JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. D5, PAGES 9283–9291, MAY 20, 1995

IN-45-TM

Error analysis for the ground-based microwave ozone measurements during STOIC

Brian J. Connor,[1] Alan Parrish,[2] Jung-Jung Tsou,[3] and M. Patrick McCormick[1]

Abstract. We present a formal error analysis and characterization of the microwave measurements made during the Stratospheric Ozone Intercomparison Campaign (STOIC). The most important error sources are found to be determination of the tropospheric opacity, the pressure-broadening coefficient of the observed line, and systematic variations in instrument response as a function of frequency ("baseline"). Net precision is 4–6% between 55 and 0.2 mbar, while accuracy is 6–10%. Resolution is 8–10 km below 3 mbar and increases to 17 km at 0.2 mbar. We show the "blind" microwave measurements from STOIC and make limited comparisons to other measurements. We use the averaging kernels of the microwave measurement to eliminate resolution and a priori effects from a comparison to SAGE II. The STOIC results and comparisons are broadly consistent with the formal analysis.

1. Introduction

A new generation of ground-based microwave spectrometers is being developed for the Network for Detection of Stratospheric Change. Among the advantages offered by these instruments are their wide range of altitude coverage and their ability to operate unattended day and night. While the technique is ideally suited to observations of the middle and upper stratosphere, it has also proven its value on the upper edge of the Antarctic ozone hole [*Connor et al.*, 1987; *de Zafra et al.*, 1989] and in the lower and middle mesosphere [*Bevilacqua et al.*, 1990; *Connor et al.*, 1994].

We report here ozone profile measurements during the Stratospheric Ozone Intercomparison Campaign (STOIC) by an instrument developed at Millitech Corporation and first deployed in the field at Table Mountain, California, immediately prior to the campaign. The instrument was subsequently in operation at Table Mountain until June 1992. The instrument, its operation, and the data analysis have been described by *Parrish et al.* [1992], and a more general presentation of the microwave technique appeared in the work of *Parrish et al.* [1988]. Some results from the 3-year series of microwave measurements at Table Mountain have appeared in the work of *Parrish et al.* [1992], *Connor et al.* [1994], and *Tsou et al.* [1995].

The current paper focuses primarily on an error analysis for the microwave measurements during STOIC, presented in section 2. Section 3 presents the microwave measurements from STOIC and illustrates application of the error analysis to the results. The bulk of the comparisons between the microwave and the other instruments appear elsewhere in this issue [*Margitan et al.*, this issue; *Komhyr et al.*, this issue; *McGee et al.*, this issue].

[1]Atmospheric Sciences Division, NASA Langley Research Center, Hampton, Virginia.
[2]Department of Physics and Astronomy, University of Massachusetts, Amherst.
[3]Lockheed Engineering and Sciences Company, Hampton, Virginia.

Paper number 94JD00413.
0148-0227/95/94JD-00413$05.00

2. Error Analysis

2.1. Definitions

The following analysis is an application of the formulation of *Rodgers* [1990]. The Rodgers error analysis has previously been applied to a variety of middle-atmosphere ozone measurements [*NASA*, 1988; *Connor and Rodgers*, 1989; *Marks and Rodgers*, 1993], and the present results may be directly compared with those. We first define the quantities to be derived.

The radiometric measurement is defined by

$$y = F(x, b) + \varepsilon_y \tag{1}$$

where y is the measured spectrum; F is the true forward function; x is the true ozone profile; b is a set of parameters such as the temperature profile and line strength, which are not to be retrieved; and ε_y is the measurement error. The retrieved profile $\hat{x}$ is given by

$$\begin{aligned} \hat{x} &= I(y, b, c) \\ \hat{x} &= T(x, b, c) \end{aligned} \tag{2}$$

where I is the inverse model, T is the transfer function which explicitly shows the dependence of $\hat{x}$ on x, and c is a set of parameters used by the inverse model but not the forward model (for example, a priori data). The sensitivity of the retrieval to the measurements is given by the contribution function D_y

$$D_y = \frac{\partial I}{\partial y} \tag{3}$$

Similarly, the sensitivity of the retrieval to the true profile is given by the averaging kernel matrix A

$$A = \frac{\partial T}{\partial x} \tag{4}$$

With these definitions we may write the following specification of the retrieval error, adapted from *Rodgers* [1990], equation (14):

$$\hat{x} - x = (A - I)(x - x_a) + D_y \Delta y_b + D_y \varepsilon_y \qquad (5)$$

Here, I is the identity matrix, x_a is the a priori used in the retrieval, and Δy_b is the effective error in y due to errors in the forward model and instrument calibration. The first term on the right-hand side arises from the smoothing of the true profile by the finite resolution measurement and from inaccuracy of the a priori data.

The second term contains errors due to data calibration and in the calculation of spectra for comparison with the data. Finally, the third term specifies errors due to noise or other instrument errors in the measured spectrum.

The following sections discuss each term of (5) at length. We first detail forward model and calibration errors (section 2.2), separately discuss errors due to the temperature profile used in the forward model (section 2.3), then take up the question of spectral measurement error (section 2.4), and lastly (section 2.5) the resolution and the effect of the a priori used in the retrieval. Section 2.6 then discusses the net error as derived from the analysis.

2.2. Forward Model and Calibration Errors

The following general procedure was employed to assess the effect of errors in the calculation of synthetic spectra and the conversion of the raw data to physical units. First, estimate or assume the magnitude of the particular error being considered (e.g., the instrument viewing angle). Second, calculate the spectral radiance error which results. Third, form the spectral covariance S_ε; for all cases the correlation between frequencies is assumed to be either 1 or 0, as appropriate. Fourth, calculate the retrieval error covariance

$$S_s = D_y S_\varepsilon D_y^T \qquad (6)$$

Finally, present the rms retrieval error, which (ignoring correlations between layers) is the square root of the diagonal elements of S_s. In the case where the spectral errors are fully correlated, this diagonal specifies the error completely since the errors at all altitudes are also fully correlated. If the spectral errors are uncorrelated, the off-diagonal elements of the profile covariance are also important. We will explore such a case in section 2.4 on measurement error.

