4.11 Four-dimensional variational assimilation without using the adjoint model

Hyo-Jong Song and Gyu-Ho Lim*
Seoul National University, Korea, Seoul

Hee-Sang Lee
Korea Meteorological Administration, Korea, Seoul

1. INTRODUCTION

The sequential data assimilation method considers only observations up to and including the time of analysis. Moreover, no matter how many analyses for various time steps were performed during the assimilation processes, the precision of the final analysis has been restricted because of the growth of background error with integration of the forecast model. The restrictions that are not fit for a high quality analysis have rarely been referred in meteorological literatures (*Cohn et al.*, 1994; *Zhu et al.*, 2003). As a method of overcoming such a restriction, *Horel and Colman* (2005) emphasized recently the application of the retrospective analysis, by which observations after the time of analysis can be accounted for the past analysis.

The 4D-Var is a well-known retrospective analysis (*Klinker et al.*, 2000). The 4D-Var finds the most probable state to be true, for the given past, current, and future observations. The maximally probable solution of the 4D-Var is the state minimizing the cost function. At each time, the minimization algorithm determines an analysis increment that is used to reduce the value of the cost function with the aid of the adjoint model, and decides the final analysis field fulfilling the minimization of the cost function under the dynamical constraints (*Daley*, 1991). However, iteration may stop before the final analysis reaches the global minimum because of the existence of the local minima, which would appear for the strong nonlinearity of a forecast model employed (*Rabier and Courtier*, 1992). Such a multiple local minima problem can be overcome by employing the computing algorithm that starts from the point of departure more close to the global minimum than the initial background (*Evensen*, 1997).

We know that the variance of an analysis derived from a background field and observations in the optimal interpolation (OI) is smaller in magnitude than that of the background field. Therefore, the analysis error variance decreases as increasing the number of time-uncorrelated observations incorporated by the OI analysis. By

* *Corresponding author address*: Gyu-Ho Lim, School of Earth and Environmental Sciences, Seoul National University, NS80, Seoul 151-742 Korea; e-mail: gyuholim@snu.ac.kr

taking into account time-uncorrelated independent observations through our new implementation technique named retrospective optimal interpolation (ROI), we will have a better starting point which leads us to find the global minimum correctly.

This paper composed of as follows. The theoretical aspect of newly developed implementation technique of the 4D-Var is explained in section 2. In order to examine and confirm the validity of our technique, we applied the method to the three-variable Lorenz model and compared it with the 4D-Var based on the gradient descent method in section 3. In section 4, the conclusions are drawn with further works to be done in future.

2. Theory

2.1. *Basic Equations for the ROI*

The vectors $\mathbf{x}_a$ and $\mathbf{x}_b$ denote the analysis state vector and the background field, respectively, for the true state of m variables at time t_0. And $\mathbf{y}_o^0$ does p observations at the same time. Then, the basic equations for the OI comprise the optimal weight matrix

$$\mathbf{W} = \mathbf{B}(\mathbf{H}_{\mathbf{x}_b}^0)^{\mathrm{T}}[\mathbf{R}^0 + \mathbf{H}_{\mathbf{x}_b}^0 \mathbf{B}(\mathbf{H}_{\mathbf{x}_b}^0)^{\mathrm{T}}]^{-1}, \qquad (1)$$

the analysis vector

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{W}[\mathbf{y}_o^0 - H^0(\mathbf{x}_b)], \qquad (2)$$

and the error covariance matrix

$$\mathbf{P}_a = (\mathbf{I} - \mathbf{W}\mathbf{H}_{\mathbf{x}_b}^0)\mathbf{B}. \qquad (3)$$

In the above expressions, the $m \times m$ matrix $\mathbf{B}$ is the background error covariance matrix and the $p \times p$ matrix $\mathbf{R}^0$ is the observation error covariance matrix at time $t = t_0$ (*Kalnay*, 2003). The forward observation operator $H^0(\mathbf{x})$ converts the background field into the first guess field for the observations at $t = t_0$, and the $p \times m$ matrix $\mathbf{H}_{\mathbf{x}_b}^0$ is the first derivative of $H^0(\mathbf{x})$ with respect to $\mathbf{x}$ at $\mathbf{x} = \mathbf{x}_b$.

