INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460392 - I

SMR/406-12

THIRD AUTUMN WORKSHOP ON ATMOSPHERIC RADIATION AND CLOUD PHYSICS
27 November - 15 December 1989

"Review of Cloud Clearing Techniques"

Jonathon R. EYRE
University of Oxford
Hooke Institute
Oxford
UNITED KINGDOM

Please note: These are preliminary notes intended for internal distribution only.

ICTP Workshop on "Atmospheric Radiation and Cloud Physics".
27 November – 15 December 1989.

J.R. Eyre *Lecture 1*

REVIEW OF CLOUD CLEARING TECHNIQUES

1. The problem

Current techniques for observing atmospheric vertical profiles of temperature and humidity from satellites are based on measurements of upwelling radiation in the infra-red and microwave spectral regions. These radiances are affected, to a greater or lesser degree, by the presence of cloud. At infra-red wavelengths the problem is acute since many types of cloud are opaque; in the microwave region clouds have a much smaller effect on the radiances, but often it is too large to be neglected. Consequently, when retrieving information on other atmospheric variables, we must be able to detect clouds which have a significant effect on the radiances and, if possible, make allowance for these effects.

In this lecture, we shall review methods which have been developed to handle the effects of cloud in the retrieval of atmospheric profiles, with particular attention to methods for "cloud-clearing", i.e. for determining the radiances which would have been measured from the same atmospheric temperature and humidity profile if no cloud had been present. Also we shall concentrate on methods developed or applicable to the TOVS (TIROS-N Operational Vertical Sounder) instruments on the TIROS-N/NOAA satellite series. A recent and more detailed review of these methods is presented by Eyre and Watts (1987).

2. Possible solutions

The first approach to the problem is technological, i.e. to sound the atmosphere at frequencies for which clouds are more transparent, such as in the microwave region. There is a move towards increasing emphasis on microwave radiometers as the technology in this field advances. However, the infra-red has a number of advantages – narrower weighting functions in the troposphere, less variable surface emissivity and higher spatial resolution. For these reasons, a combination of infra-red and microwave instruments is likely to be preferred for the foreseeable future. Also, although cloud effects in the microwave are much reduced, they are often significant; their treatment is discussed further in section 7.

Other approaches to the problem take the form of different methods for extracting information from potentially cloud-affected infra-red data, and they can be categorised as follows:

a) Cloud detection (i.e. tests to determine which measurements are affected by the presence of cloud) followed by retrieval from those data found to be "cloud-free". This approach leads to the most accurate retrievals. Unfortunately, most of the interesting weather is associated with cloudy areas, and so this is only a partial solution.

b) Cloud detection followed by a cloud clearing method in which those measured radiances which are judged to be affected by cloud are "corrected" to give clear-column radiances. These methods can be useful in partly cloudy areas.

c) Cloud detection followed by a cloud clearing method in which cloud-affected, measured radiances are not used but clear-column radiances are estimated using other data (e.g. microwave measurements and/or ancillary information).

d) Spatial filtering of the radiance field, often including some combination of methods (a) to (c).

e) Inversion directly from the cloudy radiances, by estimating parameters describing the cloud conditions simultaneously with atmospheric profile information, thus avoiding an explicit cloud-clearing process.

These approaches are now considered in more detail, except for (e) which will be discussed in Lecture 3.

3. Cloud detection

The purpose of cloud detection algorithms within a satellite sounding data system is to distinguish between cloud-free fields of view (fovs), where the measured radiances may be used directly in the inversion, and cloud-affected fovs, which require some cloud-clearing procedure before the radiances can be used. There is no single method which is efficient in detecting all types of clouds, and it is usual to use a combination of several methods. McMillin and Dean (1982) describe in detail a series of tests, some of which are now summarised:

a) Window channel too cold. This is a general indication of the presence of cloud, but can be particularly useful over the sea where the surface temperature is known a priori to within a few degrees. Also, unfrozen seas (identified by low temperatures in microwave window channels) will not give rise to brightness temperature in infra-red window channels of less than 270 K unless cloud is present.

b) Adjacent fov tests. Within a group of fovs, those which are colder than their neighbours by more than some threshold are probably cloud affected.

c) Window channel tests. Comparisons between short and long wave window channels are useful. In the daytime, cloud-affected shortwave channels (~3.7μm) will often be warmer than longwave channels (~11μm) due to reflection of solar radiation. At night, the lower cloud emissivity at short wave gives rise to the opposite effect.

d) Visible channel tests. A useful addition in the daytime, particularly over the sea, where clouds give rise to increased reflectance.

e) Inter-channel regression tests. These are useful for sounding systems with sets of weighting functions in different spectral regions. TOVS, for example, has channels for sounding tropospheric temperature in the microwave (50–60 GHz) and around 15μm and 4.3μm in the infra-red. It is possible to establish regression relations by which the radiance (or brightness temperature) expected in one region under cloud-free conditions is predicted from measured radiances in another band. Many TOVS schemes use the difference between the measured brightness temperature in MSU channel 2 (53.73 GHz, with a weighting function peaking at 700 mb) and that predicted from a group of HIRS channels as a check for cloud. McMillin and Dean (1982) also give some other inter-channel tests, including checks between longwave and shortwave HIRS channels. Most schemes use at least some of these checks.

Another class of cloud detection method makes use of simultaneously measured image data such as that available from AVHRR. The high resolution of these data make it possible to characterise the cloud field within each fov of the infra-red sounding instrument. Various levels of sophistication are possible, but simple tests to determine whether the sounder data are cloud-affected or not are potentially very valuable. For this purpose, an AVHRR cloud-detection scheme such as that described by Saunders (1986) and Saunders and Kriebel (1988) may be used.

4. Clear column radiances by correction of cloud-affected radiances

A number of schemes have been developed for estimating clear column radiances from cloud-affected radiances in adjacent fovs. The most widely used are based on the N^* method proposed by Smith (1968). The measured radiances, R_1 and R_2, in two adjacent fovs of one-channel of an infra-red radiometer can, under certain conditions, be expressed as follows:

$$R_1 = (1 - N_1)R_c + N_1 R_o$$
$$R_2 = (1 - N_2)R_c + N_2 R_o \qquad (1)$$

where R_c and R_o are the radiances appropriate to clear and completely overcast conditions respectively, and N_1 and N_2 are the effective fractional cloud coverages in fovs 1 and 2. In deriving these equations the following assumptions have been made: that the atmospheric profiles and surface characteristics in the two fovs are the same; that only one layer of cloud is present; and that the cloud top has the same height (and temperature) in both fovs. These assumptions are necessary if R_c and R_o are to be taken as constant between fovs 1 and 2. If the fractional coverages in the two fovs are different ($N_1 \neq N_2$), then equations (1) can be solved simultaneously to give the clear radiance:

$$R_c = (R_1 - N^* R_2)/(1 - N^*) \qquad (2)$$

where $N^* = N_1/N_2$. Alternatively,

$$N^* = (R_c - R_1)/(R_c - R_2) \qquad (3)$$

and so N^* can be found if we have an estimate of the clear radiance in one channel. (Some methods for deriving the clear radiance in one channel are discussed below.) Then, since N^* is equal in all channels of the radiometer, it can be used in equation (2) to find the clear radiance in other channels. For infra-red radiometers with moderately high horizontal resolution (e.g. HIRS, which has a fov spacing of about 40 km), the inherent assumptions are true sufficiently often for the method to be useful. The method is illustrated graphically in Fig. 1.

McMillin *et al.* (1973) described an application of this method for the Vertical Temperature Profile Radiometer (VTPR) on the early members of the NOAA series of satellites. The clear-column radiance in one channel required by equation (3) was obtained using data from the Scanning Radiometer (a 2-channel instrument of higher horizontal resolution) on the same satellites.

Chahine (1970) gave an analysis of the cloud-clearing problem similar to that given by Smith (1968) and later developed an alternative adjacent fov method (Chahine 1974). Equation (2) can be expressed:

$$R_c = R_1 + \eta(R_1 - R_2). \qquad (4)$$

Here

$$\eta = N_1/(N_2 - N_1) \qquad (5)$$

or

$$\eta = \{R_c(\nu') - R_1(\nu')\}/\{R_1(\nu') - R_2(\nu')\} \qquad (6)$$

where ν' is the frequency of a specially selected "cloud-sounding" channel. In an iterative approach, a first-guess profile is used with a radiative transfer model to generate $R_c(\nu')$. Equation (6) then gives η which is used in equation (4) to calculate $R_c(\nu_j)$ for other channels at frequencies ν_j. Chahine shows that the method is stable in certain cases, for example when ν' refers to a channel in the 15μm carbon dioxide band, with a weighting function peaking in the lower troposphere, and ν_j represent a set of channels in the 4.3μm carbon dioxide band. This approach is generalised by Chahine (1977) to multiple cloud layers using a group of up to four adjacent fovs. Susskind *et al.* (1984) present details of a scheme for applying the single cloud layer method (Chahine 1974) to HIRS data. HIRS/2 channel 7 at 13.4μm is used as the "cloud-sounding" channel and the 4.3μm band channels only (channels 13 to 17) are used for the temperature retrieval. In addition, when MSU data are available, a similar scheme is employed which solves for the temperature profile and cloud using only MSU and the 4.3μm band channels of HIRS.

McMillin (1978) presents another version of the adjacent fov approach which is closely related to the N* method. By writing equations (1) for two fovs and two frequencies (ν_a and ν_b), four equations are obtained which may be solved simultaneously to give:

$$R_c(\nu_a) = R_1(\nu_a) + S\{R_c(\nu_b) - R_1(\nu_b)\} \qquad (7)$$

where

$$S = \{R_1(\nu_a) - R_2(\nu_a)\}/\{R_1(\nu_b) - R_2(\nu_b)\}. \qquad (8)$$

Again these equations are valid under the same assumptions as those for which the N* method applies. McMillin shows that if several adjacent pairs are considered, those pairs for which the assumptions are valid will yield the same value of S; other pairs will tend to give different values. Only "good" values of S are then used in equation (7) to generate the clear radiance. The method still requires an estimate of the clear-column radiance in one channel.

Smith and Woolf (1976) develop a variation of the N* technique for use with HIRS/1 and SCAMS (SCAnning Microwave Spectrometer) data from Nimbus 6. In this method N* is obtained from measured radiances in all channels, infra-red and microwave, using eigenvectors of the covariance matrix of clear-column radiances (pre-calculated using a representative set of clear radiances). The microwave data play an important role here; they allow the cloud-clearing to proceed without the need for estimates from other sources of the clear radiance in one infra-red channel. From 1978 to 1980 this method formed the basis of the cloud-clearing scheme used with HIRS/2 and MSU data in the operational global retrieval system of NOAA/NESDIS (see Smith *et al.* 1979).

