A STUDY OF TOVS AND ATOVS BACKGROUND ERRORS IN RADIANCE SPACE IN A VARIATIONAL ANALYSIS SCHEME

Erik Andersson, Mike Fisher and Anthony McNally
European Centre for Medium-Range Weather Forecasts
Shinfield Park, RG2 9AX Reading, U.K.

SUMMARY

In any analysis scheme the ratio between background error and observation error fundamentally determines the weight given to the observations. Analysis schemes that use observed radiances have observation errors specified directly in terms of radiances (brightness temperatures). The background errors, on the other hand, are specified in terms of those quantities required by the background term (viz. vorticity, specific humidity and unbalanced temperature, divergence and surface pressure). Their correspondence to radiance errors is implied, and for example depends on the atmospheric state and the Jacobian of the radiative transfer model.

An accurate method to diagnose the implied background errors for TOVS and ATOVS radiances has been developed. The method can be used to help improve and tune the background error formulation. This could be especially useful in the stratosphere where there are few conventional data to compare against. It also gives an indication of the expected analysis impact of new data types such as AMSU-A and AMSU-B radiances. Maps can show the relative contribution from temperature, humidity and surface skin temperature background errors for each of the radiance channels.

The computed radiance background errors are also used to define rejection thresholds for quality control checks of TOVS and ATOVS data against model radiances. These checks are now an important part of the ATOVS data assimilation experiments being performed with the ECMWF 4D-Var scheme (see companion paper by McNally et al.). Some intermittent data problems have been found. Other data rejections may be due to rain contamination (affecting MSU-2 and AMSU-A5 for example), cirrus clouds (e.g. HIRS-3 'clear' radiances), broken lower-level clouds (e.g. HIRS-5 'clear' radiances) or improper characterisation of surface emissivity (e.g. AMSU-A5).

1. INTRODUCTION

Modern data assimilation schemes combine the 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 (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 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 improve the realism of the specified error statistics. The study of 'innovations' (i.e. observed minus background departures) is the most frequently adopted technique. It provides estimates of the sum of three components of error: observation error, background error, and forward model error. The obtained statistics are in terms of the observed quantity, i.e. radiances in the case of TOVS/ATOVS data.

In earlier schemes, before variational methods were introduced, the observed and the background information were often 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). Any observed quantity for which a meaningful model equivalent can be computed can in principle be used directly in the analysis without first being converted (or 'retrieved') into analysed quantities (*Lorenc* 1986). Satellite-measured radiances are examples of this (*Andersson et al.* 1994). For such data it has for this reason been difficult to carry out tuning based on innovation statistics.

In this paper we present a technique by which the specified background error covariance matrix can be transformed to observation space. The method gives global grid-point values of background error standard deviations in terms of observed quantities. The calculation takes into account the balance constraints (defining cross-correlations between mass and wind variables) built in to 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 TOVS and ATOVS channels.

The ATOVS channels are not currently used operationally at ECMWF. The on-going pre-operational tests with ATOVS data are described by *McNally et al.* (1999, this volume). It is hoped that the current work will help in diagnosing the data assimilation response to the new ATOVS data. It has also facilitated the development of an effective quality control procedure of the TOVS and ATOVS data against model radiances, in which the rejection thresholds are determined by the radiance background errors. As the ECMWF forecast model domain is about to be extended higher into the stratosphere (*Untch et al.* 1998, see also *Saunders et al.* 1999 in this volume) the innovation technique will be applied to some of the highest-peaking TOVS/ATOVS channels in an effort to tune the background errors of the extended model in the mid/upper stratosphere.

The method to transform background errors to radiance space is described in Section 2. A selection of results is presented in Section 3 followed by conclusions in Section 4.

2. METHOD

The current operational 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 randomisation method to diagnose the 'effective' background error variance in a 3D/4D-Var system.

2.1 Background formulation

The 3D/4D-Var minimisation 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

$$\chi = \mathbf{L}^{-1}(\mathbf{x} - \mathbf{x}_b) \tag{1}$$

where $\mathbf{L}$ is the change of variable operator such that $\mathbf{L}\mathbf{L}^T = \mathbf{B}$. The vector $\mathbf{x}_b$ 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 $J_b = 0.5\chi^T\chi$ (*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 $\mathbf{B}$ does not appear explicitly – its contents can, however, be diagnosed indirectly.

