School of Physical Sciences

Ozone concentration profile retrieval from ground-based high-resolution thermal infrared spectra

Paul van Delst

This thesis is presented as part of the
requirements for the award of the
Degree of Doctor of Philosophy
of
Curtin University of Technology

October 1996

Abstract

Simulations of the Atmospheric Emitted Radiance Interferometer (AERI), a ground-based, high-resolution infrared detection system, are used to produce retrieved atmospheric ozone concentration profiles. A line-by-line transmittance model, FASCD3P, is used for the forward model and a maximum likelihood retrieval scheme is employed for the inverse model. An *a priori* data set consisting of 83 midlatitude winter ozone sondes is used to condition the inversion. Three iterations are required to reduce the radiance residuals to less than the instrument noise. The retrieval accuaracy below 300mb is within 25% of truth. Above 300mb, variance within the *a priori* data is the dominant source of retrieval error. This is due to the number of retrieved layers (27) being higher than the amount of independent information present in the radiance spectra (~4) so much of the retrieval information above 300mb comes from the a priori data.

Acknowledgments

For all his help with my thesis work and in keeping my sometimes flagging spirits up, I would like to express my gratitude and appreciation to my supervisor, Associate Professor Mervyn J. Lynch. Thanks also go to Associate Professor Brian White for his assistance with the sometimes subtle mathematics of inversion theory.

My heartfelt thanks also goes to Professor William Smith, Dr. Tom Achtor, Dr. Hank Revercomb, Dr. Bob Knuteson, Dr. Paul Menzel and everyone at CIMSS/SSEC for the opportunity to work with the HIS group at the University of Wisconsin-Madison.

To my fiancee, Leah, and my mother and father for all their love and support, a simple thank you is not nearly enough.

Contents

List of Figures

List of Tables

1. Introduction

To accurately model the radiative, dynamical and chemical processes in the Earth's atmosphere, its composition, both spatially and temporally, must be well known. The realisation of global warming has lead to numerous studies on the impact of trace gases on the dynamic and radiative balance of the atmosphere and subsequent implication in climate change. Studies of the radiative forcing of trace gases have shown that increasing CO_2, CH_4, N_2O, and chlorofluorocarbons (CFCs) will cause a warming of the surface-troposphere system and that ozone appears to cause an overall cooling effect due to its depletion in the stratosphere by CFCs (Lacis *et al.* 1990, Hansen *et al.* 1989, Ramaswamy *et al.* 1992). There is much debate about the effect of ozone perturbations, i.e. surface warming or cooling, due mainly to the uncertainties in ozone vertical profile measurements (e.g. Ramanathan *et al.* 1985 predicts a warming effect due to decreased stratospheric ozone).

The use of general circulation models (GCMs) to estimate the radiative forcing in the presence of climate feedbacks have indicated significant warming due to trace gas greenhouse effects (Wang *et al.* 1991a; Wang *et al.* 1991b). The global temperature record over the past century does show a warming trend (Ellsaesser *et al.* 1986; Jones *et al.* 1986; Hansen and Lebedeff 1987) but this trend cannot be definitively attributed to the increase of atmospheric trace gas concentrations. Despite the apparent warming trend there are detractors to the global warming scenario (Lindzen 1990; Michaels and Stooksbury, 1992). The discrepancies in the modeled and the observed effects of trace gas warming can be attributed to a certain extent to processes in the atmosphere and ocean not yet fully understood. These include the effect on the climatic response time due to uncertainties regarding oceanic thermal inertia and mixing rates (Hansen *et al.* 1985; Wigley 1987), the effect of aerosols (natural and anthropogenic) leading to a likely cooling effect of the same order of magnitude as the calculated trace gas warming (Wang *et al.* 1986, Hansen and Lacis 1990, Kaufman *et al.* 1991, Wigley 1991, Charlson *et al.* 1992) and the climatic feedback due to clouds (Wetherald *et al.* 1988, Mitchell *et al.* 1989).

The accurate measurement of atmospheric trace gas concentration profiles is necessary for the validation of coupled atmosphere-ocean GCMs or even simple one-dimensional radiative-convective models : "There is an urgent need to acquire further data for climate model validation on both regional and global scales..." (Gates *et al.* 1990, Executive summary). No study on the radiative forcing of trace gases does not qualify its findings with the statement that high quality, comprehensive measurements of trace gas constituents are imperative. For global measurements and trends of atmospheric trace gases, the use of satellites is absolutely necessary. For regional analysis, a ground-based measurement network would be required to achieve the higher horizontal and temporal resolutions necessary.

This study concentrates on the determination of ozone vertical concentration profiles using simulated ground-based measurements of emitted downwelling atmospheric radiance. The quality of the atmospheric ozone measurement record to date is relatively poor, especially in the upper troposphere/lower stratosphere (Hansen *et al.* 1989). Ground-based infrared (IR) measurements contain more information on the vertical distribution of tropospheric ozone than current satellite-based measurement platforms, providing the means to improve the knowledge database of tropospheric ozone concentration profiles.

The ground-based IR measurements used for retrieval of ozone profiles are in the thermal longwave (980-1080cm^{-1}, 9.6μm) region of the spectrum. The spectra used in this study closely simulate, both radiometrically and in spectral resolution, those produced by the Atmospheric Emitted Radiance Interferometer (AERI) (Knuteson *et* *al.* 1991). A noise component was also added to the simulated spectra based on the nominal noise performance of the AERI. The radiance/transmittance code FASCOD3P (Clough *et al.* 1981) was used to synthesise the spectra using atmospheric temperature, water vapour and ozone profile data from ozonesondes released during the First International Satellite Cloud Climatology Program (ISCCP) Regional Experiment II / Spectral Radiation Experiment (FIRE II/SPECTRE) field experiment in Coffeyville, Kansas in November/December 1991.

The downwelling radiative transfer equation is perturbed and linearised to obtain a matrix equation relating IR radiance perturbations to ozone mixing ratio perturbations with clear sky conditions assumed. A maximum likelihood inversion scheme is employed to obtain ozone vertical profiles from the synthetic IR spectra for midlatitude winter conditions. *A priori* data used to constrain the solution come from a database of 91 ozonesondes taken from a climatological set of 408 sondes and from ozonesondes released during FIRE II/SPECTRE. The outcomes of the research demonstrate that the approach and methodology adopted are directly applicable to the retrieval of ozone mixing ratio profiles from high resolution IR measurements.

This dissertation is divided into six sections. Chapter two reviews the literature on retrieval theory as applied to IR measurements, typically for space-based temperature and water vapour profile retrievals, including the initial estimate selection procedure. Chapter three details the data manipulation in the forward model calculations, the methodology and the validation of the linearised form of the general radiative transfer equation for an upward looking sensor geometry. The application of a maximum likelihood retrieval scheme to synthetic spectra and the analysis of the retrieval products are described in chapter four. Chapter five summarises the results and describes recommendations for further research on trace gas retrieval schemes. All references cited are shown in chapter six.

2. Retrieval Theory

The mathematical representation of the physical processes involved with absorption and emission of infrared (IR) radiation in the atmosphere is given by the radiative transfer equation (RTE). For the ground-based, or uplooking, observer the downwelling RTE for wavenumber ν is given by

$$R(\nu,p_s)=\int_{p_{top}}^{p_s} B(\nu,T(p))\frac{d\tau(\nu,p)}{dp}dp, \tag{2.1}$$

where

$$\tau(\nu,p)=\exp\left[-\frac{1}{g}\int_{p}^{p_s}\sum_{i}^{Nm}k_i(\nu,p')q_i(p')dp'\right], \tag{2.2}$$

and $R(\nu,p_s)$ = measured downwelling radiance at the surface p_s,

$B(\nu,T(p))$ = Planck radiance at pressure p for atmospheric temperature T,

$\tau(\nu,p)$ = atmospheric transmittance from pressure p to the surface p_s,

$k_i(\nu,p)$ = absorption coefficient for molecular species i at pressure p,

$q_i(p)$ = mass mixing ratio of molecular species i at pressure p,

p_{top} = lowest pressure level included in retrieval,

p_s = surface pressure,

Nm = total number of molecular species, and

g = acceleration due to gravity.

The retrieval of atmospheric parameters such as temperature, *T(p)*, or absorber concentration, *q(p)*, from ground- or satellite-based IR radiometric measurements generally is a non-linear problem. The inversion of the RTE shown in (2.1) typically involves linearising it about some inital estimate. This is valid only when the differences between the initial estimate and the solution occur within a linear or quasi-linear regime. If the initial estimate is too far away from the solution, then any physical non-linear radiometric response may cause large differences between radiances obtained using forward linear models and the actual measurements. If non-linear forward models are used with an iterative solution scheme, the rate of convergence to a solution may be slow, the solutions may not converge, or they may not converge to the optimal solution. Convergence criteria for iterative solution schemes must be selected carefully. A non-linear response may still allow convergence, but the measure of convergence may no longer be valid. It is a necessary condition to have some *a priori* knowledge of the system under observation to condition the initial input and constrain the solution within a linear regime, and also to aid in the determination of the validity of the solution.

In linearising the (2.1) a Fredholm equation of the first kind is obtained, the perturbed RTE (PRTE), of the form

$$\delta R(\nu) = \int_{p_{top}}^{p_s} W(\nu,p)\delta x(p)dp\,, \qquad (2.3)$$

where $\delta R(\nu) = R(\nu) - R^o(\nu)$,

= the radiance perturbation at the surface for wavenumber ν,

with $R(\nu)$, $R^o(\nu)$ = final and initial radiances respectively,

x = atmospheric state parameter of interest,

$\delta x(p)$ = measure of parameter perturbation at pressure p, and

$W(\nu,p)$ = kernel (weighting function) operating on $\delta x(p)$ to produce $\delta R(\nu)$.

The terms superscripted by ^{o} indicate the initial estimate of the solution. The initial estimate may be climatological, statistical or iterative in origin depending on the technique used to invert (2.3). Note that the parameter x may be any atmospheric state parameter governed by (2.3), such as temperature or absorber gas concentration.

Equation (2.3) is readily expressed in matrix form as

$$\Delta\mathbf{r} = \mathbf{W}\Delta\mathbf{x} + \varepsilon, \tag{2.4}$$

where $\Delta\mathbf{r} = N\times1$ measurement perturbation vector, $\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}^o)$,

$e = N\times1$ vector of measurement errors, independent of $\Delta\mathbf{r}$,

$\Delta\mathbf{x} = L\times1$ solution perturbation vector,

$\mathbf{W} = N\times L$ transformation matrix, a linear operator,

N = number of spectral channels, and

L = number of solution levels (pressure levels).

This chapter reviews the literature on the retrieval of atmospheric state parameters from remotely acquired radiometric measurements. The methods employed in inversion of the linearised RTE will be examined and the procedures used to determine if the solution is "acceptable" and is an accurate representation of the state of the atmosphere under observation. The form of the transformation matrices used in the retrieval of atmospheric temperature and absorber concentration profiles and the methods of determining a suitable initial state estimate will also be examined.

2.1. Inversion methods

The problem is to invert the expression in (2.4) to obtain a perturbation solution vector $\Delta\hat{\mathbf{x}}$. The direct inversion of **W** is complicated in that it is nearly always ill-conditioned, i.e. it has small eigenvalues, due to the high degree of interdependence among the transformation functions for particular spectral channels. Thus any small change in $\Delta\mathbf{r}$, for example due to ε, may lead to numerical instability in the direct solution of (2.4). There is a significant body of literature dealing with the inversion of (2.4) in the context of radiative transfer. The techniques usually involve constraining the inversion to maintain the eigenvalues of the inverted matrix above a convergence threshold and/or optimising the initial estimate and retrieval using *a priori* information. Two methods widely used in atmospheric retrievals are described here - the constrained linear and maximum likelihood inversion schemes.

2.1.1. Constrained Linear Inversion

Twomey (1963) addressed the inversion of (2.4) using the method of Lagrangian multipliers, after Phillips (1961). Given that there are an infinite number of solutions to the physical state represented by (2.4), there must be some criteria applied to select which perturbation solution $\Delta\hat{\mathbf{x}}$ is "best". This may be achieved by minimising the residuals ($\mathbf{W}\Delta\hat{\mathbf{x}}$ - $\Delta\mathbf{r}$) in a least-squares sense subject to some constraint. If $S(\Delta\hat{\mathbf{x}})$ is a measure of the constraint placed upon $\Delta\hat{\mathbf{x}}$ then, using the method of Lagrangian multipliers, the function to minimise becomes

$$\left|\mathbf{W}\Delta\hat{\mathbf{x}} - \Delta\mathbf{r}\right|^2 + \gamma S(\Delta\hat{\mathbf{x}}), \qquad (2.5)$$

where ideally $S(\Delta\hat{\mathbf{x}}) = 0$ and γ is a Lagrangian multiplier. The function $S(\Delta\hat{\mathbf{x}})$ can be thought of as a penalty function, the form of which is dependent on the problem.

Deviations from a smooth solution or from an initial estimate of the solution are two common choices for $S(\Delta\hat{\mathbf{x}})$.

Twomey (1977) shows how different measures of non-smoothness may be incorporated into solution schemes where a smoothing matrix is determined according to the problem being studied. As an example, if the relationship between $\Delta\mathbf{x}$ and $\Delta\mathbf{r}$ was second order, then the squares of the third differences would be a representative penalty function because the third differences should tend to zero. Mathematically this translates to

$$S(\Delta\hat{\mathbf{x}}) = \sum_{i=4}^{L} (\Delta\hat{x}_{i-3} - 3\Delta\hat{x}_{i-2} + 3\Delta\hat{x}_{i-1} - \Delta\hat{x}_i)^2 , \tag{2.6a}$$

Equation (2.6a) can be recast in matrix form as

$$S(\Delta\hat{\mathbf{x}}) = \Delta\hat{\mathbf{x}}^T \mathbf{H} \Delta\hat{\mathbf{x}} , \tag{2.6b}$$

where $\mathbf{H}$ = L×L near-diagonal smoothing matrix

The total quantity to be minimised is now

$$\left|\mathbf{W}\Delta\hat{\mathbf{x}} - \Delta\mathbf{r}\right|^2 + \gamma\Delta\hat{\mathbf{x}}^T \mathbf{H} \Delta\hat{\mathbf{x}} . \tag{2.7}$$

Differentiating (2.7) with respect to the elements of $\Delta\hat{\mathbf{x}}$ and setting the result equal to zero provides the constrained linear inversion solution, namely

$$\Delta\hat{\mathbf{x}} = \left(\mathbf{W}^T\mathbf{W} + \gamma\mathbf{H}\right)^{-1} \mathbf{W}^T \Delta\mathbf{r} , \tag{2.8a}$$

or, using $\Delta\hat{\mathbf{x}} = \left(\hat{\mathbf{x}} - \mathbf{x}^o\right)$, we have

$$\hat{\mathbf{x}} = \mathbf{x}^o + \left(\mathbf{W}^T\mathbf{W} + \gamma\mathbf{H}\right)^{-1}\mathbf{W}^T\left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}^o)\right). \tag{2.8b}$$

Similarly, if the deviation of a solution $\Delta\hat{\mathbf{x}}$ from an initial estimate $\Delta\mathbf{x}^o$ is chosen as the quantity to be minimised, then

$$S(\Delta\hat{\mathbf{x}}) = \sum_{i=1}^{L}(\Delta\hat{x}_i - \Delta x_i^o)^2 \quad \Rightarrow \quad S(\Delta\hat{\mathbf{x}}) = \left(\Delta\hat{\mathbf{x}} - \Delta\mathbf{x}^o\right)^T\left(\Delta\hat{\mathbf{x}} - \Delta\mathbf{x}^o\right), \tag{2.9}$$

and the expression to be minimised is

$$\left|\mathbf{W}\Delta\hat{\mathbf{x}} - \Delta\mathbf{r}\right|^2 + \gamma\left(\Delta\hat{\mathbf{x}} - \Delta\mathbf{x}^o\right)^T\left(\Delta\hat{\mathbf{x}} - \Delta\mathbf{x}^o\right). \tag{2.10}$$

Following the same procedure as before, the solution

$$\Delta\hat{\mathbf{x}} = \left(\mathbf{W}^T\mathbf{W} + \gamma\mathbf{I}\right)^{-1}\left(\mathbf{W}^T\Delta\mathbf{r} + \gamma\Delta\mathbf{x}^o\right) \tag{2.11}$$

is obtained (Twomey, 1966). These solutions are robust in the sense that they allow the eigenvalues of the inverted matrix to be conditioned to reduce numerical instability. They do, however, lead to an optimisation problem regarding the choice of the Lagrangian multiplier, γ. Without *a priori* statistics, the optimum estimate of γ is chosen by performing a retrieval a number of times and using the γ value that produces a result with residuals $\left|\mathbf{W}\Delta\hat{\mathbf{x}} - \Delta\mathbf{r}\right|$ of the same size as ε, the estimated error in $\Delta\mathbf{r}$. If the residuals are larger than ε, then the solution is over-constrained and the selected γ is too large. If the residuals are smaller than ε, then the solution is under-constrained and γ is too small. In the former case, the smoothing applied by $\gamma\mathbf{H}$ will dampen the retrieved response and in the latter, some of the random error component will have been inverted also, introducing oscillations into the solution space (Twomey, 1977).

Determination of the Lagrangian multiplier, γ, by trial and error is neither an efficient method nor does it guarantee an optimal solution. King *et al.* (1979) and King (1982) incorporate the measurement error covariance into the inversion method to introduce a form of statistical optimisation of the result. The function to minimise becomes

$$(\mathbf{W}\Delta\hat{\mathbf{x}}-\Delta\mathbf{r})^T\mathbf{C}_R^{-1}(\mathbf{W}\Delta\hat{\mathbf{x}}-\Delta\mathbf{r})+\gamma\Delta\hat{\mathbf{x}}^T\mathbf{H}\Delta\hat{\mathbf{x}}, \tag{2.12}$$

and the solution becomes

$$\Delta\hat{\mathbf{x}}=\left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}+\gamma\mathbf{H}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\Delta\mathbf{r}, \tag{2.13a}$$

or as in (2.8b)

$$\hat{\mathbf{x}}=\mathbf{x}^o+\left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}+\gamma\mathbf{H}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}^o)\right), \tag{2.13b}$$

where $\mathbf{C}_R$ = measurement error covariance matrix.

For uncorrelated measurement errors,

$$\mathbf{C}_R=s^2\mathbf{I}$$

where s^2 defines the radiance error variance, and (2.13) simplifies to (2.8) with s^2 incorporated into γ. King (1982) describes a method of selecting an optimal γ value based on a positivity constraint on the solution vector but this scheme still requires a trial and error approach. Concerned with the retrieval of aerosol size distributions from optical depth measurements, the value of γ is varied until all elements of the solution vector - aerosol size - are positive, since any negative elements constitute a non-physical result. Ben-David *et al.* (1988) incorporate a solution vector positivity constraint directly into the inversion procedure leading to a solution of the form

$$\hat{\mathbf{x}} = \mathbf{x}^o + \left[\mathbf{W}^T\mathbf{W} + \gamma_s\mathbf{H}_s + \gamma_p\mathbf{H}_p\right]^{-1}\left[\mathbf{W}^T\left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}^o)\right) + \gamma_p\mathbf{h}_p\right], \qquad (2.14)$$

where $\mathbf{H}_s$ = smoothing matrix (corresponds to **H, I** in previous equations),

$\mathbf{H}_p$ = positivity constraint matrix,

$\mathbf{h}_p$ = positivity constraint vector, and

γ_s, γ_p = Lagrangian multipliers for smoothing and positivity criteria

This method, unlike that of King (1982), introduces an extra undetermined Lagrangian multiplier, rather than providing a means of determining a preferred value for a single γ.

Smith *et al.* (1985) and Hayden (1988) use constrained linear inversion in a solution scheme to retrieve temperature and water vapour concentration profiles from GOES-VAS observations. The solution algorithm employs γ values of 1.0 and 0.1 on sequential iterations dependent on the initial estimate accuracy. If the radiances are not correctly matched in the first iteration, the value of γ is decreased from 1.0 to 0.1 to sensitise the solution. No optimisation of the choice of γ is performed.

Flittner *et al.* (1993) use (2.8) to obtain total columnar ozone amounts from atmospheric optical depth measurements. Optimal selection of γ is based upon differences in the residual error between the measurement vector, $\mathbf{r}_{obs}$, and a retrieved "measurement" vector calculated from the solution vector, $\mathbf{W}\hat{\mathbf{x}}$. If the rms differences are of the order 1-3%, then a suitable γ has been found.

The magnitude of γ may also be determined from the ratio of signal-to-noise power (Foster, 1961) leading to what is termed the minimum information scheme. This result follows directly from the use of estimation theory in inversion problems and is covered in the next section.

2.1.2. Maximum Likelihood

The maximum likelihood method arrives at a particular solution from an infinite set by maximising the probability densities associated with possible solution vectors. The maximum likelihood method described here follows the Bayesian approach described by Tarantola (1987) and Eyre (1989). It must be noted that there is an abundance of literature regarding this method - the two selected above as a basis for this review are examples of the general development of the method (Tarantola) and its application to atmospheric state parameter retrieval from radiometric measurements (Eyre). An assumption implicit in the methodology is that the observating system is subject to Gaussian error statistics. This is not a necessary requirement in estimation theory but it greatly simplifies the mathematics and is a good approximation for quasi-linear systems (Tarantola, 1987).

Consider again the problem expressed in (2.4), namely

$$\Delta\mathbf{r} = \mathbf{W}\Delta\mathbf{x} + \varepsilon,$$

and a possible solution

$$\Delta\mathbf{r} = \mathbf{W}\Delta\hat{\mathbf{x}} + e, \tag{2.15}$$

where **W** maps $\Delta\mathbf{x}$ into $\Delta\mathbf{r}$ and the random errors, ε, have a Gaussian probability density, $p(\varepsilon)$. If $\Delta\hat{\mathbf{x}}$ is a good estimate of the true solution $\Delta\mathbf{x}$ then the residual probability distributions will be identical. The probability that the the measurement vector $\Delta\mathbf{r}$ is represented by the solution $\Delta\hat{\mathbf{x}}$ is then given by (Jackson and Matsu'ura, 1985)

$$p(\Delta\mathbf{r}|\Delta\hat{\mathbf{x}}) = p(e) = p(\varepsilon). \tag{2.16}$$

By maximising the probability distribution of (2.16), the most likely estimator of the solution $\Delta\mathbf{x}$ is obtained. Consider the Gaussian probability density function (PDF) for (2.16) (Tarantola, 1987),

$$p(\Delta\mathbf{r}|\Delta\hat{\mathbf{x}}) = \left((2\pi)^N |\mathbf{C}_R|\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{r}_{obs} - \mathbf{r}(\hat{\mathbf{x}}))^T \mathbf{C}_R^{-1}(\mathbf{r}_{obs} - \mathbf{r}(\hat{\mathbf{x}}))\right] \tag{2.17}$$

where $|\mathbf{C}_R|$ = determinant of the radiance error covariance matrix,

$\mathbf{r}_{obs}$ = measured radiance (including measurement error),

$\mathbf{r}(\hat{\mathbf{x}})$ = calculated radiance for $\hat{\mathbf{x}}$ from a forward model.

While optimisation of (2.17) is a simple least squares problem (i.e. minimisation of the exponent) this will not necessarily provide the most probable solution, given that the information about $\Delta\mathbf{r}$ as the solution may not be unique. The high degree of linear dependence between the columns of the matrix **W** does not allow for complete resolution of the solution vector. The incorporation of *a priori* information about the solution is required. Tarantola (1987) provides a rigorous treatment of the incorporation of *a priori* information into a retrieval scheme using Bayesian methods. The conditional PDF of the solution $\Delta\hat{\mathbf{x}}$, given the measurements $\Delta\mathbf{r}$, may be represented by using Bayes' rule

$$p(\Delta\hat{\mathbf{x}}|\Delta\mathbf{r}) = \frac{p(\Delta\mathbf{r}|\Delta\hat{\mathbf{x}})p(\Delta\hat{\mathbf{x}})}{\int p(\Delta\mathbf{r}|\Delta\mathbf{x})p(\Delta\mathbf{x})d(\Delta\mathbf{x})} = \frac{p(\Delta\mathbf{r}|\Delta\hat{\mathbf{x}})p(\Delta\hat{\mathbf{x}})}{p(\Delta\mathbf{r})}, \qquad (2.18)$$

where $p(\Delta\hat{\mathbf{x}})$ = prior PDF of $\Delta\hat{\mathbf{x}}$ (from Tarantola, 1987; Walpole and Myers, 1989).

The denominator normalises the probability over $\Delta\mathbf{x}$ space and (because $\Delta\mathbf{r}$ is a measured quantity) is equal to unity. Similarly to (2.17),

$$p(\Delta\hat{\mathbf{x}}) = \left((2\pi)^L|\mathbf{C}_X|\right)^{-1/2}\exp\left[-\frac{1}{2}\left(\hat{\mathbf{x}}-\mathbf{x}^o\right)^T\mathbf{C}_X^{-1}\left(\hat{\mathbf{x}}-\mathbf{x}^o\right)\right], \qquad (2.19)$$

where $\mathbf{C}_X$ = error covariance of *a priori* information of $\Delta\mathbf{x}$, and

$\mathbf{x}^o$ = *a priori* estimate of $\mathbf{x}$.

Substituting (2.17) and (2.19) into (2.18) gives the *a posteriori* conditional PDF of the solution incorporating *a priori* information about the variance of the true solution $\mathbf{x}$ about $\mathbf{x}^o$, namely

$$p(\Delta\hat{\mathbf{x}}|\Delta\mathbf{r}) = \left((2\pi)^L|\mathbf{C}_E|\right)^{-1/2}\exp\left[-\frac{1}{2}\left(\left(\hat{\mathbf{x}}-\mathbf{x}^o\right)^T\mathbf{C}_X^{-1}\left(\hat{\mathbf{x}}-\mathbf{x}^o\right)+\left(\mathbf{r}_{obs}-\mathbf{r}(\hat{\mathbf{x}})\right)^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\hat{\mathbf{x}})\right)\right)\right] \qquad (2.20)$$

where $\mathbf{C}_E$ = solution error covariance.

Thus by maximising the expression (2.20) (i.e. minimising the exponent) an optimal estimate of $\Delta\mathbf{x}$, and thus $\mathbf{x}$, is obtained which is the solution $\Delta\hat{\mathbf{x}}$. From (2.20), a penalty function $S(\hat{\mathbf{x}})$ may be defined, as in Tarantola (1987) or Eyre (1990), where

$$S(\hat{\mathbf{x}}) = \frac{1}{2}\left[\left(\hat{\mathbf{x}}-\mathbf{x}^o\right)^T\mathbf{C}_X^{-1}\left(\hat{\mathbf{x}}-\mathbf{x}^o\right)+\left(\mathbf{r}_{obs}-\mathbf{r}(\hat{\mathbf{x}})\right)^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\hat{\mathbf{x}})\right)\right]. \qquad (2.21)$$

Minimisation of (2.21) by differentiation with respect to the elements of $\hat{\mathbf{x}}$ gives (Eyre, 1989),

$$\left(\frac{\partial S}{\partial \hat{\mathbf{x}}}\right) = \mathbf{C}_X^{-1}\left(\hat{\mathbf{x}} - \mathbf{x}^o\right) - \mathbf{W}^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs} - \mathbf{r}(\hat{\mathbf{x}})\right) = 0, \qquad (2.22)$$

where **W** is the transformation operator, as defined in (2.4).

The general linear solution $\mathbf{r}\left(\hat{\mathbf{x}}\right)$ is given by

$$\mathbf{r}(\hat{\mathbf{x}}) = \mathbf{r}\left(\mathbf{x}^o\right) + \mathbf{W}\left(\hat{\mathbf{x}} - \mathbf{x}^o\right). \qquad (2.23)$$

Substitution of (2.23) into (2.22) provides the general linear solution arrived at by Tarantola (1987) and Eyre (1989)

$$\hat{\mathbf{x}} = \mathbf{x}^o + \left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W} + \mathbf{C}_X^{-1}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}^o)\right). \qquad (2.24)$$

This expression can be compared directly to (2.8b) for the constrained linear inversion technique. Equation (2.24) may be simplified to the same form as (2.8b) with the value of γ equal to the inverse of the signal-to-noise ratio (Chédin *et al.*, 1985; Smith *et al.*, 1972; Foster, 1961). If the errors in both **r** and **x** are uncorrelated, so that $\mathbf{C}_X = s_x^2\mathbf{I}$ and $\mathbf{C}_R = s_r^2\mathbf{I}$ (where s_x and s_r represent the sample standard deviations), then

$$\hat{\mathbf{x}} = \mathbf{x}^o + \left(\mathbf{W}^T\mathbf{W} + \frac{s_r^2}{s_x^2}\mathbf{I}\right)^{-1}\mathbf{W}^T\left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}^o)\right), \qquad (2.25)$$

which is identical to (2.8b) with the Lagrangian multiplier assuming the value of the ratio of the sample variances. This is the minimum information result.

The expression (2.24) is also the same as the result obtained from minimum variance method (Rodgers, 1976; Fleming *et al.* 1986). Tarantola (1987) emphasises that the use of minimum variance as an estimator in least-squares minimisation of (2.21) is only valid for Gaussian or near-Gaussian statistics. The minimum variance criterion may still be used for other non-Gaussian error distributions where the assumption is not strongly violated, but it may become a poor measure of the accuracy of a solution.

As with the constrained linear inversion method, the maximum likelihood solution scheme is dependent on an initial estimate which is preferably close to the true state of the atmosphere under observation. If this initial estimate, $\mathbf{x}^o$ in (2.24), is not a close representation of the solution, then the linear approximation of (2.23) is no longer valid. The selection of a suitable initial estimate may be be achieved in a number of ways but should be representative of the state of the atmosphere when the measurement was taken.

This solution scheme also may be applied using an iterative procedure where the radiometric response to a change in the atmospheric state is highly non-linear. Newtonian iteration is a common iterative method used although it is not necessarily the most efficient. Following the development of Rodgers (1976), Tarantola(1987) and Eyre (1989), an initial estimate to the solution vector may be found using the iterative Newton-Raphson equation,

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \left(\frac{\partial^2 S}{\partial \mathbf{x} \partial \mathbf{x}}\right)^{-1}_{\mathbf{x}_n} \left(\frac{\partial S}{\partial \mathbf{x}}\right)_{\mathbf{x}_n} , \qquad (2.26)$$

where the second derivative term from equation (2.21) in (2.26) is given by

$$\left(\frac{\partial^2 S}{\partial \mathbf{x} \partial \mathbf{x}}\right)^{-1}_{\mathbf{x}_n} = \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T \mathbf{C}_R^{-1} \mathbf{W}_n\right)^{-1}, \qquad (2.27)$$

which is the error covariance of the solution (Rodgers, 1976). It must be noted here that although $\mathbf{W}$ is nonlinear with respect to $\mathbf{x}$, none of the references cited deals with the derivative of $\mathbf{W}_n^T$ with respect to $\mathbf{x}_n$. Tarantola (1987) mentions that this second order term - a three dimensional operator - is hard to handle and may be discarded if either the calculated residuals are small or the forward model used to calculate $\mathbf{r}(\mathbf{x})$ is quasi-linear within the bounds of an iterative step. The regime of validity of this assumption with regard to the radiometric response of various atmospheric trace gases has not yet been investigated in the literature.

Substituting (2.22) and (2.27) into (2.26) gives the $\mathbf{x}_{n+1}$ iteration solution,

$$\begin{aligned}\mathbf{x}_{n+1} &= \mathbf{x}_n - \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\left[\mathbf{C}_X^{-1}\left(\mathbf{x}_n - \mathbf{x}^o\right) - \mathbf{W}_n^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}_n)\right)\right] \\ &= \mathbf{x}_n + \left(\mathbf{x}^o - \mathbf{x}_n\right) + \mathbf{G}_n\left[\mathbf{W}_n\left(\mathbf{x}_n - \mathbf{x}^o\right) + \left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{x}_n)\right)\right]\end{aligned}, \quad (2.28)$$

where $\mathbf{G}_n = \mathbf{C}_X\mathbf{W}_n^T\left(\mathbf{W}_n\mathbf{C}_X\mathbf{W}_n^T + \mathbf{C}_R\right)^{-1}$

$\mathbf{x}^o$ = initial estimate, and

$\mathbf{x}_n$ = nth iteration guess.

following Eyre (1989). The algebraic manipulation required to arrive at this expression is provided in Appendix A.

It is important to differentiate between the initial estimate and the iteration guess in (2.28). The matrix $\mathbf{C}_X$ describes the allowed variance of the solution about some predetermined estimate $\mathbf{x}^o$ taken from climatology or a sample of measurements of $\mathbf{x}$, whereas the iteration guess is updated for each iteration. The solution is thus constrained by the *a priori* information about the mean (Eyre, 1989; 1993 personal communication) - the essential reason for the inclusion of $\mathbf{C}_X$. If the quantities $\mathbf{x}^o$ and $\mathbf{x}_n$ are interpreted as the same quantity, then this constraint on the uniqueness of possible solutions is lost (Rodgers, 1976) and the optimal solution may not be achieved.

2.2. Retrieval schemes

The determination of the linear operator which maps a solution vector into measurement space (**W** in eqn. 2.4), or vice versa ($\mathbf{W}^{-1}$), is based on the retrieval problem formulation. Generally, this transformation operator is determined by linear regression from a sample of solution vectors and associated measurements (e.g. collocated radiosonde and radiance measurements) or from analytical derivatives of the RTE.

2.2.1. Statistical Retrieval Methods

The retrieval problem stated in (2.4) can be rewritten as

$$\Delta\mathbf{x} = \mathbf{K}\Delta\mathbf{r}\,; \qquad (2.29)$$

a linear mapping from measurement space to solution space. The transformation operator here, **K**, is analogous to $\mathbf{W}^{-1}$ from (2.4). The least squares solution for **K** from (2.29) is

$$\frac{\partial}{\partial\mathbf{K}}\left|\mathbf{K}\Delta\mathbf{r} - \Delta\mathbf{x}\right|^2 = 0 \quad \Rightarrow \quad \mathbf{K} = \Delta\mathbf{x}\Delta\mathbf{r}^T\left(\Delta\mathbf{r}\Delta\mathbf{r}^T\right)^{-1}. \qquad (2.30)$$

The operator **K** can thus be predetermined from a statistical set of spatially and temporally collocated radiosonde ($\Delta\mathbf{x}$) and radiance ($\Delta\mathbf{r}$) measurements (e.g. Smith *et al.* 1970; Lee *et al.* 1983) or a set of radiosonde measurements and their calculated brightness temperature (Hayden *et al.* 1981).

Lee *et al.* (1983) detail how the inclusion of measurement noise can stabilise simple regression retrievals for small data sets. This is an advantage for retrievals involving quantities such as trace absorbers which do not have access to large observational databases. If the least squares solution of (2.30) is conditioned with the measurement signal to noise ratio, we obtain

$$\mathbf{K} = \Delta\mathbf{x}\Delta\mathbf{r}^T\left(\Delta\mathbf{r}\Delta\mathbf{r}^T + \varepsilon^2\right)^{-1}, \tag{2.31a}$$

$$\text{where } \Delta r_i\Delta r_j + \varepsilon^2 = \begin{cases} \Delta r_i\Delta r_i\left(1+\dfrac{1}{g_i^2}\right) & \text{for i } = \text{ j}, \\ \Delta r_i\Delta r_j & \text{for i } \neq \text{ j}, \end{cases} \tag{2.31b}$$

and g_i^2 = rms signal to noise ratio for channel *i*.

This noise conditioning has a similar effect as the conditioning used in the minimum information inversion scheme described earlier. This conditioning reduces the dependence of the solution on any errors in the statistical database and on any nonlinearities between $\Delta\mathbf{r}$ and $\Delta\mathbf{x}$.

A more rigorous statistical retrieval technique is that of principal component analysis (see Wark and Fleming, 1966; Alishouse *et al.* 1967; Smith and Woolf, 1976; Lipton and Vonder Haar, 1987). The retrieval of atmospheric state parameters is particularly suited to principal component analysis because of the typically high degree of dependence between spectral radiance measurements. Measurement information which is not contributing to the retrieval (e.g. noise) may be identified and discarded.

Diagonalisation of covariance matrices of atmospheric parameters provides a set of linearly independent functions (empirical orthogonal functions) that describe the variance in the statistical set used to calculate the covariance. The amount of variance described by each eigenvector is determined from its associated eigenvalue. When ordered in terms of decreasing eigenvalue, the first few eigenvectors typically describe the majority of variance. Eigenvectors which do not contribute significant new information, relative to some threshold, can be excluded from the computation.

The retrieval problem is written in terms of linear expansions of the quantity to be retrieved and the measurements with the eigenvectors of the covariance matrices as the representation functions

$$\Delta x_j = \sum_{i=1}^{M} a_i X^*_{ji} \quad \Rightarrow \quad \Delta\mathbf{x} = \mathbf{X}^*\mathbf{a}\,, \tag{2.32a}$$

$$\Delta r_v = \sum_{u=1}^{N} b_u R^*_{vu} \quad \Rightarrow \quad \Delta\mathbf{r} = \mathbf{R}^*\mathbf{b}\,, \tag{2.32b}$$

where $\Delta\mathbf{x}$ = parameter vector to be retrieved,

$\mathbf{X}^*$ = matrix of eigenvectors of covariance matrix of $\Delta\mathbf{x}$,

$\mathbf{a}$ = principal component vector for $\Delta\mathbf{x}$,

M = no. of parameter eigenvectors (modes) retained,

$\Delta\mathbf{r}$ = measurement vector,

$\mathbf{R}^*$ = matrix of eigenvectors of covariance matrix of $\Delta\mathbf{r}$,

$\mathbf{b}$ = principal component vector for $\mathbf{r}$,

N = no. of measurement eigenvectors retained,

(Smith and Woolf, 1976; Lipton and Vonder Haar, 1987).

The principal component vectors **a** and **b** may be determined from the inverse of (2.32) by simple least squares,

$$\mathbf{a} = \left(\mathbf{X}^{*T}\mathbf{X}^{*}\right)^{-1}\mathbf{X}^{*T}\Delta\mathbf{x} = \mathbf{X}^{*T}\Delta\mathbf{x}, \qquad (2.33a)$$

$$\mathbf{b} = \left(\mathbf{R}^{*T}\mathbf{R}^{*}\right)^{-1}\mathbf{R}^{*T}\Delta\mathbf{r} = \mathbf{R}^{*T}\Delta\mathbf{r}, \qquad (2.33b)$$

and the relationship between **a** and **b** is determined by regression from a statistical set by defining a transformation matrix **D**, where

$$\mathbf{a} = \mathbf{D}\mathbf{b} \quad \Rightarrow \quad \mathbf{D} = \mathbf{A}\mathbf{B}^{T}\left(\mathbf{B}\mathbf{B}^{T}\right)^{-1}, \text{ and} \qquad (2.34)$$

where $\mathbf{A}$ = matrix of principal components **a** for each sample in the statistical set, and

$\mathbf{B}$ = matrix of principal components **b**.

The relationship between $\Delta\mathbf{x}$ and $\Delta\mathbf{r}$ may now be made directly using (2.33) and (2.34)

$$\Delta\mathbf{x} = \mathbf{X}^{*}\mathbf{D}\mathbf{R}^{*T}\Delta\mathbf{r} = \mathbf{K}\Delta\mathbf{r}. \qquad (2.35)$$

When $i = L$, the total number of pressure levels, and $u = N$, the total number of spectral channels, the result (2.35) is the same as that given in (2.30) where **K** is the standard regression coefficient operator (Smith and Woolf, 1976).

The number of modes (eigenvectors) to retain in the retrieval scheme is determined by use of those modes which account for most of the variance of the solution. Smith and Woolf (1976) discard eigenvectors beyond those that reduce the rms error (of the parameter in question) to the instrumental noise level, They do not distinguish between i and j modes setting them to the same value - ten modes are used for temperature retrievals and eight for water vapour concentration retrievals using Nimbus-6 HIRS and SCAMS data. Lipton and Vonder Haar (1987) vary the number of modes - both i and j separately - and examine the pressure-weighted average explained variance to determine optimum modal values for water vapour retrievals. The i and j values which account for the highest solution variance are selected. In their analysis using HIRS-2 data, Lipton and Vonder Haar found that only three eigenvectors were sufficient to represent all the independent information in measured radiances for water vapour retrievals.

2.2.2. Physical Retrieval Methods

The construction of transformation operators in a physical retrieval scheme is based upon physical principles of radiative transfer in the atmosphere, on how the RTE is linearised or solved, and on what parameter is to be retrieved. In atmospheric parameter retrievals, the term weighting function usually refers to the quantity $d\tau(\nu,p)/dp$. This quantity defines the sensitivity of spectral channels to radiation emanating from different layers of the atmosphere. For a particular spectral interval or channel, a weighting function value of zero at a specified pressure means there is no radiance contribution from that atmospheric layer to the radiance measured in that interval. The following reviews the construction of the transformation operator in physical retrieval schemes.

The derivative of the radiative transfer equation with repect to the parameter required to be retrieved yields physical transformation operators. For the set of linear equations,

$$\Delta\mathbf{r} = \mathbf{W}\Delta\mathbf{x},$$

the elements of **W** are the Jacobian of **r**, given by

$$W_{ij} = \frac{\partial r_j(\mathbf{x})}{\partial x_i}, \quad i = 1,2,\ldots,L; \; j = 1,2,\ldots,N, \tag{2.36}$$

where N = number of spectral channels, L = number of pressure levels.

For a simple PRTE, and considering the temperature only, we have

$$\frac{\partial R(\nu, p_s)}{\partial T(p)} = \int_0^{p_s} \frac{\partial B(\nu, T(p))}{\partial T(p)} \frac{d\tau(\nu, p)}{dp} dp. \tag{2.37}$$

Converting the measured radiances to brightness temperatures gives,

$$\frac{\partial T_B(p_s)}{\partial T(p)} \frac{\partial B(\nu, T_B(p_s))}{\partial T_B(p_s)} = \int_0^{p_s} \frac{\partial B(\nu, T(p))}{\partial T(p)} \frac{d\tau(\nu, p)}{dp} dp. \tag{2.38}$$

Rearranging (2.38) gives the dependence of the measured brightness temperature for channel j due to the atmospheric temperature at pressure level i

$$W_{ij} = \left[\frac{\partial B_j(T_i) / \partial T}{\partial B_j(T_B) / \partial T_B} \right] \left[\frac{d\tau}{dp} \right]_{ij} \tag{2.39}$$

(Susskind and Rosenberg, 1980; Chédin *et al.*, 1985).

This form of weighting function is used by Susskind and Rosenberg for temperature retrievals using TIROS-N HIRS/2 and MSU data with an iterative retrieval scheme (Chahine, 1970). Chédin *et al.* (1985) detail the use of these weighting functions in temperature retrievals using similar data employing a maximum likelihood inversion scheme.

The derivation of the Jacobian of a discrete form of the RTE, for a number of atmospheric parameters of interest (temperature, T, water vapour mixing ratio, q, surface temperature, pressure, emissivity, and cloud top pressure), is carried out by Eyre (1989) for TOVS data. The RTE used by Eyre includes a term for surface reflected radiation reaching the sensor, namely

$$R_j = \varepsilon_s B_j(T_s)\tau_j(p_s) - \int_0^{p_s} B_j(T(p)) \frac{d\tau_j(p)}{dp} dp + (1-\varepsilon_s) \int_0^{p_s} B_j(T(p)) \frac{d\tau_j^*(p)}{dp} dp, \tag{2.40}$$

where j = channel index,

ε_s = surface emissivity,

$\tau_j^*(p)$ = two-path transmittance from p to p_s and then to space,

with all other terms as previously described. The form of (2.40) is approximated in a discrete form for a channel j and L pressure levels,

$$R = \varepsilon_s B_s \tau_s + \frac{1}{2} \sum_{l=1}^{L} \left(B_l + B_{l-1}\right)\left[\tau_{l-1} - \tau_l + (1-\varepsilon_s)(\tau_l^* - \tau_{l-1}^*)\right]. \tag{2.41}$$

The differential of (2.41) with respect to the required parameter provides the transformation operator that is used in a maximum likelihood retrieval scheme using TOVS data.

The similtaneous retrieval method described by Smith and Woolf (1984) and Smith *et al.* (1985) obtains similar transformation operators as shown previously but combines the operators for temperature and absorber concentration into one matrix. Smith and Woolf (1984) detail this scheme for temperature and water vapour using TOVS radiances and Smith *et al.* (1991) extend the technique to include other absorbing constituents, such as ozone, for use with high resolution HIS radiances. A similtaneous retrieval of temperature and water vapour concentration is preferable to retrieving one parameter at a time due to the heavy radiative interdependence between temperature and water vapour concentration in the spectral channels available. The increase in retrieval accuracy by this method, as opposed to statistical and iterative methods, is attributed to this feature of the solution scheme (Smith et al, 1985).

The linearised PRTE is given by

$$\begin{aligned}\delta T_B &= \delta T(p_s)\beta^o(\nu,p_s)\tau^o(\nu,p_s) - \int_0^{p_s}\delta T(p)\beta^o(\nu,p)\frac{d\tau^o(\nu,p)}{dp}dp \\ &+ \int_0^{p_s}\delta U(p)\beta^o(\nu,p)\frac{d\tau^o(\nu,p)}{dU^o(p)}\frac{dT(p)}{dp}dp,\end{aligned} \tag{2.42}$$

where $\beta^o(\nu,p) = \left[\frac{\partial B(\nu,T^o)/\partial T(p)}{\partial B(\nu,T_B^o)/\partial T_B(p)}\right]$, and

$U(p) = \frac{1}{g}\int_0^{p_s} q(p)dp$, the absorber path length with mixing ratio $q(p)$,

and the superscript ^{o} indicates *a priori* information - the initial estimate.

The solution is recast as linear combinations of L+1 pressure basis functions, $w(p)$, for N spectral channels, namely

$$\delta T_s = a_o w_o, \tag{2.43a}$$

$$\delta U(p) = \frac{1}{g}\int_0^{p}\delta q(p')dp' = \sum_{i=1}^{K} a_i \int_0^{p} q(p')w_i(p')dp', \tag{2.43b}$$

$$\delta T(p) = -\sum_{i=K+1}^{L} a_i w_i(p), \tag{2.43c}$$

so that (2.42) becomes

$$(\delta T_B)_j = \sum_{i=0}^{L} a_i W_{ji}, \quad j = 1,2,\ldots,N, \tag{2.44}$$

where the elements of W constitute a transformation function from the solution coefficients, a, to the observables, δT_B. Using (2.42), these are given by

$$W_{0j} = \beta^o_{s,j} \tau^o_{s,j}, \tag{2.45a}$$

$$W_{ij} = \int_0^{p_s} \left[\int_0^{p} q(p') w_i(p') dp' \right] \left[\frac{\partial T}{\partial p} \frac{\partial \tau^o_j}{\partial U^o} \right] \beta^o_j dp \text{ for } 1 \le i \le K, \tag{2.45b}$$

$$W_{ij} = \int_0^{p_s} w_i \frac{\partial \tau^o_j}{\partial p} \beta^o_j dp, \text{ for } K < i \le L. \tag{2.45c}$$

The form of the basis functions used is arbitrary but in this case is the vertical derivative of transmittance which physically describe from where in the atmosphere the upwelling radiation originates for each channel. As pointed out by Hayden (1988), empirical orthogonal functions could be used in this scheme to include statistical information in the solution.

The solution (2.44) may be recast in matrix form as,

$$\Delta \mathbf{T}_B = \mathbf{W}\mathbf{a}. \tag{2.46}$$

Smith *et al.* (1985) and Hayden (1988) solve (2.46) using constrained linear inversion techniques. The value of the Lagrangian multiplier, γ, used in their solutions is selected according to experience and is not optimised.

In extending (2.42) to include more than one absorber (other than water vapour) in the retrieval scheme, Smith *et al.* (1991) group the temperature and constituent terms under the same integral, with

$$\delta T_B = \delta T(p_s)\beta^o(\nu,p_s)\tau^o(\nu,p_s) \\ -\sum_{i=1}^{Nm}\int_0^{p_s}\left\{\delta T(p)\left[\frac{dU_i^o(p)}{dp}\right] - \delta U_i(p)\left[\frac{dT(p)}{dp}\right]\right\}\beta^o(\nu,p)\frac{d\tau^o(\nu,p)}{dU_i^o(p)}dp, \tag{2.47}$$

where Nm = the number of absorbers considered, and

$$\frac{\partial\tau_i^o(\nu,p)}{\partial p} = \frac{\partial U_i^o(p)}{\partial p}\frac{d\tau_i^o(\nu,p)}{dU_i^o(p)}.$$

The solution to (2.47) detailed in Smith *et al.* (1991) retrieves atmospheric concentrations by solving for the difference of an absorber's effective temperature from the actual atmospheric temperature. The effective temperature perturbation of an absorbing gas is defined as,

$$\delta T_i(p) = \delta T(p) - \delta U_i(p)\left[\frac{dT(p)}{dU_i^o(p)}\right], \tag{2.48}$$

reducing (2.47) to

$$\delta T_B = \delta T(p_s)\beta^o(\nu,p_s)\tau^o(\nu,p_s) - \sum_{i=1}^{Nm}\int_0^{p_s}\delta T_i(p)\beta^o(\nu,p)\frac{d\tau^o(\nu,p)}{dp}dp, \tag{2.49}$$

which may be written directly in matrix form as,

$$\Delta\mathbf{T}_B = \mathbf{W}\Delta\mathbf{T}, \tag{2.50}$$

with the elements of the transformation matrix being

$$W_{0j} = \beta^{o}_{s,j}\tau^{o}_{s,j}, \tag{2.51a}$$

$$W_{lj} = -\beta^{o}_{lj}\frac{d\tau^{o}_{lj}}{dp}, \text{ for } l = 1,2,....,Nm \times L_i, \tag{2.51b}$$

where j = spectral channel index,

Nm = number of absorbers considered,

L_i = number of pressure levels retrieved for each absorber i.

A maximum likelihood inversion scheme was employed by Smith *et al.* (1991) to solve (2.50) The improvement in temperature and water vapour sounding is well documented but there is no discussion of the impact of *a priori* data on ozone retrievals. The total ozone retrieval accuracy is stated as 2.5% but with poor resolution in the troposphere.

2.3. Retrieval Analysis

The two operators that quantify the accuracy of a retrieval procedure are the retrieval error covariance and the resolution operator.

The retrieval error covariance has been defined previously for a maximum likelihood retrieval scheme (see eqn. 2.24) and is given in Rodgers (1976), Tarantola (1987), and Eyre (1989) as,

$$\mathbf{C}_E = \left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W} + \mathbf{C}_X^{-1}\right)^{-1} = \mathbf{C}_X - \mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T + \mathbf{C}_R\right)^{-1}\mathbf{W}\mathbf{C}_X, \tag{2.52}$$

where $\mathbf{C}_X$ = covariance of *a priori* information of parameter x,

$\mathbf{C}_R$ = measurement covariance,

$\mathbf{W}$ = transformation operator,

and is simply the combination of the mapping of the measurement covariance into parameter space and the *a priori* data covariance. The algebraic manipulation to arrive at the second form of $\mathbf{C}_E$ (RHS of eqn. 2.52) is shown in Appendix B. Similarly for a constrained linear inversion scheme,

$$\mathbf{C}_E = \left(\mathbf{W}^T\mathbf{W} + \gamma\mathbf{H}\right)^{-1} \text{ or } \left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W} + \gamma\mathbf{H}\right)^{-1}. \tag{2.53}$$

Here the term $\gamma\mathbf{H}$ is equivalent to the inverse of the *a priori* covariance, $\mathbf{C}_X^{-1}$.

It is usually assumed that the measurement errors are uncorrelated so that $\mathbf{C}_R$ is a simple diagonal matrix with the elements $(\mathbf{C}_R)_{ii}$ being the square of the rms measurement error. This assumption is valid in most cases although significant off-diagonal elements may be introduced into $\mathbf{C}_R$ due to sensor calibration errors (Crosby, 1989).

