

Forecast sensitivity with dropwindsonde data and targeted observations

By ZHAO-XIA PU*¹, STEPHEN J. LORD² and EUGENIA KALNAY², ¹*Environmental Modeling Center, National Centers for Environmental Prediction, Washington, DC 20233, USA*; ²*Environmental Modeling Center, National Center for Environmental Prediction, Washington DC, USA*

(Manuscript received 10 October 1997; in final form 7 April 1998)

ABSTRACT

While forecast models and analysis schemes used in numerical weather prediction have become generally very successful, there is an increasing research interest toward improving forecast skill by adding extra observations either into data sparse areas, or into regions where the verifying forecast is most sensitive to changes in the initial analysis. The latter approach is referred to as “targeting” observations. In a pioneering experiment of this type, the US Air Force launched dropwindsondes over the relatively data sparse Northeast Pacific Ocean during 1–10 February 1995. The focus of this study is the forecast sensitivity to initial analysis differences, forced by these observations by using both the adjoint method (ADJM) and quasi-inverse linear method (QILM), which are both useful for determining the targeting area where the observations are most needed. We discuss several factors that may affect the results, such as the radius of the mask for the targeted region, the basic flow and the choice of initial differences at the verification time. There are some differences between the adjoint and quasi-inverse linear sensitivity methods. With both sensitivity methods it is possible to find areas where changes in initial conditions lead to changes in the forecast. We find that these two methods are somewhat complementary: the 48-h quasi-inverse linear sensitivity is reliable in pinpointing the region of origin of a forecast difference. This is particularly useful for cases in which the ensemble forecast spread indicates a region of large uncertainty, or when a specific region requires careful forecasts. This region can be isolated with a mask and forecast differences traced back reliably. Another important application for the QILM is to trace back observed 48-h forecast errors. The 48-h adjoint sensitivity, on the other hand, is useful in pointing out areas that have maximum impact on the region of interest, but not necessarily the regions that actually led to observed differences, which are indicated more clearly by QILM. At 72 h, the linear assumption made in both methods breaks down, nevertheless the backward integrations are still very useful for pinning down all the areas that would produce changes in the regions of interest (QILM) and the areas that will produce maximum sensitivity (ADJM). Both methods can be useful for adaptive observation systems.

1. Introduction

Numerical weather prediction (NWP) is a classical initial-value problem. It has been recognized

* Corresponding author.
Environment Modeling Center, Rm 207, NOAA Space Center, 5300 Auth Road, Camp Springs MD20746, USA.

that forecasts are very sensitive to both initial errors and model errors. In recent decades, numerical weather prediction advanced rapidly through modeling research focused on advances such as the use of dynamics of the primitive equations, better numerical algorithms, advanced physics parameterization, higher resolution models, and optimal adjustment of model parameters. Considerable work also took place on objective

analysis of observations and on model initialization (reviews in Chou, 1986; Haltiner and Williams 1980). As models improved, it has been generally accepted that serious forecast failures are largely associated with analysis errors that amplify rapidly and less so with model deficiencies (Reynolds et al., 1994). In order to enhance forecast skill further, much current research has been focusing on development of advanced analysis methods such as variational data assimilation (Derber, 1987; Andersson et al., 1996) and Kalman Filtering (Cohn et al., 1994). Currently, 3-dimensional variational data assimilation (3D-VAR) has been implemented operationally at NCEP and ECMWF, and has been very successful in the efficient use of non-traditional observations such as satellite radiances (Parrish and Derber, 1992; Andersson et al., 1996). The Four-dimensional variational data assimilation (4-D VAR), a technique based on the adjoint of the forecast model and which constrains model forecasts to fit multiple-time observations has become a popular research topic in the last decade. Although it is computationally expensive (Courtier et al., 1994; Zupanski and Zupanski, 1996), operational implementation has been tested successfully at ECMWF in 1997. Some work has also focused on developing simplified 4-D VAR techniques, based on adjoint or quasi-inverse forecast sensitivity ideas, in order to improve forecast skill (Rabier et al., 1996; Pu et al., 1997a, b; Kalnay and Pu, 1998). Kalman Filtering is computationally even more expensive than 4-D VAR, so research is focused on developing computationally feasible approximations (Cohn et al., 1994; Toddling et al., 1997).

Advanced data assimilation methods have proven able to extract information more effectively from current observations. Derber and Wu (1996) made major improvements in forecast skill by replacing the assimilation of satellite sounding retrievals by the direct assimilation of radiances into the NCEP 3-D VAR scheme. There is also evidence that the analysis can be improved by the assignment of more realistic observation errors. Wu and Joo (1996) improved the analysis quality of NCEP 3-D VAR system by reducing some of the observation error variances which were evidently somewhat overestimated. Pu et al. (1997c) reduced observational error variances in areas of large forecast ensemble spread as a shortcut toward increasing the forecast error covariance in

areas where the ensemble indicates that the first guess has large uncertainties. Experiments indicated that this approach, which allowed for flow-dependent error covariances, had positive impact on the analysis and the forecasts. In a related work, Kalnay and Toth (1994) showed that minimizing the distance between observations and the first guess along the bred vectors, resulted also in positive forecast impacts.

Even though improved analysis methods can make better use of currently available observations, an obvious alternative approach to improving the initial conditions is to increase the observational data base. Currently, northern hemisphere oceans and most of the Southern Hemisphere have few in-situ observations such as rawinsondes, and instead have coverage from satellite observations (TOVS temperature soundings, cloud- and water vapor-tracked winds, surface winds from scatterometers and microwave instruments). Although the use of data from remote sensing instruments has improved substantially the accuracy of the forecasts in the Southern Hemisphere, in-situ vertical profiles of winds, temperatures and moisture such as those obtained from rawinsondes or dropwindsondes are still the most effective observations for numerical weather prediction.

There are two obvious strategies for improving the current coverage with vertical soundings: (1) add soundings uniformly in oceanic areas such as the Pacific Ocean, where few rawinsondes are available, and which is upstream of an area of interest (North America); and (2) put them only in those areas which are not only data sparse and upstream of the region of interest, but also where there is evidence of fast error growth. The second approach, denoted as an "adaptive observing strategy", is based on the evidence that error growth is flow dependent (Pu et al., 1997c), and that it is more efficient to spend the resources available for observations in the relatively small regions where fast error growth takes place. A strategy for targeted weather observations in upstream areas has been designed for the Fronts and Atlantic Storm Track Experiment (FASTEX), with additional drops of atmospheric soundings in areas showing maximum potential impact in the targeted area of interest near the British Isles. Adaptive observation strategies could also be useful for satellite observing systems, if they are designed in such a

way that they could dwell on regions where observations are most apt to be useful.

During 1–10 February 1995, the US Air Force 53rd Weather Reconnaissance Squadron C-130 aircraft flew a total 9 dropwindsonde missions over the Northeast Pacific Ocean, upstream of North America, where in situ data are sparse. The missions were designed to place the dropwindsondes in the region where, subjectively, positive forecast input would be expected over North America. In a preliminary evaluation, Lord (1996) showed that the dropwindsonde data resulted in an overall positive impact on the synoptic-scale weather forecasts. In this paper, we summarize and further analyze the impact of these data and focus on the relationship between the forecast sensitivity and targeted weather observations. As in Pu et al. (1997a, b), the forecast sensitivity to changes in the initial conditions will be investigated using the adjoint of the NCEP global tangent linear model and the quasi-inverse linear model. The forecast initial errors will be traced back from the difference between forecasts with and without dropwindsondes. These results are used to confirm the impact of the dropwindsonde observations on weather forecasts, and to recognize the relationship between the local forecast error and initial analysis impact. A local mask in the verification area is introduced and results of the forecast sensitivity in this area compared with the known analysis changes.

The organization of this paper is as following: Section 2 describes the results from US Air Force dropwindsonde experiment. The sensitivity methods using adjoint and quasi-inverse models are described in Section 3, and the use of a local mask is also discussed. In Section 4, we perform the sensitivity experiment for two cases, one representing a large positive forecast impact, and the other a null forecast impact. Additional experiments performed testing several factors which may impact the targeting area, and comments on strategies for targeting observations are presented in Section 5. Section 6 is a summary and discussion.

2. Impact of dropwindsondes on the synoptic weather forecast

During 1–10 February 1995, the United States Air Force Reserves (AFRES) 53rd Weather

Reconnaissance Squadron (WRS) flew 9 reconnaissance missions with WC-130 aircraft over the Northeast Pacific Ocean off the coast of Oregon, Washington and northern California in the region bounded by 30–60°N, 125–155°W (Fig. 1). The flight patterns for each mission were determined daily by forecasters from the National Centers for Environmental Prediction (NCEP), the Seattle Weather Forecast Office, and NOAA researchers. A total of 126 sondes was dropped over the 9 missions. Measurements of wind and thermodynamic variables were transmitted in real time to NCEP and covered the layer from approximately 300 hPa (flight level) to the surface. The number and initial time of those soundings which match to the NCEP standard analysis time are listed in Table 1. All those data will be used in the evaluation.

