TUNING THE U.K. METEOROLOGICAL OFFICE'S TOVS PROCESSING SUITE

P.D. Watts
Meteorological Office Unit, Hooke Institute,
Oxford, U.K.

ABSTRACT

It is regularly found that operational sounding systems do not give the retrieval accuracies expected from a theoretical standpoint. Partly this is because there is no error–free validation of the system available, but principally it is because the assumptions made by the particular scheme about the state of the data are not correct. Various examples of these assumptions are given and the assumption of unbiased data examined in detail. The question of how best to handle biased data in a system processing Tiros–N Operational Vertical Sounder (TOVS) data where no error–free truth is available is discussed using examples from the U.K. Meteorlogical Office TOVS processing suite. This suite has lately (1988) been provided with a comprehensive monitoring and tuning system and its performance has reached an acceptable (though still not optimal) level.

1. INTRODUCTION

In 1987 the Meteorological Office started producing retrievals of atmospheres profiles from TIROS Operational Vertical Sounder (TOVS) data using a "forecast background" inversion scheme. In this scheme (Eyre 1989) the "most likely" profile is obtained given the measurements and a background short-range forecast profile and knowledge of the error characteristics of each. The "measurements" of the scheme are limb-corrected and cloud-cleared High-resolution Infra-red Radiation Sounder (HIRS2) data and mapped and limb-corrected Microwave Sounding unit (MSU) data. Such preprocessing of the measurements and their inherent low vertical resolution mean that, even in theory, a retrieval error at mid tropospheric levels that is 0.5K lower than the background error is the best that we can expect to achieve. In practice various assumptions made in the retrieval scheme may be erroneous and although this could, in principle, mean retrievals that are better than expected, in most cases the result will be a degradation of accuracy. Experience with the Meteorological Office system is that the latter is true – it is very easy to lose the small expected improvement on the background.

In order to see where problems may lie we examine the usual minimum variance retrieval equation (see for example Rodgers 1976).

$$\hat{\mathbf{x}} - \mathbf{x^b} = \mathbf{C.K^T(K.C.K^T + E)^{-1}.(y^m - y^b)}$$

or

$$\hat{\mathbf{x}} - \mathbf{x^b} = \mathbf{W.(y^m - y^b)} \tag{1}$$

where $\hat{\mathbf{x}}$ is the solution profile, $\mathbf{x^b}$ is the background forecast profile with error covariance $\mathbf{C}$, $\mathbf{y^m}$ is the measurement vector and $\mathbf{y^b}$ the measurements expected from the background profile. We will refer to "measurements" calculated from a background profile as 'background brightness temperatures' and similarly to 'sonde brightness temperatures' for those calculated from a radiosonde. $\mathbf{E}$ is the error covariance of the measurements but also should include any error in the forward model $\mathbf{y(x)}$. $\mathbf{K}$ is the derivative of $\mathbf{y}$ with $\mathbf{x}$.

This equation represents the 'most likely' solution if

a) all errors have gaussian form with

b) covariances $\mathbf{C}$ in the background and $\mathbf{E}$ in the measurements/forward model,

c) the forward model is linear i.e. $\mathbf{K}$ is a constant, and

d) measurements, background and forward calculations are unbiased with respect to the truth.

In line with common practice and the available evidence we take a) to be satisfied. Condition b) can be very important and making incorrect assumptions about the error covariances, particularly the correlation structure, can seriously degrade the retrievals. Work on this aspect has been reported elsewhere (Watts and McNally, 1987) and is being carried further by current research at the Hooke Institute and Meteorological Office.

Condition c) is easy to test and it is found that with cloud-cleared data the forward model is quite linear with respect to the temperature profile. Strong non-linear dependencies on the humidity part of the profile are well known. It is with condition d), that the measurements and background be unbiased, that this paper is concerned.

In a completely linear processing system it is irrelevant where bias correction is carried out and therefore for practical reasons would be best done at the end. However, in a non-linear system, which almost all retrieval systems are to some extent at least, systematic errors in the measurements and background can cause random errors in the solution and consequently bias corrections must be done at the earliest possible stage. The current Meteorological Office scheme is non-linear in two aspects. The first, and largest effect, is cloud-clearing the HIRS data which is essentially a decision making process based on a comparison of HIRS and MSU data. Biases in the data may cause a different set of decisions. At the retrieval stage the non-linearity is introduced by use of different inversion matrices for different cloud-clearing routes (required because of their different error structures).

Another good reason for removing biases near to source is that they are then usually the result of fewest causes and therefore least complicated to monitor and correct successfully.

Although bias evaluation and correction is in principle a straightforward task it is complicated by lack of any strictly error–free truth. The traditional baseline of the radiosonde network introduces its own errors principally because the sonde does not measure all we need in order to calculate the radiances. Notably there is no surface skin temperature or information above ~10mb. These parts of the 'true profile' have to be invented with consequent introduction of both systematic and random differences from measurements. There are also errors and biases in the parts profile that the sonde *does* measure due to, for example, radiative effects at high levels, poor humidity sensors etc.

The rest of this paper describes a rationale for handling "real" and "introduced" biases and gives examples which, though useful, are necessarily somewhat system–dependent.

