

## Research Article

# Development of the Nonstationary Incremental Analysis Update Algorithm for Sequential Data Assimilation System

Yoo-Geun Ham,<sup>1</sup> Hyo-Jong Song,<sup>2</sup> Jaehye Jung,<sup>3</sup> and Gyu-Ho Lim<sup>3</sup>

<sup>1</sup>*Faculty of Earth Systems and Environmental Sciences, Chonnam National University, Gwangju, Republic of Korea*

<sup>2</sup>*Korea Institute of Atmospheric Prediction Systems, Seoul, Republic of Korea*

<sup>3</sup>*Seoul National University, Seoul, Republic of Korea*

Correspondence should be addressed to Hyo-Jong Song; [hj.song@kiaps.org](mailto:hj.song@kiaps.org)

Received 21 March 2016; Revised 22 August 2016; Accepted 10 October 2016

Academic Editor: Takashi Mochizuki

Copyright © 2016 Yoo-Geun Ham et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study introduces a modified version of the incremental analysis updates (IAU), called the nonstationary IAU (NIAU) method, to improve the assimilation accuracy of the IAU while keeping the continuity of the analysis. Similar to the IAU, the NIAU is designed to add analysis increments at every model time step to improve the continuity in the intermittent data assimilation. However, unlike the IAU, the NIAU procedure uses time-evolved forcing using the forward operator as corrections to the model. The solution of the NIAU is superior to that of the forward IAU, of which analysis is performed at the beginning of the time window for adding the IAU forcing, in terms of the accuracy of the analysis field. It is because, in the linear systems, the NIAU solution equals that in an intermittent data assimilation method at the end of the assimilation interval. To have the filtering property in the NIAU, a forward operator to propagate the increment is reconstructed with only dominant singular vectors. An illustration of those advantages of the NIAU is given using the simple 40-variable Lorenz model.

## 1. Introduction

Data assimilation is a mathematically rigorous procedure of combining forecast models with relatively sparse, discontinuous observations to obtain an optimal estimation of the underlying system state. The theory and practice of data assimilation are continuously evolving with several issues still needing to be fully addressed. One of these issues is related to the intermittence of certain assimilation schemes. The initialization problem relates to the discontinuities introduced in the model integration right after the model is restarted from the assimilated fields, which can destroy the intrinsic balances of the model and cause spurious high-frequency oscillations [1, 2].

To prevent these problems, Bloom et al. [2] introduced the incremental analysis updates (IAU). The IAU distributes the analysis increments over a fixed assimilation time window. It allows one to smoothly determine the influence of the observations without the use of an explicit initialization procedure, such as the use of nonlinear normal modes [3]. The IAU

behaves like a low-pass filter and dampens high-frequency oscillations [4]. This property of filtering spurious oscillations is important in significantly improving operational forecasts, because the high-frequency waves dominate most of the initial forecast errors [5]. Therefore, the IAU method has been widely used in many atmospheric and oceanic data assimilation systems [6–8]. Even though most of the atmospheric assimilation system adopted original Bloom's algorithm that the analysis is performed at the middle of the time windows for adding the IAU forcing, this type of the IAU (i.e., centered IAU) has a deficiency to be applied in the complex ocean-atmospheric coupled GCM as the computational time is increased 1.5 times than that in the intermittent method. That is, first half of the time window for adding the IAU forcing should be computed twice to calculate the observational increment and to add the increment. To avoid this time overlapping, several operational data assimilation products implemented the IAU that the analysis is performed at the beginning of the time window for the IAU (i.e., forward IAU) [6, 9, 10].

While there is an advantage in terms of this filtering property, the IAU produces assimilated fields that are less accurate than those of the intermittent assimilation procedure. It represents a suboptimal solution to the assimilation problem (as pointed out by [2]). For example, Ourmières et al. [11] have shown the root-mean-square (RMS) error using the assimilating fields updated by IAU is worse than that obtained by using an intermittent assimilation scheme.

To overcome the deficiency of the IAU, a weighting function, which is a function of time, is multiplied by the constant IAU forcing. For example, Bergemann and Reich [12] applied a hat function to decrease the IAU forcing as the difference between the increment calculation time and the increment-addition time increases. Similarly, an IAU with a weighting function based on a sinusoidal function is comparable to a digital filter that can filter out high-frequency noise without excessive dampening of the slower waves [4]. Even though there is improvement in applying weighting functions to IAU forcing, the selection of the weighting function is arbitrary and problematic, and the inconsistency between the timing for calculating the increment and that of adding the increment still exists.

With this issue in mind, this study introduces a modified version of the IAU, referred to as the nonstationary incremental analysis updates (NIAU) method as a potential alternative to the forward IAU. In NIAU, the forcing term is evolved over time using a forecasting model-derived formulation (see Section 2 for the details). Section 2 introduces the properties of the NIAU. Section 3 illustrates the experimental design, and Section 4 shows the performance of the NIAU in a Lorenz-96 model consisting of 40 variables [13]. Section 5 presents a brief summary.

## 2. Description and Characteristics of the Nonstationary IAU (NIAU) Procedure

**2.1. The IAU and the Nonstationary IAU (NIAU).** The analysis equation for the intermittent assimilation has the form  $\mathbf{x}_0^a = \mathbf{x}_0^b + \Delta\mathbf{x}_0$ , where  $\Delta\mathbf{x}_0 = \mathbf{P}_0^b \mathbf{H}^T [\mathbf{H} \mathbf{P}_0^b \mathbf{H}^T + \mathbf{R}]^{-1} [\mathbf{y} - \mathbf{H}(\mathbf{x}_0^b)]$ . The terms  $\Delta\mathbf{x}_0$ ,  $\mathbf{x}_0^a$ , and  $\mathbf{x}_0^b$  are the analysis increment, the analysis, and the background fields, respectively;  $\mathbf{y}$  is the observation; and  $\mathbf{P}_0^b$ ,  $\mathbf{R}$ , and  $\mathbf{H}$  are the background error covariance at time 0, the observational error covariance, and the observation operator, respectively.

The time differencing scheme to integrate the forecasting model is applied  $N$  times between two consecutive analysis times. Using the observational increment at time 0, the IAU is applied between two analysis times (i.e., between 0 and  $\tau = N\Delta t$ , where  $\Delta t$  is the model time step). Note that this slightly differs from that introduced in Bloom et al. [2], which performs the analysis at the middle of the time windows to add the increment. This type of the IAU is utilized in several operational data assimilation products [6, 9, 10] to avoid an overlapping period before the analysis time as mentioned in Introduction.

Then, the IAU is in the form,

$$\mathbf{x}_n^b = M_{n-1}(\mathbf{x}_{n-1}^b) + \mathbf{f}, \quad n = 1, \dots, N, \quad (1)$$

where  $\mathbf{f} = \Delta\mathbf{x}_0/N$  is the state-independent forcing of the IAU. The subscript  $n$  is defined for times 0 to  $\tau$ . The operator  $M_{n-1}$  corresponds to the nonlinear dynamics integrating from time  $(n-1)\Delta t$  to time  $n\Delta t$ , and  $\mathbf{x}_{n-1}^b$  is the background vector at time  $(n-1)\Delta t$ . Equation (1) implies that constant forcing  $\mathbf{f}$  is added at every time step.