The impact on synoptic-scale forecasts over the United States of these dropwindsonde observations was evaluated by Lord (1996), using the NCEP operational T126/L28 global spectral model and Global Data Assimilation System (GDAS), based on the spectral statistic interpolation (SSI) system (Parrish and Derber, 1992). A control run was performed by running the GDAS over the period 12 UTC 31 January 1995 — 00 UTC 14 February after removing all dropwindsonde data from the operational data files. The experimental run included all dropwindsonde data. Verifications for each run were performed by comparing 24-, 48- and 72-h forecasts with corresponding analyses over much of North America (65–125°W, 25–65°N, Fig. 1). The impact of the supplementary dropwindsonde data is summarized by a time series of the standard deviation of the 500hPa height forecast error (Fig. 2), showing that there is a large positive impact on 12 UTC 7 February and 00UTC 8 February. These impacts increase considerably with forecast hour and amount to about a 20% reduction in forecast error at 48 and 72 h. Smaller positive impacts are registered at 00UTC 1 February and 00UTC 9 February. There is little or no impact at the other times, which constitute a majority of the cases (5 out of 9). Verifications of wind fields showed positive impact, especially at 24 h (not shown). There was no impact on standard precipitation scores.

The large variability of forecast impact implies that the effectiveness of observational input to the

500 hPa Height ANALYSIS

Time Average: 12 UTC 1 February 1995 to 00 UTC 11 February

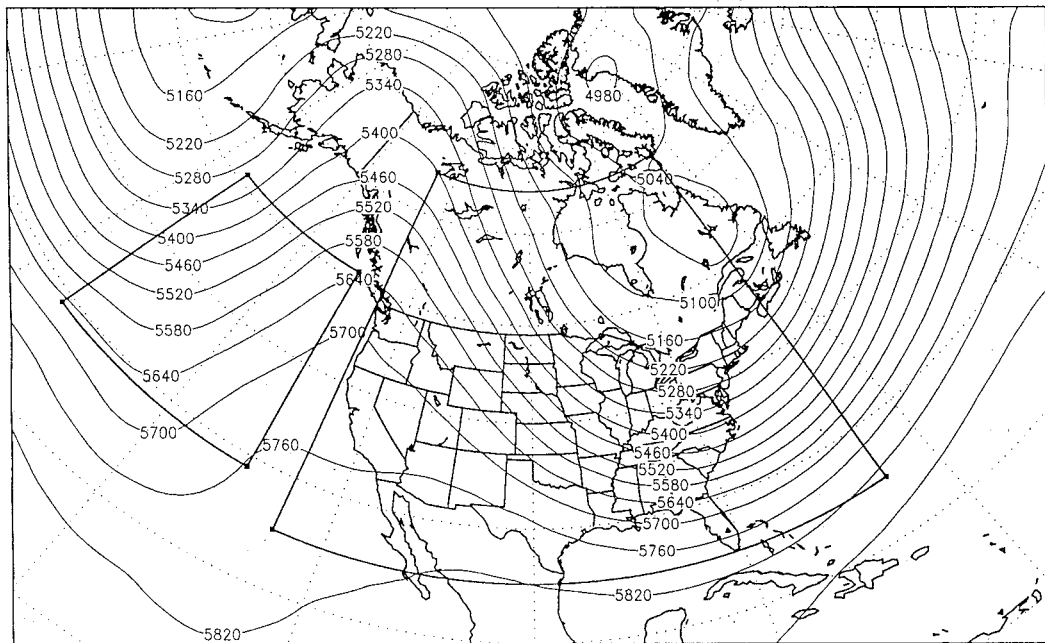


Fig. 1. Time mean 500 hpa analysis over 1–10 February 1995. The left-hand box encompasses the area over which all observations were taken; the right-hand box is the forecast verification area.

Table 1. Initial time and number of observation which used for the evaluation

Date (month/day UTC)	Number of wind soundings	Number of mass soundings
2/1 1200	7	8
2/2 1200	8	8
2/3 0000	2	2
2/4 1200	4	5
2/7 1200	6	6
2/8 0000	5	6
2/9 0000	4	8
2/10 0000	5	6

analysis has a critical dependence on the synoptic situation. Therefore, it is important to understand how the differences in the analysis affect the forecast differences, and in which situation the dropwindsondes would cause positive impact on future forecasts. In the following section, we will focus on the forecast error sensitivity analysis by

tracing the forecast differences back to the initial condition in order to understand how the dropwindsonde data relates to the forecast error.

3. Sensitivity of forecast differences to changes in initial conditions: general description of the method

3.1. Forecast sensitivity with the adjoint and the quasi-inverse methods

Assume that at time t , for a nonlinear forecast model M , the forecast differences between two forecasts started from two different initial condition X_0 and $X_0 + \delta X_0$ are defined as:

$$\delta X = M_t(X_0 + \delta X_0) - M_t(X_0). \quad (1)$$

For a perfect model and sufficiently small forecast differences and short time scales, the forecast difference can be approximated by the propagator L or tangent linear model (TLM) of

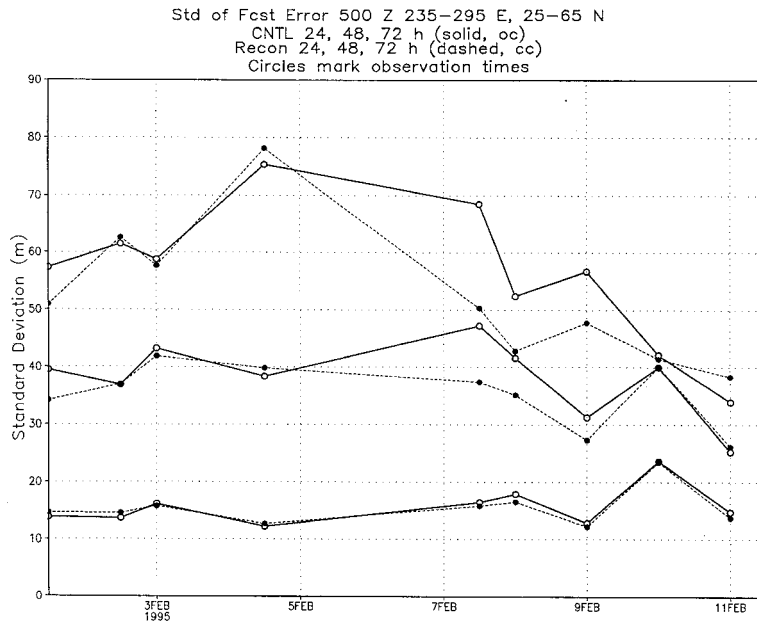


Fig. 2. Time series of the standard deviation of 500 hPa height errors over the verification area for control (solid line, open symbols) and dropwindsonde experiment (dashed line, closed symbols). The abscissa refers to the forecast initial time; observation times are marked with symbols. The bottom, middle and top pairs of lines are for 24, 48 and 72-h forecasts respectively.

the model M :

$$\delta X = L\delta X_0 \tag{2}$$

Forecast errors may be traced back to errors in the initial conditions by two methods: the adjoint (ADJM, hereafter, e.g., Rabier et al. (1996); Pu et al. (1997a)) and the quasi-inverse linear method (QILM, hereafter, e.g., Pu et al. (1997b)).

(a) adjoint method (ADJM)

As in Rabier et al. (1996) and Pu et al. 1997a, we define an error cost function (using for example a total energy norm) J :

$$\begin{aligned}
 J = \delta X^T W \delta X = & \int_0^1 \int_{\Gamma} \int_{\Gamma} \left(\nabla \Delta^{-1} \zeta \nabla \Delta^{-1} \zeta \right. \\
 & + \nabla \Delta^{-1} D \nabla \Delta^{-1} D + R_a T_r (\pi)^2 \\
 & \left. + \frac{C_p}{T_p} T^2 \right) d\Gamma \left(\frac{\partial P_r}{\partial z} \right) d\eta \\
 \approx & (L\delta X_0)^T W (L\delta X_0), \tag{3}
 \end{aligned}$$

where, $\delta X = (\zeta, D, T, \Pi)$ denotes differences

between two forecast for vorticity, divergence, temperature and natural logarithm of surface pressure respectively. T_r, P_r are the reference temperature and pressure, R_a is the gas constant for dry air, C_p is specific heat at constant pressure for dry air. Γ represents the horizontal integration domain. W is the matrix of weights defining the norm. The weights are a function of $T_r, P_r, R_a,$ and C_p . η is the vertical coordinate. The superscript T denotes the transpose of a matrix. ∇ represents the gradient operator and Δ represents the Laplacian operator.

The gradient of cost function J respect to initial condition X_0 can be computed by the adjoint operator L^T of the tangent linear model L

$$\nabla_{X_0} J = L^T W L \delta X = L^T W (M(X_0 + \delta X_0) - M(X_0)). \tag{4}$$

The gradient can therefore be obtained by integrating the adjoint model backwards once. i.e., starting from the forecast differences at final time t to the forecast initial time 0. It gives the "optimal" direction that results in the maximum decrease of the cost function for a given size perturbation.

