ASSIMILATION OF TOVS DATA : PROBLEMS AND POSSIBILITIES

J R Eyre
European Centre for Medium-range Weather Forecasts,
Shinfield Park, Reading, UK

1. INTRODUCTION

Over recent years there has been a considerable convergence in approach to the problem of assimilating satellite sounding data in numerical weather prediction (NWP). Most of the major NWP centres have moved or are planning to move towards systems through which radiance information is assimilated more or less directly into the NWP system, using methods based on the principles of variational analysis. Some of these systems are described in following papers (in this volume). The details of implementation vary from centre to centre, being subject to different operational constraints and development histories. However, the underlying principles are the same, and it is the aim of this paper to present the common aspects of the problems faced and the methods developed, and thus to serve as an introduction for the papers which follow.

Section 2 provides a brief introduction to the TIROS Operational Vertical Sounder (TOVS) instruments and their data. Section 3 considers the theoretical problems encountered in using TOVS data effectively in NWP, which have led to the adoption of the variational approach to the assimilation of radiance information, as discussed in Section 4. Section 5 describes the problems of radiance monitoring and tuning. These have been illustrated with examples from the ECMWF system but will be common to similar systems elsewhere. Section 6 discusses some future developments in operational satellite sounding data and their assimilation.

2. TOVS

TOVS is a set of three complementary instruments for sounding atmospheric temperature and humidity. It was first flown on TIROS-N launched in 1978 and has been on board each subsequent satellite in the NOAA series (NOAA-6 to NOAA-12 up to now, with one more satellite planned in the series). TOVS comprises:

- the High-resolution Infra-red Radiation Sounder (HIRS), a 20-channel filter radiometer (19 infra-red and 1 visible channel) with a horizontal resolution of about 40 km, providing information primarily on temperature in the troposphere and lower/mid stratosphere and on tropospheric humidity, but also on surface temperature, clouds and total column ozone,

- the Microwave Sounding Unit (MSU), a 4-channel microwave radiometer with a horizontal resolution of about 170 km, for sounding temperature in the troposphere and lower stratosphere, and

- the Stratospheric Sounding Unit (SSU), a 3-channel infra-red, pressure modulation radiometer with a horizontal resolution of about 200 km, for sounding temperature in the mid/upper stratosphere.

The temperature weighting functions of these channels are shown in Fig.1 and the scan patterns of HIRS and MSU for two consecutive orbits in Fig.2. For further information on these instruments, see *Smith et al.* (1979) and *Schwalb* (1978).

Global data from these instruments are processed in near real-time by NOAA/NESDIS in Washington DC. The radiance data are calibrated and earth-located, and they undergo a number of pre-processing operations including "limb correction" (i.e. adjustment to the radiances which would have been measured from a nadir view) and "cloud clearing" (i.e. detection and, if possible, adjustment for the effects of clouds on the infra-red channels). For further details, see *Smith et al.* (1979) and *McMillin and Dean* (1982). The clear-column radiances so produced are used by NOAA/NESDIS to retrieve temperature and humidity profiles, which are disseminated to users in a number of forms. The retrievals have been distributed in SATEM messages on the Global Telecommunications System (GTS) for many years, with a horizontal spacing of 500 km. More recently, data have been distributed to some users at higher horizontal resolution and in messages which also include the clear-column radiance data. At ECMWF, we are currently making use of the so-called "120 km BUFR TOVS" messages, which have recently become operational. Fig.3 illustrates the coverage obtained with these data from two satellites for a typical 6-hour period.

3. PROBLEMS WITH ASSIMILATION OF CONVENTIONAL RETRIEVALS

TOVS data have had, and continue to have, a large and positive impact on the analyses and subsequent forecasts of the state of the atmosphere in the Southern Hemisphere, where data of other types are very sparse. However, over the last 10 years, it has become increasingly difficult to demonstrate consistent positive impact from TOVS in the Northern Hemisphere. Why is this?

Since TOVS data first became available in 1978, there have been considerable improvements in NWP models and their data assimilation systems. They have yielded increasingly accurate short-range forecasts, which act as the "background" field into which new data are assimilated. This is true not only for the data-rich areas of the Northern Hemisphere continents but also for the northern oceans, through the ability of the NWP models to propagate accurate information from the continents out over the oceans with considerable skill. (This aspect is illustrated in Section 5). With an improved background field, it is necessary to treat the new data, whatever their source, with great care; erroneous observations, or good observations interpreted in an erroneous way, can easily degrade a background field which is usually accurate. This is true for all observations in principle, but in practice it has been particularly troublesome for TOVS data for the reasons explained below.

