VARIATIONAL ANALYSIS:
USE OF OBSERVATIONS - EXAMPLE OF CLEAR RADIANCES

J. Pailleux
European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, U.K.

1. INTRODUCTION

A comprehensive and practical mathematical approach of the 4 dimensional assimilation problem is presented in this volume: see Talagrand (1988). This approach uses variational techniques which minimize the distance between the model and the observations: it is called "4D-VAR" (4 dimensional variational) assimilation. The distance between the model and the observations is called "cost function" in this paper (it is sometimes called "misfit" function or "penalty" function). The technique relies on the notion of an adjoint operator which provides a very convenient tool for computing the gradient of the cost function with respect to the variables of the forecast model X (X is the "control variable" of the variational problem). Some specific aspects of the 4D-VAR problem and some simple numerical experiments are given in this volume: Courtier (1988). Both the theory and numerical results of a 4D assimilation using a vorticity equation model are fully documented in Talagrand and Courtier (1987), and in Courtier and Talagrand (1987). Several experiments made with a simple model are also described in Lorenc (1988).

In practice the variational assimilation consists in computing a cost function J(X) and its gradient with respect to X, X being the vector containing all the model variables at the initial time of the assimilation period. J(X) and Grad J are then passed to a standard minimization algorithm. The total procedure is repeated several times until an appropriate convergence is reached. Assuming the existence of a forecast model, its adjoint and an appropriate minimization scheme, the full 4D VAR problem is reduced to a 3 dimensional problem: computation of the cost function and its gradient from the observations made at the same time as the one when the model X is available.

J(X) includes not only the distance of the model X to the observations J_o, but also the distance to the first-guess J_g (the most recent 6 hour forecast) which is a similar source of information. J might include also some extra terms J_c representing physical constraints on X (Examples: geostrophic assumption in the form of a weak constraint; distance to the slow manifold).

$$J = J_o + J_g + J_c$$

To reproduce the scientific assumptions of the multivariate Optimum Interpolation (OI) analyses, J_o and J_g must be quadratic forms with O being the covariance matrix of the observation errors, and P the covariance matrix of the prediction errors (first guess errors), with a geostrophic assumption on the increments $X-X_g$:

$$J_o = (d - HX)^t \, O^{-1} \, (d-Hx) \tag{1}$$

$$J_g = (X-X_g)^t \, P^{-1} \, (X-X_g) \tag{2}$$

X_g is the first-guess (fully homogeneous to X), d the data vector, H the "postprocessing operator" which computes the equivalent of the data vector d from the model variables X: see also Talagrand (1988).

The purpose of the present paper is to discuss the problems related to the computations of J_o, J_g, and their gradients, in the context of the design and the construction of a full 3D VAR analysis (which can be transformed to a full 4D VAR system as soon as we interface it with a forecast model and its adjoint). More details are given on the observation term J_o, and also comparisons are made with the use of observations in a traditional OI analysis. The example of satellite data is highlighted and it is shown that radiances can be naturally used in a variational analysis through the radiative transfer equation and its adjoint. The control of gravity waves is addressed in another paper of the present volume: see Courtier (1988).

2. EVALUATION OF THE FIRST GUESS COST FUNCTION AND ITS GRADIENT: GENERAL PRINCIPLES

The term J_g given in equation (2) is a quadratic form built on the vector $X-X_g$ (departures from the first guess), and on the matrix P. The vector $X-X_g$ is easy to evaluate. The main practical problem comes from the inversion of the matrix P which has a size equal to the number of model variables. Let us note that it is not absolutely necessary to compute J_g in grid-point space. If the forecast model is spectral, then X is in a series of spectral coefficients, and J_g can in principle be evaluated in spectral space as well as in any other convenient space (normal mode space, gaussian grid, etc.).

The general strategy for computing J_g consists in setting up a covariance model (let us call it "structure function" model) which:

- Is simple enough to make the computation of P^{-1} feasible;

- Is "reasonable" enough so that it fits our experimental knowledge of the forecast error covariances (knowledge which can be extracted from observations).

A very simple structure function model can be based on the following assumptions: for each 3D field contained in $X-X_g$ (such as divergence, vorticity, temperature..), the correlation is homogeneous on the sphere. The vertical correlation is then the same everywhere on the globe, and the horizontal correlation is isotropic with the same scale everywhere on the globe. No restrictive assumption is made on the standard deviation of the forecast error which can be represented by fields. These fields are usually derived in an empirical way in the current operational OI assimilation schemes.