The resolution operator maps the true solution to the retrieved solution, using the form

$$\Delta\hat{x}(p) = \int_0^{p_s} R(p,p')\Delta x(p')dp'; \tag{2.54}$$

or in matrix form,

$$\Delta\hat{\mathbf{x}} = \mathbf{R}\Delta\mathbf{x} \tag{2.55}$$

This can be thought of as an averaging process that broadens the components of the true profile, $\Delta\mathbf{x}$, in the solution $\Delta\hat{\mathbf{x}}$ (Rodgers, 1990). The *rows* of **R** are typically referred to as averaging kernels or vertical resolution functions (VRFs), because they define how the retrieval system resolves the true profile in the solution.

For a maximum likelihood inversion scheme, we have

$$\begin{aligned}\Delta\hat{\mathbf{x}} &= \mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T+\mathbf{C}_R\right)^{-1}\Delta\mathbf{r}\\ &= \mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T+\mathbf{C}_R\right)^{-1}\mathbf{W}\Delta\mathbf{x}\\ &= \mathbf{R}\Delta\mathbf{x}\,;\end{aligned}$$

so,

$$\mathbf{R}=\mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T+\mathbf{C}_R\right)^{-1}\mathbf{W}=\left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}+\mathbf{C}_X^{-1}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}. \qquad (2.56)$$

Similarly, for a constrained linear inversion scheme,

$$\mathbf{R}=\left(\mathbf{W}^T\mathbf{W}+\gamma\mathbf{H}\right)^{-1}\mathbf{W}^T\mathbf{W} \text{ or } \left(\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}+\gamma\mathbf{H}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}. \qquad (2.57)$$

The relationship between the resolution and retrieval error covariance operators is evident by identifying (2.56) in (2.52),

$$\begin{aligned}\mathbf{C}_E &= \mathbf{C}_X-\mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T+\mathbf{C}_R\right)^{-1}\mathbf{W}\mathbf{C}_X\\ &= \mathbf{C}_X-\mathbf{R}\mathbf{C}_X\\ &= (\mathbf{I}-\mathbf{R})\mathbf{C}_X,\end{aligned} \qquad (2.58)$$

which may be rewritten as

$$\mathbf{I}=\mathbf{R}+\mathbf{C}_E\mathbf{C}_X^{-1}. \qquad (2.59)$$

This equation can be used to determine how many retrieval layers are resolved by the measured data and how many by the *a priori* data by taking the trace of (2.59),

$$Tr(\mathbf{I}) = Tr(\mathbf{R}) + Tr\left(\mathbf{C}_E \mathbf{C}_X^{-1}\right), \qquad (2.60)$$

where $Tr(\mathbf{I})$ = total no. of atmospheric layers to be retrieved,

$Tr(\mathbf{R})$ = no. of atmospheric layers resolved by the measured data,

$Tr\left(\mathbf{C}_E \mathbf{C}_X^{-1}\right)$ = no. of atmospheric layers resolved by the *a priori* data,

(from Tarantola, 1987). The number of resolved parameters (atmospheric layers) given by $Tr(\mathbf{R})$ also quantifies the number of independent pieces of information provided by the measured data (Huang *et al.*, 1992; Purser and Huang, 1993).

This essentially provides a measure of the effect of including *a priori* data in a retrieval scheme. This is an important practical consideration because the vertical resolution of a sounding sensor will improve or degrade depending on the vertical resolution of incorporated *a priori* data (Thompson, 1982) with the corresponding effect on the retrieval error covariance. The ideal retrieval process would minimise both the retrieval error and the spread of the VRFs for each atmospheric layer. This is not possible however, as decreasing one necessarily leads to an increase in the other. Conrath (1972) shows that small decreases in the width of VRFs can cause large increases in the retrieval error.

Huang *et al.* (1992) details a method of determining the vertical resolution of a retrieval scheme from the resolution operator using the concept of effective data density. The use of the width or spread of VRFs is not suitable in all cases as they often possess oscillatory sidelobes. The trace of the resolution operator is an overall measure of the amount of independent information available and can be used to determine the effective data density of the retrieval scheme (Purser and Huang (1993). Following Huang *et al.* the elementary expression for local data density is given by,

$$\rho_i' = \frac{R_{ii}}{\delta Z_i}, \tag{2.61}$$

where ρ_i' = local effective data density for layer i (km^{-1}),

R_{ii} = resolution operator diagonal value for atmospheric layer i,

δZ_i = height increment for layer i.

The expression (2.61) is not a good measure of the data density in that the negative sidelobes of the VRFs can cause negative values for ρ. A better estimate of the effective data density is found by smoothing (2.61) to give,

$$\rho_i = \sum_{j=1}^{L} F_{ij} R_{jj}, \tag{2.62}$$

where the smoothing operator is

$$F_{ij} = \frac{R_{ji}^2}{\sum_{k=1}^{L} R_{jk}^2 \delta Z_k} \tag{2.63}$$

(from Huang *et al.*, 1992). The vertical resolution of the retrieval scheme from which R derives is given by the reciprocal of the data density.

Huang *et al.* (1992) also explicitly derive the retrieval error covariance as the sum of two components; the null space or vertical resolution component error and the measurement noise component error, namely

$$\mathbf{C}_E = (\mathbf{I}-\mathbf{R})\mathbf{C}_X(\mathbf{I}-\mathbf{R})^T + \mathbf{G}\mathbf{C}_R\mathbf{G}^T, \tag{2.64}$$

where $\mathbf{G} = \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}$ or $\mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T + \mathbf{C}_R\right)^{-1}$,

$(\mathbf{I}-\mathbf{R})\mathbf{C}_X(\mathbf{I}-\mathbf{R})^T$ = null space component error covariance, and

$\mathbf{G}\mathbf{C}_R\mathbf{G}^T$ = measurement noise component error covariance.

The derivation of (2.64) and its equivalence with (2.52) is shown in Appendix C.

The null space or vertical resolution component error can be thought of as the contribution of variance to the solution by the *a priori* data. In those regions of the atmosphere that are transparent to measurements (e.g. isothermal layers), *a priori* data provides information which allows the retrieval scheme to determine optimal solution elements for those atmospheric regions that do not contribute to the measurement (Rodgers, 1990).

The rms retrieval error is simply the square root of the diagonals of the retrieval error covariance operator(s),

$$\varepsilon_{rms} = \sqrt{(C_E)_{ii}}\,. \tag{2.65}$$

This is a quantitative measure of the accuracy of the retrieval process. This, however, does not provide an overall assessment of the error when there are significant off-diagonal elements in the retrieval error covariance matrix. The correlation field across atmospheric layers and parameters is,

$$\rho_{ij} = \frac{(C_E)_{ij}}{\sqrt{(C_E)_{ii}(C_E)_{jj}}}. \tag{2.66}$$

A strong correlation in the retrieval errors of any two layers (or parameters) implies that they have not been independently resolved and that only some linear combination of the retrieved parameters has been resolved (Tarantola, 1987).

Retrieval accuracy is also limited by the accuracy of the parameters and the forward model used to calculate the transmittances used in the transformation matrix. Rodgers (1990) shows that the systematic error covariance contribution, $\mathbf{C}_S$, to the retrieval error covariance, $\mathbf{C}_E$, for any model parameter b (e.g. empirical constants, line strengths, etc.) with a best estimate $\hat{b}$ and a known error covariance $\mathbf{C}_b$, is

$$\mathbf{C}_S = \mathbf{R}_b \mathbf{C}_b \mathbf{R}_b^T, \tag{2.67}$$

where $\mathbf{R}_b = \left(\mathbf{C}_X^{-1} + \mathbf{W}^T \mathbf{C}_R^{-1} \mathbf{W}\right)^{-1} \mathbf{W}^T \mathbf{C}_R^{-1} \mathbf{W}_b$ or $\mathbf{C}_X \mathbf{W}^T \left(\mathbf{W} \mathbf{C}_X \mathbf{W}^T + \mathbf{C}_R\right)^{-1} \mathbf{W}_b$,

$\mathbf{W}_b$ = sensitivity of the forward model to the parameter b,

and the subscript s refers to a systematic error.

The magnitudes of known errors in forward transmittance models, and their impact in atmospheric temperature retrievals, is discussed by Ben-David *et al.* (1985). The transmittance functions used in retrievals are dependent upon line strengths, widths and shape as well as being pressure and temperature dependent. For a Lorentz lineshape, the absorption coefficient is,

$$k(\nu, p, T) = \frac{S(T)\alpha(T,p)}{\pi\left[(\nu-\nu_o)^2 + \alpha^2(T,p)\right]}, \tag{2.68}$$

where α = Lorentz halfwidth = $\alpha(T_o, p_o)\left(\frac{p}{p_o}\right)\left(\frac{T_o}{T}\right)^b$,

$S(T)$ = line strength,

ν_o = line centre,

T_o, p_o = reference temperature and pressure, and

b = exponent for the temperature dependence of the halfwidth.

Errors in the parameters describing *k* have a significant impact on the calculated transmittance functions. Ben-David *et al.* calculate that errors in the line strengths in the CO_2 14μm band can lead to errors of up to 1% in the calculated radiances. Line strength uncertainties in current spectroscopic databases used in atmospheric transmittance calculations quote errors of 5-20% in the line strengths of water vapour (6.3 and 2.7μm), ~10% for CO_2 (15μm), and ~10% and between 5-25% for O_3 (9.6μm and 2-4μm regions respectively) (Rothman *et al.*, 1987). Errors of these magnitudes can introduce errors of up to 10% in the corresponding calculated transmittance and radiance.

2.4. Initial estimate selection

It has long been recognised that the accuracy of a retrieval is dependent upon a good initial estimate to the solution. The widths of the channel weighting functions determine the depth of the atmosphere sensed by any one channel and, as weighting functions for temperature or atmospheric constituent sounding are intrinsically broad, information retrieved is a bulk measure. If the radiometric impact of the retrieved minus initial estimate profile differences in different regions of the atmosphere compensate each other, a poor initial estimate may render the retrieval procedure insensitive to errors in the basic shape of the initial estimate profile.

Procedures employed to select an initial estimate involve some method of classifying *a priori* data and then determining the best match between the radiance estimator for an *a priori* data class and the measured radiances. Overall, classification procedures are generally applicable to any parameter retrieval scheme, but should be optimised for a particular scenario (Thompson *et al.* 1985, and see, for example, McMillin, 1991). Classification is achieved by decomposing the *a priori* data into a form that allows recognition of the typical spectral features in a radiance measurement or the typical vertical features in a profile.

2.4.1. Radiance Classification

Classification in the radiance domain provides an initial estimate by matching the measured radiances to *a priori* radiances that are associated with an atmospheric profile of the quantity to be retrieved.

The radiance classification method described by Chédin and Scott (1983) for temperature retrievals using TOVS radiances involves the determination of the position of *a priori* radiance spectra, calculated from sonde profiles, in an N-dimensional radiance space. A data matrix of dimension $M \times N$ is constructed of normalised radiance vectors (or representation vectors), where M represents the number of *a priori* radiance spectra in the data set, and N the number of spectral channels, so that

$$\tilde{R}_{ij} = \frac{\left(R_{ij} - \overline{R}_j\right)}{\sum_{k=1}^{M}\left(R_{kj} - \overline{R}_j\right)^2} \quad i = 1,2,\ldots\ldots, M; \; j = 1,2,\ldots\ldots N, \qquad (2.69)$$

where R_{ij} = radiance for sample i at channel j,

$\overline{R}_j$ = average radiance at channel j.

The correlation matrix is determined from (2.69) and diagonalised using,

$$\mathbf{P} = \tilde{\mathbf{R}}^T\tilde{\mathbf{R}} = \mathbf{E}\Lambda\mathbf{E}^T, \tag{2.70}$$

where $\mathbf{P} = N\times N$ correlation matrix,

$\mathbf{E} = N\times N$ orthogonal matrix of eigenvectors,

$\Lambda = N\times N$ diagonal matrix of eigenvalues.

Note that the correlation matrix definition of (2.70) is equivalent to that of (2.66). The co-ordinates of each normalised radiance spectrum in radiance space is determined from,

$$Y_{ik} = \sum_{j=1}^{N} \tilde{R}_{ij}E_{jk} \quad i = 1,2,\ldots\ldots, M; \;\; k = 1,2,\ldots\ldots, l \;\;, \tag{2.71}$$

where $Y_{ik} = M\times l$ matrix of co-ordinate vectors for the M spectra,

l = no. of eigenvectors (modes) retained in expansion.

The number of modes retained depends on the variance described by each. In practice, two to four eigenvectors are usually enough to explain up to 99% of the variance in temperature and/or water vapour profiles. Given the co-ordinates of each radiance spectrum in radiance space, the initial estimate is selected by minimisation of the Euclidean distance between any *a priori* radiance co-ordinate to the observed radiance co-ordinate. The quantity minimised is,

$$d^2(obs, a\,priori) = \sum_{k=1}^{l}\left(Y_{obs,k} - Y_{a\,priori,k}\right)^2, \tag{2.72}$$

where the *a priori* radiance spectrum, and its associated profile, which produce the smallest d^2 is taken as the initial estimate.

Thompson *et al.* (1985) produce an initial estimate by selecting the radiance spectrum which provides the maximum inner product with the observed radiance spectrum. The inner product is a measure of how well any two radiance representation vectors are aligned in radiance space. *A priori* radiance spectra are decomposed into their eigenvector representation and the resultant vectors are scaled using,

$$V_i = \frac{(R_i - a_i)}{b_i}, \tag{2.73}$$

where V_i = scaled radiance representation vector using eigenvector i,

R_i = radiance using eigenvector i,

a_i, b_i = scaling parameters,

where the a_i and b_i are optimised for the problem. The scaling can be altered to favour particular channels/atmospheric regions. The representation by Thompson *et al.* uses the sample mean (a_i) and standard deviation (b_i). The vectors, **V**, are then normalised to their representative form, $\tilde{\mathbf{r}}$. The colinearity of any *a priori* vector, $\tilde{\mathbf{r}}_{a\,priori}$, to an observed vector, $\tilde{\mathbf{r}}_{obs}$, is then determined by their inner product,

$$\left\langle \tilde{\mathbf{r}}_{obs}, \tilde{\mathbf{r}}_{a\,priori} \right\rangle = \cos\theta, \tag{2.74}$$

where θ is a measure of the angle between $\tilde{\mathbf{r}}_{a\,priori}$ and $\tilde{\mathbf{r}}_{obs}$. Thus maximisation of the inner product of the observed spectrum representation vector with any *a priori* vector will yield an initial estimate for the atmospheric profile associated with the *a priori* representation vector. Thompson *et al.* define a critical value, an averaging limit, for the inner product result whereby the profile of any radiance vector that exceeds this limit is used to construct an average initial estimate to the retrieval. This is done to ensure that the significant features in the profile are retained while the small scale variability, which does not have a radiometric impact, is filtered out.

2.4.2. Profile Classification

Classification in profile space catagorise *a priori* profile data into classes based on their vertical structure. The derivation of vertical structure classes has been approached differently by different authors. Typical eigenvalue analysis and clustering methods have been described by Thompson *et al.* (1985) for temperature retrievals, and by Lipton et al (1986) for water vapour retrievals from satellite measured radiances. An *a priori* data covariance matrix is decomposed by eigenvalue analysis and the profile data reconstructed using a truncated series expansion

$$x_j - \overline{x}_j = \sum_{i=1}^{l} a_i E_{ij} \text{ for } j = 1,2,\ldots\ldots, L \quad , \tag{2.75}$$

where $x_j, \overline{x}_j$ = parameter and mean for atmospheric layer j,

a_i = vector of expansion coefficients (principal components),

E_{ij} = matrix of eigenvectors of covariance matrix of x,

l = no. of parameter eigenvectors (modes) retained,

L = total no. of atmospheric layers.

Thompson *et al.* (1985) define classes based on the frequency distribution of the ratio of the i^{th} expansion coefficient to the i^{th} eigenvalue for each radiosonde profile. Classes defined in this way are subjected to the same analysis and subdivided until the distribution for each sub-class appears Gaussian. Lipton *et al.* (1986) define classes based on the relative magnitudes of the expansion coefficients of the second and third eigenvector - those two eigenvectors describing the variance in the vertical structure of water vapour profiles. The classification methods differ but the results are the same - each class contains profiles of similar vertical structure. Using the terminology of Lipton *et al.*, the profiles are thus partitioned into vertical structure classes.

Another classification method (Uddstrom and Wark, 1985) directly decomposes an *a priori* profile data matrix such that each profile can be represented by a series expansion. Singular value decomposition of the data matrix gives,

$$\mathbf{A} = \mathbf{L\Sigma Y}^T, \tag{2.76}$$

where $\mathbf{A}$ = $M \times N$ matrix of N profiles specified at M levels,

$\mathbf{L}$ = $M \times M$ orthogonal matrix of left singular vectors (columns),

$\mathbf{Y}$ = $N \times N$ orthogonal matrix of right singular vectors (columns),

Σ = $M \times N$ diagonal matrix of singular values of $\mathbf{A}$, diag($\sigma_1, \sigma_2, \ldots, \sigma_M$),

and any profile (column) of $\mathbf{A}$ can be represented by the truncated series expansion,

$$\mathbf{a}_j \approx \mathbf{a}_{jL} = \sum_{k=1}^{L} \mathbf{l}_k \sigma_k y_{jk}\,, \tag{2.77}$$

where the number of terms retained, L, explains the majority of the variance in the $\mathbf{l}_k$ vectors.

The expression (2.77) allows reconstruction of the individual profiles in a data set based upon the characteristic vectors, $\mathbf{l}_1$ and $\mathbf{y}_1$, and subsequent variations, the higher order $\mathbf{l}_k$ and $\mathbf{y}_k$. To partition profiles based on these variations, it is necessary to transform (2.77) so that profiles can be represented by functions describing a particular vertical structure class (Jalickee and Ropelewski, 1979). This is done by rotating the co-ordinate system of (2.77) so that the rotated axes are better aligned with the spread of the profile representation vectors in profile space (Thompson *et al.*, 1985). The rotation transformation is described by,

$$\mathbf{a}_{jL} = \sum_{k=1}^{L} \mathbf{l}_k \sigma_k y_{jk} = \sum_{k=1}^{L} \hat{\mathbf{a}}_k \hat{y}_{jk} , \tag{2.78}$$

with the typical structure functions (TSFs), $\hat{\mathbf{a}}_k$, and the coefficients, $\hat{y}_{jk}$, given by,

$$\hat{\mathbf{a}}_k = \sum_{i=1}^{L} l_i \sigma_i \mathrm{v}_{ik} \quad ; \quad \hat{y}_{jk} = \sum_{i=1}^{L} u_{ki} y_{ji} , \tag{2.79}$$

where $\mathbf{V} = \mathbf{U}^{-1}$ and the transformation matrix $\mathbf{U}$ is determined by minimising the cross terms between the TSF coefficients. This has the effect of isolating the major contributions of any profile $\mathbf{a}_j$ to a single TSF $\hat{\mathbf{a}}_k$ - other TSFs having only a minor contribution. The *a priori* data can thus be classified according to its typical vertical structure - all the profiles in any class having similar vertical structure and features with near Gaussian distribution statistics.

Typically, once the *a priori* data has been classified, the class means are used to calculate a radiance vector which is compared with the observed radiances as outlined in §2.4.1.. There is, however, a finite probability that a profile may be misclassified which will have a systematic effect on the retrieval accuracy. The radiance classifier employed by Uddstrom and Wark (1985) uses a Bayesian discriminant function to select the optimal classified profile. Each profile class ω_i has associated with it a radiance discriminant function

$$g_i(\mathbf{r}) = \log\left[P(\omega_i|\mathbf{r})\right] \quad i = 1,2,.....,n \ , \tag{2.80}$$

where $P(\omega_i|\mathbf{r})$ = *a posteriori* conditional probability that a profile in class ω_i is classified correctly with n being the total number of classes. A Gaussian distribution is used for $P(\omega_i|\mathbf{r})$ so that (2.80) becomes,

$$g_i(\mathbf{r}) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{C}_{Ri}| - \frac{1}{2}(\mathbf{r}-\bar{\mathbf{r}}_i)^T \mathbf{C}_{Ri}^{-1}(\mathbf{r}-\bar{\mathbf{r}}_i) + \log P(\omega_i), \quad (2.81)$$

where $\mathbf{C}_{Ri}$ = radiance covariance matrix for class i,

$P(\omega_i)$ = *a priori* probability of the observed atmosphere being in class i,

N = no. of channels used.

Maximisation of (2.81) determines the class, i, to which the atmosphere described by the observed radiance vector **r** can be assigned.

3.0 Forward Problem

3.1 Data

3.1.1 *a priori* ozone concentration profile information

The *a priori* ozonesonde data used in this study is based upon nine ozonesondes from the FIRE II/SPECTRE (F2S) experiment held during November and December, 1991 in Coffeyville, Kansas (37.10N, -95.57E). A set of 83 seasonally and latitudinally matched ozonesondes (midlatitude winter) were selected from a database (DB) of 408 sondes (Woolf, 1993, private communication) to supplement the F2S ozonesonde set. The 83 database ozone profiles selected were all measured between December and March at latitudes between 30N and 60N. The location of the DB sonde stations are shown in table 3.1. The DB sondes are interpolated to 50 levels from 1000 - 2 mb. In this study both the F2S and DB sondes were re-interpolated from the surface to 2 mb at 28 levels ; the pressure layer boundaries chosen are shown in table 3.2. This provides 27 layers with the correspondence of average layer pressure to layer number shown graphically in figure 3.1. The interpolation to 28

Location	**Latitude (N)**	**Longitude (W)**	**No. of sondes**
France	44.43	1.25	14
Portugal	38.78	9.13	4
Germany	52.22	-14.12	29
Sardinia	39.25	-9.05	10
Japan	43.05	-141.33	4
Japan	36.05	-140.13	6
Japan	31.63	-130.58	2
Canada	53.30	60.37	4
Canada	53.57	113.52	3
Canada	58.75	94.07	5
U.S.A.	37.85	75.48	2

Table 3.1 Location of DB sonde stations

Level	**Pressure, *p* (mb)**	**Level**	**Pressure, *p* (mb)**
1	surface	15	300
2	950	16	250
3	900	17	200
4	850	18	150
5	800	19	100
6	750	20	75
7	700	21	50
8	650	22	25
9	600	23	15
10	550	24	10
11	500	25	7
12	450	26	5
13	400	27	3
14	350	28	2

Table 3.2 Interpolation pressure layer boundaries for ozonesonde data

levels was done primarily to provide layers of equal pressure thickness in the troposphere but also to reduce the total number of levels. The former allows direct comparison of tropospheric layer radiances and the latter reduces the distribution of limited spectral information in the vertical from a large to a smaller number of layers. Midlatitude winter temperature, water vapour, and ozone data taken from a climatological database were merged with F2S sonde data up to 2mb if the sonde top pressure was greater than 2mb. The interpolated sonde data were used in the FASCODE radiance/transmittance code to calculate synthetic radiances. The uniform layering for pressures greater than 100mb was employed to minimise the effect of varying layer thicknesses in the forward model. This is explained further in §3.1.2.

A selection of the ozone concentration perturbation percentages for the 9 F2S sondes are shown in figure 3.2(a)-(d). The perturbations were calculated by differencing temporally consecutive sondes using

$$\frac{\delta q(p)}{q_1(p)} = \frac{q_2(p) - q_1(p)}{q_1(p)},$$

where $q(p)$ = ozone concentration profile in g/kg. The conversion of the ozone volume mixing ratio in ppmv to g/kg is achieved using

$$q(g/kg) = r(ppmv)\frac{MW_{ozone}}{MW_{air}}$$

where MW_{ozone}, MW_{air} = molecular weights of ozone and air (weighted) respectively.

The perturbation peaks are generally centred in the tropopause/lower stratospheric region due to the vertical movement of the tropopause height. In this region the ozone concentration is rapidly increasing in the vertical and thus ozonesonde comparisons are particularly sensitive to this irregular vertical movement.

The temporal separation between the sondes varies from several hours (17 Nov., 1757UTC to 17 Nov., 2251UTC) to several days (22 Nov., 0246UTC to 30 Nov., 0337UTC) and is a poor measure of the variability. The greatest difference occurs for the two sondes closest in time, 17 Nov., 1757UTC and 17 Nov., 2251UTC. The original data for these two sondes is shown in figure 3.3 and 3.4. The small scale lamination features, characterised by local minima in the profile as the ozone concentration is increasing, evident in figure 3.4, and to a lesser extent 3.3, in the lower stratosphere are typical of ozone profiles for the midlatitudes in winter (Reid and Vaughan, 1991).

The rapidly changing perturbations shown in figure 3.2(b) are due to lamination features which extend over greater vertical scales than those shown in figure 3.4. The 21 Nov. 1758UTC sonde, shown in figure 3.5, has an ozone lamination which extends into the troposphere. Dobson (1973) explains similar laminar features due to a layer of tropospheric air above a layer of stratospheric air caused by tropospheric folding events, however in this case the corresponding change in the temperature lapse rate is not present. The lapse rate in figure 3.5 does alter in the same region - around 300-400mb - but the change is very slight and is typical of an ozone lamination at high latitudes.

3.1.2 Radiance spectra

Synthetic spectra were generated using version 3.0P of the FASCODE (**F**ast **A**tmospheric **S**ignature **CODE**) atmospheric transmittance code (Clough *et al.* 1981, 1986) and HITRAN92 spectroscopic database (Rothman *et al.*, 1992). FASCODE has been extensively validated against measurements (Clough *et al.*, 1988; Revercomb *et al.*, 1989c, 1991b) and other atmospheric transmittance algorithms (International Radiation Commission, 1988). The spectra generated from FASCODE were processed to simulate measured spectra from the Atmospheric Emitted Radiance Interferometer (AERI) developed at the University of Wisconsin-Madison Space Science and Engineering Centre (Knuteson *et al.*, 1991). The AERI instrument has a longwave and a shortwave band in the thermal infrared region with wavenumber limits of 500-1750cm^{-1} and 1900-3650cm^{-1} respectively with 0.482cm^{-1} wavenumber spacing. The nominal radiometric noise level for an approximately 10 minute observation (including calibration) is 0.05mW/m^2.sr.cm^{-1} for the longwave band and 0.004mW/m^2.sr.cm^{-1} for the shortwave band. For the purposes of this study, radiances were calculated in the 980-1080cm^{-1} (9.6μm) spectral region where there is a strong ozone absorption feature and the longwave radiometric noise level applied to synthetic spectra was doubled to 0.1mW/m^2.sr.cm^{-1}.

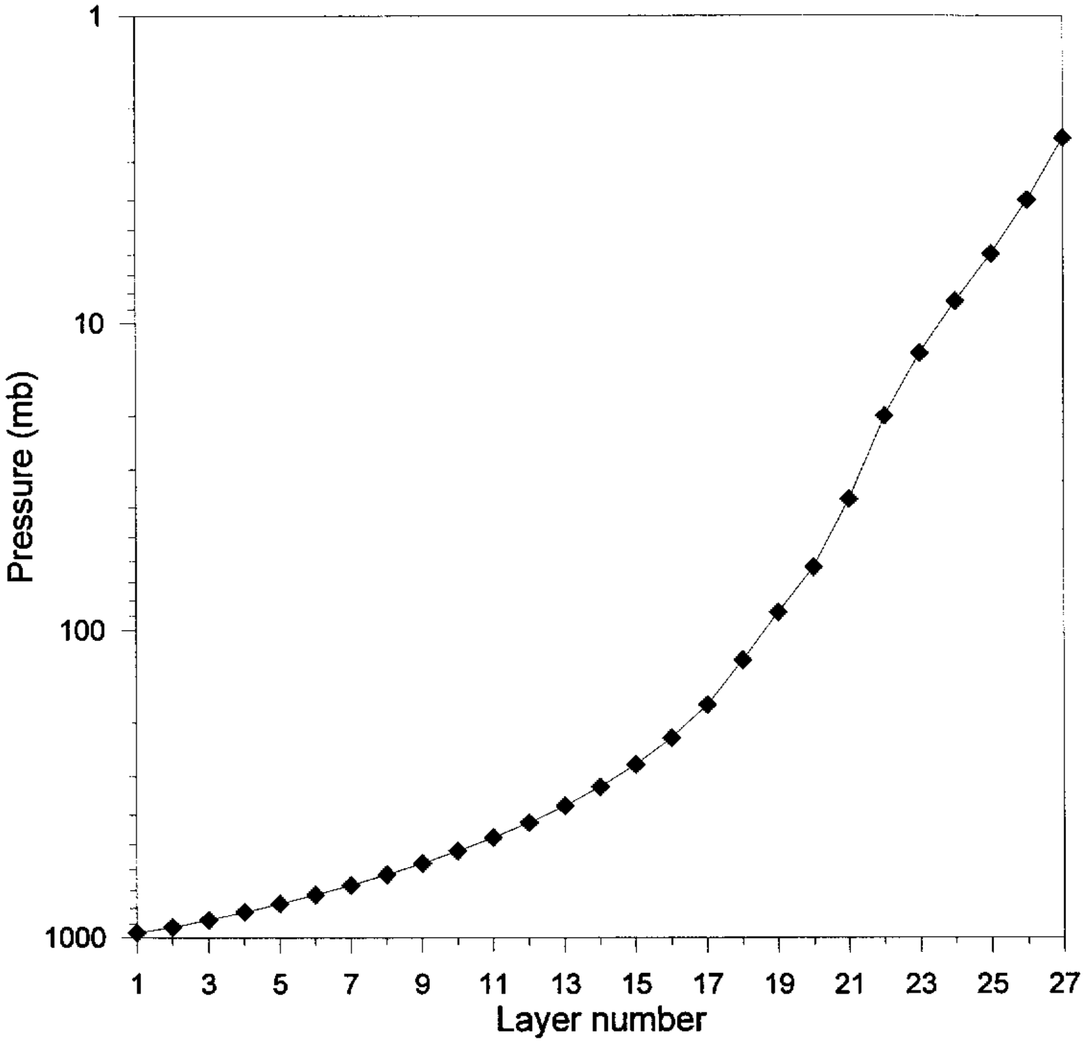
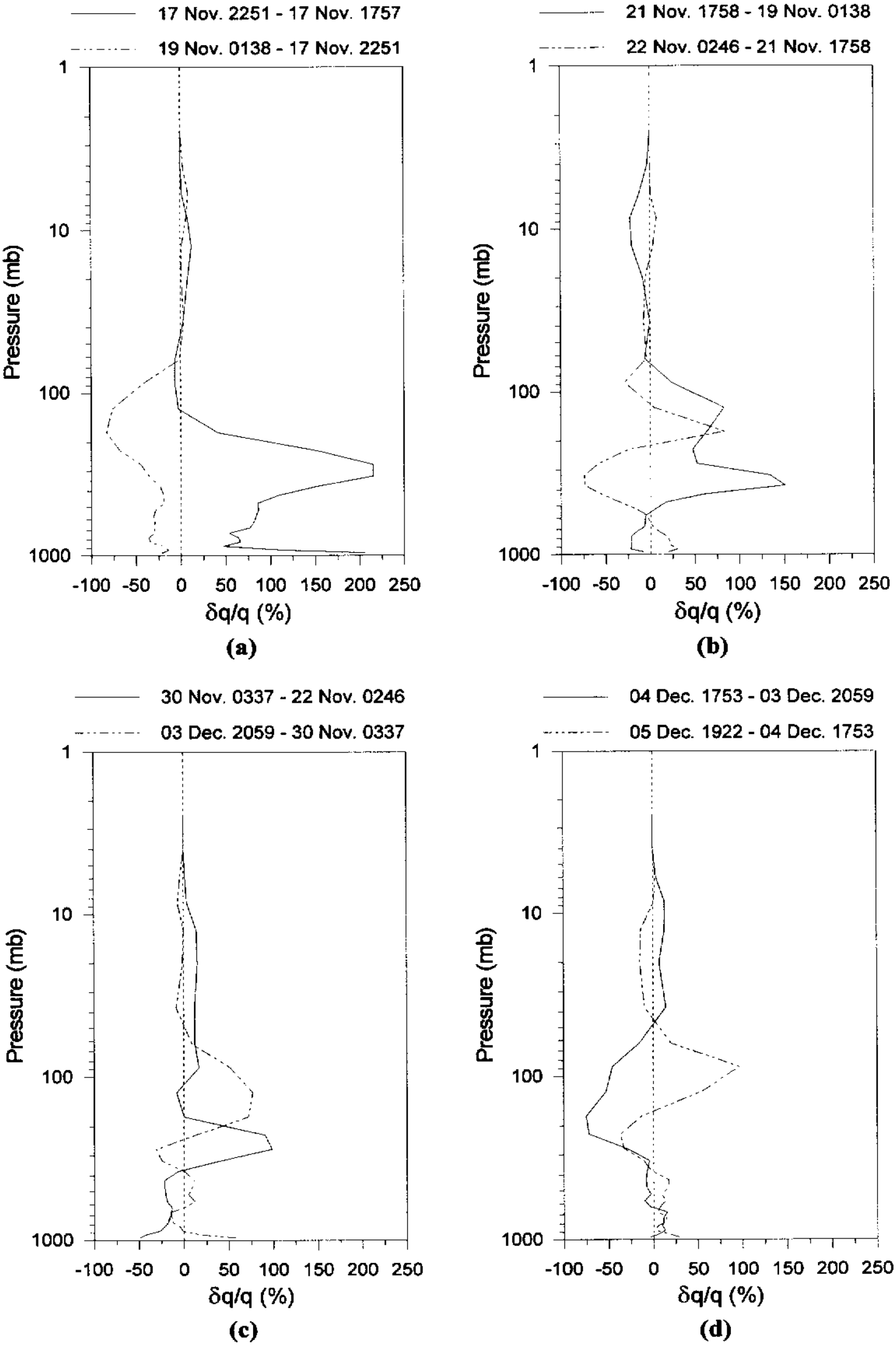


Figure 3.1 Correspondence of average layer pressure to layer number. Tropospheric layers are evenly spaced in pressure to allow direct comparison of layer radiances. Stratospheric layers are spaced to satisfy requirements in the FASCODE radiance/transmittance code.

High resolution FASCODE spectra were reduced to AERI wavenumber spacing by Fourier transforming FASCODE spectra into optical delay space and truncating the subsequent interferograms (IFGs)to the required length. Apodisation of the truncated interferograms was performed to remove the "ringing" effect introduced by the truncation. A Beer apodisation function, $B(n,N)$, was used, where

$$B(n,N) = \left[1 - \left(\frac{n}{N}\right)^2\right]^2 \qquad 1 \leq n \leq N \tag{3.1}$$

and N = total number of points retained in the interferogram. Figure 3.6 shows the apodisation function applied to a truncated interferogram. The effect of interferogram truncation and apodisation on the 9.6μm ozone radiance spectrum is shown in figure 3.7. Only the broad features of the ozone spectrum are visible together with some strong water vapour lines. Apodisation removes the spurious spectral oscillations but also degrades the spectral resolution.



(a) (b) (c) (d)

Figure 3.2 Ozone perturbation profiles from FIRE II/SPECTRE sonde set.

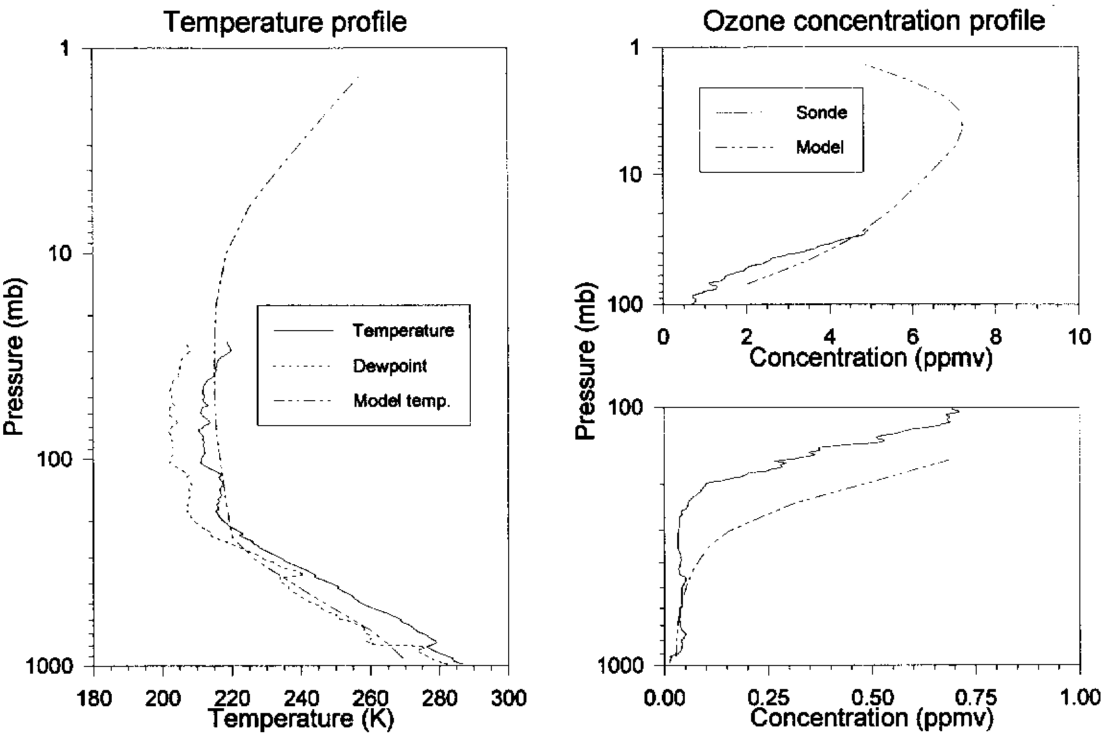


Figure 3.3 FIRE II/SPECTRE Wallops sonde, 17 Nov., 1757UTC (17NV1757 F2S) and midlatitude winter model data.

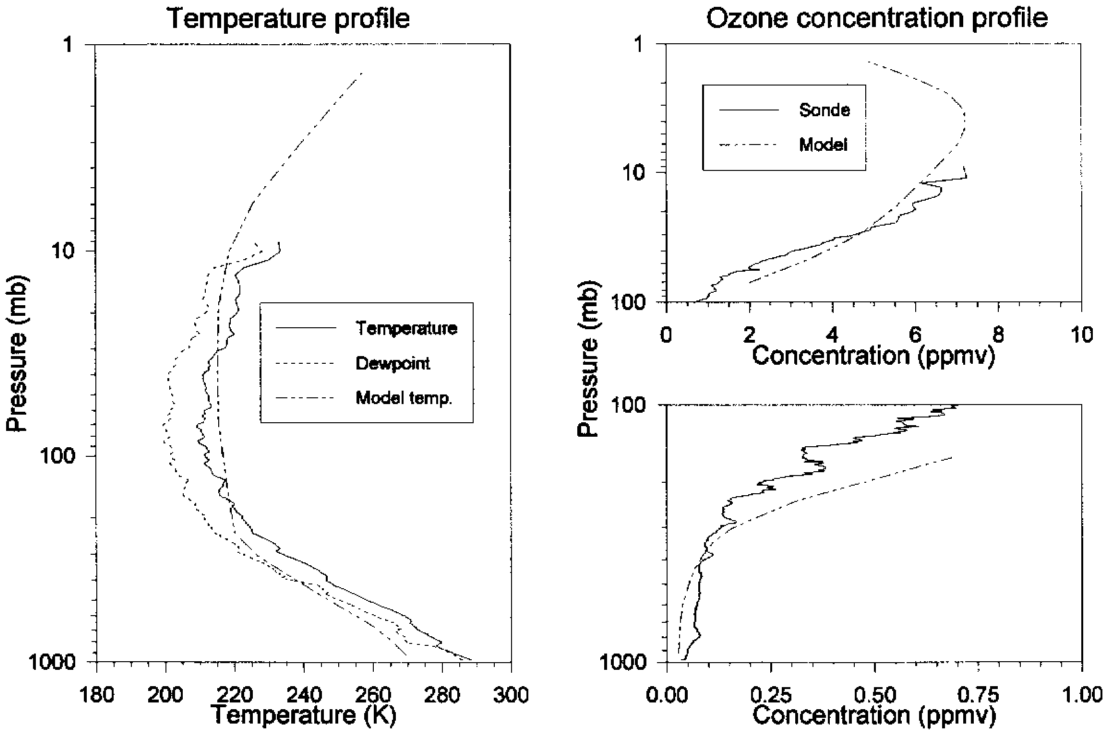


Figure 3.4 FIRE II/SPECTRE NOAA sonde, 17 Nov., 2251UTC (17NV2251 F2S) and midlatitude winter model data.

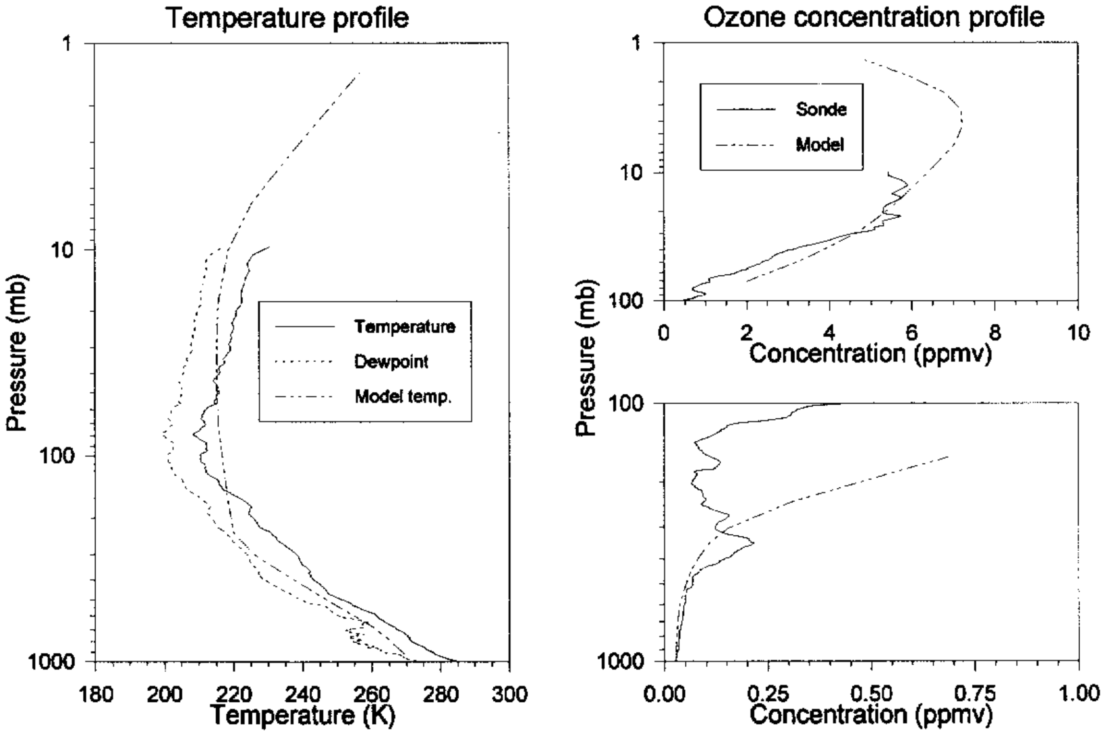


Figure 3.5 FIRE II/SPECTRE NOAA sonde, 21 Nov., 1758UTC (21NV1758 F2S) and midlatitude winter model data.

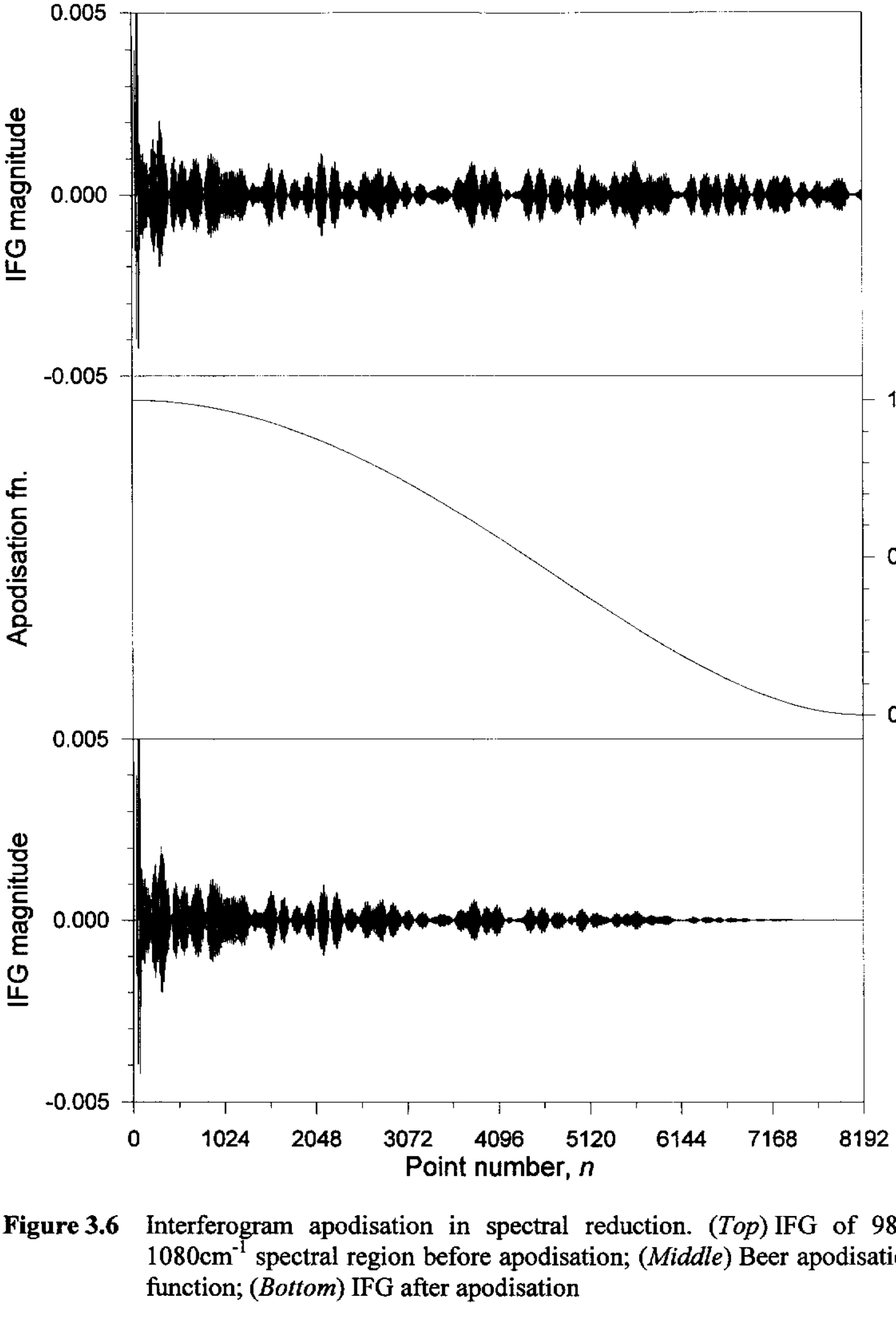


Figure 3.6 Interferogram apodisation in spectral reduction. (*Top*) IFG of 980-1080cm^{-1} spectral region before apodisation; (*Middle*) Beer apodisation function; (*Bottom*) IFG after apodisation

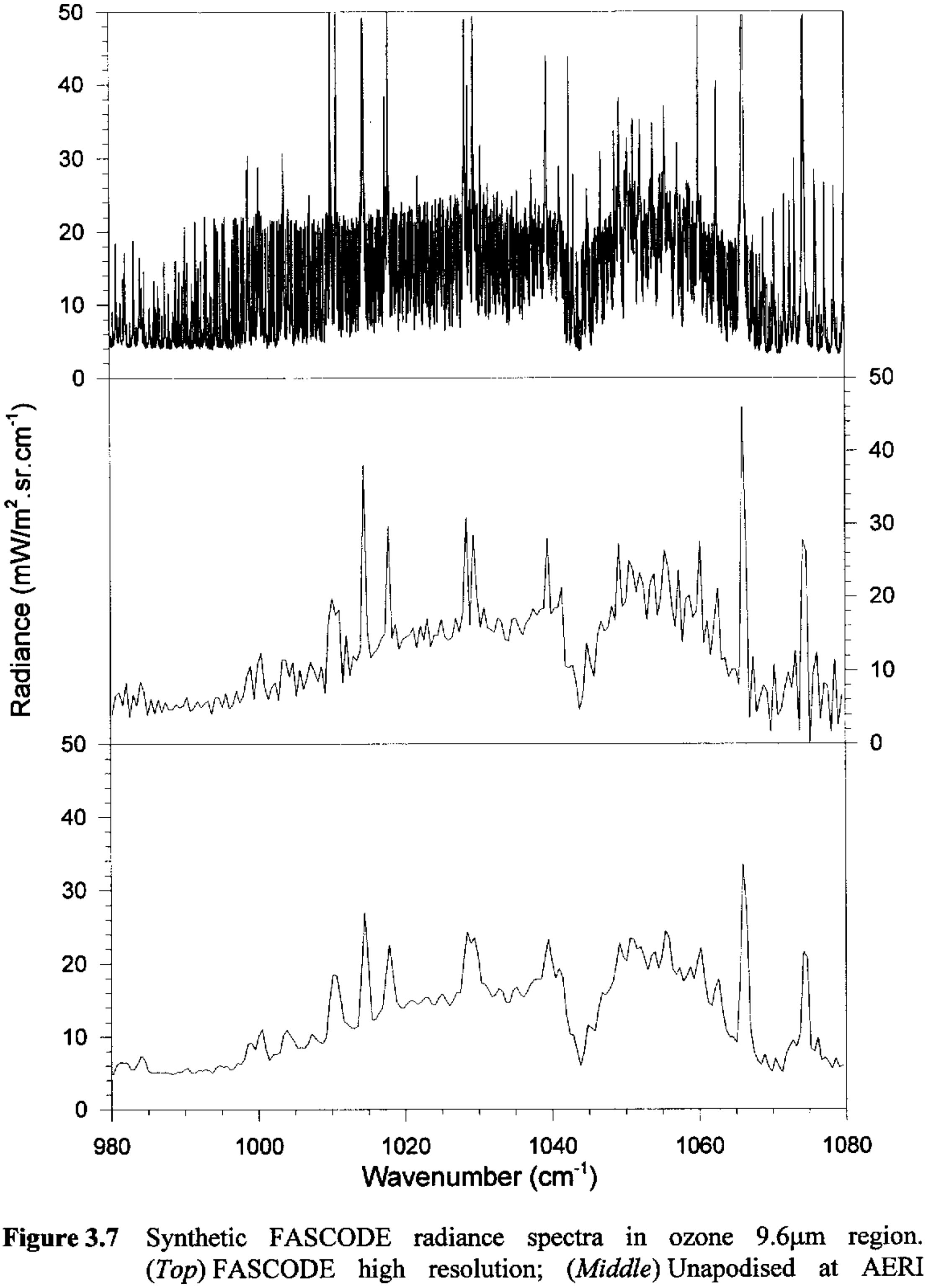


Figure 3.7 Synthetic FASCODE radiance spectra in ozone $9.6\mu m$ region. (*Top*) FASCODE high resolution; (*Middle*) Unapodised at AERI resolution; (*Bottom*) Apodised at AERI resolution.

3.2 RTE linearisation and validation

The method of linearisation of the RTE generally follows from that derived for the downlooking case by Revercomb (1991, private communication).

3.2.1 Derivation of linearised perturbed RTE

Restating the RTE for downwelling atmospheric infrared radiation (eqn. 2.1) we have,

$$R(\nu,p_s)=\int_0^{p_s} B(\nu,T(p))\frac{d\tau(\nu,p\to p_s)}{dp}dp, \qquad (3.2)$$

where $R(\nu,p_s)$ = downwelling radiance measured at the surface,

$B(\nu,T(p))$ = Planck radiance at temperature $T(p)$, and

$\tau(\nu,p\to p_s)$ = transmittance from pressure p to the surface. By definition, $\tau(\nu,p_s\to p_s)=1$.

