Uncertainties in upper stratospheric ozone trends from 1979 to 1996

D. M. Cunnold,[1] M. J. Newchurch,[2] L. E. Flynn,[3] H. J. Wang,[1] J. M. Russell,[4] R. McPeters,[5] J. M. Zawodny,[6] and L. Froidevaux[7]

Abstract. The time series of differences in coincident measurements of ozone by Stratospheric Aerosol and Gas Experiment (SAGE) and by Solar Backscattered Ultraviolet (SBUV), SBUV/2, Umkehr and Microwave Limb Sounder (MLS) are analyzed, and the slopes in the differences are calculated. SAGE ozone measurements are also compared against those by HALOE. The purpose of these comparisons is to look for statistically significant nonzero slopes which could indicate long-term calibration problems in one or more of the measurement systems. It is found that the slopes are remarkably similar between the Northern and Southern Hemisphere midlatitudes, and, apart from a few exceptions, the slopes are also similar in the tropics. Slopes of MLS-SAGE differences and HALOE-SAGE trends from approximately 1992 to 1996 have values of approximately $-0.5 \pm 0.4\%$ yr^{-1} (95% confidence limits) in Umkehr layers 7–9 (which are centered at ~37, 42, and 47 km altitude). Umkehr-SAGE slopes for 1979–1996, however, are almost all positive and in the range -0.1–0.4% yr^{-1} for Umkehr layers 4–8, while SBUV-SAGE slopes for 1979–1989 are essentially zero in layers 4–7 and 0.3–0.4% yr^{-1} in layers 8 and 9. Averaging all these results with SBUV-SAGE II slopes from 1985 to 1989, the other sensors minus SAGE slopes are most likely between 0.2 and -0.2% yr^{-1} from ~20 to 40 km altitude. The results indicate slightly negative slopes in Umkehr layers 5–7 and positive slopes in the other three layers. There thus appears to be no overall drift in the SAGE ozone measurements from 1979 to 1996, but SAGE sunrise/sunset trend differences >40 km altitude, combined with the more accurate SBUV-SAGE slopes for 1979–1989, suggest a most likely slope range of 0.4 to -0.4% yr^{-1} between 40 and 50 km altitude. SBUV/2 measurements from 1989 to 1994 have an upward trend with respect to SAGE measurements of ~0.7% yr^{-1} with some altitudinal structure; this slope exceeds the estimated 95% uncertainties on the SBUV/2 trends.

1. Introduction

Ozone decreases in the upper stratosphere since 1979 have been established in a series of papers including those based on Solar Backscattered Ultraviolet (SBUV) observations [*Hood et al.*, 1993; *Hollandsworth et al.*, 1995] and on Stratospheric Aerosol and Gas Experiment (SAGE) observations [*McCormick et al.*, 1992; *Wang et al.*, 1996]. These decreases are related to the accumulation of chlorofluorocarbons in the atmosphere [*Montzka et al.*, 1996; *Cunnold et al.*, 1997], and the decreases have been fairly successfully simulated by *Jackman et al.* [1996] and others. However, in the lower stratosphere, the ozone losses are less well understood (but see *Solomon et al.* [1996]), and their magnitude is less well established by the observa-

tions. This situation, and especially the wide range of ozone loss estimates between 15 and 20 km altitude [*World Meteorological Organization* (*WMO*), 1994], resulted in a reassessment of ozone trends by the Stratospheric Processes and their Role in Climate (SPARC) group. This paper is one of a series of papers [*Randel et al.*, 1999; M. J. Newchurch et al., unpublished data, 1998; *Cunnold et al.*, this issue; *Logan et al.*, 1999] that have resulted from that study (see also the SPARC report by *Harris et al.* [1998]).

This report focuses on the quality of the ozone measurements, and it assesses the accuracy of the upper-stratosphere ozone trend estimates. This assessment is based primarily on examining a number of time series of differences between coincident ozone measurements. Linear trends (i.e., regression lines) were fitted to each of these time series. To distinguish these trends in the sensor-versus-sensor differences from trends in ozone, throughout this paper, we refer to these linear regression fits of the coincident-measurement differences as slopes. The reason for analyzing the time series of coincident measurement differences (in contrast to discussing the trends in individual ozone time series, which are reported by M. J. Newchurch et al., unpublished data, 1998) is to eliminate the possible effects of sampling differences and more especially to remove most of the real variability in ozone from the time series (e.g., seasonal cycles and quasi-biennial oscillations). The coincident differences may then simply be fitted by linear regression slopes, as opposed to the fairly complex analytical

[1]School of Earth and Atmospheric Sciences, Georgia Tech, Atlanta.
[2]Atmospheric Sciences Department, University of Alabama, Huntsville.
[3]NOAA/National Environmental Satellite, Data and Information Services, Camp Springs, Maryland.
[4]Center for Atmospheric Science, Hampton University, Hampton, Virginia.
[5]NASA Goddard Space Flight Center, Greenbelt, Maryland.
[6]NASA Langley Research Center, Hampton, Virginia.
[7]Jet Propulsion Laboratory, Pasadena, California.

models that are used to fit the individual ozone time series. The resulting regression-slope estimates will then be more accurate indicators of instrument drift than the individual ozone trend estimates could provide.

