Review of self-sensitivity calculation in an Ensemble Kalman filter and its extended applications to forecast impacts

Junjie Liu[1] and Eugenia Kalnay[2]

Earth and Planetary Science department, U. of California, Berkeley, U. S. A.
Atmosph. and Oceanic Science department, U. of Maryland, College Park, U.S.A.

Abstract

Self-sensitivity reflects the sensitivity of analysis value to observation. As shown in Liu et al. (2009), self-sensitivity can be calculated without approximation when observation error is not correlated in ensemble Kalman filters. With self-sensitivity, observation and analysis value, analysis value change at the i^{th} observation location with deletion of the i^{th} observation can be calculated without carrying out data-denial experiments. This paper has extended the application of self-sensitivity to the forecast value change. We have derived an equation to calculate the forecast value $y_i^{f(-i)}$ without carrying out data-denial experiments. The impact of the i^{th} observation on the forecast accuracy at the i^{th} point is further defined based on $y_i^{f(-i)}$.

The results with Lorenz 40-variable model show that $y_i^{f(-i)}$ calculation based on self-sensitivity, forecast perturbations and the other quantities is a good approximation of the actual value calculated from much more expensive data-denial experiments. The impact of each observation on the forecasts at those locations based on $y_i^{f(-i)}$ can detect the bad quality observations, and reflect the actual observation impact. In a perfect model experimental setup, the observation impact calculation based on $y_i^{f(-i)}$ in primitive equation global model shows larger impact of the observations over data sparse areas, and smaller impact of the observations over data dense regions, which is consistent with Liu et al. (2009).

1. Introduction

Self-sensitivity, a quantity that is a function of the analysis error covariance and observation error covariance, indicates the sensitivity of the analysis value to the variations of the observations at the same location. The calculation of this quantity and the related diagnostics have been discussed in both variational and ensemble data assimilation frameworks (Cardinali et al., 2004; Liu et al., 2009). In variational data assimilation scheme, since the analysis error covariance is not explicitly calculated, the self-sensitivity is based on approximate analysis error covariance calculated from truncated eigenvalue decomposition (Cardinali et al., 2004). In ensemble Kalman filters (EnKF) (Evensen, 1994; Anderson, 2001; Bishop et al., 2001; Houtekamer and Mitchell, 2001; Whitaker and Hamill, 2002; Ott et al., 2004; Hunt et al., 2007), the analysis error covariance can be directly estimated from the ensemble analyses in each analysis cycle, so that the self-sensitivity calculation requires no approximation when there is no correlation in observation error covariance. Liu et al. (2009) have shown that the change in the analysis value at the i^{th} observation location with the deletion of the i^{th} observation can be computed without carrying out data-denial experiments based on this self-sensitivity, observation and analysis value in EnKFs. In this paper, we extend the application of self-sensitivity to calculate the change in the forecast value without data-denial experiments, and its applications on inferring the i^{th} observation impact on the i^{th} forecast accuracy. The observation impact we discuss here is different from the adjoint (e.g., Langland and Baker, 2004; Zhu and Gelaro, 2008; Cardinali, 2009) and the ensemble sensitivity method (Liu and Kalnay, 2008; Li et al., 2009) (see section 3).

This paper is organized as follows: section 2 reviews the self-sensitivity calculation in EnKFs and derives the equation to calculate the change in the forecast value at the i^{th} observation location with the deletion of the i^{th} observation at the analysis time. Section 3 discusses the method to calculate the impact of the i^{th} observation assimilated at the analysis time on the short-term forecast accuracy at the same location, as well as the differences between the observation impact discussed here and the observation impact from the adjoint method (e.g., Langland and Baker, 2004) and the ensemble sensitivity method (Liu and Kalnay, 2008, Li et al., 2009). In section 4, with Lorenz 40-variable model (Lorenz and Emanuel, 1998), we verify the equation derived in section 2, and calculate observation impact based on the method discussed in section 3. In section 5, with a primitive equation model and perfect experiments, we further show the observation impact on the short-term forecast accuracy based on the method discussed in section 3. Section 6 is summary and discussions.