To obtain the perturbed form of the upward looking RTE, an initial condition is defined,

$$R^o(\nu,p_s)=\int_0^{p_s} B(\nu,T^o(p))\frac{d\tau^o(\nu,p\to p_s)}{dp}dp, \qquad (3.3)$$

where the superscript ^{o} indicates an initial value of a parameter. The radiance perturbation is described by

$$\delta R(\nu,p_s)=R(\nu,p_s)-R^o(\nu,p_s). \qquad (3.4)$$

Representing the wavenumber functional dependence with the subscript ν, the perturbed radiative transfer equation (PRTE) can be written as

$$\begin{aligned}\delta R_\nu(p_s) &= \int_0^{p_s} B_\nu(T(p))\frac{d\tau_\nu(p\to p_s)}{dp}dp - \int_0^{p_s} B_\nu(T^o(p))\frac{d\tau_\nu^o(p\to p_s)}{dp}dp \\ &= \int_0^{p_s} \delta B_\nu(T(p))\frac{d\tau_\nu^o(p\to p_s)}{dp}dp - \int_0^{p_s} \delta\tau_\nu(p\to p_s)\frac{dB_\nu(T(p))}{dp}dp \\ &+ \left[B_\nu(T(p))\delta\tau_\nu(p\to p_s)\right]_0^{p_s}.\end{aligned} \tag{3.5}$$

The derivation of (3.5) from (3.3) and (3.4) is shown in Appendix D.

An expression must be found for $\delta\tau_\nu$ in terms of the mixing ratio perturbation $\delta q_i(p)$ where i labels a particular absorbing constituent. A perturbation in the mixing ratio at any pressure p will affect the atmospheric transmittance at pressures *above* the perturbation due to the modified opacity at p, whereas it will remain unchanged for pressure levels *below* the perturbation. It is therefore necessary to define a change in transmittance at pressure p due to a change in the mixing ratio of some absorber i at some other lower pressure p', where the integration limits identify the region of the atmosphere over which the transmittance perturbation is evaluated - from pressures 0 to p.

Using (2.2), the transmittance perturbation, as a first order approximation, may be written as,

$$\begin{aligned}\delta\tau_\nu(p\to p_s) &= \tau_\nu^o(p\to p_s)\delta\ln\tau_\nu(p\to p_s) \\ &= -\frac{\tau_\nu^o(p\to p_s)}{g}\int_p^{p_s}\sum_i^N k_{\nu i}(p')\delta q_i(p')dp'.\end{aligned} \tag{3.6}$$

Since at $p = 0$, $T \rightarrow 0$, so $B_\nu(T(p)) \rightarrow 0$; and from (3.6), $\delta\tau_\nu = 0$ at $p = p_s$ so (3.5) becomes,

$$\delta R_\nu(p_s) = \int_0^{p_s} \delta B_\nu(p) \frac{d\tau_\nu^o(p \rightarrow p_s)}{dp} dp - \int_0^{p_s} \delta\tau_\nu(p \rightarrow p_s) \frac{dB_\nu(T(p))}{dp} dp. \quad (3.7)$$

The first term of (3.7) can be treated as in Smith *et al.* (1991) and is concerned with the temperature perturbation δT derived from δB_ν. The second term of (3.7) describes the effect on the radiance perturbation arising from a mixing ratio perturbation δq. Working under the premise that the temperature profile $T(p)$ is a known quantity (i.e. it has already been retrieved or is otherwise known), only the second term is considered in this study.

Because the lowest pressure in the ozonesonde database used in this study is 2mb, it is necessary to define an upper boundary in the linearised PRTE. This will allow use of in situ data that is not complete at low pressures. A pressure level, p_{top}, can be defined as an intermediate boundary such that

$$\begin{aligned} \int_0^{p_s} \delta\tau_\nu(p \rightarrow p_s) \frac{dB_\nu(T(p))}{dp} dp &= \int_{p_{top}}^{p_s} \delta\tau_\nu(p \rightarrow p_s) \frac{dB_\nu(T(p))}{dp} dp \\ &+ \int_0^{p_{top}} \delta\tau_\nu(p \rightarrow p_s) \frac{dB_\nu(T(p))}{dp} dp \end{aligned} \quad (3.8)$$

The transmittance perturbation above for pressures less than p_{top} can be decomposed into

$$\delta\tau_\nu(p \rightarrow p_s) = \tau_\nu(p \rightarrow p_{top})\tau_\nu(p_{top} \rightarrow p_s) - \tau_\nu^o(p \rightarrow p_{top})\tau_\nu^o(p_{top} \rightarrow p_s), \quad (3.9)$$

where $p < p_{top}$. The problem is tailored specifically for ozone if p_{top} is chosen to be a pressure lower than that at the region of maximum absorption by ozone, in this case the stratosphere. For $p < p_{top}$, perturbations in the ozone concentration will have only a minor effect on the absolute transmittance. The effect of other species, such as water vapour, is minor because the concentrations are low in the upper stratosphere and because the spectral region of interest sits in the atmospheric window where only the water vapour continuum exists. If an assumption is made that, for the 980-1080cm^{-1} spectral region, the atmospheric transmittances for $p < p_{top}$ are *constant* and *unity* (i.e. above the region of maximum absorption by ozone), that is

$$\tau_{\nu}\left(p \rightarrow p_{top}\right) = \tau_{\nu}^{o}\left(p \rightarrow p_{top}\right) = 1 \text{ for } p < p_{top}, \tag{3.10}$$

then (3.9) can be simplified to,

$$\begin{aligned} \delta\tau_{\nu}\left(p \rightarrow p_{s}\right) &= \tau_{\nu}^{o}\left(p \rightarrow p_{top}\right)\delta\tau_{\nu}\left(p_{top} \rightarrow p_{s}\right) \\ &= \delta\tau_{\nu}\left(p_{top} \rightarrow p_{s}\right). \end{aligned} \tag{3.11}$$

The validity of this assumption expressed in (3.10) is addressed in the next section.

Replacing the expression (3.11) into the last term of (3.8) gives

$$\begin{aligned} \int_{0}^{p_{top}} \delta\tau_{\nu}\left(p \rightarrow p_{s}\right)\frac{dB_{\nu}(T(p))}{dp}dp &= \int_{0}^{p_{top}} \delta\tau_{\nu}\left(p_{top} \rightarrow p_{s}\right)\frac{dB_{\nu}(T(p))}{dp}dp \\ &= \delta\tau_{\nu}\left(p_{top} \rightarrow p_{s}\right)\int_{0}^{p_{top}} \frac{dB_{\nu}(T(p))}{dp}dp \\ &= \delta\tau_{\nu}\left(p_{top} \rightarrow p_{s}\right)B_{\nu}\left(T(p_{top})\right), \end{aligned} \tag{3.12}$$

where $B_{\nu}\left(T(p_{top})\right)$ = boundary radiance.

The full expression for (3.8) thus becomes

$$\int_0^{p_s} \delta\tau_\nu(p \to p_s)\frac{dB_\nu(T(p))}{dp}dp = \int_{p_{top}}^{p_s} \delta\tau_\nu(p \to p_s)\frac{dB_\nu(T(p))}{dp}dp + \delta\tau_\nu(p_{top} \to p_s)B_\nu(T(p_{top})) \tag{3.13}$$

Substituting (3.6) into (3.13) and removing the summation over all atmospheric molecules, that is, considering only ozone, gives

$$\begin{aligned}\delta R_\nu(\delta q, p_s) = & \int_{p_{top}}^{p_s} \frac{dB_\nu(T(p))}{dp}\left[\frac{\tau_\nu^o(p \to p_s)}{g}\int_p^{p_s} k_\nu(p')\delta q(p')dp'\right]dp \\ & + \frac{B_\nu(T(p_{top}))\tau_\nu^o(p_{top} \to p_s)}{g}\int_{p_{top}}^{p_s} k_\nu(p')\delta q(p')dp',\end{aligned} \tag{3.14}$$

which may be simplified by integration by parts to give,

$$\delta R_\nu(\delta q, p_s) = \int_{p_{top}}^{p_s} \frac{k_\nu(p)}{g}\delta q(p)\left[\int_{p_{top}}^{p} \tau_\nu^o(p' \to p_s)\frac{dB_\nu(T(p'))}{dp'}dp' + \tau_\nu^o(p_{top} \to p_s)B_\nu(T(p_{top}))\right]dp, \tag{3.15}$$

which is a linear integral equation. The algebraic manipulation to arrive at (3.15) from (3.14) is provided in Appendix E. Equation (3.15) may be expressed more clearly in the form of (2.3) as,

$$\delta R_\nu(\delta q, p_s) = \int_{p_{top}}^{p_s} W_Q\left(\tau_\nu^o(p \to p_s), T(p), k_\nu(p)\right)\,\delta q(p)dp, \text{ where} \tag{3.16a}$$

$$W_Q\left(\tau_\nu^o(p \to p_s), T(p), k_\nu(p)\right) = \frac{k_\nu(p)}{g}\left[\int_{p_{top}}^{p} \tau_\nu^o(p' \to p_s)\frac{dB_\nu(T(p'))}{dp'}dp' + \tau_\nu^o(p_{top} \to p_s)B_\nu(T(p_{top}))\right] \tag{3.16b}$$

Equation (3.16a) is a Fredholm equation of the first kind with fixed limits and (3.16b) describes the form of the weighting function or kernel - the radiative response of each pressure level to a unit ozone mass mixing ratio perturbation, namely 1g/kg. Note that in (3.16b) the quantity k_ν is the absorption coefficient for ozone only and τ_ν is the atmospheric transmittance for *all* absorbing constituents present, and also that the quantity dp has been incorporated into W.

3.2.2 Boundary pressure validation

The top pressure of the database sondes, 2mb, was selected as the upper boundary pressure, p_{top}. To determine if this p_{top} was sufficiently high enough to satisfy the assumption described in (3.10) two quantities were examined : the transmittances, both the cumulative total (all absorbers) and layer-by-layer (ozone only), and the variation of the linear radiance perturbations at the surface calculated for two upper boundary pressures; 1mb and 0.1mb. The former justifies the acceptance of (3.10) and the latter shows how, given that (3.10) is satisfied, the linear model is insensitive to higher boundary pressure, p_{top}, selections.

Cumulative and layer-by-layer transmittance calculations

The cumulative total transmittances $\tau(p \rightarrow p_s)$ up to 0.1mb for a number of selected channels in the spectral range 980-1080cm^{-1} are shown in figure 3.8. Above 2mb, the cumulative transmittances assume a near vertical profile, indicating that the layer transmittances at this point and above are close to unity. At this height the only other significant absorber in the spectral vicinity - water vapour - does not contribute to transmittance perturbations. The ozone layer-by-layer transmittances $\tau(p \rightarrow p+\Delta p)$, shown in figure 3.9, indicate that IR absorption due to ozone occurs mainly in the atmospheric region where the ozone mixing ratio is rapidly increasing, not near the ozone peak at 4-6mb. At 2mb, the minimum ozone layer transmittance for the frequencies selected is approximately 0.97, increasing to unity at 0.3-0.4mb. Thus, the radiance contribution of the atmosphere above 2mb and its impact on the total radiance perturbation at the surface must be determined.

Variation of linear radiance perturbation with boundary pressure

To determine how the selection of p_{top} affects the calculated radiance perturbation, ΔR, large ozone concentration perturbations were applied in the lower troposphere and stratosphere to the U.S. Standard (STD) atmosphere (NOAA, 1976) temperature and constituent profiles. The temperature and ozone mixing ratio profiles up to 0.1mb are shown in figure 3.10 for the STD atmosphere. We see that the 2mb boundary pressure resides just below the stratopause isothermal region. A comparison of figures 3.2 and 3.10 shows that the major ozone perturbations generally correspond to the tropopause isothermal region of the atmosphere.

The arbitrarily large ozone perturbations applied were Gaussian and of the form

$$\frac{\delta q_i(p)}{q_i^o(p)} = \frac{\delta q_i(p_o)}{q_i^o(p_o)} \exp\left[-a(p-p_o)^2\right], \tag{3.20}$$

where p_o = pressure of the perturbation maximum,

$a = \ln(2)/\alpha^2$, slope constant of Gaussian curve, and

α = halfwidth of the perturbation in mb.

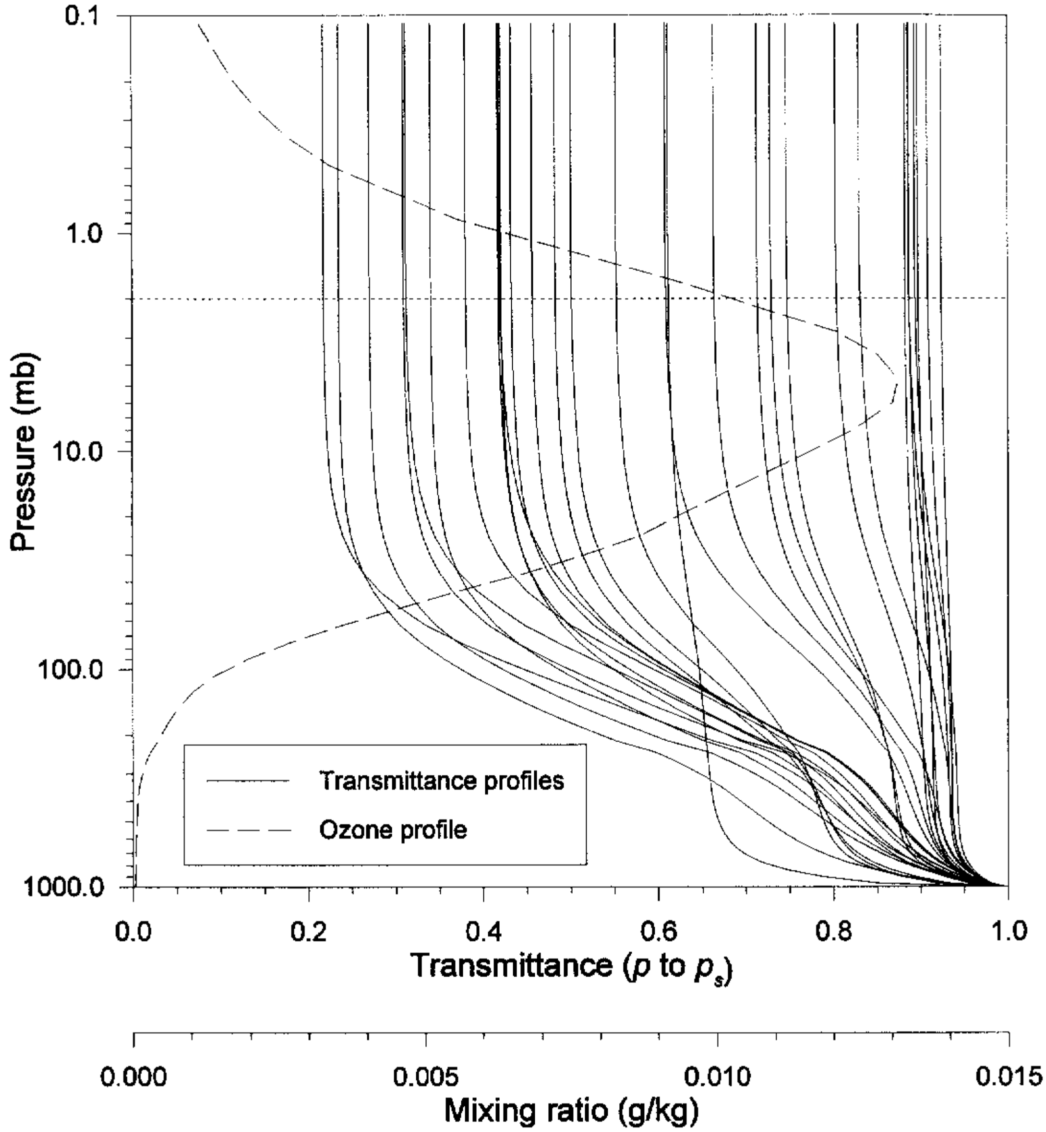


Figure 3.8 FASCODE cumulative total transmittance and ozone profiles for U.S. Standard Atmosphere (1976) to 0.1mb for selected channels. Spectral region : 980-1080cm^{-1}; $\Delta\nu = 0.482\text{cm}^{-1}$.

We use the terms α_{lo} and α_{hi} to describe the perturbation halfwidths on the low and high pressure sides of the perturbation maximum respectively. The applied perturbations were lower tropospheric (p_o = 1000mb, α_{lo} = 100mb) and middle stratospheric (p_o = 50mb, α_{lo} = 10mb, α_{hi} = 100mb). The size of the perturbations, +200%, were chosen to provide enough radiance signal to identify any discrepancies due to boundary pressure selection.

The linear radiance perturbation at the surface ($R_{STD\ perturbed}$ - R_{STD}), calculated for a tropospheric ozone perturbation, is shown in figure 3.11 and for the stratospheric ozone perturbation in figure 3.12. The differences in the radiative responses, based on our earlier comments and figure 3.8 and 3.9, are what is expected for ozone perturbations in the troposphere and stratosphere. The STD atmosphere perturbed response shows the use of one or the other (0.1mb, 2mb) top boundary pressure has negligible effect on the linear radiance perturbation - any residual signature being well below the nominal noise level of $0.1 \mathrm{mW/m^2.sr.cm^{-1}}$, confirming the relative insensitivity of the radiative response to boundary pressure. Accordingly, we conclude, the condition of (3.10) is largely satisfied within the prescribed noise levels.

3.2.3 FASCOD3P Comparisons

How well does the linear model of (3.16a) compare with the non-linear model FASCOD3P? The two models were compared using the same tropospheric and stratospheric perturbations to the STD atmosphere as described previously. The atmospheric radiative response to ozone concentration change is quasi-linear in the troposphere and highly nonlinear in the stratosphere.

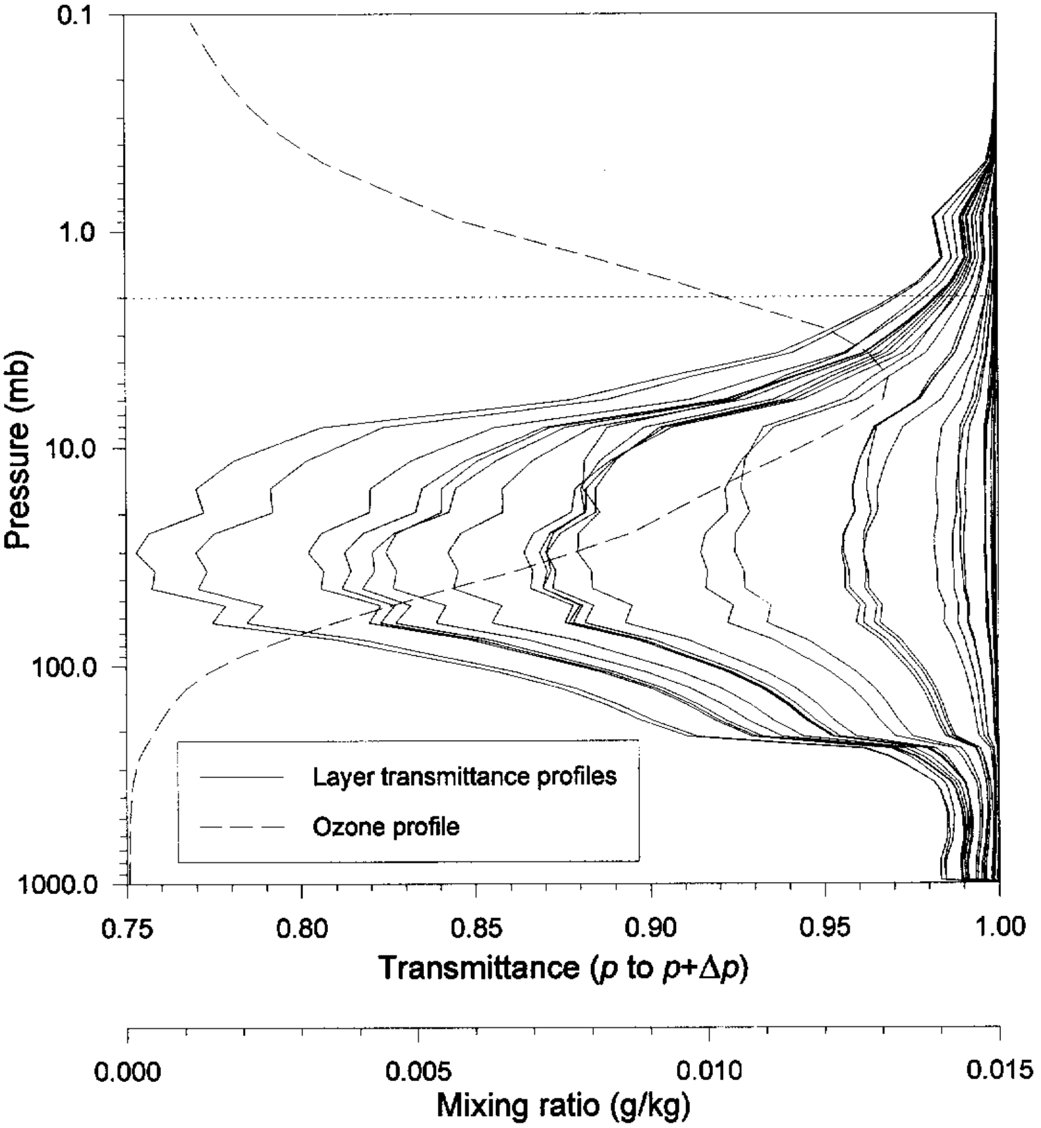
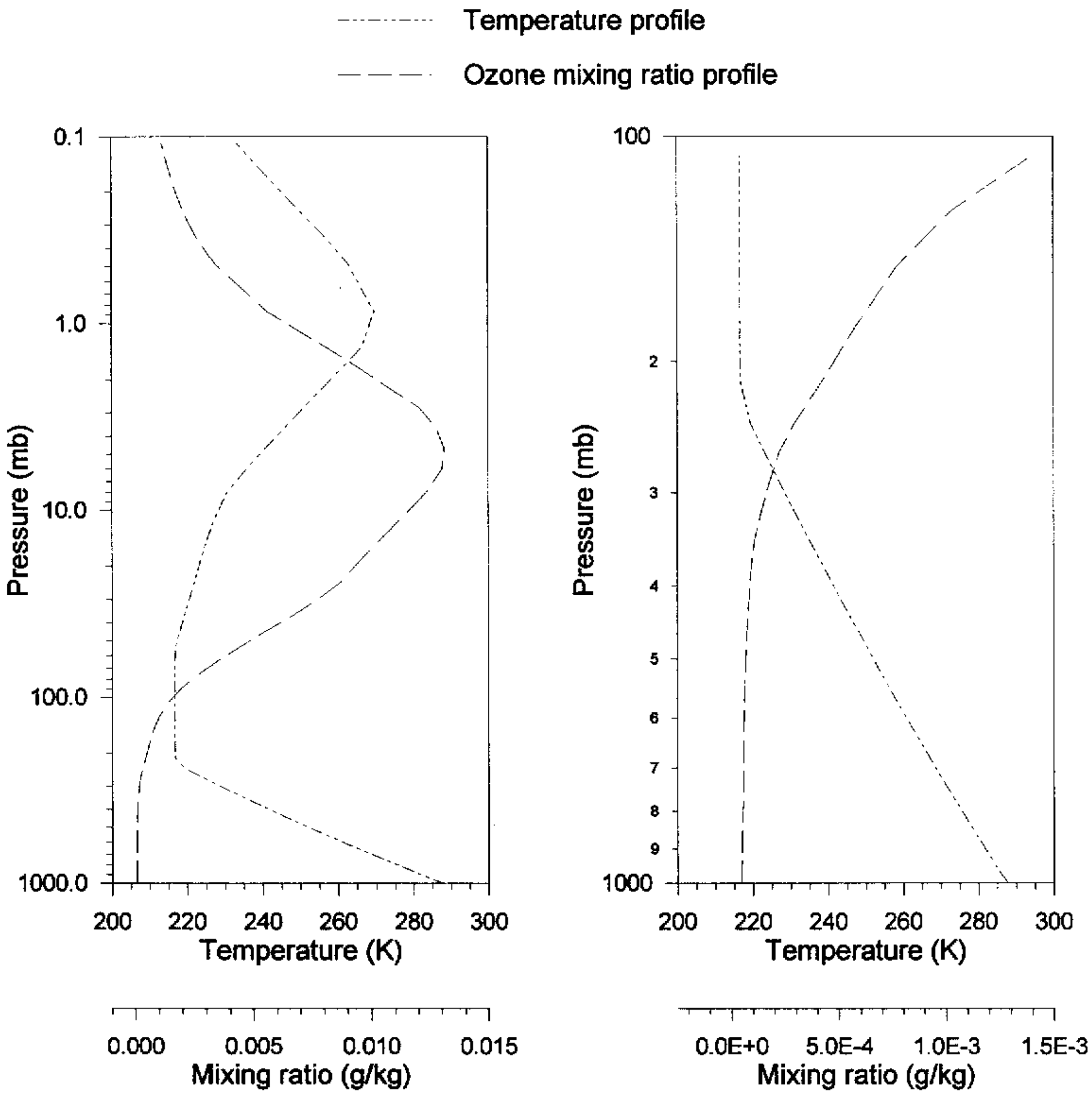


Figure 3.9 FASCODE layer ozone transmittance and ozone profiles for U.S. Standard Atmosphere (1976) to 0.1mb for selected channels. Spectral region : 980-1080cm^{-1}; $\Delta\nu = 0.482\text{cm}^{-1}$.



(a) **(b)**

Figure 3.10 **(a)** Temperature and ozone mixing ratio profiles for U.S. Standard atmosphere (1976). **(b)** Tropospheric magnification

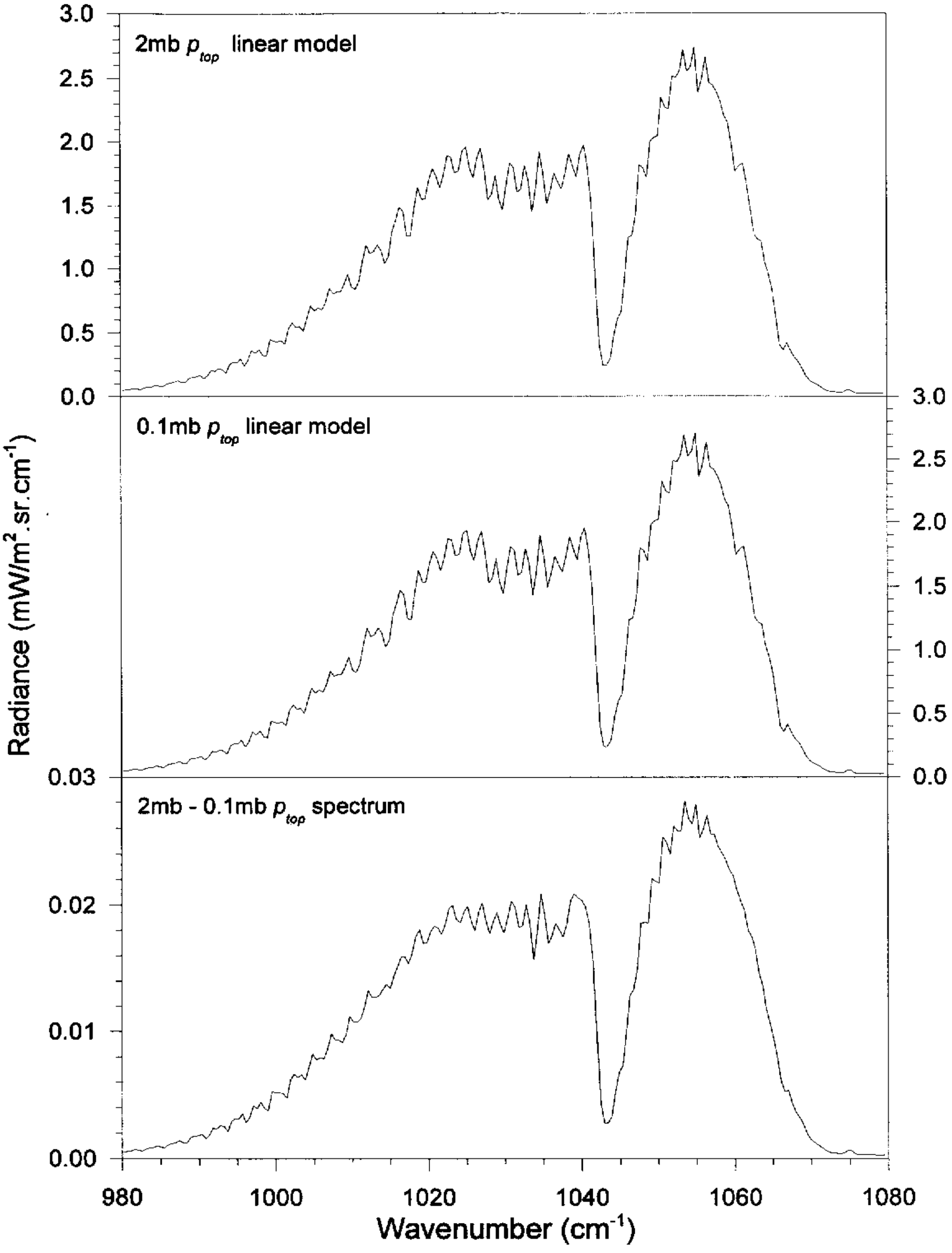


Figure 3.11 Variation in linear radiance perturbation ($R_{STD\,perturbed}$ - R_{STD}) with upper boundary pressure for a +200% Gaussian tropospheric ozone concentration perturbation at the 1000mb level to the U.S. Standard atmosphere (1976). (*Top*) 2mb upper boundary pressure, (*Middle*) 0.1mb upper boundary pressure, (*Bottom*) Difference

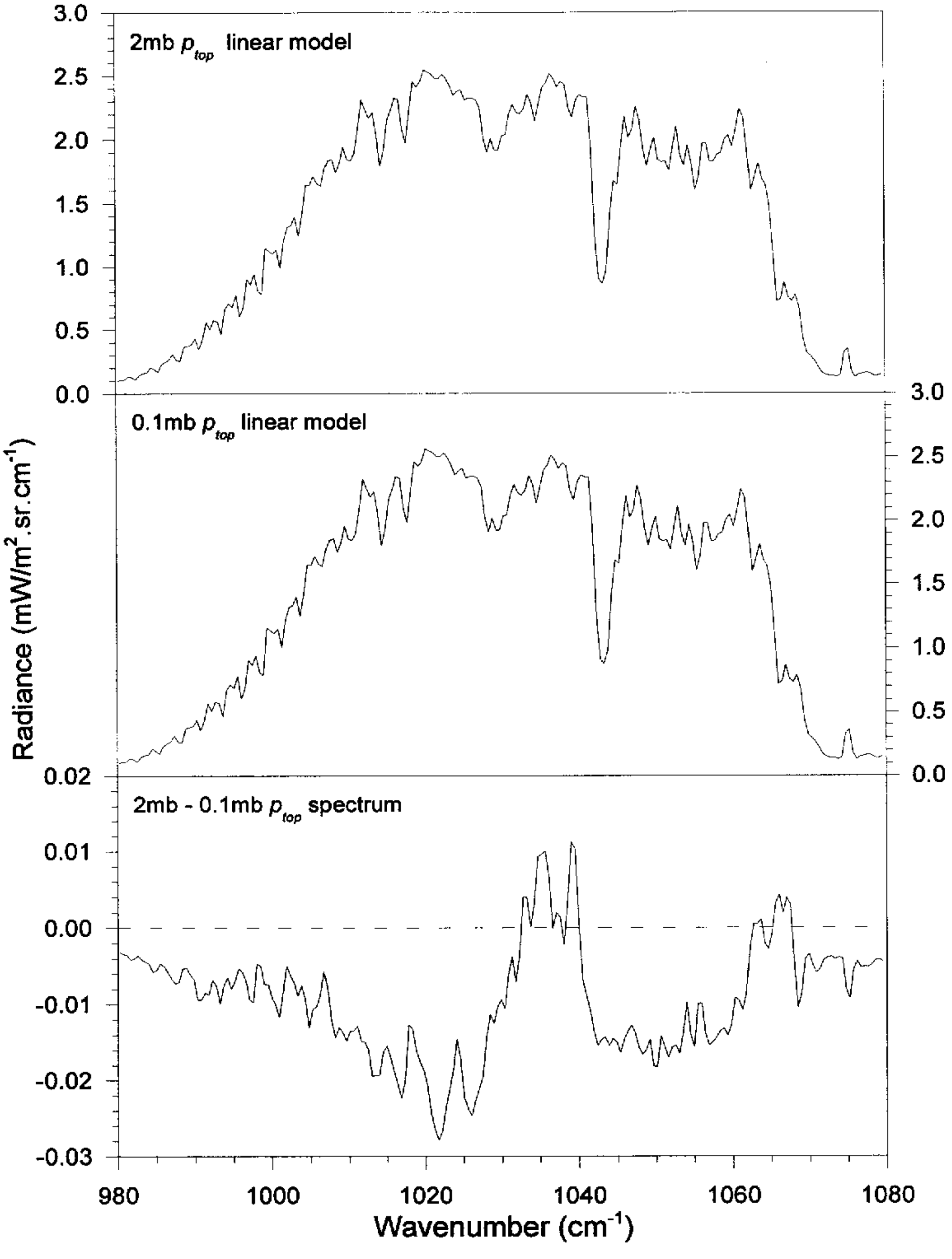


Figure 3.12 Variation in linear radiance perturbation ($R_{STD\,perturbed}$ - R_{STD}) with upper boundary pressure for a +200% Gaussian stratospheric ozone concentration perturbation at the 50mb level to the U.S. Standard atmosphere (1976). (*Top*) 2mb upper boundary pressure, (*Middle*) 0.1mb upper boundary pressure, (*Bottom*) Difference

The magnitude of the radiative response in the troposphere is driven by the temperature lapse rate so even though ozone concentrations are low, the associated radiance emitted is significant. Ozone absorption in the troposphere has a small effect on total transmittance due to its low concentration, hence transmittance perturbations for tropospheric ozone changes are small and within a linear regime. In the stratosphere, however, transmittance changes in the 980-1080cm^{-1} spectral interval are dependent almost solely on ozone due to a lack of other absorbers. Thus, changes in stratospheric ozone concentration cause large transmittance perturbations. The relationship between mixing ratio and transmittance is exponential, which results in a nonlinear radiative response.

The calculated ΔR for the linear and FASCODE models, for the tropospheric ozone perturbation to the STD atmosphere (2mb p_{top}), are shown in figure 3.13. The agreement between the two models is very good with the differences attributible to the nonlinear response component. Figure 3.14 shows the relationship between FASCODE ΔR and the Gaussian perturbation maximum fraction, $\delta q(p_o)/q^o(p_o)$, for a selection of wavenumbers in the ozone 9.6μm emission region. The wavenumbers chosen fall both in strongly and weakly emissive spectral regions. The predominantly linear tropospheric response is evident even for large ozone perturbations. Similarly, results for the stratospheric ozone concentration perturbation are shown in figures 3.15 and 3.16. The linear model greatly overestimates the FASCODE nonlinear radiance perturbations even for small stratospheric ozone perturbations for the spectral region where the emission is large (i.e. for the 1054.938cm^{-1} channel for example).

Ignoring numerical precision errors, the difference between models is due the breakdown of the assumption that the transmittance perturbations can be represented by a linear approximation. The transmittance perturbation as a limit and as a finite quantity can be expressed as,

$$\frac{d\tau_\nu}{\tau_\nu^o} = d\ln\tau_\nu \text{ , and} \tag{3.21a}$$

$$\frac{\delta\tau_\nu}{\tau_\nu^o} \approx \delta\ln\tau_\nu \tag{3.21b}$$

respectively. The validity of the approximation expressed in (3.6) requires the equivalence of (3.21a) to (3.21b) which is not valid for large $\delta\tau_\nu$. This is readily apparent when the finite difference of the logarithm of transmittances, $\delta\ln\tau_\nu$, is rearranged to give,

$$\delta\ln\tau_\nu = \ln\left[1+\frac{\delta\tau_\nu}{\tau_\nu^o}\right], \tag{3.22}$$

and thus (3.21) holds true only for $\delta\tau_\nu/\tau_\nu^o$ small (≤ 0.1).

The regime in which the approximation is within acceptable bounds can be inspected by plotting the logarithmic relationship (3.22) against $\delta\tau_\nu/\tau_\nu^o$, the LHS of equation (3.21b). This relationship is displayed in figure 3.17, which shows the approximation is satisfactory to within about 10% for $\delta\tau_\nu/\tau_\nu^o$ excursions of up to ± 20%. The FASCODE calculated transmittance perturbation, for the stratospheric ozone perturbation, is shown in figure 3.18. In the channels where ozone emission is stronger, the transmittance perturbations are quite large - in this case up to 55%. It should also be noted that, by definition, errors in the transmittance perturbation approximation will accumulate and propagate downward through the atmosphere as transmittance is a cumulative product.

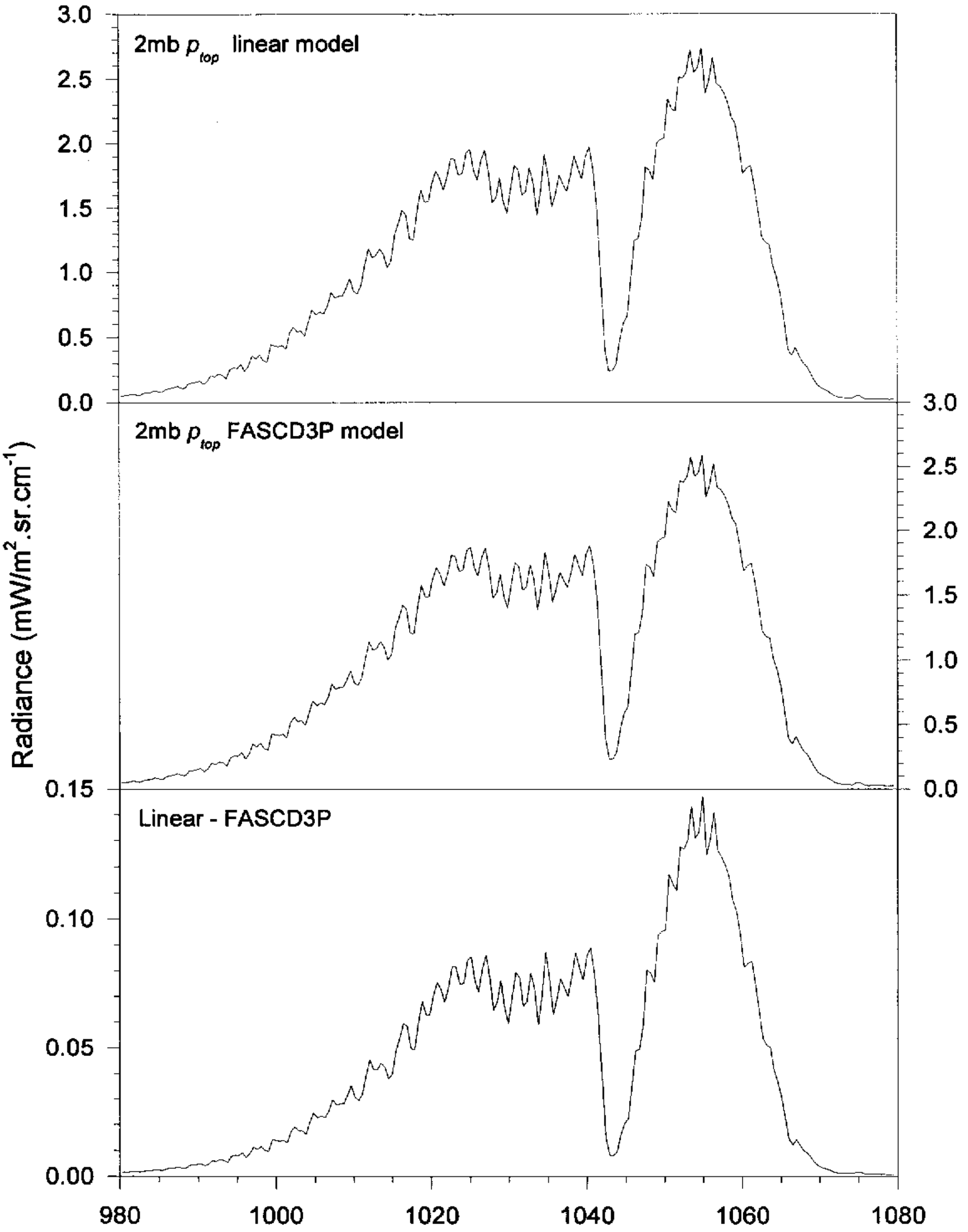


Figure 3.13 Linear and FASCOD3P radiance perturbation difference for a +200% Gaussian tropospheric ozone concentration perturbation at the 1000mb level to the U.S. Standard atmosphere (1976). (*Top*) Linear model, (*Middle*) FASCOD3P model, (*Bottom*) Difference

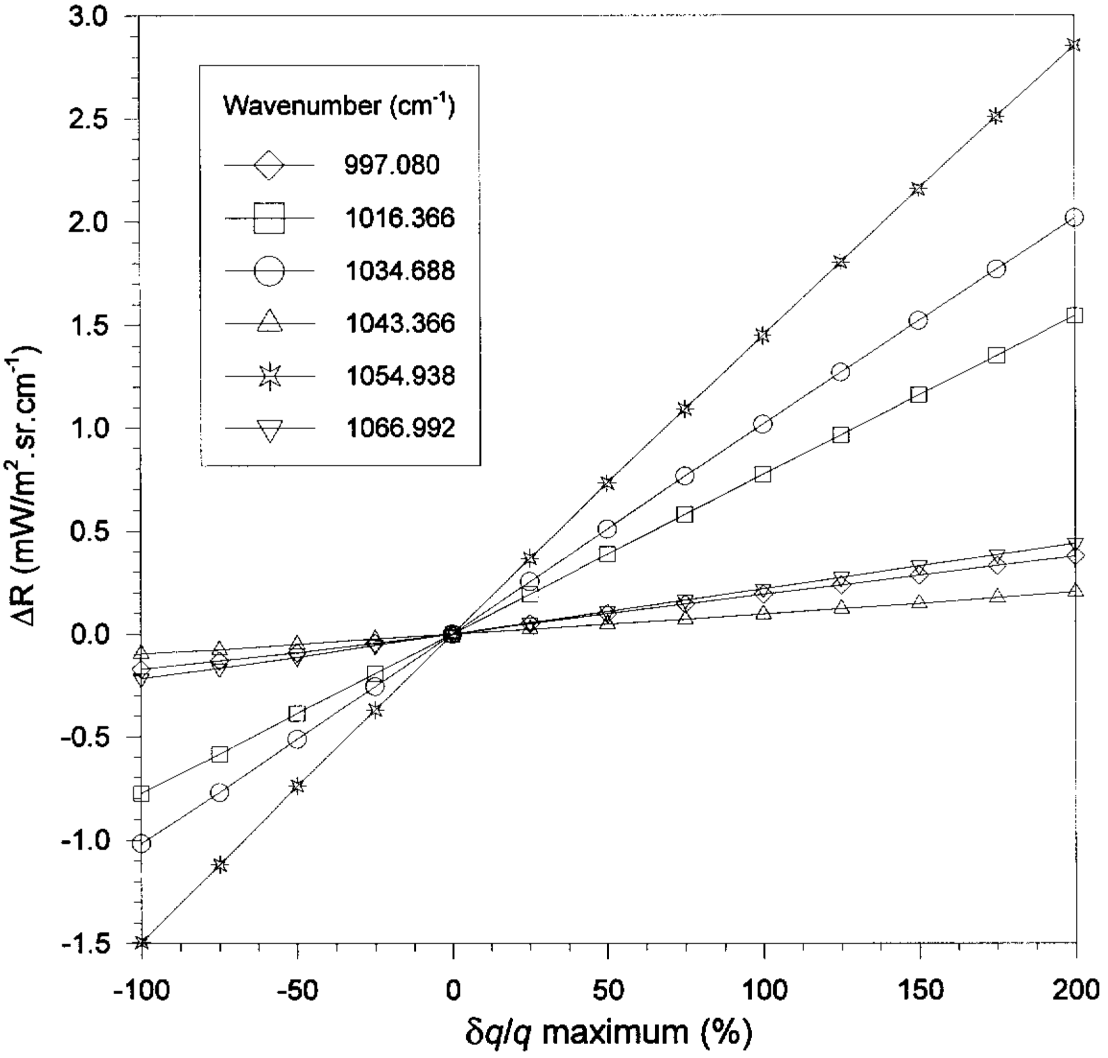


Figure 3.14 FASCOD3P radiative response curves for selected wavenumbers as a function of tropospheric ozone mixing ratio perturbation at the 1000mb level to the U.S. Standard atmosphere (1976).

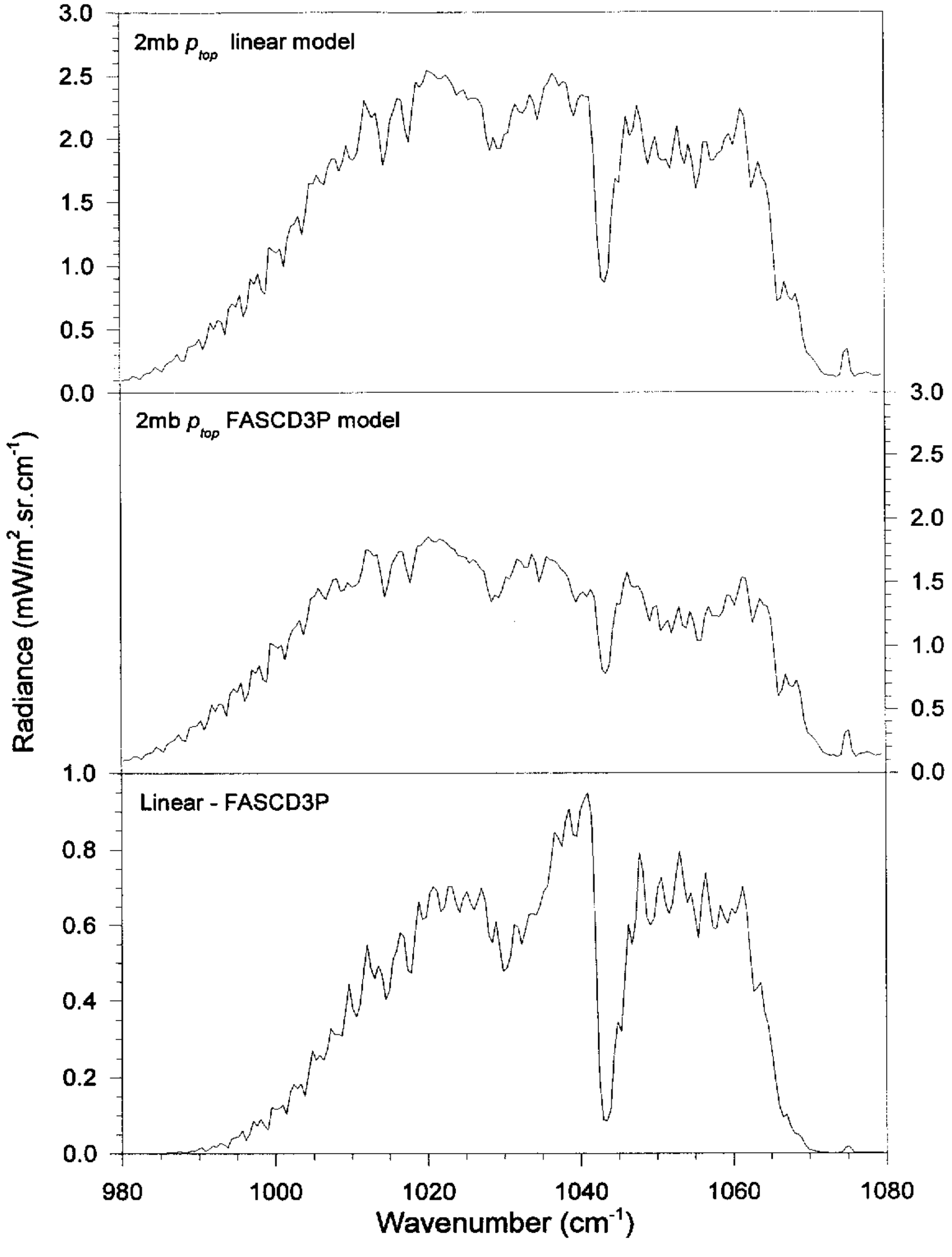


Figure 3.15 Linear and FASCOD3P radiance perturbation difference for a +200% Gaussian stratospheric ozone concentration perturbation at the 50mb level to the U.S. Standard atmosphere (1976). (*Top*) Linear model, (*Middle*) FASCOD3P model, (*Bottom*) Difference

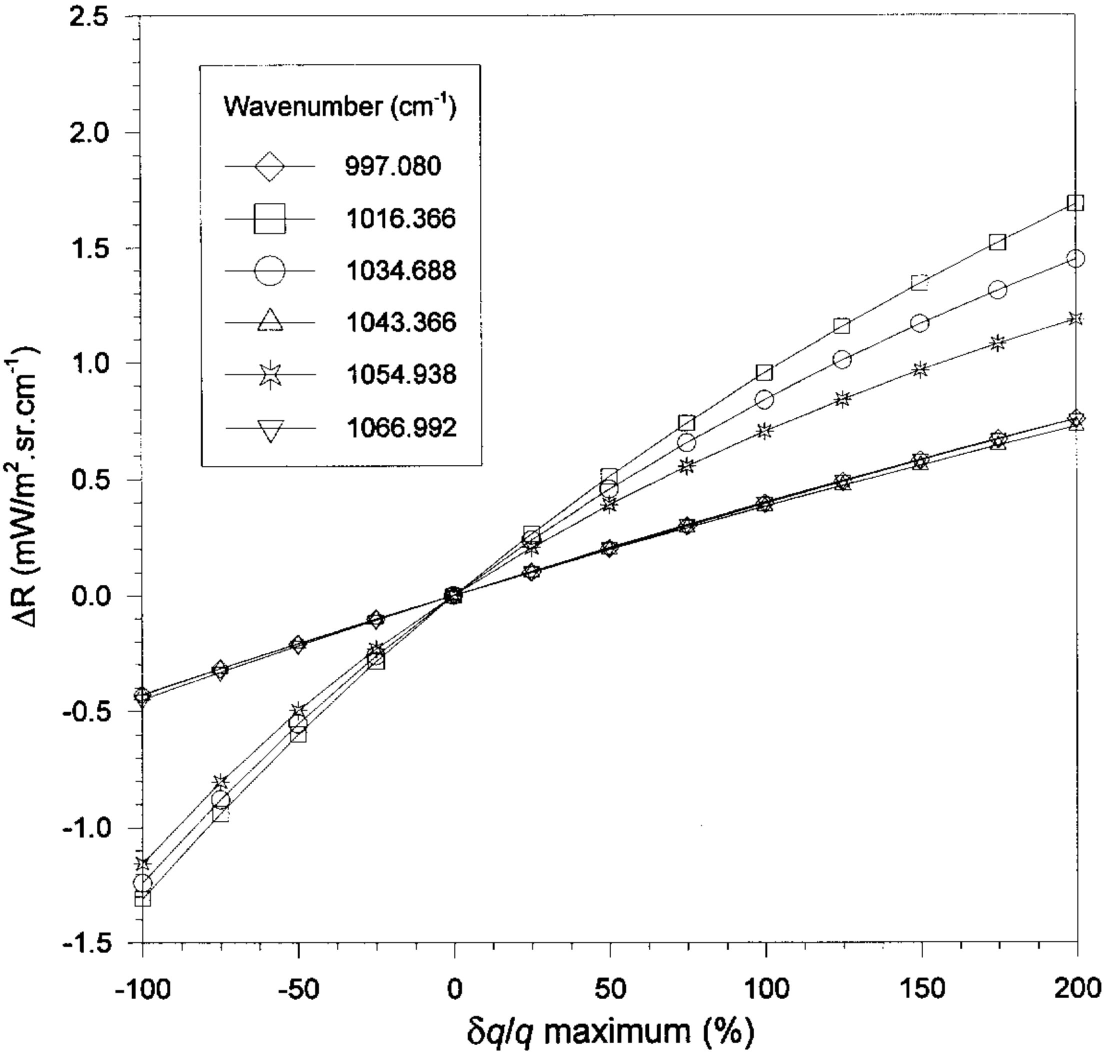


Figure 3.16 FASCOD3P radiative response curves for selected wavenumbers as a function of a stratospheric ozone mixing ratio perturbation at the 50mb level to the U.S. Standard atmosphere (1976).

