ and $\hat{\beta}(t,\tau)$. It is found that the RMSE is reduced for any lead time in a time window shorter than 30 days; in addition, the improvement is greater as

[1] IS: Initial condition ensemble with statistically estimated parameters.
[2] IC: Initial condition ensemble with climatological parameters.
[3] IP: Initial condition ensemble in perfect model.

TABLE 1. List of experiments using the L96 model. Symbols are explained in text.

Name	Ensemble methods	Model	Parameters	
IP	Initial condition	Perfect	Undefined	
IC	Initial condition	Imperfect (parameterized)	$\hat{\alpha}_c$,	$\hat{\beta}_c$
IS	Initial condition	Imperfect (parameterized)	$\hat{\alpha}(t,\tau)$,	$\hat{\beta}(t,\tau)$
IS$^-$	Initial condition	Imperfect (parameterized)	$\hat{\alpha}_c$,	$\hat{\beta}^-(t,\tau)$
IMA	Initial condition and model	Imperfect (parameterized)	$\hat{\alpha}_c$,	$\beta(t,\tau)$
IA	Initial condition	Imperfect (parameterized)	$\hat{\alpha}_c$,	$\beta(t,\tau)$

the time window becomes shorter. This suggests that the small number of samples and the high-frequency variability of the variable $\mathbf{x}$ have an insignificant effect on the degradation of the improvement rate. Moreover, the large λ at the analysis time (solid curve) indicates that the use of $\hat{\alpha}(t,\tau)$ and $\hat{\beta}(t,\tau)$ is especially effective for the reduction of the analysis error. To understand which parameter dominates in improving the forecast in IS, we carry out two additional experiments in which either $\hat{\beta}$ or $\hat{\alpha}$ is replaced with the climatology (i.e., $\hat{\beta}_c$ or $\hat{\alpha}_c$; Figs. 3b and 3c, respectively). Using $\hat{\alpha}$ and $\hat{\beta}_c$, λ is nearly zero (see Fig. 3b), whereas the set of $\hat{\alpha}_c$ and $\hat{\beta}$ results in a score that is nearly identical to that shown in Fig. 3a (see Fig. 3c); therefore, β is shown to be the parameter that affects the forecast score. This is explained by the fact that $\hat{\alpha}$ has a spectrum that is well approximated by white noise, but $\hat{\beta}$ has a spectral peak corresponding to T_p in (22) (not shown). The above results provide the baseline of the imperfect L96 model and they are used to compare the results using pEnKF in the next subsection.

b. Initial condition and model ensembles with pEnKF

By using pEnKF, we perform the combined ensemble forecast of the initial condition and the parameter. Following the results obtained in the previous subsection, α is fixed at $\hat{\alpha}_c$, and β is assimilated/perturbed with pEnKF. The ensemble forecast obtained using the imperfect model with the above setting is referred to as IMA[4] (cf. Table 1). To verify the accuracy of the estimated parameter, we first compare the parameters estimated in IMA and in statistical linear regression; the latter is denoted as $\hat{\beta}^-(t,1)$, which is similar to $\hat{\beta}(t,1)$ in the previous subsection but with the time window during the period from $t-1$ to zero (i.e., past information is used as the time window of pEnKF). We recall that $\hat{\beta}^-$ cannot be obtained unless the equations of the perfect model are known; thus, $\hat{\beta}^-$ is only used to verify pEnKF. The experiment with the parameters $\hat{\alpha}_c$ and $\hat{\beta}^-$ is referred to as IS$^-$ (see Table 1). The parameter $\beta(t,1)$ varies with a period similar to that of $\hat{\beta}^-(t,1)$, which is approximately 90 days (Fig. 4). The correlation coefficient between the two becomes 0.68. Consequently, pEnKF succeeds in the estimation of β.

Before investigating the time-averaged forecast scores, we demonstrate a typical case indicating the impact of pEnKF. Figure 5 shows the forecast and the true time series of x_1. In IC (Fig. 5a), at a lead time shorter than 5 days, all ensemble members (thin solid curves) follow the true trajectory (broken curve); at a lead time longer than 15 days, the ensemble mean (thick solid curve) comes close to the climatology because the individual member fully spreads. Significant differences between IC and IMA are found at a lead time of 8–12 days. Almost all of the ensemble members in IC fail to follow the true trajectory. In contrast, in IMA (Fig. 5b), some members can follow the true orbit, and thus the ensemble mean is closer to the true trajectory that turns to the negative value. Similar results are found in the forecast scores of the instantaneous RMSE and spread using all variables. In Fig. 5c, at a lead time of 8–12 days, the RMSE in IMA is smaller than that in IC; moreover, the spread in IMA is larger than that in IC, on account of which the ratio of the spread to the RMSE is close to unity.

Next, we investigate the time-averaged forecast scores. Figure 6a shows λ in IC ($N = 40$) and IMA ($N = 10$, $M = 4$); the total ensemble size in both experiments is 40. In the data assimilation procedure, a time window of one day is adopted in IMA. By introducing pEnKF, the forecast skill is found to improve at any time; λ is approximately 4% for a lead time up to 10 days, and it gradually decreases afterward (Fig. 6a). Moreover, in Fig. 6b, γ in IMA is close to unity unlike that in IC, indicating that the ensemble with pEnKF has a sufficiently large spread. This is attributed to the model ensemble further spreading the trajectories evolved with initial perturbations (e.g., Fig. 5b).

The improvement of the forecast using pEnKF possibly results from a combination of the following factors:

Factor A: improvement of the forecast model by the estimated parameter,

Factor B: improvement of the analysis accuracy for the state variables, caused by the improvement of the forecast model used for *first guess,* and

[4] IMA: Initial condition and model ensembles with the assimilation parameter.