The geometry and calibration of the measurement and the formulation of the forward and inverse models are fully described by *Parrish et al.* [1992a, 1988]. In outline, the observation is made by switching rapidly between two elevation angles, one near the zenith (the "reference beam") and one at 10°–27° (the "signal beam"). To minimize the effects of detector nonlinearity, the band-averaged power is equalized in the two beams by inserting a lossy dielectric sheet in the reference beam and then varying the signal beam elevation under control of a servomechanism. The observed spectrum is given by the following, adapted from equation (1) of *Parrish et al.* [1992a]:

$$T_a(\nu) = \left(\left(\frac{V_s(\nu)}{V_r(\nu)} \right)^{1+\delta(\nu)} - 1 \right) T_{\text{sys}}(\nu) \qquad (7)$$

$T_a(\nu)$ is the atmospheric signal (calculated from molecular physics and radiative transfer) modified by the losses in the instrument windows, optics, and the dielectric sheet in the reference beam. $V_s(\nu)$ and $V_r(\nu)$ are the voltages in the signal and reference beams, respectively; $\delta(\nu)$ specifies the power law for each channel, allowing correction for detector nonlinearity; it is measured in a special calibration sequence. $T_{\text{sys}}(\nu)$ is the "system temperature," which consists of the receiver equivalent noise temperature and the radiation temperature in the reference beam and is measured in a regular calibration sequence. The value derived for $T_{\text{sys}}(\nu)$ is dependent on our knowledge of the radiometric properties of the calibration targets, the optical properties of the instrument, and the emission of water vapor and oxygen in the troposphere.

We must consider errors which affect either side of (7). It will be necessary to classify each error source as fixed or variable and to decide which of them affect the measurement precision. These distinctions are not always clear-cut. For example, while error in the determination of tropospheric opacity is variable, it will depend somewhat on local weather conditions and is thus likely to be correlated from day to day during a period of stable weather. To the extent of the correlation it will not affect precision. Similarly, the transmission of the instrument optics may be considered fixed, at least until the instrument is physically disturbed, at which time it will need to be remeasured. We have taken a simple approach to these questions. Namely, errors which are unlikely to change significantly while the instrument is operating undisturbed are considered fixed. All others are variable and are assumed to contribute to measurement imprecision. This approach is likely to overestimate the random component of the error (understate the precision) since not all variable errors will change randomly day to day.

2.2.1. Variable errors. Such sources of error include beam elevation angles, radiometric temperatures of the calibration targets, the detector power laws, and the tropospheric optical depth. Error due to the variation in atmospheric path across the beam is negligible for our 2° beam and elevation angles of more than 10°. The dominant error in this category arises in determination of the tropospheric optical depth. Absorption by oxygen and water vapor in the troposphere attenuates the stratospheric signal, which is the object of measurement; thus the calculated signal from stratospheric ozone must be scaled by an appropriate factor to compensate. The tropospheric optical depth is determined as follows: First, we determine the absolute radiometric temperature of the sky at a set of n elevation angles in the far wings of the ozone line by comparing the observed signal with signals from blackbody targets at known temperatures. Second, we assume this temperature is produced entirely by emission in the troposphere and given by

$$T_{tr}(\theta_i) = T_{\text{atm}}(1 - e^{-\tau A(\theta_i)}), \quad i = 1, n \qquad (8)$$

τ is the zenith optical depth, assumed independent of frequency; A is the air mass; and T_{atm} is the effective temperature of the troposphere. Third, we solve the set of equations (8) for τ. Note that this model assumes horizontal homogeneity. T_{atm} is estimated from the surface temperature at the time of measurement, T_{out}, by $T_{\text{atm}} = T_{\text{out}} - 7$ K [*Parrish et al.*, 1988]; we estimate T_{atm} to be uncertain by ±5 K. This results in a small error, of the order of 1%. Repeated measurements of τ during relatively stable conditions show a scatter of the order of ±0.005, which is much too large to be accounted for by measurement noise but is believed to be

due to failure of the assumption of horizontal homogeneity and variations in the tropospheric temperature profile. Figure 1a shows the rms error in retrieved ozone due to random errors of ±5K in T_{atm} and ±0.005 in τ and also due to all variable errors combined. We note the net variable calibration error is 3–5% between 55 and 0.2 mbar.

2.2.2. Fixed errors. One class of errors in this category are those involved in calculation of the molecular absorption coefficient, namely, errors in the line strength and width. Line strengths of pure rotational transitions may be calculated very accurately; the primary source of error is in experimental measurement of the static dipole moment. For ozone, this uncertainty is ±1% [*Lichtenstein et al.*, 1971]; the line strength is proportional to the square of the dipole moment, so the uncertainty in line strength is ±2%. The air-broadened line width (and its temperature dependence) of the observed line (110.836 GHz) has been measured in the laboratory [*Connor and Radford*, 1986] and has a net uncertainty of ±4%.

Also included here are the effects of three approximations made in the forward model used to analyze the STOIC data. The first of these ignores the small contribution to $T_{sys}(\nu)$ of the ozone line itself, as seen in the reference beam. The second assumes emission from the troposphere, varies linearly with frequency, whereas, in fact, there is some curvature due to the oxygen line at 118 GHz. The third approximation is that the frequency response of the individual channels is a delta function. All of these approximations have been removed from our current data processing algorithm. However, the STOIC data set was fixed before that, so for present purposes, these approximations must be included as error sources. The net error due to the three approximations ranges from 2 to 4%.