2. Review of self-sensitivity calculation in EnKFs and derivation of $y_i^{f(-i)}$ calculation without carrying out data-denial experiments

As shown in Liu et al. (2009), in EnKFs, the sensitivity of the analysis vector $\mathbf{x}^a$ $(m \times 1)$ to the observation $\mathbf{y}^o$ $(p \times 1)$ is given by:

$$\mathbf{S}^o = \frac{\partial \mathbf{y}^a}{\partial \mathbf{y}^o} = \frac{1}{n-1}\mathbf{R}^{-1}(\mathbf{HX}^a)(\mathbf{HX}^a)^T \tag{1}$$

where $\mathbf{R}$ is the observation error covariance, $\mathbf{y}^a$ is the projection of the analysis vector $\mathbf{x}^a$ on the observation space. $\mathbf{HX}^a$ is the analysis ensemble perturbation matrix in the observation space, whose i^{th} column is

$$\mathbf{HX}^{ai} \cong h(\mathbf{x}^{ai}) - \frac{1}{n}\sum_{i=1}^{n} h(\mathbf{x}^{ai}) \tag{2}$$

$\mathbf{x}^{ai}$ is the i^{th} analysis ensemble member, n is the total number of ensemble analyses, and $h(\cdot)$ is the observation operator, which can be linear or nonlinear. When the observation operator is linear, the right hand side and the left hand side of Equation (2) are equal. Otherwise, the left hand side is a linear approximation of the right hand side. The diagonal elements of $\mathbf{S}^o$ are called as self-sensitivity. With uncorrelated observation errors (the covariance $\mathbf{R}$ is diagonal), self-sensitivity can be written as:

$$S_{ii}^o = \frac{\partial \mathbf{y}_i^a}{\partial \mathbf{y}_i^o} = \left(\frac{1}{n-1}\right)\frac{1}{\sigma_i^2}\sum_{j=1}^{n}(\mathbf{HX}^{aj})_i \times (\mathbf{HX}^{aj})_i \tag{3}$$

where σ_i^2 is the i^{th} observation error variance. With self-sensitivity, given the observation y_i^o and the analysis value y_i^a, the change in the analysis value y_i^a by leaving out the i^{th} observation can be computed without actually calculating $y_i^{a(-i)}$ (Cardinali et al., 2004):

$$y_i^a - y_i^{a(-i)} = S_{ii}^o(1 - S_{ii}^o)^{-1}(y_i^o - y_i^a) \tag{4}$$

$y_i^{a(-i)}$ is the analysis value at the i^{th} observation location when the i^{th} observation is not assimilated during data assimilation.

Deletion of the i^{th} observation changes the analysis value as well as the short-term forecast value. The change in the short-term forecast at the i^{th} observation location can be written as:

$$y_{i,(t|0)}^{f} - y_{i,(t|0)}^{f(-i)} = \mathbf{H}_i[M_{t|0}(\mathbf{x}^a)] - \mathbf{H}_i[M_{t|0}(\mathbf{x}^{a(-i)})] \tag{5}$$

where $y_{i(t|0)}^{f(-i)}$ is the forecast value at the i^{th} observation location when the i^{th} observation is not assimilated at the analysis time; $y_{i,(t|0)}^{f}$ is a forecast valid at time t started at the analysis t=0; $M_{t|0}(\mathbf{x}^a)$ is the short-term forecast when all the observations are assimilated at the analysis time; $M_{t|0}(\mathbf{x}^{a(-i)})$ is the short-term forecast when the i^{th} observation is left out at the analysis time; and t is the forecast length. M is the nonlinear forecast model; $\mathbf{H}_i$ is a $(1\times m)$ matrix that projects the forecast value to the i^{th} observation location. Since all the forecast states are integrated from the analysis time to time t, the subindex $t\,|\,0$ is omitted from the following equations. With Taylor expansion and the linearized model $\mathbf{M}$, Equation (5) can be further rewritten as:

$$y_i^f - y_i^{f(-i)} = \mathbf{H}_i[M(\mathbf{x}^a) - M(\mathbf{x}^{a(-i)})] \cong \mathbf{H}_i[\mathbf{M}(\mathbf{x}^a - \mathbf{x}^{a(-i)})] \tag{6}$$