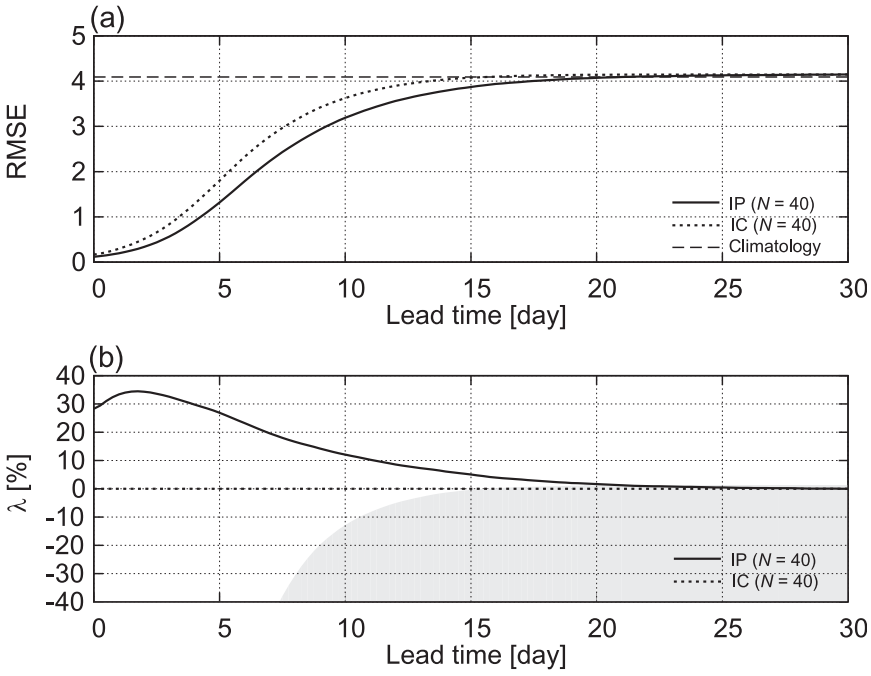

FIG. 2. The relationship between the lead time and the two scores: (a) RMSE and (b) λ. The solid and dotted curves show the score in IP and IC, respectively. The broken line in (a) shows the RMSE in the climatological forecast, and the shading in (b) denotes the area worse than the climatological forecast. The ensemble size is 40, and the reference RMSE R_r is given by the RMSE in IC ($N = 40$).

Factor C: increase of the spread among members with different parameters by introducing the model ensemble method.

To understand the relevance of each factor, an additional experiment is performed with the state variables and an unperturbed but assimilated parameter derived from IMA. This experiment, denoted IA,[5] is thus the initial condition ensemble with $N = 10$ and $M = 1$ (see Table 1). The impact of factors A and B is then evaluated by comparing IA ($N = 10$), IC ($N = 10$), and IS$^-$ ($N = 10$), and that of factor C by comparing IA ($N = 10$) and IMA ($N = 10, M = 4$). Figure 7 shows λ in IMA, IA, IC, and IS$^-$; the curve of λ in IMA is the same as that in Fig. 6a. First, the similarity between the curves of λ in IA (thin solid curve) and in IS$^-$ (dotted curve) represents the success of parameter estimation by pEnKF, as shown in Fig. 4. Note that λ in IA is higher as the lead time becomes short as compared to that in IC (broken curve). This is because of both factors A and B. Second, λ in IMA (thick solid curve) is higher than that in IA for a lead time longer than 6 days. This is explained by introducing the model ensemble (factor C). Consequently, at a lead time shorter than one week, the improvement in both the forecast model and the analysis accuracy is particularly useful for error reduction, whereas the model ensemble method is effective for a lead time longer than one week.

The above results may depend on the time scale of the model imperfections relative to the forecast period. To investigate the relationship between λ and the time scale of the model imperfections, the four experiments shown in Fig. 7 are repeated but with different values of T_p in (22). Figure 8 shows λ for $T_p = 30$ days and $T_p = 360$ days. For the model imperfections having a short time scale (Fig. 8a), λ in IA is still higher than that in IC at a short lead time but is much lower than that in IS$^-$. This indicates that the accuracy of parameter estimation is poor for $T_p = 30$ days. However, λ in IMA is higher than that in IA at a lead time longer than 6 days. Therefore, factor C dominates for the improvement of forecast with short T_p, but factors A and B do not. On the other hand, for the model imperfections having a long time scale (Fig. 8b), λ in IA is much higher than that in IC and even that in IS$^-$ at any lead time. This suggests that the parameter estimation for $T_p = 360$ days is quite successful. The model ensemble also works at a lead time longer than 6 days, as in the other cases. In conclusion, the parameter estimation contributes to reduce forecast errors with a lead time shorter than one week, whereas the model

[5] IA: Initial condition ensemble with the assimilation parameter.

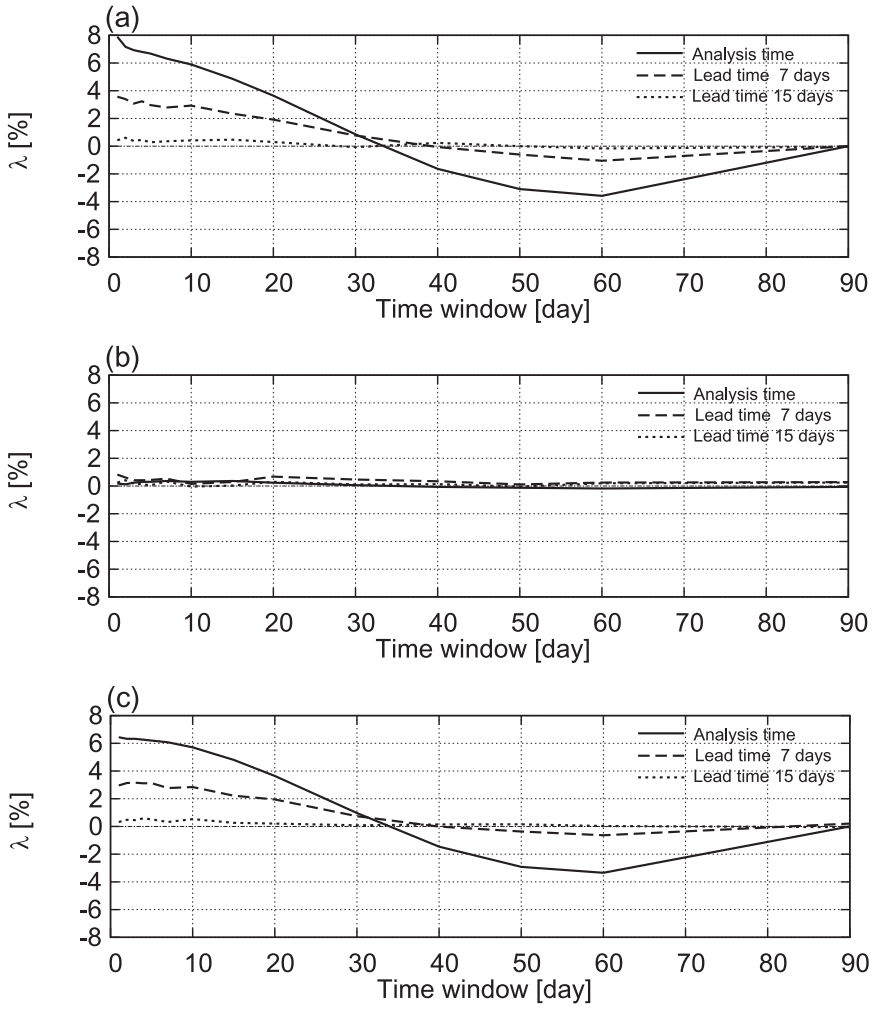

FIG. 3. The relationship between the time window and λ in three experiments: (a) IS, (b) the experiment using $\hat{\alpha}$ and $\hat{\beta}_c$, and (c) the experiment using $\hat{\alpha}_c$ and $\hat{\beta}$. The solid, broken, and dotted curves indicate the analysis time, lead time of 7 days, and lead time of 15 days, respectively. The ensemble size and the reference RMSE are the same as in Fig. 2.

ensemble is effective for improving the forecast with a longer lead time. In particular, the former is relevant when the model imperfections have a longer time scale than the forecast lead time, and the latter appears to hold regardless of T_p.