Alternatively, the proposed NIAU procedure obtains the state of time  $n\Delta t$  as

$$\mathbf{x}_n^b = M_{n-1}(\mathbf{x}_{n-1}^b) + \mathbf{M}_{n-1} \cdots \mathbf{M}_1 \mathbf{M}_0 \mathbf{f}, \quad (2)$$

where the matrix  $\mathbf{M}_{n-1}$  is the linear operator of the nonlinear model  $M$  linearized about the state vectors  $\mathbf{x}_{n-1}^b$  (i.e., Tangent Linear Operator, TLM). The main difference between NIAU and IAU is that the NIAU forcing evolves over time by using the TLM per each time step.

The solution of the NIAU at the end of each assimilation cycle is expected to be quite similar to that of the intermittent method. In an intermittent method, the state at the end of the assimilation cycle is

$$\mathbf{x}_N = M_{N-1} \cdots (M_1 (M_0 (\mathbf{x}_0^b + \Delta\mathbf{x}_0))). \quad (3)$$

This can be approximated using a first-order approximation as follows:

$$\begin{aligned} \mathbf{x}_N &\approx M_{N-1} \cdots (M_1 (M_0 (\mathbf{x}_0^b))) + \sum_{n=0}^{N-1} \mathbf{M}_n \Delta\mathbf{x}_0 \\ &= \mathbf{x}_N^b + \sum_{n=0}^{N-1} \mathbf{M}_n \Delta\mathbf{x}_0, \end{aligned} \quad (4)$$

where  $\mathbf{M}_n = (\partial M / \partial \mathbf{x})|_{\mathbf{x}_n}$  is the TLM linearized around  $\mathbf{x}_n$  and  $\mathbf{x}_n^b = M_{n-1} \cdots (M_1 (M_0 (\mathbf{x}_0^b)))$ . Under the perfect model assumption,  $\mathbf{x}_N$  is the optimal solution with all observations before the  $N$ th time step taken into account.

In the NIAU case where the analysis state is denoted with the hat function, the state at the end of the assimilation cycle is obtained from the iterative procedure

$$\begin{aligned} \hat{\mathbf{x}}_1 &= M_0(\mathbf{x}_0^b) + \widehat{\mathbf{M}}_0 \mathbf{f}, \\ \hat{\mathbf{x}}_2 &= M_1(M_0(\mathbf{x}_0^b) + \widehat{\mathbf{M}}_0 \mathbf{f}) + \widehat{\mathbf{M}}_1 \widehat{\mathbf{M}}_0 \mathbf{f} = \hat{\mathbf{x}}_2^b + 2\widehat{\mathbf{M}}_1 \widehat{\mathbf{M}}_0 \mathbf{f}, \end{aligned} \quad (5)$$

where  $\widehat{\mathbf{M}}_n = (\partial M / \partial \mathbf{x})|_{\hat{\mathbf{x}}_n}$  for  $n = 1, \dots, N$ . Hence, the NIAU value at the end of the assimilation cycle is nearly equal to

$$\hat{\mathbf{x}}_N = \mathbf{x}_N^b + \sum_{n=0}^{N-1} N \widehat{\mathbf{M}}_n \mathbf{f} = \mathbf{x}_N^b + \sum_{n=0}^{N-1} \widehat{\mathbf{M}}_n \Delta\mathbf{x}_0. \quad (6)$$

Comparing this equation with (4), the NIAU state propagates in a similar way as the first-order approximation of the propagation of the intermittent analysis.

Analogously, the first two iterations of IAU in which the analysis state is denoted with a tilde give

$$\begin{aligned} \tilde{\mathbf{x}}_2 &= M_1(M_0(\mathbf{x}_0^b) + \widetilde{\mathbf{M}}_0 \mathbf{f}) + \widetilde{\mathbf{M}}_1 \mathbf{f} \\ &= \mathbf{x}_2^b + \widetilde{\mathbf{M}}_1 \widetilde{\mathbf{M}}_0 \mathbf{f} + \widetilde{\mathbf{M}}_1 \mathbf{f}, \end{aligned} \quad (7)$$

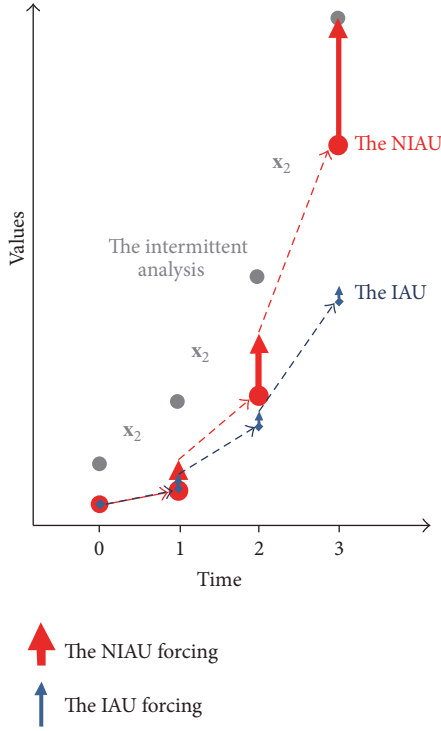


FIGURE 1: The schematic diagram of the IAU (blue), the NIAU (red), and the intermittent method (gray). It is assumed that the total time step ( $N$ ) is 3, and the system is linear.

where  $\widetilde{\mathbf{M}}_1 = (\partial M / \partial \mathbf{x})|_{\bar{\mathbf{x}}_1}$ . At the end of the assimilation cycle, the analyzed state is of the form

$$\bar{\mathbf{x}}_N = \mathbf{x}_N^b + \sum_{n=0}^{N-1} \widetilde{\mathbf{M}}_{N-1} \widetilde{\mathbf{M}}_{N-2} \cdots \widetilde{\mathbf{M}}_n \mathbf{f}, \quad (8)$$

and therefore it is quite different from the solution of the intermittent assimilation.

Figure 1 is a schematic diagram that shows that the solution of the NIAU is equivalent to that of the intermittent method at the end of the assimilation cycle. In the schematics, a simple linear system is assumed that the values are doubled at each time step. Because the IAU forcing is constant, the analysis value of the IAU is far from that of the intermittent method. Conversely, the NIAU forcing is evolved to the same degree as the state vector; therefore, the solution of the NIAU is equivalent to that of the intermittent method at the end of the assimilation cycle.