2.2 Randomisation method

Fisher and Courtier (1995) suggested that the randomisation method can be used to diagnose the effective $\mathbf{B}$-matrix in 3D/4D-Var. The method allows the calculation of a low-rank estimate of $\mathbf{B}$, in terms of model variables. In particular the diagonal of $\mathbf{B}$ (the variances) can be estimated by

$$\mathbf{B} \approx \frac{1}{N}\sum_{i=1}^{N}(\mathbf{L}\xi_i)(\mathbf{L}\xi_i)^T \tag{2}$$

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. Randomisation takes the actual J_b formulation into account, and is therefore very accurate. The variances produced by randomisation are noisier if N is small. If, in addition, the inverse spectral transform is applied to each vector $\mathbf{L}\xi_i$, we obtain background error variances on model grid points.

2.3 Transformation to observation space.

The randomisation method can be extended to compute an approximation to $\mathbf{H}\mathbf{B}\mathbf{H}^T$ – the background error in terms of *observed* quantities. The operator $\mathbf{H}$ denotes the Tangent Linear observation operators, linearized around the background state. In this context $\mathbf{H}$ is a vertical operator only, as it is applied at model grid points. For TOVS radiances, $\mathbf{H}$ represents the Jacobian of the radiative transfer model.

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

This requires that $\mathbf{H}$ can be applied to vertical profiles of model variables at model grid points. By the application of Eq. (3) maps of background error variances may be obtained in terms of each of the TOVS and ATOVS radiance channels. Using the actual $\mathbf{H}$ and $\mathbf{L}$ of the variational analysis, the only approximation in Eq. (3) is in the restricted sample size, which inevitably leads to some noise in the

estimated variances. We have found that $N = 50$, which corresponds to a 10% noise level, gives satisfactory results.

The study of $\mathbf{HBH}^T$ can be valuable as it is one of the terms that determine the weight given to observations in a variational analysis scheme, as can be seen from the analytical expression for the 3D/4D-Var solution (e.g. *Lorenc* 1986):

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{Hx}_b) \tag{4}$$

where $\mathbf{x}_a$ is the analysis, $\mathbf{x}_b$ is the background, $\mathbf{y}$ is the observations and $\mathbf{R}$ is the observation error covariance matrix. Related work in an Optimum Interpolation (OI) framework with 1D-Var can be found in the paper by *Harris et al* (1999, this volume).

3. RESULTS

The results presented in this section have been obtained by applying the RTTOV-5 radiative transfer model (*Saunders et al*, 1999, this volume) to the 50-level background error covariance matrix of the ECMWF 4D-Var. The 50-level forecast and data assimilation system (implemented operationally 10 March 1999) extends to 0.1 hPa (*Untch et al* 1998), so there is no need to extrapolate above the model's top, even for the highest TOVS/ATOVS channels. We present maps of the square-root of $\mathrm{diag}(\mathbf{HBH}^T)$, i.e. radiance background error standard deviations, for selected channels – in one case subdivided into separate temperature, humidity and surface skin temperature contributions. The calculations have been performed at spectral resolution T42. The corresponding grid has 6114 points globally, at approximately 300 km spacing. Application of radiance background errors to define quality control rejection thresholds are also presented.

3.1 Illustration of radiance background errors

The current formulation of the 3D/4D-Var $\mathbf{B}$-matrix assumes that the **correlations** are stationary and homogeneous. The error **variances**, however, may vary in both space and in time. The variation depends on the amount and the distribution of data in previous analyses (*Fisher* 1996), modified by an asymptotic error growth towards the climatological level of atmospheric variability, as obtained from the ECMWF 15-year reanalysis (ERA-15). It is therefore envisaged that the computations of $\mathbf{HBH}^T$ will need to be carried out every analysis time, i.e. every six hours. The results presented here refer to 19981210, 00 UTC. Fig. 1 shows a zonally averaged cross section of the temperature (a) and humidity (b) background errors, for that day. We can see that the temperature errors in the troposphere increase polewards from a minimum of 0.5 K on the equator to 1.5 K in polar regions. The temperature errors increase gradually with height to around 5.0 K at 0.1 hPa. The specific humidity errors are largest around 850 hPa in the tropics (nearly 2.0 g/kg), and decrease towards the poles and with altitude. The surface skin temperature background errors are set to 5.0 K over land and sea ice, and to 1.0 K over open sea.

