The information content of remote measurements of atmospheric temperature by satellite infra-red radiometry and optimum radiometer configurations

By G. PECKHAM
Heriot-Watt University, Edinburgh

(Manuscript received 18 December 1973; in revised form 11 March 1974; Communicated by Professor S. D. Smith)

SUMMARY

Infra-red radiometry from an earth satellite is becoming an important technique for the determination of the three-dimensional temperature structure of the atmosphere on a global scale. In this paper a quantitative expression for the information content of the data from a temperature sounding radiometer is derived. A simple model for the 15 μm CO₂ absorption band is used to find radiometer configurations which give maximum information. The spectral bandwidths corresponding to these optimum configurations are considerably broader than those specified in the past for radiometers of this type. The technique of selective chopping in which the radiometer contains absorbing cells filled with CO₂ is shown to make a significant increase in the information content.

1. INTRODUCTION

Continuous measurements of atmospheric temperatures are now being made by satellite-borne infra-red radiometers. Such measurements are used by meteorologists to extend their knowledge of the temperature field to regions not covered by more conventional meteorological stations and to heights greater than those reached by normal radiosondes.

The principles of such measurements have been described in many papers (e.g. Kaplan 1959, Houghton and Smith 1970, Houghton and Taylor 1973). At any frequency in the infra-red where an atmospheric constituent possesses strong absorption, the radiation intensity leaving the top of the atmosphere is a function of the distribution of the emitting gas and of the vertical temperature profile of the atmosphere. For temperature sounding, the emitting constituent should be substantially uniformly mixed. The ν₂ band of carbon dioxide at 15 μm has been most often used although some success has also been achieved with microwave radiometers operating at a wavelength of 5 mm where there is an oxygen absorption band. We shall consider only infra-red radiometers.

Such radiometers are detector-noise limited so that the signal-to-noise ratio is proportional to the width of the band of infra-red frequencies to which the instrument is sensitive. However, if this band is made too broad, the atmospheric absorption coefficient will vary significantly within the band resulting in a degraded vertical resolution in determining the atmospheric temperature profile, so that a compromise must be made. In this paper we propose that this compromise should be made in such a way that the 'information' from the experiment is maximum where 'information' is used in a quantitative technical sense as in communication theory. In order to simplify the calculation, we make use of an approximate form for the dependence of atmospheric transmission on frequency.

2. THE THEORY OF ATMOSPHERIC TEMPERATURE SOUNDING FROM SATELLITES

Let m be the mass of absorber in a vertical column of unit cross-sectional area above some particular height in the atmosphere. At a frequency ν where the absorption coefficient
is $k_s$ reciprocal mass units, the intensity of radiation emitted per unit solid angle in a vertical direction by a thin horizontal slice of this column of mass $dm$ at temperature $T$ will be

$$k_s B_s(T) dm$$

where $B_s(T)$ is the Planck function. Of this radiation a fraction $\tau_v$ will reach the top of the atmosphere where

$$\tau_v = \exp \left\{ - \int_0^m k_s dm \right\}$$

(1)

The intensity at the top of the atmosphere is

$$I_v = \int_0^M \tau_v k_s B_s(T) dm$$

where $M$ is the total mass of absorber in a vertical column of unit cross-sectional area. In terms of an altitude variable $y = -\ln(p)$ where $p$ is the pressure in atmospheres, and integrating over the frequency pass band of the radiometer

$$I = \int_{\Delta\nu} \int_{y=0}^\infty B_s(T) \left( \frac{d\tau_v}{dy} \right) dy \: dv \simeq \int_{y=0}^\infty B_{v_0}(T) K(y) dy$$

(2)

where it is assumed that $B_s(T)$ is approximately constant over the passband of the radiometer centred on $v = v_0$. The intensity $I$ is the weighted average of black body emission from the atmosphere, the weighting function describing the heights from which radiation is emitted being

$$K(y) = \int_{\Delta\nu} \left( \frac{d\tau_v}{dy} \right) dv$$

(3)

$K(y)$ is a peaked function, the height of the peak depending on the absorption coefficient. A radiometer consists of a number of channels each sensitive to different bands of frequencies and thus possessing weighting functions peaking at different heights. These functions must be accurately known to interpret the observed radiances in terms of a vertical temperature profile. They may be found from tabulated line positions and strengths by numerical integration over the spectral band to which the radiometer is sensitive. Such calculations may be combined with laboratory measurements of the transmission of carbon dioxide as described by Barnett et al. (1972).