In order to assimilate observations $\mathbf{y}_o^1$ at future time t_1 to the previous analysis $\mathbf{x}_a(t = t_0)$ which have been obtained from $\mathbf{x}_b$ and $\mathbf{y}_o^0$, we must include a forecast model $M^1(\mathbf{x})$ which evolves $\mathbf{x}_a$ from t_0 to t_1 in a future observation operator $C^1(\mathbf{x})$, which transforms $\mathbf{x}_a$ into the guesses for observations at t_1. Equations for the ROI using $\mathbf{x}_a$ as a background field at $t = t_0$ may be written as

$$\begin{aligned}
&C^1(\mathbf{x}) \equiv H^1(M^1(\mathbf{x})), \\
&\mathbf{C}_{\mathbf{x}_a}^1 = \mathbf{H}_{M^1(\mathbf{x}_a)}^1 \mathbf{M}_{\mathbf{x}_a}^1, \\
&\mathbf{W}^{(1)} = \mathbf{P}_a(\mathbf{C}_{\mathbf{x}_a}^1)^{\mathrm{T}}[\mathbf{R}^1 + \mathbf{C}_{\mathbf{x}_a}^1 \mathbf{P}_a(\mathbf{C}_{\mathbf{x}_a}^1)^{\mathrm{T}}]^{-1}, \\
&\mathbf{x}_a^{(1)} = \mathbf{x}_a + \mathbf{W}^{(1)}[\mathbf{y}_o^1 - C^1(\mathbf{x}_a)], \\
&\mathbf{P}_a^{(1)} = (\mathbf{I} - \mathbf{W}^{(1)}\mathbf{C}_{\mathbf{x}_a}^1)\mathbf{P}_a,
\end{aligned} \qquad (4)$$

in which superscripts in and without parentheses mean the number of analyses in the time domain and the observation time, respectively, for the ROI.

In implementing the above equation sets for numerical calculations, it is a real challenge to inverse the weight matrix straightly, especially when taking into account the dimension of the operational model state (about 10^6). For reducing the dimension of the inverse matrix, we employ the identity that is a variant of the Sherman-Morrison-Woodbury formula (*Kalnay*, 2003; *Bishop et al.*, 2001; *Tippett et al.*, 2003),

$$\begin{aligned}
&\mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1} \\
&= (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}
\end{aligned}, \qquad (5)$$

and letting the $m \times r^{(i-1)}$ matrix $\mathbf{S}^{(i-1)}$ be the matrix square root of $\mathbf{P}_a^{(i-1)}$ of its rank $r^{(i-1)}$. We can formulate intuitively the equation set for the ROI with a background field $\mathbf{x}_a^{(i-1)}$ at the present time t_0 and an observation vector $\mathbf{y}_o^i$ at a future time t_i. They are as followings,

$$C^i(\mathbf{x}) \equiv H^i(M^i(\mathbf{x})) \quad (6)$$

$$\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}} = \mathbf{H}^i_{M^i(\mathbf{x}_a^{(i-1)})}\mathbf{M}^i_{\mathbf{x}_a^{(i-1)}} \quad (7)$$

$$\begin{aligned}\mathbf{W}^{(i)} &= [\mathbf{I}_r + (\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{S}^{(i-1)})^{\mathrm{T}}(\mathbf{R}^i)^{-1}\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{S}^{(i-1)}]^{-1} \\ &\times(\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{S}^{(i-1)})^{\mathrm{T}}(\mathbf{R}^i)^{-1}\end{aligned} \quad (8)$$

$$\mathbf{x}_a^{(i)} = \mathbf{x}_a^{(i-1)} + \mathbf{W}^{(i)}[\mathbf{y}_o^i - C^i(\mathbf{x}_a^{(i-1)})] \quad (9)$$

$$\begin{aligned}\mathbf{P}_a^{(i)} &= \mathbf{S}^{(i-1)}[\mathbf{I}_r + (\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{S}^{(i-1)})^{\mathrm{T}}(\mathbf{R}^i)^{-1}\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{S}^{(i-1)}]^{-1} \\ &\times(\mathbf{S}^{(i-1)})^{\mathrm{T}}\end{aligned} \quad (10)$$

,where $i \geq 1$, and $\mathbf{x}_a^{(0)}$ is set to be $\mathbf{x}_a$. Figure 1 shows the process of implementing the ROI schematically.