2. THE MONITORING SYSTEM

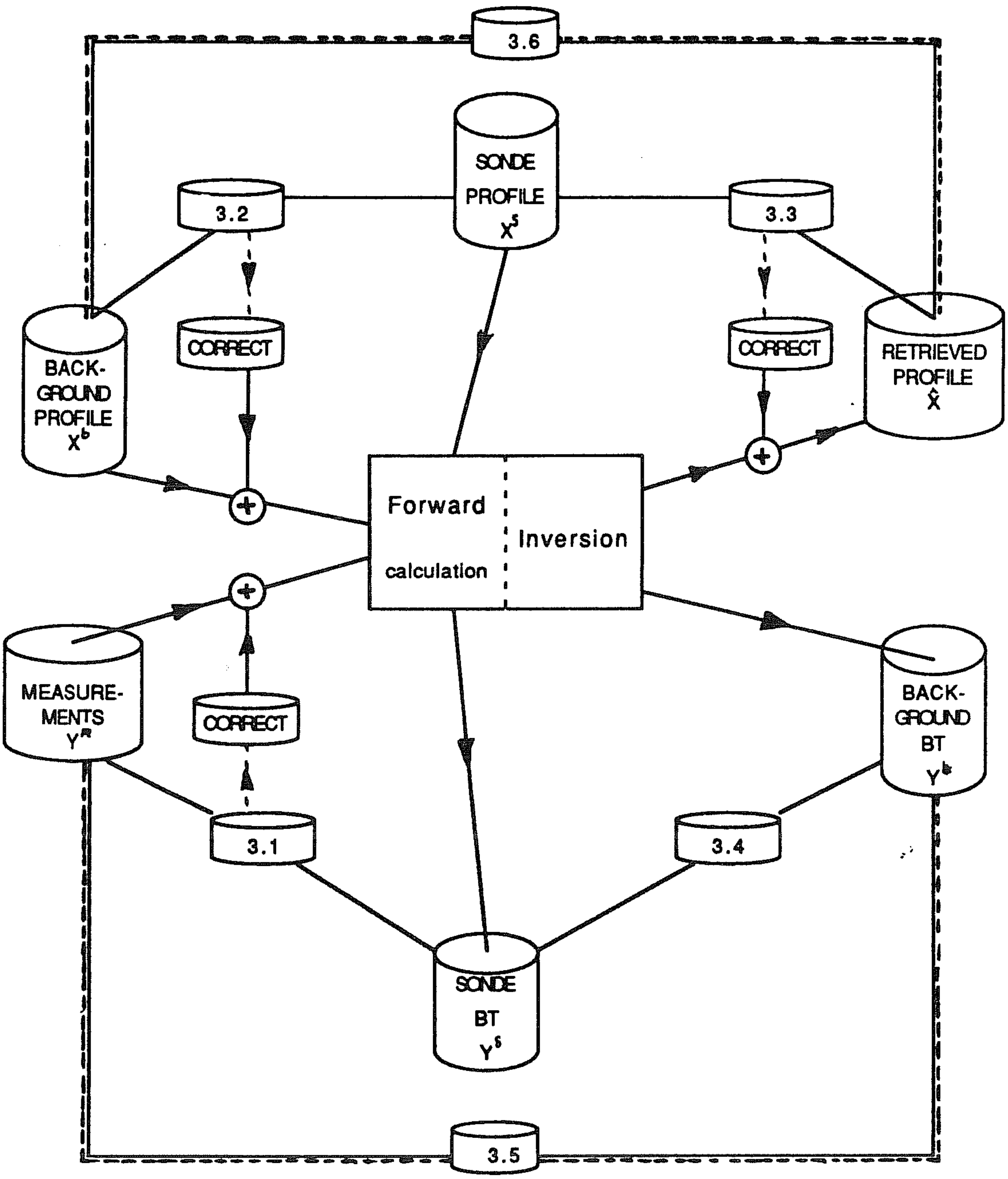
Monitoring is principally achieved through collocated radiosondes (ascents within 3 hours and 150 km of the sounding). It has also been found useful to monitor measurements and retrievals against forecast values. Figure 1 shows the Meteorological Office retrieval scheme and its associated monitoring system. The top half of the figure is essentially in profile space, the bottom in measurement space. The retrieval process flows from left to right, input brightness temperatures and background profile and output retrieval profile and background brightness temperatures. The collocated sondes and the sonde brightness temperatures are shown top and bottom respectively. The small boxes 3.1 $\rightarrow$ 3.6 on the figure indicate the six possible comparisons the monitoring system can make. The two comparisons which are apparently redundant, e.g. (measured – forecast) differences, in that they can be calculated from (measured – sonde) and (forecast – sonde), are useful because by avoiding the collocating sonde many more matches are available. The remaining boxes labelled 'correct' show the bias corrections which can be made in the system, namely the measurement bias, the background bias, and the retrieval bias. We do not correct for the retrieval bias at present believing that the other corrections have reduced it to an acceptable level. The facility, however, is available if required.

