Tellus SERIES A DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2009), 61A, 610–620
Printed in Singapore. All rights reserved

TELLUS

An efficient retrospective optimal interpolation algorithm compared with the fixed-lag Kalman smoother by assuming a perfect model

By HYO-JONG SONG and GYU-HO LIM*, *School of Earth and Environmental Sciences, Seoul National University, Seoul 151-747, Republic of Korea*

(Manuscript received 17 October 2008; in final form 8 June 2009)

ABSTRACT

We developed an efficient retrospective optimal interpolation (ROI) algorithm by which we can avoid the overlap of model integration periods, which appears in the procedure of evolving a previous analysis error covariance. If the fixed-lag Kalman smoother (FLKS) is used to determine the analysis state at the beginning of the fixed analysis window as is done for ROI, the FLKS can be considered a suboptimal version of the efficient ROI when the involved model is non-linear and has no errors. We confirm that the efficient ROI analyses are more accurate than the FLKS analyses with the increase of the analysis window size. Nevertheless, the computation costs for implementing efficient ROI are almost the same as those for FLKS. Additionally, the reduced-rank version of the efficient ROI is developed based on the accuracy-saturation property.

From the experiments using Lorenz 3-variable model, it is confirmed that the non-linearity of numerical model, which becomes stronger with the increase of the analysis window size, makes the analysis of efficient ROI more accurate than that of FLKS. Our Lorenz 40-variable experiments show that the average analysis error of the efficient ROI is smaller than that of the FLKS for an analysis window size of up to 4 d. However, the efficient ROI and the FLKS requires almost the same costs for computation. From the results of Lorenz 40-variable model experiments, it is suggested that, by using the reduced-rank formulation of the efficient ROI, we can obtain the suboptimal analysis more cost-effectively rather than FLKS.

1. Introduction

Any future atmospheric state has dynamic and physical relationships with its past atmospheric state. Observations of the future atmospheric state can offer information about the past state of the atmosphere. It was called retrospective analysis by Cohn et al. (1994) to obtain high-quality analysis data by assimilating future observations. Fixed-lag Kalman smoother (FLKS) and four-dimensional variational assimilation (4D-Var) are representatives of the retrospective analysis methods (Zhu et al., 2003; Ruggiero et al., 2006).

Cohn et al. (1994) proposed FLKS, which could produce long-term reanalysis data. The FLKS assimilates an observation into an analysis field corresponding to each past observation time within a moving analysis window. It makes an observation increment by using a Kalman filter analysis and draws the increment back to the past observation time by exploiting the adjoint model integration. Todling et al. (1998) applied the FLKS to a shallow water equation to test its various suboptimal versions. Zhu et al. (2003) implemented a retrospective data assimilation system based on the FLKS in the Goddard Earth Observing System Data Assimilation System.

The FLKS has much higher computing and memory requirements than 4D-Var. Although it requires optimization algorithms for large dimensional applications as well as the adjoint technique, the 4D-Var has been more widely used than the FLKS because of its capability of being implemented with large dimensional models (Rabier et al., 1998). The 4D-Var assumes that the cost function has its minimum value when its gradient is zero (Le Dimet and Talagrand, 1986). When the forecast model has strong non-linearity and the analysis window is too wide, the cost function may have multiple minima (Rabier and Courtier, 1992; Miller et al., 1994). If the cost function has multiple minima, the minimization process may stop after finding a local minimum near the initial state of the process (Pires et al., 1996).

Song et al. (2009) proposed the retrospective optimal interpolation (ROI) to solve the multiple-minima problem. They showed that ROI can be useful in finding the global minimum of a cost

*Corresponding author.
e-mail: gyuholim@snu.ac.kr
DOI: 10.1111/j.1600-0870.2009.00409.x

function although it depends on the analysis window size. They also showed that ROI can be implemented without using an adjoint model by adopting the perturbation method of Errico and Raeder (1999). They confirmed their hypothesis with numerical experiments using Lorenz 3-variable and 40-variable models.

Even though its development is difficult for physical schemes, the adjoint model accurately computes the gradients of cost functions independent of the validity of the linearization, and the computation is efficient. However, using the perturbation method helps us to implement ROI with ease (Zhu and Kamachi, 2000; Errico, 2003). The four-dimensional ensemble Kalman filter (4D-EnKF) proposed by Hunt et al. (2004) also allows us to obtain a retrospective analysis without the adjoint model. The four-dimensional local ensemble transform Kalman filter (4D-LETKF), which is a simplified version of the 4D-EnKF, was compared with 4D-Var using the Lorenz 40-variable model (Lorenz, 1996; Fertig et al., 2007). However, due to model nonlinearity, the 4D-LETKF analysis was shown to become more inaccurate than the 4D-Var analysis as the length of analysis window increases. Therefore, ROI is worth being proposed as a method, which can reduce analysis error due to multiple minima with circumventing the use of adjoint model.

Since ROI is basically an optimal interpolation technique, the error variance for each variable in the analysis state vector decreases after an individual ROI analysis step. This property is called the accuracy-saturation of ROI. Song et al. (2009) devised a reduced-rank formulation for ROI based on the accuracy-saturation property to reduce computation costs for implementing ROI. Using Lorenz 40-variable model, they showed that the reduced-rank formulation of ROI requires progressively less computing time than the 4D-Var for the same quality of analysis as the analysis window size increases.

The ROI analysis estimates an atmospheric field for the initial time of the analysis window every analysis step. ROI uses the previous ROI analysis as the initial condition of the model integration for calculating the guess of the next future observation field. Therefore, the time period, over which a numerical model is integrated to evolve a previous analysis error covariance, overlaps the model integration period for the next analysis step. Since such overlap hampers the efficiency of ROI, we devise an efficient ROI algorithm in which we improve the efficiency of ROI in Section 2. We find that the equations of the efficient ROI are analogous to those of the FLKS, which are described in table 1 of Cohn et al. (1994), and thus compare the efficient ROI with the FLKS in terms of accuracy and computation costs based on a mathematical reasoning. Since the efficient ROI would have the accuracy-saturation property as ROI does, we could devise a reduced-rank formulation for the efficient ROI. It can be implemented also without using tangent linear and adjoint models. The formulation for the reduced-rank version of the efficient ROI is described in Section 3. In the following section, we confirm the theoretical reasoning of Section 2 and examine the efficiency of the reduced-rank formulation described in Section 3 in comparison with the FLKS algorithm by using Lorenz 3-variable and 40-variable models.

2. Comparison of efficient ROI with the fixed-lag Kalman smoother by assuming a perfect model

2.1. *Avoiding the overlap of model integration periods to evolve analysis error covariance matrices in the ROI algorithm*

To solve the multiple-minima problem, Song et al. (2009) devised the ROI, which consists of the ROI analysis vector, the corresponding error covariance matrix, and the optimal weight matrix equations as follows:

$$\mathbf{x}^{\mathrm{roi}}_{0|n} = \mathbf{x}^{\mathrm{roi}}_{0|n-1} + \mathbf{W}^{\mathrm{roi}}_{0|n}\left\{\mathbf{y}^{\mathrm{obs}}_{n} - H_n\left[M_n\left(\mathbf{x}^{\mathrm{roi}}_{0|n-1}\right)\right]\right\}, \tag{1a}$$

$$\mathbf{P}^{\mathrm{roi}}_{0|n} = \left[\mathbf{I} - \mathbf{W}^{\mathrm{roi}}_{0|n}\mathbf{H}_n\mathbf{M}_n\right]\mathbf{P}^{\mathrm{roi}}_{0|n-1}, \tag{1b}$$

$$\mathbf{W}^{\mathrm{roi}}_{0|n} = \mathbf{P}^{\mathrm{roi}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n\mathbf{H}^{\mathrm{T}}_n\left[\mathbf{H}_n\mathbf{M}_n\mathbf{P}^{\mathrm{roi}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n\mathbf{H}^{\mathrm{T}}_n + \mathbf{R}_n\right]^{-1}, \tag{1c}$$

for $n = 0, 1, \ldots, N$. The index n stands for the observation time, and the subscript $0|n$, which is adopted from Cohn et al. (1994), means that the ROI analysis $\mathbf{x}^{\mathrm{roi}}_{0|n}$ results from the assimilation of an observation $\mathbf{y}^{\mathrm{obs}}_{n}$ at a time t_n into a background state at the initial time t_0 of the analysis window, which corresponds to the previous ROI analysis $\mathbf{x}^{\mathrm{roi}}_{0|n-1}$. The state $\mathbf{x}^{\mathrm{roi}}_{0|-1}$ and the error covariance $\mathbf{P}^{\mathrm{roi}}_{0|-1}$ are equal to the background field $\mathbf{x}^{b}_{0}$ of an atmospheric state $\mathbf{x}^{t}_{0}$ and the corresponding error covariance matrix $\mathbf{B}$, respectively. The operator $M_n(\mathbf{x})$ is a forecast model that transforms the initial model state $\mathbf{x}_0$ into a future state $\mathbf{x}_n$ at a time t_n . $M_0(\mathbf{x})$ is equivalent to an identity. An observation operator $H_n(\mathbf{x}_n)$ transforms the model state $\mathbf{x}_n$ into the observation space. The matrices $\mathbf{M}_n$ and $\mathbf{H}_n$ are the first derivatives of $M_n(\mathbf{x})$ and $H_n(\mathbf{x})$ with respect to $\mathbf{x}$. By referring to the optimal interpolation formulation, we can see that ROI is basically an optimal interpolation that assimilates subsequent observations to obtain the variance-minimum analysis (Kalnay, 2003).

