



LONDON, METEOROLOGICAL OFFICE.

Met.O. 19 Branch Memorandum No.11.

Some results on the comparison of retrieval methods - maximum probability, minimum information and Tychonov regularization. By GODDARD, J.W.F. and HUNT, R.D.

London

London, Met. Off., Met.O.19 Branch Mem.No.11,  
[1974], 31cm. Pp.9, pls.7.5 Refs.

FGZ

National Meteorological Library  
and Archive

Archive copy - reference only



Met O 19 Branch Memorandum (No 11)



0119737

Some results on the comparison of retrieval methods - maximum probability,  
minimum information and Tychonov regularization

J W F Goddard & R D Hunt

Permission to quote from this unpublished memorandum should be obtained from the  
Head of Met O 19, Meteorological Office, Bracknell, Berks, RG12 2SZ.



SOME RESULTS ON THE COMPARISON OF RETRIEVAL METHODS - MAXIMUM PROBABILITY,  
MINIMUM INFORMATION AND TYCHONOV REGULARIZATION

J W F Goddard and R D Hunt

HA/IG/12

Introduction

In order to compare three methods of retrieving vertical temperature profiles of the atmosphere from radiance measurements, large numbers of retrievals were carried out using both data produced artificially from rocket/radio-sonde ascents and data from the Nimbus E satellite. In the first case the results could be compared on a quantitative basis, whereas in the case of Nimbus E data, a more or less qualitative appraisal was possible.

One of the main results to emerge from the study was that a clear knowledge of the relative importance of the various atmospheric levels (this judged presumably by the users of the data) is essential before a choice can be made between these three retrieval methods and any others which will be considered in future work.

Retrieval methods studied

In this section the following notation will be used:

- W matrix of weighting functions of order  $4 \times 50$ , 4 being the number of Nimbus 5 channels used in the study and 50 being the usual levels at which profiles are defined (see HA/IG/8<sup>1</sup>)
- B vector of Planck profile to be retrieved from the radiance measurements (order  $50 \times 1$ ).
- R vector of measured radiances ( $4 \times 1$ )
- C matrix of atmospheric covariances ( $50 \times 50$ ) appropriate to particular latitude/season groups (see HA/IG/1<sup>2</sup>)
- $\bar{B}$  vector of mean Planck profile for particular latitude/season group ( $50 \times 1$ )
- $\bar{R}$  vector of mean measured radiances for a particular latitude/season group ( $4 \times 1$ )
- E matrix of instrumental noise covariances ( $4 \times 4$ ) assumed to be diagonal

The superscript T will denote the transpose of a matrix.



(a) Maximum probability method (referred to hereafter as method 1)

This has been described in earlier reports (eg HA/IG/3<sup>3</sup>). The solution it produces is the one of the infinite number which exist which is most likely to have occurred, this being based on statistical samples of atmospheric profiles. The equation solved is

$$B - \bar{B} = CW^T (WCW^T + E)^{-1} (R - \bar{R}) \quad (1)$$

(b) Minimum Information method (method 2)

This is a special case of method 1, with the assumption made that

(i) C is a diagonal matrix (ie temperatures at different levels of the atmosphere are uncorrelated) with a constant variance,  $\sigma_a^2$ , at each level, ie  $C = \sigma_a^2 I_{50}$  ( $I_{50}$  being the 50 x 50 unit matrix) and

(ii) E can also be expressed in a similar fashion, ie  $E = \sigma_n^2 I_4$

With these simplifying conditions, (1) can now be expressed as

$$B - \hat{B} = W^T (WW^T + \lambda I_4)^{-1} (R - \hat{R}) \quad (2)$$

where  $\lambda = \frac{\sigma_a^2}{\sigma_n^2}$  ( $\approx 10^{-3}$ ).  $\hat{B}$  and  $\hat{R}$  have been introduced

rather than  $\bar{B}$  and  $\bar{R}$  because it is no longer necessary to work in terms of deviations from the mean but, instead, using deviations from any initial guess solution, (2) can be iterated until the 'measured' radiances from one iteration are close to the actual measured radiances (within the instrumental noise limits for instance).