However, for the present purpose, an approximate form for $K(y)$ will be derived from the Elsasser model in which absorption lines are assumed to be equally spaced. This model is appropriate if the spectral bandwidth of the radiometer is sufficiently wide to include a number of absorption lines. Details are given in Appendix A. In order to examine the effects of increasing the spectral bandwidth of the radiometer, the strength of the lines as a function of frequency must be known. A narrow band weighting function (Eq. (A3)) peaks at a height in the atmosphere such that the corresponding pressure, $p_0$, is inversely proportional to the square root of the absorption line strength. $y_0 = -\ln(p_0)$ is shown as a function of frequency $v$ in Fig. 1 (data supplied by Rodgers). Although the detailed form of this dependence would have to be considered in designing a practical radiometer, since we are here only interested in describing general trends, $-\ln(p_0)$ has been assumed to be a linear function of $v$. In this case $K(y)$ may be expressed in terms of the error function (Eq. (A6)).

This equation models the essential properties necessary if an optimum configuration is to be found. If the spectral bandwidth is too small, the measurement becomes inaccurate in the presence of detector noise, while if the bandwidth is too great, the weighting functions $K$
become broader and detail in the vertical temperature profile can no longer be determined. A further advantage of this simplified form is that the search for the optimum will not be confused by the presence of local optima caused by detail in the CO₂ absorption spectrum.

The techniques described so far are appropriate for a radiometer designed to measure tropospheric temperatures. The height at which a weighting function peaks depends on the mean absorption over the spectral band accepted by the radiometer. To achieve higher weighting functions it is necessary to improve the spectral resolution so that radiation in the line centres is accepted while that in the wings is rejected. Smith and Pidgeon (1964) proposed to achieve this resolution by selective chopping with a cell of CO₂. This cell modulates radiation only where it absorbs, i.e. with appropriate pressures only in the line centres. This proposal has formed the basis of the Selective Chopper Radiometer (SCR) on Nimbus 4 (Abel et al. 1970) and Nimbus 5 (Ellis et al. 1973) and also the pressure modulated radiometer (PMR) on Nimbus F (Curtis et al. 1974).

The principle of selective chopping was implemented in different ways in these instru-
ments. In the Nimbus 4 SCR, radiation is switched at the modulation frequency between two fixed cells containing different amounts of CO₂. In the Nimbus 5 instrument measurements are made sequentially as cells containing various pressures of CO₂ are moved into the optical path, differing in this case being carried out on the ground after telemetering the information. In the Nimbus F PMR, the gas pressure in a fixed cell is varied at the modulation frequency. In this paper we allow each channel of the radiometer to contain an absorption cell filled with CO₂, the pressure being a variable parameter, in the search for an optimum configuration. Although this is closest to the Nimbus 5 SCR arrangement, the Nimbus 4 and Nimbus F arrangements can also be described in these terms.

The weighting functions (A6) include the effects of such an absorption cell. The expression for the PMR weighting functions derived by Taylor et al. (1972) may be derived from Eq. (A6) as a high altitude approximation.

The weighting functions described in Appendix A are not normalized to unit area, but include the loss in sensitivity due to the presence of a cell of absorbing gas and also the spectral bandwidth dependence of the sensitivity.

3. THE RETRIEVAL PROBLEM AND INFORMATION CONTENT OF THE MEASUREMENT

In a discrete form Eq. (2) may be written

\[ I_i = \sum_{j=1}^{n} K_{ij}B_j \]  \hspace{1cm} (4)

where \( B_j, j = 1, 2, \ldots, n \), describes an atmospheric Planck function profile and \( I_i \) describes the radiation intensity measured by the \( i \)th channel of the instrument \( (i = 1, 2, \ldots, m) \). The matrix \( K_{ij} \) is related to the \( i \)th weighting function \( K_i(y) \) by

\[ K_{ij} = hK_i(jh) \]

where \( h \) is the tabulation interval in \( y \). From now on we shall use a matrix notation so that Eq. (4) is written

\[ \mathbf{I} = \mathbf{KB} \]  \hspace{1cm} (5)