Since both $\mathbf{M}$ and $\mathbf{H}_i$ are linear, the order of these two operators can be exchanged so that:

$$y_i^f - y_i^{f(-i)} \cong \mathbf{M}[\mathbf{H}_i(\mathbf{x}^a - \mathbf{x}^{a(-i)})] = \mathbf{M}[y_i^a - y_i^{a(-i)}] \tag{7}$$

where $\mathbf{H}_i(\mathbf{x}^a - \mathbf{x}^{a(-i)})$ is the projection of the analysis value change $(\mathbf{x}^a - \mathbf{x}^{a(-i)})$ on the i^{th} observation location. Substituting equations (3) and (4) into (7), it becomes:

$$\begin{aligned} y_i^f - y_i^{f(-i)} &\cong \mathbf{M}S_{ii}^o(1-S_{ii}^o)^{-1}(y_i^o - y_i^a) \\ &= \left(\frac{1}{n-1}\right)\frac{1}{\sigma_i^2}\sum_{j=1}^{n}\mathbf{M}(\mathbf{HX}^{aj})_i \times (\mathbf{HX}^{aj})_i(1-S_{ii}^o)^{-1}(y_i^o - y_i^a) \\ &= \left(\frac{1}{n-1}\right)\frac{1}{\sigma_i^2}\sum_{j=1}^{n}(\mathbf{HX}^{fj})_i \times (\mathbf{HX}^{aj})_i(1-S_{ii}^o)^{-1}(y_i^o - y_i^a) \end{aligned} \tag{8}$$

In deriving equation (8), we exchange the order of $\mathbf{M}$ and $\mathbf{H}$ in the second step. $\mathbf{HX}^{fj}$ is the j^{th} column of the forecast ensemble perturbation matrix $\mathbf{HX}^f$, which is equal to:

$$\mathbf{HX}^{fj} \cong h(\mathbf{x}^{fj}) - \frac{1}{n}\sum_{i=1}^{n}h(\mathbf{x}^{fj}) \tag{9}$$

and $\mathbf{x}^{fj}$ is the j^{th} ensemble forecast member. From equation (8), we obtain the forecast value at the i^{th} observation location when the i^{th} observation is left out at the analysis time:

$$y_i^{f(-i)} = y_i^f - \left(\frac{1}{n-1}\right)\frac{1}{\sigma_i^2}\sum_{j=1}^{n}(\mathbf{HX}^{fj})_i \times (\mathbf{HX}^{aj})_i(1-S_{ii}^o)^{-1}(y_i^o - y_i^a) \tag{10}$$

Equations (8) and (10) show that the change in the forecast value $y_i^f - y_i^{f(-i)}$ at the i^{th} observation location by leaving out the i^{th} observation at the analysis time and $y_i^{f(-i)}$ can be approximately computed from the ensemble forecast perturbations $\mathbf{HX}^f$, the ensemble analysis perturbations $\mathbf{HX}^a$, the self-sensitivity S_{ii}^o, the observation value y_i^o and the analysis value y_i^a. The approximation of these two equations comes from two aspects: one is due to the nonlinearity in both the forecast model and the observation operator; the other is the impact of the change in the analysis values (due to the deletion of the i^{th} observation) other than the analysis at the i^{th} observation location on the forecast at the i^{th} point.

