

A simpler formulation of forecast sensitivity to observations: application to ensemble Kalman filters

By EUGENIA KALNAY^{1*}, YOICHIRO OTA², TAKEMASA MIYOSHI¹ and JUNJIE LIU³, ¹University of Maryland, College Park, MD 20742, USA; ²National Centers for Environmental Prediction, College Park, MD 20740 (detailed from Japan Meteorological Agency); ³Jet Propulsion Laboratory, Pasadena, CA, 91109

(Manuscript received 1 April 2012; in final form 25 August 2012)

ABSTRACT

We introduce a new formulation of the ensemble forecast sensitivity developed by Liu and Kalnay with a small correction from Li et al. The new formulation, like the original one, is tested on the simple Lorenz 40-variable model. We find that, except for short-range forecasts, the use of localization in the analysis, necessary in ensemble Kalman filter (EnKF) when the number of ensemble members is much smaller than the model's degrees of freedom, has a negative impact on the accuracy of the sensitivity. This is because the impact of an observation during the analysis (i.e. the analysis increment associated with the observation) is transported by the flow during the integration, and this is ignored when the ensemble sensitivity uses a fixed localization. To address this problem, we introduce two approaches that could be adapted to evolve the localization during the estimation of forecast sensitivity to the observations. The first one estimates the non-linear evolution of the initial localization but is computationally expensive. The second one moves the localization with a constant estimation of the group velocity. Both methods succeed in improving the ensemble estimations for longer forecasts.

Overall, the adjoint and ensemble forecast impact estimations give similarly accurate results for short-range forecasts, except that the new formulation gives an estimation of the fraction of observations that improve the forecast closer to that obtained by data denial (Observing System Experiments). For longer-range forecasts, they both deteriorate for different reasons. The adjoint sensitivity becomes noisy due to the forecast non-linearities not captured in the linear tangent model and the adjoint. The ensemble sensitivity becomes less accurate due to the use of a fixed localization, a problem that could be ameliorated with an evolving adaptive localization. Advantages of the new formulation include it being simpler than the original formulation and computationally more efficient and that it can be applied to other EnKF methods in addition to the local ensemble transform Kalman filter.

Keywords: Ensemble Kalman Filter, forecast sensitivity, observation impact, skill dropout, ensemble sensitivity, data assimilation

1. Introduction

Langland and Baker (2004; LB04) wrote a fundamental paper, showing how to answer the question: “Did the use of any subset of the observations make the forecast better or worse?” without having to carry out analyses and forecasts with and without assimilating those observations, as required in conventional Observing System Experiments (OSEs). Gelaro and Zhu (2009; GZ09) derived a formulation

equivalent to LB04. These estimations of forecast sensitivity to observations are based on adjoint sensitivity and have proven to be a powerful monitoring tool adopted operationally at Naval Research Laboratory (NRL) and National Aeronautics and Space Administration/Global Modeling and Assimilation Office (NASA/GMAO).

Liu and Kalnay (2008; LK08) proposed an ensemble Kalman filter (EnKF) formulation algorithm equivalent to LB04 or GZ09, without requiring the adjoint of either the forecast model or the data assimilation scheme. Li et al. (2010; LLK10) pointed out a minor error in the original LK08 formulation and noted that the cost function measuring the impact of the observations at the initial

*Corresponding author.
email: ekalnay@atmos.umd.edu

time on the forecast could be computed directly rather than through the gradient of the cost function. LK08 and LLK10 tested the ensemble sensitivity formulation within the Local Ensemble Transform Kalman Filter (LETKF; Hunt et al., 2007) coupled with the Lorenz and Emanuel (1998) 40 variables model and found results comparable to those obtained using the adjoint sensitivity, and Kunii et al. (2012) successfully applied this methodology on real observations for tropical cyclone prediction.

Here we present a formulation of the ensemble forecast sensitivity to observations based on the direct computation of the cost function, without computing its gradient. Although it is essentially equivalent to the original LK08/LLK10 formulation, it is simpler and computationally more efficient since it uses available EnKF products. Unlike the original LK08 formulation, it can be applied to the Ensemble Square Root Filter (EnSRF; Whitaker and Hamill, 2002) in which the analysis weights of the background ensemble members cannot be explicitly computed.

In addition, we find that the use of localization, necessary in EnKF when the number of ensemble members is much smaller than the number of degrees of freedom (DOF) of the model, reduces the accuracy of the ensemble forecast sensitivity. We show that two approaches to evolve the observation localization during the forecast improve the results.

Section 2 presents the new formulation and compares it with the original one. Section 3 describes the experimental design. Results from the two ensemble formulations and from the adjoint sensitivity are compared in Section 4 using the Lorenz and Emanuel (1998) model. Section 5 is a summary of the results and relative advantages and a discussion of the impact of localization on the EnKF forecast sensitivity.

2. Adjoint and ensemble formulations

Let $\bar{\mathbf{x}}_{t|0}^f$ represents a forecast started from the analysis $\bar{\mathbf{x}}_0^a$ at time 0 and verifying at time t . The overbars are used to indicate ensemble mean and are relevant only for the ensemble sensitivity formulation so that they can be ignored in the adjoint sensitivity formulation. The perceived forecast error at the verification time t from a forecast started at time 0, and verified against the analysis $\bar{\mathbf{x}}_t^a$ valid at time t , is given by $\mathbf{e}_{t|0} = \bar{\mathbf{x}}_{t|0}^f - \bar{\mathbf{x}}_t^a$. The corresponding error from the forecast started at time $t = -6$ h is given by $\mathbf{e}_{t|-6} = \bar{\mathbf{x}}_{t|-6}^f - \bar{\mathbf{x}}_t^a$ (see schematic Fig. 1). Six hours is a typical data assimilation window for numerical weather prediction (NWP), but it could be different depending on the system of our interest, for example, 1 h for mesoscale NWP, or 1 week for global ocean data assimilation. As indicated in Fig. 1, the difference between the forecast errors $\mathbf{e}_{t|0}$ and $\mathbf{e}_{t|-6}$ at