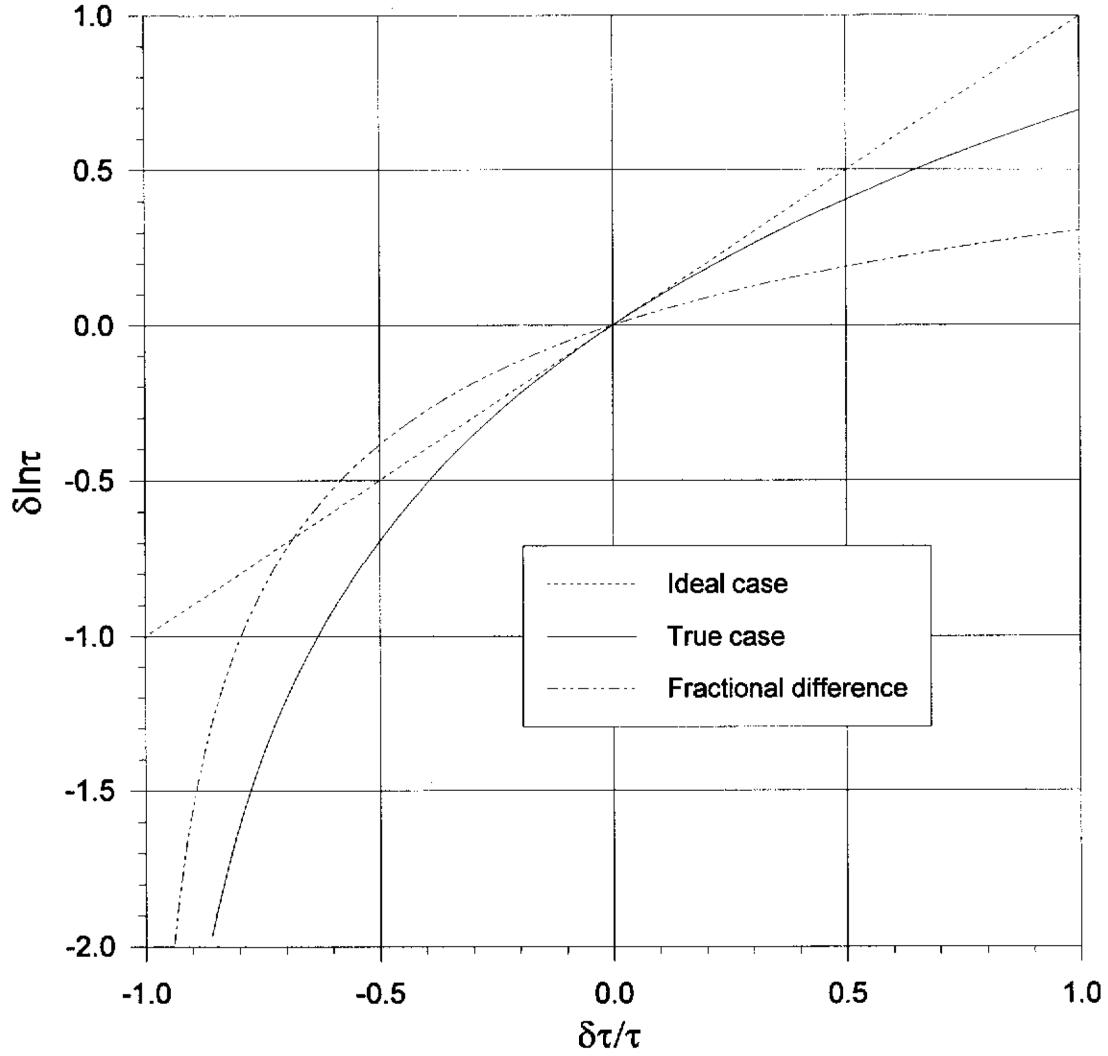


Figure 3.17 Transmittance perturbation approximation used in linear model. (············) Ideal case where the approximation is valid for all $\delta\tau/\tau$, (————) True relationship between $\delta\ln\tau$ and $\delta\tau/\tau$, (-·-·-·-·-) Fractional difference between approximation and true relationship

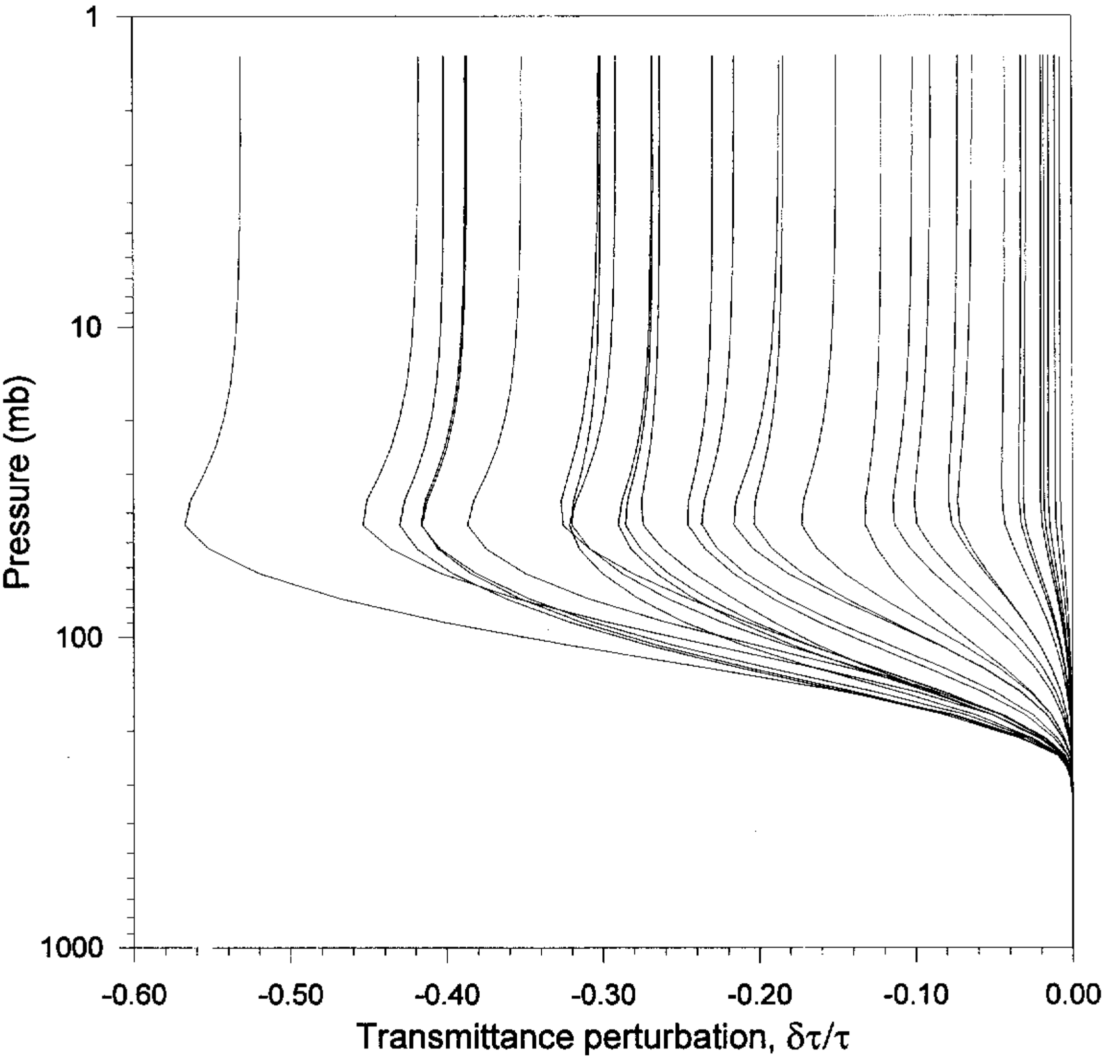


Figure 3.18 FASCOD3P transmittance perturbation for +200% Gaussian stratospheric ozone perturbation at the 50mb level to the U.S. Standard atmosphere from the surface to 1mb for selected spectral channels. Spectral region : 980-1080cm^{-1}; $\Delta\nu = 0.482$cm^{-1}

3.2.4 Ozone weighting functions

The discrete form of the ozone weighting functions of (3.16b) for a unit ozone mixing ratio perturbation (δq = 1g/kg) is given by

$$W_{Ql} = \frac{k_{\nu l}}{g}\left[\sum_{l'=L}^{l}\tau^{o}_{\nu l'}\Delta B_{\nu l'} + \tau^{o}_{\nu L}B_{\nu L}\right]\Delta p_{l}, \qquad (3.23)$$

where L = the layer where the top pressure corresponds to p_{top}. These functions will be referred to as the *absolute* weighting functions. They correspond to the modelled radiance perturbation *at the surface* due to ozone perturbations of δq_l=1g/kg for the partial atmospheric column pertaining to layer l of width Δp_l in pressure. The magnitudes of these functions at the surface are extremely large as the ozone concentrations are of the order 10^{-5}g/kg in the troposphere. When visually examining the form of the ozone weighting functions, it is more instructive to look at the radiance contribution at each pressure level p for a *relative* perturbation ($\delta q/q_o$) opposed to an absolute perturbation (δq).

The optical depth is the exponent of (2.2); the initial estimate optical depth thus being

$$\sigma^{o}_{\nu}(p \rightarrow p_s) = \frac{1}{g}\int_{p}^{p_s}\sum_{i=1}^{N}k_{\nu i}(p')q^{o}_{i}(p')dp'. \qquad (3.24)$$

For a discrete pressure interval, Δp_l, the ozone optical depth, *for that layer only*, is given by

$$\sigma^{o}_{\nu il} = \frac{1}{g}k_{\nu il}q^{o}_{il}\Delta p_{l}, \qquad (3.25)$$

where the subscript *i* is held over to distinguish between ozone only quantities (e.g. the absorption coefficient, *k*) and all-molecule quantities (e.g. the total transmittance, τ) in later equations. Using (3.25) with (3.23), the atmospheric linear *relative* weighting functions, I_{Ql}, in any layer *l* for a unit ($\delta q/q_o$ = 1, or 100%) increase in ozone is given by

$$I_{Ql} = \sigma^o_{\nu il}\left[\sum_{l'=L}^{l} \tau^o_{\nu l'}\Delta B_{\nu l'} + \tau^o_{\nu L} B_{\nu L}\right]. \tag{3.26}$$

The quantity (3.26) differs from the absolute weighting function definition of (3.23) by the factor q^o_{il}. In both cases, the response is sensitive to the pressure widths of the layers used in the forward calculation; thicker layers contributing more radiance. Uplooking relative weighting functions were calculated for the STD atmosphere using two layering schemes - a 28 layer atmosphere (see table 3.2; with an extra boundary at 1mb) and a 50 layer atmosphere. Pressure level boundaries for the 50 layer atmosphere are shown in table 3.3, with the first four levels highlighted where the layer thickness changes significantly.

The relative ozone weighting functions for the 28 layer scheme are shown in figure 3.19, overlayed on the temperature and ozone concentration profiles used in the calculation. The basic shape of the weighting functions can be separated into three vertical regions; the troposphere from the surface to 200mb, the isothermal tropopause from 200-50mb, and the stratosphere from 50mb upward. These definitions derive from the shape of the temperature profile corresponding to those regions of the atmosphere where the lapse rate is positive, zero, and negative respectively.

Level	Pressure (mb)	Level	Pressure (mb)	Level	Pressure (mb)
1	1000	18	760	35	**175**
2	990	19	740	36	150
3	980	20	720	37	125
4	970	21	700	38	100
5	960	22	680	39	75
6	950	23	660	40	50
7	940	24	640	41	30
8	930	25	620	42	**25**
9	920	26	600	43	20
10	910	27	**550**	44	15
11	900	28	500	45	10
12	**880**	29	450	46	7
13	860	30	400	47	5
14	840	31	350	48	4
15	820	32	300	49	3
16	800	33	250	50	2
17	780	34	200	51	1

Table 3.3 Interpolation pressure levels for 50 layer atmosphere.

In the troposphere, the atmospheric layers are all 50mb thick and the weighting functions are smooth. The tropospheric emission up to about 350mb decreases relatively uniformly with the temperature lapse rate since the ozone concentration profile in the troposphere does not alter significantly (see figure 3.10b). At this level, the ozone concentration begins increasing noticeably and the subsequent increased emission forces the slope of the weighting functions to turn around. In the tropopause, the temperature lapse rate is zero and the weighting function increase is due to the rapid increase in the ozone concentration alone. The discontinuity in figure 3.19 at about 90mb is due to the layer thickness changing from 50mb to 25mb (refer to the pressures of levels 18, 19, and 20 in table 3.2). A rough estimate for comparison with the 50mb thick layers would be to double the weighting function value at this discontinuity. As the temperature begins to increase, the weighting functions begin to decrease - consistent with the sign of the dB_ν/dp term in (3.16b). The stratospheric region of the atmosphere has increasing temperature and increasing ozone concentration but a decreasing radiative signature at the surface. The emission of the higher warm layers is absorbed by the cooler layers below. This decrease is also compounded by another change in the layer pressure thickness from 25mb to 10mb at aproximately 20mb.

The effect of varying pressure layer thickness is very evident in the 50 layer atmosphere relative ozone weighting functions, shown in figure 3.20. The discontinuities in figure 3.20 correspond to those levels highlighted in table 3.3. The tropospheric effect is severe with layer thickness changes from 10mb to 20mb in the boundary layer. Note, however, that in the tropopause, the layer thicknesses are uniform from approximately 200mb to 60mb and the weighting functions are increasing uniformly in tandem with the ozone profile. The structure of the 50 layer weighting functions in the upper atmosphere - most notably the minimum at 4.5mb - also indicate that the calculated response at the surface is sensitive to layer thickness variations; even at this height.

To remove the layer thickness dependence in the weighting functions, the layer radiances can be divided by the logarithm of the the pressure thickness, *d*ln*p*. This allows comparisons of "equivalent" radiances between different layers and different layering schemes. Thus "equivalenced" weighting functions can be defined from (3.26),

$$\tilde{I}_{Ql} = \frac{\sigma^o_{\nu l}}{\Delta \ln p_l}\left[\sum_{l'=L}^{l} \tau^o_{\nu l'} \Delta B_{\nu l'} + \tau^o_{\nu L} B_{\nu L}\right]. \tag{3.27}$$

The term $\frac{\sigma^o_{\nu l}}{\Delta \ln p_l}$ is no longer a function of Δp_l since from (3.25),

$$\frac{\sigma^o_{\nu l}}{\Delta \ln p_l} = \frac{k_{\nu l} q^o_{il} p_l}{g}$$

The equivalenced ozone weighting functions corresponding to figure 3.19 are shown in figure 3.21 with a magnification of the same in figure 3.22. The same result is obtained for the 50 layer atmosphere. The influence of the temperature lapse rate and ozone concentration is very evident in figure 3.22 - the changes in the gradients of the weighting functions correspond to those levels where the temperature lapse rate changes.

The equivalenced weighting functions peak in both the troposphere and stratosphere. Figure 3.22 shows that channels with a high tropospheric response have a corresponding (comparative) high response in the region 20-60mb. As the overall weighting function response increases, the stratospheric peak becomes narrower and moves downward into the tropopause where it effectively becomes incorporated into the tropospheric response. This vertically downward movement of the weighting functions corresponds to moving across the ozone spectrum from the low magnitude spectral regions (980 - 1010cm^{-1}, fig. 3.7) on the edges (more sensitive to stratospheric emission) to the high magnitude regions (1030 - 1040, 1050 - 1060cm^{-1}, fig. 3.7) in the centre (more sensitive to tropospheric emission).

Both the relative and equivalenced weighting functions are characterised by their high tropospheric sensitivity and similar shape. Two consequences of the form of the weighting functions are that a ground-based retrieval system will have a great deal of sensitivity to tropospheric ozone concentration changes but will have difficulty separating out tropospheric and stratospheric radiance information. Even though the weighting functions do exhibit a structure which allows some degree of discrimination between different atmospheric levels, the implication is that the information content of the spectra at the specified resolution is low. A retrieval scheme will require *a priori* data input to constrain the inversion and resolve all of the retrieved atmospheric layers.

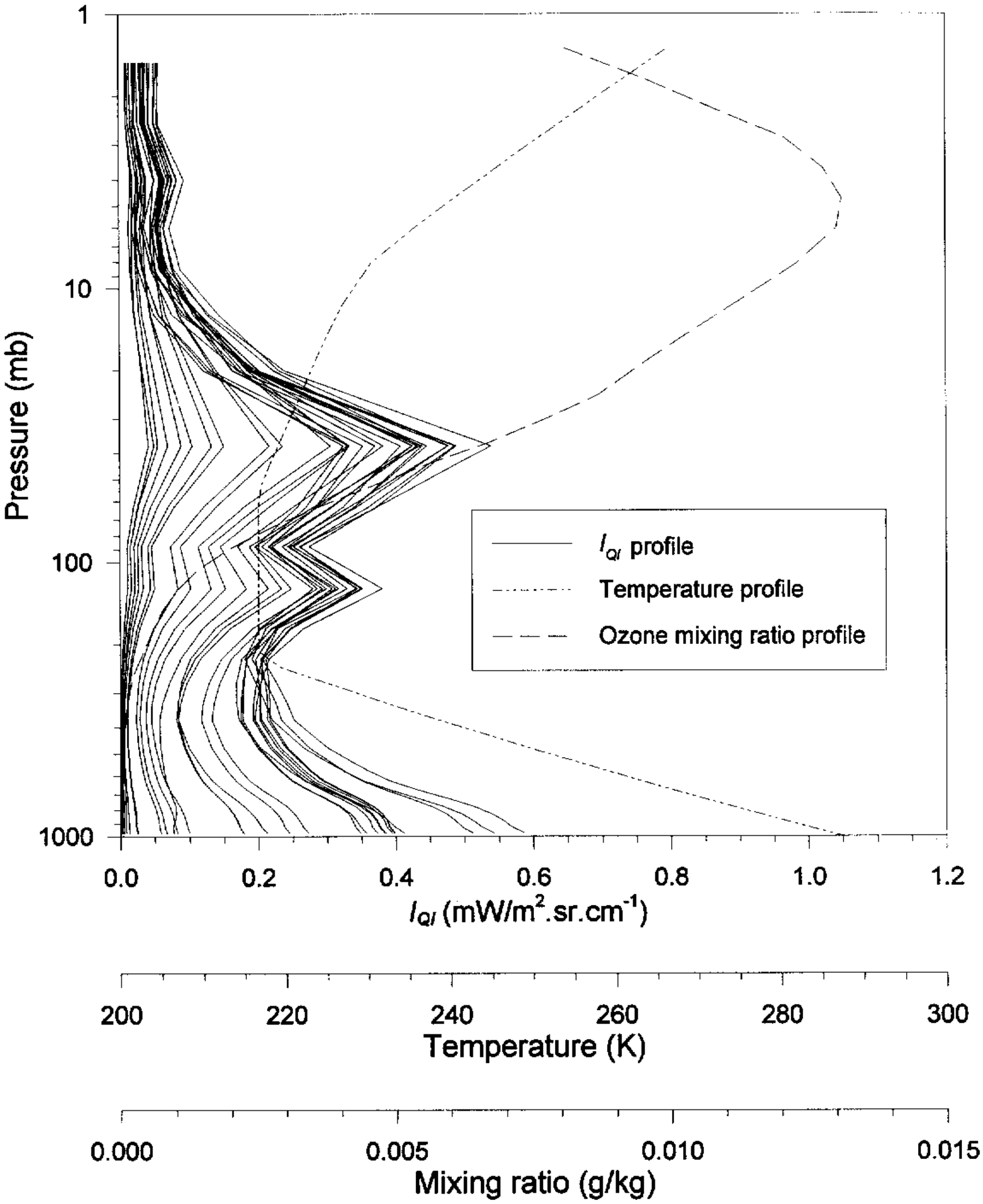


Figure 3.19 Relative ozone weighting functions, I_{Ql}, with temperature and mixing ratio profiles for a 28 layer U.S. Standard atmosphere (1976) to 1mb. 30 frequencies have been selected from the spectral region : 980-1080cm^{-1}; $\Delta\nu$ = 0.482cm^{-1}. Note the discontinuities due to layer pressure thickness changes.

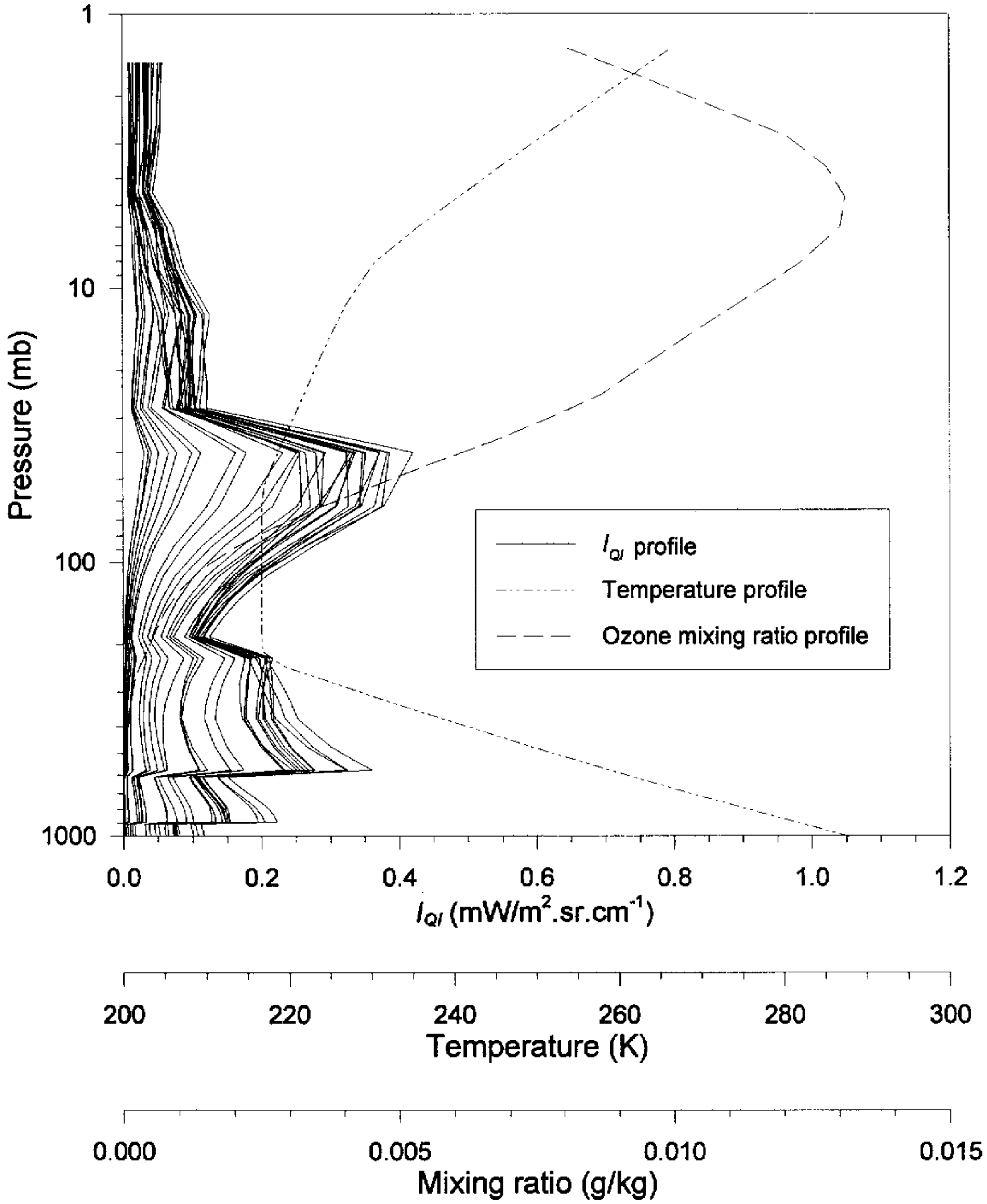


Figure 3.20 Relative ozone weighting functions, I_{Ql}, with temperature and mixing ratio profiles for a 50 layer U.S. Standard atmosphere (1976) to 1mb. 30 frequencies have been selected from the spectral region : 980-1080cm^{-1}; $\Delta\nu$ = 0.482cm^{-1}. The discontinuities due to layer pressure thickness differences are evident.

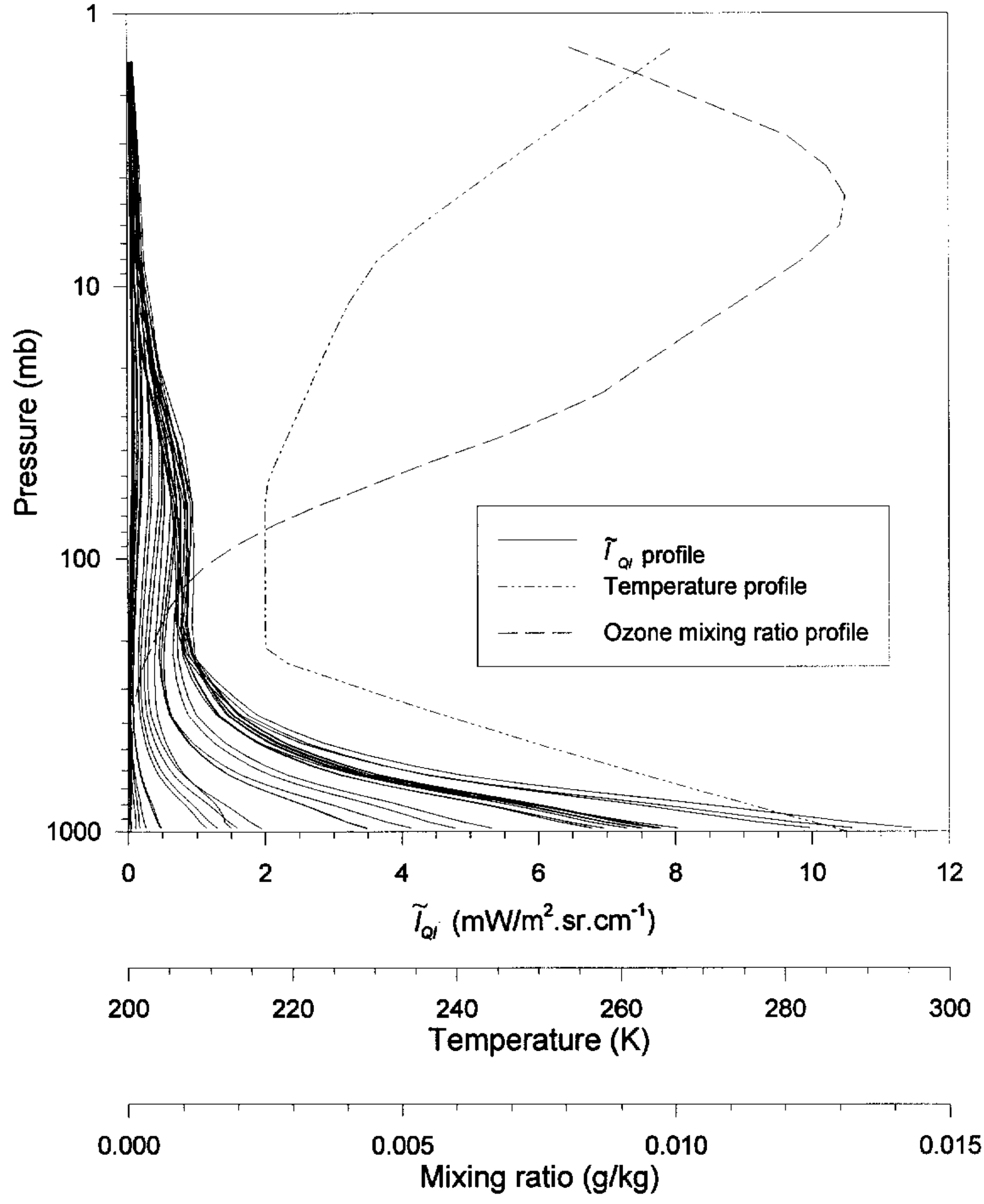


Figure 3.21 Equivalenced relative ozone weighting functions, $\tilde{I}_{Ql}$, with temperature and mixing ratio profiles for a 28 layer U.S. Standard atmosphere (1976) to 1mb. 30 frequencies have been selected from the spectral region : 980-1080cm^{-1}; $\Delta\nu = 0.482$cm^{-1}

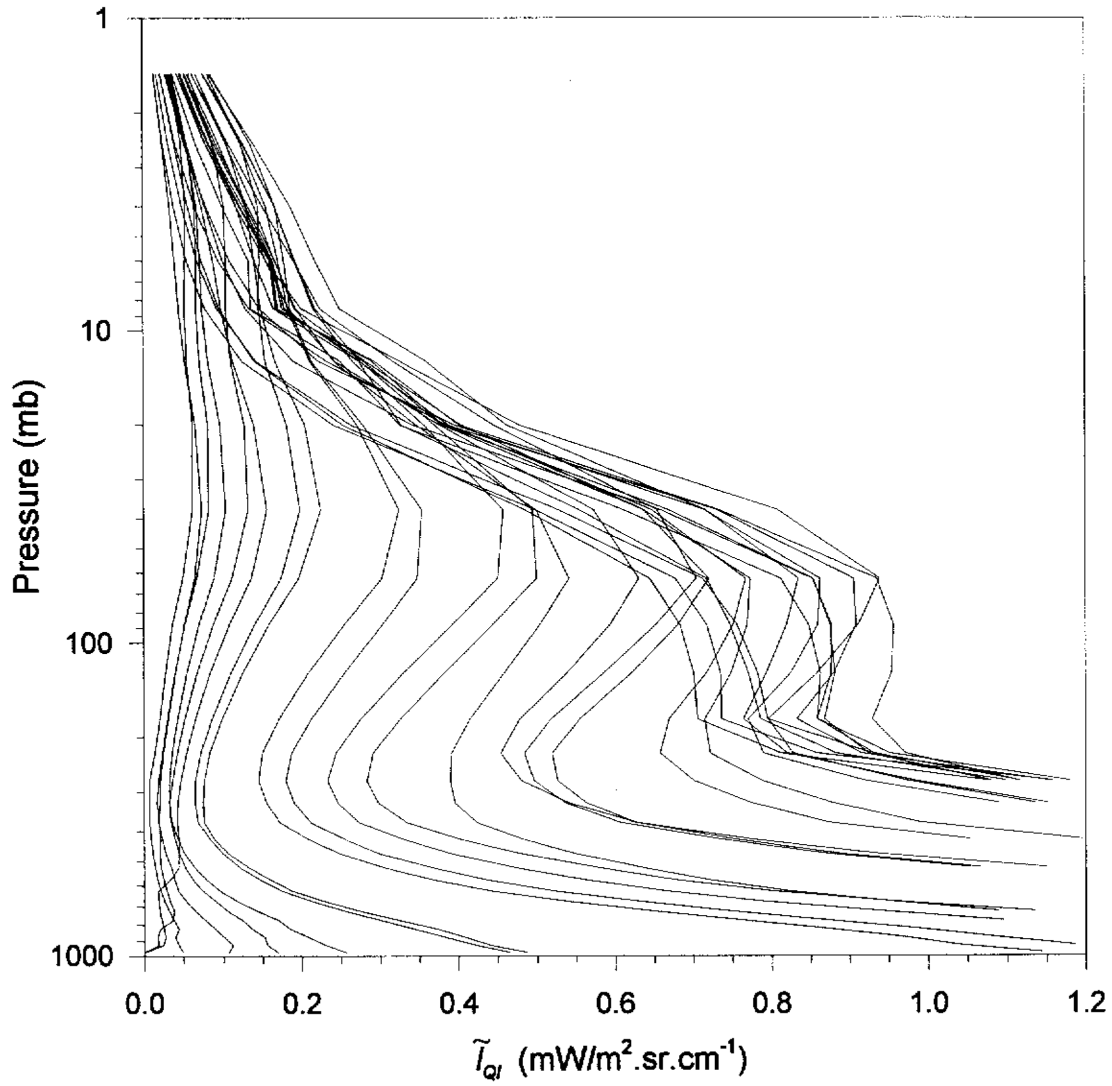


Figure 3.22 Magnification of equivalenced relative ozone weighting functions, $\tilde{I}_{Ql}$ (figure 3.21) for a 28 layer U.S. Standard atmosphere (1976) to 1mb. 30 frequencies have been selected from the spectral region : 980-1080cm^{-1}; $\Delta\nu = 0.482$cm^{-1}. Note the correspondence of the change in the weighting function gradient to the change in the temperature profile of figure 3.21.

4.0 Inverse Problem

The method used to retrieve ozone concentration profiles from synthetic FASCODE radiances was an iterative maximum likelihood inversion scheme described by,

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \left(\mathbf{q}^o - \mathbf{q}_n\right) + \mathbf{G}_n\left[\mathbf{W}_n\left(\mathbf{q}_n - \mathbf{q}^o\right) + \left(\mathbf{r}_{obs} - \mathbf{r}(\mathbf{q}_n)\right)\right], \tag{4.1}$$

where $\mathbf{G}_n = \left(\mathbf{C}_Q^{-1} + \mathbf{W}_n^T \mathbf{C}_R^{-1} \mathbf{W}_n\right)^{-1} \mathbf{W}_n^T \mathbf{C}_R^{-1}$,

$\mathbf{W}_n$ = ozone weighting functions,

$\mathbf{q}$ = ozone mixing ratio (g/kg) profile vectors,

$\mathbf{C}_Q$, $\mathbf{C}_R$ = ozone mixing ratio and radiance error covariance matrices;

the derivation of which is described in section 2.1.2. The terminology adopted to describe the different profile vectors, $\mathbf{q}$, is as follows : the initial profile constraint, $\mathbf{q}^o$, which does not change through the iterative procedure, is termed the *initial estimate* or *first guess*; the previous iteration solution, $\mathbf{q}_n$, is termed the *iteration guess*; and the profile vector, $\mathbf{q}_{n+1}$, is referred to as the *current iteration solution.* Note that, on the first iteration, the initial estimate is also the iteration guess.

As mentioned previously, unconstrained physical inversion schemes are highly unstable due to both the size of the elements in the inversion matrices and the noise in measured radiances, allowing an infinite set of possible solutions. This, coupled with the indication of low information content in the uplooking ozone weighting functions discussed in the previous chapter, requires that the solution scheme have available to it *a priori* information, in this case a database of ozonesonde profiles, and also a rigorous technique of optimisation. Both of these requirements are implicit in the maximum likelihood method.

The form of $\mathbf{G}_n$ in (4.1) is determined by comparison of the number of wavenumber channels used to the number of atmospheric layers to be retrieved. For the spectral resolution used, 0.482cm^{-1}, the number of channels in the interval 980-1080cm^{-1} is 209 while the number of retrieval layers is 27 (see table 3.2). Thus the expression used in (4.1) is computationally more efficient than the alternate form which is shown in Appendix A.

All retrieval calculations were carried out using double precision arithmetic (on a DEC 32 bit VAX processor) to avoid the introduction of significant rounding errors into the solution, particularly where the ozone concentration is low, i.e. the troposphere. Typical ozone mixing ratios in the troposphere are 10^{-4}-10^{-5}g/kg.

The computational algorithm used to invert matrices in the retrieval procedure was the Singular Value Decomposition (SVD) method taken from Forsythe et al. (1977; p228-235). An $M \times N$ matrix $\mathbf{A}$ is factorised into the form,

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T, \tag{4.2}$$

where $\mathbf{U}$ = left singular vectors, $M \times M$ orthogonal matrix,

$\mathbf{V}$ = right singular vectors, $N \times N$ orthogonal matrix, and

$\Sigma = M \times N$ diagonal matrix of singular values.

The pseudoinverse of $\mathbf{A}$ can then be found from

$$\mathbf{A}^+ = \mathbf{V}\Sigma^+\mathbf{U}^T. \tag{4.3}$$

If **A** is square and non-singular, then $\mathbf{A}^+ = \mathbf{A}^{-1}$. In all cases where matrix inversions were performed, the product $\mathbf{AA}^{-1}$ was also calculated to ensure the product was the identity matrix.

4.1 Data covariance matrices

4.1.1 Radiance error covariance

The construction of the radiance error covariance matrix was based upon the assumption that errors are uncorrelated and equal across channels. The nominal noise level of the AERI instrument, $\varepsilon = 0.1\text{mW/m}^2.\text{sr.cm}^{-1}$, was used to calculate a simple diagonal radiance error covariance matrix

$$\left(C_R\right)_{\mathrm{uv}} = \left(\varepsilon_{\mathrm{u}}\varepsilon_{\mathrm{v}}\right)_{\mathrm{u,v}=1,\dots,N} \text{ , where } \varepsilon_{\mathrm{u}}\varepsilon_{\mathrm{v}} = \begin{cases} 0.01 & \mathrm{u}=\mathrm{v} \\ 0 & \mathrm{u}\neq\mathrm{v} \end{cases}\text{, and} \qquad (4.4)$$

where N = total number of channels, and

u,v = channel indices.

The simple diagonal form of the matrix means it is easily inverted for use in equation (4.1).

4.1.2 Mixing ratio covariance

The mixing ratio covariance was calculated from the entire sonde database (F2S + DB sondes) minus the F2S sonde selected for testing the retrieval scheme. The basis for selection of the retrieval sonde is detailed in the next section. All of the sondes were interpolated to the levels given in table 3.2. The mixing ratio covariance matrix was calculated using,

$$\left(C_Q\right)_{ij} = \frac{1}{N}\sum_{k=1}^{N}\left[\left(q_k - \overline{q}\right)_i\left(q_k - \overline{q}\right)_j\right], \tag{4.5}$$

where N = total number of sondes used, and

i, j = atmospheric layer indices.

As was described in section 2.1.2, the maximum likelihood inversion scheme, as derived, uses minimum variance as an estimator in the least squares minimisation of the retrieval cost function, equation (2.21). This is valid only for Gaussian or near-Gaussian statistics. It is necessary to verify that the variation of the mixing ratio satisfies this condition. This was achieved qualitatively by examining the average difference, $\widetilde{q}$, of the direct sonde set distribution, and the skewness, ζ, or third moment, of the distribution $\left(q - \overline{q}\right)$ at every level. The skewness was then compared with the same for a set of normally distributed pseudo-random numbers with the same standard deviations as the distribution $\left(q - \overline{q}\right)$. The quantities $\widetilde{q}$ and ζ are defined as,

$$\widetilde{q}_l = \frac{1}{N}\sum_{k=1}^{N}\left(q_{kl} - \overline{q}_l\right), \tag{4.6a}$$

$$\zeta = \frac{1}{N}\sum_{k=1}^{N}\left(\frac{q_{kl} - \overline{q}_l - \widetilde{q}_l}{\sigma_l}\right)^3, \tag{4.6b}$$

where N = the total number of sondes used,

l = atmospheric layer index,

q_{kl} = ozone mixing ratio for sonde k at level l,

$\overline{q}_l$ = average ozone mixing ratio at level l, and

$$\sigma_l = \sqrt{\frac{1}{N-1}\left[\sum_{k=1}^{N}\left(q_{kl}-\overline{q}_l-\tilde{q}_l\right)^2-\frac{1}{N}\left[\sum_{k=1}^{N}\left(q_{kl}-\overline{q}_l-\tilde{q}_l\right)\right]^2\right]},$$

standard deviation at level l.

The averaged difference for the sonde set was zero to at least eight decimal places for every atmospheric level. For an ideal normal distribution ζ would also be zero.

In figure 4.1 the skewness of the random number data varies about zero indicating a symmetrical distribution. For the ozonesonde data we see that near the surface and from 500-100mb the ozone distribution has a positive skew, that is, it has a longer tail on the positive side of the zero mean. This is not unexpected as these are the two regions in the atmosphere where the ozone profile is changing most rapidly. Near the surface, ozone concentrations vary rapidly due to pollutants/photochemical smog, and the upper troposphere/lower stratosphere is where ozone concentrations rapidly increase to their stratospheric maximum. For the skew profile shown in figure 4.1 to indicate a significant deviation of the ozonesonde data set from a Gaussian distribution, it should be many times larger than the standard deviation of (4.6b) itself, which for a Gaussian distribution can be approximated by $\sqrt{15/N}$ (Press *et al.*, 1992). For the ozonesonde set (N = 83), the estimated skew standard deviation is 0.43. This compares with the maximum skew value, at 300mb, of 1.5. Based on the profile in figure 4.1, it was felt that the degree of deviation of the ozonesonde data set from a normal distribution was small enough to satisfy the condition of near-Gaussian statistics.

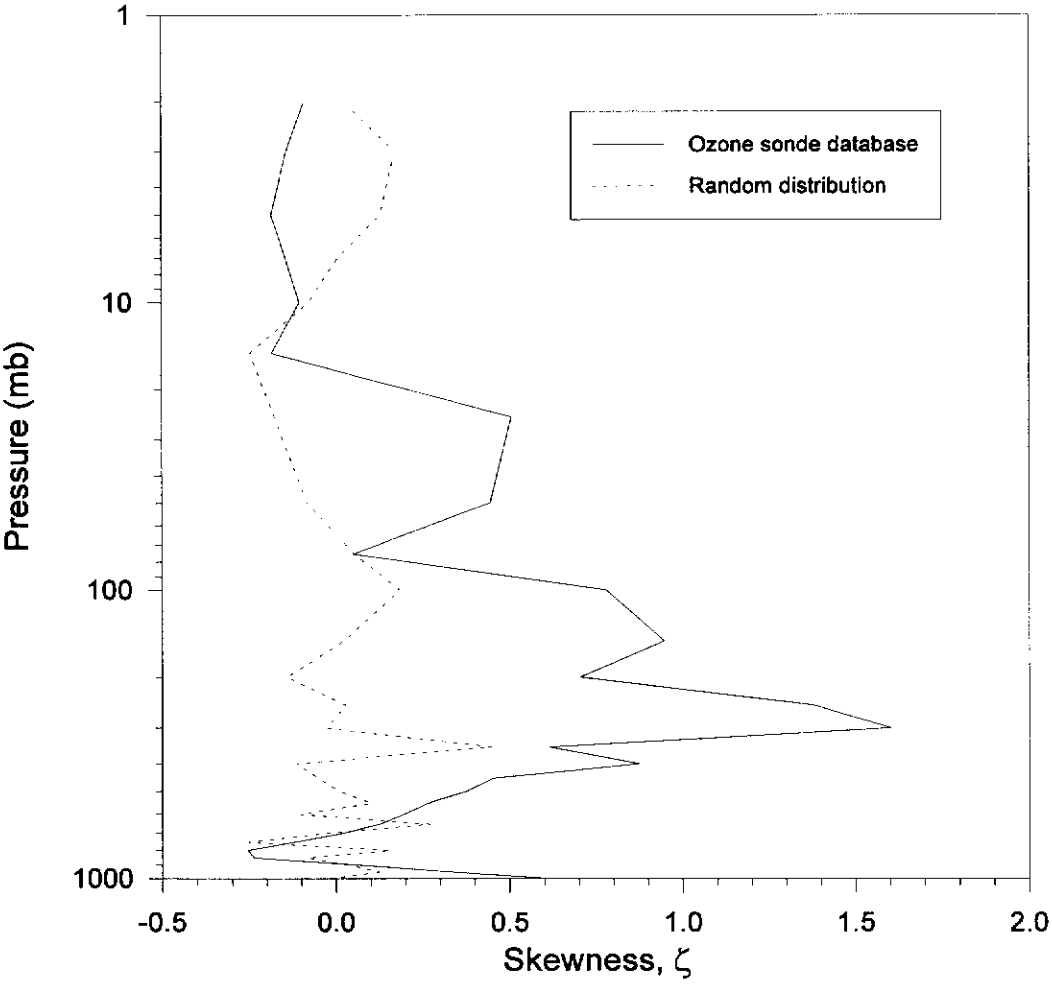


Figure 4.1 Skewness of the distribution of ozone sonde profiles, $(q-\overline{q})$. (————) Ozonesonde data, (··········) Normally distributed random number data with same standard deviation as ozonesonde data.

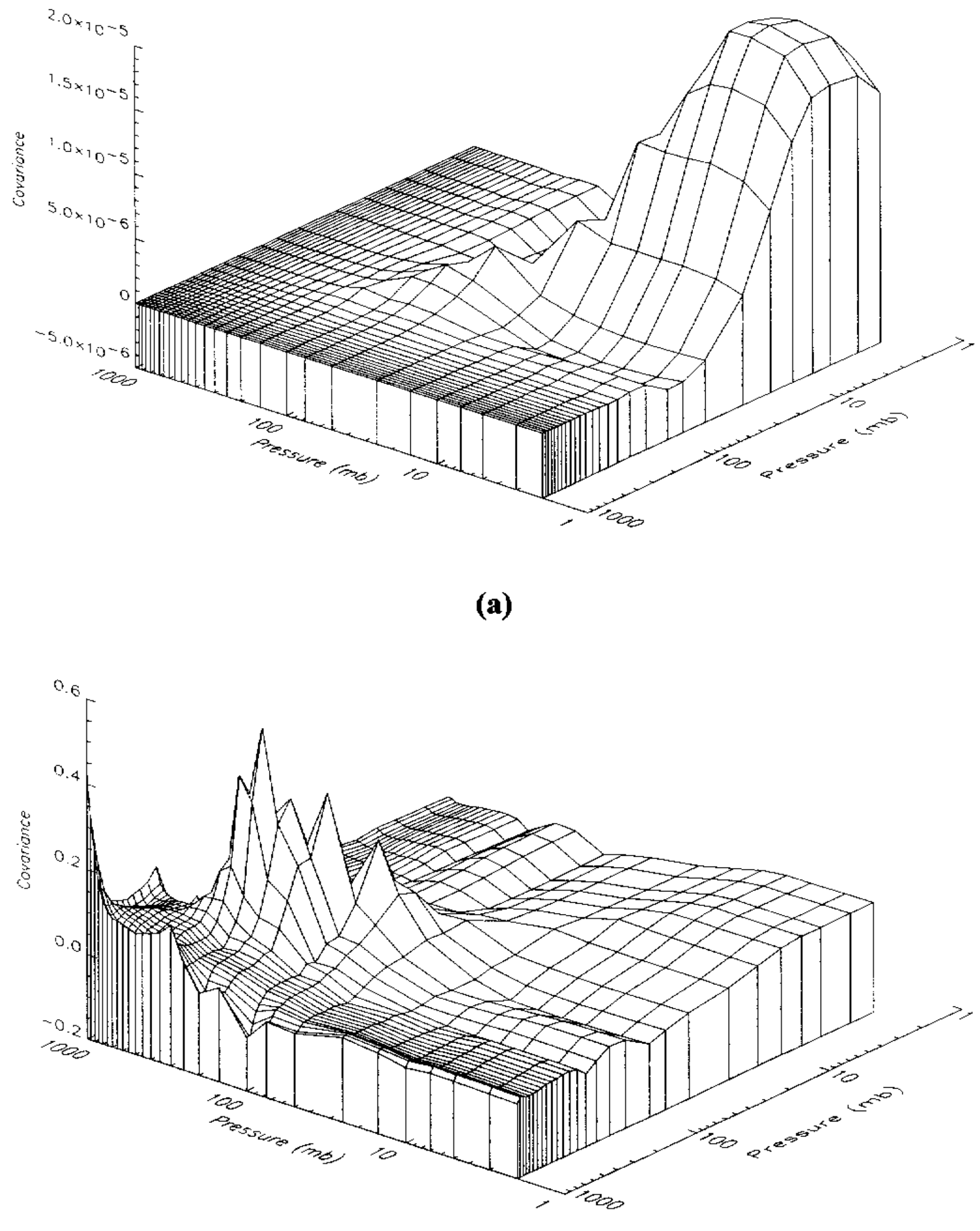
The ozone mixing ratio *perturbation* covariance matrix was also calculated using,

$$\left(\widetilde{C}_Q\right)_{ij} = \frac{1}{N}\sum_{k=1}^{N}\left[\left(\frac{q_k - \overline{q}}{\overline{q}}\right)_i \left(\frac{q_k - \overline{q}}{\overline{q}}\right)_j\right]. \tag{4.7}$$

The surfaces of (4.5) and (4.7) are shown in figure 4.2(a) and (b) respectively. The perturbation form of the covariance matrix provides more information about the relative variation of ozone concentration throughout the atmosphere. Four "regions of variability" can be identified from figure 4.2(b). The regions of high relative ozone concentration variability occur in the boundary layer and lower troposphere; namely from the surface to 700mb, and in the upper troposphere/lower stratosphere from 375mb to 60mb. The other two regions, the middle and upper troposphere and the upper stratosphere exhibit more stable ozone concentrations, especially in the upper stratosphere. The greatest relative variation occurs in that region of the atmosphere where the ozone mixing ratio is increasing rapidly - note the correspondence between the peak of figure 4.2(b) and the variation of the ozone transmittances (i.e. $d\tau/dp$) in figure 3.9 from 30-300mb.

A correlation matrix was constructed from these covariance data to explicitly determine over what vertical scales the ozone mixing ratios are correlated. The term explicit is used because an idea of the correlation of ozone amount in the vertical is self-evident from figure 4.2(a). Regions of higher variability will have lower vertical correlation and vice versa. This is confirmed in the correlation matrix shown as a contour map in figure 4.3.

The correlations between atmospheric layers is reflected in the structure of the perturbation covariance matrix. In the regions of low ozone variability, the concentrations are correlated over broad vertical scales as these are of the same extent as the region itself. These regions exhibit a bulge in the correlation contours about the diagonal in figure 4.3. Conversely, the levels of ozone concentration outside these regions - the boundary layer region and the lower stratosphere - are correlated significantly over only short vertical scales as they are subject to the greatest relative variation.



(a)

(b)

Figure 4.2 **(a)** Surface of the ozone mixing ratio covariance matrix, $\mathbf{C}_Q$. **(b)** Surface of the ozone mixing ratio perturbation covariance matrix, $\tilde{\mathbf{C}}_Q$.

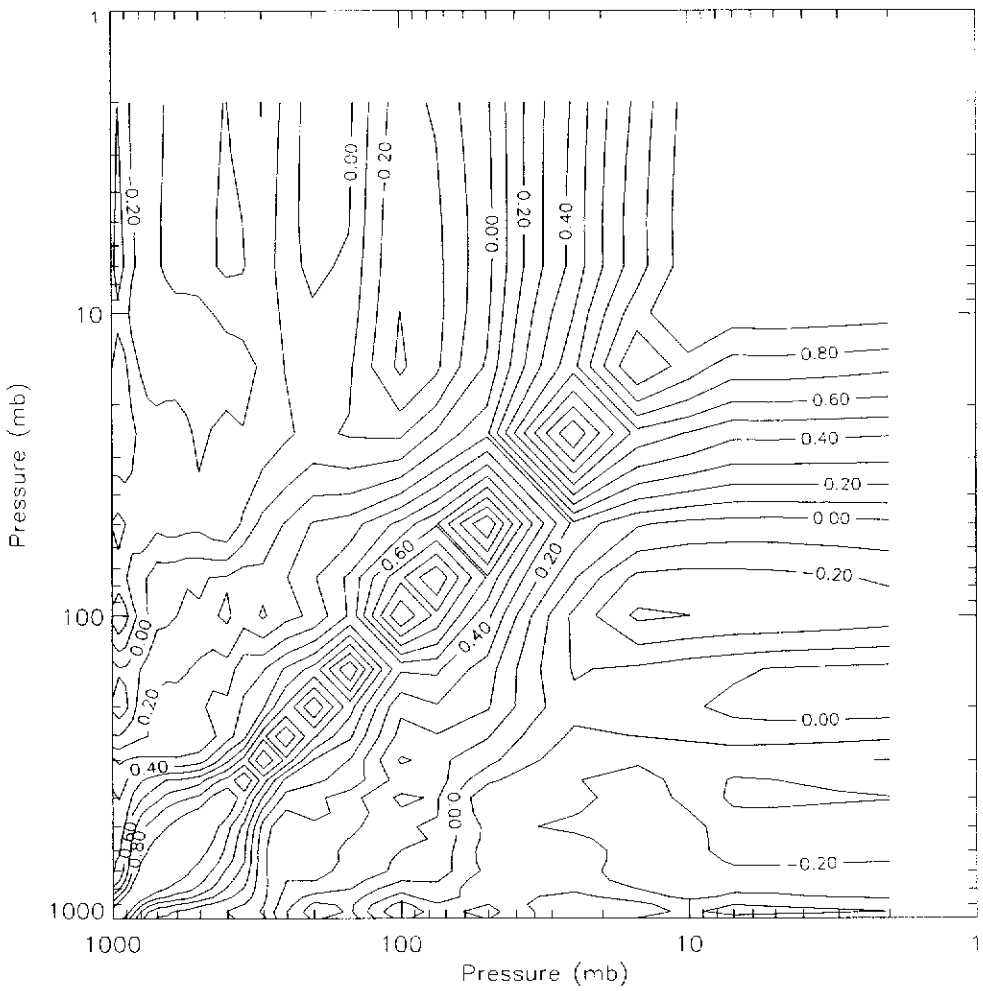


Figure 4.3 Ozone mixing ratio correlation matrix contour map. The diagonal values are 1.0 with the square features due to the discrete nature of the data.