Finally, we investigate whether pEnKF can be useful for a large observation error. The variance of the Gaussian random noise added to the observation is 1.0 for x, and $\mathbf{R}_{b,b} = 1.0$ (i.e., 250%). It turns out that pEnKF remains effective for reducing forecast errors and optimizing the spread (not shown). However, the maximum λ is approximately 2%, which is smaller than that in the previous experiments. The improvement is mainly caused by factor C as opposed to factors A and B; this result indicates that the large observation error prevents parameter estimation.

c. *Comparison between pEnKF and conventional augmented EnKF*

We test whether pEnKF has advantage in improving forecasts over EnKF, introducing the conventional state augmentation method described in section 2b, simply called the augmented EnKF. Figure 9a shows λ in IMA ($N = 10$, $M = 4$), the augmented EnKF ($N = 40$), and IS$^-$ ($N = 40$); the curve of λ in IMA is the same as that in Fig. 6a. The similarity between the curves of λ in the augmented EnKF (broken curve) and in IS$^-$ (dotted curve) represents the success of parameter estimation by the augmented EnKF. Note that λ in the augmented EnKF is higher than that in IMA (thick solid curve) for a lead time shorter than 6 days. This is because the ensemble size for the state variable in the augmented

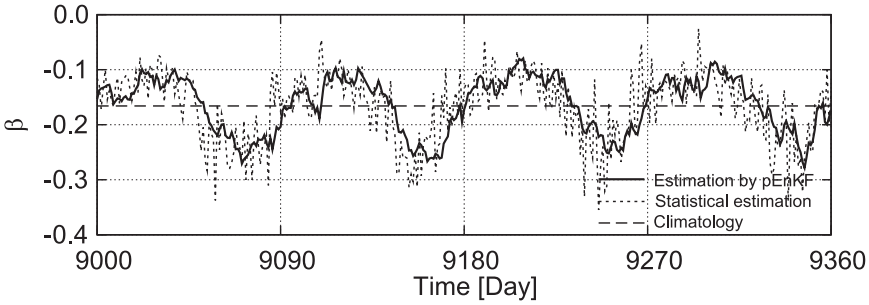

FIG. 4. Time series of the parameter β from 9000 to 9360 days. The thick solid, thin dotted, and thin broken lines respectively indicate $\beta^{-}(t, 1)$ estimated in IMA ($N = 10$, $M = 4$), $\beta(t, 1)$ estimated in the statistical linear regression, and β_c.

EnKF is larger than that in IMA (i.e., $N = 40$ in the augmented EnKF and $N = 10$ in IMA). On the other hand, λ in IMA is high for a lead time longer than one week. This is caused by the model ensemble (factor C) included in pEnKF.

The ratio of spread to RMSE, γ, in the above experiments is shown in Fig. 9b. The augmented EnKF is less useful for optimizing the spread, unlike IMA. This means that the trajectories of the model ensemble by different parameters do not spread, unlike pEnKF. Moreover, in the augmented EnKF, if we increase the spread of parameters (by increasing c_p in section 2c) to enhance the effectiveness of the model ensemble, λ decreases considerably. The large spread among parameters appears to disturb the analysis for state variables. This problem does not occur in pEnKF because of the separation of EnKF for the state variable and that for the parameter, and also by the averaging process over parameter members in EnKF for the state variable (cf. section 2c). Consequently, under the condition that the total ensemble size is prescribed, the conventional augmented EnKF is useful for improving forecasts with a lead time shorter than one week by estimating parameters, whereas pEnKF is effective for improving forecasts with a lead time longer than one week because of the model ensemble.

d. Other features of pEnKF

Since the model and parameterization, (20) and (21), are specific, further experiments are performed to demonstrate that pEnKF has generality in improving the forecast. First, the dependence of λ on the size of the parameter ensemble in IMA is investigated (Fig. 10a). We change the number of parameter members M with fixed total ensemble size at $N \times M = 40$. Note that IC ($N = 40$) is performed for $N = 40$ and $M = 1$ because parameter assimilation is not possible. For $2 \leq M \leq 5$ ($8 \leq N \leq 20$), pEnKF is shown to work effectively because λ is higher than that for $M = 1$ at any lead time (Fig. 10a). In particular, λ for the set of $N = 10$ and $M = 4$ is better than that for other values. For $M > 5$ ($N < 8$), λ remains positive at a lead time of 15 days but is almost zero or even negative at a short lead time. This is because the ensemble size of the state variables, N, is too small to improve the analysis accuracy. If we adopt a smaller ensemble size such as $N \times M = 20$, pEnKF has no advantage at a short lead time (not shown). This is also due to the insufficient number of initial members. We also investigated whether pEnKF remains valid when we increase I in (21) from $I = 8$ to $I = 40$. The result shows that pEnKF still has the advantage for improving forecasts, although λ at lead times longer than one week is smaller than that in the previous experiments (figures not shown). Consequently, for a large ensemble size of $N \times M$, which depends both on the practical computational resources and on the degree of the freedom that the model has, pEnKF has a particular advantage in improving the forecast.

The ratio of spread to RMSE, γ, is compared among the four experiments, as shown in Fig. 10b. As expected from Fig. 6b, pEnKF optimizes γ, but at a very short lead time. For example, at a lead time of 7 days, γ in IC (without pEnKF; $N = 40$, $M = 1$) is 0.81, whereas that in IMA (with pEnKF; $N = 10$, $M = 4$) is 0.94. We also investigate the impact of the covariance inflation (cf. section 2) because it can widen the spread. With a 1.5 times factor ($c_x = 0.48$) in IC, γ increases from 1.07 to 1.18 during the analysis (plus sign in Fig. 10b). In contrast, at lead times of 7 and 15 days, λ increases only by 0.07 and 0.02, respectively, which is still lower than λ in IMA ($N = 10$, $M = 4$) (X symbol and asterisk in Fig. 10b, respectively). In addition, λ decreases by 4% at a short lead time and increases only by 0%–1% at a lead time of longer than 6 days (not shown). Hence, at a lead time longer than one week, the increase in the covariance inflation factor is less useful for reducing the RMSE, although it may be slightly useful for optimizing the spread.