The pattern of the gradient of J at time zero shows the sensitivity of the forecast error to the initial conditions (Rabier et al., 1996). In order to made the changes in the initial conditions, the gradient has to be multiplied by an appropriate amplitude or a preconditioning matrix (Derber, 1987). In this paper, we specify that the ‘‘adjoint sensitivity’’ field means a preconditioned gradient, i.e., $\alpha W^{-1} \nabla_{X_0} J$, with the value of α is equal to 0.1.

(b) *Quasi-inverse linear method (QILM)*

Pu et al. (1997b) introduced a second approach: solve (2) directly by approximating the inverse of the TLM L

$$\delta X_0 = L^{-1} \delta X. \quad (5)$$

It is the solution obtained when we (approximately) trace the short-range forecast error back to initial time. Since the small dissipative terms are irreversible, we cannot invert the TLM exactly. In practice, we approximate L^{-1} by the ‘‘quasi-inverse’’ method: simply run the TLM backwards by changing the sign of Δt , while also changing the sign of the dissipative terms to avoid computational instability (Pu et al., 1997b). Because dissipation is generally small except near the surface, this approach provides quite an accurate approximation of the exact inverse of the TLM and, therefore, a good approximation to the exact solution of (2) (see details in Pu et al. (1997b)).

3.2. *Forecast model and its tangent linear and adjoint model*

The model used for the sensitivity analysis is the NCEP operational global spectral model (known as medium range forecast or MRF model) with horizontal resolution T62 and 28 vertical sigma levels. An adiabatic version of this model with a minimum set of physics (horizontal diffusion, vertical mixing and surface drag) is the basis for the adjoint and TLM and is the same as used in Pu et al. (1997a, b).

3.3. *A local mask*

In earlier studies concerning the forecast sensitivity analysis, we took the global forecast error as an initial condition and then integrated both adjoint and quasi-inverse linear model backwards

to the forecast initial time. However, for the current adaptive observation problem, the forecast area we are targeting is a specific region, and it is necessary to find the area of sensitivity affecting this area. The northern hemisphere north of 30°N as verifying area for the generation of ensemble perturbations Buizza (1994) used, which was also used for forecast sensitivity studies in Rabier et al. (1996). Langland and Rohaly (1996) and Kalnay et al. (1996) suggested the use of a local mask operator S over the area of verification:

$$\delta X|_{\text{local}} = S \delta X|_{\text{global}}. \quad (6)$$

The mask operator S gives a non-zero weighting inside the targeted region and a zero weighting outside so that forecast perturbations in the targeted area are retained, and perturbations in other regions are removed. Here we use a circular smooth-sided mask filter on the sphere based on the incomplete gamma function (Abramowitz and Stegun (1965); James Purser, personal communication) which gives the largest weight (equal to 1) in the central area of the circular mask and smoothly reduced weight to zero at the mask boundary. This distribution minimizes possible imbalances at the boundaries of the targeted region. The effect of the mask operator on the cost function is calculated by multiplying the original weighting factor W in (3) by the geographically varying mask weight. As a result, the cost function became local even though it is still calculated globally. For the QILM (5), there is no cost function but the initial global forecast perturbation is multiplied by the same mask. Therefore, the backward integration traces the forecast error from the specified region.

4. **Forecast sensitivity with dropwindsonde observations**

Two cases representative of extremes in forecast impact were chosen for this forecast sensitivity analysis. The first one (0000 UTC 8 February) had a major forecast impact over the US and the second (0000 UTC 3 February) had no impact over the US (Fig. 2). The mask (6) was applied to differences between two global forecasts, one initiated from the analysis that used all the dropwindsonde data, and the other one from the analysis without any dropwindsonde data. The masking

region (Fig. 3a) which has a radius of 1000 km and is centered at 45°N, 95°W covered a large fraction of the verification area (Fig. 1). The adjoint or quasi-inverse linear model were integrated backwards to the initial time.

4.1. 00 UTC 8 February case

We first take differences in 48-h forecasts (a “48-h window”) (for which the linear assumption used in the linear and adjointed models is fairly

valid) and then a 72-h window, for which the forecast differences begin to be non-linear but for which this methodology can still provide useful guidance for targeted observations.

Fig. 3a shows the 48-h global forecast temperature differences at sigma level 7 (about 850mb) for initial condition at 00 UTC 8 February. It shows large differences mostly over North America that move eastward with time (if compared with the 72-h forecast differences). Similar features also observed in wind field. Masked differences are

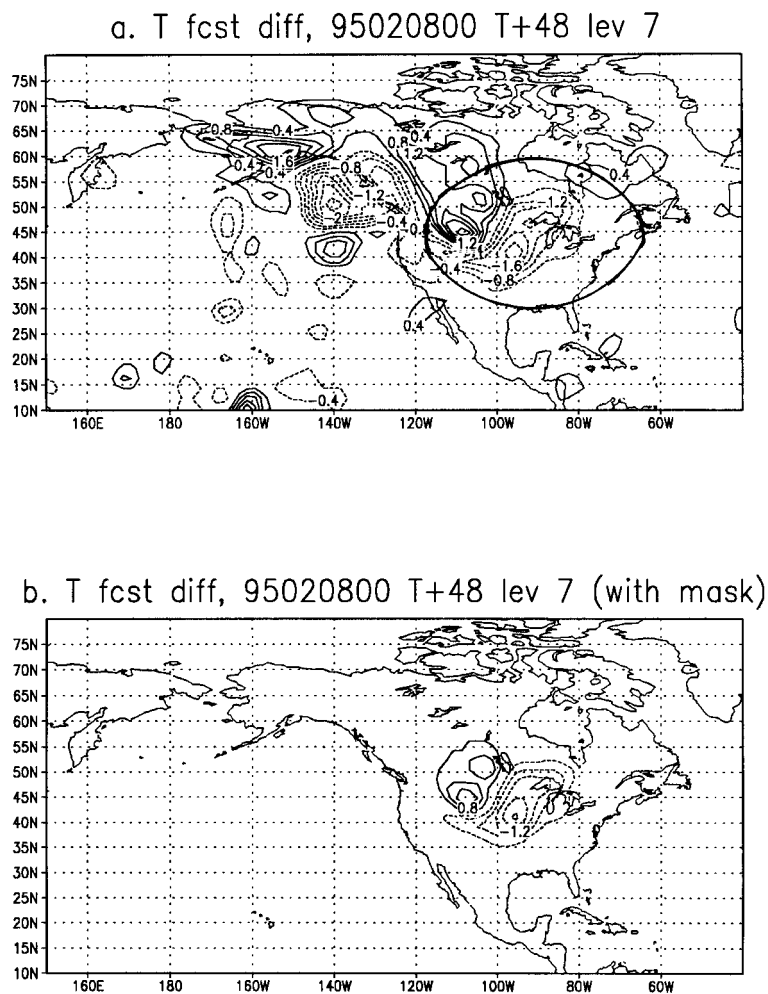


Fig. 3. (a) 48-h forecast differences at sigma level 7 (about 850 hpa) between the control and experimental forecasts at 00 UTC 8 February 1995. The contour interval is 0.4 K. There is no zero line in the contour. The area centered at (45°N, 95°N) with 1000 km radius indicates the geographical extent of the mask. (b) Same as Fig. 3(a) except for forecast differences after the mask filter.

shown in Fig. 3b. Taking these forecast differences in the mask region as initial conditions, we performed backwards integration for both of adjoint and quasi-inverse linear models.

Sensitivity field obtained from the ADJM and sensitivity perturbation obtained from QILM for temperature at the initial time (0000 UTC 8 February) (Figs. 4a, b) have some similarity with

analysis differences (Fig. 4c), with QILM having slightly better agreement. Both show maximum amplitude in the perturbation below 700 hPa around 140°W–150°W. The analysis difference and quasi-inverse perturbation have both secondary maxima in the upper troposphere, a feature which is absent from the adjoint sensitivity. The adjoint sensitivity is very compact, going from