Based on these assumptions the computation of J_g can be done through the following steps:

- In grid space, $X-X_g$ is normalized by the forecast error standard deviation;

- The correlation matrix corresponding to P is split into submatrices corresponding to different horizontal fields;

- The different correlation submatrices are then diagonal in spectral space and the corresponding parts of J_g are easy to compute.

This example of a structure function model is simple and easy to implement. More complicated models will have to be developed to fit some experimental features of the forecast error covariances better, such as different correlation scales in the tropics and in mid-latitudes.

When a scheme has been worked out to compute the cost function $J_g = (X-X_g)^t P^{-1} (X-X_g)$, the computation of the gradient J_g with respect to X is straightforward:

$$\mathrm{Grad}_X J_g = 2 P^{-1} (X-X_g)$$

3. EVALUATION OF THE OBSERVATION COST FUNCTION AND ITS GRADIENT

3.1 General principles

J_o is the quadratic form (1) built on the matrix O, the covariance matrix of the observation errors for all the data. "Observation error" has the same meaning as in the OI context, i.e. it includes also the "representativeness error". It is reasonable to assume that the observation errors associated to two different types of observation are uncorrelated, and this assumption has been made in all the operational analysis systems so far. For the pure observation error this assumption is fully justified. It is not fully justified but still reasonable for the representativenesss error. Then O will be a block-diagonal matrix if the observations are stratified by observation type and the cost-function J_o can be split in independent terms corresponding to the different observation types;

$$J_o = J_{synop} + J_{rs} + J_{airep} + J_{satem} + \ldots..$$

For one individual observation type the cost function can generally be split into independent contributions. It is reasonable to assume, for example, that the observation errors of different radiosondes are not correlated. Then the contribution to J_{rs} of each radiosonde can be computed independently:

$$J_{rs} = J_{rs1} + J_{rs2} + J_{rs3} + \ldots..$$

For each individual radiosonde, the contribution to the cost function and its gradient can be computed separately for the following parameters, as the observation errors associated with these parameters can be assumed to be uncorrelated:

- wind data;
- geopotential (or temperature, or both);
- relative humidity.

$$J_{rs1} = J_{wind} + J_{zt} + J_{rh}$$

Each term such as J_{wind} is a simple quadratic form built on the vertical covariance matrix of the observation error.

All the separations made on the computation of J_o are of course valid for the gradient of J_o with respect to the model variables, Grad J_o:

$$\text{Grad } J_o = \text{Grad } J_{synop} + \text{Grad } J_{rs} + \text{Grad } J_{airep} + \ldots..$$

$$\text{Grad } J_{rs} = \text{Grad } J_{rs1} + \text{Grad } J_{rs2} + \text{Grad } J_{rs3} + \ldots..$$

$$\text{Grad } J_{rs1} = \text{Grad } J_{wind} + \text{Grad } J_{zt} + \text{Grad }_{rh}$$

In the case of surface observations (SYNOPs) the decorrelation assumption is probably reasonable for surface pressure observation errors, but for 2 m temperatures it is likely that the representativeness error is larger than the instrument observation error. This representativeness error includes not only the local features, but also the error of the post-processing operator H (used to compute the predicted 2 m temperature from the model variables X). We can expect the representativeness error between two SYNOP temperatures to be highly correlated in some weather situations such as winter inversions over cold land surfaces (which cannot be described accurately by the forecast model).

The satellite data observation errors are often intercorrelated and the corresponding cost function computation then has to handle non-diagonal matrices (see Section 4).

3.2 <u>How to link the model to the observed data</u>

To evaluate the cost function J_o, we need first to compute the departures "observation-model" (d-HX). The quantity d_{mod} = HX is the "post-processed" data vector coming from the model, which can be compared to the observed data vector d. If the observations were providing directly all or a part of the model variables X, H would be a simple identity or projection operator. However, except on very rare occasions, no model variable is observed directly, and d_{mod} has to be computed using a more complicated post-processing operator H.

