# Realism of Sensitivity Patterns

**Lars Isaksen**

*ECMWF, Shinfield Park, Reading, UK*

## 1. Abstract

The "Key Analysis Errors" based on minimizing the two day forecast error for December 2001 and January 2002 have been calculated using three different inner product norms. Ten-day forecasts have been run to verify that the medium-range forecasts from "Key Analysis Errors" corrected analyses are better than the control forecasts for all three norms. The goal has been to investigate if "Key Analysis Errors" really are analysis errors. For the first time a systematic comparison of "Key Analysis Sensitivity Patterns" with observations has been done. During the first 24 hours the sensitivity perturbed forecasts and control forecasts are compared against observations. It is shown that forecasts starting from the "Key Analysis Errors" perturbed analyses during the first approximately 12 hours are further away from the observations than forecasts from control analyses. From 12 hours onwards the sensitivity-perturbed forecasts are closer to observations, as expected because the method is constructed to minimize 48 hour forecast errors.

## 2. Sensitivity of forecast errors to initial condition

Rabier *et al.* (1996) investigated the sensitivity of 48 hour forecast errors to the initial conditions using a tangent linear and adjoint technique similar to the one applied later in ECMWF's operational 4D-Var assimilation system (Rabier et al. 2000). The aim was to minimize a diagnostic function J that measures the difference between the 48 hour forecast and the verifying analysis (available *a posteriori,* two days after the assimilation took place).

The diagnostic function to be minimised is

$J = 1/2\langle \mathbf{P}(\mathbf{x}_T - \mathbf{x}_T^{ver.ana}), \mathbf{P}(\mathbf{x}_T - \mathbf{x}_T^{ver.ana})\rangle_W$, where T is 48 hours and $\langle \mathbf{x}, \mathbf{x}\rangle_W$ defines a final-time inner-product norm. For the often-used energy norm the definition is:

$$\langle \mathbf{x}, \mathbf{x}\rangle_W = 1/2 \int_{t_0}^{t_1} \int\int_D (u^2 + v^2 + R_d T_r (\ln P_s)^2 + (C_p / T_r) T^2) dD(\delta P_r / \delta\eta) d\eta$$

$\mathbf{x}_T$ is a 48 hour nonlinear model integration, i.e. $\mathbf{x}_T = \mathbf{M}\mathbf{x}_0$, where $\mathbf{M}$ represents the nonlinear model. $\mathbf{x}_T^{ver.ana}$ is the verifying analysis valid at time T. $\mathbf{P}$ represents a projection operator (=1 north of 30°N, otherwise zero).

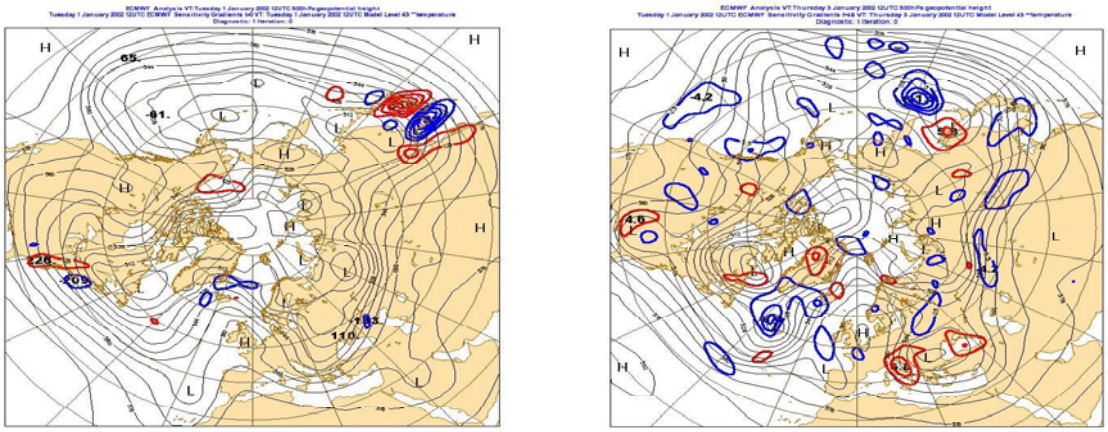
Rabier *et al.* (1996) showed that $\nabla J_T = \mathbf{P}(\mathbf{x}_T - \mathbf{x}_T^{ver.ana})$, and that under the assumptions that **the forecast model is perfect** and **the tangent-linear approximation is valid for 48 hours** this lead to $\nabla J_0 = \mathbf{R}^* \mathbf{P}(\mathbf{x}_T - \mathbf{x}_T^{ver.ana})$, where $\mathbf{R}^*$ is the adjoint of the tangent linear model, $\mathbf{R}$. So the method determines the sensitivity of forecast errors to initial condition under the specified assumptions.

The experiments performed by Rabier *et al.* (1996) showed that small sensitivity determined adjustments of the initial fields using this method could result in substantial improvement of the medium-range forecasts. They linked the sensitivity patterns to analysis errors, but it was outside the scope of their studies to prove the link.

They recommended objective verification against observations to diagnose cases of analysis shortcomings. One of the aims of this paper is to investigate that issue.

Figure 1(b) shows an example of the 48 hour temperature forecast error at model level 43 (650 hPa) overlaid the 500 hPa geopotential height field at 1200 UTC 3 January 2002. Note the large forecast errors (up to 13 °K) east of Japan. Fig 1(a) shows the energy norm based sensitivity gradient two days earlier at 1200 UTC 1 January 2002. It is seen that the large forecast errors east of Japan can be corrected by modifying the temperature field over Japan. Two days earlier, an error in the middle of the North Atlantic can be corrected by modifying the analysed temperature field near the east coast of U.S.A.

**a)** $\nabla J_0 = \mathbf{R}^* \mathbf{P}(\mathbf{x}_t^{fc} - \mathbf{x}_t^{ver.ana})$ **b)** $\nabla J_t = \mathbf{P}(\mathbf{x}_t^{fc} - \mathbf{x}_t^{ver.ana})$



**12 UTC 1 January 20021** **2 UTC 3 January 2002**

*Figure 1: Panel a shows the gradient of the objective cost function at initial time,* $\nabla J_o = \mathrm{R}^* \mathrm{P}(\mathrm{x}_\mathrm{T} - \mathrm{x}_T^{ver.ana})$*, at 1200 UTC 1 January 2002 that corresponds to the gradient 48 hours later,* $\nabla J_T = \mathrm{P}(\mathrm{x}_\mathrm{T} - \mathrm{x}_T^{ver.ana})$ *, at 1200 UTC 3 January 2002 as shown in panel b. Red and blue 2.5°K contours show positive and negative temperature gradient values for model level 43 (650 hPa). The black contours show the 500 hPa geopotential height field. Note that the perturbations at final time, T, are only present north of 30°N.*

## 3. Estimation of "Key Analysis Errors" using the adjoint technique

The sensitivity method described above identifies the **direction** towards the objective functions minima at which a perturbation would improve the two-day forecast under the simplified constraint discussed above. The **size** of the perturbation (optimal step-size) was in that study determined on a trial and error basis, though linked to the squared value of the fastest growing singular vector over 48 hours. Klinker *et al.* (1998) extended the sensitivity method by developing a way to determine both the direction and the optimal step-size of the perturbation. The method will be described briefly here.

From the sensitivity method explained above we have $J = 1/2 \langle \mathbf{P}(\mathbf{x}_T - \mathbf{x}_T^{ver.ana}), \mathbf{P}(\mathbf{x}_T - \mathbf{x}_T^{ver.ana}) \rangle_W$.

This is equivalent to $J = 1/2 (\mathbf{M}\mathbf{x}_0 - \mathbf{x}_T^{ver.ana})^T \mathbf{A}(\mathbf{M}\mathbf{x}_0 - \mathbf{x}_T^{ver.ana})$, where $\mathbf{A}$ is the matrix defining the inner product including the projection $\mathbf{P}$ on the area north of 30°N and $\mathbf{M}$ is the non-linear model.

