TECHNICAL MEMORANDUM

296

Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence

E. Andersson, M. Fisher, R. Munro and A. McNally

Research Department

November 1999

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Abstract

In any statistical data assimilation scheme the ratio between the observation and background errors fundamentally determines the weight given to the observations. The observation errors are specified directly in terms of the observable quantities. In variational data assimilation schemes these can include satellite measured radiances as well as conventional observations. The background errors, on the other hand, are specified in terms of those quantities that lead to a compact formulation of the background term (the J_b of the variational analysis), viz. balanced vorticity, unbalanced temperature, divergence and surface pressure, and specific humidity. It is not obvious how the magnitudes of these background errors can be compared with the various observation errors. Within the variational analysis the background errors in terms of observed quantities are implied, i.e. not normally computed explicitly. They depend, in general, on the J_b formulation and on the observation operators. In the case of radiance observations this involves the Jacobian of the radiative transfer model which in turn depends on the atmospheric state.

By applying the observation operators of a variational data assimilation scheme to a set of random vectors, drawn from a population whose probability density function is given by the assumed background error covariance matrix, we obtain grid-point fields of background error standard deviations for any observed quantity. These are valuable for diagnosing the response of the data assimilation system to observational data, and for tuning the specified observation and background errors in general. The calculated error standard deviations can be compared with those obtained from studies of innovation statistics (i.e. observed departures from the background). The technique has been applied to a range of observable quantities including the radiance data from both the infrared and microwave instruments of the TIROS operational vertical sounder (TOVS). We used the results for some of the higher-peaking channels to verify that the specified background errors in the recently introduced 50-level version of the ECMWF model are reasonable also in the upper stratosphere, where there are few conventional data. We also found that the operational background errors for humidity were set unrealistically large in some dry subtropical areas.

A case of poor convergence of the variational analysis was found to be due to unrealistically high background errors in terms of one of the humidity sensitive radiance channels (the Meteosat water vapour channel, similar to TOVS channel 12). Excessively large ratios between background and observation errors locally, led to larger than normal eigenvalues of the analysis Hessian - thus increasing the condition number of the minimization problem, with an associated decrease in the rate of convergence of the minimization. The mis-specification of background errors was confined to relatively small areas in the sub-tropics but affected the minimization globally.

1 Introduction

Modern data assimilation schemes combine new information from a wide variety of observations with prior information in the form of a background atmospheric state (Daley 1991). The background is in operational practice often a short-range (typically six-hour) forecast from the previous analysis. The resulting analysis is optimal only if accurate error estimates have been assigned to the observational data and to the background (Lorenc 1986). The relative weight given to the observations and to the background (and therefore the amplitude of the analysis increments) is fundamentally determined by the specified observation and background error variances.

When tuning a data assimilation system, one of the most important aspects is to specify realistic error statistics. The study of 'innovations' (i.e. observed minus background departures) is the most frequently adopted technique (Hollingsworth and Lönn-berg 1986; Lönnberg and Hollingsworth 1986). It provides estimates of the sum of observation error and background error. The statistics of innovations are usually expressed in terms of the observed quantities, i.e. temperature, wind and geopotential height in the case of radiosonde data and in terms of brightness temperatures in the case of satellite measured radiances. Such innovation statistics for 4D-Var have been studied by Järvinen (1999).

In linear statistical interpolation schemes (Lorenc 1981), the observed and the background information were presented to the analysis in terms of the same physical quantities as the analysis variables, e.g. temperature, wind, humidity and surface pressure, or a linear combination of these. The corresponding error estimates could then easily be compared with those obtained from the innovation statistics. Variational schemes, however, provide much greater flexibility in terms of observation usage (Courtier *et al.* 1998; Rabier *et al.* 1998; Andersson *et al.* 1998; Parrish and Derber 1992). Data on any observed quantity for which a meaningful model equivalent can be computed (e.g. satellite radiances) can in principle be used directly in the analysis without first being converted (or 'retrieved') into analysed quantities (Andersson *et al.* 1994; Derber and Wu 1998; McNally *et al.*, 1999). For such data it is difficult to carry out analysis system tuning based on innovation statistics because of the non-linear relation between observed variables and analysed variables.

In this paper we present a technique by which the specified background error covariance matrix ($\mathbf{B}$) can be transformed to observable quantities. The method, based on the randomization technique discussed by Fisher and Courtier (1995), gives global grid-point values of background error standard deviations in terms of observable quantities. The calculation uses the actual observation operators of the variational analysis. It takes into account the balance constraints (defining cross-correlations between mass and wind variables) built into the variational background term (Derber and Bouttier 1999) and is therefore consistent with the actual background error covariance matrix. We have computed the background error equivalents for the radiance channels of both the infrared and microwave instruments of the TIROS operational vertical sounder (TOVS). We have also applied the technique to two-metre temperature, ten-metre

wind, total column water vapour, geopotential, ozone *et-cetera*.

The four-dimensional variational data assimilation scheme (4D-Var) of the European Centre for Medium-Range Weather Forecasts (ECMWF) has been operational since November 1997 (Rabier *et al.* 1999; Mahfouf and Rabier 1999; Klinker *et al.* 1999). The work presented in this paper provides diagnostics which can help in understanding how different data types influence the analysis, given the specified **B**-matrix and a particular atmospheric state. The need for this type of diagnostic of 4D-Var was emphasized recently by four concurrent developments at ECMWF: (*i*) the extension of the forecast model and the data assimilation to the higher stratosphere (Untch *et al.* 1999); (*ii*) the introduction of data from the new generation microwave instruments (AMSU-A and B) of the NOAA-K polar orbiter (McNally 1999); (*iii*) the assimilation of ozone data and the use of the TOVS ozone channel (Hólm *et al.* 1999) and (*iv*) the experimental use of radiance data from the Meteosat water vapour channel (Munro *et al.* 1999). For each of these developments the calculations of radiance background errors described in this paper proved to be very useful, as statistics based on conventional data alone were insufficient to establish the correct magnitude of the background term. We first present some general features of the background errors in terms of observable quantities. The main focus of the paper, however, is on the particular problem of slow convergence in (*iv*) above. The problem was diagnosed and resolved, using first an estimate of the leading eigenvectors of the 4D-Var Hessian to indicate the geographic location of the problem, second, the largest eigenvalue to indicate the condition number and third, the radiance background error to indicate the reason for the poor convergence.

The method for transforming background errors to observable quantities is described in Section 2. A selection of results is presented in Section 3 followed in Section 4 by a diagnosis of the convergence problem arising in the initial experiments with Meteosat radiance data. Conclusions are drawn in Section 5.

2 Method

ECMWF's current method for analysis and background error estimation was proposed by Fisher and Courtier (1995) with further details described by Fisher (1996). In this section we expand on their suggested randomization method to diagnose the 'effective' background error variance in a variational system.

2.1 Background formulation

The minimization of the variational cost function is performed with respect to a control variable χ which is related to the model variables $\mathbf{x}$ in spectral space via a change of variable

$$\chi = \mathbf{L}^{-1}(\mathbf{x} - \mathbf{x}_\mathrm{b}) \tag{1}$$

where $\mathbf{L}$ is the change of variable operator such that $\mathbf{L}\mathbf{L}^\mathrm{T} = \mathbf{B}$, the vector $\mathbf{x}_\mathrm{b}$ is the background atmospheric state and $\mathbf{B}$ is the background error covariance matrix (both in spectral representation). With these definitions the background term is simply $J_\mathrm{b} = 0.5\chi^\mathrm{T}\chi$ (Courtier *et al.* 1998; Derber and Bouttier 1999). The operator $\mathbf{L}$ itself represents a sequence of calculations which together define $\mathbf{B}$. The resulting $\mathbf{B}$ is not calculated explicitly - its structure can, however, be diagnosed indirectly.