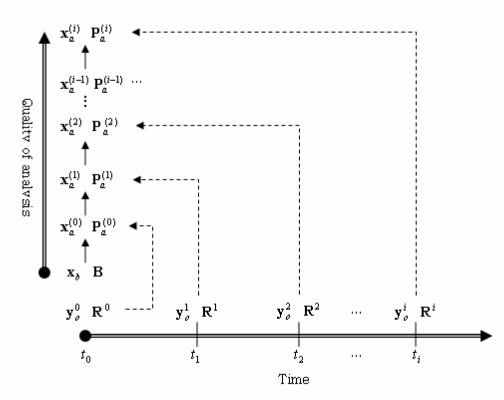

Fig. 1. A schematic procedure for the ROI implementation. The time of observations is marked with a superscript of an observation $\mathbf{y}$ and its error covariance matrix $\mathbf{R}$. The superscripts in parentheses denote the time of which observations are assimilated for the previous analysis and its error covariance matrix.

2.2. *Minimization of the Cost Function*

If we assume that the initial background $\mathbf{x}_b$ of an atmospheric state $\mathbf{x}$ at time t_0 is normally distributed around the true state $\mathbf{x}$ with the error covariance matrix $\mathbf{B}$, the probability density function (PDF) $P(\mathbf{x})$ of the state vector $\mathbf{x}$ is, referring to *Lorenc* (1986),

$$P(\mathbf{x}) \propto \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right]. \quad (11)$$

And if there are mutually independent observations $\mathbf{y}_o^0$, $\mathbf{y}_o^1$, …, and $\mathbf{y}_o^N$ at t_0, t_1, …, and t_N in order, which are Gaussian distributed with corresponding observation error covariance matrices $\mathbf{R}^0$, $\mathbf{R}^1$, …, and $\mathbf{R}^N$, the conditional joint PDF of the observations $\mathbf{y}_o^0$, $\mathbf{y}_o^1$, …, and $\mathbf{y}_o^N$ for a given $\mathbf{x}$ is

$$\begin{aligned}&P(\mathbf{y}_o^0,\mathbf{y}_o^1,\ldots,\mathbf{y}_o^N|\mathbf{x}) \propto \\ &\exp\left[-\frac{1}{2}\sum_{n=0}^{N}[\mathbf{y}_o^n - C^n(\mathbf{x})]^{\mathrm{T}}(\mathbf{R}^n)^{-1}[\mathbf{y}_o^n - C^n(\mathbf{x})]\right],\end{aligned} \quad (12)$$

where $C^0(\mathbf{x})$ is set to be $H^0(\mathbf{x})$. By the Bayesian theorem, the conditional PDF of the atmospheric state $\mathbf{x}$ for the given observations is

$$P(\mathbf{x}|\mathbf{y}_o^0,\mathbf{y}_o^1,\ldots,\mathbf{y}_o^N) \propto P(\mathbf{y}_o^0,\mathbf{y}_o^1,\ldots,\mathbf{y}_o^N|\mathbf{x})P(\mathbf{x}). \quad (13)$$

Substituting the equations (11) and (12) into the equation (13) and taking the natural logarithm of the equation will give

$$\begin{aligned}&\ln[P(\mathbf{x}|\mathbf{y}_o^0,\mathbf{y}_o^1,\ldots,\mathbf{y}_o^N)] = \\ &-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) \\ &-\frac{1}{2}\sum_{n=0}^{N}[\mathbf{y}_o^n - C^n(\mathbf{x})]^{\mathrm{T}}(\mathbf{R}^n)^{-1}[\mathbf{y}_o^n - C^n(\mathbf{x})] + const..\end{aligned} \quad (14)$$