A simplified but widely-used version of this method was presented by Smith (1980). Here, the additional piece of information required in the N^* method, i.e. the clear radiance in one channel, is obtained from the measured MSU channel 2, which acts as though it were a linear combination of cloud-free HIRS channels:

$$\hat{T} = a_o + \sum_i a_i T_i^H \qquad (9)$$

where $\hat{T}$ is the estimated value of MSU channel 2 and T_i^H are the measured brightness temperature in a group of HIRS channels. Comparison of $\hat{T}$ with the measured MSU value, T^M, is used in the cloud detection (see section 3). If cloud is detected, then the values of $\hat{T}$ in two fovs are used to calculate N^* in an approximated form of equation (3):

$$N^* = \frac{T^M - \hat{T}^1}{T^M - \hat{T}^2}, \qquad (10)$$

where the fovs are arranged such that fov 1 is the warmer (i.e. $\hat{T}^1 > \hat{T}^2$). This also forms part of the method used by the U.K. Meteorological Office (see Lecture 2).

In 1980 the NESDIS operational system was changed to incorporate a new cloud-clearing algorithm described by McMillin and Dean (1982). This algorithm is again based on the N^* method but takes great care to allow for the fact that the assumption of equal cloud height in adjacent spots is often invalid. To detect those cases in which the N^* method is applicable a series of checks is made. Firstly a more thorough treatment is given to the detection of clear areas (see section 3). In areas found to be partly cloudy, the N^* method is applied with a series of tests to check its validity, including some based on the approach of McMillin (1978). Also, N^* is calculated in two ways: from HIRS and MSU radiances, and from HIRS longwave and shortwave radiances. The latter approach is an extension of the method developed by Chahine (1974), but it does not require radiative transfer calculations as part of the algorithm. The two values obtained are required to be consistent. Finally, using the best value of N^*, the clear column radiances are calculated and another set of checks based on inter-channel regression is performed.

5. Clear column radiances estimated from other data

If cloud is detected, it is possible to estimate the clear column radiances using other data. Again, simultaneously measured microwave data are in practice the best source of information.

The method developed by the U.K. Meteorological Office follows this route when the N^* method fails. HIRS clear column brightness temperatures in cloud-affected channels are estimated from a linear combination of MSU channels. Further details are given in lecture 2.

The ψ-method developed by LMD as part of the 3I retrieval technique (Chedin *et al.* 1985) takes an analogous route whenever a HIRS channel is judged to be cloud-affected. The HIRS channel to be cleared is paired with the MSU channel with the most similar weighting function. The clear column HIRS brightness temperature is calculated using the measured MSU value and simulated HIRS and MSU brightness temperatures corresponding to an initial guess profile selected by a proximity search amongst a large library of profiles; for further details of the 3I technique, see other lectures in this series (Chedin 1989).

6. Spatial filtering methods

A few cloud clearing methods have been developed which include some aspect of spatial filtering, and they fall into two types. Firstly, there are those which apply a filter to the cloud-affected radiances as a method of cloud-clearing. Prata (1985) demonstrated the application of a Kalman filter for this purpose. Later lectures in this series (Serio 1989) also consider this problem.

Other methods have used spatial filtering to reduce the noise and to suppress rogue values in preliminary estimates of clear-column radiance. Fleming and Hill (1982) demonstrated a one-dimensional filter to remove isolated gross errors in clear-column radiances or, indeed, other geophysical data. They later extended the method to more than one dimension (Fleming and Hill 1983). Eyre and Watts (1987) developed a two-dimensional optimal estimation scheme for filtering preliminary estimates of clear-column radiances. This is described in Lecture 2.

7. Cloud problems in the microwave

Clouds attenuate microwave radiation much less than they do infra-red radiation. Microwave sounding techniques are able to penetrate most clouds and to obtain information from the atmosphere below the cloud top. However, although the effects are often small, they cannot be neglected completely. Indeed, because the radiances measured from space are affected by the full depth of the cloud (rather than just the top layer, as in the infra-red), the problem can be more complicated.

At MSU frequencies (50 60 GHz) non-precipitating cloud acts as an absorbing/emitting medium, but when precipitation-sized particles are present, scattering also becomes important. For many clouds the absorption is quite weak and leads to a perturbation on the measured brightness temperature which is almost linearly related to the total column cloud liquid water. Most TOVS processing schemes treat this as a pre-processing problem: the relatively large effect on the window channel (MSU channel 1) is used in a linear regression relation to calculate a correction for the much smaller effect on MSU channel 2. Effects on channels peaking higher are negligible. For very large amounts of cloud liquid water and/or when precipitation is present, the assumptions made with this simple, linear correction are no longer valid. Current practice is to detect such data and to avoid its use in the temperature retrieval. An algorithm to do this over the ocean was suggested by Phillips (1981); MSU data are rejected when

$$T^{MSU-2} - T^{MSU-1} < 12K \qquad (11)$$

Recent work at LMD has suggested that a larger value for this threshold (16K) may give better result.

Cloud problems with MSU data are relatively small because of the low frequency used and the large footprint of the MSU antenna ($\geq$ 110 km) much larger than most rain bands and their associated perturbations on the brightness temperature field. However, in future this will not be the case; the smaller footprints and higher frequencies range of the Advanced Microwave Sounding Unit (AMSU) to be flown on the NOAA-KLM series (see Pick 1986) will give rise to more pronounced cloud effects on the microwave data. Research is still required on methods to tackle these problems, both in terms of their effects on the retrieval of atmospheric profiles and the potential for retrieving new cloud parameters.

REFERENCES

Chahine M.T., 1970. Inverse problems in radiative transfer: a determination of atmospheric parameters. J. Atmos. Sci., 27, *960-967.*

Chahine M.T., 1974. Remote sounding of cloudy atmospheres. I: The single cloud layer. J. Atmos. Sci., 31, *238-248.*

Chahine M.T., 1977. Remote sounding of cloudy atmospheres. II: Multiple cloud formations. J. Atmos. Sci., 34, *744-757.*

Chedin A., 1989. Lectures at ICTP Workshop "Atmospheric Radiation and Cloud Physics"; Trieste; 27 November-15 December 1989.

Chedin A., Scott N.A., Wahiche C. and Moulinier P., 1985. The Improved Initialisation Inversion method: a high resolution physical method for temperature retrievals for the satellites of the TIROS-N series. J. Clim. Appl. Meteor., 24, *128-143.*

Eyre J.R. and Watts P.D., 1987. A sequential estimation approach to cloud-clearing for satellite temperature sounding. Q.J.R. Meteorol. Soc., 113, *1349-1376.*

Fleming H.E. and Hill M.L., 1982. An objective procedure for detecting and correcting errors in geophysical data: 1. One-dimensional applications. J. Geophys. Res., 87, *7312-7324.*

Fleming H.E. and Hill M.L., 1983. An objective procedure for detecting and correcting errors in geophysical data: 2. Multi-dimensional applications. J. Geophys. Res., 88, *3859-3867.*

McMillin L.M., 1978. An improved technique for obtaining clear-column radiances from cloud-contaminated radiances. Mon. Weather Rev., 106, *1590-1597.*

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

McMillin L.M., Wark D.Q., Siomkajlo J.M., Abel P.G., Werbowetzki A., Lauritson L.A., Pritchard J.A., Crosby D.S., Woolf H.M., Luebbe R.C., Weinreb M.P., Fleming H.E., Bittner F.E. and Hayden C.M., 1973. Satellite infra-red sounding from NOAA spacecraft. NOAA Tech. Rep. NESS 65.

Phillips N.A., 1981. Cloudy winter satellite temperature retrievals over the extratropical northern hemisphere oceans. Mon. Weather Rev., 109, *652-659.*

Pick D.R., 1986. The operational sounding of the lower atmosphere from satellites using millimetre waves. Proc. 16th European Microwave Conf.; Dublin; 8-12 Sept 1986; *pp. 31-41.* Microwave Exhib. Publ. Ltd. (Tunbridge Wells).

Prata A.J., 1985. Clear column radiances by optimal estimation. Tech. Proc. 2nd Int. TOVS Study Conf.; 18-22 February 1985; Igls, Austria; Ed.: W.P. Menzel; Report of CIMSS, University of Wisconsin, Madison; *pp. 201-228.*

Saunders R.W., 1986. Removal of cloud contamination from AVHRR radiances. Int. J. Remote Sensing, 7, *867-886.*

Saunders R.W. and Kriebel K.T., 1988. An improved method for detecting clear sky and cloudy radiances from AVHRR data. Int. J. Remote Sensing, 9, *123-150.*

Serio C., 1989. Lectures at ICTP Workshop "Atmospheric Radiation and Cloud Physics"; Trieste; 27 November-15 December 1989.

Smith W.L., 1968. An improved method for calculating tropospheric temperature and moisture from satellite radiometer measurements. Mon. Weather Rev., 96, *387-396.*

Smith W.L., 1980. Operational sounding algorithms. In Proc. of VAS Demonstration Sounding Workshop; NASA/GSFC; Greenbelt, Md; 15 July 1980; NASA Conf. Publ. 2157, Eds. D.L. Endres and L.W. Uccellini.

Smith W.L. and Woolf H.M., 1976. The use of eigenvectors of statistical covariance matrices for interpreting satellite sounding radiometer observations. J. Atmos. Sci., 33, *1127-1140.*

Smith W.L., Woolf H.M., Hayden C.M., Wark D.Q. and McMillin L.M., 1979. The TIROS-N operational vertical sounder. Bull. Am. Meteorol. Soc., 60, *1177-1187.*

Susskind J., Rosenfeld J., Reuter D.C. and Chahine M.T., 1984. Remote sensing of weather and climate parameters from HIRS2/MSU sounding data. J. Geophys. Res., 89, *4677-4697.*

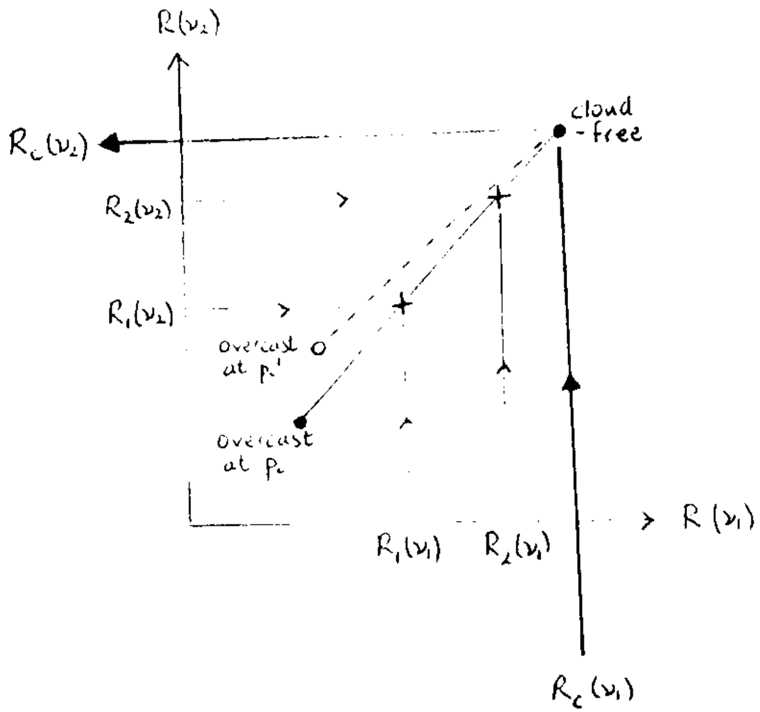

Figure 1. Illustrating the N* method

Measurements in 2 channels, ν_1 and ν_2, at 2 fovs, R_1 and R_2, give 2 points defining the line in which the cloud-free value lies. A cloud free estimate for one channel, $R_c(\nu_1)$, can then be used to find the cloud-free value in the other channel, $R_c(\nu_2)$. This assumes that cloud in both fovs are at the same pressure, p_c.

