# Statistical Information Content of Radiation Measurements used in Indirect Sensing[1]

ED R. WESTWATER AND OTTO NEALL STRAND

*Institute for Telecommunication Sciences and Aeronomy, ESSA, Boulder, Colo.*

(Manuscript received 17 October 1967)

ABSTRACT

The information content of radiation measurements used in inferring profiles is defined as a reduction in uncertainty in the estimation of a profile after the measurements are introduced. The information is shown to depend directly on the kernel of the equation of radiative transfer, the covariance matrix of experimental error, and the covariance matrix of the *a priori* statistical information. Calculations based on the minimum rms inversion method are applied to the indirect probing of the vertical temperature distribution by microwave measurements of oxygen thermal emission. Choice of optimum location of measurements is discussed and comparison of the proposed method with that of Twomey is given.

## 1. Introduction

The information content of radiation measurements has been studied previously by Twomey (1965, 1966) and Mateer (1965). These studies have shown that the information obtainable from indirect soundings is severely limited by the interdependence of the measurements themselves. In a loose sense, the "number of independent pieces of information" was taken to be the number of eigenvalues of a kernel-determined matrix which were greater than some assigned noise level. However, the information content of a signal should be judged by the *new* information that it adds to information already known. In many problems of indirect sensing, statistical information about the profile is known before any radiation measurements are made. This *a priori* knowledge is embodied in the mean and the covariance matrix, both of which can be estimated from past data (usually taken by direct measurements). Such data are currently being used in remote sensing problems in the construction of empirical orthogonal functions (Wark and Fleming, 1966; Alishouse *et al.*, 1967). The use of such data to assess the usefulness of radiation measurements in reducing the statistical variance of the unknown function is not well known.

In Section 2, we summarize recently developed inversion techniques (Strand and Westwater, 1968a, b), which lead quite naturally to a definition of information content. Applications of this quantity to remote probing of tropospheric temperature structure by the microwave emission lines of oxygen are given in Section 3. The choice of optimum frequencies is discussed, and our method of choice is compared with that of Twomey (1966).

[1] Research supported by the U. S. Army Electronics Command under Project Order 67-95892.

## 2. The minimum rms inversion method and its relation to information content

Let $f(y)$ be a continuous random function on the interval $[a,b]$. If $K(x,y)$ is a continuous function of $x$ and $y$ and if

$$\int_a^b K(x,y)f(y)dy=g(x), \tag{1}$$

then $g(x)$ is also a continuous random function. In indirect-sensing problems, it is wished to infer $f(y)$ by measuring $g(x)$ at a set of values of $x$, say $x_i$, $i=1, 2, \cdots, n$. Introducing a suitable quadrature approximation to (1) gives the matrix equation

$$\mathbf{Af}=\mathbf{g}, \tag{2}$$

where

$\mathbf{A}=(A_{ij})$, $i=1, 2, \cdots, n$; $j=1, 2, \cdots, m$,
$m$ is the number of quadrature abscissas,
$n$ is the number of observations of $g(x)$,
$A_{ij}=w_jK(x_i,y_j)$,
$y_j$= quadrature abscissas,
$x_i$= values of $x$ for which $g(x)$ is observed,
$w_j$= quadrature weight associated with $y_j$,
$f_j=f(y_j)$,
$g_i=g(x_i)$,
$\mathbf{f}=[f_1f_2\cdots f_m]^T$ is the column vector of unknown functional values (the superscript $T$ denotes matrix transposition), and
$\mathbf{g}=[g_1g_2\cdots g_n]^T$ is the column vector of values of $g(x)$.

Assume that the mean vector $\mathbf{E}(\mathbf{f})=\mathbf{f}_0$ and the covariance matrix, $\mathbf{S}_f\equiv\mathbf{E}[(\mathbf{f}-\mathbf{f}_0)(\mathbf{f}-\mathbf{f}_0)^T]$ are known. [$\mathbf{E}(\ )$ denotes the expected value operator and $\mathbf{S}_v$ denotes the covariance matrix of any vector $\mathbf{v}$]. In practice, $\mathbf{f}_0$ and $\mathbf{S}_f$ may be estimated from the past history of the function $f$. By the linearity E and the propagation rule for covariance matrices (Deutsch, 1965), we have $\mathbf{E}(\mathbf{g})=\mathbf{g}_0$ and $\mathbf{S}_g=\mathbf{AS}_f\mathbf{A}^T$. In the problem of interest here, $\mathbf{g}$ is a vector of measurements subject to error. Thus, one observes

$$\mathbf{g}_e=\mathbf{g}+\boldsymbol{\varepsilon} \tag{3}$$

instead of $\mathbf{g}$. In addition to the above assumptions, we assume 1) the errors $\epsilon_i$ are independent of $\mathbf{f}$, and hence are independent of $\mathbf{g}$; 2) the errors $\epsilon_i$ have a multivariate distribution with mean zero and known covariance $\mathbf{S}_\epsilon$; 3) the errors introduced by the quadrature approximation of Eqs. (1) and (2) are negligible with respect to $\boldsymbol{\varepsilon}$; and, 4) $\mathbf{S}_\epsilon$ and $\mathbf{S}_f$ are both nonsingular with dimensions $n\times n$ and $m\times m$, respectively.

Much of the previous work in the field has emphasized the importance of representing the solution by means of a suitable basis (Wark and Fleming, 1966; Alishouse *et al.*, 1967). This representation was necessary because previous inversion methods solved for a fixed number of parameters (usually small), the number being determined by the degree of independence of the measurements and the measurement noise level. The usual bases chosen in these methods were the eigenvectors of $\mathbf{S}_f$, arranged in decreasing order of eigenvalues. However, it was shown by Strand and Westwater (1968b) that with the minimum rms inversion method the introduction of a basis matrix to represent the solution (other than the identity in $m$ dimensions) is neither necessary nor desirable. The computational difficulties which occur in determining the eigenvectors and eigenvalues of a large matrix may be circumvented by this method. Hence, in the following, the only desired representation of the unknown function will be its values at the $m$ quadrature points. The following will summarize the results of Strand and Westwater (1968b).