The error variance of an analysis field derived by the optimal interpolation from a background field and an observation set is always smaller than that of a background field alone (Daley, 1991; Kalnay, 2003). As ROI proceeds, the error variance of the ROI analysis decreases. Accordingly, we can expect that the ROI process reinforces the validity of the tangent linear approximation at each analysis step between the background and true states,

$$H_n\left[M_n\left(\mathbf{x}^{\mathrm{roi}}_{0|n-1}\right)\right] - H_n\left[M_n(\mathbf{x}^{t}_{0})\right] \approx \mathbf{H}_n\mathbf{M}_n\left[\mathbf{x}^{\mathrm{roi}}_{0|n-1} - \mathbf{x}^{t}_{0}\right]. \tag{2}$$

If eq. (2) is valid, the weight matrix (1c) is optimal and the ROI analysis also becomes optimal with a minimum variance. Figure 1a shows the schematic procedure of ROI.

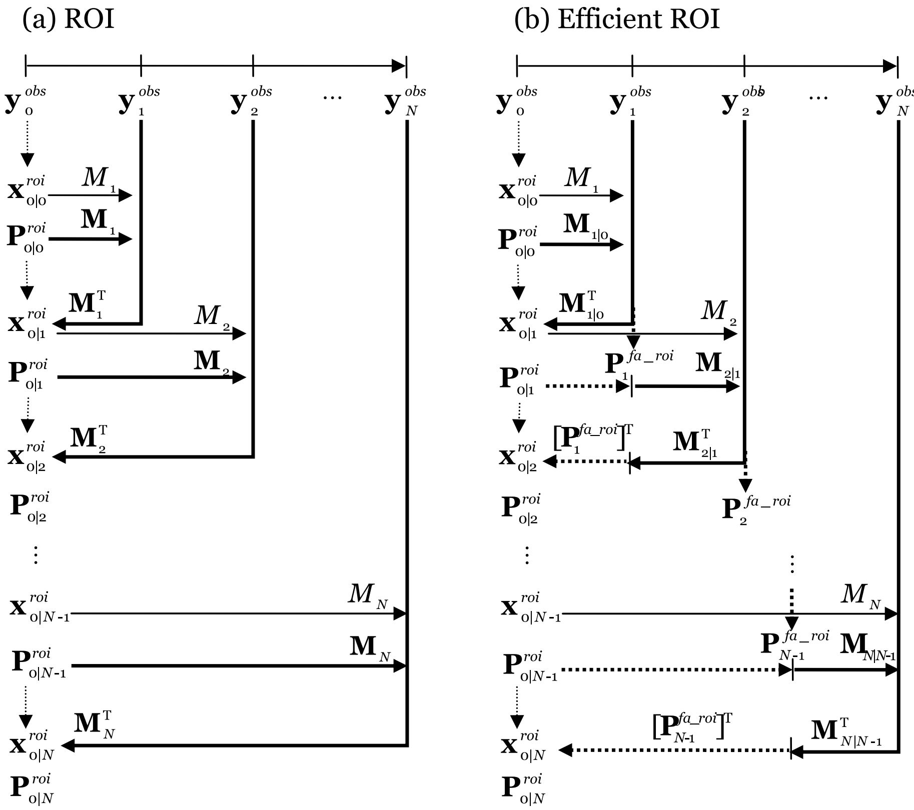

Fig. 1. The schematic procedures for implementation of (a) ROI and (b) efficient ROI. The solid arrow means that the process requires a model integration and especially the process corresponding to the bold solid arrow requires model integrations of which the number is equal to the degrees of freedom of a numerical model. The dashed arrow denotes that the process does not require model integrations.

In equation set (1), a time period for integrating a previous analysis error covariance $\mathbf{P}^{\mathrm{roi}}_{0|n-1}$ is overlapped at each analysis step as shown in Fig. 1a. In order to improve the efficiency of ROI, we devise a modified ROI algorithm in which we can eliminate such overlapping. In equation set (1), we define a matrix multiplication $\mathbf{M}_n\mathbf{P}^{\mathrm{roi}}_{0|n-1}$ as a cross error covariance $\mathbf{P}^{fa\_\mathrm{roi}}_n$ between forecast and analysis

$$\mathbf{P}^{fa\_\mathrm{roi}}_n \equiv \mathbf{M}_n\mathbf{P}^{\mathrm{roi}}_{0|n-1}. \tag{3}$$

Another matrix multiplication $\mathbf{M}_n\mathbf{P}^{\mathrm{roi}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n$ is defined as a forecast error covariance $\mathbf{P}^{f\_\mathrm{roi}}_n$

$$\mathbf{P}^{f\_\mathrm{roi}}_n \equiv \mathbf{M}_n\mathbf{P}^{\mathrm{roi}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n. \tag{4}$$

Note that we need a model integration of n observation time intervals for calculating the cross and forecast error covariance matrices. In other words, the model integration periods overlap at each analysis step.

We intend to eliminate such overlapping in calculating $\mathbf{P}^{fa\_\mathrm{roi}}_n$ and $\mathbf{P}^{f\_\mathrm{roi}}_n$. Consider a relationship,

$$\mathbf{M}_n = \mathbf{M}_{n|n-1}\mathbf{M}_{n-1}, \tag{5}$$

where the matrix $\mathbf{M}_{n|n-1}$ is the first derivative of $M_{n|n-1}(\mathbf{x})$ with respect to $\mathbf{x}$, in which $M_{n|n-1}(\mathbf{x})$ is a model operator transforming a model state $\mathbf{x}_{n-1}$ at a time t_{n-1} into a future state $\mathbf{x}_n$ of the next observation time t_n. Based on (5), we can rewrite $\mathbf{P}^{fa\_\mathrm{roi}}_n$ and $\mathbf{P}^{f\_\mathrm{roi}}_n$ as

$$\mathbf{P}^{fa\_\mathrm{roi}}_n = \mathbf{M}_{n|n-1}\mathbf{M}_{n-1}\mathbf{P}^{\mathrm{roi}}_{0|n-1}, \tag{6}$$

$$\mathbf{P}^{f\_\mathrm{roi}}_n = \mathbf{M}_{n|n-1}\mathbf{M}_{n-1}\mathbf{P}^{\mathrm{roi}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_{n-1}\mathbf{M}^{\mathrm{T}}_{n|n-1}. \tag{7}$$

By referring to eqs (1b) and (1c), eq. (6) can be rewritten as follows:

$$\begin{aligned}\mathbf{P}^{fa\_\mathrm{roi}}_n &= \mathbf{M}_{n|n-1}\mathbf{M}_{n-1}\left[\mathbf{I} - \mathbf{W}^{\mathrm{roi}}_{0|n-1}\mathbf{H}_{n-1}\mathbf{M}_{n-1}\right]\mathbf{P}^{\mathrm{roi}}_{0|n-2}\\ &= \mathbf{M}_{n|n-1}\left[\mathbf{M}_{n-1}\mathbf{P}^{\mathrm{roi}}_{0|n-2} - \mathbf{M}_{n-1}\mathbf{P}^{\mathrm{roi}}_{0|n-2}\mathbf{M}^{\mathrm{T}}_{n-1}\mathbf{H}^{\mathrm{T}}_{n-1}\right.\\ &\quad \left.\times\left[\mathbf{H}_{n-1}\mathbf{P}^{f\_\mathrm{roi}}_{n-1}\mathbf{H}^{\mathrm{T}}_{n-1} + \mathbf{R}_{n-1}\right]^{-1}\mathbf{H}_{n-1}\mathbf{M}_{n-1}\mathbf{P}^{\mathrm{roi}}_{0|n-2}\right]\\ &= \mathbf{M}_{n|n-1}\left[\mathbf{I} - \mathbf{P}^{f\_\mathrm{roi}}_{n-1}\mathbf{H}^{\mathrm{T}}_{n-1}[\mathbf{H}_{n-1}\mathbf{P}^{f\_\mathrm{roi}}_{n-1}\mathbf{H}^{\mathrm{T}}_{n-1}\right.\\ &\quad \left.+\mathbf{R}_{n-1}\right]^{-1}\mathbf{H}_{n-1}]\mathbf{M}_{n-1}\mathbf{P}^{\mathrm{roi}}_{0|n-2}.\end{aligned} \tag{8}$$