ICTP Workshop on "Atmospheric Radiation and Cloud Physics".
27 November – 15 December 1989.

J.R. Eyre *Lecture 2*

THE U.K. METEOROLOGICAL OFFICE'S TOVS PROCESSING SYSTEM AND ITS CLOUD-CLEARING SCHEME

1. Introduction

The U.K. Meteorological Office has developed an operational system for processing locally-received TOVS data for use in numerical weather prediction (NWP). The scheme is known as the Local Area Sounding System (LASS). It was developed from Version 1 of the International TOVS Processing Package (ITPP) obtained from the University of Wisconsin-Madison in 1980, although major changes have been made to the system since then.

In this lecture, LASS will be described briefly and then its cloud-clearing scheme presented in detail.

2. LASS

LASS has been running routinely in real-time since 1983 (see Turner *et al.* 1985). Charts of the 1000–500 mb thickness field derived from LASS temperature and humidity retrievals have been used by operational forecasters as a guide to their subjective analyses and forecast model intervention. More recently the data have been assimilated into the regional NWP model.

The TOVS data are processed through the following stages:

a) Ingest

Raw TIP (TIROS Information Processor) data for the European/N. Atlantic region are received directly from the satellite at the reception station at Lasham (Hampshire, U.K.). They are relayed in real-time to the U.K. Met. Office headquarters at Bracknell, where they are captured on a VAX 11/750 computer. HIRS and MSU raw data (in counts) are extracted from the TIP data stream. Each measurement is then calibrated to give a brightness temperature (equivalent black body temperature) and the set of measurements at each HIRS or MSU sounding location is assigned a latitude and longitude. All these operations are performed using ITPP software with very little modification.

b) Pre-processing

HIRS brightness temperatures are "limb-corrected", i.e. the values measured at angles off nadir are adjusted to give an estimate of the value which would have been measured by a sounding at nadir through the same atmospheric column. This is performed with a linear regression correction using measured values in other HIRS channels. MSU brightness temperatures are pre-processed in a similar way; here the regression correction accounts not only for the "limb" effect but also for the effects of variable surface emissivity, the MSU antenna response pattern and the effects of cloud liquid water. The pre-processed MSU value is thus intended to correspond to a sounding at nadir with an narrow footprint, a black Earth's surface and no cloud liquid water. Following this, MSU brightness temperatures are interpolated to the locations of the HIRS fields-of-view (fovs). Again, ITPP software is used with little modification. [Recent work has demonstrated that it would be advantageous to add a check for precipitation contamination, as described in Lecture 1, prior to interpolation.]

c) Cloud-clearing

This stage of the processing accepts pre-processed HIRS and MSU brightness temperatures at each HIRS fov and outputs for each channel, a field of clear-column brightness temperatures specified at every fov, together with a field of their expected errors. The method is described in section 3 and in more detail by Eyre and Watts (1987). It has been used operationally in LASS since 1987.

d) Inversion

Since 1987, this has been performed by linearized optimal estimation (see Rodgers 1976) using a NWP forecast profile and its expected error covariance as prior information. This can be considered as a one-dimensional, linear approximation to the direct assimilation of clear-column radiances into the NWP model. Lecture 3 explores this aspect in more detail. HIRS and MSU clear-column brightness temperatures are converted into temperature and humidity profiles. At present, inversions are performed for only one in every 3×3 HIRS fovs, a resolution of about 120 km, but the scheme is flexible in this respect.

e) Post-processing

Retrieval temperature and humidity profiles are converted to related layer quantities: thicknesses and layer water vapour contents. Also thermal winds are calculated from the local gradients of the thickness fields.

Figure 1a gives an example of a LASS product. It shows the retrieved 1000-500 mb thickness field derived from three consecutive overpasses of TOVS data. (For clarity, some of the retrievals made are not shown, but all have been used in analysing the contour lines.) Figure 1b shows, for the same occasion, the difference between the retrieval and forecast thickness fields. This is one way of representing the "forcing" of the TOVS product on the forecast background field in the NWP data assimilation. It can be used to explore both the information content of the TOVS data and any remaining quality control problems.

3. Cloud-clearing in LASS

The scheme consist of three parts:

- an algorithm to obtain preliminary estimates of clear-column brightness temperatures (and their expected errors), followed by
- a sequential filter, based on the principles of optimal estimation, to improve the values in the preliminary field
- a procedure to compensate for locally coherent biases between measured HIRS data and those estimated from MSU data.

3.1 Preliminary estimates of clear-column brightness temperatures

The clear-column HIRS brightness temperature in each channel is estimated, along with its expected error, at every HIRS spot by one of four routes:

(a) Clear conditions

Clear spots are identified by a cloud detection algorithm using a sub-set of the tests described. A linear combination of HIRS brightness temperatures is used to predict an MSU

channel 2 brightness temperature (hereafter abbreviated "MSU(2)" and similarly for other MSU and HIRS channels), using a regression relation which has a residual error of about 0.8 K for clear column brightness temperatures. (See Lecture 1, eq. 9.) The coefficients for the regression are calculated from theoretically computed brightness temperatures for a representative set of historical radiosonde profiles of the appropriate month for the European/N. Atlantic area (Watts 1984). If predicted MSU(2) is colder than measured MSU(2) by more than the residual error in the regression, then the HIRS measurements are assumed to be contaminated by cloud. Also a test is made for excessive contamination by solar radiation: if HIRS(19)-HIRS(18) is greater than 10 K or HIRS(18)–HIRS(8) is greater than 20 K then the brightness temperatures in all HIRS channels are rejected. At night, an additional check on the brightness temperature difference between channels in the 11 and 3.7 μm window regions can be used. This employs the signature of low-level water cloud through its lower emissivity in the shorter wavelength region (Eyre *et al.* 1984). If HIRS(8)–HIRS(18) exceeds 2 K then the radiances are flagged as cloud contaminated. In addition dynamic threshold tests are applied to HIRS-8 and the visible channel (HIRS-20). If a fov is colder/brighter than the warmest/darkness fov in the surrounding 7×7 box of fovs by more than a fixed threshold, it is flagged as cloud-affected.

If the data pass these tests, they are assumed to be clear and the measured (pre-processed) HIRS brightness temperatures are taken as the preliminary estimates of the clear-column values. In this case the estimated errors are comparatively low. However they are not simply the instrument noise, since our ability to determine that the measurements are truly cloud-free is limited by the sensitivity of the cloud detection process. Also, pre-processing errors for both HIRS and MSU contribute to the uncertainty. Radiometric noise, limb correction error and cloud detection sensitivity are all taken into account in the error estimation scheme. Full details of this procedure, and equivalent procedures for other cloud clearing routes, are given by Watts (1985).

(b) Partly cloudy conditions – N^* method

If cloud is detected, the clear column radiance determination is attempted by the N^* method (see Lecture 1) using the HIRS spot in question with one of its 8 adjacent spots. The most suitable partner is first identified. The partner must be colder, to ensure that the same pair of spots is not used again when the partner is itself the central spot. Also the warmer spot is the more appropriate position to assign to the clear radiance as it contains a greater proportion of cloud-free area. If more than one colder partner is found then the one which gives the best agreement between N^* values calculated separately for HIRS longwave and shortwave channels is chosen. Moreover, values of N^* greater than 0.75 and longwave-shortwave differences in N^* greater than 0.1 are not used. The N^* algorithm fails if a suitable partner cannot be found. It may also fail if the clear-column brightness temperatures do not pass certain checks. Excess contamination by reflected solar radiation is tested as described above. Also HIRS(8) is compared with a corresponding value predicted from MSU data (as explained below) and the radiances rejected if the absolute difference is excessive.

The errors in the clear-column brightness temperatures derived by the N^* method depend on the errors in the values of R_1, R_2 and N^* used in equation (2) of Lecture 1 and on the way in which this equation amplifies these errors. R_1 and R_2 contain error contributions from radiometric noise and pre-processing, both of which may be estimated. The error in N^* is determined by the errors of the terms used in equation (3) of Lecture 1, including the residual error in the prediction of MSU(2) from HIRS values. The error analysis is complicated by the correlations between the components of error in these equations, and it is difficult to assess absolute error levels accurately. However it can be seen that the error amplification through the N^* equation is very variable and becomes acute as N^* approaches unity. For this reason, the method is deemed to fail if N^* is greater than 0.75. Additional errors arise when the conditions required for the N^* algorithm, concerning the profiles and cloud cover in adjacent spots, are not satisfied. No attempt has been made to quantify these errors; they will tend to have statistical properties which are difficult to accommodate in this approach. However the quality checks described above are used to minimise the problems arising from this source.

(c) Cloudy conditions – MSU regression method

If cloud is detected and the N^* method fails, the clear-column HIRS brightness temperatures are estimated from the MSU data using a regression relation:

$$\hat{T}_j^H = b_{oj} + \sum_i b_{ij}\, T_i^M \quad (1)$$

where $\hat{T}_j^H$ is the estimated HIRS brightness temperature in channel j (j = 4 to 16), T_i^M is the measured MSU brightness temperature in channel i (i = 2 to 4) and coefficients b_{ij} are obtained by regression on brightness temperatures calculated from a large set of representative radiosonde profiles. This approach is adopted principally to obtain an estimate on every HIRS spot for the purposes of the sequential estimation which follows. Retrievals from preliminary estimates of HIRS brightness temperatures obtained in this manner would be approximately equivalent to MSU-only retrievals. However, in combination with the "HIRS–MSU bias" correction and sequential estimation schemes described below, information on the vertical profile structure can be "advected in" from nearby HIRS spots from which HIRS information is available, leading to a more effective and more nearly optimal method for using MSU information in cloudy areas.

Errors in the HIRS estimates obtained by this method arise from radiometric and preprocessing errors in the MSU data and from the residual errors in the regression. These can be assessed and give the values used in the subsequent filtering. However, the sequential estimator assumes that errors in adjacent spots are uncorrelated. This is not the case for HIRS values estimated from MSU because the errors in the regression are locally consistent – they are caused mainly by vertical profile structure which is resolved by HIRS but not by MSU. Also, the horizontal resolution of MSU is less than that of HIRS, and so one MSU spot dominates the MSU values interpolated to several HIRS spots. These problems and their solution are discussed further in section 3.3.

(d) MSU data missing

On the infrequent occasions when no MSU values have been mapped to a HIRS spot (thus precluding both the N^* method and the MSU regression method), MSU and HIRS values for a nearby HIRS spot are used and their expected errors are suitably increased. Failing this, climatological mean values are the last resort with the climatological variance supplying the expected error. Again, the reason for adopting this approach is for convenience in the subsequent filter, which requires a HIRS brightness temperature estimate and an expected error at every spot.