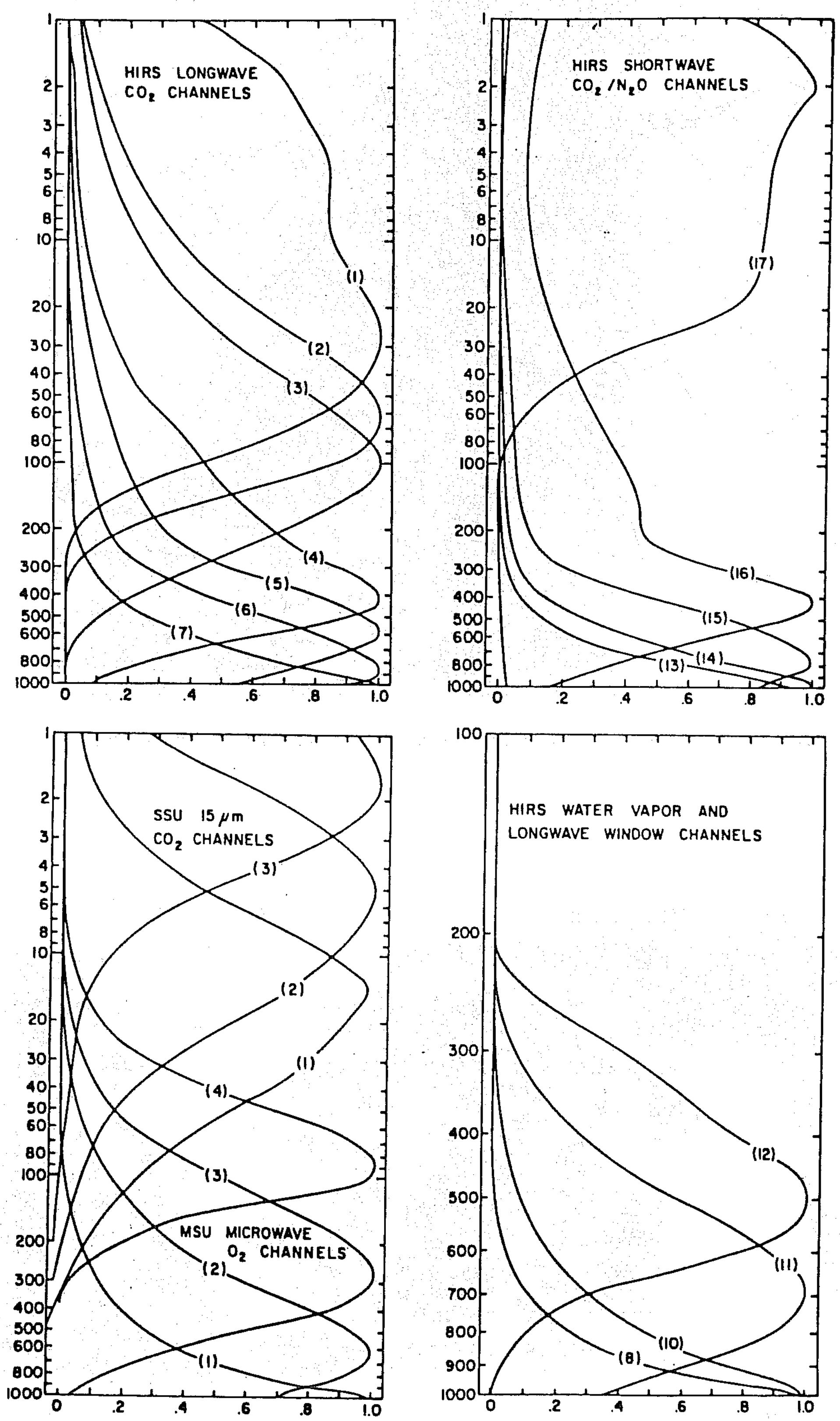

Fig. 1 TOVS weighting functions (taken from *Smith et al.*, 1979).

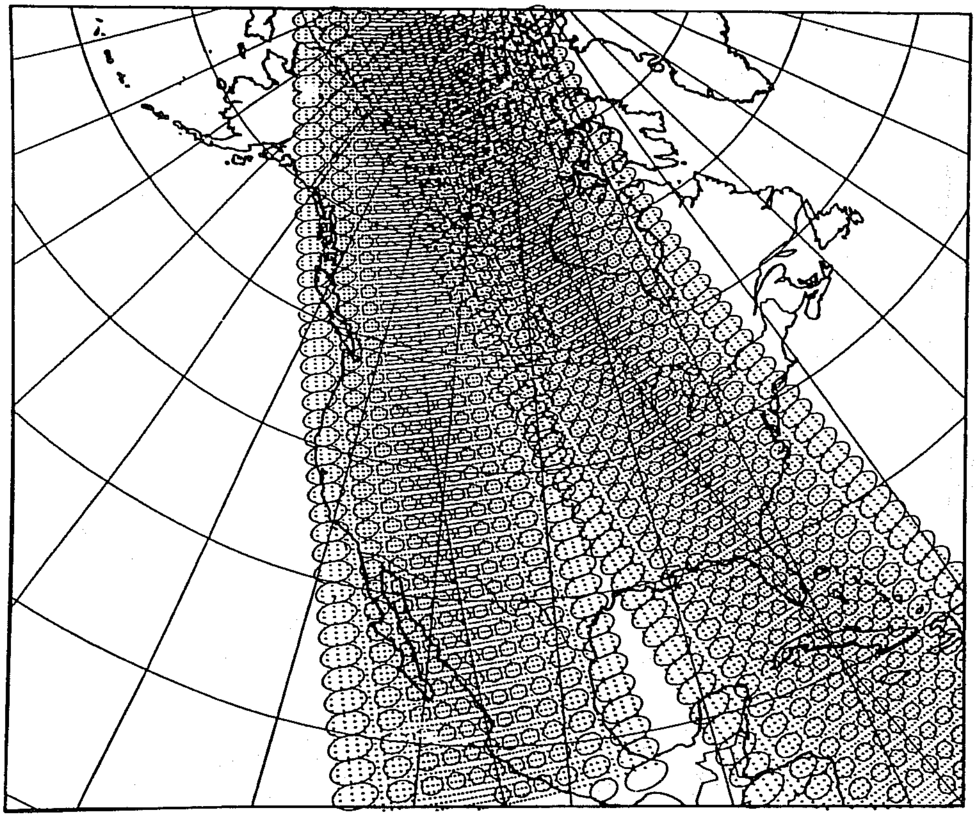

Fig. 2 HIRS and MSU scan patterns for two consecutive orbits (taken from *Smith et al.*, 1979).

