Methods for computing analysis and forecast sensitivity to observations: examples from NWP

Anthony T. Weaver
CERFACS, Toulouse

Special thanks to C. Cardinali (ECMWF) and G. Desroziers (Météo-France)

Introduction

- OSEs can provide useful information about the impact of observations on analyses and forecasts in a *particular* assimilation/forecast system.
- The number of observations assimilated on a given cycle is on the order of 10^5 to 10^6.

  OSEs (withholding different observations) can be used to study the impact of only a very small subset of these observations.
- Assimilation/forecast systems are in continuous evolution so the impact of observations needs to be regularly re-assessed.

  OSEs are expensive so feasibility of this is questionable.
- A less costly procedure is needed for regular monitoring of the impact of observations.

  This is an active area of research in NWP.

Analysis sensitivity to observations

- What is the relative influence of the observations and background state on the analysis?
- How much influence do individual observations have on the analysis?
- An influential observation

  may provide important information about an anomalous event

  may point to a data quality control problem

  may point to a problem with the assimilation system (e.g., an incorrect specification of the error covariances $\mathbf{B}$ and $\mathbf{R}$)
- The influence matrix (from ordinary least-squares) provides a quantitative diagnostic for answering these questions with statistical data assimilation systems.

The influence matrix
(Cardinali *et al.* 2004, QJRMS)

- Operational data assimilation methods are variants of statistical linear estimation.
- The (linear) analysis is given by

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b) = \mathbf{K}\mathbf{y}^o + (\mathbf{I}_n - \mathbf{K}\mathbf{H})\mathbf{x}^b$$

where

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1} \quad \text{model-space inversion}$$

$$= \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad \text{observation-space inversion}$$

- The analysis at the observation points is

$$\mathbf{y}^a = \mathbf{H}\mathbf{x}^a = \mathbf{H}\mathbf{K}\mathbf{y}^o + (\mathbf{I}_p - \mathbf{H}\mathbf{K})\mathbf{H}\mathbf{x}^b$$

The influence matrix

- The analysis at the observation points is

$$\mathbf{y}^a = \mathbf{HKy}^o + (\mathbf{I}_p - \mathbf{HK})\mathbf{Hx}^b$$

- The analysis *sensitivity* with respect to the *observations* is

$$\frac{\partial \mathbf{y}^a}{\partial \mathbf{y}^o} = \mathbf{K}^T\mathbf{H}^T$$ Observation influence

- The analysis *sensitivity* with respect to the *background equivalent of the observations* is

$$\frac{\partial \mathbf{y}^a}{\partial (\mathbf{Hx}^b)} = \mathbf{I}_p - \mathbf{K}^T\mathbf{H}^T$$ Background influence

- The *influence matrix* is defined as $\mathbf{S} = \mathbf{K}^T\mathbf{H}^T$

Properties of the influence matrix

- The diagonal elements S_{ii} of $\mathbf{S}$ are the *self-sensitivities*; the off-diagonal elements S_{ij} $(i \neq j)$ are the *cross-sensitivities*.
- If $\mathbf{R}$ is diagonal (uncorrelated observation errors) then $0 \leq S_{ii} \leq 1$.

  $S_{ii} = 0$ implies the i^{th} observation has no influence in the fit.

  $S_{ii} = 1$ implies the i^{th} observation has been fit exactly.
- The *information content* or *degrees of freedom for signal* is defined as

$$\mathrm{trace}\;(\mathbf{S}) = \sum_{i=1}^{p} S_{ii} \qquad p = \text{no. of obs.}$$

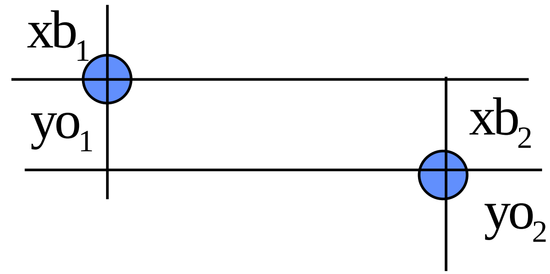
- The *Average observation Influence* (AI) is defined as

$$\text{Global} \quad \mathrm{AI} = \frac{1}{p}\sum_{i=1}^{p} S_{ii} \qquad \text{Partial} \quad \mathrm{AI} = \frac{1}{p_I}\sum_{i\in I}^{p_I} S_{ii} \qquad p_I = \text{subset of obs.}$$

- Methods that compute the *reduction of error variance* due to (subsets of) observations are closely related to the above (Desroziers *et al.* 2005; QJRMS).