We define the forecast optimal weight matrix $\mathbf{W}^{f\_\mathrm{roi}}_{n-1}$ as

$$\mathbf{W}^{f\_\mathrm{roi}}_{n-1} = \mathbf{P}^{\mathrm{f\_roi}}_{n-1}\mathbf{H}^{\mathrm{T}}_{n-1}\left[\mathbf{H}_{n-1}\mathbf{P}^{f\_\mathrm{roi}}_{n-1}\mathbf{H}^{\mathrm{T}}_{n-1} + \mathbf{R}_{n-1}\right]^{-1}, \tag{9}$$

and use the definition of $\mathbf{P}_n^{fa\_roi}$ in (4) and eqs (8) and (9) to equate $\mathbf{P}_n^{fa\_roi}$ to $\mathbf{P}_{n-1}^{fa\_roi}$

$$\mathbf{P}_n^{fa\_roi} = \mathbf{M}_{n|n-1}\left[\mathbf{I} - \mathbf{W}_{n-1}^{f\_roi}\mathbf{H}_{n-1}\right]\mathbf{P}_{n-1}^{fa\_roi}. \tag{10a}$$

We can also conclude that $\mathbf{P}_n^{f\_roi}$ can be represented with the forecast error covariance $\mathbf{P}_{n-1}^{f\_roi}$ of the previous analysis step as follows:

$$\begin{aligned}\mathbf{P}_n^{f\_roi} &= \mathbf{M}_{n|n-1}\left[\mathbf{M}_{n-1}\mathbf{P}_{0|n-2}^{\mathrm{roi}}\mathbf{M}_{n-1}^{\mathrm{T}} - \mathbf{M}_{n-1}\mathbf{P}_{0|n-2}^{\mathrm{roi}}\mathbf{M}_{n-1}^{\mathrm{T}}\mathbf{H}_{n-1}^{\mathrm{T}}\right.\\ &\quad\times\left[\mathbf{H}_{n-1}\mathbf{M}_{n-1}\mathbf{P}_{0|n-2}^{\mathrm{roi}}\mathbf{M}_{n-1}^{\mathrm{T}}\mathbf{H}_{n-1}^{\mathrm{T}} + \mathbf{R}_{n-1}\right]^{-1}\\ &\quad\left.\times\mathbf{H}_{n-1}\mathbf{M}_{n-1}\mathbf{P}_{0|n-2}^{\mathrm{roi}}\mathbf{M}_{n-1}^{\mathrm{T}}\right]\mathbf{M}_{n|n-1}^{\mathrm{T}}\\ &= \mathbf{M}_{n|n-1}\left[\mathbf{P}_{n-1}^{f\_roi} - \mathbf{P}_{n-1}^{f\_roi}\mathbf{H}_{n-1}^{\mathrm{T}}\left[\mathbf{H}_{n-1}\mathbf{P}_{n-1}^{f\_roi}\mathbf{H}_{n-1}^{\mathrm{T}} + \mathbf{R}_{n-1}\right]^{-1}\right.\\ &\quad\left.\times\mathbf{H}_{n-1}\mathbf{P}_{n-1}^{f\_roi}\right]\mathbf{M}_{n|n-1}^{\mathrm{T}}\\ &= \mathbf{M}_{n|n-1}\left[\mathbf{I} - \mathbf{W}_{n-1}^{f\_roi}\mathbf{H}_{n-1}\right]\mathbf{P}_{n-1}^{f\_roi}\mathbf{M}_{n|n-1}^{\mathrm{T}},\end{aligned} \tag{10b}$$

in which both $\mathbf{P}_{-1}^{fa\_roi}$ and $\mathbf{P}_{-1}^{f\_roi}$ are equal to the background error covariance matrix $\mathbf{B}$, and $\mathbf{M}_{0|-1}$ is equivalent to an identity. Note that, when using (10a) and (10b), we need a model integration of just one observation time interval, not of n observation time intervals, to obtain $\mathbf{P}_n^{fa\_roi}$ and $\mathbf{P}_n^{f\_roi}$. Therefore, by using (10a) and (10b) instead of (3) and (4), we can improve the efficiency of ROI. The modified ROI equation set named 'efficient ROI' hereafter is written as follows:

$$\mathbf{x}_{0|n}^{\mathrm{roi}} = \mathbf{x}_{0|n-1}^{\mathrm{roi}} + \mathbf{W}_{0|n}^{\mathrm{roi}}\left\{\mathbf{y}_n^{\mathrm{obs}} - H_n\left[M_{n|n-1}\left(M_{n-1}\left(\mathbf{x}_{0|n-1}^{\mathrm{roi}}\right)\right)\right]\right\}, \tag{10c}$$

$$\mathbf{P}_{0|n}^{\mathrm{roi}} = \mathbf{P}_{0|n-1}^{\mathrm{roi}} - \mathbf{W}_{0|n}^{\mathrm{roi}}\mathbf{H}_n\mathbf{P}_n^{fa\_roi}, \tag{10d}$$

$$\mathbf{W}_{0|n}^{\mathrm{roi}} = \left[\mathbf{P}_n^{fa\_roi}\right]^{\mathrm{T}}\mathbf{H}_n^{\mathrm{T}}\left[\mathbf{H}_n\mathbf{P}_n^{f\_roi}\mathbf{H}_n^{\mathrm{T}} + \mathbf{R}_n\right]^{-1}. \tag{10e}$$

At the expense of efficiency, however, we need the storage resource for $\mathbf{P}_n^{fa\_roi}$ and $\mathbf{P}_n^{f\_roi}$. Figure 1b shows the schematic procedure for the implementation of efficient ROI. Comparing it with Fig. 1a, we become aware that the efficient ROI eliminate the overlap of model integration periods, which appears in the procedure of evolving an analysis error covariance. Although the model integration periods to obtain a guess of an observation field are still overlapped, efficient ROI is much more cost-effective than ROI since evolving analysis error covariance matrices mainly determines the computation costs for the implementation of each method.

2.2. *Comparison of efficient ROI with the FLKS with a perfect non-linear model for one past analysis time*

FLKS was proposed by Cohn et al. (1994; table 1) as a retrospective data assimilation method to produce long-term reanalysis data. To simplify our discussion, we suppose that FLKS is applied to obtain the analysis state at the initial time of a fixed analysis window as is done for ROI, and the involved model is non-linear with no errors since the model errors are not considered in efficient ROI. By setting l and k of table 1 of Cohn et al. (1994) to be equal to the index n that stands for the observation time and the model error covariance $\mathbf{Q}_k$ to be zero, the FLKS equation set can be written as follows:

$$\mathbf{x}_{0|n}^a = \mathbf{x}_{0|n-1}^a + \mathbf{K}_{0|n}\left\{\mathbf{y}_n^{\mathrm{obs}} - H_n\left[M_{n|n-1}\left(\mathbf{x}_{n-1|n-1}^a\right)\right]\right\}, \tag{11a}$$

$$\mathbf{P}_{0|n}^a = \mathbf{P}_{0|n-1}^a - \mathbf{K}_{0|n}\mathbf{H}_n\mathbf{P}_{n,0|n-1}^{fa}, \tag{11b}$$

$$\mathbf{K}_{0|n} = \left[\mathbf{P}_{n,0|n-1}^{fa}\right]^{\mathrm{T}}\mathbf{H}_n^{\mathrm{T}}\left[\mathbf{H}_n\mathbf{P}_{n|n-1}^{f}\mathbf{H}_n^{\mathrm{T}} + \mathbf{R}_n\right]^{-1}. \tag{11c}$$

The above equations are the retrospective analysis, the retrospective analysis error covariance, and the FLKS gain. As in equation set (1) for ROI, the subscript $0|n$ means that the FLKS analysis $\mathbf{x}_{0|n}^a$ results from the assimilation of an observation $\mathbf{y}_n^{\mathrm{obs}}$ at a time t_n into a background state at the initial time t_0 of the analysis window (Cohn et al., 1994). The cross error covariance $\mathbf{P}_{n,0|n-1}^{fa}$ of the FLKS is

$$\mathbf{P}_{n,0|n-1}^{fa} = \mathbf{M}_{n|n-1}[\mathbf{I} - \mathbf{K}_{n-1|n-1}\mathbf{H}_{n-1}]\mathbf{P}_{n-1,0|n-2}^{fa}. \tag{12}$$

The Kalman filter gain $\mathbf{K}_{n-1|n-1}$ is

$$\mathbf{K}_{n-1|n-1} = \mathbf{P}_{n-1|n-2}^{f}\mathbf{H}_{n-1}^{\mathrm{T}}\left[\mathbf{H}_{n-1}\mathbf{P}_{n-1|n-2}^{f}\mathbf{H}_{n-1}^{\mathrm{T}} + \mathbf{R}_{n-1}\right]^{-1}, \tag{13}$$

and the forecast error covariance $\mathbf{P}_{n|n-1}^{f}$ of the FLKS may be expressed as

$$\mathbf{P}_{n|n-1}^{f} = \mathbf{M}_{n|n-1}[\mathbf{I} - \mathbf{K}_{n-1|n-1}\mathbf{H}_{n-1}]\mathbf{P}_{n-1|n-2}^{f}\mathbf{M}_{n|n-1}^{\mathrm{T}}. \tag{14}$$

Both $\mathbf{P}_{-1,0|-2}^{fa}$ and $\mathbf{P}_{-1|-2}^{f}$ are equal to the background error covariance matrix $\mathbf{B}$. The extended Kalman filter (EKF) analysis $\mathbf{x}_{n-1|n-1}^a$ is given by

$$\begin{aligned}\mathbf{x}_{n-1|n-1}^a &= M_{n-1|n-2}\left(\mathbf{x}_{n-2|n-2}^a\right) + \mathbf{K}_{n-1|n-1}\\ &\quad\times\left\{\mathbf{y}_{n-1}^{\mathrm{obs}} - H_{n-1}\left[M_{n-1|n-2}(\mathbf{x}_{n-2|n-2}^a)\right]\right\},\end{aligned} \tag{15}$$

where $\mathbf{x}_{-1|-1}^a$ is equal to be $\mathbf{x}_0^b$.

