# THE SENSITIVITY OF A MINIMUM VARIANCE RETRIEVAL SCHEME TO THE VALUES OF ITS PRINCIPAL PARAMETERS

P D Watts[1] and A P McNally[2]

Robert Hooke Institute for Cooperative Atmospheric Research,
Clarendon Laboratory, Oxford, U.K.

[1] Meteorological Office Unit
[2] Department of Atmospheric, Oceanic and Planetary Physics

## 1. INTRODUCTION

The U.K. Met Office has recently implemented a TOVS retrieval scheme which uses a forecast profile as a first guess in a minimum variance retrieval of temperature and humidity (Eyre et al, 1986). An initial estimate of the profile with known error covariance and satellite measurements with known error covariance are combined with weights that minimise the expected error in the retrieved profile at all levels. Preliminary results with the scheme (as gauged by comparisons with collocated radiosondes) show that we now have retrievals of significantly higher accuracy than those from a previously used regression scheme (which effectively used a climatological first guess). However, comparison of the forecast with the radiosonde shows that the first guess and the retrieval are of comparable accuracy, suggesting that the retrieval is providing no additional information. Admittedly this validation is principally over N.W. Europe where the 6-12 hour forecast may be expected to be very good and any improvement difficult to make. Nevertheless, if we are modelling the problem correctly the minimum variance solution should be an improvement over the first guess. The purpose of this work was to find our sensitivity to errors in the 'model' thereby establishing why the theoretical improvement over the guess profile is not realised and also to obtain a better 'model'.

## 2. THEORY AND METHOD

The minimum variance solution is a linear one and assumes a linear forward problem:

$$\mathbf{Y} - \mathbf{Y_o} = \mathbf{K}.(\mathbf{X_t} - \mathbf{X_o}) \qquad (1)$$

where $\mathbf{Y}$ and $\mathbf{X_t}$ are the brightness temperature measurements and true atmospheric temperature profile and $\mathbf{Y_o}$ and $\mathbf{X_o}$ are the forecast values. $\mathbf{K}$ are brightness temperature derivatives $\mathbf{dY/dX}$ ('incremental weighting functions') and are

assumed, for now, to be independent of $\mathbf{X}$. Note that $\mathbf{X}$ includes both temperature and humidity profiles which are to be retrieved simultaneously. If the measurements $\mathbf{Y}$ and forecast $\mathbf{X}$ have error covariances $\mathbf{E}$ and $\mathbf{C}$ respectively, then the minimum variance solution, $\hat{\mathbf{X}}$, is

$$\hat{\mathbf{X}} = \mathbf{X_o} + (\mathbf{K}.\mathbf{C})^{\mathbf{T}}(\mathbf{K}.\mathbf{C}.\mathbf{K}^{\mathbf{T}} + \mathbf{E})^{-1}.(\mathbf{Y} - \mathbf{Y_o}) \tag{2}$$

$$\text{or} \qquad \hat{\mathbf{X}} = \mathbf{X_o} + \mathbf{W}.(\mathbf{Y} - \mathbf{Y_o})$$

(See Rodgers, 1976). The expected value $\mathbf{S}$ of $(\hat{\mathbf{X}} - \mathbf{X_t}).(\hat{\mathbf{X}} - \mathbf{X_t})^{\mathbf{T}}$ , may be written,

$$\mathbf{S} = \mathbf{C} - \mathbf{W}.\mathbf{K}.\mathbf{C} \tag{3}$$

The diagonals $S_{jj}$ are the expected error variances of the retrieval at the profile levels and the ratio $S_{jj}/C_{jj}$ gives the reduction in variance expected (this ratio is called the fractional unexplained variance, FUV). The matrices $\mathbf{C}$ , $\mathbf{K}$ and $\mathbf{E}$ constitute the 'model' discussed in the introduction and it is the sensitivity of the retrieval to these parameters with which we are concerned. This sensitivity is gauged by calculating the error covariance $\mathbf{S}'$ expected when the true conditions are $\mathbf{C}'$,$\mathbf{E}'$ and $\mathbf{K}'$ but $\mathbf{C}$,$\mathbf{E}$ and $\mathbf{K}$ have been used to derive $\mathbf{W}$. $\mathbf{S}'$ may be shown to be:

$$\mathbf{S}' = \mathbf{C}' - \mathbf{W}.\mathbf{K}'.\mathbf{C}' - (\mathbf{W}.\mathbf{K}'.\mathbf{C}')^{\mathbf{T}} + \mathbf{W}.(\mathbf{K}'.\mathbf{C}'.\mathbf{K}'^{\mathbf{T}} + \mathbf{E}').\mathbf{W}^{\mathbf{T}} \tag{4}$$

which reduces to equation 3 when $\mathbf{C}' \rightarrow \mathbf{C}$, $\mathbf{E}' \rightarrow \mathbf{E}$ and $\mathbf{K}' \rightarrow \mathbf{K}$. This equation can be used to study the theoretical sensitivity of the scheme to incorrect assumptions about the principal parameters, mainly by examining the diagonals $S'_{jj}$.

