Tests of a scheme for regression retrieval and time–space interpolation of stratospheric temperature from satellite measurements

By D. R. JACKSON*, R. S. HARWOOD and E. RENSHAW

1 Department of Meteorology, University of Edinburgh; 2 Department of Statistics, University of Edinburgh

(Received 5 June 1989; revised 2 April 1990)

SUMMARY

A scheme to retrieve and analyse stratospheric temperature from satellite measurements is tested. Because of the lack of 'ground truth' in the stratosphere, the 'true' atmosphere is represented by an atmosphere simulated by a numerical model. Simulated observations are calculated by computing the radiance that would be observed from the 'true' atmosphere by a satellite instrument. The radiances are then retrieved and analysed and the resultant analyses compared with the corresponding 'true' fields. The tests are made using output from a day when a sudden warming was present. The retrievals are made by using a multiple linear regression model which regresses radiances against Planck function. The corresponding temperatures are then analysed on a grid using a linear time–space interpolation scheme.

The retrieval scheme is seen not to perform so well within the area of the sudden warming as outside it. However, this may be expected as the vertical structure within the sudden warming is generally too small to be resolved by a satellite instrument. The analysis scheme analyses the stratospheric field well, even in the area of a sudden warming. The effect of varying the distance radius used to select observations for the scheme is also discussed.

1. INTRODUCTION

Electromagnetic radiation leaving the top of the atmosphere carries information about the distribution of temperature and of the emitting gases. If measurements are made at wavelengths at which the emission is by a gas of known mixing ratio such as carbon dioxide, then some details of the temperature distribution may be deduced (Kaplan 1959): this is a principle exploited in several remote-sensing satellites. However, the process of deducing the temperature structure from measurements is not necessarily straightforward. Usually there are two aspects to the problem, although it is possible to devise procedures in which they are combined. The two aspects are: 'retrieval', in which a single temperature profile is deduced from a more-or-less instantaneous set of measurements; and 'analysis' in which the state of the atmosphere at a given instant is deduced on a regularly spaced grid of points from the retrieved profiles, which are asymptotic and distributed according to the shifting satellite orbit. Without further information the retrieval problem is under-constrained because in general an infinite number of atmospheric profiles can yield the same finite set of measurements. Moreover the analysis problem can suffer from aliasing difficulties. Aliasing occurs in all Fourier analyses of discrete data: the time period of the data imposes a limit on the highest resolvable frequency, and hence any higher frequency present will be analysed falsely within the range of the lower, resolvable, frequencies.

The aim of this paper is to evaluate the performance of a retrieval—analysis scheme for obtaining stratospheric temperatures from the TIROS Operational Vertical Sounder (TOVS) instrument (Schwalb 1978; Smith et al. 1979) on the TIROS-N series of polar-orbiting satellites. TOVS comprises three sounders, of which the Stratospheric Sounding Unit (SSU) (see Miller et al. (1980), and references therein) is of most relevance to stratospheric studies.

The retrieval scheme which we have tested is based on a regression model similar to that used by the Meteorological Office (Pick and Brownscombe 1981), the main

* Present affiliation: Department of Meteorology, University of Reading.
difference being that their scheme uses the measured radiances to give thicknesses of fairly thick layers of atmosphere, whereas ours gives the temperature profile at 31 pressure levels from 0·2 to 570 mb. Many previous tests of retrieval schemes have compared retrievals with coincident rocketsonde measurements. Nash and Brownscombe (1983) and Pick and Brownscombe (1981) tested the Stratospheric Sounding Unit (SSU) on TOVS; Barnett et al. (1975) tested the Selective Chopper Radiometer (SCR) on the Nimbus-5 satellite. Whilst the chief purpose of those tests was to assess the performance of the satellite instrument, here we pay particular attention to the regression–retrieval scheme itself. We examine the representativeness of the data-sets used to calculate the regression coefficients, and test the ability to retrieve temperature in various atmospheric conditions.

The time–space analysis method is that used operationally by the Meteorological Office, so that this paper will help users of those analyses to evaluate the confidence which can be placed in them. The analysis scheme gives each observation a time and distance weight which decreases the further the observation is from the gridpoint or analysis time. Only observations lying within a specified time and distance (called 'search radii') of the gridpoint and analysis time are used in the scheme. Most tests of the scheme are initially made using idealized fields, and then on a field simulated by a numerical model. We concentrate chiefly on the way the quality of the analysis changes when the search radii are changed.

A difficulty in testing such schemes using real observations is the absence of adequate 'ground-truth' observations. Rocketsondes observe temperature in the upper stratosphere but rocket flights are infrequent and badly spaced. Radiosondes observe in the lower stratosphere and give better global coverage, but even so there are few observations made over the oceans or in the southern hemisphere. Tests which have been done with real data compare retrieved profiles with coincident rocketsonde measurements (Nash and Brownscombe 1983; Pick and Brownscombe 1981; Barnett et al. 1975). Furthermore it is generally even more difficult to make comparisons with ground truth for fields analysed from satellite measurements than it is for retrieved profiles. Possible difficulties with analyses have been reported by Al-Ajmi et al. (1985) and by Clough et al. (1985). The former paper gives evidence that there may be temporal variations too rapid for proper resolution, and the latter provides evidence that the vertical temperature structure is not always adequately resolved. Accordingly we have chosen to test the schemes in a simulation experiment which uses an atmosphere calculated in a numerical model. Simulated observations are calculated by computing the radiances which would be observed from this model atmosphere by a TOVS-like instrument, including the effects of instrumental noise. These radiances are then retrieved and analysed and the resultant analyses compared with the corresponding model fields.

Section 2 gives details of the model and of the simulation method. Section 3 describes the retrieval scheme and the method of obtaining the regression coefficients, together with the results of tests of the retrieval scheme. Section 4 contains a description of the analysis scheme. Results of tests of the analysis scheme on analytical fields also appear in section 4, whilst results of tests made on fields calculated in a numerical model appear in section 5. Conclusions appear in section 6.