2.2 Randomization method

Fisher and Courtier (1995) suggested that a randomization method can be used to diagnose the effective variances of the $\mathbf{B}$-matrix in 3D-Var. The method allows the calculation of a low-rank estimate $\hat{\mathbf{B}}$, of $\mathbf{B}$, in terms of model variables in grid point space (subscript g). In particular the variances can be estimated by the diagonal of

$$\mathbf{B}_\mathrm{g} \approx \hat{\mathbf{B}}_\mathrm{g} = \frac{1}{N}\sum_{i=1}^{N}(\mathcal{S}^{-1}\mathbf{L}\xi_i)(\mathcal{S}^{-1}\mathbf{L}\xi_i)^\mathrm{T} \tag{2}$$

where ξ_i is a set of N random vectors in control-variable space, whose elements are independent and drawn from a population with zero mean and unit Gaussian variance, and $\mathcal{S}^{-1}$ represents the inverse spectral transform. This randomization estimate takes the actual J_b formulation into account. The variances produced by randomization are noisy if N is small. We have found that N=50 gives satisfactory results. With Gaussian statistics and a sample size N the randomization noise in the estimated standard deviations is $1/\sqrt{2N} = 10\%$ (Barlow 1989, page 78).

The described method applies to 3D-Var, as in Eq. (2) no account is taken of the evolution of background errors in time. In a 4D-Var system it provides a diagnosis of the effective background errors at the beginning of the assimilation period (which currently is 6 hours).

2.3 Transformation to observable quantities

The randomization method can be extended to compute an approximation to the diagonal of $\mathbf{H}\mathbf{B}\mathbf{H}^\mathrm{T}$, the background error in terms of observable quantities, where the operator $\mathbf{H}$ denotes the tangent linear observation operators, linearized around the background state, $\mathbf{x}_\mathrm{b}$. For TOVS radiances, $\mathbf{H}$ includes the tangent linear of the radiative transfer model. Here, as in standard notation (Ide *et al.* 1997), $\mathbf{H}$ also includes the inverse spectral transform, $\mathcal{S}^{-1}$.

$$\mathbf{HBH}^{\mathrm{T}} \approx \mathbf{H}\hat{\mathbf{B}}\mathbf{H}^{\mathrm{T}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{HL}\xi_i)(\mathbf{HL}\xi_i)^{\mathrm{T}} \tag{3}$$

The calculations in Eq. (3) provide maps of background error variances in each of the TOVS radiance channels, for example. Using the actual operators $\mathbf{H}$ and $\mathbf{L}$ of the variational analysis, the main limitation of Eq. (3) is in the restricted sample size, which inevitably leads to some noise in the estimated variances. The term $\mathbf{HBH}^{\mathrm{T}}$ is one of the terms that determine the weight given to observations in a variational analysis scheme, as shown in the analytical expression for the 3D-Var solution (e.g. Lorenc 1986):

$$\mathbf{x}_{\mathrm{a}} - \mathbf{x}_{\mathrm{b}} = \mathbf{BH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{Hx}_{\mathrm{b}}) \tag{4}$$

where $\mathbf{x}_{\mathrm{a}}$ is the analysis, $\mathbf{x}_{\mathrm{b}}$ is the background, $\mathbf{y}$ is the vector of observations and $\mathbf{R}$ is the observation error covariance matrix.

3 Estimates of the background error in TOVS channels

In this section we estimate the background error in TOVS sounding channels. The results have been obtained by applying the observation operators (Courtier *et al.* 1998) to the background error covariance matrix of the ECMWF 50-level data assimilation system operational from March 1999 onwards. The observation operator for satellite radiances is the so called RTTOV-5 radiative transfer model described by Saunders *et al.* (1999). The 50-level forecast and data assimilation system extends to 0.1 hPa (Untch *et al.* 1999), and so provides a background estimate for all TOVS channels, including the highest channels. For surface skin temperature, which is not included in the standard B-matrix, we set a constant error of 5 K over land and ice, and 1 K over open sea. The results are shown as cross sections and maps of the square root of diag($\mathbf{H}\hat{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$), i.e. the standard deviation of background error in terms of a few selected observed quantities. The calculations have been performed at spectral resolution T42. The corresponding grid has 6114 points globally, at approximately 300 km spacing. Higher resolution would be desirable because of better land/sea definition, for example, but would currently be too computationally expensive for operational use.

3.1 Background errors for model variables

The current formulation of the 3D/4D-Var B-matrix assumes that the *correlations* are stationary and homogeneous. The error *variances*, however, vary both in space and in time depending on the volume and the distribution of data in previous analyses (Fisher 1996). In the absence

of observations the error grows asymptotically towards the climatological level of atmospheric variability, as obtained from the ECMWF 15-year reanalysis (ERA-15, Gibson *et al.* 1997). Therefore the diagnostic computations of $\mathbf{H}\hat{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$ are relevant for a particular analysis time. The results presented here refer to 1 May 1999 at 00 UTC. Figures 1 (a) and (b) show zonally averaged meridional cross sections of temperature (K) and specific humidity (gkg^{-1}) background errors for that day, computed using Eq. (2). We can see that the temperature errors in the troposphere increase polewards from a minimum of around 0.6 K on the equator to 1.6 K in high latitudes. The temperature errors increase gradually with height to around 2.5 K at all latitudes in the upper stratosphere. The assumed temperature errors are probably unrealistically small near land surfaces. The specific humidity errors (Fig. 1b) are largest around 800 hPa in the tropics where they reach 2.0 gkg^{-1}, and decrease towards the poles and with altitude. (Note that this latter cross section is restricted to the interval between 1000 and 100 hPa).

3.2 Background errors for conventional data

As examples of background errors in terms of quantities measured by conventional data, we show in Fig. 1 (c) u-component of wind (ms^{-1}) and (d) geopotential height (m), respectively. The wind errors show a distinct maximum at tropospheric jet level of 4 to 5 ms^{-1} and the level of maximum wind error follows the general slope of the tropopause. Wind errors have a minimum around 40 hPa and increase quasi-linearly with the logarithm of pressure from that level upwards. For geopotential height the observation operator is a function of surface pressure and a vertical integral of temperature and specific humidity (Appendix A of Courtier *et al.* 1998). The cross section (Fig. 1d) shows geopotential background errors generally around 10 m in the tropics, increasing to just over 35 m in the upper troposphere in high latitudes. There is a rapid increase in geopotential background errors with altitude in the upper stratosphere.

Horizontal maps of background error show the influence of the data distribution used in earlier analyses. Geopotential height at the lowest model level, for example, shows lower error over Europe and North America than over the oceans (Fig. 2) due predominantly to the high density of surface pressure observations over land. Noise due to the limited sample size of the randomization is noticeable especially in the Southern Oceans.

These diagnosed background errors can for conventional data be compared with those obtained through study of innovations (Järvinen 1999). The preliminary indications of such comparison are that the specified background errors in 4D-Var are approximately 30 % too large in most parts of the troposphere. A future retuning of the specified statistics is expected to reduce this discrepancy. Work in this area is in progress.

3.3 Background errors for radiances

Examples of the application of Eq. (3) with **H** being the radiance operators for the channels AMSU-A11 and HIRS-6 are shown in Figs. 3 and 4, respectively. AMSU-A11 is a stratospheric temperature sounding channel and HIRS-6 is sensitive to tropospheric temperature and humidity as well as to surface skin temperature. For a detailed description of the characteristics of the various TOVS instruments (HIRS, MSU and SSU) see Smith *et al.* (1979) and for the new TOVS microwave instruments (AMSU-A and B) see Klaes (1995). AMSU-A11 peaks around 20 hPa where the temperature background errors are 1.0 to 1.2 K in the tropics and around 1.4 K in mid and high latitudes (see Fig. 1). The background errors in terms of the vertically relatively broad AMSU-A11 brightness temperature (Fig. 3) are 0.3 K in the tropics and 0.5 to 0.8 K in mid and high latitudes. The low AMSU-A11 background errors are an indication of significant compensation in the vertical of temperature background errors within the layers spanned by the AMSU-11 weighting function. From Fig. 3 and similar plots for the other temperature sounding AMSU channels we can conclude that these data will have larger impact on the analysis at mid and high latitudes than in the tropics.