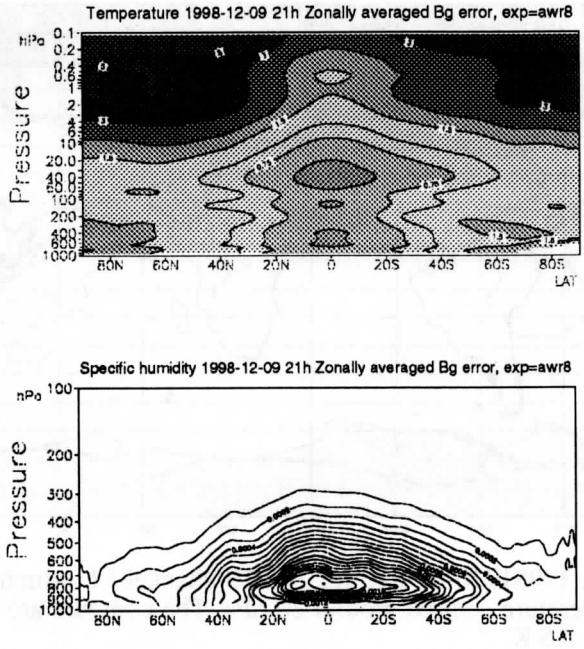

Figure 1: Zonally averaged background error (19981210-00 UTC) for temperature (K), top panel, and specific humidity (g/g), lower panel. Plotted contours are 0.5, 0.75, 1.0, 1.5, 2.0, 3.0, 4.0 and 5.0 K, and every 0.0001 kg/kg, respectively. Note that the vertical extent is reduced to the interval 1015-100 hPa in the lower panel.

The result of the application of Eq. (3) is here exemplified by AMSU-A11 (Fig. 2) which is a stratospheric temperature sounding channel and by HIRS-6 (Fig. 3), which is sensitive to tropospheric temperature and humidity as well as to surface skin temperature. AMSU-A11 peaks around 20 hPa where the temperature background errors are 0.5 to 0.8 K in the tropics and around 1.5 K in mid and high latitudes (see Fig. 1). The corresponding figures in terms of the vertically relatively broad AMSU-A11 brightness temperature 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 is an indication of significant compensation of temperature background errors (i.e. negative correlation) within the layers spanned by the AMSU-11 weighting function. From Fig. 2 and similar plots for the other temperature sounding AMSU channels we can conclude that these data will have their largest impact on the analysis at mid and high latitudes.

The background errors in terms of HIRS-6 (Fig. 3) vary over sea from less than 0.3 in the driest subtropical 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 seems to be larger the higher the altitude (Rockies and Andes for example) and the drier the atmosphere (Arabia and Australia, for example) but it may also be high where humidity variations dominate (e.g. Amazons). 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 the other two variables set to zero. The result is shown in Fig. 4 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 temperature 3D/4D-

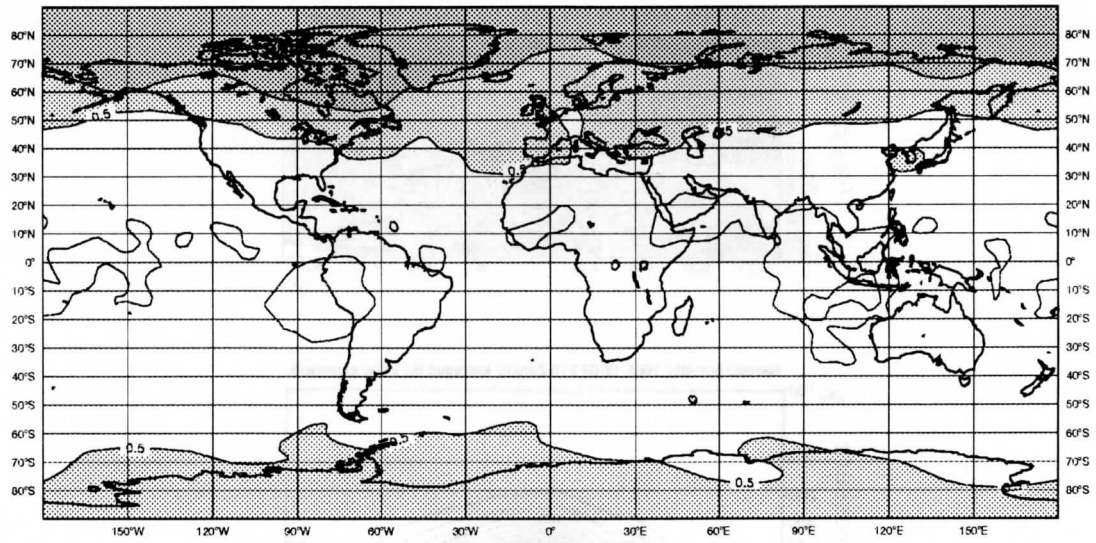

Figure 2: Background error, 19981210-00 UTC for AMSU channel 11, which primarily is sensitive to stratospheric temperature centred around 20 hPa. The contours are 0.3, 0.5 and 0.7 K, with shading starting at 0.5 K.