To clarify the difference between efficient ROI and FLKS, we compared the FLKS and efficient ROI equations in Table 1. We can see that, except for the underlined parts of the retrospective analysis equations, the equations for the efficient ROI are the same as those for the FLKS. We would call the underlined part in the retrospective analysis of efficient ROI, corresponding to the EKF analysis of the FLKS, 'ROI forecast'. Efficient ROI can be considered as a variant of FLKS, which uses a forecast starting with a retrospective analysis as a guess of an observation field, instead of a forecast from the EKF analysis. Since the model integration periods for obtaining a guess of an observation field, not for evolving an analysis error covariance, are overlapped in the efficient ROI, the efficient ROI has higher computation costs than the FLKS. However, the dimension m of the model state is

Table 1. The equation sets of FLKS and efficient ROI

	FLKS	Efficient ROI
Retrospective analysis	$\mathbf{x}^a_{0\|n} = \mathbf{x}^a_{0\|n-1} + \mathbf{K}_{0\|n}\{\mathbf{y}^{obs}_n - H_n[M_{n\|n-1}(\underline{\mathbf{x}^a_{n-1\|n-1}})]\}$	$\mathbf{x}^{roi}_{0\|n} = \mathbf{x}^{roi}_{0\|n-1} + \mathbf{W}^{roi}_{0\|n}\{\mathbf{y}^{obs}_n - H_n[M_{n\|n-1}(\underline{M_{n-1}(\mathbf{x}^{roi}_{0\|n-1})})]\}$
Kalman filter analysis (ROI forecast)	$\mathbf{x}^a_{n-1\|n-1} = M_{n-1\|n-2}(\mathbf{x}^a_{n-2\|n-2}) + \mathbf{K}_{n-1\|n-1}\{\mathbf{y}^{obs}_{n-1} - H_{n-1}[M_{n-1\|n-2}(\mathbf{x}^a_{n-2\|n-2})]\}$	$M_{n-1}(\mathbf{x}^{roi}_{0\|n-1})$
Retrospective analysis error covariance	$\mathbf{P}^a_{0\|n} = \mathbf{P}^a_{0\|n-1} - \mathbf{K}_{0\|n}\,\mathbf{H}_n\,\mathbf{P}^{fa}_{n,0\|n-1}$	$\mathbf{P}^{roi}_{0\|n} = \mathbf{P}^{roi}_{0\|n-1} - \mathbf{W}^{roi}_{0\|n}\mathbf{H}_n\mathbf{P}^{fa\_roi}_n$
Cross error covariance	$\mathbf{P}^{fa}_{n,0\|n-1} = \mathbf{M}_{n\|n-1}[\mathbf{I} - \mathbf{K}_{n-1\|n-1}\mathbf{H}_{n-1}]\mathbf{P}^{fa}_{n-1,0\|n-2}$	$\mathbf{P}^{fa\_roi}_n = \mathbf{M}_{n\|n-1}[\mathbf{I} - \mathbf{W}^{f\_roi}_{n-1}\mathbf{H}_{n-1}]\mathbf{P}^{fa\_roi}_{n-1}$
Forecast error covariance	$\mathbf{P}^f_{n\|n-1} = \mathbf{M}_{n\|n-1}[\mathbf{I} - \mathbf{K}_{n-1\|n-1}\mathbf{H}_{n-1}]\mathbf{P}^f_{n-1\|n-2}\mathbf{M}^T_{n\|n-1}$	$\mathbf{P}^{f\_roi}_n = \mathbf{M}_{n\|n-1}[\mathbf{I} - \mathbf{W}^{f\_roi}_{n-1}\mathbf{H}_{n-1}]\mathbf{P}^{f\_roi}_{n-1}\mathbf{M}^T_{n\|n-1}$
Retrospective analysis gain	$\mathbf{K}_{0\|n} = [\mathbf{P}^{fa}_{n,0\|n-1}]^T\mathbf{H}^T_n[\mathbf{H}_n\mathbf{P}^f_{n\|n-1}\mathbf{H}^T_n + \mathbf{R}_n]^{-1}$	$\mathbf{W}^{roi}_{0\|n} = [\mathbf{P}^{fa\_roi}_n]^T\mathbf{H}^T_n[\mathbf{H}_n\mathbf{P}^{f\_roi}_n\mathbf{H}^T_n + \mathbf{R}_n]^{-1}$
Kalman filter gain (forecast optimal weight)	$\mathbf{K}_{n-1\|n-1} = \mathbf{P}^f_{n-1\|n-2}\mathbf{H}^T_{n-1}[\mathbf{H}_{n-1}\mathbf{P}^f_{n-1\|n-2}\mathbf{H}^T_{n-1} + \mathbf{R}_{n-1}]^{-1}$	$\mathbf{W}^{f\_roi}_{n-1} = \mathbf{P}^{f\_roi}_{n-1}\mathbf{H}^T_{n-1}[\mathbf{H}_{n-1}\mathbf{P}^{f\_roi}_{n-1}\mathbf{H}^T_{n-1} + \mathbf{R}_{n-1}]^{-1}$

Note: The FLKS is assumed to produce the analysis state at the initial time of a fixed analysis window as in the case of ROI, and the involved model is non-linear and has no errors. The FLKS equation set is obtained by setting l and k of table 1 of Cohn et al. (1994) to be equal to the index n that stands for the observation time and the model error covariance to be zero.

huge relative to the number N of future observation times in an analysis window. Therefore, the computation costs to implement efficient ROI are almost the same as those for FLKS.

Zhu et al. (2003) showed that forecasts produced from retrospective analyses match observations better than forecasts from filter analyses. Based on the experimental results by Zhu et al. (2003), we anticipate that the efficient ROI will produce more accurate results than the FLKS. In addition to their experimental results, we will theorize our hypothesis.

From the EKF analysis in (15) and the efficient ROI analysis in (10c), we know that $\mathbf{x}^a_{0|0}$ is equal to $\mathbf{x}^{roi}_{0|0}$. The second analysis $\mathbf{x}^{roi}_{0|1}$ of efficient ROI is written as

$$\mathbf{x}^{roi}_{0|1} = \mathbf{x}^{roi}_{0|0} + \mathbf{P}^{roi}_{0|0}\mathbf{M}^T_1\mathbf{H}^T_1\left[\mathbf{H}_1\mathbf{M}_1\mathbf{P}^{roi}_{0|0}\mathbf{M}^T_1\mathbf{H}^T_1 + \mathbf{R}_1\right]^{-1} \times \left\{\mathbf{y}^{obs}_1 - H_1[M_{1|0}(\mathbf{x}^{roi}_{0|0})]\right\}. \quad (16)$$

The second EKF analysis $\mathbf{x}^a_{1|1}$ is

$$\mathbf{x}^a_{1|1} = M_{1|0}(\mathbf{x}^a_{0|0}) + \mathbf{P}^f_{1|0}\mathbf{H}^T_1\left[\mathbf{H}_1\mathbf{P}^f_{1|0}\mathbf{H}^T_1 + \mathbf{R}_1\right]^{-1} \times \left\{\mathbf{y}^{obs}_1 - H_1\left[M_{1|0}(\mathbf{x}^a_{0|0})\right]\right\}. \quad (17)$$

Since we know that $\mathbf{P}^f_{1|0} = \mathbf{M}_1\mathbf{P}^{roi}_{0|0}\mathbf{M}^T_1$, eq. (16) can be rewritten as follows:

$$\mathbf{x}^a_{1|1} = M_1\left(\mathbf{x}^{roi}_{0|0}\right) + \mathbf{M}_1\left[\mathbf{x}^{roi}_{0|1} - \mathbf{x}^{roi}_{0|0}\right]. \quad (18)$$