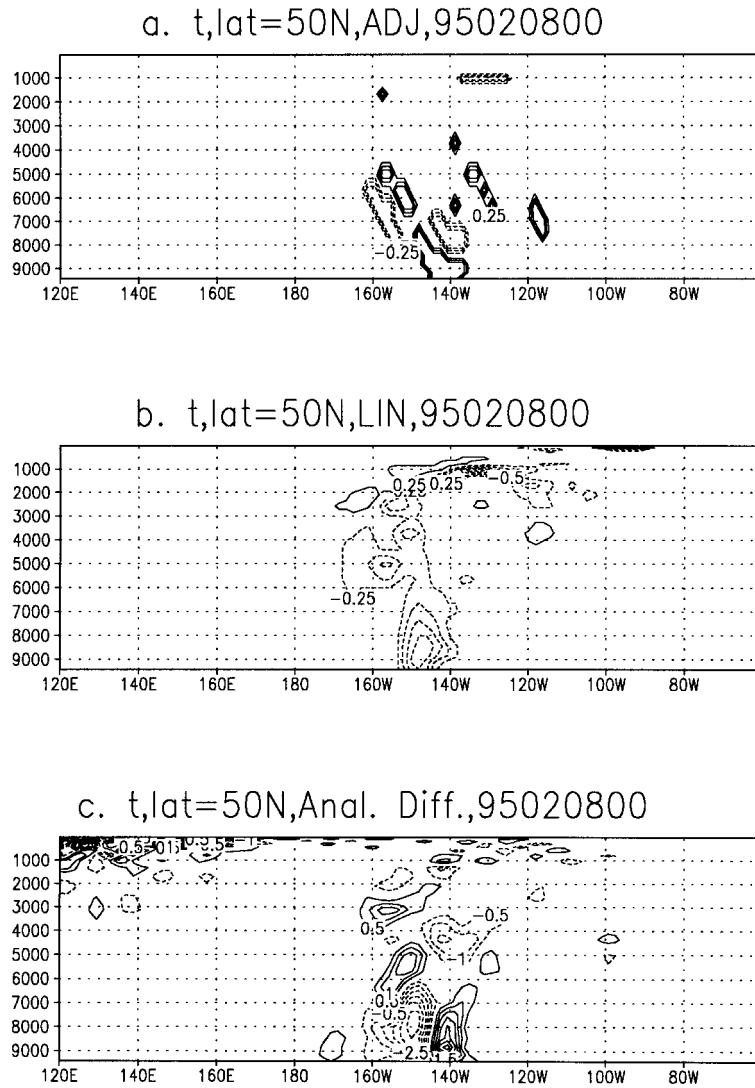


Fig. 4. Vertical cross section at 50°N of the sensitivity perturbation for temperature calculated from 48-h masked forecast differences by the adjoint method (a), by the quasi-inverse linear method (b), and the analysis differences at 00UTC 8 February 1995 (c). The contour interval is 0.25 K in (a) and (b), and 0.5 K in (c). Fig. 4. (d)–(f) are same as Fig. 4(a)–(c) except for vorticity field. The contour interval is $0.5e-5/s$ in (d) and (e), and $1.e-5/s$ in (f).

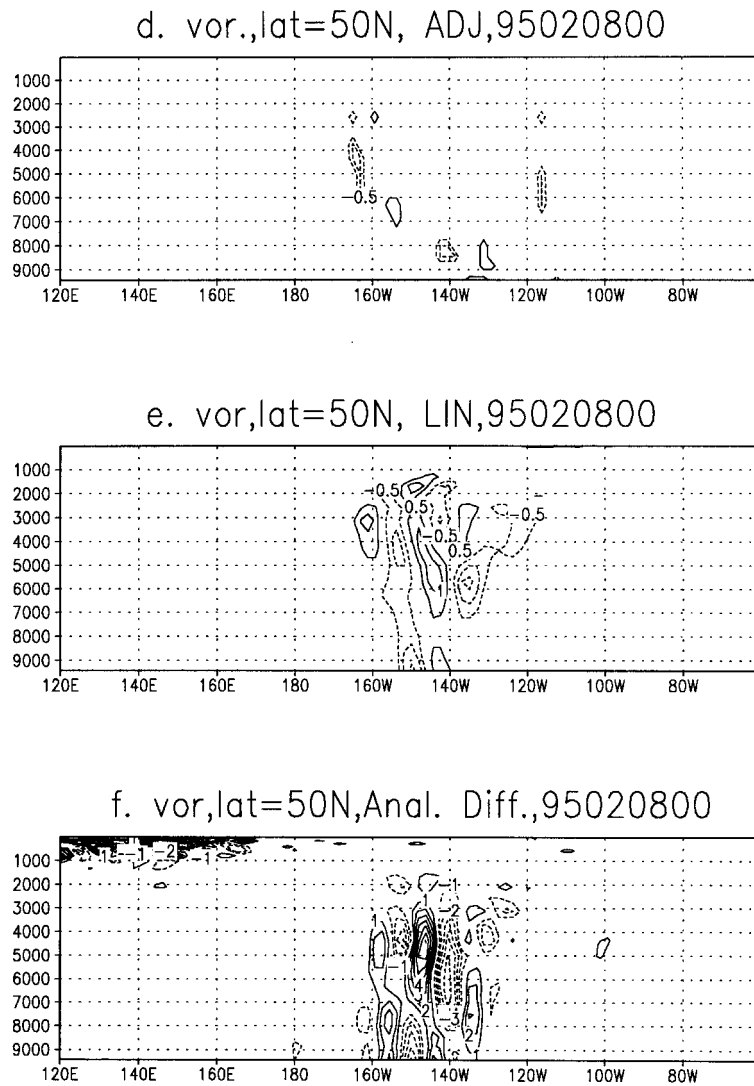


Fig. 4. (Continued).

essentially zero to the maximum value very sharply when compared with the linear inverse or the analysis differences, and it shows an extremely strong baroclinic tilt at low levels, also observed in other studies (Buizza and Palmer, 1995). The amplitudes in the quasi-inverse linear temperature differences are smaller than those of the analysis by a factor of about 3. For the vorticity differences (Fig. 4d-f), the agreement with analysis differences reasonably good for QILM, although the amplitudes are again about a factor of 3 smaller.

Both differences show a maximum perturbation centered around 145°W , with low and upper level maxima. The adjoint sensitivity in the vorticity is quite different from the other two, with little overlap in the area of perturbation of the analysis or the linear sensitivity, and the magnitude is an order of magnitude smaller than for the temperature.

The same experiment was also performed for 72-h forecasts, starting from the differences between the forecasts with and without the drop-

windsondes, masking the differences in the same region as in Fig. 3a with the mask radius 1000 km, and integrating both the quasi-inverse linear and adjoint model back to initial time. The results (not shown) are qualitatively very similar to those of 48 hrs, except that the sensitivity fields are slightly more extended in the horizontal and vertical.

The adjoint sensitivity, like the singular vectors with which it is closely associated, is strongly dependent on the choice of the norm (Rabier et al., 1996; Palmer et al., 1998). The energy norm is known to produce very small amplitude wind or vorticity perturbations, which are far from quasi-geostrophic balance (Szunyogh et al., 1997). Our results confirm these earlier studies, and indicate that the low weights given to the vorticity compared to the temperature in the energy norm defining the cost function (3) do not result in realistic sensitivity patterns. Because the vorticity fields are relatively small (comparing with the temperature) in adjoint sensitivities, and in order to display the sensitivity field more clearly with the different methods, in most of the following experiments we will show the sensitivity field for the temperature field only and only discuss vorticity in the text.

Considering that the dropwindsonde is a vertical sounder throughout the troposphere, we define the vertically-averaged squared vorticity and squared temperature fields from model levels 7–18 (sigma 0.845–0.210) to represent the sensitivity signal. Fig. 5a–c show the vertically averaged squared temperature differences field for the quasi-inverse linear (Fig. 5b) and adjoint (Fig. 5a) at initial time for 48-h window calculations. Fig. 5c showed the corresponding analysis difference at initial time. There is excellent agreement between the quasi-linear and adjoint sensitivities for temperature; they both indicate an area south of Alaska that corresponds closely to the area of largest analysis differences, as well as smaller difference centers on the West Coast of North America and Northern Canada not apparent in the analysis. For the squared vorticity, however, the adjoint sensitivity does not have much geographical overlap with the analysis differences (not shown).

Fig. 6a–b show the sensitivity fields for the squared temperature for a 72-h windows. By 72 h, the linear assumption used for both the quasi-inverse and adjoint linear models are clearly violated; nevertheless the results, though less compact, are still encouraging. The areas pointed out

by both 48 and 72-h quasi-inverse both include the area south of the Gulf of Alaska where the analysis differences are maximum (Fig. 5c). However, at this longer range both the linear and the adjoint sensitivities indicate an area further upstream centered at the dateline and 35N, which is not apparent in the analysis differences (Fig. 5c). In the squared vorticity differences, the adjoint has smaller amplitudes, and does not have the corresponding central Pacific upstream area. The fact that both methods agree in pointing out additional areas upstream suggests that additional sondes dropped as far west as the dateline would have further improved the 72-h forecast over the US.

4.2. 00 UTC 3 February case

Although 0000 UTC 3 February is a case of virtually no impact over the US region (Fig. 2) and there were only two dropwindsondes at this analysis time, the impact of the data on the analysis and forecast in area where the soundings were taken is still large for both wind field and surface pressure field (figure not shown) compared to other times. As with the case of large impact of 8 February, we computed the forecast sensitivity starting from the 48-h forecasts with and without dropwindsondes, and masked their difference over the US, with a 1000 km mask centered at 45°N, 95°W, as in Fig. 3a.