The influence matrix: a simple example

- Consider 2 observations (yo_1, yo_2) and 2 parameters (x_1, x_2) with background (xb_1, xb_2)



$$\mathbf{H} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{R} = \sigma_o^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{B} = \sigma_b^2 \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix}$$

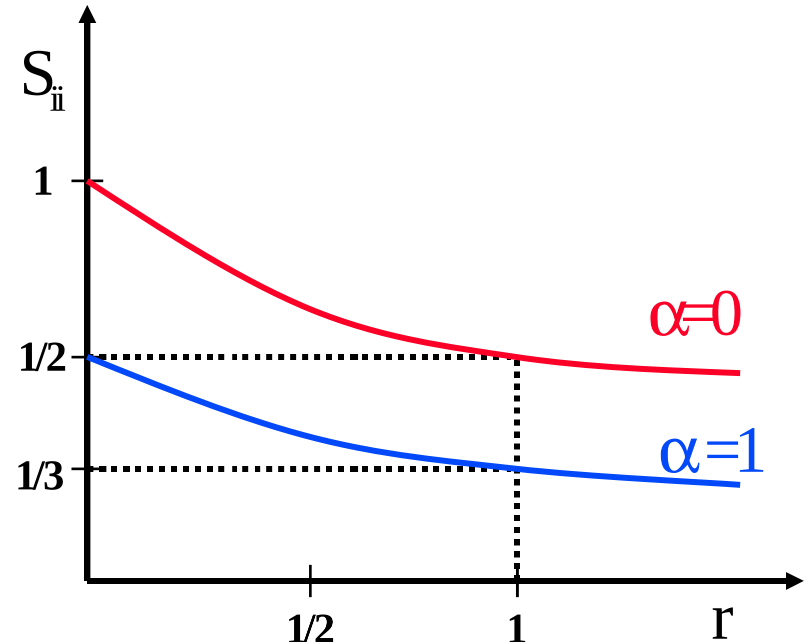
$$\mathbf{S} = \frac{1}{r^2 + 2r + 1 - \alpha^2} \begin{pmatrix} r + 1 - \alpha^2 & -\alpha r \\ -\alpha r & r + 1 - \alpha^2 \end{pmatrix}$$

where $r = \sigma_o^2 / \sigma_b^2$

The influence matrix: a simple example

$$\mathbf{S} = \frac{1}{r^2 + 2r + 1 - \alpha^2}\begin{pmatrix} r + 1 - \alpha^2 & -\alpha r \\ -\alpha r & r + 1 - \alpha^2 \end{pmatrix}$$

$$r = \sigma_o^2 / \sigma_b^2$$



- $r \rightarrow \infty$: $S_{ii}=0$ Worthless observation/perfect back
- =1, r = 1 : $S_{ii}=1/3$ Perfect correlation, back. + obs. of equal accuracy
- =1, r = 0 : $S_{ii}=1/2$ Perfect correlation, perfect obs.
- =0, r = 1 : $S_{ii}=1/2$ No correlation, back. + obs. of equal accuracy
- =0, r = 0 : $S_{ii}=1$ No correlation, perfect obs./worthless background

The important points

- High background-error correlations can significantly reduce the observation influence (self-sensitivity) in favour of background influence and surrounding observation influence.
- For observations that are sparse (relative to the background-error correlation), the observation influence (self-sensitivity) is determined by their accuracy relative to the background.

Computing the self-sensitivities

- To compute the self-sensitivities requires estimating the analysis error variances (the diagonal of the inverse of the Hessian matrix).
- For assimilation methods based on reduced-rank covariance formulations, this computation should be straightforward.
- For full-rank covariance formulations used with variational methods, the computation is more complex.