For $i = 2, \ldots, n-1$, the analysis $\mathbf{x}^{roi}_{0|i}$ of the efficient ROI is written as

$$\mathbf{x}^{roi}_{0|i} = \mathbf{x}^{roi}_{0|i-1} + \mathbf{P}^{roi}_{0|i-1}\mathbf{M}^T_i\mathbf{H}^T_i\left[\mathbf{H}_i\mathbf{M}_i\mathbf{P}^{roi}_{0|i-1}\mathbf{M}^T_i\mathbf{H}^T_i + \mathbf{R}_i\right]^{-1} \times \left\{\mathbf{y}^{obs}_i - H_i\left[M_{i|i-1}\left(M_{i-1}\left(\mathbf{x}^{roi}_{0|i-1}\right)\right)\right]\right\}, \quad (19)$$

and the EKF analysis $\mathbf{x}^a_{i|i}$ is

$$\mathbf{x}^a_{i|i} = M_{i|i-1}\left(\mathbf{x}^a_{i-1|i-1}\right) + \mathbf{P}^f_{i|i-1}\mathbf{H}^T_i\left[\mathbf{H}_i\mathbf{P}^f_{i|i-1}\mathbf{H}^T_i + \mathbf{R}_i\right]^{-1} \times \left\{\mathbf{y}^{obs}_i - H_i\left[M_{i|i-1}\left(\mathbf{x}^a_{i-1|i-1}\right)\right]\right\}. \quad (20)$$

If the following relationship,

$$\mathbf{x}^a_{i-1|i-1} = M_{i-1}\left(\mathbf{x}^{roi}_{0|i-2}\right) + \mathbf{M}_{i-1}\left[\mathbf{x}^{roi}_{0|i-1} - \mathbf{x}^{roi}_{0|i-2}\right],$$

is valid, by referring to the relation $\mathbf{P}^f_{i|i-1} = \mathbf{M}_i\mathbf{P}^{roi}_{0|i-1}\mathbf{M}^T_i$, we can say that

$$\mathbf{x}^a_{i|i} = M_i\left(\mathbf{x}^{roi}_{0|i-1}\right) + \mathbf{M}_i\left[\mathbf{x}^{roi}_{0|i} - \mathbf{x}^{roi}_{0|i-1}\right]. \quad (21)$$

For the above equation, we assumed the following tangent linear relationship,

$$M_{i-1}\left(\mathbf{x}^{roi}_{0|i-1}\right) = M_{i-1}\left(\mathbf{x}^{roi}_{0|i-2}\right) + \mathbf{M}_{i-1}\left[\mathbf{x}^{roi}_{0|i-1} - \mathbf{x}^{roi}_{0|i-2}\right]. \quad (22)$$

Inductively, we conclude that the EKF analysis $\mathbf{x}^a_{n-1|n-1}$ can be written as

$$\mathbf{x}^a_{n-1|n-1} = M_{n-1}\left(\mathbf{x}^{roi}_{0|n-2}\right) + \mathbf{M}_{n-1}\left[\mathbf{x}^{roi}_{0|n-1} - \mathbf{x}^{roi}_{0|n-2}\right]. \quad (23)$$

Note assumption (22) that is necessary for deriving (23) cannot guarantee the equality of the EKF analysis $\mathbf{x}^a_{n-1|n-1}$ to the ROI forecast $M_{n-1}(\mathbf{x}^{roi}_{0|n-1})$ since i does not include n. Equation (23) implies that the EKF analysis $\mathbf{x}^a_{n-1|n-1}$ is a tangent linear approximation of an ROI forecast $M_{n-1}(\mathbf{x}^{roi}_{0|n-1})$, which is integrated starting with an ROI analysis $\mathbf{x}^{roi}_{0|n-1}$. We can say that the FLKS is a suboptimal version of the efficient ROI, which uses the tangent linear approximation $M_{n-1}(\mathbf{x}^{roi}_{0|n-2}) + \mathbf{M}_{n-1}[\mathbf{x}^{roi}_{0|n-1} - \mathbf{x}^{roi}_{0|n-2}]$

as a guess of an observation field, which differs from an ROI forecast $M_{n-1}(\mathbf{x}^{\mathrm{roi}}_{0|n-1})$.

As the analysis progresses, the difference between $M_{n-1}(\mathbf{x}^{\mathrm{roi}}_{0|n-1})$ and $M_{n-1}(\mathbf{x}^{\mathrm{roi}}_{0|n-2}) + \mathbf{M}_{n-1}[\mathbf{x}^{\mathrm{roi}}_{0|n-1} - \mathbf{x}^{\mathrm{roi}}_{0|n-2}]$ increases since the non-linearity of the model operator becomes stronger. When FLKS is applied for the analysis state at the beginning of the fixed analysis window, the efficient ROI analysis is more accurate than the FLKS analysis as the analysis window size increases when the involved model is non-linear and has no errors. However, the above reasoning will not work for the analysis window too wide to satisfy the assumption (2).

3. Reduced-rank formulation of efficient ROI without using an adjoint model

Equation (1) points out that the analysis errors of ROI decrease as the analysis step progresses. Song et al. (2009) set a criterion defined as error tolerance to select the analysis components that will be analysed in the ROI procedure and to neglect all errors that are smaller in magnitude than the so-called error tolerance. The transform matrix $\mathbf{S}_{0|n-1}$ plays the role of filtering the accuracy-saturated components of the analysis vector of which the error variance is smaller in magnitude than the predefined error tolerance. By using the k by l transform matrix $\mathbf{S}_{0|n-1}$ for $k < l$, the reduced-rank formulation of ROI is written as follows:

$$\mathbf{s}^b_{0|n-1} = \mathbf{S}_{0|n-1}\mathbf{s}^a_{0|n-1}, \tag{24a}$$

$$\begin{aligned}\mathbf{s}^a_{0|n} &= \mathbf{s}^b_{0|n-1} + \left[\left(\mathbf{S}_{0|n-1}\mathbf{P}^s_{0|n-1}\mathbf{S}^{\mathrm{T}}_{0|n-1}\right)^{-1}\right.\\ &\quad \left. + \mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n\mathbf{H}^{\mathrm{T}}_n\mathbf{R}^{-1}_n\mathbf{H}_n\mathbf{M}_n\mathbf{T}_{0|n-1}\right]^{-1}\mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n\mathbf{H}^{\mathrm{T}}_n\mathbf{R}^{-1}_n\\ &\quad \times \left[\mathbf{y}^{\mathrm{obs}}_n - H^s_n\left(\mathbf{s}^b_{0|n-1}\right)\right],\end{aligned} \tag{24b}$$

$$\begin{aligned}\mathbf{P}^s_{0|n} &= \left[\left(\mathbf{S}_{0|n-1}\mathbf{P}^s_{0|n-1}\mathbf{S}^{\mathrm{T}}_{0|n-1}\right)^{-1}\right.\\ &\quad \left. + \mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_n\mathbf{H}^{\mathrm{T}}_n\mathbf{R}^{-1}_n\mathbf{H}_n\mathbf{M}_n\mathbf{T}_{0|n-1}\right]^{-1},\end{aligned} \tag{24c}$$

$$\mathbf{x}^{\mathrm{roi}}_{0|n} = \mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{s}^a_{0|n} + \mathbf{T}^{\mathrm{T}}_{0|n-2}\mathbf{n}_{0|n-1}, \tag{24d}$$

where $\mathbf{S}_{0|-1}$ is an $m \times m$ identity matrix and l is equal to m at the initial analysis step. When $\mathbf{s}^a_{0|-1}$ is equal to the background state $\mathbf{x}^b_0$ having m components and $\mathbf{T}_{0|n-1} = \prod_{i=n}^{0}\mathbf{S}_{0|i-1}$, the null space of the analysis field is defined as $\mathbf{n}_{0|n-1} = \sum_{i=0}^{n}\mathbf{T}^{\mathrm{T}}_{0|i-2}[\mathbf{s}^a_{0|i-1} - \mathbf{S}^{\mathrm{T}}_{0|i-1}\mathbf{s}^b_{0|i-1}]$ with an $m \times m$ identity matrix $\mathbf{T}_{0|-2}$ and $H^s_n(\mathbf{s}^b_{0|n-1}) = H_n[M_n(\mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{s}^b_{0|n-1} + \mathbf{n}_{0|n-1})]$.