If $\mathbf{x}$ is a solution with maximum probability, the cost function as defined below would have a minimum value.

$$\begin{aligned}J_N(\mathbf{x}) &\equiv \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) \\ &+\frac{1}{2}\sum_{n=0}^{N}[\mathbf{y}_o^n - C^n(\mathbf{x})]^{\mathrm{T}}(\mathbf{R}^n)^{-1}[\mathbf{y}_o^n - C^n(\mathbf{x})]\end{aligned}, \quad (15)$$

in which the subscript of the cost function means the number of future observation vectors except for the present one. Note that multiple local minima may be possible in (15) for the highly nonlinear observation operator $C^n(\mathbf{x})$. Hence, the value of $\mathbf{x}$ which makes the gradient of $J_N(\mathbf{x})$ zero may not be always the maximum probability solution.

By exploiting the tangent linear relationship between the better guess $\mathbf{x}_a^{(N)}$ and the true solution, we can approximate the above cost function into a quadratic equation:

$$\begin{aligned} J_N(\mathbf{x}) \approx & \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) \\ & +\frac{1}{2}\sum_{n=0}^{N}[\mathbf{y}_o^n - C^n(\mathbf{x}_a^{(N)}) - \mathbf{C}^n_{\mathbf{x}_a^{(N)}}(\mathbf{x}-\mathbf{x}_a^{(N)})]^{\mathrm{T}} \\ & \times(\mathbf{R}^n)^{-1}[\mathbf{y}_o^n - C^n(\mathbf{x}_a^{(N)}) - \mathbf{C}^n_{\mathbf{x}_a^{(N)}}(\mathbf{x}-\mathbf{x}_a^{(N)})]. \end{aligned} \quad (16)$$

Because the minimum of the quadratic equation is unique in general, the value of $\mathbf{x}$ which produces the zero gradient of the cost function is the maximum probability solution.

By the characteristic of the OI that gives the analysis with minimum variance at each analysis step, the weight on the observational increments decreases as the number of analyses in the ROI increases. Although the period of model integration increases with larger n, therefore, we can assume that $\mathbf{C}^n_{\mathbf{x}_a^{(N)}} = \mathbf{C}^n_{\mathbf{x}_a^{(n-1)}}$ and $C^n(\mathbf{x}_a^{(N)}) = C^n(\mathbf{x}_a^{(n-1)}) + \mathbf{C}^n_{\mathbf{x}_a^{(n-1)}}[\mathbf{x}_a^{(N)} - \mathbf{x}_a^{(n-1)}]$. If this assumption is valid, the gradient of the cost function may be written as

$$\begin{aligned} \nabla J_N(\mathbf{x}) = & \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) - \\ & \sum_{n=0}^{N}(\mathbf{C}^n_{\mathbf{x}_a^{(n-1)}})(\mathbf{R}^n)^{-1}[\mathbf{y}_o^n - C^n(\mathbf{x}_a^{(n-1)}) - \mathbf{C}^n_{\mathbf{x}_a^{(n-1)}}(\mathbf{x}-\mathbf{x}_a^{(n-1)})] \end{aligned}$$

(17)

With equation (5), we can change the analysis error covariance matrix (10) into

$$\mathbf{P}_a^{(i)} = [(\mathbf{P}_a^{(i-1)})^{-1} + (\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}})^{\mathrm{T}}(\mathbf{R}^i)^{-1}\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}]^{-1}, \quad (18)$$

and the weight matrix (8) into

$$\begin{aligned} \mathbf{W}^{(i)} = & [(\mathbf{P}_a^{(i-1)})^{-1} + (\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}})^{\mathrm{T}}(\mathbf{R}^i)^{-1}\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}]^{-1} \\ & \times(\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}})^{\mathrm{T}}(\mathbf{R}^i)^{-1} \end{aligned}. \quad (19)$$

And from the equation (9), (18), and (19), $\nabla J_N(\mathbf{x}_a^{(N)}) = \nabla J_{N-1}(\mathbf{x}_a^{(N-1)}) = \ldots = \nabla J_0(\mathbf{x}_a^{(0)}) = 0$ (*Kalnay*, 2003), that is, the gradient of the cost function for $\mathbf{x}_a^{(N)}$ becomes zero, which suggests the successful finding of the most probable atmospheric state only by the ROI. Therefore, it is not necessary any longer to perform the minimization process of the cost function.