Let $\boldsymbol{\eta}=\mathbf{f}-\mathbf{f}_0$, $\mathbf{h}=\mathbf{g}-\mathbf{g}_0=\mathbf{A}(\mathbf{f}-\mathbf{f}_0)=\mathbf{A}\boldsymbol{\eta}$ and let the reduced observed data (with respect to the mean) be $\mathbf{h}_e$, where

$$\mathbf{h}_e=\mathbf{A}\boldsymbol{\eta}+\boldsymbol{\varepsilon}. \tag{4}$$

The estimate of $\boldsymbol{\eta}$, $\hat{\boldsymbol{\eta}}$, is determined as a linear combination of the data

$$\hat{\boldsymbol{\eta}}=\mathbf{Bh}_e, \tag{5}$$

where $\mathbf{B}$ is an $m\times n$ matrix to be determined. The matrix $\mathbf{B}$ is uniquely determined by requiring that the fit to $\boldsymbol{\eta}$ be the best on the average in the mean-square sense, i.e., that $\mathbf{E}\{(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta})^T(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta})\}$ is minimized with respect to $\mathbf{B}$. Here, the expected value is taken over the joint probability distribution of $\mathbf{f}$ and $\boldsymbol{\varepsilon}$. This requirement leads to the optimum linear unbiased estimate of $\boldsymbol{\eta}$ as

$$\hat{\boldsymbol{\eta}}=\mathbf{S}_f\mathbf{A}^T\mathbf{H}^{-1}\mathbf{h}_e, \tag{6a}$$

where

$$\mathbf{H}=\mathbf{S}_\epsilon+\mathbf{AS}_f\mathbf{A}^T,$$

or alternatively

$$\hat{\boldsymbol{\eta}}=\mathbf{X}^{-1}\mathbf{A}^T\mathbf{S}_\epsilon^{-1}\mathbf{h}_e, \tag{6b}$$

where

$$\mathbf{X}=\mathbf{S}_f^{-1}+\mathbf{A}^T\mathbf{S}_\epsilon^{-1}\mathbf{A}.$$

The equivalence of (6a) and (6b) follows from the identity

$$\mathbf{A}^T\mathbf{S}_\epsilon^{-1}\mathbf{H}=\mathbf{XS}_f\mathbf{A}^T, \tag{7}$$

which relates the $n\times n$ matrix $\mathbf{H}$ to the $m\times m$ matrix $\mathbf{X}$. Similar results have been obtained by C. D. Rodgers[2] of Oxford University. The covariance matrix of the solution, $\mathbf{S}_{\hat{\eta}-\eta}$, is given by

$$\mathbf{S}_{\hat{\eta}-\eta}=\mathbf{X}^{-1}. \tag{8}$$

A convenient overall quality criterion is the sum of the diagonal elements of $\mathbf{S}_{\hat{\eta}-\eta}$, i.e., $\mathrm{Tr}(\mathbf{S}_{\hat{\eta}-\eta})$ where Tr( ) denotes the trace. Thus,

$$\mathrm{Tr}\{\mathbf{S}_{\hat{\eta}-\eta}\}=\mathrm{Tr}\mathbf{X}^{-1}. \tag{9}$$

It can be shown that the trace of $\mathbf{S}_{\hat{\eta}-\eta}$ is $m$ times the expected mean-square error.

The statistical information added by the radiation measurements may be determined by comparison of (9) with the trace of $\mathbf{S}_f$. In the absence of any measurements the best estimate of $\mathbf{f}$ is $\mathbf{f}=\mathbf{f}_0(\boldsymbol{\eta}=0)$ with an overall variance $\mathrm{Tr}\mathbf{S}_f$. Adding measurements modifies the estimate to (6), reduces the overall variance to $\mathrm{Tr}\mathbf{X}^{-1}$, and reduces the variance by $\mathrm{Tr}(\mathbf{S}_f-\mathbf{X}^{-1})$. Useful quantities for judging the information are

$$R=\mathrm{Tr}(\mathbf{S}_f-\mathbf{X}^{-1}), \tag{10}$$

$$F=\frac{\mathrm{Tr}(\mathbf{S}_f-\mathbf{X}^{-1})}{\mathrm{Tr}\mathbf{S}_f}, \tag{11}$$

$$U=\left(\frac{\mathrm{Tr}\mathbf{X}^{-1}}{m}\right)^{\frac{1}{2}}, \tag{12}$$

where $R$ is the total reduction in variance, $F$ is the fractional reduction, and $U$ is the standard deviation per point. From this point of view, the maximum reduction in variance yielded by $m$ independent errorless measurements would be $\mathrm{Tr}\mathbf{S}_f$. The information contained in $n(<m)$ independent errorless measurements is easily obtained by letting $\mathbf{S}_\epsilon\to 0$ in (6), (7), and (8) as

$$R=\mathrm{Tr}\{\mathbf{S}_f-\mathbf{S}_f\mathbf{A}^T(\mathbf{AS}_f\mathbf{A}^T)^{-1}\mathbf{AS}_f\}. \tag{13}$$

In any practical remote-sensing experiment, information is limited in two ways: 1) the number of independent measurements that can be taken is small, and 2) measurement errors are always present. The usefulness of the experiment can be judged by determining

[2] Dr. Rodgers has graciously shown us his results for comparison of inversion methods.

$R$, $F$ and $U$ as functions of $\mathbf{S}_f$, $\mathbf{S}_\epsilon$ and $\mathbf{A}$. If $F$ is used, the theoretical maximum achievable information is $F=1$ and occurs when $n=m$ and $\mathbf{S}_\epsilon=0$.

The quantity $R$(or $F$) may also be used to study the optimum placement of measurements. The optimum placement is achieved, given that the number $n$ is fixed, when the function $R$(or $F$) achieves its maximum over the range of $x$. The optimum set derived in this way can differ considerably from the set as chosen by the method of Twomey (1966). This is shown in the following section by a physically-reasonable numerical example. Since $R$ depends on $\mathbf{S}_f$, the optimum set also depends on $\mathbf{S}_f$. This implies that large seasonal or geographic variations could influence the optimum choice.

## 3. Application to indirect temperature sensing

The possibilities of inferring the vertical temperature profile from measurements of microwave thermal emission by oxygen are well known. Meeks and Lilley (1963) discussed the determination of the gross temperature profile from 40 km to sea level by satellite observations; Westwater (1965) discussed determining the tropospheric temperature profile by ground-based techniques. The calculations reported here are intended to illustrate the usefulness of the quality criteria given in Section 2 in planning indirect-sensing experiments and to indicate the accuracy that oxygen thermal emission measurements can yield.