2. SIMULATION OF ATMOSPHERE AND OBSERVATIONS

(a) The stratosphere–mesosphere model

The retrieval and analysis schemes are tested below in a simulation experiment which uses an atmosphere calculated in a numerical model. Although the model gives a
reasonable representation of the 'true' atmosphere, it is not capable of reproducing certain phenomena, such as tides, which exist in the real stratosphere. The model used is the Meteorological Office stratosphere–mesosphere multi-level model (Fisher 1987) based on the primitive equations. These equations are solved to fourth-order accuracy in the horizontal, and to second-order accuracy both in the vertical and in time, using energy-conserving 'box' type finite differences and leapfrog integration. The model utilizes a regular grid in spherical coordinates with gridpoints at intervals of 5° in latitude and longitude and 33 levels between 100 and 0.001 mb, which are equally spaced in log pressure, and are approximately 2 km apart. To avoid having to represent the troposphere, a lower boundary condition is imposed near the tropopause, namely the geopotential height of the 100 mb surface specified from analysed observations.

Our simulation uses one day's output at 1-hour intervals from a run with lower boundary heights corresponding to 18 Jan. 1987. On that day a 'sudden warming' was present in the modelled northern hemisphere. Such warmings cause large and rapid changes to the temperature structure of the stratosphere—there may be temperature rises of the order of 50 K over a few days. The phenomenon provides the most stringent circumstances for testing both the retrieval and analysis schemes.

(b) Satellite observation pattern

The TIROS-N satellite views the atmosphere by scanning from one side of the vertical to the other at 8 scan angles. Adjacent observations are then averaged in blocks of 4 so that the 16 observations of two successive scans are combined to give 4 'superobservations' at effective angles of −30°, −10°, 10° and 30° degrees from the vertical. Figure 1 shows the superobservations made by a TIROS-N series satellite (NOAA-7) in a 24-hour period. There is almost global data coverage, though some areas in the subtropics are free of observations. Model temperatures are interpolated linearly in time and space between these observation points, and the interpolated temperatures are used to calculate radiances that the satellite would 'observe'. In these tests it is assumed that there is no data loss due to calibration sequences.

Figure 1. 'Superobservation' points of the TOVS instrument of NOAA-7 for 18 Jan. 1987. Observations are made at the 4 scan angles −30°, −10°, 10° and 30° from the local vertical.
(c) Radiative transfer

As mentioned in the introduction, the radiance measured by the satellite is a function of atmospheric temperature and of the distribution of the emitting gas. Radiation transfer theory on which temperature sounding is based (Kaplan 1959; Houghton and Smith 1970) relates the spectral radiance, $R_{\nu}$, of an instrument channel centred at wavenumber $\nu$ to the Planck function, and hence to temperature. With negligible transmission from the earth’s surface this theory gives

$$R_{\nu} = \int_{0}^{\infty} K(y) B_{\nu}(T) dy.$$  \hspace{1cm} (1)

Here $y$ is a vertical coordinate given by $y = -\ln(p/p_0)$ where $p$ is pressure and $p_0$ is a reference pressure, $K(y)$ is the channel weighting function, and $B_{\nu}(T)$ is the Planck function at temperature $T$ given by

$$B_{\nu}(T) = c_1 \nu^2/[\exp(c_2 \nu/T) - 1]$$  \hspace{1cm} (2)

where $c_1 = 1.19096 \times 10^{-5}$ mW m$^{-2}$ cm$^4$ ster$^{-1}$ and $c_2 = 1.43879$ cm K.

The weighting functions depend upon the angle of view. However, to simplify the calculations, we have simulated all radiances using the weighting function appropriate to the vertical view. There is thus the possibility of an extra source of error in the real case arising from a misrepresentation of the radiative transfer for the slant path which is not simulated in the present study. The transmission profile, and hence weighting function, for each channel was calculated using a numerical technique developed by McMillin and Fleming (1976): first, transmission profiles are calculated for a small number of representative and extreme atmospheres using the line-by-line method (Drayson 1966), and then these pre-computed profiles are interpolated to any arbitrary temperature profile. ‘Observed’ radiances can thus be calculated by evaluating Eq. (1) and then adding a randomly generated number to simulate the radiometric noise of the instrument.

The largest contribution to the radiance comes from pressure levels close to the peak of the weighting function. Indeed, channel radiance may be considered to give a measure of the temperature of a layer 10–15 km thick situated about the peak of the weighting function. The weighting functions for the 8 channels used in the retrieval scheme are shown in Fig. 2. Table 1 shows the pressure level at each channel’s weighting function peak and its central wavenumber $\nu$.

3. Retrieval

(a) The regression model

The Planck function, and hence temperature, is retrieved from radiance measurements by regression. The temperature is calculated from the Planck function under the assumption that the Planck function has been calculated at a reference wavenumber $U$, which here is chosen to be 668 cm$^{-1}$. The regression model assumes that the Planck function is linearly related to the radiances, which have also been standardized to the reference wavenumber $U$. This is done because the standardized radiances are more linearly related to the Planck function at 668 cm$^{-1}$ than are the ‘observed’ radiances. The standardized radiance, $\chi_j$, of channel $j$ is easily calculated from $R_j$, the measured radiance of channel $j$ via

$$\chi_j = c_1 U^2/[\exp(c_2 U/r) - 1]$$  \hspace{1cm} (3)

where $r = c_2 \nu^2/[\exp(c_2 \nu/R_j)$. 