Eq. (5) allows the radiation intensities \( \mathbf{I} \) to be calculated from the atmospheric profile \( \mathbf{B} \). The inverse problem of retrieving \( \mathbf{B} \) from the measured \( \mathbf{I} \) is more difficult. Typically, there might be 6 channels in a radiometer so that \( \mathbf{I} \) has 6 components, but \( \mathbf{B} \) might have as many as 50 components depending on the tabulation interval. Eq. (5) is then very much underdetermined and more information must be supplied. Rodgers (1970) has suggested using the known statistical properties of the atmosphere to choose the most probable atmospheric profile consistent with the observed intensities. Maximum probability estimates of this type are discussed by Liebelt (1967, chapter 5). If \( \mathbf{C} \) is the covariance matrix of atmospheric profiles and \( \mathbf{E} \) the covariance matrix of the measured intensities, for Gaussian distributions, the most probable estimate of \( \mathbf{B} \) is \( \mathbf{B}^* \) where

\[ \mathbf{B}^* - \overline{\mathbf{B}} = \mathbf{CK}^T(\mathbf{E} + \mathbf{KCK}^T)^{-1}(\mathbf{I} - \overline{\mathbf{I}}) \]  \hspace{1cm} (6)

\( \overline{\mathbf{B}} \) is the mean profile and \( \overline{\mathbf{I}} = \mathbf{KB} \).

The covariance of the estimate \( \mathbf{B}^* \) is \( \mathbf{C}^* \) where

\[ \mathbf{C}^{*-1} = \mathbf{C}^{-1} + \mathbf{K}^T\mathbf{E}^{-1}\mathbf{K} \]  \hspace{1cm} (7)

The fact that the measured intensities \( \mathbf{I} \) have provided information about the atmospheric profile is reflected in the fact that \( \mathbf{C}^* \) will be smaller than \( \mathbf{C} \) assuming some suitable measure of the magnitude of a covariance matrix is available.
Information theory (Wiener 1948, Feinstein 1958) provides a quantitative expression for the information content $q$ of a measurement of a quantity $x$ in terms of the prior probability distribution $p(x)$ and the distribution $p^*(x)$ of the estimate of $x$ after the measurement

$$q = \int p^*(x) \log \{p^*(x)\} dx - \int p(x) \log \{p(x)\} dx.$$  

(8)

The logarithm may have base $e$ giving a 'natural' unit of information or base 2 giving information in 'bits' or the equivalent number of binary choices. We shall use base 2. For a multivariate Gaussian distribution, Eq. (8) becomes

$$q = \frac{1}{2} \log_2 |C| - \log_2 |C^*|$$  

(9)

where $|C|$ is the determinant of $C$.

From Eq. (7), the information content of the measurement is

$$q = \frac{1}{2} \log_2 \{|U + CK^TE^{-1}K|\}$$  

(10)

where $U$ is the unit matrix.

An alternative expression may be obtained by considering the covariance matrix $D$ of the intensity profiles $I$. Before measurement

$$D = KCK^T$$  

(11)

After the measurement

$$D^*^{-1} = D^{-1} + E^{-1}$$  

(12)

and

$$q = \frac{1}{2} \log_2 \{|D| - \log_2 |D^*|\}$$

or

$$q = \frac{1}{2} \log_2 \{|U + KCK^TE^{-1}|\}$$  

(13)

That expressions (10) and (13) are equivalent may be seen as follows. If $A$ and $B$ are rectangular matrices of dimension $n$ by $m$ and $m$ by $n$ respectively then the $n$ eigenvalues of $(AB)$ are identical to the $m$ eigenvalues of $(BA)$ but for the addition of $n-m$ zeros where we assume $m < n$. If we identify $CK^TE^{-1}$ with $A$ and $K$ with $B$ the determinants in expressions (10) and (13) are both equal to

$$(1 + \lambda_1)(1 + \lambda_2) \ldots (1 + \lambda_m)$$

where $\lambda_1, \lambda_2 \ldots \lambda_m$ are the $m$ eigenvalues of $BA = KCK^TE^{-1}$.

Expression (13) is preferred for numerical calculations as the dimension of the determinant is smaller.