Microwave radiometric measurements are usually expressed as an equivalent emission temperature or effective antenna temperature (Shklovsky, 1960). This effective temperature is both the frequency average over the bandwidth of the receiver and the weighted directional average over the radiometer's antenna pattern of radiation received from all frequencies and directions. The uni-directional, monochromatic radiation from any infinitesimal solid angle is expressed as the brightness temperature $T_b(\nu)$ at frequency $\nu$. Thus,

TABLE 1. Mean temperature and pressure for Denver, February ($h$ is height above surface in km).

| Downward | | | Upward | | |
|---|---|---|---|---|---|
| $s=10-h$ (km) | $\bar{T}(s)$ (°K) | $\bar{P}(s)$ (mb) | $h$ (km) | $\bar{T}(h)$ (°K) | $\bar{P}(h)$ (mb) |
| 0.060 | 217.315 | 201.659 | 0.000 | 267.956 | 831.423 |
| 0.314 | 217.666 | 209.799 | 0.140 | 269.781 | 816.770 |
| 0.759 | 218.282 | 224.897 | 0.416 | 269.474 | 788.533 |
| 1.378 | 219.138 | 247.703 | 0.723 | 268.102 | 758.244 |
| 2.145 | 222.138 | 278.691 | 0.943 | 266.806 | 737.167 |
| 3.029 | 226.598 | 318.933 | 1.094 | 265.866 | 722.982 |
| 3.994 | 232.577 | 368.134 | 1.462 | 263.574 | 689.408 |
| 5.000 | 239.914 | 425.960 | 2.000 | 260.161 | 642.664 |
| 6.006 | 247.069 | 490.838 | 2.538 | 256.857 | 598.547 |
| 6.971 | 253.704 | 560.410 | 2.906 | 254.517 | 569.811 |
| 7.855 | 259.270 | 630.462 | 3.328 | 251.729 | 538.201 |
| 8.622 | 264.095 | 696.903 | 4.615 | 242.720 | 450.398 |
| 9.241 | 267.891 | 754.769 | 6.500 | 228.973 | 342.680 |
| 9.686 | 269.588 | 798.907 | 8.385 | 219.466 | 257.053 |
| 9.940 | 268.740 | 825.097 | 9.672 | 217.686 | 210.282 |

the brightness temperature may be regarded as the antenna temperature of an idealized radiometer which only accepts radiation from a single direction and frequency. The brightness temperature observed looking vertically through an atmosphere of thickness $H$ is given by

$$T_b(\nu)=T_b^0(\nu)\exp[-\tau_\nu(0,H)]+\int_0^H T(h)\alpha_\nu(h)\exp[-\tau_\nu(0,h)]dh, \quad (14)$$

where

$$\tau_\nu(0,h)=\int_0^h \alpha_\nu(h)dh,$$

$\alpha_\nu(h)$ is the absorption per unit length, $T(h)$ temperature, $h$ the distance from radiometer, and $T_b^0(\nu)$ the unattenuated brightness temperature from external sources.

The microwave absorption coefficient $\alpha_\nu$ is due to water vapor and oxygen. At frequencies near 60 GHz, the fractional contribution due to water vapor is small and will be neglected here. For humid locations the introduction of a model atmosphere to account for the wet component might be justified (Dutton and Bean, 1965). The absorption due to oxygen can be calculated as a function of temperature and pressure from the Van Vleck equation (Van Vleck, 1947). The major uncertainty in these calculations is the pressure dependence of the oxygen line widths (Meeks and Lilley, 1963). The calculations here are based on a quadratic expansion of line width as a function of pressure, with constants derived from a least-squares fit to the data of Artman (1953); the line widths are assumed to have a temperature dependence of $T^{-0.85}$. The details of the line width analysis are given by Westwater and Strand (1967). In the height region of interest here (0–10 km) Doppler and Zeeman broadening are negligible.

The temperature mean and covariance matrices were obtained by averaging 5 years of February radiosonde data (163 soundings) taken at Denver, Colo. If we denote the mean temperature at the quadrature point $h_i$ by $\bar{T}(h_i)$ and the $i$, $j$ element of the covariance matrix by $(\mathbf{S}_T)_{ij}\equiv\mathbf{S}_T(h_i,h_j)$, then

$$\bar{T}(h_i)=\frac{1}{N}\sum_{\rho=1}^{N} T_\rho(h_i), \quad (15)$$

and

$$\mathbf{S}_T(h_i,h_j)=\frac{1}{N-1}\sum_{\rho=1}^{N}[T_\rho(h_i)-\bar{T}(h_i)]\times[T_\rho(h_j)-\bar{T}(h_j)], \quad (16)$$

where $N$ is the number of pieces of data and $\rho$ is an index for each member of the sample. Table 1 gives the mean $T$ and $P$ at the Gauss-Legendre quadrature heights

TABLE 2. Temperature covariance matrix for downward inversion for Denver, February. $(\mathbf{S}_T)_{ij}=\mathbf{S}_T(s_i,s_j)$ [°K²] and $s_i$ are the quadrature distances (km) given in Table 1.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 35.50 | 33.10 | 26.59 | 15.06 | 0.00 | −13.24 | −18.91 | −21.80 | −22.54 | −22.76 | −24.80 | −26.39 | −27.44 | −25.87 | −15.46 |
| | 32.49 | 27.03 | 16.14 | 1.44 | −11.60 | −17.49 | −20.67 | −21.20 | −21.25 | −23.12 | −24.63 | −25.33 | −23.74 | −14.07 |
| | | 25.70 | 17.42 | 4.59 | −7.13 | −12.89 | −16.34 | −16.64 | −16.72 | −18.27 | −19.78 | −20.24 | −18.65 | −11.42 |
| | | | 16.18 | 8.58 | .28 | −4.63 | −7.74 | −8.23 | −8.44 | −9.41 | −10.55 | −10.82 | −9.21 | −5.79 |
| | | | | 12.19 | 11.30 | 8.98 | 6.65 | 5.83 | 5.29 | 4.89 | 4.61 | 4.55 | 4.96 | 2.25 |
| | | | | | 21.24 | 22.25 | 20.76 | 19.82 | 19.16 | 20.08 | 20.47 | 20.34 | 19.35 | 12.21 |
| | | | | | | 27.55 | 28.25 | 27.75 | 27.09 | 27.77 | 28.36 | 28.34 | 27.02 | 18.41 |
| | | | | | | | 32.12 | 31.93 | 31.00 | 31.28 | 32.10 | 32.42 | 30.78 | 21.07 |
| | | $(\mathbf{S}_T)_{ij}=(\mathbf{S}_T)_{ji}$ | | | | | | 33.85 | 33.54 | 33.54 | 34.04 | 34.38 | 32.48 | 22.81 |
| | | | | | | | | | 35.22 | 35.47 | 35.75 | 36.19 | 34.27 | 24.80 |
| | | | | | | | | | | 39.46 | 40.49 | ·40.93 | 38.76 | 28.29 |
| | | | | | | | | | | | 45.62 | 47.43 | 44.77 | 31.98 |
| | | | | | | | | | | | | 51.39 | 49.17 | 35.14 |
| | | | | | | | | | | | | | 50.98 | 36.61 |
| | | | | | | | | | | | | | | 36.07 |