A first order approximation of cost function change ($\delta J$) as a function of an increment ($\delta \mathbf{x}_0$) added to the analysis $\mathbf{x}_0$ can be found (see Klinker *et al.* (1998) for further details).

$\delta J$ is maximized under the $\langle \delta x_0, \delta x_0 \rangle$ constraint N is a constant.

This gives $\delta \mathbf{x}_0 = 1/(2\lambda)\nabla J_C$, with $\lambda^2 = 1/(4N)\nabla J_C^T C \nabla J_C$, where C defined the initial time norm.

So this determines the amplitude of the optimal perturbation under the idealized assumptions used.

The process is then iterated a number of times, each time from the improved estimate of the previous minimization, in many aspects similar to 4D-Var minimization. Klinker *et al.* (1998) tried to determine the optimal number of iterations. They determined the optimal perturbation in the sense that it was the perturbation that fitted observations best and thus was describing the analysis errors. They stated that this "optimal" perturbation estimated the "Key Analysis Errors". In their case study only three energy norm based iterations moved the analysis significantly closer to observations than the control analysis. In Klinker *et al.* (1998) it is discussed why more iterations not necessarily provide a better description of analysis errors. Main reasons are too simple description of forecast errors and simplifications in the tangent-linear model, especially in the physical processes. Klinker *et al.* (1998) show that increasing the number of iterations will amplify noise in the initial perturbation due to contracting directions of the tangent-linear operator. It is also shown that enstrophy norm based perturbations minimize the two day forecast error but have a very different structure than energy norm based perturbations. The enstrophy norm based perturbations fits the observations significantly worse than the energy norm based patterns. In this paper we have further investigated this dependency of the chosen norm on the structure of sensitivity patterns. It is seen from the equations above that the initial norm (C metric) plays an important role in both the determination of the patterns and the optimal step-size. Based on these limitations it is questionable if the "Key Analysis Errors" really estimate the part of the analysis error that is largely responsible for the short-range forecast errors.

Figure 2a shows the temperature perturbation after one iteration, based on the above method, for the case study presented in Figure 1. It is seen that Figure 1a and Figure 2a have the same patterns. Figure 1a just gives the gradient whereas Figure 2a gives the perturbation size based on the optimal step-size calculated using the method described above. Figure 2b shows the perturbations after three iterations, the so-called "Key Analysis Errors". The additional iterations maintain the overall structure introduced by the first iteration but also adds finer scale structures and new larger scale perturbations, i.e. over South-East Europe. The large (up to 13°K) errors east of Japan are corrected by modifying the temperature field over Japan by 1.5°K, i.e. a ten times error growth in 48 hours.

**Temperature perturbation after 1 iteration**

**Temperature perturbation after 3 iterations = Key analysis errors**

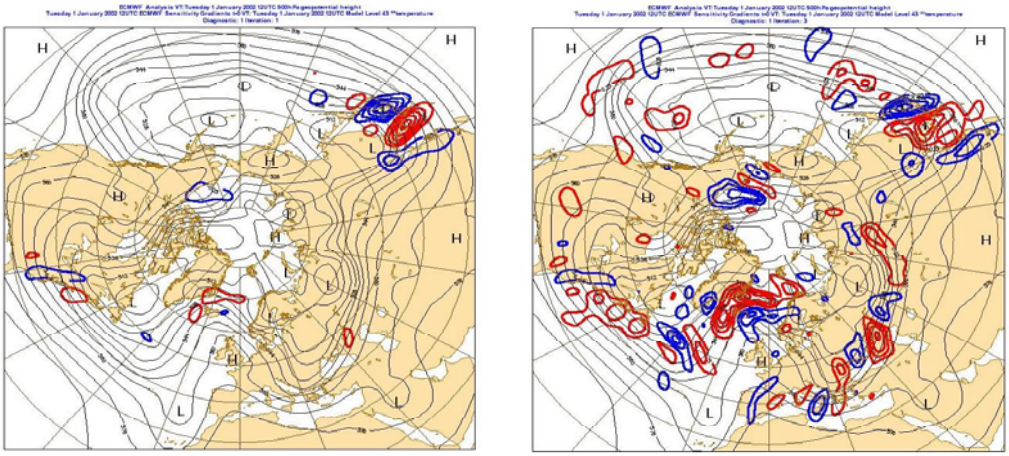


*Figure 2: Panel a shows the temperature perturbation on 1200 UTC 3 January 2002 after one iteration based on Klinker et al. (1998) that corresponds to the gradient 48 hours later, as shown in Figure 1 panel b. Red and blue 0.25°K contours show positive and negative temperature gradient values for model level 43 (650 hPa). The black contours show the 500 hPa geopotential height field. Panel b shows the temperature perturbations (using 0.25°K contours like in panel a) after three iterations, the so-called "Key Analysis Errors".*

## 4. Calculation of "Key Analysis Errors" for various inner product norms

In this paper, we present a systematic comparison of "Key Analysis Errors" with observations. We argue that, unless background error variance is systematically underestimated in the ECMWF system (in which case a simple reduction of background error should benefit the analysis), an analysis modified using "Key Analysis Errors" should fit the observations better than the unperturbed analysis.

"Key Analysis Errors" for December 2001 and January 2002 were calculated using three different norms at initial time. For each choice of norm ten day forecasts were run to verify that the medium-range forecasts from analyses modified using "Key Analysis Errors" resulted in better forecasts than the control forecasts. During the first 24 hours the modified forecasts and the control forecasts were compared with observations. We show below that forecasts starting from the modified analyses during the first approximately 12 hours are further away from the observations than forecasts from control analyses. From 12 hours onwards the forecasts from the perturbed analyses are closer to observations. The latter is expected because the method is using additional future information, i.e. the day two forecast error, as an essential part of the minimization process.

As discussed in the previous section, "Key Analysis Errors" are calculated by applying a few steps of minimization to an objective cost function that measures the difference between a two-day forecast and a verifying analysis. This requires routines to calculate inner products with respect to the initial-time norm, and routines to multiply vectors by the symmetric square root of the matrix that defines the norm, and by the symmetric square-root of its inverse. In the case of the Hessian norm, this necessitates the use of an approximate Hessian. We use the same approximation as is used to precondition the minimization. That is, the Hessian is replaced by:

$$\tilde{\mathbf{H}} = \mathbf{B}^{-1/2}\tilde{\mathbf{Y}}\mathbf{B}^{-1/2} \tag{1}$$

where:

$$\tilde{\mathbf{Y}} = \mathbf{I} + \sum_{k=1}^{K}(\lambda_k - 1)\mathbf{v}_k\mathbf{v}_k^T . \tag{2}$$

Here, $\lambda_k$ and $\mathbf{v}_k$ are the leading eigenvalues and eigenvectors of the Hessian in the space of the control vector of the 4D-Var minimization. The approximate Hessian is easy to invert and to square root. In this study, we use $K$=100. A more detailed description of the approximate Hessian can be found in Barkmeijer et al. (1998) and Barkmeijer et al. (1999).

"Key Analysis Errors" have also been calculated using the simpler $J_b$-norm. It is based on the following definition of $\tilde{\mathbf{H}}$:

$$\tilde{\mathbf{H}} = \mathbf{B}^{-1} \tag{3}$$

This $J_b$ formulation has no influence from observations. Similarly, the effect of truncating the sum in equation (2) is to make the "Key Analysis Errors" calculated using either the approximate Hessian norm or the $J_b$ norm more similar. It has been noted that they become almost identical in data sparse regions. This is because the leading eigenvectors of the Hessian tend to be concentrated in regions of dense observations.