(c) Tychonov Regularisation method (method 3)

The Tychonov Regularisation technique, described for example by Bellman<sup>4</sup>, is useful for solving matrix equations of the form  $Ax = y$ , where A is an ill-conditioned matrix. Rather than minimise the quadratic form  $(Ax-y, Ax-y)$  which may not produce a stable solution, the expression  $(Ax-y, Ax-y) + f(x)$  is minimised, where  $f(x)$  is some function chosen to improve the stability of the problem.



In relation to the retrieval situation, where the matrix  $W^T W$  is very ill-conditioned, the most useful form of the stabilising function is the quadratic expression  $\lambda ((B - \hat{B}) - D, (B - \hat{B}) - D)$  where  $\hat{B}$  is, as with method 2, some initial guess and  $D$  is an estimate of  $(B - \hat{B})$ . We then need to minimise the expression

$$(W(B - \hat{B}) - (R - \hat{R}), W(B - \hat{B}) - (R - \hat{R})) + \lambda ((B - \hat{B}) - D, (B - \hat{B}) - D) \quad (3)$$

the minimum value being

$$(B - \hat{B}) = (W^T W + \lambda I_{50})^{-1} (W^T (R - \hat{R}) + \lambda D) \quad (4)$$

which can be solved iteratively and where  $D$  is zero for the first iteration and subsequently assumes the value of L.H.S. from the previous iteration. Because of the matrix identity  $W^T (W W^T + \lambda I_4)^{-1} = (W^T W + \lambda I_{50})^{-1} W^T$ , this solution differs from method 2 only in the inclusion of  $\lambda D$  on the R.H.S. of equation (4). The choice of  $\lambda$  is more or less one of trial and error, however physical reasoning leads one to choose a similar value to that used in method 2, namely  $10^{-3}$ .

#### Data used

The data used in the comparison study were:-

1. Sample of 207 rocket-sonde/radio-sonde profiles in the latitude-season group winter  $50^\circ - 70^\circ$  (see HA/IG/1) converted into 'measured' radiances using the weighting functions for the  $B_{12}$ ,  $B_{23}$ ,  $B_{34}$  and  $A_1$  channels on Nimbus E and adding random noise.
2. Sample of 25 rocket-sonde/radio-sonde profiles in the latitude-season group winter  $30^\circ - 50^\circ$  also converted into 'measured' radiances.
3. Radiances from the  $B_{12}$ ,  $B_{23}$ ,  $B_{34}$  and  $A_1$  channels of Nimbus E for the dates 12-14 February 1973.

Hereafter, these data will be referred to as data 1, 2 and 3 respectively.

#### Results using data 1 and 2

Table 1 shows some of the results of the retrievals from all three methods using data 1. As the sample retrieved was not independent, the results of the retrievals could be compared directly with the true values. With methods 1 and 2, results are stated using different values of  $\lambda$  in equations (2) and (4) and also starting with initial guesses other than the sample mean. From these results and from those using data 2, the following comments can be made about the various methods.



(a) Dependence on  $\lambda$  (methods 2 and 3)

It is apparent that with method 2, setting  $\lambda = 0$  produces worse results than with  $\lambda$  given some realistic value, although the exact value of  $\lambda$  seems to be unimportant, (certainly to within an order of magnitude). There is only a small improvement in the results when  $\lambda$  is increased from  $10^{-3}$  to  $5 \times 10^{-3}$ , but, as can be seen from table 1, an improvement in standard error for example of the 30-10 mb thickness of 4.6 decametres (dm) is caused by increasing  $\lambda$  from 0 to  $10^{-3}$ . The number of iterations required to produce a solution also depends on  $\lambda$ . With  $\lambda = 10^{-3}$  one iteration is usually all that is needed. Increasing  $\lambda$  has the effect of increasing the number of iterations required although two or three sufficed on most occasions.

With method 3, no results could be obtained with  $\lambda = 0$ , this due to the fact that the matrix  $(W^T W)^{-1}$  in equation (4) contains some terms of the order of  $10^{25}$ . However, the ill-conditioned nature of the matrix is improved considerably with the addition of a small  $\lambda$  onto the diagonal terms. With  $\lambda = 10^{-4}$ , the inverted matrix contains terms mainly of the order  $10^2$  or  $10^3$  and with this value of  $\lambda$  the best results with this data are produced. However, apart from the 10-3 mb and 30-10 mb thicknesses, increasing  $\lambda$  from  $10^{-4}$  to  $10^{-2}$  only produced a small deterioration in the retrieved profiles. The 30-10 mb layer showed a deterioration in standard error of 6.0 dm. Again, only one iteration is necessary with data 1 and the optimum value of  $\lambda$ , but more are required as  $\lambda$  is increased.