used in the downward inversion and at the three-interval Gauss-Radau quadrature heights for the upward inversion. The covariance matrix for downward inversion is given in Table 2. These matrix elements are associated with height as

$$(\mathbf{S}_T)_{ij}\equiv \mathbf{S}_T(s_i,s_j)\equiv \mathbf{S}_T(10-h_i,\ 10-h_j),$$

where $s_k$ is the distance from the radiometer.

For many ground-based probing schemes, the value of the unknown $\mathbf{f}$ at the surface can usually be measured directly. This constrained point can be used to modify the statistical estimation and its uncertainty as follows. First, the constraint can be incorporated directly into the integral equation by using a quadrature formula (such as Gauss-Radau), which uses the value $f_1$ of the function at 0, directly, i.e.,

$$\int_0^H K(x,y)f(y)dy=\sum_{j=1}^{m} w_jK(x,y_j)f(y_j)$$

$$=w_1K(x,0)f_1+\sum_{j=2}^{m} w_jK(z,y_j)f(y_j). \quad (17)$$

By subtracting $w_1K(x,0)f_1$ from the reduced measured quantity $h_e$, a matrix equation to be solved for the $(m-1)$ components of the function $\mathbf{f}$ is obtained. Second, knowing $f_1$ reduces the uncertainty in all the other functional values. The new covariance matrix $\mathbf{S}_f{}^{(c)}$, of dimension $(m-1)\times(m-1)$, has elements

$$\mathbf{S}_{ij}{}^{(c)}=\mathbf{S}_{ij}-\frac{\mathbf{S}_{1i}\mathbf{S}_{1j}}{\mathbf{S}_{11}},\quad i,\ j=2,\ 3,\ \cdots,\ m,$$

where

$$\mathbf{S}_f\equiv(\mathbf{S}_{ij})\quad\text{and}\quad \mathbf{S}_f{}^{(c)}=(\mathbf{S}_{ij}{}^{(c)}). \quad (18)$$

For convenience, the matrix $\mathbf{S}_f{}^{(c)}$ will be referred to in the following as the constrained covariance matrix. Furthermore, instead of the mean $f_0$ as the best *a priori* estimate of $\mathbf{f}$, the effect of knowing the first value $f_1$ modifies this optimum *a priori* estimate to $\hat{\mathbf{f}}$, where

$$\hat{f}_i=f_{0i}+\frac{S_{i1}}{S_{11}}(f_1-f_{01}),\quad i=2,\ 3,\ \cdots,\ m. \quad (19)$$

Eqs. (18) and (19) may be derived from linear regressions of the $(m-1)$ functional values $f_2,\ f_3,\ \cdots f_m$ as functions of the surface value $f_1$ (Westwater and Strand, 1967).

The unconstrained and constrained "upward" covariance matrices for the Denver February temperature structure are shown in Tables 3 and 4, respectively.

TABLE 3. Temperature covariance matrix for upward unconstrained inversion for Denver, February. $(\mathbf{S}_T)_{ij}=\mathbf{S}_T(h_i,h_j)$ [°K²] and $h_i$ are the quadrature heights (km) given in Table 1.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 35.49 | 36.36 | 35.51 | 34.39 | 33.56 | 32.89 | 30.85 | 28.41 | 25.15 | 23.89 | 23.43 | 20.25 | 15.14 | −2.74 | −12.39 |
| | 46.54 | 46.24 | 44.50 | 43.30 | 42.38 | 39.66 | 36.11 | 31.85 | 30.47 | 30.22 | 27.30 | 20.90 | −3.94 | −19.63 |
| | | 50.74 | 49.92 | 48.80 | 47.83 | 44.82 | 40.65 | 35.67 | 34.13 | 33.91 | 30.91 | 23.71 | −4.95 | −23.01 |
| | | | 51.19 | 50.33 | 49.46 | 46.49 | 42.17 | 37.05 | 35.54 | 35.39 | 32.10 | 24.57 | −6.00 | −24.12 |
| | | | | 49.94 | 49.19 | 46.55 | 42.34 | 37.28 | 35.78 | 35.59 | 32.29 | 24.93 | −5.94 | −24.04 |
| | | | | | 48.59 | 46.18 | 42.13 | 37.14 | 35.61 | 35.42 | 32.15 | 24.97 | −5.97 | −23.99 |
| | $(\mathbf{S}_T)_{ij}=(\mathbf{S}_T)_{ji}$ | | | | | 44.77 | 41.53 | 36.93 | 35.36 | 35.17 | 31.97 | 24.88 | −5.92 | −23.64 |
| | | | | | | | 40.34 | 36.70 | 35.16 | 34.93 | 31.56 | 24.65 | −5.32 | −22.79 |
| | | | | | | | | 35.41 | 34.39 | 34.16 | 30.37 | 23.39 | −4.52 | −20.87 |
| | | | | | | | | | 34.11 | 34.20 | 30.49 | 23.29 | −4.15 | −20.34 |
| | | | | | | | | | | 34.82 | 31.28 | 23.68 | −4.01 | −20.23 |
| | | | | | | | | | | | 31.59 | 24.37 | −3.12 | −20.09 |
| | | | | | | | | | | | | 24.47 | 1.38 | −14.57 |
| | | | | | | | | | | | | | 12.96 | 11.21 |
| | | | | | | | | | | | | | | 31.85 |

TABLE 4. Temperature covariance matrix for upward constrained inversion for Denver, February. $(\mathbf{S}_T^{(c)})_{ij}=\mathbf{S}_T^{(c)}(h_{i+1},h_{j+1})$ [°K²] and $h_i$ are the quadrature heights (km) given in Table 1.