We also test the dependence of λ on the time window, τ. It turns out that IMA with shorter τ results in the better forecast (not shown). The result of IS in section 4a

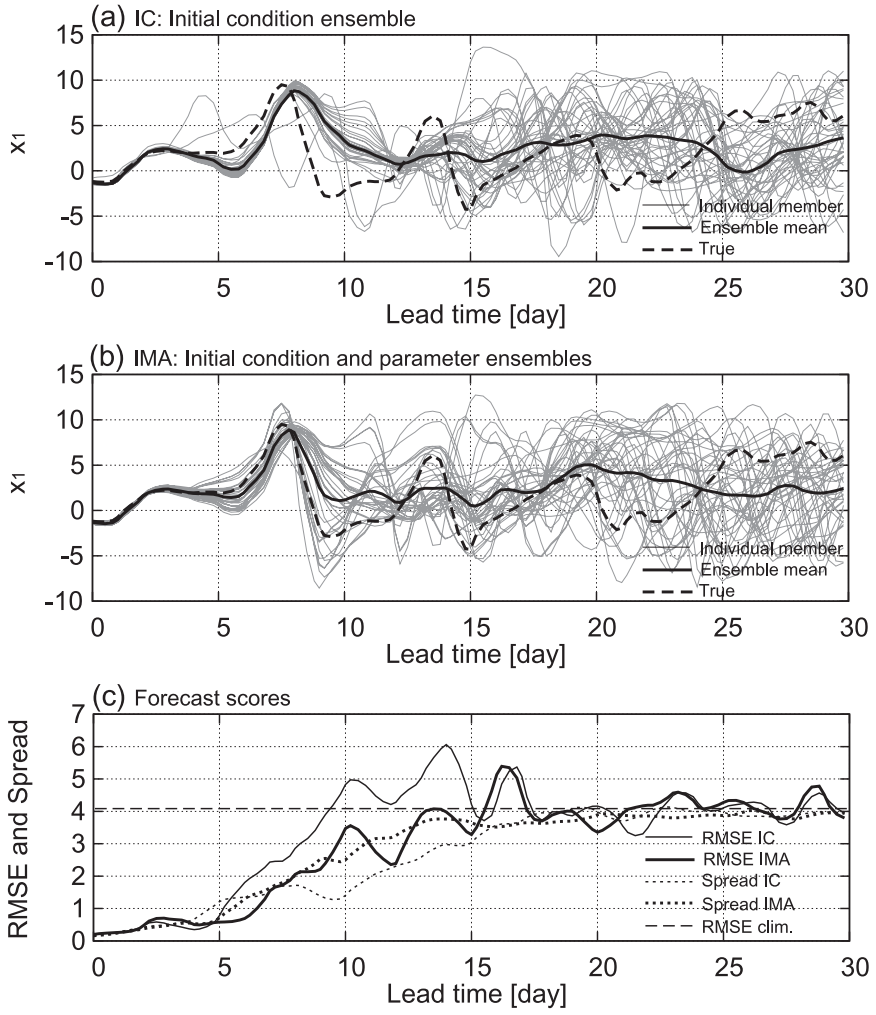

FIG. 5. The forecast and the true time series of x_1 from 6450 to 6480 days in (a) IC ($N = 40$) and (b) IMA ($N = 10$, $M = 4$). The thin solid, thick solid, and thick broken curves indicate the individual member, ensemble mean, and true trajectory, respectively. (c) The instantaneous scores of the RMSE and spread using all variables. The thick and thin solid curves indicate the RMSE in IMA and in IC, and the thick and thin dotted curves indicate the spread in IMA and in IC, respectively. The broken line shows the RMSE in the climatological forecast. The horizontal axis indicates the lead time.

(i.e., that statistically estimated parameters with short τ improve the forecast) supports the above finding; moreover, it suggests that the time window in pEnKF affects the parameter estimation.

5. Application to AGCM

In this section, we present preliminary results of pEnKF applied to an AGCM in order to verify the generality of the results obtained in section 4. A full description of this work will be reported in a forthcoming paper; therefore, in this section we present only a description of the relevance of pEnKF in a complex atmospheric model. The model used is the global spectral model developed at the Center for Climate System Research (CCSR), University of Tokyo, the National Institute for Environmental Studies (NIES), and the Frontier Research Center for Global Change (FRCGC), called the CCSR–NIES–FRCGC AGCM (Hasumi and Emori 2004). We choose a horizontal resolution of T21 and 11 vertical layers.

A single integration by this model but not the real observation is regarded as a true state, as in the experiments using the L96 model. The perfect model is defined by the model in which six parameters in the cumulus convection and cloud schemes (cf. Table 2) are varied by referring to the red noise having a decorrelation time of

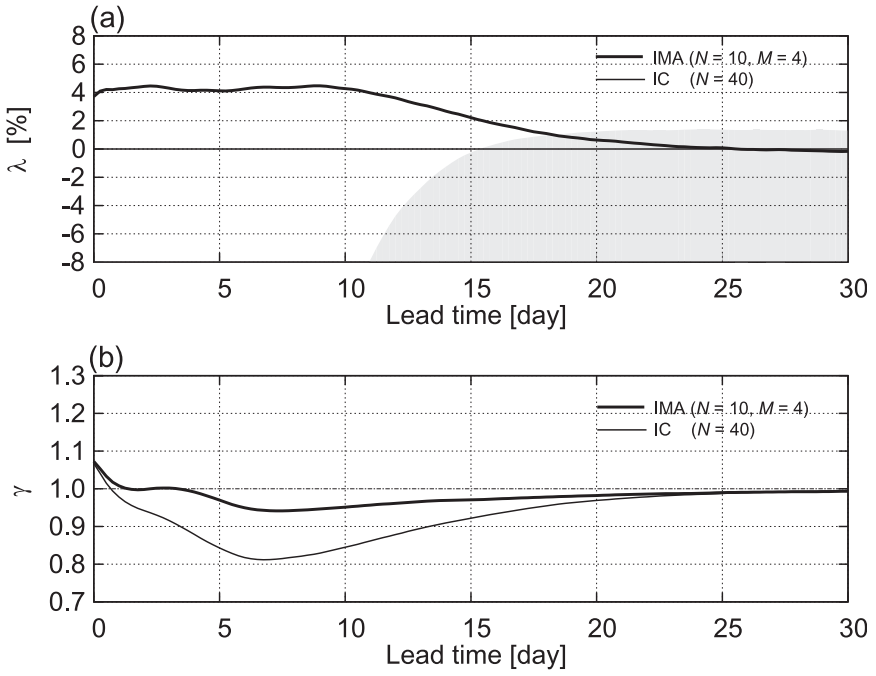

FIG. 6. The relationship between the lead time and the two scores: (a) λ and (b) γ. The thick and thin solid lines indicate the score in IMA ($N = 10$, $M = 4$) and IC ($N = 40$), respectively. Shading is as in Fig. 2b.