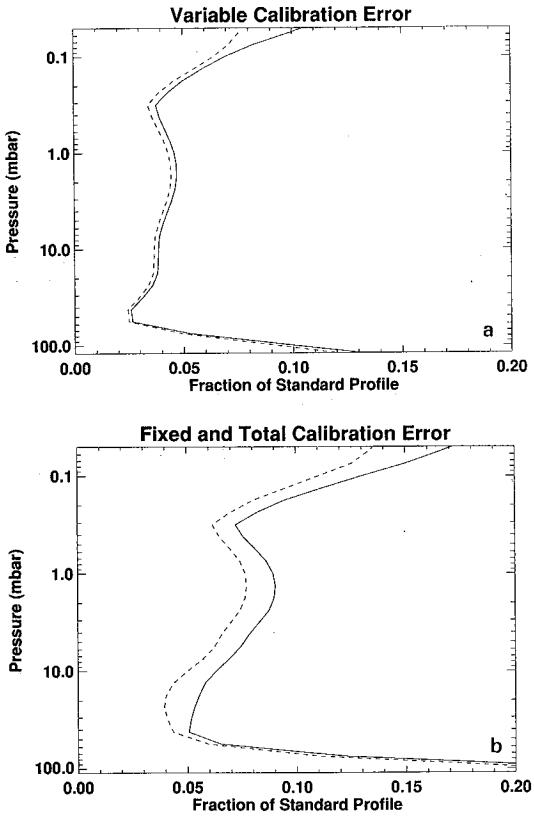

Figure 1. Retrieval error due to forward-model and calibration error. (a) Error due to tropospheric opacity determination only (dashes). Total of all variable errors (solid). (b) Fixed errors only (dashes); fixed and variable combined (solid).

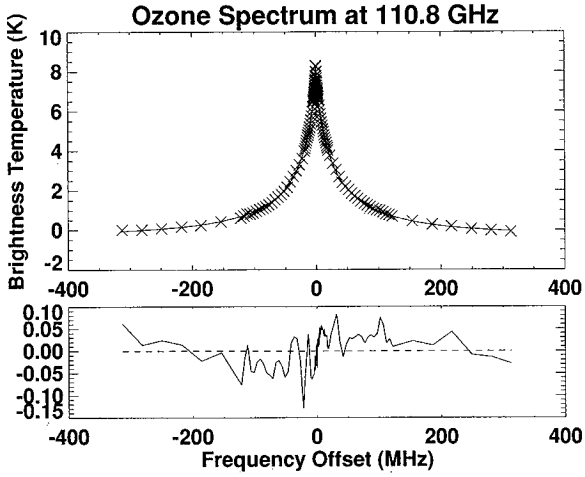

Figure 2. Ozone spectrum recorded July 30, 1989. The top panel shows the measured spectrum (X values) superimposed with the spectrum calculated from the retrieved profile shown in Figure 9 (solid curve). The bottom panel shows the difference between the measured and the calculated spectra.

The last important error source in this category is the nonlinear variation of instrument frequency response referred to as baseline. It arises from reflections in the optics and the resulting interference between forward and reverse traveling waves. The relative phase of the interfering waves varies with frequency, with the net result being an approximately sinusoidal output signal. Baseline error is something of an imponderable for ground-based microwave measurements, because the continuum underlying the observed line cannot be directly observed. We have estimated its effect as follows: First, we inspect the residuals of the spectrum calculated from the retrieved profile subtracted from the observed spectrum and note the amplitude and wavelength of any sinusoidal pattern. For the STOIC data these numbers are roughly 0.05 K and 600 MHz. A typical spectrum, in which this pattern is visible, is shown in Figure 2. We then assume there is a sinusoidal baseline of this amplitude and wavelength centered in the band (which is the worst case phase for such a signal) and calculate the resulting error. In the present case, we derive an error of about 4% at 55 mbar and <1% above 10 mbar.

Finally, we include uncertainty in the transmission of the instrument optics and the systematic component of uncertainty in the detector response. These are both small, making a combined contribution to the error budget of about 1%.

Figure 1b shows the rms error in retrieved ozone due to the combined effects of all the errors treated as fixed. Also shown in Figure 1b is the combined error due to fixed and variable forward model and calibration errors. This combined calibration error ranges from 5 to 10% between 55 and 0.2 mbar.

2.3. Temperature Profile Errors

Error in the temperature profile input to the forward model is a type of variable forward-model error. We have chosen to

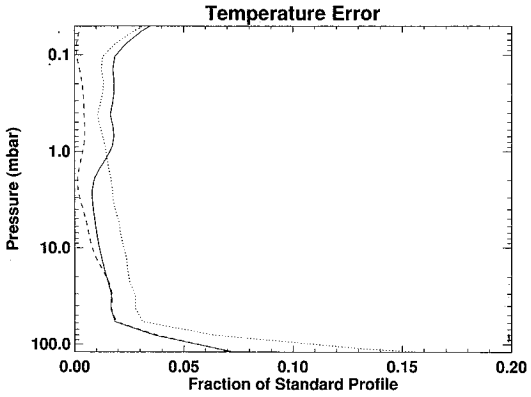

Figure 3. Retrieval error due to error in atmospheric temperature. Case 1 (dots): 5 K random, uncorrelated error. Case 2 (solid): random error with 8-km correlation length; 2 K below 1 mbar, 5 K above. Case 3 (dashes): 2 K uniform error.

treat it separately, however, both because particular interest is often expressed in the effect of atmospheric temperature on the retrievals and also because a special procedure was developed to model it, which could as well be applied to other remote sensing problems.

The following is a variant of the Rodgers procedure for forward-model errors. We define the temperature sensitivity matrix

$$D_T = \frac{\partial \hat{x}}{\partial T} \tag{9}$$

where T is the atmospheric temperature profile. D_T is calculated a column at a time by perturbing a single layer of a standard temperature profile and determining its effect on the retrieval. The hydrostatic effects of the temperature perturbation are included.

Then the covariance of the retrieval due to temperature errors is

$$S_{s,T} = D_T S_T D_T^T \tag{10}$$

where $S_{s,T}$ is the covariance of the atmospheric temperature. This approach is attractive because of its generality. It allows evaluation of the effect of temperature errors with arbitrary cross correlations simply by constructing the appropriate matrix S_T and applying (10).

Figure 3 shows the rms error due to temperature (square root of the diagonal of $S_{s,T}$) for three cases. In all cases the temperature profile is specified on a pressure grid with spacing of 0.125 in $\log_{10}P$ (≈2-km spacing in altitude). Case 1 is random, uncorrelated 5 K error. Case 2 is a random error of 2K below 1 mbar and 5K above, with an 8-km correlation length throughout. Case 3 is a constant error of 2 K. Case 2 is considered most realistic for our operational retrievals, where we use the National Meteorological Center analyses to approximately 1 mbar and the CIRA climatology above [*Barnett and Corney*, 1985]. It may be seen that for that case, rms retrieval errors are 1–2%. We have adopted the results of case 2 as our standard estimate of uncertainty due to temperature. During STOIC, sonde and lidar temperatures were available for nearly every day, so temperature errors may well have been smaller.