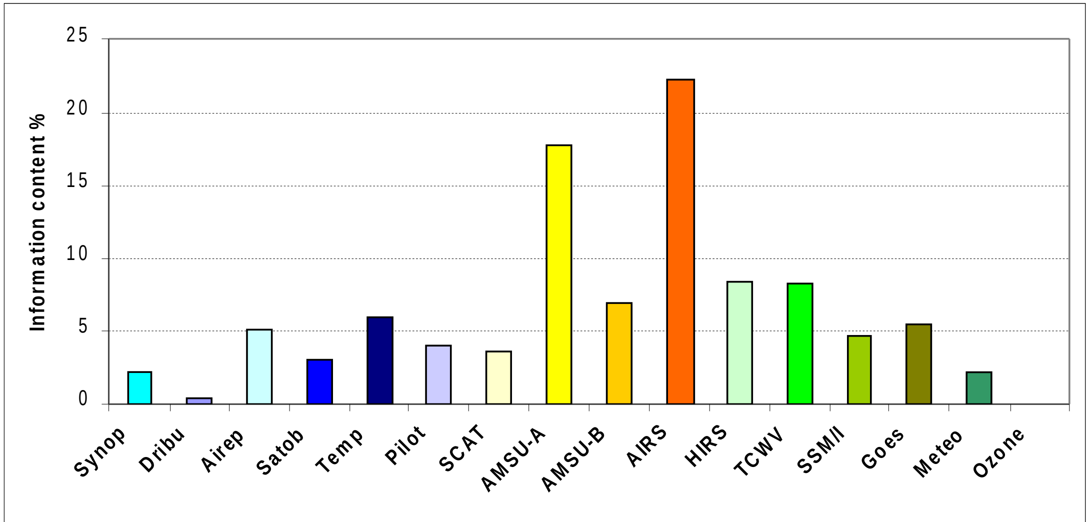
    Approximate methods have been developed using a combined randomization and Lanczos algorithm (Cardinali *et al.* 2004; QJRMS).

    Randomization methods are efficient if only the trace (GAI or PAI) is desired (Desrozier *et al.* 2005; QJRMS).

Example of analysis sensitivity in the ECMWF 4D-Var system

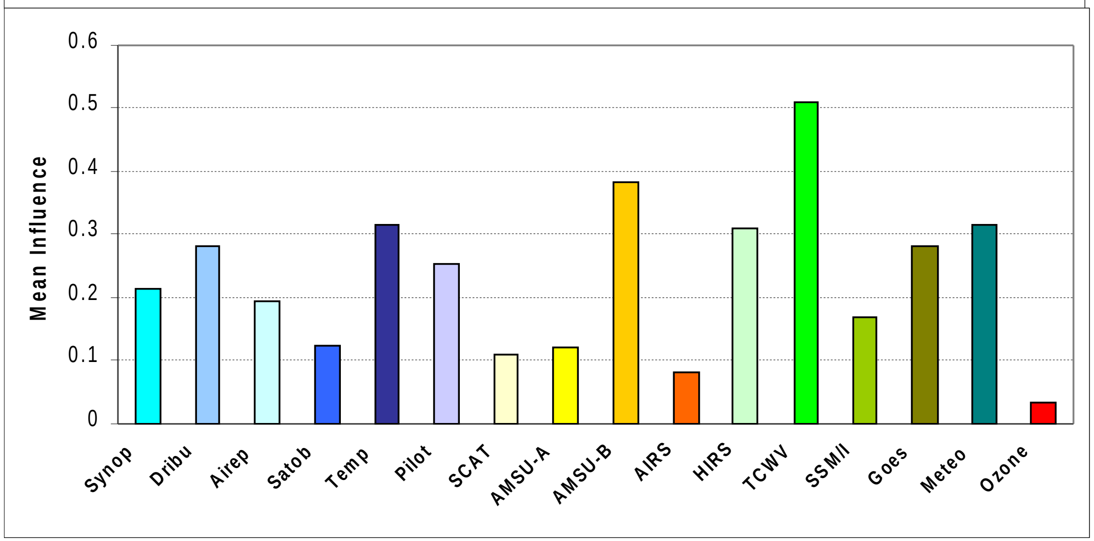
From the 2003 ECMWF operational system (Cardinali 2004)

Global Average Observation Influence: GAI=15%



Information content (%)