The estimated deviation of profile Planck function, $\Delta \hat{B}_i$, from the mean is written as a linear combination of the deviation of the radiances from the mean radiance, namely

$$\Delta \hat{B}_i = a_{0,i} + \sum_{j=1}^{N} (\chi_j - \bar{\chi})a_{j,i} \quad (i = 1, \ldots, 31).$$

The $a_{0,i}$ and $a_{j,i}$ are predetermined in advance by least squares as described in section 3(b) below. If there is no bias in the system then, of course, $a_{0,i}$ would be zero. It has been shown (e.g. Eyre 1987) that this retrieval scheme is mathematically equivalent to other schemes which retrieve temperature by a minimum variance solution of the radiative transfer equation. The mean Planck function and the mean standardized radiance, $\bar{\chi}$, are calculated from a set of rocketsonde temperature measurements (this data-set is also described in section 3(b)). $N$ is the total number of channels and $\chi_j$ is the observed standardized radiance. Retrievals are performed at 31 pressure levels between 0-2 and 570 mb.
The retrieved Planck function, \( \hat{B}_i \), for level \( i \) is then given by
\[
\hat{B}_i = B_i + \Delta B_i \quad (i = 1, \ldots, 31)
\]
where \( \overline{B}_i \) is the mean Planck function for pressure level \( i \). Calculation of temperature from the retrieved Planck function is straightforward.

**(b) Calculation of regression coefficients**

The regression analysis is based on a data-set of 1200 temperature profiles. Each profile is calculated using a combination of a radiosonde measurement and a quasi-coincident rocketsonde measurement. Standardized radiances are calculated from these measurements using Eqs. (1) and (3). The data are divided into 7 zones (Table 2) according to the latitude and season of each rocketsonde measurement, and regression coefficients are calculated for each of the zones. This is done to restrict the range of atmospheric conditions over which the regression analyses are applied. Note that in using calculated radiances to determine the regression coefficients, we are implicitly assuming that the weighting functions in Eq. (1) are known correctly (i.e., that the 'forward model' is accurate). In practice this may not be the case, leading to potential errors (see e.g., Nash and Brownscombe 1983) of a type beyond the scope of the present investigation.

**TABLE 2. LATITUDE-SEASON ZONES FOR WHICH REGRESSION COEFFICIENTS ARE CALCULATED. 'WINTER' IS THE SIX MONTHS BETWEEN OCTOBER AND MARCH FOR THE NORTHERN HEMISPHERE OR BETWEEN APRIL AND SEPTEMBER FOR THE SOUTHERN HEMISPHERE. 'SUMMER' IS THE SIX MONTHS BETWEEN APRIL AND SEPTEMBER FOR THE NORTHERN HEMISPHERE OR BETWEEN OCTOBER AND MARCH FOR THE SOUTHERN HEMISPHERE**

<table>
<thead>
<tr>
<th>Zone</th>
<th>Latitude–season</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>70°–90°, winter</td>
</tr>
<tr>
<td>2</td>
<td>50°–70°, winter</td>
</tr>
<tr>
<td>3</td>
<td>30°–50°, winter</td>
</tr>
<tr>
<td>4</td>
<td>30°N–30°S, all seasons</td>
</tr>
<tr>
<td>5</td>
<td>30°–50°, summer</td>
</tr>
<tr>
<td>6</td>
<td>50°–70°, summer</td>
</tr>
<tr>
<td>7</td>
<td>70°–90°, summer</td>
</tr>
</tbody>
</table>

(i) **Discussion of the sonde data-set.** The data-set used to calculate the regression coefficients contains rocketsonde measurements of the real atmosphere. However, in this paper the retrieval and analysis schemes are tested in a simulation experiment which uses an atmosphere calculated in a numerical model. It is important that the model adequately reproduces the real atmosphere observed by rocketsondes. This is examined by comparing the means and standard deviations of the two data-sets.

We consider first the southern hemisphere, zone 5. The model field in this zone is similar to the climatology of the sonde data. Figure 3(a) shows means and standard deviations of sonde and model data-sets in zone 5. The dashed lines show the mean of the sonde temperatures together with departures of one standard deviation. The solid lines show the corresponding statistics for the model data-set on 18 Jan. 1987 in this zone. Both sonde and model standard deviations are low, indicating the lack of variation in the summer stratosphere. Moreover, at most pressure levels the means and standard
deviations of both data-sets are close to one another, implying that the model field in zone 5 is similar to that observed by rocketsondes.

Whilst sonde and model statistics are similar for zone 5, the same is not true for zones 1, 2 and 3. On 18 Jan. 1987 the model reproduces a sudden warming in these latter zones. In Table 2 ‘winter’ is defined as the 6-month period between October and March for the northern hemisphere or between April and September for the southern hemisphere. Within such a large time period only a small proportion (if any) of rocketsondes will observe a sudden warming, thus it is likely that sonde data in zones 1, 2 and 3 will not be representative of sudden warming conditions. Figure 3(b) is similar to Fig. 3(a) except that data for zone 1 are shown. As one might expect in a winter stratosphere, both sonde and model data-sets have high standard deviations. The two mean profiles differ greatly at a number of pressure levels. Between 3 and 50 mb the model mean temperature is more than one standard deviation of the sonde data away from the sonde mean temperature. Thus model conditions in zone 1 differ greatly from
those observed by rocketsonde. This is also true in zone 2 (and to a lesser extent in zone 3). It is important to test the retrieval scheme in a sudden warming precisely because of this extreme difference between sudden warming conditions and mean sonde values. If the scheme performs well in a sudden warming, then it is reasonable to conclude that it will perform well in most conditions.

These results have important implications for the method of retrieval by regression, since the sonde data-set used to calculate the regression coefficients is used widely. As this data-set appears to be unrepresentative of sudden warming conditions, future retrieval research might profitably use data-sets which are taken from a forecast model rather than from climatology. Other approaches in the context of tropospheric retrievals, have been suggested by Uddstrom and Wark (1985) and Chedin et al. (1985). The former use a 'stratified climatology' of several atmospheric classes; the radiances are used to identify from which class to take the appropriate regression coefficients (or their equivalents). The latter use a somewhat related 'library search' technique.

(ii) Testing the regression model. It is an assumption of the regression model that the Planck functions in the profile are linearly related to the measured radiances. This is suggested by the form of the radiative transfer equation (1), provided the weighting functions are not temperature dependent. To test the validity of this assumption we have produced plots of Planck function at level $k$ against radiance for channel $j$ for a number of $k,j$ pairs (not shown). No evidence of a need to depart from the linear model was found.