(b) Dependence on initial guess (methods 2 and 3)

Data 1 was retrieved by methods 2 and 3 using both the sample mean and the mean for the tropical sample as initial guess. Some retrievals were also carried out with method 2 using nearly isothermal atmospheres (ie isothermal up to about 0.8 mb with a mean profile above) as initial guesses. In general, the accuracy of the retrievals (in terms of thickness values) is not greatly dependant on the initial guess, although, as would be expected,



the best retrievals were obtained with the sample mean as starting solution. These results agree essentially with those described by Smith, Woolf and Fleming<sup>5</sup>. The retrievals performed with the nearly isothermal atmosphere starting solution were only slightly worse in the layer between about 50 mb and 1 mb. Figure 1 shows a profile in the winter 30°-50° sample retrieved by method (2) using the sample mean and, secondly, the isothermal atmosphere as initial guess.

(c) Relative Accuracy of layer retrievals

Due to the configuration of the weighting functions, certain thicknesses are retrieved better than others with all three methods. Using method 1, the 3-1 mb layer is retrieved well, but with methods 2 and 3 this layer produces some of the worst results. The 100-30 mb layer is retrieved well with methods 2 and 3, probably because this region is mainly within the range of only one weighting function. This is borne out by the fact that the 10-3 mb thickness is retrieved appreciably better with data 1 and method 2 when the  $B_{23}$  channel is omitted leaving the layer mainly within the range of the  $B_{34}$  channel. (It should be added that the 3-1 mb and 30-10 mb thickness retrievals are worsened by this). Another point, discussed in HA/IG/3, concerns the 'cancelling out' effect of positive and negative errors of retrieved temperature over large layers. Hence the 100-1 mb and 300-30 mb thicknesses have, in percentage terms and sometimes even in absolute terms, smaller errors than the constituent layers. For example with data 1, method 2 and  $\lambda = 10^{-3}$ , the 100-1 mb layer has a standard error of 11.5 dm whereas the 3-1 mb layer has an error of 15.4 dm.

(d) Comparisons between methods

Generally speaking, using data 1 and 2, methods 2 and 3 compare favourably with method 1 for the large layers, but for the smaller thicknesses, the statistical information implicit in method 1 gives it an advantage. The 100-1 mb layer, for example, has a standard error of 18.5 dm, 11.5 dm and 11.3 dm respectively with methods 1, 2 and 3 while the same figures for the 3-1 mb layer are 5.5 dm, 15.4 dm and 14.8 dm.

Although there is no clear advantage, method 2 produces slightly better retrievals than method 3 with this data.



### Results with data 3

Figures 2, 3 and 4 show the objectively analysed mean daily 100-1 mb layer for the northern hemisphere for the 13 February 1973 retrieved employing methods 1, 2 and 3 respectively and with the same field for the 12 February used as a background field. Clearly the major features north of about  $30^{\circ}\text{N}$  are reproduced similarly in all of the charts, the chief difference being in the shape and axis of the large trough centred near  $50^{\circ}\text{E}$  longitude. Close to the equator all the charts are rough, that produced by method 2 being the roughest. Method 1 produces values about 20 dm higher than the others over almost the whole hemisphere. Other thickness fields show biases in the other direction.

Figures 5 and 6 show the 100-30 mb chart for the same day produced by methods 1 and 2. The former is certainly smoother and looks more realistic (see HA/IG/8 for comparisons between method 1 charts and those produced from radio-sonde data). One of the worst examples of a poorly retrieved field was the 30-10 mb field for the southern hemisphere on the 13 February, produced by methods 2 and 3. Totally unrealistic charts were produced whereas method 1 gives a very smooth field for the same level. These results tend to confirm those discussed in previous sections, method 1 being more advantageous for smaller layers.

Table 2 shows the number of observations which were unretrievable for data on 12 February. (These arise when  $B-\hat{B}$  or  $B-\bar{B}$  is large and negative, making the final B negative).