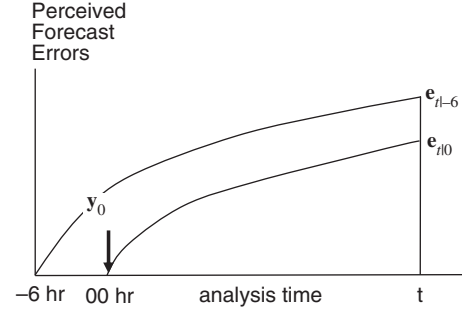


Fig. 1. Schematic of the perceived forecast error $\mathbf{e}_{t|0}$ verified against the analysis at the verification time t from two forecasts started from the analysis at $t = 0$ h, and from the analysis at $t = -6$ h $\mathbf{e}_{t|-6}$. Since the forecast started at $t = -6$ h serves as first guess for the analysis at $t = 0$ h, the only difference between the two forecasts is the assimilation of the observations \mathbf{y}_0 at $t = 0$ h. Adapted from Langland and Baker (2004).

verification time t is only due to the observations \mathbf{y}_0 assimilated at time 0 that change the background $\bar{\mathbf{x}}_{0|-6}^b$ by the analysis increment

$$\delta\bar{\mathbf{x}}_0 = (\bar{\mathbf{x}}_0^a - \bar{\mathbf{x}}_{0|-6}^b) = \mathbf{K}\delta\mathbf{y}_0. \quad (1)$$

Here \mathbf{K} is the gain matrix that defines the data assimilation algorithm, $\delta\mathbf{y}_0 = (y - H(\bar{\mathbf{x}}_{0|-6}^b))$ is the observational increment with respect to the first guess, and H is the non-linear observation operator. We indicated with a superscript b the 6-h forecast started from the analysis at time -6 h and used as background at time 0.

LB04 introduced a cost function to measure the impact of the observations at time 0 on the forecast at time t as the difference between the squares of the forecast errors with and without assimilating the observations \mathbf{y}_0 :

$$J = \Delta e^2 = (e_{t|0}^2 - e_{t|-6}^2) = [\mathbf{e}_{t|0}^T \mathbf{C} \mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^T \mathbf{C} \mathbf{e}_{t|-6}], \quad (2)$$

where the matrix of weights \mathbf{C} defines the squared norm to be used (dry total energy in the case of LB04 and GZ09).

Here we follow the suggestion of LLK10 and compute the sensitivity to observations directly from the cost function eq. (2) rather than from a Taylor series approximation and assume that the forecast length is short enough to allow the use of the linear tangent model $(\bar{\mathbf{x}}_{t|0}^f - \bar{\mathbf{x}}_{t|-6}^f) \approx \mathbf{M}(\bar{\mathbf{x}}_0^a - \bar{\mathbf{x}}_{0|-6}^b)$:

$$\begin{aligned} \Delta e^2 &= \mathbf{e}_{t|0}^T \mathbf{C} \mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^T \mathbf{C} \mathbf{e}_{t|-6} = (\mathbf{e}_{t|0} - \mathbf{e}_{t|-6})^T \mathbf{C} (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}) \\ &= (\bar{\mathbf{x}}_{t|0}^f - \bar{\mathbf{x}}_{t|-6}^f)^T \mathbf{C} (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}) \approx [\mathbf{M}(\bar{\mathbf{x}}_0^a - \bar{\mathbf{x}}_{0|-6}^b)]^T \\ &\quad \mathbf{C} (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}) = [\mathbf{M}\mathbf{K}\delta\mathbf{y}_0]^T \mathbf{C} (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}). \end{aligned} \quad (3)$$

LB04 and GZ09 solve eq. (3) by using the adjoint approach (cf. eq. (7) in LB04 and eq. (7) in GZ09):

$$\Delta e^2 \approx \delta \mathbf{y}_0^T \mathbf{K}^T \mathbf{M}^T C (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}), \quad (4)$$

thus requiring the adjoint of both the model (\mathbf{M}^T) and of the data assimilation (\mathbf{K}^T).

In their corrected ensemble formulation, LLK10 used the LETKF formulation (Hunt et al., 2007) to rewrite eq. (3) as

$$\begin{aligned} \Delta e^2 &\approx \delta \mathbf{y}_0^T \tilde{\mathbf{K}}^T \mathbf{X}_{t|-6}^T C (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}) \\ &\approx \delta \mathbf{y}_0^T \tilde{\mathbf{K}}^T \mathbf{X}_{t|-6}^T C (2\mathbf{e}_{t|-6} + \mathbf{X}_{t|-6}^f \tilde{\mathbf{K}} \delta \mathbf{y}_0) \end{aligned} \quad (5)$$

[cf. eq. (9b) in LLK10]. Here $\tilde{\mathbf{K}} = [(K-1)\mathbf{I} + (\mathbf{H}\mathbf{X}^b)^T \mathbf{R}^{-1} \mathbf{H}\mathbf{X}^b]^{-1} (\mathbf{H}\mathbf{X}^b)^T \mathbf{R}^{-1}$ is an intermediate matrix used in the LETKF, K is the number of ensemble members, and $\mathbf{X}_{t|-6}^f$ is a matrix whose columns are the forecast ensemble perturbations with respect to the mean, started from the analysis at time -6 h and valid at time t . Equation (5) requires estimating the Kalman gain matrix at each grid point and is therefore computationally burdensome.