It is of interest to investigate if a certain amount of effort could be saved by using less than 8 channels in the regression. Also, it is of interest to know what the information content of each channel is (Peckham 1974). Accordingly, we have produced scatter plots of radiances in pairs of channels (not shown) to discover the degree of correlation between them. Several channels show a high degree of association, but a test of regression models using less than 8 channels in all cases produced significantly higher r.m.s. errors than the 8-channel model, which has, in consequence, been used in the rest of the work here.

(c) Retrieval results

In this section we test the retrieval scheme by retrieving the profiles corresponding to the simulated observations. Temperatures are retrieved using Eqs. (4) and (5), and biases and r.m.s. retrieval errors are calculated under the assumption that the interpolated model temperature is the 'true' temperature. The standard deviation of the retrievals is also calculated. This is, of course, equal to the square root of the difference between the square of the r.m.s. error and the square of the bias.

For all regression zones the r.m.s. retrieval error at pressures greater than that of the peak of the bottommost weighting function (300 mb) is generally higher than at 300 mb, and the r.m.s. error at pressures less than that of the topmost weighting function peak (1.5 mb) is generally larger than the r.m.s. error at 1.5 mb. This is of course to be expected, as away from the region containing the weighting-function peaks the satellite measurements provide little information about the temperature.

As anticipated, errors in the zones containing a sudden warming (zones 1 and 2) are high. Difficulties arise because of the small-scale vertical temperature structure present in a sudden warming and because the sonde measurements in these zones were made in conditions different from those in the sudden warming (see, for example, Fig. 3(b)). Figure 4(a) shows the r.m.s. error (solid), bias (dashed) and standard deviation (dotted) profiles for zone 1. For pressures higher than 1.5 mb the error is between 2 and 6 K. The
Figure 4. R.m.s. temperature retrieval error (solid); bias (dashed) and standard deviation (dotted) profiles (K). (a) zone 1 (70° to 90° latitude, winter); (b) zone 4 (30°N to 30°S, all seasons); (c) zone 5 (30° to 50° latitude, summer); (d) zone 6 (50° to 70° latitude, summer).
Figure 4. Continued.
standard deviation has nearly the same values as the r.m.s. errors at most levels and hence the bias is small, not exceeding 2 K. The corresponding profiles for zone 2 (not shown) are similar to those of zone 1. In zone 3 (also not shown) the standard deviation is close to the r.m.s. errors but their values are smaller than in zones 1 and 2. In zone 4 the r.m.s. error profile (Fig. 4(b)) has a zig-zag behaviour. This is due to biases in the retrievals, since peaks in the error profile occur at the same pressure levels as high values in the bias profile. For example, at 7 mb the r.m.s. error is 4-33 K and the bias is −4-13 K and at 100 mb the r.m.s. error is 5-61 K and the bias is −5-20 K. The shape of the bias profile is closely related to the difference between the means for model and sondes and therefore seems to be related to an unrealistic structure in the modelled tropics. The low standard deviation of the sonde measurements used to calculate the regression coefficients (not shown) constrains retrievals to the mean of the sonde temperatures. R.m.s. errors for zone 5 (Fig. 4(e)) are lower than for zones 1 to 3 and the profile has a zig-zag pattern, which is also due to bias in the retrievals. However, these zig-zags are much smaller than for the corresponding profiles in zone 4, the largest bias value being −2-55 K (compared to −5-20 K in zone 4). This is probably because the sonde temperatures in zone 5 (Fig. 3(a)) are more representative of zone-5 model conditions on 18 Jan. 87. In regression zones 6 to 7, where there was little wave activity, r.m.s. retrieval errors are low. The zone-6 r.m.s. error profile (Fig. 4(d)), for example, does not exceed 3 K. The standard deviation of the retrieved temperatures is close to the r.m.s. error and biases are much smaller than in zone 4. This is because, although sonde standard deviations in zones 6 and 7 (not shown) are small, the sonde measurements are more representative of model conditions in zones 6 and 7 than corresponding sonde measurements are of model conditions in zone 4. The biases of the retrievals in zones 6 and 7 are also slightly smaller than the zone 5 bias because, although the sonde measurements represent model conditions well in all 3 zones, the sonde standard deviation is larger in zones 6 and 7 and so the retrievals are less constrained to the mean sonde profile.

It is of interest to note that for the middle-latitude and high-latitude zones, the standard deviations are low. This is of significance for the computation of thermal winds, which, being proportional to horizontal temperature gradients, are unaffected by bias except possibly at transitions between zones.

Cross-sections. The high retrieval errors in zones 1 and 2 are partly attributable to the large vertical temperature gradients present in a sudden warming. This is illustrated by plotting a cross-section of the model temperature field along half a satellite orbit’s observation points and comparing it with a corresponding cross-section of retrieved temperatures.

The half-orbit we considered starts at the equator and passes through the region affected by the sudden warming before returning to the equator about 50 minutes later. Observations are made at 189 points. Broadly speaking, observations 1 to 81 and 153 to 189 are made outside the sudden warming region (in zones 3 and 4) and observations 82 to 152 are made within the sudden warming (in zones 1, 2 and 3).

Figure 5(a) shows a cross-section of the model temperature field between 100 and 0-2 mb and for observations 1 to 189. Within the region of the sudden warming there is a deep low centred at 15 mb and observation number 125. Above this there is a region of high temperature, with peaks at 1-5 mb (observation 100), 0-4 mb (observation 125), and 0-4 mb (observation 177). Below the highs is a strong temperature gradient of approximately 4 K km⁻¹.

Figure 5(b) shows the corresponding cross-section of the retrieved temperatures. Both outside and within the sudden warming retrieved values are less than model values at pressures below 0-5 mb. A good example of this is the low at observation 97. Figure
Figure 5. 'Cross-section' of temperature (K) at 189 observation points along a satellite orbit. (a) model; (b) retrieved values; (c) difference between model and retrieved values.
5(c) shows that the difference between model and retrieved temperatures at observation 97 exceeds 10 K. This suggests that the retrieval is biased at pressures less than 0.5 mb because these pressure levels are far away from weighting function peaks.

