## 6.3A Estimation of observation impact without adjoint in an ensemble Kalman filter

Junjie Liu[1*], E. Kalnay[1]
[1]University of Maryland, College Park, MD

### 1. Introduction

In recent years, operational Numerical Weather Prediction centers are assimilating more satellite observations, such as the kilo-channel Advanced InfraRed Satellite (AIRS) and the Infrared Atmospheric Sounding Interferometer (IASI), in addition to in situ observations. Statistically, the assimilation of new observations improves, on average, the accuracy of short-range forecasts (e.g. Joiner *et al.* 2004). However, the value added to the forecast by various observations depends on the instrument type, observation type, and observation locations, as well as the presence of other observations. The knowledge of the impact that different observations have on the analyses and forecasts is important to better use the observations which have large impact on the forecasts, and to avoid using observations which have no impact or even negative impact on the forecasts.

Traditionally, the observation impact has been estimated by carrying out experiments in which part of observations used in the control experiment were excluded in the data-denial experiments (e.g., Zapotocny *et al.* 2000). However, this requires much computational time since a new analysis/forecast experiment has to be carried out for any subset of observations that needs to be evaluated. Langland and Baker (2004, LB hereafter) proposed an adjoint-based procedure to assess the observation impact on short-range forecasts without carrying out data-denial experiments. This adjoint-based procedure can evaluate the impact of any or all observations used in the data assimilation and forecast system on a selected measure of short-range forecast error. In addition, it can be used as a diagnostic tool to monitor the quality of observations, showing which observations make the analysis or the forecast worse, and can also give an estimate of the relative importance of observations from different sources. Following a similar procedure as LB, Zhu and Gelaro (2007) developed the adjoint of the Grid point Statistical Interpolation (GSI) assimilation system to estimate observation impact in a near-operational data assimilation system. They showed that their procedure provides accurate assessments of the forecast sensitivity with respect to most of the observations assimilated. This adjoint procedure to estimate observation impact requires the adjoint of the forecast model which is complicated to develop and not always available.

In this paper, we propose an ensemble-based sensitivity method to assess the observation impact as in LB but without using the adjoint model. It is different from the ensemble sensitivity method proposed by Ancell and Hakim (2007) as discussed in section 2. We compare the observation impact calculated from the ensemble sensitivity method we propose with that from the adjoint method, and further compare the impacts from both methods with the actual forecast error reduction due to assimilation of these observations in Lorenz-40 variable model (Lorenz and Emanuel, 1998). The paper is organized as follows: the derivation of the formula is given in section 2, and an alternative formula derivation is

[1*] *Corresponding author address, 2113 CSS, College Park, MD. Email: jjliu@atmos.umd.edu*

briefly given in Appendix B. The experimental design is presented in section 3, and the results are in section 4. Section 5 contains a summary and conclusions.

## 2. Derivation of the ensemble sensitivity method to calculate observation impact without using the adjoint model

### 2.1 The sensitivity of forecast errors to observations

Following LB, in order to study the observation impact on the reduction of forecast errors, we first calculate the sensitivity of a cost function at time t to the observations assimilated at time t=00hr (Figure 1). LB defined a cost function at time t as the energy forecast error norm difference between the forecasts from initial condition at 00hr (at a time when observations $\mathbf{y}_0^o$ were assimilated) and from initial conditions at -6hr that did not benefit from the use of the observations $\mathbf{y}_0^o$. Without loss of generality, and since we will test our calculation procedure in Lorenz-40 variable model, we define a cost function as the difference of squared forecast errors between the forecasts started at 00hr and -6hr and verified at time t:

$$J=\frac{1}{2}(\boldsymbol{\varepsilon}_{t|0}^T\boldsymbol{\varepsilon}_{t|0}-\boldsymbol{\varepsilon}_{t|-6}^T\boldsymbol{\varepsilon}_{t|-6})=\frac{1}{2}(\boldsymbol{\varepsilon}_{t|0}^T+\boldsymbol{\varepsilon}_{t|-6}^T)(\boldsymbol{\varepsilon}_{t|0}-\boldsymbol{\varepsilon}_{t|-6})$$
(1)

where $\boldsymbol{\varepsilon}_{t|0}=\overline{\mathbf{x}}_{t|0}^f-\overline{\mathbf{x}}_t^a$, and $\boldsymbol{\varepsilon}_{t|-6}=\overline{\mathbf{x}}_{t|-6}^f-\overline{\mathbf{x}}_t^a$, $\overline{\mathbf{x}}_t^a$ is the verification analysis at time t[2]. The purpose of the following derivation is to rewrite equation (1) as function of observations assimilated at 00hr, which does not depend on the choice of the verification state. We follow Bishop's (2007) notation, with the first sub-index indicating the verification time, and the second sub-index, separated by a vertical bar, indicating the time of the initial conditions of a forecast or forecast error, so that $\overline{\mathbf{x}}_{t|0}^f$ and $\overline{\mathbf{x}}_{t|-6}^f$ are the ensemble mean forecasts valid at time t, initialized at 00hr and -6hr respectively. Here, $K$ is the number of ensemble members, $\overline{\mathbf{x}}_{t|0}^f=\frac{1}{K}\sum_{i=1}^{K}M_{t|0}(\mathbf{x}_0^{ai})$,

