Adjoint diagnostics of data assimilation systems

Carla Cardinali

European Centre for Medium-Range Weather Forecast,
Reading, UK

Summary

Adjoint diagnostic tools are here presented, in particular the observation influence in the analysis and the observation impact on the forecast error are described and their use to monitor the assimilation and forecast system. I

In this paper, the corresponding concepts to the influence matrix used in ordinary least-squares applications for monitoring statistical multiple-regression analyses have been derived in the context of linear statistical data assimilation in Numerical Weather Prediction. Results show that, in the January 2008 operational system, 7% of the global influence is due to the assimilated observations in any one analysis, and the complementary 93% is the influence of the prior (background) information, a short-range forecast containing information from earlier assimilated observations. Low-influence data points usually occur in data-rich areas, while high-influence data points are in data-sparse areas or in dynamically active regions. Background error correlations also play an important role: High correlation diminishes the observation influence and amplifies the importance of the surrounding real and pseudo observations (prior information in observation space). Incorrect specifications of background and observation error covariance matrices can be identified, interpreted and better understood by the use of influence matrix diagnostics for the variety of observation types and observed variables used in the data assimilation system.

This paper also describes the use of forecast sensitivity to observations as a diagnostic tool to monitor the observation impact on the quality of the short range forecasts (typically 24 hour). Overall, the assimilated observations decrease the forecast error. However, locally some poor performances are detected that are related either to the data quality, the sub-optimality of the data assimilation system or biases in the model. It is also found that some synoptic situation can deteriorate the quality of certain measurements or can induce some local weather variability over small areas that the assimilation system cannot correctly resolve. Finally, the performance of the current operational version (2009) of the data assimilation system for the last four months of 2008 shows a consistent overall positive impact of the observations.

1. Introduction

Over the last decade, data assimilation schemes have evolved towards very sophisticated systems, such as the four-dimensional variational system (4D-Var) (Rabier *et al.* 2000) that operates at the European Centre for Medium-Range Weather Forecasts (ECMWF). The scheme handles a large variety of both space and surface-based meteorological observations. It combines the observations with prior (or background) information on the atmospheric state and uses a comprehensive (linearized) forecast model to ensure that the observations are given a dynamically realistic, as well as statistically likely response in the analysis. Effective performance monitoring of such a complex system, with an order of 10^8 degrees of freedom and more than 10^7 observations per 12-hour assimilation cycle, has become an absolute necessity.

The assessment of each observation contribution to the analysis is among the most challenging diagnostics in data assimilation and numerical weather prediction. Methods have been derived to measure the observational influence in data assimilation schemes (Purser and Hung 1993, Cardinali *et al*. 2004, Fisher 2003, and Chapnick *et al*. 2004). These techniques show how the influence is assigned during the assimilation procedure, which partition is given to the observation and which is given to the

background or pseudo-observation. They therefore provide a indication of the robustness of the fit between model and observations and allow some tuning of the weights assigned in the assimilation system. Measures of the observational influence are useful for understanding the DA scheme itself: How large is the influence of the latest data on the analysis and how much influence is due to the background? How much would the analysis change if one single influential observation were removed? How much information is extracted from the available data?

To answer the questions it is necessary to turn to the diagnostic methods that have been developed for monitoring statistical multiple regression analyses and 4D-Var is a special case of the Generalized Least Square (GLS) problem (Talagrand, 1997) for weighted regression, thoroughly investigated in the statistical literature.

The structure of many regression data sets makes effective diagnosis and fitting a delicate matter. In robust (resistant) regression, one specific issue is to provide protection against distortion by anomalous data. In fact, a single unusual observation can heavily distort the results of ordinary (non-robust) LS regression (Hoaglin *et al.* 1983). Unusual or influential data points are not necessarily bad data points: they may contain some of the most useful sample information. For practical data analysis, it helps to judge such effects quantitatively. A convenient diagnostic measures the effect of a (small) change in the observation y_i on the corresponding predicted (estimated) value $\hat{y}_i$. In LS regression this involves a straightforward calculation: any change in y_i has a proportional impact on $\hat{y}_i$. The desired information is available in the diagonal of the *hat matrix* (Velleman and Welsh, 1981), which gives the estimated values $\hat{y}_i$ as a linear combination of the observed values y_i. The term *hat matrix* was introduced by J.W. Tukey (Tukey, 1972) because the matrix maps the observation vector $\boldsymbol{y}$ into $\hat{\boldsymbol{y}}$, but it is also referred to as the *influence matrix* since its elements indicate the data influence on the regression fit of the data. The matrix elements have also been referred to as the *leverage* of the data points: in case of high *leverage* a unit y-value will highly disturb the fit (Hoaglin and Welsh, 1978). Concepts related to the influence matrix also provide diagnostics on the change that would occur by leaving one data point out, and the effective information content (degrees of freedom for signal) in the data.

Recently, adjoint-based observation sensitivity techniques have been used (Baker and Daley 2000, Langland and Baker 2004, Cardinali and Buizza, 2004, Morneau *et al*., 2006, Xu and Langlang, 2006, Zhu and Gelaro 2008, Cardinali 2009) to measure the observation contribution to the forecast, where the observation impact is evaluated with respect to a scalar function representing the short-range forecast error. In general, the adjoint methodology can be used to estimate the sensitivity measure with respect to any parameter of importance of the assimilation system. Very recently, Daescu (2008) derived a sensitivity equation of an unconstrained variational data assimilation system from the first order necessary condition with respect to the main input parameters: observation, background and their error covariance matrices. The paper provides the theoretical framework for further diagnostic tool development not only to evaluate the observation impact on the forecast but also the impact of the other analysis parameters. Sensitivity to background covariance matrix can help in evaluating the correct specification of the background weight and their correlation. Limitations and weaknesses of the covariance matrices are well known, several assumptions and simplifications are made to derive them. Desroziers and Ivanov (2001) and Chapnik et al. (2006) discussed the importance of diagnosing and tuning the error variances in a data assimilation scheme.

The adjoint-based observation sensitivity technique measures the impact of observations when the entire observation dataset is present in the assimilation system and also it measures the response of a single forecast metric to all perturbations of the observing system. It provides the impact of all observations assimilated at a single analysis time.

The adjoint-based technique is restricted by the tangent linear assumption, valid up to 3 days. Furthermore, a simplified adjoint model is usually used to carry the forecast error information backwards, which limits further the validity of the linear assumption, and therefore restricts the use of the diagnostic to a typical forecast range of 24-48 hours. One implication to use a simplified adjoint model is that the analysis uncertainties obtained throughout the adjoint integration can be incorrect if the propagating back signal is weak (Isakseen et al., 2005).

In this paper, first the influence matrix diagnostic for ordinary least-squares regression is explained in Section 2. In Section 3 the corresponding concepts for linear statistical DA schemes are derived. It will be shown that observational influence and background influence complement each other. In Section 4, the theoretical background of the forecast sensitivity (observation and background), the numerical solution and the calculation of the forecast error contribution from observations are shown. Section 5 some illustrations and selected examples related to data influence on the analysis and on the forecast are presented. Conclusions are drawn in Section 6.

2. Classical statistical definitions of influence matrix and self-sensitivity

The ordinary linear regression model can be written

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.1}$$

where $\mathbf{y}$ is an $m\times1$ vector for the response variable (predictand); $\mathbf{X}$ is an $m\times q$ matrix of q predictors; $\boldsymbol{\beta}$ is a $q\times1$ vector of parameters to be estimated (the regression coefficients) and $\boldsymbol{\varepsilon}$ is an $m\times1$ vector of errors (or fluctuations) with expectation $E(\boldsymbol{\varepsilon})=0$ and covariance $var(\boldsymbol{\varepsilon})=\sigma^2\mathbf{I}_m$ (that is, uncorrelated observation errors). In fitting the model (2.1) by LS, the number of observations m has to be greater than the number of parameters q in order to have a well-posed problem, and $\mathbf{X}$ is assumed to have full rank q.

The LS method provides the solution of the regression equation as $\boldsymbol{\beta}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$. The fitted (or estimated) response vector $\hat{\mathbf{y}}$ is thus:

$$\hat{\mathbf{y}} = \mathbf{S}\mathbf{y} \tag{2.2}$$

where

$$\mathbf{S} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \tag{2.3}$$

is the *m×m influence matrix* (or hat matrix). It is easily seen that

$$\mathbf{S} = \frac{\partial\hat{\mathbf{y}}}{\partial\mathbf{y}} \tag{2.4}$$

and that

$$S_{ij} = \frac{\partial \hat{y}_j}{\partial y_i}$$
$$S_{ii} = \frac{\partial \hat{y}_i}{\partial y_i} \qquad 2.5$$

for the off-diagonal ($i \neq j$) and the diagonal ($i=j$) elements, respectively. Thus, S_{ij} is the rate of change of $\hat{y}_i$ with respect to y_j variations. The diagonal element S_{ii}, instead, measures the rate of change of the regression estimate $\hat{y}_i$ with respect to variations in the corresponding observation y_i. For this reason the *self-sensitivity* (or self-influence, or leverage) of the ith data point is the ith diagonal element S_{ii}, while an off-diagonal element is a *cross-sensitivity* diagnostic between two data points.

