USE OF RADIANCES IN 3D/4D VARIATIONAL DATA ASSIMILATION

E. Andersson, J. Pailleux*, J-N. Thépaut, J.R. Eyre,
A.P. McNally, G.A. Kelly and P. Courtier

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

*Current affiliation, Météo-France, Toulouse, France

ABSTRACT

The direct use of TOVS radiances in a three/four-dimensional variational data assimilation scheme is described. The scheme uses a fast radiative transfer model and its adjoint. Global mass, wind and humidity fields are retrieved/analysed simultaneously under certain balance constraints. This ensures horizontal consistency and controls the amount of gravity waves actually in the analysis. The scheme thus performs retrieval, analysis and initialisation in one step and thus one can achieve an optimal combination of the information contained in the radiances, the conventional data and in the background.

1. INTRODUCTION

The TOVS (TIROS-N Operational Vertical Sounder) instrument on the NOAA series of polar orbiting satellites measures radiances that are sensitive to the temperature and humidity of the atmosphere (*Smith et al.*, 1979). These data have traditionally been used in Numerical Weather Prediction (NWP) in the form of temperature and humidity profiles, obtained through an inversion process known as "retrieval". The retrieved profiles are then used together with all other observations to form an analysis which after initialisation is used as the starting point for a forecast.

This paper describes the direct use of TOVS radiances in a variational analysis scheme. In a three-dimensional formulation (3D-Var), horizontal as well as vertical consistency in the use of the radiances is ensured. Mass and wind (and humidity) are analysed simultaneously under certain balance constraints which control the level of gravity waves actually in the analysis. The scheme thus combines retrieval, analysis and initialisation in one step.

In a four-dimensional formulation (4D-Var), time consistency, as given by the evolution of the forecast model and its adjoint, acts as an additional constraint.

Section 2 presents the background to the development of a 3D/4D variational analysis scheme from the point of view of radiance data. The variational method is described briefly in section 3. In section 4 we show results from the validation of the system when applied to a simplified retrieval/analysis problem with an analytical solution. We present results from a global three-dimensional analysis of TOVS radiances in section 5 and

compare with an OI analysis which had used retrieved profiles. In section 6 we demonstrate that the four-dimensional formulation is able to infer additional information from the dynamics of the forecast model. We conclude by discussing important areas for future work.

2. BACKGROUND

The conversion from radiances to profiles of temperature and humidity is an ill-posed problem unless additional "background" information is used (*Rodgers*, 1976; *Eyre and Lorenc*, 1989). Retrieval schemes can use statistical background information (e.g. *Reale et al.*, 1986) or can use the radiances themselves to pick a background profile from a library of representative atmospheric profiles (e.g. *Chédin and Scott*, 1985). Even with very sophisticated techniques it is unavoidable that biases in the choice of background profile can contribute to the retrieval error (*Flobert et al.*, 1991).

The retrieved profiles generated by NOAA/NESDIS (National Environmental Satellite, Data and Information Service) (*Fleming et al.*, 1986), called SATEMs here, are used in many operational NWP data assimilation schemes, as for example in the Optimum Interpolation (OI) scheme at ECMWF (*Kelly and Pailleux*, 1988; *Kelly et al.*, 1991).

Impact studies have shown that SATEM data have a large positive impact on the quality of southern hemisphere forecasts and a neutral or in some cases negative impact on northern hemisphere forecasts (*Halem et al.*, 1982; *Gallimore and Johnson*, 1986; *Andersson et al.*, 1991). The difficulty in demonstrating a consistent positive impact in the northern hemisphere indicates a less than optimal use of the data. The explanation is that there are structures in the atmospheric vertical profiles that the TOVS radiances cannot see because of the broadness of their weighting functions. The SATEMs contain a statistical or climatological component, which sometimes degrade rather than improve the analyses (*Eyre*, 1987). This shows up as very systematic, air-mass dependent biases in the SATEM data (*Andersson et al.*, 1991).

Because of these problems interest in recent years has shifted towards schemes which use an NWP forecast model to provide the background information (e.g. *Eyre*, 1989). The NWP model provides a better estimate of the atmospheric state than either a library of profiles or by pure statistical information. Such schemes are operational at ECMWF (*Eyre et al.*, 1992), at U.K. Meteorological Office (*Turner et al.*, 1985) and are in pre-operational tests at NMC (National Meteorological Centre), Washington. The work of *Eyre et al.* is very closely related to the work presented in this paper. Their scheme performs the retrieval in a variational formulation profile by profile, and can be seen as a one-dimensional analysis of TOVS radiances. Much of the development involved for TOVS is common between 3D/4D-Var and the one-dimensional scheme, including the radiative

transfer model and the radiance monitoring/tuning system. To emphasize the link between the two projects, the one-dimensional scheme is known as "1D-Var".

The advantage of these variational schemes is the use of more accurate background information (a six-hour forecast, interpolated to the TOVS locations). For a one-dimensional scheme this introduces a correlation between the observation error and the background error, which may cause problems when the same six-hour forecast is used again as the background for the subsequent analysis (as discussed by *Eyre et al.*, 1992).

The problem disappears in a three or four-dimensional variational (3D/4D-Var) scheme, where retrieval and analysis are combined. Radiances are then used directly together with all other observations and the background field. The potential benefits from the direct use of radiances in a variational analysis was a strong motivation for the joint effort between ECMWF and Météo-France to develop the "IFS/Arpège" Integrated Forecasting System (*Pailleux*, 1990).

A three-dimensional scheme which combines retrieval and analysis has already been successful in an OI analysis scheme (*Durand*, 1985; *Durand and Pierrard*, 1989), but was limited by the assumption of a linear link between the radiances and the analysed quantities. A nonlinear three-dimensional variational scheme was proposed by *Hoffman* (1983) and applied to simulated TOVS data by *Hoffman and Nehrkorn* (1989). They were successful in producing three-dimensional retrievals and indicated the possibility of including conventional observations in their scheme. The scheme was computationally expensive.

Le Dimet and Talagrand (1986) suggested the adjoint technique as a means of making large variational problems more tractable. The adjoint technique makes global 3D/4D variational assimilation feasible (e.g. *Courtier and Talagrand*, 1990; *Thépaut and Courtier*, 1991). *Thépaut and Moll* (1990) developed an exact adjoint of a TOVS radiative transfer model and showed that the computational cost of variational retrievals with the adjoint technique was manageable.

In this paper we apply the adjoint technique to the variational retrieval/analysis problem in three/four dimensions. Global variational analyses at spectral resolution T63, 19 vertical levels, have been produced and compared with ECMWF OI analyses. We demonstrate the advantages in an NWP system of the 3D/4D variational approach to the assimilation of TOVS radiances and show the increased flexibility this method offers.

3. METHOD

3.1 Variational formulation

We follow the general variational approach to the assimilation of data into an NWP system (*Lorenc*, 1986; *Talagrand*, 1988) by minimizing the cost-function $J(x)$ with respect to the atmospheric state x, where $J(x)$ measures the degree of mis-fit to the observations and to the background information. If the errors involved have Gaussian distributions, then the optimal penalty function is a sum of quadratic terms:

$$J(x) = (x - x_b)^T B^{-1} (x - x_b) + [y - H(x)]^T O^{-1} [y - H(x)] \quad (1)$$

Where x_b is the background with estimated error covariance B, y represents the observations with estimated error covariance O, and H is the observation operator (or "forward" operator) which computes model equivalents of the observed quantities at the observation points. The matrix O should in addition to the observation error also include the representativeness error, i.e. the error in the forward operator.

Eq.(1) applies to a wide range of problems. It has the same form in one as well as three and four-dimensional applications. In the case of TOVS radiances, H specifically represents the radiative transfer model that calculates radiances from the state vector of the forecast model. In 4D-Var, H includes a model integration from the time of the background to the time of the observation (see *Thépaut et al.*, 1992). 3D-Var is thus, in theory as well as in practice, equivalent to a 4D-Var without model integration.

The solution in the linear case, is given by:

$$x_a - x_b = BH'^T (H'PH'^T + O)^{-1} [d - H(x_b)] \quad (2)$$

where $x_a - x_b$ represents the analysis increment (analysis minus background). H' is the Jacobian of H and contains the partial derivatives of H with respect to the elements of x. In the nonlinear case, the solution can be found through e.g. Newtonian iteration (*Rodgers*, 1976; *Eyre*, 1989) or through the adjoint technique as described by *Le Dimet and Talagrand* (1986) and applied to the retrieval problem by *Thépaut and Moll* (1990). Details particular to the IFS/Arpège implementation of the adjoint technique have been documented in *Pailleux et al.* (1991).

At the minimum, the second derivative of the cost function. $J''(x)$ represents the inverse of the analysis error covariance:

$$J''(x) = A^{-1} = B^{-1} + H'^{T}O^{-1}H' \tag{3}$$

Note that this latter expression is approximate, since terms involving the gradient of $H'(x)$ have been omitted.

3.2 Simultaneous analysis of TOVS and other observations

Use R and O for the TOVS observation operator and the observation error covariance, respectively. Similarly, use H and Q for any other observations. (H could represent the vertical integration of the hydrostatic equation to form model equivalents of radiosonde height measurements, for example.)

The expression for the analysis error (A) is (from Eq. (3)):

$$A = \frac{1}{B^{-1} + R'^{T}O^{-1}R' + H'^{T}Q^{-1}H'} \tag{4}$$

The equivalent analysis can be achieved in two steps.
First a "retrieval" step:

$$A_1 = \frac{1}{B^{-1} + R'^{T}O^{-1}R'} \tag{5}$$

Followed by an "analysis" step:

$$A_2 = \frac{1}{A_1^{-1} + H'^{T}Q^{-1}H'} \tag{6}$$

We note that $A_2 = A$, provided the appropriate retrieval error covariance matrix (A_1) is passed with every TOVS point. Although this is possible in theory, practical considerations (huge data volumes) lead to simplifications of the error statistics passed from the retrieval to the analysis step. In the ECMWF implementation of 1D-Var, for example, a set of three covariance matrices is used, one for each retrieval type: clear, partly cloudy and cloudy.

In 3D-Var, on the other hand, $\boldsymbol{O}$ is specified in terms of radiances and the appropriate error covariance in terms of the control variable ($\boldsymbol{R'^T O^{-1} R'}$) appears implicitly, complete with the correct spatial variation.

The arguments so far have assumed that the observation operators are linear. In the nonlinear case, there are the additional effects of $\boldsymbol{R}$ and $\boldsymbol{H}$ (and $\boldsymbol{R'}$ and $\boldsymbol{H'}$) changing with $\boldsymbol{x}$, during the iterative minimisation process. Analysis changes in surface pressure introduced by conventional observations would for example lead to a modified $\boldsymbol{R}$, which would affect the TOVS retrievals in the vicinity.

In principle, not only do the retrievals become more accurate during the minimisation, but also their error estimates become more accurate. This will produce a more optimal combination of information from TOVS and other data types, than is possible in the traditional approach with separate retrieval and analysis steps.

3.3 3D/4D-Var as a retrieval scheme

The retrieval algorithm of 3D/4D-Var is to a large extent based on the operational 1D-Var algorithm. It uses the same fast radiative transfer model (*Eyre*, 1991), the same observation error statistics, the same bias correction (*Eyre*, 1992) and the same quality control (*Eyre et al.*, 1992). The brightness temperatures used are the global cloud-cleared data generated by NESDIS and disseminated as the so called "120 km BUFR TOVS" data set. These data have already undergone substantial pre-processing at NESDIS (see *Smith et al.*, 1979) followed by cloud-clearing (*McMillin and Dean*, 1982; *Reale et al.*, 1986).

The scheme is "physical" in that radiative transfer calculations are carried out at every TOVS point at every iteration. It is also "interactive" with the data assimilation cycle because it uses model data as background information. It retrieves temperature and humidity simultaneously. It is nonlinear and iterative.

Cloud-cleared radiances are used in the experiments reported here, although the scheme can in principle be applied to "raw" radiances to retrieve cloud parameters like fractional cloud cover and cloud height, as demonstrated in one dimension by *Eyre* (1989). However, this dramatically increases the nonlinearity of the inversion problem (*Eyre*, 1989; *Hoffman and Nehrkorn*, 1989).

In addition to these one-dimensional aspects, 3D-Var provides substantial new possibilities from its three-dimensional formulation. Primarily:

- Horizontal consistency is ensured
- All data types are used simultaneously
- Mass-wind balance constraints are imposed

These three points act as additional constraints on the inversion process. They introduce further redundancy to the estimation problem and will reduce the possibility of retrieval errors.

Furthermore, in 4D-Var, the evolution in time of the background error covariances are taken into account implicitly through the dynamics of the forecast model and its adjoint.

3.4 TOVS observation operator

The computation of the TOVS observation cost function is organized like that for conventional data (*Vasiljevic et al.*, 1992), with the addition that, for TOVS, horizontal observation error correlations are taken into account (*Pailleux*, 1989; 1992). However, data from different satellites are assumed to be uncorrelated. Data from different retrieval types are also assumed to be uncorrelated. Thus it has been possible to separate the observations in several de-correlated sets. With large numbers of TOVS it becomes necessary to split the sets further. Technical reasons limit the number of members in each set to a couple of hundred (currently we use 192). The radiance observation error correlation function is specified to be Gaussian with a length scale of 350 km.