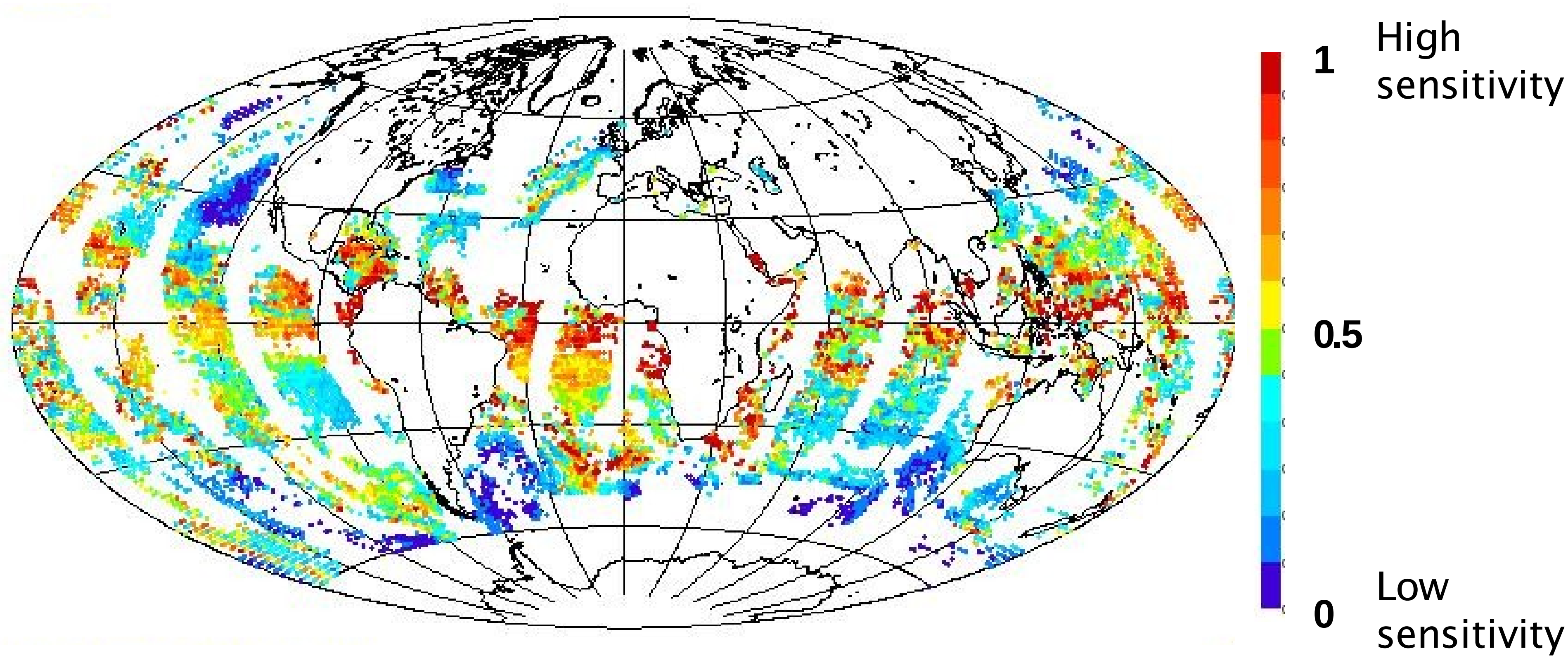
Global Average Background Influence: 100%-GAI=85%



Mean influence

Example of analysis sensitivity in the ECMWF 4D-Var system

Analysis sensitivity to Total Column Water Vapour (TCWV) observations



(Cardinali 2004)

Forecast sensitivity to observations
(e.g., Langland and Baker, 2004; Tellus)

- Define a scalar measure of forecast error (a cost function *J*); e.g., the "energy" of the difference between forecast and verifying analysis in a given region.
- The forecast-error *sensitivity* with respect to the *analysis* is