The background errors in terms of HIRS-6 (Fig. 4) vary over sea from less than 0.3 in the driest sub-tropical air to 0.7 at high latitudes, where the temperature variability is larger, and in a band along the Equator, where the humidity variability is larger. Over land the error is larger the higher the altitude (the Rockies and the Andes for example) and the drier the atmosphere (Sahara and Australia, for example) but it may also be high where humidity variations dominate (e.g. Amazonas). In order to quantify more precisely the relative contribution from each of the three variables: 1) air temperature, 2) humidity and 3) surface skin temperature, we have repeated the calculations three times, each time with the background error for two of the variables set to zero. The result is shown in Fig. 5 for temperature (top), humidity (middle) and surface skin temperature (lower panel), in terms of relative contributions to the total HIRS-6 background error, i.e. the sum of the three panels is equal to one everywhere. We can see that HIRS-6 observations primarily will influence the 3D/4D-Var temperature analysis in the oceanic areas polewards of 30° N and 30° S and in the driest areas of the subtropics. It will primarily influence the humidity analysis in the tropics. The influence of the surface is greatest where the elevation is high and/or the atmospheric humidity is low. The figure illustrates how 3D/4D-Var partitions the information from HIRS-6 observations between temperature and humidity analysis increments, given a particular **B**-matrix and a particular atmospheric state.

The global r.m.s. of background error for HIRS, MSU, SSU, AMSU-A and AMSU-B channels is summarized in Fig. 6. Surface sensing channels generally show larger background error (e.g. HIRS-8, AMSU-A1, A2, A3 and A15, and B1) as do most humidity sensing channels (HIRS-10, 11, 12 and AMSU-B2, B3, B4, B5). Pure temperature channels show larger background error the higher in the stratosphere they peak, as can be clearly seen in the ranges from AMSU-A7 to A14, MSU-3, 4, SSU 1, 2, 3 and HIRS-4, 3, 2, 1. Broader weighting functions (e.g. HIRS-4 and MSU-3) tend to correspond to lower background error, due to compensation of

temperature background errors in the vertical. The calculated range of background errors for the stratospheric temperature sensing AMSU channels (A10 to A14) compares reasonably well with statistics deduced from study of innovations (Table 1). The innovation statistics include the observation error and ought to be larger than the diagnosed background errors. Estimates of the AMSU observation errors (from Saunders *et al.* 1999) are also given in the table. Table 1 thus indicates that the current background errors for AMSU-A10 and A11 may be overestimated while the results for AMSU-A12, A13 and A14 have given us some confidence also in the upper-stratospheric part of the operational **B**-matrix.

Table 1: Comparison between standard deviations of innovation statistics and diagnosed background errors for stratospheric temperature sensing AMSU channels, for 5 to 10 December 1998. The sample size is 124,888. Observation errors are from Saunders *et al.* (1999).

AMSU channel number	A10	A11	A12	A13	A14
Peak of weighting function (hPa)	40-70	20-40	8-20	4-8	1-4
Estimated observation errors (K)	0.35	0.35	0.50	0.75	1.10
Diagnosed background errors (K)	0.36	0.42	0.53	0.72	1.13
St.dev. of innovations (K)	0.34	0.42	0.79	1.21	1.57

3.4 Background errors for a humidity sounding channel

The TOVS radiance observation operator is nearly linear with respect to temperature (Saunders *et al.* 1999) but is considerably less linear with respect to humidity. As a result, the effective background errors for the humidity sensitive channels are strongly dependent on the actual background state. We illustrate this non-linear sensitivity by making linearised calculations of the changes in HIRS-12 radiance resulting from a given change in the humidity profile, where the humidity change is made first in a dry atmosphere and then in a wet atmosphere.

We have calculated the weighted sensitivity $(\partial T_r/\partial q_j)\Delta q_j$ of the HIRS-12 brightness temperature T_r to a change in humidity Δq_j (at each level j) given by $\Delta q_j = 0.1 q_{sat}(T_j)$. The weighted sensitivity of T_r in the dry atmosphere is O(1 K) and is shown by the open bars in Fig. 7. By contrast, in an atmosphere with the same temperature structure (T_j) but with a much moister equatorial humidity profile, the profile of sensitivity is quite different and is much smaller, O(0.2 K), confined to the upper troposphere, as shown by the black bars in Fig. 7. In other words, for a dry sub-tropical profile the open bars in Fig. 7 show that a 10 % increase in relative humidity anywhere in mid-troposphere will give a large, O(1 K), change in HIRS-12. By contrast, for a moist equatorial profile, the black bars in Fig. 7 show that a 10 % increase in relative humidity anywhere below 300 hPa will give almost no change in HIRS-12. In such a moist atmosphere, HIRS-12 is sensitive mainly to the humidity above 300 hPa, because humidity changes at lower levels are masked by the humidity in the overlying layers.

This non-linear behaviour creates a very complicated field of background error for HIRS-12. An example is shown in Fig. 8(a) for parts of the South Atlantic. The dry and the moist

profiles of Fig. 7 are indicated in Fig. 8(a) by 'D' and 'M', respectively. Lowest background errors tend to occur where the humidity is high as at point 'M', and vice versa at point 'D', in accordance with the sensitivities shown in Fig. 7. These characteristics of the simulated HIRS-12 data in combination with crudely specified background errors for specific humidity in dry regions can in some cases result in very large estimated background errors for HIRS-12. The maximum in Fig. 8(a) is 17.3 K (over Sahara) and there have been cases in excess of 30 K (not shown). Statistics of HIRS-12 innovations indicate that such very large background errors are spurious and should be corrected by improving the specification of humidity background errors. It is conceivable also that further improvements to the radiative transfer model could reduce its large sensitivity to small humidity perturbations in dry atmospheres. Further work on these issues is required.

4 A case of poor 4D-Var convergence

The successful direct assimilation of TOVS radiance data in 4D-Var is currently being extended to other observations, such as the radiance data from the Meteosat water vapour sensor. There are presently two Meteosat platforms in geostationary orbit (over 0° West and 63° East, respectively) each providing full disc coverage every 60 minutes. The characteristics of the Meteosat water vapour channel are very similar to those of HIRS channel 12, i.e. the channel is predominantly sensitive to upper tropospheric humidity. Ongoing experiments with 4D-Var use frequent radiance data from Meteosat, with the aim of improving the tropical analysis of wind (by exploring 4D-Var's ability to track humidity features, Daley 1995) as well as to improve the humidity analysis itself. Preliminary results of this work have been documented by Munro *et al.* (1999).

In our initial experimentation with the Meteosat water vapour channel we found that the addition of Meteosat radiance data altered the analysis in areas far away from that observed by the Meteosat satellites. Analysis differences in excess of 1 K appeared in parts of the South Pacific and in the Arctic, for example. Significant analysis differences, that were not due to differences in data rejection, were also found in data dense areas such as North America. These unexpected results turned out to be symptoms of insufficient convergence of the minimization. The minimization was performed as in the operational 4D-Var (Klinker *et al.* 1999) i.e. 50 iterations with simplified physics, followed by 20 iterations with full physics (Mahfouf and Rabier 1999), using the M1QN3 quasi-Newton algorithm (Gilbert and Lemarechal 1989). The same number of iterations was performed whether Meteosat data were used or not. The spurious analysis differences were reduced in experiments with increased number of iterations. The observed behaviour was a clear indication that the conditioning of the 4D-Var problem had been seriously degraded by the inclusion of Meteosat radiance data.