The forward operator $\boldsymbol{H}$, is the product of all the operations necessary to go from the control variable $\boldsymbol{x}$ to model radiances at observation points. The operator $\boldsymbol{H}$ is continuous in $\boldsymbol{x}$. It may be linear or nonlinear and it ought to be differentiable in general but it does not have to be differentiable for *all* values of $\boldsymbol{x}$. Linear interpolation is for example differentiable between model levels but not differentiable exactly *at* a model level.

The chain of operators in the TOVS forward and adjoint calculations is shown schematically in Fig. 1. It starts with a change of variable from the control variable to model spectral variables (*Heckley et al.*, 1992), followed by the inverse spectral transforms to obtain grid-point data of temperature and specific humidity on the model's Gaussian grid. The model grid-point data are then interpolated in the horizontal with a 12-point bi-cubic interpolator and in the vertical to 40 pressure levels, assuming linearity of $\boldsymbol{T}$ and $\boldsymbol{q}$ in $\boldsymbol{p}$. The radiative transfer model is formulated in terms of $\boldsymbol{T}$ and $\ln \boldsymbol{q}$ on 40 fixed pressure levels from 1000 to 0.1 hPa. The calculations are carried out from $\boldsymbol{p} = \boldsymbol{p}(z_m)$ to 0.1 hPa, where z_m is the elevation of the TOVS locations, as given in the

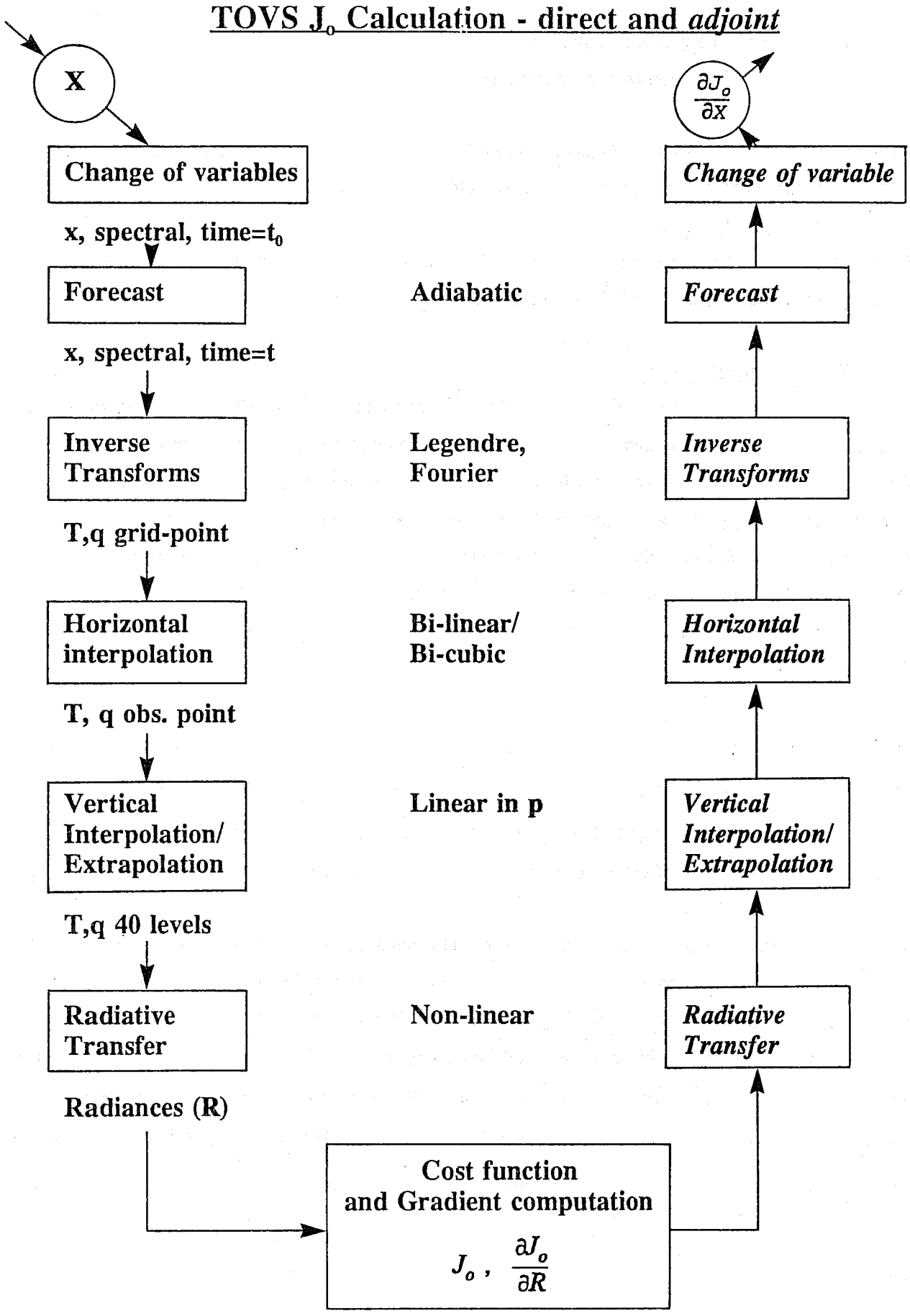

Fig.1 Schematic of the calculation of the observation cost function (J_o) for TOVS radiances - direct (on the left) and adjoint (on the right).

TOVS reports. The radiative transfer model also uses T_s and T_2m. As the control variable is currently limited to model level quantities we must either extrapolate the available information to 0.1 hPa or bring in auxiliary information from an external source. We have chosen to bring in the 1D-Var retrieved temperature at the surface (T_s) and in the stratosphere above the top of the model (7.3 hPa). We use the temperature at the lowest model level in place of T_2m. The reasoning behind this choice will be discussed in section 4.4.

For humidity we discard the model variables above 300 hPa. They are replaced with a constant value of q above 70 hPa and extrapolated according to an empirical power-law between 300 and 70 hPa.

Once model radiances have been computed, the cost function and its gradient with respect to radiances can be calculated. Then the adjoint operators are applied in the reverse order (Fig.1) to yield the gradient of the cost-function with respect to the control variable. The J_o computation is followed by the computation of J_b and its gradient (*Heckley et al.*, 1992), and the whole procedure is repeated until the minimisation has reached convergence, or the maximum number of iterations has been reached.

4. VALIDATION

4.1 Forward and adjoint operators

The forward calculation for J_o was validated by carrying out radiance calculations at all TOVS locations for a given six-hour forecast. RMS and bias of the observed departures from the calculated radiances were compared with the result of operational radiance computations (with the same radiative transfer model).

The adjoint calculations of Fig. 1 were validated with the so called gradient test, which tests that the adjoint is fully consistent with the forward operators. In the gradient test, a test value (t) is computed as the ratio between a perturbation of the cost-function and its first order Taylor expansion:

$$t = \lim_{\delta x \to 0} \frac{J(x+\delta x) - J(x)}{\langle \nabla J, \delta x \rangle} = 1 \qquad (7)$$
$$\delta x = \alpha x$$

Repeatedly decreasing α one order of magnitude, printing t at each iteration should show t approaching one, by one order of magnitude at a time, provided $J(x)$ is approximately linear over the interval $[x,x+\delta x]$. An example is shown in the table below. In this case $J(x)$ represents a global multi-variate analysis of TOVS

radiances. The near linear behaviour of t over a wide range of α (initially as well as after minimisation) proves that the coded adjoint is the proper adjoint.

α	test (t) initially	test (t) after minimisation
10^{-2}	-.36241129	-.47398104
10^{-3}	0.52705531	0.83170738
10^{-4}	0.95177748	0.97291877
10^{-5}	0.99468148	0.99579516
10^{-6}	0.99946326	0.99954898
10^{-7}	0.99994596	0.99995526
10^{-8}	0.99999457	1.00001444
10^{-9}	0.99999729	1.00017900

4.2 An example with a simple analytical solution

Consider the univariate temperature analysis of *one* radiance datum at *one* location. Assume that temperature forecast errors are constant in the vertical. With the linear approximation the temperature analysis increment $(T_a - T_b)$, is given by (from Eq.2):

$$T_a - T_b = BR'^T (R'BR'^T + O)^{-1} [y - R(x)] \quad (8)$$

where:

R' is $\partial R/\partial T_i$, and T_i is temperature at model level i

B is the temperature covariance matrix

O is the observation error variance of the channel in question

$R'BR'^T$ is the radiance equivalent of the forecast error variance

For the vertical distribution of the analysis increment (ignoring the amplitude), Eq.(8) simplifies to:

$$T_a - T_b \propto BR'^T \quad (9)$$

The matrix R' represents the sensitivity of a given TOVS channel to the temperature at discrete model levels. The matrix B describes how the observation increment is spread in the vertical. In the first validation

experiment we simplify further and assume a diagonal $\boldsymbol{B}$. This de-couples the analysis in the vertical, and we obtain a two-dimensional analysis. Hence, the analysis increment at each level should be proportional to $\boldsymbol{R}'$.

Carrying out a full T63 3D-Var analysis, with the above simplifications, using the channel MSU-2 only, gives the vertical profile of analysis increment given in Fig. 2b. Comparing with $\boldsymbol{R}'$ for MSU-2 as shown in Fig. 2a we see that the analysis indeed gives the expected solution.

With a non-diagonal vertical correlation matrix, the theoretical analysis increment at a given level depends on $\boldsymbol{R}'$ at all other levels, according to Eq.(9). The theoretical analysis increment with the full $\boldsymbol{B}$ matrix is shown in Fig. 3a (thick line - the thin line represents 1D-Var and will be discussed in section 5.3) to be compared with the actual analysis increment in Fig. 3b. The two are virtually identical, which fully validates the vertical aspects of 3D-Var.

4.3 Simple multivariate example

The multivariate J_b introduces a balanced coupling between mass and wind (Heckley et al. 1992). In our simple case of analysing just one datum of MSU-2 at latitude 60 South we obtain the result shown in Fig. 4. The temperature increment changes sign near 250 hPa, in agreement with Fig. 3a. The increments are isotropic (as seen from the circular contours) and close to geostrophic, in accordance with the specification of J_b.

4.4 On the sensitivity of some channels to model level temperature

A row of the matrix $\boldsymbol{R}'$ contains the partial derivatives of a given channel with respect to temperature at model levels. It can be interpreted as the sensitivity of the observation to the analysed variable. Fig. 2a therefore represents the "signal" and Fig. 3 the "response" (for MSU-2). Because the scheme is nonlinear, the signal depends on the atmospheric state. The curves presented here are strictly speaking valid for a point at (-60,-170), 19920516-12 UT.

Note that these curves will show larger sensitivity to temperatures that represent thicker model layers, i.e. they are not directly comparable with the normal TOVS weighting functions shown in Fig. 5. The bottom two points in the diagrams represent T_{2m} and T_s.