approximately three months. These parameters are called true parameters. Note that the parameters are fixed in space. The imperfect model is defined by the same model except that the true parameters are unknown. Observations are generated by adding a Gaussian random noise with a 10% magnitude of climatological standard deviations to the true state. State variables in the AGCM include the zonal and meridional winds, temperature, specific humidity, and surface pressure. They are assimilated by using the imperfect model every 6 h and every three horizontal grids. The standard deviations of the observation errors are of the same magnitude as the abovementioned Gaussian random noise. Moreover, the six parameters are estimated every 2 days. As in the previous section, we compare two experiments of IC ($N = 128$) using the climatological mean values of the parameters and IMA ($N = 32$, $M = 4$) using parameters estimated by pEnKF.

Figure 11 shows the time series of two parameters of dqrat and rhmcrt. For each parameter, the estimated time series in IMA well resembles the true time series except for the initial one month when pEnKF is adapting to the model. To assess the validity of the small parameter ensemble of $M = 4$, we performed additional experiments with fewer parameters, for which we selected two sets of parameters of (dqrat, rhmcrt) and (b_1, preczh).

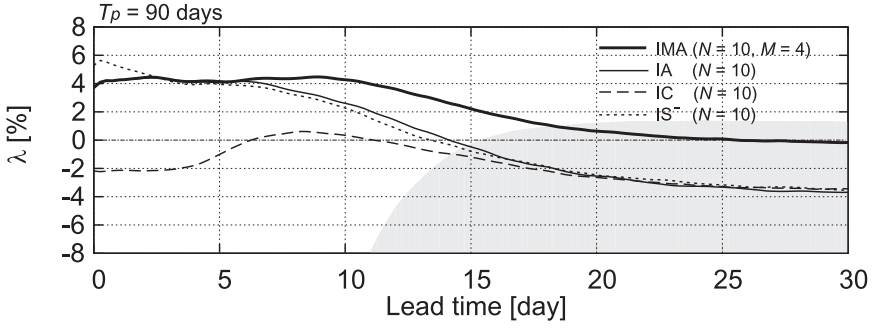

FIG. 7. The relationship between the lead time and λ. The thick solid, thin solid, broken, and dotted curves indicate λ in IMA ($N = 10$, $M = 4$), IA ($N = 10$), IC ($N = 10$), and IS$^-$ ($N = 10$), respectively. Shading is as in Fig. 2b.

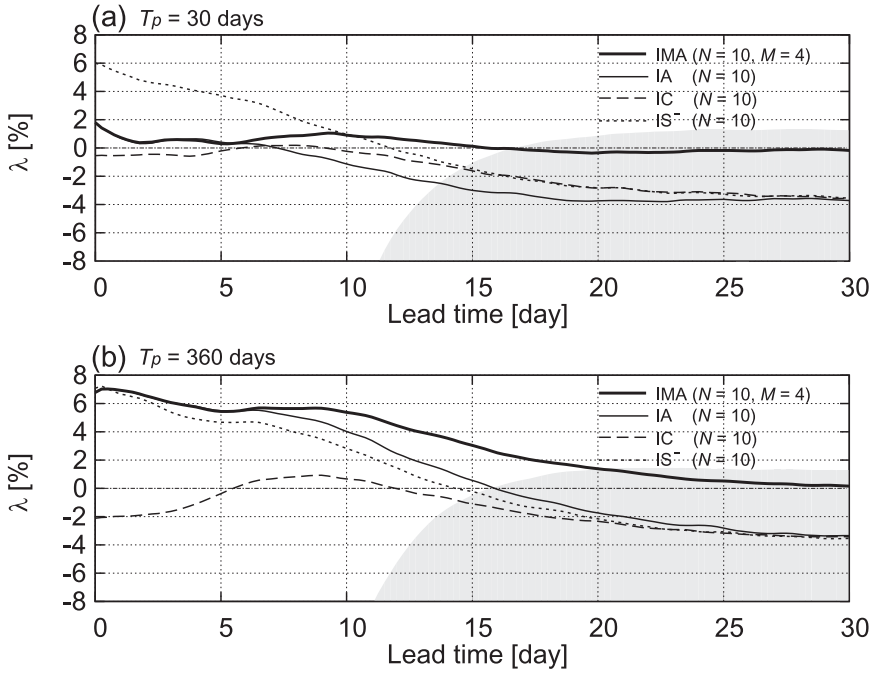

FIG. 8. As in Fig. 7, but for (a) $T_p = 30$ days and (b) $T_p = 360$ days.

These experiments are referred to as EXP1 and EXP2 and compared with IMA. Another experiment (EXP3) was also carried out with a larger M ($M = 16$). The correlation coefficients between the true and estimated parameters (Table 3) show that the parameters dqrat and rhmcrt are well estimated in EXP1 and EXP3, indicating that these parameters are robust. Table 3 also presents the correlation for b_1 and preczh; they are not