To devise the reduced-rank formulation for efficient ROI, we should be able to equate $\mathbf{P}^{fa\_\mathrm{roi}}_n$ and $\mathbf{P}^{f\_\mathrm{roi}}_n$ to the analysis error covariance $\mathbf{P}^{\mathrm{roi}}_{0|n-1}$. Therefore, we use eqs (6) and (7) instead of (10a) and (10b) for obtaining $\mathbf{P}^{fa\_\mathrm{roi}}_n$ and $\mathbf{P}^{f\_\mathrm{roi}}_n$ with some manipulations to avoid the overlap of the model integration periods. By using a variant of the Sherman–Woodbury–Morrison formula (Kalnay, 2003), we can write the reduced-rank version of eqs (10c) and (10d) as follows:

$$\begin{aligned}\mathbf{s}^a_{0|n} &= \mathbf{s}^b_{0|n-1} + \left[\left(\mathbf{S}_{0|n-1}\mathbf{P}^s_{0|n-1}\mathbf{S}^{\mathrm{T}}_{0|n-1}\right)^{-1}\right.\\ &\quad \left. + \mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_{n-1}\mathbf{M}^{\mathrm{T}}_{n|n-1}\mathbf{H}^{\mathrm{T}}_n\mathbf{R}^{-1}_n\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{M}_{n-1}\mathbf{T}_{0|n-1}\right]^{-1}\\ &\quad \times \mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_{n-1}\mathbf{M}^{\mathrm{T}}_{n|n-1}\mathbf{H}^{\mathrm{T}}_n\mathbf{R}^{-1}_n\left[\mathbf{y}^{\mathrm{obs}}_n - H^s_n\left(\mathbf{s}^b_{0|n-1}\right)\right],\end{aligned} \tag{25a}$$

$$\begin{aligned}\mathbf{P}^s_{0|n} &= \left[\left(\mathbf{S}_{0|n-1}\mathbf{P}^s_{0|n-1}\mathbf{S}^{\mathrm{T}}_{0|n-1}\right)^{-1}\right.\\ &\quad \left. + \mathbf{T}^{\mathrm{T}}_{0|n-1}\mathbf{M}^{\mathrm{T}}_{n-1}\mathbf{M}^{\mathrm{T}}_{n|n-1}\mathbf{H}^{\mathrm{T}}_n\mathbf{R}^{-1}_n\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{M}_{n-1}\mathbf{T}_{0|n-1}\right]^{-1}.\end{aligned} \tag{25b}$$

In the above equation, the time period required to evolve the matrix $\mathbf{T}_{0|n-1}$ overlaps the model integration period of the next analysis step.

From the definition of $\mathbf{T}_{0|n-1}$, we know that the column vectors of $\mathbf{T}_{0|n-1}$ belong to those of $\mathbf{T}_{0|n-2}$. Therefore, we can obtain the result of the matrix multiplication $\mathbf{M}_{n-1}\mathbf{T}_{0|n-1}$ by simply selecting the column vectors of $\mathbf{M}_{n-1}\mathbf{T}_{0|n-2}$ corresponding to $\mathbf{T}_{0|n-1}$. We name the resulting matrix multiplication a matrix $\mathbf{T}^{fs}_{0|n-2}$. Based on (5), we can rewrite the matrix multiplication $\mathbf{M}_n\mathbf{T}_{0|n-1}$ as follows:

$$\begin{aligned}\mathbf{M}_n\mathbf{T}_{0|n-1} &= \mathbf{M}_{n|n-1}\mathbf{M}_{n-1}\mathbf{T}_{0|n-1}\\ &= \mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}.\end{aligned} \tag{26}$$

In the k-dimensional subspace, the analysis and the corresponding error covariance equations are changed into

$$\begin{aligned}\mathbf{s}^a_{0|n} &= \mathbf{s}^b_{0|n-1} + \left[\left(\mathbf{S}_{0|n-1}\mathbf{P}^s_{0|n-1}\mathbf{S}^{\mathrm{T}}_{0|n-1}\right)^{-1}\right.\\ &\quad \left. + \left(\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}\right)^{\mathrm{T}}\mathbf{R}^{-1}_n\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}\right]^{-1}\\ &\quad \times \left(\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}\right)^{\mathrm{T}}\mathbf{R}^{-1}_n\left[\mathbf{y}^{\mathrm{obs}}_n - H^s_n\left(\mathbf{s}^b_{0|n-1}\right)\right],\end{aligned} \tag{27a}$$

$$\begin{aligned}\mathbf{P}^s_{0|n} &= \left[\left(\mathbf{S}_{0|n-1}\mathbf{P}^s_{0|n-1}\mathbf{S}^{\mathrm{T}}_{0|n-1}\right)^{-1}\right.\\ &\quad \left. + \left(\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}\right)^{\mathrm{T}}\mathbf{R}^{-1}_n\mathbf{H}_n\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}\right]^{-1}.\end{aligned} \tag{27b}$$

If the error tolerance is set to be zero, the reduced-rank formulation of efficient ROI is exactly the same as the efficient ROI formulation (10). By using equation set (27) and eqs (24a) and (24d), we can avoid the overlap of the model integration periods when evolving the transformation matrix $\mathbf{T}_{0|n-1}$ for the implementation of the reduced-rank formulation. We need the model integration of just one observation time interval instead of n observation time intervals in obtaining $\mathbf{M}_n\mathbf{T}_{0|n-1}$ that is equal to $\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}$. However, we need extra storage resource for the matrix.

In (27), the matrix multiplication $\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}$ requires a tangent linear model. However, the perturbation method by Errico and Raeder (1999) allows us to implement efficient ROI without using the tangent linear model. Suppose that a state vector $\mathbf{x}$ is decomposed into a reference state $\mathbf{x}^r$ and a deviation field $\delta\mathbf{x}$

from $\mathbf{x}^r$. A non-linear and differentiable function $f(\mathbf{x})$ is operating on $\mathbf{x}$. By adopting a scale factor α much smaller than 1, we can approximate the multiplication of the Jacobian $.\frac{\partial f}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{x}^r}$ and the deviation $\delta\mathbf{x}$ irrespective of the Euclidean norm of $\delta\mathbf{x}$, as described by Errico and Raeder (1999)

$$\left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^r} \delta\mathbf{x} \approx \frac{f(\mathbf{x}^r + \alpha\delta\mathbf{x}) - f(\mathbf{x}^r)}{\alpha}. \tag{28}$$

If we apply (28) to the matrix multiplication $\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}$, we obtain the following equation,

$$\begin{aligned}\mathbf{M}_{n|n-1}\mathbf{t}^{fs\_r}_{0|n-2} = \tfrac{1}{\alpha_r}\big\{&M_{n|n-1}\big[M_{n-1}\big(\mathbf{x}^{\mathrm{roi}}_{0|n-1}\big) + \alpha_r\mathbf{t}^{fs\_r}_{0|n-2}\big] \\ &- M_{n|n-1}\big[M_{n-1}(\mathbf{x}^{\mathrm{roi}}_{0|n-1})\big]\big\},\end{aligned} \tag{29}$$

where $\mathbf{t}^{fs\_r}_{0|n-2}$ is the rth column vector of the matrix $\mathbf{T}^{fs}_{0|n-2}$, which is scaled by the factor α_r, and $r = 1, 2, \ldots, k$. By using (29), we can obtain the matrix multiplication $\mathbf{M}_{n|n-1}\mathbf{T}^{fs}_{0|n-2}$, that is $\mathbf{M}_n$ $\mathbf{T}_{0|n-1}$, and circumvent the use of an adjoint model to implement the reduced-rank formulation of the efficient ROI, since obviously eq. (27) does not require the calculation of the product of the adjoint model and any matrix. However, using the perturbation method or using an adjoint model is a choice. Although the scale factors α_r are adopted to eliminate errors originating from higher-order terms in the tangent linear approximation (28), the result of (28) may not be exactly the same as that of the tangent linear model.

4. Numerical experiments using simple models

4.1. *Comparison of efficient ROI with FLKS applied to a perfect model*

In this subsection, we confirm that the analysis of efficient ROI is more accurate than that of FLKS. We assume that FLKS is applied to obtain the analysis state at the initial time of a fixed analysis window, and the involved model is non-linear and has no errors to simplify our discussion. The dynamical model for our experiment is the Lorenz 3-variable model (Lorenz, 1963). The parameter set and the method to generate the background state (error variance = 1) and the observations (error variance = 0.25) are the same as in Song et al. (2009). While in their 3-variable model experiments they analysed only one component for easy display of cost functions, in the present experiment all three components are included because we need not to draw cost functions. In order to examine the dependence of the skills of both methods on the analysis window size, we change the number of observation times with a fixed observation time interval of 0.2.

We calculate the root-mean-squared (rms) analysis errors from FLKS and ROI analyses, which assimilated all observation sets in each analysis window. We average the analysis errors of 60 realizations to define the average analysis error. The average analysis error as a function of the analysis window size is shown in Fig. 2a for both FLKS (dashed) and efficient ROI (solid), respectively. The average analysis error of the FLKS becomes larger than that of the efficient ROI when the analysis window size is wider than 1.2. While the efficient ROI can improve the analysis quality by using farther observation data sets in analysis windows wider than 1.2, FLKS produces poorer results in spite of assimilating more observation sets in that case.