Hoaglin and Welsh (1978) discuss some properties of the influence matrix. The diagonal elements satisfy

$$0 \leq S_{ii} \leq 1 \qquad i = 1, 2, ..., m \qquad 2.6$$

as **S** is a symmetric and idempotent projection matrix ($\mathbf{S}=\mathbf{S}^2$). The covariance of the error in the estimate $\hat{\mathbf{y}}$, and the covariance of the residual $\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}$ are related to **S** by

$$\mathrm{var}(\hat{\mathbf{y}}) = \sigma^2 \mathbf{S}$$
$$\mathrm{var}(\mathbf{r}) = \sigma^2 (\mathrm{I_m} - \mathbf{S}) \qquad 2.7$$

The trace of the influence matrix is

$$tr(\mathbf{S}) = \sum_{i=1}^{m} S_{ii} = q = rank(\mathbf{S}) \qquad 2.8$$

(in fact **S** has m eigenvalues equals to 1 and m-q zeros). Thus, the trace is equal to the number of parameters. The trace can be interpreted as the amount of information extracted from the observations or *degrees of freedom for signal* (Wahba *et al.* 1995). The complementary trace, $tr(\mathbf{I} - \mathbf{S}) = m - tr(\mathbf{S})$, on the other hand, is the *degree of freedom for noise*, or simply the degree of freedom (*df*) of the error variance, widely used for model checking (F test).

A zero self-sensitivity $S_{ii}=0$ indicates that the ith observation has had no influence at all in the fit, while $S_{ii}=1$ indicates that an entire degree of freedom (effectively one parameter) has been devoted to fitting just that data point. The average self-sensitivity value is q/m and an individual element S_{ii} is considered 'large' if its value is greater than three times the average (Velleman and Welsh, 1981). By a symmetrical argument a self-sensitivity value that is less than one-third of the average is considered 'small'.

Furthermore, the change in the estimate that occurs when the ith observation is deleted is

$$\hat{y}_i - \hat{y}_i^{(-i)} = \frac{S_{ii}}{(1 - S_{ii})} r_i \qquad 2.9$$

where $\hat{y}_i^{(-i)}$ is the LS estimate of y_i obtained by leaving-out the ith observation of the vector **y** and the ith row of the matrix **X**. The method is useful to assess the quality of the analysis by using the discarded observation, but impractical for large systems. The formula shows that the impact of deleting $(y_i, \mathbf{x}_i)$ on $\hat{y}_i$ can be computed by knowing only the residual r_i and the diagonal element S_{ii} - the nearer the self-sensitivity S_{ii} is to one, the more impact on the estimate $\hat{y}_i$. A related result concerns the so-called cross-validation (CV) score: that is, the LS objective function obtained when each data point is in turn deleted (Whaba, 1990, theorem 4.2.1):

$$\sum_{i=1}^{m}(y_i - \hat{y}_i^{(-i)})^2 = \sum_{i=1}^{m}\frac{(y_i - \hat{y}_i)^2}{(1-S_{ii})^2} \qquad 2.10$$

This theorem shows that the CV score can be computed by relying on the all-data estimate $\hat{\mathbf{y}}$ and the self-sensitivities, without actually performing m separate LS regressions on the leaving-out-one samples. Moreover, equation (2.9) shows that self-sensitivities can be used to compute the change in the estimate by the leaving out one observation.

The definitions of influence matrix (2.4) and self-sensitivity (2.5) are rather general and can be applied also to non-LS and nonparametric statistics. In spline regression, for example, the interpretation remains essentially the same as in ordinary linear regression and most of the results, like the CV-theorem above, still apply. In this context, Craven and Wahba (1979) proposed the generalized-CV score, replacing in (2.10) S_{ii} by the mean tr(**S**)/q. For further applications of influence diagnostics beyond usual LS regression (and further references) see Ye (1998) and Shen *et al*. (2002). The notions related to the influence matrix that we have introduced here will in the following section be derived in the context of a statistical analysis scheme used for data assimilation in numerical weather prediction (NWP).

3. Observational influence and self-sensitivity for a DA scheme

3.1. Linear statistical estimation in Numerical Weather Prediction

Data assimilation systems for NWP provide estimates of the atmospheric state **x** by combining meteorological observations **y** with prior (or background) information $\mathbf{x}_b$. A simple Bayesian Normal model provides the solution as the posterior expectation for **x**, given **y** and $\mathbf{x}_b$. The same solution can be achieved from a classical *frequentist* approach, based on a statistical linear analysis scheme providing the Best Linear Unbiased Estimate (Talagrand, 1997) of **x**, given **y** and $\mathbf{x}_b$. The optimal GLS solution to the analysis problem (see Lorenc, 1986) can be written

$$\mathbf{x}_a = \mathbf{K}\mathbf{y} + (\mathbf{I}_n - \mathbf{K}\mathbf{H})\mathbf{x}_b \qquad 3.1$$

The vector $\mathbf{x}_a$ is the 'analysis'. The gain matrix **K** ($n \times p$) takes into account the respective accuracies of the background vector $\mathbf{x}_b$ and the observation vector **y** as defined by the $n \times n$ covariance matrix **B** and the $p \times p$ covariance matrix **R**, with

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1} \qquad 3.2$$

Here, **H** is a $p \times n$ matrix interpolating the background fields to the observation locations, and transforming the model variables to observed quantities (e.g. radiative transfer calculations transforming the models temperature, humidity and ozone into brightness temperatures as observed by several satellite instruments). In the 4D-Var context introduced below, **H** is defined to include also the propagation in time of the atmospheric state vector to the observation times using a forecast model.

Substituting (3.2) into (3.1) and projecting the analysis estimate onto the observation space, the estimate becomes

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{x}_a = \mathbf{H}\mathbf{K}\mathbf{y} + (\mathbf{I}_p - \mathbf{H}\mathbf{K})\mathbf{H}\mathbf{x}_b \tag{3.3}$$

It can be seen that the analysis state in observation space ($\mathbf{Hx_a}$) is defined as a sum of the background (in observation space, $\mathbf{Hx_b}$) and the observations **y**, weighted by the $p \times p$ square matrices $\mathbf{I} - \mathbf{HK}$ and $\mathbf{HK}$, respectively.

Equation (3.3) is the analogue of (2.1), except for the last term on the right hand side. In this case, for each unknown component of $\mathbf{H}\boldsymbol{x}$, there are two data values: a real and a 'pseudo' observation. The additional term in (3.3) includes these pseudo-observations, representing prior knowledge provided by the observation-space background $\mathbf{Hx_b}$. From (3.3) and (2.4), the analysis sensitivity with respect to the observations is obtained

$$\mathbf{S} = \frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{y}} = \mathbf{K}^T\mathbf{H}^T \tag{3.4}$$

Similarly, the analysis sensitivity with respect to the background (in observation space) is given by

$$\frac{\partial \hat{\mathbf{y}}}{\partial(\mathbf{H}\mathbf{x}_\mathbf{b})} = \mathbf{I} - \mathbf{K}^T\mathbf{H}^T = \mathbf{I}_\mathrm{p} - \mathbf{S} \tag{3.5}$$

We focus here on the expressions (3.4) and (3.5). The influence matrix for the weighted regression DA scheme is actually more complex (see Appendix 1), but it obscures the dichotomy of the sensitivities between data and model in observation space.

The (projected) background influence is complementary to the observation influence. For example, if the self-sensitivity with respect to the ith observation is S_{ii}, the sensitivity with respect the background projected at the same variable, location and time will be simply $1\text{-}S_{ii}$. It also follows that the complementary trace, $\mathrm{tr}(\mathbf{I}-\mathbf{S}) = p - \mathrm{tr}(\mathbf{S})$, is not the *df* for noise but for background, instead. That is the weight given to prior information, to be compared to the observational weight $\mathrm{tr}(\mathbf{S})$. These are the main differences with respect to standard LS regression. Note that the different observations can have different units, so that the units of the cross-sensitivities are the corresponding unit ratios. Self-sensitivities, however, are pure numbers (no units) as in standard regression. Finally, as long as **R** is diagonal, (2.6) is assured, but for more general non-diagonal **R**-matrices it is easy to find counter-examples to that property.