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9.29 | 9.86 | 9.27 | 8.92 | 8.68 | 8.05 | 7.00 | 6.08 | 5.99 | 6.22 | 6.55 | 5.39 | −1.13 | − 6.94 |
| | 15.21 | 15.51 | 15.22 | 14.92 | 13.95 | 12.22 | 10.51 | 10.23 | 10.47 | 10.65 | 8.56 | −2.21 | −10.61 |
| | | 17.87 | 17.81 | 17.59 | 16.60 | 14.64 | 12.68 | 12.39 | 12.69 | 12.48 | 9.90 | −3.34 | −12.11 |
| | | | 18.21 | 18.09 | 17.38 | 15.47 | 13.50 | 13.19 | 13.43 | 13.14 | 10.61 | −3.35 | −12.32 |
| | | | | 18.11 | 17.59 | 15.80 | 13.83 | 13.47 | 13.71 | 13.38 | 10.94 | −3.43 | −12.51 |
| | | | | | 17.95 | 16.83 | 15.07 | 14.59 | 14.80 | 14.37 | 11.72 | −3.54 | −12.87 |
| | $(\mathbf{S}_T^{(c)})_{ij}=(\mathbf{S}_T^{(c)})_{ji}$ | | | | | 17.60 | 16.57 | 16.04 | 16.17 | 15.35 | 12.53 | −3.13 | −12.87 |
| | | | | | | | 17.59 | 17.46 | 17.56 | 16.02 | 12.66 | −2.58 | −12.09 |
| | | | | | | | | 18.03 | 18.43 | 16.86 | 13.10 | −2.31 | −12.00 |
| | | | | | | | | | 19.35 | 17.91 | 13.68 | −2.20 | −12.05 |
| | | | | | | | | | | 20.04 | 15.73 | −1.56 | −13.02 |
| | | | | | | | | | | | 18.01 | 2.55 | − 9.28 |
| | | | | | | | | | | | | 12.75 | 10.25 |
| | | | | | | | | | | | | | 27.52 |

Here $(\mathbf{S}_T)_{ij}=\mathbf{S}_T(h_i,h_j)$ and $(\mathbf{S}_T^{(c)})_{ij}=\mathbf{S}_T^{(c)}(h_{i+1},\ h_{j+1})$, where the $h_k$ are the quadrature heights given in Table 1. The surface constraint reduces the matrix elements describing the lower atmosphere but has little effect on the upper levels. The trace is reduced from 572.81 (°K)² to 247.52 (°K)².

To study the information added by oxygen emission measurements, calculations were made for five frequencies and several choices of measurement error for both upward and downward inversion over the 10-km height interval. The kernel was determined from the mean temperature and pressure data given in Table 1. Since the kernel is temperature-dependent, an inversion scheme based on the linear methods given here would necessarily be iterative (Westwater,1965). The brightness temperatures and optical depths for the mean profile are given in Table 5. In the upward calculations, the emission above 10 km is neglected; in the downward case, the ground is assumed to be a blackbody radiating at the surface temperature $T_0$. An example of a solution covariance matrix is given in Table 6 for the downward inversion with each of the five measurements having an assumed rms error of 0.1K in the brightness temperature. This matrix is to be compared with Table 2 to show the reduction in uncertainty of the profile estimation by the introduction of the five measurements.

TABLE 5. Calculated brightness temperature, $T_b$ and optical depths $\tau$, for Denver, February, mean profile; $T_0=267.956$K.

| | Upward | | Downward | |
|---|---|---|---|---|
| $\nu$ (GHz) | $T_b$ (°K) | $\tau$ | $T_b$ (°K) | $T_b+T_0e^{-\tau}$ (°K) |
| 51.2 | 100.699 | 0.514 | 98.665 | 259.461 |
| 53.3 | 195.022 | 1.448 | 184.798 | 248.055 |
| 55.0 | 260.669 | 4.668 | 226.924 | 229.459 |
| 57.3 | 268.089 | 14.674 | 219.516 | 219.516 |
| 61.193059 | 268.633 | 20.677 | 218.731 | 218.731 |

The meaning of a 15×15 covariance matrix is difficult to present in simple form. A rough estimate of the standard deviation to be expected at each quadrature height is given by the square root of the corresponding diagonal element of the covariance matrix. These quantities are plotted as functions of height in Figs. 1–3. In all cases shown, the measurement error covariance matrix is scalar, $\mathbf{S}_\epsilon=\sigma_\epsilon^2\mathbf{I}$, where $\sigma_\epsilon$ is the standard deviation in °K and $\mathbf{I}$ is the $n\times n$ identity matrix. The reduction of the total error in the temperature profile by various choices of $\sigma_\epsilon$ is given in Table 7. For example, in the case of upward constrained inversion with brightness temperature errors of 1K in each of the five measurements, the variance is reduced by $R=178.13$ (°K)², the fractional reduction $F=0.72$, and the remaining uncertainty per point $U$ is 2.1K. It is apparent that many experimental possibilities can be investigated with the preceding methods.

TABLE 6. Solution covariance matrix for downward inversion for Denver, February. $\sigma_\epsilon=0.1$K; $\mathbf{X}_{ij}^{-1}=\mathbf{X}^{-1}(s_i,s_j)$ [°K²]; and $s_i$ are the quadrature distances (km) given in Table 1.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.70 | 0.41 | −0.76 | −0.81 | −0.11 | 0.32 | 0.42 | 0.43 | 0.07 | −0.11 | −0.17 | −0.10 | −0.36 | −0.43 | −0.23 |
| | 0.57 | −0.32 | −0.61 | −0.16 | 0.26 | 0.28 | 0.21 | 0.04 | −0.03 | −0.10 | −0.09 | −0.13 | −0.24 | −0.10 |
| | | 0.81 | 0.13 | −0.33 | −0.21 | −0.03 | −0.01 | 0.19 | 0.17 | 0.09 | −0.09 | −0.06 | −0.14 | −0.24 |
| | | | 1.38 | 0.19 | −0.60 | −0.57 | −0.26 | 0.01 | 0.14 | 0.17 | 0.06 | 0.14 | 0.45 | 0.54 |
| | | | | 1.26 | −0.05 | −0.52 | −0.56 | −0.44 | −0.20 | −0.04 | 0.40 | 0.68 | 0.77 | 0.19 |
| | | | | | 1.47 | 0.38 | −0.72 | −0.99 | −0.79 | 0.12 | 0.64 | 0.71 | 0.50 | 0.32 |
| | | | | | | 1.87 | 0.28 | −0.18 | −0.34 | −0.23 | −0.13 | −0.23 | −0.16 | 0.22 |
| | | | | | | | 1.65 | 0.93 | 0.10 | −0.66 | −0.84 | −0.87 | −0.78 | −0.67 |
| | | | | | | | | 1.84 | 1.21 | −0.27 | −1.24 | −1.56 | −1.69 | −1.28 |
| | $(\mathbf{X}^{-1})_{ij}=(\mathbf{X}^{-1})_{ji}$ | | | | | | | | 2.15 | 0.47 | −1.19 | −1.76 | −1.94 | −1.19 |
| | | | | | | | | | | 1.87 | 0.28 | −0.75 | −1.19 | −0.67 |
| | | | | | | | | | | | 2.08 | 1.95 | 1.01 | −0.08 |
| | | | | | | | | | | | | 3.64 | 3.08 | 1.14 |
| | | | | | | | | | | | | | 6.37 | 3.59 |
| | | | | | | | | | | | | | | 10.74 |