2.4. Measurement Error

We assess the effect of errors in the spectral measurement (the third term on the right-hand side of (5)) by a procedure similar to that used on the forward-model and calibration errors. We calculate the profile error covariance due to measurement error, S_m, by

$$S_m = D_y S_\varepsilon D_y^T \tag{11}$$

The diagonal elements of S_ε are estimated from the residuals of the fit of the calculated to measured spectra. These residuals are larger than a theoretical noise calculation would indicate, except in the narrowest channels for a short observation. The residuals also show signs of correlation between nearby channels (Figure 2), although the pattern of the residuals is by no means repeatable. In principle, one can include the effects of partially correlated spectral errors by incorporating appropriate off-diagonal elements in S_ε. However, we have insufficient information about any such correlations to be able to do that; further, assuming the errors are uncorrelated is conservative since there is less information in the spectrum if the errors are uncorrelated. For both those reasons we assume S_ε is diagonal.

Figure 4a shows the rms retrieval error for three cases: daily average measurements for both day and night and a single (20-min) daytime measurement. Several points should be noted. First, the errors range from 1 to 3% between 55 and 0.2 mbar. Second, the errors increase very rapidly below

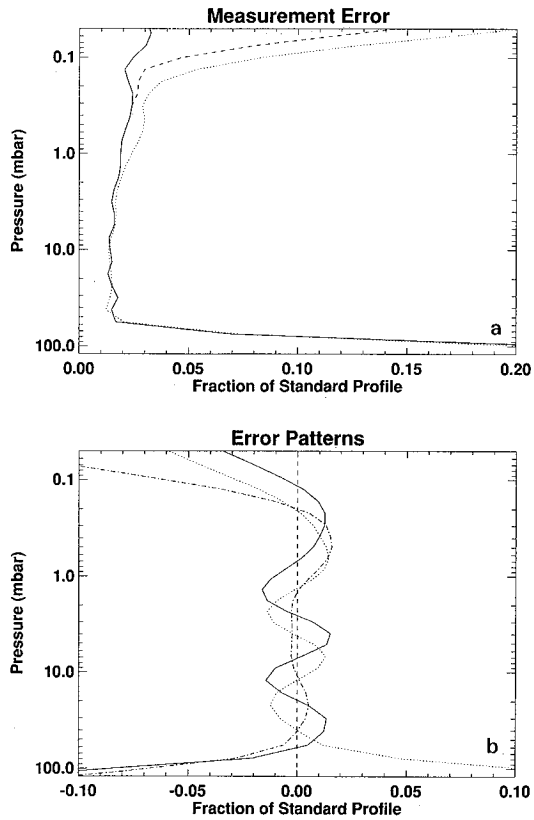

Figure 4. (a) Retrieval error due to measurement noise. Daytime, single measurement (dots). Daytime, daily average (dashes). Night, daily average (solid). (b) Three largest error patterns corresponding to daytime, daily average retrieval (dashed curve in Figure 4a).

55 mbar, both because the ozone signal decreases rapidly at lower altitudes and because the remaining signal is not fully contained in the instrument band pass. Third, the relative effect of noise in the lower mesosphere is much less at night, simply because there is more ozone and thus more signal. Finally, the improvement in precision obtained by making a day-long measurement instead of a 20-min measurement is quite small. This is due to two factors. As mentioned above, the spectral errors do not decrease with time as expected theoretically but much more slowly. Also, precision and resolution are not independent but may be exchanged one for the other by adjustments to the retrieval parameters [*Brillet*, 1989]. Our operational retrieval is geared to improve the resolution as the spectral signal-to-noise ratio improves. We shall see in section 2.5 that the resolution is significantly better for the daily average measurements.

Because the spectral errors are uncorrelated, we must also consider the off-diagonal elements of S_m. The eigenvectors of S_m have been called by *Rodgers* [1990] error patterns and represent independent components of the error. The three most important error patterns for the daytime, daily average retrieval are shown in Figure 4b. From these, we see that the characteristic errors between 55 and 0.2 mbar are oscillatory in nature, with length scale of roughly a factor 3 in pressure and amplitude about 2%.

2.5. Resolution

The rows of the matrix A defined by (4) are the measurement averaging kernels, which show explicitly how the atmospheric profile is smoothed in the measurement and retrieval process. From (5)

$$\hat{x} = x_a + A(x - x_a) + \text{error terms} \qquad (12)$$

so the retrieved profile in the absence of other significant errors is the a priori profile plus the product of A and the difference between the true and the a priori profiles. Figure 5a shows the averaging kernels for the daily average measurement. At the nominal altitude for each curve, the retrieved profile is the average of the true profile at all altitudes weighted by the averaging kernel. Resolution is formally defined as the full width at half maximum of the kernels; however, this is meaningful only if the kernels are well peaked and centered fairly closely on their nominal altitude. Figure 5b shows the resolution, so defined, for both the daily average and the single observations. The resolution of the daily average measurement is in the range 8–10 km between 55 and 3 mbar and increases to 17 km at 0.2 mbar. Below 55 and above 0.2 mbar the resolution concept gradually loses meaning as the measurement is progressively less able to respond to the true profile. It may also be seen that the resolution of the single measurement is 1–2 km coarser everywhere. The smoothing of the averaging kernels introduces an error, which will be important in any comparison between measurements made at different resolutions. The magnitude of this error can only be estimated if the statistical variability of the true atmosphere is known, especially on fine vertical scales. However, it is possible to eliminate this smoothing error from a comparison of measurements by using the averaging kernels. In comparing a high-resolution measurement (or a model calculation) with a low-resolution measurement, one computes $x_a + A(x_m - x_a)$, where x_m is the high-resolution measurement, and compares the result to