Fig. 7 shows the sensitivity initial differences calculated from the 48-h forecast differences for the vertically averaged squared temperature. The quasi-inverse linear method shows only a small sensitivity area overlapping the regional of analyzed temperature differences. The adjoint places an area of temperature sensitivity south of the analysis difference, and another stronger area of sensitivity in NW Canada as does the quasi-inverse linear sensitivity.

When the sensitivities are computed from 72-h forecasts, results are similar to those of 48-h. Even at 72-h the linear method indicates correctly that the small forecast differences over the US are due in part to changes over the Gulf of Alaska, but it also suggests that larger sensitivity comes from other regions such as the West Coast, the Arctic region north of Alaska, and the subtropical Pacific. The adjoint approach shows no sensitivity in the Gulf of Alaska, indicating that initial analysis differences in that area would not grow as they

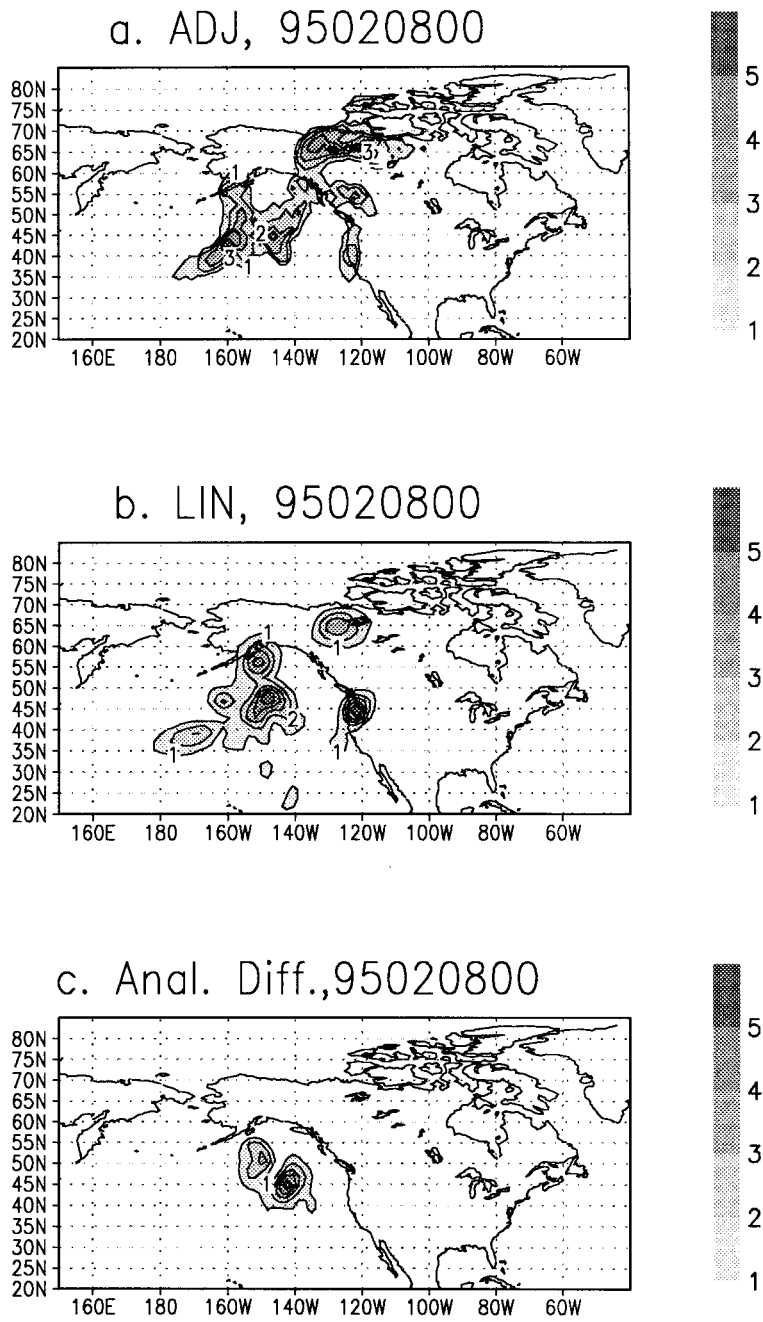


Fig. 5. The vertically average sensitivity field for squared temperature over sigma levels 7–18 at 00 UTC 8 February 1995, calculated from 48-h forecast differences by the adjoint method (a), by the quasi-inverse linear method (b), and the analysis differences (c). The contour interval is $1.e-1K^*K$ in (a) and (b), and $1K^*K$ in (c).

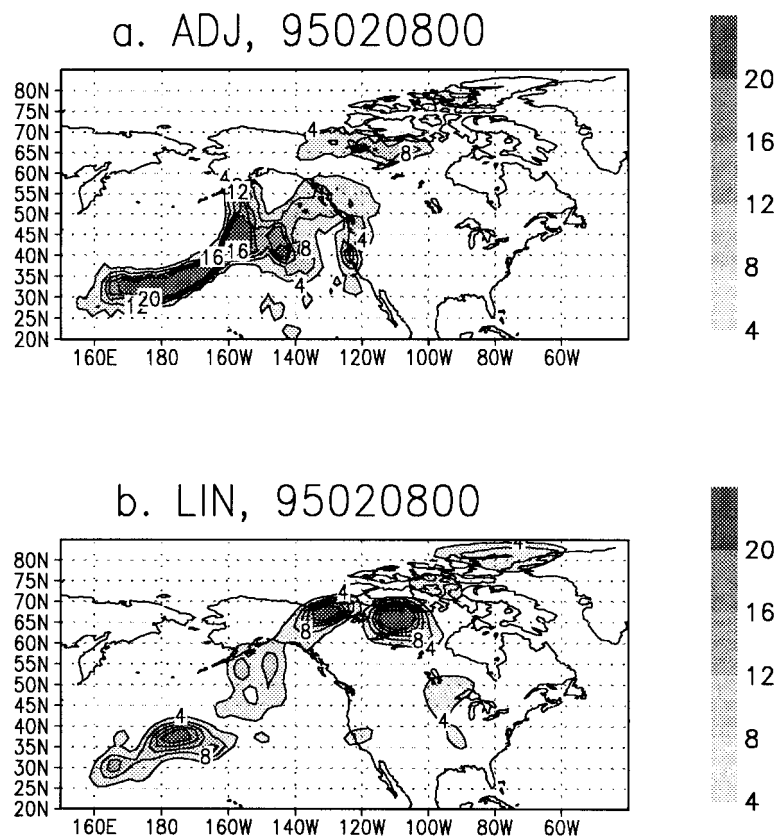


Fig. 6. Same as Fig. 5(a)–(b), except the sensitivity field calculated from 72-h forecast differences. The contour interval is $4 \times 10^{-1} \text{K}^2$.

evolved over the US. The 72-h adjoint result also shows areas of sensitivity North of Alaska and in the subtropical Pacific and agrees well with the 48-h adjoint sensitivity.

The forecast differences (not shown) show that the temperature differences for the case of 3 February are indeed quite small over the US, and larger over the Northern Canada. Vorticity differences had a similar pattern (not shown). To investigate their sensitivities, we repeated the backward experiments with a 500 km mask centered over the Northern Canada (65°N , 105°W). The 48-h sensitivity results (Figs. 8a, b) show that for the area of Northern Canada, where there were larger forecast impacts than over the US, the temperature differences in the linear sensitivity are consistent with observed differences (Fig. 7c). The adjoint sensitivity is relatively small in magnitude but the sensitivity area still shows good agreement

with the linear sensitivity. The excellent agreement between the sensitivity and analysis differences indicate that the dropwindsonde data on this day strong impact on the forecast in Northern Canada rather than on the US region where they had a null impact.

The different sensitivity results suggest that we need to further explore the impact of different choices (e.g., the mask region, sensitivity methods, etc.) on the sensitivity computation. In next section, we will further discuss the above results and test several other factors which could have an effect on the sensitivity calculation.