	S. Hem	N. Hem
Method 1	28	5
Method 2	397	558
Method 3	484	667
Total no of obs	1945	1855

Table 2

Reasons for the much smaller number of observations lost with the maximum probability method are discussed in the next section.



### Other comments

The study of retrievals performed with data 3 reveals an important difference between method 1 and the others. Because of the iterative techniques employed in methods 2 and 3, the final solutions must give rise to the same radiances as those actually measured. The maximum probability method, however, has no such guarantee and, although for data which is statistically similar to the original sample the final solution will very closely obey the above criterion, some of the solutions produced with the real data have radiances several units different from those actually measured. In practical terms, this appears to be an advantageous property as the roughnesses in the measured radiances tend to be smoothed out by the retrieval process, and profiles which are at least realistic are produced. Methods 2 and 3 showed that there were a not insignificant number of observations which, when retrieved, converged quickly to totally unrealistic profiles (some containing negative Planck radiances as mentioned above) irrespective of the initial guess. Further work is being done to see whether this property of the minimum information method is improved by smoothing the radiances before retrieval rather than relying on the analysis to smooth the results.

Arising from these points, work was done with the minimum information method to observe the effects on the retrieved solution of small measured radiance charges. Because of the overlapping nature of the weighting functions and the fact that the final solution must have exactly the same radiances as those measured, it was found, not unexpectedly, that relatively small charges in radiances measured on any one channel (say about 3 radiance units) have a significant effect on the final solution. Figure 7 illustrates this point. A relative charge between the two adjacent channels of about 9 units was sufficient to produce a totally unrealistic profile.

### Conclusions

This work has emphasised two main points concerning the choosing of the best retrieval method:-



- (a) The choice depends on the configuration of the weighting functions for the particular instrument under consideration and
- (b) It also depends on the relative importance attached to various atmospheric levels and the kind of vertical resolution sought.

Other retrieval methods (including direct regression of radiance measurements against objectively analysed thickness fields derived from conventional data) will also need to be considered.

#### References

1. HA/IG/8                      The production of mean daily radiance charts and high-level thickness charts from Nimbus E data; R.D. Hunt; Met O 19 internal report, 1974.
2. HA/IG/1                      Vertical temperature profiles of the troposphere and stratosphere on punched cards; B.R. Barwell and G.C. Hoskin Met O 19 internal report, 1972.
3. HA/IG/3                      Thickness retrievals using various sets of weighting functions and values of instrumental noise; R.D. Hunt Met O 19 internal report, 1973.
4. Bellman, R                    Introduction to matrix analysis pp 362. McGraw-Hill.
5. Smith, W.L.  
    Woolf, H.M. and  
    Fleming, H.E.                Retrieval of atmospheric temperature profiles from satellite measurements for dynamical forecasting. Journal of Appl. Met, Vol.11, pp 113-122, February 1972.



METHOD 1

METHOD 2 - STANDARD ERRORS

Layer	Method 1 Standard Error	Sample mean as 1st guess			Tropical mean 1st guess $\lambda = 10^3$	as 1st guess $\lambda = 10^{-3}$ B <sub>23</sub> missing
		$\lambda = 0$	$\lambda = 10^{-3}$	$\lambda = 5 \times 10^{-3}$		
100-1 mb	18.5	11.5	11.5	11.4	11.1	14.4
300-30	10.3	10.8	9.7	9.5	12.6	10.5
3-1	5.5	14.5	15.4	15.8	16.5	19.2
10-3	7.9	14.4	12.0	11.2	15.9	8.6
30-10	7.2	14.4	9.8	8.4	12.6	14.2
100-30	11.0	5.2	4.2	4.3	4.8	4.7
300-100	12.4	11.2	11.0	10.9	13.3	11.1

METHOD - 3 STANDARD ERRORS

Layer	Sample mean as initial guess			$\lambda = 10^{-2}$	Tropical mean 1st guess $\lambda = 10^{-4}$
	$\lambda = 10^{-4}$	$\lambda = 5 \times 10^{-4}$	$\lambda = 10^{-3}$		
100-1 mb	11.3	12.1	12.3	12.4	11.0
300-30	10.4	10.5	10.9	12.5	14.1
3-1	14.8	14.9	14.8	14.4	16.3
10-3	13.6	13.9	15.0	16.7	16.3
30-10	12.6	14.0	15.3	18.6	13.0
100-30	4.9	4.7	5.1	6.3	5.0
300-100	11.2	11.1	11.3	11.6	13.9

Table 1 Standard errors (retrieved - true) of thickness retrievals using methods 1, 2 and 3 and data 1 (sample of 207 winter 50° - 70° cases).  
(Units - decametres).