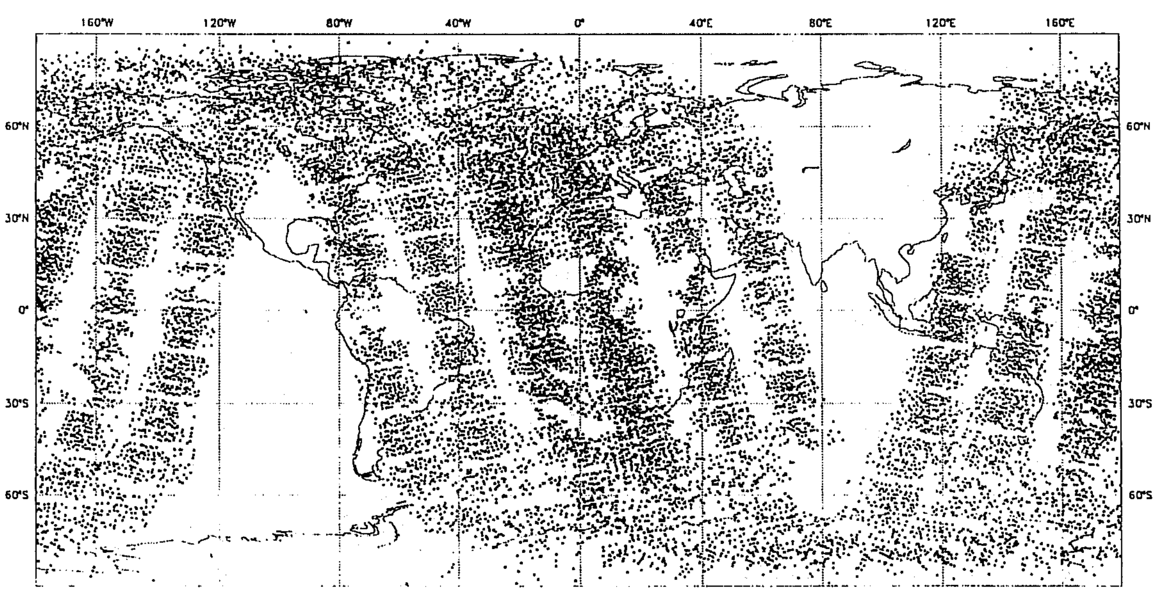

Fig. 3 Coverage of TOVS radiance data for a typical 6-hour period (15-21 UTC, 1 September 1992).

The NWP assimilation problems in relation to TOVS data may be considered of two types, which we shall call here "practical" and "fundamental". Practical problems are those whereby the data processing at either the satellite data centre (in this case NOAA/NESDIS) or at the NWP centre falls short of what might be achieved. Such problems are often related to various aspects of quality control, either in the pre-processing of the data or the temperature/humidity retrieval methods at NESDIS, or in the application of these retrievals by users. Such problems have received considerable attention over the years, with the result that the quality of present-day products and the methods for their application have improved considerably. These problems are important and continuing attention to them is required, but it is not intended to discuss them further here.

The fundamental problems are related to the peculiar nature of the satellite observations themselves and of the information they contain. They are caused by the intrinsically poor vertical resolution of the TOVS system, and this manifests itself in the so-called "background-dependence" problem. This is discussed in detail elsewhere (see *Eyre*, 1987; *Eyre*, 1989a), and only the main aspects are repeated here.

Let us consider the characteristics of the error in a vertical profile of temperature and humidity retrieved from TOVS data. x represents the profile, and superscripts t, b and r denote the true state, the prior ("background") estimate and the retrieved estimate respectively. y represents a vector of measured radiances. $H\{x\}$ represents the radiances corresponding to state x, i.e. calculated using a radiative transfer (RT) model. Let us assume that the errors in the measurements and in the RT calculations are unbiased and that they have a combined random error, ε. Then the RT or "forward" problem can be expressed by the following linearized approximation (which is adequate for this error analysis):

$$y - H\{x^b\} = H'\{x^b\}.(x^t - x^b) + \varepsilon \,, \tag{3.1}$$

where $H'\{x\} = \nabla_x H\{x\}$, the gradient of $H\{x\}$ with respect to x, and may be thought of as containing the "generalized weighting functions". The linearized retrieval or "inverse" problem can be expressed in general as follows:

$$x^r - x^b = W.(y - H\{x^b\}) \,, \tag{3.2}$$

where W is the inversion operator.

Combining (3.1) and (3.2), we obtain:

$$x^r - x^t = (I - W.H').(x^b - x^t) + W.\varepsilon \,. \tag{3.3}$$

The left-hand side is the retrieval error and the terms of the right-hand side are contributions to it, the first originating from the background error, $(x^b - x^t)$, and the second from the measurement error, ε. The second term is easy to understand; the measurement errors are mapped into the retrieval error through the inversion operator. However it is the presence of the first term which makes satellite sounding retrievals different in character from most other meteorological observations and also more difficult to assimilate into an NWP system. It is why they are not pure observations but "pseudo-observations", containing both observed and unobserved information. The background for the retrieval, x^b, will tend to have errors significantly correlated in the horizontal and vertical, and part of this pattern will therefore be transferred to the retrieval error. Note that this well-known property of satellite retrievals originates primarily in the retrieval process and not in the radiance measurements themselves.

If $(I - W.H')$ were small, then the problem would not be significant. However for TOVS, $(I - W.H')$ is not small, primarily as a result of the broad weighting functions and hence the poor vertical resolution of the sounding system. Attempts to force it to be small, through suitable choice of W, only result in unacceptable amplification of the measurement error (see *Rodgers*, 1976).

It has been demonstrated that this theoretical problem is also significant in practice. Examples have been found where the NESDIS operational retrievals differ significantly from the ECMWF short-range forecast, but where there is no significant difference between the measured radiances and those computed from the forecast. In such cases, the discrepancy is an artefact of the retrieval process not a characteristic of the radiance data.

Conventional retrievals, where the background for the retrieval is provided by an independent system such as in the NESDIS operational scheme, are difficult to use properly in NWP because the observed and unobserved parts of the retrieval cannot be decoupled. It is therefore more attractive to take the radiances themselves and to use the NWP system itself to provide the background. However, this does not fully solve the problem; if the NWP model is used to generate a "forecast-first-guess" retrieval, which is then passed to the data assimilation system, the background-dependence can now be understood and controlled more easily, but it has not disappeared. Moreover, we have now introduced the so-called "incest" problem, whereby part of the forecast error is fed back to the system as though it were new observations. Our attempts to address and control this problem at ECMWF are described elsewhere (Eyre et al. 1993). It is only through full three- or four-dimensional assimilation of radiance information that this problem can be overcome (see *Andersson et al.*, 1993).