New estimates of $\mathbf{C}$ , $\mathbf{K}$ and $\mathbf{E}$ were obtained and used both in the simulation described above and to perform retrievals on real data. The source of the new estimates was a data set collected over three months of collocated forecast, measurement and radiosonde data. These data were split into two sets for the study, independent data being (arbitrarily) the even days of the month and the dependent data being the odd days. The dependent data were used to calculate the matrices and the retrievals were performed on the independent data. To ease description of the results the following convention is adopted. A result described as P1$\diamond$W(P2) indicates an estimator W derived using a matrix **P2** was used where the

true conditions were actually described by **P1**. For example, Ep$\diamond$W(E) describes the result of using a **W** derived using **E** with measurements that actually have an error covariance **Ep**. Normally there will be three simulations:

- $P \diamond W(P)$—the expected result with the original parameter,
- $P' \diamond W(P)$—how we may expect to 'suffer' given the wrong parameter in the operator,
- $P' \diamond W(P')$—how well we can expect to do with 'correct' parameters. (We expect $P'$ to be nearer the truth than $P$)

The following three sections deal with each matrix parameter in turn, firstly describing how the new estimate was obtained, secondly giving the results of the simulation experiments and finally the results using real data.

## 3. C - FORECAST ERROR COVARIANCE

### 3.1 C: Derivation

The original forecast error covariance matrix, denoted hereafter by C1, was obtained from the Meteorological Office Forecasting Research Branch and was calculated by comparing 12–hour forecasts with the next verifying analysis. A single matrix was supplied to describe all conditions. Certainly one problem with this method is that it will tend to underestimate forecast errors, especially in data–sparse areas, because the analysis uses the forecast as a background field. An attempt was made to allow for this with a simple multiplier. We may also suppose that the error characteristics of the forecast model will at least be a function of location, and that this approach may not assess correctly the inter–level correlations of error.

The new C matrix, denoted by C2, was estimated by comparing the forecast profile with the 'true' collocated radiosonde profile from the dependent data. The behaviour of C derived in this way when the collocation distance is reduced is not strong so we can be reasonably sure that forecast errors are not being seriously overestimated. Three C matrices were obtained; C2 from all the dependent collocation data, CW from collocations West of 14 Deg W representing an area of relatively high forecast error and CE from collocations East of 14 Deg W, an area of low forecast error. CE resembled the original C1 in size whereas the elements were much larger in CW. However, the correlation of forecast errors implied by CE, CW and C2 were very similar and generally higher than those in C1.

### 3.2 C: Simulations

The result C2◇ W(C1), is shown in figure 1. Compared with C1◇W(C1) it is apparent that even with the wrong operator, guess errors as described by C2 are easier to correct in the retrieval (at most levels) than those described by C1. Using the correct estimator, C2◇W(C2), the result is still better but only markedly so near the surface and above 400 mb. The C2 guess is easier to improve upon than C1 because the errors are more strongly correlated between levels. Measurements are able to 'see' errors that are highly correlated over a vertical range comparable to, or greater than, the weighting function width.

A second experiment was done with matrices CW and CE and the simulation suggested there is no advantage in such a stratification. Again the reason appears to be the similarity of the error correlations in the two areas. If there are areas or times where the forecast has significantly different error *correlations* then there may be scope for further improvement.

### 3.3 C: Real data

The result with real data equivalent to the first simulation is given in figure 2 showing, following the notation, Ind◇W(C1) and Ind◇W(C2) i.e. the result of using Ws calculated from the matrices C1 and C2 both applied to the independent data. As in the simulation there is some gain around 1000mb and above 300mb, though the latter is not significant below 150mb. There appears to be an increased error near the surface (1.5m $T_s$) which may be caused by the lack of a skin temperature measurement in the sonde report. Notice that at all levels with C1 the FUV is close to or greater than unity. The use of C2 at least reduces this to make it a useful retrieval around 1000mb. The effects on humidity retrieval is, in accordance with the simulation, negligible.

In summary, the C matrix derived from collocated radiosonde data appears to be a better description of the guess error than C1 and improves the retrieval accuracy. That this method of deriving C is subject to unwanted collocation errors appears not to be a serious objection as the behaviour of C with decreasing collocation distance is not strong. As expected from simulation, simple stratification by areas of high and low forecast error is not an advantage.

## 4.1 E: Derivation

The error covariance of the measurements is poorly known although we have reasonable bounds on the diagonal elements (see Watts, 1984). Errors in different channels are almost certainly correlated (by preprocessing, cloud–clearing procedures etc.). The measurement error is taken to refer to the term $(\mathbf{Y} - \mathbf{Y_o})$ and consequently should include any random error in the radiative transfer calculations. The original **E** matrix used in the scheme was diagonal implying no correlations between errors in any of the channels. The values for the diagonals were chosen to lie between two limits, a minimum defined by the radiometric noise (very low in HIRS channels) and a maximum determined by routine comparisons of measured brightness temperatures and brightness temperatures calculated from collocated radiosondes.