**2.2. Linear Analysis.** In this subsection, we compare the solution of the NIAU with that of the IAU for linear dynamics. Following Polavarapu et al. [4], we can give the linear system as

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \mathbf{A}\mathbf{x} + \mathbf{f}(t), \\ \mathbf{x}(0) &= \mathbf{x}_0^b, \end{aligned} \quad (9)$$

where  $\mathbf{x}$  is the model state vector,  $\mathbf{A}$  represents the linear dynamics of the system, and  $\mathbf{f}(t)$  is state-dependent forcing. The solution of this linear system is

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}_0^b + \int_0^t e^{\mathbf{A}(t-s)} \mathbf{f}(s) ds. \quad (10)$$

For simplicity, we assume the analysis is performed at time  $t = 0$ , so that

$$\mathbf{x}_0^a = \mathbf{x}_0^b + \Delta \mathbf{x}_0. \quad (11)$$

In an intermittent assimilation scheme, the analysis increment is added entirely at  $t = 0$ ; therefore, the solution at time  $t$  is

$$\mathbf{x}^{\text{INT}}(t) = e^{\mathbf{A}t} \mathbf{x}_0^b + e^{\mathbf{A}t} \Delta \mathbf{x}_0, \quad t \geq 0. \quad (12)$$

The IAU and the NIAU procedures add the increment at each time step. Therefore, their forcing term can be written as a term added gradually from time 0 to time  $\tau$ ,

$$\mathbf{f}(t) = \frac{g(t)}{\bar{g}} \Delta \mathbf{x}_0, \quad (13)$$

$$\bar{g} = \int_0^\tau g(s) ds,$$

where  $g(t)$  is an adequate weighting factor applied between time 0 and time  $\tau$ . The IAU and the NIAU solutions can be written as

$$\begin{aligned} \mathbf{x}^{\text{NIAU}}(t) &= \mathbf{x}^{\text{IAU}}(t) \\ &= e^{\mathbf{A}t} \mathbf{x}_0^b + \left[ \frac{1}{\bar{g}} \int_0^t e^{\mathbf{A}(t-s)} g(s) ds \right] \Delta \mathbf{x}_0. \end{aligned} \quad (14)$$

To examine the filtering responses of the IAU and the NIAU, we proceed as in Bloom et al. [2] and decompose the model dynamics  $\mathbf{A}$  into its eigenmodes; that is, we use the decomposition

$$\mathbf{A}\mathbf{E} = \mathbf{\Lambda}\mathbf{E}, \quad (15)$$

where each diagonal entry of  $\mathbf{\Lambda}$  is the complex eigenvalue  $\lambda_j$  and each column of  $\mathbf{E}$  is the corresponding eigenvector. Under this decomposition, the intermittent solution in (12) becomes

$$\hat{\mathbf{x}}^{\text{INT}}(t) = e^{\lambda_j t} \hat{\mathbf{b}}_j + e^{\lambda_j t} \hat{\mathbf{d}}_j, \quad (16)$$

where  $\hat{\mathbf{x}}^{\text{INT}}(t) = \mathbf{E}^{-1} \mathbf{x}^{\text{INT}}(t)$ ,  $\hat{\mathbf{b}} = \mathbf{E}^{-1} \mathbf{x}_0^b$ , and  $\hat{\mathbf{d}} = \mathbf{E}^{-1} \Delta \mathbf{x}_0$ . We can derive the solution for the IAU or the NIAU in mode space as follows:

$$\hat{\mathbf{x}}^{\text{NIAU}}(t) = \hat{\mathbf{x}}^{\text{IAU}}(t) = e^{\lambda_j t} \hat{\mathbf{b}}_j + e^{\lambda_j t} \gamma(\lambda_j, t) \hat{\mathbf{d}}_j, \quad (17)$$

$$\gamma(\lambda_j, t) = \frac{1}{\bar{g}} \int_0^t e^{-\lambda_j s} g(s) ds, \quad (18)$$

respectively. The only difference between the IAU and the NIAU is the factor,  $g(s)$ . In typical IAU applications, this is

constant, whereas, in the NIAU applications it is the forward operator; that is,

$$g_{\text{IAU}}(t) = 1, \quad (19)$$

$$g_{\text{NIAU}}(t) = e^{\lambda_j t}, \quad (20)$$

respectively. Therefore, the response functions of the IAU and the NIAU are

$$\gamma_{\text{IAU}}(\lambda_j, t) = \frac{1}{g_{\text{IAU}}} \int_0^t e^{-\lambda_j s} ds, \quad (21)$$

$$\gamma_{\text{NIAU}}(\lambda_j, t) = \frac{1}{g_{\text{NIAU}}} \int_0^t ds, \quad (22)$$

respectively. This shows that the response function of the NIAU is independent of the frequency or the growth rate of each mode, implying that the NIAU does not have filtering properties. To overcome this, we provide an alternative method to include the filtering property in the NIAU by modifying the TLM in calculating the increment of the NIAU in Section 2.3.

Finally, the solution of the intermittent assimilation and the NIAU is

$$\begin{aligned} \hat{\mathbf{x}}^{\text{INT}}(t) &= e^{\lambda_j t} \hat{\mathbf{b}}_j + e^{\lambda_j t} \hat{\mathbf{d}}_j, \\ \hat{\mathbf{x}}^{\text{NIAU}}(t) &= e^{\lambda_j t} \hat{\mathbf{b}}_j + \frac{t}{\tau} e^{\lambda_j t} \hat{\mathbf{d}}_j. \end{aligned} \quad (23)$$

Consistently, the NIAU solution at the end of each assimilation period  $\tau$  is equivalent to that of the intermittent assimilation method, which shows at the same time that the NIAU is improved in terms of continuity of the analysis field.

**2.3. Filtering Property in the NIAU.** In this subsection, we will introduce a filtering property to the NIAU while keeping the advantages of its accuracy. The main goal is to filter out the unphysical components (e.g., inertia-gravity waves in the numerical weather prediction) in the NIAU increment, which might ruin the forecast [14]. For this purpose, the TLM to propagate single time step ( $\mathbf{M}_n$ ) in (2) is decomposed using the singular value decomposition (SVD) as follows [15–19]:

$$\mathbf{M}_n = \mathbf{U} \Sigma \mathbf{V}^T. \quad (24)$$

Each column in the matrix  $\mathbf{U}$  ( $\mathbf{V}$ ) denotes the initial (final) singular vector (SV) and the diagonal components in  $\Sigma$  denote the matrix with singular values. Only the dominant SVs are used to reconstruct the TLM, and this reconstructed TLM is utilized to propagate the NIAU increment. This idea is similar to the Tikhonov filter in Hansen [20] or Johnson et al. [21] to dampen the increments projected onto the SVs with relatively small singular values to filter out the small spatial scale features. Note that  $\mathbf{M}_n$  is designed to propagate a single time step in this study, and the SVD is performed at each time step.

Adding increments only projected onto dominant SVs possibly reduces the unwanted artificial noises while keeping the quality of the analysis. Palmer et al. [19] showed that

the evolved dominant SVs (SVs with larger singular values) are projected well onto the forecast errors, while the evolved minor SVs (SVs with smaller singular values) are less projected on the forecast errors. They calculated the projection amplitude of the 48-hour forecast error onto each singular vector and showed that it is in proportion to the singular value itself. This means that the increments projected minor SVs would not be efficient to reduce the forecast errors; therefore, it would not affect the analysis quality significantly.

In addition, SVs with relatively larger spatial scales can be well constrained by the observation [19]. This implies that the increments projected onto the SVs with larger spatial scales are less probable to contain the signal related to the artificial high-frequency waves. On the other hand, SVs with relatively smaller spatial scales are poorly constrained by the observation; it is determined more by the background field [22]. This means that the artificial waves that are not relevant to the observed ones can be projected onto the SVs with smaller spatial scales. We will show that the spatial scales of dominant SVs tend to be larger than that of the minor SVs in Section 4.2, implying increments projected onto dominant SVs are well constrained by the observation to suppress the artificial waves.

One can still argue that dominant singular vectors still include high-frequency variations; therefore, the high-frequency variability is not entirely removed by the SVD method. As the dominant SVs have a physical meaning [23], one may want to keep the high-frequency variability that has physical meaning (e.g., diurnal cycle, severe storm development, and high-frequency oceanic eddies). Therefore, it should be noted that the SVD method in this study focuses on filtering out the unphysical artificial waves, not all the high-frequency variability.