Inserting (3.2) into (3.4), we obtain

$$\mathbf{S} = \mathbf{R}^{-1}\mathbf{H}(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T \tag{3.6}$$

As $(\mathbf{B}^{-1}+\mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H})^{-1}$ is equal to the analysis error covariance matrix **A**, we can also write $\mathbf{S}=\mathbf{R}^{-1}\mathbf{H}\mathbf{A}\mathbf{H}^{T}$.

3.2. An idealized case, for illustration

Assume there are two observations, each coincident with a point of the background - that is $\mathbf{H}=\mathbf{I}_2$. Assume the error of the background at the two locations have correlation, that is $B=\begin{pmatrix}\sigma_b^2 & \alpha \\ \alpha & \sigma_b^2\end{pmatrix}$, with variance σ_b^2, and that $R=\sigma_o^2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. For this simple case S is obtained with

$$S_{11}=S_{22}=\frac{r+1-\alpha^2}{r^2+2r+1-\alpha^2} \qquad 3.7$$

$$S_{12}=S_{21}=\frac{\alpha r}{r^2+2r+1-\alpha^2} \qquad 3.8$$

where $r=\sigma_o^2/\sigma_b^2$. We can see that if the observations are very close (compared to the scale-length of the background error correlation), i.e $\alpha \sim 1$, then

$$S_{11}=S_{22}=S_{12}=S_{21}\simeq\frac{1}{r+2} \qquad 3.9$$

Furthermore, if $\sigma_b=\sigma_o$, that is r=1, we have three pieces of information with equal accuracy and $S_{11}=S_{22}=1/3$. The background sensitivity at both locations is $1-S_{11}=1-S_{22}=2/3$. If the observation is much more accurate than the background ($\sigma_b>>\sigma_o$), that is r~0, then both observations have influence $S_{11}=S_{22}=1/2$, and the background sensitivities are $1-S_{11}=1-S_{22}=1/2$.

We now turn to the dependence on the background-error correlation α, for the case $\sigma_b=\sigma_o$ (r=1). We have

$$S_{11}=S_{22}=\frac{2-\alpha^2}{4-\alpha^2} \qquad 3.10$$

$$S_{12}=S_{21}=\frac{\alpha}{4-\alpha^2} \qquad 3.11$$

If the locations are far apart, such that $\alpha\sim 0$, we obtain $S_{11}=S_{22}=1/2$, the background sensitivity is also ½ and $S_{12}=S_{21}=0$. We can conclude that where observations are sparse, S_{ii} and the background-sensitivity are determined by their relative accuracies (*r*) and the off-diagonal terms are small (indicating that surrounding observations have small influence). Conversely, where observations are dense, S_{ii} tends to be small, the background-sensitivities tend to be large and the off-diagonal terms are also large.

It is also convenient to summarize the case $\sigma_b=\sigma_o$ (r=1) by showing the projected analysis at location *1*

$$\hat{y}_1 = \frac{1}{4-\alpha^2}\left[(2-\alpha^2)y_1 + 2x_1 - \alpha(x_2 - y_2)\right] \qquad 3.12$$

The estimate $\hat{y}_1$ depends on y_1, x_1 and an additional term due to the second observation. We see that, with a diagonal **R**, the observational contribution is generally devalued with respect to the background because a group of correlated background values count more than the single observation [$\alpha\rightarrow\pm$ 1, ($2-\alpha^2$)$\rightarrow$1]. From the expression above we also see that the contribution from the second observation is increasing with the correlation's absolute value, implying a larger contribution due to the background x_2 and observation y_2 nearby observation y_1.

Note that Eq.(2.9) can be applied to quantify how much the analysis at a given observation location would change by deleting the observation itself. The change depends only on the self-sensitivity and the residual value at that location. Note also that the tr(**S**) provides estimates of the information content of the data with respect to the background, and is equal to the degrees of freedom for signal as studied by e.g. Purser and Huang, 1993, Rabier *et al.* (2002) and Fourrié and Thépaut (2003) in the context to remote-sensing retrieval applications. In particular, Rabier *et al.* (2002) use a "data resolution matrix", which is basically the same as the influence matrix. Recently Fisher (2003) computed an estimate of the global tr(**S**) by using the Bay *et al.*(1996) method, without explicitly computing the individual elements S_{ii}. Comparison with Fisher's (2003) estimate has provided validation of our method, in terms of the global trace.

4. Forecast sensitivity to the observations

Baker and Daley (2000) derived the forecast sensitivity equation with respect to the observations in the context of variational data assimilation. Let us consider a scalar J-function of the forecast error. Then, the sensitivity of J with respect to the observations can be written using a simple derivative chain as:

$$\frac{\partial J}{\partial \mathbf{y}} = \frac{\partial J}{\partial \mathbf{x}_a}\frac{\partial \mathbf{x}_a}{\partial \mathbf{y}} \qquad 4.1$$

$\partial J/\partial \mathbf{x}_a$ is the sensitivity to forecast error to initial condition $\mathbf{x}_a$ (Rabier et al. 1996, Gelaro et al., 1998). From (3.1) the sensitivity of the analysis system with respect to the observations and the background can be derived from:

$$\frac{\partial \mathbf{x}_a}{\partial \mathbf{y}} = \mathbf{K}^T$$
$$\frac{\partial \mathbf{x}_a}{\partial \mathbf{x}_b} = \mathbf{I} - \mathbf{H}^T\mathbf{K}^T \qquad 4.2$$

By using (4.2) and (3.2) the forecast sensitivity to the observations becomes:

$$\frac{\partial J}{\partial \mathbf{y}} = \mathbf{K}^T\frac{\partial J}{\partial \mathbf{x}_a} = \mathbf{R}^{-1}\mathbf{H}(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\frac{\partial J}{\partial \mathbf{x}_a} \qquad 4.3$$

A third (or second) order sensitivity gradient needs to be considered in 4.3 because only superior orders than first contain the information related to the forecast error. In fact, the first order one only contains information on the sub-optimality of the assimilation system (Cardinali 2009). To compute the third order sensitivity gradient, two forecasts of length f starting from x_a and length g starting from $\mathbf{x}_b$ have to be considered. Both forecasts verify at time t. Following Langland and Baker (2004) and Errico (2007) the third order sensitivity gradient is defined as

$$\frac{\partial J}{\partial \mathbf{x}_a} = \frac{\partial J_f}{\partial x_a} + \frac{\partial J_g}{\partial x_b} \qquad 4.4$$

Where J_f and J_g are a quadratic measure of the two forecast errors ($\mathbf{x}_t$ the verifying analysis, taken here as the truth), and **C** is a matrix of dry energy norm weighting coefficients. It is clear from (4.4) that the adjoint model maps the sensitivity (with respect to the forecast) of J_f along the trajectory *f* and the sensitivity of J_g along the trajectory *g*.

Once the forecast sensitivity is computed (see Cardinali 2009 for details), the variation δJ of the forecast error expressed by J can be found by rearranging (3.1) and by using the adjoint property for the linear operator:

$$\delta J = \left\langle \frac{\partial J}{\partial \mathbf{x}_a}, \delta \mathbf{x}_a \right\rangle = \left\langle \frac{\partial J}{\partial \mathbf{x}_a}, \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b) \right\rangle = \left\langle \mathbf{K}^T \frac{\partial J}{\partial \mathbf{x}_a}, \mathbf{y} - \mathbf{H}\mathbf{x}_b \right\rangle = \left\langle \mathbf{K}^T \frac{\partial J}{\partial \mathbf{x}_a}, \delta \mathbf{y} \right\rangle = \left\langle \frac{\partial J}{\partial \mathbf{y}}, \delta \mathbf{y} \right\rangle \qquad 4.5$$

where $\delta \mathbf{x}_a = \mathbf{x}_a - \mathbf{x}_b$ are the analysis increments and $\delta \mathbf{y} = \mathbf{y} - \mathbf{H}\mathbf{x}_b$ is the innovation vector. The sensitivity gradient $\partial J / \partial \mathbf{x}_a$ is valid at the starting time of the 4D-Var window (typically 09 and 21 UTC for the 12h 4D-Var set-up used at ECMWF). As for **K**, its adjoint $\mathbf{K}^T$ incorporates the temporal dimension, and the δ**y** innovations are distributed over the 12-hour window. The variation of the forecast error due to a specific measurement can be summed up over time and space in different subsets to compute the average contribution of different component of the observing system to the forecast error. For example, the contribution of all AMSU-A satellite instruments, *s*, and channels, *i*, over time T will be:

$$\delta J_{AMSU-A} = \sum_{\substack{s \in S \\ t \in T}} \sum_{i \in channel} \delta J^s{}_{it}$$

The forecast error contribution can be gathered over different subsets that can represent a specific observation type, a specific vertical or horizontal domain, or a particular meteorological variable.