We see from Fig. 2a that MSU-2 is most sensitive to temperature at model levels 10 to 13 (400 to 690 hPa) as expected, and more surprisingly that although this channel peaks in the mid-troposphere it is more sensitive

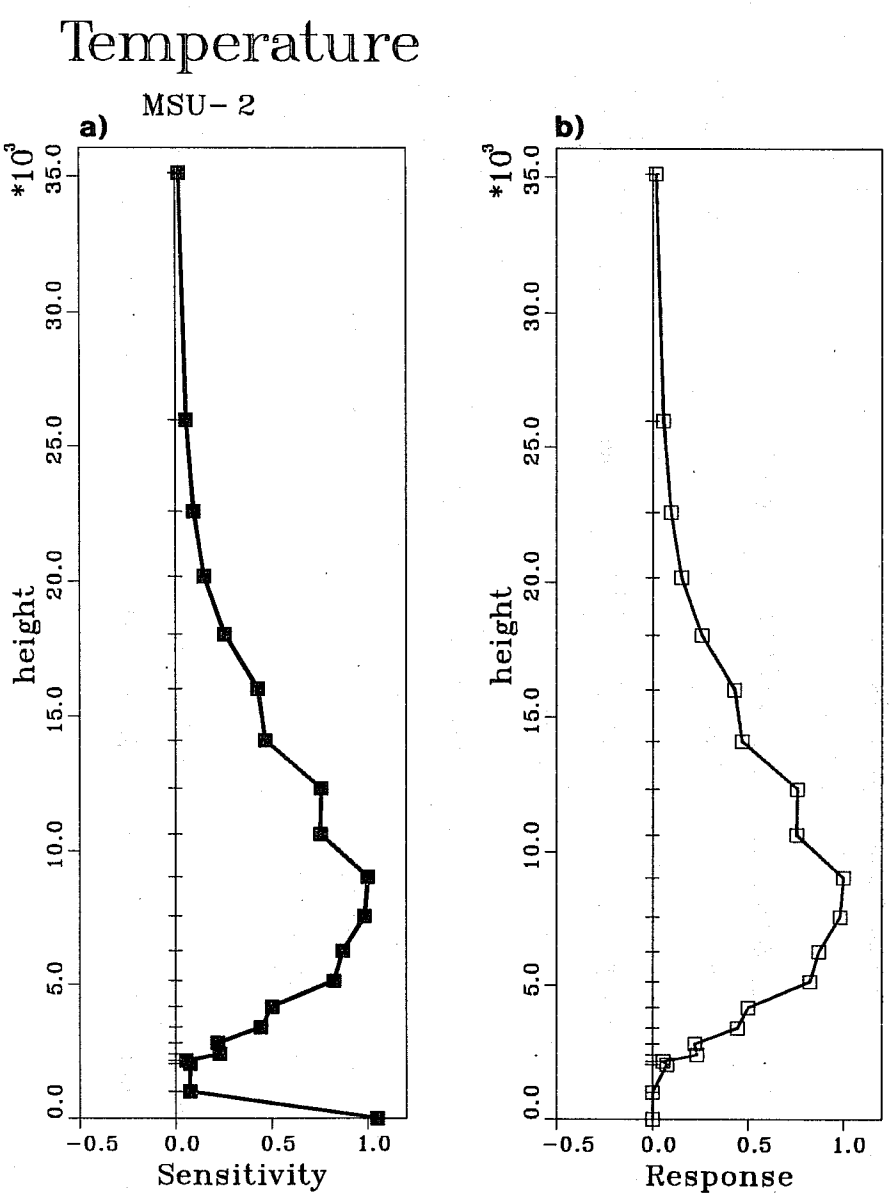

Fig.2 Shows for MSU-2: a) "Sensitivity", i.e. $R' = \partial R / \partial T_l$, the partial derivative of a TOVS channel with respect to model level temperatures and b) "Response", i.e. analysis increment in a 2D analysis. The two lowest points on the curves represent T_s and T_{2m}.

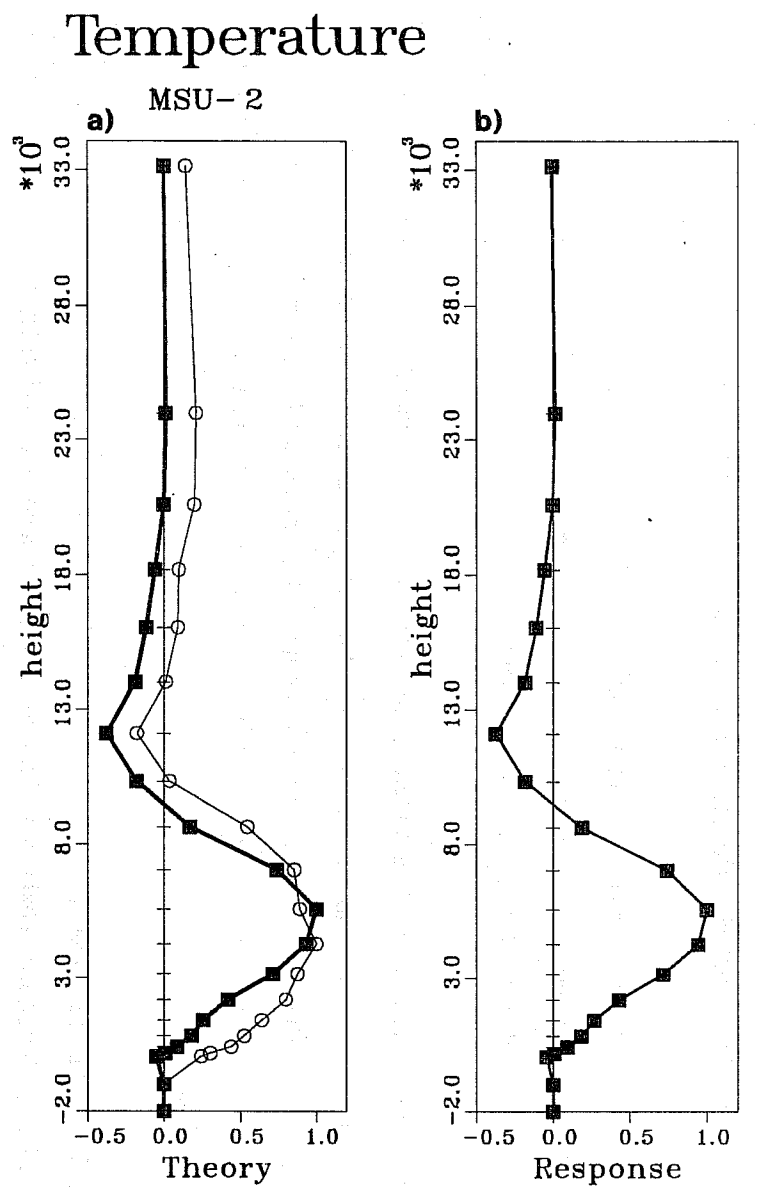

Fig.3 Like Fig.2 showing a) Theoretical response (3D-Varthick, 1D-Varthin line) and b) actual response in a 3D-VAR analysis.

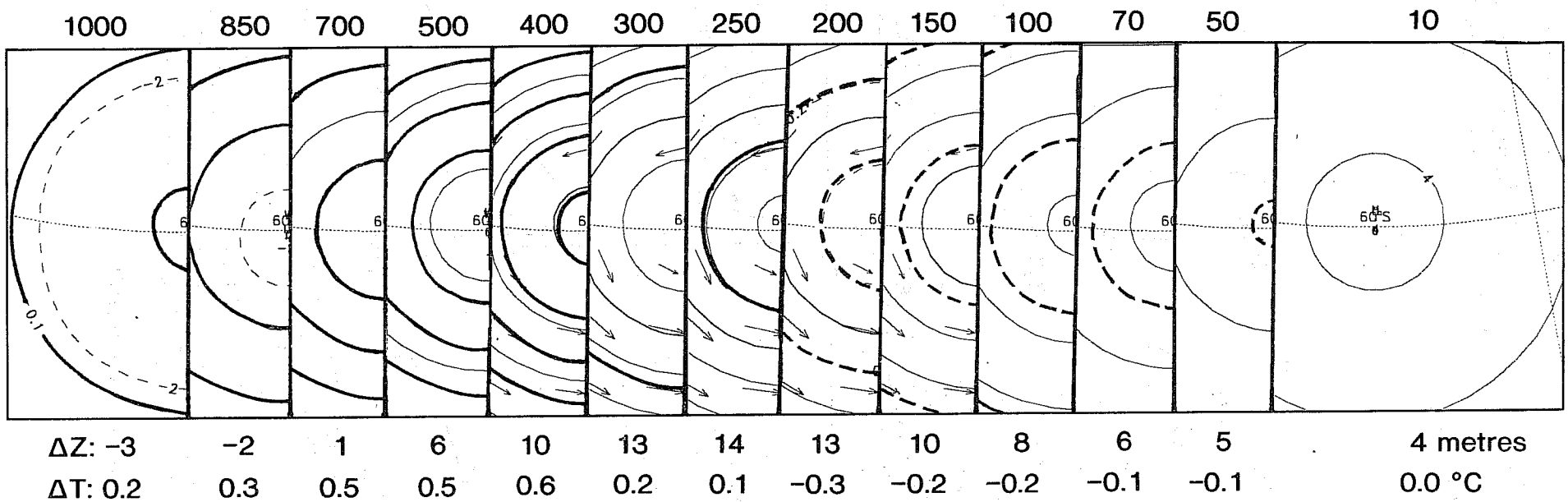

Fig.4 Analysis increments post-processed to pressure levels from 1000 to 10 hPa, as indicated, of a 3D-VAR T63 analysis of one MSU-2 datum. The plot shows temperature, height and wind. Wind arrows are not plotted at all levels because they were to small. The magnitude of the height and temperature increments are indicated below the figures.

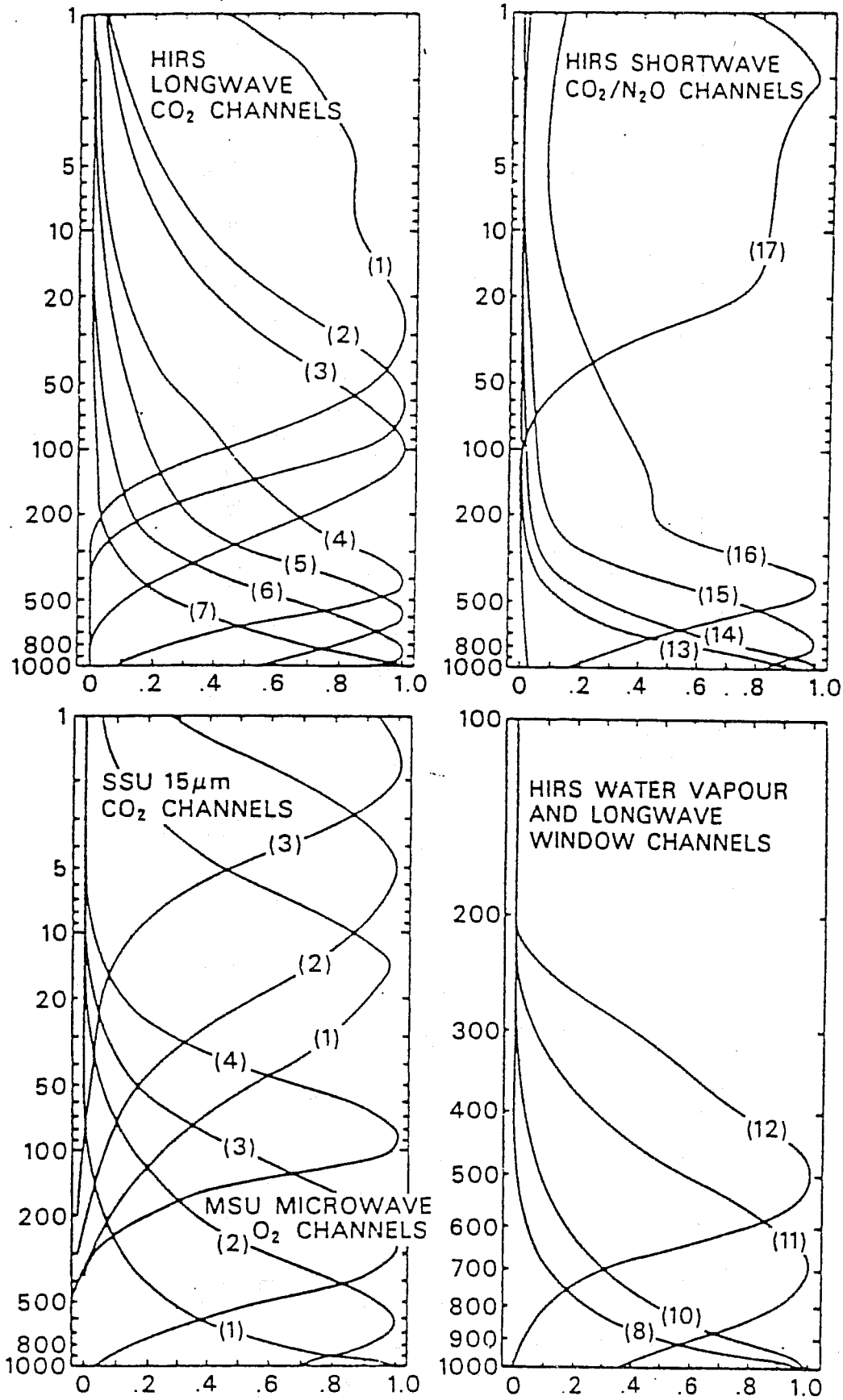

Fig.5 TOVS weighting functions, from (Smith et al. 1979).

to T_s than it is to any upper level temperature. For HIRS-14 (Fig.6a), which peaks lower, at level 13 (690 hPa), the sensitivity to T_s is more than four times larger than to any upper level temperature. This shows that there is potentially a problem in our formulation as our 3D-Var is unable to adjust T_s in order to fit the observed radiances. It is possible in principle to include T_s in the control variable but for technical reasons this is not the case in practice at present. Instead we have chosen to use the retrieved T_s from 1D-Var, which should be more consistent with the observed radiances than the model surface temperature.

We can also use the graphs of $\boldsymbol{R}$ to investigate the impact of a regression extrapolation of temperature above the top of the model (7.3 hPa) (as mentioned in section 3.4). Fig. 7a shows HIRS-1 and Fig. 8a shows HIRS-4 (Fig.7b and 8b show the theoretical analysis response and will be discussed in section 5.3). We can clearly see that the curves are distorted, especially HIRS-1 (Fig.7a), in comparison with the original weighting functions in Fig. 5. The signal above the model top has been zeroed and added back to the model levels according to the adjoint of the extrapolation. This is clearly not satisfactory. We have therefore abandoned the idea of extrapolating the control variable and to use instead the 1D-Var retrieved temperature above the model top.