### 3. Application of NIAU to the Lorenz-96 Model

We use the Lorenz-96 model [13, 24] to evaluate the performance of NIAU, whose equation set is

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F, \quad (25)$$

where  $j = 1, \dots, 40$ . The 40 grid points circulate over an arbitrary longitudinal band, with a grid point representing  $9^\circ$  longitudinal line. The external forcing  $F$  is set to 8 and the time step  $\Delta t$  is 0.003125 units. The coefficients of the first (quadratic) and the second (diffusion) terms in (25) are set to be unity for one unit to signify the dissipative decay time, approximately 5 days [13]. Thus the selection of  $\Delta t$  in this study makes a physical time step 0.375 hours.

The true state is obtained from 5 years of model integration. It is sometimes called the Observation System Simulation Experiments (OSSE) [25, 26]. Because the values are obtained from the model integration after a sufficient spin-up period (5 years), it is balanced and constrained by model dynamics. Observations are collected at every grid point by adding uncorrelated random noise from a Gaussian distribution (zero mean and unit variance) to the true state. The main set of experiments discussed below uses observations

taken every 12 hours, even though experiments with different observation frequencies were performed. Using the 2.5-day forecast of the 2000 ensembles, whose initial conditions use a Gaussian distribution around the true value, an initial background error covariance matrix  $\mathbf{P}_0^a$  is constructed. To minimize the impact of the sampling error, we repeated this procedure for 50 times for both analysis and forecast experiments; therefore, the total number of assimilation experiments is 50 for each scheme.

To confirm the differences between the methods that are significant, we performed a significant test using a bootstrap method [27]. That is, among forecast experiments using the IAU, we randomly select the 50 sets of forecast experiments by allowing the overlapping. Then, the averaged RMSE in the selected 50 sets of experiments is calculated. It is repeated 1000 times; then the upper and lower bounds of the two-tailed 95% confidence level are defined as the upper and lower 25th RMSE, respectively. Based on the confidence level, it is checked whether the RMSE in the NIAU is significantly different from that in the IAU.

The sequential methods, such as the 3DVAR, the extended Kalman filter (EKF), and the ensemble Kalman filter (EnKF), all easily exhibit shocks due to the analysis increment. Among the category of sophisticated sequential methods, we select the EKF as a test for a comparison of the IAU and the NIAU. For the NIAU with SVD method, we performed sensitivity test with different numbers of SVs from 31 to 40. That is, (4) is applied for the NIAU, and  $\mathbf{M}_n$  is reconstructed with 31 dominant SVs (i.e., hereafter, NIAU31) to all possible SVs (i.e., 40 SVs) which is equivalent to full  $\mathbf{M}_n$ . Most of the results in this study are from the IAU and the NIAU applied forward in time. This type of approach slightly differs from that introduced in Bloom et al. [2], which performs the analysis at the middle of the time windows to add the increment. The assimilation is performed at every 0 and 12 hr and the analysis increment for the NIAU and the IAU at 0 hr is gradually added from 0 to 12 hr time window.

## 4. Experimental Results

**4.1. The Assimilation Quality of the EKF, the IAU, and the NIAU.** In this subsection, we compare the analysis discontinuity and the accuracy of the EKF (i.e., the intermittent method), the IAU, and the NIAU with full TLM. First, to measure the degree of shock (or jump) in state vectors due to assimilation, we examine the time tendencies associated with each assimilation procedure. Figure 2 shows the root-mean-square (RMS) of the difference in the time tendencies between the assimilated states and the true states for the IAU, the NIAU, and the EKF methods. Note that the time tendency at a target time is defined as the difference between the state at a target time step and the state at a previous time step. Figure 2 shows the averaged time tendency difference in all experiments. This tendency shows how the underlying assimilation procedure induces an abrupt change in the integration as compared to the true state evolution. In all methods, these values are stabilized roughly after day 3, while the states are not optimally adjusted to the observations during the early

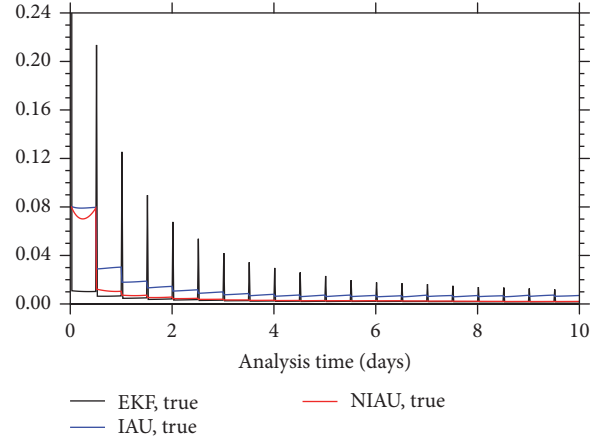


FIGURE 2: Variable-averaged RMS time tendency difference between assimilated variables and true states using the EKF (black line), the IAU (blue line), and the NIAU methods (red line). Note that a 12-hour assimilation interval is applied.

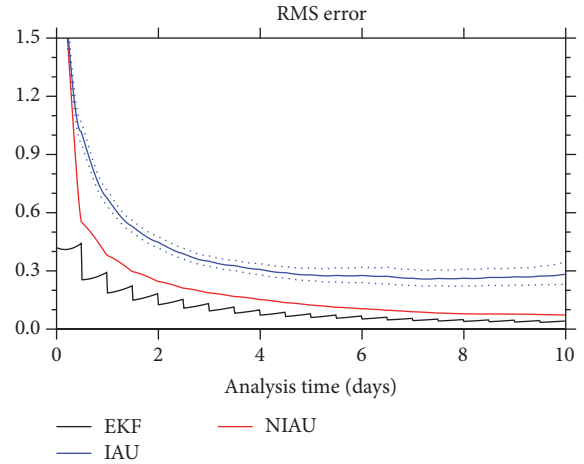


FIGURE 3: Variable-averaged RMS error (RMSE) obtained from the EKF (black line), the IAU (blue line), and the NIAU methods (red line). Note that a 12-hour assimilation interval is applied, and the dotted line denotes the 95% confidence level of the IAU using a bootstrap method.

stages of the assimilation. The EKF results have large discontinuity at every analysis time. In particular, RMS of the time tendency in the EKF is extremely large at the initial time step, while that in the NIAU or the IAU does not show any sudden jumps. This clearly shows the effectiveness of the NIAU or the IAU as a method inserting the analysis increment in the gradual manner. This is expected, because the EKF is designed to add a correction to the state instantaneously. However, the time tendency differences of the IAU and the NIAU methods are small and continuous over time. In particular, the discontinuity in the NIAU method is smaller than that in the IAU method, especially for the early part of the period.

To assess the accuracy of the analysis field in each method, Figure 3 shows the 40-variable-averaged RMS error (RMSE) of the analysis for the three assimilation methods. During

the early period of the assimilation, the RMSE is generally higher in all methods because it starts from an arbitrary initial condition. In the intermittent assimilation, the RMSE is smaller than the others from the beginning, implying that the adjustment to the observation in the intermittent method is relatively fast. This is due to the fact that the increment is added at the earlier time step in the EKF than other schemes. This denotes that the integrated increments in the EKF work to improve the analysis to some extent. That is, the increment at time  $T$  is added at the same time in the EKF; however, that in the NIAU and IAU is slowly added from time  $T$  to  $T + 12$  hours. The short spin-up time is a critical issue in some cases like an assimilation of a severe storm with radar observations, when the analysis field should be quickly adjusted to the radar observation within several analysis cycles, and no previous observation is available [14]. The NIAU still has similar problems; however, the RMSE is significantly smaller than that in the IAU from the beginning with 95% confidence level, which means that the spin-up time of the NIAU is systematically faster than that of the IAU. This is because the NIAU forcing is dynamically evolved, and the NIAU solution adjusts more quickly to the true state than does the IAU solution. The NIAU method has an advantage in reaching a stable solution with less spin-up time than in the IAU.

