

NASA TM X- 55857

INVERSE PROBLEMS IN RADIATIVE TRANSFER: A REVIEW

BY

BARNEY J. CONRATH

FACILITY FORM 602	N 67 - 32849	(THRU)
	(ACCESSION NUMBER)	1
	35	(CODE)
	(PAGES)	33
	TMX-55857	(CATEGORY)
	(NASA CR OR TMX OR AD NUMBER)	

GPO PRICE \$ _____

CFSTI PRICE(S) \$ _____

Hard copy (HC) 3.00

Microfiche (MF) 1.65

FEBRUARY 1967

ff 653 July 65



GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

**INVERSE PROBLEMS IN RADIATIVE
TRANSFER: A REVIEW**

by

Barney J. Conrath

February 1967

**GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland**

PRECEDING PAGE BLANK NOT FILMED.

INVERSE PROBLEMS IN RADIATIVE
TRANSFER: A REVIEW

by
Barney J. Conrath

ABSTRACT

An elementary introduction to the inverse problem of radiative transfer as applied to remote measurements of infrared radiation from planetary atmospheres is presented, along with a review of work which has been done on the problem. Particular attention is given to the problems of inferring vertical temperature profiles and water vapor distributions in the earth's atmosphere. The principle methods which have been developed for solving these problems are discussed briefly. Examples of applications of temperature and water vapor inversions to both synthetic data and data taken with an IRIS instrument at balloon altitudes are presented.

CONTENTS

	<u>Page</u>
ABSTRACT	iii
INTRODUCTION	1
General	1
Physical Description of the Problem	1
Mathematical Formulation	2
METHODS DEVELOPED FOR TEMPERATURE INVERSION	4
Linear Methods	4
Expansion Functions	8
Nonlinear Method of King	10
Temperature Inversion in Cloudy Atmospheres	14
CONSTITUENT INVERSION	14
SUMMARY	17
ACKNOWLEDGMENTS	18
REFERENCES	18

INVERSE PROBLEMS IN RADIATIVE

TRANSFER: A REVIEW

INTRODUCTION

General

In the direct problem of radiative transfer as applied to a planetary atmosphere, the vertical structure of the atmosphere is specified along with the energy sources, and the radiation field is calculated. In particular the spectral intensity of the upwelling electromagnetic radiation at the effective top of the atmosphere can be predicted. In the inverse problem the outgoing intensity is assumed to be obtained observationally, and an effort is made to infer information on the vertical atmospheric structure. In the present paper we shall consider a narrowly restricted class of the inverse problem; inference of vertical temperature profiles and vertical constituent distributions from remote measurements of thermally emitted radiation.

The primary purposes of the paper are to provide an introduction to the subject and summarize the present level of development. The problem of inferring the vertical temperature profile will first be considered, and a review of the basic methods which have been developed will be given. The important questions of the accuracy of the inferences will be discussed, and examples of applications to both synthetic and real data will be presented. Finally we shall discuss some recent work on the problem of remotely inferring the distribution of water vapor in the lower terrestrial atmosphere.

Physical Description of the Problem

As an example, let us consider the problem of inferring the temperature profile in the lower atmosphere of the earth by observing the spectral intensity in the 15μ CO_2 absorption band with an instrument located above the effective upper boundary of the atmosphere. The observations consist of a spectral scan from the band center out into the band wings. From a knowledge of the absorptivity of CO_2 as a function of wavenumber, we can obtain the relative optical depths. With the additional information that CO_2 is uniformly mixed with a known mixing ratio in the part of the atmosphere we are considering, we can relate the optical depths at the various wavenumbers to actual pressure levels in the atmosphere. In other words we can calculate the transmissivity of the atmosphere from a given pressure level to the top as a function of wavenumber.

In Figure 1 we have plotted the pressure level at which the transmissivity drops to $1/e$. This figure corresponds to a resolution equivalent to 0.1 cm^{-1} which is beyond the reach of present satellite instrumentation, but serves to illustrate our point. In scanning from the band center toward the wing we see progressively deeper into the atmosphere. Thus, from measurements at selected wavenumbers we should expect to be able to infer information on the atmospheric temperature at a number of different pressure levels, and in this way piece together a temperature profile.

Now the intensity at a given wavenumber, instead of depending on the temperature at a single level, actually contains contributions over a number of levels on either side of the e-folding pressure. To isolate the temperature at a given level we would have to subtract out contributions from all other layers which is equivalent to differentiation of the data. Thus, we might intuitively expect to encounter problems with the stability of our solutions.

Mathematical Formulation

The solution to the equation of radiative transfer for a nonscattering planetary atmosphere in local thermodynamic equilibrium can be written in the well known integral form

$$I(\nu) = I_s(\nu) \tau(\nu, x_s) - \int_0^{x_s} B[\nu, T(x)] \frac{\partial \tau(\nu, x)}{\partial x} dx \quad (1)$$

where $I(\nu)$ is the spectral specific intensity emerging at the top of the atmosphere, $B(\nu, T)$ is the Planck intensity at wavenumber ν and temperature T , x is any independent variable which is a single valued function of pressure, and $\tau(\nu, x)$ is the transmissivity at wavenumber ν of the column of atmospheric gas between level x and the top of the atmosphere. The subscript s refers to the planetary surface. For simplicity, we shall assume that our measurements are from sufficiently opaque parts of the spectrum that the boundary term in (1) can be neglected.

We see from (1) that the outgoing intensity can be interpreted as a weighted mean of the Planck functions associated with each layer of the atmosphere. The weighting function is just $\partial \tau / \partial x$. Typical weighting functions for the $15\mu \text{ CO}_2$ band are shown in Figure 2. These functions were calculated by Kunde (private communication) for an instrumental response function 5 cm^{-1} in width. From this figure we can obtain an idea of the relative contribution of each atmospheric

layer to the outgoing intensity at each wavenumber. The "width" of the weighting functions is determined primarily by the way in which the optically active gas is distributed and for pure monochromatic radiation is ~ 1.4 atmospheric scale heights at the $1/e$ points. Averaging over a finite spectral interval in which the absorption coefficient is changing will tend to increase this width somewhat. The spectral resolution required depends on the number of pieces of information we are trying to retrieve and the levels of the atmospheres at which information is sought. In general by increasing the resolution we can reach higher levels in the atmosphere. A resolution equivalent to about 5 cm^{-1} seems to be the lower practical limit for inferring five or six pieces of information in the troposphere and stratosphere. Relatively broad band measurements could be used to infer limited information on the troposphere.

The temperature inversion problem can now be stated as follows; given $I(\nu)$ and $\partial\tau/\partial x$, solve (1) for $T(x)$. In its present form, (1) is a nonlinear integral equation in $T(x)$. The problem can be simplified however by making use of the fact that the Planck function is a slowly varying function of wavenumber in the spectral region and at the temperatures we are considering. Hence we can linearize (1) by using an approximation of the form

$$B[\nu, T] \simeq a(\nu) + b(\nu) B[\nu_0, T] \quad (2)$$

where ν_0 is some reference wavenumber near the center of our spectral measurements. Letting $B[\nu_0, T(x)] \equiv B(x)$, (1) becomes

$$g(\nu) = \int_0^{x_s} B(x) K(\nu, x) dx \quad (3)$$

where

$$g(\nu) \equiv \frac{I(\nu) - a(\nu)}{b(\nu)}$$

$$K(\nu, x) \equiv \frac{\partial\tau(\nu, x)}{\partial x}$$

We can now regard $B(x)$ as our unknown function since there is a one-to-one correspondence between $B(x)$ and $T(x)$. Equation (3) is a linear Fredholm integral equation of the first kind. Such equations are notorious for their instability, the stability depending on the nature of the kernel $K(\nu, x)$.

We have already indicated the physical nature of the instability. Some insight into the mathematical nature of the instability can be gained from the following consideration. Suppose $B_0(x)$ represents the solution to (3). Now add to the solution a term of the form $\sin 2\pi fx$, and insert this into the right hand side of (3), carry out the integration and compare the result with $g(\nu)$. If the frequency f is sufficiently high so $\sin 2\pi fx$ changes rapidly compared to $K(\nu, x)$ for all ν considered, then adjacent negative and positive half cycles tend to cancel, and the contribution of this term will be small. Hence, the total integral will differ only slightly from $g(\nu)$. Thus, we see that if $g(\nu)$ is imperfectly known as will be the case in practice, wildly oscillating solutions may result in place of the relatively smooth solution sought.

The problem is further compounded by the fact that we will have available measurements corresponding to $g(\nu)$ only at say N discrete values of ν . Thus, even with perfect data, we cannot hope to obtain a unique solution without imposing additional constraints upon the problem. In other words, we must specify an interpolation between the measured data points.

We may summarize the problem as being one in which a solution $B(x)$ to (3) is sought from an imperfect specification of $g(\nu)$ at N discrete values of ν . Any attempted solution must specify in effect an interpolation of $g(\nu)$, and a means whereby the basically unstable solution can be objectively smoothed.

METHODS DEVELOPED FOR TEMPERATURE INVERSION

Linear Methods

Perhaps the most direct way to attack the problem is to attempt to expand the Planck function $B(x)$ in terms of some function set $F_j(x)$. Such an expansion must of course be truncated at a number of terms equal to or less than the number of independent measurements we have available to us. On substitution of the expansion into (3) we obtain

$$g(\nu) \simeq \sum_{j=1}^M a_j \int_0^{x_s} F_j(x) K(\nu, x) dx \quad (4)$$

where the a_j 's are the expansion coefficients to be determined. Thus, in choosing the set $F_j(x)$ we have implicitly specified the functional form of the interpolation for $g(\nu)$. If we like we can regard the process of finding the a_j 's as a curve fitting of (4) through the measured points $g_i \equiv g(\nu_i)$. We obtain the set of linear equations

$$g_i = \sum_{j=1}^M A_{ij} a_j \quad \begin{array}{l} i = 1, 2, \dots, N \\ N \geq M \end{array} \quad (5)$$

where we have let

$$A_{ij} = \int_0^{x_s} F_j(x) K(\nu_i, x) dx$$

For the sake of simplicity we shall assume $N = M$, but what follows applies to the more general case also. It is convenient to write (5) in matrix notation

$$\mathbf{g} = \mathbf{A} \mathbf{a} \quad (6)$$

We can now obtain a direct solution by solving (5) or (6) in the form

$$\mathbf{a} = \mathbf{A}^{-1} \mathbf{g} \quad (7)$$

Such a solution applied to noisy data usually leads to catastrophic results, because the matrix A is generally ill conditioned. This is equivalent to saying that the equations (5) do not possess a high degree of independence. This in turn can be traced back to the fact that the relative contributions from the various levels in the atmosphere are highly overlapping from one measured intensity to another.