5. Results

5.1. Observation Influence or Analysis sensitivity to observations

The diagonal elements of the influence matrix have been computed for the operational 4D-Var assimilation system. The calculations in Eq(3.15) have been carried out on 91 model levels at T255 spectral truncation. The observation departures ($\mathbf{y}$-$\mathbf{H}\mathbf{x}_b$) were calculated by comparing the observations with a 12-hour forecast integration at T511 resolution. The counts of assimilated observations for each

main observation type are given in Table.1. A large proportion of the used data is provided by satellite systems (Thépaut and Andersson 2003): ASCAT and QuikSCAT near-surface winds, AMV (Meteosat, Goes, GMS and MODIS) cloud-drift winds, AMSU-A and HIRS infrared radiances, SSMI microwave imager, GOES and METEOSAT water-vapour radiances, AMSU-B and MHS microwave sounder radiances, AIRS and IASI infrared sounder radiances, GPS-RO radio occultation, MTSATIMG infrared imager radiances, AMSE radiometer radiances and Ozone data. The remainder are surface-based observing systems (see WMO 1996).

Table 1: Observation type, 20090901-12 UTC.

Type of Data	Description
SYNOP	Surface Observations from land and ship stations: measuring p_s, T, RH , u and v
SCAT	Satellite microwave scatterometer: derived measurement is u and v at the ocean surface
DRIBU	Drifting buoy measuring p_s, T, RH, u and v
Meteosat AMV	Satellite cloud drift winds (European)
Goes AMV	Satellite cloud drift winds (American)
GMS AMV	Satellite cloud drift winds (Japanese)
MODIS AMV	Polar Satellite cloud drift winds
TEMP	Radiosondes from land and ship measuring p_s, T, RH , u and v
AMSUA	Satellite microwave sounder radiances
PILOT	Sondes and Wind profiler measuring u and v
AIRS	Satellite infrared radiances
IASI	Satellite infrared radiances
HIRS	Satellite infrared radiances
MSG	METEOSAT Second Generation [satellite]
AIREP	Aircraft measurements of T, u and v
SSMI	Satellite microwave imager radiances
MHS	Satellite microwave radiances
AMSU-B	Satellite microwave radiances
GPS-RO	Satellite Radio Occultation
GOES	Geostationary satellite infrared sounder radiances
OZONE	Satellite ozone retrieval
METEOSAT-R	Geostationary satellite infrared sounder radiances
GOES-R	Geostationary satellite infrared sounder radiances
MTSATIMG	

5.1.1. Self-sensitivity examples

Self-sensitivities for SYNOP surface pressure observations are shown in Fig.1. Each box indicates the observation influence at the observation location. Low-influence data points have large background influence (see 3.4 and 3.5), which is the case in data-rich areas such as North America and Europe (observation influence ~ 0.2). In data-sparse areas individual observations have larger influence: in the Polar regions, where there are only few isolated observations, $S_{ii} \sim 1$ and the background has small influence on the analysis.

In dynamically active areas (Fig.1: North Atlantic cyclone genesis), several fairly isolated observations have large influence on the analysis. This is due to the evolution of the background-error covariance matrix as propagated by the forecast model in 4D-Var (Thépaut *et al.* 1993, 1996). As a result, the data assimilation scheme can fit these observations more closely.

Similar features can be seen in Fig.2 showing the influence of u-component wind observations for aircraft (AIREP), radiosonde (TEMP), wind profiler and Atmospheric Motion Vector data between 300 and 200 hPa. Isolated flight tracks over Atlantic and Pacific oceans show larger influences than measurements over data-dense areas over America and Europe.

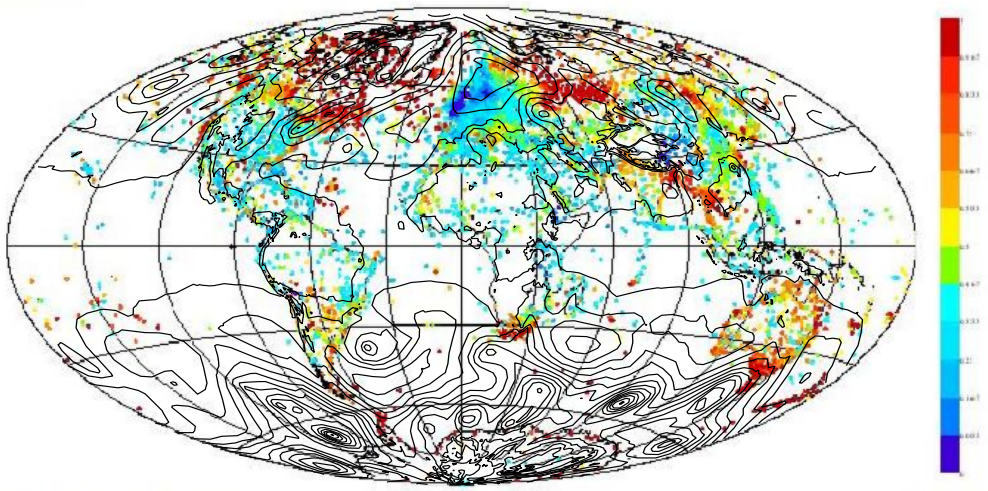

Figure 1: Synop surface pressure observation influence and MSLP field superimposed.

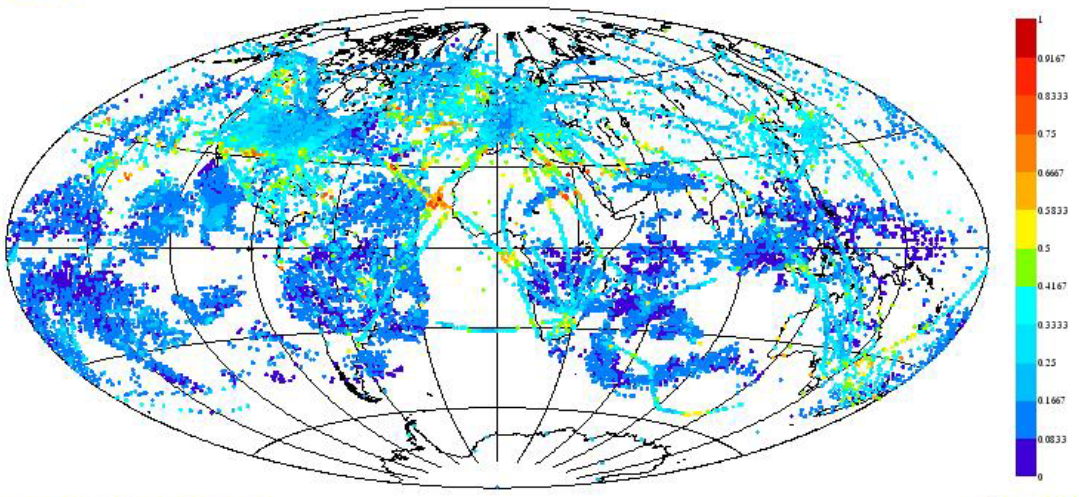

Figure 2: Observation influence of u component of wind for Aircraft, Radiosonde, Profiler and AMV

5.1.2. Trace diagnostic

We define the Global Average Influence (*GAI*) as the globally averaged observation influence. It is given by

$$GAI = \frac{tr(\mathbf{S})}{p} \qquad 5.1$$

where p is the total number of observations. In our experiment we found that GI=0.07. Consequently, the average background global influence to the analysis at observation points is equal to 0.93 (see 3.5). Another index of interest is the Partial Influence (*PAI*) for any selected subset of data

$$PAI = \frac{\sum_{i \in I} S_{ii}}{p_I} \qquad 5.2$$

where p_I is the number of data in subset I. The subset I can represent a specific observation type, a specific vertical or horizontal domain, a particular meteorological variable, for example. Figure 2 shows the GAI for the ECMWF operational system (January 2008).