Beyond the initial adjustment time for the observations, the RMSE values are stabilized for all methods. The RMSE of the EKF is smaller than that of both the NIAU and the IAU. This is expected since the EKF is adding all the increment at the right timing; the analysis value is closer to the true state than that in the IAU or NIAU. It is worthwhile to note that the RMS error in the EKF is better than that in the IAU or NIAU even though adding all the increments in the EKF can cause an initial shock that acts to degrade the quality of the EKF. This implies that time consistency issue (i.e., adding the increment at the right timing) is likely to be much important than the balance issue to keep the accuracy of the analysis. On the other hand, both the NIAU and the IAU are suboptimal assimilation methods that the increment is added with delay in time. Moreover, the RMSE of the NIAU is systematically smaller than that of the IAU. This clearly shows that the NIAU is better than the IAU in terms of the accuracy of the analysis.

The improvement in the accuracy in the NIAU is present regardless of the assimilation interval. To investigate the sensitivity of the accuracy with respect to the assimilation interval, the 10-day average (from day 0 to day 10) RMSEs of assimilated states for the three assimilation methods are reviewed as shown in Figure 4. It is clear that the RMSE of the NIAU is always smaller than that of the IAU method with 95% confidence level, even though it is slightly greater than that of the EKF. The advantage in RMSE becomes larger as the assimilation interval gets longer. Conversely, the superiority of any of the assimilation methods seems to disappear as the frequency of the analysis updates is increased.

#### 4.2. The Use of Singular Vectors (SVs) for the NIAU Increment.

To examine a filtering property in the NIAU using SVs, it is necessary to check the property of the SVs in the TLM. Figure 5 shows the power spectrum analysis of the left and right SVs obtained from the TLM,  $\mathbf{M}_n$ , that evolves a

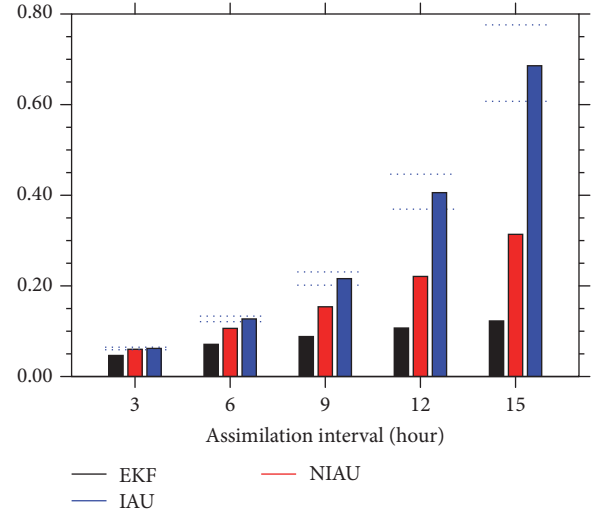


FIGURE 4: The time averaged (from day 0 to day 10) RMSE with respect to the assimilation interval using the EKF (black bar), the NIAU (red bar), and the IAU (blue bar) methods. Note that the dotted line denotes the 95% confidence level of the IAU using a bootstrap method.

perturbation at a time step  $t_n$  to  $t_{n+1}$ . The time averaged power of reconstructed TLMs to propagate single time step is illustrated. Note that the power spectrum analysis is applied in the spatial domain. It is clear that the SVs with larger singular values tend to have higher power at a smaller wavenumber in the spatial domain, while those with smaller singular values have higher power at a larger wavenumber. For example, the most dominant SV has the greatest power around the 8-grid point wavelength (wavenumber 5), while the SV with the smallest singular value (i.e., the 40th SV) has the highest power at less than one 3-grid point wavelength. This implies that the NIAU increments projected onto the dominant SVs would have a relatively lower-wavenumber power (i.e., larger spatial scale), while that projected onto the minor SVs would mainly contribute to the growth of the higher-wavenumber power (i.e., smaller spatial scale). Therefore, the reconstruction of the TLM with dominant SVs (or without minor SVs) would suppress higher-wavenumber features. Assuming that higher-wavenumber motion contributes to the variability of higher frequency in time domain ([28]; refer to Figure 1.1 therein), the calculation of the NIAU increment using the reconstructed TLM without minor SVs would have filtering properties like the IAU. In addition, SVs with relatively lower-wavenumber motions can be well constrained by the observation [19], which implies that the artificial waves that are not relevant to the observed one can be projected onto the SVs with higher-wavenumber motions.

To examine whether the removal of the minor SVs in the TLM contributes to suppressing the high-frequency powers in the temporal domain, the probability distribution function (PDF) of the deviation of the high-frequency (i.e., higher than 1.5 hours) power in the EKF, the IAU, the NIAU, and the NIAU with the reconstructed TLM from that in the true state are shown in Figure 6(a). For the comparison, we add the results