The propagation of the random error in the measurements of g can easily be analyzed and can be expressed in the form

$$\frac{\delta a}{a} \approx G \frac{\delta g}{g} \quad (8)$$

where the error amplification factor G is in general dependent on the function set employed and the number of terms in the expansion. A typical value for a 5 term expansion in terms of line segments is $G \sim 10^4$. Obviously such an approach is impractical because of the extremely high demands placed on the accuracy with which g must be determined. In order to obtain a with an accuracy of only 10%, it would be necessary to reduce the experimental random error to less than one part in 10^5 .

Since a direct inversion of data containing what might normally be considered reasonable errors leads in general to wildly oscillating solutions, it becomes necessary to introduce some type of smoothing process. Perhaps the most promising method of smoothing is that utilized by Wark and Fleming (1966) which employs a method of solution initially suggested by Phillips (1962) and extended by Twomey (1963).

The method can be viewed essentially as follows. Assume we have N measurements g_i with a certain probable error associated with them. Now let us allow ourselves the freedom of modifying each g_i by some additive amount ϵ_i . Of course as soon as we do this we have available to us once again an infinite number of solutions, and we must impose some additional constraints. One obvious constraint to impose is that the ϵ_i 's should not exceed the probable error of our measurements. We still have an infinite number of solutions available to us, but they are now members of a rather restricted set. From this set of solutions we pick that solution which minimizes some quadratic form of the a_j 's

$$Q = \sum_{i,j} H_{ij} a_i a_j \quad (9)$$

where the coefficients H_{ij} are determined by the type of smoothing constraint which we wish to impose. The specific quadratic form chosen depends on our a priori knowledge of the form of the solution. For example we may wish to

minimize the mean square curvature of the solution or the mean square deviation from the mean or perhaps the mean square deviation from some trial solution based on a climatological mean temperature profile.

The smoothed solution $\mathbf{a}^{(s)}$ can be expressed in the form

$$\mathbf{a}^{(s)} = (\tilde{\mathbf{A}}\mathbf{A} + \gamma\mathbf{H})^{-1} \tilde{\mathbf{A}}\mathbf{g} \quad (10)$$

where $\tilde{\mathbf{A}}$ is the transpose of \mathbf{A} , \mathbf{H} is a matrix whose elements are the coefficients H_{ij} , and γ is a Lagrange multiplier. The value of γ can in principle be related to the probable error in the data. In practice it is established through trial and error. In general increasing γ increases the amount of smoothing. For $\gamma = 0$, (10) reduces to the direct inversion (7) and $\mathbf{a}^{(s)} = \mathbf{a}$. Thus, if we like we can view (10) as a means of introducing a specified type of smoothing with the degree of smoothing being conveniently controlled by the value of a single scalar parameter. Arithmetically, the effect of increasing γ can be viewed as simply making the matrix to be inverted more diagonally dominant and hence better conditioned. This follows from the fact that for most reasonable types of smoothing \mathbf{H} is strongly diagonally dominant.

An example of this type of smoothed solution is given in Figure 3 where data from a balloon borne IRIS instrument obtained near Palestine, Texas, 8 May 1966 have been employed. The function set used consists of four line segments requiring the determination of five coefficients. The random error of the measurements is uncertain, but is estimated at about 2%. The curve marked $\gamma = 0$ is the unsmoothed solution. The curve for a value of γ giving the best fit is shown along with a solution employing a relatively large value of γ , resulting in an oversmoothing. The constraint employed was minimization of the second differences of the amplitudes at the break points.

The question of the true information content of the data is an important one. We would like to ascertain how many independent pieces of information we can expect to infer from data of a given accuracy. Assume that using N intensity measurements we have successfully determined N coefficients of an expansion for the temperature profile. We can now substitute the inferred profile back into the equation of transfer and use the resulting relation for interpolation or extrapolation. Now assume we have an additional measurement available to us. We can use our extrapolation formula to predict the value of the additional measurement, and if the predicted value agrees with the measured value to within the probable error of the measurement, then we cannot hope to

improve on our solution by adding an additional term to the expansion. Indeed, any attempt at inferring an additional expansion coefficient can be expected to increase the instability.

While we are limited to the number of independent pieces of information which we can obtain, additional information can be used in a redundant sense to improve the effective signal-to-noise ratio. To realize a significant improvement, one must have available a number of additional measurements since the effective signal-to-noise ratio is approximately proportional to the square root of the redundancy.

The number of independent pieces of information which can be inferred from a given set of data depends on the degree of independence of the kernel functions and the signal-to-noise ratio. Twomey (1966) has provided an elegant mathematical method of analyzing this problem. It appears that at most about 5 independent pieces of information can be inferred from present state-of-the-art satellite borne instruments.

Expansion Functions

Since we have a limited number of expansion terms available to us, it is important to choose a function set which gives us the most rapid convergence possible. Most of the classical orthogonal function sets have been tried, and of these the best appears to be trigonometric functions. Straight line segments can be made to fit well, but in general there is no good a priori way of choosing the break points.

In the case of temperature profiles in the earth's atmosphere, we have available to us a great deal of information on the statistical behavior of the profile for a given season and latitude. A convenient means of incorporating such information into our inversion technique is provided by the use of empirical orthogonal functions. Application of this technique to the temperature inversion problem has been made by Alishouse, et al. (1966).

Let us assume we can treat the profile $B(x)$ as a stochastic function whose covariance is given by

$$R(x, x') = \langle [B(x) - \langle B(x) \rangle] \cdot [B(x') - \langle B(x') \rangle] \rangle \quad (11)$$

where the angular brackets denote ensemble averaging. It can be shown (cf. Obukhov, 1960) that the eigenfunctions of $R(x, x')$ taken as an integral

operator

$$\int R(x, x') \phi_n(x') dx' = \lambda_n \phi_n(x) \quad (12)$$

form the most natural set for expansion of members of the ensemble in the following sense. Let $b(x) \equiv B(x) - \langle B(x) \rangle$ be approximated by N terms of an expansion in an orthogonal function set and let this approximation be denoted by $b^{(N)}(x)$. Then define a mean square error for the ensemble by

$$\sigma_N^2 = \int \langle [b(x) - b^{(N)}(x)]^2 \rangle dx \quad (13)$$

Of all possible sets of orthogonal functions, σ_N^2 is a minimum for the set $\phi_n(x)$ obtained from (12) and ordered by decreasing eigenvalues λ_n . Thus we would expect the set $\phi_n(x)$ to provide a good set for use in our inversion problem, providing we can accurately estimate $R(x, x')$.

To gain some insight into the physical interpretation of the ϕ_n 's we make use of the relation

$$R(x, x') = \sum_{n=1}^{\infty} \lambda_n \phi_n(x) \phi_n(x') \quad (14)$$

which is derivable from (12) and the orthogonality properties of the ϕ 's. $R(x, x')$ is the relative correlation of the derivation from the mean at levels x and x' . The first term in (14) is the first approximation to $R(x, x')$ and contains the overall gross features of the function. Thus $\phi_1(x)$ reflects the large scale statistical relationship of the derivations from the mean. The higher order functions are measures of the correlation of finer structure.

In Figure 4 we show two functions, the first order eigenfunction and the fifth order eigenfunction for an ensemble of twenty-five temperature profiles narrowly restricted in geographic location and time. The two functions shown

serve to illustrate the way in which the structure increases as we go to higher eigenfunctions. By looking at ϕ_1 , we see that for this ensemble when we have a tropopause cooler than the mean, it is statistically related to both surface temperatures and stratospheric temperatures warmer than the mean.

As a numerical experiment we have chosen one member of the ensemble and calculated synthetic "data" for five points in the 15μ CO_2 band. These data were then used in an inversion employing a five term expansion in the empirical orthogonal function set. The results are shown in Figure 5. The solution displays an amazing amount of structure for a five term expansion. We next considered a profile which was not a member of the ensemble. The resulting inversion is shown in Figure 6 where we see the agreement is not as good this time, particularly at the tropopause. While the deviation from the mean of the second sounding was no greater than that for the first sounding, the behavior of the second at the tropopause was somewhat atypical which apparently accounts for the poorer fit.

In employing empirical orthogonal functions, we are restricting ourselves in such a way that we can generally not hope to recover unusual profiles. Some profiles will be represented better than others. In other words, if we are going to employ statistical methods we must expect statistical results. The degree of success of the method depends on how accurately we can estimate $R(x, x')$. This is a subject of current investigation.

Nonlinear Method of King

We shall now consider a somewhat different approach to the temperature inversion problem developed by King (1964). In this method it is assumed that (1) can be written in the form

$$I(s) = \int_0^{\infty} B(x) e^{-sx} s dx \quad (15)$$

for a channel sufficiently opaque that the boundary term may be neglected. The parameter s is a function of wavenumber. Form (15) assumes the transmission function can be represented by the analytical form $\tau(\nu, p) = \exp[-s(\nu) x(p)]$. This form is valid for purely monochromatic radiation, but is only approximate for transmissivities averaged over finite band widths. It has been found, however, that transmissivities in the 15μ CO_2 band can be fit fairly well by this analytic form over a limited wavenumber range.

If (15) is assumed to be valid, the outgoing intensity is formally just the Laplace transform of the Planck function profile. Unless $I(s)$ can be expressed in some analytic form whose Laplace transform is known, this helps us very little since numerical inversion of Laplace transforms is in general an unstable process (Bellman, et al., 1966) as we might expect since we have done nothing to change the physical nature of the problem. Let us turn the problem around and assume some physically reasonable form for $B(x)$, calculate $I(s)$ and attempt to fit the resulting expression to the data just as we did in the linear method.