In the new formulation of the forecast sensitivity in a deterministic EnKF framework, we start from eq. (3) and use the Kalman gain matrix $\mathbf{K} = \frac{1}{K-1} \mathbf{X}_0^a \mathbf{X}_0^{aT} \mathbf{H}^T \mathbf{R}^{-1}$ and the matrix $\mathbf{Y}_0^a = \mathbf{H}\mathbf{X}_0^a$ composed of analysis ensemble perturbations in observation space. Here $\mathbf{P}_0^a = \frac{1}{K-1} \mathbf{X}_0^a \mathbf{X}_0^{aT}$ is the analysis ensemble error covariance. We can then write $\mathbf{M}\mathbf{K} = \frac{1}{K-1} \mathbf{M}\mathbf{X}_0^a \mathbf{X}_0^{aT} \mathbf{H}^T \mathbf{R}^{-1} \approx \frac{1}{K-1} \mathbf{X}_{t|0}^f \mathbf{Y}_0^{aT} \mathbf{R}^{-1}$ so that the cost function eq. (3) becomes

$$\Delta e^2 \approx \frac{1}{K-1} \delta \mathbf{y}_0^T \mathbf{R}^{-1} \mathbf{Y}_0^a \mathbf{X}_{t|0}^{fT} C (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}). \quad (6)$$

Each column k of the perturbation forecast matrix $\mathbf{X}_{t|0}^{fT}$ can be computed with the full non-linear model $[M(x_0^{a(k)}) - M(x_0^{a(j)})]$. The new formulation eq. (6) uses analysis ensemble products rather than the Kalman gain and is more efficient than the original LK08/LLK10 formulation eq. (5). Unlike eq. (5), the new formulation eq. (6) does not introduce an approximation of $(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$, and it uses $\mathbf{M}\mathbf{X}_0^a \approx \mathbf{X}_{t|0}^f$ rather than $\mathbf{M}\mathbf{X}_{0|-6}^b \approx \mathbf{X}_{t|-6}^f$ used in eq. (5). As a result, although the two formulations are essentially equivalent, eq. (6) should be slightly more accurate than eq. (5). In addition, this formulation can be applied to EnKF methods other than the LETKF.

3. Experimental design

We use the Lorenz and Emanuel (1998) 40 variables model for the experiments. This model is governed by the equation:

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F, \quad (7)$$

where j is an index of the variables and $1 \leq j \leq 40$ with periodic boundary conditions. As in LK08, we allow for model errors by using $F=8.0$ for the truth run and $F=7.6$ for the data assimilation cycle and the forecast. The time integration is performed with the fourth-order Runge-Kutta scheme with time step 0.01 in non-dimensional time. Data assimilation with the LETKF is performed every 0.05 time units, estimated by Lorenz and Emanuel (1998), to be equivalent to 6 h when $F=8.0$. Observations are created at every grid point on each analysis time by adding independent Gaussian random errors with a standard deviation (SD) of 0.2 to the truth run. Like the experiment in LK08, observations at the 11th grid point have larger random errors (SD of 0.8), but we still use 0.2 as the observation errors for these observations in the assimilation process. We run 10 sets of 14,600 data assimilation cycles with different truth runs. In each set, we verify over the last 14,360 cycles.

We performed observation impact estimations using the three methods LB04/GZ09 (eq. 4 using adjoint formulation, denoted 4-ADJ), LK08/LLK10 (eq. 5, denoted 5-OLD) and the new ensemble formulation (eq. 6, denoted 6-NEW). For the adjoint formulation, the gradient of the cost function is constructed at forecast time and brought back to the initial time by the adjoint model as in LB04/GZ09. In the Runge-Kutta time scheme, the non-linear trajectories are used at every time step and substep of the time integration for the adjoint operation. The adjoint code is validated by the test described in Appendix. The same background covariance (full matrix with the same covariance localization as in the EnKF) is constructed from the EnKF perturbations and used in the computation of \mathbf{K}^T .

Since EnKF requires space localization when the number of ensemble members is much smaller than the number of DOF of the model, comparisons are made both with and without localization. For the experiments with localization, we applied a Gaussian localization function with an e-folding scale of $3\sqrt{2}$ grid points. The same localization is also applied to the observation impact estimation (Kunii et al., 2012). For the LK08/LLK10 method (5-OLD),

observation impacts are computed on each grid point j and observation l as

$$(\Delta e^2)_{j,l} = (\delta y_0)_l \left[\tilde{\mathbf{K}} \left(\mathbf{X}_{t-l-6}^{\text{T}} \right)_j C_{jj} \left(2 \left(\mathbf{e}_{t-l-6} \right)_j + \left(\mathbf{X}_{t-l-6}^{\text{T}} \right)_j \tilde{\mathbf{K}} \delta y_0 \right)_l \right] \quad (8)$$

The same localization as in the LETKF analysis is applied in the computation of $\tilde{\mathbf{K}}$ so that it does not explicitly appear in eq. (8). For eq. (6) (the NEW formulation), this becomes

$$(\Delta e^2)_{j,l} = \frac{1}{K-1} (\delta y_0)_l \left[\rho_j \mathbf{R}^{-1} \mathbf{Y}_0^a \left(\mathbf{X}_{t-l-6}^{\text{T}} \right)_j C_{jj} \left(\mathbf{e}_{t-l-6} + \mathbf{e}_{t-l-6} \right)_j \right] \quad (9)$$

where ρ_j is the localization function on grid point j . Both eq. (8) and eq. (9) can be simply summed up to get the impact of each observation. The localization function is fixed during the analysis and forecast times except as noted in Section 4.3. Multiplicative inflation (Anderson, 2001) is applied with a tuned parameter of 1.15². The observation impact estimations are computed at the forecast times equivalent to 6 h and 0.5, 1, 2, 3, 5 and 7 d.

4. Results

In Section 4.1, experiments are made with an ensemble of 40 members (same as the number of DOF of the model) so that no spatial localization is needed. In Section 4.2, we run the ensemble with only 10 members, more representative of operational applications for which the number of ensemble members is much smaller than the model DOF, so that localization is required. Since the results deteriorate with the use of localization, in Section 4.3 we introduce two methods that ameliorate this problem in longer forecasts, and compare the percentage observations that improve the forecasts in each of the methods with that obtained with a much more expensive OSE. In Section 4.4, we compare how well the different methods can detect “bad” observations with much higher error variance than specified in the assimilation system.