If the forecast model is spectral, X will contain a set of spectral coefficients and H will contain the inverse spectral transforms. Although it is possible and more exact to compute the model variables at each observation point directly from the spectral coefficients, it is faster in practice to compute the values on the gaussian grid. H contains then a horizontal integration from a grid to an observation point: the simplest one is the bilinear interpolation from the four nearest grid points. The horizontal part of the operator H is then common to all observation types.

In the vertical, depending on the observation type, H contains various mathematical operators such as:

- Ordinary vertical interpolations (e.g. interpolation of the model wind profile to the pressure level of an observed wind);

- Integrations involving both temperature T and specific humidity q profiles (e.g.: computation of a geopotential height from model variables, to be compared with a radiosonde datum;

- Radiative transfer computations for radiance data. (see Section 4).

Let us note that the vertical part of the different operators H is not necessarily linear. The possibility of using in a consistent way data which are linked to the model variables through non linear operators is actually a major advantage of the variational analysis (see also Lorenc (1988)). Any data d can be used provided they can be related to the model variables X by a mathematical operator H which is well defined, accurate enough and "differentiable". The gradient computations involving H must be physically meaningful, it is why a kind of differentiability property is needed.

3.3 Use of the adjoint of the operator H for the gradient computations

H being the operator computing d_{mod} from X, its adjoint H* is by definition the operator computing $Grad_X J$ from $Grad_{d_{mod}} J$, for any function J.

For the variational analysis, we need to compute the gradient of the distance to the observations J_o with respect to the model variables X.

$$Grad_X J_o = H^*(Grad_{d_{mod}} J_o)$$

If the operator H is decomposed in a series of simple operators H_1, H_2, ... H_n ($H = H_n \ldots . H_2.H_1$), the chain rule allows the computation of H*:

$$H^* = H_1^*.H_2^* \ldots .H_n^*$$

For one given observation, the post-processing operator H can be decomposed in a horizontal operator H_h and a vertical one H_v:

$$H = H_v.H_h$$

Then the gradient of the contribution J of one observation (or a subset of observations) to the cost function can be computed through the following steps:

- computation of the gradient with respect to the vector d-HX corresponding to one observation:

  As $J(X) = (d-HX)^t\ O^{-1}\ (d-HX)$,

  then $Grad_{d-HX} J = 2\ O^{-1}\ (d-HX)$

- computation of the gradient with respect to $d_{mod} = HX$:

  $Grad_{d_{mod}} J = - Grad_{d-HX} J$

- computation of the gradient with respect to $H_h X$:
  $H_h X$ is the profile of model variables at the observation point

  $$\mathrm{Grad}_{H_h X} J = H_v^* (\mathrm{Grad}_{HX} J)$$

- computation of the gradient with respect to X:
  we just multiply the gradient with respect to $H_h X$ by the ad hoc coefficients of the horizontal bilinear interpolation H_h in order to get the contribution to the gradient with respect to X at the four nearest grid points. This simple coefficient multiplication is indeed the adjoint of H_h. These contributions are then added to the general variable $\mathrm{Grad}_X J$.

3.4 A simple example of the non linear operator H: use of 2 metre observations

The 2 metre temperature and relative humidity are observed with a good instrumental accuracy at all the surface SYNOP meteorological stations. However, only a very limited use of these data is made in the current ECMWF analysis scheme (not used at all in the mass and wind analysis), because they are difficult to link with the model prognostic variables in the OI context.

The operators H associated to both the 2 m temperature observation (T^{2m}) and the 2 m relative humidity (RH^{2m}) have been studied by Vasiljevic (pers. comm.), as well as the gradient with respect to the model variables of the corresponding cost functions:

$$J_o = \left(\frac{T^{2m} - T^{2m}_{mod}}{\sigma_o(T)}\right)^2 \text{ for } T^{2m} \;;$$

$$= \left(\frac{RH^{2m} - RH^{2m}_{mod}}{\sigma_o(RH)}\right)^2 \text{ for } RH^{2m}.$$

a) For the temperature, the ECMWF operational post-processing operator H involves the model variables at the surface and at the first model level (about 30 m in the ECMWF model). It is not linear, it relies on the flux computations and on the Charnock formula and it is complicated to express analytically. The gradient computations ($Grad_X J_o$) are almost impossible to handle analytically. The only practical way seems to code the adjoint operator by applying the chain rule (see 3.3).

b) For the relative humidity, the ECMWF operational post processing operator H is simpler:

$$RH^{2m}_{mod} = \frac{q_1}{q_s(p_s, T_1)}$$

The relative humidity at 2 metres is just assumed to be equal to the one obtained at the first model level, dividing the specific humidity Q_1 by the saturated specific humidity q_s. q_s can be expressed as a function of two other model variables p_s and T_1. q_s contains an exponential function and is not linear, so H is also non linear.