4.1 Preconditioning

The current 4D-Var is preconditioned with respect to the J_b term (Courtier *et al.* 1998; Derber and Bouttier 1999) as suggested by Lorenc (1988). The J_o term (the observation term of the variational cost function) is, however, not preconditioned. Introduction of additional data will therefore tend to deteriorate the conditioning of 4D-Var (Courtier *et al.* 1994). The ideal preconditioning would be to use the Hessian $\mathbf{J}''$ of the analysis cost function, as demonstrated by Thépaut and Moll (1990) in a problem with small dimension. The rate of convergence of the minimization is largely determined by the condition number $\mathcal{C}$ of $\mathbf{J}''$, i.e. by the ratio of its largest and smallest eigenvalues: $\mathcal{C} = \lambda_{max}(\mathbf{J}'')/\lambda_{min}(\mathbf{J}'')$. The Hessian is the sum of the Hessian matrices of J_b and J_o: $\mathbf{J}'' = \mathbf{J}_b'' + \mathbf{J}_o''$, where $\mathbf{J}_b''$, by virtue of the choice of control variable, is equal to the identity matrix, $\mathbf{J}_b'' = \mathbf{I}$, and $\mathbf{J}_o'' = \mathbf{L}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{L}$, where $\mathbf{L}$ (the change of variable operator) has been defined in Eq.(1). Thus,

$$\mathbf{J}'' = \mathbf{J}_b'' + \mathbf{J}_o'' = \mathbf{I} + \mathbf{L}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{L} \tag{5}$$

where $\mathbf{J}_o''$ is a positive semi-definite matrix, whose eigenvalues are consequently all non-negative. Thus, we have $\lambda(\mathbf{J}'') = 1 + \lambda(\mathbf{J}_o'')$. Furthermore, if the number of observations is smaller than the dimension of the control variable, then $\lambda_{min}(\mathbf{J}_o'') = 0$ and $\lambda_{min}(\mathbf{J}'') = 1$, so that in that case $\mathcal{C} = \lambda_{max}(\mathbf{J}'')$. It is important to note that the conditioning does not depend on the observed data values but it does depend on the specified observation and background errors, as well as on the amount and the distribution of data.

In order to clarify formally how the condition number of the variational analysis depends on the observation and background errors and on the data density we have devised a small theoretical example. We consider the analysis of two grid points. There are n uncorrelated observations at each gridpoint, all with the observation error σ_o, i.e. $\mathbf{R}^{-1} = \mathbf{I}/\sigma_o^2$. The observation operator $\mathbf{H}$ is then a $2n$ by 2 rectangular matrix with n rows equal to (1 0) and n rows equal to (0 1) and the product $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} = n\mathbf{I}/\sigma_o^2$. The background at each point is assumed to have an error σ_b with correlation α between the two points, thus defining $\mathbf{B}$. The operator $\mathbf{L}$ can easily be found as it is the symmetric square root of $\mathbf{B}$. With these definitions we compute the Hessian matrix using Eq. (5) and we obtain the condition number $\mathcal{C}$ as the ratio between its two eigenvalues,

$$\mathcal{C} = \frac{\sigma_b^2(1+\alpha) + \sigma_o^2/n}{\sigma_b^2(1-\alpha) + \sigma_o^2/n} \tag{6}$$

which, if the grid points are close, i.e. $\alpha \approx 1$, becomes

$$\mathcal{C} = 2n(\sigma_b/\sigma_o)^2 + 1 \tag{7}$$

Equation (7) shows that the conditioning of the variational problem deteriorates with increasing data density (larger n) and with increasing ratio between background and observation error.

This is the analogous problem to the ill-conditioning encountered in Optimum Interpolation (Lorenc 1981).

In the light of this theoretical result we can now return to the investigation of the original problem. The Meteosat humidity channel observation error was set to 2 K, which was felt to be a cautious yet reasonable value - certainly not small enough to cause the degraded convergence. As experiments with reduced data density were only partly successful we turned to a study of radiance background errors.

The Meteosat water vapour channel is very similar to HIRS-12 and we may use the results presented for HIRS-12 in the previous section, as a proxy for the Meteosat channel. Figure 8(a) showed that the strong non-linearity in the HIRS-12 sensitivity to humidity perturbations locally produced very large radiance background errors. With such a large discrepancy between assumed background errors (> 10 K locally) and observation errors (2 K) the analysis is required to draw very closely to the Meteosat radiance observations, with possible implications for the condition number.

4.2 Eigenvectors of the analysis Hessian

The leading eigenvectors and eigenvalues of the 4D-Var Hessian are routinely computed as part of the esimation of the analysis error (Fisher and Courtier 1995), using the combined Lanczos and conjugate gradient algorithm. The Hessian is the matrix of second derivatives of the variational cost function, and it can be shown that the Hessian (at $\mathbf{x}_a$) is equal to the inverse of the analysis error (Rodgers 1976; Gauthier 1992; Rabier and Courtier 1992; Courtier *et al.* 1994). In our experiments we found that the eigenvalue associated with the leading eigenvector of the analysis Hessian was much larger in the analysis with Meteosat data than without: 4495 compared to 2229 (Table 2). As the condition number is the ratio between the largest and the smallest eigenvalues, this result provided confirmation that the inclusion of Meteosat data had indeed deteriorated the conditioning of the minimization problem.

Table 2: Leading eigenvalues of the 4D-Var Hessian

	Original q Bg Error	Modified q Bg Error
without Meteosat data	2229	2208
with Meteosat data	4495	2232

The structure of the leading eigenvectors provided information on the origin of the problem. Figure 8(c) shows the amplitude in specific humidity at model level 37 (approximately 500 hPa) of the leading eigenvector of the analysis Hessian with Meteosat data. Its maximum amplitude is located near the west coast of Africa, and coincides with one of the maxima in the radiance background error shown in Fig. 8(a). The radiance background errors that locally appeared to be unrealistically large may therefore be the reason for the poor convergence.

To test this hypothesis a simple modification to limit the humidity background errors in very dry areas was developed. The original formulation has been described by Rabier *et al.* (1998). A decision was made to arbitrarily limit the background error to a maximum of 125 % of the background specific humidity above 800 hPa. In so doing the general features of the humidity background error were unaltered, but the extremely high values of radiance background errors were reduced, as shown in Fig. 8(b). The local maximum with the value 10.4 K was reduced to 6.4 K. The success of this minor alteration of the humidity errors is most clearly seen in the reduction of the leading eigenvalue. It was reduced from 4495 to 2232 in the analysis with Meteosat data (Table 2). The new lower value is similar to the leading eigenvalue in an analysis without Meteosat data, which in this case is 2208. We can therefore expect that with the modified humidity background error, the convergence rate will be unaffected by the addition of Meteosat data. This has indeed proved to be the case. The spurious analysis differences that were observed far away from the Meteosat data in the original experiments were removed in experiments with the modification. Only small-amplitude large-scale analysis differences remained outside the Meteosat area.

4.3 Summary

In this section we have demonstrated how the leading eigenvectors of the 4D-Var Hessian can be used to locate geographically the cause of degraded 4D-Var convergence. Maps of $\mathbf{H}\hat{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$, computed with the randomization method described in section 2, revealed an area of mis-specified background errors, in terms of one satellite-measured radiance channel, which coincided with the location of the leading eigenvector. The leading eigenvalue of the Hessian was used to determine the condition number of the 4D-Var minimization problem.

5 Conclusions

A method has been developed to diagnose the variances of background errors for observable quantities which are non-linearly related to the analysis variables. The observation operators (linearized around the background model state) are applied to a set of random vectors drawn from a population with the probability density function of the background error covariance matrix. One thereby obtains global maps of the effective background error for all observed quantities used by the variational analysis scheme - both conventional data and satellite measured radiances, for example.

We have demonstrated how the computed radiance background errors helped explain a case of poor convergence in 4D-Var, where a small alteration to the humidity background errors resolved the problem. We showed that the modified humidity background errors reduced the leading eigenvalue of the cost function Hessian and the rate of convergence of the minimization

improved. The leading eigenvectors of the Hessian were used to pin-point the geographical location of the cause of the problem.

The diagnosed background errors have also been used to tune the setting of observation and/or background errors, by comparison with innovation (observation minus background) statistics. This has proved to be useful as the ECMWF model domain has recently been extended higher into the stratosphere and as new satellite measured radiance observations such as those from the new TOVS microwave instrument AMSU-A have been introduced. The method has also been used (in Fig 5) to quantify how the information from radiance observations (which may be sensitive to both temperature and humidity) is partitioned between analysis increments in temperature, humidity and surface skin temperature, respectively.