Essentially the same procedure was used for the new **E** matrix using the radiosonde profiles from the dependent collocation data but the full covariance of errors was derived rather than just the variances in each channel. Three **E** matrices were derived because the 'clear' radiances are obtained through three distinct routes namely: clear FOVs where no cloud is detected, FOVs where an adjacent FOV has been used to estimate the clear radiances ($N^*$ method), and FOVs where cloud conditions preclude the $N^*$ method and the HIRS brightness temperatures are obtained by regression on MSU data (for details see Eyre and Watts 1987). The **E** matrices thus derived have different characteristics both in the size of the elements and the correlations. The clear **E** is most like the diagonal **E** originally used with low off–diagonal values. The $N^*$ **E** has larger off–diagonals with some strong negative correlations. Strongest correlations (up to 0.5) are found for **E** in the HIRS estimated from MSU.

## 4.2 E: Simulations

A sample simulation result, that for the $E_{msu}$ (HIRS–MSU regression), is shown in figure 3. The three results shown are,

- solid line; $E_{diag} \diamond W(E_{diag})$ i.e. our originally assumed expected error,
- dashed line; $E_{msu} \diamond W(E_{diag})$ i.e. what we may expect when the diagonal **E** is used on real data and
- dotted line; $E_{msu} \diamond W(E_{msu})$ representing the best we may expect given our best estimate of the measurement error characteristics.

The detrimental effect of using $E_{diag}$ on MSU–regression data is substantial throughout the profile, and affects both temperature and humidity. A smaller effect is found for 'clear' and $N^*$ cases. In most cases the accuracy of the profile is mostly restored when the correct **E** is used though there is always some loss because the non–zero correlations in the new **E** matrices imply less information. Interestingly, the size of effect is much larger than found with the **C** matrix simulations and is actually capable of pushing the FUV up to and over 1.0 which is what is found with the original routine scheme. Use of the correct **E** reduces this to $\approx$ 0.8 or less representing a useful retrieval.

### 4.3 E: Real data

Results analogous to the simulations just described but using the real independent data were obtained and the sample MSU–regression shown in figure 4. Here the three lines are,

- solid; $Ind\Diamond W(E_{diag}, C1)$,
- dashed; $Ind\Diamond W(E_{diag}, C2)$ and
- dotted; $Ind\Diamond W(E_{msu}, C2)$.

In this case the using the new **E** much reduces the retrieval error in temperature and significantly reduces it in humidity. In temperature the FUV reduction is $\approx$ 0.25 at maximum decreasing to 0.05 at the tropopause. More importantly, the FUV is reduced to less than 1.0 and, considering the large proportion of soundings that this route provides ($\approx$ 80%), this represents a considerable improvement.

In the clear cases for temperature retrieval any change was small (0.1 or 0.2 FUV) but positive. In the humidity the gain was larger though still not close to the simulated effect. With $N^*$, more gain was made at most levels in temperature though far less than the simulations suggested. The effect on humidity was as in the clear case.

In summary, the new **E** matrices are substantially different from the original diagonal **E** and differ also according to the cloud–clearing route used. The simulated effect of these **E** matrices is large and, at least in the case of MSU–regression FOVs, borne out in real retrievals. This case, because of the large proportion of soundings it constitutes, is the most important. For clears and $N^*$s the effect is smaller than simulated though still positive and useful. The relatively small numbers of these cases available when generating the **E** matrices may mean the estimates are not close to the best possible and we feel that a larger impact, in line with that made on the MSU–regression FOVs, should be possible.

## 5. K - INCREMENTAL WEIGHTING FUNCTIONS

### 5.1 K: Derivation

The original **K** matrix was the mean of the incremental weighting functions calculated for 100 randomly selected mid–latitude profiles. The standard deviation about this mean is found to be fairly high (5–10% of the mean value), i.e. there is a significant dependency on the profile – the problem is non–linear. In order that our model be correct, **K** should be evaluated at the true profile or, this being impossible, at the guess profile, $\mathbf{K_o}$. **W** would then be calculated using $\mathbf{K_o}$. This procedure would be computationally expensive and would only be justified if the retrieval were found to be relatively sensitive to the exact value of **K**.

The new **K** was obtained by randomly extracting a profile from the mid–latitude set. This is not an improved **K** of course but it allows us to gauge the sensitivity to this part of the model.

### 5.2 K: Simulations

Equation 4 is derived on the basis that the retrieval can be written;

$$(\hat{\mathbf{X}} - \mathbf{X_o}) = \mathbf{W}.(\mathbf{Y} - \mathbf{Y_o}) = \mathbf{W}.\mathbf{K}'.(\mathbf{X_t} - \mathbf{X_o}) + \mathbf{W}.\epsilon \qquad (5)$$

Where $\mathbf{X_t}$ is the true profile and $\epsilon$ is a measurement noise vector. In practise the guess brightness temperatures are calculated with a full radiative transfer routine using approximately the correct physics, the measurements by definition use the correct physics. We may use $\mathbf{K}'$ to simulate the difference between the true atmosphere and the linearised model, **K**, used in **W**.