3. BIAS ESTIMATION

The three profiles – background, sonde and retrieval – may be written as $\mathbf{x^b}$, $\mathbf{x^s}$ and $\hat{\mathbf{x}}$ respectively and similarly for the measurements $\mathbf{y^b}$, $\mathbf{y^s}$ and $\mathbf{y^m}$ where the background and sonde brightness temperatures are calculations via the radiative transfer model $\mathbf{y}(\mathbf{x})$ from the relevant profile. A bias in a vector quantity will be denoted by prefacing it with a $\mathbf{b}$; thus $\mathbf{bx^s}$ is the bias error in the sonde. To facilitate discussion we also assume one can write $\mathbf{K.bx}$ to state how a profile bias is mapped into a measurement bias, and similarly $\mathbf{W.by}$ for a measurement bias mapped by the inverse operator into a profile bias. These are approximations,however, because they may not be linear operations.

It is convenient to 'invent' true profiles and measurements $\mathbf{x^t}$ and $\mathbf{y^t}$ neither of which are available but serve to define the various biases. Thus, for example, $\mathbf{bx^b} = \mathbf{x^b} - \mathbf{x^t}$

Figure 1. Meteorological Office TOVS processing and monitoring system



BT = Brightness temperature

gives the background bias. It is these biases, as compared with truth, which are required. Unfortunately we do not have access to $\mathbf{x^t}$ or $\mathbf{y^t}$ and we can only perform comparisons such as $\mathbf{x^b}-\mathbf{x^s}$ and hope to extract the relevant information. It is helpful, though, to examine the comparisons with truth to establish the contributing factors.

1. $\mathbf{bx^s} = \mathbf{x^s}-\mathbf{x^t}$ is any bias error in the sonde profile. The principal source of error here will be at high levels (<15mb) were the sonde record is incomplete. In the Meteorological Office system values for these levels are obtained by extrapolation on lower levels using a regression relation. This extrapolation is subject to very obvious seasonally varying bias errors. Radiosonde humidities are also subject to large errors at high levels, and large biases in HIRS channel–12 (water vapour channel) are supporting evidence of this.

2. $\mathbf{bx^b} = \mathbf{x^b} - \mathbf{x^t}$. The background profile is extrapolated to high levels in a similar way to the sonde and will likewise have similar errors. There are also likely to be biases due to deficiencies in the forecast model.

3. $\mathbf{by^m} = \mathbf{y^m} - \mathbf{y^t}$. There are several sources of bias in the measurements arising from the calibration, limb-correction and cloud-clearing processes. The last appears to be the biggest effect as HIRS biases are steadily negative and only large for channels peaking low in the atmosphere.

4. $\mathbf{by^s} = \mathbf{y^s} - \mathbf{y^t} = \mathbf{by^\tau} + \mathbf{K.bx^s}$. Biases in the sonde brightness temperatures arise from errors in the forward calculation, $\mathbf{by^\tau}$, and errors in the sonde profile mapped into measurement space, $\mathbf{K.bx^s}$. Errors in the forward calculation may arise from transmittance errors or errors of quadrature in the radiative transfer integration of a discrete profile (e.g. consistently smoothing the tropopause would give a warm bias).

5. $\mathbf{by^b} = \mathbf{y^b} - \mathbf{y^t} = \mathbf{by^\tau} + \mathbf{K.bx^b}$. Biases in background brightness temperatures will be from forward model bias and from the background profile bias mapped into measurement space.

6. $\mathbf{b\hat{x}} = \mathbf{\hat{x}}-\mathbf{x^t} = \mathbf{W.K.bx^b} + \mathbf{W.(by^m}-\mathbf{by^\tau)}$. The bias in the retrieved profile is a result of biases in the background profile mapped first into measurement space and then back to profile space, and also as a result of residual forward model and

measurement biases mapped into profile space.

That completes a description of comparisons with true profiles which we cannot in fact make. The six comparisons between sonde, background and measured values we can make and which are combinations of the above, are used to infer the desired biases where possible. To reiterate, the biases required are the background profile bias $\mathbf{bx^b}$, the combined measurement/forward model bias $(\mathbf{by^m} - \mathbf{by^\tau})$ and the retrieval bias $\mathbf{b\hat{x}}$. Sections 3.1–3.6 describe the available comparisons starting with the three which enable us to estimate the required biases.

3.1 Comparison of measurements and sonde brightness temperatures

$$\mathbf{by^m} - \mathbf{by^s} = \mathbf{by^m} - \mathbf{by^\tau} - \mathbf{K.bx^s}$$

Notice that this is the required measurement bias with a complicating contribution from radiosonde errors. If it is known that $\mathbf{K.bx^s}$ is significant then it should be estimated and added to the observed bias. An example is given later (section 4(a)) where there is a large sonde bias at high levels. If this effect is not allowed for then the measurement bias correction will introduce it into the retrieval system where it will cause a bias in the retrieved profile. An unfortunate consequence is that this may actually make the apparent retrieval bias smaller whereas in reality the retrieval is simply closer to an erroneous sonde.

It is not necessary to try to separate the two effects $\mathbf{by^m}$ and $\mathbf{by^\tau}$ since correcting $\mathbf{y^m}$ for the former and $\mathbf{y^b}$ for the latter is equivalent to correcting either of them by $(\mathbf{by^m} - \mathbf{by^\tau})$.

3.2 Comparison of background and sonde profiles

$$\mathbf{bx^b} - \mathbf{bx^s}$$

is the quantity obtained and $\mathbf{bx^b}$ is the quantity required. Again the sonde error complicates interpretation and since the background is likely to have similar errors to the sonde (especially in upper level extrapolation) the terms tend to cancel.