5. The impact of some choices in the computation of the forecast sensitivity

First, as we have seen in previous section, in a case of large forecast impact (e.g., over the US

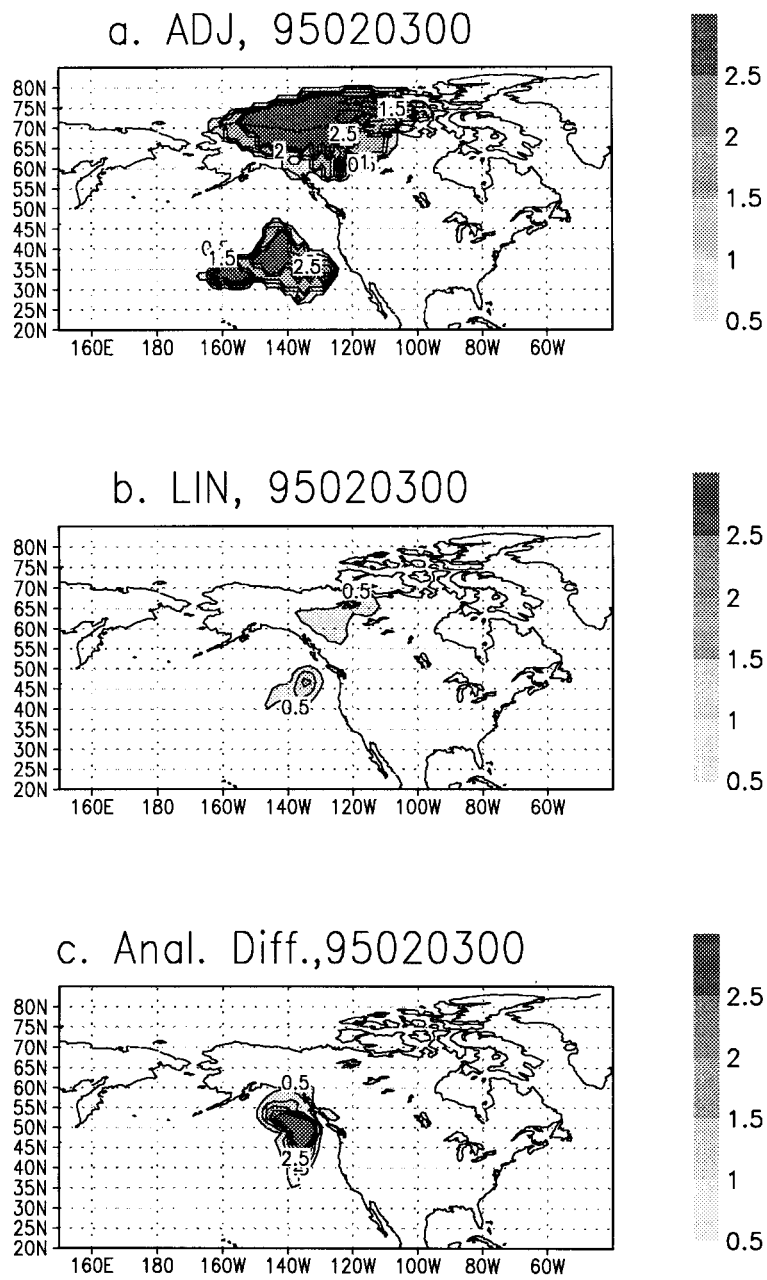


Fig. 7. Same as Fig. 5, except for 00 UTC 3 February 1995 case. The contour interval is $0.5e-1K^*K$ in (a) and (b), and $0.5K^*K$ in (c).

region at 00UTC 8 February and the Northern Canada area at 00UTC 3 February), the adjoint sensitivity and quasi-inverse linear sensitivity point out the same sensitivity region, but in a case

of low impact (for example, over the US region at 00UTC 3 February), the sensitivity area calculated from two methods can look quite different. In our earlier studies concerning the forecast error sensit-

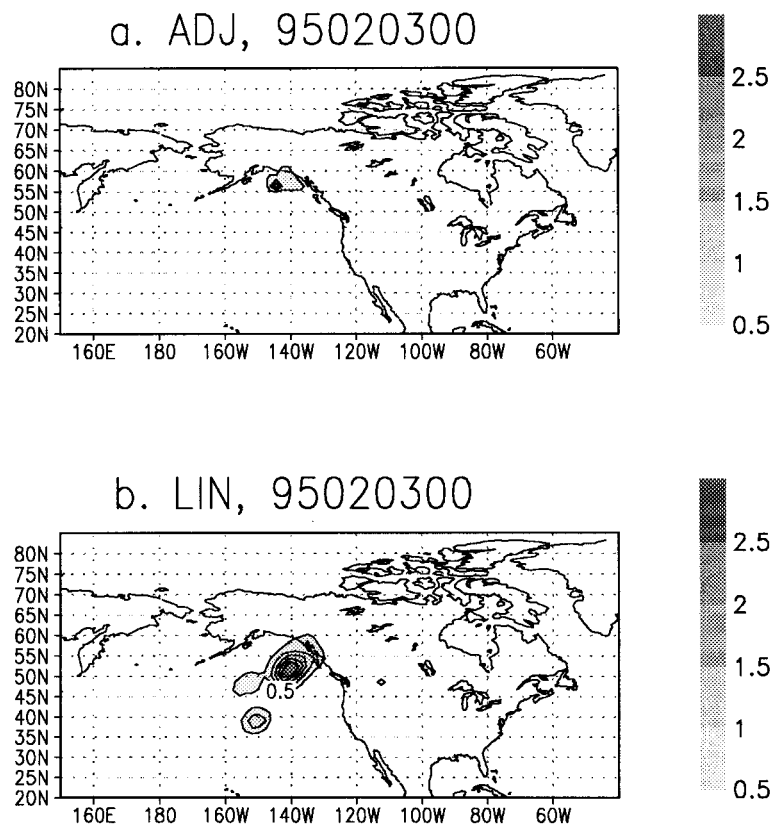


Fig. 8. Same as Fig.7(a)–(b), except the sensitivity field calculated from 48-h forecast differences with a mask filter centered at (105°W, 65°N) with 500 km radius for 00 UTC 3 February case.

ivity (Pu et al., a, b), we pointed out that the sensitivity patterns computed from the two methods have different properties. The adjoint sensitivity is dominated by the fastest growing singular vectors and as such depends very strongly on the type of norm used (Palmer et al., 1998). If the energy norm is used, as in the majority of studies, including the present, it tends to have a very strongly baroclinic structure, and, as pointed out by Szunyogh et al. (1997), the temperature and velocity perturbations are very far from quasi-geostrophic balance, with the vorticity being at least an order of magnitude smaller than could be expected from a state of balance. The quasi-inverse linear sensitivity is in much better geostrophic balance, but (unlike the adjoint sensitivity) contains both growing and decaying errors, and therefore tends to have larger amplitudes than the adjoint sensitivity. These properties help to explain

the results obtained so far, and in particular why they were different in the cases of large and small forecast impact. For large impact the two methods show the same large sensitivity area because the changes in the analysis were fast growing errors. For small impact, the quasi-inverse linear indicates the area where the forecast differences originated but which lead to small forecast changes, and the adjoint sensitivity shows other areas where drop-windsondes would have had larger effects. The results suggest that the two methods are complementary, and that they could both be used in future experiments for comparison and decision making.

We have seen that by 72 h, both methods extend the area of sensitivity well beyond the area of maximum difference in the analysis. For example, in the case of large impact (8 February), both methods show large sensitivities over Canada and

in the central Pacific, further north and west from the area where the sondes were dropped. These differences between the sensitivities and the true analysis differences could be due to several possible causes: (a) the choice of a regional mask, which necessarily distorts the initial balance of the forecast differences; (b) the linear assumption does not hold by 3 days; (c) the lack of proper physical parameterizations in the quasi-inverse linear and in the adjoint models; (d) the basic flow solution that we have taken in our experiments can contain significant errors and affect the sensitivities.

In this section we explore further some of the possible causes and how they affect the sensitivities. We use the 0000UTC 8 February case for these tests.

5.1. Impact of the mask size

We first test the impact of the size of the mask radius. For this purpose, we take a 72-h forecast window and compare three different experiments: (1) calculate the initial sensitivity from the forecast difference in the US region, with a mask centered at 45°N, 95°W and with a radius of 500 km; (2) same as (1), but with a mask radius of 1000 km (same case discussed in Section 4); and (3) calculate the initial sensitivity from the *global* forecast differences; Fig. 9 shows the sensitivities measured by the vertically averaged squared vorticity for the quasi-inverse linear sensitivities. Very similar patterns of sensitivity also observed for the adjoint sensitivities (figure not shown). The results are encouraging: the two methods agree quite well in the areas identified as sensitive from the global initial differences. When the regional masks are applied, both sensitivities shrink, as expected, but the adjoint shrinks faster, both in area and in magnitude. As we saw before, the QILM points more clearly to the area where the true analysis differences are located, and shows good agreement with the area of analysis differences. The reduction in sensitivity area as the mask is reduced and the consistency among the three areas of sensitivity indicates that the negative effects of using a regional mask are not large, and that it should be possible to determine the sensitivity area for a given region using, for example, the local ensemble spread to define the initial uncertainties (Kalnay et al., 1996).

5.2. Impact of the basic flow

The linear tangent model and its adjoint are defined with respect to a basic nonlinear flow, which we assume is known with sufficient accuracy. In real time situations, the best estimate of the evolution of the atmosphere at the verification time (i.e., the analysis) is not available 2 or 3 days in advance for sensitivity calculations. Therefore, we are forced to use a substitute estimate, i.e., a nonlinear forecast. In these experiments, in order to simulate a real time situation, we have used the forecast from the best analysis (with the dropwindsondes) as a basic flow.