Figures 2 and 3 illustrate the results of the preliminary cloud-clearing for a typical pass of data. Figure 2 is an AVHRR image showing qualitatively the location and structure of the cloud fields. Super-imposed on this image are the locations of HIRS fields-of-view. Figure 3

is a character map indicating the outcome of the preliminary cloud-clearing procedure at each spot for this pass.

3.2 Sequential estimation filter

The sequential estimation filter is an approximation to an optimal method for improving the preliminary estimates of clear-column brightness temperature by combining them with estimates obtained from additional information and weighting them according to their expected errors. Thus for one radiometer channel i, if we have i independent "observations" of clear-column brightness temperature x_i, with errors of Gaussian distribution and variance σ_i^2, we can combine the observations to give the best estimate of the clear-column brightness temperature

$$\hat{T} = \left(\sum_i \frac{1}{\sigma_i^2}\right)^{-1} \sum_i \frac{T_i}{\sigma_i^2} \quad (2)$$

with variance

$$\hat{\sigma}^2 = \left(\sum_i \frac{1}{\sigma_i^2}\right)^{-1} \quad (3)$$

The additional information could in principle come from a number of sources (e.g. AVHRR data). In this scheme it comes from estimates of clear-column brightness temperatures in nearby fovs with their errors suitably increased to allow for their horizontal separation (i.e. for the fact that they are really estimates for different locations).

A simple and efficient way of approximating this approach is through a two-dimensional sequential estimation filter, as illustrated in Figure 4. The filter runs along the scan lines in alternate directions as shown. At spot (m,n) we combine the initial estimate of brightness temperature $T_{m,n}(-)$, and its expected error variance $\sigma^2_{m,n}(-)$, with improved estimates from the preceding spot on the same line, $T_{m,n-1}(+)_F$, and the equivalent spot on the preceding line, $T_{m-1,n}(+)_F$, and we obtain an improved estimate at (m,n), $T_{m,n}(+)_F$, and its expected error variance, $\sigma^2_{m,n}(+)_F$:

$$\frac{T_{m,n}(+)_F}{\sigma^2_{m,n}(+)_F} = \frac{\mu\, T_{m,n}(-)}{\sigma^2_{m,n}(-)} + \frac{\lambda\, T_{m,n-1}(+)_F}{\sigma^2_{m,n-1}(+)_F + \beta^2} + \frac{\lambda\, T_{m-1,n}(+)_F}{\sigma^2_{m-1,n}(+)_F + \beta^2} \quad (4)$$

$$\frac{1}{\sigma^2_{m,n}(+)_F} = \frac{\mu}{\sigma^2_{m,n}(-)} + \frac{\lambda}{\sigma^2_{m,n-1}(+)_F + \beta^2} + \frac{\lambda}{\sigma^2_{m-1,n}(+)_F + \beta^2} \quad (5)$$

$(-)$ and $(+)$ indicate "before" and "after" filtering respectively. Parameters μ, λ and β^2 determine the weights given to the data and are discussed further below. Subscript F indicates that this is the "forward" filtering process. A similar process is run along the pass in the opposite direction to give "backward" estimates: $T_{m,n}(+)_B$ and $\sigma^2_{m,n}(+)_B$. The forward and backward estimates are then combined to give the final estimate and its expected error:

$$\frac{T_{m,n}(+)}{\sigma^2_{m,n}(+)} = \frac{T_{m,n}(+)_F}{\sigma^2_{m,n}(+)_F} + \frac{T_{m,n}(+)_B}{\sigma^2_{m,n}(+)_B} \quad (6)$$

$$\frac{1}{\sigma^2_{m,n}(+)} = \frac{1}{\sigma^2_{m,n}(+)_F} + \frac{1}{\sigma^2_{m,n}(+)_B} \quad (7)$$

If care is taken to ensure that the forward and backward filters process a given line from opposite ends, then the combined filter advects information almost uniformly from all directions.

The choice of a sequential estimation approach, rather than some other method of objective analysis, was influenced by the availability of data on a regular grid. The technique described here is not a truly "optimal" method: only data from adjacent spots are used directly, and information from other nearby spots has influence only indirectly through the sequential nature of the method. In common with other objective analysis methods, it is necessary to consider the trade-off between this loss of optimality and any advantages of the technique. The main advantage is the computational speed through the use of only adjacent spots. To use all the available information "optimally" would require knowledge of appropriate error covariances between all pairs of spots, and these cannot be assessed accurately for this problem. They are however effectively included, albeit rather crudely, through the filtering.

The filtering parameters μ, λ and β^2 determine the weights given to different pieces of information by the sequential estimator. They are discussed in detail by Eyre and Watts (1987), but their basic roles are as follows:

μ: It can be shown theoretically that μ=0.5 for consistent results. It allows for the fact that ... m, n) is used twice in estimating the improved value at its own location.

λ: This is included principally to compensate for the overuse of data from neighbouring fovs. A normal value of λ=0.4 has been shown empirically to give good results. It is also used to prevent filtering occurring across expected steps in the brightness temperature field, i.e. at coastlines for window channels. Here λ is reduced in proportion to the fraction of the measured radiance which comes from the surface.

β^2: This is included to represent the increase in error variance for one fov when it is used as an estimate for its neighbour. It is related to the expected horizontal variations in the cloud-free brightness temperature field. It is highest for window channels and lowest for stratospheric channels.

3.3 The HIRS-MSU bias problem

HIRS values estimated from MSU data contain errors which are highly correlated locally. Straightforward use of the sequential estimator on a field containing these data produces unsatisfactory results; a bias remains in the final field and the error is under-estimated. It is desirable, therefore, to compensate for these biases if possible. At spots where HIRS brightness temperatures are obtained by the clear or N^* route, we can also obtain values by the MSU regression route and hence derive estimates of the bias between the two. Because the bias is locally coherent, particularly within areas of the same air mass, these sparse estimates of "HIRS-MSU bias" (i.e. the bias between measured HIRS brightness temperatures and those estimated from MSU) can be filtered using a similar sequential estimator to obtain values for the bias field at all spots. The filtered values of bias are then used to adjust brightness temperatures predicted from MSU and to make them consistent with neighbouring values obtained by the clear or N^* route. The effect of this procedure is not only to produce more horizontally consistent fields but also to advect information on vertical structure from clear and partly cloudy spots into surrounding areas where only preliminary estimates from MSU are available.

Fig. 5 illustrates the effect of the filtering for a single scan line of HIRS data. It shows the preliminary estimates and the final estimates, all with their expected errors. The

estimated HIRS-MSU biases for the same line are shown in Fig. 6. The sparse preliminary bias estimates are used to obtain the final filtered bias estimates at all spots, which are then applied as corrections to HIRS values estimated from MSU and are included in the final estimates shown in Fig. 5. The removal of local biases between HIRS estimates from MSU and by other routes can be seen. In both Fig. 5 and Fig. 6 it should be remembered that the filtering is a two-dimensional procedure and also uses information from neighbouring scan lines (not shown).

4. Validation using AVHRR data

Assessing the effectiveness of a cloud-clearing scheme is not a simple matter because independent measurements of "true" clear-column radiances are usually not available. One approach to this problem – the one which was used in validating the cloud-clearing scheme described above – is to use coincident AVHRR data (see Watts and Eyre 1986). AVHRR has a resolution of 1.1 km and 5 spectral channels in visible and infra-red window regions. Its high spatial resolution allows cloud to be detected with much greater accuracy than is possible with TOVS. The nominal HIRS fov can be co-located with AVHRR to within about 1 km (Aoki 1984; Lloyd *et al.* 1985). This allows the AVHRR pixels associated with each HIRS fov to be determined.

At a quantitative level, it is very instructive to overlay the ellipses corresponding to the HIRS fovs on the corresponding AVHRR image, as illustrated in Fig. 2. If information on the cloud detection and clearing tests is added, inspection of the image yields useful diagnostics on problems which the scheme. This approach was very helpful during the development of this TOVS cloud-clearing scheme.

For quantitative validation, it is possible to compare the cloud-cleared brightness temperature in a HIRS window channel (channel 8) with the average brightness temperatures in cloud-free AVHRR pixels lying within the HIRS ellipse for AVHRR channels 4 and 5, making small corrections for the differences in their spectral responses. Various techniques are available for detecting cloud contamination of AVHRR radiances including radiance threshold, spatial coherence and inter-channel tests. We have used the scheme described by Saunders (1986) which generates a cloud mask at pixel resolution. In this way it is possible to test statistically the effects of changes to the cloud-clearing scheme on the quality of clear-column radiances in HIRS channel 8. Since this is a channel most sensitive to the effects of cloud, it is reasonable to assume that it is a good measure of the performance of the cloud-clearing scheme as a whole (see Eyre and Watts 1987).

39-89

REFERENCES

Aoki T., 1984. Clear radiance retrieval of HIRS channels with the use of AVHRR data. Tech. Proc. 1st Int. TOVS Study Conf.; 29 August-2 September 1983; Igls, Austria; Report from CIMSS, University of Wisconsin-Madison, Ed.: W.P. Menzel; *pp. 1-9.*

Eyre J.R., Brownscombe J.L. and Allam R.J., 1984. Detection of fog at night using AVHRR imagery. Meteorol. Mag., 113, *266-271.*

Eyre J.R. and Watts P.D., 1987. A sequential estimation approach to cloud-clearing for satellite temperature sounding. Q.J.R. Meteorol. Soc., 113, *1349-1376.*

Lloyd P.E., Barnett J.J. and Eyre J.R., 1985. Investigations of AVHRR data to improve TOVS retrievals. Tech. Proc. 2nd Int. TOVS Study Conf.; 18-22 February 1985; Igls, Austria; Report from CIMSS, University of Wisconsin-Madison, Ed.: W.P. Menzel; *pp. 162-176.*

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

Saunders R.W., 1986. Removal of cloud contamination from AVHRR radiances. Int. J. Remote Sensing, 7, *867-886.*

Turner J., Eyre J.R., Jerrett D. and McCallum E., 1985. The HERMES system. Meteorol. Mag., 114, *161-173.*

Watts P.D., 1984. A study of local area synthetic coefficients for use in the LASS system. Meteorological Office internal report. Met.O.19 Branch Memo 76.

Watts P.D., 1985. An improved TOVS cloud clearing scheme-preliminary report. Meteorological Office internal report. Met.O.19 Branch Memo 81.