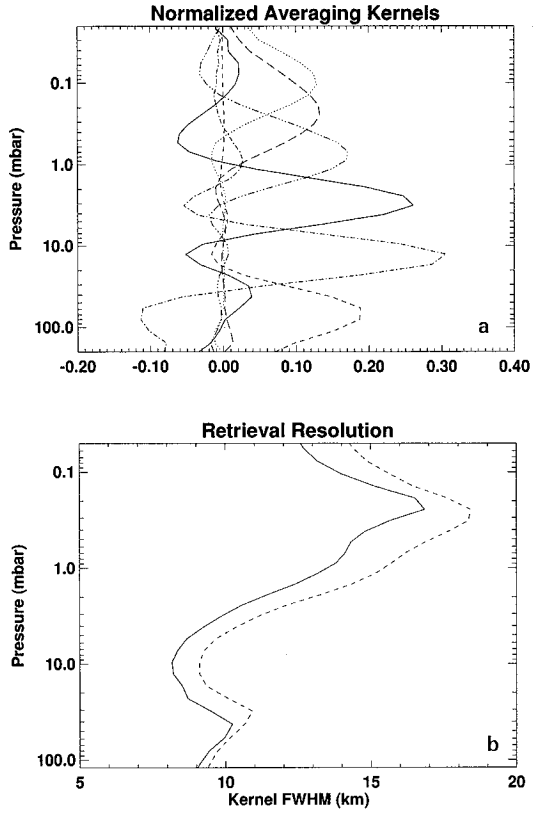

Figure 5. (a) Averaging kernels for the daily average measurement. Nominal pressures (millibars) for the kernels shown: 55 (dashes), 12 (dot-dash), 3 (solid), 0.8 (dash-dot-dot-dot), 0.2 (long dashes), 0.05 (dots). (b) Resolution of the daily average retrieval (solid) and single measurement (dashes).

$\hat{x}$. An example of this type of comparison is given in section 3. Because this source of error can be unambiguously eliminated from comparisons with models or other measurements, it is not included in the net error estimates of section 2.6.

It is also important in this context to examine the influence of the a priori on the retrieval. *Rodgers* [1990] has shown that if the eigenvectors of A are used as basis vectors for the atmospheric profile, the eigenvalue corresponding to a given vector explicitly gives the fraction of that vector's coefficient which comes from the true profile. In other words, eigenvectors of A which have eigenvalues ~1 represent components of the true profile which are well measured, while eigenvectors which have eigenvalues ~0 represent components which depend heavily on the a priori. Figure 6 shows four of the eigenvectors of A and their corresponding eigenvalues. This shows explicitly that vertical scales finer than the measurement resolution are measured poorly, i.e., come from the a priori (Figure 6d), while vertical structures broader than the resolution have very little dependence on the a priori (Figures 6a and 6b). Figure 6c shows an intermediate case where 70% of the information comes from the real profile and 30% from the a priori.

Alternatively, we may rewrite (12) as

$$\hat{x} = Ax + (I - A)x_a + \text{error terms} \qquad (13)$$

expressing the retrieval as a linear combination of the true and the a priori profiles. A rough idea of the net contribution

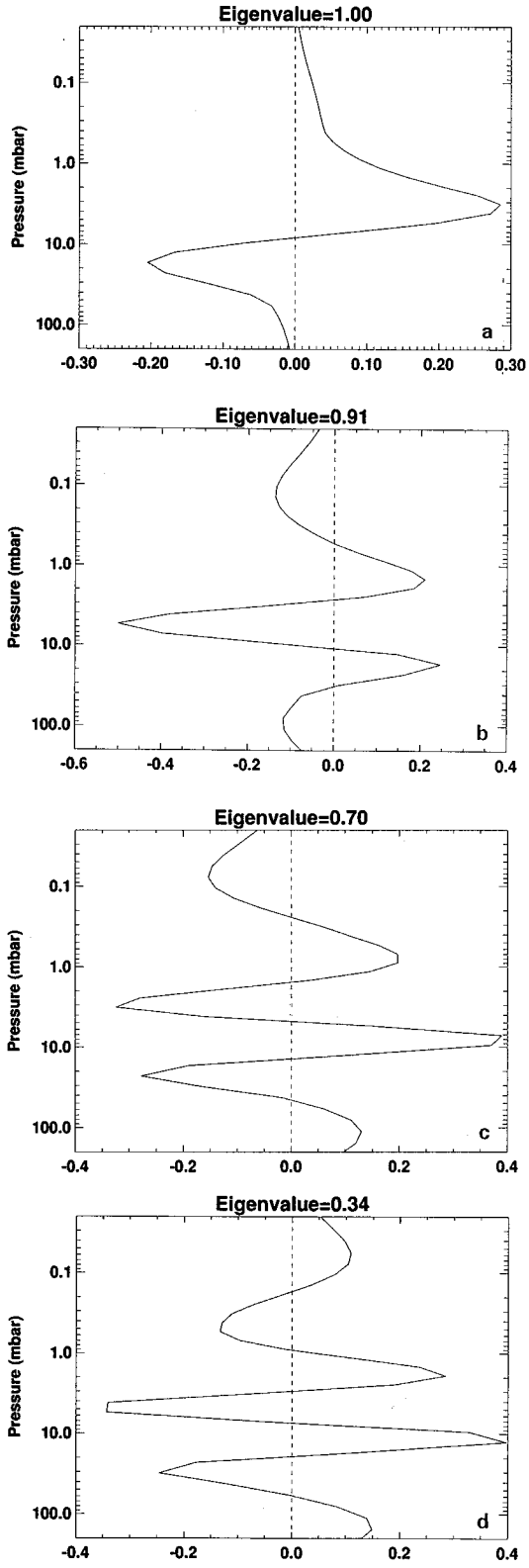

Figure 6. Selected eigenvalues of the averaging kernel matrix sampled in Figure 5a.

of the a priori may be gained by examining $(I - A)_j x_a / \hat{x}_j$, where $(I - A)_j$ is the jth row of $(I - A)$. This quantity is shown in Figure 7. It may be seen that the contribution of the a priori to the retrieval, defined in this way, is typically 5%.