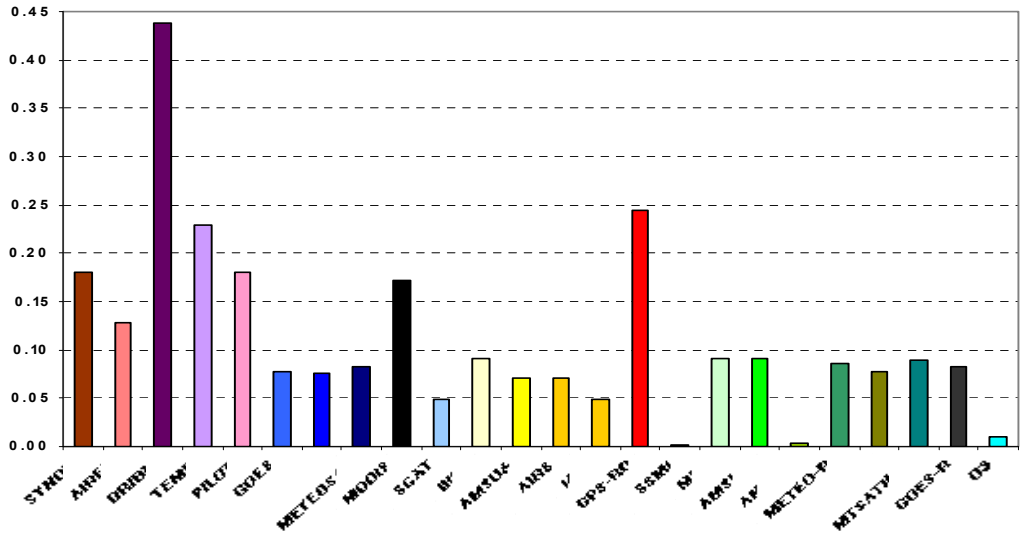

Figure 3: Average observation influence for all observation type assimilated. ECMWF January 2008

5.1.3. Information content

In Section 2 we showed that tr(**S**) can be interpreted as a measure of the amount of information extracted from the observations. In fact, in non-parametric statistics, tr(**S**) measures the 'equivalent number of parameters' or *degrees of freedom for signal*. Having obtained values of all the diagonal elements of **S** (see Cardinali *et al* 2004 for details) we can now obtain reliable estimates of the information content in any subset of the observational data. In Figure 4 we illustrate this in one example. The figure shows the information content for all main observation types. We see that AIRS radiances are the most informative data type, followed by IASI, AMSU-A and GPS-RO data. The information content of AIREP, TEMP and SYNOP is the higher for conventional observations comparable to SCAT and HIRS data. DRIBU information content is small but on the contrary (Fig.3)

the mean influence is very large. This implies that the DFS is affected by the low number of observations.

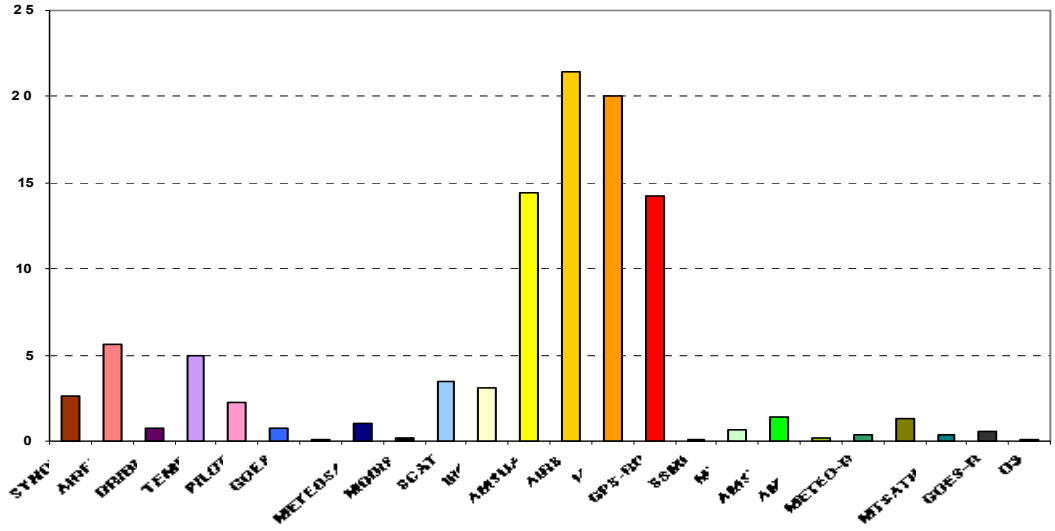

Figure 4: Information content or DFS(%) for observation type. ECMWF January 2008.

5.2. Forecast sensitivity to observations

The forecast sensitivity to the observation (FSO) has been computed from 15 June to 15 July 2006 at 00 and 12 UTC based on the forecast error calculated for the *control* experiment of the Observing System Experiments (OSE) performed by Kelly and Thépaut (2007). The FSO calculation (4.4) has been carried out on 60 model levels and with a horizontal truncation of T159 to match with the OSE final inner loop resolution and also based on both 00 and 12 UTC forecast error (only the 00 impact is shown). As for the OSE, the observation departures were computed at T511 (model trajectory resolution, Rabier *et al*. 2000). All the experiments were performed using the third order sensitivity gradient defined in section 4 and based on the 24 hour forecast error. The sensitivity to the humidity initial condition is obtained as a secondary effect due to the adjoint of the linearized moist physical processes used in the sensitivity gradient calculation (Lopez and Moreau 2005, Tompkins and Janisková 2004, Janisková *et al*. 2002) which accounts for the dependency of the forecast error at the verification time due to the humidity errors in the initial conditions. The energy norm diagnostic function was computed from the OSE *control* (using all available observations) experiment forecast error.

5.2.1. Global impact

The global observation performance over this month, as described in 4.5, is summarized in Fig.5. Negative (positive) values correspond to a decrease (increase) of forecast error due to a specific observation type. The range of the results accuracy is estimated to be ~16%, therefore small negative and small positive values should be regarded quantitatively as neutral observation impact. Nonetheless, degradation observed in the error range can bring useful information on the possible causes affecting the data performance on the forecast, as will be shown.

The largest error decrease is due to AMSU-A (four satellites) and AIRS radiances followed by SYNOP (mainly surface pressure), AIREP and DRIBU (mainly surface pressure) conventional

observations. Good error reduction is also observed from SCAT (Quikscat and ERS scatterometer) and AMSU-B radiance observations. An increase of forecast error is caused by AMVs (Atmospheric Motion Vector) from geostationary satellites. Some degradation is also observed from PILOT observations.

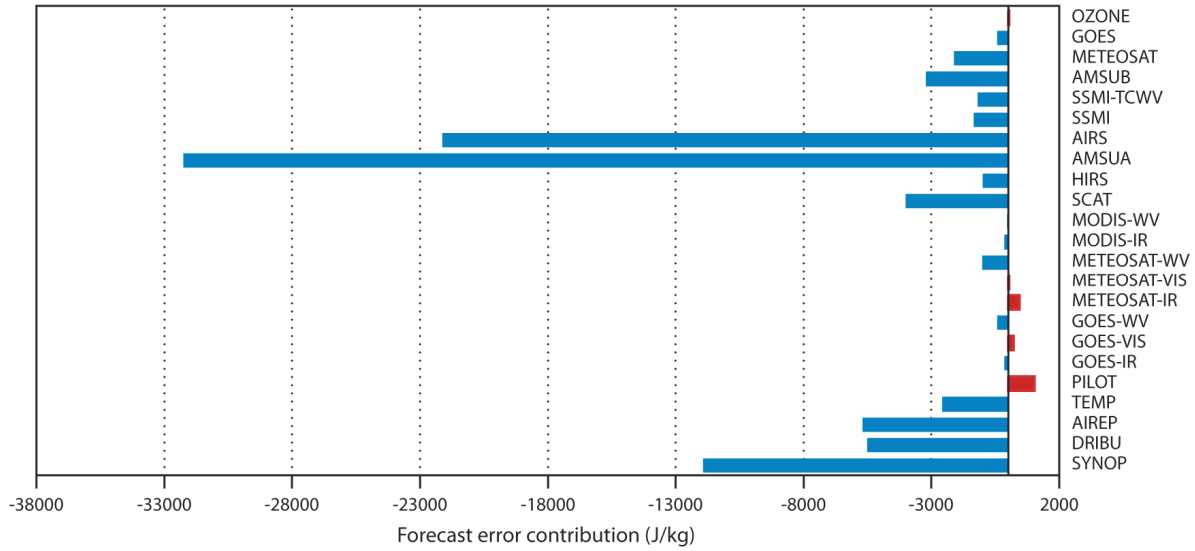

Figure 5: 24-hour forecast error contribution (third order sensitivity gradient) in J/kg of the components (types) of the observing system in summer 2006. Negative (positive) values correspond to a decrease (increase) in the energy norm of forecast error.