4.2 Initial estimate selection

The ozone sonde profile selected for evaluating the retrieval scheme, the F2S NOAA sonde released on 17th November, 1991 at 2251UTC (F2S 17NV2251, shown in figure 3.4), was initially chosen because of the temporal proximity of the F2S NASA Wallops sonde released at 1757UTC the same day. It was mistakenly thought that the 1757UTC sonde would provide a suitable initial estimate for retrieval of the 2251UTC sonde. The procedure subsequently employed in the selection of the initial estimate for initiating the retrieval process was that detailed by Chédin and Scott (1983); minimisation of the distance between the coordinates of radiance spectra in radiance space described in section 2.4.1.

The radiance spectrum for every sonde profile in the database was calculated, using the F2S 17NV2251 sonde temperature and water vapour profile (assumed known quantities) to permit construction of a normalised radiance covariance matrix, $\tilde{\mathbf{R}}$, from which a radiance correlation matrix, **P**, was calculated for the 980-1080cm^{-1} ozone spectral region. The F2S 17NV2251 sonde was not included in the covariance calculation. The radiance correlation matrix and contour map are displayed in figures 4.4(a) and (b) respectively, and show the high correlation between channel radiances.

The high correlation between radiances from channels sensitive to different regions of the atmosphere is another indication of the sparsity of independent information about the vertical distribution of ozone in the IR spectrum. As was expected, due to the structure of the weighting functions shown in the previous chapter, radiance from those channels sensitive to stratospheric ozone emission is highly correlated with the radiance from tropospheric channels. The lowest correlation shown in figure 4.4(a) between channels is ~0.70-0.75 occurring in the spectral regions 1050-1060cm^{-1} (more sensitive to troposphere), and 1075-1080cm^{-1} (more sensitive to stratosphere).

The correlation matrix was diagonalised as in (2.70)

$$\mathbf{P} = \mathbf{E}\Lambda\mathbf{E}^T,$$

to obtain the co-ordinates of each spectrum in radiance space, Y_{ik}, from (2.71),

$$Y_{ik} = \sum_{j=1}^{209} \tilde{R}_{ij} E_{jk} \quad i = 1,2,...,91; \; k = 1,2,...,l \quad , \tag{4.8}$$

where the number of channels is 209, the number of spectra used 91 (83 DB sondes and 8 F2S sondes), and l the number of eigenvectors retained.

The synthetic ("observed") spectrum, with which we will test the retrieval procedure, was constructed by adding a Gaussian noise spectrum with an rms value 0.1mW/m^2.sr.cm^{-1} to the calculated radiance using the F2S 17NV2251 ozonesonde. The eigenvectors obtained from the calculated radiances were used to determine the co-ordinates of the observed spectrum. Thus comparisons of proximity of radiance spectra included the effect of radiometric noise in the "observed" spectrum. The spectrum with added noise is shown in figure 4.5 with a second order polynomial fit included for later residual radiance comparison.

Diagonalisation of the radiance correlation matrix revealed that the first three correlation eigenvectors, which are shown in figure 4.6, accounted for 99.9% of the variance in the synthetic radiances. The explained variance and associated eigenvalues for the first three eigenvectors are shown in table 4.1. The size of the first eigenvalue, the variance it explains, and the fact that only three eigenvectors out of 209 were required, are consistent with the low independent information content of the spectra as indicated by the high values of correlation between channels shown in figures 4.4(a) and (b).

Examination of the structure of the correlation matrix and its eigenvectors also provides an understanding of the distribution of tropospheric and stratospheric information in the ozone radiance spectrum. The correlation between tropospheric and stratospheric ozone sounding channels would be expected to be low due to the different mechanisms controlling ozone emission in each region, as described previously. Examination of figure 4.4 shows that there are lower correlations between sounding channels in the spectral regions 980-1000cm^{-1} and 1065-1080cm^{-1} and those from 1000-1040cm^{-1} and 1045-1065cm^{-1}. The former channels are more sensitive to the stratosphere and the latter the troposphere. Note that even though the correlations between these spectral regions are *relatively* low, they are still significant at ≈0.7-0.9 as opposed to 0.9-1.0 so there can be no definitive isolation of particular sounding channels sensitive to *only* a particular region of the atmosphere. This is expected considering the shape of the weighting functions shown in figure 3.21.

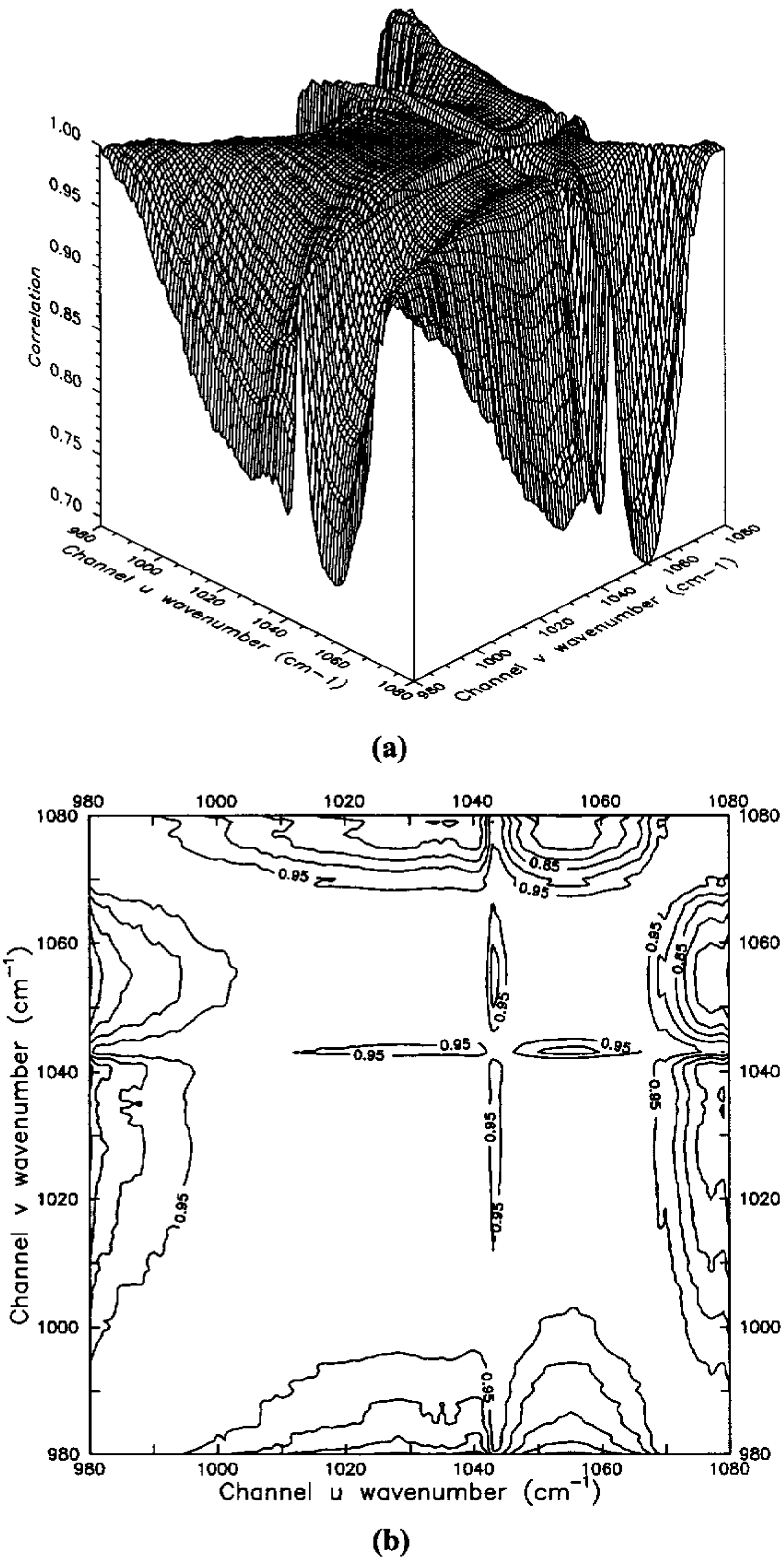


Figure 4.4 Radiance correlation matrix for radiances calculated from sonde set. **(a)** Correlation matrix surface. **(b)** Contour map.

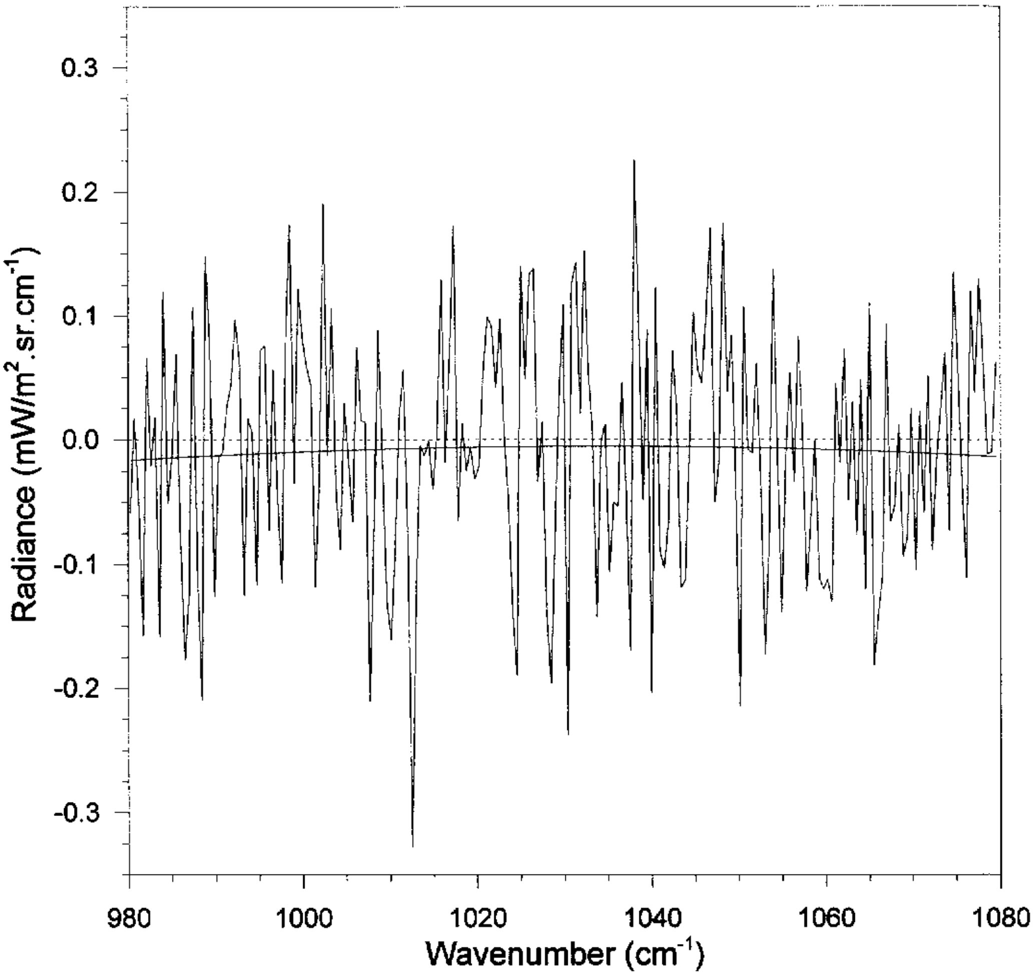


Figure 4.5 Gaussian noise spectrum, rms 0.1mW/m^2.sr.cm^{-1}, added to F2S 17NV2251 calculated radiance spectrum. A second order polynomial fit through the spectrum is included to show the mean radiance level for the noise distribution in this spectral region.

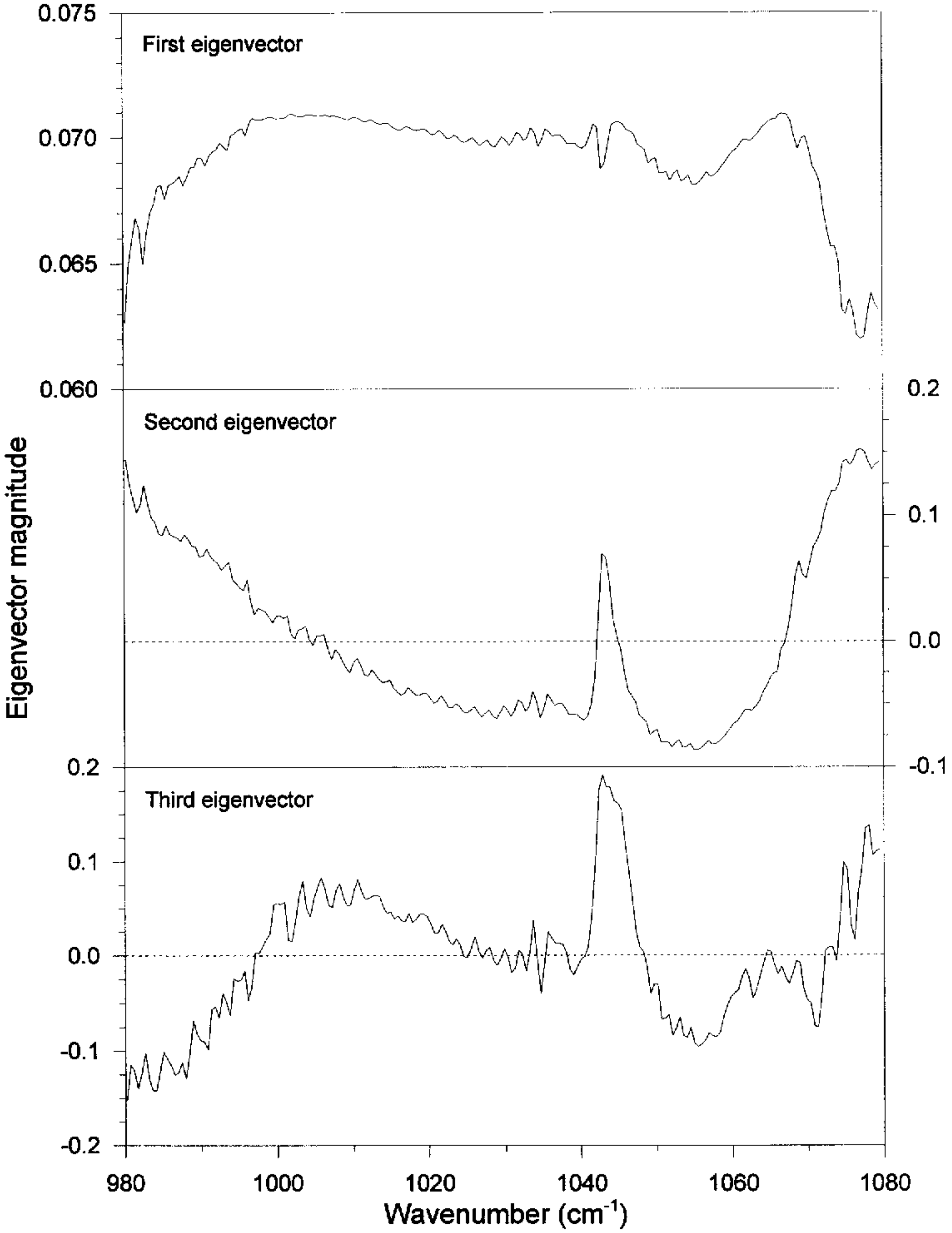


Figure 4.6 First three eigenvectors of the correlation matrix of synthetic radiances calculated from the ozone sonde dataset.

Eigenvector	Eigenvalue	Explained variance (%)
1	198.488	94.97
2	9.987	4.78
3	0.376	0.18
Total	**208.853**	**99.93**

Table 4.1 Eigenvalues and explained variance for first three eigenvectors of synthetic radiance correlation matrix

Despite the high correlations among the spectral channels, comparison of the radiance differences arising from tropospheric and stratospheric ozone perturbations (see figures 3.11 and 3.12) indicate that certain channels are more sensitive to ozone emission from different regions of the atmosphere. A qualititative division of channels in the ozone 980-1080cm^{-1} spectral region into those mostly sensitive to tropospheric emission and those mostly sensitive to stratospheric emission is possible by inspection of the first two correlation eigenvectors (figure 4.6). The channels near the peak values of the first eigenvector can be thought as spectral regions where the relative tropospheric and stratospheric sensitivities are similar. These regions correspond to where the second eigenvector is close to or equal to zero. Thus the zero line of the second correlation eigenvector can be thought of as a separator such that channels whose second eigenvector value is greater than zero have greater stratospheric sensitivity and those with values less than zero have greater tropospheric sensitivity.

A number of equivalenced weighting functions (from figure 3.21) taken from the spectral regions corresponding to positive (983.580, 1078.081cm^{-1}), zero (1003.831, 1044.331, 1067.956cm^{-1}) and negative (1030.881, 1054.456cm^{-1}) second eigenvector values are shown in figure 4.7(a), (b), and (c) respectively. The interdependence between the troposphere and stratosphere is still evident as all the weighting functions are double peaked. However there are relative differences between the magnitudes of the tropospheric and stratospheric responses for the three spectral regions.

The first two co-ordinates of each sonde radiance spectrum were calculated using (4.8) and the distances between all pairs of points determined. The vector scatter plot of all the radiance spectra is shown in figure 4.8. The synthetic ("observed") radiance spectrum is well separated from the other F2S sonde radiances, emphasising how poor the criterion of temporal proximity can be in initial estimate selection. The closest radiance coordinate is that belonging to the spectrum calculated from the sonde DB 7735511C. This ozone sonde profile was selected as the best initial estimate for the retrieval of the ozone profile from the calculated radiance ("observed") spectrum.

The ozone profiles of the F2S 17NV2251 and DB 7735511C sondes are shown in figure 4.9 with the associated perturbation. Even with the best radiance estimate, the ozone mixing ratio perturbations still peak at approximately 50% in both the troposphere and the stratosphere. It is evident that the large differences from 30-300mb are due to a lower tropopause level for the DB 7735511C sonde compared to the "true" profile so matching the F2S 17NV2251 temperature profile to the DB 7735511C ozone profile introduces a meteorological inconsistency. This inconsistency does not disqualify DB 7735511C as an initial estimate as the radiative response due to the negative ozone perturbation from 30-300mb is, to a degree, offset by the radiative response due to the positive ozone perturbation from 300mb to the surface because of the high level of interdependence across channels. Despite this, the *shape* of the tropospheric profiles are comparatively well matched from 400mb to the surface, even though there is an offset. The increased ozone concentration gradient from 800mb to the surface is well represented in the DB 7735511C ozone profile (figure 4.9(b)). This suggests that the radiance classification method used is more sensitive to the basic structure of the tropospheric ozone profile than the entire column profile; that is, most of the radiometric information in a ground-based, uplooking ozone spectrum describes the tropospheric ozone distribution.

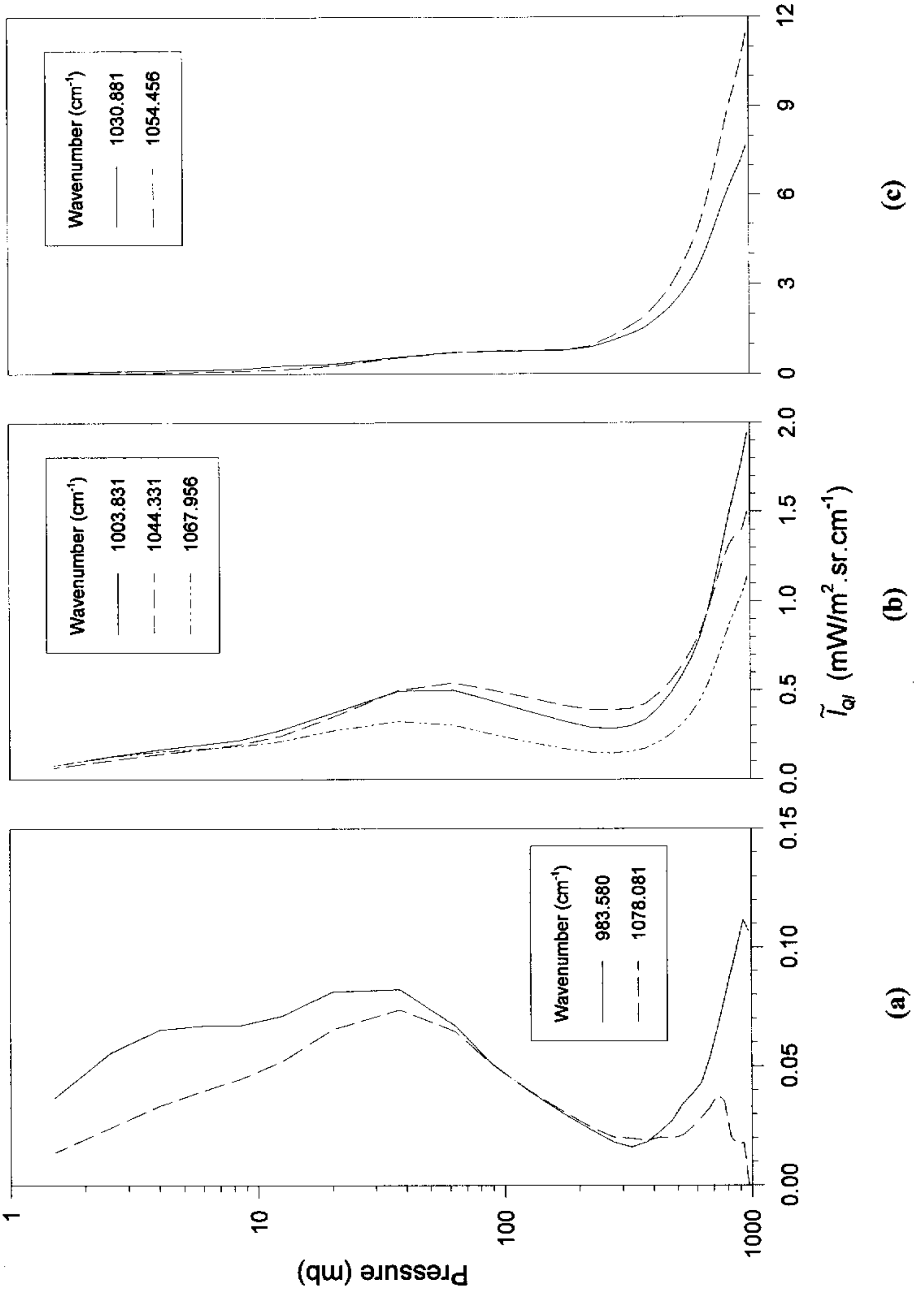


Figure 4.7 Selected equivalenced weighting functions from spectral regions with different vertical sensitivity. **(a)** Weighting functions primarily sensitive to stratospheric emission. **(b)** Weighting functions with effective equal sensitivity to both tropospheric and stratospheric emission. **(c)** Weighting functions primarily sensitive to tropospheric emission.

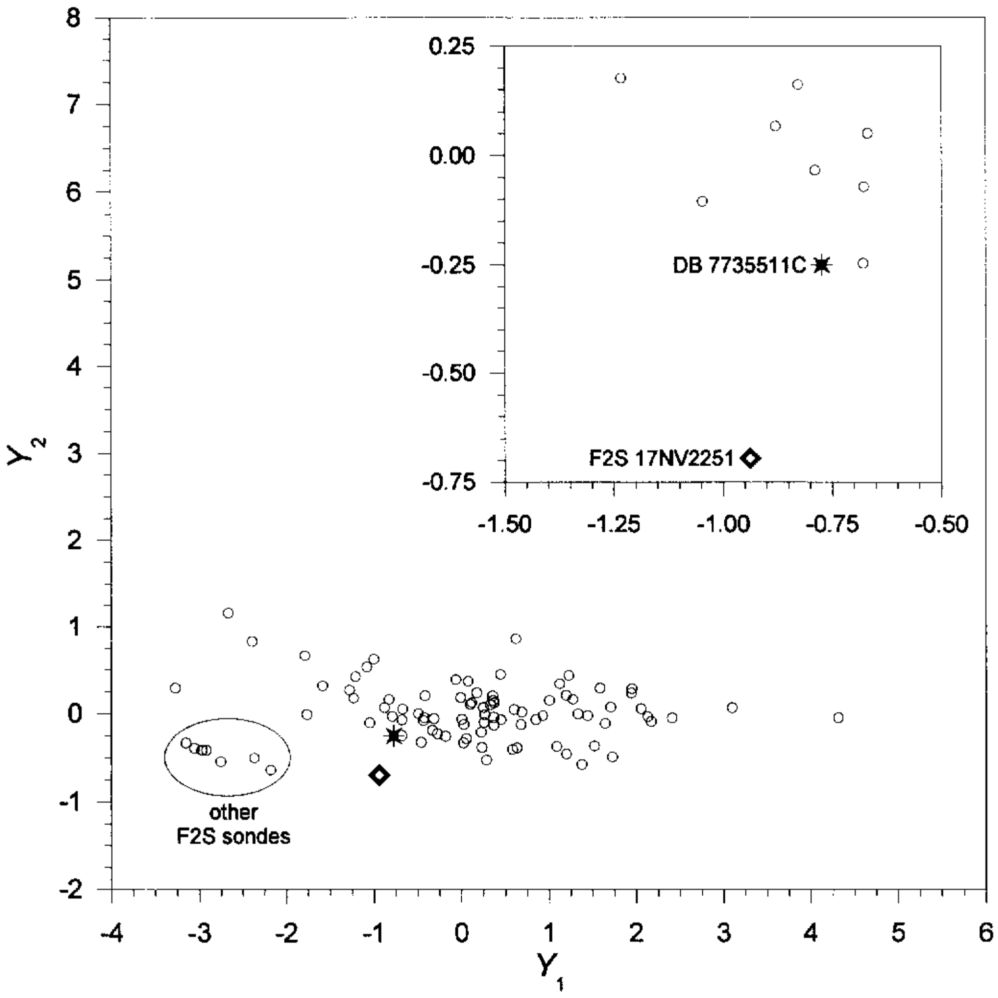
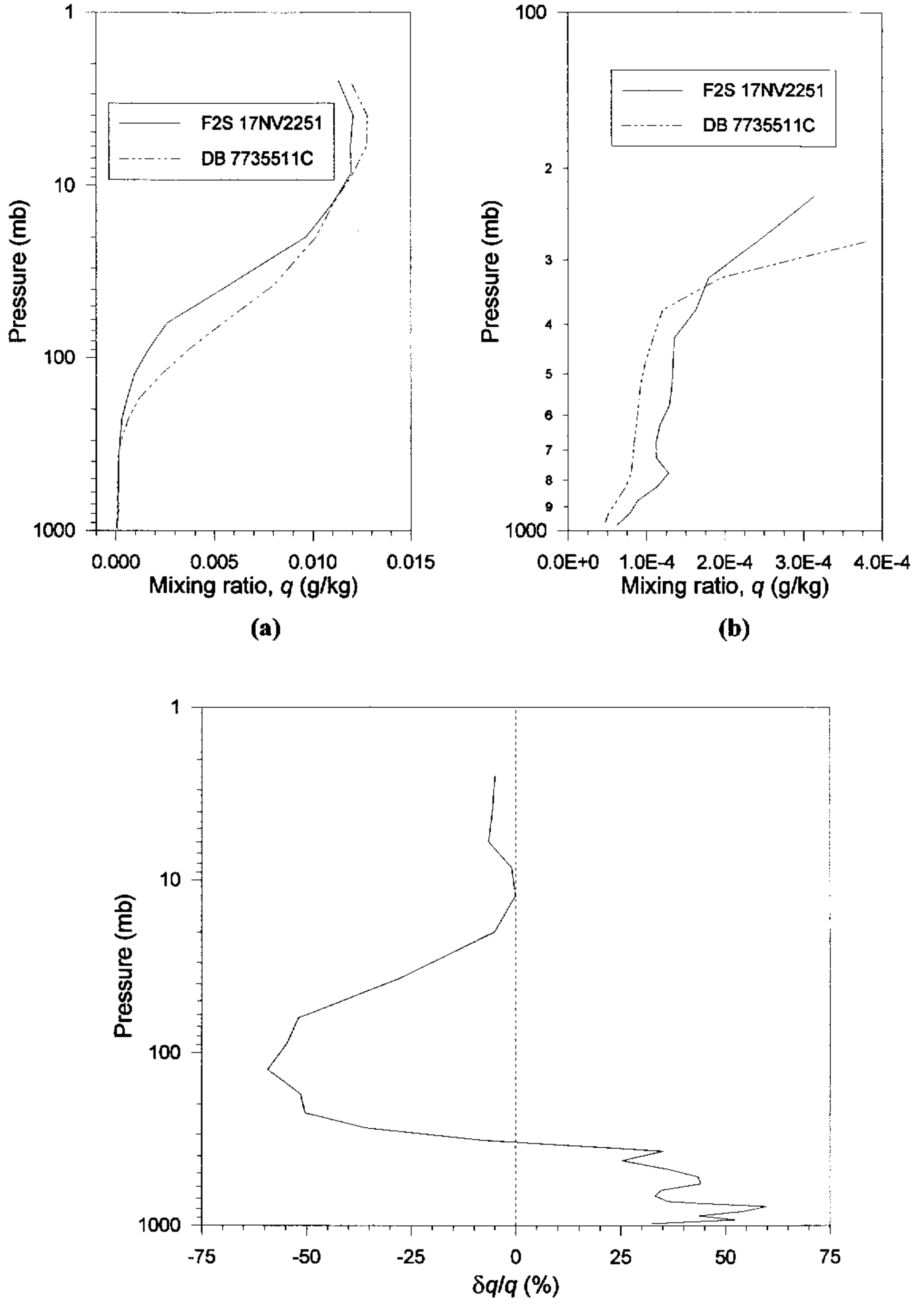


Figure 4.8 Vector scatter plot of the synthetic radiance spectra in the first two co-ordinates of radiance space (see eqn. 4.8) showing the synthetic observed spectrum ◇ (calculated from the F2S 17NV2251 sonde) and its closest neighbour in radiance space ✳ (calculated from the DB 7735511C sonde). The other 90 radiance spectra are represented by ○. The 8 F2S sondes are shown to cluster relatively tightly. (*Inset*) Magnification of the plot in the vicinity of the calculated ("observed") spectrum and the closest initial estimate.



(c)

Figure 4.9 (a) Ozone mixing ratio profiles, $q(p)$, for the "true" profile, F2S 17NV2251 sonde, q_{F2S}, and the best initial estimate, DB 7735511C sonde, q_{DB}. (b) Tropospheric magnification. (c) Mixing ratio perturbation, $(q_{\mathrm{F2S}}-q_{\mathrm{DB}})/q_{\mathrm{DB}}$

4.3 Ozone profile retrieval

4.3.1 Initial estimate weighting functions

The relative weighting functions calculated from the initial estimate profile are shown in 4.10. They differ from those of the U.S. Standard atmosphere shown in figure 3.21 due to the increase in the ozone concentration occurring at a lower pressure, the increased height of the isothermal layer, which is also thinner than for the U.S. Standard atmosphere, and a decrease in boundary layer ozone concentration.

The surface feature - a sharp decrease in the weighing functions - once again indicates the sensitivity of the radiative response to the variability of boundary layer ozone concentration. Small changes in already low ozone concentrations produce large radiative responses due to the tropospheric temperature. The temperature and ozone profiles of the F2S 17NV2251 sonde, before interpolation, are shown in figure 3.3. An upper level temperature inversion, also corresponding to an ozone concentration peak, is apparent at approximately 800mb. After interpolation to 27 layers the temperature inversion is effectively removed due to its small vertical scale, but the ozone profile, corresponding to this temperature inversion, still shows the decrease from 800mb to the surface. The initial estimate sonde, DB 7735511C, also has this similar ozone profile feature in the boundary layer. It is this change in the rate of decrease in ozone concentration which causes the weighting function decrease at the surface as seen in figure 4.10. This is shown clearly in figure 4.11(a) where the interpolated DB 7735511C ozone profile is shown with a representative weighting function. The ozone profile begins decreasing rapidly from 775mb to the surface and this decrease is reflected in the weighting function by a smaller radiative response. The U.S. Standard atmosphere ozone profile does not exhibit this boundary layer ozone profile feature; thus, the representative weight shown in figure 4.11(b) increases uniformly to the surface.

The peak in the initial estimate weighting functions at 300mb (fig. 4.10) is due to the meteorological inconsistency between the synthetic retrieved temperature profile (i.e. the F2S 17NV2251 temperature profile) and the initial estimate ozone profile (DB7735511C). The tropopausal region of the DB 735511C sonde corresponds to the upper tropospheric region of the F2S 17NV2251 sonde (fig. 4.9). The rapid ozone increase associated with the tropopause and lower stratosphere is situated in the upper troposphere where the radiative response is in balance between increasing ozone concentration and absorption. In this case, because of the profile inconsistency, the peak of ozone absorption is occurring lower in the atmosphere in the isothermal layer.

4.3.2 Ozone retrieval results

In retrieving the ozone profile from the synthetic ("observed") radiance spectrum calculated using the F2S 17NV2251 sonde with the DB 7735511C sonde as an initial estimate, only three iterations were required to reduce the magnitude of the residual radiances to the level of the noise added to the spectrum.

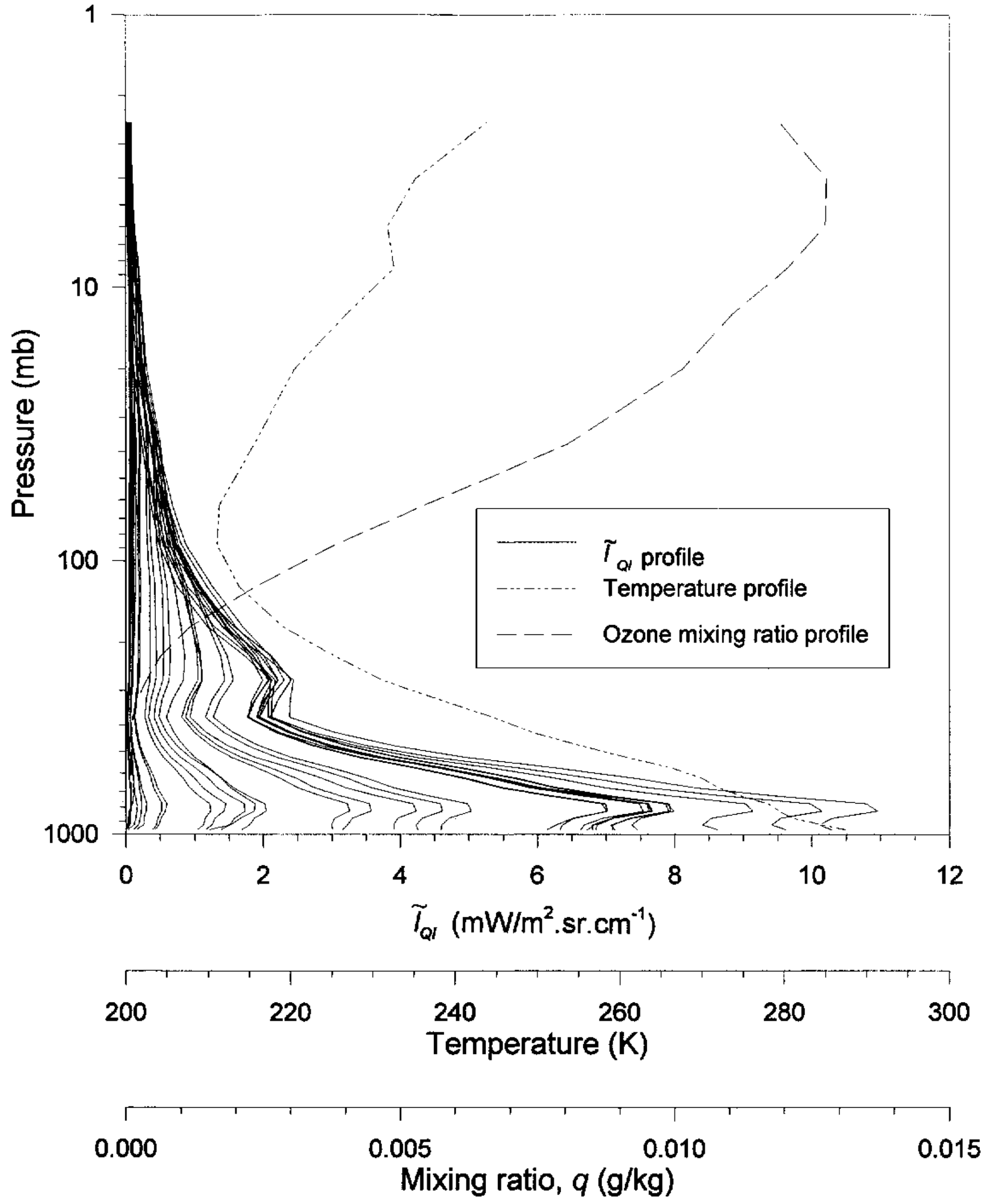
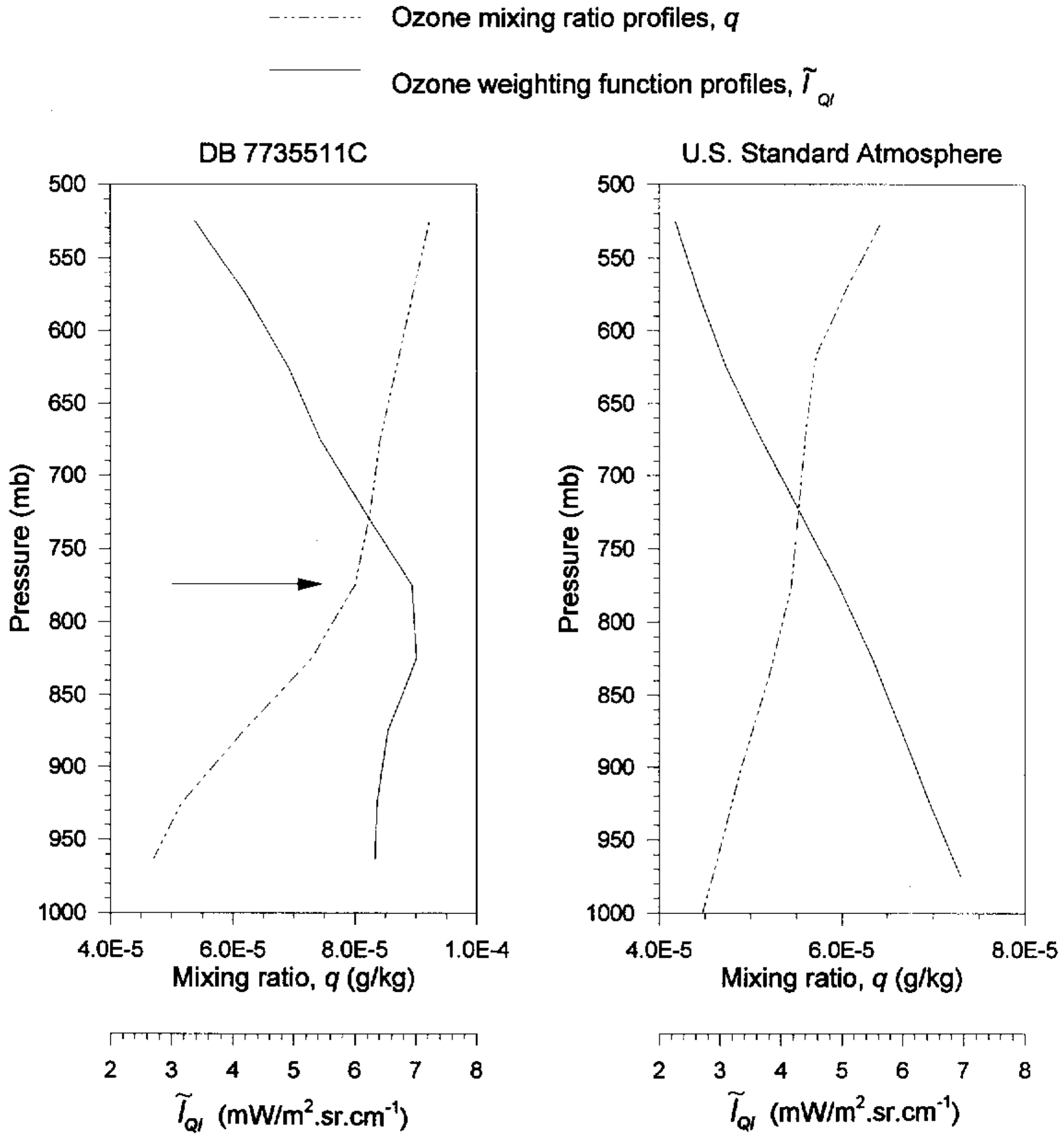


Figure 4.10 Equivalenced relative ozone weighting functions, $\tilde{I}_{Ql}$, with temperature and mixing ratio profiles for 27 layer atmosphere using the DB 7735511C sonde. The 30 weighting functions shown were selected from spectral region : 980-1080cm^{-1}; $\Delta\nu = 0.482$cm^{-1}



(a) **(b)**

Figure 4.11 Comparison of sonde DB 7735511C and U.S. Standard atmosphere tropospheric ozone profiles, q, and representative equivalenced relative weighting functions, $\tilde{I}_{Ql}$. **(a)** DB 7735511C ozone profile and weighting function showing the correspondence of the ozone profile and weighting function magnitude decrease commencing at 775mb (*arrow*). **(b)** U.S. Standard atmosphere ozone profile and weighting function. The weighting function does not display any surface feature due to the smoother ozone profile to 1000mb. Note change of scale in mixing ratios.

First iteration

The radiance difference between the synthetic radiance spectrum, $\mathbf{r}_{obs}$, and retrieval iteration guess spectrum, $\mathbf{r}(\mathbf{q}_n)$, is shown in figure 4.12. The residual signal conforms with the expected response, given the ozone perturbation profile shown in figure 4.9. The positive perturbation in the troposphere is producing a positive radiance difference in the channels which are more sensitive to tropospheric emission (1020-1040, 1050-1060cm^{-1}) and the negative perturbation in the stratosphere giving the negative response in channels predominantly sensitive to stratospheric emission (980-1000, 1070-1080cm^{-1}). The transition between negative and positive radiance perturbations in figure 4.12 is relatively sharp because the temperature profile and the initial estimate ozone profile mismatch essentially causes a negative *lower stratospheric* ozone perturbation to be treated as a negative *upper tropospheric* perturbation. Thus, the upper tropospheric emission in the initial estimate radiance spectrum is greater than that for the "observed" spectrum. This situation is reflected in the channels sensitive to radiance perturbations in this region, roughly 1000-1010cm^{-1} and 1065-1070cm^{-1}.

The first iteration ozone profile retrieval from the synthetic radiance spectrum calculated from sonde F2S 17NV2251 is shown in figure 4.13 using equation (4.1). The most notable change is the correction of the ozone profile between approximately 30-300mb where good agreement between the initial estimate and first iteration retrieval is achieved. The ozone increase has been shifted vertically to conform with the temperature profile. The tropopausal region identified from both the temperature and ozone profiles now correspond to the same pressure - about 100mb. The lower and mid-stratospheric regions appear well retrieved on the first iteration. Note the retrieval performance is relatively poor in an absolute sense in the upper stratosphere at the peak of ozone concentration. This is not unexpected because the radiative sensitivity to ozone concentration changes in this region is small compared to the rest of the atmosphere and the nominal noise level. The retrieval scheme is obtaining more upper stratospheric profile information from the *a priori* data rather than from the synthetic radiance spectrum.

A magnification of the retrieval in the troposphere is shown in figure 4.14. From 400mb down, the retrieval performs well, capturing the ozone decrease from 800mb to the surface. The small scale peak in the true ozone profile at 800mb was not captured but the rate of change of ozone at the surface matches closely that for the true profile. The shape of the weighting functions for ozone suggest that resolving small structural features, such as the peak at 800mb in the true profile, would be difficult. The large scale correction to the ozone profile from 30mb to the surface varies about the true profile intersecting it near 750, 400, and 100mb (figure 4.13 and 4.14). The retrieval underestimate between 100 and 400mb is radiometrically compensated by the overestimate in mixing ratio between 2 and 100mb.

The retrieval of the major perturbation features is particularly evident if we view the retrieval as a perturbation about the initial estimate as shown in figure 4.15. The differences that remain are in the upper stratosphere (2-20mb), the upper troposphere (100-300mb) and the boundary layer regions (surface-700mb). The effectiveness of a second iteration is dependent on the radiative signal obtained from ozone concentration differences that remain between the true profile and the retrieval. The magnitude of the upper tropospheric perturbation was not fully recovered and the remaining difference is large enough to have a measurable radiative impact. However the radiative contribution due to the remaining perturbations in the upper stratosphere and boundary layer regions are small and nearing the limit of radiometric detection.

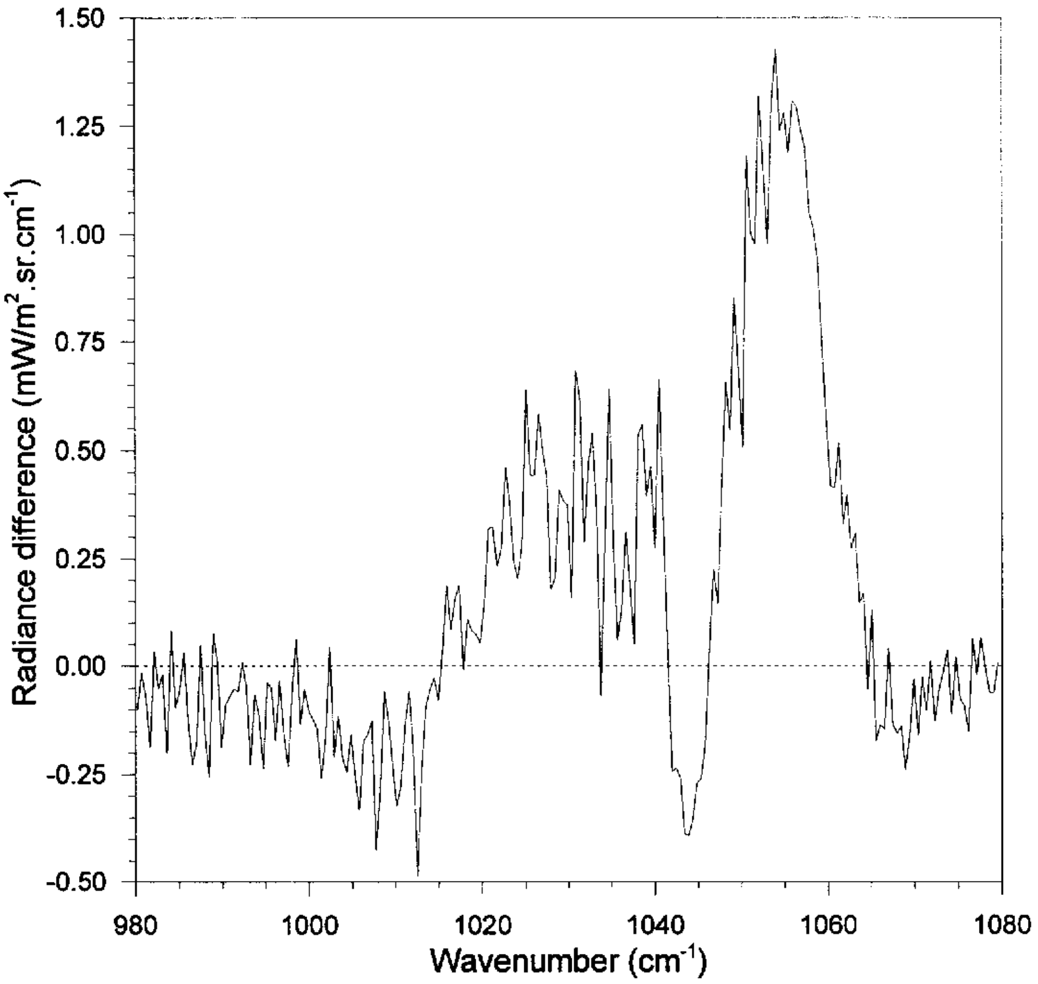


Figure 4.12 Initial estimate residual radiance spectrum, $\mathbf{r}_{obs}$ - $\mathbf{r}(\mathbf{q}^o)$. Radiance difference between the synthetic radiance spectrum, calculated from F2S 17NV2251 sonde, and the first iteration guess spectrum, calculated from the DB 7735511C sonde. Note the initial estimate and first iteration guess are the same.