4.1. Analysis with 40 ensemble members without localization

Table 1 shows the average forecast error reduction amount of each experiment. For completeness, our results include estimates of forecast error reduction compared with values obtained from (1) the truth run and (2) the analysis cycles, as the truth is not known in real forecast applications. Because the number of ensemble members is as large as the number of DOF of the model, spatial localization is not required and, therefore, not applied here. The verification results against the analysis have very similar means and SDs as the verifications against the truth. The adjoint formulation (eq. 4) and the new ensemble formulation (eq. 6) both estimate well the average impacts throughout the forecast range, with eq. (6) being the most accurate at longer forecast ranges but with much higher SD. The old ensemble formulation (eq. 5) estimates similar, sometimes slightly better, observation impacts up to 3-d forecast but tends to underestimate the average impact for longer forecast times.

Figure 2 shows a skill score (SS) of the corresponding time–mean root mean square error (RMSE) of the different estimations of the total forecast error reduction verified against the true error reduction:

$$\text{SS}_{\text{estimated}} = \frac{\left(\sum_{\text{time}} (\Delta e_{\text{true}}^2)^2 \right)^{1/2} - \left[\sum_{\text{time}} (\Delta e_{\text{true}}^2 - \Delta e_{\text{estimated}}^2)^2 \right]^{1/2}}{\left(\sum_{\text{time}} (\Delta e_{\text{true}}^2)^2 \right)^{1/2}},$$

so that if the estimated reduction is always zero, $\text{SS} = 0$, and if the estimated reduction always agrees with the true reduction (eq. 2), $\text{SS} = 1$. Note that in Figs. 2–4, the skill peaks at 2–3 d because the forecast error is estimated against the analysis (affected by observation errors) and verified against the truth. It is only after a day or two that the forecast errors become large compared to the analysis errors. The performance of the adjoint and ensemble estimations is essentially identical for the first 2 d of the forecast, but the estimation by equation (4-ADJ) is somewhat better at 3- and 5-d forecasts, with the new ensemble

Table 1. Forecast error reduction of each formulation averaged over time and over 10 experiments. Truth and analysis show the average forecast error reduction verified against the truth, and against the analysis, respectively. Values in the parenthesis show the standard deviation of the 10 experiments. All values are multiplied by 100. No localization was used in the LETKF

	6 hr	12 hr	1 d	2 d	3 d	5 d	7 d
Truth	−24 (0)	−33 (0)	−57 (0)	−122 (3)	−218 (8)	−606 (29)	−1171 (52)
Analysis	−27 (0)	−32 (0)	−53 (0)	−116 (3)	−211 (8)	−593 (29)	−1152 (51)
(4-ADJ)	−23 (0)	−28 (0)	−46 (0)	−101 (2)	−181 (6)	−467 (22)	−840 (82)
(5-OLD)	−27 (0)	−33 (0)	−54 (0)	−121 (2)	−218 (7)	−421 (21)	−260 (27)
(6-NEW)	−31 (0)	−37 (0)	−60 (1)	−128 (3)	−220 (9)	−564 (41)	−1167 (86)

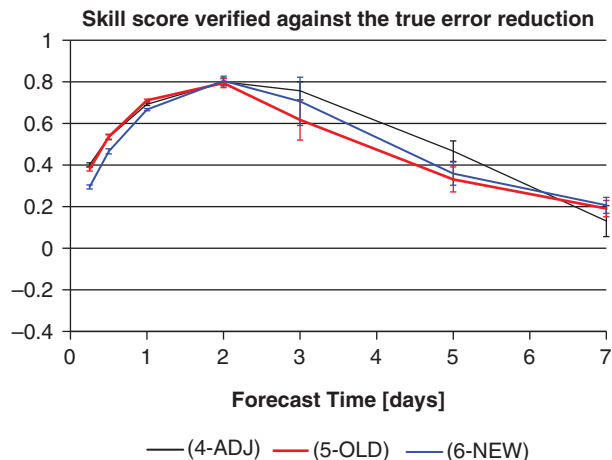


Fig. 2. Skill score of the estimation of the total forecast error reduction verified against the true forecast error reduction. Black, red and blue lines show the results of equations (4-ADJ), (5-OLD) and (6-NEW). The vertical error bars show the standard deviation of the 10 experiments. No localization was used in the LETKF.

formulation (6-NEW) better than the old one (5-OLD). The adjoint formulation (4-ADJ) becomes increasingly inaccurate and noisy beyond 6 d because, as shown in the Appendix, the linearization assumption made in the adjoint model does not hold anymore.

4.2. Analysis with 10 ensemble members and fixed localization

Figure 3 and Table 2 are the same as Fig. 2 and Table 1 but with only 10 ensemble members and now introducing localization in the analysis. As a benchmark, no localization was applied to the ensemble-based observation impact computations, shown in Fig. 3 as (5-OLD) Noloc and

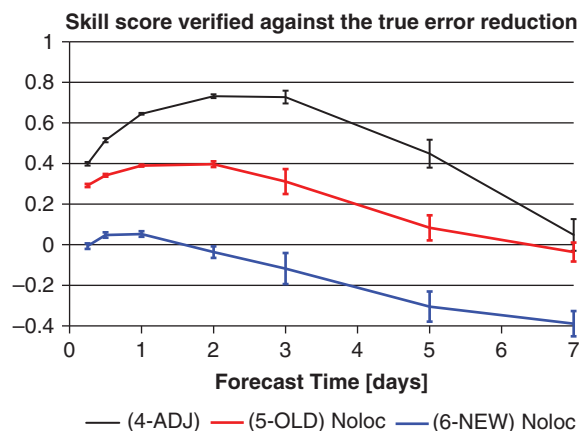


Fig. 3. Same as Fig. 2 but with only 10 members and using localization in the analysis. No localization was applied for the ensemble-based observation impact computations.