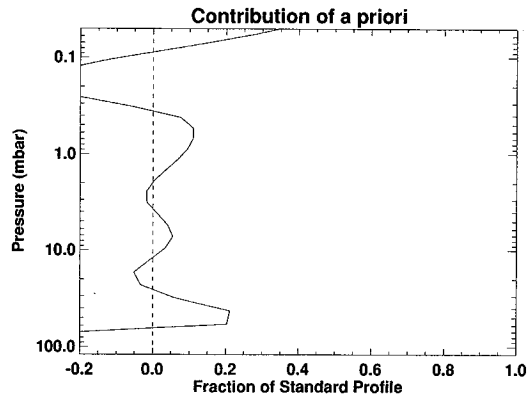

Figure 7. Relative contribution of the a priori to the daily average retrieval, as defined in the text.

2.6. Net Error and Discussion

Figure 8 shows the net error as derived from the analysis, excluding smoothing error (as discussed in section 2.5). The curves shown are appropriate for nighttime. During the day, errors in the mesosphere are somewhat larger. The dashed curve shows net precision, including contributions from the variable forward model and calibration errors, temperature profile errors, and spectral measurement errors. The solid curve is accuracy, which includes in addition fixed-calibration errors, forward-model approximations, instrument baseline, and errors in molecular parameters. It may be seen that precision ranges from 4 to 6% between 55 and 0.2 mbar, while accuracy is 6–10% over the same pressures.

The results of the error analysis are summarized in Table 1. We consider the most important error sources to be the molecular parameters, the tropospheric opacity measurement, and the instrument baseline (the errors due to forward-model approximations can be entirely eliminated and have been in our more recent work). Taking these in turn, the molecular parameters (primarily the line-broadening coefficient) cause an error which is truly fixed and so can be eliminated for some purposes by studying relative variations. The line-broadening coefficient is already fairly well known (±4%), so a very accurate laboratory measurement would be needed to improve the situation.

The uncertainty due to tropospheric opacity could be much reduced by more detailed knowledge of the distribu-

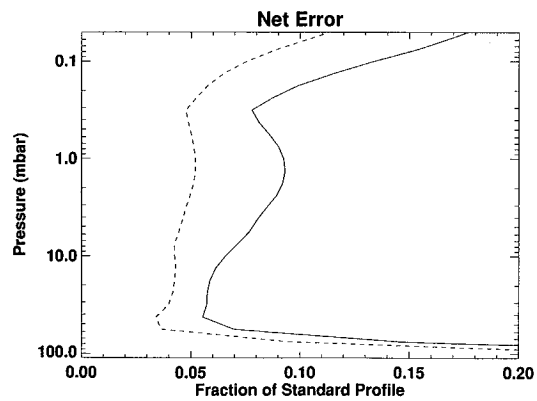

Figure 8. Net precision (dashes) and accuracy (solid).

Table 1. Precision, Accuracy, and Resolution for Stratospheric Ozone Intercomparison Campaign

		Variable Errors, %					Fixed Errors, %					
P, mbar	Approximation z (km)	Opacity	Other Calibration	Temperature	Measurement	Precision, %	Spec.	Calibration	FM Approximation	Baseline	Accuracy, %	Resolution, km
55	20	3	1	2	2	4	3	...	3	4	7	10
12	30	4	1	1	2	4	4	1	2	2	6	8
3	40	4	1	1	2	5	6	1	3	1	8	10
0.8	50	4	2	2	2	5	6	1	4	...	9	14
0.2	60	5	2	2	2*	6	7	1	4	...	10	17

*3% daytime.

tion of water vapor and temperature with height, which would allow proper modeling of the tropospheric emission. High vertical resolution would not be required since the quantity of interest is the total opacity. It is worth investigating the feasibility of using ground-based infrared measurements for this purpose.

Finally, the instrument baseline can become a significant source of error at altitudes below about 10 mbar and, furthermore, one which is very difficult to estimate. While during a short campaign such as STOIC, when the instrument is left undisturbed and observing conditions are not too variable, this error will be roughly constant (and so has been considered "fixed" here), it will most certainly not be constant in general. We believe the primary cause of the problem in the present system to be reflections from the window of the receiver dewar, and we are investigating use of a flexible window mount which could be moved so as to vary the reflection phase randomly and thus minimize the residual baseline structure.

3. Stratospheric Ozone Intercomparison Campaign (STOIC) Results

The microwave measurements from STOIC are presented in Figures 9 and 10. These are the blind data, as submitted to the coordinator during the campaign. As discussed by *Margitan et al.* [this issue], the revised microwave data set was only slightly different, and the two sets are equivalent for all practical purposes. Figure 9 shows a typical profile, for July 30, with error bars representing the absolute accuracy from Table 1. Figure 10 is a contour plot of profiles for the entire period. This figure reiterates the observation of Margitan et al. that there was very little real variability in the middle to upper stratosphere during the campaign and shows that the same conclusion applies in the lower and middle mesosphere. (Note that only nighttime data are included in the STOIC data set to make measurements by the various instruments as nearly simultaneous as possible. Inclusion of daytime measurements would reveal a pronounced diurnal variation above 1 mbar [*Connor et al.*, 1994].)

Figure 11 shows the percent difference of the microwave measurements from the STOIC reference profile [*Margitan et al.*, this issue], along with error bars indicating the total accuracy. At all altitudes the error bars overlap zero difference. Considerable caution is needed in interpreting this plot, both because the microwave measurements are themselves included in the reference and because the composition of the reference changes with height. Nevertheless, some observations are worthwhile. First, comparisons of the microwave data to the STOIC reference are most meaningful between 20 and 32 or 33 km simply because, in that range, there are approximately twice as many independent measurements included as at higher altitudes (because of the three series of sonde flights). Above 35 km the reference is dominated by the lidars and the microwave itself (since there were relatively few Stratospheric Aerosol and Gas Experiment (SAGE II) and ROCOZ measurements); indeed above 45 km, where the Goddard Space Flight Center (GSFC) lidar made only a few measurements, the microwave accounts for roughly 40% of the measurements in the reference. The

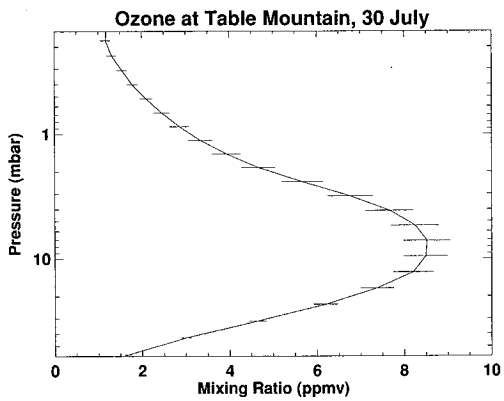

Figure 9. Microwave ozone measurement of July 30, 1989, retrieved from the spectrum of Figure 2. The error bars are measurement accuracy from Table 1.