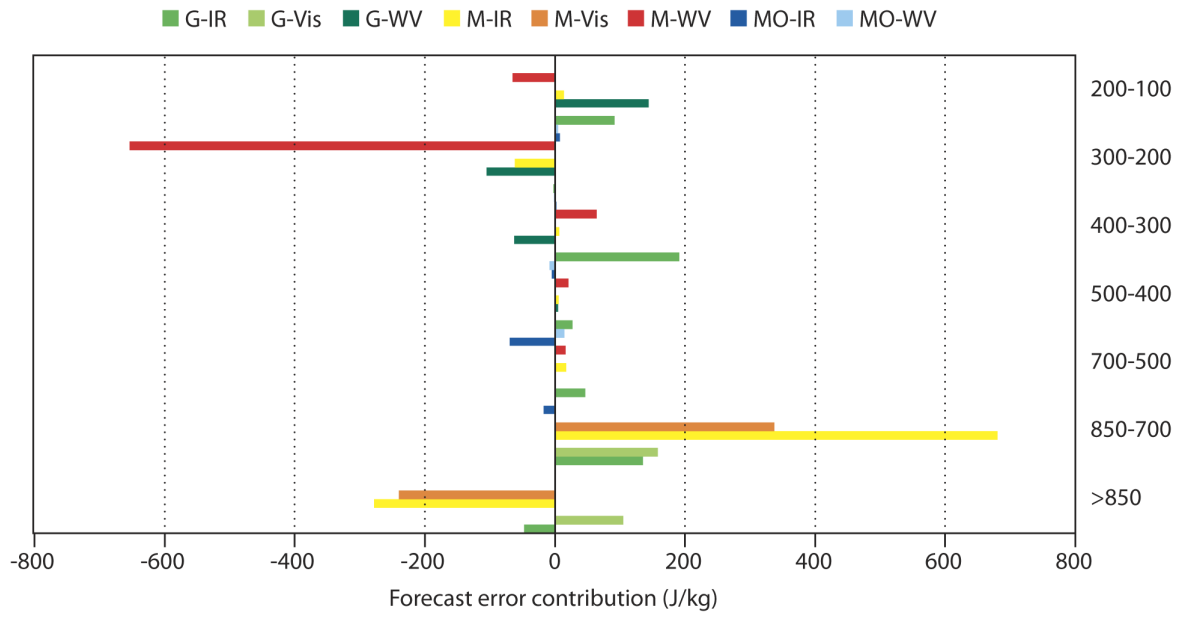

Figure 6: Forecast error contribution (third order sensitivity gradient) of the observed u-component of the wind on pressure levels and grouped by satellite types: GOES (G, two satellites GOES-8 and 9), METEOSAT (M, two satellite METEOSAT-7 and 8) and MODIS (MO, two satellites: Terra and Aqua) and by frequency bands: infrared (IR), visible (V) and water vapour (WV). Negative (positive) values correspond to a decrease (increase) of forecast error

5.2.2. AMV

A more detailed diagnostic of the forecast error contribution from AMVs is shown in Fig.6. The contribution to the forecast error of the observed u-wind component is grouped by pressure levels, satellite types, such as GOES (G, two geostationary satellites GOES-8 and 9), METEOSAT (M, two geostationary satellites METEOSAT-7 and 8) and MODIS polar instruments (MO, MODIS Terra and Aqua), and by frequency bands: infrared (IR), visible (V) and water vapour (WV). The largest degradation is due to the visible and infrared frequency band at levels below 700 hPa, (mainly at 850 hPa) from METEOSAT (to a larger extent) and from the GOES satellites.

The geographical locations of the degradation are shown in Fig 7 which displays the 00 UTC forecast error contribution of the visible and infrared bands between 1000 and 700 hPa accumulated over the summer month. The largest degradation is found over the southern equatorial band, in particular over the Atlantic (area-1) and Indian ocean (area-2) where the METEOSAT satellites are located, followed by the one over the West Pacific (area-3) where GOES is operated. In the Indian Ocean, a well established Indian Monsoon circulation was taking place, characterized by a strong low level wind from South-East towards the Indian continent. Such a situation is not well represented by the model that tends to reinforce too much the low level circulation. The degradation due to the AMV in the area-2 is therefore likely attributed to a model bias. On the contrary, over the South of the Atlantic ocean (area-1) due to the presence of semi-permanent anti-cyclone circulation in the tropical band, the associated large scale subsidence reinforces the trade inversion with a subsequent suppression of deep clouds (around 30 degrees), leaving only the shallow ones. This synoptic situation has implication with the methodology applied by the data provider to measure the height of the top of the clouds, resulting in a degradation of the data quality. Similar synoptic situation to area-2 is also noticed in area-3, it is therefore believed that, even for this case, the degradation is attributed to the data quality.

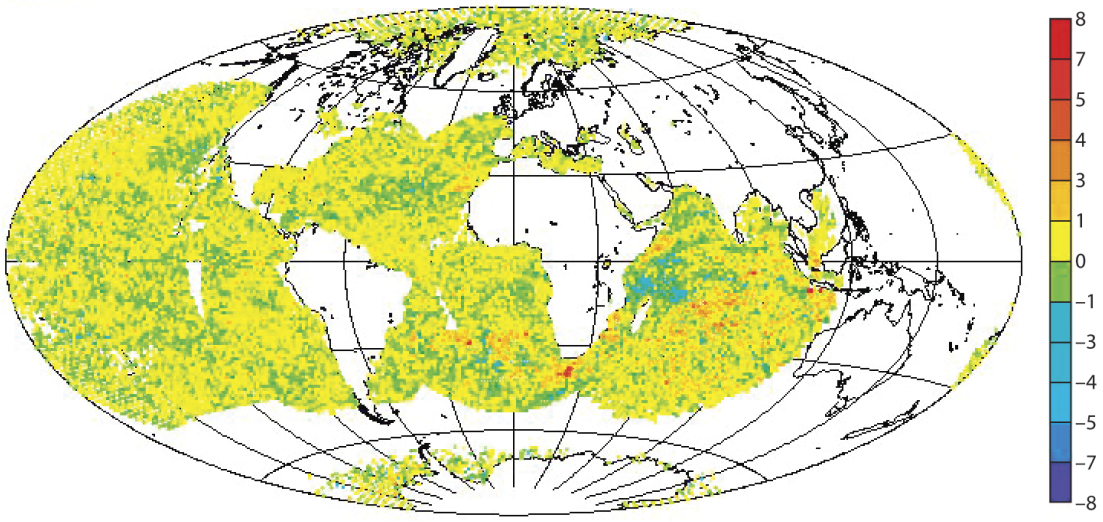

Figure 7: 00 UTC forecast error contribution (J/kg) (third order sensitivity gradient) of the observed u-component of the wind between 700 and 1000 hPa from GOES and METEOSAT visible wavelength bands accumulated over one month in summer 2006. Negative (positive) values correspond to a decrease (increase) of forecast error.

Figure 5 shows also a forecast error increase due to PILOT observations (Table 1). The geographical display of the forecast error for PILOT observations (not shown) indicates that the degradation was coming from the American wind profilers. Problems with the American wind profilers at low levels (below 700 hPa) were known in spring time due to bird migration contamination (Wilczak *et al*. 1995). But other meteorological situations also produce a contamination of profiler measurements (Ackley *et al*. 1998), one of which is the limitation of the local horizontal atmospheric uniformity assumption that must be satisfied to have a correct mean wind measure. Meteorological conditions in which short spatial and temporal scales of variability have amplitudes as large as the mean, as for example in the presence of a CBL (Convective Boundary Layer) and severe storms, limit the horizontal wind measurement. It was effectively found that the CBL-activity was rather high for this period as can be seen from the large height of the boundary layer at the station locations, averaged among all profiler stations (not shown). It was also found that both CAPE and TCWV compared with the ERA climatology (Uppala *et al*. 2005) indicated larger CAPE and humidity advection from the Gulf of Mexico in areas where wind profilers are located (not shown). Together, high TCWV and CAPE, triggered the convection activity. The lessons learnt with wind profilers is that their impact on the forecast can change quite a lot given the meteorological situations, therefore monitoring their impact on forecast skill, on a daily basis, would allow a more efficient screening of the contaminated measurements.

5.3. Analysis and Forecast Sensitivity comparisons

In this section the observation influence is compared to the observation impact on the 24 hour forecast error. The ECMWF operational system in September 2008 is used and both sensitivities are computed at the resolution T255 for 91 model levels. The model trajectory resolution is T511. Figure 8 shows the information content (8 a) and the variation of the forecast impact (8 b) due to all different observation types assimilated. Both impacts are shown in percentage. It is evident that the two impacts are quite similar. In particular, the figures show the degree of the dependency of the two different measures to $\mathbf{K}^{\mathrm{T}}$. In fact, when Eq. 3.4 and Eq. 4.5 are compared, we can see that whilst the observation influence depends only by $\mathbf{K}^{\mathrm{T}}$, the forecast error variation contains two extra-terms, the forecast error (sensitivity gradient) and the innovation vector. Therefore, the differences between the two plots would depend on these two components. AMSU-A data with the largest contribution on the forecast error reduction have smaller influence in the analysis than AIRS which is considered the most influential observation type. Similar is for IASI when compared to AMSU-A. It can be possible that the contribution in the forecast error decreases with respect to the analysis observation influence due to presence of model biases. Viceversa small observation influence and large contribution in the forecast could be due to some sub-optimality of the assimilation system that prevent the extraction of all the information from the data.