The usual third simulation $K'\Diamond W(K')$ was not done because of the symmetry of the problem; the result of $K'\Diamond W(K')$ should be approximately $K\Diamond W(K)$. Five cases were tried but since the results were very similar only one is shown here in figure 5. The temperature retrieval is moderately affected in both cases with 0.05–0.1 K increase in error or 0.05 FUV throughout the profile. The dramatic effect is in the humidity retrieval which is degraded typically by 5-10% relative humidity or up to 0.4 FUV. It demonstrates the non-linearity of constituent retrieval; **K** is strongly dependent on the humidity part of the profile for water vapour channels.

### 5.3 K: Real data

It appears necessary to calculate **K** and therefore **W** at the true profile if the humidity retrieval is to be anywhere near the optimum result. This is clearly

impossible in practise and the best that can be done is to evaluate $\mathbf{K}$ at $\mathbf{X}_o$. This approach, however, is tending towards a fully non–linear scheme which is planned for the next generation of retrieval schemes and beyond the scope of this study. More in keeping with the approach adopted here is to see if there is some simple *ad hoc* stratification that provides a $\mathbf{K}$ closer to the true value than the climatological one used currently. Most of the non–linearity of the problem arises from humidity profile and, because a high temperature indicates a high absolute humidity, we stratified $\mathbf{K}$ according to temperature. The effect of using this stratification on the independent real data set is not shown but was almost totally negligible. It is possible that the collocated radiosonde humidity is not a good ground truth for testing retrievals against but a substantial effect was seen with the $\mathbf{E}$ changes. It therefore appears that a simple stratification is not adequate and a full on–line calculation of $\mathbf{K}$ may be required.

## 6. CONCLUSIONS

Significant improvements can be made to the performance of the retrieval scheme by using error covariances that are estimated from real data. A guess error covariance was calculated from forecast profiles and collocated radiosondes and was found to give results broadly in agreement with simulations; improvements around 1000 mb and above 300 mb. Also in agreement with simulations, stratification into high and low error forecast areas proved not to be profitable. The effect of more realistic measurement error covariances was more marked, probably because the original (diagonal) estimate was so crude. Stratification into cloud–clearing routes was found to be necessary. The gain for clear and partly cloudy FOVs did not match the simulated effect and we feel there is perhaps some statistical problem remaining here. Attempts at simple stratification of the $\mathbf{K}$ matrix proved useless, and improvement here awaits further investigation in fully non–linear retrieval schemes.

The new $\mathbf{C}$ and $\mathbf{E}$ matrices have been adopted by the routine scheme and work will continue on their refinement, especially the estimation of $\mathbf{C}$ since it is probable that the forecast guess will remain central to routine processing of sounding data in the Met Office for the foreseeable future.

## REFERENCES

EYRE, J.R., PESCOD, R.W., WATTS, P.D., LLOYD, P.E., ADAMS, W. AND ALLAM, R.J.(1986), TOVS retrievals in the U.K.: progress and plans. *Tech. Proc. 3rd Int. TOVS study conf.; Madison, Wisconsin, USA; 13–19 Au-*

*gust 1986; pp 60–91. Report of CIMSS, University of Wisconsin, Madison. Editor: W. P. Menzel.*

EYRE, J.R. AND WATTS, P.D. (1987), A sequential approach to cloud–clearing for satellite temperature sounding. *Q.J.R.Meteorol.Soc.* **113**, *1349–1376*

RODGERS, C.D. (1976), Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics and Space Physics*,**14**, *609–624*

WATTS, P.D. (1984), A study of local area synthetic coefficients for use in the LASS system. *Met O 19 Branch Memorandum No. 76, Meteorological Office, Bracknell, U.K.*