The computations of background errors for observed quantities have been carried out operationally every analysis cycle (i.e. every six hours), since May 1999. The computational cost is relatively small. The reasons why the computations have to be repeated every cycle are: (*i*) the specified background errors depend on the data distribution in previous analyses and (*ii*) non-linear effects of the observation operators will influence the results, as demonstrated for HIRS channel 12 in this paper. The radiance background errors are currently used operationally to define rejection limits in the so-called background check, which checks the measured radiances against radiances computed from the background atmospheric fields.

The validity of the results presented in this paper is restricted to the beginning of the 4D-Var assimilation period, and to 3D-Var. However, in future work the evolution in time of the effective background errors will be studied. The current method will be extended such that each of the random vectors in Eq. (2) are evolved in time by the tangent linear forecast model. Preliminary results (Andersson and Fisher 1999) show that flow-dependent background errors can be generated with such an approach, at a reasonable cost.

Acknowledgements

We are grateful to Anthony Hollingsworth for suggesting the scope of the paper, for improving its presentation and for valuable discussions that helped clarify the results. Roger Saunders and Graeme Kelly provided valuable insight on the AMSU characteristics and the radiative tranfer model. We thank Adrian Simmons and François Bouttier for their comments on the manuscript. Jocelyn Williams skillfully improved the figures.

References

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J., McNally, A. P., Kelly, G. and Courtier, P. 1994 Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **120,** 627–653

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Branković, C., Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Théput, J-N. and Viterbo, P. 1998 The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part III: Experimental results. *Q. J. R. Meteorol. Soc.*, **124,** 1831–1860

Andersson, E. and Fisher, M. 1999 Background errors for observed quantities and their propagation in time. Proc. ECMWF Workshop on "Diagnosis of data assimilation systems", Reading, 1-4 November 1998, 81–90

Barlow, R. J. 1989 *Statistics: a guide to the use of statistical methods in the physical sciences.*, John Wiley & Sons Ltd.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994 A strategy for operational implementation of 4D-Var, using an incremental approach *Q. J. R. Meteorol. Soc.*, **120,** 1367–1387

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. 1998 The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.*, **124,** 1783–1808

Derber, J. C. and Wu, W. 1998 The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Weather Rev.*, **126,** 2287–2299

Derber, J. and Bouttier, F. 1999 A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus.* **51A,** 195–221

Daley, R. 1991 *Atmospheric data analysis.*, Cambridge atmospheric and space sciences series, Cambridge University Press

Daley, R. 1995 Estimating the wind field from chemical constituent observations: Experiments with a one-dimensional extended Kalman filter. *Mon. Weather Rev.*, **123,** 181–198

Fisher, M. and Courtier, P. 1995 Estimating the covariance matrices of analysis and forecast error in variational data assimilation, ECMWF Tech. Memo. 220 (Available from ECMWF)

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

Gibson, J. K., Kallberg, P., Uppala, S., Hernandez, A., Nomura, A. and Serrano, E. 1997 *ERA Project report series. 1: ERA description*, available by anonymous ftp from www.ecmwf.int, cd pub/era/era15, get description.ps.gz

Gilbert, J. C. and Lemaréchal, C. 1989 Some numerical experiments with variable storage quasi-Newton algorithms. *Math. Prog.*, **B25,** 407–435

Gauthier, P. 1992 Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus*, **44A**, 2-17

Hollingsworth, A. and Lönnberg, P. 1986 The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136

Hólm, E. V., Untch, A., Simmons, A., Saunders, R., Bouttier, F. and Andersson, E. 1999 Multivariate ozone assimilation in four-dimensional data assimilation. Proc. SODA Workshop on "Chemical Data Assimilation", deBilt,9-10 December 1998, 89-94

Ide, K., Courtier, P., Ghil, M. and Lorenc, A. C. 1997 Unified notation for data assimilation: operational, sequential and variational. *J. Met. Soc. Japan*, **75**, 181–189

Järvinen, H. 1999 A study of the innovation and residual sequences in variational data assimilation. Proc. ECMWF Workshop on "Diagnosis of data assimilation systems", Reading, 1-4 November 1998, 91–99

Klaes, K. D. 1995 Preparations for ATOVS data processing in Europe. Tech. Proc. 8th International TOVS Study Conf., Queenstown, New Zealand, 5-11 April 1995, 247–258

Klinker, E., Rabier, F., Kelly, G. and Mahfouf, J. F. 1999 The ECMWF operational implementation of four dimensional variational assimilation. Part III: Experimental results and diagnostics with operational configuration. *Q. J. R. Meteorol. Soc.*, In press

Lönnberg, P. and Hollingsworth, A. 1986 The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, **38A**, 137–161

Lorenc, A. C. 1981 A global three-dimensional multivariate statistical interpolation scheme. *Mon. Weather Rev.*, **109**, 701–721

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

Lorenc, A. C. 1988 Optimal nonlinear objective analysis. *Q. J. R. Meteorol. Soc.*, **114**, 205–240

Mahfouf, J. F. and Rabier, F. 1999 The ECMWF operational implementation of four dimensional variational assimilation. Part III: Experimental results with improved physics. *Q. J. R. Meteorol. Soc.*, In press

McNally, A. P., Derber, J. C., Wu, W. and Katz, B. B. 1999 The use of TOVS level-1B radiances in the NCEP SSI analysis system. *Q. J. R. Meteorol. Soc.*, In press

McNally, A. 1999 Assimilation of AMSU radiances at ECMWF. Tech. Proc. 10th International TOVS Study Conf., Boulder, Colorado, 27 Jan - 2 Feb 1999

Munro, R., Kelly, G., Rohn, M. and Saunders, R. 1999 Assimilation of geostationary water vapour radiance data at ECMWF. Tech. Proc. 10th International TOVS Study Conf., Boulder, Colorado, 27 Jan - 2 Feb 1999

Parrish, D. F. and Derber, J. C. 1992 The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Weather Rev.*, **120,** 1747–1763

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

Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F. 1998 The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124,** 1809–1829

Rabier, F., Järvinen, J., Klinker, E., Mahfouf, J. F. and Simmons, A. 1999 The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, In press

Rodgers, C., D. 1976 Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys. Space Phys.* **14,** 609–624

Saunders, R., Matricardi, M. and Brunel, P. 1999 An improved fast radiative transfer model for assimilation of satellite radiance observations. *Q. J. R. Meteorol. Soc.*, **125,** 1407–1426

Smith, W. L., Woolf, H. M., Hayden, C. M., Wark, D. Q, and McMillin, L. M. 1979 The TIROS-N operational vertical sounder. *Bull. Amer. Met. Soc.*, **60,** 1177–1197

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

Untch, A., Simmons, A., Hortal, M., Jakob, C. and colleagues 1999 Increased stratospheric resolution in the ECMWF forecasting system. Proc. SODA Workshop on "Chemical data assimilation", deBilt 9-10 December 1998