(6-NEW) Noloc. Compared with the results of 40 ensemble members without localization, the accuracies of both ensemble sensitivity methods in terms of RMSE are strongly degraded relative to the adjoint sensitivity, especially for (6-NEW) Noloc. This suggests that with much fewer ensemble members compared to the number of DOF of the model, spatial localization is also required in the computation of the observation impacts. Fig. 4 shows a graph similar to Fig. 3, but now with fixed spatial localization (as in eqs. 8 and 9), denoted (5-OLD) Fxloc and (6-NEW) Fxloc, respectively. The RMSEs of both ensemble-based estimations are improved with the fixed localization.

Table 2 includes the average forecast error reduction of these experiments, as well as the new moving localization methods introduced in Section 4.3. Although the RMSE of the estimates is reduced by applying the localization, the average estimate of (6-NEW) with fixed localization (Fxloc) tends to underestimate the impacts especially for the longer forecast times. Results using (5-OLD) have the wrong sign, so clearly the new formulation eq. (6) outperforms eq. (5) in the estimation of the mean when the ensemble size is smaller than the model DOF.

4.3. Analysis with 10 ensemble members and forecast sensitivity with moving localization

The analysis increment for any observation (i.e. the impact of the observation) evolves through the forecast. The localization function should include this evolution, otherwise the ensemble sensitivity will miss most of the observation impacts on longer forecast times. Unfortunately, the localization used in eqs. (8) and (9) is fixed so that it should become less accurate for longer forecasts.

The problem of how to evolve optimally the localization function in the computation of the forecast sensitivity is not trivial, and solving it is beyond the scope of this paper. Here we show two approximations that could be used in real systems.

The first approximation is to move the localization function (a Gaussian localization function centred on the location of observation l) with the non-linear incremental evolution (denoted NL-loc). The localization function of the observation l is added to the ensemble mean analysis field, and the evolution of the increment (G_l) is forecasted as

$$G_l(\bar{\mathbf{x}}_{t|0}^f) = \frac{M(\bar{\mathbf{x}}^a + \alpha \mathbf{f}_l) - \bar{\mathbf{x}}_{t|0}^f}{\alpha}, \quad (10)$$

where α is small positive real number set here at 0.01. We take the absolute value of eq. (10) and smooth it by using the fourth-order Shapiro (1970) filter in order to

Table 2. Same as Table 1 but with 10 ensemble members and localization when indicated

	6 h	12 h	1 d	2 d	3 d	5 d	7 d
Truth	−23 (0)	−32 (0)	−56 (0)	−121 (2)	−213 (6)	−582 (27)	−1144 (47)
Analysis	−25 (0)	−31 (0)	−51 (0)	−115 (2)	−206 (6)	−570 (27)	−1126 (46)
(4-ADJ)	−21 (0)	−25 (0)	−39 (0)	−77 (1)	−132 (4)	−344 (17)	−627 (52)
(5-OLD) Noloc	−12 (0)	−14 (0)	−24 (1)	−50 (3)	−69 (4)	136 (22)	1163 (65)
(6-NEW) Noloc	−32 (0)	−39 (0)	−62 (1)	−133 (4)	−227 (6)	−574 (41)	−1220 (80)
(5-OLD) Fxloc	−25 (0)	−30 (0)	−47 (0)	−89 (2)	−117 (4)	−9 (9)	486 (29)
(6-NEW) Fxloc	−28 (0)	−33 (0)	−50 (0)	−86 (2)	−106 (4)	−125 (11)	−162 (29)
(6-NEW) NL-loc	−24 (0)	−24 (0)	−47 (0)	−107 (2)	−178 (4)	−422 (27)	−788 (57)
(6-NEW) CL-loc	−28 (0)	−33 (0)	−54 (0)	−99 (2)	−151 (5)	−284 (18)	−462 (30)

remove 2-gridlength waves. Then we normalize it with the area integral of the localization function and limit the maximum and minimum values not exceeding 1 and 0. This value is used as the localization function of observation l for the impact estimations. Fig. 5a shows an example of the evolution of the localization function with the NL-loc method. Figs. 4 and 6 and Table 2 include results of the experiment with this moving localization. Both the RMSE and the average estimation are improved by applying the moving localization.

In a realistic geophysical data assimilation problem, using eq. (10) would be prohibitive because it requires a forecast for each observation. For this reason, we also tested a much simpler method to evolve the localization function. In this method, the localization centre is translated with a constant speed, $+0.6$ grid points per day, equal to the climatological group velocity of the dominant

wavenumbers 6–8. The group velocity is estimated from the modulation envelope of the space–time correlation using the method of Yoon et al. (2010). Figure 5b shows the evolution of the localization for the same case as Figure 5a but with the simpler constant linear method (CL-loc). Figure 4 and Table 2 also include the results of CL-loc, which, although simpler, is also clearly better than the fixed localization (Fxloc).

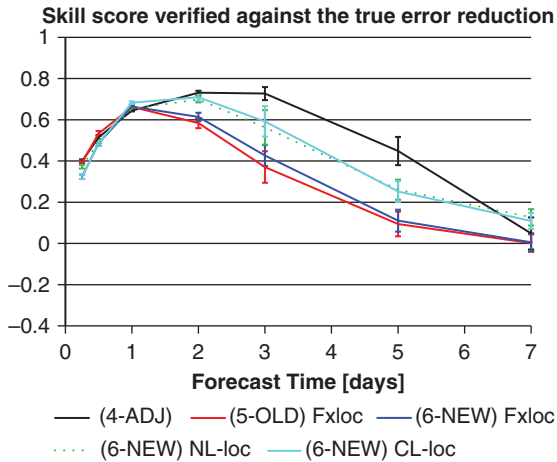


Fig. 4. As per Fig. 3 but now with localization used in the observation impact computations. Red and blue lines show the results of equations (5-OLD) and (6-NEW) with a fixed localization (Fxloc) function as used in the LETKF analysis. Dotted green and light blue lines show the results of equation (6-NEW) with localization functions moving with non-linear evolution (NL-loc) and constant speed equal to the group velocity (CL-loc).