4. VARIATIONAL ASSIMILATION OF TOVS RADIANCES

The basic principles of variational data assimilation have been introduced in preceding papers in this volume (*Courtier*, 1994; *Pailleux*, 1994). We present here only how this approach applies to the assimilation of satellite sounding radiances:

We minimize a cost function of the form:

$$J(x) = \tfrac{1}{2}(y - H\{x\})^T.(O+F)^{-1}.(y - H\{x\}) + \tfrac{1}{2}(x - x^b)^T.B^{-1}.(x - x^b) + J_o + J_c\ , \qquad (4.1)$$

where x^b is now a background state provided by a NWP model and B is the covariance of background error. $H\{x\}$ is the "observation operator" linking the NWP model variables, x, to the TOVS radiance observations. O and F are the covariances of error in the measurements and in the observation operator respectively. The first term is the cost representing the fit to the radiance data and the second to the background. The third term represents terms similar to the first measuring the fit to other observations, if any, and the fourth term represents other physical or dynamical constraints, if any.

Minimization of $J(x)$ requires the calculation of its gradient, $\nabla_x J(x)$:

$$\nabla_x J(x) \ = -H'\{x\}^T.(O+F)^{-1}.(y - H\{x\}) + B^{-1}.(x - x^b) + \nabla_x J_o + \nabla_x J_c. \qquad (4.2)$$

These equations are general with respect to the dimension of the field represented by x. x can be a vertical profile at one horizontal location. At ECMWF, we have developed such a system for TOVS data which we call one dimensional variational analysis (1DVAR), and it is described in a following paper (*McNally et al.*, 1994). However x can also be the three-dimensional field describing the full atmospheric state, and this is how TOVS data are used (along with other observations) in ECMWF's three-dimensional variational analysis (3DVAR) system (see *Andersson et al.*, 1994).

For 1DVAR, the Jacobian matrix H' has ~70x20 elements, and so it can be calculated explicitly and stored. This allows the Hessian matrix (the second derivative of $J(x)$) to be calculated and eq.(4.1) to be minimized with a Newtonian scheme. For 3DVAR, H' is too large to be computed and stored. Eq.(4.2) is evaluated instead by operating on $(O+F)^{-1}.(y-H\{x\})$ with the adjoint of the observation operator. Hence, the adjoint of the RT model is required, and this has also been developed at ECMWF.

An alternative way of visualizing the assimilation of satellite radiance information through such methods is shown in Fig.4.

For TOVS radiances, the major part of the observation operator is the RT model, which computes the brightness temperature expected in each TOVS channel given an adequate description of the atmospheric profile and surface conditions. The RT model used at ECMWF is described by *Eyre* (1991). To simulate the TOVS clear-column brightness temperatures, as supplied by NESDIS, it requires as input the temperature and humidity profiles, the surface pressure and skin temperatures, and the ozone amount.