3. The observation impact on the forecast accuracy

Equation (10) shows that $y_i^{f(-i)}$ can be calculated without carrying out a data denial experiment. With both $y_i^{f(-i)}$ and y_i^f known, and a verification state $y_{i,t}^v$ valid at time t available, we can calculate the quadratic forecast error change J_i at the i^{th} observation location due to deletion of that observation. J_i is defined as:

$$J_i = (y_i^f - y_{i,t}^v)^2 - (y_i^{f(-i)} - y_{i,t}^v)^2 \tag{11}$$

which reflects the impact of the i^{th} observation on the forecast accuracy at the i^{th} point. When the i^{th} observation improves the forecast at the point i, J_i will be negative, otherwise, it will be positive. With J_i, the impact of a group of observations can be calculated by summing J_i over these observations. J_i can also be used in observation data thinning by deleting the observations that make the forecast worse.

The observation impact defined in equation (11) is different from the adjoint method (e.g., Langland and Baker, 2004; Zhu and Gelaro, 2008; Cardinali, 2009) and the ensemble sensitivity method (Liu and Kalnay, 2008, Li et al., 2009). The cost function (equation (11)) in this study is defined as the error difference between $y_i^{f(-i)}$ and y_i^f, which shows the forecast error change at the i^{th} observation location if the i^{th} observation is not assimilated at t = 0, as shown in the left panel of Figure 1.

The cost functions defined in Langland and Baker (2004) and Liu and Kalnay (2008) are the error difference between the forecasts integrated from the analysis at t = 0 and t = 6, which examines the contribution of all the observations assimilated at t = 0 to the reduction of the forecast error at time t, as shown in the right panel of Figure 1.

Furthermore, the cost function examined in this study directly reflects the impact of the observation of interest, while the cost functions in the other two studies have to be rewritten as function of the observations assimilated at t = 0.

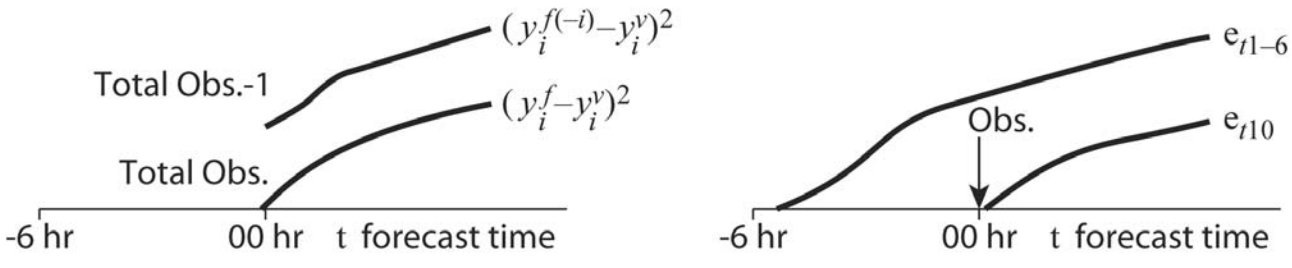

Figure 1: Left panel: schematic plot of the forecast error change at the i^{th} observation location by leaving out the i^{th} observation at the analysis time; right panel: schematic plot of the time relationship of the observation impact on the reduction of forecast errors at time t. (After Liu and Kalnay, 2008 Fig 1.)

4. Validation of $y_i^{f(-i)}$ calculation without data denial experiments and observation impact experiments with Lorenz 40-variable model

4.1. Lorenz 40-variable model and experimental setup

The Lorenz 40-variable model is governed by:

$$\frac{d}{dt}x_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F \tag{12}$$

The variables (x_j, j=1, L, J) represent "meteorological" variables on a "latitude circle" with periodic boundary conditions. As in Lorenz and Emanuel (1998), J is chosen to be 40. 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. Observations are simulated by adding Gaussian random perturbations (with a standard deviation equal to 0.2) to the nature run.

The data assimilation scheme we use is the Local Ensemble Transform Kalman Filter (LETKF), which is one type of EnKF especially efficient for parallel computing (see Hunt et al., 2007 for a detailed description of this method). Since F has different values in the nature run and in the forecast run during data assimilation, the multiplicative covariance inflation method (Anderson and Anderson, 1999) has been applied to account for model error in addition to sampling errors. The covariance inflation factor is fixed to be 1.3 in this study, which means that the background error covariance $\mathbf{P}^b$ is multiplied by 1.3 in each data assimilation cycle.