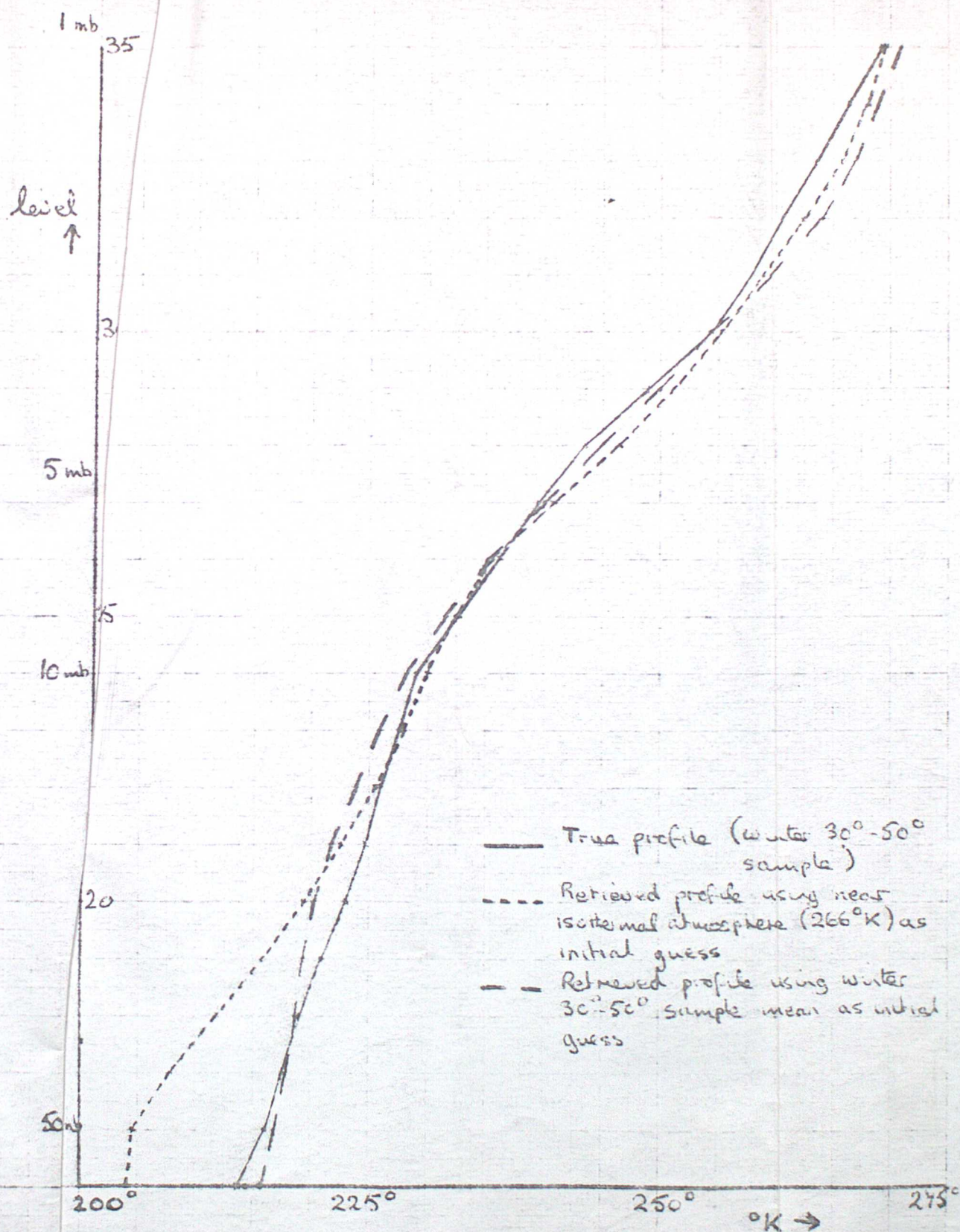


Figure 1 Comparison of retrievals using different initial guesses with minimum information method