2.3. *Differential factor*

Suppose that a nonlinear function $f(\mathbf{x})$ is analytic like the numerical weather prediction model, which expresses a set of partial differential equations as polynomial functions approximately. When a state $\mathbf{x}$ is represented by the sum of a reference state $\mathbf{x}_r$ and a deviation $\delta\mathbf{x}$ from the $\mathbf{x}_r$, by Taylor expansion,

$$\begin{aligned} f(\mathbf{x}_r+\delta\mathbf{x}) = & f(\mathbf{x}_r) + \left.\frac{\partial f}{\partial \mathbf{x}}\right|_{x=x_r}(\delta\mathbf{x}) \\ & + \frac{1}{2}\left.\frac{\partial^2 f}{\partial \mathbf{x}^2}\right|_{x=x_r}(\delta\mathbf{x})^2 + \sum_{j=3}^{\infty}\left[\frac{1}{j!}\left.\frac{\partial^j f}{\partial \mathbf{x}^j}\right|_{x=x_r}(\delta\mathbf{x})^j\right]. \end{aligned} \quad (20)$$

By introducing a differential factor α into (20),

$$\begin{aligned} \frac{1}{\alpha}[f(\mathbf{x}_r+\alpha\delta\mathbf{x}) - f(\mathbf{x}_r)] = & \left.\frac{\partial f}{\partial \mathbf{x}}\right|_{x=x_r}(\delta\mathbf{x}) \\ & + \frac{\alpha}{2}\left.\frac{\partial^2 f}{\partial \mathbf{x}^2}\right|_{x=x_r}(\delta\mathbf{x})^2 + \sum_{j=3}^{\infty}\left[\frac{\alpha^{j-1}}{j!}\left.\frac{\partial^j f}{\partial \mathbf{x}^j}\right|_{x=x_r}(\delta\mathbf{x})^j\right]. \end{aligned} \quad (21)$$

And when α is much smaller than 1 (*Barkmeijer et al.*, 1998), we get

$$\left.\frac{\partial f}{\partial \mathbf{x}}\right|_{x=x_r}(\delta\mathbf{x}) = \frac{1}{\alpha}[f(\mathbf{x}_r+\alpha\delta\mathbf{x}) - f(\mathbf{x}_r)]\,. \quad (22)$$

If the equation (22) is applied to the matrix $\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{s}_k^{(i-1)}$ containing the tangent linear model in (8), the matrix is given by

$$\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{s}_k^{(i-1)} = \frac{1}{\alpha_k}[C^i(\mathbf{x}_a^{(i-1)} + \alpha_k \mathbf{s}_k^{(i-1)}) - C^i(\mathbf{x}_a^{(i-1)})], \quad (23)$$

where the vector $\mathbf{s}_k^{(i-1)}$ is the k th column of the matrix square root $\mathbf{S}_k^{(i-1)}$. Consequently, by exploiting the equation (23), we may calculate the analysis error covariance and the optimal weight matrix using the original nonlinear model and obtain the maximum probability solution by the ROI without using the adjoint model, which, in most cases, should be revised whenever the forecast model is changed. The flowchart for the ROI utilizing the differential factor is given in Appendix A.