In all the experiments, we use 40 ensemble members. In verifying the calculation of $y_i^{f(-i)}$ based on equation (10), we calculate it in two ways. One is to leave out each observation in turn to get $\mathbf{x}^{a(-i)}$ using the same background forecasts after full observation data assimilation, and then calculate $y_i^{f(-i)}$ based on $y_i^{f(-i)} = \mathbf{H}(M(\mathbf{x}^{a(-i)}))_i$. The forecast length is 4 times of the assimilation forecast length. The other way is to calculate it from equation (10), using ensemble forecast perturbations $\mathbf{HX}^f$, self-sensitivity and the other quantities without calculating $\mathbf{x}^{a(-i)}$ from data denial experiments.

4.2. Results

Figure 2 shows $y_i^f - \mathbf{H}(M(\mathbf{x}^{a(-i)}))_i \left(i = 1, \cdots, 40\right)$ (solid line with closed circles), and $y_i^f - y_i^{f(-i)}$ (solid line with plus signs) calculated from equation (10) at one analysis time. These two quantities collocate with each other over most of the points, and with very small difference over the other points. This indicates that $y_i^{f(-i)}$ based on equation (10) is a good approximation of $\mathbf{H}(M(\mathbf{x}^{a(-i)}))_i$ obtained from data denial experiments.

Since $y_i^{f(-i)}$ can be calculated from equation (10) without carrying out data denial experiments, the impact of the i^{th} observation on the forecast accuracy at that point can be obtained by comparing the forecast error difference between y_i^f and $y_i^{f(-i)}$, as defined in equation (11). When the i^{th} observation improves the forecast at that point, the forecast error gets smaller with the assimilation of the i^{th} observation, and J_i will be negative. Otherwise, J_i will be positive. To show whether this method can detect the bad quality observations as the adjoint method (Langland and Baker, 2004) and the ensemble sensitivity method (Liu and Kalnay, 2008, Li et al., 2009), we design an experiment that the 11th observation has 4 times larger random error standard deviation than the other points. The observation impact is calculated from equation (11). Figure 3 shows the observation impact averaged over the last 500 analysis cycles of the 1000 total assimilation cycles. It reveals that the 11th observation makes the forecast worse, and the other points improve the forecast. This indicates that this method can detect the bad quality observations. Furthermore, $y_i^{f(-i)}$ from data denial experiments (solid line with closed circles) and $y_i^{f(-i)}$ based on equation (10) (solid line with plus signs) show similar observation impact, which further verify that the calculation of $y_i^{f(-i)}$ based on (10) is a good approximation of the actual value.

As any other observation impact study, different verification states may give different observation impact estimation. Figure 4 shows that the observation impact signal, both negative and positive, becomes stronger when the truth is used as verification state. The observation impact calculation method discussed here may be more sensitive to the verification state than the other methods (e.g., Langland and Baker, 2004; Liu and Kalnay, 2008; Li et al., 2009), since this observation impact formula (equation (11)) calculates the impact of the observation on the forecast only at that observation location, which may give a smaller signal than the impact on all grid points as calculated from other methods (e.g., Langland and Baker, 2004; Liu and Kalnay, 2008; Li et al., 2009). However, since the observation impact calculation based on equation (11) requires little computational time, and examines different aspects of the observation impact than other methods as shown in section 3, it is an efficient tool to complement the estimation of observation impact from other methods.