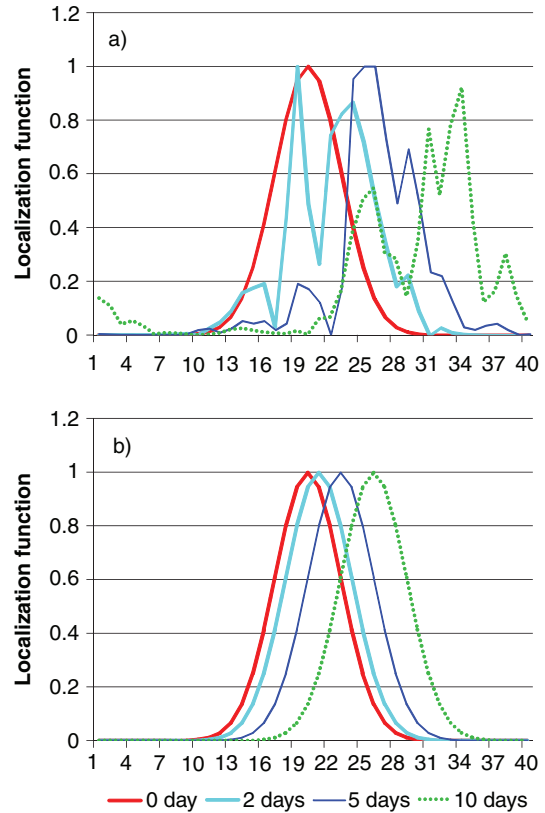


Fig. 5. Example of the evolution of the localization function with two methods. (a) Localization function with the non-linear incremental evolution (NL-loc); and (b) localization function moving with constant group velocity (CL-loc).

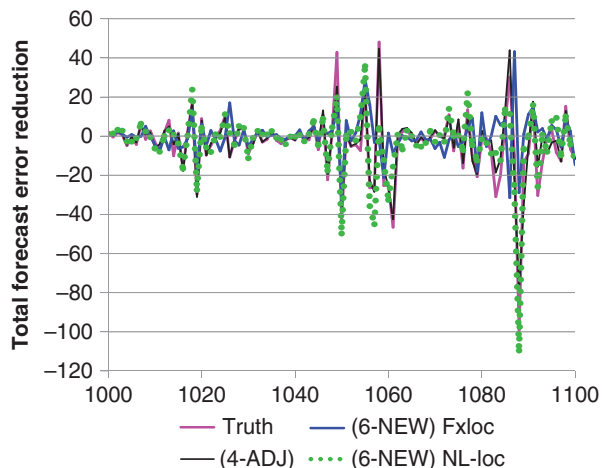


Fig. 6. Time series of the total forecast error reduction at 5-d forecast from cycle 1000 to 1100. Pink lines show the true forecast error reduction, the black line is the adjoint estimation, the blue line is the ensemble estimation with fixed localization and the dotted green line is the ensemble estimation with non-linear evolution of the localization. A total of 40 observations are assimilated on each analysis time.

Figure 6 compares a time series of the total forecast error reductions of (4-ADJ), (6-NEW) Fxloc with fixed localization and (6-NEW) NL-loc with moving localization of non-linear incremental evolution at 5 d with the true reduction of error. Ensemble-based estimates generally capture the large peaks (degradations and improvements of the forecast). However, (6-NEW) Fxloc sometimes misses a peak, but the ensemble-based estimations with moving

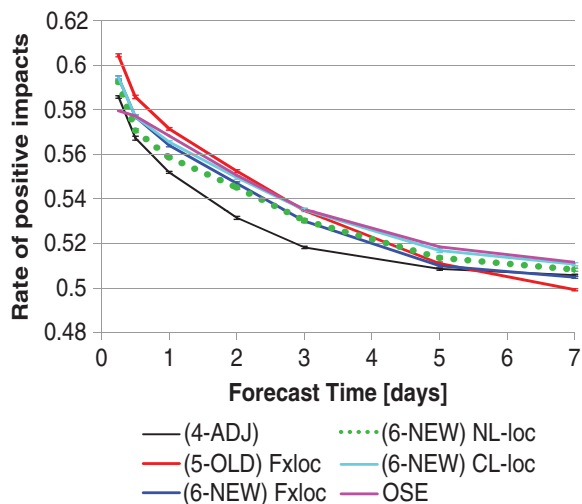


Fig. 7. Fraction of the observations that are estimated to decrease the forecast error in each method. Purple solid line shows the result of data denial experiments, denoted as an Observing System Experiment (OSE).

localization (6-NEW) NL-loc capture the time series very well.

Figure 7 shows the fraction of the observations that are estimated to decrease the forecast error in each method. As a reference, we also performed the much more computationally expensive data denial experiments of each observation at each analysis time and derived the actual rate of positive impacts (plotted as OSE). For short-range forecasts, the adjoint tends to underestimate the percentage of observations providing a positive impact, whereas the EnKF-based estimations have larger rate and are in better agreement with the data denial experiments. The fraction of positive impacts of (5-OLD) Fxloc is the largest up to 2-d forecast but becomes less than a half at 7-d forecast. Other methods converge to just above 0.5 (the value expected when forecasts have no more skill, and observations errors are Gaussian with zero bias) for long-range forecasts with (6-NEW) NL-loc and (6-NEW) CL-loc being closer to the OSE result.