5. RESULTS OF A THREE-DIMENSIONAL EXPERIMENT

In the following section, we will compare the result of a T63 three-dimensional analysis of TOVS radiances plus conventional data with our best estimate of the truth at the same resolution - a T63 OI analysis using all available data, including 1D-Var retrievals.

The 3D-Var analysis used radiances from 4,433 TOVS spots, in total 46,205 pieces of information from channels HIRS-1 to 7, 10 to 15 and MSU-2 to 4. In order to minimize the effect of forecast errors in surface temperature, only the seven most high-peaking channels were used over land (HIRS-1 to 4 and 12, MSU-3 and 4), and from cloudy spots only the cloud insensitive channels were used (HIRS-1 to 3, MSU 2 to 4). The situation was 890209-12 UT, for which we had received "cloud-cleared" radiances from NESDIS, Washington, with nearly complete global data coverage.

We used the multi-variate J_b, assuming 10 % of the background error in the gravity modes and 90 % in the Rossby modes, as defined by a projection of the control variable onto the Hough modes of the background. The control variable was in terms of the mass-variable $\boldsymbol{P}$, vorticity, divergence and specific humidity, all in spectral space projected on the eigen-vectors of the vertical correlation matrix, as described in *Heckley et al.*

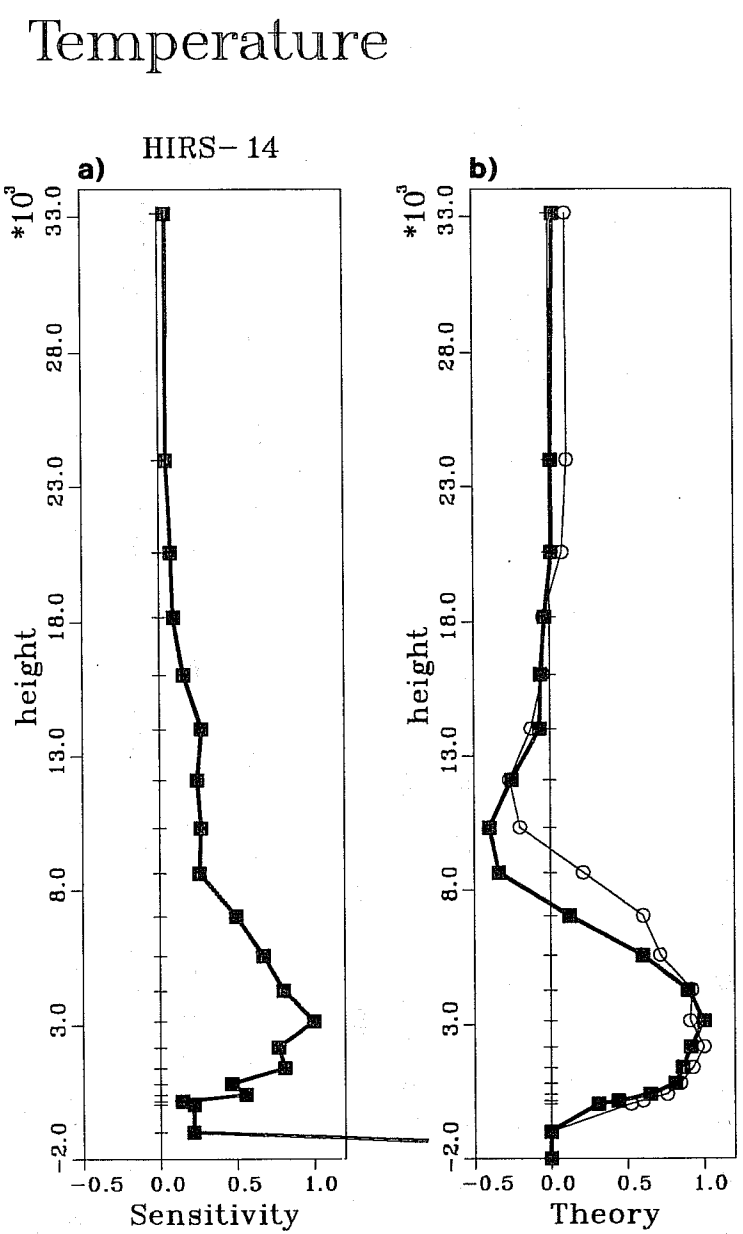

Fig.6 a) "Sensitivity" (as in Fig.2) and b) theoretical analysis increment (as in Fig.3), for HIRS-14.

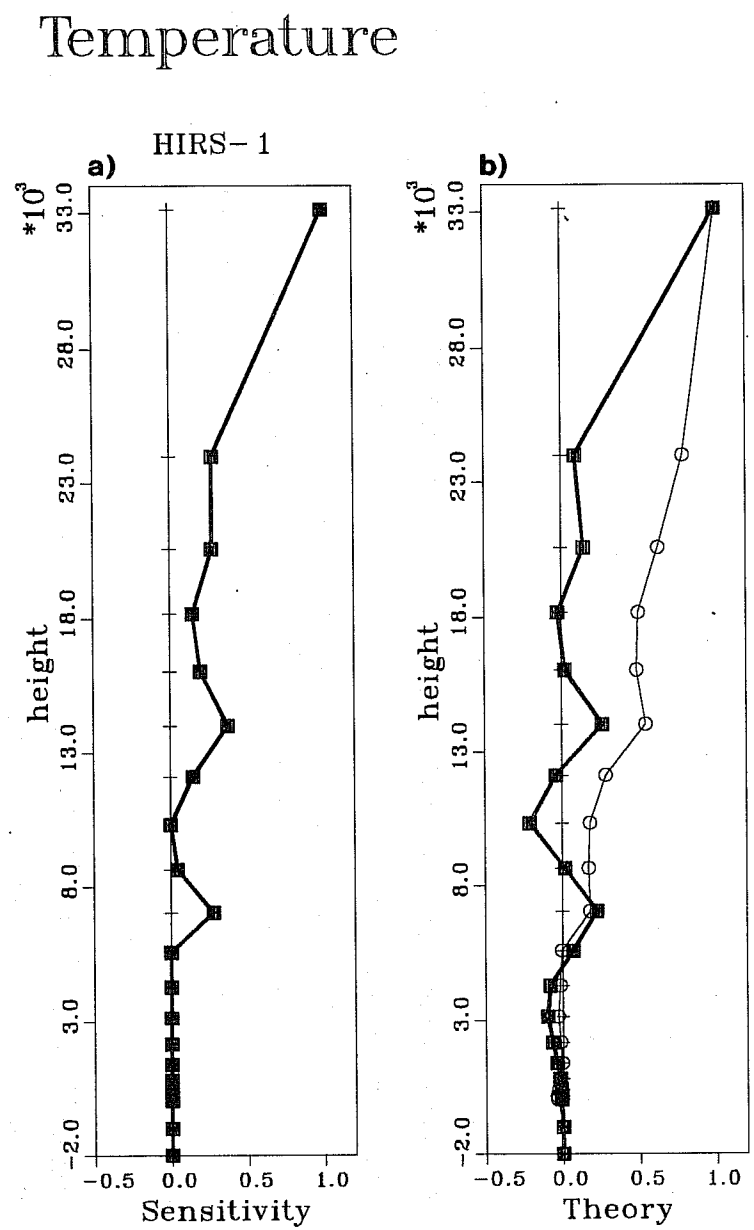

Fig.7 As Fig.6, for HIRS-1.

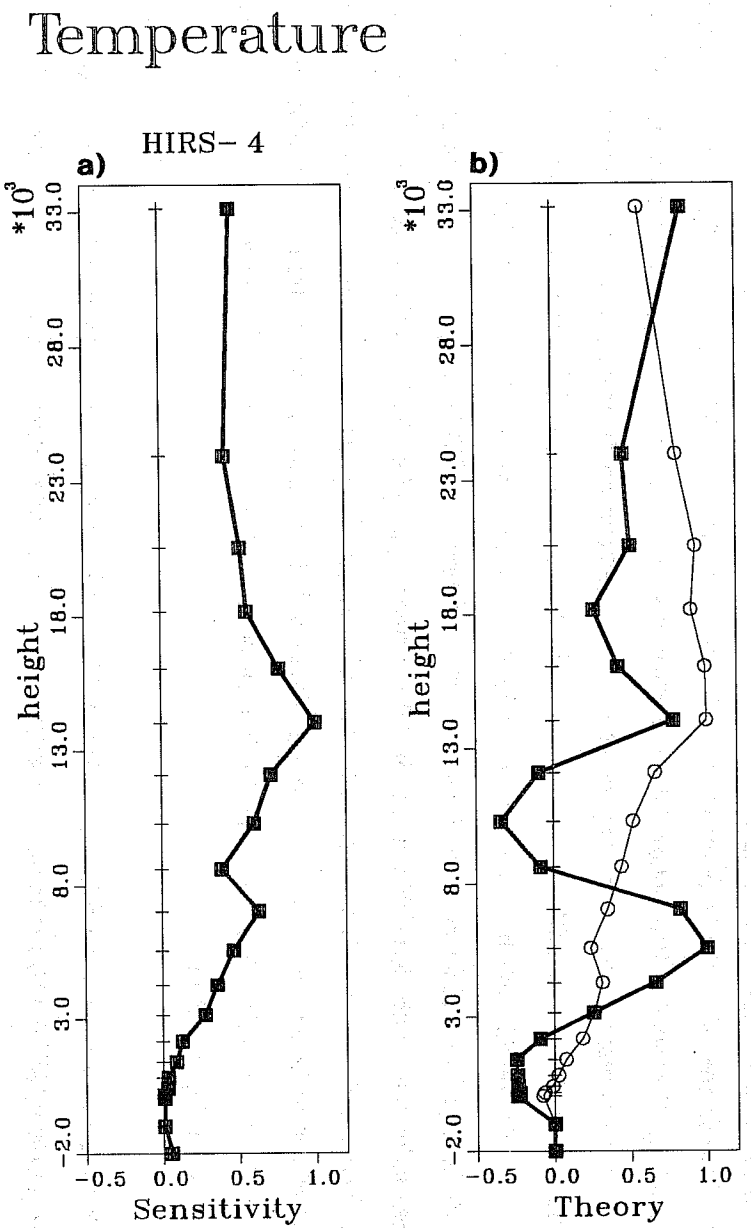

Fig.8 As Fig.6, for HIRS-4.

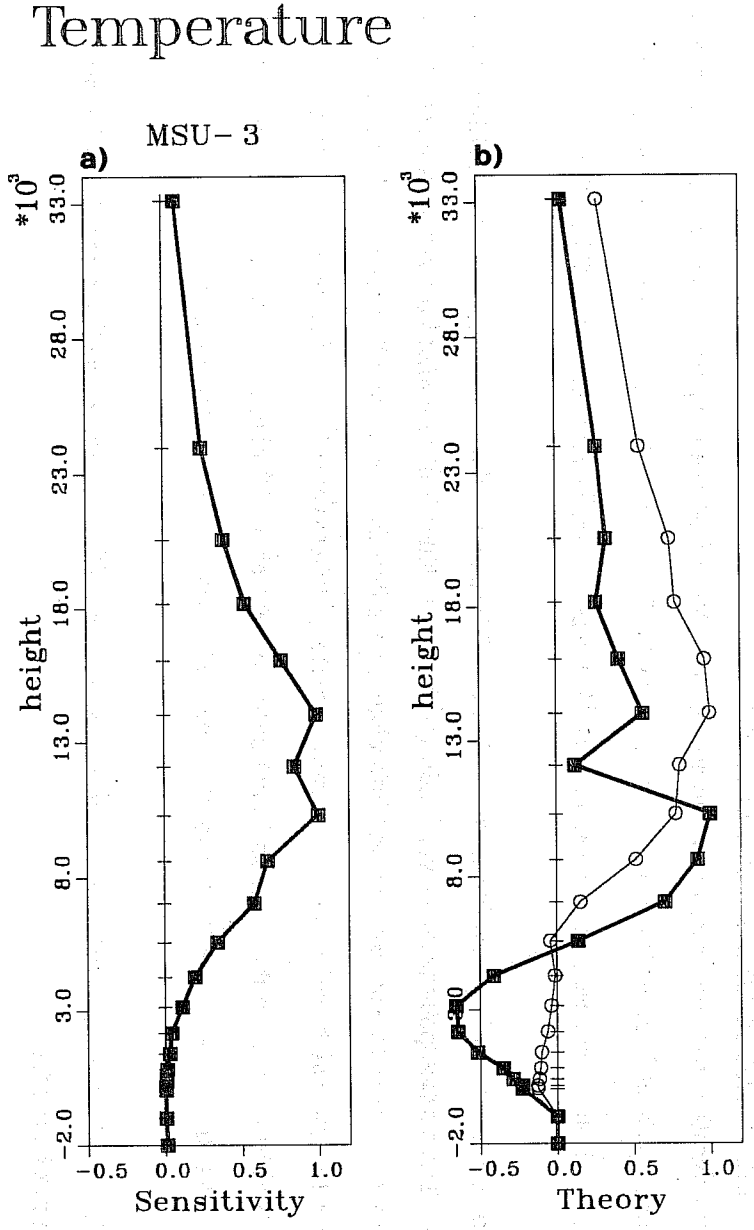

Fig.9 As Fig.6, for MSU-3.