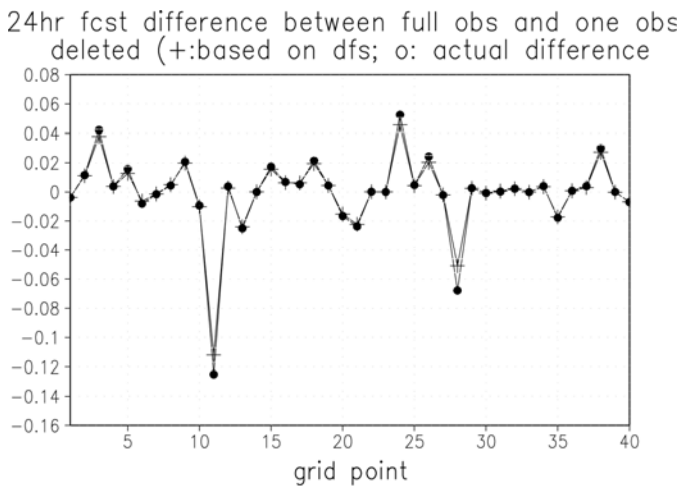

Figure 2: The difference between the forecast $y_i^f\ (i=1,\cdots,40)$ that initialized with the full observation analysis and $y_i^{f(-i)}\ (i=1,\cdots,40)$ that initialized with the analysis that the i^{th} observation is deleted at the analysis time. Solid line with plus signs: $y_i^{f(-i)}$ is calculated from equation.(10) without actually carrying out data denial experiments; solid line with closed circles: $y_i^{f(-i)}\ (i=1,\cdots,40)$ is the forecast initialized with the analysis that each observation deleted in turn at the analysis time.

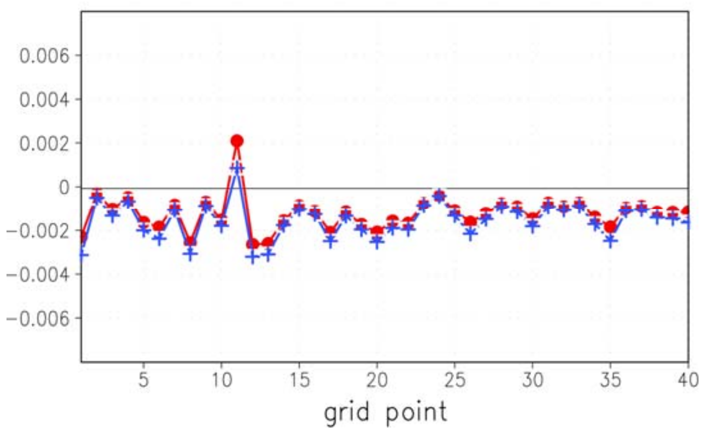

Figure 3: Time average of the impact of each observation on the forecast error reduction (equation (11)) at the observation location. The observation error at the 11th point is four times larger than the other points. Solid line with plus signs: $y_i^{f(-i)}$ is calculated from equation (10) without actually carrying out data denial experiments; solid line with closed circles: $y_i^{f(-i)}\ (i=1,\cdots,40)$ is the forecast initialized with the analysis that each observation deleted in turn at the analysis time. The verification state is the analysis at the verification time.

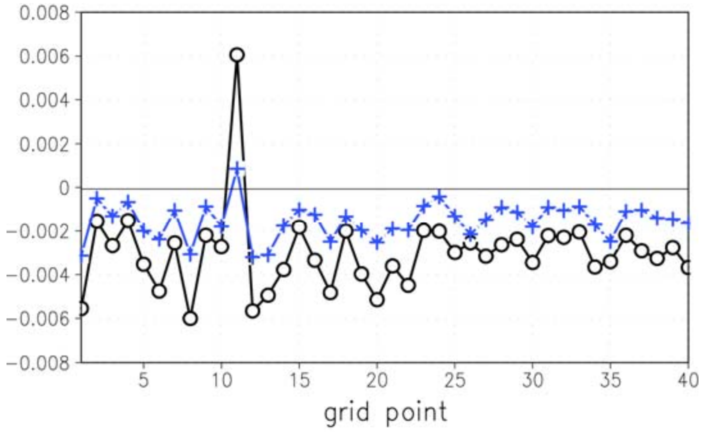

Figure 4 The impact of the verification state on the calculation of the observation impact on the forecast error reduction. Solid line with plus signs is the same as in Figure 3; Solid line with open circles is the same as the line with plus signs except that the verification state is from the true state.