4. Optimum radiometer configurations

Conrath (1972) considered the effects of choice of spectral intervals and number of channels on the vertical resolution of a temperature sounding instrument. Following Backus and Gilbert (1967, 1968, 1970) he defines a vertical resolution in terms of an averaging kernel which in our notation is the matrix $A$ obtained by substituting Eq. (5) in Eq. (6) which may then be written

$$(B^* - \overline{B}) = A(B - \overline{B})$$  

(14)

Because $A$ has non-zero off-diagonal elements, the estimate of the radiance $B_i^*$ at a particular height $i$ is affected by the true radiance $B_j$ over a range of heights $j$. Backus and Gilbert's resolution $\omega_i$ measures the width of $A_{ij}$ considered as a function of $j$. Conrath considers more general inversion procedures than the maximum probability estimate of Eq. (6). He tries to choose both spectral intervals for the various channels of the radiometer and inversion procedures so as to minimize the resolution $\omega_i$. A problem arises in that as $\omega_i$ is made smaller
so the variance of the estimate $B^*_\nu$ is increased and some arbitrary compromise has to be made. If this criterion is to be used to design a radiometer another choice that has to be made is the desired variation of $\omega_i$ with height. Presumably $\omega_i$ would be chosen to match the scale of atmospheric structure. Note that such information is present in the atmospheric covariance matrix of the statistical theory presented here.

Information theory provides an alternative method for choosing an optimum radiometer design which avoids the arbitrary compromises in Conrath’s approach. It is proposed that the optimum radiometer configuration is that which maximizes the information $q$ defined by Eq. (13). How this procedure works in practise may be seen in the case of a 2-layer atmosphere each layer of which is supposed to emit radiation in a different spectral band. We consider three configurations of a 2-channel radiometer: configuration 1 in which the channels are sensitive to radiation from different layers; configuration 2 in which one channel is sensitive to both layers and the other to one layer only; and configuration 3 in which both channels are sensitive to both layers. The matrices $K$ for these three configurations are

$K_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad K_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad K_3 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

(We assume that the instrument is of filter type rather than dispersive so that the two channels may look at the same spectral interval.) The channels are supposed to be independent and have noise variance $\varepsilon$ and we assume an atmospheric covariance matrix in which each layer has the same variance $\sigma$.

$$
\mathbf{E} = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \sigma & \mu \\ \mu & \sigma \end{pmatrix}
$$

With these assumptions Eq. (13) may be evaluated for the three configurations. The configuration which gives maximum information depends on the covariance $\mu$ and instrumental noise $\varepsilon$. The areas in $\varepsilon - \mu$ space within which particular configurations are optimum are shown in Fig. 2. These results agree with intuitive ideas. If the layers are highly correlated ($\mu \sim \sigma$) there is no point in measuring them independently and one might as well increase the signal-to-noise by increasing the bandwidth of each channel to include both spectral intervals (configuration 3). If the layers are anticorrelated ($\mu \sim -\sigma$) no information would be obtained if the channels were sensitive to both layers and one must revert to configuration 1 in spite of the decreased signal-to-noise. In between is a region where configuration 2 provides most information.

![Figure 2. Values of the instrument noise and of the atmospheric covariance for which different configurations of a two-channel radiometer give optimum performance.](image-url)
We now apply these ideas to more realistic situations using measured atmospheric covariances and the simplified weighting functions described in Appendix A.

5. Atmospheric Covariance

An atmospheric covariance matrix appropriate to temperate latitudes was supplied by Rodgers. The dimension of the matrix was 50 and the tabulation interval was 0.2 in y. The matrix included a component for ground temperature variation. So far we have ignored radiation from the ground which may in fact form a significant fraction of the signal for channels with low weighting functions. We shall assume that the ground radiation is accurately known from other measurements (e.g. a radiometer operating in the 10 μm atmospheric window) so that it may be subtracted from the signals. The covariance matrix is modified by this knowledge of ground temperature.

Consideration of the form of Eq. (7) for an accurate ground viewing channel leads to the result that the row and column of \( C^{-1} \) referring to ground values must be deleted to obtain the modified inverse matrix. In terms of the matrix itself

\[
C'_{ij} = C_{ij} - C_{0i}C_{0j}/C_{00}
\]  

(15)

where \( C'_{ij} \) is an element of the modified matrix \( i, j = 1, 2, \ldots, n \), \( C_{ij} \) the original matrix and the suffix 0 refers to ground values.