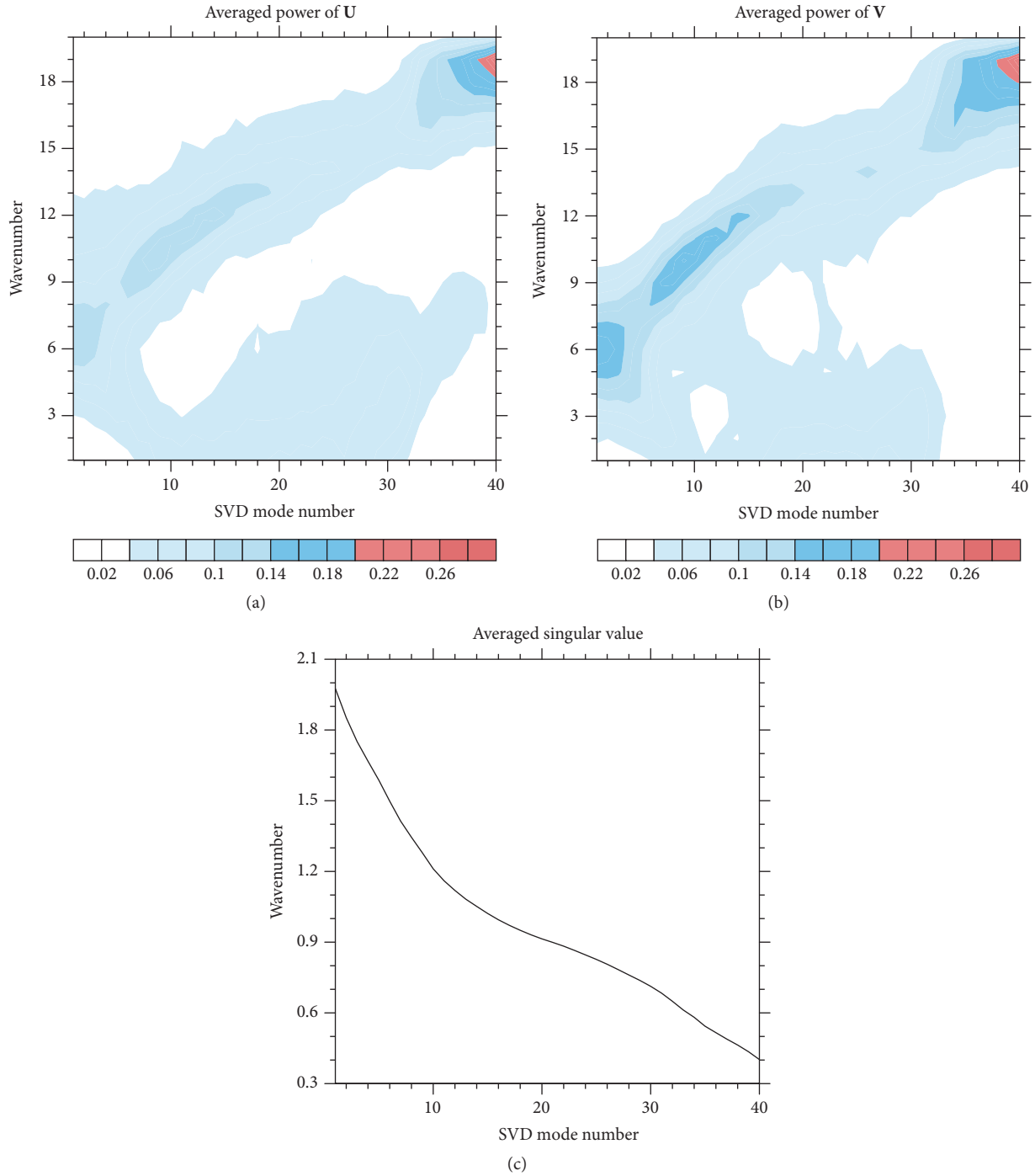


FIGURE 5: The time averaged (from day 0 to day 10) power spectrum analysis in (a) the left and (b) the right singular vectors and (c) the corresponding singular values in the TLM  $\mathbf{M}_n$  that evolves a perturbation at a time step  $t_n$  to  $t_{n+1}$ . Note a singular value on the y-axis right panel is a linear perturbation growth rate when the perturbation is grown linearly for 12 hours.

using the centered IAU, in which the analysis is performed at the middle of the time window for adding the IAU forcing. Note that the NIAU and IAU in previous figures (Figures 2–4) are applied forward in time, so the computational burden in the centered IAU is 1.5 times greater. The deviation has been normalized by the high-frequency power of true state. The lower and upper end of the bar denote 15% and 85% of the

cumulative distribution function, respectively, and a median is denoted as a line crossing the bar. To assess the sensitivity of the number of SVs for the reconstruction of the TLM, a series of experiments using 31–39 dominant SVs are performed. The time-averaged high-frequency power in the forward IAU is systematically suppressed from that in the true state, supporting that the forward IAU has a filtering property. However,

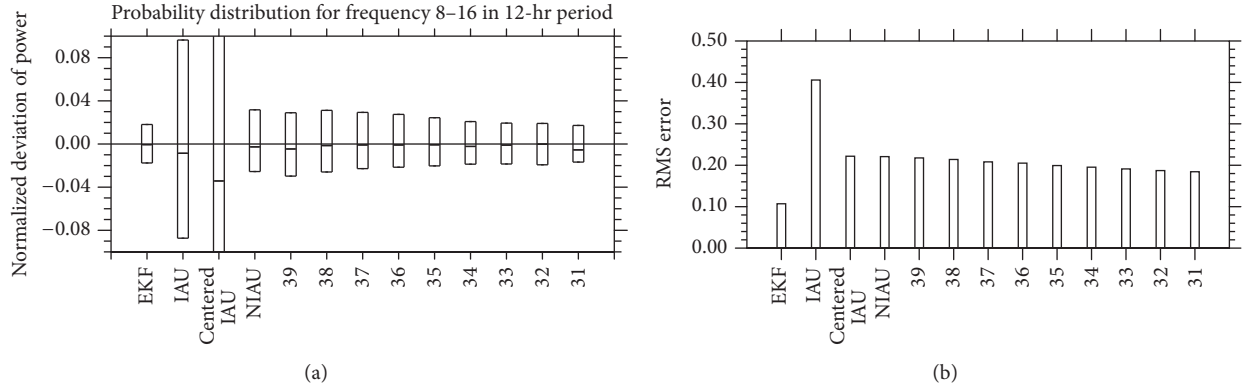


FIGURE 6: (a) Probability distribution function (PDF) of the deviation of the high-frequency (i.e., higher than 1.5 hours) power in each method (the EKF, the IAU, the NIAU, and the NIAU with the reconstructed TLM) from that of the true state, normalized by the high-frequency power of truth. The lower and upper end of the bar denote 15% and 85% of the cumulative distribution function, respectively, with the median denoted as a line crossing the bar. The numbers in the x-axis denote the number of SVs used to reconstruct the TLM for the calculation of the NIAU increment. (b) denotes the time average (from day 0 to day 10) RMSE in the EKF, the IAU, the NIAU, and the NIAU with the reconstructed TLM.

the error bar range is surprisingly wider than other methods. Even though the time-averaged high-frequency power in the centered IAU is further suppressed than the forward IAU, the deviation is extremely large. This means that the IAU generally has filtering properties; however, in an individual case, the high-frequency power of the IAU can be too much suppressed or insufficiently suppressed than that in the true state. This clearly indicates that the filtering property in the IAU is not optimal at least under the experimental setting of this study.

The high-frequency power in the NIAU with full TLM is greater than that in the IAU. However, the range of the error bar is significantly reduced from that of both the forward and the centered IAU, implying that the time-varying high-frequency fluctuation in the true state is relatively well simulated in the NIAU. In addition, the assimilation results in some of the NIAU without minor SVs show that the time-averaged high-frequency power is successfully decreased to some extent. For example, the NIAU with 31 SVs exhibits similar degrees of the time-averaged high-frequency power to that in the IAU, showing that the removal of the minor SVs in the TLM for calculating the NIAU increment successfully suppresses the high-frequency variability. This implies that the NIAU with some modification of the TLM to propagate the NIAU increment can successfully suppress the high-frequency variations. However, note that some of NIAU experiments do not show the suppression of the high-frequency variability. This might imply that the removal of the few (e.g., two or three) lowest SVs may not be enough to suppress the high-frequency variability (for NIAU with 38 SVs or 37 SVs). Or, it is probable that less than 31 SVs would be sufficient to filter out the high-frequency waves. The main point of this section is that the usage of the SVs can be one of the possible ways, but not the detailed strategies for selecting optimal SVs, to filter the high-frequency waves. We leave this point as a future work.

The removal of the minor SVs is also beneficial in terms of the accuracy. Figure 6(b) shows the averaged RMSE in the EKF, the forward IAU, the centered IAU, and the NIAU with the reconstructed TLM. The RMSE in the NIAU is even similar to that in the centered IAU, even though the computational time in the NIAU is significantly less than that in the centered IAU, which shows the benefits of NIAU developed in this study. It is clear that the RMSE in the NIAU gradually decreases as the number of SVs to reconstruct the TLM in the NIAU increment calculation decreases, and the RMSE in the NIAU with 31 SVs is best among all methods except for the intermittent method. The reason for the improvement in the NIAU using the reconstructed TLM can be as follows. First, the suppression of the high-frequency variability would be beneficial on the short-time forecast, giving a background state for analysis that is more accurate. Second, the difference in the state between the analysis and the true state may result in the difference in the basic state for the TLM, and this difference might be mostly projected onto the minor SVs.