5. The observation impact calculation in a primitive equation global model

The tests in the Lorenz 40-variable model show that short-term forecast value change can be calculated without carrying out data denial experiments, and this can be used in calculating the impact of the observation on the forecast accuracy at that location. In this section, with Observing System Simulation Experiments (OSSEs) setup, we examine the impact of observations on the forecast accuracy in a primitive equation global model. As in Liu et al. (2009), we follow a "perfect model" OSSE setup, (e.g., Lord et al. 1997), and use the Simplified Parameterizations primitivE Equation DYnamics (SPEEDY, Molteni, 2003) model, which is a global atmospheric model with 96x48 grid points in the horizontal and 7 vertical levels in a sigma coordinate. The observations are obtained at rawinsondes locations (closed circles in Figure 6 in Liu et al. 2009) with error standard deviation equal to 30% of natural variability (Figure 5 in Liu et al. 2009). At each assimilation cycle, we randomly pick out 30 observation locations, with every 6 of which have 4 times larger random error standard deviation of one type of observations (*u, v, T, q, ps*). We carry out one and a half month data assimilation cycles, and the results shown here are the average over the last one month.

Based on equations (10) and (11), we calculate the impact of every observation on the 12-hr forecast error change at that observation location in each assimilation cycle. Figure 5 shows the time average of the observation impact on the 12-hr forecast error change summed over all the vertical levels for zonal wind observations (top panel) and specific humidity observations (bottom panel). It indicates that, on a time average, almost all the observations have positive impact on the forecast accuracy, even though we randomly add larger random error to a small portion of observation in each assimilation cycle. The observations along the coast and in data sparse area have much larger impact on the forecast accuracy change than data dense region, which is consistent with the findings of Liu et al. (2009) that the observations at data sparse area are more important. The impact of specific humidity observations (bottom panel of Figure 5) on the 12hr forecast shows similar features as the zonal wind observations.

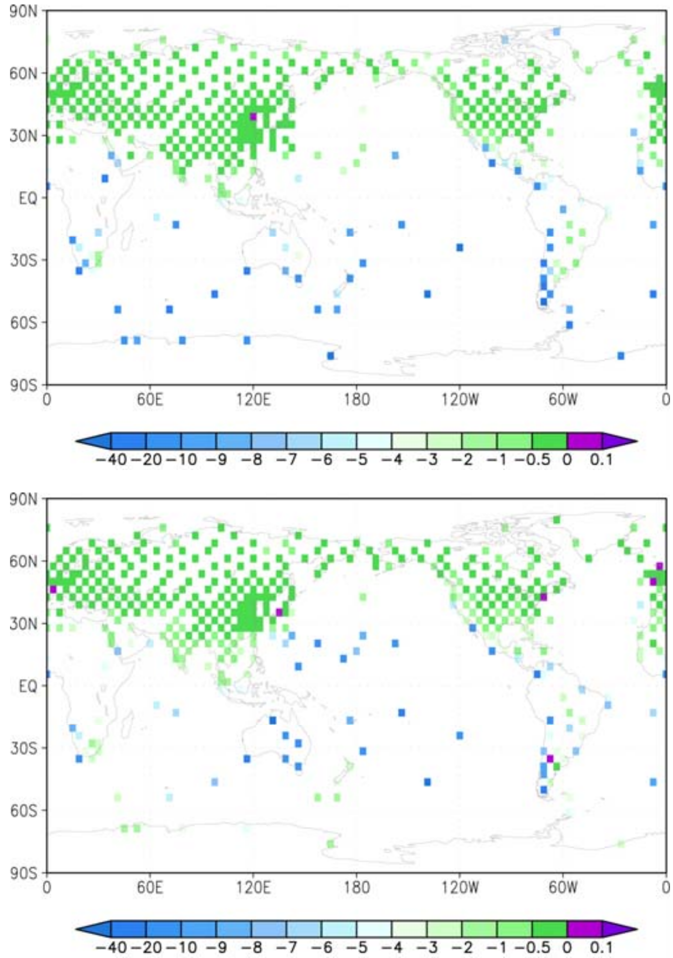

Figure 5 Time average of the observation impact on the 12hr forecast error change as defined in equation (11) *summed over all the vertical levels for zonal wind (top panel, unit: m^2/s^2) and specific humidity (bottom panel, unit: 1.0e-7kg^2/kg^2). The verification state is the analysis valid at the forecast time.*