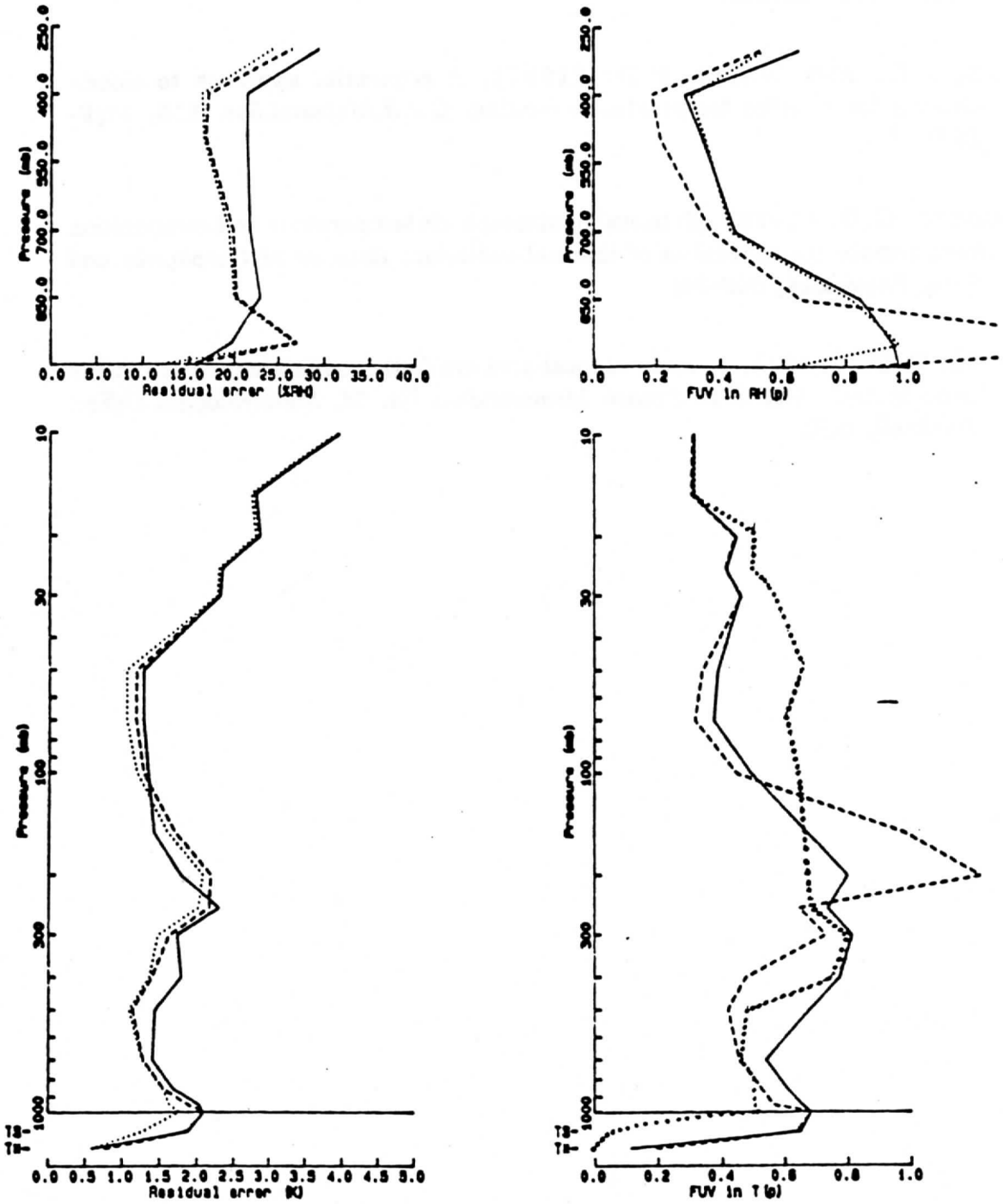


Figure 1. C-matrix Simulation

C1◇W(C1) ——————

C2◇W(C1) - - - - - - -

C2◇W(C2) .....