## 5. Design of assimilation experiments and method for validation against observations

ECMWF's 4D-Var assimilation system (Rabier *et al*. 2000) version 25r1 operational April 2002 to January 2003 was used for the experimentation. The incremental version (Courtier *et al*. 1994) was used with both

outer-loop and inner-loop calculations performed at $T_L159L60$. Forecast error calculations were performed at T42L60. The setup is similar to ECMWF's operational assimilation system in 2002, except that in the operational system the outer-loop calculations are performed at $T_L511L60$.

4D-Var assimilation experiments were performed for the period 1-31 December 2001 and 1-31 January 2002. For both periods, daily runs were performed to calculate "Key Analysis Errors" from the 48-hour forecast error using the method described in Klinker *et al*. (1998). The calculations were performed using three different metrics: the energy norm, the $J_b$ norm and the approximate Hessian norm. The energy norm calculations used 1200 UTC as the starting point, whereas the $J_b$ and approximate Hessian norm calculations used 0300 UTC as the starting point, (the latter because the approximate Hessian norm calculations were influenced by observations during the 4D-Var assimilation window, from 0300 UTC to 1500 UTC). The structure and realism of the sensitivity patterns can be compared without problems despite the difference in starting point because additional control forecasts were performed from 0300 UTC.

Forecast verification scores (Figure 3) confirm that modifying the initial conditions using energy norm "Key Analysis Errors" improves the 2-7 day forecasts. Forecasts are improved to an approximately similar degree for perturbations calculated using each of the three initial norms (not shown), even though the initial perturbations are very different. The forecast improvement is sustained beyond 2 days because the error patterns that have grown most during the first two days have been substantially reduced and delays the errors growth on subsequent forecast days. It should be kept in mind that this delayed error growth could be obtained in an infinite number of ways because the perturbations are determined by the inner product norm used. In this paper it is investigated if the sensitivity pattern calculated from three different meteorologically well justifiable inner product norms represent analysis errors.

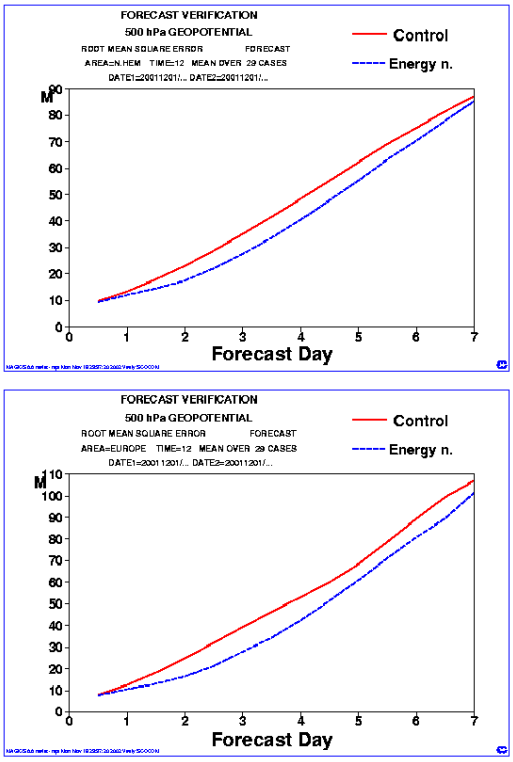



*Figure 3: Panel a: Forecast verification root-mean-square errors for Northern Hemisphere extra-tropics 500 hPa geopotential height during 29 days in December 2001. Red curves show control forecast scores from standard analysis. Blue curves show forecasts from analysis modified by energy norm based "Key Analysis error" perturbations. Panel b: like panel a, but for Europe.*

A methodology similar to the one used in Klinker *et al*. (1998) was applied to investigate the realism of the various sensitivity patterns. As in Klinker *et al*. (1998) we judge the quality of the analysis by its ability to fit observations. In addition, short-range (24 hour) forecasts were performed from standard control analyses and from analyses modified by the three types of "Key Analysis Error" perturbations. These forecasts included comparison against observations at the correct time and location, using the same code that is used to calculate observation-minus-model departures in 4D-Var.

## 6. The impact of initial-time norm on the general structure of sensitivity patterns

We have investigated how the general structure of the "Key Analysis Errors" depends on the chosen norm. Figure 4, panels (b), (c) and (d) show the January 2002 average "Key Analysis Errors" for temperature at model level 42 (in the 60-level configuration, approximately 620 hPa) for the experiments described in the previous section. Figure 4(a) shows the RMS of the Eady index for January 2002 calculated for the 300 hPa to 850 hPa layer using the Hoskins and Valdez (1990) definition. It is seen that the energy-norm sensitivities contain more small-scale features and shorter frequency wave train patterns than the $J_b$ and approximate Hessian norm sensitivities. The average $J_b$ and approximate Hessian norm sensitivities show a very similar pattern, the main difference being larger amplitudes for the approximate Hessian norm sensitivities, and a tendency to avoid data-rich land areas.

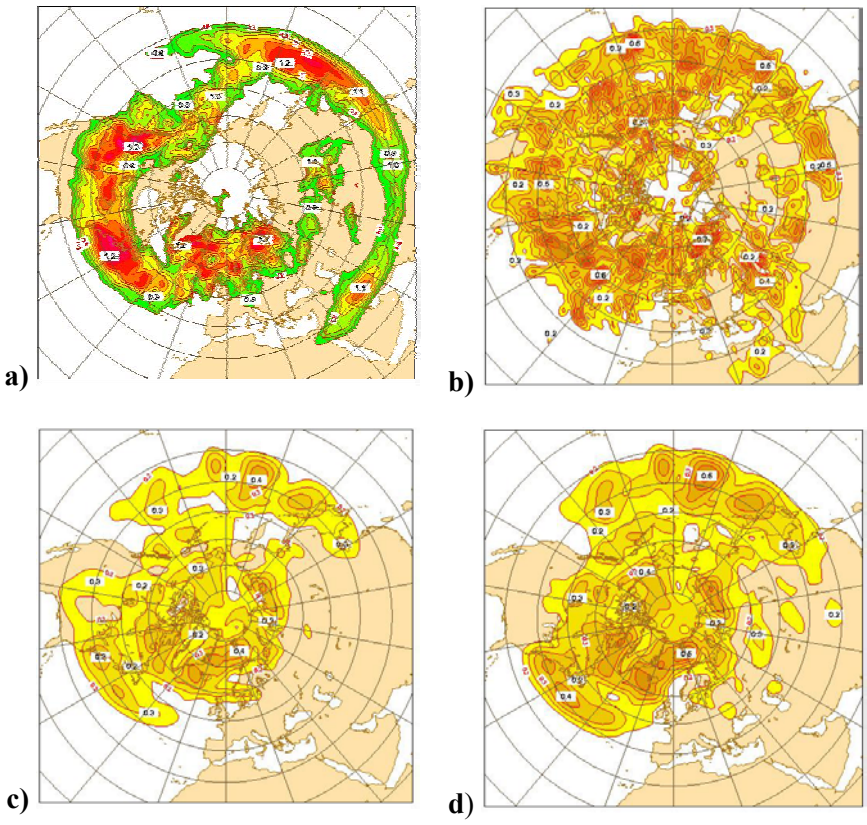



*Figure 4: Root-Mean-Square values based on 1200 UTC fields from 29 days in January 2002. Panel a: Eady index calculated from 300 hPa and 850 hPa fields. Panel b: "Key Analysis Errors" calculated using the total energy nom. Panel c: "Key Analysis Errors" for the $J_b$ norm. Panel d: "Key Analysis Errors" for the approximate Hessian norm.*

The energy norm sensitivities are clearly located in the areas with high Eady index, where the troposphere is most baroclinic. This feature has earlier been discussed by Buizza and Palmer (1995) and Buizza *et al.* (1993). The $J_b$ and the approximate Hessian norm sensitivities typically have largest amplitude in the regions

just south or north of the most unstable regions. They seem to adjust the baroclinic instability by modifying the temperature of the air masses north and south of the frontal regions.