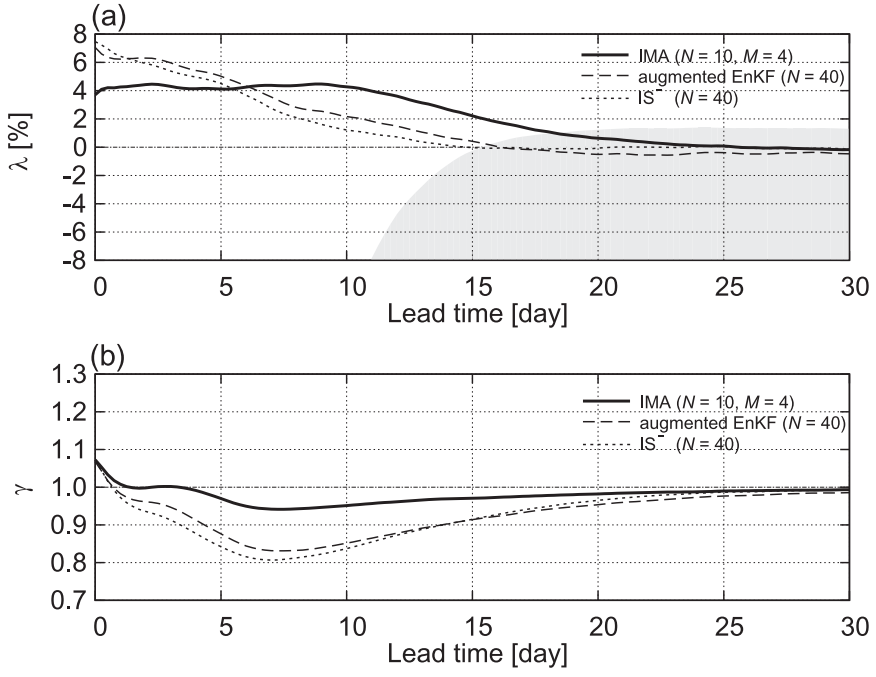

FIG. 9. The relationship between the lead time and the two scores: (a) λ and (b) γ. The thick solid, broken, and dotted curves indicate the score in IMA ($N = 10$, $M = 4$), the conventional augmented EnKF ($N = 40$), and IS^- ($N = 40$), respectively. Shading is as in Fig. 2b.

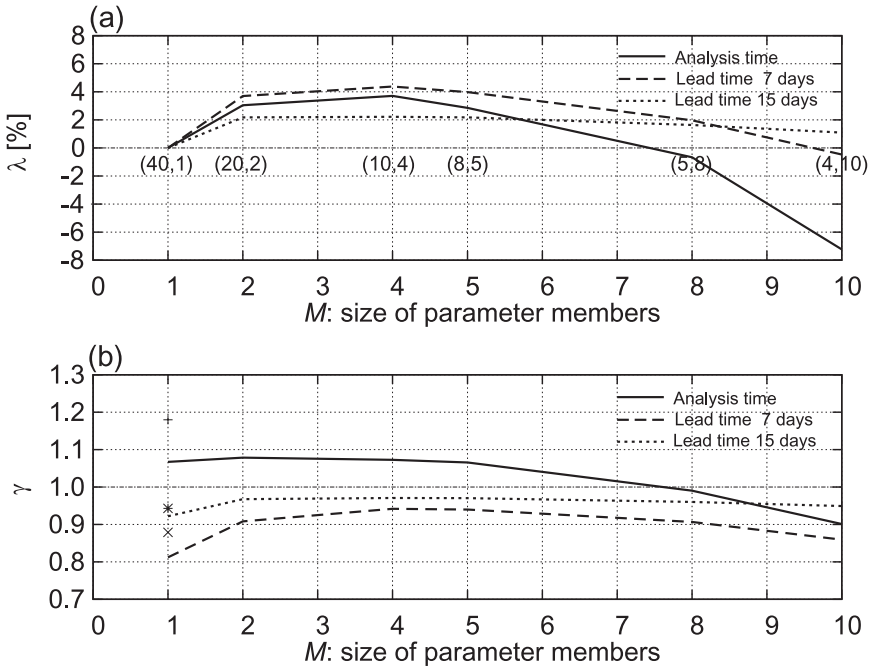

FIG. 10. The relationship between the size of the parameter ensemble and the two scores: (a) λ and (b) γ. The total ensemble size is fixed at $N \times M = 40$. The notation (n, m) in (a) represents the set of n initial members and m parameter members. In both IC and IMA (or in IC using $c_x = 0.48$), the solid (or plus symbol), broken (or $\times$ symbol), and dotted curves (or *) indicate the analysis time, lead time of 7 days, and lead time of 15 days, respectively.

well estimated in IMA but are better estimated when M is increased. One of them (preczh) can be successfully estimated in EXP2 as well. These results indicate that the performance of the parameter estimation in pEnKF depends on the number of parameters and the size of the parameter ensemble. However, our conclusion based on IMA will not be seriously affected because the two parameters playing an important role in improving the forecast do not heavily depend on the choice of M and the number of parameters.

We then performed the one-month forecasts starting from the seven initial dates in December, January, and February. The 500-hPa temperature and the 850-hPa specific humidity, which are averaged over the last 22 days, are shown in Fig. 12a for IC (they are very similar in IMA). The temporal RMS error for these variables in IC and IMA and the difference (latter minus former) are also shown in Fig. 12. It is clearly seen that the temperature error is reduced over the North Pacific, Mediterranean Sea, and tropical Atlantic (contour in Fig. 12d). The forecast error of the specific humidity is also reduced in low-latitude regions (shading in Fig. 12d). These results are obtained because of the improvement in the parameterization due to pEnKF. The parameter ensemble is also useful for improving the 850-hPa specific humidity but is not as effective as that in the L96 experiment. Finally, instead of using the pEnKF with $M = 16$, we performed the one-month forecasts with $M = 4$ but the initial 16-parameter perturbations were prepared by adding the Gaussian random noise to the analysis. This forecast is to confirm whether the small parameter ensemble of $M = 4$ can represent the probability distribution.

TABLE 2. List of parameters selected in the experiments using AGCM.

Abbreviation	Scheme type	Description
b_1	Cloud process	Efficiency for the indirect effect
dqrat	Cloud process	PDF width for the subgrid water
tsice	Cloud process	Critical temperature for cloud liquid–ice partition t_s (if $t \geq t_s$, only water cloud)
twice	Cloud process	Critical temperature for cloud liquid–ice partition T_w (if $T \leq T_w$, only ice cloud)
preczh	Cumulus convection	Vertical profile of precipitation (scale height)
rhmcrt	Cumulus convection	Critical relative humidity