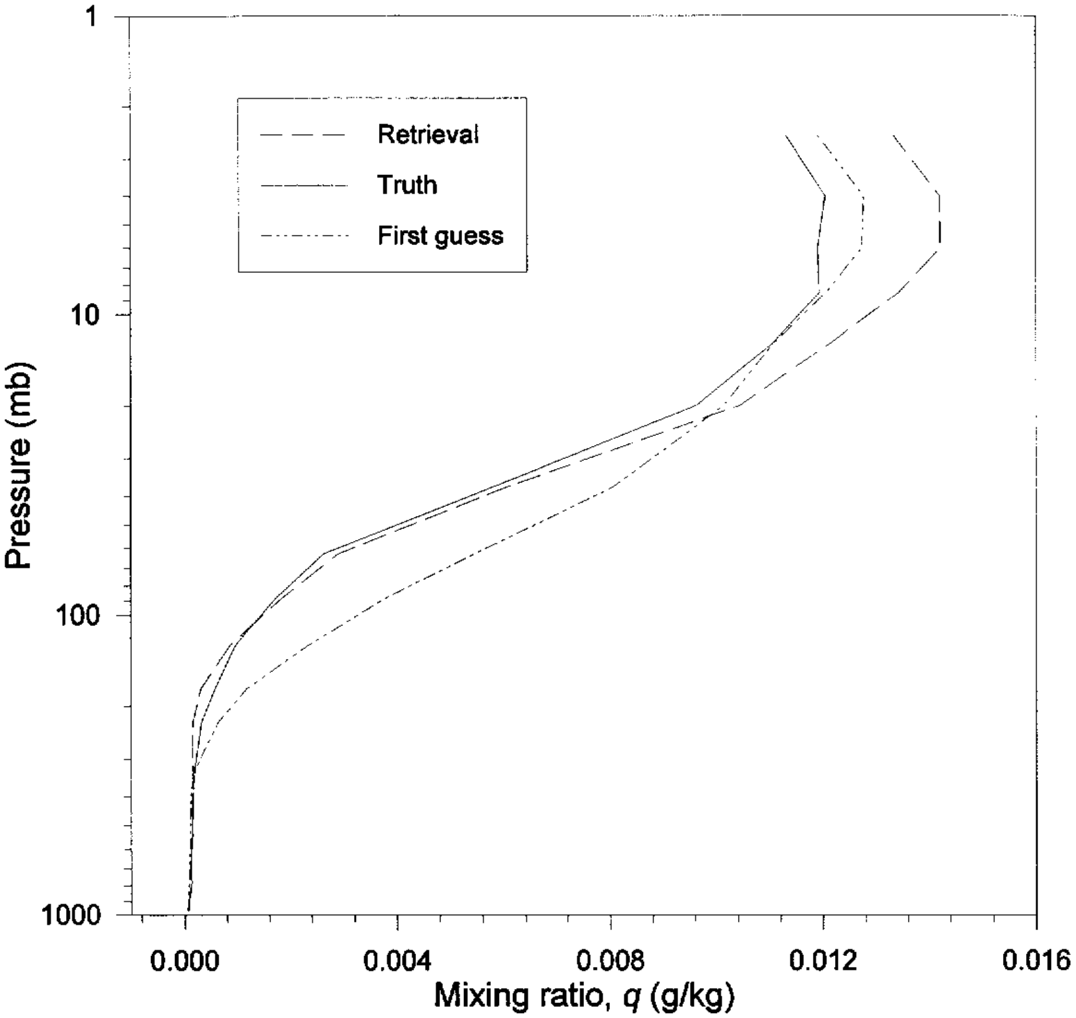


Figure 4.13 First iteration ozone retrieval. *True profile* : F2S 17NV2251. *Iteration guess profile* : DB 7735511C. *Initial estimate profile* : DB 7735511C

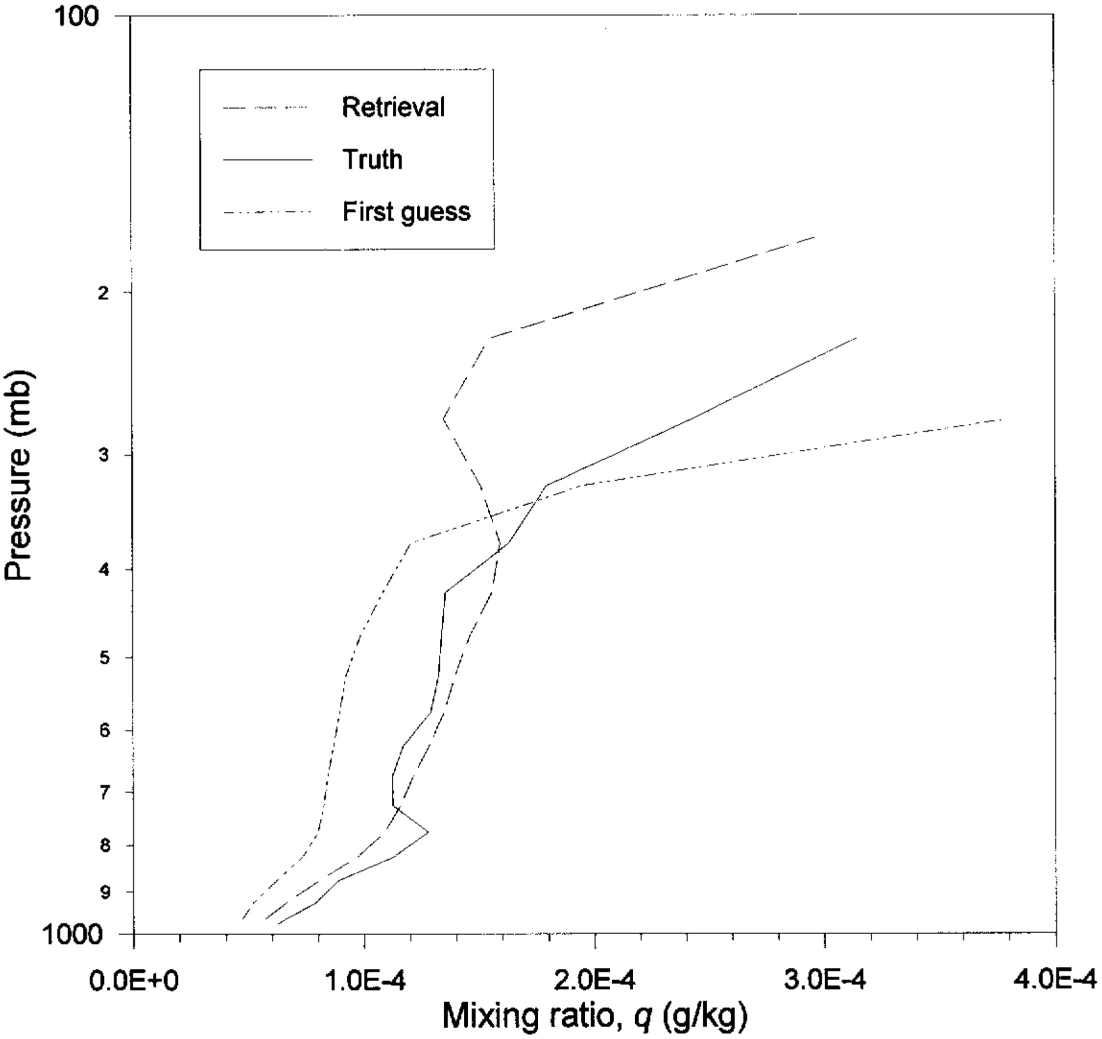


Figure 4.14 First iteration ozone retrieval - tropospheric magnification. *True profile* : F2S 17NV2251. *Iteration guess profile* : DB 7735511C. *Initial estimate profile* : DB 7735511C

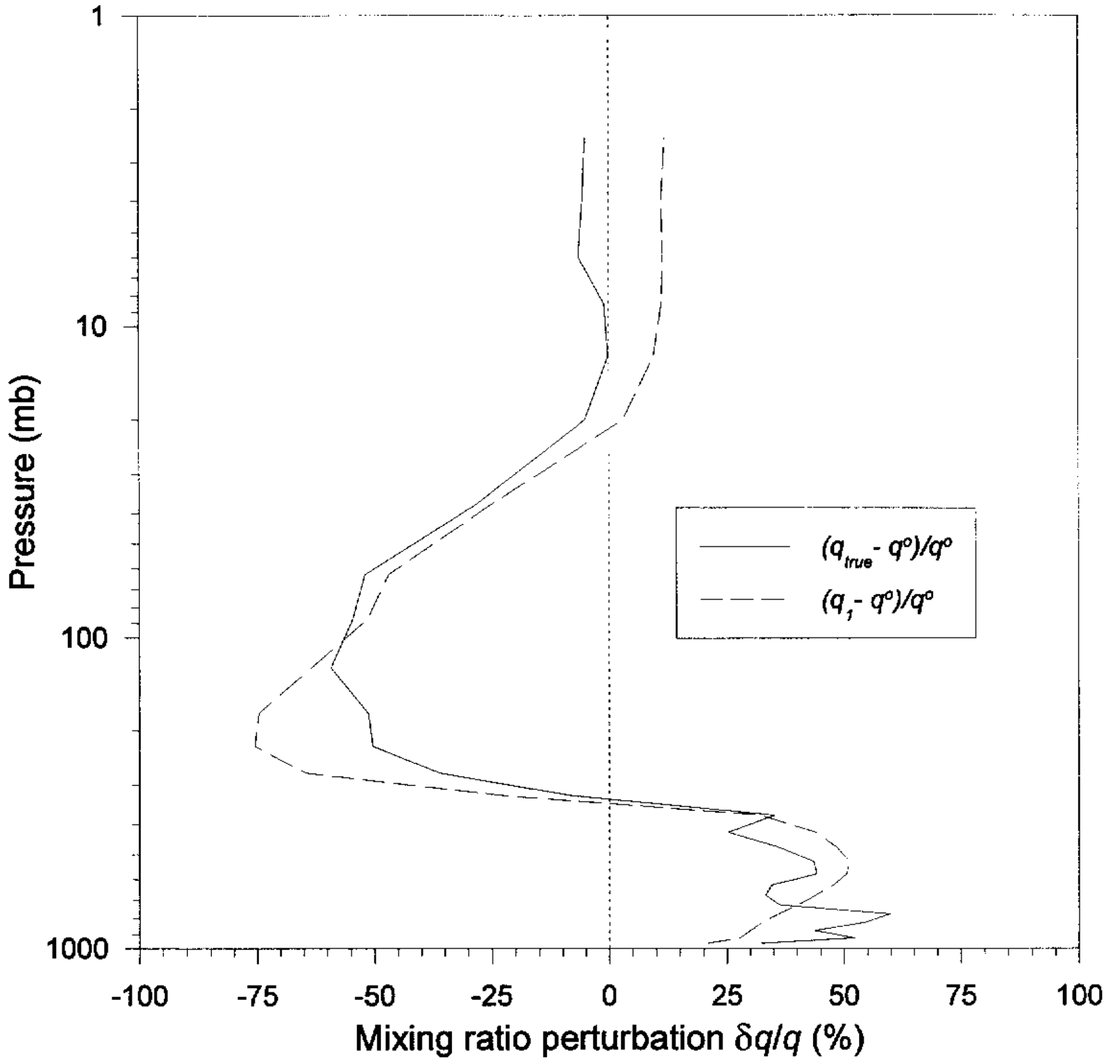


Figure 4.15 True and first iteration retrieved ozone perturbations about initial estimate. *True profile*, q_{true} : F2S 17NV2251. *Initial estimate profile*, q^o : DB 7735511C

Second iteration

The procedure used to identify if a further iteration was required was based on the two retrieval products. The first was the magnitude of the residual radiance spectrum (i.e. between the synthetic spectrum and the radiance spectrum calculated from the first iteration retrieval). This difference, $\mathbf{r}_{obs}$ - $\mathbf{r}(\mathbf{q}_1)$, where $\mathbf{q}_1$ is the first iteration retrieved ozone profile vector, is shown in figure 4.16. This may be compared to the initial radiance difference spectrum, figure 4.12. If the residual radiance spectrum was noticeably greater than the nominal sensor noise level ($0.1\text{mW/m}^2\text{.sr.cm}^{-1}$) and contained identifiable spectral features, a further iteration was performed. As can be seen from figure 4.16, these conditions are satisfied. The second criterion for a further iteration was the comparison of the absolute value of the ozone profile differences, $|\mathbf{q}_{n+1} - \mathbf{q}_n|$, to the rms retrieval error. If the term differences were greater than the retrieval rms error - taken from the diagonal of retrieval error covariance matrix of equation (2.52) - then another retrieval iteration was performed. The absolute term difference and rms retrieval error for the first iteration are shown in figure 4.17. The first iteration term difference is larger than the rms retrieval error in the region 30-300mb where the major ozone profile adjustment occurred.

The second iteration retrieval (figure 4.18) is still characterised by the approximately 15% difference in the upper stratospheric ozone mixing ratio profile. The retrieval however has moved closer to the true profile in the upper stratosphere and now matches the true profile near 20mb. The retrieval scheme will always perform poorly in this region due to the lack of upper stratospheric information (small magnitude weighting function; see figure 4.10) in an uplooking ozone spectrum. An ozone perturbation of 15% in the upper stratosphere does not provide a radiative contribution which is above the noise level. Thus, the retrieval performance, in an absolute sense, appears poor between 2-10mb due to the spectral differences not containing information about comparatively small stratospheric perturbations and thus profile information comes from the *a priori* data - constraining the solution within a physical regime.

A magnification of the tropospheric region of the second iteration retrieval is shown in figure 4.19. This figure should be compared to the tropospheric profile from the first iteration, figure 4.14. The tropospheric retrieval performs exceptionally well, fitting the true profile closely and correcting for the large differences from 100-400mb which existed after the first iteration. This is very evident in figure 4.20 where the perturbation retrieval has closely matched the true perturbation. Examination of this region in figure 4.18 (and comparison with figure 4.15) indicates that this overshoot has effectively been shifted upwards from the upper troposphere into the lower stratosphere. The tropospheric retrieval still does not retrieve the ozone peak observed at approximately 800mb. As commented earlier, at that pressure, the vertical structure of this feature cannot be resolved by the retrieval scheme.

The second iteration retrieval shows how the low information content of the ozone radiance spectrum gives rise to a compensation of radiance contributions from different regions in the atmosphere. It might initially be expected that the ozone peak at 800mb would be quite visible in the residual radiances. However, because the weighting functions are similar in shape in the troposphere, the retrieval scheme cannot adequately resolve the differences between a peak of small vertical scale at 800mb, or the mixing ratio difference on a broader vertical scale from 200-500mb (see figure 4.19). Thus these two regions effectively compensate each other.

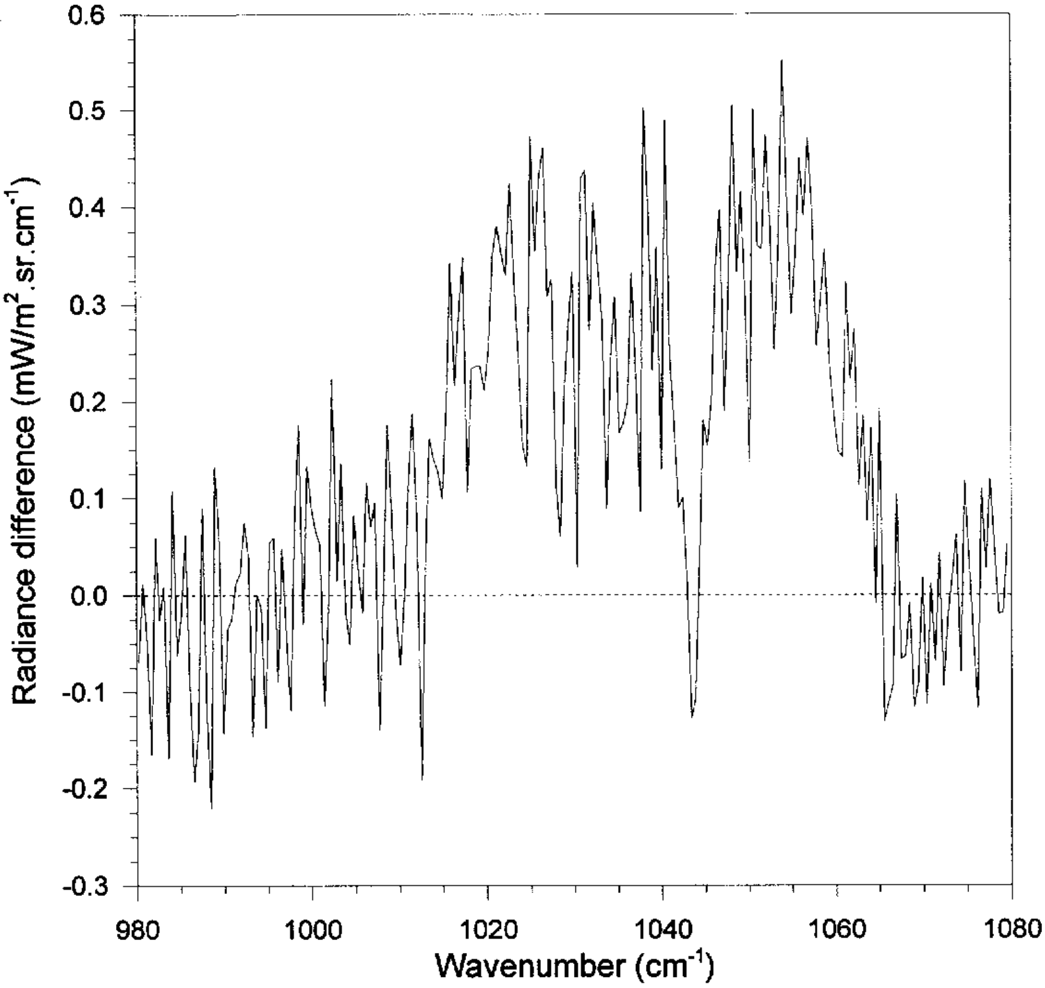


Figure 4.16 First iteration residual radiance spectrum, $\mathbf{r}_{obs}$ - $\mathbf{r}(\mathbf{q}_1)$. Radiance difference between synthetic observed radiance spectrum calculated from F2S 17NV2251 sonde, and the second iteration guess spectrum calculated from the first iteration retrieval.

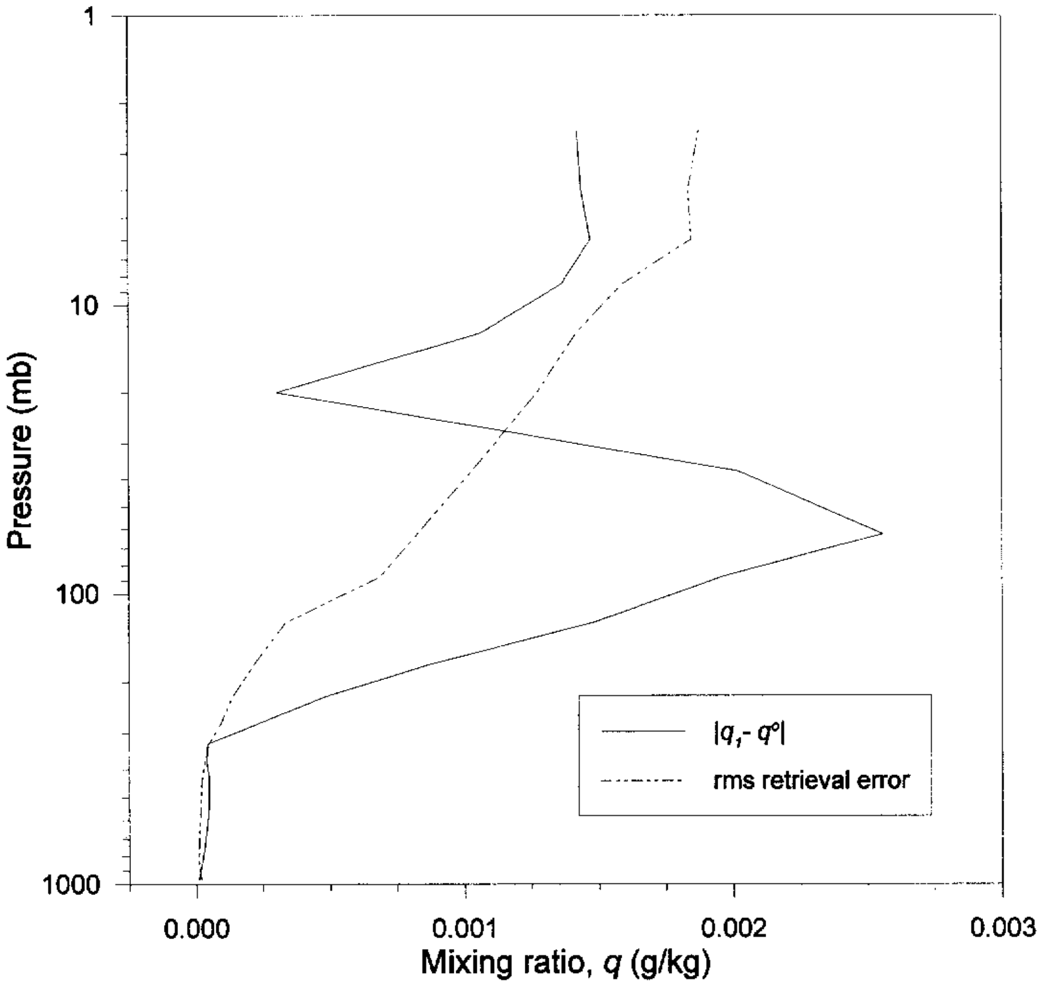


Figure 4.17 First iteration term difference, $|q_1 - q^o|$, comparison with absolute rms retrieval error.

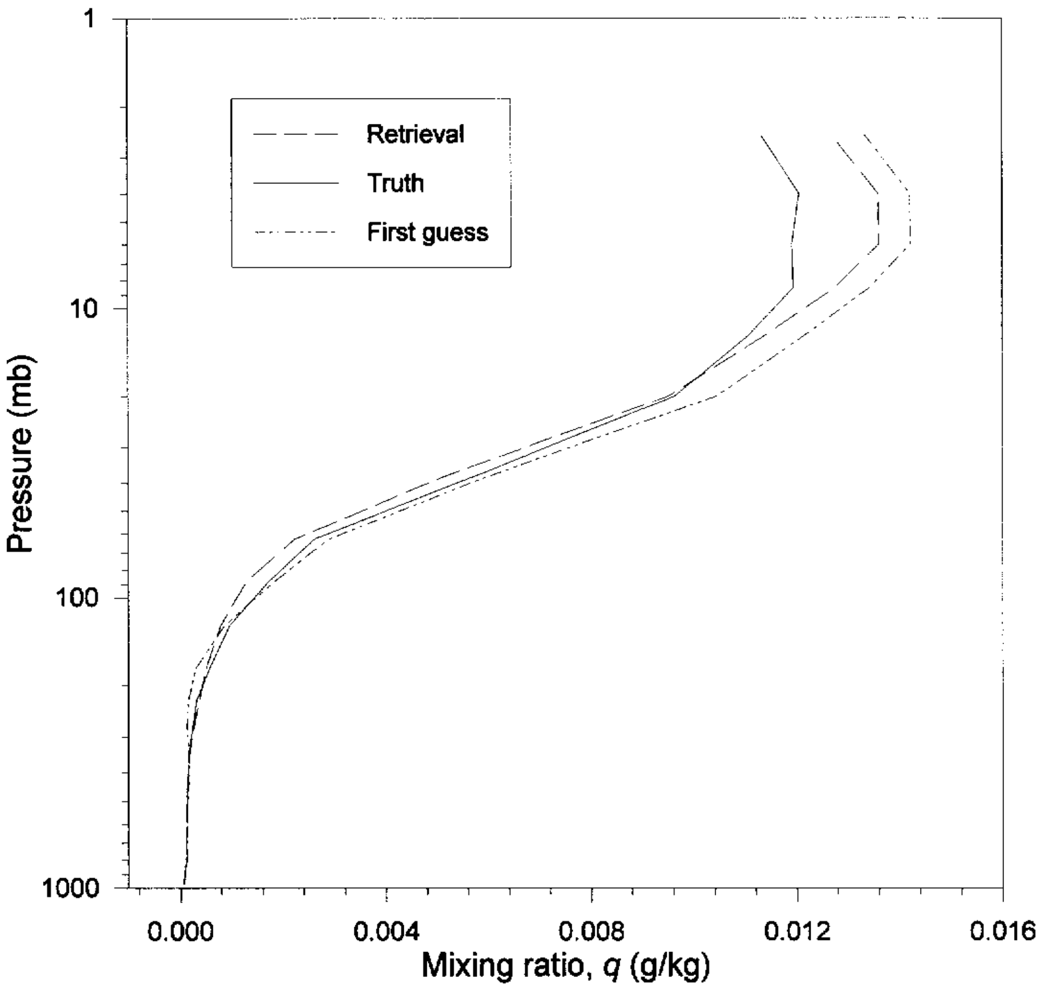


Figure 4.18 Second iteration ozone retrieval. *True profile* : F2S 17NV2251. *Iteration guess profile* : First iteration retrieval. *Initial estimate profile* : DB 7735511C

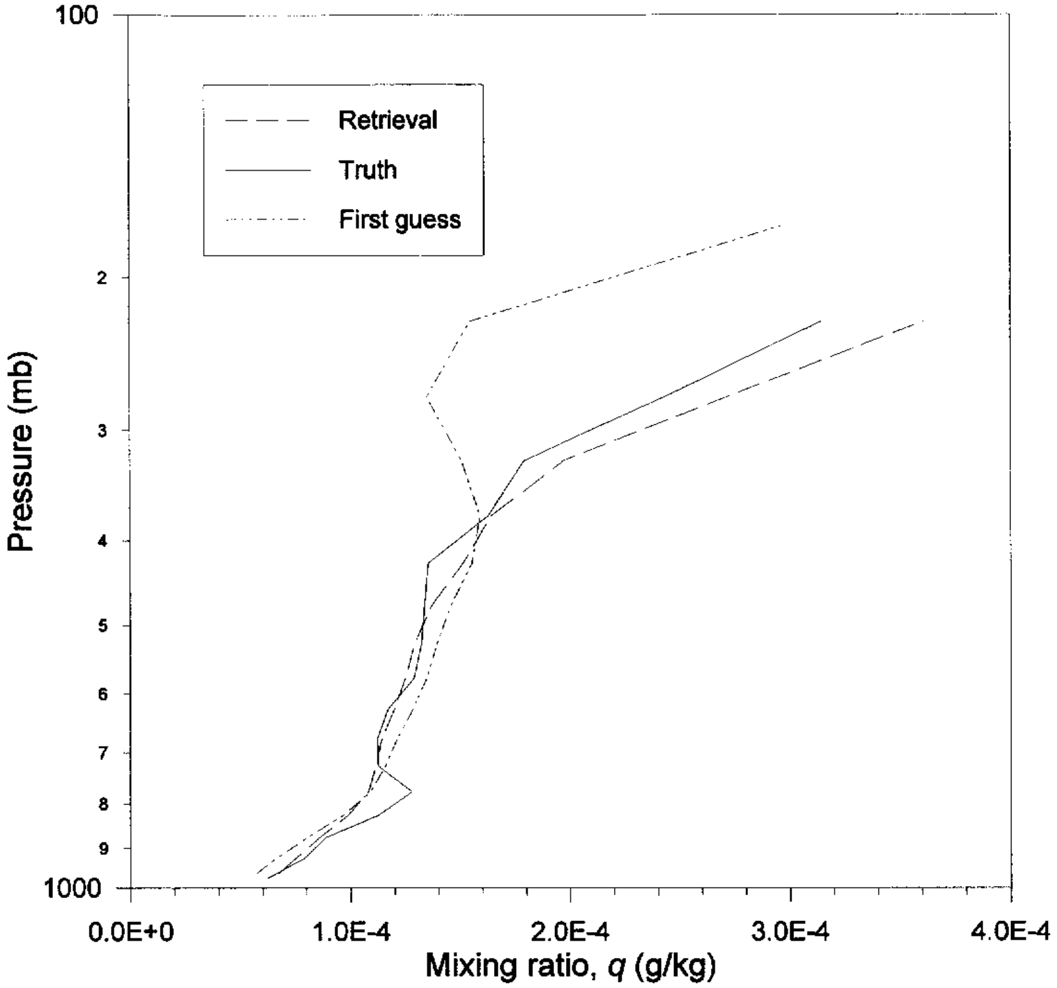


Figure 4.19 Second iteration ozone retrieval - tropospheric magnification. *True profile* : F2S 17NV2251. *Iteration guess profile* : First iteration retrieval. *Initial estimate profile* : DB 7735511C

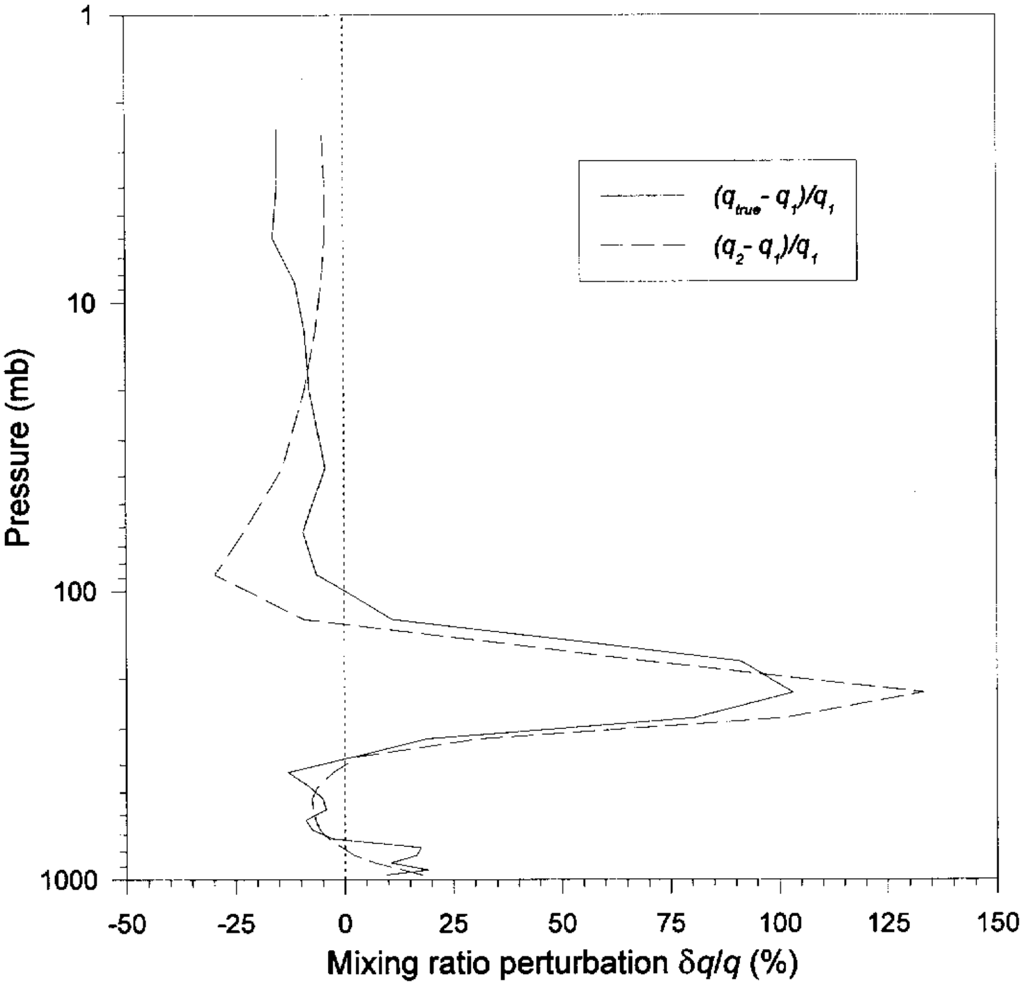


Figure 4.20 True and second iteration retrieved ozone perturbations about first iteration retrieval. *True profile*, q_{true} : F2S 17NV2251. The major perturbation structure has been retrieved well with good performance in the troposphere.

Third iteration

The residual radiance spectrum for the second iteration retrieved ozone profile, shown in figure 4.21, closely resembles the Gaussian noise spectrum added to the F2S 17NV2251 calculated spectrum. There is no recognisable ozone spectral structure in the residual radiance, however comparison of figures 4.5 and 4.21 shows there are differences in the second order polynomials fitted to the spectra. The shape of this polynomial fit in figure 4.21 suggests some small amount of ozone profile information in the residual radiance in that it peaks in the region 1040-1060cm^{-1}.

This is confirmed by viewing the difference between figure 4.5 and 4.21, shown in figure 4.22. The ozone spectral structure is evident, but the magnitudes are below the rms noise level. It is clear that nearly all of the stratospheric information has been extracted from the spectrum after two iterations - the radiance differences in the spectral regions 980-1000cm^{-1} and 1070-1080cm^{-1} are effectively zero. What is suggested by figure 4.22 is that only upper tropospheric (1000-1015cm^{-1}) and lower tropospheric (the peak at 1055cm^{-1}) information remains. This is consistent with the differences between the true ozone profile and the second iteration retrieval. The decrease in figure 4.22 from 1015-1040cm^{-1}, also indicates that the middle tropospheric region has been retrieved satisfactorily.

The residual radiances alone suggest that a third iteration is not required, but the second iteration term differences are greater than the rms retrieval error in the 200-300mb region (figure 4.23). A third iteration was carried to examine how the retrieval scheme would perform with a very noisy, small radiance signal and to update the weighting functions with the latest retrieved profile.

The perturbation retrieval is shown figure 4.24. The absolute retrieved and iteration guess profiles are difficult to distinguish and are not included here. The small radiance signal was not able to provide enough information to retrieve the gross structure of the remaining perturbation, but the small corrections made brought the term differences well below the rms retrieval error level. Note also that the corrections made in the third iteration are consistent with the compensation observed in the second iteration retrieval at 200mb. The upper tropospheric ozone concentrations have been redistributed to reduce this difference. This is consistent with the radiance difference shown in figure 4.22. The remaining radiative signature has the form of a tropospheric signal due to an upper tropospheric perturbation (1000-1015cm^{-1}) and boundary layer perturbation (1055cm^{-1}).

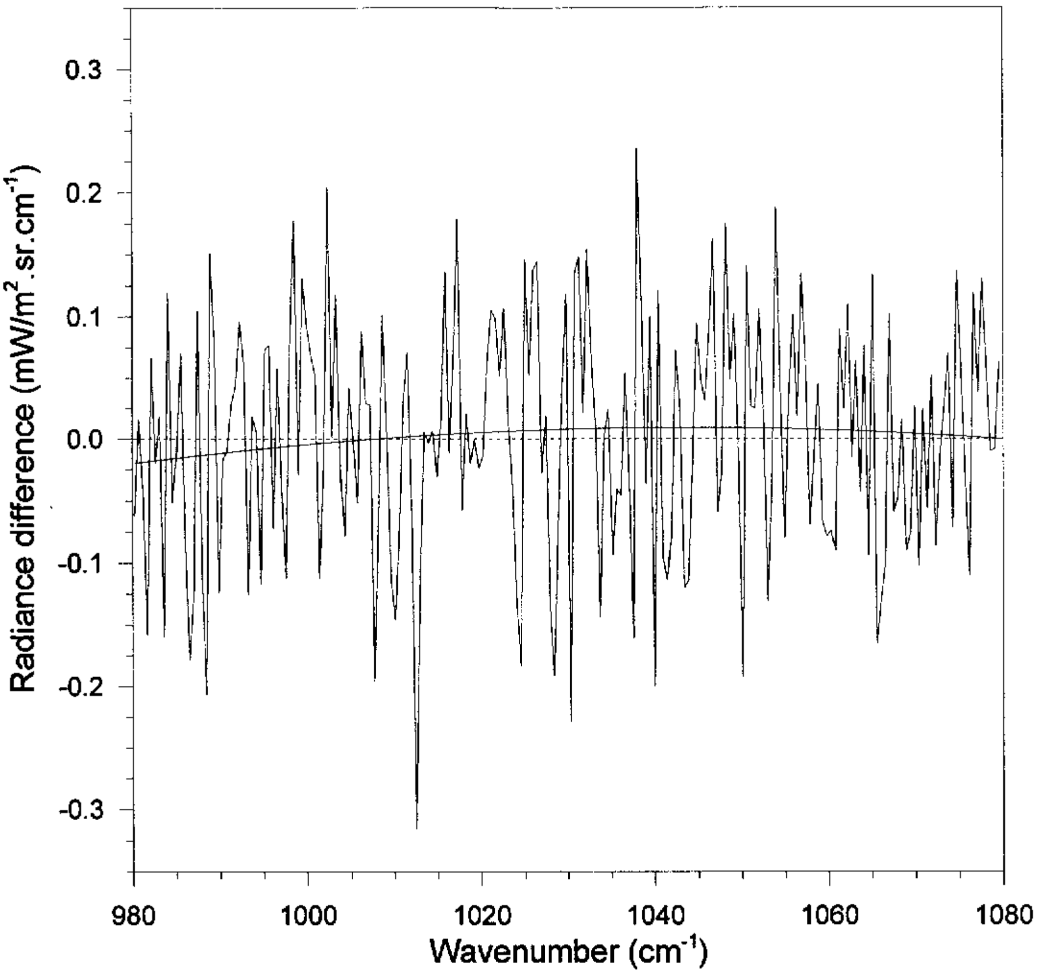


Figure 4.21 Second iteration residual radiance spectrum, $\mathbf{r}_{obs}$ - $\mathbf{r}(\mathbf{q}_2)$. Radiance differences between the synthetic ("observed") radiance spectrum, calculated from F2S 17NV2251 sonde, and the third iteration guess spectrum calculated from the second iteration retrieval. The spectrum closely resembles that of the Gaussian noise spectrum added to the F2S 17NV2251 sonde spectrum shown in figure 4.5.

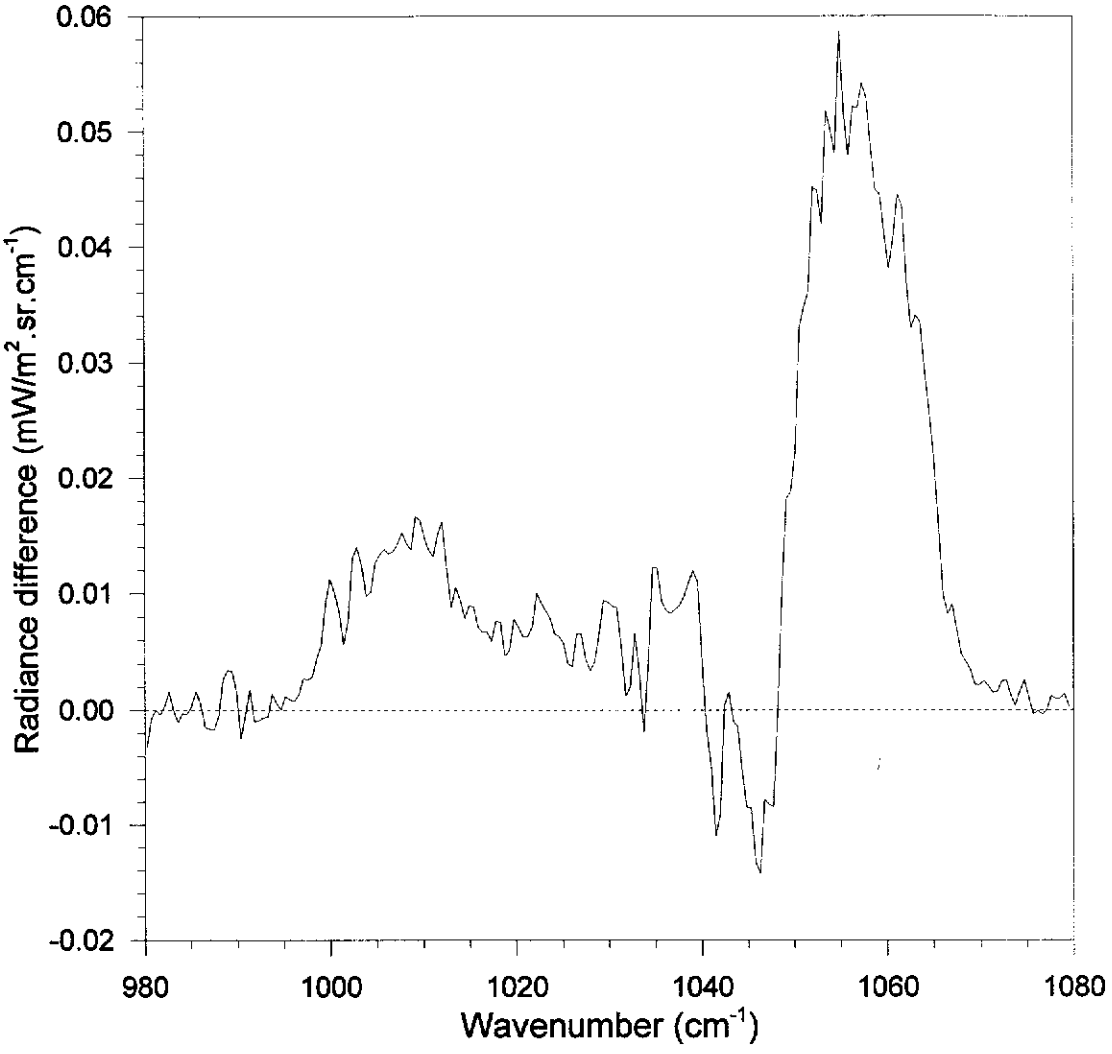


Figure 4.22 Difference between second iteration residual radiance spectrum, $\mathbf{r}_{obs}$ - $\mathbf{r}(\mathbf{q}_2)$ (figure 4.20), and the Gaussian noise spectrum (rms 0.1mW/m^2.sr.cm^{-1}) added to the F2S 17NV2251 sonde calculated radiance (figure 4.5). The ozone spectral structure is evident, but is below the rms noise level

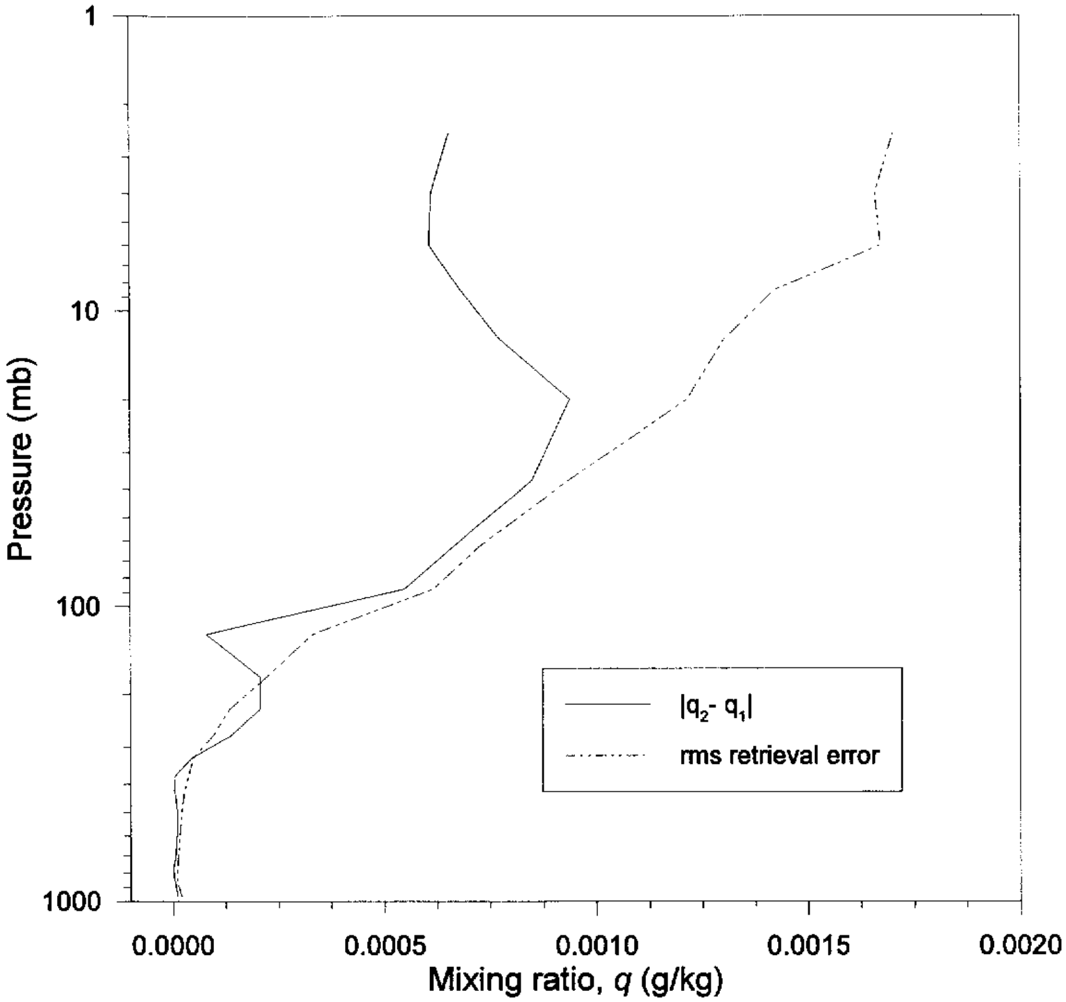


Figure 4.23 Second iteration term difference, $|q_2 - q_1|$, comparison with absolute rms retrieval error.

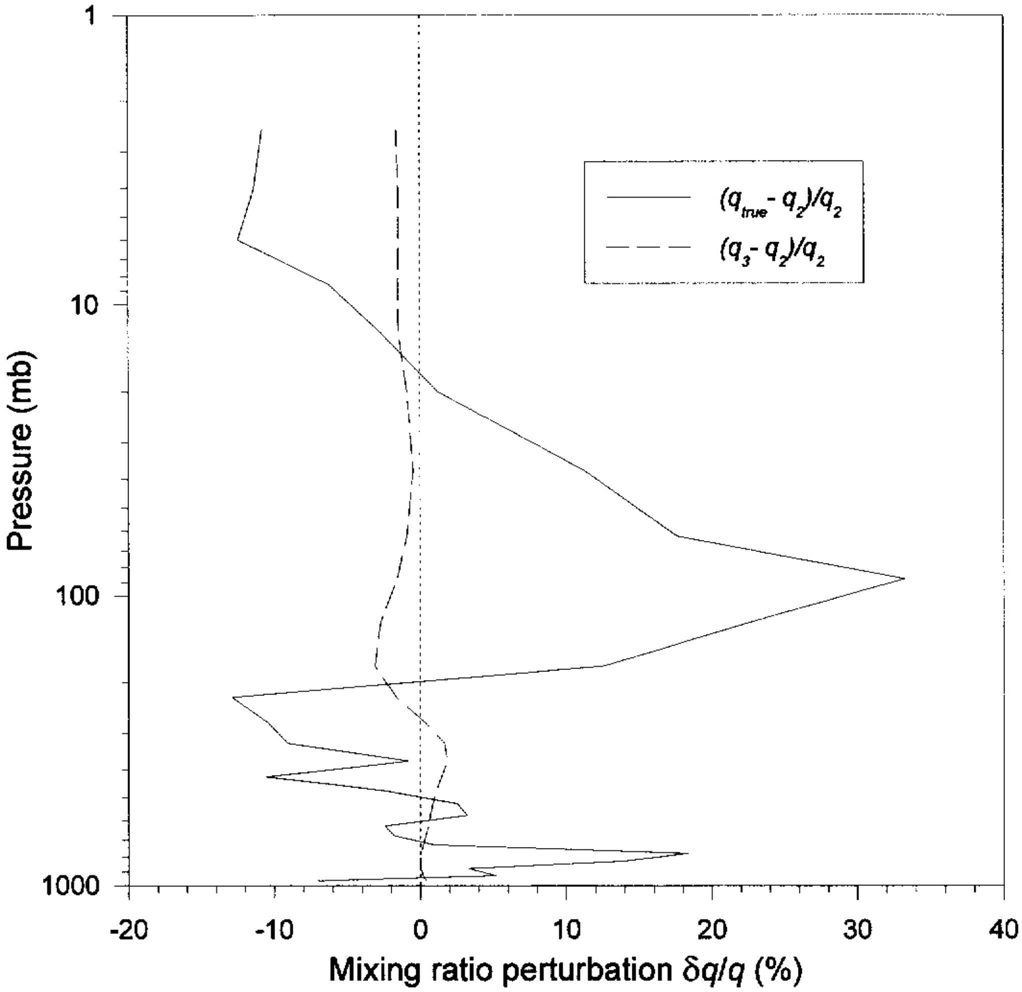


Figure 4.24 True and third iteration retrieved ozone perturbations about second iteration retrieval. *True profile*, q_{true} : F2S 17NV2251. The remaining vertically broad stratospheric perturbation has not been retrieved due to poor sensitivity. The tropospheric fine scale structure was also not picked up by the retrieval scheme. Note the redistribution of ozone in the upper troposphere in this iteration, consistent with the remaining differences observed after the second iteration retrieval.

4.3.3 Ozone retrieval analysis

In analysing retrieval errors, it is instructive to look first at the vertical resolution of retrieved layers. As described previously, the ozone mixing ratio solution vector, $\hat{\mathbf{q}}$, is related to the true ozone mixing ratio, **q**, by

$$\hat{\mathbf{q}} = \mathbf{R}\mathbf{q}$$

where **R** is the resolution matrix. The number layers resolved in the retrieval scheme by the measured data are obtained by calculating the sum of the diagonal elements of **R** (Tarantola, 1987). This value has also been termed the effective degrees of constraint, or the number of independent precise measurements, imposed on the observed data (Huang et al., 1992; Purser and Huang, 1993). It should be noted that the term "number of layers" does not define particular, discrete layers in the retrieved profile but refers to the amount of ozone profile information supplied by the measured data in the retrieval scheme; information that is distributed throughout profile space. The number of layers resolved by the *a priori* data can be obtained by simply subtracting the trace of **R** from the total number of layers used in the retrieval scheme. The number of layers resolved by the measurements and by the *a priori* data for each of the three iterations is shown in table 4.2. Effectively, only four independent pieces of information about the vertical distribution of ozone are obtained from the 209 channel observed radiance spectrum. This high degree of interdependence between channel radiances is expected considering the similar shapes of the ozone weighting functions.

Iteration	No. of independent pieces of information resolved by "observed" data	No. of independent pieces of information resolved by *a priori* data
1	4.11	22.89
2	4.06	22.94
3	4.02	22.98

Table 4.2 Effective degrees of constraint of observed and *a priori* data to 27 layer ozone profile retrieval scheme. Numbers rounded to 2 decimal places.

The small amount of independent information in the observed radiances effectively means no layer is completely resolved by the measurements alone. To reduce the dependence of the retrieval on *a priori* data, the number of atmospheric layers retrieved should not be much greater than the number of independent pieces of information resolved by the observed data. The dependence of the retrieval scheme on the *a priori* data means that the error characteristics of the retrievals will reflect the statistics of the *a priori* data set. The statistics describing the variance of the *a priori* data - the absolute and relative standard deviation profiles - are shown in figure 4.25(a) and (b). The two plots are simply the square roots of the diagonal elements of the covariance matrices shown in figures 4.2(a) and (b).

The structure of the variance of the *a priori* data is mirrored in the rms retrieval errors. Retrieved ozone amounts in atmospheric regions of high ozone variability have a larger associated error than regions of low variability. The retrieval error is calculated from equation (2.52); the square root of the diagonal elements of the retrieval error covariance operator. The retrieval error profiles for the first and second iterations are shown in figure 4.26(a) and (b) respectively. The third iteration error profiles are similar to those of the second iteration. The similarity in the shapes of figure 4.25(b) and the percentage errors in figures 4.26(a) and (b) demonstrate that the major source of retrieval error is the null space error due to the poor vertical resolution of the retrieval scheme. The strong dependence of the rms retrieval error on the *a priori* data variance is clear. The large peak in the first iteration percentage error (figure 4.27(a)) is a consequence of the large adjustment required in that region to correctly position the tropopause and is not present in subsequent iterations.

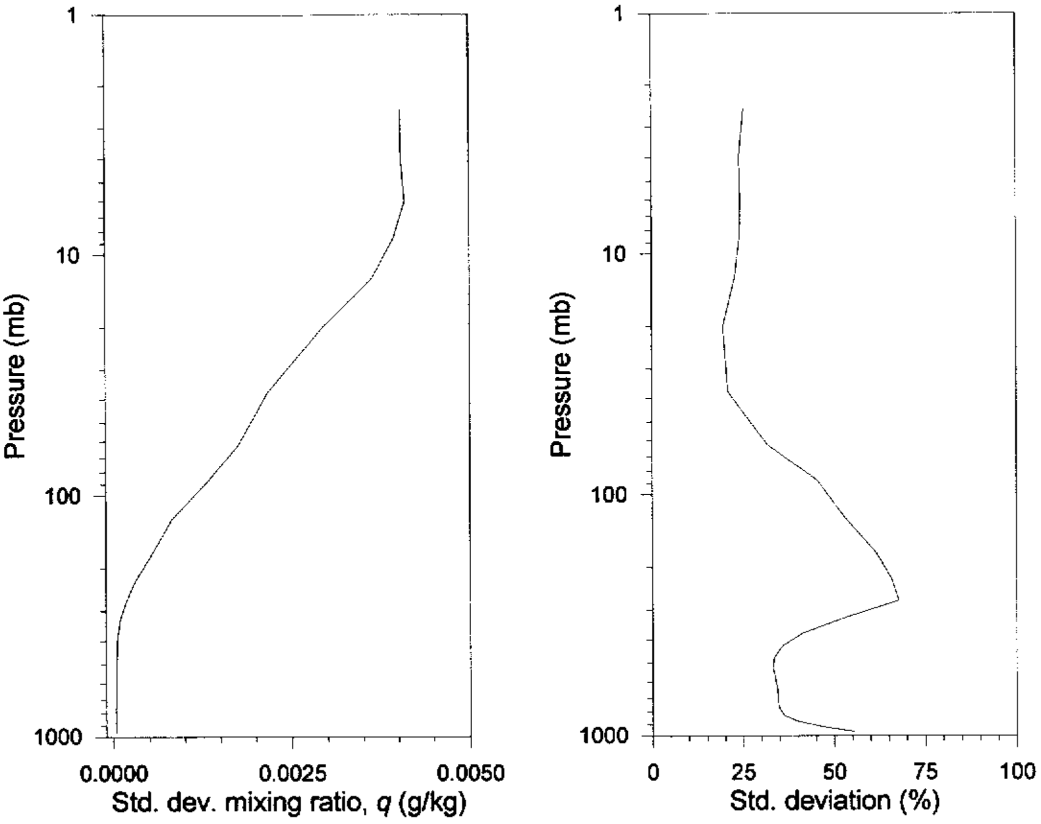
The differences between the *a priori* variance and retrieval errors can be explained by determining the vertical resolution of the retrieval scheme. This was done using the method described in Huang et al. (1992), where the vertical resolution is given by the reciprocal of the local effective data density, ρ. This method allows determination of vertical resolution even when the vertical resolution functions (VRFs) - the *rows* of the resolution operator, **R** - have large oscillatory sidelobes. The layer by later VRFs for the first iteration retrieval are shown in Appendix F. They behave ideally from layer 1 (surface) to about layer 14 (325mb), having a single peaked Gaussian-like form. At layer 15 (275mb) and above they begin oscillating on the high pressure side of the peak value. Simple inspection of the spread of the VRFs will therefore not provide an assessment of the vertical resolution of the retrieval scheme in those regions where the oscillations occur.