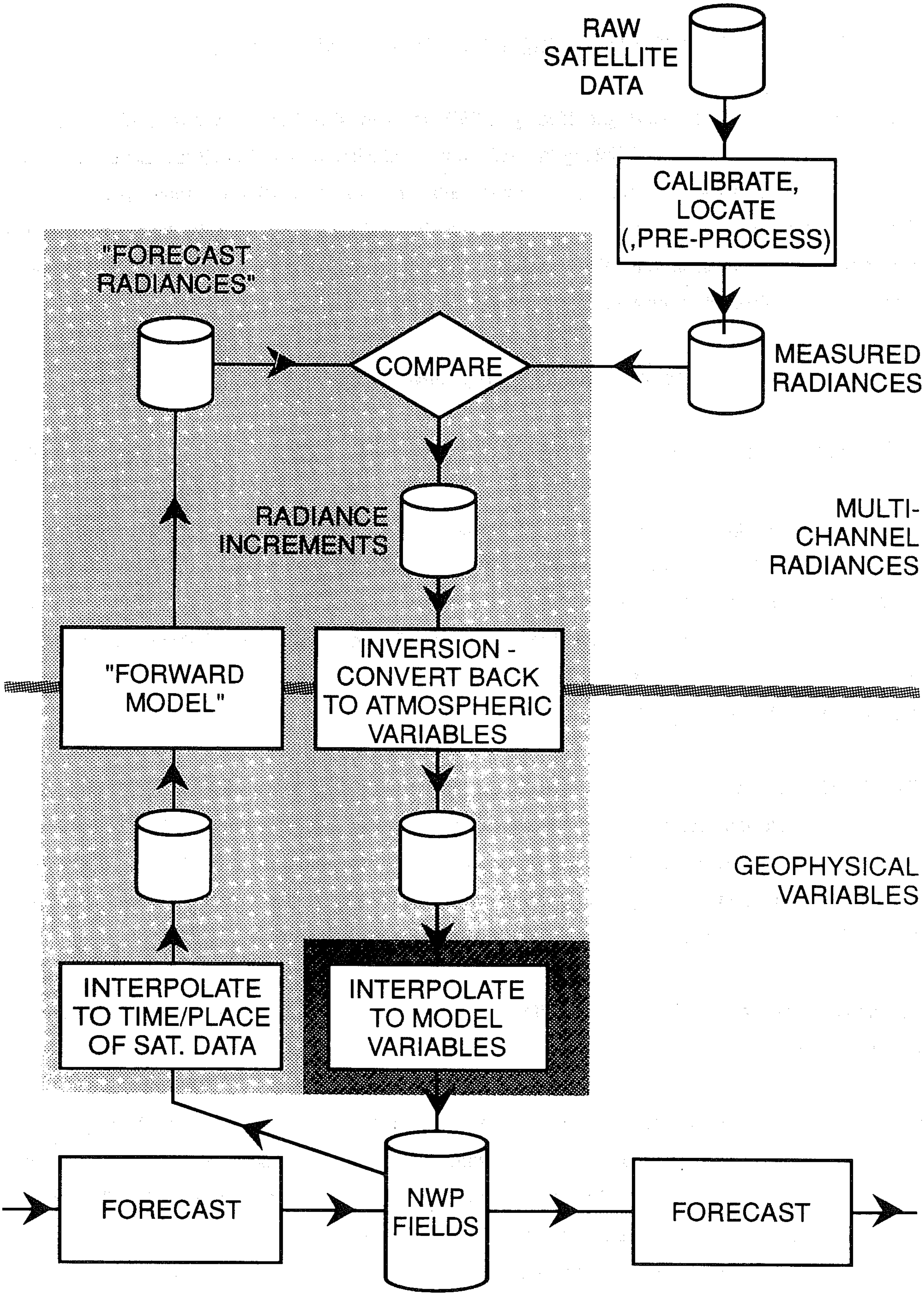

Fig. 4 Schematic illustration of the assimilation of satellite sounding radiances into a NWP system.

Compared with most other observations, the TOVS operator is not only more complex but also more nonlinear. The ability of the variational approach to handle such nonlinearities was one of the motivations for its development at ECMWF. The TOVS RT model is summarized in Fig.5. It can be seen that several of the steps are strongly nonlinear. For infra-red channels, the conversion of temperature profile to Planck function profile is nonlinear, but this is largely compensated by the inverse operation when radiance is converted to brightness temperature. Because of this compensation, the relationship between brightness temperature and temperature profile is only weakly nonlinear. The main nonlinearity is between brightness temperature and humidity profile, arising mainly through the exponential relation between optical depth and transmittance.

5. RADIANCE MONITORING AND TUNING

NESDIS clear-column brightness temperatures are received operationally at ECMWF. For each measured radiance, an equivalent "forecast" brightness temperature is calculated. For each 6-hour cycle of the NWP system, short-range forecast fields (+3-hour, +6-hour and +9-hour) are interpolated in time and space to the location of each TOVS sounding, and a brightness temperature is computed for each channel using the RT model. As can be seen from Fig.3, this yields many thousands of measured-forecast brightness temperature pairs in each channel every day. These data are archived and used to calculate statistics of measured-minus-forecast departures. The global mean and standard deviation in each channel are monitored every day, as an early warning for problems with the data. These departures are also used, typically each month, to monitor and correct for biases between measured and forecast brightness temperatures when they are applied in the 1DVAR or 3DVAR schemes.

Bias correction is an essential part of the system, as the bias is often comparable or larger than the error in the short-range forecast temperature field which we are attempting to correct during the data assimilation process. This is illustrated in the following figures. The calculated brightness temperatures in most channels suffer from a bias (with respect to the measured values) with a significant variation in bias from equator to pole. The bias results mainly from deficiencies in the RT scheme (see below). At present the problem is most acute for MSU channel 3, for which a map of monthly mean bias is shown in Fig.6. The measured-minus-calculated bias varies from about 1 K in the tropics to about -1 K in the Antarctic. The corresponding map of local standard deviation of the same data is shown in Fig.7. Over most of the world, it is in the range 0.3-0.5 K, i.e. smaller than the variation in bias. The variance of the difference can be interpreted as the sum of the variances in all the sources of error, i.e. the error in the measurements (after pre-processing), the error in the RT model and the error in the forecast. Therefore we can see that, for this channel at least, the bias in the observation departure exceeds the error in the forecast temperature (when mapped into brightness temperature space). The problem is most acute for MSU channel 3, but it is generally present for all the critical channels sounding the temperature of the troposphere and lower stratosphere. Fig.8 shows the standard deviation plot for HIRS channel 15, a temperature sounding channel