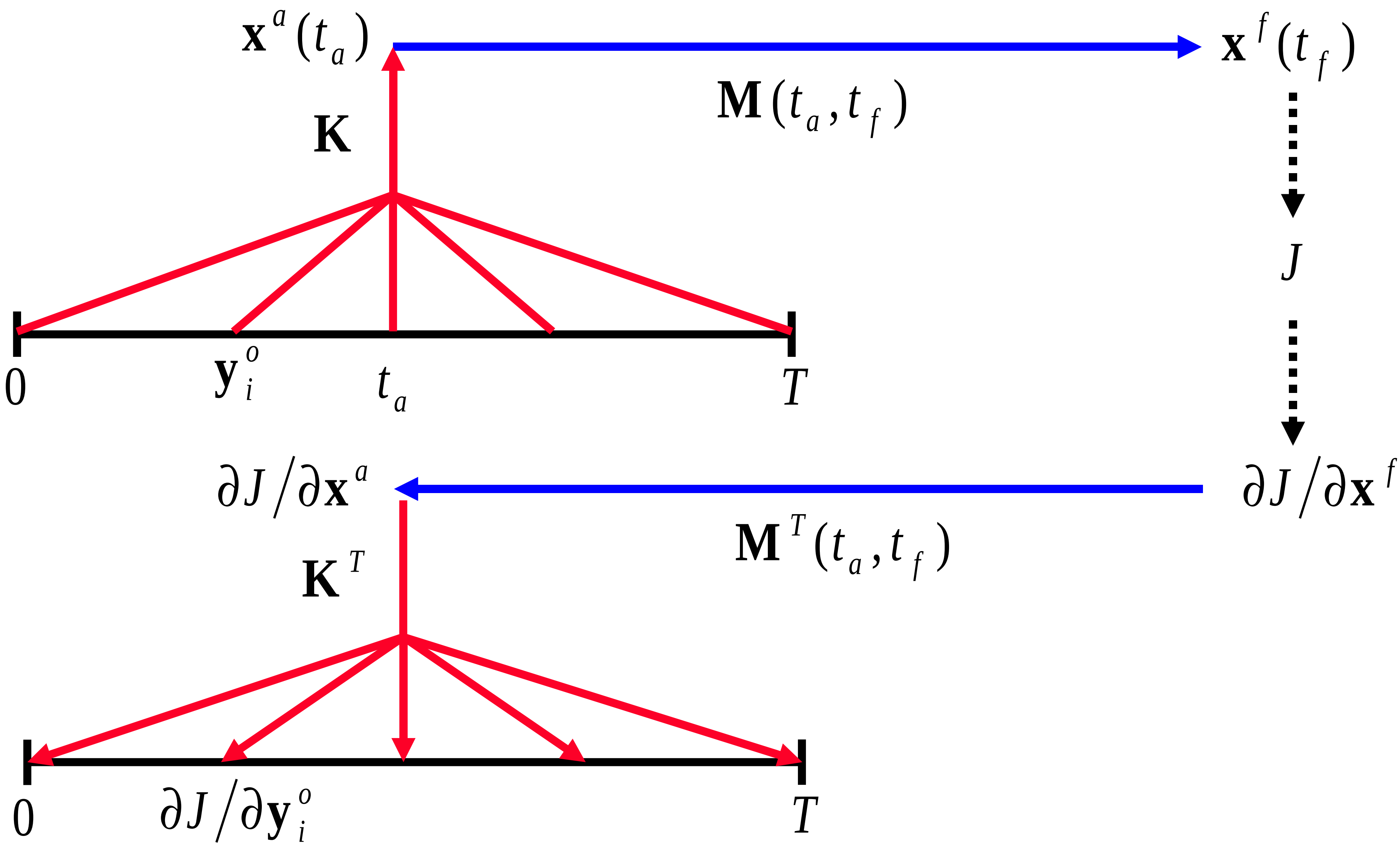
$$\frac{\partial J}{\partial \mathbf{x}^a} = \frac{\partial \mathbf{x}^f}{\partial \mathbf{x}^a} \frac{\partial J}{\partial \mathbf{x}^f} = \mathbf{M}^T \frac{\partial J}{\partial \mathbf{x}^f}$$

- The forecast-error *sensitivity* with respect to the *observations* is

$$\frac{\partial J}{\partial \mathbf{y}^o} = \frac{\partial \mathbf{x}^a}{\partial \mathbf{y}^o} \frac{\partial J}{\partial \mathbf{x}^a} = \mathbf{K}^T \frac{\partial J}{\partial \mathbf{x}^a}$$

- The forecast-error sensitivity can thus be computed using the *adjoint of the forecast model* and the *adjoint of the data assimilation method (gain matrix)*.

Forecast sensitivity to observations



How to compute $\mathbf{K}^T \partial J / \partial \mathbf{x}^a$ in Var systems?

$$\mathbf{K}^T = (\mathbf{HBH} + \mathbf{R})^{-1}\mathbf{HB}$$ observation-space (PSAS) inversion

$$= \mathbf{R}^{-1}\mathbf{H}(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}$$ model-space (Var) inversion

$$\mathbf{A}$$

- In operational Var systems, $\mathbf{A}$ is ~10^6x10^6 and $\mathbf{B}$, $\mathbf{R}$ and $\mathbf{H}$ are only available in operator-form, so direct inversion of $\mathbf{A}$ is not possible.
- The linear system $\mathbf{A}^{-1}\mathbf{z} = \partial J / \partial \mathbf{x}^a$

  can be solved efficiently using a conjugate gradient method.
- Note that the variational analysis itself is obtained by solving the same linear system but with a different right-hand side.
- Then compute $\partial J / \partial \mathbf{y}^o = \mathbf{R}^{-1}\mathbf{H}\,\mathbf{z}$
- An analogous procedure exists for observation-space formulations.