3. Numerical Experiments

3.1. *Experiment Design*

In order to ascertain the skill of the ROI, we carried out an observing system simulation experiment on the Lorenz model (*Lorenz*, 1963; *Miller et al.*, 1994):

$$\begin{aligned} \frac{dX}{dt} &= -\sigma X + \sigma Y, \\ \frac{dY}{dt} &= -XZ + rX - Y, \\ \frac{dZ}{dt} &= XY - bZ. \end{aligned} \quad (24)$$

The initial background is formed by adding the normally distributed random values, with the zero mean and the identity matrix $\mathbf{I}$ as its error covariance matrix, to the true model state (X_t, Y_t, Z_t) resulting from integrations of the model for 2000 time steps that start from $(0,1,0)$. The parameters σ, b, and r are set to be 10, 28, and $8/3$, which drives the system into a chaotic regime. Assimilated observations at nine time steps from t_0 to t_8 have normal distribution around a true state at the corresponding time with error covariance equal to its background. The time interval of 0.25 dimensionless time units was adopted as done by *Miller et al.* (1994) for their experiments.

3.2. *Results of Experiments*

In order to examine the proper functioning of our technique, a randomly generated error, of which the root-mean-squared (RMS) value for three model variables is 1, was evolved by exploiting the differential factor (DF), and compared with the exact result evolved by the tangent linear model (TLM). In Figure 2, the deviation of a result of the DF from that of the TLM decreases to be about zero as the differential factor goes from 1 to 0.001. We will set the factor to be 0.001 in the following experiments. And we want to examine how the total variance of the analysis changes as the number of analyses increases in the ROI. As clear in Figure 3, the final analysis with minimum variance is to be a preferred point of departure for deciding the global minimum without failure.

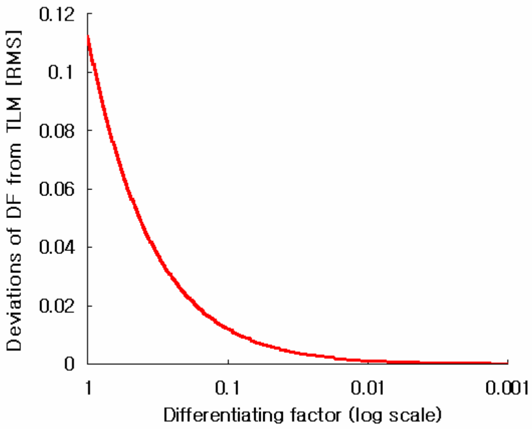

Fig. 2. Deviations in the evolution of errors by the

differential factor (DF) from that by the tangent linear model (TLM). The deviations are RMS values for three model variables.

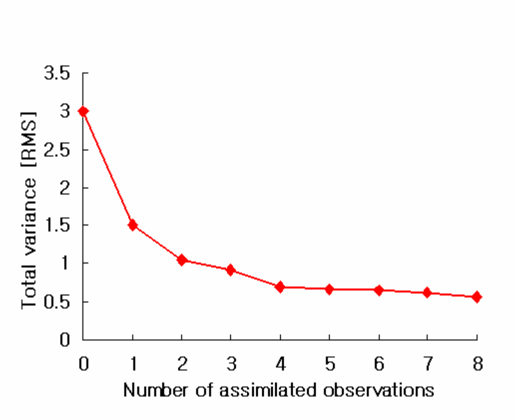

Fig. 3. Reduction of the total variance for the three model variables in the ROI based on the average of 30 times repetition. Refer to the main text and Figure 1 for the configuration of observations distributed in the assimilation window.

The cost function used in our experiments is same as the equation (15) of which $N = 8$. The minimization algorithm for the 4D-Var in the experiments is the gradient descent method. The procedure starts from the initial background, in which the step size varies inversely on the curvature of the cost function on the direction to be marched. The method furnishes the state vector which makes the gradient of the function zero. The analysis by the gradient descent method seems to be trapped in the convex of a local minimum near the initial background in Figure 4. From no longer decrease in the magnitude of the cost function against continuous iterations after reaching about 850 (Figure 4a) and the zero gradient of the function at the end of iteration (Figure 5), we are positive as to the existence of the local minimum.