Watts P.D. and Eyre J.R., 1986. Validating a TOVS cloud-clearing scheme using AVHRR data. Tech. Proc. 3rd Int. TOVS Study Conf.; Madison, Wisconsin; 13-19 August 1986; Report of CIMSS, University of Wisconsin-Madison, Ed.: W.P. Menzel; *pp. 312-331.*

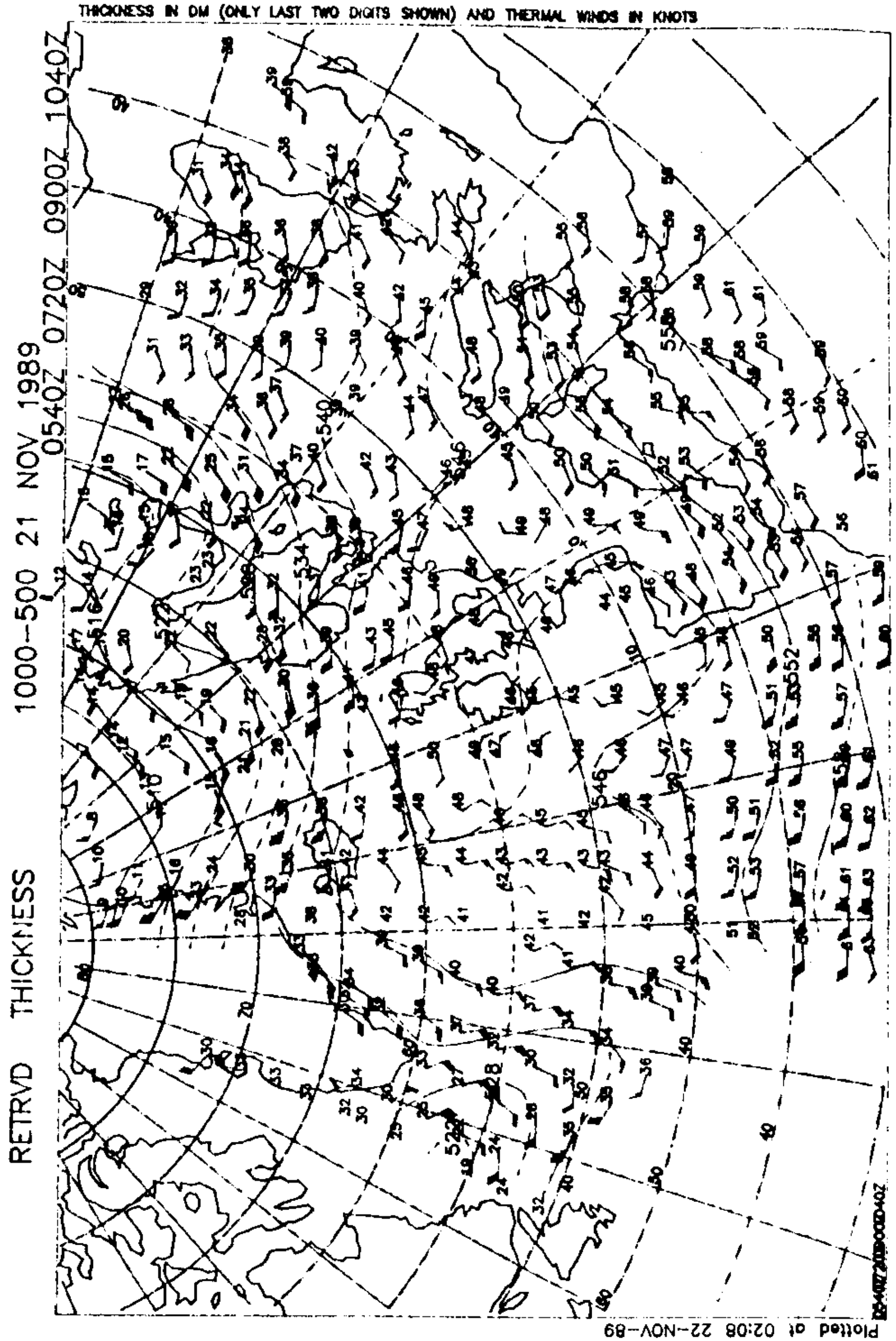

Figure 1(a). Example of LASS product. 1000-500 mb thickness field from 4 consecutive NOAA-10 orbits. Values in decametres (with leading "5" omitted for individual points).

Figure 1(b). As Fig. 1(a), except [illegible] fields. Values in decametres. Lines: solid = positive, dashed = negative, dotted = zero.

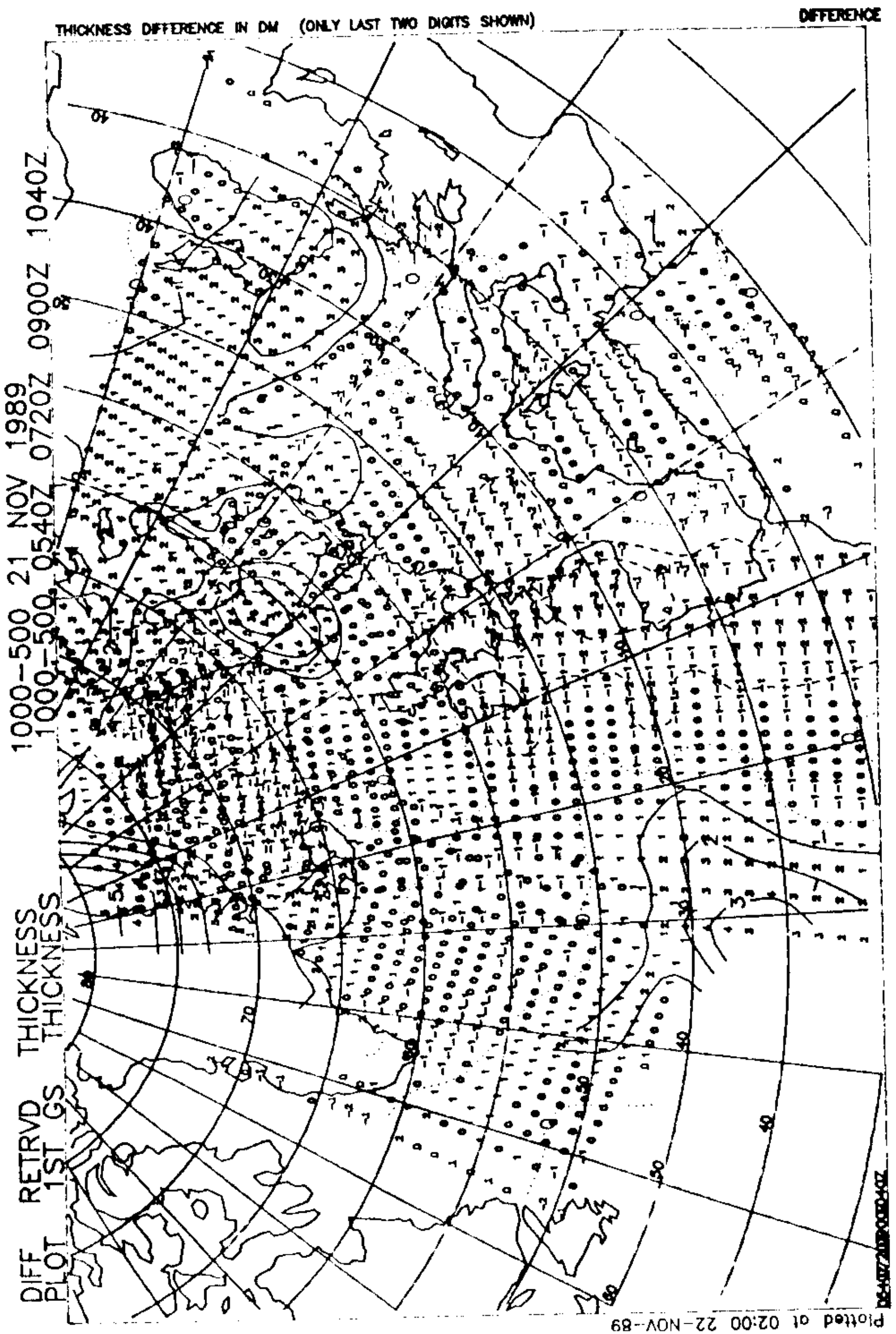

Figure 2 AVHRR channel 1 (visible) image and HIRS field-of-view locations. Part of overpass at 13 GMT on 16 April 1985

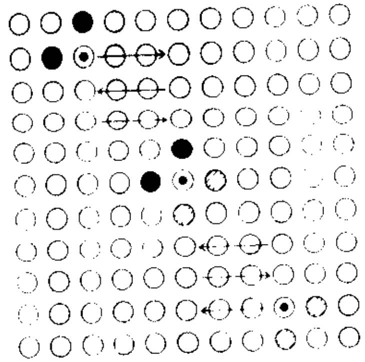

Figure 4 Illustrating the application of the sequential estimation scheme. The solid spots at the top indicate the HIRS spots used in the 'forward' estimation of the adjacent dotted spot. The arrows indicate the direction of progression of the sequential estimator. The corresponding pattern at the bottom is for the equivalent 'backward' filter. The spots marked in the centre illustrate the symmetric nature of the combined forward and backward filters.

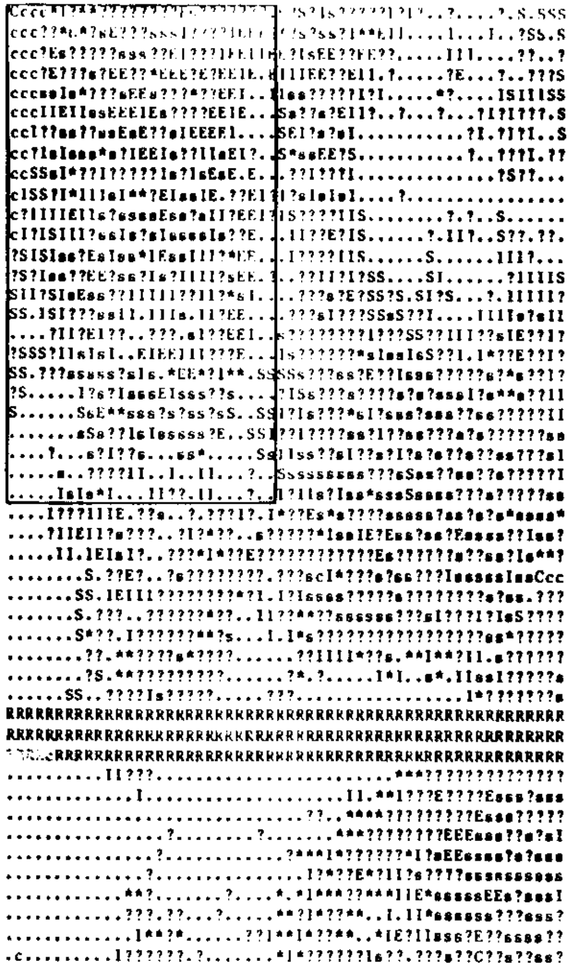

Figure 3. Summary of preliminary cloud clearing for overpass at 13 GMT on 16 April 1985. Box gives position of the image shown in Fig. 2. HIRS line numbers are given on the right. Preliminary cloud-clearing routes are symbolized as follows

- . Clear
- * N^* method
- R MSU regression: HIRS data missing (calibration line)
- S MSU regression: excessive solar reflection
- I MSU regression: long-wave and short wave N^*s inconsistent
- s MSU regression: excessive solar reflection after N^* route
- E MSU regression: rogue value of N^*
- ? MSU regression: no suitable pair for N^* method
- • MSU regression: low cloud test
- c No MSU value: search nearby spots
- C No MSU value: use climatological values

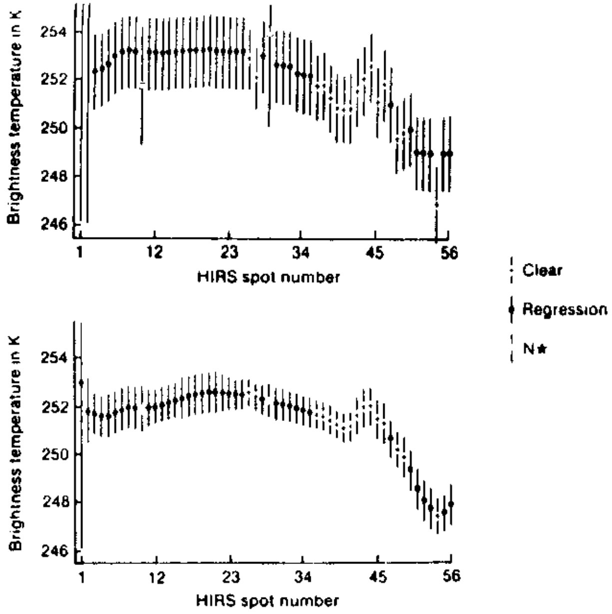

Figure 5. Cloud-clearing results for one scan line (line 8 in Fig. 3) for HIRS channel 6. Symbols indicate the preliminary cloud-clearing route. Upper and lower graphs show before and after filtering respectively. Values in K.