How does J vary with respect to the observations?

- We can use the adjoint sensitivity gradient to estimate the (first-order) observation impact on forecast error:

$$\delta J = \left(\partial J/\partial \mathbf{x}^f\right)^T \delta\mathbf{x}^f = \left(\partial J/\partial \mathbf{x}^f\right)^T \mathbf{M}(t_a, t_f)\, \delta\mathbf{x}^a \quad \text{in model space}$$

$$= \left(\mathbf{M}(t_a, t_f)^T\, \partial J/\partial \mathbf{x}^f\right)^T \delta\mathbf{x}^a = \left(\partial J/\partial \mathbf{x}^a\right)^T (\mathbf{x}^a - \mathbf{x}^b)$$

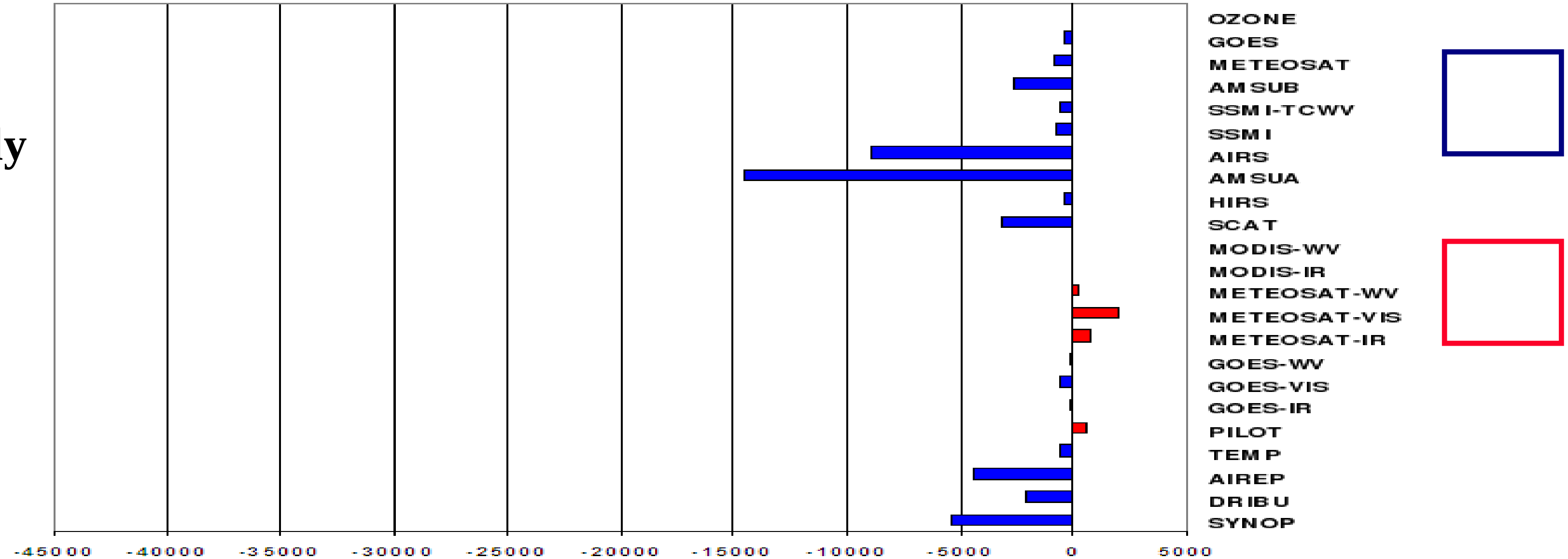
$$= \left(\partial J/\partial \mathbf{x}^a\right)^T \mathbf{K}(\mathbf{y}^o - \mathbf{H}\mathbf{x}^b) = \left(\mathbf{K}^T \partial J/\partial \mathbf{x}^a\right)^T (\mathbf{y}^o - \mathbf{H}\mathbf{x}^b)$$

$$= \left(\partial J/\partial \mathbf{y}^o\right)^T (\mathbf{y}^o - \mathbf{H}\mathbf{x}^b) = \left(\partial J/\partial \mathbf{y}^o\right)^T \delta\mathbf{y}^o \quad \text{in observation space}$$

- We can use the last expression to examine the contributions to δJ from subsets of the observations (innovations).