It is convenient to integrate (15) by parts

$$I(s) = B(0) + \int_0^{\infty} \frac{dB}{dx} e^{-sx} dx \quad (16)$$

where $B(0)$ is the Planck intensity at the top of the atmosphere. Let us assume we can represent $B(x)$ as a set of n isothermal slabs so we can write

$$\frac{dB(x)}{dx} = \sum_{j=1}^n \Delta B_j \delta(x - x_j) \quad (17)$$

where ΔB_j is the change in B at the slab boundary x_j . Instead of specifying the x_j 's a priori as is done in the linear methods when line segments are used, they are treated as unknowns which are to be determined from the measurements along with the ΔB_j 's. Substitution of (17) into (16) gives

$$I(s) = B(0) + \sum_{j=1}^n \Delta B_j e^{-sx_j} \quad (18)$$

Thus, we are requiring that our data be fit by a series of exponentials with the free parameters $B(0)$, n values of x_j 's, and n values of ΔB_j 's. If we have measurements of I at $2n + 1$ discrete values of s , then (18) becomes a set of $2n + 1$ nonlinear equations in $2n + 1$ unknowns.

There exists an algorithm for solving such a set of equations (cf. Lanczos, 1964). This algorithm requires values of $I(s)$ at integral values of s . Physically this means that data must be sampled at integrally spaced values of the absorption coefficient. (To apply the algorithm directly, it is necessary to specify $B(0)$ by some other means. This point will be discussed further below.)

In order that our solution be physically meaningful, the x_j 's must be real and positive. If we get one or more x_j 's which do not meet this criterion, then we must conclude that there is no atmosphere containing n isothermal slabs which can produce the values of $I(s)$ at the measured points. This may be due either to noisy data or simply that n isothermal slabs are not an adequate representation of the true thermal profile.

For the thermal inversion problem being considered, it has been found empirically that the ΔB_j associated with a negative or imaginary x_j is frequently one or two orders of magnitude smaller than the remaining ΔB_j 's. When this occurs, the term in (18) involving the unacceptable x_j can be dropped and the resulting $n - 1$ isothermal slabs taken as a smoothed solution.

In practice, the atmospheric temperature profile can be approximated somewhat better by ramps than slabs. The use of ramps can be implemented by integrating (16) by parts to obtain

$$sI(s) = sB(0) + B'(0) + \int_0^{\infty} \frac{d^2 B}{dx^2} e^{-sx} dx \quad (19)$$

We assume the vertical profile can be approximated by

$$\frac{d^2 B}{dx^2} = \sum_{j=1}^n \Delta B_j' \delta(x - x_j) \quad (20)$$

where the $\Delta B_j'$'s are the discontinuous changes in slope. This results in

$$sI(s) = sB(0) + B'(0) + \sum_{j=1}^n \Delta B_j' e^{-sx_j} \quad (21)$$

We still require only $2n + 1$ measurements to determine the $2n + 2$ parameters since the additional relation required is the $s = 0$ case which is physically equivalent to the requirement that the temperature be finite at arbitrarily large optical depths. The techniques described for the isothermal slab model can now be applied.

When the nonlinear method was first developed, the upper boundary conditions were specified by what amounted to a guess. However, better techniques have now been developed. For example a differentiation of (21) with respect to s eliminates $B'(0)$

$$I(s) + sI'(s) = B(0) - \int_0^{\infty} x \frac{d^2 B}{dx^2} e^{-sx} dx \quad (22)$$

A second differentiation with respect to s eliminates $B(0)$

$$sI''(s) + 2I'(s) = \int_0^{\infty} x^2 \frac{d^2 B}{dx^2} e^{-sx} dx \quad (23)$$

In principle we could apply the algorithm directly to (23). In practice it is obviously highly undesirable to try to obtain $I'(s)$ and especially $I''(s)$ from noisy data. The method currently being employed by King involves use of (22) with $2n + 1$ measurements in the 15μ CO_2 band and an additional measurement in the atmospheric window. The value of $B(0)$ is chosen by an iterative procedure to make the temperature profile agree at the surface with the surface temperature inferred from the window measurement.

Because of the nonlinear nature of this method, error analysis is more difficult than in the linear method. The best approach seems to be application of the method to synthetic data from a wide variety of model atmospheres. Random noise can be applied to these "data" to simulate instrumentation effects. A program of this type is currently being carried out by King. Preliminary results indicate that with realistic satellite instruments one can hope to infer only two ramps consistently and in rare instances three ramps. It should be noted that a two ramp atmosphere corresponds to five independent pieces of information, which is essentially the same conclusion reached for the linear method.

We must consider whether or not a two or three ramp temperature profile provides useful information for the earth's atmosphere. Figure 7 shows a three ramp inversion of a tropical-type atmosphere, and the agreement is found to be fairly good. However, a three ramp inversion of a midlatitude profile does not give a particularly good fit in the stratosphere as shown in Figure 8. The reason for this is simply that a profile of this type cannot be well represented by three ramps.

The nonlinear method has the advantage of objective smoothing of noisy data in the sense that when noise results in a spectrum which is physically inconsistent with a temperature profile of n ramps a spectrum resulting from a smaller number of ramps is substituted. It has the disadvantage that the exponential kernel required is not capable of giving a perfect fit of actual transmissivities. It is not possible to incorporate a priori knowledge of the earth's atmosphere into the formulation. The method may be better suited for applications to radiometric data from the atmospheres of other planets where little a priori information is available. An example of an application to synthetic data for a model Martian atmosphere is shown in Figure 9.

Temperature Inversion in Cloudy Atmospheres

In the discussions given so far we have implicitly assumed that we are dealing with a completely cloud-free atmosphere. The formulations given above are equally applicable for situations in which a homogeneous cloud layer exists throughout the field of view of the measuring instrument. Cases will frequently occur in which the field of view is only partially cloud filled or contains cloud layers at several different altitudes. No well formulated treatment of this problem has yet been presented in the literature although techniques are currently being developed by Smith (private communication).

CONSTITUENT INVERSION

Once we have obtained the temperature profile from an absorption band of a uniformly mixed atmospheric constituent, we can in principle infer the vertical distribution of nonuniformly mixed optically active gases which are in local thermodynamic equilibrium with the other atmospheric constituents.

For purposes of formulating this problem, it is convenient to use the integration by parts form of the transfer equation

$$I(\nu) = B(\nu, 0) + \int_0^x \frac{\partial B}{\partial x} [\nu, T(x)] \tau [\nu, u(x)] dx \quad (24)$$

where $u(x)$ is the mass of the optically active gas between level x and the top of the atmosphere. Given measurements of $I(\nu)$ and $T(x)$, (24) is a nonlinear integral equation for the unknown function $u(x)$.

Since τ is generally a strong function of ν unlike the Planck function in the temperature inversion problem, a somewhat different approach is required. One possible approach suggested by King (1964) is to do the temperature inversion in both a CO_2 band and an absorption band of the gas whose vertical distribution is to be determined. The former will give temperature as a function of pressure and the latter will give temperature as a function of the absorber mass of the gas whose distribution is unknown. The temperature can then be eliminated from the two relations to obtain absorber mass as a function of pressure. The possibility of applying this technique to measurements in the rotational bands of water vapor is currently being investigated by Yamamoto (private communication). Another method involving the iterative comparison of calculated and measured intensities in relatively broad spectral bands is being pursued at ESSA (Smith, private communication). The object of the method is to obtain relatively crude two-parameter water vapor and temperature distributions in the lower troposphere.

Still another approach is being pursued in our laboratory. The method involves an attempt at direct linearization of (24). It is assumed that the transmission function appearing in (24) can be approximated by

$$\tau[\nu, u(x)] \simeq \tau[\nu, u_0(x)] + \left(\frac{\partial \tau}{\partial u}\right)_{u_0(x)} \delta u(x) \quad (25)$$

where $u_0(x)$ is some initial guess at the absorber distribution. When (25) is substituted into (24) we obtain an integral equation of the first kind in the correction function $\delta u(x)$

$$I(\nu) - I_0(\nu) = \int_0^{x_s} K_0(\nu, x) \delta u(x) dx \quad (26)$$

where $I_0(\nu)$ is the calculated intensity from the distribution $u_0(x)$ and the kernel is given by

$$K_0(\nu, x) = \frac{\partial B}{\partial x}(\nu, x) \left(\frac{\partial \tau}{\partial u}\right)_{u_0(x)} \quad (27)$$

The techniques developed for the temperature inversion problem can now be applied to (26). The process can be iterated as many times as desired. However, when $I(\nu) - I_0(\nu)$ becomes smaller than the probable error of the measurements nothing further can be gained by iteration.

Application of this method to the problem of obtaining water vapor distributions from the 6.3μ H_2O band has been made. The results of application of this method to synthetic data from a model atmosphere are shown in Figure 10. For application to the water vapor problem, $\delta(\log u^*)$ is used rather than δu where u^* is the reduced absorber mass. The curvature of the transmissivity as a function of $\log u^*$ is small in the region of interest so even though the correction $\delta(\log u^*)$ may be relatively large, (25) is still a fairly good approximation.

For this calculation, the transmissivities of Möller and Raschke (1964) were used. The temperature profile used in the model and the resulting kernel functions are shown in Figure 11. A two parameter water vapor distribution was assumed. In terms of the mixing ratio

$$q(p) = q_s \left(\frac{p}{p_s} \right)^k \quad (28)$$

where q_s and k are the parameters to be determined.

The method has also been applied to data obtained with an IRIS instrument flown on a balloon at Palestine, Texas, on 8 May 1966. An average of thirteen spectra distributed throughout the day was employed. The inferred distribution, which can be regarded as an average over the day, is shown in Figure 12. Radiosonde measurements taken at the stations nearest Palestine are shown for comparison.

Although only water vapor has been mentioned in detail here, the method should also be applicable to O_3 in the earth's atmosphere.

More sophisticated representations of $u(x)$ than those indicated above can of course be employed. However, the more degrees of freedom $u(x)$ is allowed, the more severe the stability problem will become. One obvious approach would be the application of empirical orthogonal functions. Whether or not water vapor is statistically stable enough to make the use of such functions practical is a subject of investigation.