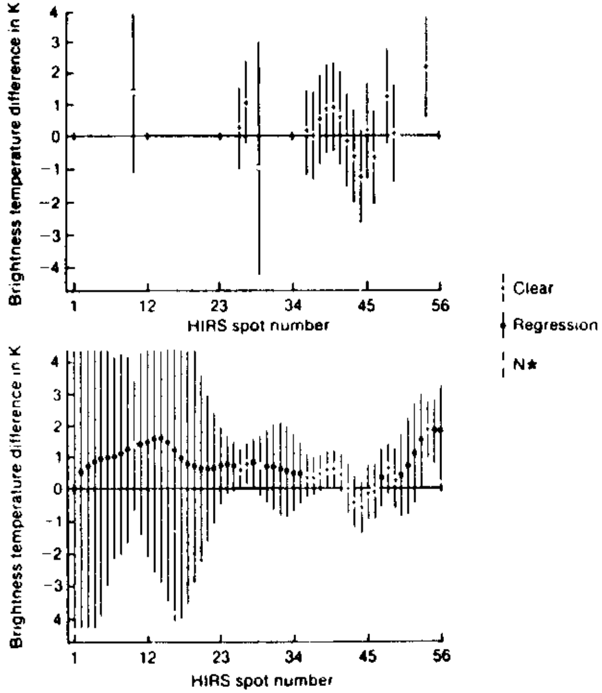

Figure 6. As Fig. 5 except for the HIRS − MSU bias

ICTP Workshop on "Atmospheric Radiation and Cloud Physics".
27 November – 15 December 1989.

J.R. Eyre

Lecture 3

DIRECT INVERSION OF CLOUD-AFFECTED RADIANCES

1. Introduction

Most satellite sounding retrieval schemes treat the cloud-clearing problem and the inversion problem separately; the measured radiances are first checked and, if necessary, adjusted for the effects of cloud, then the clear column radiances are inverted to give profiles of temperature and humidity. However, some schemes have now been developed which attempt to combine these two processes either partially or totally.

The TOVS retrieval scheme presented by Susskind *et al.* (1984), which is described briefly in Lecture 2, applies the cloud-clearing and inversion steps in an iterative loop, with considerable feedback between them, to retrieve a full range of atmospheric profile, surface and cloud parameters. Smith *et al.* (1985) developed a physical method for TOVS in which all atmospheric profile and surface parameters are retrieved simultaneously, and Huang and Smith (1986) extended this to include cloud-top pressure and cloud amount in the list of simultaneously retrieved variables. A simultaneous retrieval including cloud parameters has recently been combined with the principles of nonlinear optimal estimation to give a method for extracting information from potentially cloud-affected TOVS radiances for use in numerical weather prediction (Eyre 1989a). In this lecture, this method and its theoretical basis will be described.

2. The general inversion problem

The "forward" or radiative transfer problem for atmospheric sounding can in general be considered as that of calculating a vector of multi-channel radiances $\mathbf{y}\{\mathbf{x}\}$, where $\mathbf{x}$ is a discrete representation of the atmospheric state. It contains the atmospheric temperature and humidity profiles and any other atmospheric variables which affect significantly the radiation emitted to space. This will include surface parameters and, in general, parameters from which the effects of cloud on the out-going radiances may be calculated. The inverse problem is that of estimating $\mathbf{x}$, given a set of measurements $\mathbf{y}^m$. In most of the following discussion, the problem will be treated as one-dimensional, i.e. $\mathbf{y}^m$ will be multi-channel radiances at one location and $\mathbf{x}$ will be a description of the atmosphere in the vertical at that location. However the theory presented is not restricted to one-dimension; in principle, $\mathbf{x}$ could be a 3-dimensional description of the atmospheric state and $\mathbf{y}^m$ a field of multi-channel radiances.

With this approach, the cloud parameters no longer play such a special role in the retrieval problem – they become just some amongst a large number of parameters required for an adequate description of the atmosphere. However, the inclusion of cloud in the problem does have important practical implications for the method of solution because it makes the problem <u>highly nonlinear</u>. The degree of linearity of the problem is determined by examining the gradient of the forward problem:

$$\mathbf{K}\{\mathbf{x}\} \;=\; d\mathbf{y}\{\mathbf{x}\}/d\mathbf{x} \qquad (1)$$

In general $\mathbf{K}$ is a function of $\mathbf{x}$, but in the special case that $\mathbf{K}$ is independent of $\mathbf{x}$ the problem is linear. When $\mathbf{x}$ includes cloud parameters, $\mathbf{K}\{\mathbf{x}\}$ becomes highly dependent on the values of these parameters. This can be appreciated by noting that $\mathbf{K}\{\mathbf{x}\}$ is closely linked with the conventional definition of the weighting function – in the linear limit they are identical. Fig. 1 illustrates the effect of a cloud layer on a single temperature weighting function. It can be seen that the presence of cloud has a profound effect on the shape of the weighting function, and this is what we mean when we say that clouds make the forward problem and inversion problem highly nonlinear.

3. Bayes Theorem

Before considering the optimal solution to the inversion problem, let us introduce some relevant concepts from probability theory.

It is often useful to think of our knowledge of a variable in terms of a probability density function (PDF): let $P(x)$ express the probability that a scalar variable has a value x. If we know the estimates of x have mean value x_o and errors which, statistically, are normally distributed (Gaussian) with standard derivation σ, then we can say that x has a probability described by a PDF,

$$P(x) \;\alpha\; \exp\,[-\frac{1}{2}\,(x-x_o)^2/\sigma^2]\,. \qquad (2)$$

When considering a vector variable $\mathbf{x}$, the equivalent equation is:

$$P(\mathbf{x}) \;\alpha\; \exp\,[-\frac{1}{2}\,(\mathbf{x}-\mathbf{x}_o)^T.\mathbf{C}^{-1}.(\mathbf{x}-\mathbf{x}_o)] \qquad (3)$$

where $\mathbf{C}$ is the error covariance about the mean value $\mathbf{x}_o$, and ^{T} and $^{-1}$ denote matrix transpose and inverse operations respectively.

A very powerful tool in probability theory is BAYES THEOREM, which helps us to manipulate the probabilities of two states or events occurring together. Let us consider two states, x and y. The probability that x and y will occur <u>together</u> is called their JOINT PROBABILITY and is denoted $P(x,y)$. The probability that x will occur <u>when</u> y occurs is called the CONDITIONAL PROBABILITY of x given y, and it is denoted $P(x\mid y)$. It is evident that the two are related by

$$P(x,y) \;=\; P(x\mid y)\;P(y) \qquad (4)$$

i.e. that the probability of x and y occurring together is the probability that x occurs when y occurs multiplied by the probability that y occurs.

We can also interchange x and y:

$$P(x,y) \;=\; P(y\mid x)\;P(x) \qquad (5)$$

Combining Eqs. (4) and (5):

$$P(x\mid y) \;=\; \frac{P(y\mid x)\;P(x)}{P(y)} \qquad (6)$$

This is BAYES THEOREM.

4. The maximum probability solution to the inversion problem

Our problem can be stated simply: given the measurements $\mathbf{y}^m$, what is the most probable value of the atmospheric state vector $\mathbf{x}$? We solve this by maximising the conditional probability of $\mathbf{x}$ given $\mathbf{y}^m$, i.e.

$$P(\mathbf{x} \mid \mathbf{y}^m) \;=\; \text{maximum} \tag{7}$$

We apply Bayes theorem in the form,

$$P(\mathbf{x} \mid \mathbf{y}^m) \;\alpha\; P(\mathbf{y}^m \mid \mathbf{x}) \; P(\mathbf{x}) \,. \tag{8}$$

Here we have taken $P(\mathbf{y}^m)$, the prior probability of making a measurement $\mathbf{y}^m$ to be constant (over the range of values allowed by the instrument).

$P(\mathbf{y}^m \mid \mathbf{x})$ is the probability that we shall make a measurement $\mathbf{y}^m$ when the atmospheric state is $\mathbf{x}$. If, as above, we represent the forward problem by the general expression $\mathbf{y}\{\mathbf{x}\}$, then for error-free measurements $P(\mathbf{y}^m \mid \mathbf{x})$ would be a delta function peaking at $\mathbf{y}^m = \mathbf{y}\{\mathbf{x}\}$. However, the measurements will contain errors which we shall assume are Gaussian with covariance $\mathbf{E}$. Then the PDF becomes (cf. Eq. (3))

$$P(\mathbf{y}^m \mid \mathbf{x}) \;\alpha\; \exp\left[-\frac{1}{2}\,(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}\})^T.\mathbf{E}^{-1}.(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}\})\right] . \tag{9}$$

$P(\mathbf{x})$ contains our information on $\mathbf{x}$ prior to making any measurement. This information may come from a number of sources. For example, we may use a forecast profile from a NWP model (along with some estimate of its probable error characteristics) or we may have climatological information such as the climatological mean profile and its covariance about the mean. We shall call such data BACKGROUND INFORMATION and denote it by a background profile $\mathbf{x}^b$ and its error covariance $\mathbf{C}$. Then, for normally distributed background errors, the prior probability of the profile having a value $\mathbf{x}$ is given by (cf. Eq. (3)):

$$P(\mathbf{x}) \;\alpha\; \exp\left[-\frac{1}{2}\,(\mathbf{x} - \mathbf{x}^b)^T.\mathbf{C}^{-1}.(\mathbf{x} - \mathbf{x}^b)\right] . \tag{10}$$

It is more convenient to maximise the logarithm of Eq. (7) rather than Eq. (7) itself; substituting from Eqs. (9) and (10) and taking the log gives

$$ln\,\{P(\mathbf{x} \mid \mathbf{y}^m)\} \;=\; -\frac{1}{2}\,(\mathbf{x} - \mathbf{x}^b)^T\,.\,\mathbf{C}^{-1}.(\mathbf{x} - \mathbf{x}^b) \;-\; \frac{1}{2}(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}\})^T.\mathbf{E}^{-1}.(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}\}) + \text{constant} \tag{11}$$

It can therefore be seen that the maximum probability solution for information with Gaussian errors corresponds to the problem of minimizing a quadratic "cost" or "penalty" function of the form:

$$J(\mathbf{x}) \;=\; \frac{1}{2}\,(\mathbf{x} - \mathbf{x}^b)^T.\mathbf{C}^{-1}.(\mathbf{x} - \mathbf{x}^b) \;+\; \frac{1}{2}(\mathbf{y}^m - y\{\mathbf{x}\})^T.\mathbf{E}^{-1}.(\mathbf{y}^m - y\{\mathbf{x}\}) \tag{12}$$

It would also be possible to add additional terms to the cost function to represent other dynamical or physical constraints on the solution (e.g. that the profile should not be super-saturated or super-adiabatic).