TABLE 7. Solution error vs error in brightness temperature for five frequencies of Table 5. $S_\epsilon=\sigma_\epsilon^2 I$ [°K]². All values in (°K)².

| | $TrS_T$ | TrX⁻¹ $\sigma_\epsilon=0$ | $\sigma_\epsilon=0.01$ | $\sigma_\epsilon=0.1$ | $\sigma_\epsilon=0.5$ | $\sigma_\epsilon=1.0$ | $\sigma_\epsilon=1.5$ | $\sigma_\epsilon=2.0$ |
|---|---|---|---|---|---|---|---|---|
| Downward | 495.56 | 19.68 | 21.63 | 38.39 | 66.61 | 91.45 | 114.10 | 138.23 |
| Upward | 572.81 | 21.83 | 27.54 | 44.17 | 63.68 | 78.63 | 91.98 | 104.07 |
| Upward constrained | 247.52 | 18.27 | 23.19 | 37.85 | 55.54 | 69.29 | 81.53 | 92.79 |

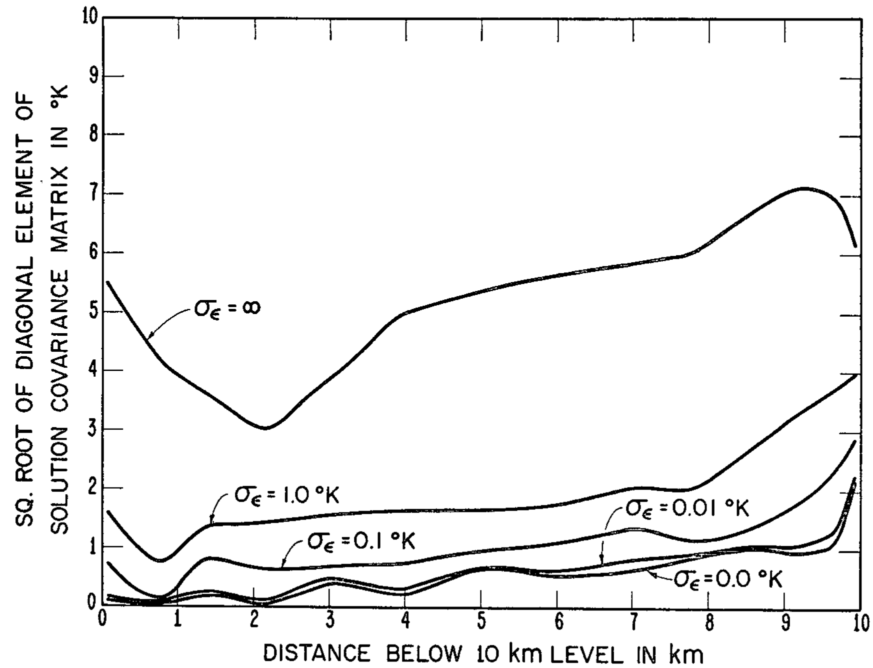


FIG. 1. Square root of diagonal element of solution covariance matrix (°K) vs distance below 10-km level for various values of $\sigma_\epsilon$. Downward inversion, five frequencies, Denver, February.

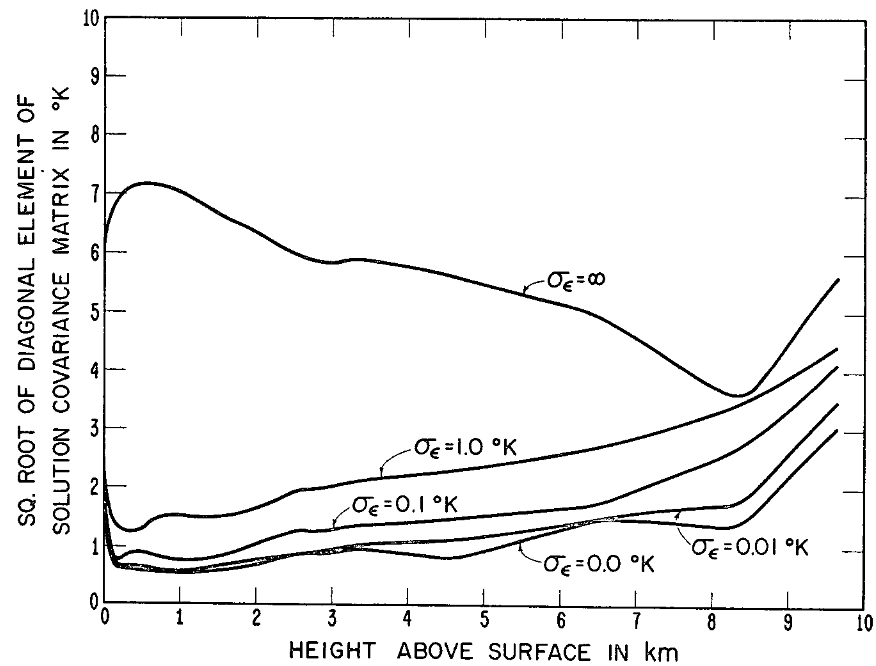


FIG. 2. Same as Fig. 1 except for height above surface and unconstrained upward inversion.