As expected, the vertical and horizontal structure of the $J_b$ norm perturbations and approximate Hessian norm perturbations are linked to the form of the structure functions. Likewise there are strong links between the structure of energy norm perturbations and the singular vectors (see Gelaro *et al*. 1998). It is also clear from Figure 5 that the power spectra of the temperature and vorticity fields for "Key Analysis Errors" at model level 42. For reference, the spectra of the temperature and vorticity analysis fields are also shown. The analysis fields are at $T_L159$ resolution and the perturbations are calculated at T42 resolution. For both temperature and vorticity, energy norm perturbations have largest amplitude at smaller scales than for perturbations calculated using the $J_b$ norm and the approximate Hessian norm. The $J_b$ norm and approximate Hessian norm perturbations peak at total wave number 5-10 for temperature, and at total wave number 15-30 for vorticity. This is in agreement with the power spectra for the total fields. From the vertical profiles shown in Figure 5c-d it is seen that the $J_b$ norm and approximate Hessian norm perturbations have larger vorticity perturbations and smaller temperature perturbations than the energy norm perturbations. The relative size of temperature versus vorticity amplitudes for the energy norm is determined by the scaling constants in the definition of the norm.

The approximate Hessian norm perturbations do not differ much from the $J_b$ based norm perturbations. This may be because significantly more than 100 eigenvectors may be required to account for the smaller-scale influence of observations in the analysis.

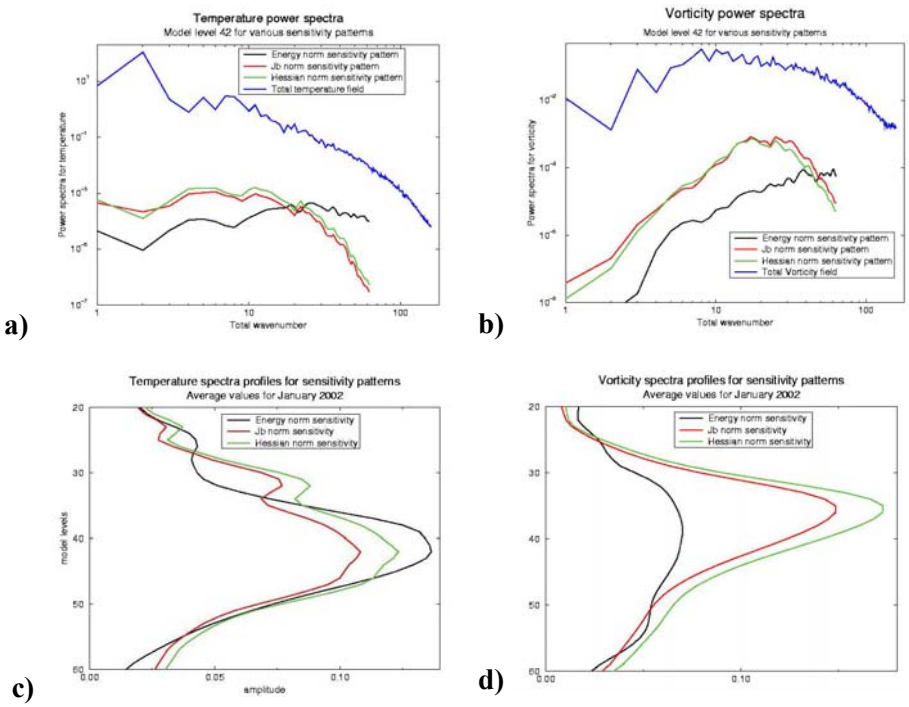



*Figure 5: Power spectra for temperature (K) (panel a) and vorticity ($s^{-1}$) (panel b) at model level 42 (near 620 hPa) valid at 12 UTC 1 January 2002. Black curves: energy-norm key analysis error. Red curves: $J_b$-norm key analysis error. Green curves: approximate Hessian-norm key analysis error. Also shown are the power spectra for the analysed temperature and vorticity fields. Panel c shows the amplitude of the temperature component of the "Key Analysis Errors" for model levels 20 to 60 (from around 35 hPa to the surface). The colour coding of curves is the same as in panel a. Panel d: Similar to panel (c) but for the vorticity component ($s^{-1}$), scaled by a factor 100000.*

Figure 6 shows the vertical and horizontal structure of energy norm, $J_b$ norm and approximate Hessian norm perturbation patterns on a specific day (1200 UTC 1 January 2002) for a region near Japan. The results for this individual case and other cases not shown reflect the average statistical results described above. "Key Analysis Errors" calculated using the $J_b$ and approximate Hessian norms have similar structure, but have larger amplitude for the $J_b$ norm than for the approximate Hessian norm. "Key Analysis Errors" calculated using the energy norm have much finer scale, and show wave-train features, both vertically and horizontally. All three flavours of key analysis error have vertical tilts, most pronounced for the energy norm sensitivities. It is evident (especially from the horizontal patterns) that the energy norm perturbations may locally have the opposite sign to the $J_b$ norm or approximate Hessian norm perturbations (e.g. over Eastern Japan). Figure 7 shows a vertical cross section of the potential temperature analysis field along the same east-west axis as used in figure 6. It is seen that the main perturbations and vertical tilt does take place in the baroclinic region near the western boarder of Figures 6 and 7.