We tested 3 estimates of the basic flow: (1) the forecast starting from the initial analysis at 0000 UTC 8 February 1995 using the dropwindsonde data; (2) the forecast starting from the control initial analysis without the dropwindsonde data; (3) 6-h analyses over the forecast period 0000 UTC 8 February 1995 to 0000 UTC 11 February 1995, which is our best estimate of the atmosphere. For QILM, changing the forecast shows little sensitivity with respect to the impact of the dropwindsondes on the basic forecast (not shown). However, when we use the analysis as the basic flow, a second large sensitivity area appears over Northern Canada, which was not clearly present in the other two experiments (figures not shown). These results show that the basic flow has some effect on the initial sensitivity perturbations, which are to some extent “basic-flow dependent”. With the ADJM, the impact of the basic flow on sensitivity perturbations is stronger. The temperature adjoint sensitivity (Fig. 10) shows better agreement with the analysis differences when the analysis is used as basic flow (Fig. 10c), indicating that a more accurate basic flow will improve the location of the large sensitivity area. However, these experiments suggest that 48–72 h forecasts may be sufficiently to indicate realistic sensitivity areas.

5.3. Impact of the initial differences

Lord (1996) compared the relative impact of mass only and wind only dropsonde data and found that winds forced larger forecast impact. For the case of 0000 UTC 08 February 1995 we test the ability of the wind data to recover temperature (mass) data by using the vorticity differences

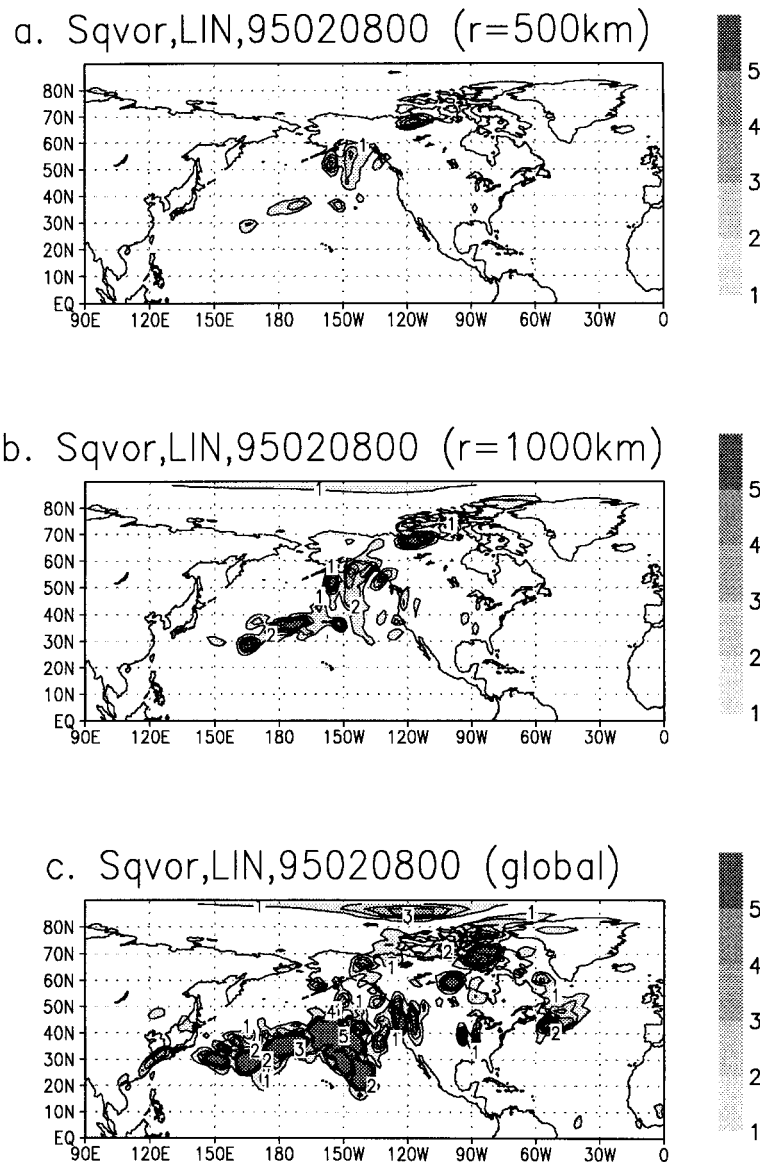


Fig. 9. The vertically average QILM sensitivity field for squared vorticity over sigma 7–18 at 00 UTC 08 February 1995. Calculated from 72-h forecast differences over (a) the US region, mask is centered (95°W , 45°N) with a radius of 500 km, (b) the US region, mask is centered (95°W , 45°N) with a radius of 1000 km, and (c) the global region. The contour interval is $1.e-10/\text{s}/\text{s}$.

between two forecasts (with and without the dropwindsondes) at the verification time as the initial condition for the sensitivity calculation, and zeroing out all other components (temperature, divergence and pressure). This calculation is equivalent

to keeping only the first term of the energy norm in (3) and excluding all the other terms. The results with both of QILM and ADJM (figure not shown) that by using the vorticity differences only, we get almost the same

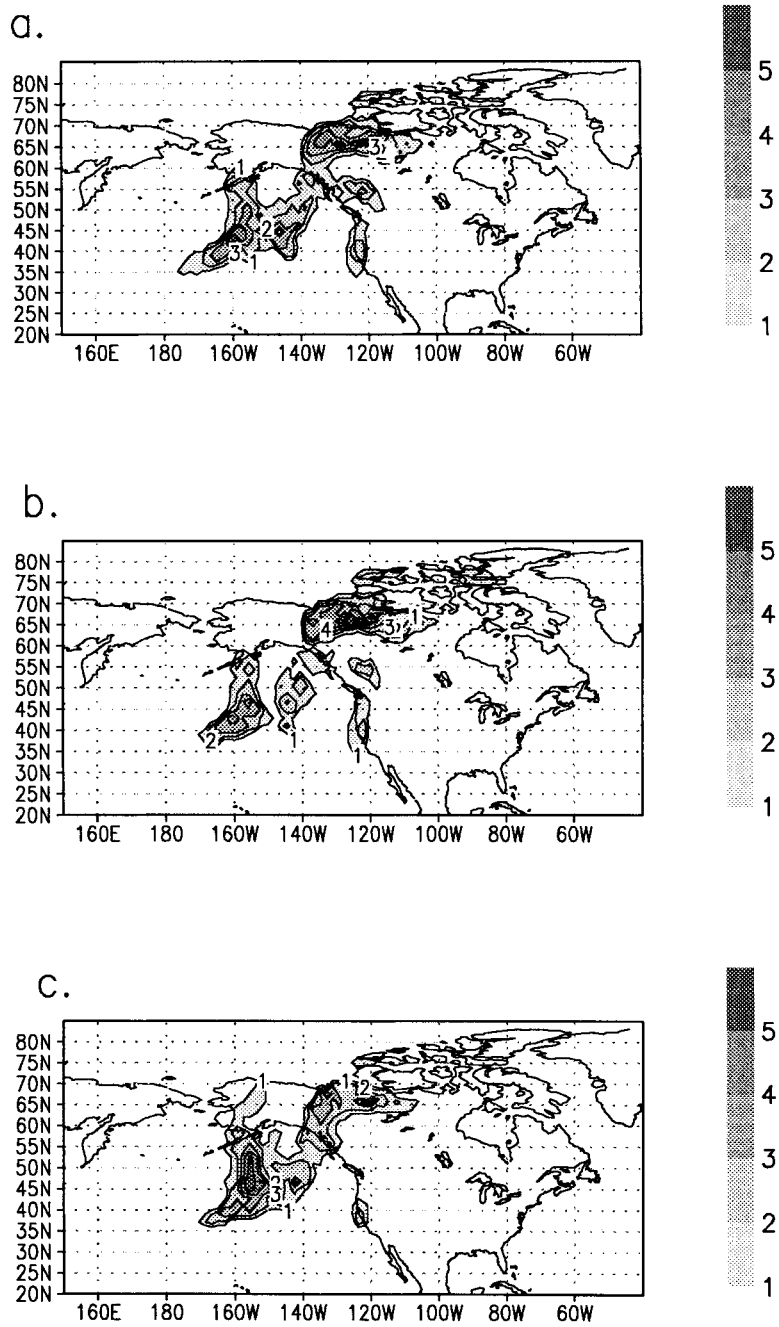


Fig. 10. The vertically average sensitivity fields for squared temperature (unit: $1.e-1 K^*K$) over sigma 7-18 at 00 UTC 08 February 1995, calculated by adjoint method from 48-h forecast differences with a mask filter centered at (95°W, 45°N) with 1000 km radius. The different basic flows: (a) the forecasts which started from the initial analysis with dropwindsonde data at 00 UTC 8 February 1995, (b) the forecast started from the control initial analysis without the dropwindsonde data, and (c) every 6-h analysis over the forecast period 00 UTC 08 February 1995 to 00 UTC 10 February 1995 were used for the calculations.

geographic sensitivity as when we used forecast differences for all components, except that the magnitude is smaller. Even though at the initial (verification) time only vorticity differences are available, at the end of the backward integration the sensitivity features for temperature are also recovered (compared with Fig. 5).