SUMMARY

The inverse problem of radiative transfer for planetary atmospheres has been reviewed with emphasis placed on the temperature and water vapor inversion problems for the terrestrial atmosphere. The general instability of the problem is the main factor to be dealt with in formulating a method of solution. No entirely satisfactory method has been developed for the solution of the problem. However, a few methods showing some promise have been formulated.

The linear temperature inversion method developed by Wark and his associates seems most promising for applications to the problem of the acquisition of vertical temperature profiles in the terrestrial atmosphere on a global basis. When used in conjunction with empirical orthogonal functions, the method allows the use of a maximum of a priori information on the vertical atmospheric structure. However, since the method is essentially statistical, we must expect results of a statistical nature. We cannot hope to recover profiles which may be strongly atypical. The success of the method will depend largely on our ability to estimate satisfactorily the temperature covariance. Best results can be expected for geographical areas and seasons for which the atmospheric structure is stable in a statistical sense. It would appear that the problem of calculating covariance functions requires further work.

The only constraint imposed in the nonlinear method of King is that the temperature profile be represented by a series of ramps. It does not utilize additional a priori information, and the smoothing procedure is objective. Because of the nonlinear nature of the method, error analysis is somewhat difficult. The best approach seems to be by way of application of the method to a wide variety of climatological profiles. Such an analysis is currently being undertaken by King and his associates. Preliminary results indicate two ramp inferences can generally be expected. It would appear that this method may be most applicable to other planetary atmospheres for which a minimum amount of a priori information is available and only gross structure is sought.

Recently the problem of inferring vertical water vapor distribution has been given some consideration. Work completed thus far indicates at least the gross features of the vertical distribution can be obtained from data taken with instruments such as the IRIS. The maximum amount of information which can be extracted is a subject of current investigations. No detailed analysis of the problem of ozone inversion has been made to our knowledge. However, methods developed for the water vapor inversion problem should be applicable, at least in principle.

The usefulness of information obtained on the vertical structure of the earth's atmosphere on a global basis will ultimately have to be determined by the meteorologists who wish to make use of the data. While analyses based on model atmospheres and data obtained from balloon flights at one or a few locations is useful in the development of techniques, only after satellite data covering a wide variety of seasons and geographic locations becomes available can the practicality of the approach be established.

ACKNOWLEDGMENTS

The author is indebted to V. Kunde, NASA/GSFC, and to S. Roland Drayson, University of Michigan for providing the CO₂ transmissivities employed in this paper. Dr. Edwin Florance of the GCA Corporation, kindly provided the examples of nonlinear inversions.

REFERENCES

- Alishouse, J. C., L. J. Crone, H. E. Fleming, F. L. Van Cleef, and D. Q. Wark, A Discussion of Empirical Orthogonal Functions and Their Application to Vertical Temperature Profiles, to appear in Tellus (1966).
- Bellman, R., R. E. Kalaba, and J. A. Lockett, Numerical Inversion of the Laplace Transform, American Elsevier Publishing Company, New York (1966).
- King, Jean I. F., Inversion by Slabs of Varying Thickness, J. of Atmos. Sci., 21, 324-326 (1964).
- King, Jean I. F., Meteorological Inferences from Satellite Radiometry. I, J. of Atmos. Sci., 20, 245-250 (1963).
- Möller, F. and E. Raschke, Evaluation of TIROS III Radiation Data, Final Report, NASA Research Grant NsG-305, March 1964.
- Obukhov, A. M., The Statistically Orthogonal Expansion of Empirical Functions, Izvestiya, Seriya Geofizicheskaya, Akademiya Nauk, SSSR, No. 3, 432-439 (1960). (English translation by the AGU, Nov. 1960, 288-291).
- Phillips, D. L., A Technique for the Numerical Solution of Certain Integral Equations of the First Kind, J. Association for Computing Machinery, 9, 84-97 (1962).

Twomey, S., On the Numerical Solution of Fredholm Integral Equations of the First Kind by Inversion of the Linear System Produced by Quadrature, J. Association for Computing Machinery, 10, 97-101 (1963).

Twomey, S., Indirect Measurements of Atmospheric Temperature Profiles from Satellites: II. Mathematical Aspects of the Inversion Problem, Monthly Weather Review, 94, 351-361 (1966).

Wark, D. Q., and H. E. Fleming, Indirect Measurements of Atmospheric Temperature Profiles from Satellites: I. Introduction, Monthly Weather Review, 94, 351-361 (1966).

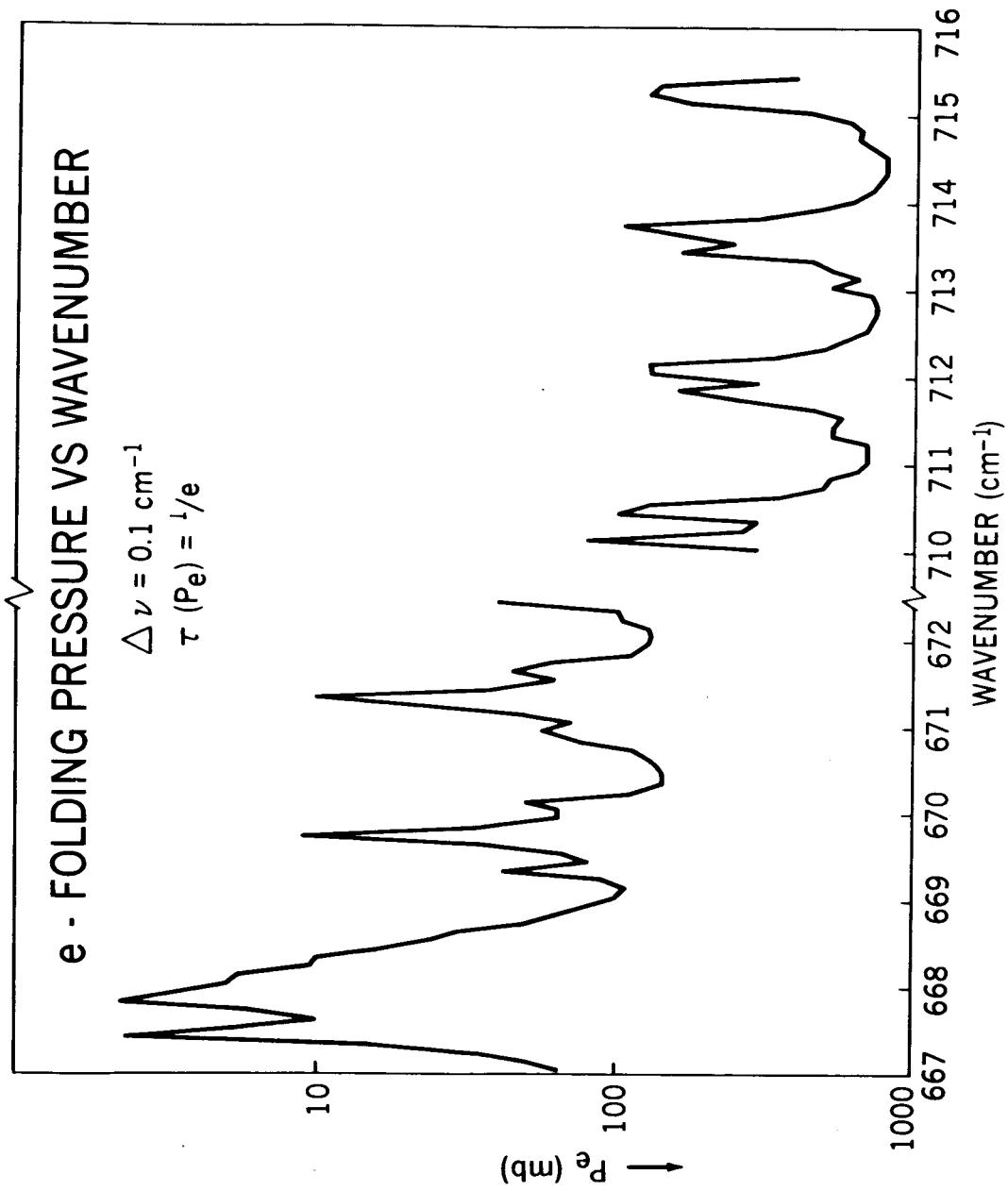


Figure 1. Fifteen micron carbon dioxide absorption band e-folding pressure levels