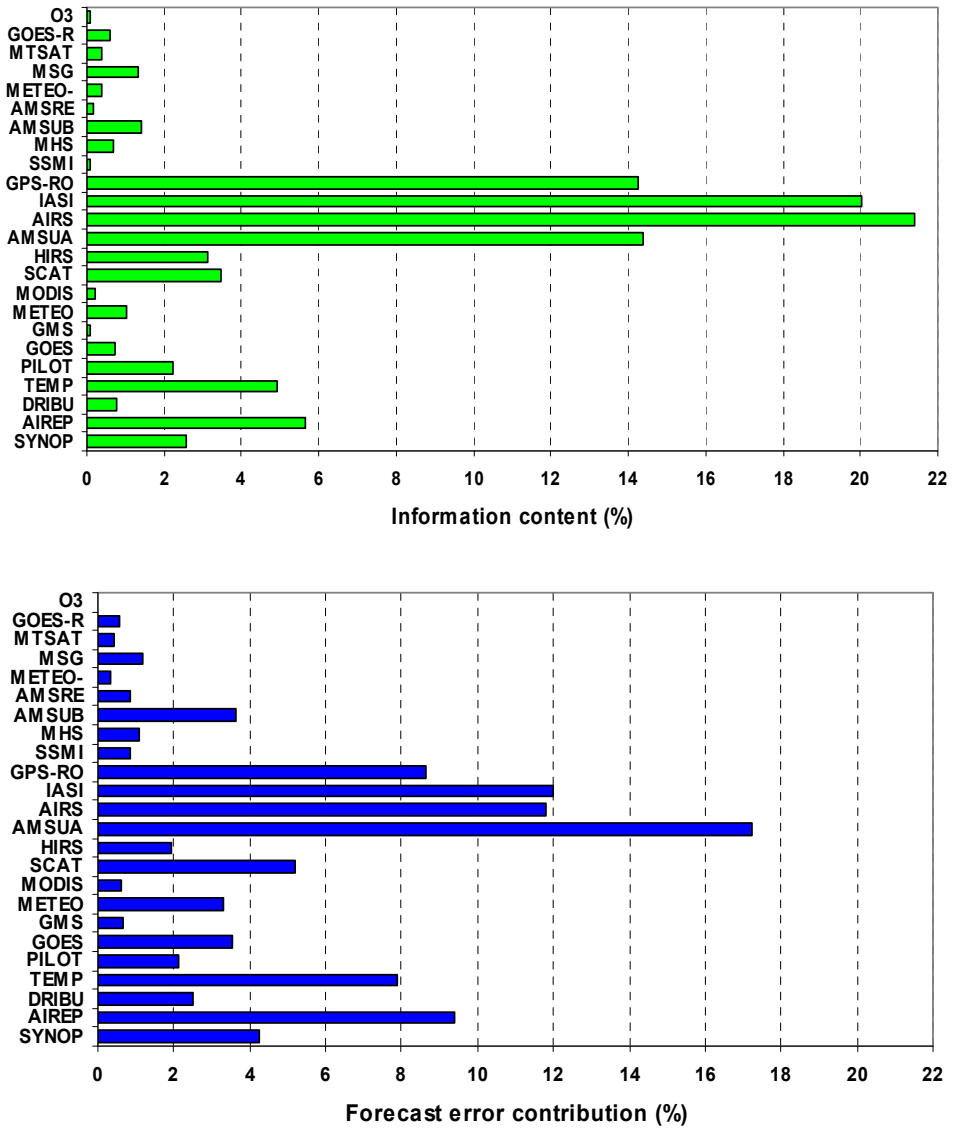

Figure 8: Information content (a) and Forecast error contribution in percentage for all data type assimilated in September 2008 at ECMWF.

6. Conclusions

Over the last few years, the potential of using derived adjoint-based diagnostic tools has been increasingly exploited.

The influence matrix is a well-known concept in multi-variate linear regression, where it is used to identify influential data and to predict the impact on the estimates of removing individual data from the regression. In this paper we have derived the influence matrix in the context of linear statistical analysis schemes, as used for data assimilation of meteorological observations in numerical weather prediction (Lorenc 1986). The self-sensitivity provides a quantitative measure of the observation influence in the analysis. In robust regression, it is expected that the data have similar self-sensitivity (sometimes called leverage) - that is, they exert similar influence in estimating the regression line. Disproportionate data influence on the regression estimate can have different reasons: First, there is

the inevitable occurrence of incorrect data. Second, influential data points may be legitimately occurring extreme observations. However, even if such data often contain valuable information, it is constructive to determine to which extent the estimate depends on these data. Moreover, diagnostics may reveal other patterns e.g. that the estimates are based primarily on a specific sub-set of the data rather than on the majority of the data. In the context of 4D-Var there are many components that together determine the influence given to any one particular observation. First there is the specified observation error covariance **R**, which is usually well known and obtained simply from tabulated values. Second, there is the background error covariance **B**, which is specified in terms of transformed variables that are most suitable to describe a large proportion of the actual background error covariance. The implied covariance in terms of the observable quantities is not immediately available for inspection, but it determines the analysis weight given to the data. Third, the dynamics and the physics of the forecast model propagate the covariance in time, and modify it according to local error growth in the prediction. The influence is further modulated by data density. We showed examples for surface pressure and conventional wind observations indicating that low influence data points occur in data-rich areas while high influence data points are in data-sparse regions or in dynamically active areas. Background error correlations also play an important role. In fact, very high correlations drastically lessen the observation influence (it is halved in the idealized example presented in Section 3.2) in favour of background influence and amplify the influence of the surrounding observations.

In this study the global observation influence per assimilation cycle has been found to be 7%, and consequently the background influence is 93%. Thus, on average the observation influence is low compared to the influence of the background (the prior). However, it must be taken into account that the background contains observation information from the previous analysis cycles. The theoretical information content (the degrees of freedom for signal) for each of the main observation types was also calculated. It was found that AIRS, IASI and AMSU-A radiance data provide the most information to the analysis followed by GPS-RO. AIREP, TEMP and SYNOP is the higher for conventional observations and then SCAT and HIRS data.

Self-sensitivities provide an objective diagnostic on the performance of the assimilation system. They could be used in observation quality control to protect against distortion by anomalous data (however this aspect has not been explored within the current study). In fact, the leaving-out-one observation, that is not practical for large system dimension, uses the discarded observation to assess the quality of the analysis. It has been shown (Eq. 2.9) that Self-sensitivities provide a similar diagnostic without performing separate least square regressions. Self-sensitivities also provide indication on model and observation error specification and tuning. Incorrect specifications can be identified, interpreted and better understood through observation influence diagnostics, partitioned e.g. by observation types, variable, levels, and regions.

In the near future more satellite data will be used and likely be thinned. Thinning has to be performed either to reduce the observation error spatial correlation (Bormann *et al.* 2003) or to reduce the computational cost of the assimilation. The observation influence provides an objective way of selecting observations dependent on their local influence on the analysis estimate to be used in conjunction with forecast impact assessments. It would be interesting to see if a small sub group of very influential data (i.e. satellite observations) have the same impact in the forecast than the full amount of data. If this is the case, a dynamical thinning can be thought that selects, every assimilation cycle, the most influent partition of a particular remote sensing instrument measurements, from

information based on the previous cycle. Clearly, it can be assumed that components of the observing network remains constant and background error variances remained almost unchanged for close assimilation cycles.

The forecast sensitivity to the observation has been based on the forecast error of the *control* experiment from observing system experiments that have been performed at ECMWF (2007). Forecast sensitivity to observations can only be used to diagnose the impact on the short-range forecast, namely 24 to 48 hours, given the use of a simplified adjoint of the data assimilation system and the implied linearity assumption. On the other hand, the use of FSO allows the identification of potential problems and directs further investigations. the global impact of observations was found to be positive and the forecast errors decrease for almost all data type. Problems have been noticed with Atmospheric Motion Vectors mainly derived from visible and infrared wavelength bands (and for low-level winds). Problems with conventional observations, wind profilers was mainly due to the local synoptic situation. Wind profiler measurements were corrupted by the presence of strong convection activity in the boundary layer.

Over the four months most recent period examined in autumn 2008, the impact of all type of observations on the short-range forecast has impressively increased and it has been shown that microwave satellite measurements (AMSU-A) are responsible for the 18% of the forecast error reduction, infrared measurements (AIRS and IASI) for the 12% and the 10% of error reduction is due to radio occultation observations. Conventional observations (surface pressure, vertical profiler and aircraft) are as well decreasing the forecast error being responsible for an average reduction of 6%.