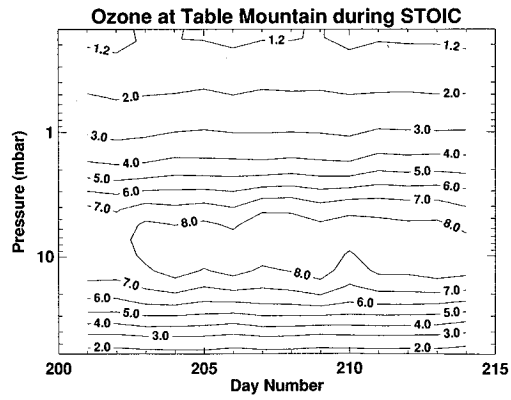

Figure 10. Contours derived from all the nighttime microwave measurements during the Stratospheric Ozone Intercomparison Campaign (STOIC).

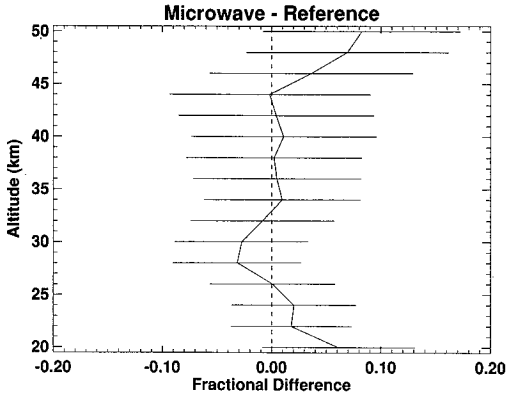

Figure 11. The mean difference between the microwave measurements and the STOIC reference profile. The error bars show the microwave accuracy from Table 1.

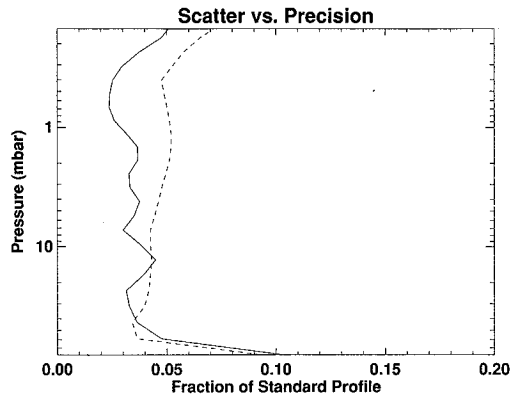

Figure 13. The rms scatter in the microwave STOIC measurements (solid) and the predicted precision from Table 1 (dashed).

departure of the microwave from the reference in that region simply reflects a divergence from the Jet Propulsion Laboratory (JPL) lidar measurements. On the other hand, the oscillatory structure at 20–32 km is suggestive of an error in the microwave average.

Figure 12 shows a time series of the microwave measurements at 24, 28, and 32 km plotted against the mean of all other measurements (note that these mean values are distinct from the daily averages of Margitan et al. because they do not include the microwave data). The range 20–32 km was chosen because it is where the largest number of independent measurements were made, as discussed above. There appear to be some coherent day-to-day variations in the two curves at the 5–10% level. The correlation coefficient of the two is in the range 0.4–0.5 for 24–32 km, though <0.1 at 20 km.

The rms variability of the microwave profiles is plotted as the solid curve in Figure 13, along with the expected precision. It may be seen that agreement is fairly good, with both curves having similar shape and magnitude. Nevertheless, the variability might be expected to be greater than the precision, since the former includes real atmospheric changes as well as instrument effects, while in fact the converse is true. We believe we have understated short-term measurement precision by including a number of error sources which are likely to vary slowly, as discussed in section 2.2.

We now examine the comparison between the microwave results and SAGE II. Figure 14 shows the mean fractional difference between the two measurements for the 3 days of SAGE II overpasses. The solid curve is the direct comparison between retrieved profiles, with the microwave interpolated to the 1-km SAGE II resolution; the dashed line shows the comparison after the SAGE II profiles have been convolved with the microwave averaging kernels. The two measurements agree at the 5–10% level throughout the range 60–0.5 mbar. The convolution of SAGE II with the averaging kernels smoothes out the fine structure in the difference without changing its broad vertical structure. The microwave agrees with the convolved SAGE profiles within 6% at

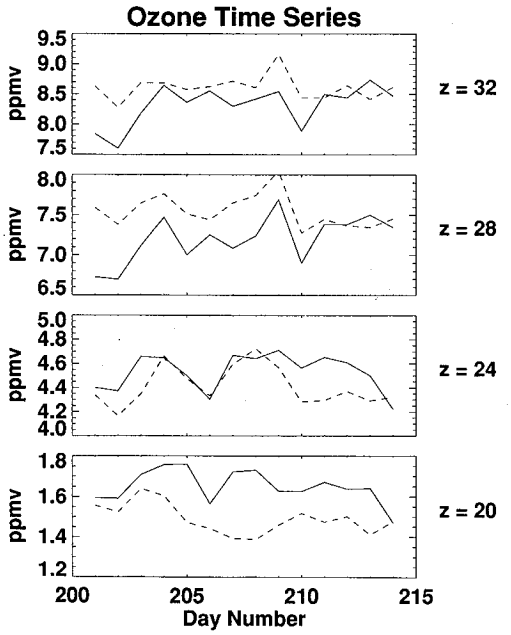

Figure 12. Time series comparison of ozone at 20, 24, 28, and 32 km. The solid curve is the microwave; the dashed line is the average of all other measurements.