$\overline{\mathbf{x}}_{t|-6}^f=\frac{1}{K}\sum_{i=1}^{K}M_{t|-6}(\mathbf{x}_{-6}^{ai})$, and $M_{t|0}$ and $M_{t|-6}$ represent the integration with the nonlinear model initialized with the analysis at 00hr and -6hr respectively. Substituting the definitions of $\boldsymbol{\varepsilon}_{t|0}$ and $\boldsymbol{\varepsilon}_{t|-6}$ into the above equation, the cost function can be written as:

$$J=\frac{1}{2}(2\boldsymbol{\varepsilon}_{t|-6}+\overline{\mathbf{x}}_{t|0}^f-\overline{\mathbf{x}}_{t|-6}^f)^T(\overline{\mathbf{x}}_{t|0}^f-\overline{\mathbf{x}}_{t|-6}^f) \quad (2)$$

In the following derivation, we aim to express the forecast difference $(\overline{\mathbf{x}}_{t|0}^f-\overline{\mathbf{x}}_{t|-6}^f)$ valid at time t as a function of the observation increments $\mathbf{v}_0=\mathbf{y}_0^o-\overline{\mathbf{y}}_{0|-6}^b$ at 00hr (Figure 1), so that the sensitivity of the cost function to the observations $\frac{\partial J}{\partial \mathbf{v}_0}$ will be a function of the observations at time 00hr. $\overline{\mathbf{y}}_{0|-6}^b=h(\overline{\mathbf{x}}_{0|-6}^b)$ is the

[2] The forecast length should be short enough that perturbations grow linearly, and long enough that the forecast error is much larger than the error of the verification analysis state. Following LB, we choose t=24hr. The verification analysis state can be either from the same or from a different analysis system as long as it is more accurate than the forecasts $\overline{\mathbf{x}}_{t|-6}^f$ and $\overline{\mathbf{x}}_{t|0}^f$.

prediction of the observations at t=00hr, with $h(\cdot)$ the nonlinear observation operator.

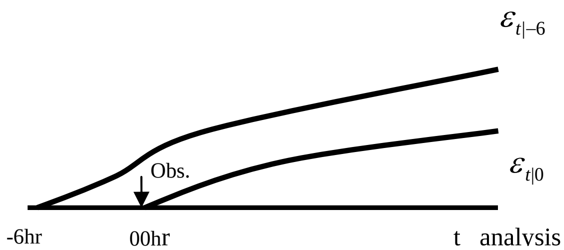


Figure 1 Schematic plot of the time relationship of the observation impact on the forecast error at time t. (After Langland and Baker, 2004, Fig 1.)

Following Hunt *et al.* (2007), the $i^{th}$ analysis ensemble member $\mathbf{x}_0^{ai}$ can be written as:

$$\mathbf{x}_0^{ai} = \overline{\mathbf{x}}_{0|-6}^{b} + \mathbf{X}_{0|-6}^{b}\tilde{\mathbf{K}}_0\mathbf{v}_0 + \delta\mathbf{x}_0^{ai} \qquad (3)$$

where $\mathbf{X}_{0|-6}^{b} = \left[\delta\mathbf{x}_{0|-6}^{b1} \quad | \quad \cdots \quad | \quad \delta\mathbf{x}_{0|-6}^{bK}\right]$ is a matrix whose $K$ columns are background ensemble perturbations with the $i^{th}$ column given by $\delta\mathbf{x}_{0|-6}^{bi} = \mathbf{x}_{0|-6}^{bi} - \overline{\mathbf{x}}_{0|-6}^{b}$. $\tilde{\mathbf{K}}_0 = \tilde{\mathbf{P}}_0^{a}\mathbf{Y}_0^{bT}\mathbf{R}_0^{-1}$ and $\tilde{\mathbf{P}}_0^{a} = \left[(K-1)\mathbf{I} + \mathbf{Y}_0^{bT}\mathbf{R}_0^{-1}\mathbf{Y}_0^{b}\right]^{-1}$ are the Kalman gain matrix and analysis error covariance matrix in the ensemble subspace spanned by ensemble forecasts. The $i^{th}$ column of the analysis perturbation matrix $\mathbf{X}_0^{a} = [(K-1)\tilde{\mathbf{P}}_0^{a}]^{\frac{1}{2}}$ is the analysis perturbation vector $\delta\mathbf{x}_0^{ai} = \mathbf{x}_0^{ai} - \overline{\mathbf{x}}_0^{a}$. $\mathbf{Y}_0^{b}$ is a matrix whose $i^{th}$ column is equal to $h(\mathbf{x}_{0|-6}^{bi}) - h(\overline{\mathbf{x}}_{0|-6}^{b})$. $\mathbf{R}_0$ is the observation error covariance. An over-bar represents an average over $K$ ensemble member, a tilde indicates that a vector or matrix is represented in the subspace of ensemble forecasts, and $\delta$ represents the difference between an ensemble member and the ensemble mean.