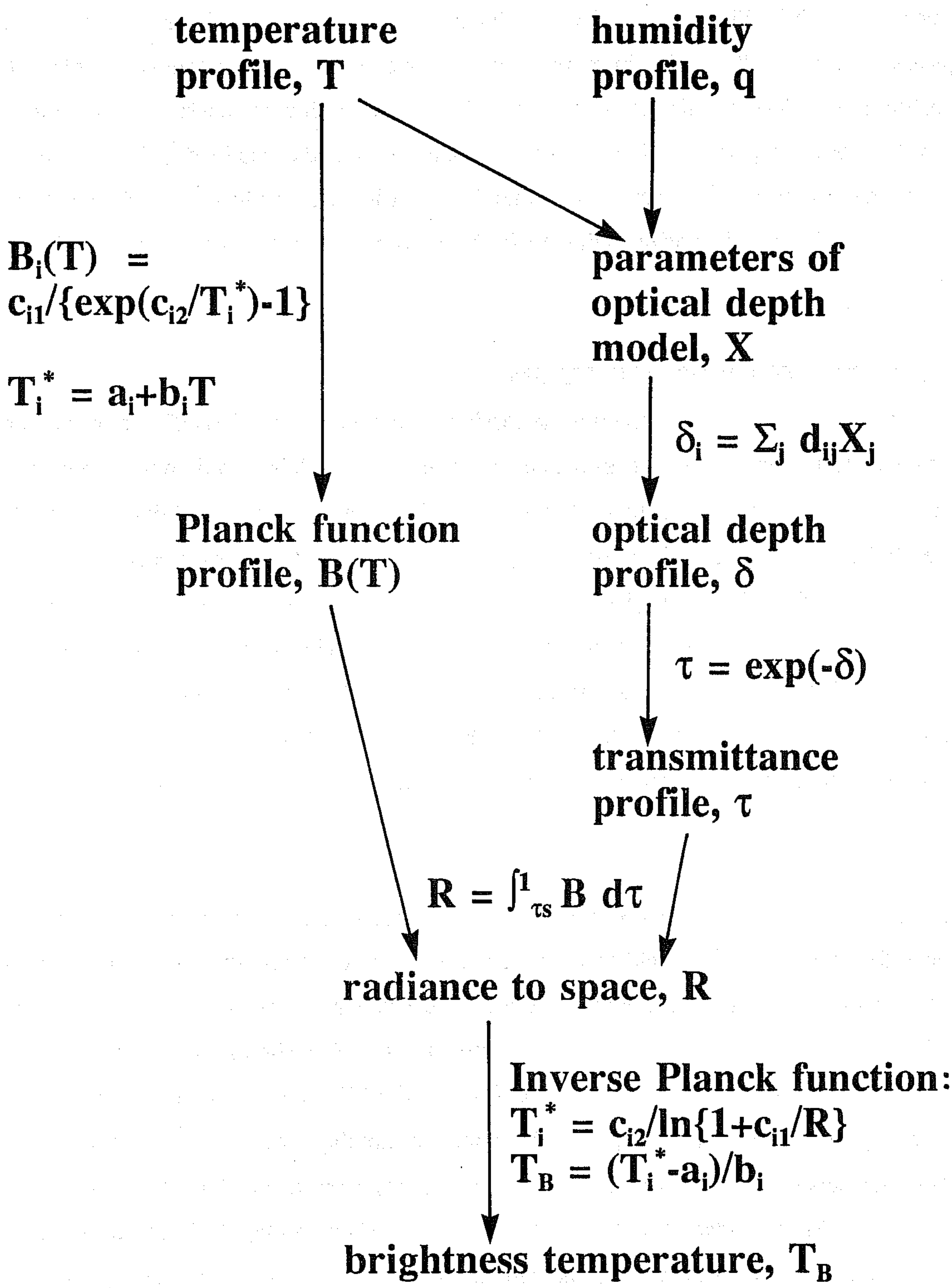

Fig. 5 Schematic summary of the TOVS radiative transfer model illustrating the source of nonlinearity.

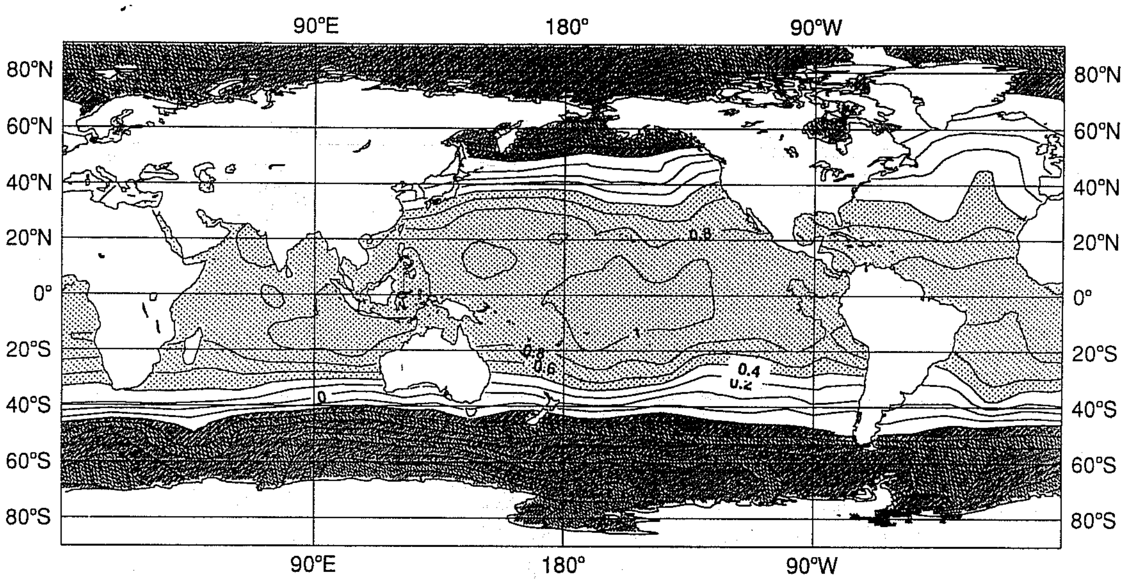

Fig. 6 Mean difference between measured and forecast brightness temperatures before bias correction. NOAA-11, MSU channel 3, October 1992.
Contour interval = 0.2 K; light shading > 0.4 K; dark shading < -0.4 K.

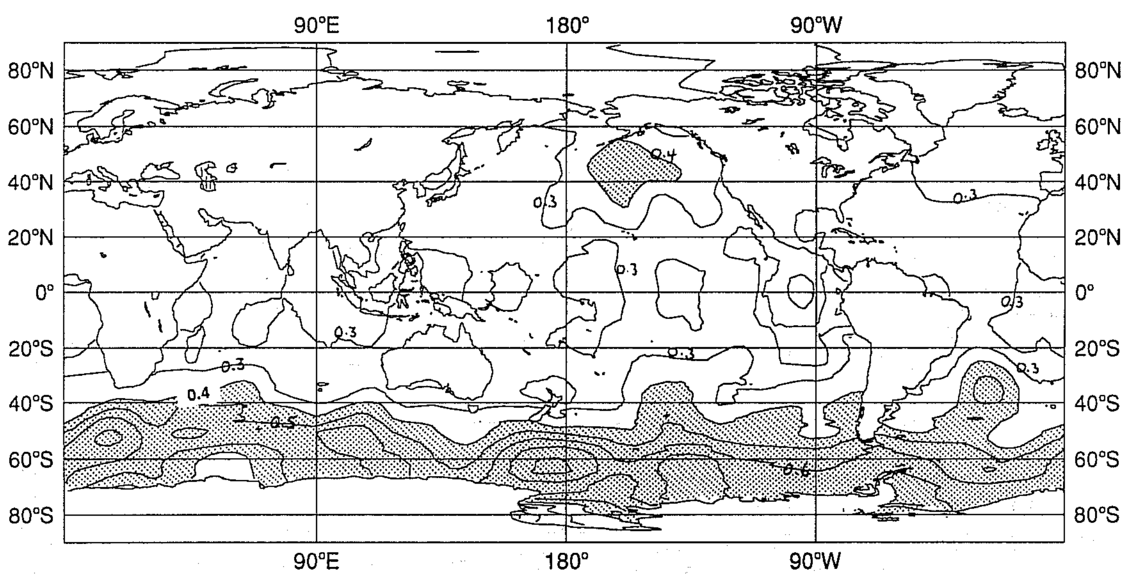

Fig. 7 Local standard deviation of the difference between measured and forecast brightness temperatures. NOAA-11, MSU channel 3, October 1992.
Contour interval = 0.1 K; shading > 0.2 K.