6. Multi-channel Radiometers

In order to calculate the information content, \( q \), from expression (13), we must make some assumptions about the covariance matrix \( E \) of the measured intensities and the way this matrix varies as the instrument configuration is changed. Firstly, we assume that the different channels of the radiometer are independent so that \( E \) becomes diagonal and secondly that the variance of each channel expressed in radiance units is identical. (Note that variations due to spectral bandwidth and internal absorption cells are already included in \( K \).) This seems reasonable if the channels are detector-noise limited and have similar optical losses.

We next assume that the variance is proportional to \( m \) where \( m \) is the number of channels. This would be appropriate for a time shared or sequential filter radiometer where the time available for each channel is proportional to \( 1/m \), or for a parallel radiometer of fixed size where the area of the entrance aperture for each channel is proportional to \( 1/m \) if the detector noise variance is also proportional to area. This assumption would not be appropriate for radiometers making their spectral selection by gratings or by interferometry. However, operational instruments are of the filter type (e.g. the Vertical Temperature Profile Radiometers on NOAA satellites).

With these assumptions

\[
E = maU
\]  

(16)

where \( U \) is the unit matrix. It is not essential to assume that each channel has equal variance and in fact it might be advantageous to spend more time measuring one channel so reducing its variance at the expense of the other channels, but this is not considered further here.

The value for \( a \) has been taken from the practical performance of the selective chopper radiometer of Nimbus 4. The signal integration time is an important factor in determining the noise. Ideally this would be chosen to match the horizontal scale of atmospheric structure for a given satellite velocity. However, we have chosen an arbitrary 20 seconds which is the
time taken for the satellite to cross the Nimbus 4 SCR field of view (120 km). Normalized to a 1 cm\(^{-1}\) spectral bandwidth

\[ g^\ddagger = 0.26 \text{ mW m}^{-2} \text{ster}^{-1}/\text{cm}^{-1}. \]

Calculations were repeated with \( g^\ddagger \) equal to one tenth of this value. Various schemes for defining spectral intervals were considered. The most obvious is to choose frequencies defining a spectral interval \( v_1 < v < v_2 \) and since the linear approximation in Fig. 1 is symmetrical about the \( Q \) branch, a second spectral interval \( v'_2 < v < v'_1 \) may be added where \( v' \) is the reflection of \( v \) about the \( Q \) branch. This scheme requires three parameters per channel to be optimised i.e. \( v_1, v_2 \) and \( p_c \), the pressure of gas in the absorption cell. A simpler scheme allowed all frequencies in the band \( v_1 < v < v'_1 \) and so required only two parameters \( v_1 \) and \( p_c \) (see Fig. 1). Trial calculations showed that the information \( q \) provided by an optimised radiometer of the latter type was only marginally less than that provided by the first type. Since the work in finding an optimum increases rapidly with the number of variable parameters, the two parameter scheme was used in most cases.

Powell's (1964) programme for finding the minimum of a function of many variables was used to find values of the parameters which gave a maximum for the determinant in Eq. (13). Calculations were made for 2, 4, 6 and 8-channel radiometers for two values of the instrument noise, \( \varepsilon \). Since radiation from the ground cannot in practice be allowed to dominate the signal, \( v_1 \) was restricted to a band such that \( p_1 < 1 \) (Eq. (A6)). A plot of the information \( q \) from the optimized radiometer against the number of channels is shown in Figure 3.

7. DISCUSSION

It can be seen from Fig. 3 that the information content increases rapidly at first as the number of channels is increased, but eventually saturates so that there is little advantage in any further increase in the number of channels. For the high noise case this occurs at about \( m = 6 \). For the low noise case saturation does not set in until \( m > 8 \). This behaviour is in agreement with Conrath's (1972) conclusions.

The optimum configuration in the high-noise case was used as a starting value in the search for the optimum in the low-noise case. It was found that the low-noise optimum configuration was then almost identical with that for high noise. Thus the optimum configuration was not sensitive to the particular value chosen for the instrument noise although the information content was considerably increased by a decrease in instrument noise.