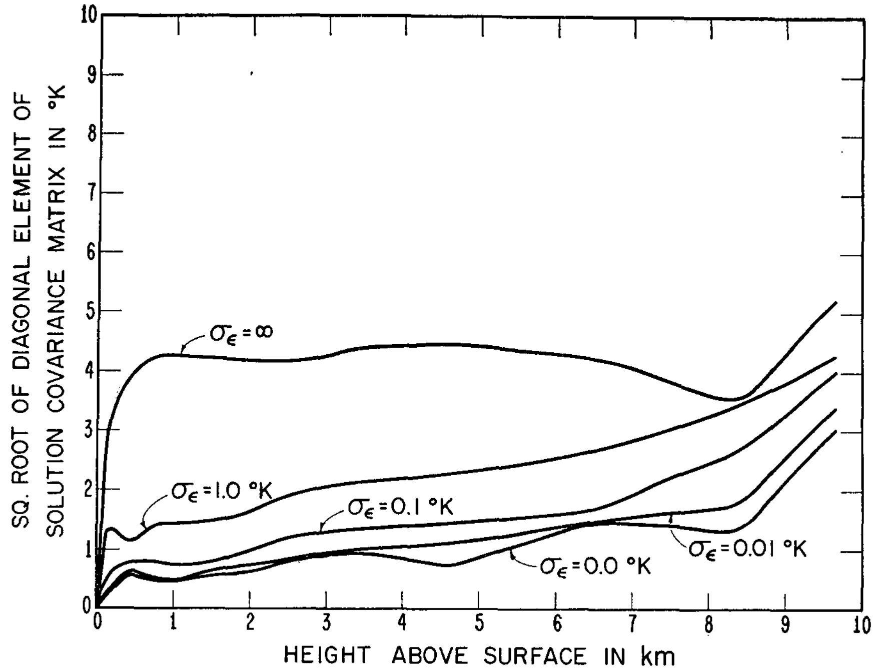


FIG. 3. Same as Fig. 2 except for constrained upward inversion.

## 4. Interdependence and optimum location of measurements

Twomey (1966) has given an explicit procedure for determining the optimum location of measurements. This method orders a set of measurements by systematically eliminating the most redundant of the set and is based entirely on the properties of the kernel. However, when the information contained in the *a priori* statistics, $\mathbf{S}_f$, is introduced, an additional correlation between radiation measurements and $\mathbf{S}_f$ occurs. Thus, if we define the optimum subset of a large set of measurements as the subset which yields a minimum $\mathrm{Tr}\mathbf{X}^{-1}$ (i.e., minimum expected mean-square error), the order-

TABLE 8. Solution variance $\mathrm{Tr}\mathbf{X}^{-1}$ [°K²] resulting from successive introduction of measurements for Denver, February. $\mathrm{Tr}\mathbf{S}_f^{(D)} = 495.56$ (°K)², $\mathrm{Tr}\mathbf{S}_f^{(UU)} = 572.81$ (°K)², $\mathrm{Tr}\mathbf{S}_f^{(UC)} = 247.52$ (°K)² where $D$ is downward, $UU$ upward unconstrained, $UC$ upward constrained; frequency is in GHz and $\sigma_\epsilon$ is the rms measurement error.

| | | $\sigma_\epsilon$ (°K) 0.0 | 0.1 | 0.5 | 1.0 |
|---|---|---|---|---|---|
| 1 Freq (1) | D | 292.01 | 292.11 | 294.51 | 301.66 |
| | UU | 136.26 | 136.36 | 138.74 | 146.01 |
| | UC | 119.31 | 119.47 | 123.09 | 133.20 |
| 2 Freq (1, 2) | D | 109.63 | 110.03 | 119.39 | 145.33 |
| | UU | 65.50 | 65.89 | 74.18 | 92.93 |
| | UC | 58.78 | 59.31 | 69.10 | 85.28 |
| 3 Freq (1, 2, 3) | D | 47.15 | 48.94 | 73.13 | 100.77 |
| | UU | 50.10 | 53.57 | 71.11 | 88.90 |
| | UC | 40.51 | 42.82 | 58.76 | 73.28 |
| 4 Freq (1, 2, 3, 4) | D | 32.40 | 41.68 | 71.89 | 97.93 |
| | UU | 36.18 | 46.28 | 65.62 | 80.89 |
| | UC | 29.07 | 39.98 | 57.45 | 71.54 |
| 5 Freq (1, 2, 3, 4, 5) | D | 19.68 | 38.39 | 66.61 | 91.45 |
| | UU | 21.83 | 44.17 | 63.68 | 78.63 |
| | UC | 18.27 | 37.85 | 55.54 | 69.29 |

Twomey's downward frequency ranking (GHz)
(1) 61.193059
(2) 55.0
(3) 53.3
(4) 57.3
(5) 51.2

Twomey's upward unconstrained frequency ranking (GHz)
(1) 61.193059
(2) 53.3
(3) 57.3
(4) 55.0
(5) 51.2

Twomey's upward constrained frequency ranking (GHz)
(1) 61.193059
(2) 55.0
(3) 53.3
(4) 57.3
(5) 51.2

ing of sets will differ, in general from that obtained by Twomey's method. In addition, the dependence of observations is described differently here. Instead of an effective "number of independent pieces of information," all measurements will reduce the variance of the solution and are considered as information. The dependence of measurements is seen as a law of diminishing returns and adding additional dependent measurements reduces the variance very little. However, if one is willing to pay the cost, the variance can be reduced to any arbitrarily small level by proper selection of a large number of observations.

The effect of dependence of measurements was studied for the microwave inversion problem as follows. Kernels corresponding to the five frequencies of Table 5 were ranked in the order of decreasing strength by the method of Twomey. The solution variance was calculated for a number of assumed experimental errors as each of these observations was successively added. These results are shown in Table 8. In the errorless cases, each successive measurement reduces the variance by substantial amounts; when errors are introduced, the reduction of variance after the first two or three measurements are added is small. The non-optimality of Twomey's method of measurement ranking is evident since cases occur when the addition of a channel reduces the variance more than its predecessor.

The optimum subset of a large number of possible measurements will depend on the kernel and the covariance matrices of the statistics and experimental error. In general, the optimum will also depend on the number of elements in the subset. The rankings of the five frequencies for one-frequency and two-frequency optimums are shown in Tables 9 and 10, respectively, for the upward constrained case. Note that the ranking changes for different choices of experimental error, and that all rankings differ from those based from Twomey's scheme which is based only on the kernel.