To find the most probable value of $\mathbf{x}$, we may either maximise Eq. (12) or find where the gradient of $J(\mathbf{x})$ with respect to $\mathbf{x}$ is zero:

$$\mathbf{J}'(\mathbf{x}) \;=\; \mathbf{C}^{-1}\,.\,(\mathbf{x} - \mathbf{x}^b) \;-\; \mathbf{K}\{\mathbf{x}\}^T\,.\,\mathbf{E}^{-1}\,.\,(\mathbf{y}^m - y\{\mathbf{x}\}) \;=\; 0 \tag{13}$$

5. The linear problem

For the linear problem, where $\mathbf{K}$ is independent of $\mathbf{x}$, we may write

$$\mathbf{y}\{\mathbf{x}\} \;=\; \mathbf{y}\{\mathbf{x}^b\} \;+\; \mathbf{K}\,.\,(\mathbf{x} - \mathbf{x}^b)\,. \tag{14}$$

In the weakly nonlinear case, in which $\mathbf{K}$ is a function of $\mathbf{x}$, but varies only a little over the domain which represents all reasonable departures of $\mathbf{x}$ from $\mathbf{x}^b$, we may also use Eq. (13) in which $\mathbf{K} = \mathbf{K}\{\mathbf{x}^b\}$.

Substituting Eq. (14) into (13), we can find an analytic solution

$$\mathbf{x} \;=\; \mathbf{x}^b \;+\; (\mathbf{C}^{-1} \;+\; \mathbf{K}^T.\mathbf{E}^{-1}.\mathbf{K})^{-1}\,.\,\mathbf{K}^T.\mathbf{E}^{-1}.(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}^b\})\,. \tag{15}$$

Matrix manipulation gives an equivalent formula which is computationally more efficient when the dimension of $\mathbf{x}$ exceeds that of $\mathbf{y}^m$:

$$\mathbf{x} \;=\; \mathbf{x}^b \;+\; \mathbf{C}.\mathbf{K}^T.(\mathbf{K}.\mathbf{C}.\mathbf{K}^T + \mathbf{E})^{-1}\,.\,(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}^b\})\,. \tag{16}$$

This is the familiar minimum variance solution which is also the maximum probability solution in the case of Gaussian statistics (see, for example, Rodgers 1976).

When performing a retrieval from cloud-cleared radiances, the problem is usually sufficiently linear for this form of the solution to be adequate. Indeed, it is possible to pre-compute the inverse matrix, $\mathbf{W} = \mathbf{C}.\mathbf{K}^T.(\mathbf{K}.\mathbf{C}.\mathbf{K}^T + \mathbf{E})^{-1}$, for a small number of atmospheric and surface conditions, scan angles, etc., and then to choose a value of $\mathbf{W}$ appropriate to $\mathbf{x}^b$ by "nearest neighbour" methods or by interpolation.

However, experience with TOVS retrievals from cloud-cleared radiances has revealed other problems with the measurements after pre-processing and cloud-clearing; the cleared radiances often contain large errors and errors which are non-Gaussian (see Watts and Eyre 1986) and correlated between channels (Watts and McNally 1988). In general, it is clear that the error characteristics after cloud-clearing are complicated and are not accommodated properly by the simple formulation of Eq. (16).

6. The nonlinear problem

Even when using cloud-cleared radiances the problem is nonlinear, particularly with respect to the retrieval of humidity, and ideally should be treated as such. When attempting to retrieve from cloudy radiances, the problem is highly nonlinear, i.e. $\mathbf{K}$ changes rapidly as a function of some of the elements of $\mathbf{x}$. We must, therefore, seek an alternative approach to the solution of Eq. (13). Newtonian iteration yields the following solution: if we guess a vector $\mathbf{x}_n$, then the solution is found through the iteration:

$$\mathbf{x}_{n+1} \;=\; \mathbf{x}_n \;-\; \mathbf{J}''(\mathbf{x}_n)^{-1}\,.\,\mathbf{J}'(\mathbf{x}_n) \tag{17}$$

where, by differentiation of Eq. (13), and neglecting second order terms,

$$\mathbf{J}''(\mathbf{x})^{-1} \simeq \mathbf{S}_n = (\mathbf{C}^{-1} + \mathbf{K}_n^T . \mathbf{E}^{-1} . \mathbf{K}_n)^{-1} , \tag{18}$$

and $\mathbf{K}_n = \mathbf{K}\{\mathbf{x}_n\}$.

The matrix $\mathbf{S}_n$ is interesting in its own right, as it represents the error covariance of the retrieval (see Rodgers 1976). When the measurement vector is shorter than the profile vector, $\mathbf{S}_n$ is more efficiently computed in its equivalent form:

$$\mathbf{S}_n = \mathbf{C} - \mathbf{C}.\mathbf{K}_n^T.(\mathbf{K}_n.\mathbf{C}.\mathbf{K}_n^T + \mathbf{E})^{-1}.\mathbf{K}_n.\mathbf{C} \tag{19}$$

Iteration of these equations proceeds until convergence, i.e. until the increment $(\mathbf{x}_{n+1} - \mathbf{x}_n)$ is acceptably small.

By matrix manipulation, we arrive at another formulation which is computationally more efficient:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + (\mathbf{x}^b - \mathbf{x}_n) + \mathbf{W}_n . [\mathbf{y}^m - \mathbf{y}\{\mathbf{x}_n\} - \mathbf{K}_n . (\mathbf{x}^b - \mathbf{x}_n)]. \tag{20}$$

where

$$\mathbf{W}_n = \mathbf{C}.\mathbf{K}_n{}^T.(\mathbf{K}_n.\mathbf{C}.\mathbf{K}_n{}^T + \mathbf{E})^{-1} \tag{21}$$

In practice, the problem including cloud is so nonlinear that the iteration of Eq. (20) tends to be unstable and the iteration must be "damped". Technical details of this are given by Eyre (1989a).

Iteration of the damped equivalent of Eq. (20) proceeds until convergence. At this point we should also find that, if we substitute $\mathbf{x}$ back into the radiative transfer model, the departures $(\mathbf{y}^m - \mathbf{y}\{\mathbf{x}\})$ are of the order of the measurement error in all channels. If this is not the case, it suggests a problem with either the measurements or the forward model, such as more complex cloud conditions than the forward model allows for. This offers a natural and, in practice, a powerful mechanism for quality control.

Substitution of the final profile $\mathbf{x}$ into Eq. (19) provides an estimate of the expected retrieval error covariance, $\mathbf{S}$. This, together with the prior error covariance $\mathbf{C}$, can be used when assessing the relative weight to be given to the retrieved profile when it is used in the subsequent NWP data assimilation.

7. Application to cloud-affected TOVS radiances

The method described above has been applied to "raw" TOVS radiances - they have undergone no pre-processing (other than the mapping of MSU data to HIRS fovs) and no cloud-clearing. The profile vector $\mathbf{x}$ includes all the variables required for an adequate repr .tion of the radiative transfer problem:

- temperature profile (surface to 0.1 mb),
- humidity profile (surface to 300 mb),
- surface skin temperature and surface pressure,
- microwave surface emissivity,
- cloud-top pressure and fractional cloud amount.

The method requires not only a fast radiative transfer model for TOVS channels to calculate $\mathbf{y}\{\mathbf{x}\}$ but also a fast gradient model to calculate $\mathbf{K}\{\mathbf{x}\}$. The TOVSRAD model (Eyre 1984), developed from routines in the International TOVS Processing Package which closely follow Weinreb *et al.* (1981), has been further extended to calculate the derivatives of brightness temperature with respect to all the elements of the profile vector (see Eyre 1989a). Cloud-affected radiances are calculated as follows:

$$R = (1 - N_c)\ R^{cl} + N_c\ R^o(p_c) \tag{22}$$

where R^{cl} is the clear column radiance, $R^o(p_c)$ the overcast radiance for cloud-top p_c, and N_c the "effective" fractional cloud amount. For opaque cloud, N_c represents the true cloud amount, while for semi-transparent cloud it is the product of the fractional cloud cover and the cloud "emissivity" (i.e. one minus cloud transmittance).

The matrix $\mathbf{E}$ represents the "measurement error" covariance. However it is important that it contains contributions not only from the expected radiometric noise but also from expected errors in the radiative transfer calculations, which will often dominate.

The method has been tested in the context of a TOVS retrieval scheme which provides information primarily for input to a NWP model. Therefore the background profile $\mathbf{x}^b$ is taken from a short-range forecast interpolated in time and space to the location of the TOVS data, and its error covariance $\mathbf{C}$ is constructed by assessing (with some difficulty) the expected errors in such a forecast profiles (see Eyre 1989a).

The forecast model does not (currently) provide estimates of cloud parameters but this does not matter; background cloud parameters are set to mean values (e.g. 50% cover at 600 mb) and their expected errors set very large. Thus cloud is effectively unconstrained by the background and the retrieved values are governed by the data. (Microwave surface emissivity is treated similarly).

The iterative retrieval through Eq. (20) must be started from some initial profile. For most parameters it is convenient (though not essential) to use the background profile, since an accurate initial point will speed convergence. An accurate guess at the cloud parameters is also found to be important; these are obtained not from the background values but using the method described by Eyre and Menzel (1989). Note that, in an optimal method such as this, the solution is a function of the background vector (and its error covariance) through Eq. (13). However it is <u>not</u> a function of the parameters used to initialize the iteration, unless the cost function (Eq. (12)) has multiple minima.

8. Expected error characteristics

As discussed above, if we find a profile vector $\mathbf{x}$ which satisfies Eq. (13), then its error covariance $\mathbf{S}$ is given by Eq. (19). The roots of the diagonal elements of $\mathbf{S}$ may be interpreted as the standard deviations of retrieval error. By evaluating $\mathbf{S}$ for different values of $\mathbf{x}$, we can study some of the error characteristics of the inversion scheme. One should not place too much emphasis on the absolute values of retrieval error, since they are highly dependent on the values of background error (for which a typical covariance matrix has been taken for illustrative purposes). More attention should be paid to the ratio of retrieval to background error, as this is related to the information which the radiance data supply to the NWP system. Of particular interest here is the behaviour of $\mathbf{S}$ as a function of those variables which make the problem highly nonlinear, namely the cloud-top pressure and cloud amount.