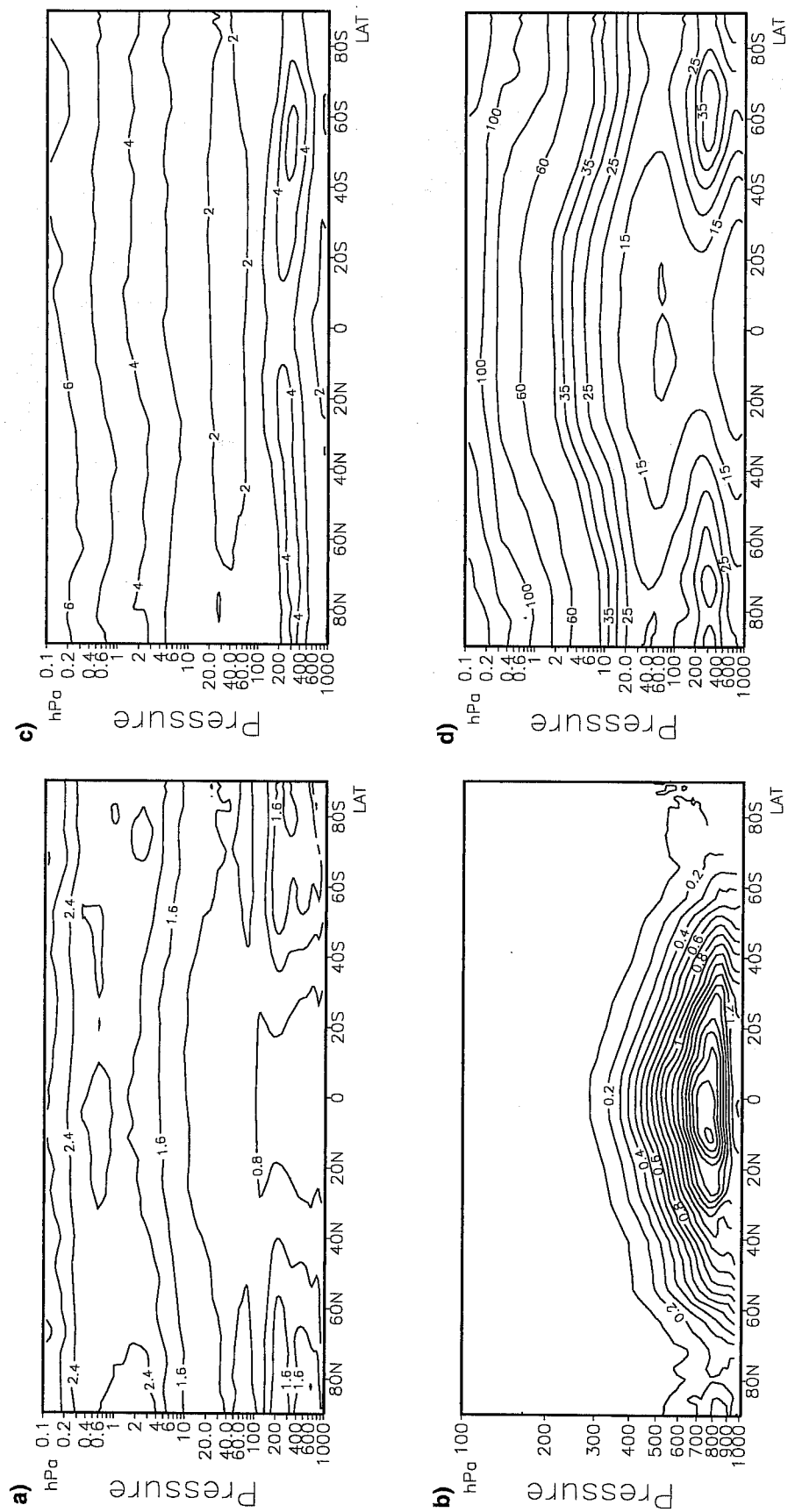

Fig. 1 Zonally averaged background error (19990501-00 UTC) for (a) temperature (0.4K), (b) specific humidity (0.1 g/kg), (c) u-component of wind (1.0 m/s) and (d) geopotential height (contoured every 5m to 40m, then every 20m). The vertical scale is logarithmic in pressure. Note that the vertical extent is reduced to the interval 1015 to 100 hPa in (b).

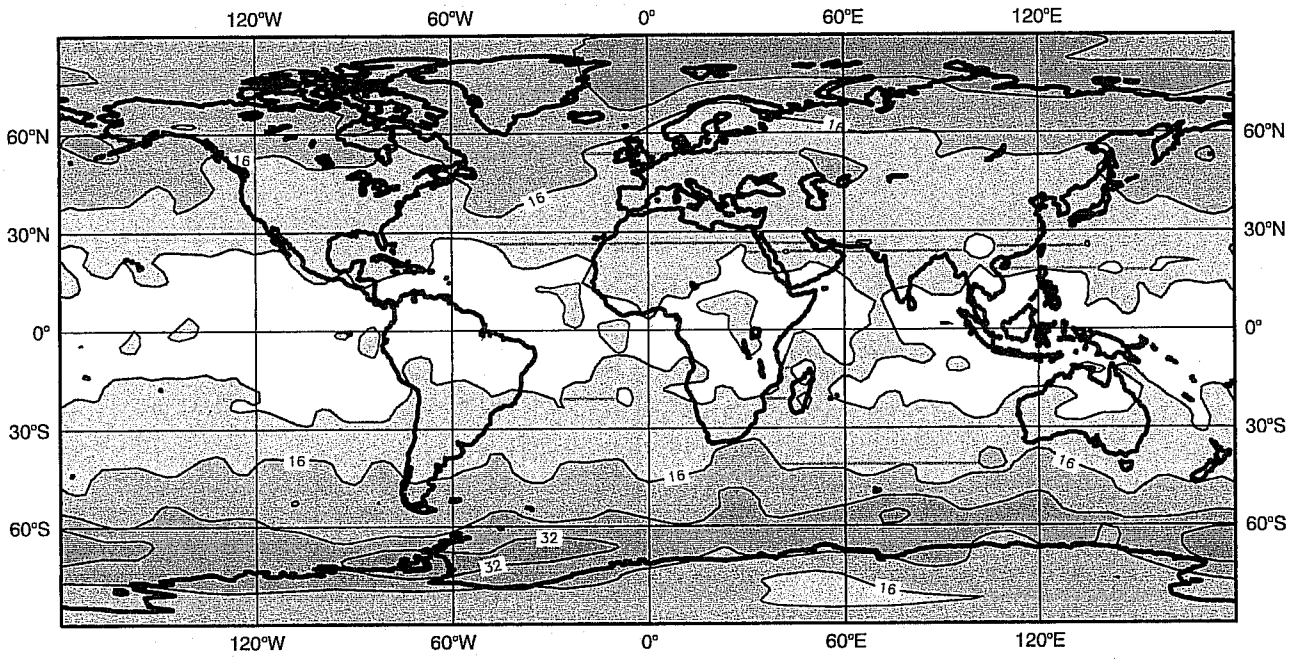

Fig. 2 Lowest model level background error in terms of geopotential height, 19990501-00 UTC, with a contour interval of 8 m. Shading starts at 8 m.