The vertical resolution and data density for the first iteration retrieval are shown in figure 4.27(a) and (b). Most of the information is concentrated in the troposphere, particularly in the boundary layer. The higher data density in the troposphere is responsible for the reduction in the retrieval error profile from that in the *a priori* data statistics in the troposphere. The percentage standard deviation of the *a priori* data in the troposphere varies from 55% at the surface, 35% from 400-800mb, and 70% around 200mb (figure 4.25(b)). These errors are reduced to 30%, 10-15%, and 40% respectively (figure 4.26(b)) after two iterations due to relatively good vertical resolution in the lower and middle troposphere. The characteristics of the *a priori* statistics will always be present in the retrieval error but the information from the measurements reduces their magnitude in the troposphere. Higher up, at 100-200mb, the vertical resolution begins to approximate the depth of the entire atmosphere. Here contributions from all levels of the atmosphere are included in a retrieved quantity and the retrieval scheme cannot resolve these layers.

The relative contributions of the null space and measurement errors can be determined using the analysis detailed in Huang et al. (1992). The total retrieval error covariance is composed of the null space error covariance and the measurement error covariance as shown in eqn. (2.64). In looking at the magnitudes of each component, the diagonal element covariances were used rather than their square root as summing the rms component errors is mathematically inconsistent with the total rms retrieval error. The variances of the calculated null space and measurement errors for the first and second iterations as a percentage of the total retrieval error variance are shown in figure 4.28(a) and (b). These plots explicitly show the dependence of the retrieval scheme on the *a priori* data.

In the troposphere, up to 90% of the retrieval variance is due to poor vertical resolution - even in those regions where the actual rms error is low (400-800mb) and the impact of the *a priori* data is comparatively low due to the higher data density (figure 4.28(b)). It is here that most of the retrieval information comes from the observation and thus the vertical resolution error component is the highest.

Higher up in the atmosphere, the contribution of the measurement error variance increases as the radiative signal from those regions decreases. This is accompanied by a decrease in the null space error variance as nearly all the information required to resolve these higher layers comes from the *a priori* data. The stratospheric measurement error variance is greater on the second iteration as the first iteration residual radiance contained very little information about stratospheric ozone.



(a) **(b)**

Figure 4.25 Error characteristics of the mean ozone mixing ratio profile from *a priori* data (considering the mean profile as a virtual measurement). **(a)** Standard deviation of *a priori* data. **(b)** Standard deviation as a percentage of the mean profile.

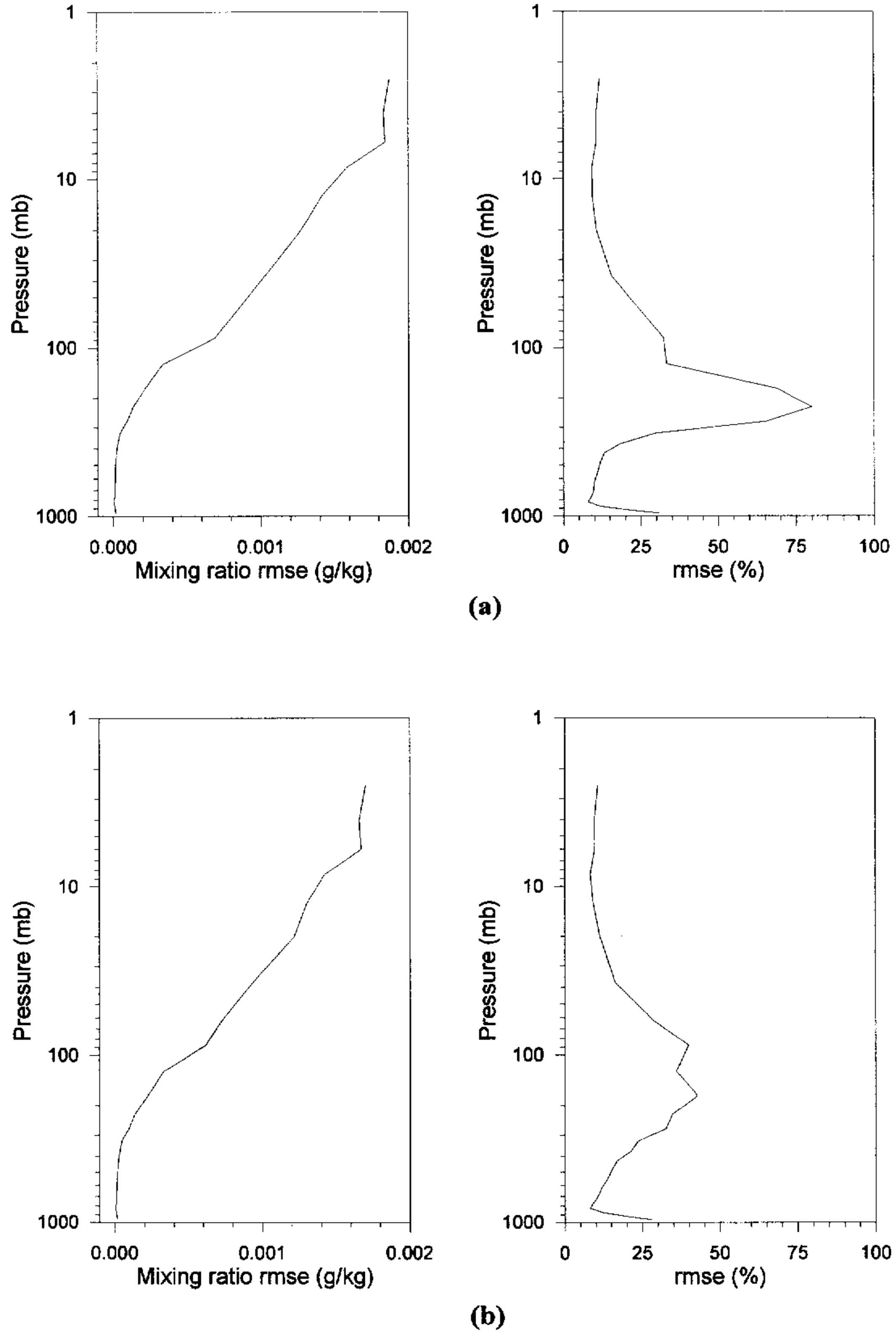
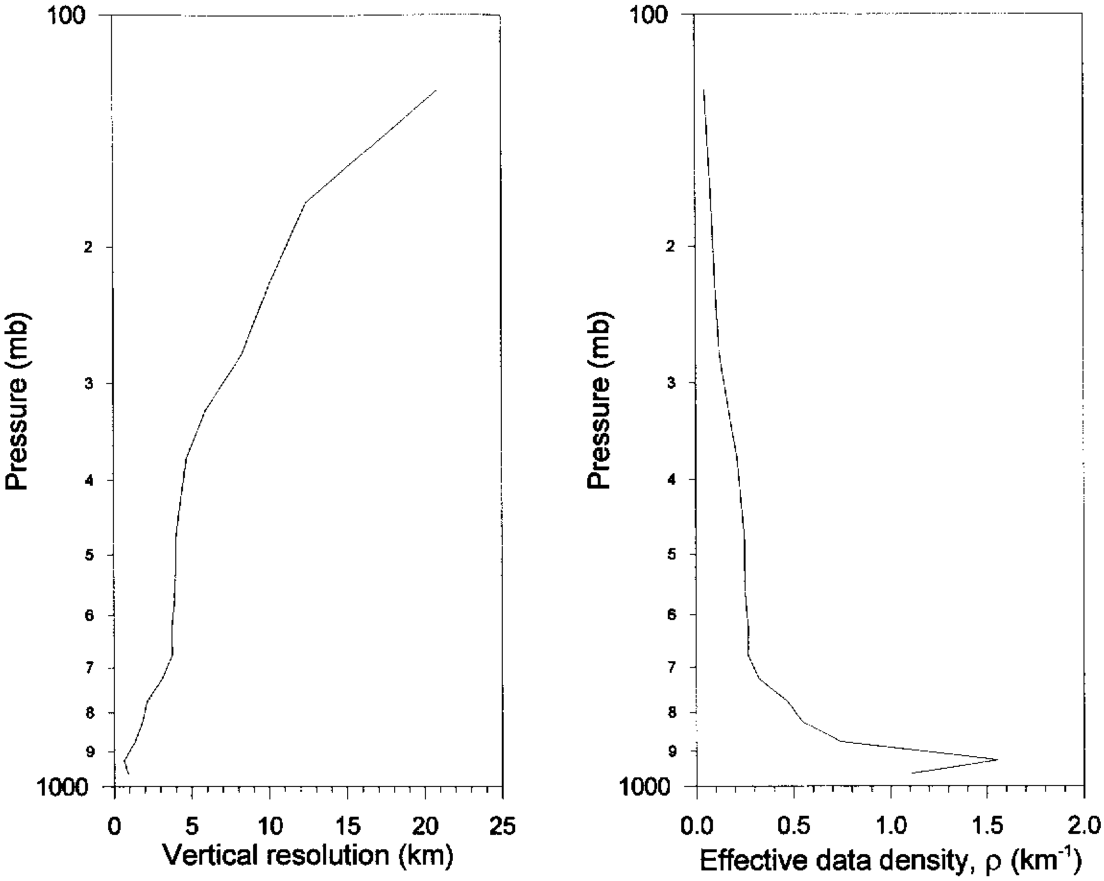
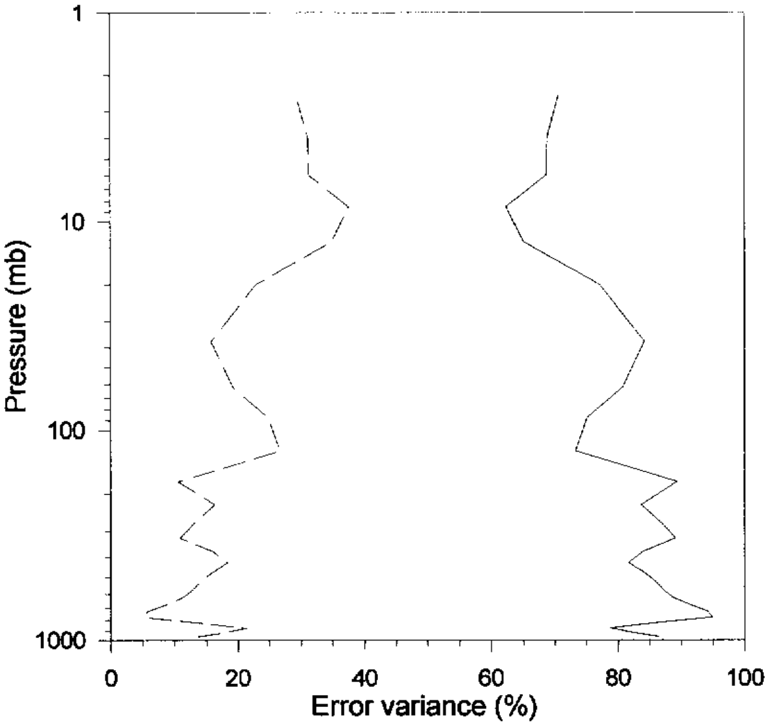


Figure 4.26 rms retrieval error (rmse) profiles. **(a)** First iteration absolute rms error (*left*) and percentage rms error (*right*). **(b)** Second iteration absolute rms error (*left*) and percentage rms error (*right*)

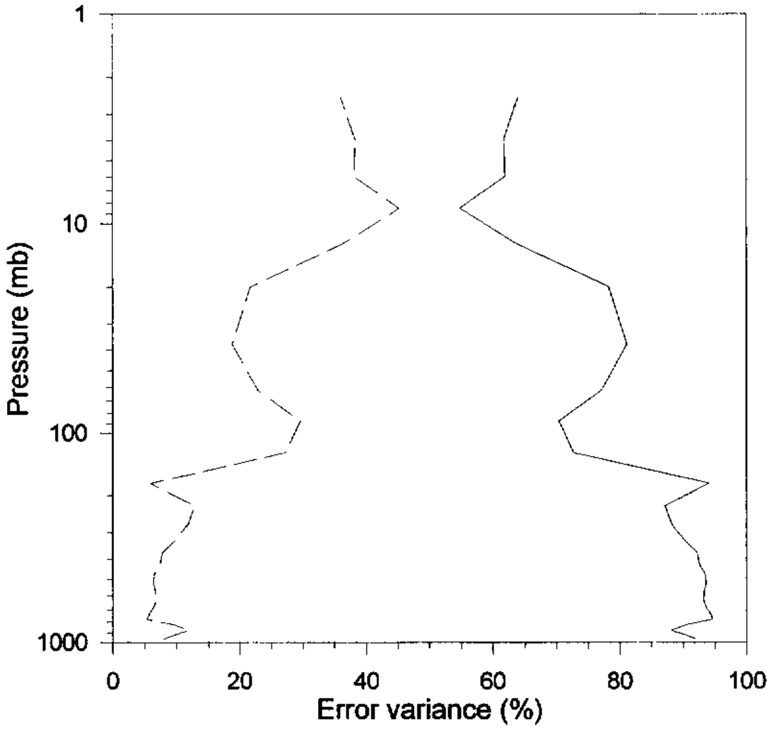


(a) **(b)**

Figure 4.27 **(a)** Vertical resolution in km and **(b)** effective data density for first iteration retrieval. Note that above 100mb the vertical resolution in (a) becomes increasing large and is not shown above this level.



(a)



(b)

Figure 4.28 Percentage error variances - diagonals of the component error covariance matrices compared to the total error covariance; (————) Null space error variance; (— — — —) Measurement error variance. **(a)** First iteration. **(b)** Second iteration.

5. Conclusions and Recommendations

The performance of the ground-based ozone profile retrieval scheme described is determined by the information content of the uplooking ozone 980-1080cm^{-1} spectrum. For the entire atmospheric column (surface - 2mb), effectively only four independent pieces of information describing the vertical distribution of ozone are available. Most of this applies to tropospheric ozone as the retrieval data density in the stratosphere is very low. This deficiency of information produces a high dependence on *a priori* data.

The ozone retrieval scheme presented performs well in the troposphere. Vertical resolution ranges from 1km in the boundary layer to 10-15km at 200mb. The retrieval accuracies vary from 30% at the surface, 10-15% in the middle troposphere, and 40% in the upper troposphere. In the lower stratosphere, the retrieval scheme was particularly successful on the first iteration in correcting the inconsistency that existed between the initial estimate ozone profile and the true temperature profile. The effect of this disparity on the weighting functions and the first iteration result was large. In the middle and upper stratosphere, nearly all the retrieval information is obtained from the *a priori* data. In this case, the upper stratospheric ozone concentration profiles could not be retrieved from the synthetic radiance data. It must be pointed out that the upper stratospheric perturbations typical in this study - about 20-25% - are for mid-latitude winter conditions. Perturbations of this magnitude in the middle-upper stratosphere have little radiative impact at the surface. Higher latitude ozone profiles typically can have large (+100%) perturbations in the middle stratospheric regions due to tropospheric folding. The radiative signature of these perturbations will increase the radiative response in the more transparent regions of the ozone spectrum (980-1010, 1070-1080cm^{-1}) adding information about the upper stratospheric ozone distribution.

The sensitivity of this retrieval scheme to tropospheric ozone is far better than existing ground-based ozone profiling techniques. Umkehr ozone retrievals from Dobson spectrophotometer data are typically determined for 12 layers with the only about three layers encompassing the troposphere (depending on the meteorology). The other main instruments that measure ozone routinely are space-based and almost exclusively work in the UV region (e.g. SBUV). Remote sensing of ozone in the UV region provides reliable results at levels above the ozone maximum, but below this level results are uncertain even though there is general agreement with ozonesonde observations (DeLuisi, private communication, 1993). The use of ground-based thermal IR for tropospheric ozone concentration profile retrieval and UV ozone retrieval schemes for stratospheric ozone has the capacity to provide information on ozone concentrations for the entire atmospheric column. Coupled with the fact that continuous observations in the infrared are possible - as occurred with the AERI instrument during the FIRE II/SPECTRE experiment - ground-based ozone profiling can provide temporal evolution sequences of vertical ozone distributions in the troposphere.

A number of issues need to be addressed in more detail if we are to progress to operational retrievals of ozone using ground-based IR instruments. The main areas requiring further attention are

- Impact on the retrieval of errors in temperature and water vapour profiles,
- FASCOD3P transmittance model errors,
- *a priori* data dependence, and
- dependence of retrieval information on spectral resolution.

In calculating the synthetic observed ozone spectra, no analysis of the effect of uncertainties in the temperature or water vapour profiles were included. An error analysis of FASCOD3P model errors (e.g. line intensity and halfwidth uncertainties, continuum models) was also not included as all spectra used were calculated and radiance differences due to the factors mentioned above are consistently reproduced. An indication of the radiance differences in the 980-1080cm^{-1} which are observed between calculated and measured spectra are shown in figure 5.1. Shown are the differences between FASCOD3P and AERI unapodised spectra measured on November 18, 1991 during the FIRE II/SPECTRE experiment . The FASCOD3P spectra were calculated using input temperature and water vapour profiles from sondes released close to the time the radiometric observations were made. The ozone profile data was taken from climatology. Radiance differences of the order 1-2mW/m^2.sr.cm^{-1} are typical. The major differences are due to the overall background radiance level and water vapour lines.

The "baseline" level differences are due to both temperature and water vapour continuum differences but the largest contribution is from temperature. Small uncertainties in temperature cause large changes in both the Planck radiance and also in the magnitude of the water vapour lines. For example, a 1°C change in the tropospheric temperature profile produces about a 2mW/m^2.st.cm^{-1} change in the Planck radiance in the ozone longwave region at the surface, and up to a 5mW/m^2.sr.cm^{-1} increase in the strength of the water vapour present from 980-1080cm^{-1}. The high sensitivity of the retrieval scheme in the troposphere makes it imperative that the vertical profiles and errors of temperature and, to lesser extent, water vapour are as small as possible; also, they should be incorporated into the ozone retrieval scheme.

The impact of temperature uncertainties on initial estimate selection also needs to be investigated. In this study, the shape of the tropospheric initial estimate matched well with the shape of the true profile but not in its magnitude. This was offset by large differences in the lower stratospheric ozone concentrations. This radiative compensation (arising from different atmospheric regions) in the ozone spectrum may cause difficulty due to the naturally high temporal variability of ozone. A technique of inital estimate selection, which attempts to determine the best estimate through structure analysis (such as that detailed by Uddstrom and Wark (1985)), should be assessed.

The dependence of the retrieval scheme on *a priori* statistics because of the low information content of the 980-1080cm^{-1} ozone spectrum has been demonstrated. This dependence is, to some extent, a function of the number of layers used in the retrieval. The number of levels chosen in this study, 28, were selected from the 50 levels that the *a priori* data was defined at so that the pressure thickness of the layers in the troposphere were equal - to allow direct comparison of layer radiances - and also acceptable in terms of FASCODE execution. This gave a number of retrieval layers that was much larger than the number of independent pieces of information contained in the calculated spectra so that the ozone profile information in the synthetic observed spectra are distributed over a comparatively large number of levels. This, in effect, makes the inversion highly under-determined requiring stabilisation through *a priori* information. Reducing the number of levels to a value closer to that of the number of independent pieces of information in the measured spectra would decrease the dependence of the retrieval on the *a priori* data.

The amount of independent information available can be increased by higher spectral resolution measurements. As shown in Fleming and Barnes (1989), retrieval accuracy improves in proportion to the square root of the number of channels used -

the more channels used, the lower the retrieval error. However, for an interferometric instrument such as the AERI, the time required for an observation with a given noise level is inversely proportional to the square of the resolution (GPHIS Phase A Study, 1990). Thus a balance must be struck between information content and noise performance. A small increase in the amount of independent information by increasing the spectral resolution will be accompanied by an increase in the noise levels. Recent work by Clough et al. (1995) on ozone retrievals from simulated radiances observed from space has shown that, using realistic noise estimates, ozone retrieval errors decrease, to a limit, with higher spectral resolution.

The dependence of ozone retrievals on *a priori* statistics may be reduced by decomposing the weighting functions as detailed in Thompson (1992) for the temperature retrieval problem. The transformation operator, **W**, is recast using singular value decompostion. In this way the components which contribute the inherent instability in inversion schemes can be eliminated minimising the need for *a priori* data to contrain the solution - in effect a principal components type of analysis but on physical rather than statistical data. The applicability of this solution scheme to ozone retrievals needs to be investigated. This type of analysis also has importance for the trace gas species (such as CH_4 and CO) for which there is sparse or non-existant *a priori* vertical profile information. Clough et al. (1995), employing an iterative form of the minimum information method (similar to eqn. 2.13b), reduce the dependence of ozone retrievals on *a priori* information by using the variance of the *a priori* data to determine the value of the Lagrangian multiplier. This does not provide rigorous optimisation of each retrieval iteration, however it allows inclusion of information regarding the physical variance of ozone profile concentrations without strict assumptions about the nature of the a priori data, i.e. that the data is distributed normally and that the ensemble mean is representative of the profile being retrieved.

The ground-based ozone retrieval scheme presented in this study demonstrates the potential of ground-based IR monitoring of ozone, particularly in the troposphere. Improved knowledge of the vertical distribution of ozone and the dynamics of ozone transport on both a regional and global scale are needed for refining the parameterisation of climate models. Also, a much better understanding of the radiative balance of the atmosphere is required for global warming studies. The urgent need for more detailed tropospheric ozone profiles for climate model validation can be met by the retrieval scheme presented here using existing high spectral resolution instrumentation.

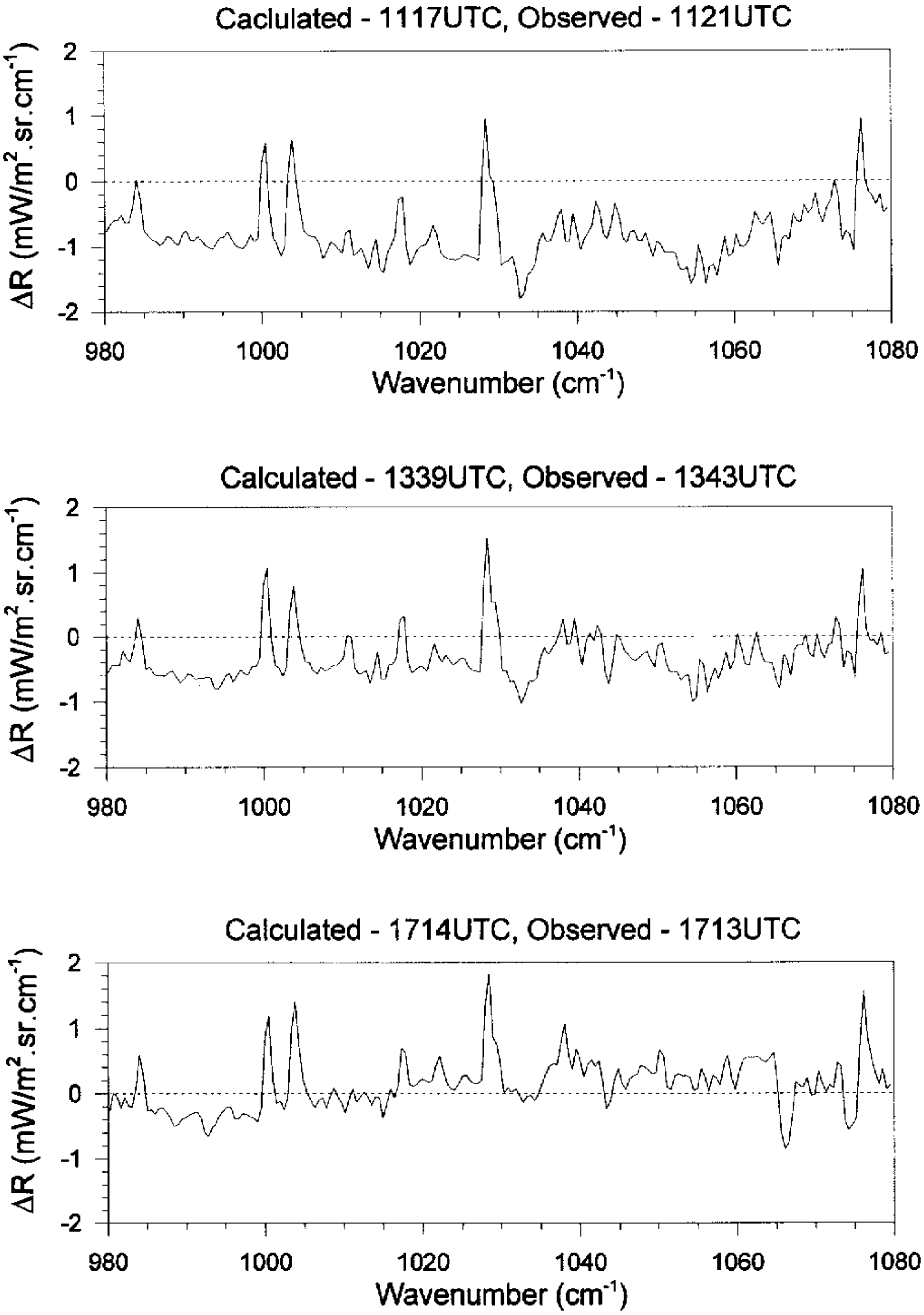


Figure 5.1 Unapodised difference spectra; FASCOD3P - AERI observation. AERI Band 1, SPECTRE Coffeyville 18 Nov. 1991. FASCOD3P calculation used temperature and water vapour profile data from CLASS sonde released at a time closest to AERI observation.

6. References

Alishouse, J.C., Crone, L.J., Fleming, H.E., van Cleef, F.L., and Wark, D.Q. (1967), A discussion of empirical orthogonal functions and their application to vertical temperature profiles, *Tellus*, **19**;477-481

Ben-David, A, Herman, B.M., and Reagan, J.A. (1985), Accuracy limitations in satellite temperature retrievals in the 14.0μm CO_2 band, Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and M.T. Chahine (eds.), A. Deepak Publishing, Hampton, Virginia, p233-245

Ben-David, A., Herman, B.M., and Reagan, J.A. (1988), Inverse problem and the pseudoempirical orthogonal function method of solution. 1: Theory, *Applied Optics*, **27**;1235-1242

Ben-David, A., Herman, B.M., and Reagan, J.A. (1988), Inverse problem and the pseudoempirical orthogonal function method of solution. 2: Use, *Applied Optics*, **27**;1243-1254

Brasseur, G., and Granier, C. (1992), Mount Pinatubo aerosols, chlorofluorocarbons, and ozone depletion, *Science*, **257**;1239-1241

Cess, R.D., Potter, G.L., Blanchet, J.P., Boer, G.J., Del Genio, A.D., Déqué, M., Dymnikov, V., Galin, V., Gates, W.L., Ghan, S.J., Kiehl, J.T., Lacis, A.A., Le Treut, H., Li, Z.-X., Liang, X.-Z., McAvaney, B.J., Meleshko, V.P., Mitchell, J.F.B., Morcrette, J.-J., Randall, D.A., Rikus, L., Roeckner, E., Royer, J.F., Schlese, U., Sheinin, D.A., Slingo, A., Sokolov, A.P., Taylor, K.E., Washington, W.M., Wetherald, R.T., Yagai, I., and Zhang, M.-H. (1990), Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models, *Journal of Geophysical Research*, **95**;16601-16615

Chahine, M.T. (1968), Determination of the temperature profile in an atmosphere from its outgoing radiance, *Journal of the Optical Society of America*, **58**;1634

Chahine, M.T. (1970), Inverse problems in radiative transfer: Determination of atmospheric parameters, *Journal of the Atmospheric Sciences*, **27**;960-967

Charlson, R.J., Schwartz, S.E., Hales, J.M., Cess, R.D., Coakley, J.A., Hansen, J.E., and Hofmann, D.J. (1992), Climate forcing by anthropogenic aerosols, *Science*, **255**;423-430

Chédin, A., and Scott, N.A. (1983), Improved Initialisation Inversion procedure (3I), Technical Proceedings of the First International TOVS Conference, Igls, Austria, p14-79

Chédin, A., Scott, N.A., Wahiche, C., and Moulinier, P. (1985), The Improved Initialisation Inversion method: A high resolution physical method for temperature retrieval from satellites of the TIROS-N series, *Journal of Climate and Applied Meteorology*, **24**;128-143

Chédin, A, Scott, N.A., Husson, N., Flobert, J.F., Levy, C., and Moine, P. (1988), Satellite meteorology and atmospheric spectroscopy: Recent progress in earth remote sensing from the satellites of the TIROS-N series, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **40**;257-273

CIMSS Proposal (1988a), High Resolution Interferometer Sounder (HIS) Phase II, Cooperative Institute for Meteorological Satellite Studies, SSEC, University of Wisconsin-Madison

CIMSS Proposal (1988b), Atmospheric Trace Gas Vertical Concentration Profiling - Source, Sink, and Flux Observations, Cooperative Institute for Meteorological Satellite Studies, SSEC, University of Wisconsin-Madison

Clark, W.C. ed. (1982), Carbon Dioxode Review : 1982, Clarendon Press, New York

Clough, S.A., Kneizys, F.X., Rothman, L.S., and Gallery, W.O. (1981), Atmospheric spectral transmittance and radiance: FASCOD1B, *SPIE*, **277**; 152-166

Clough, S.A., Kneizys, F.X., Shettle, E.P., and Anderson, G.P. (1986), Atmospheric radiance and transmittance: FASCOD2, Proceedings of the Sixth Conference on Atmospheric Radiation, Williamsburg, Virginia, p141-144

Clough, S.A., Worsham, R.D., Smith, W.L., Revercomb, H.E., Knuteson, R.O., Woolf, H.M., Anderson, G.P., Hoke, M.L., and Kneizys, F.X. (1988), Validation of FASCODE calculations with HIS spectral radiance measurements, International Radiation Symposium, Lille, France

Connor, B.J. and Rodgers, C.D. (1989), A comparison of retrieval methods: Optimal estimation, onion-peeling, and a combination of the two, RSRM'87: Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and J.S. Theon (eds.), A. Deepak Publishing, Hampton, Virginia, p189-202

Conrath, B.J. (1969), On the estimation of relative humidity profiles from medium-resolution infrared spectra obtained from a satellite, *Journal of Geophysical Research*, **74**;3347-3361

Conrath, B.J. (1972), Vertical resolution of temperature profiles obtained from remote radiation measurements, *Journal of the Atmospheric Sciences*, **29**;1262-1271

Deepak, A. ed. (1977), Inversion Methods in Atmospheric Remote Sounding, Academic Press, New York

Deepak, A. ed. (1985), Advances in Remote Sensing Retrieval Methods, A. Deepak Publishing, Hampton Virginia

DeLuisi, J.J., Longenecker, D.U., Mateer, C.L. and Wuebbles, D.J. (1989), An analysis of northern-latitude Umkehr measurements corrected for stratospheric aerosols for 1979-1986, *Journal of Geophysical Research*, **94**;9837-9846

Dobson, G.M.B (1973), The laminated structure of the ozone in the atmosphere, *Quarterly Journal of the Royal Meteorological Society*, **99**;599-607

Ellsaesser, H.W., MacCracken, M.C., Walton, J.J., and Grotch, S.L. (1986), Global climate trends as revealed by the recorded data, *Reviews of Geophysics*, **24**;745-792

Eyre, J.R. (1987), On systematic errors in satellite sounding products and their climatological mean values, *Quarterly Journal of the Royal Meteorological Society*, **113**;279-292

Eyre, J.R. (1989), Inversion of cloudy satellite sounding radiances by nonlinear optiml estimation. I: Theory and simulation for TOVS, *Quarterly Journal of the Royal Meteorological Society*, **115**;1001-1026

Eyre, J.R. (1990), The information content of data from satellite sounding systems: A simulation study, *Quarterly Journal of the Royal Meteorological Society*, **116**;401-434

Fleming, H.E., Crosby, D.S., and Neuendorffer, A.C. (1986), Correction of satellite temperature retrieval errors due to errors in atmospheric transmittances, *Journal of Climate and Applied Meteorology*, **25**;869-882

Fleming, H.E. and Barnes, L.L. (1989), Satellite tomographic remote sensing by frequency scanning, RSRM'87: Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and J.S. Theon (eds.), A. Deepak Publishing, Hampton, Virginia, p35-50

Flittner, D.E., Herman, B.M., Thome, K.J., Simpson, J.M., and Reagan, J.A. (1993), Total ozone and aerosol optical depths inferred from radiometric measurements in the Chappius absorption band, *Journal of the Atmospheric Sciences*, **50**;1113-1121

Forsythe, G.E., Malcolm, M.A., and Moler, C.B. (1977), Computer Methods for Mathematical Computations, Prentice Hall Inc., Eaglewood Cliffs, New Jersey

Foster, M. (1961), An application of Weiner-Kolmogorov smoothing theory to matrix inversion, *Journal of the Society of Industrial and Applied Mathematics*, **9**;387-392

Fuelberg, H.E., and Meyer, P.J. (1986), An analysis of mesoscale VAS retrievals using statistical structure functions, *Journal of Climate and Applied Meteorology*, **25**;59-76

Gille, J.C. and Bailey, P.L. (1989), Vertical resolution and error components of ozone retrievals from measurements of infrared limb emission, RSRM'87: Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and J.S. Theon (eds.), A. Deepak Publishing, Hampton, Virginia, p311-321

Hansen, J., Russell, G., Lacis, A.A., Fung, I., Rind, D., and Stone, P. (1985), Climate response times: Dependence on climate sensitivity and ocean mixing, *Science*, **229**;857-859

Hansen, J. and Lebedeff, S. (1987), Global trends of measured surface air temperature, *Journal of Geophysical Research*, **92**;13345-13372

Hansen, J., Fung, I, Lacis, A.A., Rind, D., Lebedeff, S., Ruedy, R., and Russell, G. (1988), Global climate changes as forecast by Goddard Institute for Space Studies three-dimensional model, *Journal of Geophysical Research*, **93**;9341-9364

Hansen, J., Lacis, A.A., and Prather, M. (1989), Greenhouse effect of chlorofluorocarbons and other trace gases, *Journal of Geophysical Research*, **94**;16417-16421

Hansen, J.E. and Lacis, A.A. (1990), Sun and dust versus greenhouse gases: an assessment of their relative roles in global climate change, *Nature*, **346**;713-719

Hayden, C.M., Smith, W.L., and Woolf, H.M. (1981), Determination of moisture from NOAA polar orbiting satellite sounding radiances, *Journal of Applied Meteorology*, **20**;450-466

Hayden, C.M. (1988), GOES-VAS similtaneous temperature-moisture retrieval algorithm, *Journal of Applied Meteorology*, **27**;705-733

Houghton, J.T., Jenkins, G.J., and Ephraums, J.J. eds. (1990), Climate change : The IPCC scientific assessment, Cambridge University Press, Cambridge

Huang, H-L., Smith, W.L., and Woolf, H.M. (1992), Vertical resolution and accuracy of atmospheric infrared sounding spectrometers, *Journal of Applied Meteorology*, **31**;265-274

International Radiation Commission (1988), Report on the Intercomparison of Transmittance and Radiance Algorithms, A. Chédin, H. Fischer, K. Kunzi, D. Spankuch, and N.A. Scott (eds.), Campaign and Workshop, University of Maryland, 12-14 March 1986

Idso, S.B., and Balling, R.C. (1991), Surface air temperature response to increasing global industrial productivity: A beneficial greenhouse effect?, *Theoretical and Applied Climatology*, **44**;37-41

Jackson, D.D., and Matsu'ura, M. (1985), A bayesian approach to nonlinear inversion, *Journal of Geophysical Research*, **90**;581-591

Jalickee, J.B. and Ropelewski, C.F. (1979), An objective analysis of the boundary-layer thermodynamic structure during GATE. Part I: Method, *Monthly Weather Review*, **107**;68-76

Jones, P.D., Wigley, T.M.L., and Wright, P.B. (1986), Global temperature variations between 1861 and 1984, *Nature*, **322**;430-434

Kaufman, Y.J., Fraser, R.S., and Mahoney, R.L. (1991), Fossil fuel and biomass burning effect on climate - heating or cooling?, *Journal of Climate*, **4**;578-588

Kiehl, J.T., and Solomon, S. (1986), On the radiative balance of the stratosphere, *Journal of the Atmospheric Sciences*, **43**;1525-1534

Kiehl, J.T., and Boville, B.T. (1988), The radiative-dynamical response of a stratospheric-tropospheric general circulation model to changes in ozone, *Journal of the Atmospheric Scinces*, **45**;1798-1817

King, M.D., Byrne, D.M., Herman, B.M., and Reagan, J.A. (1978), Aerosol size distribution obtained by inversion of spectral optical depth measurements, *Journal o fthe Atmospheric Sciences*, **35**;2153-2167

King, M.D. (1982), Sensitivity of constrained linear inversions to the selection of the Lagrange multiplier, *Journal of the Atmospheric Sciences*, **39**;1356-1369

Knuteson, R.O., Revercomb, H.E., Smith, W.L., Buijs, H., Dedecker, R., Howell, H.B., Woolf, H.M., and Ackerman, S.A. (1991), Realization of a low-cost ground-based temperature, humidity, and trace gas profiling system. Seventh Symposium on Meteorological Observations and Instrumentation, New Orleans, Louisiana, American Meteorological Society, Boston, Massachusetts, January 13-18

Lacis, A.A., Wuebbles, D.J., and Logan, J.A. (1990), radiative forcing of climate by changes in the vertical distribution of ozone, *Journal of Geophysical Research*, **95**;9971-9981

Lee, T., Chesters, D., and Mostek, A. (1983), The impact of conventional surface data upon VAS regression retrievals in the lower troposphere, *Journal of Climate and Applied Meteorology*, **22**;1853-1874

Lindzen, R.S. (1990), Some coolness concerning global warming, *Bulletin of the American Meteoroloical Society*, **71**;288-299

Lipton, A.E., Hillger, D.W., and Vonder Haar, T.H. (1986), Water vapour vertical profile structures retrieved from satellite data via classification and discrimination, *Monthly Weather Review*, **114**;1103-1111

Lipton, A.L., and Vonder Haar, T.H. (1987), Retrieval of water vapour profiles via principal components: Options and their implications, *Journal of Climate and Applied Meteorology*, **26**;1038-1042

Mateer, C.L. (1965), On the information content of Umkehr observations, *Journal of the Atmospheric Sciences*, **22**;370-381

McMillin, L.M. (1991), Evaluation of a classification method for retrieving atmospheric temperatures from satellite measurements, *Journal of Applied Meteorology*, **30**;432-446

Michaels, P.J., and Stooksbury, D.E. (1992), Global warming: A reduced threat?, *Bulletin of the American Meteorological Society*, **73**;1563-1577

Mitchell, J.F.B., Senior, C.A., and Ingram, W.J. (1989), CO_2 and climate: a missing feedback?, *Nature*, **341**;132-134

Neuendorffer, A.C. (1989), The feasibility of 1-deg/1-km tropospheric retrievals, RSRM'87: Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and J.S. Theon (eds.), A. Deepak Publishing, Hampton, Virginia, p323-332

National Oceanic and Atmospheric Administration, United States Committee on Extension to the Standard Atmosphere (1976), U.S. Standard Atmosphere, Washington D.C., NOAA

Phillips, D.L. (1962), A technique for the numerical solution of certain integral equations of the first kind, *Journal of the Association for Computing Machinery*, **9**;84-97

Purser, R.J., and Huang, H-L. (1993), Estimating effective data density in a satellite retrieval or an objective analysis, *Journal of Applied Meteorology*, **32**;1092-1107

Ramanathan, V. (1975), Greenhouse effect due to chlorofluorocarbons : Climatic implication, *Science*, **190**;50-52

Ramanathan, V., Cicerone, R.J., Singh, H.B., and Kiehl, J.T. (1985), Trace gas trends and their potential role in climate change, *Journal of Geophysical Research*, **90**;5547-5566

Ramanathan, V., Callis, L., Cess, R.D., Hansen, J., Isaken, I., Kuhn, W., Lacis, A.A., Luther, F., Mahlman, J., Reck, R., and Schlesinger, M. (1987), Climate-chemical interactions and effects of changing atmospheric gases, *Reviews of Geophysics*, **25**;1441-1482

Ramanathan, V., Barkstrom, B.B., and Harrison, E.F. (1989), Climate and the Earth's radiation budget, *Physics Today*, **42**;22-32

Ramaswamy, V., Schwarzkopf, M.D., and Shine, K.P (1992), Radiative forcing of climate from halocarbon-induced global stratospheric ozone loss, *Nature*, **355**;810-812

Reid, S.J., and Vaughan, G. (1991), Lamination in ozone profiles in the lower stratosphere, *Quarterly Journal of te Royal Meteorological Society*, **117**;825-844

Revercomb, H.E., Buijs, H., Howell, H.B., LaPorte, D.D., Smith, W.L., and Sromovsky, L.A. (1988), Radiometric calibration of IR Fourier transform spectrometers: Solution to a problem with the High-resolution Interferometer Sounder, *Applied Optics*, **27**;3210-3218

Revercomb, H.E., Buijs, H., Howell, H.B., Knuteson, R.O., LaPorte, D.D., Smith, W.L., Sromovsky, L.A., and Woolf, H.M. (1989a), Radiometric calibration of IR interferometers: Experience from the High-resolution Interferometer

Sounder (HIS) aircraft instrument, RSRM'87: Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and J.S. Theon (eds.), A. Deepak Publishing, Hampton, Virginia, p89-101

Revercomb, H.E., Smith, W.L., Sromovsky, L.A., Knuteson, R.O., Buijs, H., LaPorte, D.D., and Howell, H.B. (1989b), Radiometrically accurate FTS for atmospheric emission observations, *SPIE*, **1145**;70-78

Revercomb, H.E., Smith, W.L., Knuteson, R.O., Woolf, H.M., and Howell, H.B. (1989c), Comparisons of FASCODE spectra with HIS observations, Proceedings of the Annual Review Conference on Atmospheric Transmission Models, AFGL, 6-7 June 1989

Revercomb, H.E., Knuteson, R.O., Smith, W.L., Woolf, H.M., and Howell, H.B. (1990), Spectroscopic inferences from HIS measurements of atmospheric thermal emission. Optical Remote Sensing of the Atmosphere, Incline Village, Nevada, 12-15 February 1990

Revercomb, H.E., Smith, W.L., Knuteson, R.O., Best, F.A., LaPorte, D.D., Silverman, S.H., and Chase, S.H. (1991a), The NASA Geo-Platform High-resolution Interferometer Sounder (GP-HIS): Advanced instrumentation for meteorology and atmospheric chemistry. Seventh Symposium on Meteorological Observations and Instrumentation, New Orleans, Louisiana, American Meteorological Society, Boston, Massachusetts, 13-18 January 1991

Revercomb, H.E., Smith, W.L., Knuteson, R.O., Woolf, H.M., and Howell, H.B. (1991b), Comparisons of FASOCDE spectra with HIS observations II.

Proceedings of the AFGL Annual Review Conference on Atmospheric Transmission Models, AFGL, 11-12 June 1991

Rind, D., Chiou, E.-W., Chu, W., Larsen, J., Oltmans, S., Lerner, J., McCormick, M.P., and McMaster, L. (1991), Positive water vapour feedback in climate models confirmed by satellite data, *Nature*, **349**;500-503

Rodgers, C.D. (1971), Some theoretical aspects of remote sounding in the Earth's atmosphere, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **11**;767-777

Rodgers, C.D. (1976a), Retrieval of atmospheric temperature and composition from remote measurement of thermal radiation, *Review of Geophysics and Space Physics*, **14**;609-624

Rodgers, C.D. (1976b), The vertical resolution of remotely sensed temperature profiles with a priori statistics, *Journal of the Atmospheric Sciences*, **33**;707-709

Rodgers, C.D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements, *Journal of Geophysical Research*, **95**;5587-5595

Rothman, L.S., Gamache, R.R., Goldman, A., Brown, L.R., Toht, R.A., Pickett, H.M., Poynter, R.L., Flaud, J.M., Camy-Peyret, C., Barbe, A., Husson, N., Rinsland, C.P., and Smith, M.A.H. (1987), The HITRAN database: 1986 edition, *Applied Optics*, **26**;4058-4097

Rothman, L.S. (1988), The spectroscopic database for high resolution atmospheric modeling, *SPIE*, **928**;79-93

Rothman, L.S., Gamache, R.R., Tipping, R.H., Rinsland, C.P., Smith, M.A.H., Benner, D.C., Devi, V.M., Flaud, J.-M., Cmay-Peyret, C., Perrin, A., Massie, S.T., Brown, L.R., and Toth, R.A. (1992), The HITRAN molecular database: Editions of 1991 and 1992, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **48**;469-507

Scott, N.A., Chédin, A., Achard, V., Bonnet, B., Cheruy, F., Claud, C., Escobar., J., Husson, N., Rieu, H., Tahani, Y., and Tournier, B. (1991), Recent advances in the 3D thermodynamic analysis of the earth system through the "3I" algorithm. Extension to the second generation vertical sounders., Proceedings of the Sixth International TOVS Study Conference, May 1-6, Airlie, Virginia, p425-467

Smith, W.L. (1970), Iterative solution of the radiative transfer equation for the temperature and absorbing gas profile of an atmosphere, *Applied Optics*, **9**;1993-1999

Smith, W.L., Woolf, H.M., and Jacob, W.J. (1970), A regression method for obtaining real-time temperature and geopotential height profiles from satellite spectrometer measurements and its application to Nimbus-3 SIRS observations, *Monthly Weather Review*, **98**;582-603

Smith, W.L., and Howell, H.B. (1971), Vertical distributions of atmospheric water vapour from satellite infrared spectrometer measurements, *Journal of Applied Meteorology*, **10**;1026-1034

Smith, W.L., Woolf, H.M. and Fleming, H.E. (1972), Retrieval of atmsopheric temperature profiles from satellite measurements for dynamical forecasting, *Journal of Applied Meteorology*, **11**;113-122

Smith, W.L., and Woolf, H.M. (1976), The use of eigenvectors of statistical covariance matrices for interpreting satellite sounder radiometer observations, *Journal of the Atmospheric Sciences*, **33**;1127-1140

Smith, W.L., Woolf, H.M., Hayden, C.M., Wark, D.Q., and McMillan, L.M. (1979), TIROS-N operational vertical sounder, *Bulletin of the Americam Meteorological Society*, **60**;1177-1187

Smith, W.L., Howell, H.B., and Woolf, H.M. (1979), The use of interferometric radiance measurements for sounding the atmosphere, *Journal of the Atmospheric Sciences*, **36**;566-575

Smith, W.L. (1983a), Passive radiometry for vertical sounding from meteorological satellites, *Applied Optics*, **22**;2641-2643

Smith, W.L. (1983b), The retrieval of atmospheric profiles from VAS geostationary radiance observations, *Journal of the Atmospheric Sciences*, **40**;2025-2035

Smith, W.L, Revercomb, H.E., Howell, H.B., and Woolf, H.M. (1983), HIS - a satellite instrument to observer temperature and moisture profiles with high vertical resolution, Proceedings of the Fifth Conference on Atmospheric Radiation, Baltimore, Maryland, p1-9

Smith, W.L., and Woolf, H.M. (1984), Improved vertical soundings from an amalgamation of polar and geostationary radiance observations, Proceedings of the Conference on Satellite/Remote Sensing and Applications, Clearwater Beach, Florida, p45-48

Smith, W.L., Woolf, H.M., and Schreiner, A.J. (1985), Similtaneous retrieval of surface atmospheric parameters: a physical and analytically direct approach, Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and M.T. Chahine (eds.), A. Deepak Publishing, Hampton, Virginia, p221-230

Smith, W.L., Woolf, H.M., Howell, H.B., Huang, H.-L., and Revercomb, H.E. (1989), The similtaneous retrieval of atmospheric temperature and water vapour profiles - application to measurements with the high spectral resolution interferometer sounder (HIS), RSRM'87: Advances in Remote Sensing Retrieval Methods, A. Deepak, H.E. Fleming, and J.S. Theon (eds.), A. Deepak Publishing, Hampton, Virginia, p189-202

Smith, W.L., Revercomb, H.E., Howell, H.B., Woolf, H.M., Knuteson, R.O., Decker, R.G., Lynch, M.J., Westwater, E.R., Strauch, R.G., Moran, K.P., Stankov, B., Falls, M.J., Jordan, J., Dabbert, W.F., McBeth, R., Albright, G., Paneitz, C., Wright, G., May, P.T., and Decker, M.T. (1990), GAPEX: A ground based atmospheric profiling experiment, *Bulletin of the American Meteorological Society*, **71**;310-318

Smith, W.L., H.M. Woolf, and H.E. Revercomb (1991), Linear simultaneous solution for temperature and absorbing constituent profiles from radiance spectra, *Applied Optics*, **30**;1117-1123

Stewart, K., Carli, B., Chédin, A., and Fischer, H. (1987), Atmospheric Sounding and Trace Gas Measurements with Interferometers from Geostationary Satellites, ESA Report

Strand, O.N., and Westwater, E.R. (1968), Statistical estimation of the numerical solution of a Fredholm integral equation of the first kind, *Journal of the Association for Computing Machinery*, **15**;100-114

Susskind, J., and Searl, J.E. (1978), Synthetic atmospheric transmittance spectra near 15- and 4.3μm, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **19**;195-215

Susskind, J. and Rosenberg, A. (1980), Temperature retrievals from TIROS-N, Remote Sensing of Atmospheres and Oceans, A. Deepak (ed), Academic Press, New York, p45-64

Susskind, J., Rosenfield, J., and Reuter, D. (1983), An accurate radiative transfer model for use in the direct physical inverison of HIRS2 and MSU temperature sounding data, *Journal of Geophysical Research*, **88**;8550-8568

Susskind, J., Rosenfield, J., Reuter, D., and Chahine, M.T. (1984), Remote sensing of weather and climate parameters from HIRS2/MSU on TIROS-N, *Journal of Geophysical Research*, **89**;4677-4697

Tarantola, A. and Valette, B. (1982), Generalised nonlinear inverse problems solved using the least-squares criterion, *Reviews of Geophysics and Space Physics*, **20**;219-232

Tarantola, A. (1987), Inverse problem theory: Methods for data fitting and model parameter estimation, Elsevier Science, New York

Thompson, O.E. (1982), HIRS-AMTS satellite sounding system test - theoretical and empirical resolving power, *Journal of Applied Meteorology*, **21**;1550-1561

Thompson, O.E., Goldberg, M.D., and Dazlich, D.A. (1985), Pattern recognition in the satellite temperature retrieval problem, *Journal of Climate and Applied Meteorology*, **24**;30-48

Thompson, O.E. (1992), Regularizing the satellite temperature-retrieval problem through singular-value decomposition of the radiative transfer physics, *Monthly Weather Review*, **120**;2314-2328

Twomey, S. (1977), Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements, Elsevier Scientific Publishing, Amsterdam