Figure 2 shows the retrieval errors expected for two cases – cloud-free conditions and full cloud cover at 500 mb – along with the background profile errors, for typical mid-latitude conditions. It shows that, even in cloud-free conditions, the retrieval errors represent only a moderate reduction from the background errors. This arises because the background errors themselves have been assumed quite small (for the temperature profile), typical of 12-hour forecast errors for a regional NWP model operating in northern hemisphere mid-latitudes. Also the inter-level correlations of background error are quite weak and the TOVS weighting functions relatively broad. The combination of these facts limits the information content of the TOVS radiances with respect to the NWP analysis. Nevertheless, the ratio of retrieval to background error represents a decrease in error variance which would be beneficial if obtained in practice.

The retrieval errors for the cloudy case shown in Figure 2 demonstrate that, for levels well below the cloud top, retrieval performance degrades to behaviour typical of an MSU-only system. However, at and just above the cloud-top, the retrieval error is reduced a little. This effect is demonstrated more clearly in Figure 3, which shows 500 mb temperature retrieval error as a function of cloud-top pressure and cloud amount. Errors are presented here as fractional unexplained variances, i.e. the variance of retrieval error divided by the variance of the prior (or background or forecast) error. These results show that, when treated optimally, the effects of cloud on infra-red radiances can improve some aspects of the retrieval performance. This is not observed in retrieval schemes where cloud-clearing and inversion are separated; in these, cloud can only degrade the retrievals.

The crosses in Figure 2 indicate the retrieval error in the cloud-free case when the background errors of cloud parameters are set to zero (i.e. when we know *a priori* that we have a cloud-free case). This shows that, except for surface skin temperature, retrieval performance is only degraded a little by lack of prior knowledge of the cloud conditions.

Figures 4 and 5 show plots for the error in retrieved cloud parameters as a function of these parameters. These plots demonstrate that cloud parameters may be retrieved with reasonable accuracy under most conditions, and particularly well for significant amounts of high cloud. Understandably, cloud-top pressure errors increase at low cloud amount, as do cloud amount errors for low-level cloud. In these cases, cloud uncertainties have little effect on the measured radiances and so do not degrade unduly the retrieval of other parameters.

9. Implementation

The retrieval method described above has been tested on limited amounts of TOVS data. One case study is presented by Eyre (1989b), showing encouraging results. Work has been continuing recently on the testing and refining of this scheme in a joint project between the U.K. Meteorological Office, Oxford University and ECMWF. Another retrieval scheme based on the method presented here has been implemented at AES in Canada, again with encouraging initial results (see Steenbergen *et al.* 1989).

Although early results are promising, it has still to demonstrated whether such methods can yield consistently better results than retrieval schemes with a separate cloud-clearing stage. Two factors are likely to determine whether such schemes can be used operationally in the near future:

- The radiative transfer models used must be sufficiently accurate, particularly with respect to their representation of cloud, and (a related point) their errors must be adequately characterised.

- If these potential scientific problems are not too serious, the computational cost for real-time use will need to be considered; the method involves for each sounding several computations of TOVS radiances and their derivatives together with some moderately large matrix manipulations.

In the longer term, the approach presented here looks very attractive, because of its complete generality (and hence its applicability to data from other sounding systems) and because it is compatible with other recent developments in the assimilation of satellite data into NWP models.

40-89

REFERENCES

Eyre J.R., 1984. User guide to TOVSRAD. A program for calculating synthetic HIRS-2 and MSU equivalent blackbody temperatures. Met.O.19 Branch Memorandum 75, Meteorological Office, Bracknell.

Eyre J.R., 1989a. Inversion of cloudy satellite soundings by nonlinear optimal estimation: theory and simulation for TOVS. Q.J.R. Meteorol. Soc., 115, *1001-1036.*

Eyre J.R., 1989b. Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation: application to TOVS data. Q.J.R. Meteorol. Soc. 115, *1027-1037.*

Eyre J.R. and Menzel W.P., 1989. Retrieval of cloud parameters from satellite sounder data: a simulation study. J. Appl. Meteorol. 28, *267-275.*

Huang H.-L. A. and Smith W.L., 1986. An extension of the simultaneous TOVS retrieval algorithm the inclusion of cloud. Tech. Proc. 3rd Int. TOVS Study Conf.; Madison, Wisconsin; 13-19 August 1986; Report of CIMSS, University of Wisconsin-Madison, Ed.: W.P. Menzel; *118 131.*

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

Steenbergen D., Chouinard C. and Greaves B.T., 1989. Preliminary results from using nonlinear optimal estimation retrievals in the Canadian regional analysis scheme. To appear in Tech. Proc. 5th Int. TOVS Study Conf.; Toulouse; 24-28 July 1989; Report of CIMSS, University of Wisconsin-Madison.

Susskind J., Rosenfield J., Reuter D. and Chahine M.T. 1984. Remote sensing of weather and climate parameters from HIRS2/MSU on TIROS-N. J. Geophys. Res., 89, *4677 4697.*

Watts P.D. and Eyre J.R., 1986. Validating a TOVS cloud-clearing scheme using AVHRR data. Tech. Proc. 3rd Int. TOVS Study Conf.; Madison, Wisconsin; 13-19 August 1986; Report of CIMSS, University of Wisconsin-Madison, Ed.: W.P. Menzel; *pp. 312 331.*

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

Weinreb M.P., Fleming, H.E., McMillin L.M. and Neuendorffer A.C., 1981. Transmittances for the TIROS Operational Vertical Sounder. NOAA Tech. Rep. NESS 85.

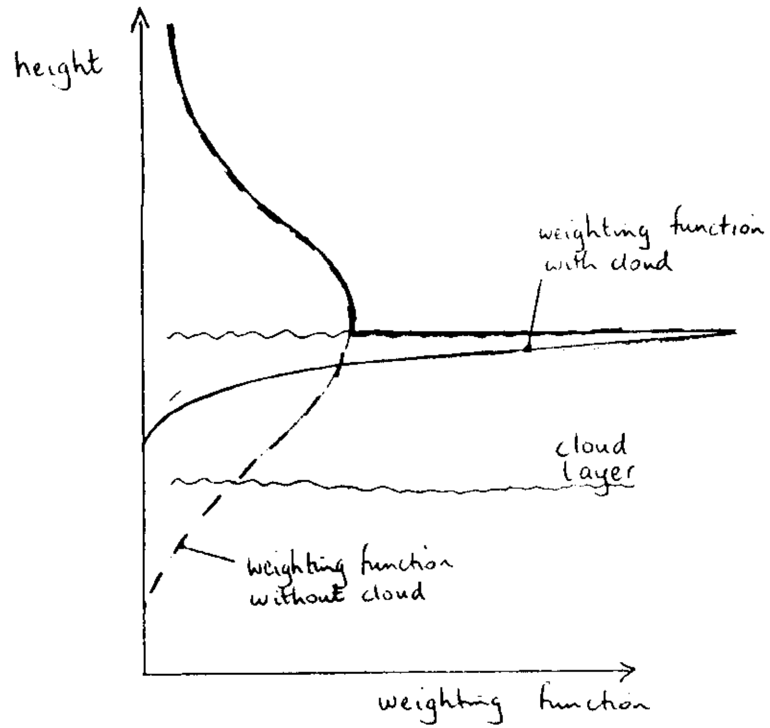

Figure 1. Illustrating the effect of cloud on an infra-red weighting function.

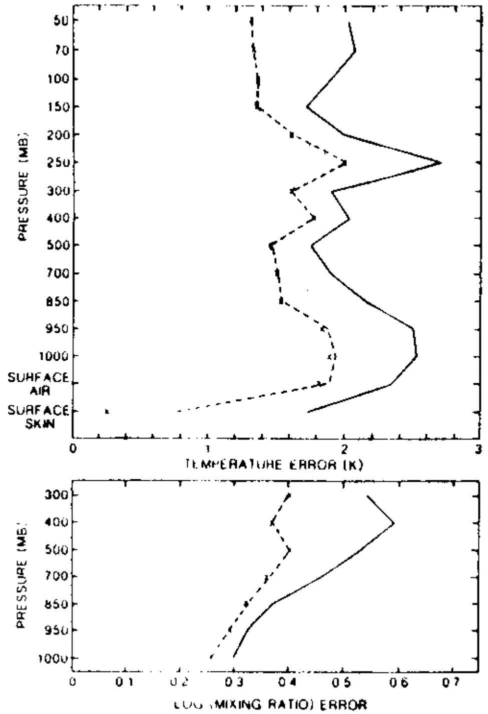

Figure 2. Theoretical r.m.s. errors. —— background errors, - - - retrieval errors for cloud-free case, ··· retrieval errors for complete cloud cover at 500 mb. × represents retrieval errors for cloud-free case which is known *a priori* to be cloud free.

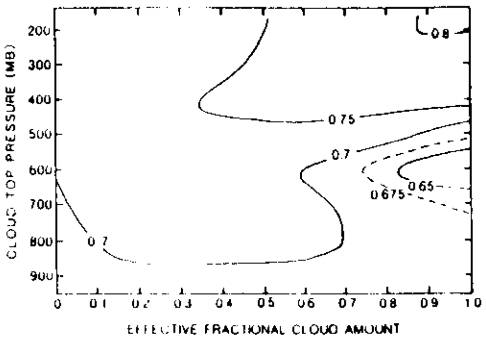

Figure 3. Theoretical errors in retrieved 500 mb temperature expressed as fractional unexplained variance as a function of cloud conditions.

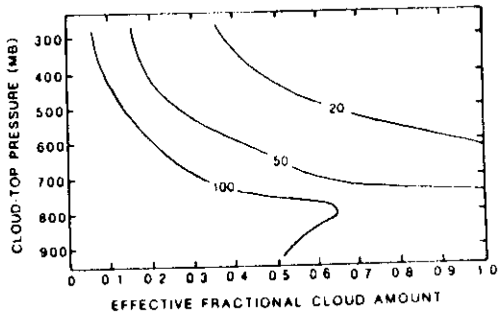

Figure 4. Theoretical r.m.s. errors in mb in retrieved cloud-top pressure as a function of cloud conditions.

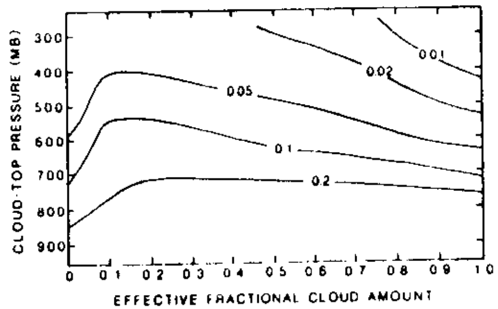

Figure 5. Theoretical r.m.s. errors in retrieved effective fractional cloud amount as a function of cloud conditions.