As a check, quite different starting values were chosen and the high-noise case for

![Figure 3](image-url)
\( m = 4 \) was repeated. A different set of parameters was obtained although the information content of the new optimum arrangement agreed closely with the previous value. The spectral responses of the 4 channels in these two cases are shown in Fig. 4 which also includes a configuration with 3 parameters per channel again giving only a slight change in the information \( q \). It seems probable that there are many configurations which will give values for \( q \) close to the optimum. However, the value of \( q \) itself for an optimum configuration is well defined.

The spectral bandwidths shown in Fig. 4 are considerably greater than those usually employed in temperature sounding radiometers (e.g. 10 cm\(^{-1} \) for SCR, but PMR does employ broad band filters). The information from an instrument similar to Nimbus 4 SCR, but with double passbands (one on each side of the \( Q \) branch), was calculated and found to be well below the optimum for a 7-channel radiometer (see Fig. 3. The selective chopping channels in this instrument were treated as 2 channels to conform to the scheme adopted in this paper).

\[
P^*_c = (p_1^2 - p_2^2)^t.
\]

Figure 4. Filter passbands for 3 optimum configurations of a 4-channel radiometer giving similar information \( q \). The shaded areas represent the fraction of radiation absorbed in the CO\(_2\) cells and \( p^*_c = (p_1^2 - p_2^2)^t \).

In order to assess the value of selective chopping in increasing the information \( q \), optimum configurations were found with the cell pressure in the 3-parameter scheme constrained to be zero. The values of \( q \) are shown in Fig. 3. These values do fall significantly below the curve for an optimum configuration, particularly for the 6 and 8-channel cases.

The weighting functions are usually plotted to show the range of heights within which the atmospheric temperature can be reliably retrieved and to give some idea of the vertical resolution of the instrument. Where weighting functions overlap substantially, as in the case of selective chopping, differences may be shown to give a clearer display of the resolution. In the present case, the weighting functions for the optimum configurations are very broad and overlap each other considerably so that little impression of the vertical resolution is conveyed by a direct plot, nor is it apparent just which pairs of functions should be subtracted to show the improved resolution obtained by differencing. Accordingly methods
were sought to define linear combinations of the weighting functions which would convey this information.

A set of weighting functions $K'$ which are linear combinations of the original functions $K$ may be defined by

$$K' = T^T K.$$ 

A radiometer with channels responding to atmospheric radiation with the combined weighting function would record signals

$$I' = T^T I$$

(unequal the original function, some elements of $K'$ may be negative, as may elements of $I'$).

The transforming matrix $T$ may be uniquely defined by insisting that (1) the combined signals $I'$ are statistically independent and have variance $\varepsilon$ and (2) the combined weighting functions $K'$ are orthogonal. Condition (1) implies that $T$ is an orthogonal matrix and (2) that

$$KK'^T = \Lambda$$

where $\Lambda$ is a diagonal matrix. The columns of $T$ are then the eigenvectors of $(KK^T)$ and the elements of $\Lambda$ the eigenvalues.

The orthogonalized weighting functions for an optimized 4-channel radiometer are shown in Fig. 5(a), those for an optimized 6-channel instrument in 5(b), for the Nimbus 4 SCR in 5(c) and for an optimized configuration with no CO$_2$ absorption cells in Fig. 5(d). The plotted weighting functions $K''$ have been normalized so that

$$h \sum_i (K''_{ij})^2 = 1$$

($h$ is the tabulation interval in $y$)

or

$$K''_{ij} = K_{ij}(h\lambda_i)^{-1/2}$$

A detectivity $D_1 = (\lambda_i/h\varepsilon)^{1/2}$ may be defined so that if the functions in Fig. 5 represent radiance profiles in the atmosphere, the unit for the horizontal scale now being mW m$^{-2}$ ster$^{-1}$/cm$^{-2}$, the signal-to-noise ratio for the radiance measurement $I'_i$ will be $D_1$. The detectivities $D_1$ are marked on the curves in Fig. 5. The wide optical bandwidths of the optimized radiometers give much larger values for their detectivities than for the Nimbus 4 SCR configuration.