In the same way, we also test the ability of the temperature data to recover wind data by using the temperature differences between two forecasts (with and without the dropwindsondes) at the verification time as the initial condition for the sensitivity calculation, and zeroing out all other components (vorticity, divergence and pressure). The results showed that we recovered much of the information at initial time for both temperature and vorticity field by QILM (compared with Fig. 5b and Fig. 9b). However, the calculation with ADJM showed a different conclusion: the temperature field is recovered over a small area and with a small magnitude, but the recovered vorticity field is essentially zero. The results indicated that only the QILM can produce reasonable results if only the temperature information is used.

6. Summary and conclusions

In this study, we used analyzes and forecasts from a dropwindsonde experiment performed over the northeast Pacific in February 1995 (Lord, 1996) to perform a forecast sensitivity study, and to provide guidance toward the development of a methodology for adaptive observation systems. In an adaptive observing system, the crucial question is how to determine the target area where the data are most needed. In this study we already knew the location of the additional data, so we studied two cases, one with a large impact from the dropwindsondes over the US, and the second with minimal impact, in order to assess whether forecast sensitivity experiments could distinguish between the two cases.

Two methods of estimating forecast sensitivity from a backward integration of forecast differences over a region of interest (the US in our case), using either the ADJM or QILM. The basic conclusions of the study are the following.

- Both ADJM and QILM should be a useful tool for adaptive observations in order to determine the areas where the data are most needed.

Using the energy norm, the adjoint sensitivity is much more baroclinic (tilted) than either the quasi-inverse or the actual analysis differences. The energy-norm adjoint sensitivity has relatively very small wind or vorticity differences, and is far from quasi-geostrophic balance. The quasi-inverse sensitivity tends to be larger amplitude and more in balance.

- The adjoint sensitivity determines areas that will result in maximum impact over the area of interest. The quasi-inverse sensitivity answers more accurately the question: given a forecast difference (such as the large spread of ensemble forecast members over the area of interest), where did the difference come from? As a result the two methods are complementary, and both should be used together.

- The methodology used was generally successful: Taking forecast differences and applying a local mask over the region of interest to start the backward integrations worked well, without negative effects due to spatial truncation. Both methods identified the region in the North Pacific where the sondes were dropped in the high impact case from the 48-h forecast differences. They both indicated that for the 72-h forecast it would have been even more effective to launch the sondes further west over the Pacific Ocean.

- Although by 72-h the assumption of linearity breaks down, the sensitivity results covered a larger area but otherwise continued to be useful for this forecast length. It was found that the size of the local region and the basic flow used for the linear and adjoint integrations had some impact but did not dominate the signal.

- In the case of low impact from the dropwindsondes, the ADJM successfully indicated low sensitivity in the region of the drops, whereas the QILM was able to pinpoint better where the actually observed small forecast differences came from. Both methods pointed out other areas (such as further south from the region of the drops) which in this case would have been a more useful location for additional observations.

- In the sensitivity calculation, the accuracy of the basic flow has an impact on the results. The more accurate basic flow tends to lead to more accurate sensitivity information.

- In real time applications, forecasts impacts are not available, and alternative verification time perturbations must be chosen. Ensemble forecast

differences, especially among ensemble members showing a larger forecast difference would be a good choice.

- Experiments using different initial perturbations such as using only the vorticity information, indicate that the geographical location from the sensitivity calculations is still good. With the temperature information only the QILM gives reasonable results.

Many of the techniques developed during this study were already adopted during the FASTEX experiment. We will report the final FASTEX evaluation results in the near future.

7. Acknowledgments

We would like to express our thanks to Majors Jon Talbot and Bob Katz and flight crews of the

AFRES 53rd WRS in carrying out this experiment, to M. Shapiro, B. Colman and the staffs of the Seattle Forecast Office for their assistance, and to the NCEP Hydrometeorological Prediction Center in providing suggest for the experiment. The contributions of G. Frederick and Radian Corporation in providing dropsonde and coordinating the experiment were invaluable. We also would like to thank two anonymous reviewers for their valuable comments on the manuscript. The first author is supported by UCAR Visiting Scientist Program.

REFERENCES

- Abramowitz, M. and Stegun, I. A. 1965. *Handbook of mathematical function*. Dover Publications, INC., New York, pp. 1046.
- Andersson, E., Courtier, P., Gaffard, C., Haseler, J., Rabier, F., Uden, P. and Vasiljevic, D. 1996. 3D-Var—the new operational analyses scheme. *ECMWF Newsletter* **71**, 2–5.
- Buizza, R. 1994. Sensitivity of optimal unstable structures. *Q. J. R. Meteorol. Soc.* **120**, 429–451.
- Buizza, R. and Palmer, T. N. 1995. The singular-vector structure of the atmospheric global circulation. *J. of Atmos. Sci.* **52**, 1434–1456.
- Courtier, P., Thepaut, J.-N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4-D VAR using an incremental approach. *Q. J. R. Meteorol. Soc.* **122**, 1367–1387.
- Chou, J. 1986. *Long-range numerical weather prediction*. Meteorological Press, Beijing, 353 pp (in Chinese).
- Cohn, S. E., N. S. Sivakumaran and R. Todling, 1994. A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.* **122**, 2838–2867.
- Derber, J. C. 1987. Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.* **115**, 998–1008.
- Derber, J. C. and W.-S. Wu, 1996. The use of cloud-cleared radiances in the NCEP's SSI analysis system. *11th Conf. on numerical weather prediction*. Norfolk, VA, American Meteorological Society, pp. 236–237.
- Haltiner, G. J. and Williams, R. T. 1980. *Numerical prediction and dynamical meteorology*. Wiley, Now York, pp. 477.
- Kalnay, E. and Toth, Z. 1994. Removing growing errors in the analysis. *10th Conf. on numerical weather prediction*. Portland, OR, American Meteorological Society, pp. 212–215.
- Kalnay, E., Toth, Z., Pu, Z.-X. and Lord, S. 1996. Targeting weather observation to locations where they are most needed. *CAS/JSC working group on numerical experimentation*. WMO/ICSU/IOC world climate research programme. World Meteorological Organization (WMO), Report No. 23, pp. 1.20–1.21.
- Kalnay, E. and Pu, Z.-X. 1998. Application of the quasi-inverse method to accelerate 4-D VAR. *12th Conf. on numerical weather prediction*. Phoenix, AZ, American Meteorological Society, pp. 41–42.
- Langland, R. H. and Rohaly, G. D. 1996. Analysis error and adjoint sensitivity in prediction of a north Atlantic frontal cyclone. *11th Conf. on numerical weather prediction*. Norfolk, VA, American Meteorological Society, pp. 150–152.
- Lord, S. 1996. The impact on synoptic-scale forecasts over the united states of dropwindsonde observations taken in the northeast pacific. *11th Conf. on numerical weather prediction*. Norfolk, VA, American Meteorological Society, pp. 70–71.
- Palmer, T. N., Gelaro, R., Barkmeijer, J. and Buizza, R. 1998. Singular vector, metrics and adaptive observations. *J. of Atmos. Sci.* **55**, 633–653.
- Parrish, D. F. and Derber, J. C. 1992. The National Meteorological Center's global spectral statistical interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763.
- Pu, Z.-X., Kalnay, E., Derber, J. C. and Sela, J. 1997a. Using forecast sensitivity patterns to improve the future forecast skill. *Q. J. R. Meteorol. Soc.* **123**, 1035–1053.

- Pu, Z.-X., Kalnay, E., Sela, J. and Szunyogh, I. 1997b. Sensitivity of forecast error to initial conditions with a quasi-inverse linear method. *Mon. Wea. Rev.* **125**, 2479–2503.
- Pu, Z.-X., Kalnay, E., Parrish, D., Wu, W. and Toth, Z. 1997c. The use of bred vectors in NCEP 3-D variational analysis system, *Weather and Forecasting* **12**, 689–695.
- Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. 1996. Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.* **122**, 121–150.
- Reynolds, C., Webster, P. and Kalnay, E. 1994. Random error growth in NMC's global forecasts. *Mon. Wea. Rev.* **122**, 1281–1305.
- Szunyogh, I., Kalnay, E. and Toth, Z. 1997. A comparison of Lyapunov vectors and optimal vectors in a low resolution GCM. *Tellus* **49A**, 200–227.
- Toddling, R., Cohn, S. E. and Sivakumaran, N. S. 1997. Suboptimal schemes for retrospective data assimilation based on the fixed-lag Kalman smoother. DAO Office Note 97-04. Available at Data Assimilation Office, Goddard Space Flight Center, Greenbelt, Maryland.
- Wu, W.-S. and Joo, S.-W. 1996. The change of the observational errors in the NCEP SSI analysis, *11th Conf. on numerical weather prediction*. Norfolk, VA, American Meteorological Society, pp. 84–85.
- Zupanski, M. and Zupanski, D. 1996. A quasi-operational application of a regional 4-dimensional variational data assimilation. *11th Conf. on numerical weather prediction*. Norfolk, VA, American Meteorological Society, pp. 94–95.