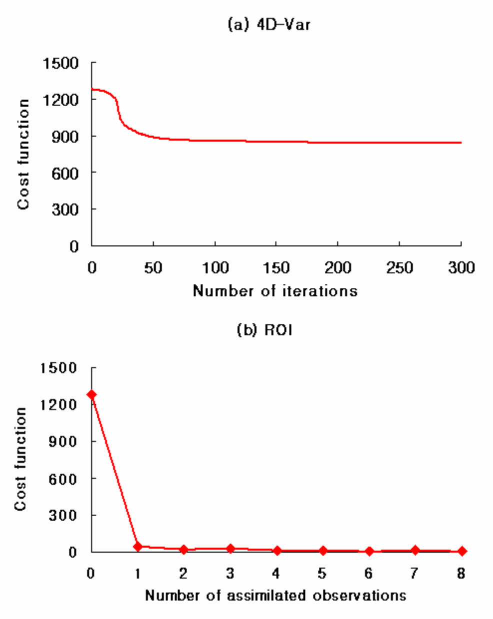

Fig. 4. The cost functions: (a) for the 4D-Var and (b) for the ROI.

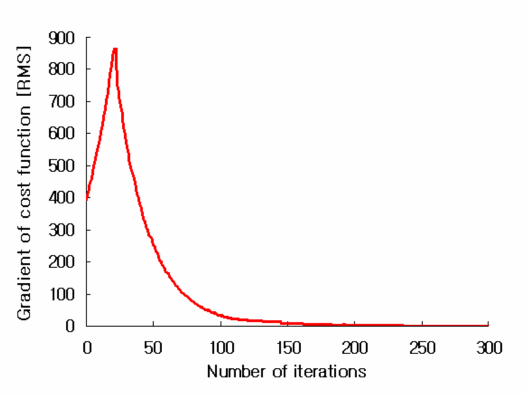

Fig. 5. The RMS gradient of the cost function for the 4D-Var in Figure 4.

Under these circumstances, the final value of the cost function by the ROI is 11.88 in Figure 4b. We should remember that there are many new technologies about the minimization of the cost function, and the method used for our calculation is a rudimentary one (*Klinker et al.*, 2000). We can say at least that the ROI works normally in spite of a local minimum near the background. To check whether it is necessary to perform the minimization process any more or not, we applied the minimization process to the final analysis which has resulted from the ROI. As evident in Figure 6, the cost function does not reduce any more explicitly. As the validity of the tangent linear assumption for the observation operator obviously depends on the observation time interval, we examined the proper functioning of the ROI and displayed the affirmative results in Figure 7.

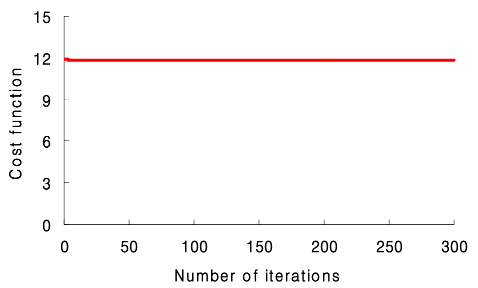

Fig. 6. Change of the value of the cost function during the minimizing process starting from the final analysis which has resulted from the ROI.

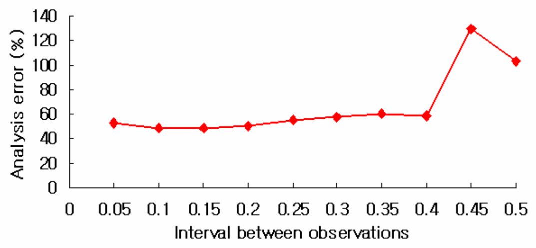

Fig. 7. Percentages of the RMS analysis error divided by the initial RMS background error based on the average of 30 times repetition.

4. Conclusion

In this paper, we proposed the theoretical basis of the ROI, which allowed us to determine the most probable atmospheric state at a given time by incorporating the observations after the analysis time, and confirmed that the technique gives the solution minimizing the cost function for the simple three-variable Lorenz model.