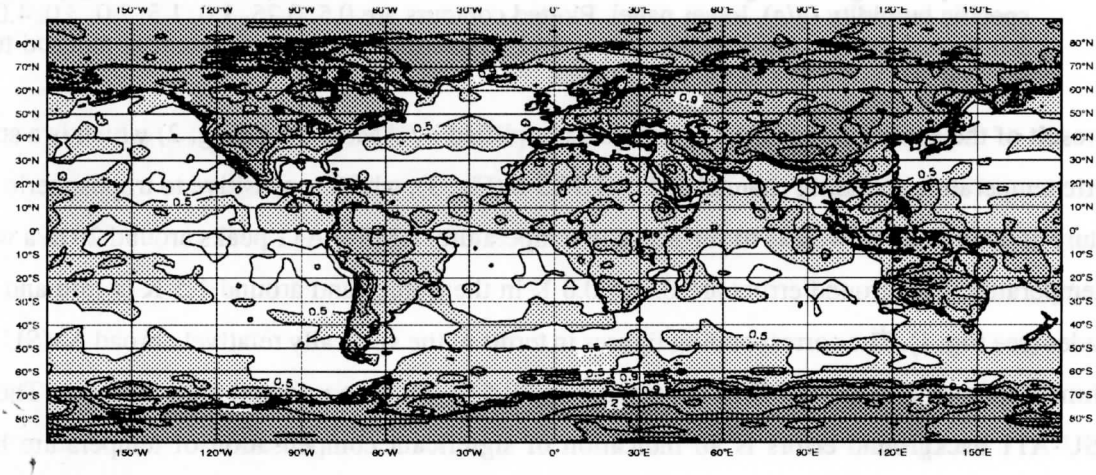

Figure 3: Background error, 19981210-00 UTC for HIRS channel 6, which is sensitive to tropospheric temperature and humidity as well as to surface skin temperature. The contours are 0.3, 0.5, 0.7, 0.8, 1.2, 2.0, 3.0, 4.0 and 5.0 K, with shading starting at 0.5 K.

Var 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 splits the information from HIRS-6 observations between temperature and humidity analysis increments. Surface sensitive channels cannot be used over land as there is currently no skin temperature analysis in operational 3D/4D-Var. A surface skin temperature analysis at TOVS locations has, however, been developed recently and is being tested in the ATOVS assimilation experiments (*McNally et al*, 1999, this volume).

The global r.m.s. of background error for HIRS, MSU, SSU, AMSU-A and AMSU-B channels is summarized in Fig. 5. 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