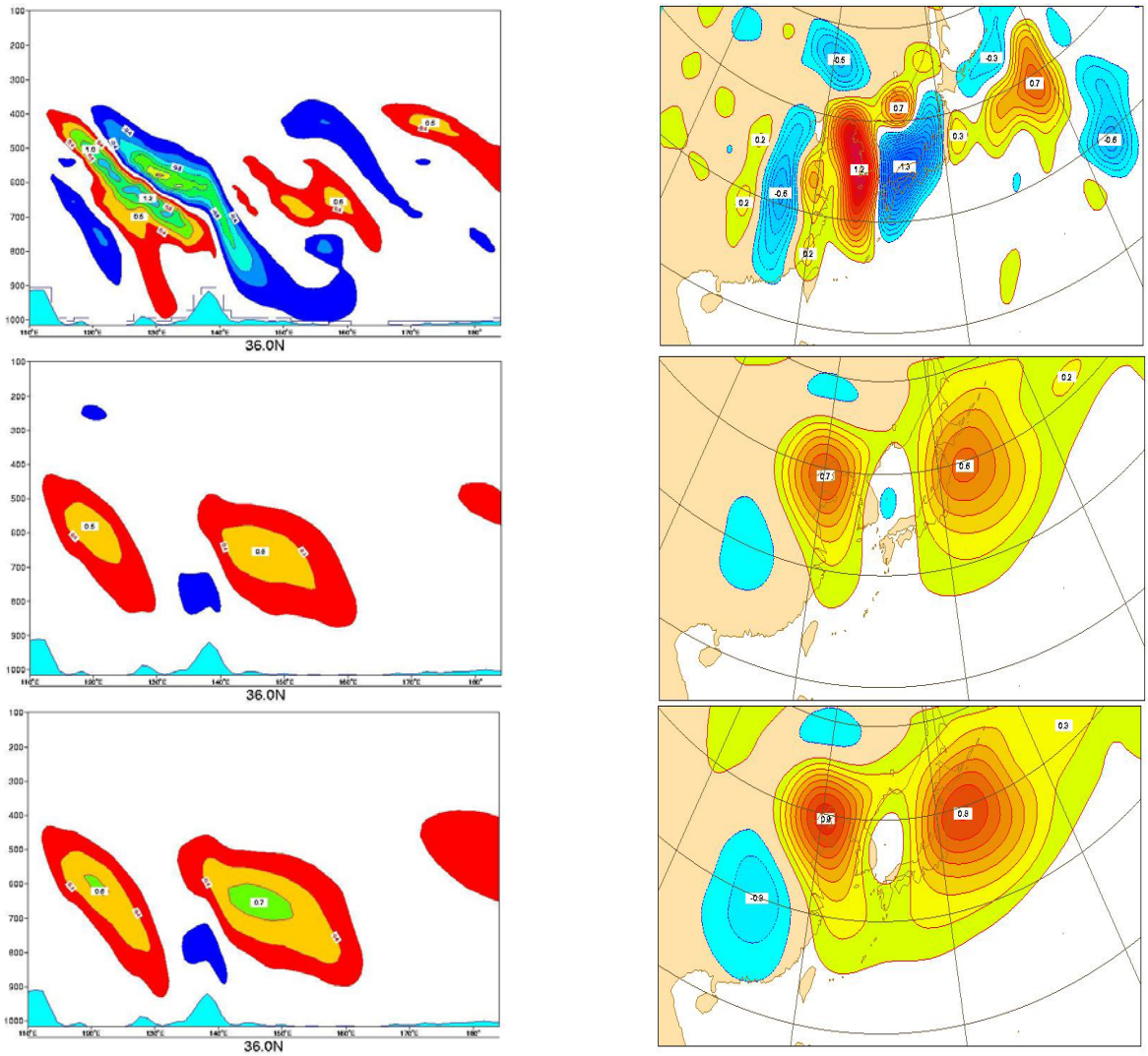



*Figure 6: The left column shows cross-sections of temperature for "Key Analysis Errors" near Japan along latitude 36N valid 1200 UTC 1 January 2002. The right column shows the corresponding temperature perturbations at model level 42 (620 hPa). The top row is for energy norm, the middle row for the $J_b$ norm and the bottom row for the approximate Hessian norm.*

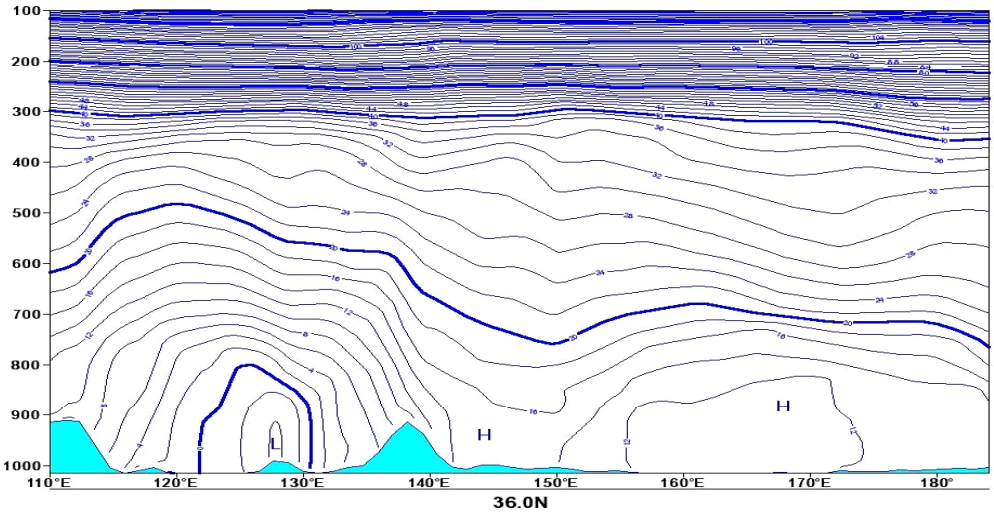



*Figure 7: Potential temperature east-west cross section along the axis used in figure 6*

Figure 8(top row) shows an example of temperature and vorticity perturbations at model level 43 (650 hPa) from the energy norm based "Key Analysis Error" calculations at 1200 UTC on a specific day in January 2002. Figure 8(bottom row) shows the Hessian norm based perturbations. Note the large difference in the structure and sign of the perturbations. This is a typical example from the two-month assimilation experiment. It is another indication that the size, structure and sign of the initial perturbations are determined by the inner product norm used. Another fairly typical example from another day in January 2002 is presented in Figure 9 that shows how the amplitude of temperature and vorticity perturbations can be very different for respectively the energy based norm and the Hessian based norm.

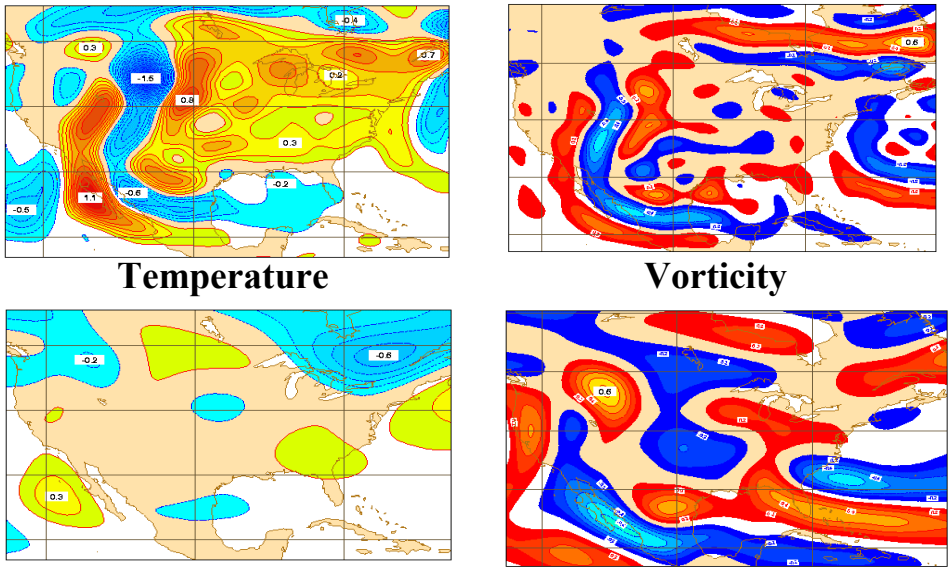



*Figure 8: Top row: Temperature and vorticity perturbations at model level 43 (650 hPa) from the energy norm based "Key Analysis Error" calculations at 1200 UTC on a day in January 2002. Bottom row: Like top row, but for Hessian norm based perturbations. Note the large difference in the structure and sign of the perturbations.*

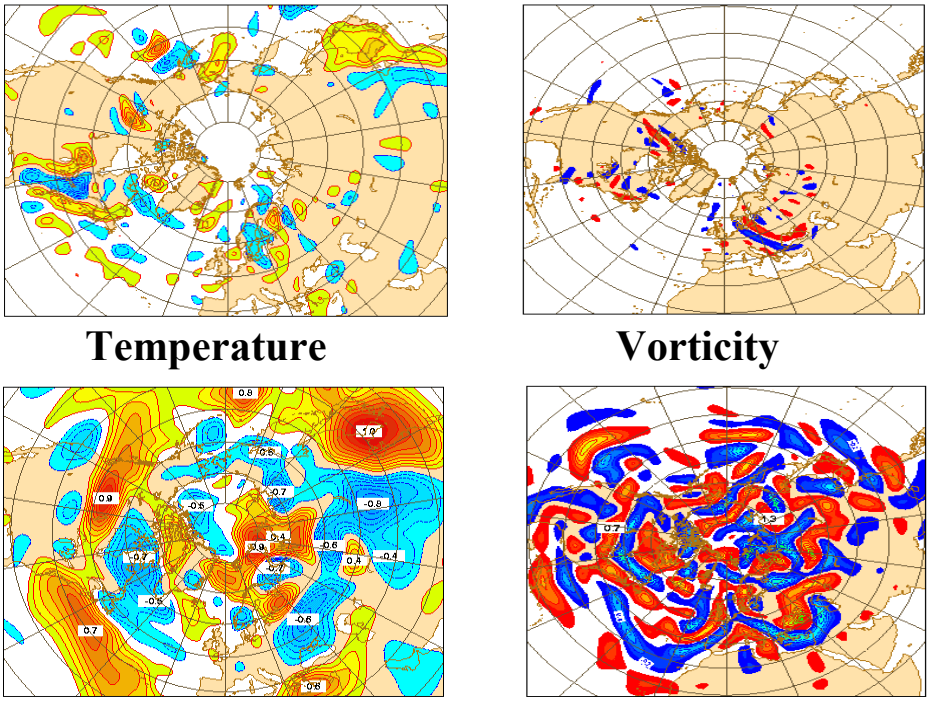



*Figure 9: Like Figure 8, but for another day in January 2002. Note the large difference in amplitude for the energy norm and Hessian norm based patterns.*