There may be two distinct cases where correction is required. Firstly, if there is a significant residual $\mathbf{bx^b} - \mathbf{bx^s}$ and no reason to suspect a sonde error we may assume it is a background error. Secondly, if there is a known sonde error but the residual is near zero then we can assume the background has the same error.

3.3 <u>Comparison of retrieved and sonde profiles</u>

$$\mathbf{b\hat{x}} - \mathbf{bx^s} \;=\; \mathbf{W.K.bx^b} \;+\; \mathbf{W.(by^m - by^r)} \;-\; \mathbf{bx^s}$$

is the bias difference observed between sonde and retrieved profile. As usual the sonde error should be estimated and removed to obtain the required bias estimate. If what remains is unacceptably large it can be corrected but it is really indicating in this case that the other bias corrections are unsatisfactory. It is also a resultant bias so that it will change if the inverse operator $\mathbf{W}$ changes or is profile dependent (i.e. a non–linear retrieval). With the previous Meteorological Office retrieval scheme, which used a regression based on climatological data, $\mathbf{bx^b}$ was quite large and a retrieval bias correction was necessary. With the new forecast–based background, the problem should be minimal.

3.4 <u>Comparison of background and sonde brightness temperatures</u>

$$\mathbf{by^b} - \mathbf{by^s} \;=\; \mathbf{K.bx^b} \;-\; \mathbf{K.bx^s}$$

It is not obvious how this comparison could be useful as it reflects comparison 3.2 but in measurement space.

3.5 <u>Comparison of measurements and background brightness temperatures</u>

$$\mathbf{by^b} - \mathbf{by^m} \;=\; \mathbf{by^r} \;+\; \mathbf{K.bx^b} - \mathbf{by^m}$$

This comparison is very similar to 3.1 and could be used to obtain the measurement bias but is complicated by the presence of a background bias term. However, many "perfect" collocations are available and good statistics are obtained. If there are reasons for suspecting one of the sources will vary according to some parameter, then these data provide a good base for investigation. For example a strong limb-dependent bias was found in NOAA–10 HIRS data by examining this difference across the scan. Given that $\mathbf{bx^b}$ can be expected not to vary with scan position we can conclude the variation must be in $\mathbf{by^\tau}$ or $\mathbf{by^m}$. Many other sources of variable biases could be investigated in this way but care must be taken over interpretation of results.

3.6 Comparison of background and retrieved profiles

$$\mathbf{bx^b} - \mathbf{b\hat{x}} = \mathbf{bx^b} - \mathbf{W.K.bx^b} - \mathbf{W.(by^\tau} - \mathbf{by^m)}$$

This is a very difficult comparison to interpret and has the complication that the first two terms are not independent.

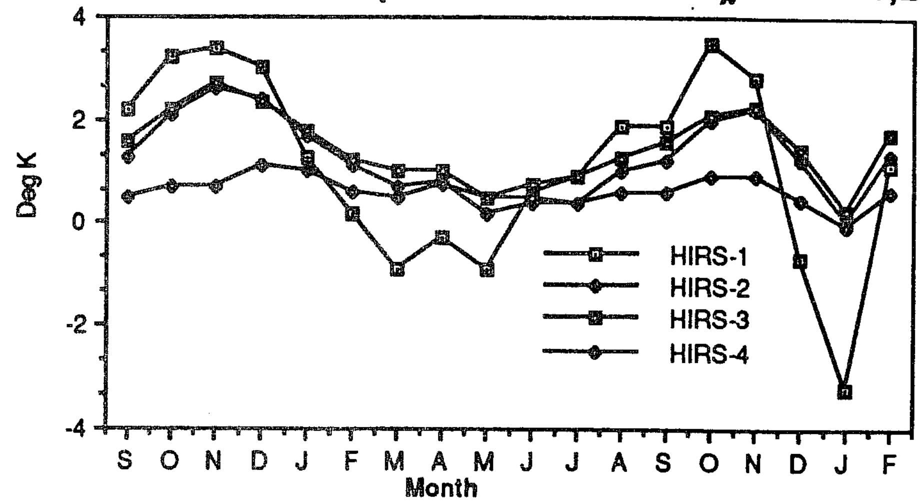
4. EXAMPLES

Some examples are given to illustrate points made above. Figure 2 is a sequence of monthly means over a year or so of the measurement – sonde brightness temperature difference for NOAA–9 from the collocation system. Profiles of the six comparisons described above are given in figure 3 for three months of NOAA-10 data. Consequently the figures cannot be compared. Working through the channels in figure 2:

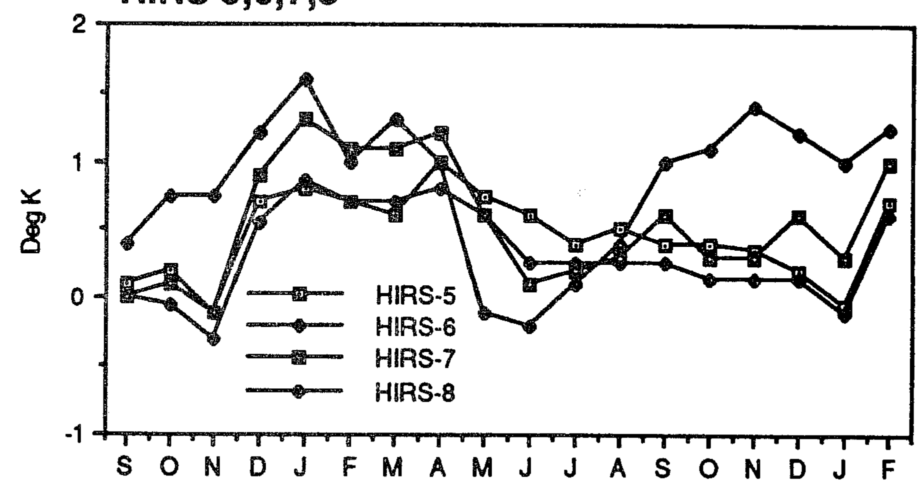
a) **The stratospheric channels** show a varying bias with near zero values by May or June and maximum values in November. The cyclic pattern is evident in HIRS–17,1,2 and 3 with decreasing amplitude. It can often be seen in HIRS–16 and MSU–4 but is not evident here. (NOAA–9 has no working HIRS–16.) We can be fairly sure this is the profile extrapolation error, $\mathbf{bx^s}$, and so when estimating the measurement bias we should ignore it. Supporting evidence that this is the cause can be seen from the channel biases of figure 3.1. The background profile is currently extrapolated from 50 mb, i.e. well below the starting point for sonde extrapolation. $\mathbf{bx^b}$ may therefore be expected to be larger than $\mathbf{bx^s}$. This can be seen in the sonde–background bias in channels 1,2,3 and 4 shown in figure 3.1. The question is what is the measurement/forward model bias underlying the seasonal variation. We may perhaps assume that the regression extrapolation will be unbiased over a year since it is based on global 12–monthly data. The residual is then the average

NOAA-9 SEP'85 - FEB'87

Fig 2. Monthly Brightness temperature biases. (Sonde - measured). HIRS 1,2,3,4



HIRS 5,6,7,8



HIRS 10,11,12

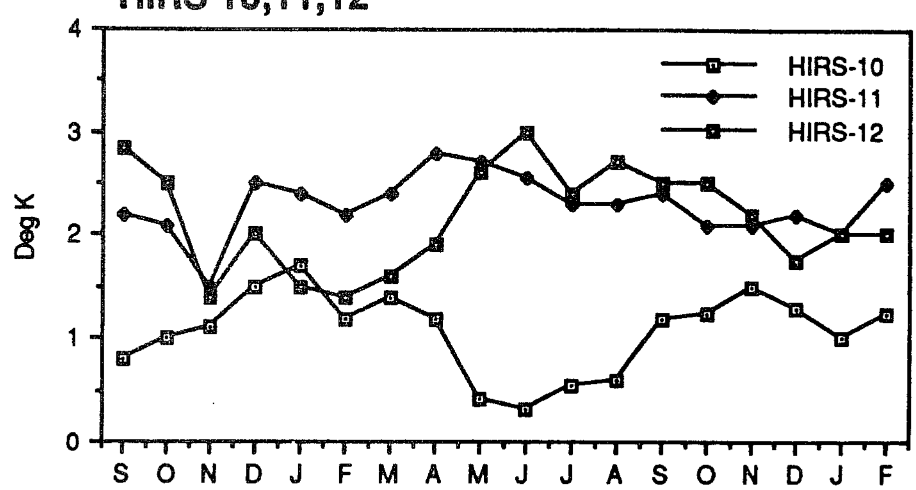
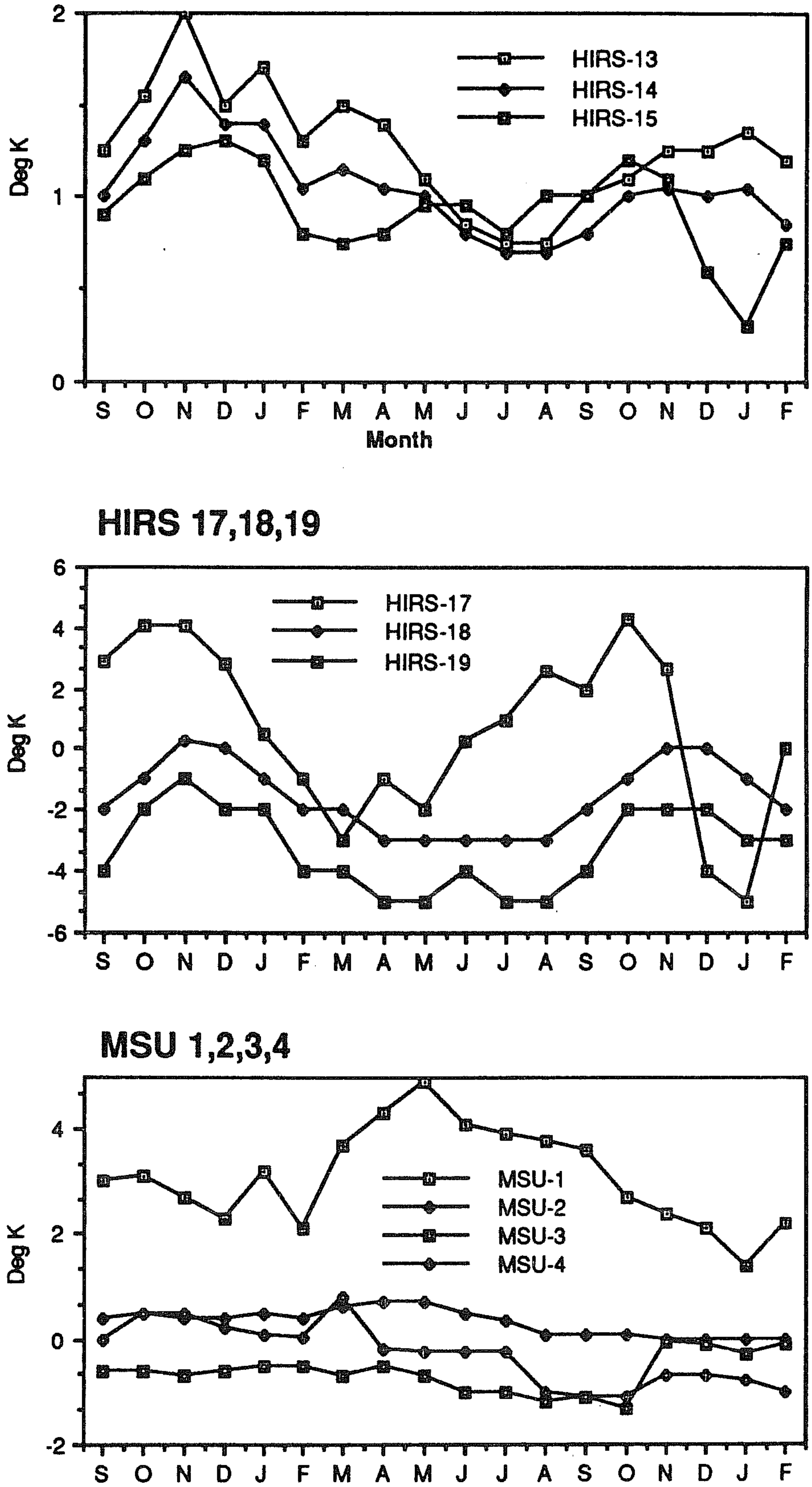