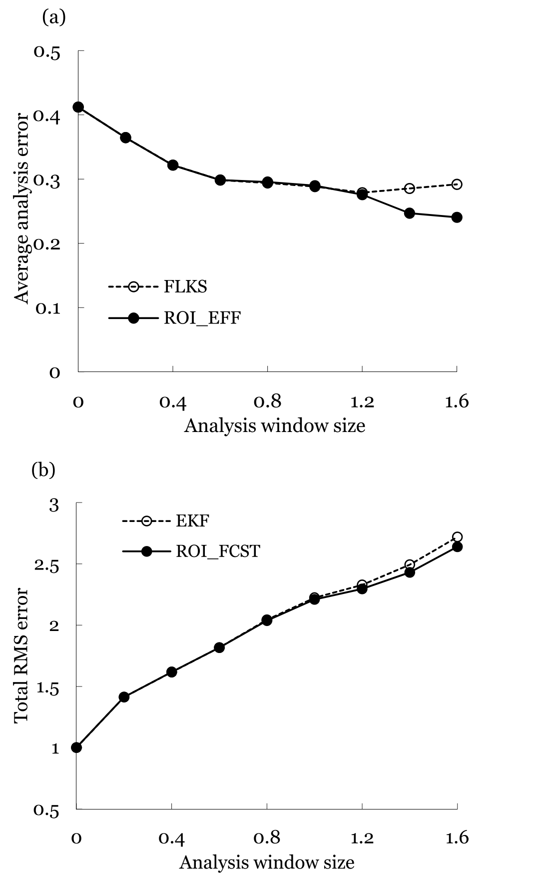

Fig. 2. (a) The average analysis error as a function of the analysis window size for FLKS (dashed) and efficient ROI (solid) and (b) the total rms error for the EKF analysis (dashed) and the ROI forecast (solid). The observation time interval is 0.2 in Lorenz 3-variable model experiments.

At once, we calculate the rms errors of both EKF analysis and ROI forecast every analysis step, and sum up the rms errors respectively to define 'total rms error'. The results show how a FLKS analysis is affected by using an EKF analysis instead of an ROI forecast as a guess of an observation field. When the analysis window size is wider than 1.2, the total rms error of the EKF analysis becomes larger than that of the ROI forecast as shown in Fig. 2b. This coincides with the results of the average analysis error in Fig. 2a. As the observation set to be assimilated gets farther away from the initial time of the analysis window, the time period of model integration to produce a guess of an observation field becomes longer and the non-linearity of the model operator makes stronger. Since, following the theoretical reasoning of Section 2, an EKF analysis is a tangent linear approximation to an ROI forecast, the ROI forecast becomes more accurate than the EKF analysis as the non-linearity of the model operator increases. As a result, the analysis quality of the efficient ROI becomes better than that of the FLKS with the increase of the analysis window size.

4.2. *Comparison of the reduced-rank version of efficient ROI with FLKS and ROI in terms of accuracy and efficiency*

The degrees of freedom of Lorenz 3-variable model are too small to compare the efficiency of each method presented in here. In this subsection, we use the Lorenz-96 model that has 40 variables (Lorenz, 1996). The model considers 40 grid points spaced equally along a constant latitude circle. The equation set for the model is,

$$\mathrm{d}\mathbf{x}_j/\mathrm{d}t = (\mathbf{x}_{j+1} - \mathbf{x}_{j-2})\mathbf{x}_{j-1} - \mathbf{x}_j + F, \tag{30}$$

where j (= 1, ..., 40) denotes the longitudinal coordinate. We implement (30) numerically with the fourth-order Runge–Kutta scheme. We set the external forcing F to 8, and the time step Δt to 0.0125, which roughly corresponds to one and half hours in a global weather model (Lorenz and Emmanuel, 1998).

We integrate a randomly generated initial condition for spin-up during 1 yr to obtain a true state. By adding a Gaussian-distributed random error of which the standard deviation is 0.25 to every component of the true state, we construct an ensemble, which consists of 50 members. The number of ensemble members is sufficient in that the degrees of freedom of the model are 40. Considering the doubling time of errors in Lorenz 40-variable model is 2.1 d, we integrate the ensemble members during 2 d to generate the ensemble forecasts (Lorenz and Emmanuel, 1998). We construct a background error covariance matrix **B** by using 50 deviation fields obtained from the subtraction of the mean of the ensemble forecasts from each ensemble members (Zhang, 2005). By adding a random noise vector with a Gaussian distribution (mean = 0, covariance = **B**) to the true state that has been integrated during 2 d, we produce a background state. We randomly select 20 observation positions from the 40 grid points and fix them for all experiments, and make observations every 6 h by adding uncorrelated random noises having Gaussian distributions (mean = 0, variance = 1) to the components of true state corresponding to the observation positions. We change the number of observation times with the fixed observation time interval to vary the analysis window size. The same background state and observations are used for efficient ROI, FLKS and ROI. As in Section 4.1, it is assumed that FLKS is applied to determine the analysis state at the beginning of a fixed analysis window as is done for ROI, and the involved model is non-linear and has no errors.

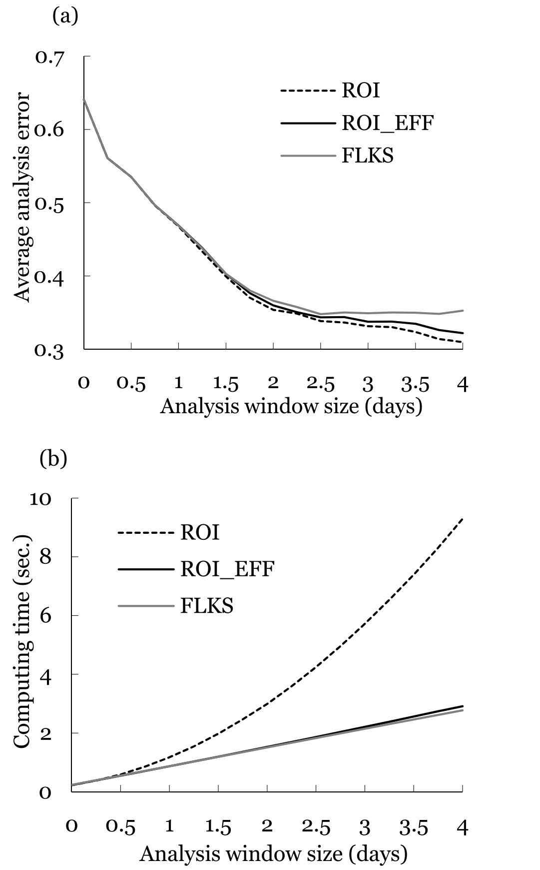

Fig. 3. (a) The average analysis error and (b) the computing time as functions of the analysis window size for ROI (dashed black), efficient ROI (solid black) and FLKS (solid grey) in Lorenz 40-variable model experiments. The error tolerance for ROI and efficient ROI is set to be zero. The computing time is in units of seconds.

To evaluate the efficient ROI analysis, in comparison with the FLKS and ROI analyses, which assimilated all observation sets in each analysis window, we calculate the average analysis error and the computing time required to obtain those analyses. We measure the analysis accuracy as the ratio of the average analysis error to the standard deviation of observation error, which is the inverse of the average analysis error because the standard deviation of the observation error is equal to 1. We estimate the relative accuracy by dividing the analysis accuracy of each ROI method by that of FLKS. The relative computing time is defined as the computing time of each ROI method divided by that of FLKS.

We show the average analysis error as a function of the analysis window size in Fig. 3a for the ROI analysis (dashed black), the efficient ROI analysis (solid black), and the FLKS analysis (solid grey). In this case, the ROI error tolerance was set to be zero (refer to Section 3). Irrespective of the size of the analysis window, the average analysis error of the efficient ROI is almost the same as that of the ROI (Fig. 3a). However, as the analysis window size increases, the analysis error of efficient ROI becomes slightly larger than that of ROI. In the ROI algorithm, a basic state for tangent linear model is a previous ROI analysis. However, in the efficient ROI algorithm, the tangent linear model is approximately implemented by eq. (5) as follows:

$$\mathbf{M}_n\left(\mathbf{x}^{\mathrm{roi}}_{0|n-1}\right) \approx \mathbf{M}_{n|n-1}\left[M_{n-1}\left(\mathbf{x}^{\mathrm{roi}}_{0|n-1}\right)\right]\mathbf{M}_{n-1|n-2} \left[M_{n-2}\left(\mathbf{x}^{\mathrm{roi}}_{0|n-2}\right)\right]\cdots\mathbf{M}_{1|0}\left(\mathbf{x}^{\mathrm{roi}}_{0|0}\right). \quad (31)$$

When the error tolerance is set to be zero, the result of tangent linear model in (29) is actually obtained by the right-hand side of (31). This approximation causes the slightly higher error of the efficient ROI analysis. When the analysis window size is wider than 2.5 d, the average analysis error of the efficient ROI decreases while that of the FLKS increases. This result supports the mathematical reasoning of Section 2, suggesting that FLKS is a suboptimal version of efficient ROI. This result is also consistent with that shown in Fig. 2.

The efficient ROI has much lower computation costs contrary to the ROI (Fig. 3b). The difference between the computing times for the ROI and for the efficient ROI increases as the analysis window size expands in time. Note that if the observation time interval over the analysis window is regular and is considered to be a unit forecast period, the model running times to evolve previous analysis error covariance matrices in the original ROI formulation (1) are proportional to $N(N + 1)$. In the efficient ROI and the FLKS, however, we need to calculate the cross and forecast error covariance where the model simulation times are proportional to N. Since evolving analysis error covariance matrices mainly determines the computation costs for the implementation of each method, the cost for the ROI behaves as a quadratic function of the analysis window size and those for the efficient ROI and the FLKS have the forms of linear functions as shown in Fig. 3b. Even though not shown in here, the results from fitting the lines of Fig. 3b with quadratic or linear functions confirm this reasoning. While offering a more accurate analysis than the FLKS, the efficient ROI requires almost the same computation costs as the FLKS (Fig. 3b).