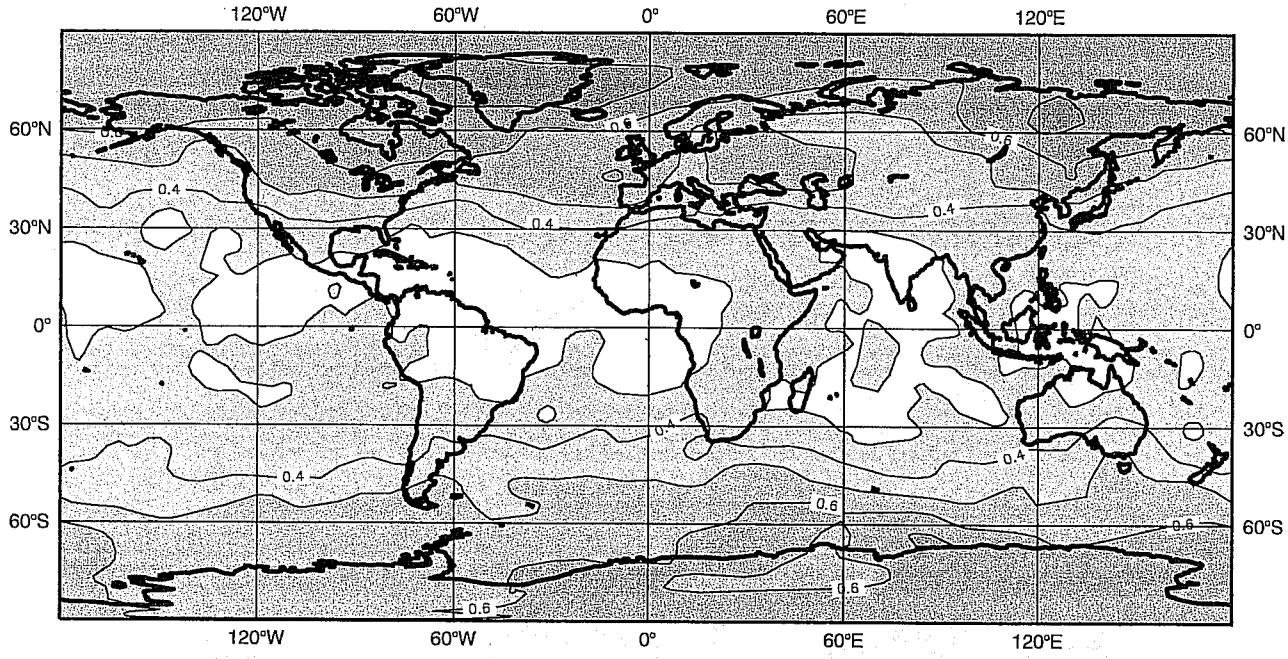

Fig. 3 Background error, 19990501-00 UTC for AMSU channel A11, which is sensitive primarily to stratospheric temperature between 10 and 40 hPa. The contour interval is 0.1 K, with shading starting at 0.3 K.

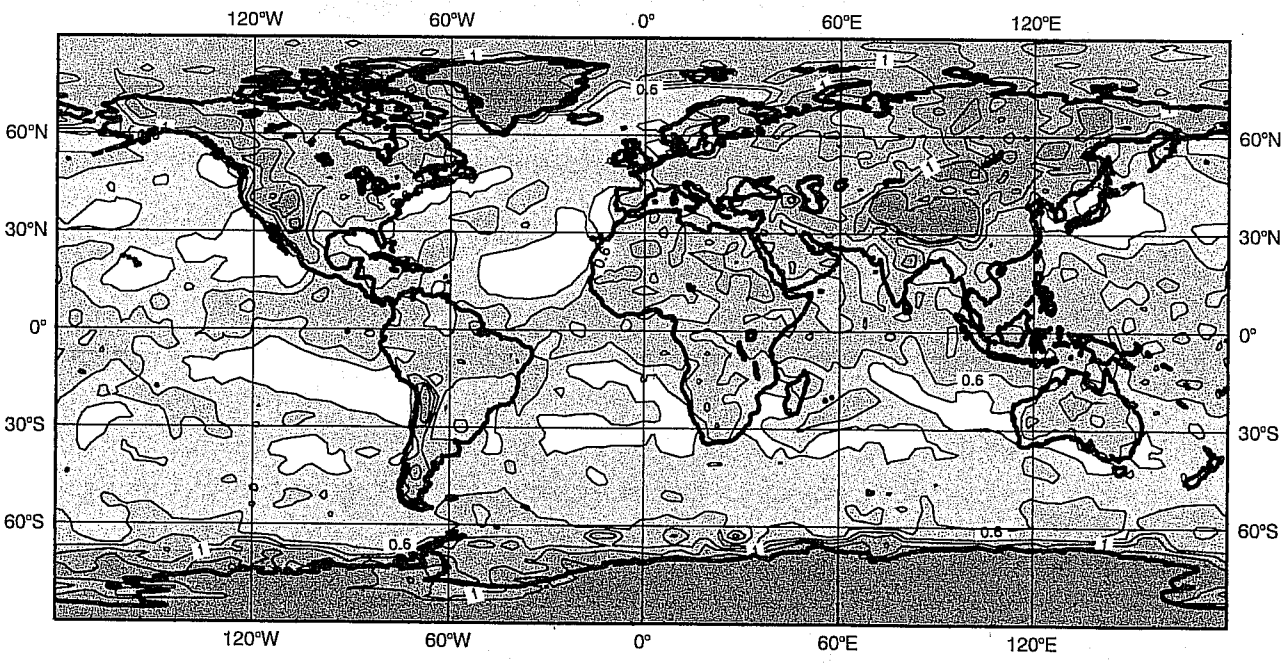

Fig. 4 Background error, 19990501-00 UTC for HIRS channel 6, which is sensitive to tropospheric temperature and humidity as well as to surface skin temperature. The contour interval is 0.2 K, with shading starting at 0.4 K.

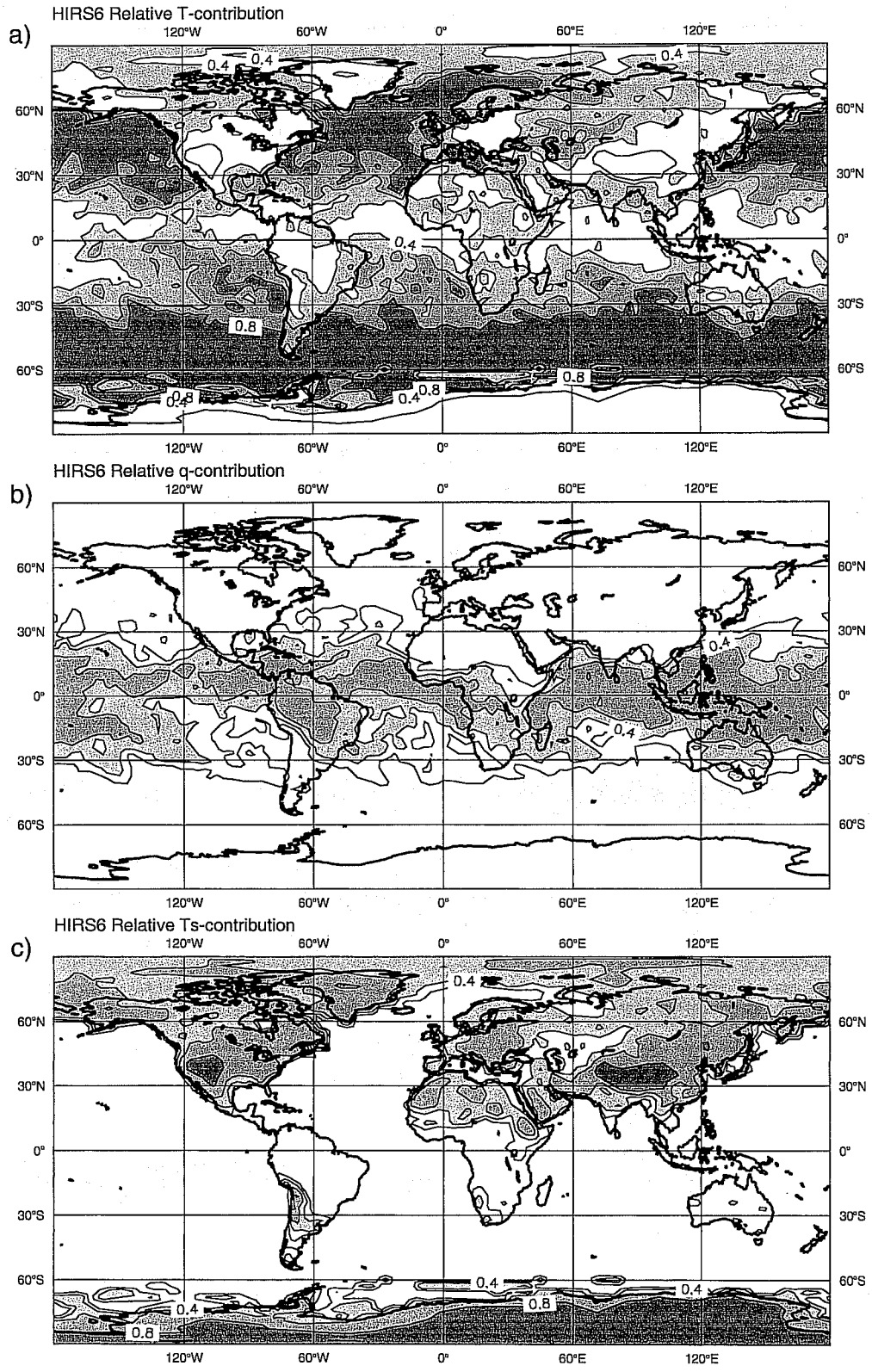

Fig. 5 Relative contribution to the HIRS channel 6 background error from, respectively, air temperature (top), humidity (middle) and surface skin temperature (lower panel). The contours are 0.2, 0.4, 0.6 and 0.8, with shading starting at 0.4. The sum of the three charts is by construction equal to one everywhere.