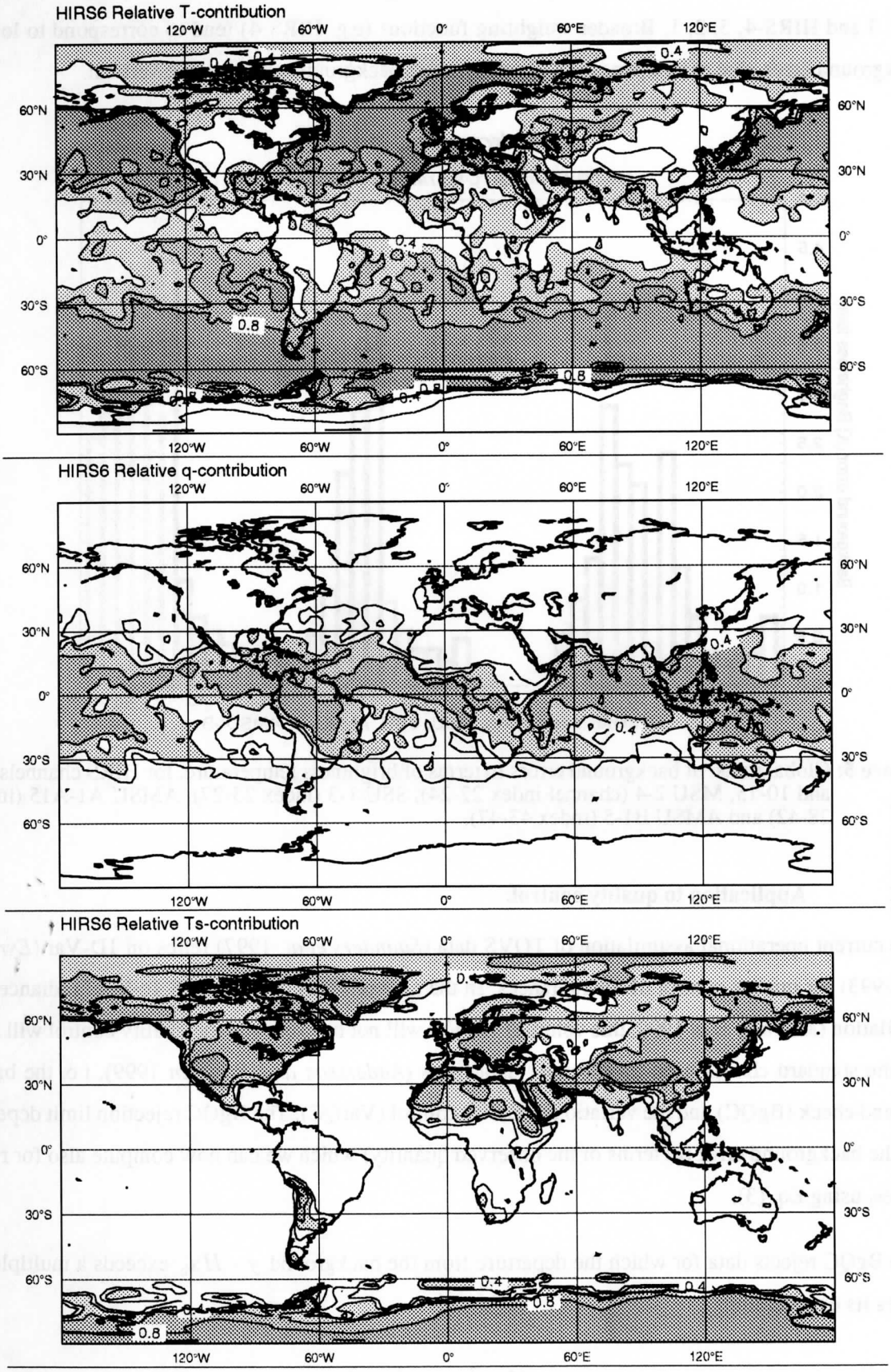

Figure 4: Relative contribution to the HIRS channel 6 background error (19981210-00 UTC) from, respectively, air temperature (top), humidity (middle) and surface skin temperature (lower panel). The contours are 0.2, 0.4, 0.6 and 0.8, with shading starting at 0.4. The sum of the three charts is by construction equal to one everywhere.

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) tend to correspond to lower background error, due to compensation of temperature background errors in the vertical.