## 7. Comparison of "Key Analysis Errors" with observations

Short-range forecasts from a better analysis can be expected to show smaller departures from reliable quality observations. We have compared the quality of forecasts from initial conditions modified using the three types of key analysis error described above. Forecasts were run for 24 hours using the ECMWF assimilation system in screening mode. This allowed the forecasts to be compared with observations using the standard 4D-Var observation operators. The proper space interpolation is used and forecast fields are used at proper time, avoiding interpolation of model fields in the time domain. Forecasts were run every day for the two-month assimilation experiments, making it possible to make statistical evaluation based on large numbers of observations, covering many different synoptic situations. Figures 10-13 show some of the results from this evaluation. Figure 10 shows observation-minus-forecast statistics for three observing systems: Meridional wind component for American profilers (top row), DRIBU wind speed (middle row) and QuikSCAT wind speed (bottom row). All these high quality observing systems measure continuously or at hourly intervals, making it possible to evaluate the forecast quality every hour during the first 24 hours of the forecasts. The red bars show statistics for the forecasts run from the normal analyses, whereas the black curves represent the forecasts run from the perturbed analyses. The perturbed analyses are designed to improve the day-two forecast (and have been proved to do so), so it is not surprising that towards the end of the 24 hour forecast window the perturbed forecasts are better in agreement with observations. But, during the first 12 hours or so, the sensitivity forecasts from all three types of sensitivity-perturbed analyses are further away from observations than the control analyses. For wind observations this is most pronounced for the approximate Hessian norm perturbation forecasts, as is to be expected from Figure 5(d). Figure 11 shows similar statistics for three additional observing systems: SYNOP surface pressure (top row), SSM/I wind speed (middle row) and DRIBU surface pressure (bottom row). Figure 12 shows a more detailed view of the observation-minus-forecast RMSE statistics for the zonal wind component of American profilers. The yellow bars show statistics for forecasts performed from initial conditions modified using energy norm based key analysis error perturbations. Green bars show statistics for forecasts performed from unperturbed control analyses. Red

bars show the RMSE difference between control forecasts (green) and perturbed forecasts (yellow) multiplied by 50. Figure 12b is similar to Figure 12a, but shows forecast errors ($\sigma_b = \sqrt{\text{"observation} - \text{forecast"}^2 - \sigma_o^2}$). A constant observation error, $\delta_0$, of 1.9 m/s has been subtracted in these calculations. This makes it possible to show the forecast error component separately which results in a more pronounced signal. Figure 13 is similar to Figure 12 but shows the Hessian norm statistics. Even though the differences between control forecasts and perturbed forecasts are small they are statistically significant

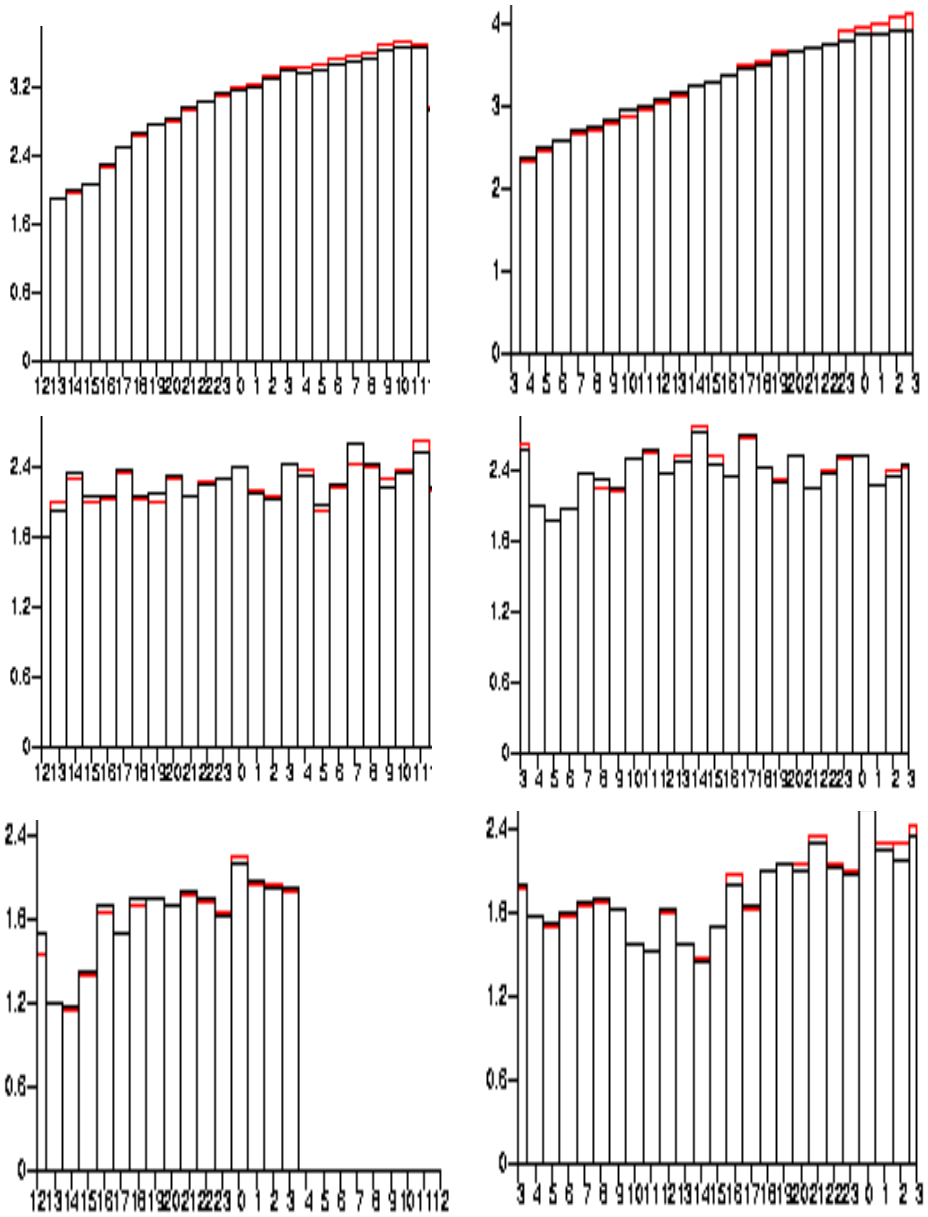


*Figure 10: Observation minus forecast statistics as function of forecast time in hourly time slots. Black bars show statistics for forecasts performed from initial conditions modified using key analysis error perturbations. Red bars show statistics for forecasts performed from unperturbed control analyses. Top row: Statistics for meridional wind component from American profilers for forecasts modified using energy-norm "Key Analysis Errors" (left) and approximate Hessian-norm "Key Analysis Errors" (right). Middle row: Like first row, but for DRIBU wind speed. Bottom row: Like first row, but for QuikSCAT wind speed.*