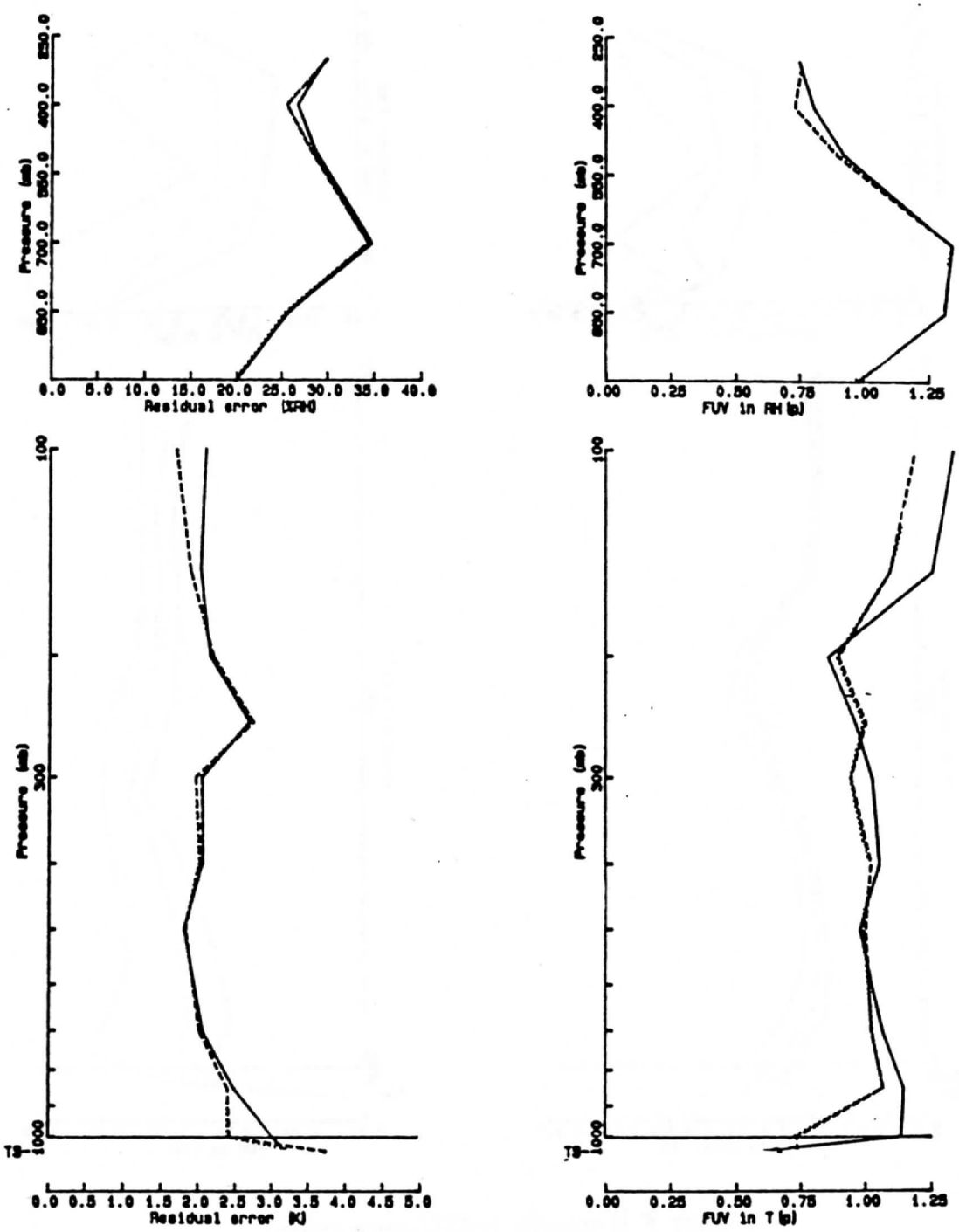


Figure 2. C-matrix Real data

Ind<>W(C1) ——————

Ind<>W(C2) - - - - - - - -

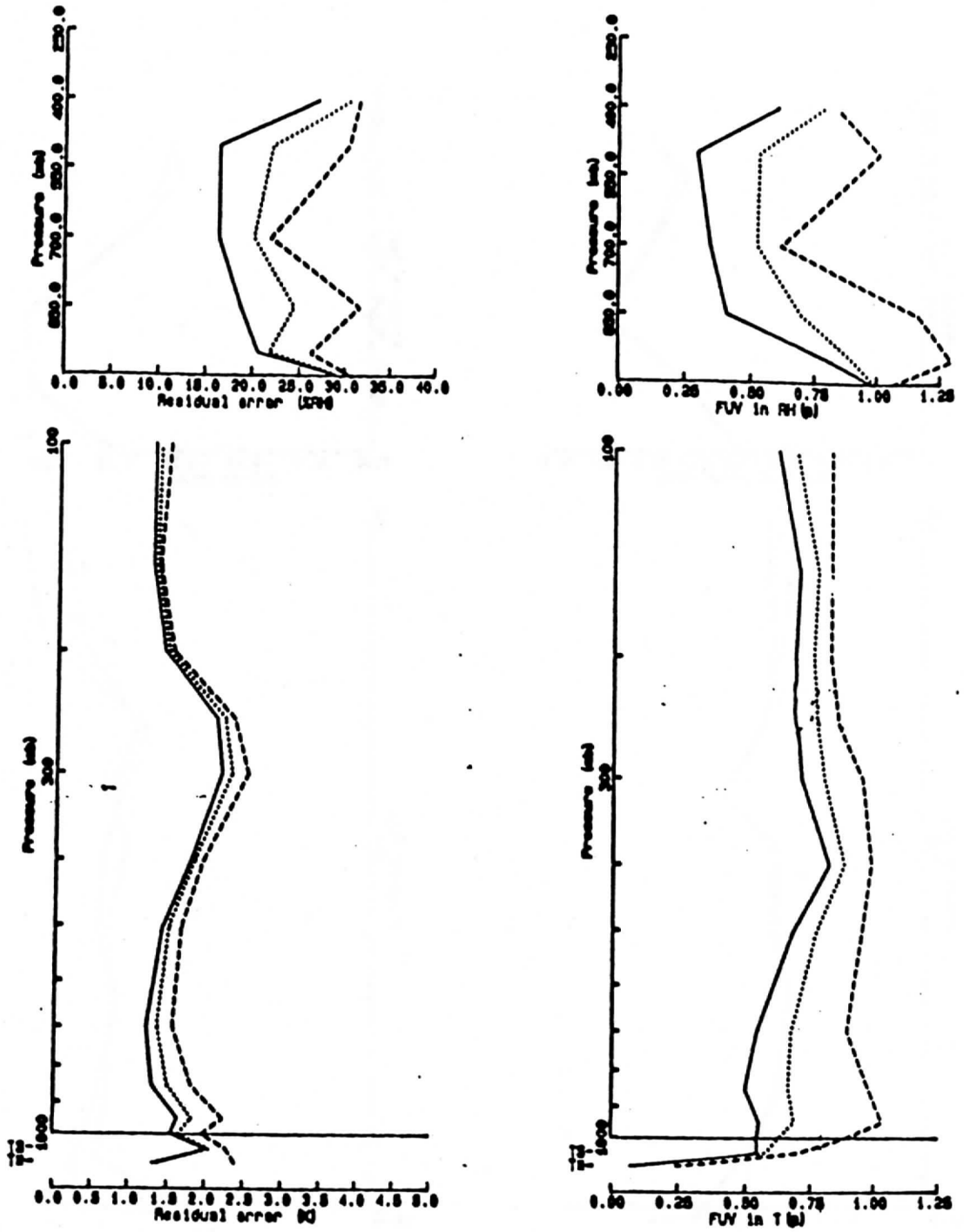


Figure 3. E-matrix (MSU-regression)

Simulation