Even though the ROI does not demand the adjoint model, we are not free from the computing burden for the application of our method to operational forecast models because of the number of model integrations equal to the rank of the background error covariance matrix. However, such a burden could be mitigated by developing a parallel procedure for implementing our method and also by reducing the rank of the error covariance at each analysis step based on the feature of the variance-minimizing OI. The method for making the ROI efficient will be described concretely and verified in our future works.

Appendix A: A Flowchart for the ROI

The flowchart in Figure A1 shows a procedure in which the time-uncorrelated observations distributed in the assimilation window as shown in the Figure 1 are incorporated into an analysis $\mathbf{x}_a^{(N)}$ with its error covariance matrix $\mathbf{P}_a^{(N)}$. The meaning of all indices is the same as that in section 2. We use the equation sets (6) to (10) for the ROI, and specially the differential factor in the equation (23) for calculating the matrix $\mathbf{C}^i_{\mathbf{x}_a^{(i-1)}}\mathbf{s}_k^{(i-1)}$ in the analysis error covariance and the weight matrix.

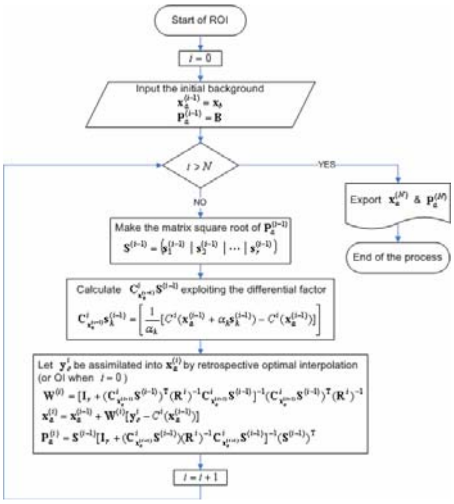

Fig. A1. A flowchart for the ROI process. Differential factors are employed for the calculation of the matrix $\mathbf{C}^{i}_{\mathbf{x}_a^{(i-1)}}\mathbf{s}_k^{(i-1)}$. Refer to the main text and Figure 1 for the configuration of analysis fields and observations in the iterations.

Acknowledgments.

This research was supported by the International Cooperation Research Program of the Ministry of Science & Technology and in part by the Brain Korea 21 Project and the Seoul Science Fellowship. We thank Baek-Min Kim for his valuable comments that helped to clarify and improve the paper.

References

Barkmeijer J., M. Van Gijzen, and F. Bouttier (1998), Singular vectors and estimates of the analysis error covariance metric, *Q. J. R. Meteorol. Soc.*, *124*, 1695-1713.

Bishop, C. H. and B. J., Etherton, and S. J. Majumdar (2001), Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, *Mon. Weather Rev.*, *129*, 420-436.

Cohn, S. E., N. S. Sivakumaran, and R. Todling (1994), A fixed-lat Kalman smoother for retrospective data assimilation, *Mon. Weather Rev.*, *122*, 2838-2867.

Daley R. (1991), *Atmospheric Data Analysis*, 457pp, Cambridge Univ. Press, New York.

Evensen, G. (1997), Advanced data assimilation in strongly nonlinear dynamical systems, *Mon. Weather Rev.*, *125*, 1342-1354.

Horel, J., and B. Colman (2005), Real-time and retrospective mesoscale objective analyses, *Bull. Am. Meteorol. Soc.*, *86*, 1477-1480.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, 341pp, Cambridge Univ. Press, New York.

Klinker, E., F. Rabier, G. Kelly, and J. F. Mahfouf (2000), The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration, *Q. J. R. Meteorol. Soc.*, *126*, 1191-1215.

Lorenc, A. C. (1986), Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, *112*, 1177-1194.

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

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

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

Tippett, M. K., J. L. Anderson, C. H. Bishop (2003), Ensemble square root filters, *Mon. Weather Rev.*, *131*, 1485-1490.

Zhu, Y., R. Todling, J. Guo, S. E. Cohn, I. M. Navon, and Y. Tyan (2003), The GEOS-3 retrospective data assimilation system: The 6-hour lag case, *Mon. Weather Rev.*, *131*, 2129-2150.