For the ROI, the relative computing time exceeds the relative accuracy as the analysis window size increases (Fig. 4a). For the efficient ROI, the relative computing times tend to increase due to the overlap of model integration periods in obtaining the guess of an observation field with the increase of the analysis window size (Fig. 4b). However, the relative accuracy exceeds the relative computing time for the analysis window of which sizes are narrower than 1 d or wider than 3 d, while both relative values are similar for the other analysis windows. The results for wide analysis window are directly related to the reasoning of Section 2. When the analysis window size is narrow, the computing times required for calculating the cross and forecast analysis error covariance matrices in all three methods are the same. However, the FLKS may demand additional burdens such as matrix multiplication and inverse for obtaining the EKF analysis. Such extra costs make the FLKS inefficient relative to both ROI methods for narrow analysis windows.

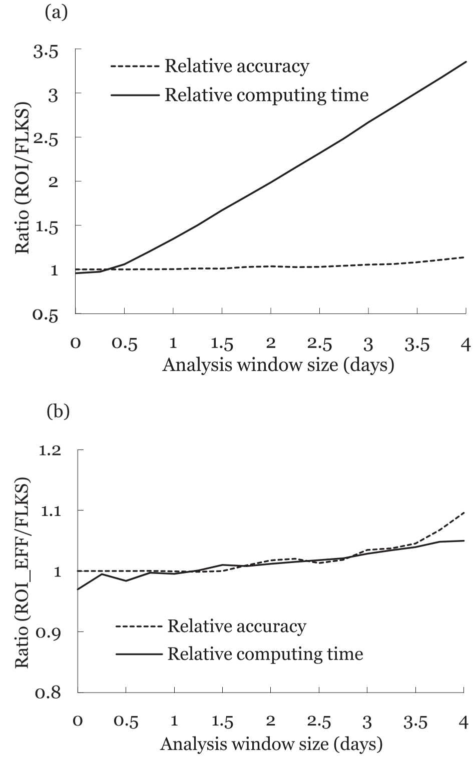

Fig. 4. The relative accuracy (dashed) and the relative computing time (solid) as functions of the analysis window size for (a) ROI and (b) efficient ROI. They are the analysis accuracy and the computing time for both ROI methods divided by those of FLKS, respectively.

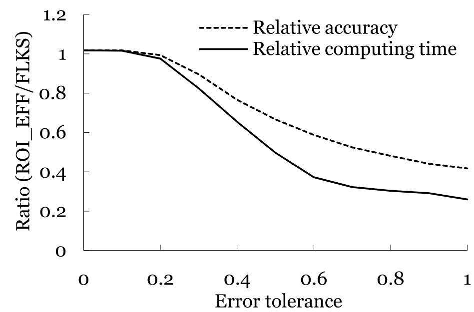

Fig. 5. The relative accuracy (dashed) and the relative computing time (solid) as functions of error tolerance for efficient ROI for the 2-d analysis windows.

The relative accuracy and the relative computing time depend on the magnitude of the error tolerance as shown in Fig. 5. For the 2-d analysis window, both relative accuracy and relative computing time decrease as error tolerance increases. When the error tolerance is larger than 0.3, relative accuracy is smaller than 1, which means the accuracy of the efficient ROI analysis is worse than that of the FLKS analysis. However, the relative computing time is smaller than the relative accuracy. By using the reduced-rank formulation of the efficient ROI, we can obtain the suboptimal analysis more cost-effectively rather than FLKS.

5. Conclusions

ROI is useful in avoiding the multiple-minima problem of 4D-Var without using an adjoint model (Song et al., 2009). In the ROI equation set, however, the model integration periods to evolve a previous analysis error covariance overlap at each analysis step. To improve the efficiency of the ROI algorithm, we devised a modified ROI algorithm that can avoid such overlapping and named it efficient ROI. When FLKS is used to obtain the analysis state at the initial time of a fixed analysis window, it is a suboptimal version of the efficient ROI since the FLKS uses a tangent linear approximation of a forecast stemming from a retrospective analysis for guessing an observation field. Therefore, the efficient ROI produces more accurate results than the FLKS with the increase of the analysis window size. The computation cost for implementing the efficient ROI nevertheless is almost the same as that for the FLKS.

From the experiments using Lorenz 3-variable model, we confirmed that the analysis quality of the efficient ROI becomes better than that of the FLKS as the analysis window size increases. To compare the functionality of efficient ROI with FLKS and ROI in terms of accuracy and computation costs, we carried out Lorenz 40-variable experiments. As the analysis window size increases, the analysis quality of the efficient ROI become better than that of the FLKS, even though the efficient ROI has almost the same computation costs as the FLKS. By increasing the error tolerance, we can get a suboptimal analysis with lower computation costs than for the FLKS.

In this study, the analysis window size was controlled by the number of observation times. However, it can also vary with the observation time interval. We would face such situations when assimilating the data observed with different time intervals. In future, the ability of the efficient ROI will be examined with increasing the observation time interval for the fixed number of observations. Also, for the analysis window too wide to fulfil the optimality of efficient ROI, we will compare the accuracies of the efficient ROI and the FLKS. The model errors are not considered in the present formulation of efficient ROI. In future, the efficient ROI will be generalized to incorporate model errors.

Since efficient ROI requires huge resources in computing and memory as FLKS does, we should use its reduced-rank formulation for application with realistic models, which typically have over a million degrees of freedom. In the reduced-rank formulation, however, only the components that belong to the analysis subspace are analysed. Since the accuracy-saturated components may be correlated to the components of the analysis subspace, the correction concentrated on a portion of the full analysis field may cause an imbalance among the variables of the analysis field. We will determine whether this problem is crucial to the quality of the efficient ROI analysis or not and try to find a proper solution for the potential problem.

6. Acknowledgments

This work was supported by the Korea Meteorological Administration Research and Development Program under Grant CATER 2006-2201. The first author acknowledges financial support from the Brain Korea 21 Project and the Seoul Science Fellowship. We especially thank two anonymous reviewers for their invaluable comments, which contributed to clarifying the problem proposed by this paper and the technical detail to solve it.

References

Cohn, S. E., Sivakumaran, N. S. and Todling, R. 1994. A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.* **122**, 2838–2867.

Daley, R. 1991. *Atmospheric Data Analysis*. Cambridge University Press, New York.

Errico, R. M. 2003. The workshop on applications of adjoint models in dynamic meteorology. *Bull. Am. Meteorol. Soc.* **84**, 795–798.

Errico, R. M. and Raeder, K. D. 1999. An examination of the accuracy of the linearization of a mesoscale model with moist physics. *Quart. J. R. Meteorol. Soc.* **125**, 169–195.

Fertig, E. J., Harlim, J. and Hunt, B. R. 2007. A comparative study of 4D-VAR and a 4D Ensemble Kalman Filter: perfect model simulations with Lorenz-96. *Tellus* **59A**, 96–100.

Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J. and co-authors. 2004. Four-dimensional ensemble Kalman filtering. *Tellus* **56A**, 273–277.

Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, New York.

Le Dimet, F.-X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* **38A**, 97–110.

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

Lorenz, E. N. 1996. Predictability—a problem partly solved. In: *Proceedings on Predictability*, held at ECMWF on 4–8 September 1995. Reading, UK, 18pp.

Lorenz, E. N. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations: simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.

Miller, R. N., Ghil, M. and Gauthiez, F. 1994. Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.* **51**, 1037–1056.

Pires, C., Vautard, R. and Talagrand, O. 1996. On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus* **48A**, 96–121.

Rabier, F. and Courtier, P. 1992. Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. R. Meteorol. Soc.* **118**, 649–672.

Rabier, F., Thépaut, J.-N. and Courtier, P. 1998. Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quart. J. R. Meteorol. Soc.* **124**, 1861–1887.

Ruggiero, F. H., Michalakes, J., Nehrkorn, T., Modica, G. D. and Zou, X. 2006. Development and tests of a new distributed-memory MM5 adjoint. *J. Atmos. Ocean Sci.* **23**, 424–436.

Song, H.-J., Lim, G.-H., Lee, D.-I. and Lee, H.-S. 2009. Comparison of retrospective optimal interpolation with four-dimensional variational assimilation. *Tellus* **61A**, 428–437.

Todling, R., Cohn, S. E. and Sivakumaran, N. S. 1998. Suboptimal schemes for retrospective data assimilation based on the fixed-lag Kalman smoother. *Mon. Wea. Rev.* **126**, 247–259.

Zhang, F. 2005. Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Mon. Wea. Rev.* **133**, 2876–2893.

Zhu, J. and Kamachi, M. 2000. The role of time step size in numerical stability of tangent linear models. *Mon. Wea. Rev.* **128**, 1562–1572.

Zhu, Y., Todling, R., Guo, J., Cohn, S. E., Navon, I. M. and co-author. 2003 The GEOS-3 retrospective data assimilation system: the 6-hour lag case. *Mon. Wea. Rev.* **131**, 2129–2150.