$E_{diag} \diamond W(E_{diag})$ ——————

$E_{msu} \diamond W(E_{diag})$ - - - - - - -

$E_{msu} \diamond W(E_{msu})$ ..... .....

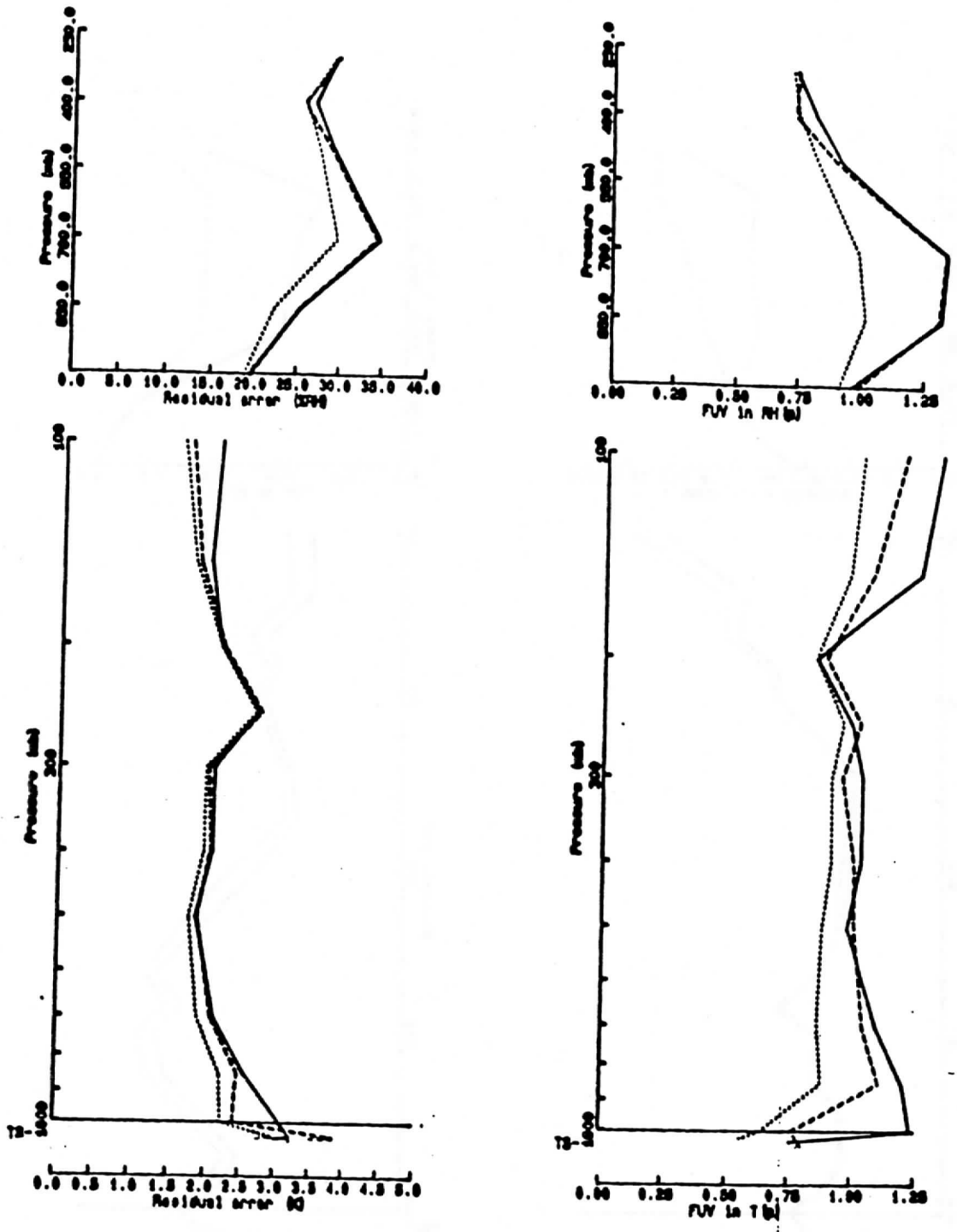

Figure 4. E-matrix (MSU-regression)