FIGURE 2

100-1 mb THICKNESS

13th FEBRUARY 1973

MAXIMUM PROBABILITY METHOD

(UNITS - DECAMETRES)

Polar Stereographic Projection  
MET. O. C. 1968, D. 11-122



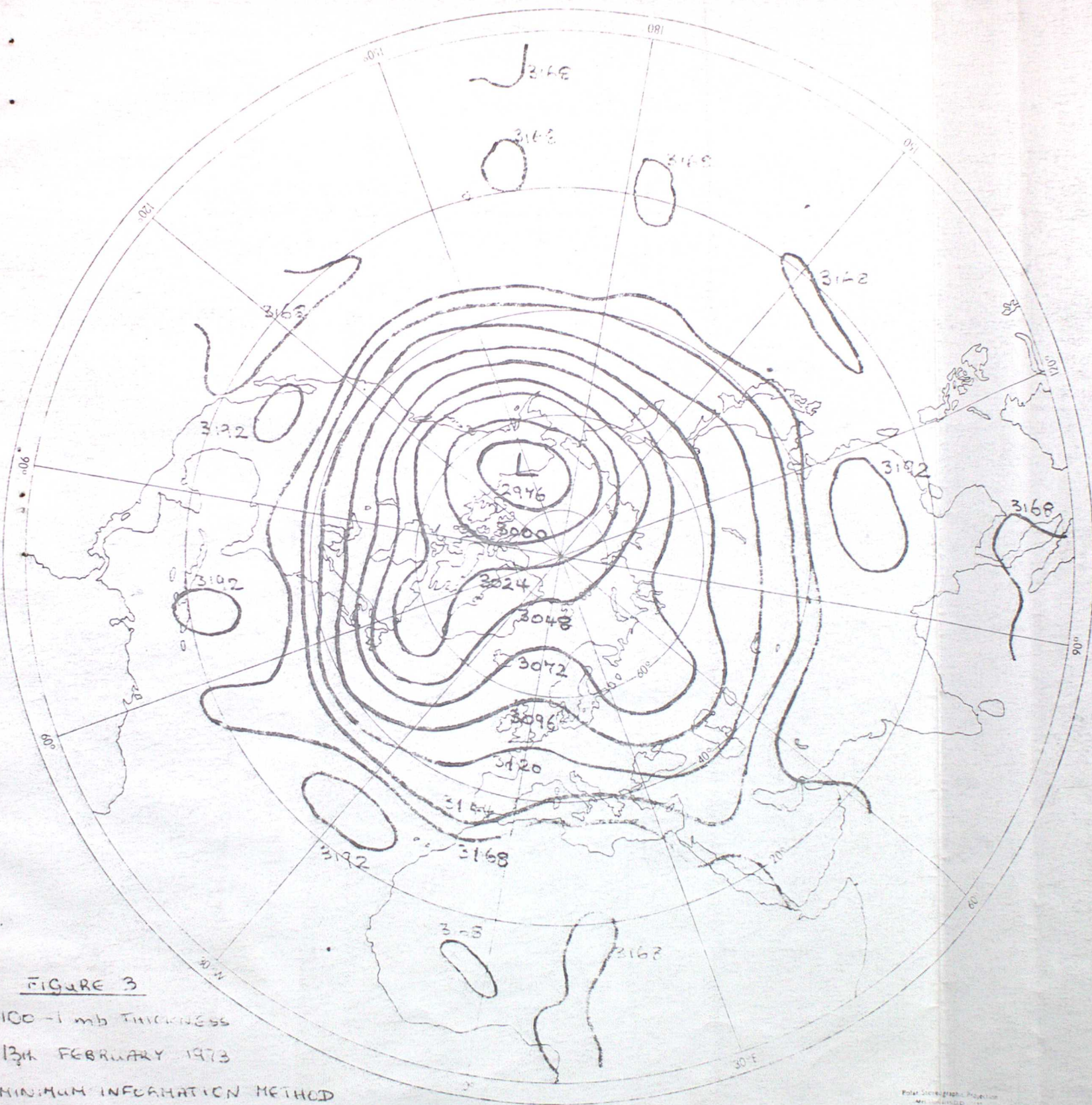


FIGURE 3

100-mb THICKNESSES

13th FEBRUARY 1973

MINIMUM INFORMATION METHOD

(UNITS - DECAMETERS)