At pressures greater than 0.5 mb the model field is well retrieved outside the sudden warming region (observations 1 to 81 and 153 to 189)—differences between model and retrieved fields are generally less than 5 K, which is small compared with the temperature variation over the northern hemisphere (e.g. 45 K at 5 mb). However, within the sudden warming the field is not well retrieved. The model low temperature centred at 15 mb (observation 125) has been retrieved in the right place, but its retrieved depth is around 10 K greater than its model magnitude. The model maximum at 1.5 mb (observation 100) is well reproduced, but the nearby maximum at 0.4 mb (observation 125) is underestimated by 15 K. The retrieved temperature gradient between the highs and the major low is not as strong as in the model field, especially between 3 and 1 mb and observations 109 to 141—differences between model and retrieved fields are greater than 20 K in places. This emphasizes the difficulty in retrieving such vertical structure.

4. **TIME-SPACE INTERPOLATION**

   (a) **The method**

In the introduction we stated that the problem of inferring temperature from satellite measurements has two parts, namely 'retrieval' which was discussed in section 3 and 'analysis', or 'interpolation', which forms the subject of this and the next section.

Figure 6 illustrates the problem of 'analysing' observations on a grid. The figure shows satellite observations made on 18 Jan. 1987 between -20° and +20° latitude, -160° and -120° longitude and 1020 and 1340 GMT. The grid has a spacing of 5° in both latitude and longitude and we wish to interpolate the observations to the space-time gridpoints.

In the interpolation scheme under test, each observation is given a time and distance weight. The weights decrease the further an observation is away from the gridpoint and the analysis time. Here the weights are chosen to be linear; schemes tested using cosine

![Figure 6](image)

Figure 6. TOVS observation points on 18 Jan. 1987 between -20° and +20° latitude, -160° and -120° longitude, and 1020 and 1340 GMT. Also plotted is the model grid with a spacing of 5°.
and negative exponential weights produced essentially similar results. Only observations which lie within a specified time or distance of the analysis time or the gridpoint (called 'search radii') are employed.

The time-weight for the kth observation point within the search time interval is given by

$$w_t(k) = \begin{cases} \frac{(r_t - |t - t_0(k)|)}{r_t} & |t - t_0(k)| \leq r_t \\ 0 & |t - t_0(k)| > r_t \end{cases}$$

where \( r_t \) is the search radius, \( t \) is the analysis time and \( t_0(k) \) is the time of the kth observation within the search radius. Similarly, the distance-weight for the kth observation within the distance search circle is given by

$$w_d(k) = \begin{cases} \frac{(r_d - d(k))}{r_d} & d(k) \leq r_d \\ 0 & d(k) > r_d \end{cases}$$

where \( r_d \) is the distance search radius and \( d(k) \) is the physical distance between the gridpoint and the observation point. Using a standard geometric argument, \( d(k) \) is expressed as

$$d(k) = a[\cos \rho \cos \lambda - \cos \rho_0 \cos \lambda_0]^2 + (\cos \rho \sin \lambda - \cos \rho_0 \sin \lambda_0)^2 + (\sin \rho - \sin \rho_0)^2]^{1/2}$$

where \((\rho, \lambda)\) is the latitude and longitude of the gridpoint and \((\rho_0, \lambda_0)\) is the latitude and longitude of the kth observation point within the search radius.

The combined weight \( w(k) \) is taken to be the product of the time and distance weights:

$$w(k) = w_t(k) w_d(k).$$

We now form the normalized weights:

$$w'(k) = \frac{w(k)}{\sum_{j=1}^{n} w(j)}$$

since we require weights which sum to 1, where \( n \) is the number of observations within both time and distance radii. The interpolated value, \( \hat{T} \), at time \( t \) and gridpoint \((\rho, \lambda)\) is then expressed as the linear sum of the products of weight, \( w'(k) \), and observed value, \( T_0(k) \) for all observations within the search radii, i.e.

$$\hat{T} = \sum_{k=1}^{n} w'(k) T_0(k).$$

(b) Tests with analytical radiance fields

The sizes of the time and distance radii selected for the interpolation scheme are important. If a large radius is chosen, then the analysis may be oversmoothed. The magnitude of this oversmoothing depends on the temporal and spatial variability of the field. The task of simulating radiances from model temperatures and then retrieving and interpolating them requires considerable computer time. Accordingly the effect of varying the search radii used in the interpolation scheme was initially tested using an analytically calculated radiance field. Such tests are computationally inexpensive, and thus can be repeated using a large variety of time and distance radii. In addition, they involve no retrieval errors, since we are solely concerned with interpolating radiances from satellite observation points to grid points. In this section we describe these tests.
The radiance field is expressed as a sum of spherical harmonics; radiances are
easily calculated at either grid or 'observation' points. Satellite 'observation' points are
determined using a model of a polar-orbiting satellite orbit and for computational
Economy are calculated with time intervals larger than those of the real satellite.
Observations are calculated at the nadir point only, and the time interval is 256s. The
'observations' are then interpolated in time and space using the time–space interpolation
scheme, and compared with the grid-point analytical radiance field.

Tests were made on a number of fields. We present results for a field which resembles
the stratosphere during a northern hemisphere winter in that there is a small-amplitude
wavenumber-1 wave in the southern hemisphere and tropics and there are higher-
amplitude wavenumber-1 and wavenumber-2 waves in the middle and high latitudes of
the northern hemisphere. To provide a stringent test of the scheme we have chosen high
wave speeds. The wavenumber-1 wave, for example, completes a latitude circle in 5
days.

Results for a variety of time and distance radii are shown in Table 3. Reducing the
distance radius clearly decreases the r.m.s. error, provided the distance radius is greater
than 2000 km. In addition, detailed consideration of the original and interpolated fields
(not shown) reveals that reduction of the size of the distance radius leads to an improve-
ment in the estimate of the model field's wavenumber-1 and wavenumber-2 features. In
particular:

1. The analysis made with a time radius of 12 hours and a distance radius of 12000 km
fails to estimate the wavenumber-1 and wavenumber-2 behaviour of the model field.