Based on equation (3), the forecast $\overline{\mathbf{x}}_{t|0}^{f} = \frac{1}{K}\sum_{i=1}^{K} M_{t|0}(\mathbf{x}_0^{ai})$ initialized at t=00hr is written as:

$$\overline{\mathbf{x}}_{t|0}^{f} = \frac{1}{K}\sum_{i=1}^{K} M_{t|0}(\overline{\mathbf{x}}_{0|-6}^{b} + \mathbf{X}_{0|-6}^{b}\tilde{\mathbf{K}}_0\mathbf{v}_0 + \delta\mathbf{x}_0^{ai}) \quad (4)$$

Note that although in the following derivation we make a linearization, the actual computation does not require either the tangent linear or adjoint model. We linearize equation (4) around the background mean state $\overline{\mathbf{x}}_{0|-6}^{b}$, and define the tangent linear model as $\mathbf{M}_{t|0}$, so that:

$$\overline{\mathbf{x}}_{t|0}^{f} \cong M_{t|0}(\overline{\mathbf{x}}_{0|-6}^{b}) + \frac{1}{K}\sum_{i=1}^{K}\mathbf{M}_{t|0}(\mathbf{X}_{0|-6}^{b}\tilde{\mathbf{K}}_0\mathbf{v}_0 + \delta\mathbf{x}_0^{ai})$$

(5)

From Hunt *et al.* (2007), $\delta\mathbf{x}_0^{ai} = \mathbf{X}_{0|-6}^{b}\delta\mathbf{w}_0^{ai}$, where $\delta\mathbf{w}_0^{ai}$ is the $i^{th}$ column of a symmetric $K$ by $K$ matrix of weight perturbations $\mathbf{W}_0^{a} = [(K-1)\tilde{\mathbf{P}}_0^{a}]^{1/2}$ with elements $\delta w_0^{aji}$.

Substituting $\delta\mathbf{x}_0^{ai} = \mathbf{X}_{0|-6}^{b}\delta\mathbf{w}_0^{ai}$ into equation (5),

$$\overline{\mathbf{x}}_{t|0}^{f} \cong M_{t|0}(\overline{\mathbf{x}}_{0|-6}^{b}) + \frac{1}{K}\sum_{i=1}^{K}\mathbf{M}_{t|0}\left[\mathbf{X}_{0|-6}^{b}(\tilde{\mathbf{K}}_0\mathbf{v}_0 + \delta\mathbf{w}_0^{ai})\right]$$

(6)

We define

$$\overline{\mathbf{x}}^{f}_{t|-6} = M_{t|0}(\overline{\mathbf{x}}^{b}_{0|-6}) \cong \frac{1}{K}\sum_{i=1}^{K} M_{t|-6}(\mathbf{x}^{ai}_{-6})$$

. In addition, we assume that t is short enough that perturbations grow linearly, so that $\mathbf{X}^{f}_{t|-6} = \mathbf{M}_{t|0}\mathbf{X}^{b}_{0|-6}$, where $\mathbf{X}^{f}_{t|-6} = \left[\delta\mathbf{x}^{f1}_{t|-6} \quad | \quad \cdots \quad | \quad \delta\mathbf{x}^{fK}_{t|-6}\right]$ is a matrix whose $K$ columns are forecast ensemble perturbations with the $i^{th}$ column $\delta\mathbf{x}^{fi}_{t|-6} = \mathbf{x}^{fi}_{t|-6} - \overline{\mathbf{x}}^{f}_{t|-6}$. Then, equation (6) becomes:

$$\begin{aligned}\overline{\mathbf{x}}^{f}_{t|0} &= \overline{\mathbf{x}}^{b}_{t|-6} + \frac{1}{K}\sum_{i=1}^{K}\mathbf{X}^{f}_{t|-6}(\tilde{\mathbf{K}}_0\mathbf{v}_0 + \delta\mathbf{w}^{ai}_0) \\ &= \overline{\mathbf{x}}^{b}_{t|-6} + \mathbf{X}^{f}_{t|-6}\tilde{\mathbf{K}}_0\mathbf{v}_0 + \frac{1}{K}\sum_{j=1}^{K}\mathbf{X}^{fj}_{t|-6}\sum_{i=1}^{K}\delta w^{aji}_0\end{aligned} \tag{7}$$

The last term in equation (7) vanishes, because the perturbation weights summed over either the $K$ columns or the $K$ rows are equal to one: $\sum_{j=1}^{K}\delta w^{aji}_0 = \sum_{i=1}^{K}\delta w^{aji}_0 = 1$ (Appendix A) and the average of the forecast ensemble perturbations is equal to zero. Therefore, we can write

$$\boldsymbol{\varepsilon}_{t|0} - \boldsymbol{\varepsilon}_{t|-6} = \overline{\mathbf{x}}^{f}_{t|0} - \overline{\mathbf{x}}^{f}_{t|-6} = \mathbf{X}^{f}_{t|-6}\tilde{\mathbf{K}}_0\mathbf{v}_0 \tag{8}$$

Similarly,

$$\begin{aligned}\boldsymbol{\varepsilon}_{t|0} + \boldsymbol{\varepsilon}_{t|-6} &= \overline{\mathbf{x}}^{f}_{t|0} - \overline{\mathbf{x}}^{a}_{t} + \overline{\mathbf{x}}^{f}_{t|-6} - \overline{\mathbf{x}}^{a}_{t} \\ &= \overline{\mathbf{x}}^{f}_{t|-6} - \overline{\mathbf{x}}^{a}_{t} + \overline{\mathbf{x}}^{f}_{t|-6} - \overline{\mathbf{x}}^{a}_{t} + \overline{\mathbf{x}}^{f}_{t|0} - \overline{\mathbf{x}}^{f}_{t|-6} \\ &= 2\boldsymbol{\varepsilon}_{t|-6} + \mathbf{X}^{f}_{t|-6}\tilde{\mathbf{K}}_0\mathbf{v}_0\end{aligned} \tag{9}$$