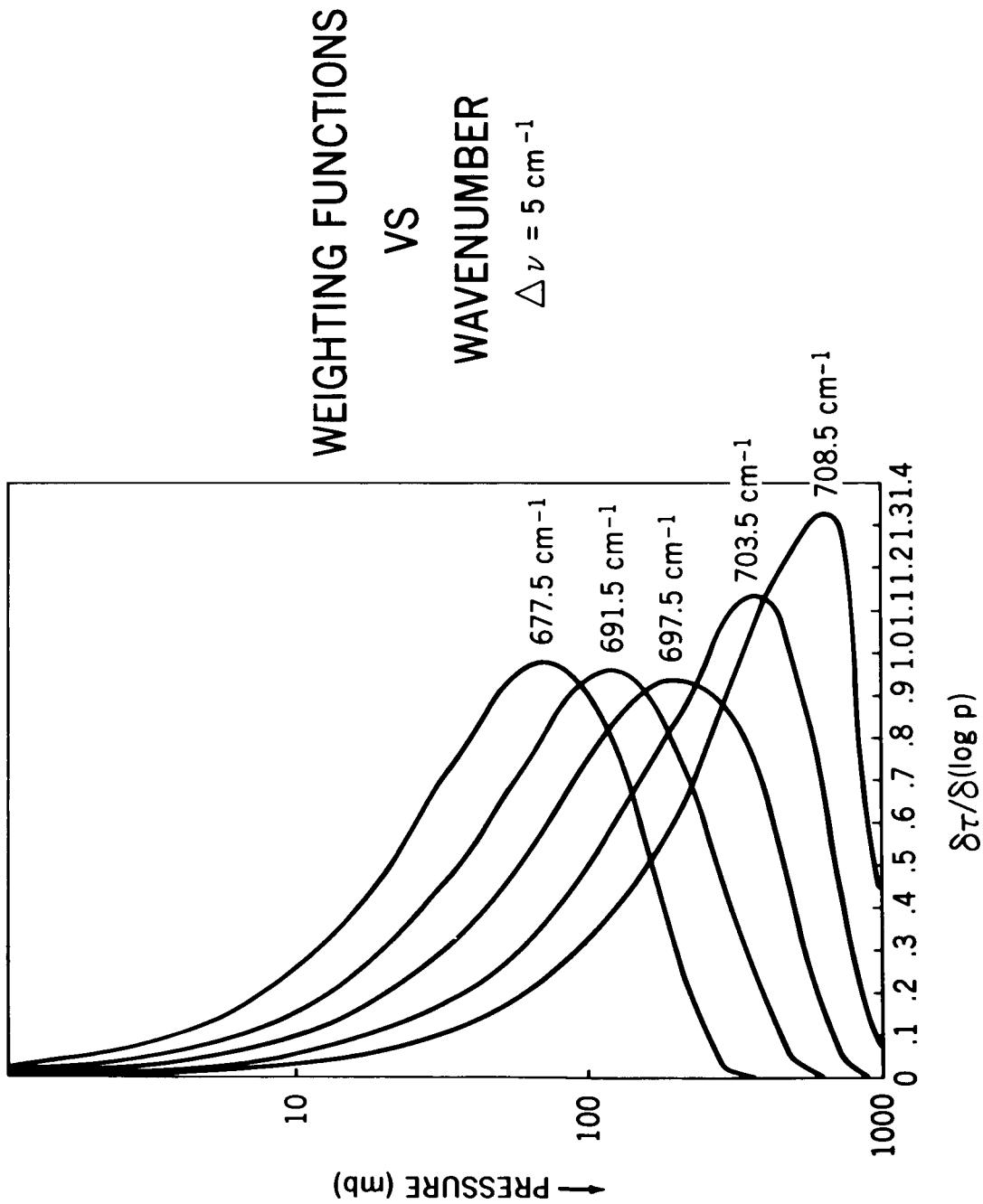


Figure 2. Fifteen micron carbon dioxide absorption band weighting functions. The 5 cm^{-1} wide spectral intervals were chosen with mid-points at the indicated wavenumbers.

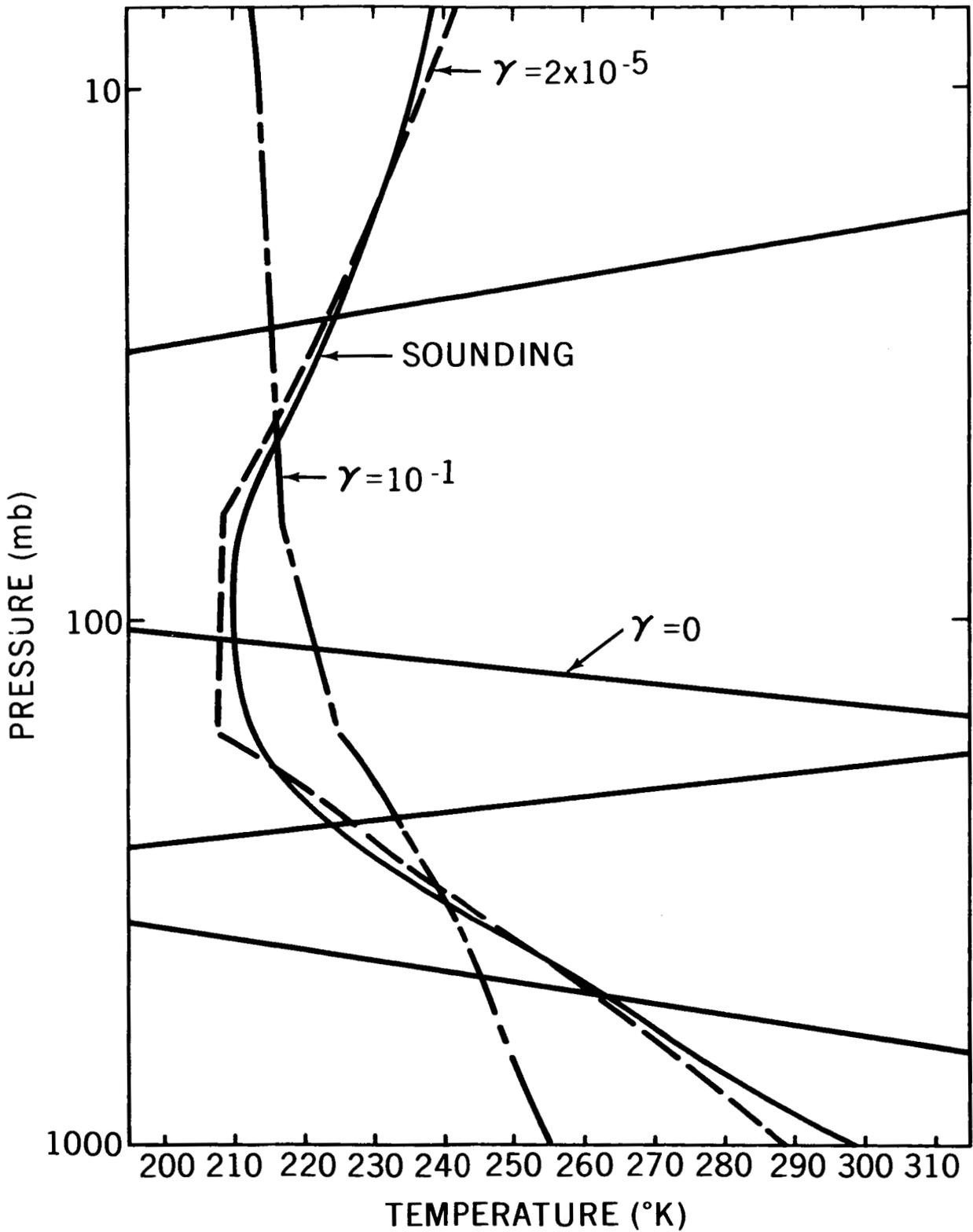


Figure 3. Temperature inversion of IRIS balloon flight data. The effects of the Lagrangian multiplier γ on the inversion is demonstrated.

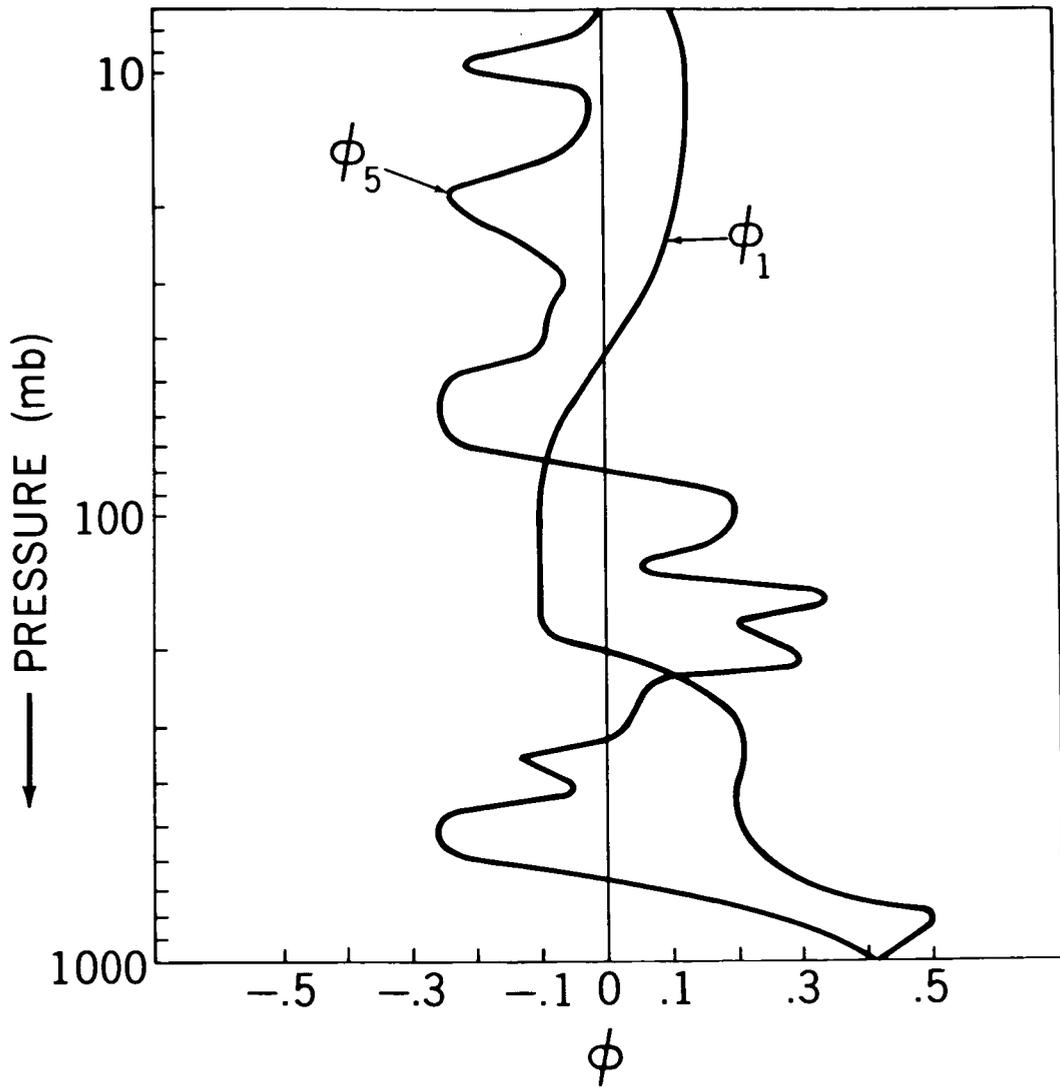


Figure 4. Two members of a set of empirical orthogonal functions. The function set was constructed from an ensemble of twenty-five temperature profiles