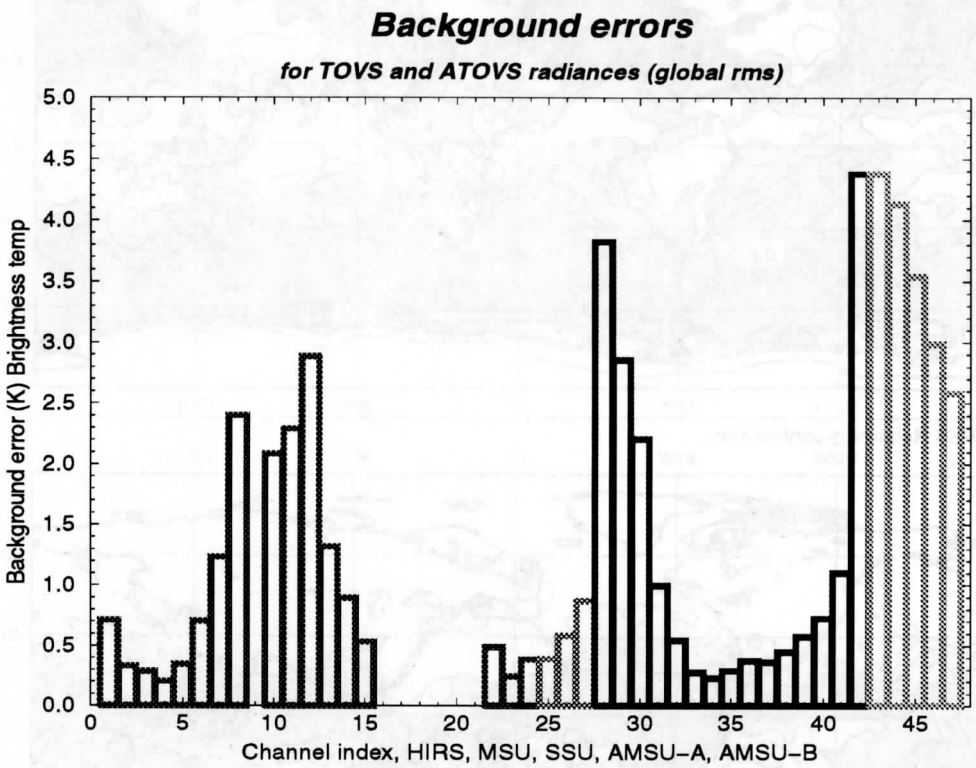

Figure 5: Global r.m.s. of background errors in terms of brightness temperature, for HIRS channels 1-8 and 10-15, MSU 2-4 (channel index 22-24), SSU 1-3 (index 25-27), AMSU A1-A15 (index 28-42) and AMSU B1-5 (index 43-47).

3.2 Application to quality control.

The current operational assimilation of TOVS data (*Saunders et al.* 1997) relies on 1D-Var (*Eyre et al.* 1993) for quality control of the radiances. In the context of TOVS/ATOVS 1c-level radiance assimilation (*McNally et al* 1999, this volume) 1D-Var will not be used, and the quality control will rely on the standard checks used for all other data types (*Andersson and Järvinen* 1999), i.e. the background check (BgQC) and the variational quality control (VarQC). The BgQC rejection limit depends on the background error (in terms of the observed quantity) which we can now compute also for radiances, using Eq. (3).

The BgQC rejects data for which the departure from the background $y - H\mathbf{x}_b$ exceeds a multiple α times its expectation:

$$(y - H\mathbf{x}_b)^2 > \alpha(\sigma_b^2 + \sigma_o^2) \tag{5}$$

where σ_b and σ_o are the background and observation error standard deviations, respectively.

BgQC aims at removing the tails of the data distribution. Fig. 6 shows an example for HIRS channel 3. The figure shows a histogram of observed departures (observed radiance minus radiances computed

from the background) for HIRS channel 3 1c-radiances, accumulated over a four-day period. The BgQC-rejected data are highlighted in bold. We can see that the check has successfully removed the tails of the distribution. More data are rejected on the cold side of the distribution than on the warm side, which could be a sign of remaining cloud contamination affecting some of the data. The horizontal distribution of the rejected data is shown in Fig. 7. We can see that many 'cold rejections' (triangles) occur in areas of deep convection (Indonesia, Amazons and in the Tropical convergence zone in the Atlantic and the East Pacific). The 'warm rejections' (circles) are almost exclusively in the North Atlantic area, suggesting a regional bias in either the data or the model, at this time. The figure also shows that there has been a data problem with a section of one orbit in the southern Atlantic. Data rejections in other channels (not shown) may be due to rain contamination (affecting MSU-2 and AMSU-A5 for example), broken lower-level clouds (e.g. HIRS-5) or improper characterisation of surface emissivity (e.g. AMSU-A5).

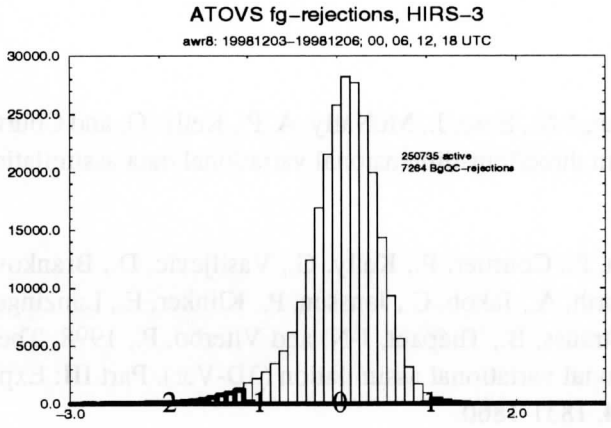

Figure 6: Histogram of observation-minus-background departures of HIRS channel 3 brightness temperatures accumulated over a four-day period (19981203 to 19981206). Data rejected by BgQC are shown in bold. The sample is 250,735 of which 7,264 data were rejected by BgQC.