(1992). The forecast error variances vary in the vertical but were for mass and wind (not for specific humidity) for technical reasons constant in the horizontal.

5.1 Convergence and control of gravity waves

The minimisation was continued for 50 iterations. Fig. 10a shows the evolution of the two terms of the cost function. J_b is zero initially because the starting point of the minimisation is equal to the background (an initialized six-hour forecast). It appears that the convergence is saturated after about 35 iterations.

Fig. 10b shows the evolution during the minimisation of the energy in the gravity waves of the first five vertical modes. The starting point contains very little gravity waves since it is an initialized six-hour forecast. As the minimisation draws to data, some of the information goes into the gravity part. However, it is clear from the diagram that the multi-variate balance formulation of J_b effectively controls the amount of gravity waves present in the analysis so that after an initial over-shoot (at about 10 iterations) a stable level is reached. No other means than J_b was used to control the amount of gravity waves in this experiment.

Experience with the NMC scheme by *Parrish and Derber* (1992) showed that a separate initialisation step should not be needed.

5.2 Comparison of analysis increments

Fig. 11a shows the 500 hPa height and wind increments of the 3D-Var analysis. Fig. 11b shows the corresponding OI analysis. Conventional data dominate the analyses over land. In the Atlantic, where TOVS data are more important, there is a very good correspondence between the two analyses. All features occur in both analyses with only small differences in their magnitude. There is also good correspondence in the Pacific and in the Polar regions but here the 3D-Var increments appear stronger. A 32 metre increment in the OI (Central Pacific), for example, corresponds to a 47 metre increment in 3D-VAR. It is also noticeable that three localized increments with closed circulation in OI increments [at (45N,150W), (45N,165W) and (37N,150E)] do not appear in 3D-VAR or have become part of more large-scale patterns. Furthermore, there seems to be a bias between the radiances and the background field (a six-hour forecast) in the sub-tropical part of the eastern Pacific. This may be due to the fact that TOVS data had not been used in the operational assimilation that produced the background field.

At 50 hPa, Fig. 12a and 12b, there are many similarities between the two analyses, especially over the oceanic areas. It appears that 3D-Var in most parts extracts the same information from the radiances as does 1D-Var

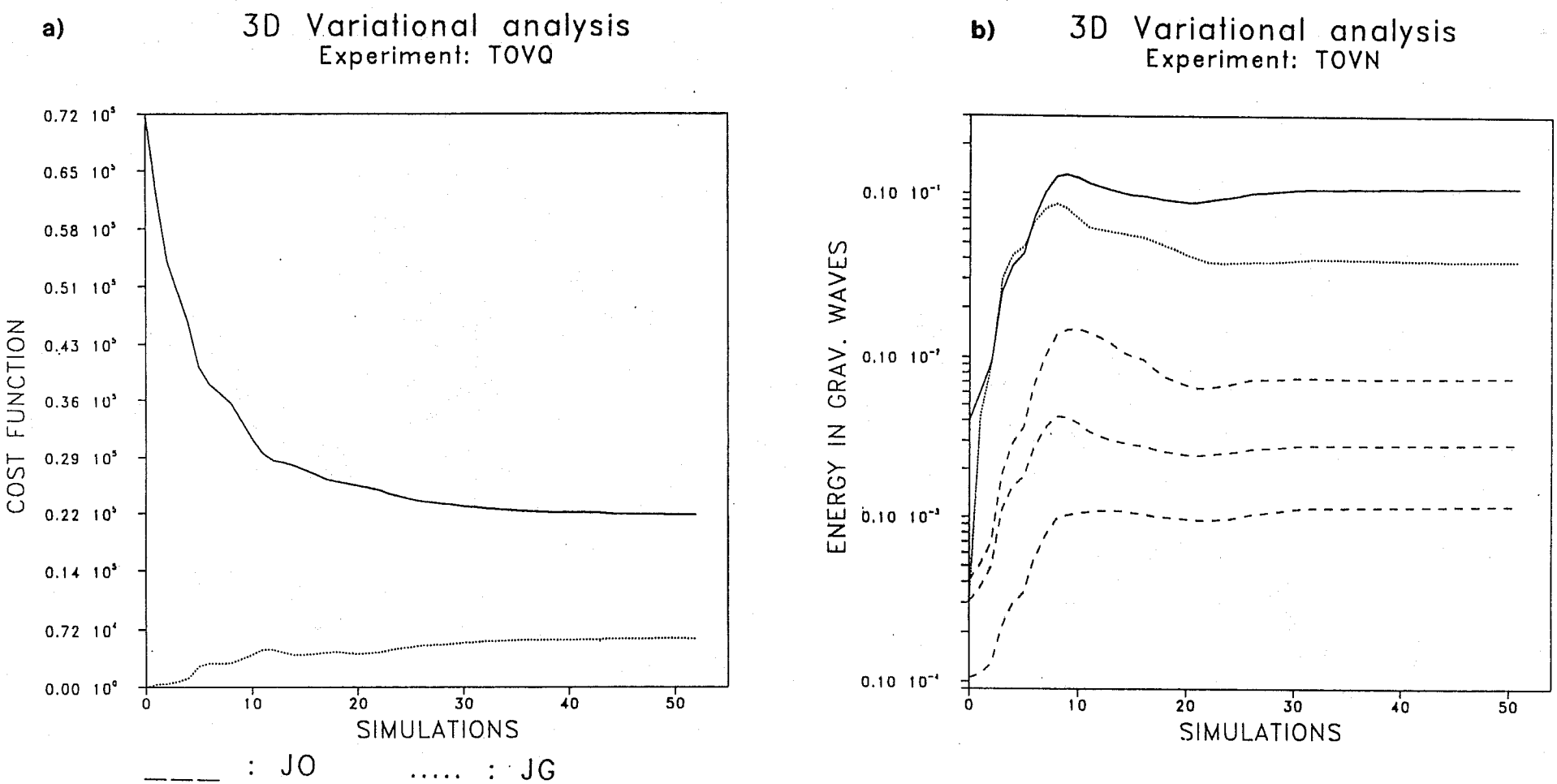

Fig.10 Evolution during the minimisation of (a) the cost functions (J_o full line and J_b dotted) and (b) energy in gravity modes 1 to 5, as indicated.

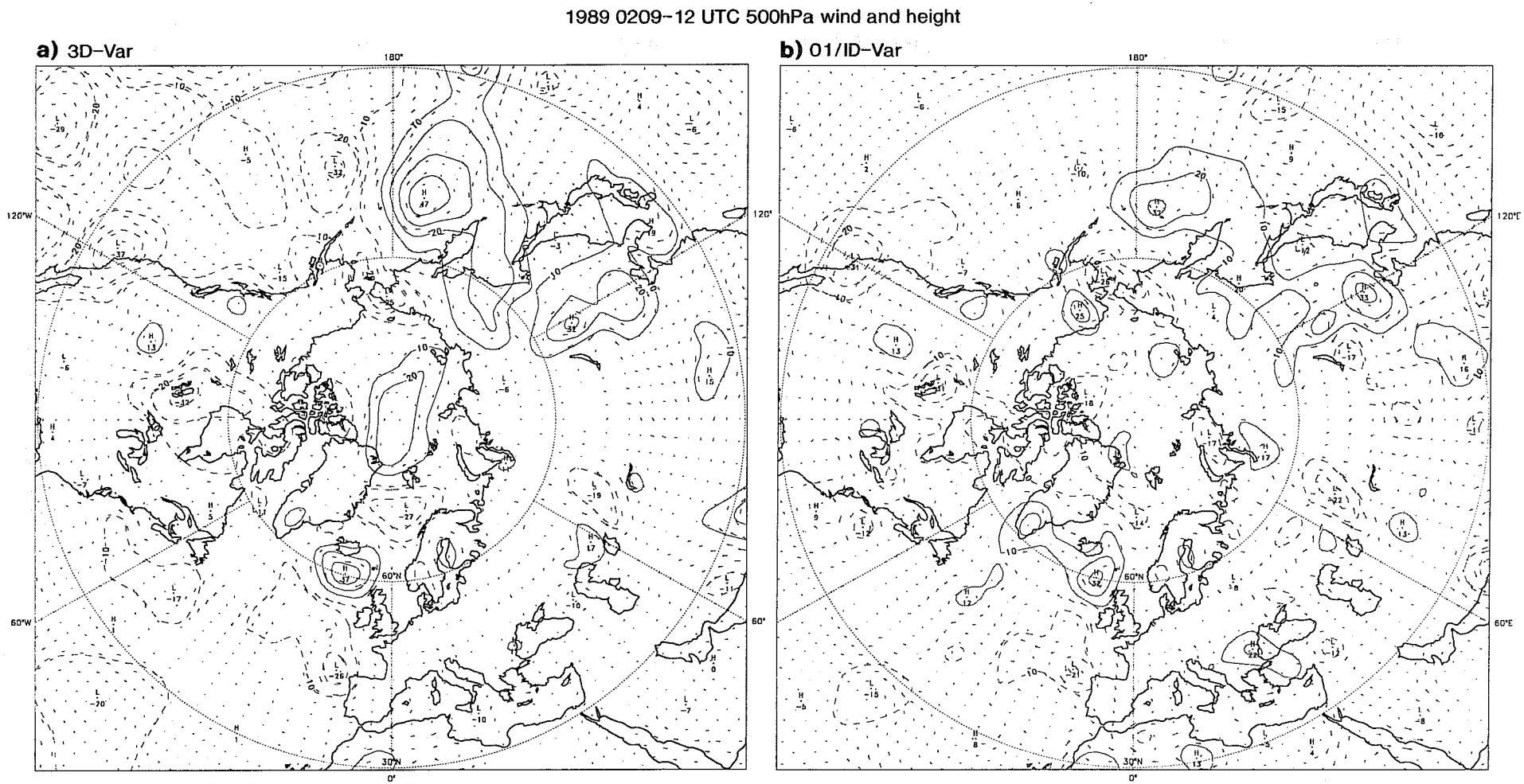

Fig.11 Analysis increment of geopotential height and wind at 500 hPa, 890209-12 UT, with a contour interval of 10 metres, negative dashed. a) is 3D-Var and b) is OI/1D-VAR.

1989 0209-12 UTC 50hPa wind and height

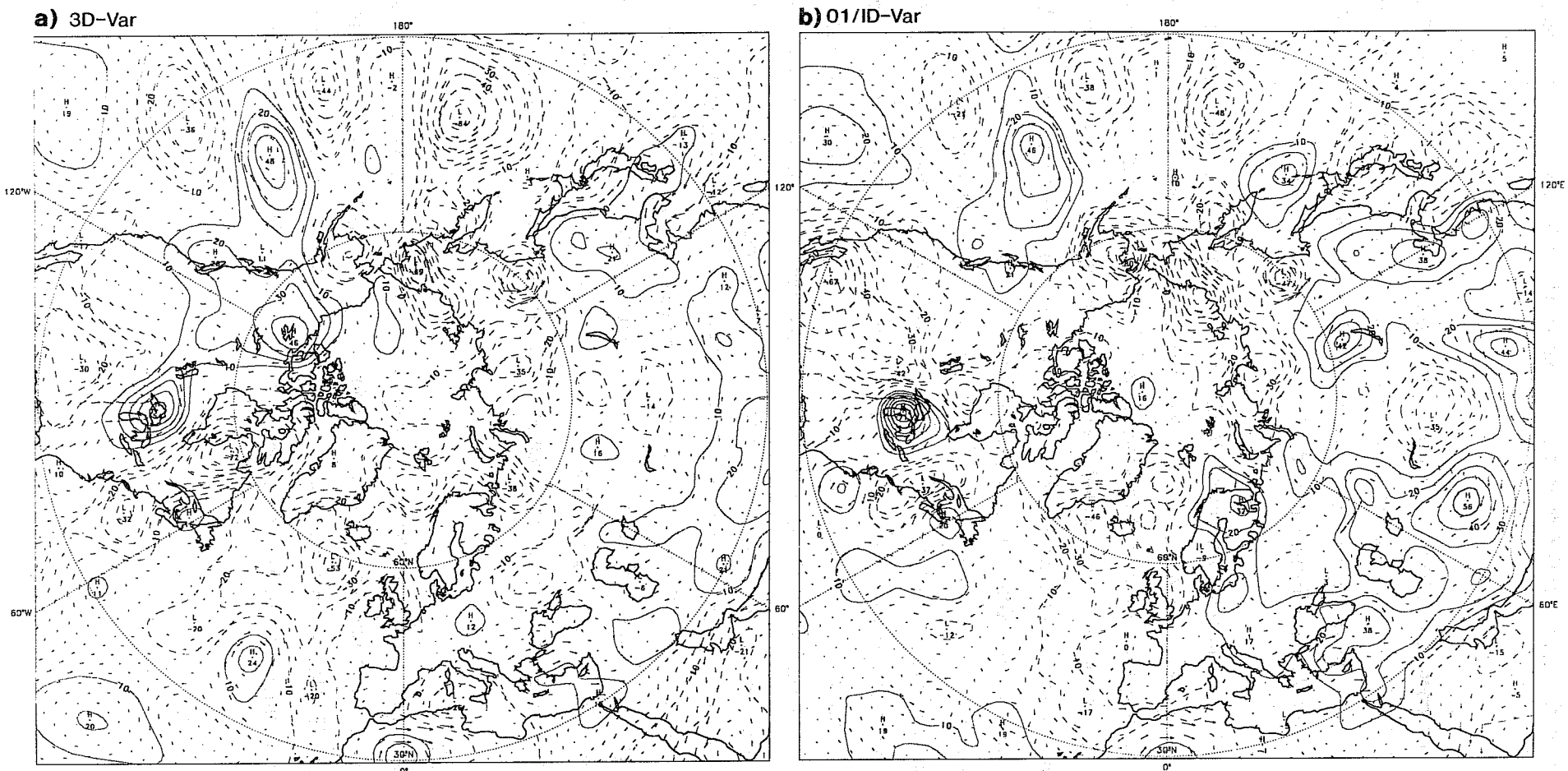

Fig.12 As Fig.11, for 50 hPa.

plus OI, but with a larger amplitude in 3D-Var. The main differences appear over Asia, where a general bias between the radiosondes and the TOVS radiances is a complication which has slightly different impact in the two schemes.