As H is simple enough in this case, the gradient computations can be made analytically. RH^{2m}_{mod} is a function of only three model variables: p_s (surface pressure), T_1 and q_1 (temperature and specific humidity at the first model level). So only $\frac{\partial J_o}{\partial p_s}$, $\frac{\partial J_o}{\partial T_1}$ and $\frac{\partial J_o}{\partial q_1}$ are not equal to zero among the components of $Grad_X J_o$.

The computation of these gradient components about different basic states (p_s, T_1, q_1) confirms the non-linearity: the values of the components are highly sensitive to the basic state. It confirms also that $\frac{\partial J_o}{\partial q_1}$ is large compared to $\frac{\partial J_o}{\partial T_1}$, which is itself large compared to $\frac{\partial J_o}{\partial p_s}$. This simply reflects that when we compute RH^{2m}_{mod}, most of the information is coming from q_1 and very little is coming from p_s.

However, we can see in this particular example that there is no distinction between "mass and wind" analysis and "relative humidity" analysis in the variational scheme. We have the model variables X on one side, and the observations on the other side. Each observation contributes to the cost function J_o and to the gradient $Grad_X J_o$. Its contribution to $Grad_x J_o$ has a non-zero component with respect to all the model variables which are used in the post-processing operator. Then a 2 m relative humidity observation contributes to "update" q_1, but also the mass variables T_1 and p_s (although this last update is very small). And of course, though the term J_g, RH^{2m} will contribute also to update indirectly other model variables (e.g. T and q at other model levels).

Another remark is that the variational framework is flexible enough in order to allow a consistent use of 2 metre observations. It is consistent with the post processing operators of the model and it avoids for example the need for a vertical correlation (of forecast errors) between the 2 metre level and the other model levels.

4. USE OF SATELLITE DATA IN THE VARIATIONAL ANALYSIS

4.1 Interface between satellite data and variational analysis

In the ECMWF operational analysis and in most of the operational systems, the information obtained from the polar orbiting satellites is used as SATEM retrieved soundings in the OI analysis. The SATEM observations contain both thickness data and Precipitable Water Content (PWC) data. For more details about the interface between the SATEMs and the current ECMWF analysis system, see Pailleux (1986) and Kelly and Pailleux (1988); see also Pailleux et al. (1988), a companion paper in the present volume.

However, the SATEMs are not the genuine observed quantities, but an "interface" produced by a specific "retrieval" technique. An important part of the SATEM information is coming from the retrieval technique (an inversion algorithm and possibly initial profile to start the inversion with). The retrieval technique is also probably responsible for a large part of the correlation between observation errors of different SATEM profiles.

In the variational context, the natural way is to avoid interfaces such as SATEMs and to use data which are as close as possible to the observed quantity. For satellite data the real observed quantities are raw radiances. However raw radiances are strongly affected by clouds, so their corresponding operator H would involve the model clouds to a large extent, and the cloud quality of the current models is not good enough to rely on them. One safe solution consists in trying to use clear radiances in the variational analysis, after an appropriate cloud clearing algorithm which does not involve the model variables X at all.

Eyre (1987) chose a different approach in which the cloud-clearing algorithm is integrated with the retrieval procedure in a single variational problem. This approach is in principle better, especially if it is generalized to the 3D and 4D context, as better clouds mean better cleared radiances and better analyses. If the clouds are "over-contained" in such a minimization scheme (kept close to a preliminary evaluation), Eyre's approach is almost equivalent to the approach using clear radiances.

The radiances are another good example of data which are linked to the model variables through an operator H_v which is not linear and rather complicated: H_v is the radiative transfer equation. And radiances can still be used in the variational analysis, as H_v satisfies the requirements specified at the end of Section 3.2.