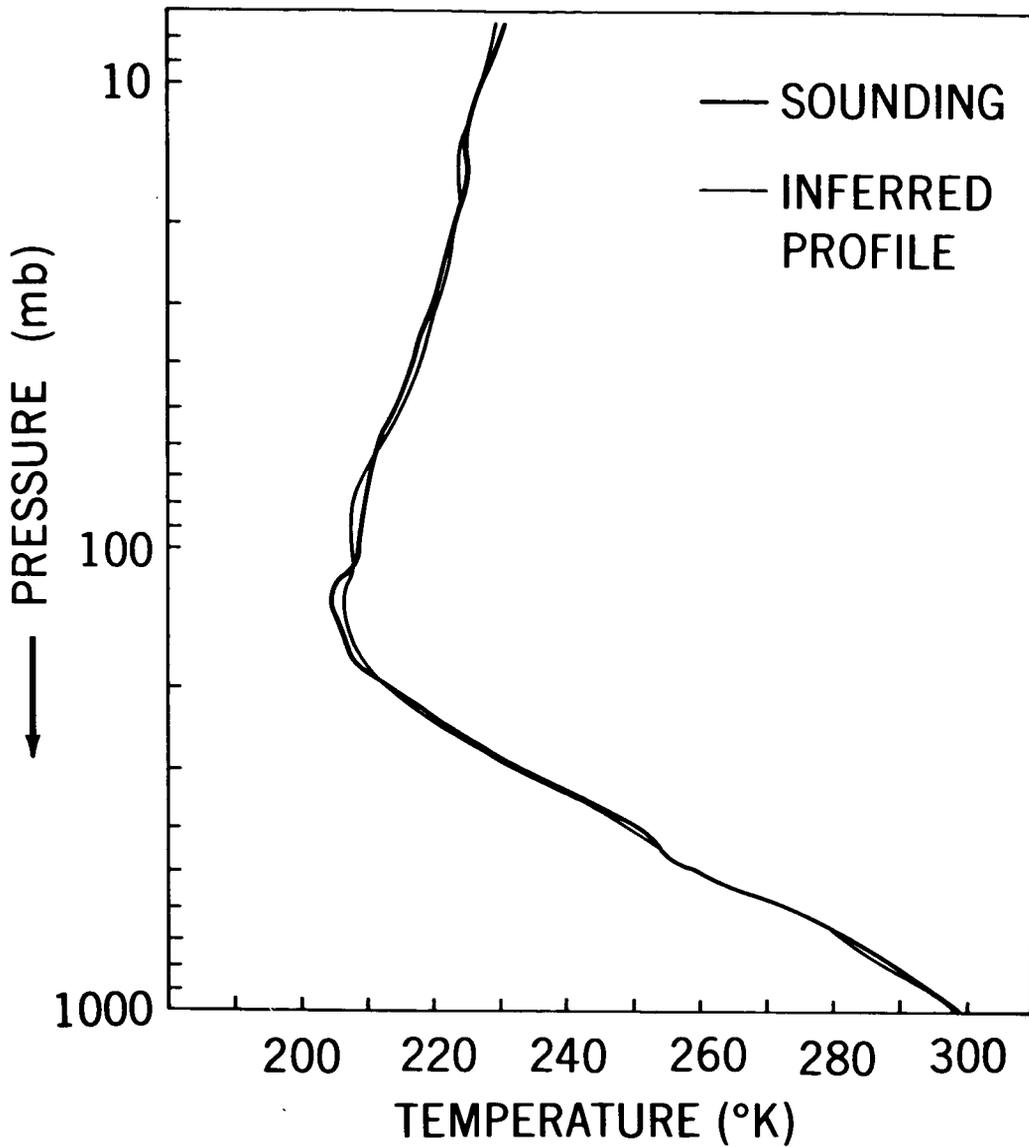


Figure 5. Synthetic data inversion in terms of empirical orthogonal functions. The sounding employed was a member of the ensemble used in constructing the function set

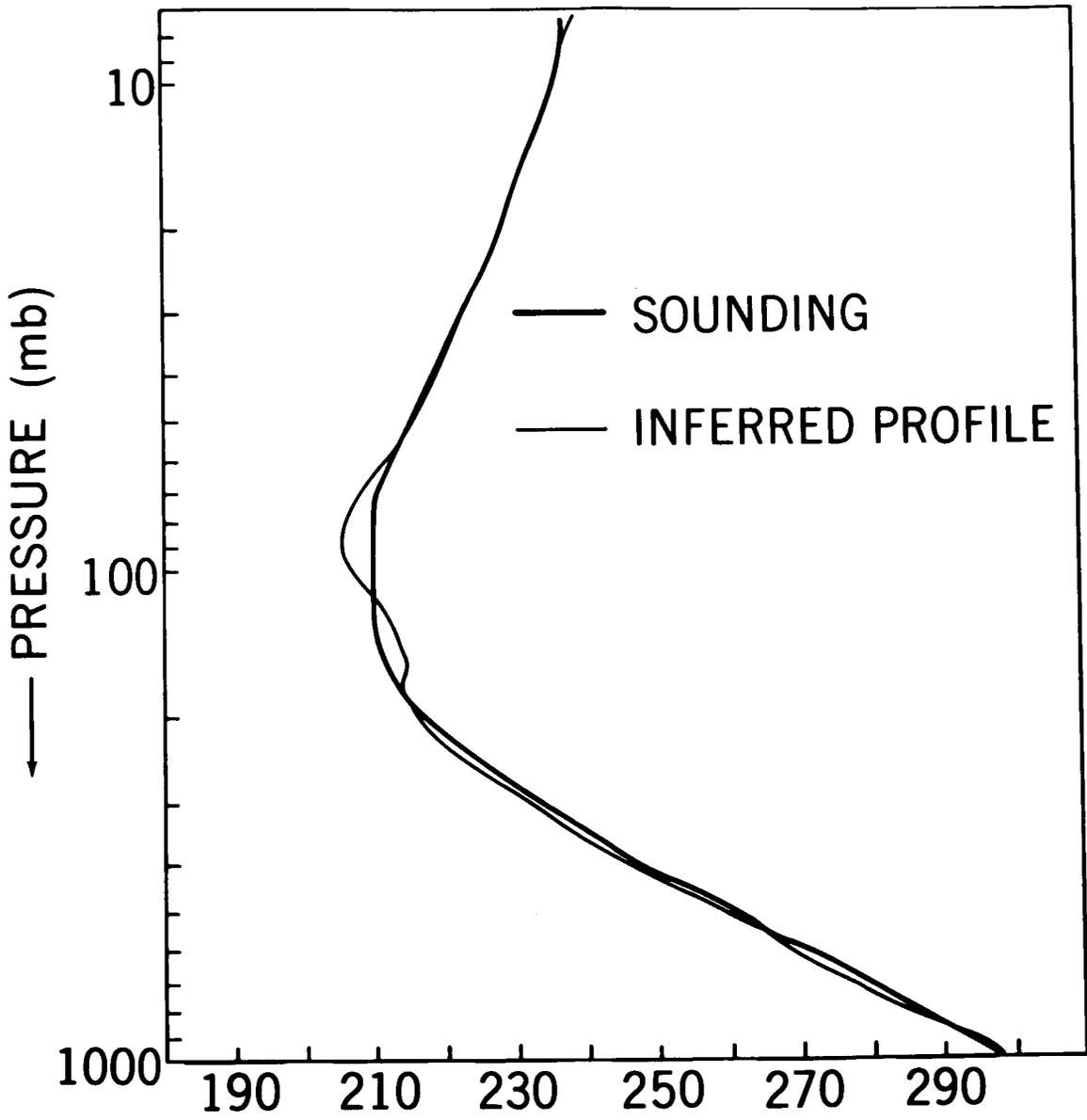


Figure 6. Synthetic data inversion in terms of empirical orthogonal functions. The sounding employed was not a member of the ensemble used in constructing the function set

NONLINEAR INVERSION OF SYNTHETIC DATA

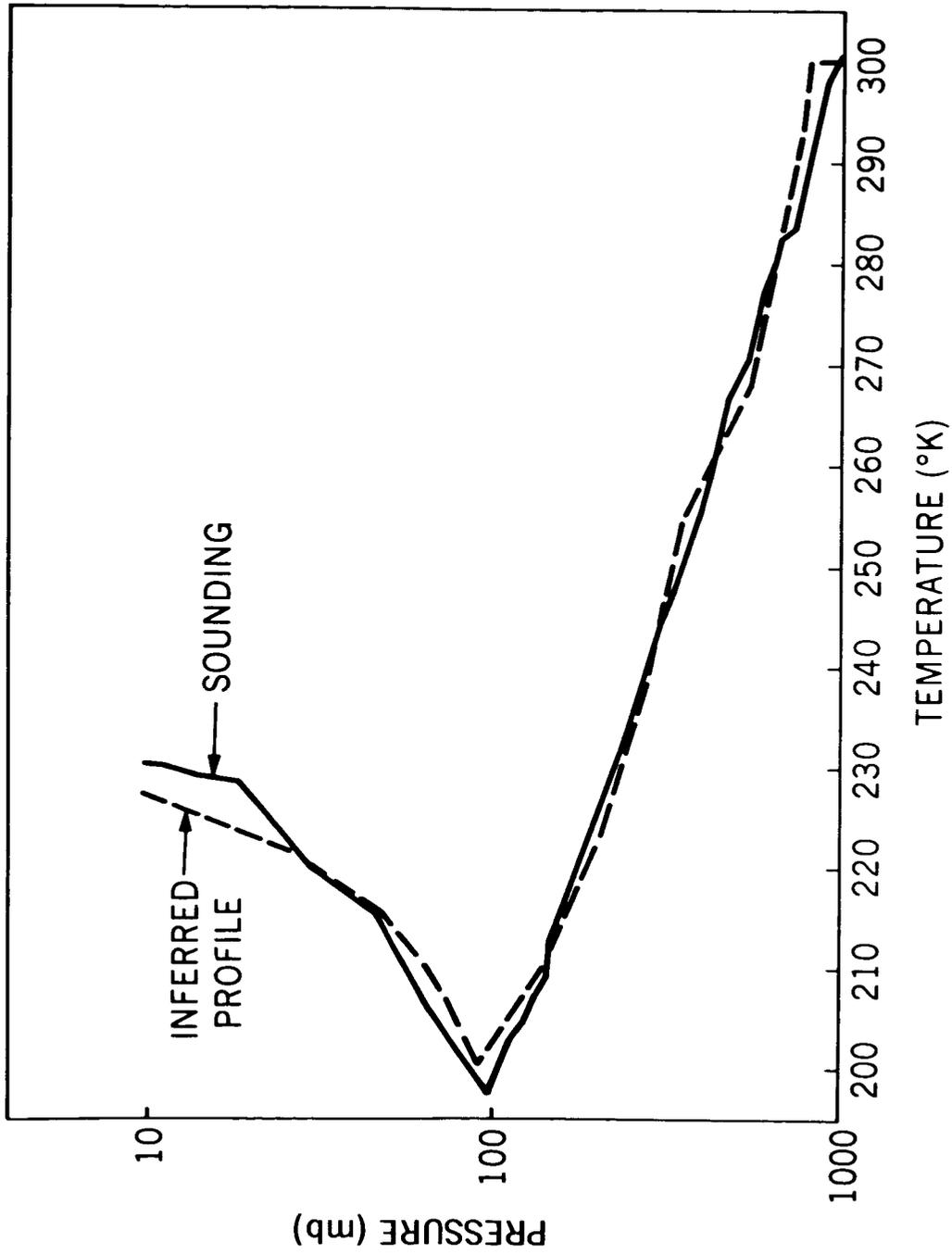


Figure 7. An application of the nonlinear inversion method to synthetic data. This illustrates the behavior of the method when applied to a tropical profile