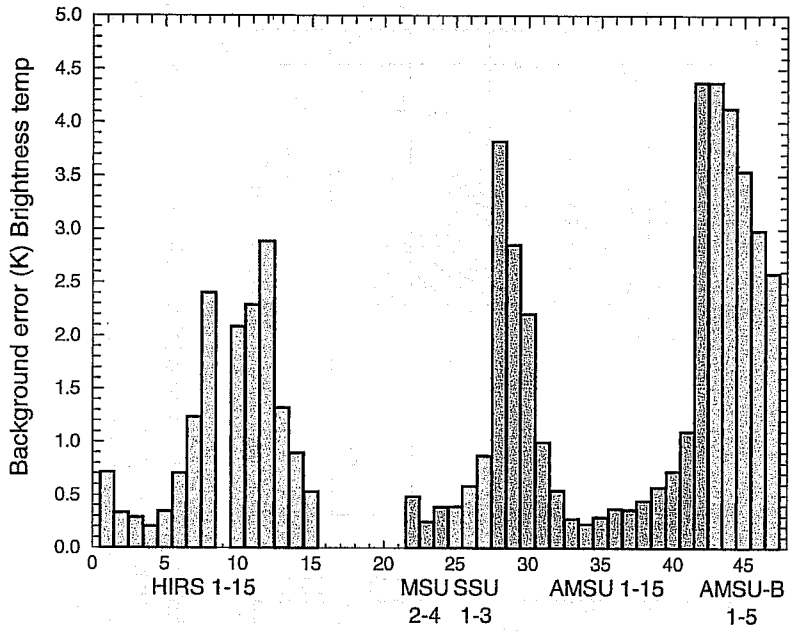

Fig. 6 Global r.m.s. of background errors in terms of brightness temperature (K), for HIRS channels 1-8 and 10-15, MSU 2-4, SSU 1-3, AMSU A1-A15 and AMSU B1-B5. Channels HIRS-9, HIRS-16 to 20 and MSU-1 have not been computed.

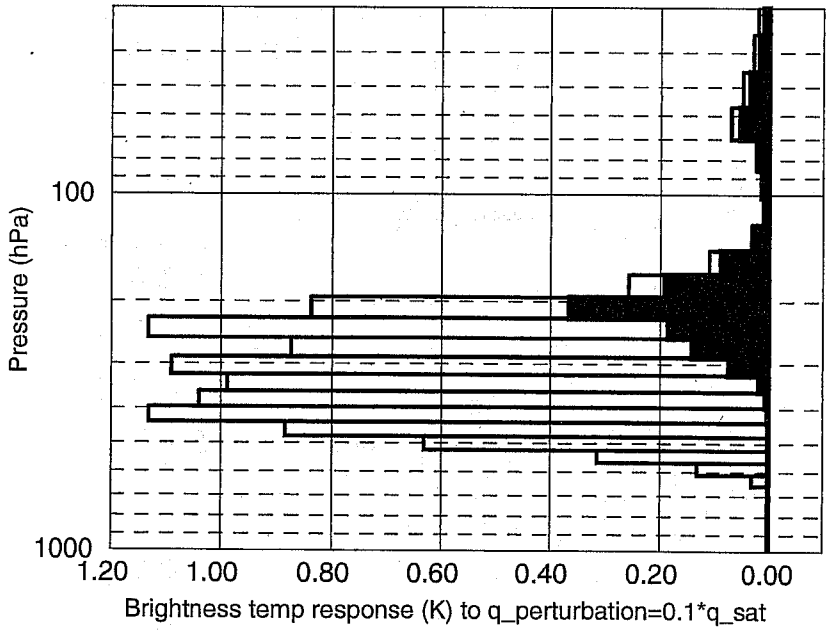

Fig. 7 The weighted sensitivity i.e. (partial T_r/partial q_j)*Delta q_j in TOVS channel HIRS-12 brightness temperature (K), for humidity perturbations Delta q_j=0.1*q_sat(T_j). Open bars correspond to the dry profile, the location of which is marked 'D' in Fig. 8(a), and filled bars correspond to the moist profile, marked 'M'. The linearization temperature profile T_j is the same in both calculations in order to ensure equal profiles of humidity perturbations. The vertical scale is logarithmic in pressure. Each bar represents one layer in the 43-layer representation used by RTTOV-5.

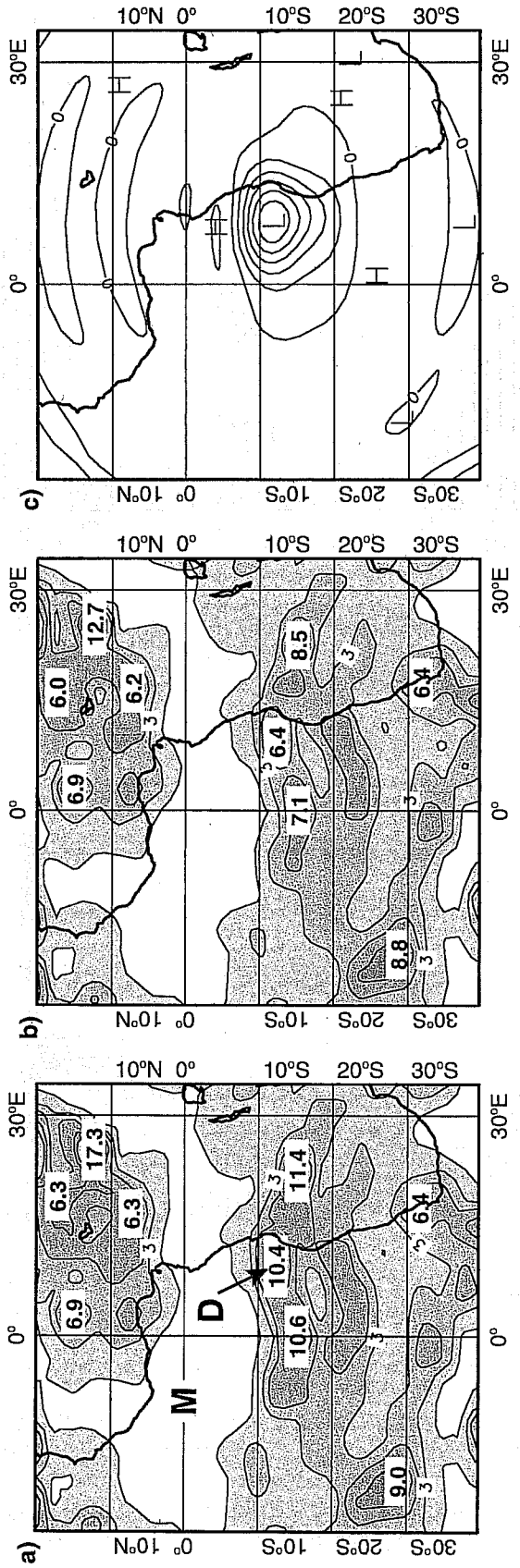

Fig. 8 Background error standard deviation for HIRS channel 12, 19990501-00 UTC, for parts of the South Atlantic, before (a) and after (b) the modification of humidity background errors. Contours are drawn for 2, 3, 4, 6 and 8 K brightness temperature, and maxima exceeding 6 K are indicated. Panel (c) shows the amplitude in specific humidity at model level 37 (approximately 500 hPa) of the leading eigenvector of the analysis Hessian with Meteosat data.. The markers 'D' and 'M' in (a) locate the dry and the humid profiles, respectively, of Fig. 7.