6. Summary and discussion

Self-sensitivity indicates the sensitivity of analysis to observations, which is a function of analysis error covariance and observation error covariance. Since EnKFs explicitly calculate analysis error covariance in each assimilation cycle, self-sensitivity can be calculated without approximation when the observation errors have no correlation, as shown in Liu et al. (2009). With self-sensitivity, observation and analysis value, the analysis value change at the i^{th} observation location with the deletion of the i^{th} observation can be obtained without carrying out data denial experiments. In this paper, we extend this property to the forecast values, deriving an equation that calculates $y_i^{f(-i)}$ without carrying out data denial experiments. Based on this forecast value, a cost function that measures the impact of the observation on the forecast accuracy at that observation location is defined. The observation impact defined in this paper is different from Langland and Baker (2004), Liu and Kalnay (2008) and Li et al. (2009), because it measures the impact on the forecast when some observations are deleted during data assimilation, while Langland and Baker (2004), Liu and Kalnay (2008) and Li et al. (2009) examine the impact of the observations when all the observations are assimilated simultaneously. This method is also different from the observation impact study by Cardinali (2009), which utilizes analysis sensitivity in a chain rule derivation of the adjoint observation impact formula.

With the Lorenz 40-variable model, we have shown that the forecast value change at the i^{th} observation location based on equation (10) is a good approximation of the actual forecast value change obtained from data denial experiments. Furthermore, the observation impact based on the forecast value obtained from equation (10) can detect the bad quality observations, and reflect the actual impact of each observation on the forecast at those observation locations.

Perfect model experiments with a global primitive equation model show that the observation impact obtained from equation (11) have larger impact over data sparse area and smaller impact over data dense region, as shown in Liu et al. (2009).

References

Anderson JL. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.

Anderson JL and Anderson SL. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127,** 2741–2758.

Bishop CH, Etherton B and Majumdar SJ. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420-436.

Cardinali C, Pezzulli S and Andersson E. 2004. Influence-matrix diagnostic of a data assimilation system. *Quart. J. Roy. Meteor. Soc.* **130**, 2767-2786

Cardinali C. 2009. Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.* **135**, 239-250.

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

Houtekamer PL, and Mitchell HL. 2001. A sequential Ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123-137.

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

Langland RH and Baker NL. 2004. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56a**, 189-201.

Li H., Liu J., and Kalnay E. 2009. Correction on "Observation impact without adjoint model in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc (submitted)*

Liu J. and Kalnay E. 2008. Estimating observation impact study without adjoint model in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc*, **134**, 1327-1335.

Liu, J., E. Kalnay, T. Miyoshi, and C. Cardinali, 2009: Analysis sensitivity calculation within an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc. (In press)*

Lord, S. J., E. Kalnay, R. Daley, G. D. Emmit, and R. Atlas: 1997, Using OSSEs in the design of the future generation of integrated observing system. 1st Symposium on Integrated Observation Systems, AMS

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

Molteni F. 2003. Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multi-decadal experiments. *Climate Dyn.*, **20**, 175-191.

Ott, E, Hunt BR, Szunyogh I, Zimin AV, Kostelich EJ, Corazza M, Kalnay E, Patil DJ, and Yorke JA. 2004. A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus*, **56A**, 415-428

Whitaker JS, and Hamill TM. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1913-1924.

Zhu, Y. and Gelaro, R., 2008. Observation sensitivity calculations using the adjoint of the Gridpoint Statistical Interpolation (GSI) analysis system. Mon. Wea. Rev. **136**, 335-351