4.2 Computation of J_o and its gradient for clear radiances

H_v is a direct radiative transfer model T_r tuned to compute the radiances of the different channels available on the satellite instruments (mainly TOVS for the moment). We apply the adjoint T_r^* of the radiative transfer equation, to derive the gradient with respect to the model quantities from the gradient with respect to the radiances: see Moll (1988) and also Betout (1988) in these proceedings. The gradient computations have also been carried out by Eyre (1987) who used a simplification of the analytical computation. As J_o is expressed in the radiance space, by the quadratic form $(R-R_{mod})^t\, O^{-1}\, (R-R_{mod})$, the computation of J_o requires the derivation of some terms of the jacobian matrix $\left(\left(\frac{\partial R_i}{\partial T_j}\right)\right)$, i being a channel number and j a level number (level used in the direct radiative transfer model). Table 1 shows a few terms of this jacobian matrix, computed from the two approaches, Moll's approach consisting in coding T_r^* (PM) and Eyre's approach performing an analytical evaluation (JE). The results are displayed for the derivatives of HIRS channels 2 to 4 with respect to the 50 hPa temperature, and for three different reference profiles: the standard atmosphere profile, a mid-latitude oceanic profile

		$\frac{\partial HIRS2}{\partial T50}$	$\frac{\partial HIRS3}{\partial T50}$	$\frac{\partial HIRS4}{\partial T50}$
Standard atmosphere	JE	0.1109	0.0901	0.0444
	PM	0.1151	0.0943	0.0426
Oceanic mid-latitudes profile	JE	0.1111	0.0908	0.0448
	PM	0.1168	0.0946	0.0440
Cold Scandinavian profile with stratospheric vortex	JE	0.0916	0.0711	0.0321
	PM	0.1133	0.0853	0.0325

Table 1: One example of gradient computations involving the radiative transfer equation.

The computations are made by two approaches:

- <u>JE</u>: John Eyre's approach
- <u>PM</u>: Patrick Moll's approach
  (see text)

(taken from a radiosonde observation of weather ship P), and an "abnormal profile" (taken from a radiosonde observation made in Scandinavia) characterized by a cold air mass and a strong stratospheric vortex.

The comparisons between the two approaches show generally a good agreement. The results show also the impact of the profile dependency on the gradient (non linearity of the radiative transfer equation). Especially for the cold Scandinavian air mass the gradients $\frac{\partial R_i}{\partial T_j}$ are smaller than for the two other profiles. These gradients can be interpreted as a kind of objective estimate of the amount of information provided by the radiances to the temperature analysis. This amount of information is situation dependent, and it is a potential advantage for the variational analysis compared to the current OI, that it is able to take this situation-dependency into account. In other words, because the radiative transfer equation is non-linear, the use of this variational technique should be a significant improvement compared to the OI technique: the main reason is that the gradient computations will be made about the genuine model profile; only profile-dependent statistics (with an infinity of profile types!) could achieve a similar description of the link between radiances and model variables. Most of the operational OI analysis systems are using SATEMs with "fixed" statistics for SATEM errors, then the weight given to the SATEM is not dependent on the situation.

4.3 Specific problems in the computation of J_o for satellite data

The radiance contribution to the cost function J_o is the quadratic form $(R-R_{mod})^t\ O^{-1}\ (R-R_{mod})$. As O contains not only the radiance measurement error but also the error due to the cloud-clearing and the error due to the operator T_r when computing R_{mod}, then the radiance errors are correlated in the horizontal, and O is not diagonal (and might be large). A similar problem with the size of the matrix would occur if we would try to use SATEMs in the variational analysis. This problem is very similar to the one described in the computation of J_g (section 2). It can be treated in a way which is similar, at least for the first step: for one given set of satellite data O is first split in horizontal and vertical correlation, and the cost function is split in different terms J_{eigi}, each of them corresponding to one eigen vector of the vertical correlation matrix of satellite errors. "Vertical" correlation means "inter-channel" correlation in the radiance context.

$$J_o = J_{eig1} + J_{eig2} + J_{eig3} + \ldots..$$

Each term J_{eigi} has to be computed by inverting the horizontal correlation matrix. If this matrix is too large, one possibility is to split the satellite data subset according to observation time until the size is manageable in memory: the data of one orbit would be split into segments in a way similar to the one used in most of the satellite retrieval procedures.