Equation (1) is then written as:

$$\begin{aligned}J &= \frac{1}{2}(\boldsymbol{\varepsilon}^{T}_{t|0} + \boldsymbol{\varepsilon}^{T}_{t|-6})(\boldsymbol{\varepsilon}_{t|0} - \boldsymbol{\varepsilon}_{t|-6}) \\ &= \frac{1}{2}[2\boldsymbol{\varepsilon}_{t|-6} + \mathbf{X}^{f}_{t|-6}\tilde{\mathbf{K}}_0\mathbf{v}_0]^{T}\mathbf{X}^{f}_{t|-6}\tilde{\mathbf{K}}_0\mathbf{v}_0\end{aligned} \tag{10}$$

If the model is nonlinear, equation (10) is a linear approximation of equation (1). Since the error $\boldsymbol{\varepsilon}_{t|-6}$ is not correlated with the observations assimilated at t=00hr, the sensitivity of the forecast error to the observation increments can be written as:

$$\frac{\partial J}{\partial \mathbf{v}_0} = \left[\tilde{\mathbf{K}}^{T}_0\mathbf{X}^{fT}_{t|-6}\right]\left[\boldsymbol{\varepsilon}_{t|-6} + \mathbf{X}^{f}_{t|-6}\tilde{\mathbf{K}}_0\mathbf{v}_0\right] \tag{11}$$

Note that the sensitivity of the cost function $J$ to the observation increments (equation (11)) does not require the adjoint model. When an independent analysis state verifying at time t and forecast perturbations initialized at -6hr and verifying at time t are available at 00hr, such as in a reanalysis mode, equation (11) can be calculated along with the data assimilation. This ensemble sensitivity method is different from Ancell and Hakim (2007) and Torn and Hakim (2007), who also proposed a method to calculate the forecast sensitivity to the observations without using adjoint model. In their approach, the sensitivity is a function of the inverse of the analysis error covariance, and it is a linear regression of analysis errors onto a given forecast metric. In Appendix B, we give another derivation of the sensitivity of the cost function $J$ to the observations without linearization, which gives results indistinguishable from those calculated from equation (11).

**2.2 Observation impact on forecasts**

As discussed in LB, the forecast sensitivity can be used to examine the actual observation impact on the forecast. Using the forecast sensitivity $\frac{\partial J}{\partial \mathbf{v}_0}$, the observation impact on the forecast can be written as:

$$J = \frac{1}{2}(\boldsymbol{\varepsilon}_{t|0}^T \boldsymbol{\varepsilon}_{t|0} - \boldsymbol{\varepsilon}_{t|-6}^T \boldsymbol{\varepsilon}_{t|-6}) = \left\langle \mathbf{v}_0, \frac{\partial J}{\partial \mathbf{v}_0} \right\rangle \quad (12)$$

Equation (12) expresses the forecast error difference as a function of observation increments $\mathbf{v}_0$. When the assimilated observations improve the forecast at time t, the forecast error is reduced, and the value calculated from equation (12) will be negative. When the assimilated observations degrade the forecast, the value calculated from equation (12) will be positive. Furthermore, if the observation errors between subsets of observations are not correlated, the cost function $J$ can be expressed as a sum of $J^l$:

$$J = \sum_{l=1}^{L} J^l = \sum_{l=1}^{L} \mathbf{v}_0^l \cdot \frac{\partial J}{\partial \mathbf{v}_0^l} \quad (13)$$

where $J^l$ is the observational impact caused by the $l^{th}$ subset of the observations $\mathbf{y}_0^o = \left[(\mathbf{y}_0^{o1})^T, \quad \cdots \quad, (\mathbf{y}_0^{oL})^T\right]^T$ and $\mathbf{v}_0^l = \mathbf{y}_0^{ol} - h^l(\overline{\mathbf{x}}_{0|-6}^b)$. Based on equation (13), we can calculate the observation impact on the forecasts from any subset of observations without conducting data denial experiments, and can also compare the importance of observations from different sources.

In the following sections, we will compare the observation impact calculated from ensemble sensitivity method we proposed with that using the adjoint method and compare both results with the actual observation impact calculated from equation (1). Further we will examine whether the ensemble sensitivity method we proposed can actually detect bad observations whose errors do not satisfy the Gaussian assumption $\varepsilon^o \to N(0,R)$ made in the assimilation. These experiments are carried out in the Lorenz-40 variable model.

**3. Experimental design**

The Lorenz 40-variable model is governed by the following equation:

$$\frac{d}{dt} x_j = (x_{j+1} - x_{j-2}) x_{j-1} - x_j + F \quad (14)$$

The variables ($x_j$, j=1…J) represent a “meteorological” variable on a “latitude circle” with periodic boundary conditions. As in Lorenz and Emanuel (1998), J is chosen to be equal to 40. The time step is 0.05, which corresponds to a 6-hour integration interval. F is the external forcing, which is 8 for the nature run, and 7.6 for the forecast, allowing for some model error in the system.