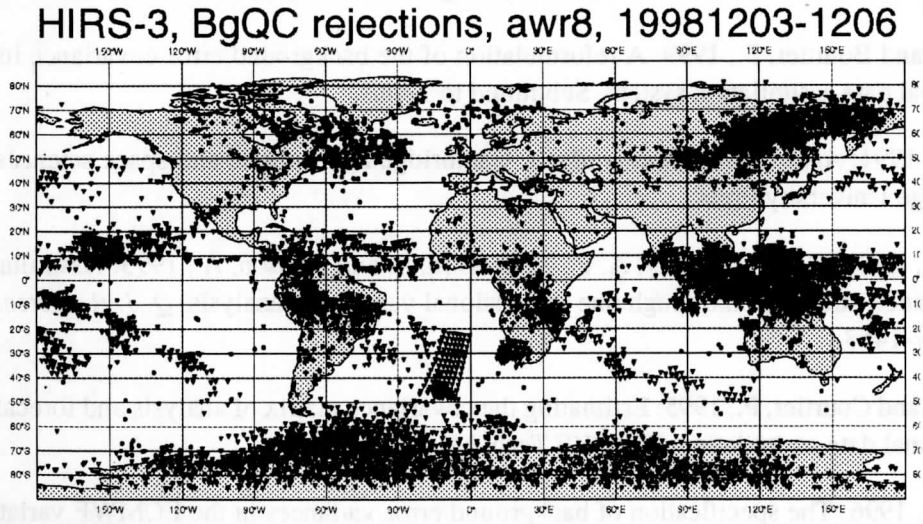

Figure 7: BgQC rejections of HIRS-3 data in the period 19981203 to 19981206. Triangles show negative departures and filled circles show positive departures from the background. The number of rejections is 7,264, corresponding to 2.9% of the sample.

4. CONCLUSIONS

A method to diagnose background errors for **observed quantities** has been developed. The observation operators (linearized around the background model state) are applied to a set of random vectors drawn from a population with the p.d.f. of the background error covariance matrix. The method can be used to tune the setting of observation and/or background errors, by comparison with innovation statistics (observation minus background). This may be especially useful as the ECMWF model domain recently has been extended higher into the stratosphere and new observations such as AMSU-A and B are introduced.

The radiance background errors are also used to define rejection limits in BgQC – the check of measured radiances against radiances computed from the background atmospheric fields. Such checks now constitute an important part of the experimental assimilation of TOVS and ATOVS 1C-radiances, at ECMWF (*McNally et al*, 1999, this volume).

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., Brankovic, C., Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Thépaut, 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 Järvinen, H., 1999: Variational quality control. *Q. J. R. Meteorol. Soc.* **125**, 697-722.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, 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. and Bouttier, F., 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. Submitted to *Tellus*.

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

Eyre, J. R., Kelly, G. A., McNally, A. P., Andersson, E. and Persson, A., 1993: Assimilation of TOVS radiance information through one-dimensional variational analysis. *Q. J. R. Meteorol. Soc.*, **119,** 1427-1463.

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

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.

Järvinen, H. and Undén, P., 1997: Observation screening and first guess quality control in the ECMWF

3D-Var data assimilation system. ECMWF Tech. Memo. 236. Available from ECMWF.

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

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.

Saunders, R. W., Andersson, E., Kelly, G., McNally, A. and Harris, B., 1999: The direct assimilation of TOVS radiances at ECMWF. Tech. Proc. 9th International TOVS Study Conf., Igls, Austria, 20-26 February 1997, 417-428.

Saunders, R., Andersson, E., Kelly, G., Munro, R. and Harris, B., 1999: Recent developments at ECMWF in the assimilation of TOVS radiances. In this volume.

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

TECHNICAL PROCEEDINGS OF THE TENTH INTERNATIONAL ATOVS STUDY CONFERENCE

Boulder, Colorado

27 January - 2 February 1999

Edited by

J. Le Marshall and J.D. Jasper

Bureau of Meteorology Research Centre, Melbourne, Australia

Published by

Bureau of Meteorology Research Centre

PO Box 1289K, GPO Melbourne, Vic., 3001, Australia

December 1999