In the southern hemisphere, where TOVS provides the dominant data source, 3D-Var (Fig.13a) picks up all the major features of the OI analysis (Fig.13b), again with somewhat larger amplitude than 1D-Var/OI. The agreement is very good with one exception: a 90 metre height increment in the 3D-Var analysis at (-2·,-50·), which corresponds to a 24 metre increment in the OI.

This increment is in the baroclinic area of a 956 hPa cyclone (see Fig.14, where surface fronts have been drawn). A closer look at the 3D-VAR temperature increments level by level (not shown) indicates that the cold air in the lower troposphere in the 3D-Var analysis has advanced further than in the first-guess. Positive increments aloft indicate that the warm occluded air ahead of the surface warm front is warmer and has reach further around the low (towards the pole) than in the first-guess. Qualitatively this amounts to a displacement of the front by the analysis. *Kelly* (1978) suggested the zero contour of 500 hPa vorticity as an objective indicator of the position of upper-air fronts. A comparison of vorticity maps for the analysis and the background indeed showed that the 3D-Var had moved the front forward a distance of 70 to 100 km. Fig. 14b shows vertical profiles of the analysis and the first-guess at point A, as indicated in Fig. 14a. Point A is situated just behind the surface cold front. The radiances have cooled the lower part of the troposphere by up to 7 degrees (at 750 hPa) and warmed the lower part of the stratosphere, in so doing lowered the tropopause from 250 to 400 hPa. The change in upper troposphere humidity is also consistent with a colder air-mass. At point B (Fig.14c), just ahead of the surface warm front, there has been a general warming of the profile. In summary, it seems that the unusually large 3D-VAR increment has occurred in an area were there has been a slight phase error in the first-guess. The analysis increments are consistent with moving the frontal system forward. This gives us some confidence that the large 3D-Var analysis increment in this area can be realistic. However, there is necessarily little vertical structure in the increments, and they are barotropic by design of J_b.

5.3 Discussion

The results presented indicate that the 3D-Var retrieval/analysis of TOVS radiances works as anticipated. The analysis increments in general agree with those produced by 1D-Var/OI. The size of the increments are somewhat larger in 3D-Var, which partly reflects the high degree of confidence we put on measured radiances relative to SATEM retrievals. The 1D-Var retrievals are also given a lower weight in OI to account for their

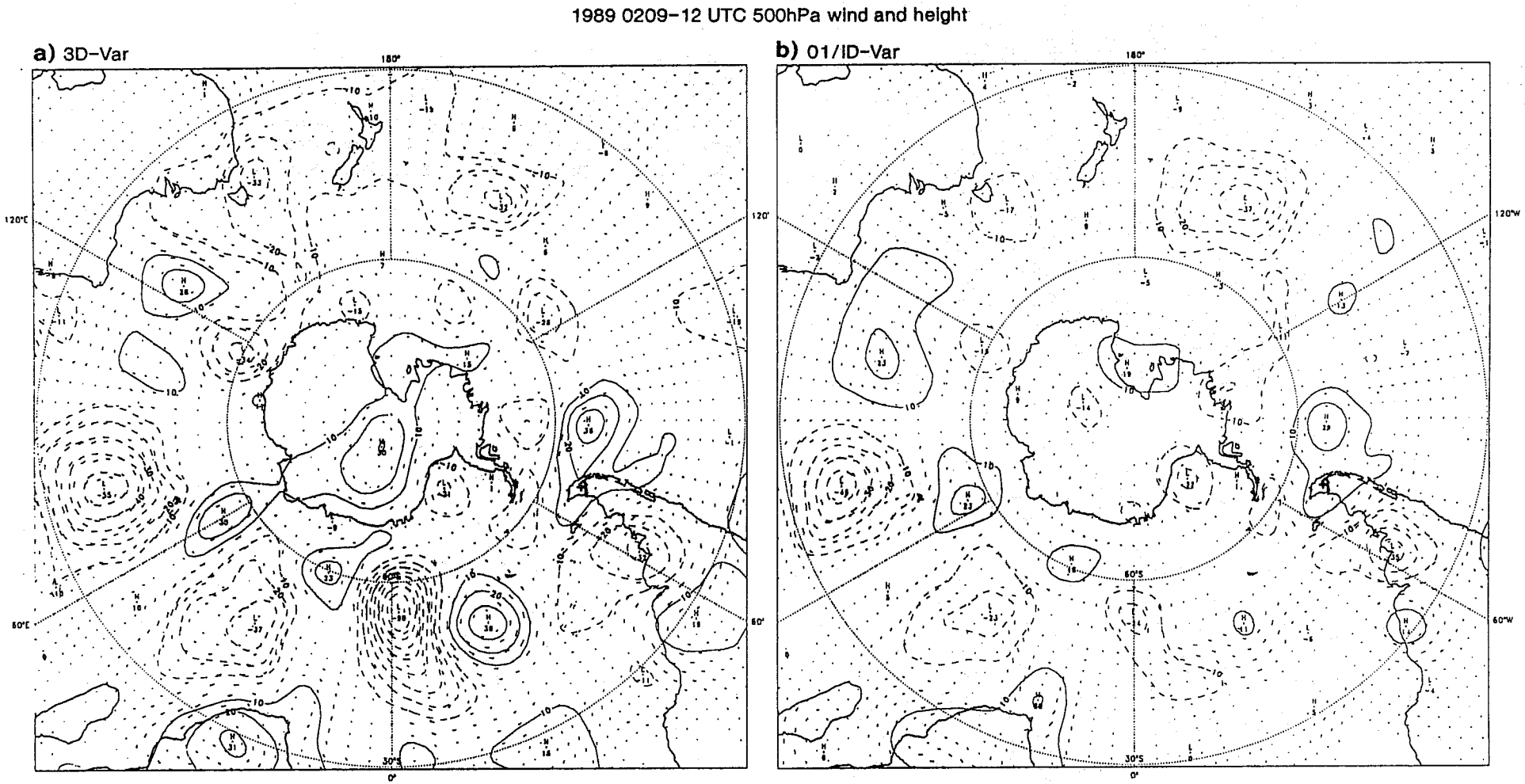

Fig.13 As Fig.11, for southern hemisphere.

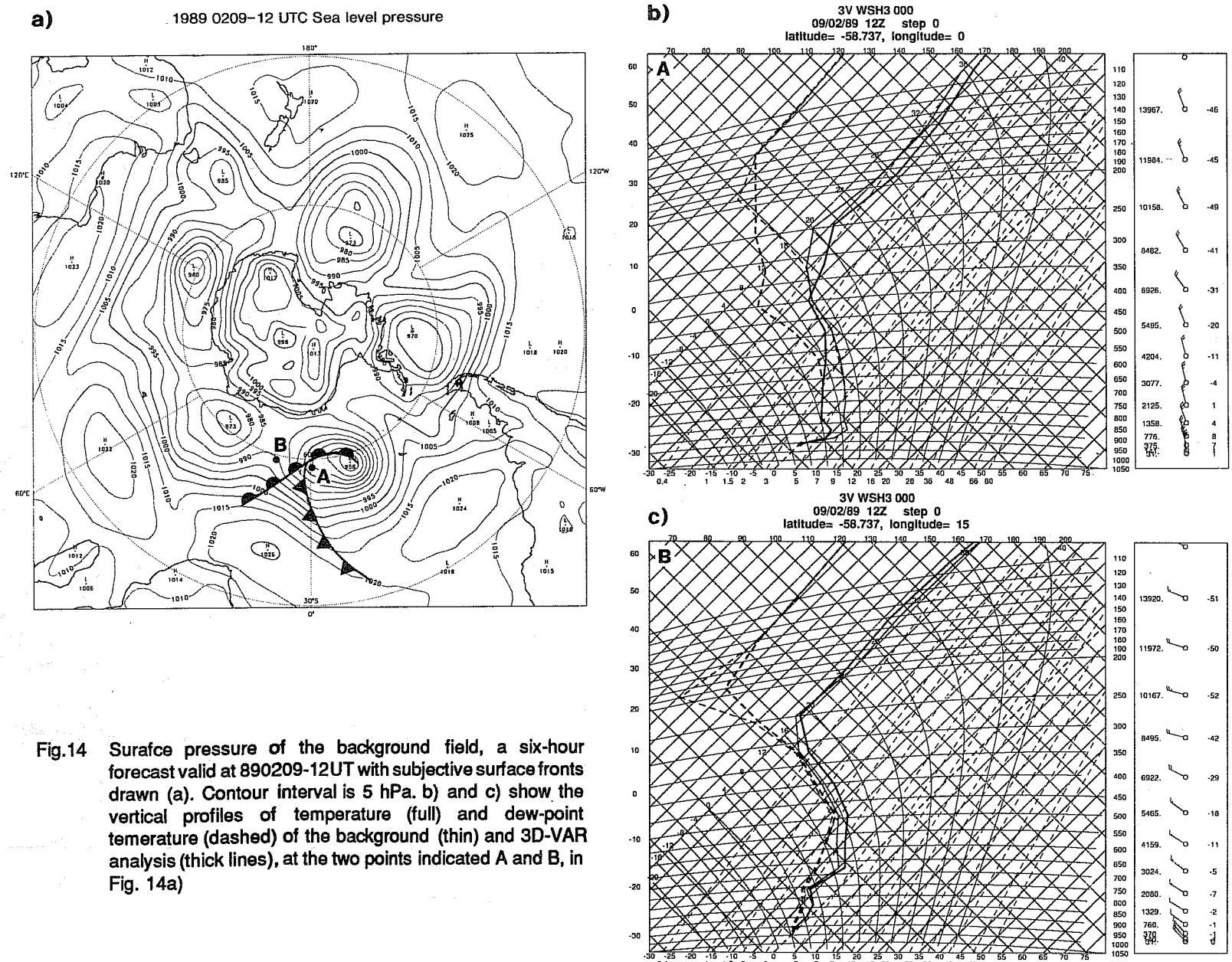

Fig.14 Surafce pressure of the background field, a six-hour forecast valid at 890209-12UT with subjective surface fronts drawn (a). Contour interval is 5 hPa. b) and c) show the vertical profiles of temperature (full) and dew-point temerature (dashed) of the background (thin) and 3D-VAR analysis (thick lines), at the two points indicated A and B, in Fig. 14a)

correlation with the background error. In addition, 3D-Var assumes larger temperature forecast errors than 1D-Var, a discrepancy that needs to be reconciled.

An important factor for the analysis of TOVS data is the specification of the vertical correlations for temperature. Since the vertical resolution of the satellite information is less than the resolution of the forecast model, the finer details in the vertical of the analysis increments are imposed by the specified prediction error covariances (*Lönnberg*, 1989). Lönnberg found that very noisy analysis increments could result from an improper specification of observation and forecast error statistics. Large oscillations in the temperature increments on model levels occurred in the ECMWF OI analysis when layer-mean temperatures (T_v) of thicker layers were used, i.e. T_v(1000-500) instead of T_v(1000-700) and T_v(700-500).

Similarly, oscillation in the temperature increments on model levels can occur when radiances with broad weighting functions are used directly, as in 3D-Var. This can be seen in the right hand panels of the figures presented in sections 4.2 and 4.4. The diagrams show the theoretical analysis increment for a given channel (when used on its own) as given by Eq.9. The thicker line refers to the vertical correlations used in 3D-Var (these are equal to those used by ECMWF OI over oceans at 45 N) and the thinner line is 1D-Var. The 1D-Var correlations were derived directly from temperature data (in the U.K. Meteorological Office model), *Eyre et al.* (1992), whereas the 3D-Var (and OI) correlations were derived from height data, fitted to a continuous representation from which temperature correlations were obtained as the first derivative in $\ln p$.