Twomey, S. (1963), On the numerical solution of Fredholm integral eqations of the first kind by the inversion of the linear system produced by quadrature, *Journal of the Association for Computing Machinery*, **10**;97-101

Twomey, S. (1965), The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements, *Journal of the Franklin Institute*, **279**;95-109

Twomey, S. (1966), Indirect measurements of atmospheric temperature profiles from satellites: II. Mathematical aspects of the inversion problem, *Monthly Weather Review*, **94**;363

Uddstrom, M.J. and Wark, D.Q. (1985), A classification system for satellite temperature retrievals, *Journal of Climate and Applied Meteorology*, **24**;16-29

Walpole, R.E. and Myers, R.H. (1993), Probability and Statistics for Engineers and Scientists, 5th. ed., New York, MacMillan

Wang, W-C., Wuebbles, D.J., Washington, W.M., Isaacs, R.G., and Molnar, G. (1986), Trace gases and other potential perturbations to global climate, *Review of Geophysics*, **24**;110-140

Wang, W.-C., Dudek, M.P., Liang, X.-Z., and Kiehl, J.T. (1991a), Inadequacy of effective CO_2 as a proxy in simulating the greenhouse effect of other radiatively active gases, *Nature*, **350**;573-577

Wang, W.-C., Shi, G.-Y., and Kiehl, J.T. (1991b), Incorporation of the thermal radiative effect of CH_4, N_2O, CF_2Cl_2, and $CFCL_3$ into the National Center for Atmospheric Research Community Climate Model, *Journal of Geophysical Research*, **96**;9097-9103

Wark, D.Q. and Fleming, H.E. (1966), Indirect measurementsof atmospheric temperature profiles from satellites: I. Introduction, *Monthly Weather Review*, **94**;351-362

Washington, W.M., and Meehl, G.A. (1989), Climate sensitivity due to increased CO_2: experiments with a coupled atmosphere and ocean general circulation model, *Climate Dynamics*, **4**;1-38

Weinreb, M.P., and Crosby, D.S. (1972), Optimisation of spectral intervals for remote sensing of atmospheric temperature profiles, *Remote Sensing Environment*, **2**;193-201

Wetherald, R.T. and Manabe, S. (1988), Cloud feedback processes in a general-circulation model, *Journal of the Atmospheric Sciences*, **45**;1397-1415

Wigley, T.M.L. (1987), Relative contributions of different trace gases to the greenhouse effect, *Climate Monitor*, **16**;14-28

Wigley, T.M.L. (1991), Could reducing fossil-fuel emissions cause global warming?, *Nature*, **349**;503-506

WMO (1990), Report of the International Ozone Trends Panel 1988, World Meteorological Organisation Global Ozone Research and Monitoring Project - Report #18, WMO, Washinton, DC

WMO (1992), Scientic Assessment of Ozone Depletion: 1991, World Meteorological Organisation Global Ozone Research and Monitoring Project - Report #25, WMO, Washington, DC

Wuebbles, D.J., Luther, F.M., and Penner, J.E. (1983), Effect of coupled anthropogenic perturbations on stratospheric ozone, *Journal of Geophysical Research*, **88**;1444-1456

Appendices

A : Derivation of iteration equation for maximum likelihood inversion scheme

We have $\mathbf{x}_{n+1} = \mathbf{x}_n - \left(\frac{\partial^2 S}{\partial\mathbf{x}\partial\mathbf{x}}\right)_{\mathbf{x}_n}^{-1} \left(\frac{\partial S}{\partial\mathbf{x}}\right)_{\mathbf{x}_n}$ (A.1)

where $\left(\frac{\partial S}{\partial\mathbf{x}}\right)_{\mathbf{x}_n} = \mathbf{C}_X^{-1}\left(\mathbf{x}-\mathbf{x}^o\right) - \mathbf{W}_n^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right)$ (A.2)

and $\left(\frac{\partial^2 S}{\partial\mathbf{x}\partial\mathbf{x}}\right)_{\mathbf{x}_n}^{-1} = \left(\mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n + \mathbf{C}_X^{-1}\right)^{-1}$ (A.3)

Replacing (A.2) and (A.3) into (A.1) gives

$$
\begin{aligned}
\mathbf{x}_{n+1} &= \mathbf{x}_n - \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\left[\mathbf{C}_X^{-1}\left(\mathbf{x}_n-\mathbf{x}^o\right) - \mathbf{W}_n^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right)\right] \\
&= \mathbf{x}_n - \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{C}_X^{-1}\left(\mathbf{x}_n-\mathbf{x}^o\right) \\
&\qquad + \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right) \\
&= \mathbf{x}_n - \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{C}_X^{-1}\left(\mathbf{x}_n-\mathbf{x}^o\right) \\
&\qquad - \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\left(\mathbf{x}_n-\mathbf{x}^o\right) \\
&\qquad + \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\left(\mathbf{x}_n-\mathbf{x}^o\right) \\
&\qquad + \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right) \\
&= \mathbf{x}_n - \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)\left(\mathbf{x}_n-\mathbf{x}^o\right) \\
&\qquad + \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\left(\mathbf{x}_n-\mathbf{x}^o\right) \\
&\qquad + \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right) \\
&= \mathbf{x}_n - \left(\mathbf{x}_n-\mathbf{x}^o\right) + \left(\mathbf{C}_X^{-1} + \mathbf{W}_n^T\mathbf{C}_R^{-1}\mathbf{W}_n\right)^{-1}\mathbf{W}_n^T\mathbf{C}_R^{-1}\left[\mathbf{W}_n\left(\mathbf{x}_n-\mathbf{x}^o\right) + \left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right)\right] \\
&= \mathbf{x}_n + \left(\mathbf{x}^o-\mathbf{x}_n\right) + \mathbf{C}_X\mathbf{W}_n^T\left(\mathbf{W}_n^T\mathbf{C}_X\mathbf{W}_n + \mathbf{C}_R\right)^{-1}\left[\mathbf{W}_n\left(\mathbf{x}_n-\mathbf{x}^o\right) + \left(\mathbf{r}_{obs}-\mathbf{r}(\mathbf{x}_n)\right)\right]
\end{aligned}
$$

(A.4)

B : Manipulation of retrieval error covariance operator

Given $\mathbf{C}_E = \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}$ (B.1)

and the matrix identity $\left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1} = \mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T + \mathbf{C}_R\right)^{-1}$ (B.2)

$$\begin{aligned}
\mathbf{C}_E &= \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1} \\
&= \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\left[\left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)\mathbf{C}_X - \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\mathbf{C}_X\right] \\
&= \mathbf{C}_X - \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\mathbf{C}_X \\
&= \mathbf{C}_X - \mathbf{C}_X\mathbf{W}^T\left(\mathbf{W}\mathbf{C}_X\mathbf{W}^T + \mathbf{C}_R\right)^{-1}\mathbf{W}\mathbf{C}_X
\end{aligned} \qquad \text{(B.3)}$$

C : Derivation of retrieval error covariance operator described by Huang et al. (1992) and Rodgers (1990)

The linear one step maximum likelihood solution to

$$\Delta\mathbf{r} = \mathbf{W}\Delta\mathbf{x} \tag{C.1}$$

is given by

$$\begin{aligned}\Delta\hat{\mathbf{x}} &= \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}(\Delta\mathbf{r} + \varepsilon) \\ &= \mathbf{G}(\Delta\mathbf{r} + \varepsilon) \\ &= \mathbf{G}\Delta\mathbf{r} + \mathbf{G}\varepsilon \\ &= \mathbf{R}\Delta\mathbf{x} + \mathbf{G}\varepsilon\end{aligned} \tag{C.2}$$

where **R** = **GW**, the resolution operator, and the measurement error, ε, is stated explicitly. The definition of error covariance gives

$$\mathbf{C}_E = \left\langle(\Delta\mathbf{x} - \Delta\hat{\mathbf{x}})(\Delta\mathbf{x} - \Delta\hat{\mathbf{x}})^T\right\rangle \tag{C.3}$$

Substituting the expressions (C.1) and (C.2) into (C.3)

$$\begin{aligned}\mathbf{C}_E &= (\Delta\mathbf{x} - \mathbf{R}\Delta\mathbf{x} - \mathbf{G}\varepsilon)(\Delta\mathbf{x} - \mathbf{R}\Delta\mathbf{x} - \mathbf{G}\varepsilon)^T \\ &= [(\mathbf{I} - \mathbf{R})\Delta\mathbf{x} - \mathbf{G}\varepsilon][(\mathbf{I} - \mathbf{R})\Delta\mathbf{x} - \mathbf{G}\varepsilon]^T \\ &= (\mathbf{I} - \mathbf{R})\Delta\mathbf{x}\Delta\mathbf{x}^T(\mathbf{I} - \mathbf{R})^T - (\mathbf{I} - \mathbf{R})\Delta\mathbf{x}\varepsilon^T\mathbf{G}^T - \mathbf{G}\varepsilon\Delta\mathbf{x}^T(\mathbf{I} - \mathbf{R})^T + \mathbf{G}\varepsilon\varepsilon^T\mathbf{G}^T\end{aligned}$$

The assumption is of uncorrelated errors so that $\langle\Delta\mathbf{x}\varepsilon^T\rangle = \langle\varepsilon\Delta\mathbf{x}^T\rangle = \mathbf{0}$, so that

$$\begin{aligned}\mathbf{C}_E &= (\mathbf{I} - \mathbf{R})\Delta\mathbf{x}\Delta\mathbf{x}^T(\mathbf{I} - \mathbf{R})^T + \mathbf{G}\varepsilon\varepsilon^T\mathbf{G}^T \\ &= (\mathbf{I} - \mathbf{R})\mathbf{C}_X(\mathbf{I} - \mathbf{R})^T + \mathbf{G}\mathbf{C}_R\mathbf{G}^T\end{aligned} \tag{C.4}$$

The equivalence of (C.4) with

$$\mathbf{C}_E = \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1} \tag{C.5}$$

is shown by using the relation,

$$\mathbf{I} = \mathbf{R} + \mathbf{C}_E\mathbf{C}_X^{-1} \text{ or } \mathbf{I} - \mathbf{R} = \mathbf{C}_E\mathbf{C}_X^{-1} \tag{C.6}$$

and assuming (C.5) is true. Substituting (C.6) into (C.4) and expanding the operator **G**,

$$\begin{aligned}
\mathbf{C}_E &= \mathbf{C}_E\mathbf{C}_X^{-1}\mathbf{C}_X\left(\mathbf{C}_E\mathbf{C}_X^{-1}\right)^T + \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{C}_R\left[\left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\right]^T \\
&= \mathbf{C}_E\mathbf{C}_X^{-1}\mathbf{C}_E + \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1} \\
&= \mathbf{C}_E\mathbf{C}_X^{-1}\mathbf{C}_E + \mathbf{C}_E\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\mathbf{C}_E \\
&= \left(\mathbf{C}_E\mathbf{C}_X^{-1} + \mathbf{C}_E\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)\mathbf{C}_E
\end{aligned}$$

therefore

$$\begin{aligned}
\mathbf{I} &= \mathbf{C}_E\mathbf{C}_X^{-1} + \mathbf{C}_E\mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W} \\
&= \mathbf{C}_E\left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)
\end{aligned}$$

and thus

$$\mathbf{C}_E = \left(\mathbf{C}_X^{-1} + \mathbf{W}^T\mathbf{C}_R^{-1}\mathbf{W}\right)^{-1}$$

D : Derivation of PRTE

The RTE for downwelling atmospheric infrared radiation can be written as,

$$R_\nu = \int_0^{p_s} B_\nu(T(p)) \frac{d\tau_\nu(p)}{dp} dp . \tag{D.1}$$

The radiance perturbation is described by,

$$\delta R_\nu = R_\nu - R_\nu^o$$

so that one obtains,

$$\delta R_\nu = \int_0^{p_s} B_\nu(T(p)) \frac{d\tau_\nu(p)}{dp} dp - \int_0^{p_s} B_\nu\left(T^o(p)\right) \frac{d\tau_\nu^o(p)}{dp} dp \tag{D.2}$$

Equation (D.2) can be simplified by expanding perturbation terms, giving,

$$\begin{aligned} \delta R_\nu &= \int_0^{p_s} B_\nu\left(T(p)\right) \frac{d\left[\tau_\nu^o(p) + \delta\tau_\nu(p)\right]}{dp} dp - \int_0^{p_s} \left[B_\nu\left(T(p)\right) - \delta B_\nu\left(T(p)\right)\right] \frac{d\tau_\nu^o(p)}{dp} dp \\ &= \int_0^{p_s} B_\nu\left(T(p)\right) \frac{d\left[\delta\tau_\nu(p)\right]}{dp} dp + \int_0^{p_s} \delta B_\nu\left(T(p)\right) \frac{d\tau_\nu^o(p)}{dp} dp \end{aligned} \tag{D.3}$$

The first term of (D.3) is simplified by integration by parts. Defining

$u = B_\nu\left(T(p)\right) \Rightarrow du = dB_\nu\left(T(p)\right)$, and

$d\upsilon = d\left[\delta\tau_\nu(p)\right] \Rightarrow \upsilon = \delta\tau_\nu(p)$

and using the integration by parts formulation

$$\int_a^b u\,dv = \left[uv\right]_a^b - \int_a^b v\,du\,,$$

then (D.3) becomes, with rearrangement of terms,

$$\begin{aligned}\delta R_\nu = {} & \int_0^{p_s} \delta B_\nu\big(T(p)\big)\frac{d\tau_\nu^o(p)}{dp}\,dp - \int_0^{p_s} \delta\tau_\nu(p)\frac{dB_\nu\big(T(p)\big)}{dp}\,dp \\ & + \Big[B_\nu\big(T(p)\big)\delta\tau_\nu(p)\Big]_0^{p_s}\end{aligned} \quad . \tag{D.4}$$

E : Integration by parts of the PRTE

Given

$$\delta R_\nu(\delta q_i, p_s) = \int_{p_{top}}^{p_s} \frac{dB_\nu(T(p))}{dp}\left[\frac{\tau_\nu^o(p)}{g}\int_p^{p_s} k_{\nu i}(p')\delta q_i(p')dp'\right]dp + \frac{B_\nu(T(p_{top}))\tau_\nu^o(p_{top})}{g}\int_{p_{top}}^{p_s} k_{\nu i}(p')\delta q_i(p')dp' \tag{E.1}$$

the first term can be simplified by integration by parts. Defining

$$u = \frac{1}{g}\int_p^{p_s} k_{\nu i}(p')\delta q_i(p')dp'$$

then

$$\frac{du}{dp} = \left(\frac{1}{g}\right)\frac{d}{dp}\int_p^{p_s} k_{\nu i}(p')\delta q_i(p')dp' = \left(\frac{1}{g}\right)(-1)k_{\nu i}(p)\delta q_i(p) \tag{E.2}$$

therefore

$$du = -\frac{1}{g}k_{\nu i}(p)\delta q_i(p)dp \tag{E.3}$$

(E.2) is obtained using the Fundamental Theorem of Calculus which states if

$$g(x) = \int_x^a f(t)dt$$

then

$$g'(x) = \frac{d}{dx}\int_{x}^{a} f(t)dt = -f(x)$$

Similarly, let

$$d\upsilon = \tau_{\nu}^{o}(p)\frac{dB_{\nu}(T(p))}{dp}dp \Rightarrow \upsilon = \int_{p_{top}}^{p} \tau_{\nu}^{o}(p')\frac{dB_{\nu}(T(p'))}{dp'}dp' \qquad \text{(E.4)}$$

Integration by parts takes the form

$$\int_{a}^{b} u\,d\upsilon = [u\upsilon]_{a}^{b} - \int_{a}^{b} \upsilon\,du$$

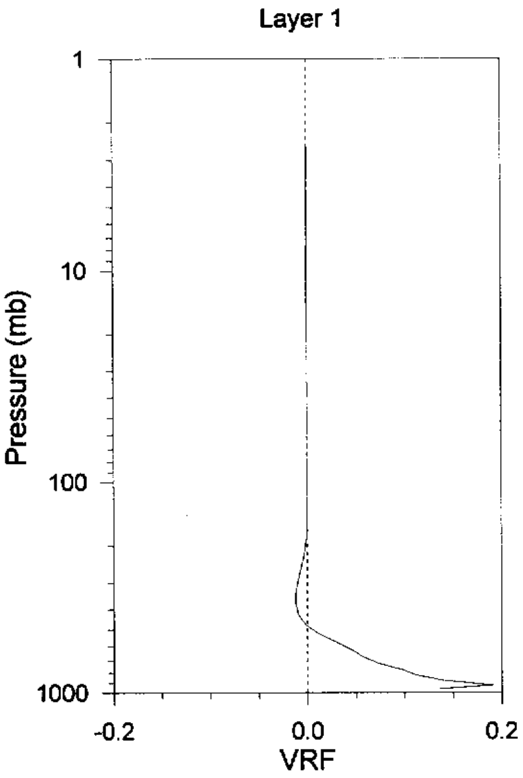
so (E.1) becomes

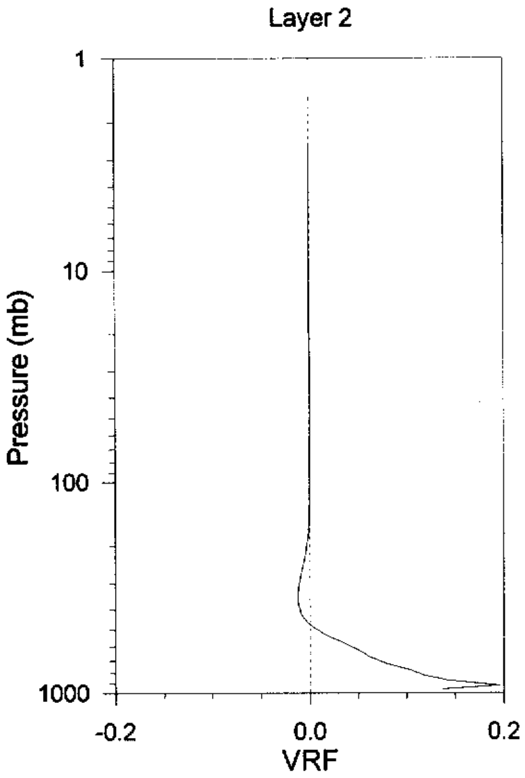
$$\delta R_{\nu}(\delta q_i, p_s) = \left[\frac{1}{g}\int_{p}^{p_s} k_{\nu i}(p')\delta q_i(p')dp' \int_{p_{top}}^{p} \tau_{\nu}^{o}(p')\frac{dB_{\nu}(T(p'))}{dp'}dp'\right]_{p_{top}}^{p_s}$$
$$+\frac{1}{g}\int_{p_{top}}^{p_s} k_{\nu i}(p)\delta q_i(p)\left[\int_{p_{top}}^{p} \tau_{\nu}^{o}(p')\frac{dB_{\nu}(T(p'))}{dp'}dp'\right]dp$$

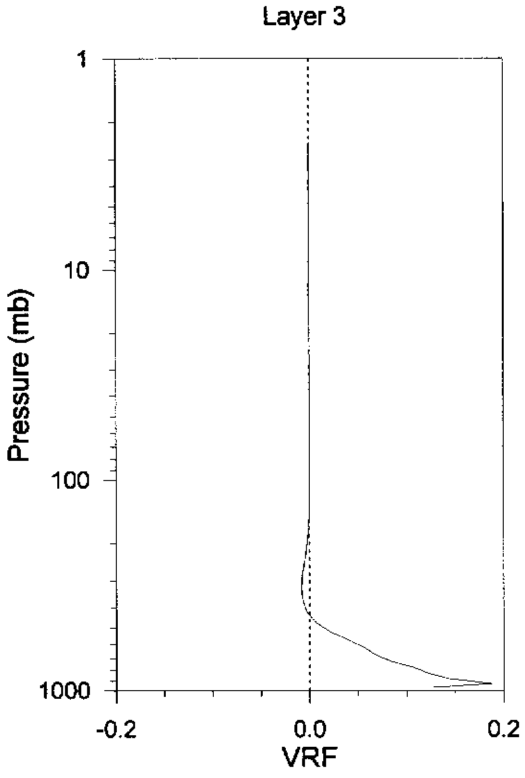
where the first term is necessarily zero giving

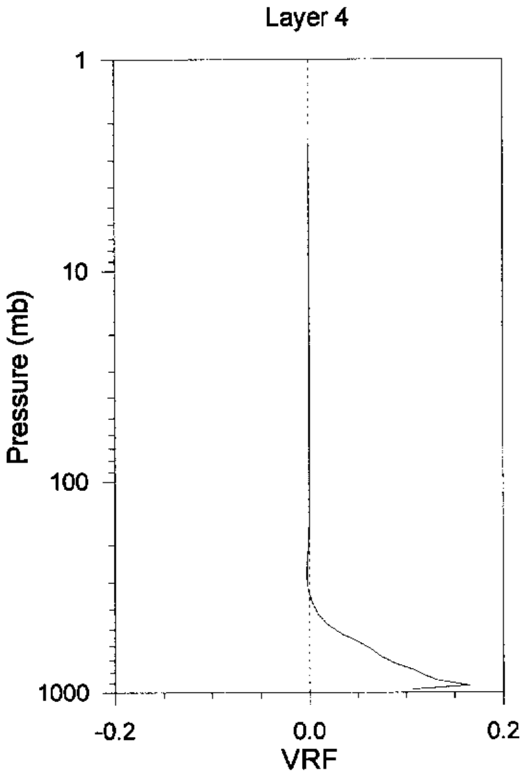
$$\boxed{\delta R_{\nu}(\delta q_i, p_s) = \frac{1}{g}\int_{p_{top}}^{p_s} k_{\nu i}(p)\delta q_i(p)\left[\int_{p_{top}}^{p} \tau_{\nu}^{o}(p')\frac{dB_{\nu}(T(p'))}{dp'}dp'\right]dp \qquad \text{(E.5)}}$$

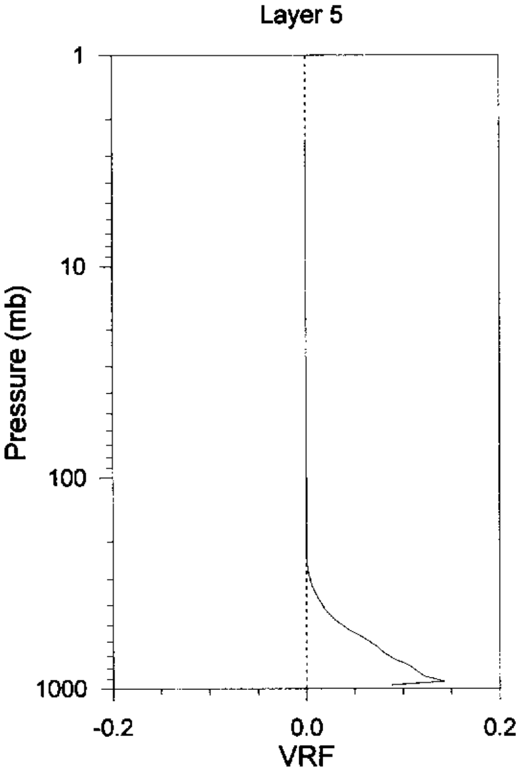
F : Ozone vertical resolution functions by layer **(Note scale changes throughout)**

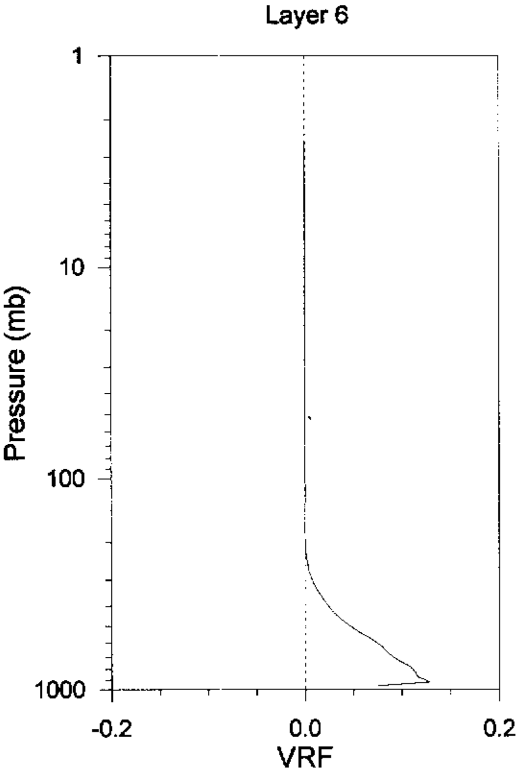

Layer 1

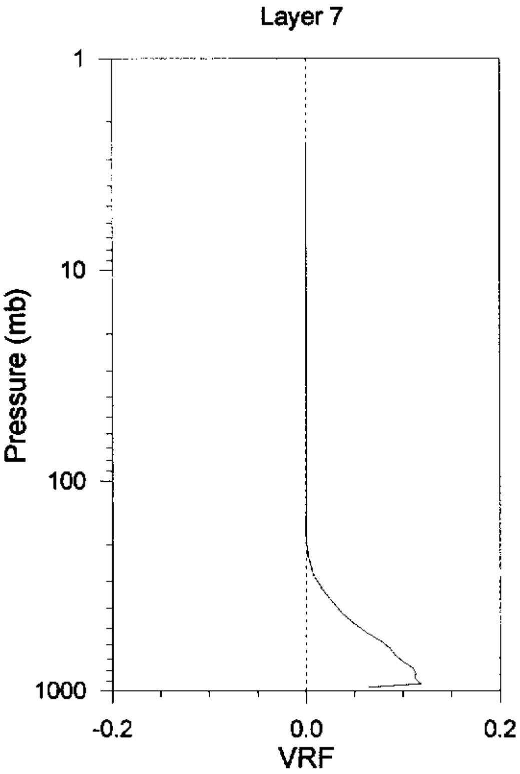

Layer 2

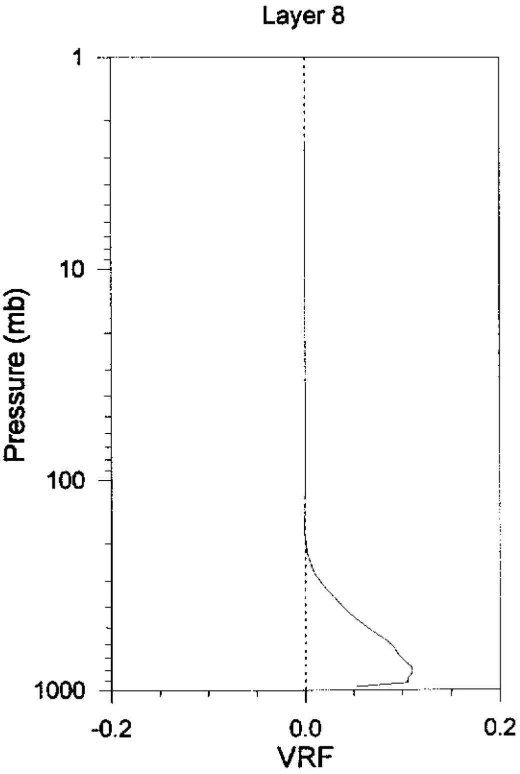

Layer 3

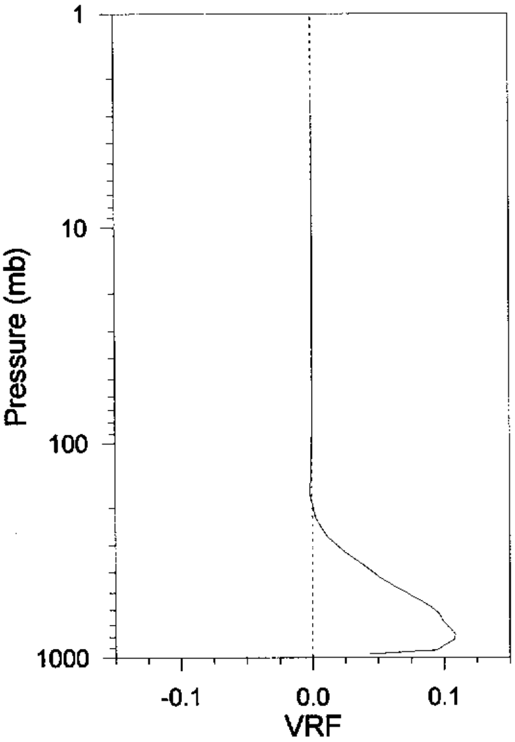

Layer 4

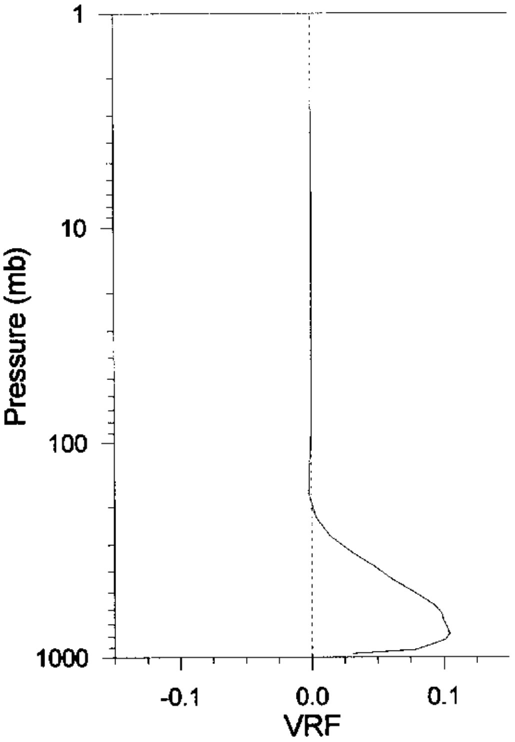

Layer 5

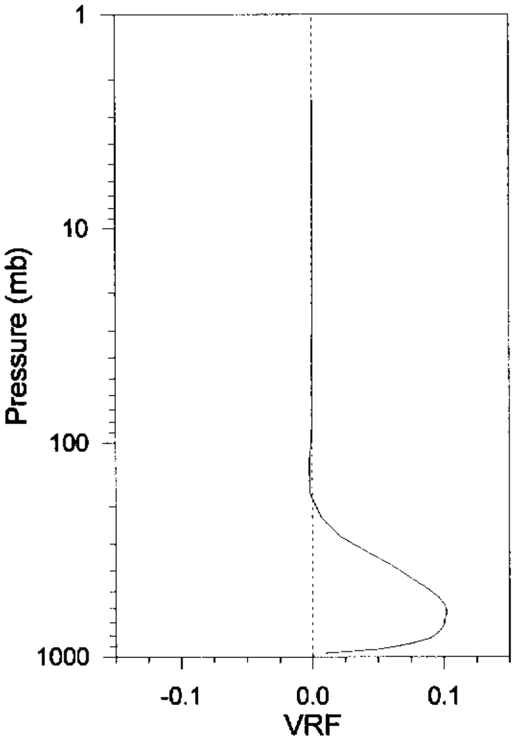

Layer 6

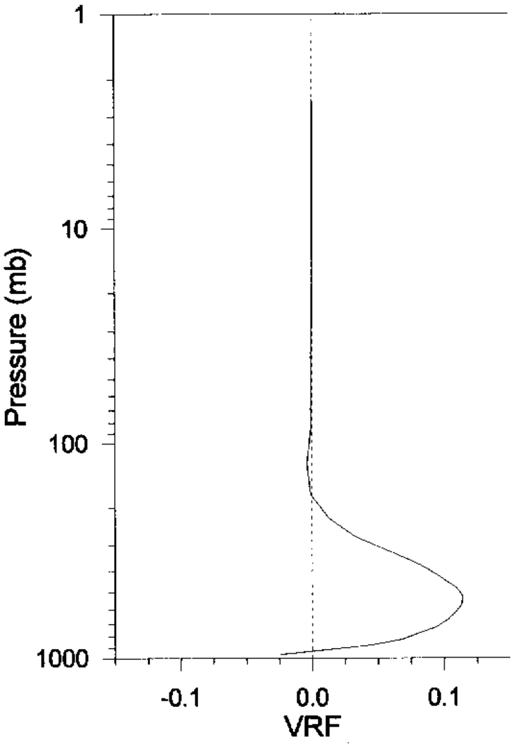

Layer 7

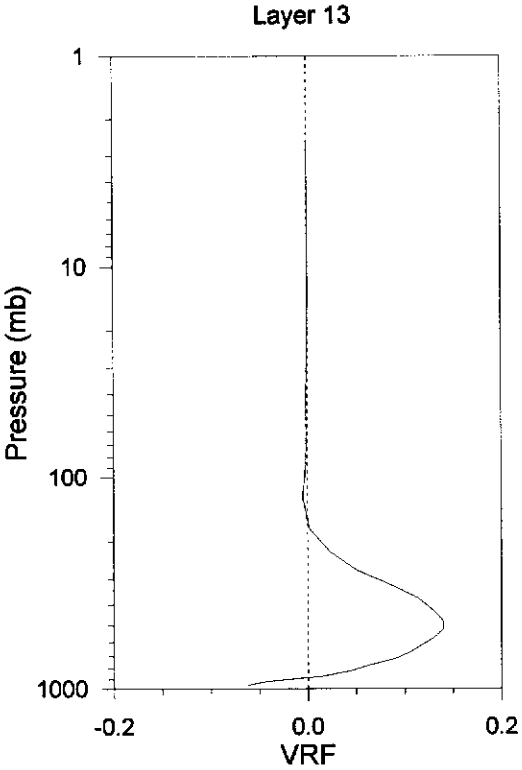

Layer 8

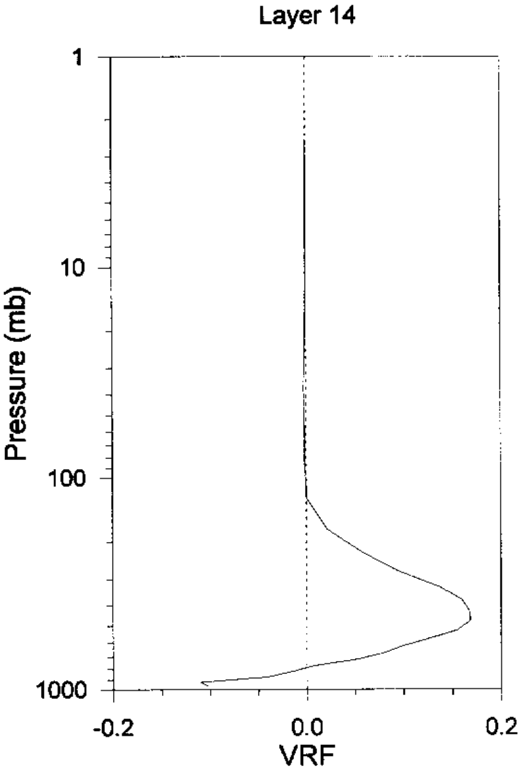

Layer 9

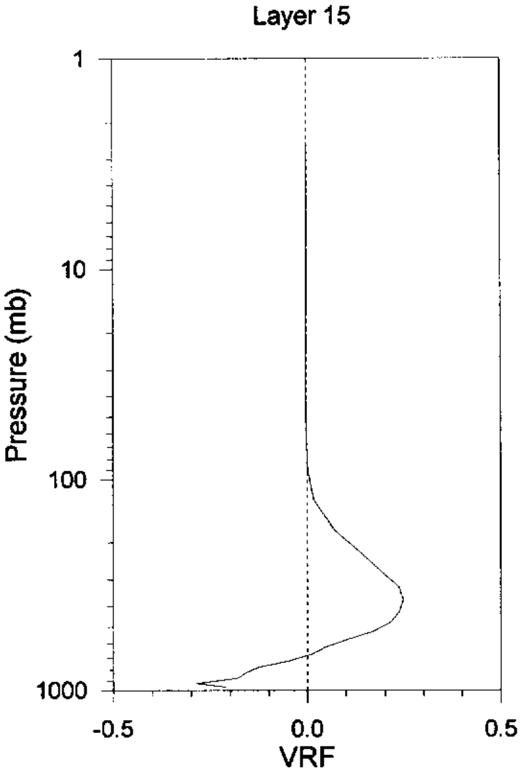

Layer 10

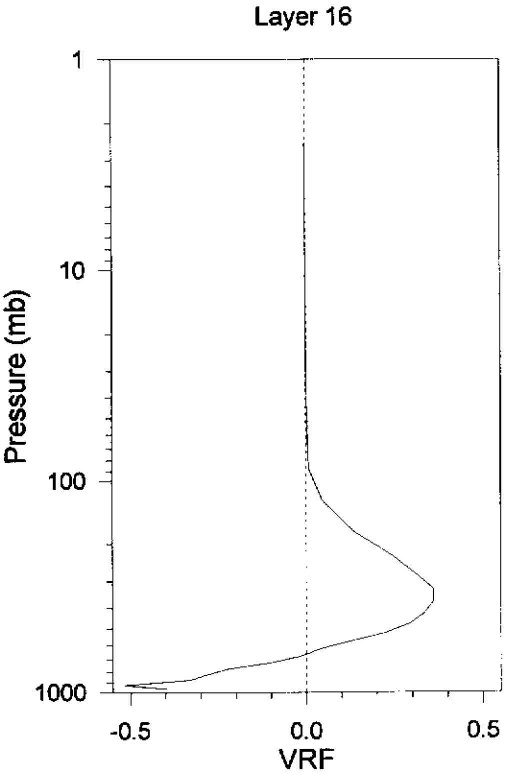

Layer 11

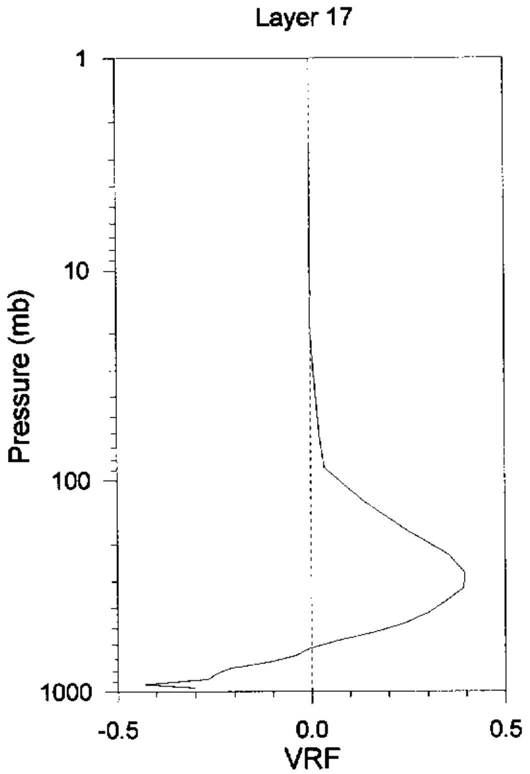

Layer 12

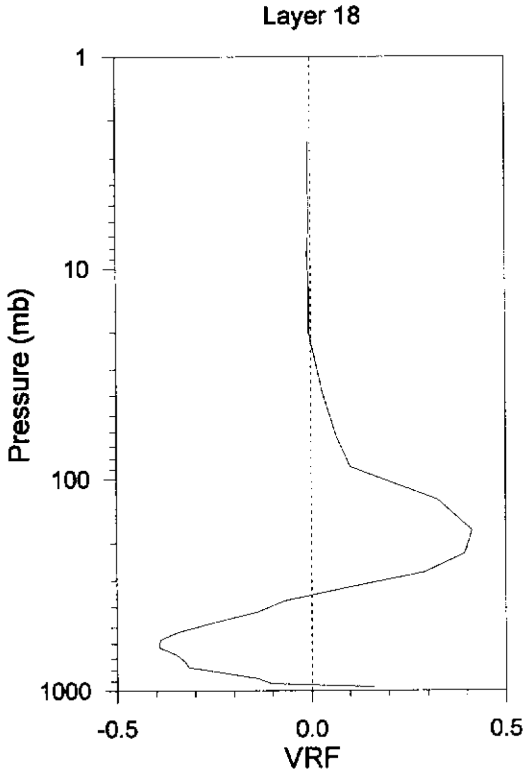

Layer 13

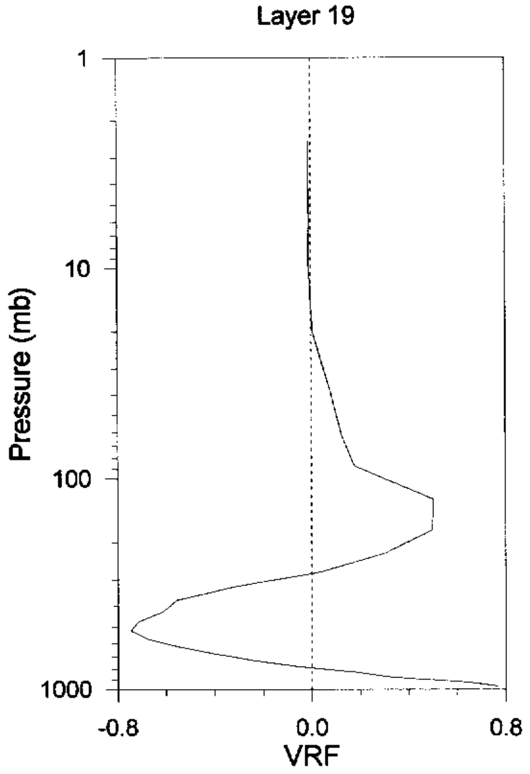

Layer 14

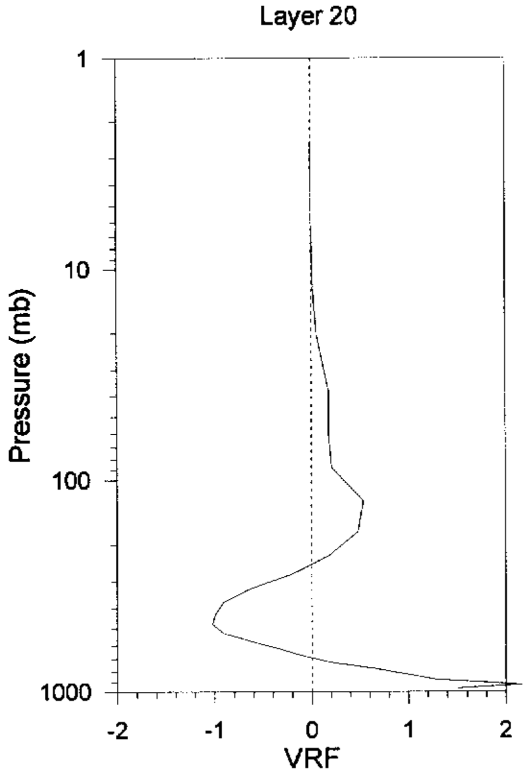

Layer 15

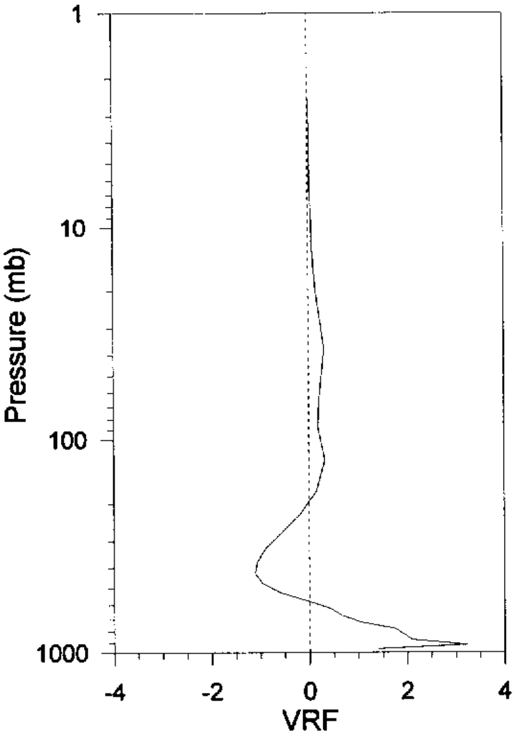

Layer 16

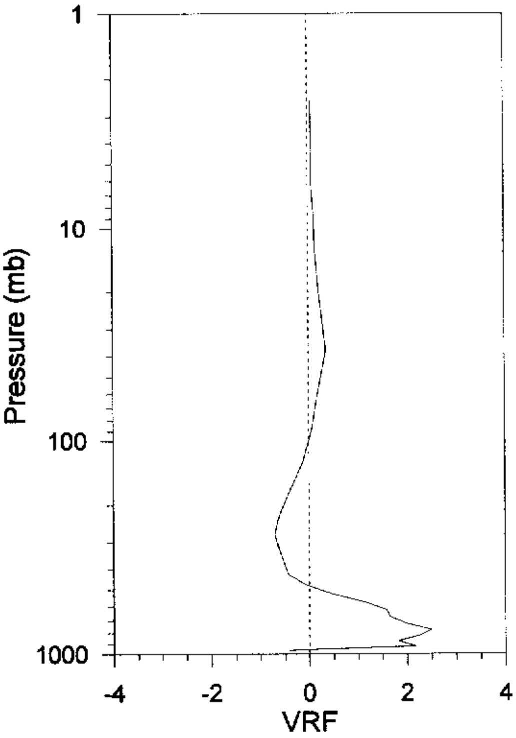

Layer 17
Pressure (mb)
1
10
100
1000
-0.5
0.0
0.5
VRF

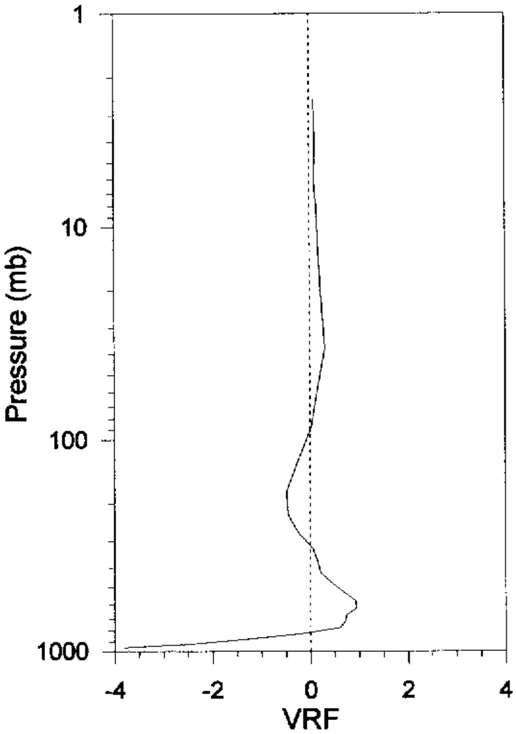

Layer 18
Pressure (mb)
1
10
100
1000
-0.5
0.0
0.5
VRF

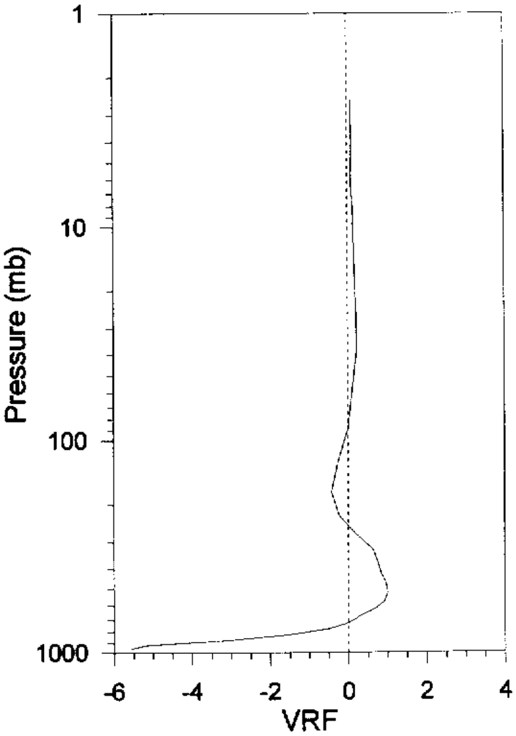

Layer 19
Pressure (mb)
1
10
100
1000
-0.8
0.0
0.8
VRF

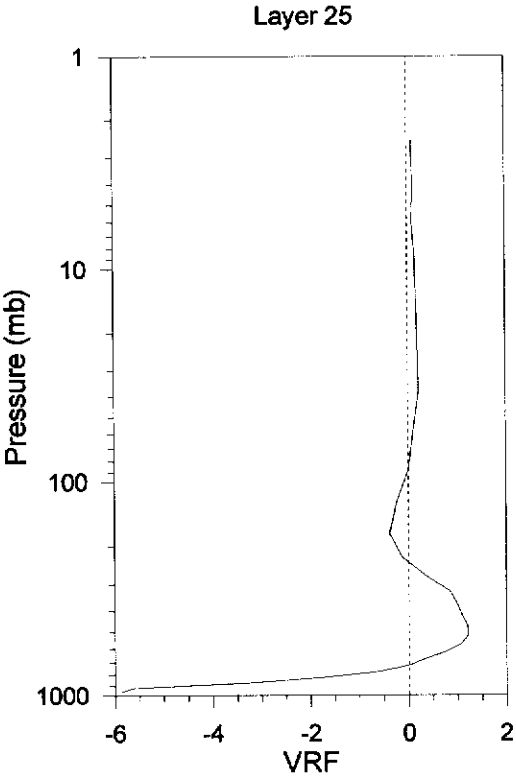

Layer 20
Pressure (mb)
1
10
100
1000
-2
-1
0
1
2
VRF

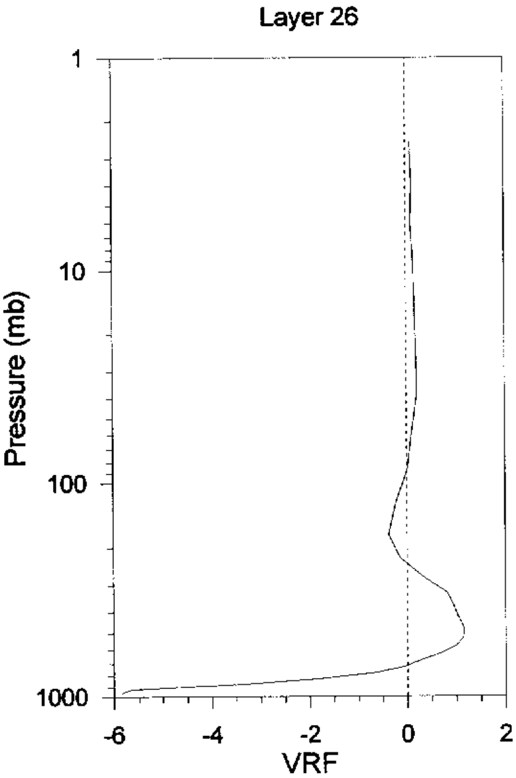

Layer 21

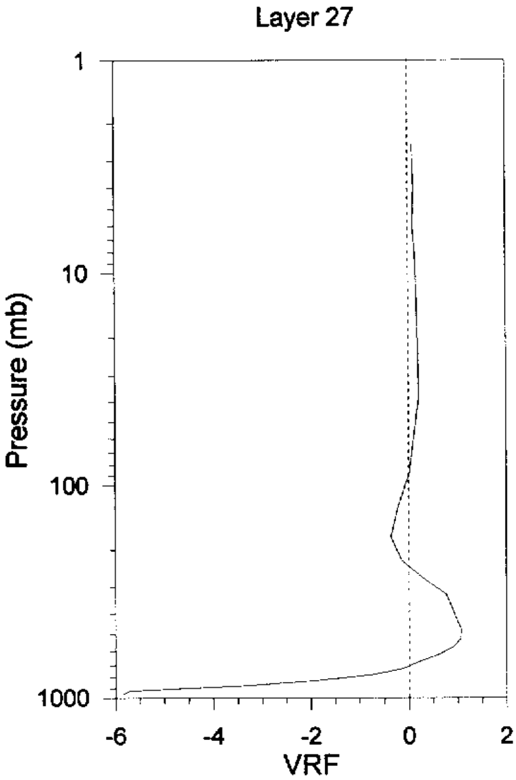

Layer 22


Layer 23


Layer 24


Layer 25
Pressure (mb)
1
10
100
1000
-6
-4
-2
0
2
VRF


Layer 26
Pressure (mb)
1
10
100
1000
-6
-4
-2
0
2
VRF


Layer 27
Pressure (mb)
1
10
100
1000
-6
-4
-2
0
2
VRF