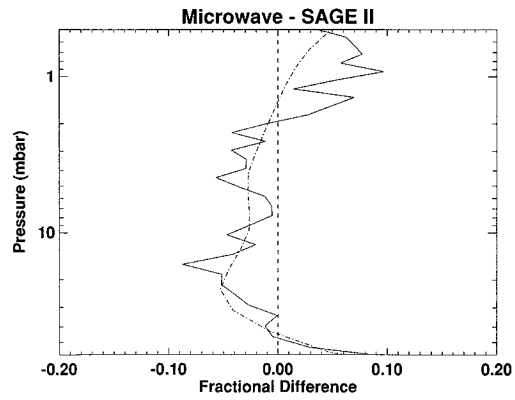

Figure 14. The mean difference between the microwave and the Stratospheric Aerosol and Gas Experiment (SAGE II) measurements of July 23, 24, and 25 (solid). The dashed-dotted curve is the difference after convolution of the SAGE II measurements with the microwave averaging kernels.

all altitudes within the microwave error bars. Because that is so and because of the small number of profiles included in this average (three), we do not believe there is any significance to the vertical structure in the comparison.

4. Conclusion

In conclusion the formal analysis appears representative of the STOIC results. More extensive comparisons of the 3-year series of microwave measurements at Table Mountain with the colocated lidar and with SAGE II overpasses appear in the work of *Tsou et al.* [1995]. These provide a more stringent test of the analysis and of the value of the microwave technique for long-term stratospheric monitoring.

Acknowledgments. We gratefully acknowledge the technical and logistical assistance of William B. Ricketts of JPL and of the entire staff of the JPL Table Mountain Facility. The work was supported by the NASA Upper Atmosphere Research Program.

References

Barnett, J. J., and M. E. Corney, Middle atmosphere reference model derived from satellite data, *MAP Handb.*, *16*, 47–85, 1985.

Bevilacqua, R. M., D. F. Strobel, M. E. Summers, J. J. Olivero, and M. Allen, The seasonal variation of water vapor and ozone in the upper mesosphere: Implications for vertical transport and ozone photochemistry, *J. Geophys. Res.*, *95*, 883–893, 1990.

Brillet, J., A theoretical study of ozone measurements made with ground-based microwave sensors, *J. Geophys. Res.*, *94*, 12,833–12,850, 1989.

Connor, B. J., and H. E. Radford, Pressure broadening of millimeter-wave ozone lines by atmospheric gases, *J. Mol. Spectrosc.*, *117*, 15–29, 1986.

Connor, B. J., and C. D. Rodgers, A comparison of retrieval methods: Optimal estimation, onion-peeling, and a combination of the two, in *Advances in Remote Sensing Retrieval Methods*, edited by A. Deepak, H. E. Fleming, and J. S. Theon, pp. 271–281, A. Deepak, Hampton, Va., 1989.

Connor, B. J., J. W. Barrett, A. Parrish, P. M. Solomon, R. L. de Zafra, and M. Jaramillo, Ozone over McMurdo Station, Antarctica, austral spring 1986: Altitude profiles for the middle and upper stratosphere, *J. Geophys. Res.*, *92*, 13,221–13,230, 1987.

Connor, B. J., D. E. Siskind, J. J. Tsou, A. Parrish, and E. E. Remsberg, Ground-based microwave observations of ozone in the upper stratosphere and mesosphere, *J. Geophys. Res.*, *99*, 16,757–16,770, 1994.

de Zafra, R., M. Jaramillo, J. Barrett, L. K. Emmons, P. Solomon, and A. Parrish, New observations of a large concentration of ClO in the springtime lower stratosphere over Antarctica and its implications for ozone-depleting chemistry, *J. Geophys. Res.*, *94*, 11,423–11,428, 1989.

Komhyr, W. D., B. J. Connor, I. S. McDermid, T. J. McGee, A. D. Parrish, and J. J. Margitan, Comparison of STOIC 1989 ground-based lidar, microwave spectrometer, and Dobson spectrophotometer Umkehr ozone profiles with ozone profiles from balloon-borne ECC ozone sondes, *J. Geophys. Res.*, this issue.

Lichtenstein, M., J. J. Gallagher, and S. A. Clough, Millimeter wave spectrum of ozone, *J. Mol. Spectrosc.*, *40*, 10–26, 1971.

Margitan, J., et al., Stratospheric Ozone Intercomparison Campaign (STOIC): Overview, *J. Geophys. Res.*, this issue.

Marks, C. J., and C. D. Rodgers, A retrieval method for atmospheric composition from limb emission measurements, *J. Geophys. Res.*, *98*, 14,939–14,953, 1993.

McGee, T. J., R. A. Ferrare, D. N. Whiteman, J. J. Butler, J. F. Burris, and M. A. Owens, Lidar measurements of stratospheric ozone during the STOIC campaign, *J. Geophys. Res.*, this issue.

National Aeronautics and Space Administration (NASA), Earth Science and Applications Division, in Report of the International Ozone Trends Panel, chap. 3, *WMO Rep. 18*, 1988.

Parrish, A., R. L. de Zafra, P. M. Solomon, and J. W. Barrett, A ground-based technique for millimeter-wave spectroscopic observations of stratospheric trace constituents, *Radio Sci.*, *23*(2), 106–118, 1988.

Parrish, A., B. J. Connor, J. J. Tsou, I. S. McDermid, and W. P. Chu, Ground-based microwave monitoring of stratospheric ozone, *J. Geophys. Res.*, *97*, 2541–2546, 1992a.

Rodgers, C. D., Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*, 5587–5595, 1990.

Tsou, J. J., B. J. Connor, A. Parrish, I. S. McDermid, and W. P. Chu, Ground-based microwave monitoring of middle atmospheric ozone: Comparison to lidar and SAGE II satellite observations, *J. Geophys. Res.*, *100*, 3005–3016, 1995.

B. J. Connor and M. P. McCormick, Atmospheric Sciences Division, Mail Stop 401B, NASA Langley Research Center, Hampton, VA 23681-0001.

A. Parrish, Department of Physics and Astronomy, University of Massachusetts, Amherst, MA 01003.

J.-J. Tsou, Lockheed Engineering and Sciences Company, Hampton, VA 23665.

(Received May 28, 1993; revised January 11, 1994;
accepted January 11, 1994.)