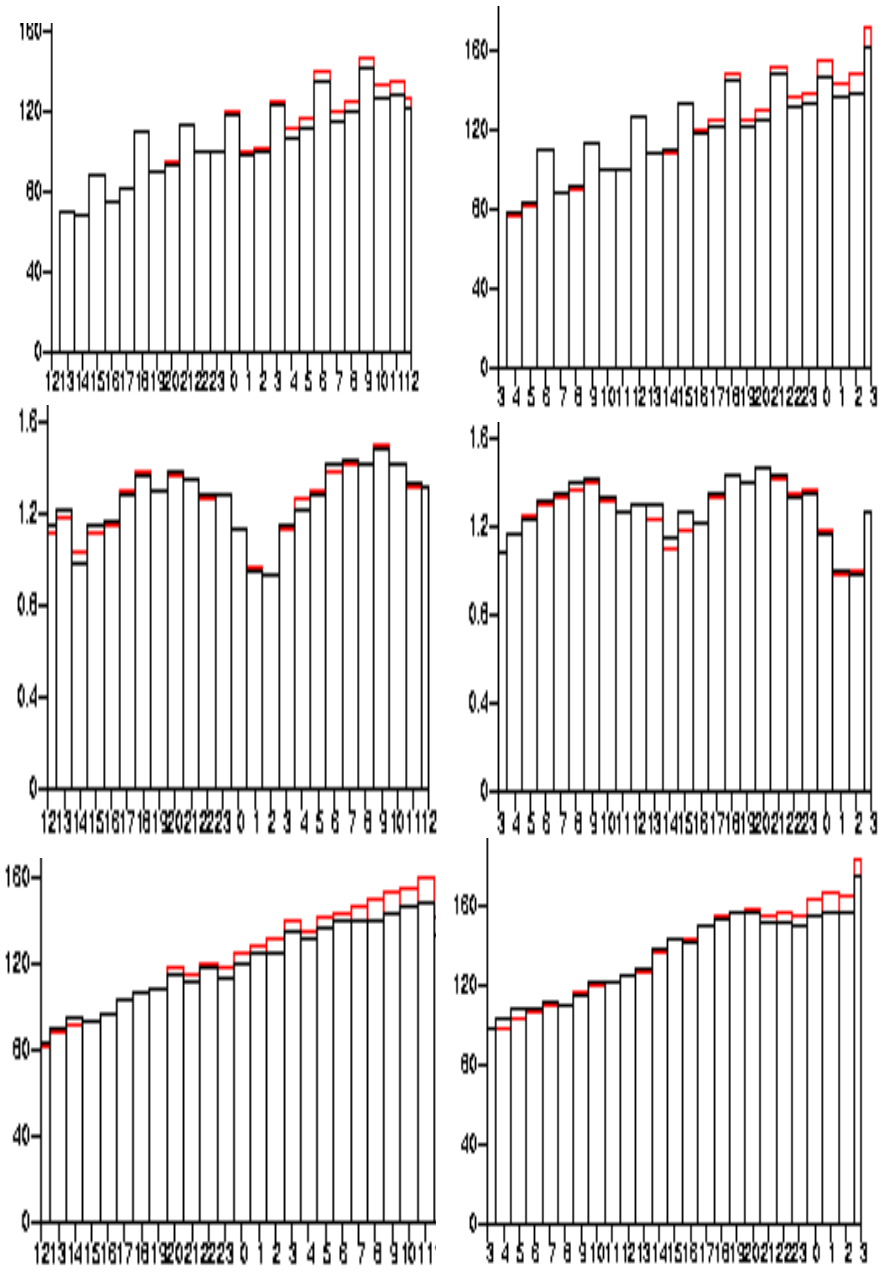


*Figure 11: Similar to figure 12. Top row: Statistics for SYNOP surface pressure. Middle row: Statistics for SSM/I wind speed. Bottom row: Statistics for DBIBU surface pressure.*

due the large sample size of approximately 47000 observations per time slot. The 95% significance level corresponds to 0.25 on the left-hand vertical scale for the red bars, which is achieved for almost the entire initial 12 forecast hours.

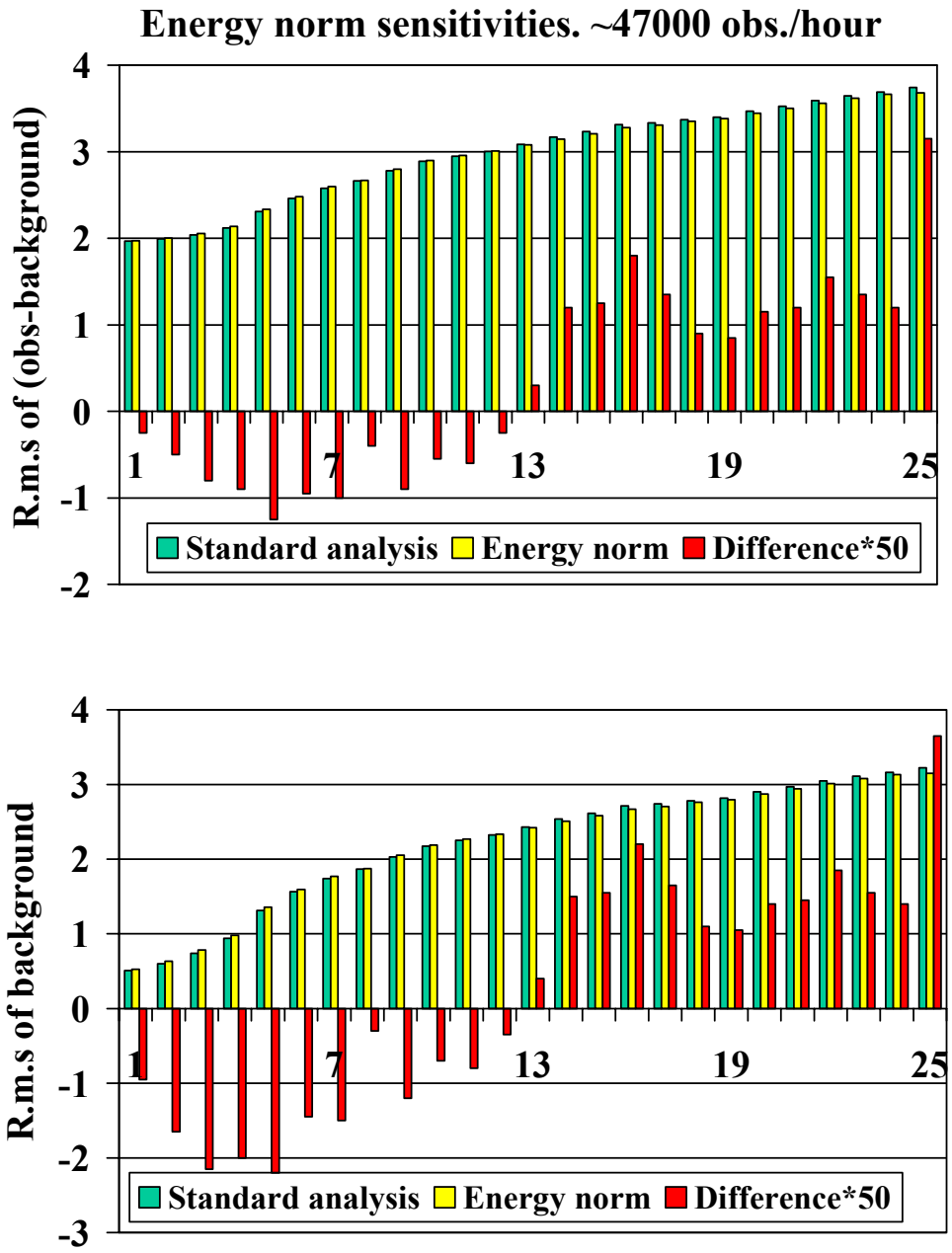



*Figure 12: Panel a: Observation minus forecast RMSE statistics as function of forecast time in hourly time slots for zonal wind component from American profilers. Yellow bars show statistics for forecasts performed from initial conditions modified using key analysis error perturbations. Green bars show statistics for forecasts performed from unperturbed control analyses. Red bars show the RMSE difference between control forecasts (green) and perturbed forecasts (yellow) multiplied by 50. Panel b is similar to Figure 12a, but shows forecast errors ($\sigma_b = \sqrt{"observation - forecast"^2 - \sigma_o^2}$). A constant observation error, $\sigma_o$, of 1.9 m/s has been subtracted in these calculations.*