The 1D-Var correlations are significantly broader than those used by 3D-Var and give rise to a broader response in MSU-2 (Fig.3a) HIRS-14 (Fig.6b) and HIRS-1 (Fig.7b). For HIRS-4 (Fig.8b) and MSU-3 (Fig.9b), which have very broad weighting functions, we see that the sharper correlation functions of 3D-Var give rise to oscillations in the temperature increments, much like the examples of *Lönnberg* (1989). Oscillations do not occur with the 1D-Var vertical correlations.

It is not clear whether this feature of the 3D-Var correlations causes problems when a number of channels are used simultaneously, but it should be kept in mind. There is no good reason why 1D-Var and 3D-Var should use different forecast error statistics. A common formulation which is good for temperature as well as height (and vorticity and divergence) is needed.

6. RESULTS OF A FOUR-DIMENSIONAL EXPERIMENT

A 4D-Var experiment was carried out in order to find out to what extent the dynamics of the forecast model, in the absence of J_b, could provide wind increments in response to TOVS information solely on the mass field.

The background information is generally much less important in 4D-Var than in 3D-Var. The three-dimensional analysis problem is under-determined unless background information is provided. In four dimensions, a strong constraint is posed by the evolution of the forecast model; the model trajectory has to stay close to the observations over a period of time, which makes it possible to ignore J_b at low resolutions.

The experiment presented here used TOVS from 18,208 locations over a 24 hour period as well as most conventional data (SYNOP-SHIPs, DRIBU, TEMP, PILOT, AIREP and SATOB). The resolution of the spectral model was T42-L19 (140 000 degrees of freedom). The model was used in its adiabatic version, and a combination of a weak constraint term on the tendencies of the energy in the gravity components of the model solution and a normal mode initialisation scheme was applied in the assimilation process in order to control noise.

A second assimilation was run with only conventional data so that the difference between the two assimilations would show the impact of TOVS in the presence of all other data.

The J_o computation was split into one-hour time slots, but is otherwise as described in section 3.4 and Fig. 1. The forward operator in the 4D case involves a model integration to the time of the observations, and in the adjoint part there is the corresponding adjoint integration to obtain the gradient of the cost function with respect to the initial time.

The minimisation was terminated at 30 iterations, although the cost function was still decreasing. Previous experiments had shown that in the absence of a background there was a tendency to draw to the data too much at the final stages of the minimisation and noise was generated. After 30 iterations the cost function had been divided by a factor of six.

6.1 Impact on the mass field

Fig. 15 represents the 500 hPa height field difference between the 4D-Var experiment using conventional and TOVS data and the 4D-Var analysis using conventional observations only. This level roughly corresponds to the peak of HIRS channels 4, 5 and 15 and MSU channel 2. The impact of TOVS radiances is mainly located

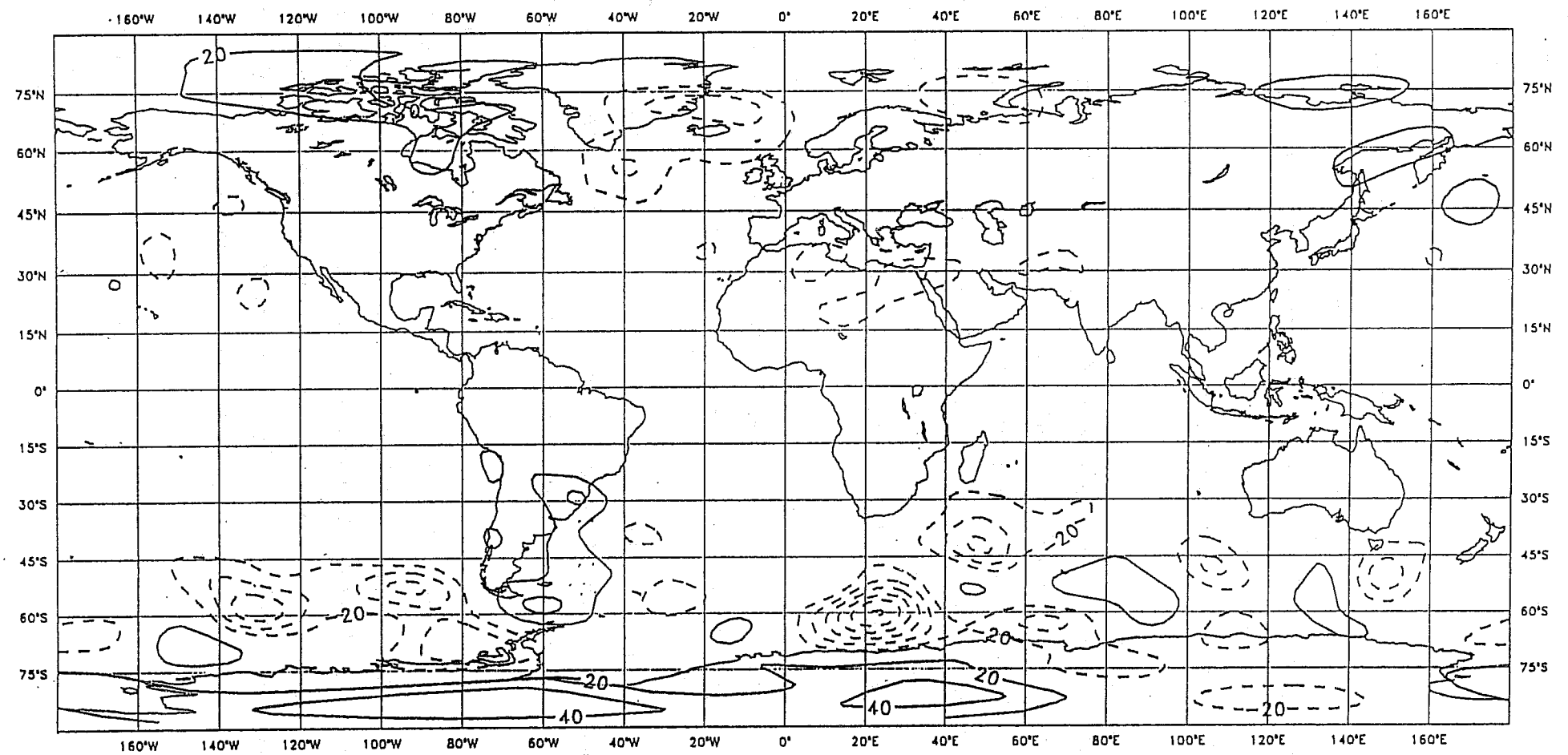

Fig.15 Geopotential differences at 500 hPa at the end of the assimilation period (890210-12UT) between 4D-Var performed with conventional observations plus TOVS radiances and 4D-Var performed with conventional observations only. Negative differences are dashed, and the contour interval 30 metres.

in the mid-latitudes of the southern hemisphere, and more particularly over the oceans where the differences reach more than 100 m. Bearing in mind the lack of conventional data in the southern hemisphere, the horizontal distribution of the impact is as expected.

It is notable that the magnitude of the increments agrees with that of the 3D-Var experiment presented in Fig. 13, although J_b, which is dominant in the 3D case, is absent in the 4D experiment. Also the position of the increments show a great deal of agreement with the 3D experiment, bearing in mind that there is a 24 hour time difference between the two analyses. The increments are generally positive in the ridges and negative in the troughs (not shown), they hence tend to amplify the flow in the analysis.

6.2 Impact on the wind field

It is more interesting to look at the wind increments inferred by the use of the radiances (Fig.16). When comparing with Fig. 15, one can notice that in the mid-latitudes the wind differences are geostrophically related to the geopotential differences, as expected from balanced fields. This shows that the information brought by the additional observations is dynamically consistent with the model solution. In other words, the mass-wind equilibrium information, mainly enforced by the presence of a constraint term on the gravity mode tendency and a normal mode initialisation before the model integration, has been properly transferred by the dynamics.

Wind differences are also noticeable in the tropics. The location of the largest wind difference patterns correspond quite well to the locations of the humidity increments (not shown). This raised the question whether there had been an impact on the wind analysis from the humidity information in the radiances.

To investigate this wind-humidity coupling, a 4D-Var experiment using TOVS data had to be rerun but excluding TOVS humidity channels HIRS-11 and 12. The new resulting 500 hPa wind difference is shown in Fig. 17. A large part of the wind increments has been wiped out, both in the tropics and in the mid-latitudes.

Channels HIRS-11 and 12 are rather sensitive to temperature which is very variable in the mid-latitudes. In the tropics where the temperature field is fairly flat, the comparison between Fig. 17 and Fig. 16 clearly confirms that a large part of the wind increments is due to radiance measurements in the humidity channels. Since in our experiments the humidity field behaves like a passive tracer (neglecting the effect of humidity on temperature in the model equations), this result, intrinsically linked to the four-dimensional nature of the assimilation, shows a nice example of how the dynamics of the model is able to infer information on an un-observed component of the flow (the wind) from remotely sensed information on the humidity field.

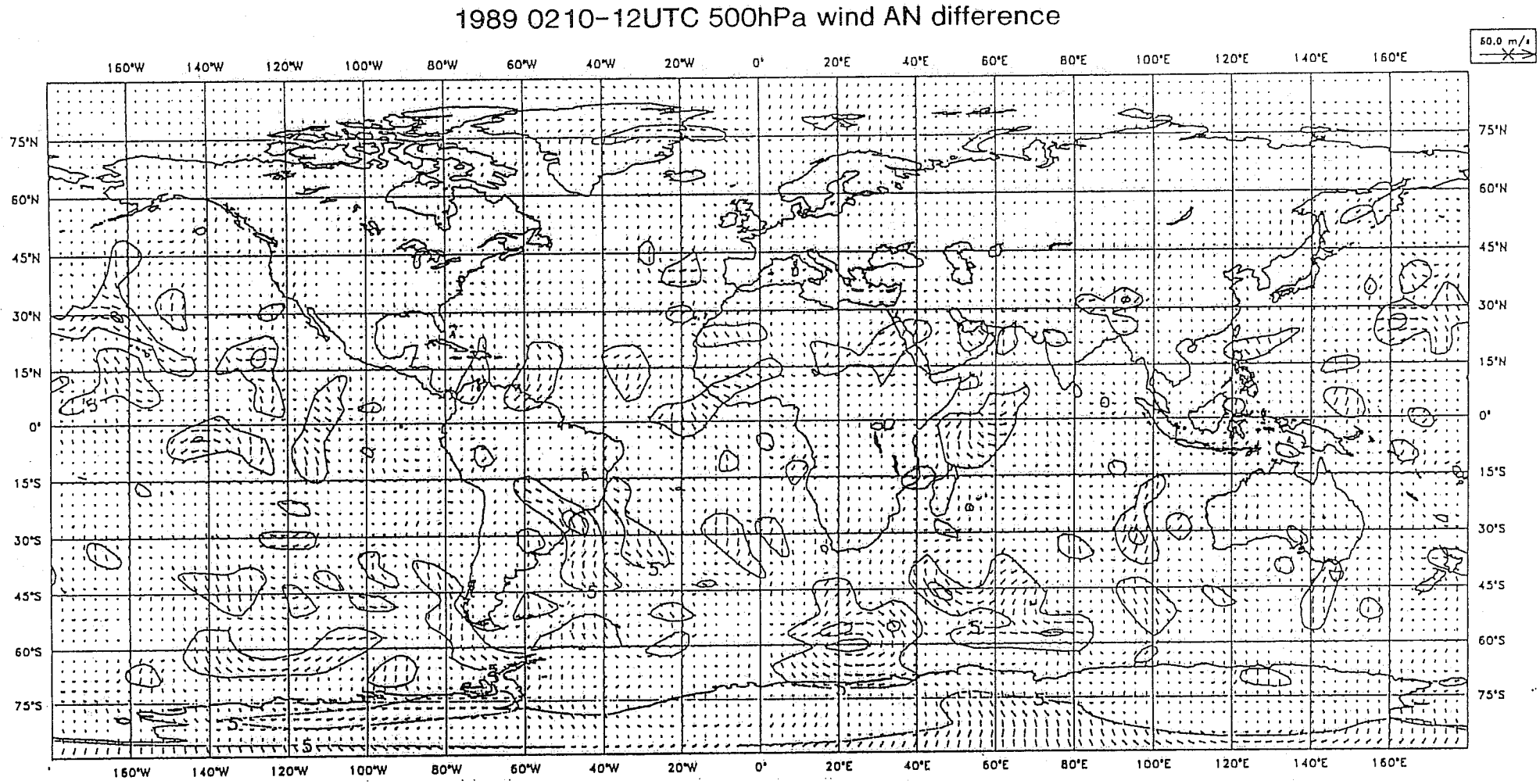

Fig.16 As Fig.15 but for the wind field.