Following LB, we estimate the impact of the observations assimilated at 00hr on the forecast valid at t=24hr, so that in equation (1) the cost function is defined as the forecast error difference between a 24-hour forecast (initialized at 00hr) and a 30-hour forecast (initialized at -6hr). The error difference between these two forecasts is solely due to the assimilation of the observations at 00hr in the initial conditions of 24-hour forecast. The observations are observed at every grid point. We present experiments with “normal”, “larger random error” and “biased observation” cases. In the normal case, the assumed observation error standard deviation 0.2 does represent the actual error statistics for every observation obtained from the nature run plus a Gaussian random perturbation. In the “larger random error” case, the observation at a single grid point (the 11th grid point) has four times larger random error standard deviation than the other observations. However, in the data assimilation process, we still use the error standard deviation 0.2 to represent the error statistics for every observation, including the 11th grid point. Such an experiment simulates real cases when some observations may have larger (or smaller) random errors than assumed in the data assimilation system. In addition, real observations may also have biases, something especially common when we assimilate satellite observations (e.g., Derber and Wu, 1998). Therefore, in the "biased" case experiment, we include a bias equal to 0.5 in the observation at the 11th grid point, but still assume that the observation is unbiased during data assimilation.

We run each experiment for 7500 analysis cycles with the Local Ensemble Transform Kalman Filter (LETKF; Ott *et al.* 2004; Hunt *et al.* 2007) data assimilation scheme. The time average statistics shown in the next section are the average over the last 7000 analysis cycles. Throughout these experiments, we check whether our ensemble sensitivity method is comparable with the adjoint method of LB in assessing the observation impact on the forecast error, and compare the ability of both methods to detect poor quality observations. Unlike the variational assimilation approach, which needs to develop the adjoint of the assimilation system (Langland and Baker, 2004), the calculation of observation impact in an Ensemble Kalman filter (EnKF) does not need the adjoint of the assimilation system. This is true for both the adjoint method and the ensemble sensitivity method because the transpose of the Kalman gain matrix can be directly obtained. We will discuss the results case by case in the next section.

## 4. Results

### 4.1 Normal case

Figure 2 shows observation impact calculated from the adjoint method (grey line with plus signs), the ensemble sensitivity method (grey line with closed circles) and the actual spatially summed forecast error difference (black line with open circles) between the analysis cycles 5700 and 5780 for the "normal" case. It shows that the observation impact calculated from the ensemble sensitivity method is similar to the result from adjoint sensitivity method, and both methods succeed in capturing most of the actual forecast improvement due to assimilation of the observations at 00hr.

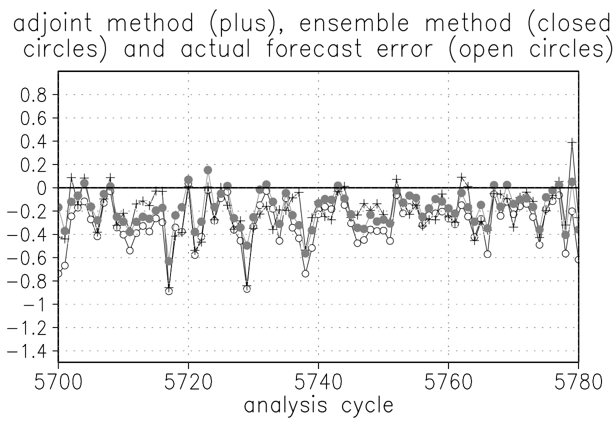


Figure 2 Evolution of spatially summed forecast error differences and the observation impact for the normal case between analysis cycle 5700 and 5780. Black line with open circles: the actual forecast error difference between 24-hour forecast and the 30-hour forecast; black line with plus signs: the observation impact calculated from the adjoint method; grey line with closed circles: the observation impact calculated from the ensemble method; zero line: no impact.

### 4.2 Larger random error case

When the observation has four times larger random error standard deviation at the 11th grid point than at the other locations, both the ensemble sensitivity method and the adjoint method show that the assimilation of this observation increases the forecast error (Figure 3). The signal from ensemble sensitivity method at the 11th observation location is larger than that of the adjoint method, but elsewhere, the observation impact calculated from both methods has similar values. Conversely, if an observation has consistently smaller observation error than assumed, its contribution to the reduction of forecast errors is larger than other observations (not shown). It is interesting to note that the observations at adjacent locations (e.g., 12th grid point) have larger impact in improving the forecast accuracy than at the other locations due to the larger weights given to these observations through the larger background uncertainty estimated along with the LETKF. Snapshots of the spatially summed impact show that the observation impact calculated from both methods reflects the actual forecast error difference (Figure 4) even when one of the observations has erroneous error statistics. Because of the poor quality of the observation at the 11th observation location, the domain summed observation impact has some large spikes (Figure 4).

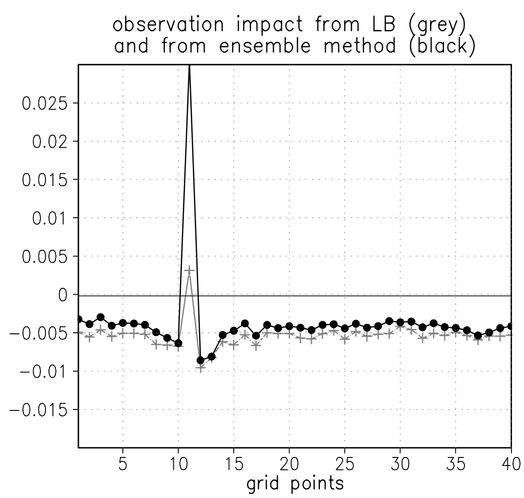


Figure 3 Time average (over the last 7000 analysis cycles) of the contribution to the reduction of the forecast error from each observation location (four times larger random error at the 11th grid point). Black line with closed circles is from ensemble method, and the grey line with plus signs is from adjoint method, and the black solid line is zero line.