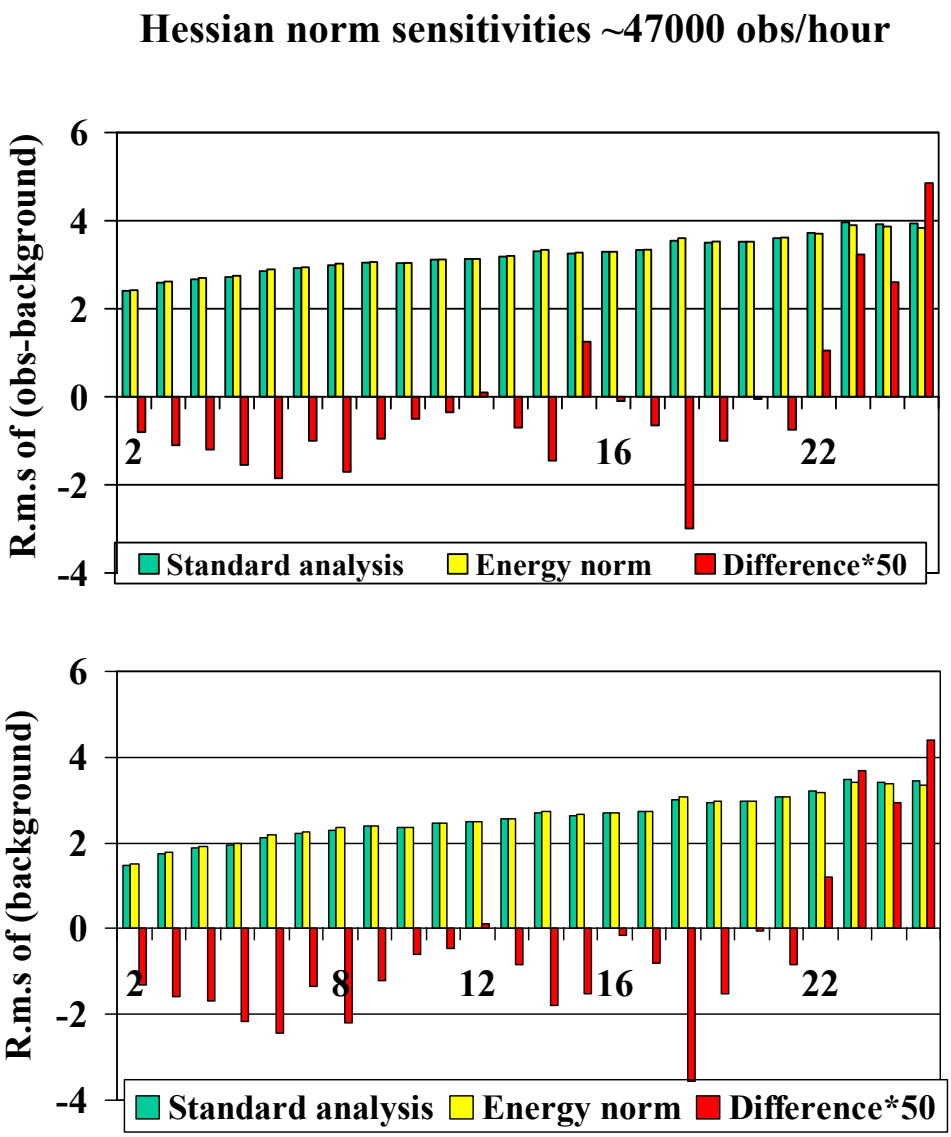



*Figure 13: Similar to Figure 12, but for approximate Hessian norm sensitivities.*

Seven observing systems have been presented here. We found that for almost all observing systems used in this study, the control analysis forecasts are more in agreement with observations during the first 12 hours. This can be taken as an indication that the "Key Analysis Error" perturbations that "improve" the control analyses do not represent observable analysis errors, but rather artificial adjustments to the initial state that, due to the formulation of the objective cost function, are bound to improve the 2-5 day forecast. The fact that very different perturbations (*viz*. perturbations generated using the energy norm and the approximate Hessian norm) are capable of producing similar improvements to the two-day forecasts is a clear indication that "Key Analysis" method is unable to identify errors in the analysis unequivocally.

There is a close link between the subspace within which the background error covariance matrix is allowed to be flow dependent in the Reduced Rank Kalman Filter method and the subspace explored by the approximate Hessian norm "Key Analysis Errors" (Fisher and Andersson 2002). The "Key Analysis Error" method applied here has the very favourable advantage that the day-two forecast error is known and used as part of the determination of the optimal perturbation sizes along with the various directions within the subspace. Even with this advantage, the perturbations do not seem to represent analysis errors well.

## 8. Discussion and conclusions

In this study it has been shown that the sensitivity patterns depend strongly on the norm used. It is seen that the energy norm sensitivities are smaller scale and often very different in structure than $J_b$ or Hessian norm sensitivities. The energy norms are more closely associated with baroclinic regions than seen for $J_b$ or Hessian norms. The $J_b$ and Hessian norms give rather similar sensitivity patterns. This is most likely because this study only used 100 eigenvectors to estimate the Hessian.

It is shown that the forecasts from all three types of sensitivity pattern modified analyses are often further away from observations during the first 12 hours than for the control forecasts. From approximately 12 forecast hours onwards the sensitivity forecasts are closer to the observations than the control forecast – as expected due to the 48-hour optimization time.

These results facilitate a better understanding of the poor Reduced Rank Kalman Filter performance: it seems to confirm other investigations (Fisher *et al.* (2003)) that a 100 dimensional subspace is not sufficient to describe analysis errors, at least when the search directions are determined by the approximate Hessian used in this paper. It is also questionable to use sensitivity patterns and "Key Analysis Errors" to decide where to launch targeted observations, restructuring of observing networks and estimating the benefit of new satellite instruments.

## 9. Acknowledgements

This work has only been possible due to help from Jan Barkmeijer (then ECMWF, now KNMI) in developing the software required to perform approximate Hessian based "Key Analysis Error" calculations. Discussions with Mike Fisher (ECMWF) about the relevance of this study in relation to evaluation of the Reduced Rank Kalman Filter performance have been greatly appreciated by the author.

## 10. References

Barkmeijer, J., R. Buizza. and T.N. Palmer, 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc*. **125,** 2333-2351.

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

Buizza R., J. Tribbia, F. Molteni and T.N. Palmer, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus*, **45A**, 388-407.

Buizza, R. and T.N. Palmer, 1995: The Singular-Vector Structure of the Atmospheric Global Circulation. Journal of the Atmospheric Sciences, **52**, No. 9, pp. 1434-1456

Courtier, P., J.-N. Thepaut and A. Hollingsworth, 1994: A strategy for operational implementation of {4D-Var}, using an incremental approach. *Q. J. R. Meteorol. Soc*. **120,** 1367-1388.

Fisher, M. and E. Andersson, 2002: Developments in 4D-Var and Kalman Filtering. *ECMWF Tech. Memo.*, **347,** pp. 36.

Gelaro, R., R. Buizza, Palmer, T.N. and E. Klinker, 1998: Sensitivity Analysis of Forecast Errors and the Construction of Optimal Perturbations Using Singular Vectors. J. Atmos. Sci., **55,** 1012-1037.

Hoskins, B.J. and P.J. Valdez, 1990: On the existence of storm tracks. J. Atmos. Sci., **47**, 1854-1864.

Klinker, E., F. Rabier and R. Gelaro, 1998: Estimation of key analysis errors using the adjoint technique. *Q. J. R. Meteorol. Soc*. **124,** 1909-1933.

Rabier, F., E. Klinker, P. Courtier and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc*. **122,** 121-150.

Rabier, F., Jarvinen, H. Klinker, E., Mahfouf, J.-F. and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc*. **126,** 1143-1170.