2. Changing the distance radius from 12000 to 5000 km reduces the r.m.s. analysis
error by about 50%, and the resultant analysis determines most of the model field's
wavenumber-1 behaviour, and also some of its wavenumber-2 behaviour.

3. Changing the distance radius from 5000 to 3000 km again reduces the r.m.s. analysis
error by about 50% and both wavenumber-1 and wavenumber-2 features of the model
field are better estimated. The estimates are improved even further when the distance is
reduced from 3000 to 2000 km.

TABLE 3. R.M.S. ANALYSIS ERRORS IN RADIANCE UNITS
(mW m⁻² cm ster⁻¹)

<table>
<thead>
<tr>
<th>( r_o ) (km)</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>12000</td>
<td>3.35</td>
<td>3.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td></td>
<td>1.67</td>
<td>1.67</td>
<td>1.73</td>
</tr>
<tr>
<td>3000</td>
<td></td>
<td>0.93</td>
<td>0.93</td>
<td>1.01</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td>0.64</td>
<td>0.64</td>
<td>0.73</td>
</tr>
<tr>
<td>1500</td>
<td></td>
<td>0.67</td>
<td>0.66</td>
<td>0.73</td>
</tr>
</tbody>
</table>

When a distance radius of 1500 km is used, the r.m.s. error is slightly higher than with
a distance radius of 2000 km. All wavenumber-1 and wavenumber-2 features are well
estimated but the analysed field is less smooth than the model field because in many
cases there is only one observation within the search radii. In tests where the distance
radius is reduced to less than about 1200 km there are a number of gridpoints which have
no observations within their corresponding search radii (note that when observations are
simulated at actual TIROS-N orbit locations—see below—this missing observation problem
occurs with distance radii of less than 1000 km). Changing the time radius has little effect
on the r.m.s. error. Tests made using a 5000 km distance radius showed that changing
the time radius from 12 to 18 hours has negligible effect on the r.m.s. error, and that this error increases only slightly when the time radius is further raised to 24 hours. Similar conclusions may be drawn from tests made using distance radii of 3000 and 2000 km.

Other tests were made using fields which contained higher wavenumber waves. R.m.s. errors changed with varying time and distance radii in a similar manner to that described above. Despite the fact that the analytical fields only roughly mimic the real atmosphere, one can conclude that, within the range of radii tested, r.m.s. errors decrease when the distance radius is reduced to the limit where some analysis points have no corresponding observations. In addition, despite the high wave speeds of the analytical field, the size of the time radius appears to have little impact on the r.m.s. error.

5. Interpolation of Retrieved Temperatures

In this section we present the results of the tests of the combined retrieval–analysis scheme using the numerical model atmosphere. The simulated observations and retrievals are based on 24 hours of model fields forced from below using the observed geopotential heights for 18 Jan. 1987. The 5 mb temperatures developed at the beginning, middle and end of that period by the model are shown in Fig. 7. During the 24-hour period there is a cooling by about 10 K in a region which extends from a position near 50°N, 300°E to the north pole, indicating that the sudden warming which is developed by the model around this date has passed its peak at this height. In the tests which follow, the retrieval–analysis scheme is used to attempt to retrieve the field at the centre of this period, namely 1200 GMT.

Retrieved temperatures from 18 Jan. 1987 are interpolated using three combinations of search radii. One analysis is made with a time radius of 6 hours and a distance radius of 2000 km, while another uses a time radius of 12 hours and a distance radius of 1000 km. The third analysis uses the search radii employed in the Meterological Office’s operational stratospheric analysis scheme, namely 12 hours and 500 km. When analyses are performed with distance radii of 2000 and 1000 km every gridpoint has at least one observation within its corresponding search radii. However, when a distance radius of 500 km is used there are a number of gridpoints which have no observations within their search radii. These missing values are filled by linearly interpolating the nearest ‘good’ observations from gridpoints east and west of those with no observations, and then smoothing the whole field with a 3-point smoother using weights of the form (0.25, 0.5, 0.25). The temperatures are interpolated in the model grid and to the analysis time of 1200 GMT. For computational reasons, analyses are made only at selected pressure levels. These include some close to the SSU weighting-function peaks (1.5, 5.0 and 15.0 mb), two in between these peaks (3.0 and 10.0 mb), and two outside the range of the weighting function peaks (0.2 and 25.0 mb).

(a) Root-mean-square (r.m.s.) errors

The r.m.s. error of the combined retrieval and analysis process is calculated, under the assumption that the model temperature at 1200 GMT is the ‘true’ temperature. The biases of the estimated temperatures are also calculated. Both r.m.s. errors and biases were calculated for every one of the 7 latitude–season retrieval zones, and also for all latitudes. R.m.s. errors for the analyses made using distance radii of 2000 and 1000 km are shown in Table 4(a); and those for analyses made using distance radii of 1000 and 500 km are shown in Table 4(b).

Inspection of Table 4(a) shows that the r.m.s. errors at 0.2 mb are generally higher than for any other pressure level. There are two possible explanations for this. Firstly,
Figure 7. Temperature field on 18 Jan. 1987 at 5 mb. The northern hemisphere field is plotted on the right and the southern hemisphere field on the left. The contour spacing is 5 K. (a) model field at 0000 GMT; (b) as (a), except the field at 1200 GMT is plotted; (c) as (a), except the field at 2400 GMT is plotted.