Fig 2 cont. Monthly Brightness temperature biases. (Sonde - measured). HIRS 13,14,15



HIRS 17,18,19

MSU 1,2,3,4

Fig 3. Profile comparisons SEPT-NOV 1987

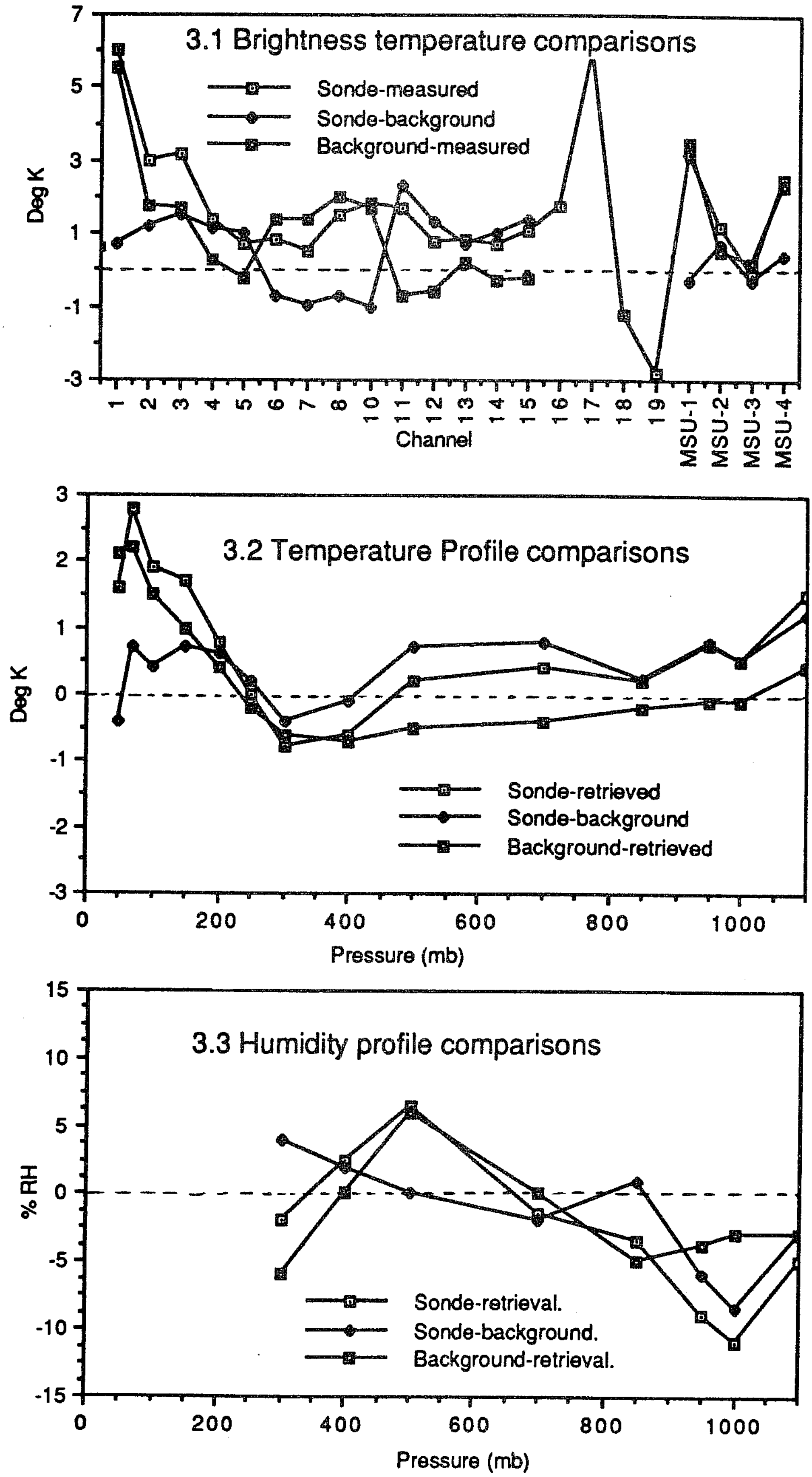
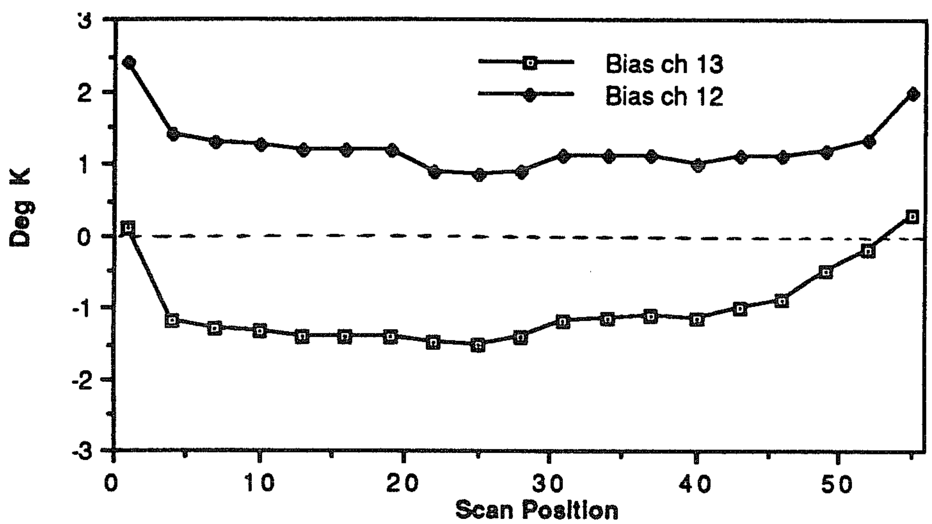
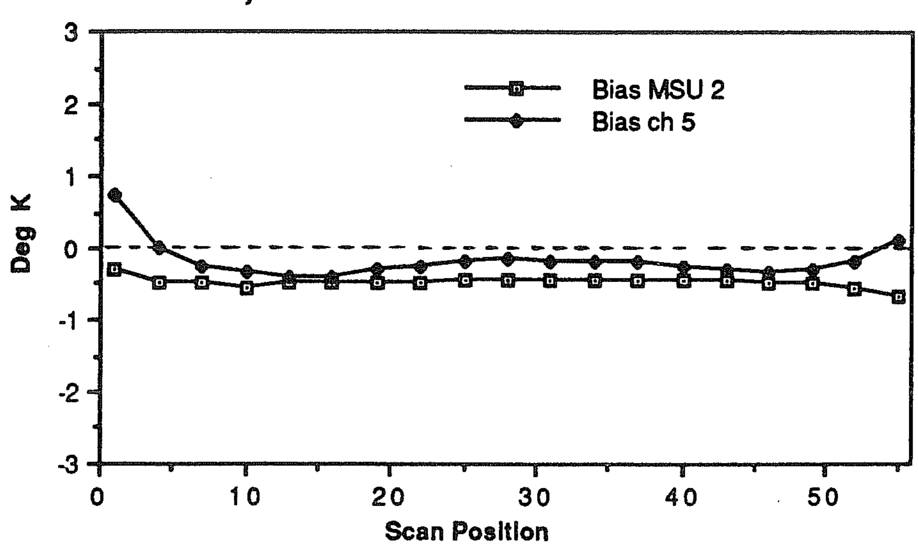


Fig. 4 Scan dependent Brightness temperature biases. (Measured-background)

HIRS-12, HIRS-13



MSU 2, HIRS 5



value over the year and represents the ($\mathbf{by}^r - \mathbf{by}^m$) term. The obvious problem that it requires 6–9 months at least to establish the cyclic pattern after a new satellite goes operational may be solved by assuming the pattern (though not the value of the ($\mathbf{by}^r - \mathbf{by}^m$) term) will be the same as for the previous satellite.

b) **The HIRS longwave tropospheric** sounding channels 5–8 also show a pattern which emerges as the channel peaks become lower in the atmosphere. Again the effect is seasonal though not as smooth as that in stratospheric channels. This is almost certainly due to residual cloud contamination and represents the changing cloud amounts through the year. The effect is largest in the winter months when much more cloud is generally present. This bias *is* a measurement bias. It may be desirable to 'follow' the bias on a monthly basis rather than simply to take the yearly or long period mean value, especially in the lowest channels. Another effect that may be of importance with the longwave channels is the seasonal variation of ozone amount. This is expected to have a maximum signal in HIRS–5 of ≈ 0.7 K peak to peak (see Eyre, 1986) and is a forward model error. There may also be a sonde error present but this is difficult to establish; certainly the fact that three channels approach zero in the summer suggest there is nothing serious.