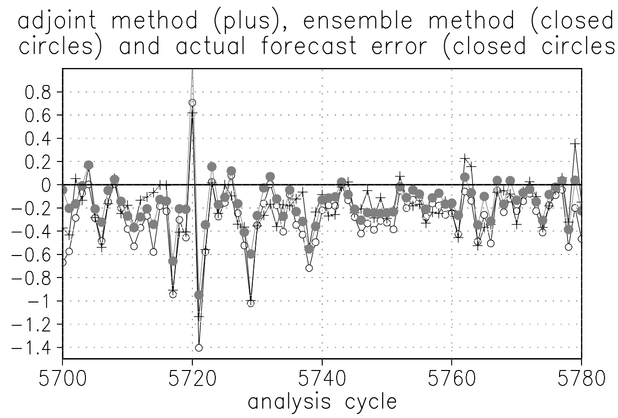


Figure 4 Evolution of spatially summed forecast error differences and the observation impact from the larger random error case between analysis cycle 5700 and 5780. The notations are the same as in Figure 2.

### 4.3 Biased case

When the 11th observation location has a bias, the ensemble sensitivity method (Figure 5) indicates, like the adjoint method, that the assimilation of this observation increases the forecast error. Again, the negative impact from assimilation of this observation makes the positive impact (reduction of forecast error) of assimilating the adjacent observations larger.

These examples show that the ensemble sensitivity method gives observation impact similar with that from adjoint method, and both methods accurately reflects the actual forecast error reduction due to assimilation of the observations at 00hr. Like the adjoint method, the ensemble method can detect observations that have poor quality either with larger random error or bias, and the signal detected by the ensemble sensitivity method is stronger. When we reduce the observation coverage, the same conclusion is still valid (not shown).

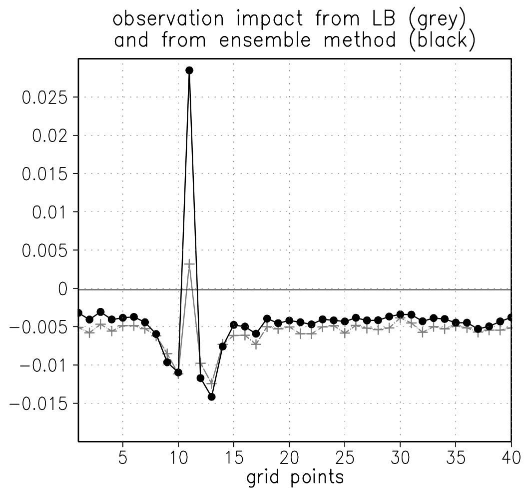


Figure 5 The biased case with a bias equal to 0.5 at 11th grid point. The notations are the same as in Figure 3.

## 5. Summary

The observations are the central information introduced into the forecast system during data assimilation. However, the quality and impact of observations is always different due to the magnitude of observation error, observation locations and model dynamics. Accurately monitoring the quality and impact of the observations assimilated in the system can help to delete the observations that routinely degrade the forecast, and to better use the observations that have larger impact on the forecast than the other observations.

In the past, monitoring the quality of observations has been based on observational increments, but we have found that the observation sensitivity approach is more effective in detecting poor observations than the monitoring of observational increments (not shown). In this paper, following LB, we proposed an ensemble sensitivity method to measure observation impact on the reduction of forecast error due to assimilation of observations. Unlike the adjoint method by LB, the ensemble sensitivity method we proposed does not need the adjoint model. We compared the ensemble sensitivity method to the adjoint method using Lorenz-40 variable model. The results show that the ensemble sensitivity method gets similar observation impact as the adjoint method, and both reflects most of the actual forecast improvement. Both methods can detect the “bad” observations that are of poor quality, with either larger random errors or with biases, and the ensemble sensitivity method shows a stronger signal in such scenarios. Like the adjoint method by LB, this method can be applied in the observation quality control as well as to compare the importance of different type observations. With the verification analysis available, either from the same or a different analysis system, this method could be routinely calculated within an ensemble Kalman filter, thus providing a powerful tool to understand cases of forecast failure and to tune the observation error statistics.

## Acknowledgements

We are very grateful to Dr. Kayo Ide, Dr. Shu-Chih Yang and Dr. Istvan Szunyogh for their suggestions and discussions, and to Dr. Ron Gelaro for his encouragement. This work was supported by a NASA grant through NNG04GK29G.

## References

Ancell, B. and Hakim, G. J., 2007: Comparing adjoint and ensemble sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.* (submitted).

Bishop, C. H, 2007: Introduction to data assimilation research at NRL and flow adaptive error covariance localization. http://www.weatherchaos.umd.edu/workshop/Bishop_UMD_workshop.pdf

Derber, J. C. and Wu, W.-S., 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.,* **126,** 2287-2299.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I., 2007: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter. *Physics D*., **230**, 112-126.