<table>
<thead>
<tr>
<th>Zone</th>
<th>0-2</th>
<th>1-5</th>
<th>3-0</th>
<th>5-0</th>
<th>10-0</th>
<th>15-0</th>
<th>25-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5-88/7-59</td>
<td>4-28/4-14</td>
<td>3-36/4-07</td>
<td>5-96/4-37</td>
<td>6-33/2-73</td>
<td>6-78/2-46</td>
<td>7-82/3-36</td>
</tr>
<tr>
<td>2</td>
<td>8-54/7-63</td>
<td>4-73/4-71</td>
<td>4-88/2-87</td>
<td>5-49/2-38</td>
<td>4-94/2-35</td>
<td>5-13/2-76</td>
<td>4-37/3-43</td>
</tr>
<tr>
<td>3</td>
<td>5-27/6-13</td>
<td>4-31/3-88</td>
<td>2-72/2-24</td>
<td>3-16/2-46</td>
<td>2-94/2-21</td>
<td>2-09/1-24</td>
<td>1-98/1-89</td>
</tr>
<tr>
<td>4</td>
<td>4-69/5-16</td>
<td>4-01/3-95</td>
<td>3-21/3-05</td>
<td>1-96/1-99</td>
<td>3-01/3-03</td>
<td>1-74/1-78</td>
<td>2-63/2-52</td>
</tr>
<tr>
<td>5</td>
<td>3-47/3-98</td>
<td>2-08/1-97</td>
<td>1-44/1-58</td>
<td>1-35/1-23</td>
<td>1-31/1-26</td>
<td>0-97/0-83</td>
<td>1-27/1-17</td>
</tr>
<tr>
<td>6</td>
<td>3-74/4-75</td>
<td>3-26/3-82</td>
<td>1-98/2-10</td>
<td>1-49/1-58</td>
<td>0-80/0-91</td>
<td>0-65/0-67</td>
<td>1-22/1-33</td>
</tr>
<tr>
<td>7</td>
<td>3-50/3-56</td>
<td>1-60/1-98</td>
<td>0-89/0-97</td>
<td>0-67/0-59</td>
<td>0-59/0-64</td>
<td>0-50/0-63</td>
<td>2-12/1-78</td>
</tr>
<tr>
<td>Global</td>
<td>5-25/5-62</td>
<td>3-73/3-71</td>
<td>2-98/2-67</td>
<td>3-19/2-28</td>
<td>3-38/2-32</td>
<td>3-11/1-71</td>
<td>3-54/2-41</td>
</tr>
</tbody>
</table>

TABLE 4(b). R.M.S. ERRORS OF THE COMBINED RETRIEVAL AND ANALYSIS IN DEGREES K. ERRORS FOR THE ANALYSIS MADE USING A TIME RADIUS OF 12 HOURS AND A DISTANCE RADIUS OF 1000 km ARE SHOWN TO THE LEFT OF THE SLASH, ERRORS FOR THE ANALYSIS MADE USING A TIME RADIUS OF 12 HOURS AND A DISTANCE RADIUS OF 500 km ARE SHOWN TO THE RIGHT OF THE SLASH

<table>
<thead>
<tr>
<th>Zone</th>
<th>0-2</th>
<th>1-5</th>
<th>3-0</th>
<th>5-0</th>
<th>10-0</th>
<th>15-0</th>
<th>25-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7-59/8-17</td>
<td>4-14/4-26</td>
<td>4-07/4-74</td>
<td>4-37/4-66</td>
<td>2-73/2-76</td>
<td>2-46/1-82</td>
<td>3-36/2-65</td>
</tr>
<tr>
<td>2</td>
<td>7-63/7-32</td>
<td>4-71/4-80</td>
<td>2-87/2-73</td>
<td>2-38/2-02</td>
<td>2-35/2-48</td>
<td>2-76/2-39</td>
<td>3-43/3-30</td>
</tr>
<tr>
<td>3</td>
<td>6-13/6-45</td>
<td>3-88/3-83</td>
<td>2-24/2-29</td>
<td>2-46/2-37</td>
<td>2-21/2-17</td>
<td>1-24/1-14</td>
<td>1-89/1-96</td>
</tr>
<tr>
<td>4</td>
<td>5-16/5-20</td>
<td>3-95/3-95</td>
<td>3-05/3-02</td>
<td>1-99/2-08</td>
<td>3-03/3-14</td>
<td>1-78/1-90</td>
<td>2-52/2-46</td>
</tr>
<tr>
<td>5</td>
<td>3-98/4-50</td>
<td>1-97/2-00</td>
<td>1-58/1-67</td>
<td>1-23/1-24</td>
<td>1-26/1-28</td>
<td>0-83/0-82</td>
<td>1-17/1-17</td>
</tr>
<tr>
<td>6</td>
<td>4-75/5-06</td>
<td>3-82/3-94</td>
<td>2-10/2-08</td>
<td>1-58/1-58</td>
<td>0-91/0-98</td>
<td>0-67/0-72</td>
<td>1-33/1-43</td>
</tr>
<tr>
<td>7</td>
<td>3-56/3-81</td>
<td>1-98/1-98</td>
<td>0-97/1-21</td>
<td>0-59/0-73</td>
<td>0-64/0-73</td>
<td>0-63/0-73</td>
<td>1-78/1-68</td>
</tr>
<tr>
<td>Global</td>
<td>5-62/5-79</td>
<td>3-71/3-74</td>
<td>2-67/2-81</td>
<td>2-28/2-33</td>
<td>2-32/2-39</td>
<td>1-71/1-59</td>
<td>2-41/2-26</td>
</tr>
</tbody>
</table>

0.2 mb is far away from a weighting-function peak so one would expect the retrieval at this level to be poor. Secondly, the 0.2 mb field (unlike the fields at the other 6 levels) has considerable small-scale structure that the interpolation scheme might have difficulty resolving. R.m.s. errors at 0.2 mb are higher using a 1000 km distance radius than when using a 2000 km radius. This is the opposite of the results at other levels and tends to suggest that the high error is due to poor retrievals.

The model fields at 1.5, 3.0, 5.0, 10.0, 15.0 and 25.0 mb are quite similar to each other. With a distance radius of 2000 km, global r.m.s. errors range between 2.98 and 3.73 K. Reducing the distance radius to 1000 km reduces the global r.m.s. error to between 1.71 and 3.71 K. Only at 1.5 mb is the global r.m.s. error not noticeably reduced when the distance radius is reduced to 1000 km. An inspection of the errors in each zone reveals that when the distance radius is reduced from 2000 km to 1000 km, the largest fall in r.m.s. error generally occurs in the region of the sudden warming. Away from the sudden warming r.m.s. errors change little when the distance radius is changed. This means that we require a small distance radius when analysing the high spatial variability within a sudden warming, but that a larger distance radius is adequate to analyse the field elsewhere.