NONLINEAR INVERSION OF STANDARD ATMOSPHERE

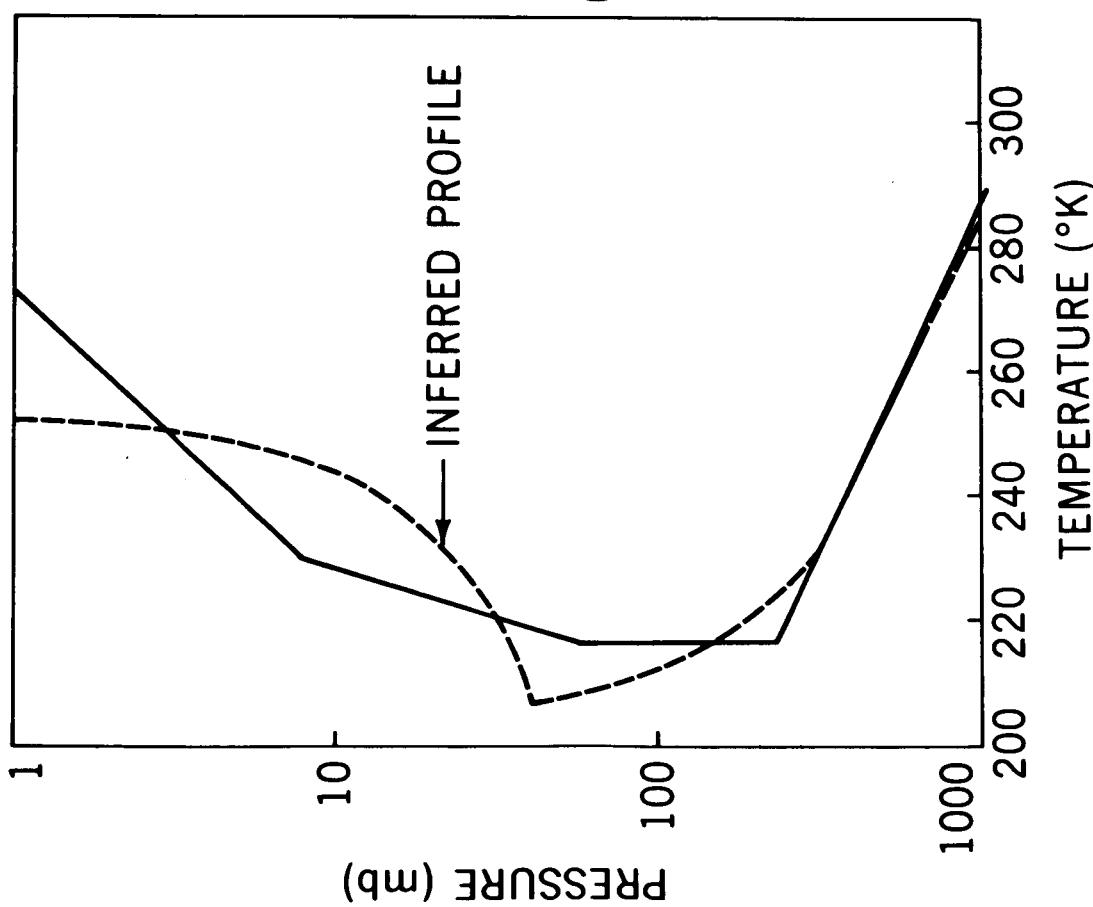


Figure 8. An application of the nonlinear inversion method to synthetic data. The behavior of the method when applied to a midlatitude profile is illustrated

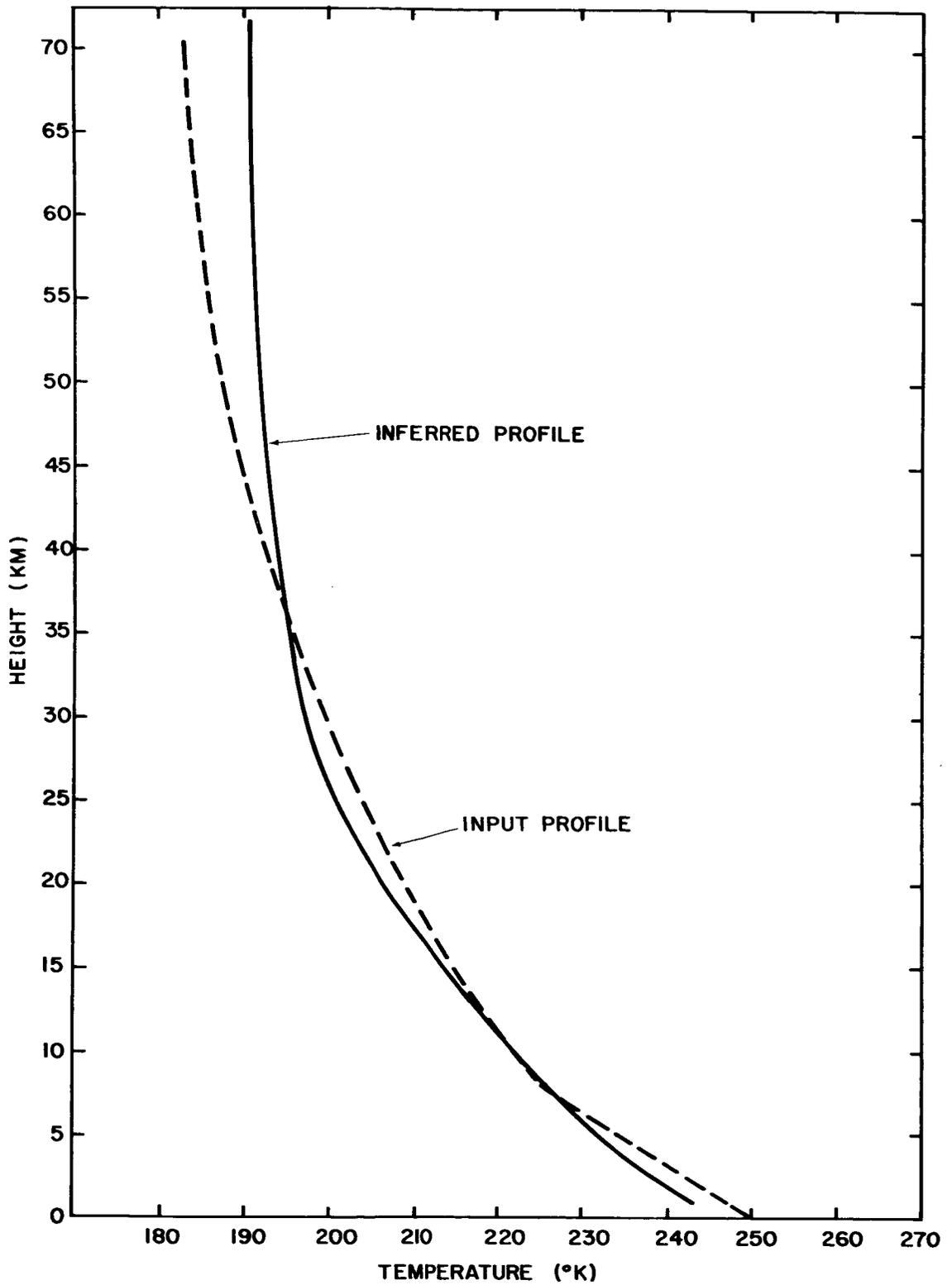


Figure 9. Nonlinear inversion of a model Martian atmosphere. IRIS-type data was synthesized numerically