Joiner, J., Poli, P., Frank, D., and Liu, H. C., 2004: Detection of cloud-affected AIRS channels using an adjacent-pixel approach. *Quart. J. Roy. Meteor. Soc.,* **130**, 1469-1487.

Langland, R. H. and Baker, N. L. 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56a**, 189-201.

Liu, J., 2007: Applications of the LETKF to adaptive observations, analysis sensitivity, observation impact, and assimilation of Moisture. Ph. D thesis, University of Maryland.

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

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A., 2004: A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus*, **56A**, 415-428.

Torn, R. D., and Hakim, G. J., 2007: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.* (submitted)

Zapotocny, T. H., Menzel, W. P., Nelson, J. P., and Jung, J. A., 2002: An impact study of five remotely sensed and five in situ data types in the eta data assimilation system. *Weather and forecasting,* **17,** 263-285.

Zhu, Y. and Gelaro, R., 2007: Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. *Mon. Wea. Rev.* (accepted).

## APPENDIX A: Perturbation weights averaged over the ensemble

The following derivation is based on Hunt *et al.* (2007). We define a column vector of K ones: $\mathbf{v} = (1, \; 1, \; \cdots, \; 1)^T$. $\mathbf{v}$ is an eigenvector of $\tilde{\mathbf{P}}^a$ with eigenvalue $(K-1)^{-1}$: $(\tilde{\mathbf{P}}^a)^{-1}\mathbf{v} = \left[(K-1)\mathbf{I} + (\mathbf{Y}^b)^T \mathbf{R}^{-1}\mathbf{Y}^b\right]\mathbf{v} = (K-1)\mathbf{v}$ because the sum of the columns of $\mathbf{Y}^b$ is zero. Therefore,

$$\frac{1}{K-1}(\tilde{\mathbf{P}}^a)^{-1}\mathbf{v} = \mathbf{v} \tag{A1}$$

In addition, the matrix of analysis weights is given by $\mathbf{W}_0^a = [(K-1)\tilde{\mathbf{P}}_0^a]^{1/2}$, so that

$$\mathbf{W}_0^a \mathbf{W}_0^{aT} = (K-1)\tilde{\mathbf{P}}_0^a \tag{A2}$$

Multiplying both sides by the vector $\mathbf{v}$, we get $\mathbf{W}_0^a \mathbf{W}_0^{aT}\mathbf{v} = \mathbf{v}$, so that $\mathbf{v}$ is an eigenvector of $\mathbf{W}_0^a \mathbf{W}_0^{aT}$ matrix with eigenvalue equal to 1. Based on the properties of a symmetric matrix, $\mathbf{v}$ is also an eigenvector of $\mathbf{W}_0^a$ matrix with the eigenvalue equal to 1:

$$\mathbf{W}_0^a \mathbf{v} = \mathbf{v} \tag{A3}$$

Since $\mathbf{v}$ is a column vector of K ones, $\sum_{i=1}^{K} \delta w_0^{a(ji)} = 1$, where $\delta w_0^{a(ji)}$ is an element of the $\mathbf{W}_0^a$. $\mathbf{W}_0^a$ is a symmetric matrix, therefore, we have the following equation:

$$\sum_{j=1}^{K} \delta w_0^{a(ji)} = \sum_{i=1}^{K} \delta w_0^{a(ji)} = 1 \tag{A4}$$

## Appendix B: Another derivation of the sensitivity of forecast errors to observations

Unlike the derivation in section 2, this derivation does not use linearization. Instead, it assumes the mean forecast at time t can be obtained with the same weights as obtained in the ensemble analyses. More details are in Liu (2007).

We first express the $i^{th}$ analysis ensemble member at time 00hr as a linear combination of the six-hour ensemble forecasts initialized at -6hr. Based on Hunt *et al.* (2007), the $i^{th}$ analysis ensemble member is given by:

$$\mathbf{x}_0^{ai} = \overline{\mathbf{x}}_{0|-6}^{b} + \mathbf{X}_{0|-6}^{b}\mathbf{w}_0^{ai} \tag{B1}$$

where $\mathbf{w}_0^{ai} = \overline{\mathbf{w}}_0^{a} + \delta\mathbf{w}_0^{ai}$ is a vector of $K$ dimension, whose elements are $w_0^{aji} = \overline{w}_0^{aj} + \delta w_0^{aji}$, $j = 1, \cdots, K$. $\overline{\mathbf{w}}_0^{a} = \tilde{\mathbf{K}}_0 \mathbf{v}_0$ is the mean weighting vector, and the perturbation weighting vector $\delta\mathbf{w}_0^{ai}$ is the $i^{th}$ column of the $K$ by $K$ matrix $\mathbf{W}_0^{a} = [(K-1)\tilde{\mathbf{P}}_0^{a}]^{1/2}$ with the elements $\delta w_0^{aji}$. After expanding the terms on the right hand side of equation (B1) based on the definitions of each term and combining the same term together, we get:

$$\mathbf{x}_0^{ai} = \sum_{j=1}^{K} \mathbf{x}_{0|-6}^{bj}\left(\overline{w}_0^{aj} - \overline{\overline{w}}^{a} + \delta w_0^{aji}\right) \tag{B2}$$

where $\overline{\overline{w}}_0^{a} = \frac{1}{K}\sum_{j=1}^{K}\overline{w}_0^{aj}$, and $\overline{w}_0^{aj}$ is the $j^{th}$ element of the mean weight vector $\overline{\mathbf{w}}_0^{a}$.