The determination of an optimum set of frequencies from an ensemble large enough to adequately cover the entire oxygen band would be a large computational chore. If $N$ frequencies suffice to cover the band and one wishes to determine an optimum $M$ of them, then $N!/(N-M)!M!$ trace computations must be compared to determine a minimum. In view of the high cost of microwave radiometers, however, such calculations might be in order.

TABLE 9. One-frequency expected mean-square errors $TrX^{-1}$ [$°K^2$] for Denver, February; upward constrained inversion. Ranking according to minimum trace criterion is enclosed in parentheses. $TrS_f = 274.52$ $(°K)^2$ and $\sigma_\epsilon$ is the rms measurement error.

| Twomey's ranking | Frequency (GHz) | $\sigma_\epsilon$(°K) 0.0 | 0.5 | 1.0 |
|---|---|---|---|---|
| 1 | 61.193059 | 119.31 (5) | 123.09 (5) | 133.20 (4) |
| 2 | 55.0 | 78.09 (2) | 81.70 (2) | 91.68 (1) |
| 3 | 53.3 | 74.30 (1) | 80.73 (1) | 97.43 (2) |
| 4 | 57.3 | 105.45 (4) | 109.03 (4) | 118.78 (3) |
| 5 | 51.2 | 79.44 (3) | 101.46 (3) | 142.66 (5) |

TABLE 10. Two-frequency expected mean-square errors $TrX^{-1}$ [$°K^2$] for Denver, February; upward constrained inversion. Ranking according to minimum; trace criterion is enclosed in parentheses. $TrS_f = 247.52$ $(°K)^2$ and $\sigma_\epsilon$ is the rms measurement error.

| Twomey's ranking | Frequency (GHz) | $\sigma_\epsilon$(°K) 0.0 | 0.5 | 1.0 |
|---|---|---|---|---|
| 1<br>2 | 61.193059<br>55.0 | 58.78 (9) | 69.10 (4) | 85.28 (4) |
| 1<br>3 | 61.193059<br>53.3 | 56.40 (7) | 65.14 (2) | 82.61 (3) |
| 1<br>4 | 61.193059<br>57.3 | 75.89 (10) | 106.30 (10) | 116.63 (10) |
| 1<br>5 | 61.193059<br>51.2 | 58.58 (8) | 77.66 (8) | 103.86 (9) |
| 2<br>3 | 55.0<br>53.3 | 50.54 (2) | 66.23 (3) | 79.95 (1) |
| 2<br>4 | 55.0<br>57.3 | 55.48 (5) | 69.53 (5) | 85.77 (6) |
| 2<br>5 | 55.0<br>51.2 | 51.12 (3) | 69.99 (6) | 85.34 (5) |
| 3<br>4 | 53.3<br>57.3 | 54.01 (4) | 63.54 (1) | 80.69 (2) |
| 3<br>5 | 53.3<br>51.2 | 49.38 (1) | 78.62 (9) | 93.53 (7) |
| 4<br>5 | 57.3<br>51.2 | 55.79 (6) | 74.30 (7) | 97.39 (8) |

## 5. Conclusion

It was shown that the information content of radiation measurements used for remote probing can be defined with reference to *a priori* information. Comparison of the solution quality criterion, $TrX^{-1}$ ($m$ times the mean-square error), with the trace of the *a priori* covariance matrix, $TrS_f$, can be used to judge the information content. An optimum set of measurements can be defined as the set which minimizes $TrX^{-1}$, and is shown to depend on the *a priori* knowledge, the measurement covariance matrix, the kernel of the equation, and the order (number of elements) of the set.

Calculations of information obtainable from microwave measurements of oxygen thermal emission were carried out for a radiometer measuring upwelling and downwelling radiation from a 10-km height interval in the troposphere. Out of an initial choice of five frequencies a one-frequency and two-frequency optimum set was obtained, and comparisons were made with the optimum set obtained by the method of Twomey. The two methods differ considerably in optimum ranking.

REFERENCES

Alishouse, J. C., L. J. Crone, H. E. Fleming, F. L. Van Cleef and D. Q. Wark, 1967: A discussion of empirical orthogonal functions and their application to vertical temperature profiles. *Tellus*, **19**, 477–482.

Artman, J. O., 1953: Absorption of microwaves by oxygen in the millimeter wavelength region. Columbia Radiation Laboratory Rept., 84 pp.

Deutsch, R., 1965: *Estimation Theory*. Englewood Cliffs, N. J., Prentice-Hall, Inc., 267 pp.

Dutton, E. J., and B. R. Bean, 1965: The biexponential nature of tropospheric gaseous absorption of radio waves. *Radio Sci., J. Res. Natl. Bur. Std.*, **69D**, 885–891.

Mateer, C. L., 1965: On the information content of Umkehr observations. *J. Atmos. Sci.*, **22**, 370–381.

Meeks, M. L., and A. E. Lilley, 1963: The microwave spectrum of oxygen in the earth's atmosphere. *J. Geophys. Res.*, **68**, 1683–1703.

Shklovsky, I. S., 1960: *Cosmic Radio Waves*. Harvard University Press, 444 pp.

Strand, O. N., and E. R. Westwater, 1968a: The statistical estimation of the numerical solution of a Fredholm integral equation of the first kind. *J. Assoc. Comp. Mach.*, **15**, 100–115.

——, and ——, 1968b: Minimum-rms estimation of the numerical solution of a Fredholm integral equation of the first kind. *Siam J. Numerical Anal.* (in press).

Twomey, S., 1965: The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements. *J. Franklin Inst.*, **279**, 95–109.

——, 1966: Indirect measurements of atmospheric temperature profiles from satellites: II. Mathematical aspects of the inversion problem. *Mon. Wea. Rev.*, **94**, 363–366.

Van Vleck, J. H., 1947: The absorption of microwaves by oxygen. *Phys. Rev.*, **71**, 413–424.

Wark, D. Q., and H. E. Fleming, 1966: Indirect measurements of atmospheric temperature profiles from satellites: I. Introduction. *Mon. Wea. Rev.*, **94**, 351–362.

Westwater, E. R., 1965: Ground-based passive probing using the microwave spectrum of oxygen. *Radio Sci., J. Res. Natl. Bur. Std.*, **69D**, 1201–1211.

——, and O. N. Strand, 1967: Application of statistical estimation techniques to ground-based passive probing of the tropospheric temperature structure. ESSA Tech. Rept. IER 37-ITSA 37, 87 pp.