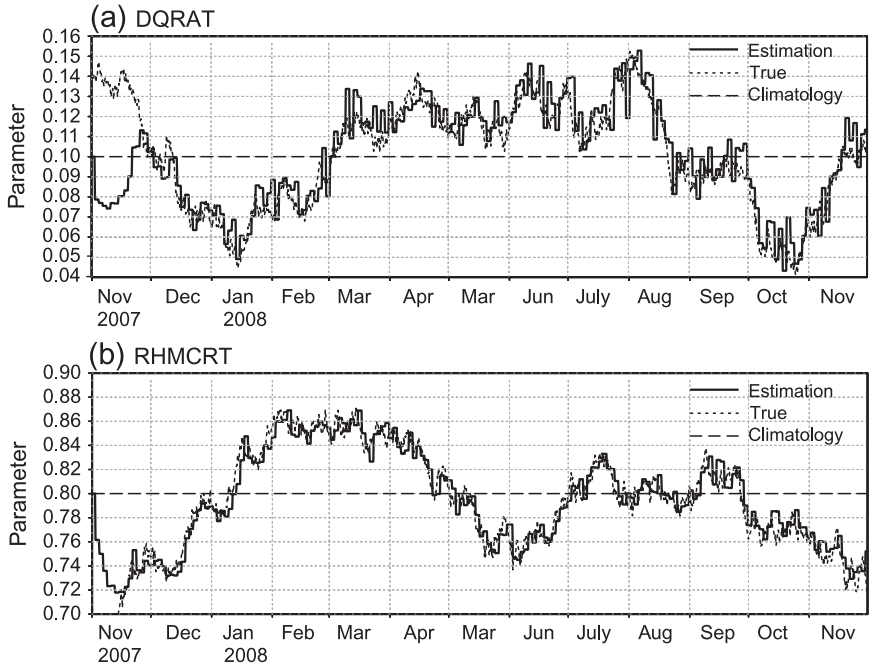

FIG. 11. Time series of the parameters (a) dqrat and (b) rhmcrt from 1 Nov 2007 to 30 Nov 2008. The solid, dotted, and broken lines indicate the parameter estimated in IMA ($N = 32$, $M = 4$), true parameter, and climatological parameter, respectively.

It turns out that the improvement rate of the forecast with this ensemble was quite close to that in IMA. The major reason as to why the small ensemble is sufficient for representing the probability distribution may be the pseudo-observations that are generated by the identical model only by perturbing parameters and thereby prevent the large ensemble spread among different attractors. If we use real observations, the forecast may reveal higher sensitivity to M. This issue is currently ongoing with higher-resolution AGCMs.

6. Summary and discussion

a. Summary

We have introduced a method (pEnKF) based on EnKF with the state augmentation method to reduce the forecast error in the extended-range to one-month forecast. The main features of pEnKF are the combined ensemble forecast of the initial condition and the model parameter, and the adaptive estimation of the time-varying parameter for the parameter ensemble. We first validate pEnKF in the imperfect L96 model constructed by parameterizing the small-scale variable of the perfect model. The results indicate the success of the parameter estimation, the reduction in the ensemble-mean forecast error, and the optimization of the ensemble spread. It is found that the time-dependent parameter estimation contributes to reduce forecast errors with a lead time shorter than one week, whereas the model ensemble is effective for improving forecasts with a longer lead time. The former works well when the model imperfections have a longer time scale than the forecast lead time, while the latter appears to hold in any time scale. Preliminary results using a low-resolution AGCM that implements pEnKF support some of the above findings: the success of the parameter estimation and the reduction in the ensemble-mean forecast error.

b. Discussion

We proposed pEnKF on the basis of the serial EnSRF, which is a particular implementation of EnKF. However, pEnKF can be applied to any type of EnKF. In fact, we found that pEnKF based on the local ensemble

TABLE 3. Correlation coefficient between the true parameter and estimated parameter for the period from 1 Dec 2007 to 31 Mar 2008.

Experiment	IMA	EXP1	EXP2	EXP3
Parameter	All	(dqrat, rhmcrt)	(b_1, preczh)	All
M	4	4	4	16
dqrat	0.92	0.89	—	0.87
rhmcrt	0.98	0.98	—	0.97
b_1	0.36	—	−0.12	0.54
preczh	0.21	—	0.74	0.69

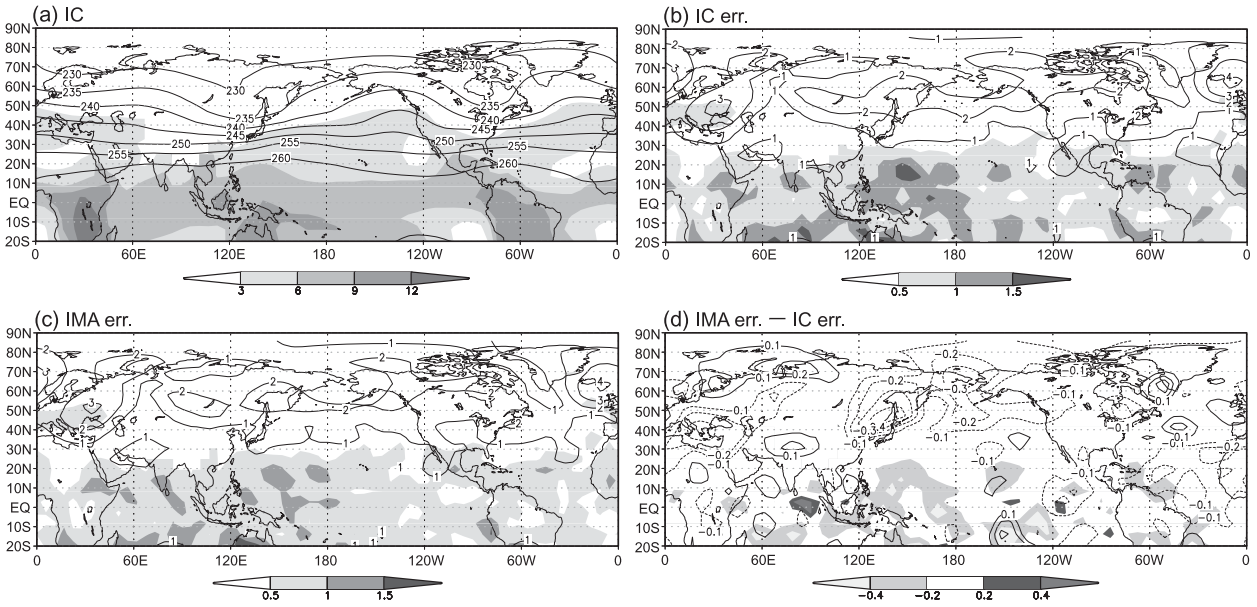

FIG. 12. (a) The 500-hPa temperature (contour, interval 5 K) and the 850-hPa specific humidity (shading, g kg^{-1}). (b) The temporal RMS errors for the temperature (contour, interval 1 K) and specific humidity (shading, g kg^{-1}) in IC ($N = 128$). (c) As in (b), but for IMA ($N = 32$, $M = 4$). (d) Values for (c) minus (b) (contour: interval is 0.1 K, shading: g kg^{-1}). All the quantities are averaged over the last 22 days of the one-month forecasts, starting from the seven initial dates in December–February.

transform Kalman filter (LETKF; Hunt et al. 2007; Miyoshi and Yamane 2007; Miyoshi et al. 2007) also works as expected. The choice of the type of EnKF is mainly important with regard to the computational efficiency rather than the concept of pEnKF.