We estimate the $i^{th}$ ensemble forecast at time t initialized at t=00hr with the ensemble forecasts initialized at t=-6hr using the same weights as at the analysis time:

$$\mathbf{x}_{t|0}^{i} = \sum_{j=1}^{K} \mathbf{x}_{t|-6}^{fj}(\overline{w}_0^{aj} - \overline{\overline{w}}^{a} + \delta w_0^{aij}) + \mathbf{e}_{t|0}^{i} \tag{B3}$$

where $\mathbf{e}_{t|0}^{i}$ represent the error from this approximation. We take an ensemble average of these forecasts initialized at t=00h, so the mean forecast initialized at 00hr can be written as:

$$\overline{\mathbf{x}}_{t|0}^{f} = \overline{\mathbf{x}}_{t|-6}^{f} + \sum_{j=1}^{K} \mathbf{x}_{t|-6}^{fj} \delta\overline{w}_{0}^{aj} + \overline{\mathbf{e}}_{t|0} \tag{B4}$$

In getting equation (4), we use the relationship $\sum_{j=1}^{K} \delta w_{0}^{a(ji)} = \sum_{i=1}^{K} \delta w_{0}^{a(ji)} = 1$. We note that, although very small, the error $\overline{\mathbf{e}}_{t|0} = \overline{\mathbf{x}}_{t|0}^{f} - \overline{\mathbf{x}}_{t|-6}^{f} - \sum_{j=1}^{K} \mathbf{x}_{t|-6}^{fj} \delta\overline{w}_{0}^{aj}$ cannot be neglected in order to obtain accurate observation sensitivity. After expressing $\delta\overline{\mathbf{w}}_{0}^{a}$ with elements $\delta\overline{w}_{0}^{aj}\,(\mathrm{j}=1, \ \cdots \ K)$ in terms of the increments $\mathbf{v}_{0}$, equation (B4) can be written as

$$\overline{\mathbf{x}}_{t|0}^{f} = \overline{\mathbf{x}}_{t|-6}^{f} + \mathbf{X}_{t|-6}^{f} \delta\widetilde{\mathbf{K}}_{0} \mathbf{v}_{0} + \overline{\mathbf{e}}_{t|0} \tag{B5}$$

where $\mathbf{X}_{t|-6}^{f} = \left[\mathbf{x}_{t|-6}^{f1} \ \ | \ \ \cdots \ \ | \ \ \mathbf{x}_{t|-6}^{fK}\right]$ is a matrix whose $K$ columns are background ensemble forecasts. Note that this notation is different from that in section 2. $\delta\widetilde{\mathbf{K}}_{0}$ is a $K$ by $P$ matrix whose elements are $\delta\widetilde{K}_{0}^{jp}$, and $\delta\widetilde{K}_{0}^{jp} = \widetilde{K}_{0}^{jp} - \overline{\widetilde{K}}_{0}^{p}$, $\overline{\widetilde{K}}_{0}^{p} = \frac{1}{K}\sum_{j=1}^{K} \widetilde{K}_{0}^{jp}$, $\widetilde{K}_{0}^{jp}$ is an element of $K$ by $P$ matrix $\widetilde{\mathbf{K}}_{0}$.

Based on equation (B5), the cost function is written as:

$$\begin{aligned} J &= \frac{1}{2}(2\boldsymbol{\varepsilon}_{t|-6}^{T} + \overline{\mathbf{x}}_{t|0}^{f} - \overline{\mathbf{x}}_{t|-6}^{f})(\overline{\mathbf{x}}_{t|0}^{f} - \overline{\mathbf{x}}_{t|-6}^{f}) \\ &= \frac{1}{2}[2\boldsymbol{\varepsilon}_{t|-6} + \mathbf{X}_{t|-6}^{f} \delta\widetilde{\mathbf{K}}_{0} \mathbf{v}_{0} + \overline{\mathbf{e}}_{t|0}]^{T} (\mathbf{X}_{t|-6}^{f} \delta\widetilde{\mathbf{K}}_{0} \mathbf{v}_{0} + \overline{\mathbf{e}}_{t|0}) \end{aligned} \tag{B6}$$

Since the error $\boldsymbol{\varepsilon}_{t|-6}$ is not correlated with the observations assimilated at t=00hr, the sensitivity of the forecast error to observations is written as:

$$\frac{\partial J}{\partial \mathbf{v}_0} = \left[\delta\tilde{\mathbf{K}}_0^T \mathbf{X}_{t|-6}^{fT}\right]\left[\boldsymbol{\varepsilon}_{t|-6} + \mathbf{X}_{t|-6}^f \delta\tilde{\mathbf{K}}_0 \mathbf{v}_0 + \overline{\mathbf{e}}_{t|0}\right] \quad \text{(B7)}$$

The sensitivity to the observations in equation (B7) can be directly calculated based on the weighting function from data assimilation at 00hr, the observation increment at 00hr, and the ensemble forecast initialized at -6hr. This derivation neglects the correlation between $\overline{\mathbf{e}}_{t|0}$ and the observations assimilated at t=00hr. If we replace $\overline{\mathbf{e}}_{t|0}$ from (B4) into equation (B7) and linearize as in section 2, we get equation (11). Both (B7) and (11) yield indistinguishable results.