Polar Stereographic Projection  
Meters/Decimeter = 10









FIGURE 5

100 - 30 mb THICKNESS

13th FEBRUARY 1973

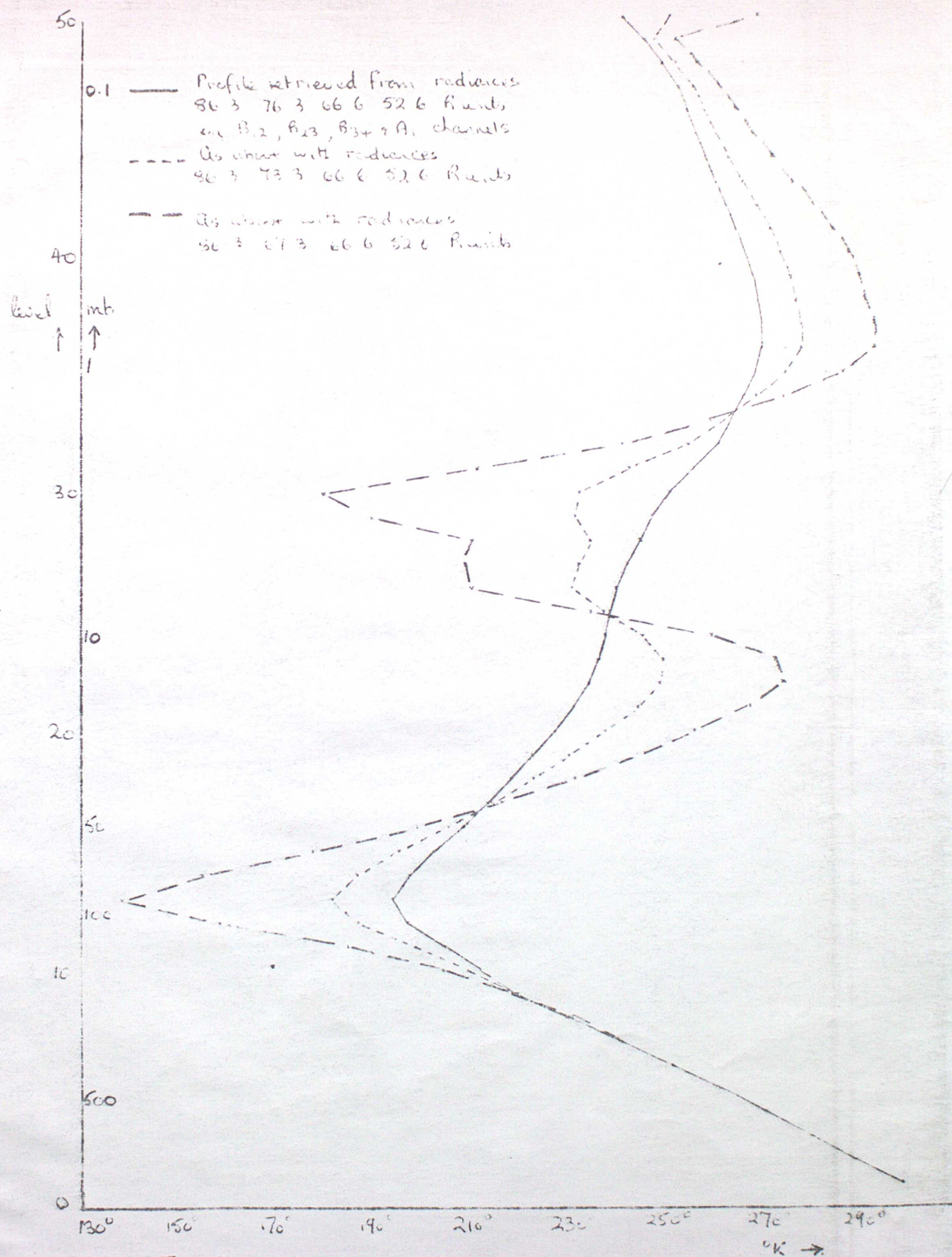
MAXIMUM PROBABILITY METHOD

(UNITS - DECAMETRES)









**Figure 7** Profiles retrieved from radiance measurements with alterations to the B<sub>23</sub> channel radiances.