Real data

$Ind\Diamond W(E_{diag}, C1)$ ————

$Ind\Diamond W(E_{diag}, C2)$ - - - - - - -

$Ind\Diamond W(E_{msu}, C2)$ ·····

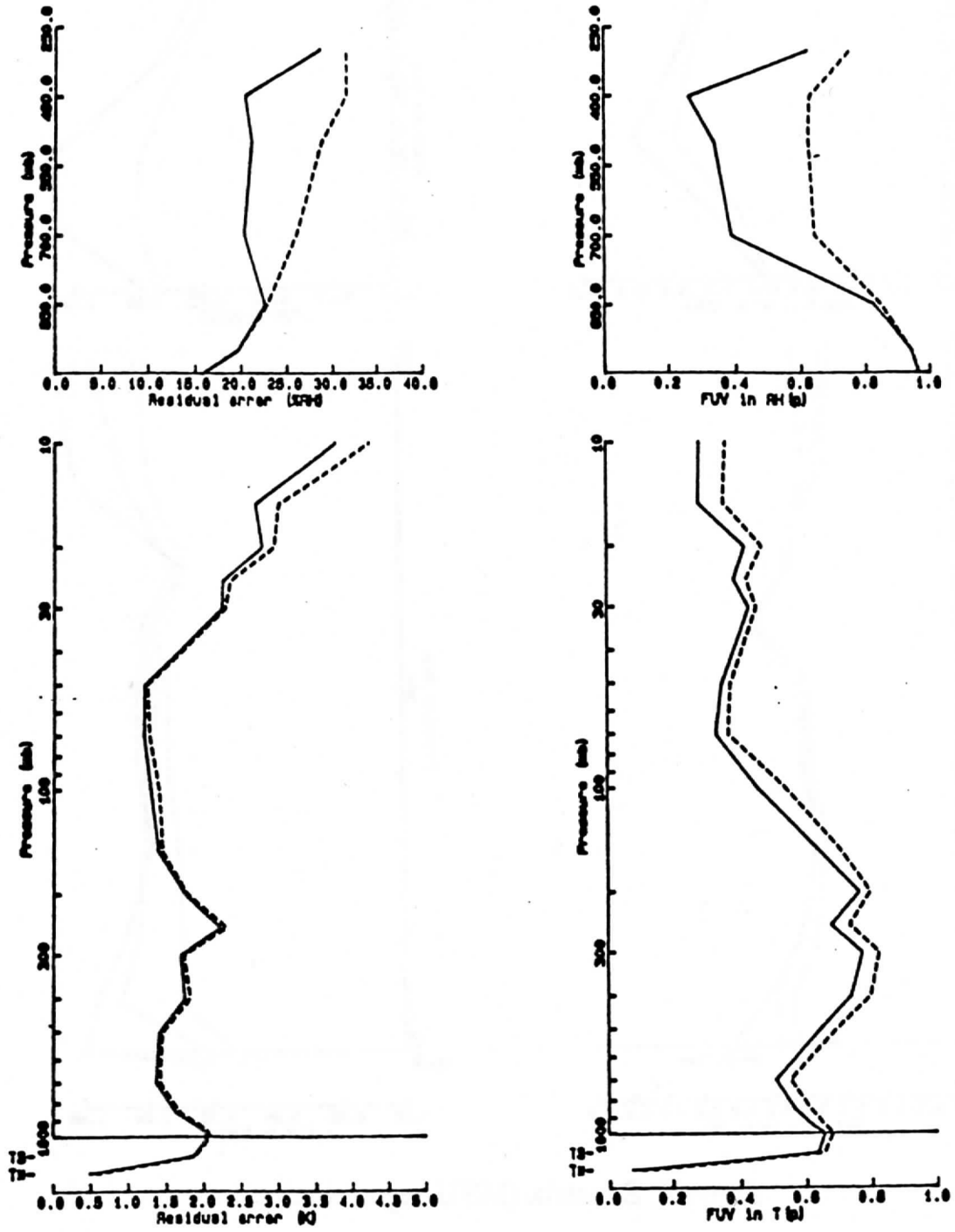


Figure 5. K-matrix Simulation

K◇W(K) ——————

K'◇W(K) - - - - - - - -

# The Technical Proceedings of

# The Fourth International TOVS Study Conference

Igls, Austria

March 16-22, 1988

Edited by

W. P. Menzel

Cooperative Institute for Meteorological Satellite Studies
Space Science and Engineering Center
University of Wisconsin
1225 West Dayton Street
Madison, Wisconsin 53706
(608) 262-0544

October 1988