The observation influence is compared to the observation impact on the 24 hour forecast error. The ECMWF operational system in September 2008 is used and both sensitivities are computed. The two sensitivities impact is quite similar showing strong influence in the analysis and the short range forecast of the Kalman gain.

Acknowledgements

The author thanks Mike Fisher, Erik Andersson and Fernando Prates for valuable discussion and suggestions. Thank to Jan Haseler and Sami Saarinen for the technical support.

References

Ackley M., R. Chadwick J. Cogan, C. Crosiar, F. Eaton, K.Gage, E. Gossard, R. Lucci, F. Merceret, W. Neff, M. Ralph, R. Strauch, D. Van de Kamp, B. Weber, A. White, 1998: U.S. Wind Profilers: A review. FCM-R14-1998

Andersson, E. and Fisher, M. 1999: Background errors for observed quantities and their propagation in time. Proc. ECMWF Workshop on "Diagnosis of Data Assimilation Systems", Reading, U.K., 1-4 Nov. 1998, 81—90.

Andersson, E., Fisher, M. Munro, R. and McNally, A, 2000: Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc.* **126,** 1455—1472.

Bay, Z., Fahey, M. and Golub, G. H.1996: Some large scale matrix computation problems. *J. Comput. Appl. Math*., **74**, 21—89.

Baker N.L. and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation targeting problem. *Q. J. R. Meteorol. Soc*, **126**,1431-1454

Bormann, N., Saarinen, S., Kelly G. and Thépaut, J-N. 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.*, **131,** 706—718.

Cardinali, C., S. Pezzulli and E. Andersson, 2004: Influence matrix dagnostics of a data assimilation system. *Q. J. R. Meteorol. Soc.*, **130,** 2767—2786

Cardinali, C., and R. Buizza: 2004. Observation sensitivity to the analysis and the forecast: a case study during ATreC targeting campaign. Proceedings of the First THORPEX International Science Symposium, 6-10 December 2004, Montreal, Canada, WMO TD 1237 WWRP/THORPEX N. 6.

Cardinali, C, 2009: Monitoring the forecast impact on the short-range forecast. *Q. J. R. Meteorol. Soc.*, **135,** 239—250.

Chapnik, B., G.Desnrozier, F. Rabier and O. Talagrand, 2006: Diagnosis and tuning of observation error in a quasi-operational data assimilation setting. *Q. J. R. Meteorol. Soc.*, **132,** 543—565.

Courtier, P., Andersson, E., Heckley, W., Pailleux,, J., Vasiljevic, Hamrud, D. M., Hollingsworth, A., Rabier, F. and Fisher, M., 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.* **124,** 1783—1807.

Craven, P., and Wahba, G., 1979: Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math.*, **31,** 377—403.

Daescu, D.N., 2008: On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation. *Mon. Wea. Rev.*, **136**, 3050-3065.

Errico, R., 2007: Interpretation of an adjoint-derived observational impact measure. *Tellus*, **59A**, 273-276.

Fisher, M., 1996: The specification of background error variances in the ECMWF variational analysis system. Proc. ECMWF workshop on "Non-linear aspects of data assimilation", Reading, 9-11 September 1996, 645—652.

Fisher, M., 2003: Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. *ECMWF Tech. Memo.*, **397,** pp 18.

Fisher, M. and Courtier, P., 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF Tech Memo.*, **220,** pp 26.

Fisher, M. and Andersson, E., 2001: Developments in 4D-Var and Kalman Filtering. *ECMWF Tech Memo.*, **347,** pp 36.

Fourrié, N. and Thépaut, J-N., 2003: Evaluation of the AIRS near-real-time channel selection for application to Numerical Weather Prediction. *Q. J. R. Meteorol. Soc*, **129**, 2425-2439.

Gelaro R, Buizza, R., Palmer T.N. and Klinker E., 1998: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, **55**, 1012—1037.

Kelly, G., and Thépaut, J-N 2007: Evaluation of the impact of the space component of the Global Observation System through Observing System Experiments. ECMWF Newsletter No 113, 16-28.

Hoaglin, D. C., Mosteller, F. and Tukey J.W., 1982. Understanding Robust and Exploratory Data Analysis. *Wiley Series in Probability and Statistics*

Isaksen L., M. Fisher, E. Andersson, and J. Barkmeijer 2005: The structure and realism of sensitivity perturbations and their interpretation as 'Key Analysis'. *Q. J. R. Meteorol. Soc*, **131**, 3053—3078.

Hoaglin, D. C., and Welsch, R. E. 1978: The hat matrix in regression and ANOVA. *The American Statisticians*, **32**, 17—22 and *Corrigenda* **32**, 146.

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

Lorenc, A., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112,** 1177—1194.

Lopez, P. and E. Moreau, 2005: A convection scheme for data assimilation: Description and initial tests. *Q.J.R.Meteorol.Soc*., **131**, 409—436

Janiskova, M., J.-J. M. J.-F. Mahfouf, and F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation, *Q. J. R. Meteorol. Soc*., **128**, 1505—1527.

Purser, R. J. and Huang, H.-L.,1993: Estimating Effective Data Density in a Satellite Retrieval or an Objective Analysis. *J. Appl. Meteorol.*, **32,** 1092—1107.

Rabier, F. and Courtier, P., 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.*, **118,** 649—672.

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

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

Rabier, F., Fourrié, N., Chafaï D.and Prunet, P., 2002: Channel selection methods for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.*, **128,** 1011—1027.

Shen, X., Huang H., and Cressie N., 2002: Nonparametric hypothesis testing for a spatial signal. *J.Am.Stat.Ass.*, **97**, 1122—1140

Talagrand, O., 1997: Assimilation of observations, an Introduction. *J. Meteorol. Soc. Japan*, **Vol 75,** N.1B,191—209.

Talagrand, O., 1999: A posteriori evaluation and verification of the analysis and assimilation algorithms. Proceedings of Workshop on diagnosis of data assimilation systems, ECMWF, Reading, England.

Talagrand O., 2002: A posteriori validation of assimilation algorithms. Proceding of NATO Advanced Study Institute on Data Assimilation for the Earth System, Acquafreda, Maratea, Italy.

Thépaut, J.N., Hoffman, R.N. and Courtier, P., 1993: Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.*, **121,** 3393—3414.

Tompkins, A.M. and M. Janiskova: 2004, A cloud scheme for data assimilation: Description and initial tests. *Q.J.R.Meteorol.Soc*,**130**, 2495—2517

Thépaut, J.N., Courtier, P., Belaud G. and Lemaître, G., 1996: Dynamical structure functions in four-dimensional variational assimilation: A case study. *Q. J. R. Meteorol. Soc.*, **122,** 535—561.

Thépaut, J.N. and Andersson, E. 2003: Assimilation of high-resolution satellite data. *ECMWF Newsletter*, **97,** 6—12.

Tukey, J. W. 1972: Data analysis, Computational and Mathematics. *Q. of Applied mathematics*, **30**, 51—65

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J. 2005: The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.*, **131**, 2961—3012.

Velleman, P. F., and Welsch, R. E., 1981: Efficient computing of regression diagnostics. *The American Statisticians,* **35**, 234—242.

Wahba, G., 1990: Spline models for observational data. SIAM, CBMS-NSF, *Regional Conference Series in Applied Mathematic,* **59,** pp 165

Wahba, G., Johnson, D.R., Gao F. and Gong, J. 1995: Adaptive tuning of numerical weather prediction models: Randomized GCV in three- and four-dimensional data assimilation. *Mon. Wea. Rev.*, **123,** 3358—3369.

Wilczak, J., R. Strauch, F. Ralph, B. Weber, D. Merritt, J. Jordan, D. Wolfe, L. Lewis, D. Wuertz, J. Gaynor, S. McLaughlin, R. Rogers, A. Riddle, and T. Dye, 1995: Contamination of Wind Profiler Data by Migrating Birds: Characteristics of Corrupted Data and Potential Solutions. *J. Atmos. Oceanic Technol*., **12**, 449-467.

WMO, 1996: Guide to meteorological instruments and methods of observation. Sixth Edition. WMO-No.8. Geneva, Switzerland.

Ye J., 1998: On measuring and correcting the effect of data mining and model selection. *J. Am. Stat. Ass.,* **93**, 120—131

Zhu, Y. and Gelaro, R., 2008: Observation Sensitivity Calculations Using the Adjoint of the Gridpoint Statistical Interpolation (GSI) Analysis System. *Mon. Wea. Rev.*, **136,** 335—351.