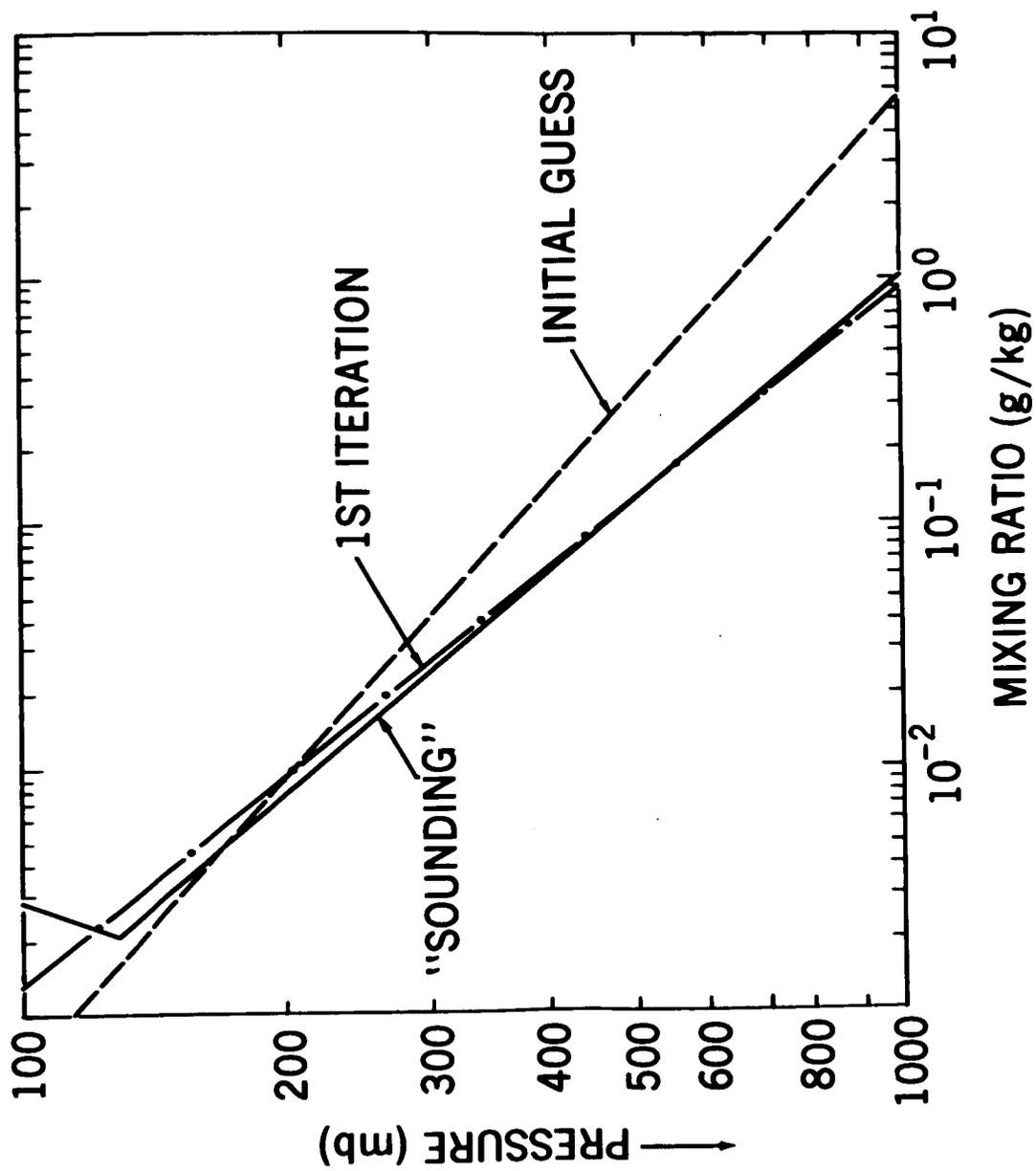


Figure 10. Water vapor inversion of synthetic data. The "sounding" employed is an idealized model

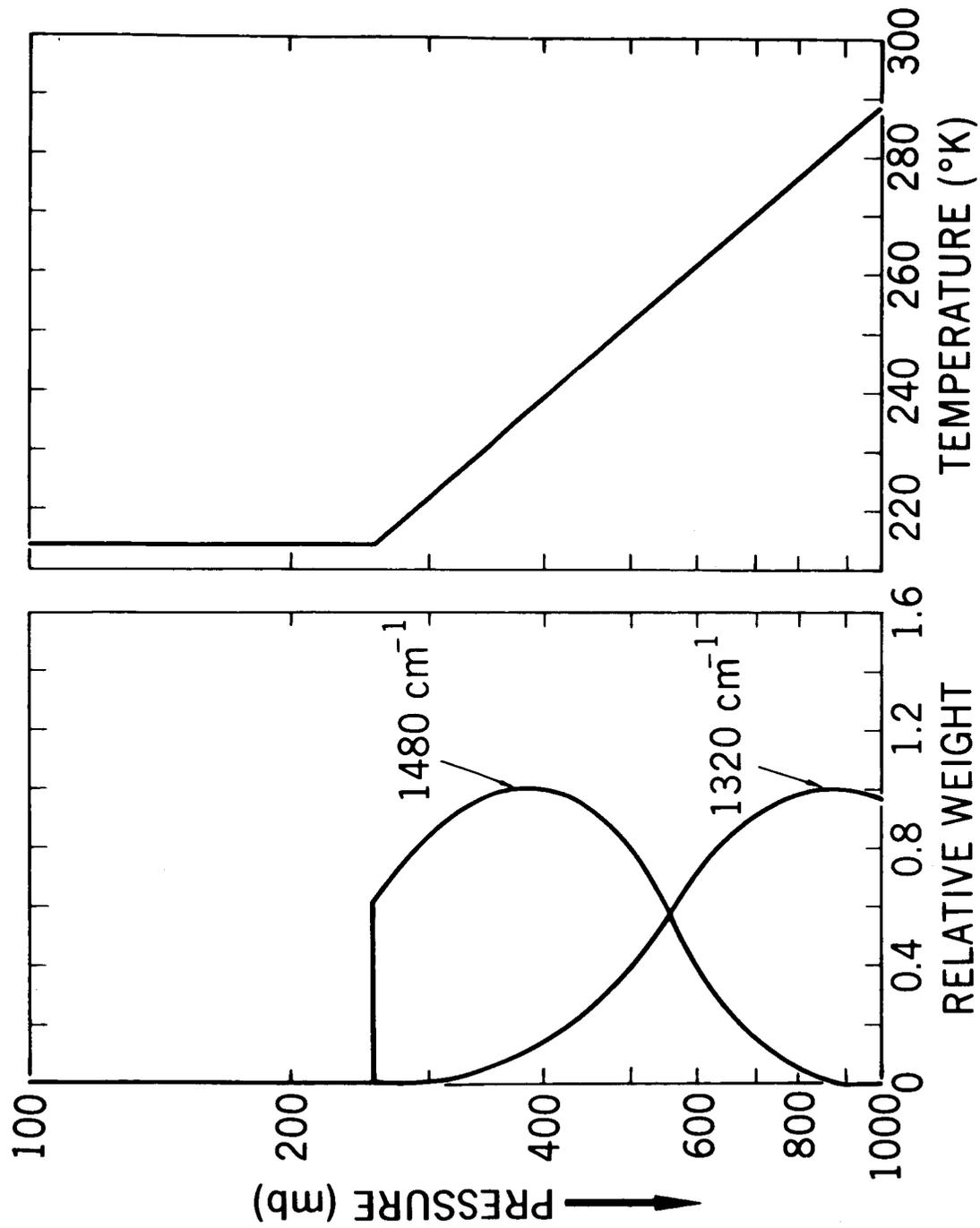


Figure 11. Weighting functions employed in the water vapor inversion shown in Figure 9. The temperature profile employed in the model is also shown

BALLOON DATA INVERSION OF WATER VAPOR

AVE. OF 14 AVE. OF 13 SPECTRA

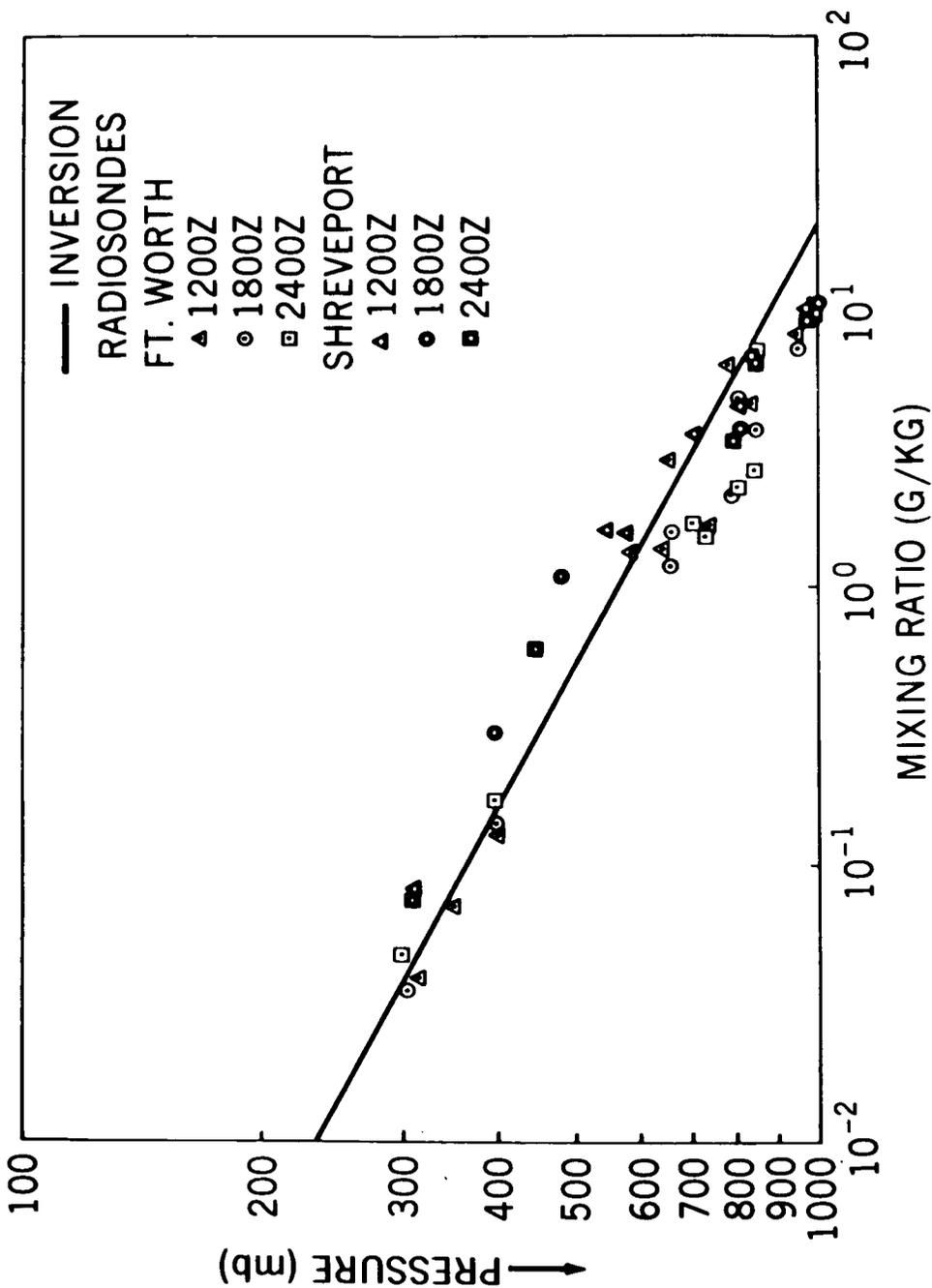


Figure 12. Water vapor inversion of IRIS balloon data. Radiosonde data from nearby stations are shown for comparison