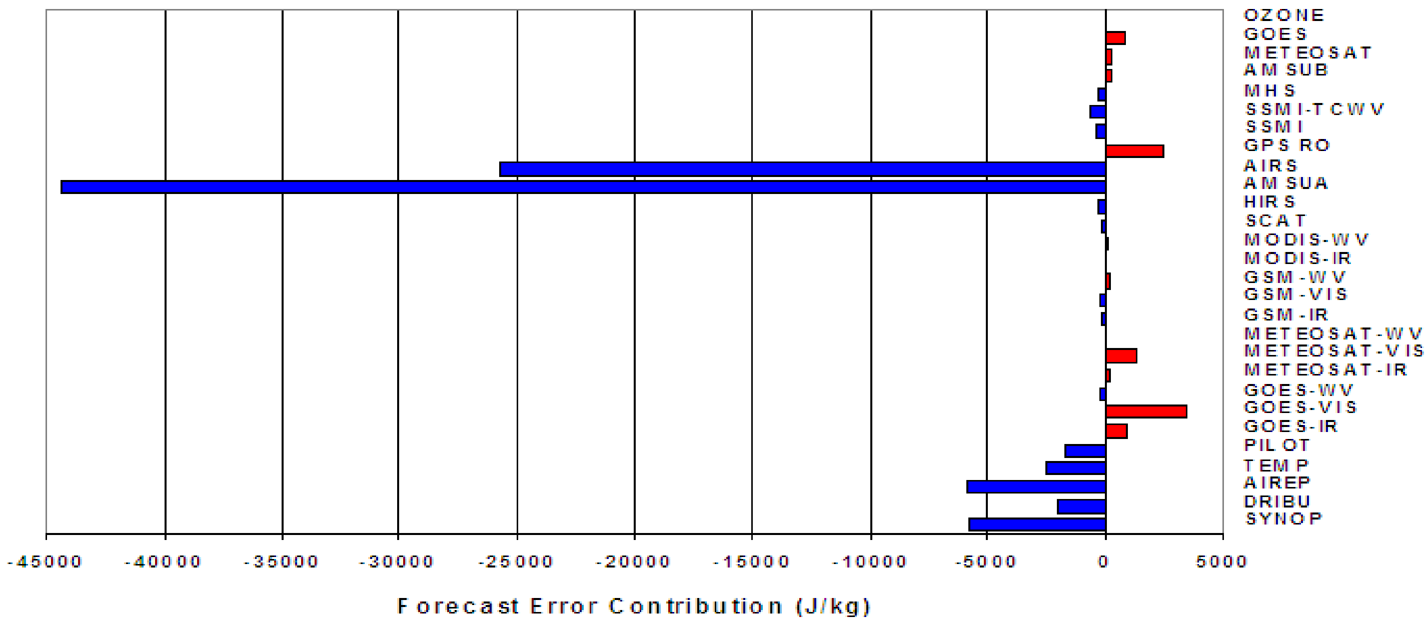
Monitoring forecast sensitivity in the ECMWF 4D-Var System