The combined forecast of the initial condition and model ensembles in pEnKF can settle on different attractors, as described in section 2. This has an advantage and a disadvantage in comparison with only the initial condition ensemble. The advantage is that the optimum forecast spread of the probability density distribution can be constructed by considering possible differences between the nature and model attractors; this can also lead to a reduction in the forecast error. The disadvantage is that the accuracy of analysis of the state variable can be reduced by using ensembles with different attractors unlike the conventional EnKF. In pEnKF, this disadvantage is overcome by the separation of EnKF for the state variable and that for the parameter, and by the average process over parameter members in EnKF for the state variable (13). The above separation also helps the parameter estimation by using the average process over initial condition members in EnKF for the parameter (14). This is because of reducing the difference between forecasts due to inaccurate initial conditions and also extracting the difference between forecasts by using different parameters.

We used a sinusoidal (periodic) model error in the experiments using the L96 model, in which the time scale but not the periodicity is crucial for the behavior of pEnKF. This is ascertained by experiments using a nonperiodic model error, in which the parameters b and c are varied by red noise having a decorrelation time of approximately 17 days (denoted as RN17; roughly $T_p =$ 90 days) and 72 days (RN72; roughly $T_p = 360$ days). They still support the effectiveness of pEnKF, but the improvement rate λ is lower than that in the sinusoidal model error; the maximum value is only $\lambda = 2\%$. This suggests that the model error with a time scale shorter than the forecast lead time, which is due to the red noise, reduces the effectiveness of pEnKF. The impact of the parameter ensemble does not change for either type of model error; in contrast, the parameter estimation is successful in RN72 but fails in RN17. Even if we select a longer time window, the parameter estimation in RN17 is not improved. Hence, we may have to modify pEnKF for such a model error. This could be done in two ways. First, a state variable having a longer lead time could be used in EnKF for the parameter; this may lead to a higher correlation between the state variable and parameter, and a parameter optimized for the forecast of a longer lead time can be estimated. Second, the tendency of the state variable (Rodwell and Palmer 2007) in EnKF for the parameter could be used; this may lead to a higher correlation between the state variable and the parameter.

It is found that in the L96 experiments, the effectiveness of pEnKF does not depend on the initial value and perturbation of parameters. However, the parameter

perturbation tends to maintain the initial one. This may have negatively affected the parameter ensemble in the experiments using a more complex model. Adding a random term may solve the above problem, but it may also lead to other problems (e.g., an increase in sampling errors).

The advantages of pEnKF may not be supported in the following cases: the state variable and the parameter are less correlated and are nonunique, the observation for the state variable is poor, and the phenomena involving the model error cannot be simulated in the model. Some of the above problems may be solved by using a different method instead of the state augmentation method [e.g., the methods proposed by Dee and da Silva (1998) and Gillijns and De Moor (2007)]. We will clarify these problems by AGCM experiments using a realistic observation and will modify pEnKF in a forthcoming paper.

Acknowledgments. The authors thank two anonymous reviewers for their useful comments. The corresponding author (HK) is also grateful to Profs. T. Miyoshi and K. Yamazaki for their stimulating discussions. This work is partially supported by a Grant-in-Aid of the Japanese Ministry of Education, Sports, Culture, Science and Technology.

REFERENCES

Anderson, J. L., 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.

Annan, J. D., and J. C. Hargreaves, 2004: Efficient parameter estimation for a highly chaotic system. *Tellus,* **56A,** 520–526.

——, D. J. Lunt, J. C. Hargreaves, and P. J. Valdes, 2005: Parameter estimation in an atmospheric GCM using the ensemble Kalman filter. *Nonlinear Processes Geophys.,* **12,** 363–371.

Baek, S.-J., B. R. Hunt, E. Kalnay, E. Ott, and I. Szunyogh, 2006: Local ensemble Kalman filtering in the presence of model bias. *Tellus,* **58A,** 293–306.

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75,** 257–288.

Dee, D., and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.,* **124,** 269–295.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

Fertig, E. J., and Coauthors, 2009: Observation bias correction with an ensemble Kalman filter. *Tellus,* **61A,** 210–226.

Fujita, T., D. J. Stensrud, and D. C. Dowell, 2007: Surface data assimilation using an ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.,* **135,** 1846–1868.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Gillijns, S., and B. De Moor, 2007: Model error estimation in ensemble data assimilation. *Nonlinear Processes Geophys.,* **14,** 59–71.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129,** 2776–2790.

Hasumi, H., and S. Emori, 2004: K-1 coupled model (MIROC) description. K-1 Tech. Rep. 1, Center for Climate System Research, University of Tokyo, 34 pp.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D,* **230,** 112–126.

Kondrashov, D., C. J. Sun, and M. Ghil, 2008: Data assimilation for a coupled ocean-atmosphere model. Part II: Parameter estimation. *Mon. Wea. Rev.,* **136,** 5062–5076.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20,** 130–141.

——, 1996: Predictability: A problem partly solved. *Proc. Seminar on Predictability,* Vol. 1, Reading, Berkshire, United Kingdom, ECMWF, 1–18.

——, and K. A. Emanuel, 1998: Optimal sites for supplementary observation sites: Simulation with a small model. *J. Atmos. Sci.,* **55,** 399–414.

Miyoshi, T., and S. Yamane, 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Mon. Wea. Rev.,* **135,** 3841–3861.

——, ——, and T. Enomoto, 2007: Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA,* **3,** 89–92.

Orrell, D., 2003: Model error and predictability over different timescales in the Lorenz '96 systems. *J. Atmos. Sci.,* **60,** 2219–2228.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.,* **133,** 129–146.

Roulston, M. S., and L. A. Smith, 2003: Combining dynamical and statistical ensembles. *Tellus,* **55A,** 16–30.

Smith, L. A., 2000: Disentangling uncertainty and error: On the predictability of nonlinear systems. *Nonlinear Dynamics and Statistics,* A. Mees, Ed., Birkhauser, 31–64.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130,** 1913–1924.

Ziehmann, C., 2000: Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational models. *Tellus,* **52A,** 280–299.

Zupanski, D., and M. Zupanski, 2006: Model error estimation employing ensemble data assimilation approach. *Mon. Wea. Rev.,* **134,** 1337–1354.