4.4. Detection of bad observations

Finally, we compare the ability of the adjoint- and ensemble-based impact estimates to detect flawed observations useful for realistic applications. In this subsection, we show the results of 10-member ensemble experiments.

Figure 8 shows the average impact estimates of each method at each observation point on 5-d forecasts. Estimates are derived from the last 1000 (Fig. 8a) and the last 100 assimilation cycles (Fig. 8b). For a large sample of 1000, both the adjoint and the two ensemble-based methods can clearly detect the erroneous observation at the 11th grid point (Section 3) with very similar results. With only 100 samples, both methods still detect a peak at the 11th grid point, but at other points the estimates are much noisier, especially for the adjoint approach. In practical applications, the number of samples (assimilation cycles) can be of the order of 100–1000. This result suggests that all three methods are able to detect bad observations even for relatively small samples, but the adjoint approach may be noisier.

5. Summary and conclusions

In this article, we introduce a new formulation of the ensemble forecast sensitivity developed by Liu and Kalnay (2008) with the small correction of Li et al. (2010). The two formulations are compared with the adjoint forecast sensitivity using the same analysis with the Lorenz and Emanuel (1998) 40-variables model. For the first 2 d, the three formulations are essentially identical. From 3 to 5 d, the adjoint formulation is the most accurate, but it becomes the worst at longer time scales, when the linear approximation made in the adjoint formulation breaks down.

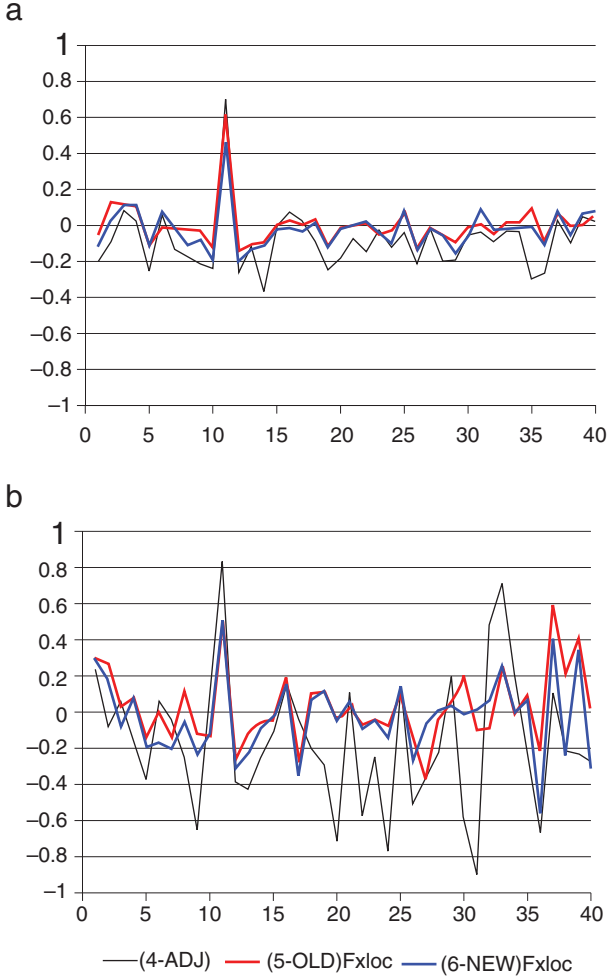


Fig. 8. Time average of observation impact estimates at each point on 5-d forecast with (a) 1000 and (b) 100 samples, using the adjoint approach eq. (4), the LLK10 formulation eq. (5) and the new ensemble formulation eq. (6), both with fixed localization. The observation at grid point 11 has an error variance of 0.8 rather than the assumed value of 0.2.

We find that the use of localization in the analysis, necessary in EnKF when the number of ensemble members is much smaller than the model’s DOF, has a negative impact on the accuracy of the ensemble sensitivity. This is not surprising since the impact of an observation during the analysis (i.e. the analysis increment associated with the observation) is transported by the flow during the integration, and this is ignored with a fixed localization. To address this problem, we also introduce two approaches that could be adapted to evolve the localization function during the estimation of forecast sensitivity to the observations. The first method estimates the non-linear evolution of the initial localization and is very expensive. The second one moves the localization with a constant estimation of

the group velocity. Both methods succeed in significantly improving the estimations for longer forecasts. When compared with the fraction of observations that reduce the forecast errors obtained running with and without each observation (as in an OSE) for short-range forecasts, the adjoint approach tends to underestimate the percentage of positive impact of the observations, and the new ensemble formulation gives impacts closer to the OSEs.

In summary, the adjoint and ensemble forecast impact estimations give similarly accurate results for short-range forecasts, except that the new formulation estimation of the fraction of observations that reduce forecast errors is closer to that obtained with OSEs. For longer-range forecasts, they both deteriorate for different reasons. The adjoint sensitivity becomes noisy due to the fact that the adjoint model is based on the tangent linear model and, therefore, cannot capture forecast non-linearities that become large with the forecast length. The ensemble sensitivity becomes less accurate due to the use of fixed localization, a problem that could be ameliorated with an evolving adaptive localization method (Bishop and Hodyss, 2009a; 2009b). Advantages of the new ensemble formulation are that it is simpler, more computationally efficient, and that it can be applied to other EnKF methods, and not just the LETKF. It has been implemented and tested with real forecasts at National Centers for Environmental Prediction (NCEP) using the EnSRF (Whitaker and Hamill, 2002).

6. Appendix

Validation tests of the tangent linear and adjoint models, and impact of non-linear forecast error growth.