c) **HIRS water vapour** channels 10–12 show a mixture of behaviours. HIRS–10, peaking very near the surface, is almost a copy of HIRS–8 showing that the cloud effect is largest. The values are marginally higher and this may be a sonde humidity error or a transmittance error. HIRS–11 and 12 show no cloud signal but neither do they have a pattern of their own (though in similar plots for NOAA-10 a distinct pattern is present, unlike the cloud pattern). The biases are high and it is not clear how to assign them. Evidence from figure 3.3, that background minus sonde humidity biases of $\approx 10\%$ at some levels indicate a problem with one or the other, suggests that it is probable the large measurement biases in channels 10–12 are primarily from the sonde. For more discussion of likely sonde humidity problems see Eyre (1983).

d) **Shortwave HIRS channels** 13–15 show the cloud pattern but not so strongly as the longwave channels. There is clearly a significant underlying bias in these channels which is unlikely to be from the sonde. Possible bias effects include transmittance errors, surface emissivity variation and reflected sunlight though the latter effect has the wrong sign to explain the effect.

e) **HIRS channels 17–19 and MSU–1** are not used in the inversion scheme as yet and are not discussed fully here. Suffice to say HIRS–17 has the stratospheric error pattern as seen in 1–3 but is noisier and obviously has another bias superimposed (probably caused by fluorescence). Channels–18 and 19 have the cloud pattern but with a strong negative bias which plausibly may be solar contamination. MSU–1 has large errors probably associated with the variable emissivity of the surface.

f) **MSU channels 2–4.** MSU–2 and 3 are well behaved and have low biases. Without cloud effects, little contribution from above 15 mb and little response to humidity what is left is almost certainly $\mathbf{by}^{\tau}$ or $\mathbf{by}^{\mathbf{m}}$. From figure 2 the same can be said for MSU–4 though we might have expected some effect from the sonde extrapolation in this channel.

Errors in the background profile $\mathbf{bx}^{\mathbf{b}}$ are best found from figures 3.2 and 3.3. Clearly the background is cold by $\approx$ 0.5–1.0 K in the troposphere where we may expect no significant sonde errors. The background humidity is also markedly different from the sonde but the background brightness temperatures and measurements (figure 3.1) appear to be more in agreement. This tends to confirm the problem with sonde humidities. As expected the background is not biased very differently from the sonde in the stratosphere though clearly both are biased with respect to the measurements. So we have no measure of the stratospheric $\mathbf{bx}^{\mathbf{b}}$ apart from inferences from the measurements. Currently under consideration is the role which the Stratospheric Sounding Unit (SSU) may be able to play in aiding the extrapolation of profiles to high levels.

Figure 4 is a plot of background brightness temperatures – measurements, $\mathbf{by}^{\mathbf{b}} - \mathbf{by}^{\mathbf{m}}$, according to scan position for one month of NOAA–10 data. Notice the strong dependence on position especially at the edges of the scan. The mean difference across the scan reflects mean background *and* mean measurement errors. The Meteorological Office scheme now routinely monitors this behaviour and a scan-dependent correction to the measurements is made in addition to the mean correction.

5. CONCLUSION

Of all the sources of error in a TOVS retrieval system, that of biases in the input data is conceptually the easiest to handle. Nevertheless great care must be taken when interpreting the biases obtained from a monitoring system using, for example, the radiosonde network. This is because the "truth" here will also be error-prone. The guiding principle is that a bias should only be corrected if it is not the result of the monitoring system. Similarly a bias in the data may be "hidden" by similar biases in the "truth". The question also arises as to whether one should aim to correct towards the unattainable truth at all, rather than correcting to the baseline set by the sonde network. We have assumed the former philosophy in this work. The Meteorological Office system now has a comprehensive monitoring and correction system for biases and its performance is improved but still significantly sub–optimal. Other parts of the system such as assumptions about error characteristics are now under scrutiny and may lead to further improvements.

REFERENCES

Eyre J.R., 1983: Comparisons between measured and simulated TOVS brightness temperatures. Met O 19 Branch memorandum No. 68,Meteorlogical Office, Bracknell.

Eyre J.R., 1986: The effect of ozone on HIRS brightness temperatures Met O 19 Branch memorandum No. 84. Meteorlogical Office, Bracknell.

Eyre J.R., 1989: TOVS retrievals at the U.K. Meteorological Office. To be presented at ECMWF/EUMETSAT Workshop on "Use of Satellite Data in Operational Numerical Weather Prediction"; 8–12 May 1989; Reading, U.K.

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

Watts P.D. and A.P. McNally, 1988: The sensitivity of a minimum variance scheme to the values of its principal parameters. Tech. Proc. 4th Int. TOVS Study Conf.; Igls, Austria; 16–22 March 1988; Report of CIMSS, University of Wisconsin-Madison; Ed.: W.P. Menzel; pp. 399–412.