5. CONCLUSION

Compared to the OI scheme, the 3D variational analysis is an approach which provides many potential improvements: better use of observations especially when they are linked to the model variables through a complicated operator, possibility of using more data. In addition the need of a data selection algorithm disappears as the analysis is treated in one global variational problem: this is likely to be beneficial for the 3D consistency of the analysed fields. The possibility of introducing more physical constraints through extra terms in the cost function is a very flexible framework for further developments.

When these potential improvements are confirmed by the experimentation with a full 3D variational analysis, they will justify operational implementation. In an operational context, the variational analysis would be run in addition to the traditional OI analysis (rather than in replacement). The OI analysis will be used for two main purposes:

- To provide an initial point for the minimization scheme, which is close to the final solution;

- To perform a quality control of the data.

To start the minimization from a point X close to the final solution might be a very important practical aspect, as it is likely that both the memory size devoted to the minimization algorithm, and the number of iterations will be critical factors for the operational feasibility.

For both scientific and practical reasons it might also be more convenient to run this preliminary OI analysis without satellite data. This would at least provide the framework for a better quality control of satellite data. Assuming most of the quality control is performed in the OI analysis, an extra level of quality control has still to be envisaged in the variational step for the following reasons:

- This extra quality control is necessary if we use in the variational analysis some data which do not enter the OI analysis (likely example: radiance data);

- The variational context provides a framework to perform extra quality control checks which are more difficult to set up in the traditional OI (e.g. quality control checks using a sequence of model values at the data points for different steps of the minimization algorithm).

Following the idea discussed in Gandin (1988), all the quality control information of the preliminary OI has to be kept in order to leave all possibilities open in the variational quality control step. For example, an observation rejected by the OI must still have the possibility to be used in the variational analysis.

Finally, a very important advantage of the 3D variational analysis is that it can be generalized easily to a 4D variational assimilation, just by an appropriate interface with a forecast model and its adjoint. The computer resources will drive the choice of the length of the assimilation period and of the maximum number of iterations which can be afforded for operational implementation. A 4D assimilation would be a very important step for numerical weather forecasting as it would be the first time that all the information available for the meteorological analysis is merged in one consistent step: observations, first guess, dynamics of the atmosphere through the primitive equations. All the preliminary studies performed so far indicate that this goal can be achieved.

References

Betout, P., 1988: Current work at French DMN: Kalman filter, variational analysis and satellite data impact study. Proceedings of the ECMWF seminar on data assimilation and the use of satellite data. Reading 5-9 September 1988.

Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation - Part 2: Numerical results. Quart.J.R.Met.Soc., 113, 1329-1368.

Courtier, P., 1988; Four dimensional variational assimilation: shallow water results and development of a baroclinic system. ECMWF seminar proceedings on data assimilation and the use of satellite data. Reading, 5-9 September 1988.

Eyre, J., 1987: Inversion of cloudy satellite radiances by non-linear optimal estimation: theory and simulation for TOVS. Quart.J.R.Met.Soc., December 1987.

Gandin, L., 1988: Comprehensive quality control. ECMWF seminar proceedings on data assimilation and the use of satellite data. Reading, 5-9 September 1988.

Kelly, G.A. and J. Pailleux, 1988: Use of satellite vertical sounder data in the ECMWF analysis system. ECMWF Technical Memorandum No. 143.

Lorenc, A., 1988: Optimal non-linear objective analysis. Quart.J.R.Met.Soc., 112, 1177-1194.

Moll, P., 1988; Inversion variationelle de radiances observées. Mémoire de DEA. Imprimé par Université de Paris 6 (juin 1988).

Pailleux, J., 1986: The impact of satellite data on global numerical weather prediction. Proceedings of the Dundee Summer School (August-September 1986). Remote Sensing Applications in Meteorology and Climatology. Edited by R.A. Vaughan. NATO ASI Series.

Pailleux, J., G. Kelly, J-F. Flobert and E. Andersson, 1988: Current work at ECMWF on the use of satellite data. Proceedings of the ECMWF seminar on data assimilation and the use of satellite data. Reading, 5-9 September 1988.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part 1: Theory. Quart.J.R.Met.Soc., 113, 1313-1328.

Talagrand, O., 1988; Four dimensional variational assimilation. ECMWF seminar proceedings on data assimilation and the use of satellite data. Reading, 5-9 September 1988.