The shapes of the orthogonalized weighting functions may also have some significance. The optimized configurations give more regular shapes with equal intervals between maxima and minima. We may suppose that the $m$-channel radiometer is capable of determining the first $m$ coefficients in an expansion of the atmospheric radiance profile in orthogonal functions the first $m$ of which are chosen to coincide with the functions $K$. This problem of representing a function by a truncated series arises in numerical analysis and functions such as the Tschebyscheff polynomials which are efficient for this purpose share the property of oscillating between regularly spaced maxima and minima.

The technique of selective chopping was devised to give information about the high atmosphere. To illustrate this, in Table 1 are listed the variances of the average radiances over slabs of atmosphere thickness 2 in the log-pressure scale both before and after measurement with the instrument configurations used for Fig. 5(b), (c) and (d).

At very high levels, (8–10), the measurement provides little direct information about the atmospheric profile, the slight reduction in variance arising from covariance with lower levels. At high levels, (6–8, 4–6), the configurations containing CO$_2$ (b and c) both give a smaller variance than the configuration without CO$_2$ absorption cells (d). The optimum configuration (b) gives the smallest variance at all levels except the lowest.
Figure 5. Orthogonalized weighting functions (the value of the detectivity, $D_n$, is marked on each curve).
TABLE 1. Atmospheric variances (mW m\(^{-2}\) ster\(^{-1}\)/cm\(^{-1}\))^2

<table>
<thead>
<tr>
<th>level (y = -\ln(p))</th>
<th>prior variance</th>
<th>optimum 6-channel (b)</th>
<th>Nimbus 4 (c)</th>
<th>Optimum 6-channel without CO(_2) (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-10</td>
<td>107.678</td>
<td>70.338</td>
<td>81.709</td>
<td>87.899</td>
</tr>
<tr>
<td>6-8</td>
<td>40.816</td>
<td>5.775</td>
<td>9.193</td>
<td>11.721</td>
</tr>
<tr>
<td>4-6</td>
<td>21.450</td>
<td>0.538</td>
<td>1.907</td>
<td>4.383</td>
</tr>
<tr>
<td>2-4</td>
<td>4.247</td>
<td>0.034</td>
<td>0.057</td>
<td>0.181</td>
</tr>
<tr>
<td>0-2</td>
<td>6.584</td>
<td>0.024</td>
<td>0.010</td>
<td>0.009</td>
</tr>
</tbody>
</table>

8. Conclusions

A new criterion has been proposed for evaluating the performance of remote atmospheric temperature sounding radiometers. An optimum configuration for a radiometer is that which maximizes the information content \(q\) of the measurement as defined by Eq. (13). Approximate expressions for the atmospheric transmission in the 15 \(\mu m\) CO\(_2\) absorption band have been used to find optimum configurations for filter radiometers incorporating the selective chopping principle. It was found that little extra information resulted from increasing the number of channels above some value depending on the instrument noise. The information \(q\) from an optimised configuration was well defined, but several configurations could give similar values for \(q\). All optimized configurations had broad spectral passbands each occupying a substantial part of the CO\(_2\) absorption band. Narrow passband radiometers such as SCR gave considerably lower values for \(q\). The concept of orthogonal weighting functions can give some insight into these results. Substantial loss of information occurred in the case of a pure filter radiometer containing no CO\(_2\) absorption cells. It was shown that selective chopping is a valuable technique in obtaining information about the high atmosphere. Although the results presented here are based on a fairly gross approximation to the CO\(_2\) absorption spectrum, the general nature of the conclusions is unlikely to change with more accurate calculations. Other factors such as the detailed structure of the CO\(_2\) absorption band and the presence of overlapping ozone bands must be considered in the choice of spectral intervals for a practical radiometer. The difficulties associated with the presence of cloud in the real atmosphere have not been discussed here.

Acknowledgments

My interest in this problem arose from my involvement in the Nimbus 4 and 5 Selective Chopper Radiometer experiments which were jointly developed by groups led by Professor S. D. Smith at Heriot-Watt University and by Dr. J. T. Houghton at Oxford. I am grateful to Dr. C. D. Rodgers for supplying the atmospheric covariance data and the CO\(_2\) absorption data in Fig. 1.

References


Conrath, B. J.


Curtis, P. D., Houghton, J. T., Peskett, G. D. and Rodgers, C. D.


Feinstein, A.


Goody, R. M.


Houghton, J. T. and Smith, S. D.