**4.3. The Forecasting Experiments.** Next, we evaluate the performance of each method by examining the quality of the forecasts issued from their analysis. Figure 7 shows the 10-day averaged forecast RMSE for each method. Note that a single ensemble member is used to assess the forecast skill of the analysis fields from each assimilation method. While there is a small RMSE within the assimilation period shown in Figure 3, the forecast RMSE is also relatively small with the EKF and the NIAU methods. The similar RMSE in the forecast between the EKF and the NIAU implies that both the accuracy and the balanced state without initial shock are crucial for the skillful forecast. That is, even though there is the relatively huge initial shock in the EKF, the forecast RMSE is similar to the NIAU, which emphasizes the importance of the time consistency in the EKF to guarantee the accuracy of the analysis. On the other hand, even though the NIAU lacks accuracy

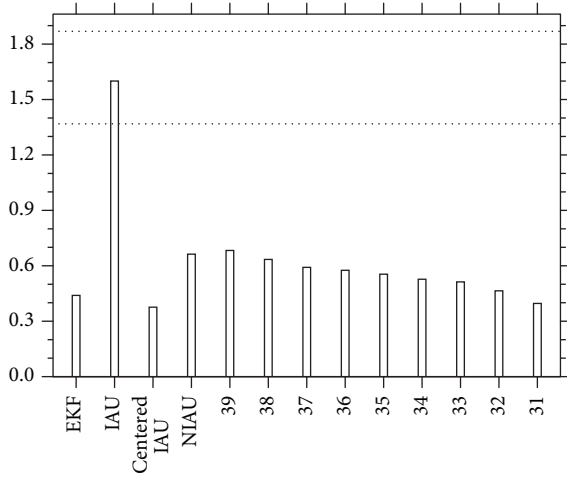


FIGURE 7: The time averaged (from day 0 to day 10) forecast RMS error (RMSE) initialized using the EKF, the IAU (forward), the centered IAU, the NIAU, and the NIAU with the reconstructed TLM. Note that the initial condition is obtained after a 10-day assimilation for each case, which is the same as in the assimilation experiment. The dotted line denotes the 95% confidence level of the forecast RMSE in the IAU (forward).

as the increment is not added in right timing, the analysis with less initial shock acts to improve the forecast skill.

However, the RMSE for the forward IAU method is significantly large compared to the NIAU with 95% confidence level using a bootstrap method. This shows that the NIAU method significantly improves forecasting skill in comparison to the forward IAU method, when keeping the time-evolved forcing method in the analysis. It is interesting that the forecast RMSE tends to be reduced as the number of SVs for the NIAU increment calculation is decreased. For example, the 10-day averaged forecast RMSE in the NIAU is about 0.7, while it is slightly smaller than 0.4 in the NIAU with the TLM reconstructed by 31 SVs (NIAU 31). Even the forecast RMSE of NIAU 31 at day 10 is smaller than that of the EKF (not shown here). This is consistent with the previous result that the NIAU with the reconstructed TLM would be beneficial compared to the NIAU in terms of the accuracy and the removal of the high-frequency power. The forecast RMSE in the centered IAU is similar to that in the NIAU with 31 SVs, and RMSEs in both experiments are even smaller than that in the intermittent assimilation, which emphasizes the benefits of the NIAU or the IAU. It is, therefore, supported that the NIAU with the reconstructed SVs is able to successfully combine both advantages of the high-frequency filtering of the IAU and the analysis accuracy of the intermittent analysis approach.

## 5. Summary

This study introduces a modified version of the incremental analysis updates (IAU), called the nonstationary IAU (NIAU) method, to improve the assimilation accuracy of the forward IAU. The advantages of NIAU come from the theory that the forcing evolves in time by using a forward operator. It is shown that the solution of the NIAU is equivalent to that

of the intermittent assimilation method at the end of each assimilation cycle for a linear system, while the solution of the forward IAU is different. Applying the NIAU to the Lorenz-96 model, we show that its performance is superior to that of the forward IAU in terms of accuracy and discontinuity. For a filtering property in the NIAU, we reconstruct the TLM for the calculation of the NIAU increment using the singular vectors. The reconstructed TLM with dominant SVs suppresses the high-frequency power. In addition, it is found that the analysis quality in terms of the accuracy in the NIAU with the reconstructed TLM is also superior to the NIAU. This implies that the formulation of the forward operator to propagate the analysis increment is important in determining the analysis quality of the NIAU.

The development and maintenance of the TLM in an operational data assimilation system except for four-dimensional data assimilation system are hard to justify, especially for use in only the NIAU. In addition, there is additional computation cost by using the TLM in NIAU. Therefore, we propose some options for obtaining the time-evolving NIAU forcing using ensemble members, which will significantly reduce the costs and difficulties of utilizing the TLM in the NIAU.

One option is that the NIAU forcing is defined as the difference between two realizations as follows:

$$\begin{aligned} & \mathbf{M}_n \mathbf{M}_{n-1} \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{f} \\ & \approx M_n (M_{n-1} \cdots (M_2 (M_1 (\mathbf{x}_0^a + \mathbf{f})))) \\ & - M_n (M_{n-1} \cdots (M_2 (M_1 (\mathbf{x}_0^a)))). \end{aligned} \quad (26)$$

The NIAU forcing is defined as the difference between the integrations with the analysis values and those with the analysis values plus the IAU forcing. By incorporating this method into the ensemble-based data assimilation method, there would be no additional cost of the NIAU beyond that of the IAU or intermittent method.

Another option is to obtain the NIAU forcing using ensemble covariance. The increment ( $\Delta \mathbf{x}_0$ ) can be denoted within the ensemble-based assimilation method as follows:

$$\begin{aligned} \Delta \mathbf{x}_0 &= [\mathbf{x}_0^b] [\mathbf{x}_0^{b'}]^T \mathbf{H}^T [\mathbf{H} [\mathbf{x}_0^b] [\mathbf{x}_0^{b'}]^T \mathbf{H}^T + \mathbf{R}]^{-1} \\ &\cdot [\mathbf{y} - \mathbf{H} (\mathbf{x}_0^b)], \end{aligned} \quad (27)$$

where  $\mathbf{x}_0^b$  contains the column vectors of the background ensemble perturbations. The NIAU forcing can be derived using the ensemble covariance as follows:

$$\begin{aligned} & \mathbf{M}_n \mathbf{M}_{n-1} \cdots \mathbf{M}_2 \mathbf{M}_1 \mathbf{f} \\ & \approx \frac{[\mathbf{x}_n^b] [\mathbf{x}_0^b]^T \mathbf{H}^T [\mathbf{H} [\mathbf{x}_0^b] [\mathbf{x}_0^b]^T \mathbf{H}^T + \mathbf{R}]^{-1} [\mathbf{y} - \mathbf{H} (\mathbf{x}_0^b)]}{N}. \end{aligned} \quad (28)$$

The time-evolved NIAU forcing can be calculated using the ensemble covariance without the TLM, enabling the application of this method to the complex operational assimilation system.