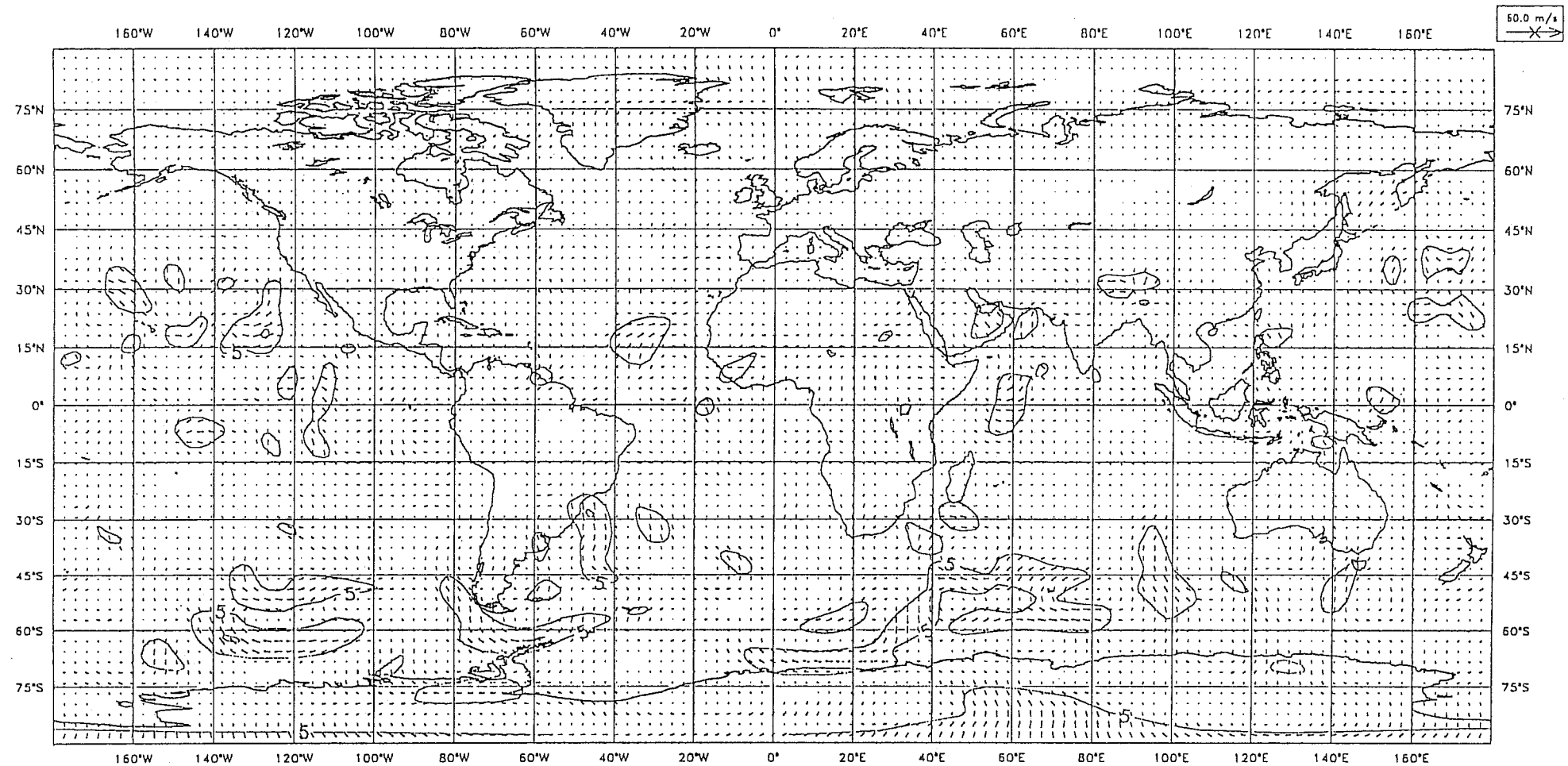

Fig.17 As Fig.16 but humidity TOVS channels HIRS-11 and 12 have been excluded from the assimilation using TOVS radiances.

7. CONCLUSION

A three/four-dimensional variational analysis scheme, which uses radiances directly, has been developed. The scheme has been validated and shown to work well. It successfully combines retrieval and analysis, and produces balanced fields.

The first results show a good agreement with OI analyses in oceanic areas. We expect improvements in the near future from the introduction of horizontally varying background error covariances (variances as well as vertical correlations). The specification of height vertical correlations (in 3D/4D-Var) should be improved so that the temperature correlations they imply agree better with those obtained directly from statistics of temperature data.

Future work will show whether or not there is a significant positive impact on forecast quality from the direct use of radiances. Experiments will also be carried out to determine which sub-set of channels to use over land and sea, and to study the impact of the radiances on the tropical wind-field. TOVS data in the tropics are currently not used in ECMWF operational analysis. With a more appropriate tropical mass/wind balance imposed in 3D-Var, we can expect improvements.

There is also a positive impact expected from the use of HIRS channels 11 and 12 on the humidity analysis. Statistics of TOVS minus forecast differences have shown systematic biases in the current ECMWF analyses (*Eyre*, 1992), especially in the sub-tropics, which the radiance data should be able to correct.

The 4D-Var experiments have shown that the evolution of the model over a 24-hour period acts as an additional constraint on the retrieval/analysis. It was shown that 4D-Var is able to infer information on the tropical wind field from the use of the humidity sensitive TOVS channels in an adiabatic experiment. Another important point (studied by *Thépaut et al.*, 1992) is the vertical transport of information in 4D-Var. The vertical structure functions implicitly used by 4D-Var are non-separable and flow dependent, as opposed to those specified in OI and in 3D-Var.

ACKNOWLEDGEMENTS

The authors wish to thank the members of the IFS/Arpège team at Météo-France and at ECMWF, especially Drs. W. Heckley, M. Hamrud and P. Undén. The manuscript was greatly improved by Dr. A. Hollingsworth's helpful comments. We are also very grateful to NOAA/NESDIS, Washington, for providing magnetic tapes with the cloud-cleared radiances and to INRIA (Institut National de Recherche en Informatique et en Automatique), Le Chesnay, France, for providing the minimisation algorithm (M1QN3).

REFERENCES

Andersson, E., A. Hollingsworth, G. Kelly, P. Lönnberg, J. Pailleux and Z. Zhang, 1991: Global observing system experiments on operational statistical retrievals of satellite sounding data. Mon.Wea.Rev., 119, 1851-1864.

Chédin, A. and N.A. Scott, 1985: Initialization of the radiative transfer equation inversion problem from a pattern recognition type approach. Application to the satellites of the TIROS-N series. Advances in Remote Sensing Retrieval Methods, Academic Press, A. Deepak Ed., 495-515.

Courtier, P. and O. Talagrand, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow- water equations. Tellus, 42A, 531-549.

Durand, Y., 1985: The use of satellite data in the French high resolution analysis. Proc. of the ECMWF Workshop on High Resolution Analysis, Reading, 24-26 June 1985, 89-128.

Durand, Y. and M.C. Pierrard, 1989: Use and impact of satellite data in the PERIDOT system. Proc. of the ECMWF/EUMETSAT Workshop on Use of Satellite Data in Operational Numerical Weather Prediction: 1989-1993, Reading, 8-12 May 1989, Vol.1, 107-138.

Eyre, J.R., 1987: On systematic errors in satellite sounding products and their climatological mean values. Q.J.R.Meteorol.Soc., 113, 279-292.

Eyre, J.R., 1989: Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. Q.J.R.Meteorol.Soc., 115, 1001-1037.

Eyre, J.R., 1991: A fast radiative transfer model for satellite sounding systems. ECMWF Tech.Memo.176, pp 28.

Eyre, J.R., G.A. Kelly, A.P. McNally, E. Andersson and A. Persson, 1992: Assimilation of TOVS radiance information through one-dimensional variational analysis. Submitted to Q.J.R.Meteorol.Soc. Also available as ECMWF Tech.Memo.187, pp 31.

Eyre, J.R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. ECMWF Tech.Memo.186, pp 27.

Eyre, J.R. and A.C. Lorenc, 1989: Direct use of satellite sounding radiances in numerical weather prediction. Meteorol.Mag., 118, 13-16.

Fleming, H.E., M.D. Goldberg and D.S. Crosby, 1986: Minimum variance simultaneous retrieval of temperature and water vapor from satellite radiance measurements. Proc. of Second Conference on Satellite Meteorology - Remote Sensing and Apllications, Williamsburg, Amer.Meteorol.Soc., 20-23.

Flobert, J-F., E. Andersson, A. Chédin, A. Hollingsworth, G. Kelly, J. Pailleux and N.A. Scott, 1991: Global data assimilation and forecast experiments with the Improved Initialization Inversion method for satellite soundings. Mon.Wea.Rev., 119, 1881-1914.

Gallimore R.G. and D.R. Johnson, 1986: A case study of GWE satellite data impact on GLA assimilation analysis of two ocean cyclones. Mon.Wea.Rev., 114, 2016-2032.

Halem, J., E. Kalnay, W.E. Baker and R. Atlas, 1982: An assessment of the FGGE satellite observing system during SOP-1. Bull.Am.Meteorol.Soc., 63, 407-429.

Heckley, W.A., P. Courtier, J. Pailleux and E. Andersson, 1992: On the use of background information in the variational analysis at ECMWF. In this volume.

Hoffman, R.N., 1983: Three-dimensional inversion of satellite observed radiances: A proposal. Preprints, Fifth Conf. on Atmospheric Radiation, Baltimore, Maryland, Am.Meteorol.Soc., 43-46.

Hoffman, R.N. and T. Nehrkorn, 1989: A simulation test of three-dimensional temperature retrievals. Mon.Wea.Rev., 117, 473-494.

Kelly, G.A.M., 1978: Interpretation of satellite cloud mosaics for southern hemisphere analysis and reference level specification. Mon.Wea.Rev., 106, 870-889.

Kelly, G. and J. Pailleux, 1988: Use of satellite vertical sounder data in the ECMWF analysis system. ECMWF Tech.Memo.143, pp 46.

Kelly, G., E. Andersson, A. Hollingsworth, P. Lönnberg, J. Pailleux and Z. Zhang, 1991: Quality control of operational physical retrievals of satellite sounding data. Mon.Wea.Rev., 119, 1866-1880.

Le Dimet, F.X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus, 38A, 97-110.

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

Lönnberg, P., 1989: Quality control and filtering of satellite data. Proc. ECMWF/EUMETSAT Workshop on the Use of Satellite Data in Operational Numerical Weather Prediction: 1989-1993. Reading, 8-12 May 1989, Vol.1, 61-80.

McMillin, L.M. and C. Dean, 1982: Evaluation of a new operational technique for producing clear radiances. J.Appl.Meteorol., 21, 1005-1014.

Pailleux, J., 1989: Design of a variational analysis: Organisation and main scientific points. Computation of the distance to the observations. ECMWF Tech.Memo.150, pp 23.

Pailleux, J., 1990: A global variational assimilation scheme and its application for using TOVS radiances. Preprints, WMO Int.Symp. on Assimilation of Observations in Meteorology and Oceanography, Clermont-Ferrand, 9-13 July 1990, WMO Report, 325-328.

Pailleux, J., 1992: Organization of 3D-Var within the Arpège/IFS cooperation. Future plans at Météo-France. In this volume.

Pailleux, J., W. Heckley, D. Vasiljevic, J-N. Thépaut, F. Rabier, C. Cardinali and E. Andersson, 1991: Development of a variational assimilation system. ECMWF Tech.Memo.179, pp 51.

Parrish, D.I. and J.C. Derber, 1992: The National Meteorological Centre's spectral statistical interpolation analysis system. Mon.Wea.Rev., 120, 1747-1763.

Reale, A.L., D.G. Gray, M.W. Chalfront, A. Swaroop and A. Nappi, 1986: Higher resolution operational satellite retrievals. Preprints, 2nd Conf. on Satellite Meteorology/Remote Sensing and Applications, Williamsburg, 13-16 May 1986, Am.Meteorol.Soc, 16-19.

Rodgers, C.D., 1976: Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. Rev.Geoph.Space.Phys. 14, 609-624.

Smith, W.L., H.M. Woolf, C.M. Hayden, D.Q. Wark and L.M. McMillin, 1979: The TIROS-N operational vertical sounder. Bull.Am.Meteorol.Soc., 60, 1177-1187.

Talagrand, O., 1988: Four-dimensional variational assimilation. Proc. of ECMWF Seminar on Data Assimilation and the use of satellite data, Reading, 5-9 September 1988, Vol II, 1-30.

Thépaut, J.N. and P. Moll, 1990: Variational inversion of simulated TOVS radiances using the adjoint technique. Q.J.R.Meteorol.Soc., 116, 1425-1448.

Thépaut, J.N. and P. Courtier, 1991: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. Q.J.R.Meteorol.Soc., 117, 1225-1254.

Thépaut, J.N., P. Courtier and R.N. Hoffman, 1992: Use of dynamical information in a 4D variational assimilation. In this volume.

Turner, J., J.R. Eyre, D. Jerrett and E. McCallum, 1985: The HERMES system. Meteorol.Mag., 114, 161-173.

Vasiljevic, D., C. Cardinali and P. Undén, 1992: ECMWF 3D variational assimilation of conventional observations. In this volume.