Houghton, J. T. and Taylor, F. W.


Kaplan, L. D.


Liebelt, P. B.

1967 An introduction to optimal estimation, Addison-Wesley.

Powell, M. J. D.


Rodgers, C. D.


Smith, S. D. and Pidgeon, C. R.


Taylor, F. W., Houghton, J. T., Peskett, G. D., Rodgers, C. D. and Williamson, E. J.


Wiener, N.


APPENDIX A

APPROXIMATE WEIGHTING FUNCTIONS

For a single collision-broadened spectral line of intensity \( S \) centred at frequency \( \nu_0 \), the absorption coefficient is

\[
k_\nu = S \nu_0 \rho \pi^{-1} \left( (\nu - \nu_0)^2 + \gamma_0^2 \rho^2 \right)^{-1}
\]

(A1)

where \( \rho \) is the pressure in atmospheres and \( \gamma_0 \) is the line half-width at one atmosphere pressure. We assume an isothermal atmosphere.

The radiometer accepts a band of frequencies which will include a number of absorption lines. The Elsasser model (see e.g. Goody 1964) in which lines are assumed equally spaced may be used to approximate the CO₂ absorption spectrum in this case. In the strong line limit this model gives for the average transmission over the band

\[
\tau = 1 - \text{erf} \left( \frac{p}{\sqrt{2} \rho_0} \right)
\]

(A2)
where \( p_0^2 = \delta^2/(\pi SM) \), \( \delta \) being the line spacing, \( M \) the total mass of absorber in a vertical column of the atmosphere of unit cross-sectional area and \( \text{erf}(x) = (2/\sqrt{\pi}) \int_0^x e^{-t^2} \, dt \) (the error function).

The weighting function in this case is

\[
K(y) = (2/\pi)^{3/2} (p/p_0) \exp \left\{ -(p^2/2p_0^2) \right\}
\]  
(A3)

peaking at \( p = p_0 \), with a half width 1.80 in \( y \) (Houghton and Smith 1970).

The effects of an absorption cell containing \( \text{CO}_2 \) within the instrument may be taken into account by replacing \( p \) in (A2) by \( p_e \) where

\[
p_e^2 = p^2 + 2(1 + \beta)(M_c/M)p_c
\]

Here \( M_c \) is the mass of gas per unit cross-section of the cell, \( p_c \) the pressure of \( \text{CO}_2 \) within the cell and \( \beta \) the self broadening coefficient allowing for the greater line width in pure \( \text{CO}_2 \) compared with dilute \( \text{CO}_2 \) in the atmosphere.

On differentiating, the weighting function becomes

\[
K(y) = (2/\pi)^{3/2} (p^2/p_e p_0) \exp \left\{ -(p_e^2/2p_0^2) \right\}
\]  
(A5)

To allow for the effects of a broad spectral bandwidth, (A5) is to be integrated over frequency with the assumption that \( -\ln(p_0) = bv + \text{constant} \) (see Fig. 1).

\[
K(y) = (p^2/bp_0^2) \left\{ \text{erf}(p_e/\sqrt{2p_1}) - \text{erf}(p_e/\sqrt{2p_2}) \right\}
\]  
(A6)

The integral is from \( v_1 \) to \( v_2 \) where \( p_1 \) is the value of \( p_0 \) corresponding to \( v_1 \) and \( p_2 \) that corresponding to \( v_2 \).

Expression (A3) has been normalized to unit area, but (A5) and (A6) have been left unnormalized so as to allow for the effects on signal level of including a cell of absorbing gas and of variations in the spectral bandwidth. We shall take \( K(y) \) to be normalized to unit area for a standard instrument of spectral bandwidth \( (v_2 - v_1) \) equal to 1 cm\(^{-1}\) and with no absorbing gas \( (p_e = p) \). The value of \( b \) from Fig. 1 is 0.05/cm\(^{-1}\).

In the linear approximation, the spectrum is assumed to be symmetrical about the \( Q \) branch at 668 cm\(^{-1}\) so that, if necessary, the interval must be divided and (A6) applied to each linear part. In addition, a term of the form (A5) is included to describe the high altitude contribution from the very strongly absorbing \( Q \) branch over a band of width 1.25 cm\(^{-1}\).