The successful application of the NIAU method to a simple model suggests that the NIAU may systematically improve the quality of the assimilation systems for more complex forecasting models like the general circulation models (GCMs) using real observations. In particular, in a nonlinear system like the GCMs, the NIAU is superior in terms of accuracy and may be more robust, because the NIAU can generate the forcing that considers nonlinear evolution. However, there is also a limitation that the increment for the NIAU is added forward in time in this study. To minimize the difference between the timing for calculating the increment and that of adding the increment, the analysis can be performed in the middle of the time window that adds the increment in original IAU from Bergemann and Reich [12]. This centered IAU is not applicable for the current version of the NIAU as it is hard to propagate the increment backward in time from the center of the time window to the beginning of the time window. Therefore, it should be noted that the current version of the NIAU might be only one of the alternative ways to replace the forward IAU (i.e., do analysis at the beginning of the time window), which it is utilized in several operational data assimilation products [6, 9, 10]. To propagation of the increment backward in time, the NIAU is worthwhile to be incorporated into the ensemble Kalman filter (EnKF) or 4DVAR, which can propagate the perturbation backward in time while keeping the model constraint.

## Competing Interests

The authors declare that there is no conflict of interest regarding the publication of this article.

## Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1A6A1A03012647). Yoo-Geun Ham is supported by the Korean Meteorological Administration Research and Development Program under Grant KMIPA2015-2114. Hyo-Jong Song is supported by the R&D project on the Development of Global Numerical Weather Prediction Systems of Korea Institute of Atmospheric Prediction Systems (KIAPS) funded by Korea Meteorological Administration (KMA).

## References

- [1] W. E. Baker, S. C. Bloom, J. S. Woollen et al., "Experiments with a three-dimensional statistical objective analysis scheme using FGGE data," *Monthly Weather Review*, vol. 115, no. 1, pp. 272–296, 1987.
- [2] S. C. Bloom, L. L. Takacs, A. M. Da Silva, and D. Ledvina, "Data assimilation using incremental analysis updates," *Monthly Weather Review*, vol. 124, no. 6, pp. 1256–1271, 1996.
- [3] R. Daley, "Variational non-linear normal mode initialization," *Tellus*, vol. 30, no. 3, pp. 201–218, 1978.
- [4] S. Polavarapu, S. Ren, A. M. Clayton, D. Sankey, and Y. Rochon, "On the relationship between incremental analysis updating and incremental digital filtering," *Monthly Weather Review*, vol. 132, no. 10, pp. 2495–2502, 2004.
- [5] E. Kalnay and S.-C. Yang, "Accelerating the spin-up of ensemble Kalman filtering," *Quarterly Journal of the Royal Meteorological Society*, vol. 136, no. 651, pp. 1644–1651, 2010.
- [6] M. A. Balmaseda, A. Vidard, and D. L. T. Anderson, "The ECMWF ocean analysis system: ORA-S3," *Monthly Weather Review*, vol. 136, no. 8, pp. 3018–3034, 2008.
- [7] J. Cummings, L. Bertino, P. Brasseur et al., "Ocean data assimilation systems for GODAE," *Oceanography*, vol. 22, no. 3, pp. 96–109, 2009.
- [8] M. M. Rienecker, M. J. Suarez, R. Gelaro et al., "MERRA: NASA's modern-era retrospective analysis for research and applications," *Journal of Climate*, vol. 24, no. 14, pp. 3624–3648, 2011.
- [9] M. A. Balmaseda, K. Mogensen, and A. T. Weaver, "Evaluation of the ECMWF ocean reanalysis system ORAS4," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 674, pp. 1132–1161, 2013.
- [10] C. L. Keppenne, M. M. Rienecker, J. P. Jacob, and R. Kovach, "Error covariance modeling in the GMAO ocean ensemble Kalman filter," *Monthly Weather Review*, vol. 136, no. 8, pp. 2964–2982, 2008.
- [11] Y. Ourmières, J.-M. Brankart, L. Berline, P. Brasseur, and J. Verron, "Incremental analysis update implementation into a sequential ocean data assimilation system," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 12, pp. 1729–1744, 2006.
- [12] K. Bergemann and S. Reich, "A mollified ensemble Kalman filter," *Quarterly Journal of the Royal Meteorological Society*, vol. 136, no. 651, pp. 1636–1643, 2010.
- [13] E. N. Lorenz, "Predictability—a problem partly solved," in *Proceedings of the ECMWF Seminar on Predictability*, Shingfield Park, UK, September 1995.
- [14] E. Kalnay, *Atmospheric Modeling, Data assimilation, and Predictability*, Cambridge University Press, Cambridge, UK, 2003.
- [15] Y.-G. Ham and I.-S. Kang, "Growing-error correction of ensemble Kalman filter using empirical singular vectors," *Quarterly Journal of the Royal Meteorological Society*, vol. 136, no. 653, pp. 2051–2060, 2010.
- [16] R. Kleeman, Y. Tang, and A. M. Moore, "The calculation of climatically relevant singular vectors in the presence of weather noise as applied to the ENSO problem," *Journal of the Atmospheric Sciences*, vol. 60, no. 23, pp. 2856–2868, 2003.
- [17] J.-S. Kug, Y.-G. Ham, M. Kimoto, F.-F. Jin, and I.-S. Kang, "New approach for optimal perturbation method in ensemble climate prediction with empirical singular vector," *Climate Dynamics*, vol. 35, no. 2, pp. 331–340, 2010.
- [18] A. M. Moore, J. Zavala-Garay, Y. Tang et al., "Optimal forcing patterns for coupled models of ENSO," *Journal of Climate*, vol. 19, no. 18, pp. 4683–4699, 2006.
- [19] T. N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza, "Singular vectors, metrics, and adaptive observations," *Journal of the Atmospheric Sciences*, vol. 55, no. 4, pp. 633–653, 1998.
- [20] P. C. Hansen, "The L-curve and its use in the numerical treatment of inverse problems," in *Computational Inverse Problems in Electrocardiology*, P. Johnstone, Ed., pp. 119–142, WIT Press, Southampton, UK, 2001.
- [21] C. Johnson, B. J. Hoskins, and N. K. Nichols, "A singular vector perspective of 4D-Var: filtering and interpolation," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 605, pp. 1–19, 2005.

- [22] A. Hollingsworth, "Objective analysis for numerical weather prediction," in *Proceedings of the Short and Medium Range Numerical Weather Prediction. Collected Papers WMO/IUGG NWP Symposium*, pp. 11–59, Meteorological Society of Japan, Tokyo, Japan, 1987.
- [23] R. Buizza and T. N. Palmer, "The singular-vector structure of the atmospheric global circulation," *Journal of the Atmospheric Sciences*, vol. 52, no. 9, pp. 1434–1456, 1995.
- [24] M. Leutbecher, "Adaptive observations, the Hessian metric and singular vectors," in *Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, Shinfield, UK, September 2003.
- [25] A. Caya, J. Sun, and C. Snyder, "A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation," *Monthly Weather Review*, vol. 133, pp. 3081–3094, 2005.
- [26] M. Xue, M. Tong, and K. K. Droegemeier, "An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 1, pp. 46–66, 2006.
- [27] B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans*, Society for Industrial and Applied Mathematics, 1982.
- [28] R. Daley, *Atmospheric Data Analysis*, Cambridge University Press, Cambridge, UK, 1991.