A.1 Test of the tangent linear model

To test the adjoint code, we first tested a tangent linear code of the Lorenz 96 model. If we have a basic state \mathbf{x}_0 and a perturbation \mathbf{x}' , the evolution of perturbation can be derived using the non-linear model M and tangent linear model \mathbf{M} .

$$\mathbf{x}'_{nt} = [M(\mathbf{x}_0 + \alpha\mathbf{x}') - M(\mathbf{x}_0)]/\alpha, \text{ and } \mathbf{x}'_{it} = \mathbf{M}\mathbf{x}',$$

where $|\alpha| \ll 1$. A similarity index (SI), equivalent to the pattern correlation, is computed between the two perturbations:

$$\begin{aligned} SI &= \frac{(\mathbf{x}'_{nt}, \mathbf{x}'_{it})}{\|\mathbf{x}'_{nt}\| \|\mathbf{x}'_{it}\|} = \frac{[M(\mathbf{x}_0 + \alpha\mathbf{x}') - M(\mathbf{x}_0)]^T [\mathbf{M}\mathbf{x}']}{\|M(\mathbf{x}_0 + \alpha\mathbf{x}') - M(\mathbf{x}_0)\| \|\mathbf{M}\mathbf{x}'\|} \\ &= 1 - O(\alpha^2). \end{aligned} \quad (\text{A1})$$

If tangent linear model \mathbf{M} is correct and α is small enough, we should observe such a dependence on α , since SI is the cosine of the angle between the two vectors. \mathbf{x}_0 is

Table A1. Similarity indices between non-linear and tangent linear evolution of the perturbations (1 day evolution) with various perturbation amplitudes

Amplitude	SI
10^{-5}	0.9999999999953
10^{-4}	0.9999999994920
10^{-3}	0.9999999423196
10^{-2}	0.9999937996438
10^{-1}	0.9994241704626

picked from a truth run and \mathbf{x}' is generated from random numbers with normal distribution. Table A1 shows the SI with different orders of magnitude values of α after evolving for 1 d. The relation eq. (A1) is clearly verified; thus, the tangent linear code is validated.

A.2 Impact of non-linear error growth

To examine the accuracy of the tangent linear approximation in the forecast experiments, SI was also computed with larger perturbations, namely, α values of 0.2, 0.5, 1.0 and 2.0, corresponding to forecast RMSEs (verified against the truth) approximately observed after 1, 3, 5 and 8 d. In each setting, the results are derived by averaging SI from 1000 realizations. Figure A1 shows the result. At least for the short-range forecast (up to 3-d forecast), the tangent linear model (and thus the adjoint) is accurate and validated for use in this study. Due to the developing non-linear effect, the tangent linear approximation as well as the adjoint becomes less accurate for longer forecasts.

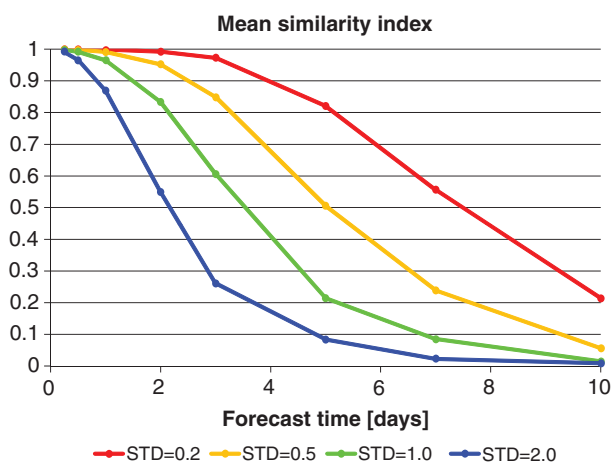


Fig. A1. Similarity indices between the non-linear and tangent linear evolution of the perturbations. Standard deviation of the perturbations are 0.2 (red), 0.5 (yellow), 1.0 (green) and 2.0 (blue). Values are averaged over 1000 realizations.

A.3 Test of the adjoint model

The adjoint code was validated through computing

$$\left| \frac{(\mathbf{M}\mathbf{x}')^T(\mathbf{M}\mathbf{x}') - \mathbf{x}'^T[\mathbf{M}^T(\mathbf{M}\mathbf{x}')] }{(\mathbf{M}\mathbf{x}')^T(\mathbf{M}\mathbf{x}')} \right|, \quad (\text{A2})$$

where \mathbf{M} and \mathbf{M}^T are tangent linear model and its adjoint operator. This value should be of the order of the truncation error of the computation. We verified that this value was of the order of $O(10^{-15})$ for double-precision real computations.

References

- Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.
- Bishop, C. H. and Hodyss, D. 2009a. Ensemble covariances adaptively localized with ECO-RAP. Part 1: tests on simple error models. *Tellus* **61A**, 84–96.
- Bishop, C. H. and Hodyss, D. 2009b. Ensemble covariances adaptively localized with ECO-RAP. Part 2: a strategy for the atmosphere. *Tellus* **61A**, 97–111.
- Gelaro, R. and Zhu, Y. 2009. Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus* **61A**, 179–193.
- Hunt, B. R., Kostelich, E. J. and Szunyogh, I. 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D* **230**, 112–126.
- Kunii, M., Miyoshi, T. and Kalnay, E. 2012. Estimating impact of real observations in regional numerical weather prediction using an ensemble Kalman filter. *Mon. Wea. Rev.* **140**, 1975–1987.
- 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.
- Li, H., Liu, J. and Kalnay, E. 2010. Correction of ‘Estimating observation impact without adjoint model in an ensemble Kalman filter’. *Quart. J. Roy. Meteor. Soc.* **136**, 1652–1654.
- Liu, J. and Kalnay, E. 2008. Estimating observation impact without adjoint model in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.* **134**, 1327–1335.
- Lorenz, E. N. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations. *J. Atmos. Sci.* **55**, 399–414.
- Shapiro, R. 1970. Smoothing, filtering, and boundary effects. *Rev. Geophys. Space Phys.* **8**, 359–387.
- Yoon, Y., Ott, E. and Szunyogh, I. 2010. On the propagation of information and the use of localization in ensemble Kalman filtering. *J. Atmos. Sci.* **67**, 3823–3834.
- Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1913–1924.