15 June - 15 July
Summer 2006



24h OSE FcE
Cycle 31R2
T511T95T159
L60

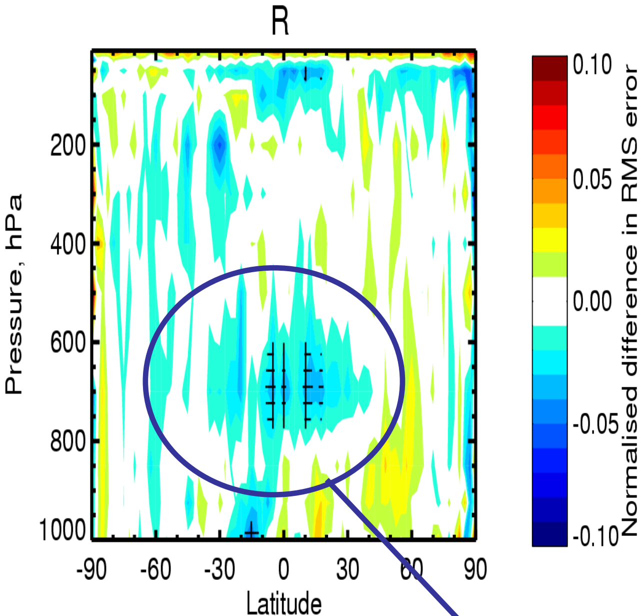
5 January -
12 February
Winter 2007



Example of 10-day forecast sensitivity in the ECMWF 4D-Var system

Traditional OSE

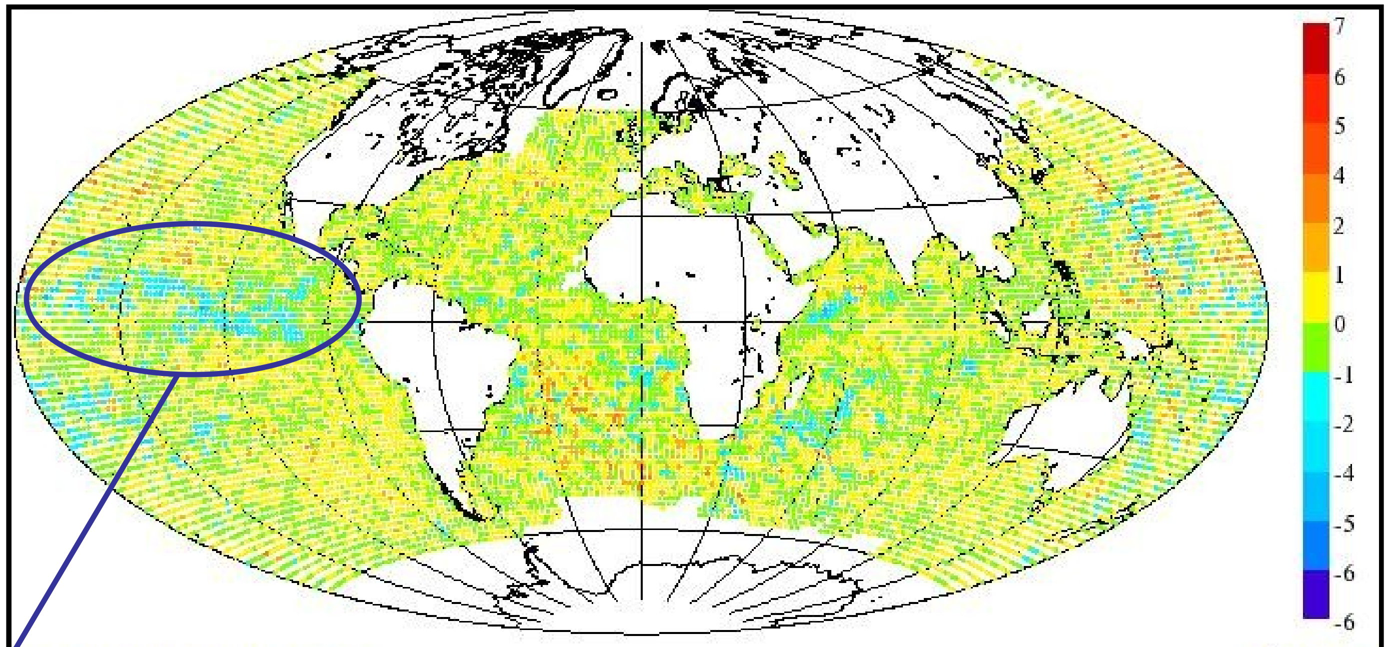
RAIN – NORAIN impact on relative humidity forecast (T + 48h) error



(Geer et al. 2007)

TCWV observation impact on forecast error

δJ



[J / kg]

(Cardinali 2007)

OSE and sensitivity experiment show consistency

Conclusions (1)

- The information matrix provides quantitative information about the impact of observations on analyses generated by statistical data assimilation methods.
- Important applications of the self-sensitivities (diagonal elements) of the information matrix include:

  Monitoring the influence of different data-sets/types.
  Identifying problems in the assimilation system (**B** and **R** specification).
  Quality control (effect of leaving one observation out).
  Data thinning.
- Practical methods have been developed in NWP to estimate the self-sensitivities (individual elements or trace).

  Randomization methods (Desroziers *et al.* 2005; QJRMS).
  Lanczos methods (Cardinali *et al.* 2004; QJRMS).

Conclusions (2)

- Practical methods for estimating forecast sensitivity to observations have been developed in NWP. The fundamental ingredients are:

  the adjoint of the forecast model (also needed for 4D-Var);

  the adjoint of the data assimilation method – gain matrix (also needed for the information matrix).

- Forecast error is defined precisely via a cost function.

  Different cost functions will give different sensitivities.

- In NWP, forecast sensitivity experiments have been used to identify model-bias and data-quality problems.

- The techniques described here are complementary to traditional OSEs/OSSEs which are more appropriate for longer range forecasts (when the linear assumption breaks down) and testing impact on forecasts (rather than forecast error).