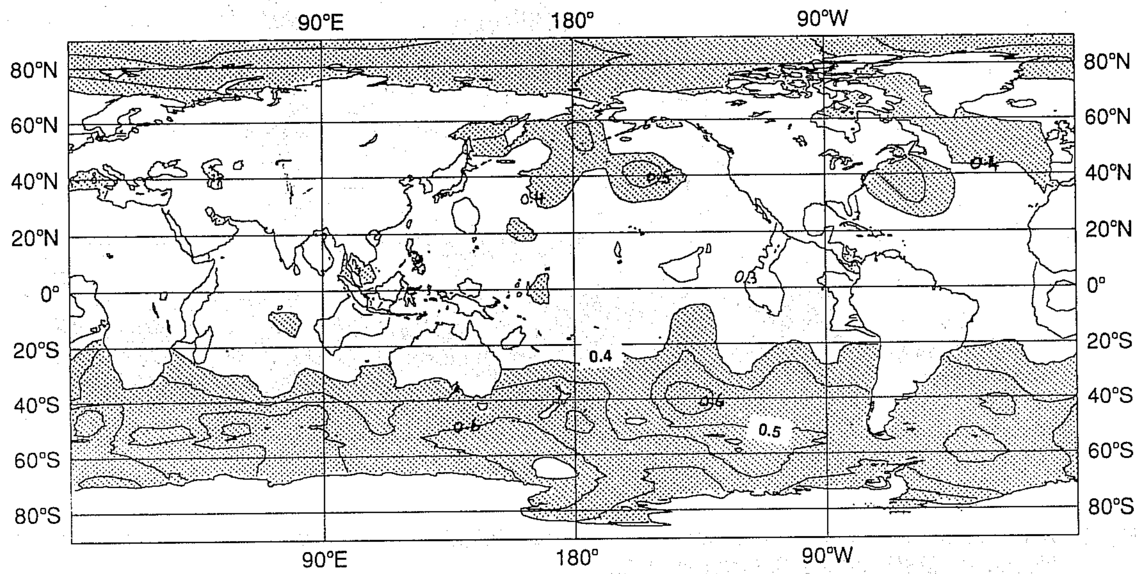

Fig. 8 Local standard deviation of the difference between measured and forecast brightness temperatures. NOAA-11, HIRS channel 15, October 1992.
Contour interval = 0.1 K; shading > 0.2 K.

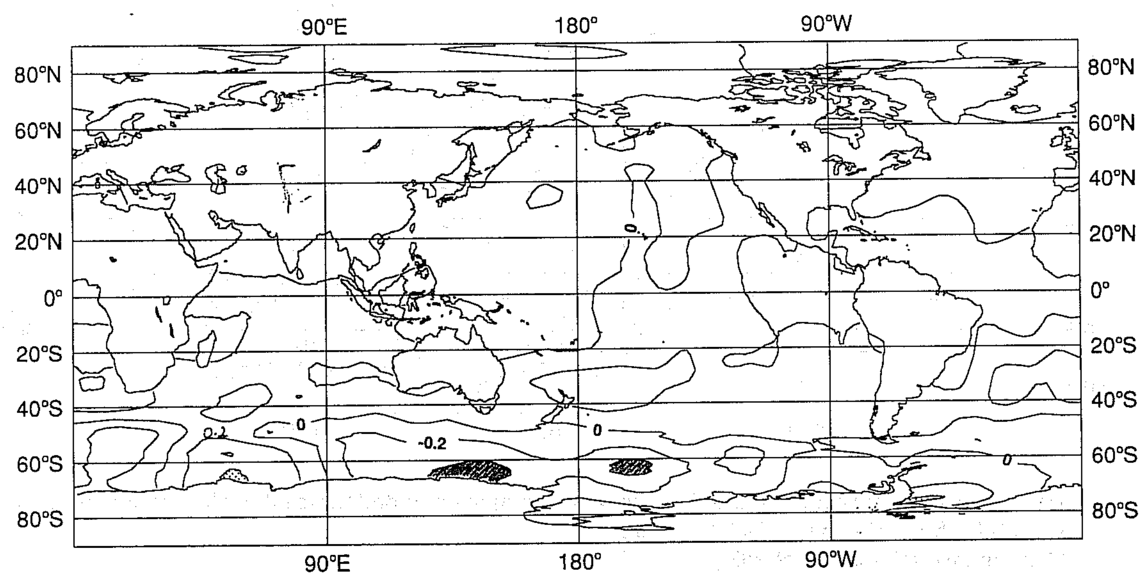

Fig. 9 As Fig.6, but after bias correction.

peaking around 700 hPa. The bias correction problem is therefore important, and the data must be corrected if they are to be used successfully.

In passing, we draw attention not only to the very low values of the standard deviations in Figs. 7 and 8 but also to their relative low spatial variation. This confirms the remarks in Section 3 concerning the accuracy of short-range forecasts, both for data-rich and data-sparse areas.

Biases may have various sources. Most arise from problems in the RT modelling (and often from the spectroscopic data underlying them). However others may be present in the measurements, arising from calibration errors or from systematic errors in the pre-processing of the data. It is not essential that we separate these errors, as we are primarily concerned to correct the bias in the **difference** between measured and calculated brightness temperatures. Biases can also be present in the NWP fields and lead to apparent biases in the calculated brightness temperatures. We would prefer not to tune these out before assimilation; it is desirable to use the data to correct these biases in the NWP fields through the assimilation process.