Table 4(b) reveals that changing the distance radius from 1000 to 500 km and interpolating between gridpoints for those with no nearby observations produces little change in the r.m.s. error, and maps of fields analysed with a 500 km distance radius (not shown) are little different from corresponding fields analysed using a 1000 km distance radius. This suggests that the use of a 500 km distance radius is unnecessary, as
comparably good analyses can be obtained using a distance radius of 1000 km, without problems caused by missing observations.

(b) **Comparison of maps at 5 mb**

To study the performance of the interpolation scheme more closely, we compare model and analysed fields at 5 mb. The analysed fields used in the comparison are those estimated using distance radii of 2000 and 1000 km. The results at 5 mb provide a good example of the improvement in the analysis caused by reducing the distance radius from 2000 to 1000 km. The 5 mb model field (Fig. 7(b)) shows there is strong wavenumber-1 activity in the middle and upper latitudes of the northern hemisphere. There is a major high situated between 40° and 90°N and 90°E and 90°W with two peaks at 70°W, 70°N (266 K) and at 80°N, 30°E (262 K). There is also a large low situated between 50° and 70°N, and 30°E and 30°W, having a minimum value of 216 K. The high and low are separated by a region of strong temperature gradients. Outside the middle and upper latitudes in the northern hemisphere the flow is generally zonal.

A comparison of two analyses of the 5 mb field confirms that use of a smaller distance radius reduces the r.m.s. error. Figure 8(a) shows the 5 mb field analysed with a time radius of 6 hours and a distance radius of 2000 km; model field features have been reproduced adequately at the correct geographical location. However Fig. 8(b) (which shows the difference between model and analysed fields) reveals that the major high in the northern hemisphere has been underestimated by up to 14 K. In addition, the major low has been overestimated by 10 K and in consequence the large gradient between the major high and low has been poorly estimated. The interpolation made with a distance radius 1000 km (Fig. 8(c)) produces a much better estimate of this large temperature gradient. In addition, the estimated magnitudes of the major high and low are improved. Figure 8(d) shows that the underestimate of the major high has been reduced to between 6 and 8 K and that the major low has been estimated to within 3 K of the model value. Again comparison (not shown) between maps of analyses made with a 1000 km distance radius and of analyses made with a 500 km distance radius with gap-filling reveals that they are very similar.

We conclude that there is little advantage in the current operational distance radius of 500 km, which requires gap-filling where orbits fall far from gridpoints, over the 1000 km radius, which requires gap-filling only where observations are missing because of calibration sequences or drop-outs. On the other hand, the performance would be degraded by increasing the radius beyond 1000 km.

6. **Conclusions**

We have demonstrated that, given an accurate forward model, the stratospheric temperature field can be satisfactorily (generally within 2-5 K in the stratosphere) estimated using a retrieval–analysis scheme which uses the techniques of multiple linear regression and linear time–space interpolation. Although the scheme was tested using only one model field, this conclusion is robust, since the model field contains a sudden warming, and so provides a stringent test of both retrieval and analysis schemes.

Temperature retrieval errors were highest within the region of a sudden warming. This is because the vertical temperature structure in the sudden warming was too small to be ‘observed’ by the satellite instrument, and also because the data used to calculate the regression coefficients were inevitably not representative of sudden warming conditions. The former errors, due to small-scale vertical structure, are inherent in the method of observation rather than the method of temperature retrieval. In contrast, the
Figure 8. Temperature field for 1200 cm at 18 Jan. 1987 at 5 mb. The northern hemisphere field is plotted on the right and the southern hemisphere field on the left. The contour spacing is 5 K. Dashed contours represent negative values. (a) analysed values obtained using a time radius of 6 hours and a distance radius of 2000 km; (b) difference between model field (Fig. 7(b)) and field in (a); (c) as (a), except the analysis is made with a time radius of 12 hours and a distance radius of 1000 km; (d) difference between model field (Fig. 7(b)) and field in (c).
latter errors are caused by dissimilarities between the sonde data-set and sudden warming conditions. These errors can be described as 'first guess errors' since the mean of the data-set can be thought of as a first guess estimate of the retrieved temperature. Ideally one would wish to reduce such errors by using a first guess based on a model forecast rather than on climatology, and future research may involve the development of a retrieval scheme which uses for the a priori information the output at the previous analysis time from a numerical weather prediction model. At present such models do not produce operational forecasts for the stratosphere, but it is expected that a number of such models will be extended up to the stratosphere in the near future.

The time-space interpolation scheme was tested using both idealized radiances fields and a stratospheric temperature field which included a sudden warming. Most features of these fields were satisfactorily reproduced, even those of high temporal variability such as a 5-day planetary wave, or of high spatial variability such as exist in a sudden warming. It was found that reducing the size of distance radius in the scheme decreased analysis errors, but that varying the size of the time radius altered the error values only very slightly. A future refinement of the tests would involve the simulation of atmospheric tides (which are not calculated in the model) and the examination of possible biases in the analysis when data from certain satellite orbits are missing. Further research should also involve the comparison of the time-space interpolation scheme with other analysis schemes. As this time-space interpolation scheme produces essentially a weighted average of a set of observations, maximum values in the field are always underestimated and minimum values are always overestimated. Other analysis methods may not have this problem, and thus a comparison with other techniques, such as sequential estimation (Rodgers 1976) would be valuable. However, the time-space interpolation method used in this paper has the advantage that it is easy to understand, is computationally efficient and, as has been demonstrated, produces good results. Moreover it is versatile, since it can be used to analyse any sort of satellite data (e.g. composition measurements, temperature, radiance) easily and effectively without the need to estimate the noise characteristics of the observations.

ACKNOWLEDGEMENTS

DRJ was supported by the Natural Environment Research Council during the course of this work. We would like to thank the Meteorological Office for their co-operation, especially Alan O'Neill and Mike Bailey.
REFERENCES


Schwalb, A. 1978 The TIROS-N/NOAA A-G satellite series. NOAA Technical Memo. NES95