The TOVS bias correction scheme used at ECMWF is described in detail by *Eyre* (1992). Briefly, a spatially-varying correction is applied to each channel using a regression relation in which measured brightness temperatures in selected channels act as predictors. The correction to channel i is:

$$\Delta T_i = a_{0i} + \Sigma_i \, a_{ji} \, T_j \,, \qquad (5.1)$$

where T_j is the brightness temperature in channel j and the summation is over selected channels (currently MSU channels 2, 3 and 4). Coefficients a_{ji} are obtained by linear regression on the data base of measured-minus-forecast brightness temperatures described above. The data are subjected to strict quality control tests, and only those data close to an active radiosonde station are used in order to reduce unwanted adjustments for biases originating in the NWP system (which are likely to be largest in data-sparse areas).

This simple scheme has been found to work satisfactorily; for MSU channel 3, the field of residual biases after correction is shown in Fig.9. For all the channels sensitive mainly to tropospheric and mid-stratospheric temperature, the residuals after bias correction are similarly low. However this is not true for the water vapour channels, HIRS channel 10-12, nor for those temperature sounding channels significantly sensitive to water vapour (HIRS channels 5-7). For these channels, residual patterns of large amplitude are found, and these can be attributed to biases in the NWP model's humidity field. Fig.10 illustrates this for HIRS channel 11; north of 20°N the radiance data are assimilated through the 1DVAR system, and so the residuals are small, but south of 20°N where TOVS humidity information were not (at the time of these data) assimilated, biases are large. The amplitudes of -3 K to +3 K represent biases in the monthly mean specific humidity of up to 30% in the mid/upper troposphere. (For the data shown, the bias correction method has been applied but has had little effect on these patterns.)

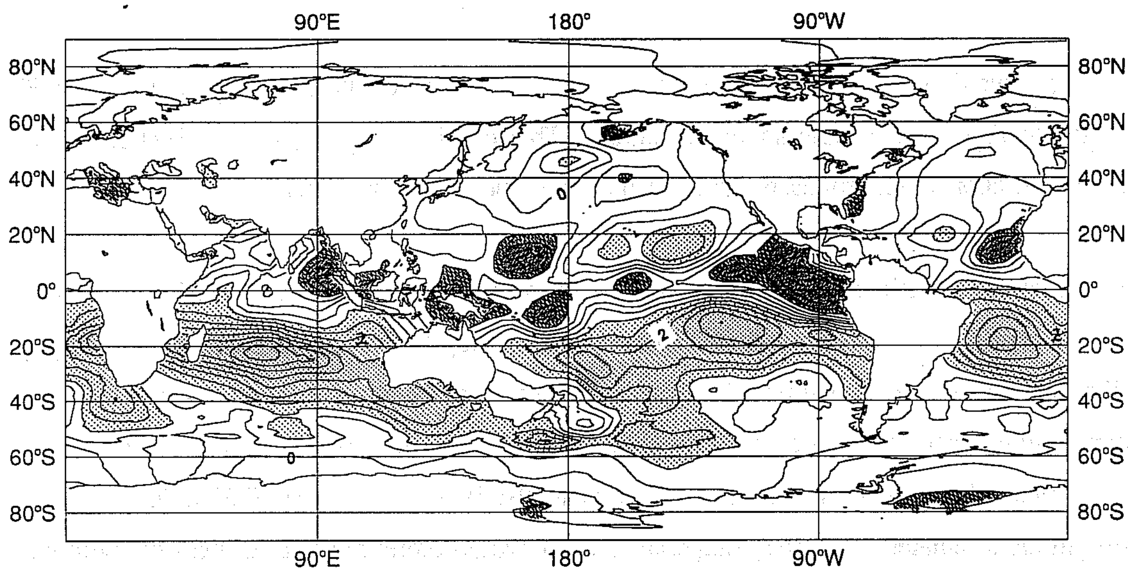

Fig. 10 Mean difference between measured and forecast brightness temperatures (after bias correction). NOAA-11, HIRS channel 11, July 1992.
Contour interval = 0.5 K; light shading > 1.0 K; dark shading < -1.0 K.

REFERENCES

Andersson, E, J Pailleux, J-N Thépaut, J R Eyre, A P McNally, G A Kelly and P Courtier, 1994. Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. Proc ECMWF/EUMETSAT Seminar on "Developments in the use of satellite data in numerical weather prediction"; Reading, UK; 6-10 Sept 1993; ECMWF Report.

Courtier, P, 1994. Introduction to numerical weather prediction data assimilation methods. Proc ECMWF/EUMETSAT Seminar on "Developments in the use of satellite data in numerical weather prediction"; Reading, UK; 6-10 Sept 1993; ECMWF Report.

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, 1989a. Research on satellite soundings at the UK Meteorological Office. Proc ECMWF Seminar on "Data assimilation and the use of satellite data"; Reading, UK; 5-9 Sept 1988; ECMWF Report, 169-179.

Eyre, J R, 1989b. 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.

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

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

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

McNally, A P, G A Kelly, J R Eyre, and E Andersson, 1994. Experiments using one-dimensional variational analysis of TOVS data at ECMWF. Proc ECMWF/EUMETSAT Seminar on "Developments in the use of satellite data in numerical weather prediction"; Reading, UK; 6-10 Sept 1993; ECMWF Report.

Pailleux, J, 1994. Assimilation of satellite data using variational techniques. Proc ECMWF/EUMETSAT Seminar on "Developments in the use of satellite data in numerical weather prediction"; Reading, UK; 6-10 Sept 1993; ECMWF Report.

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

Schwalb, A, 1978. The TIROS-N/NOAA A-G satellite series. NOAA Tech Memo NESS 95.

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.