Because SAGE provides the most extensive (in time and altitude range) satellite data set, all differences are expressed as [other instrument minus SAGE)/SAGE] times 100% except where otherwise indicated. In this paper, we report comparisons made against NIMBUS-7 SBUV, NOAA-11 SBUV/2, and Dobson Umkehr observations. Upper Atmosphere Research Satellite (UARS) Microwave Limb Sounder (MLS) and Halogen Occultation Experiment (HALOE) ozone measurements are also examined to assess the level of the constraints on long-term trends in ozone provided by these instruments. The only extensive long-term data on ozone profiles not considered here are the ozonesonde records. These observations are discussed in the SPARC report [see *Russell et al.*, 1998]. This omission is consistent with the focussing of this paper on ozone trends in the upper stratosphere. However, *Logan et al.* [1999] and *Cunnold et al.* [this issue] discuss ozonesonde results in the lower atmosphere.

This paper will first briefly discuss the possible sources of systematic trend error in each of the measurement systems, followed by the discussion of the slopes in the coincident differences of ozone measurements. A major objective of this study is to assess possible uncertainties in the ozone trends derived from individual satellite instruments, especially those from SAGE, which are caused by instrumental deterioration and resulting calibration uncertainties. This error assessment is intended to support the upper stratosphere ozone trend results given by M. J. Newchurch et al. (unpublished data, 1998).

2. Sources of Trend Uncertainty

In this section, instrumental sources of systematic uncertainty in the ozone trends are discussed. We focus on those sources of potential error which are likely to dominate the ozone trend uncertainties. Our approach differs slightly from that in SPARC Chapter 1 [*Hofmann et al.*, 1998] in that several of our uncertainty estimates are more directly related to previously published results. Moreover some of the uncertainties discussed by *Hofmann et al.* [1998] are omitted from this section because they are accounted for elsewhere in our analysis. For example, inaccuracies in the measurements from a single Umkehr instrument will mostly be captured in the differences between the measurements when an ensemble of Umkehr instruments are used. Throughout this paper, all the error bars discussed and shown are 95% confidence limits.

2.1. SAGE Version 5.96 Measurements

SAGE II retrieval version 5.96 was used in these studies. It still suffers from some of the limitations discussed by *Wang et al.* [1996]; these limitations contribute to the variability of the ozone observations. Compared to version 5.93 used by Wang et al., version 5.96 contains improvements in NO_2 sunrise retrievals and in the derived long-term trends in NO_2, but these changes are expected to result in only very minor changes in the ozone trends. More important are the changes made for the 5.96 retrievals in order to improve the separation of ozone and aerosols. These changes resulted from an improved representation of aerosols produced by the incorporation of a more realistic simulation of typical aerosol size distributions and the wavelength dependence of the extinction by aerosols.

The improvement is most evident during high atmospheric aerosol loading conditions (e.g., 1991–1993) below 20 km altitude [see *Cunnold et al.*, this issue]. Above 20 km altitude the changes in ozone are small, and they should typically produce negligible changes in upper-stratospheric ozone trends. Nevertheless, aerosol-ozone separation remains difficult under high aerosol loading conditions, and ~1 year of SAGE ozone data have been removed in all the trend studies between 20 and 30 km altitude following the Mt. Pinatubo eruption (see *Cunnold et al.* [this issue] and *Russell et al.* [1998] for more details). More specifically, below 30 km altitude, data from June 1991 through December 1991 have all been eliminated. The period of rejected data is longer with decreasing altitude and latitude. At latitudes equatorward of 35°, data are only used after June 1992 at 21 mbar, after December 1992 at 32 mbar, and after June 1993 at 46 mbar. At latitudes poleward of 35°, data are only used after June 1992 at 32 mbar and after December 1992 at 46 mbar.

As illustrated by *Cunnold et al.* [this issue], after removing SAGE ozone data at 46.4 mbar from mid-1991 to mid-1993 at midlatitudes and until the beginning of 1994 in the tropics, comparisons of SAGE and MLS measurements in the lower stratosphere suggest that the aerosol contamination has essentially been removed. The filtering that has been applied is found to be closely equivalent to removing ozone data when local aerosol extinctions exceed 3×10^{-3} km^{-1} at 525 nm [*Cunnold et al.*, this issue]. In order to estimate any potential residual error, SAGE ozone trends have been compared against those calculated using an aerosol extinction cut off of 2×10^{-3} km^{-1}. Restricting the discussion to altitudes above 19 km (the bottom of Umkehr layer 4), the trend differences for the 1979–1996 period are very small and only approach 0.1% at 22 km altitude in the tropics. For the UARS data period, 1992–1996, however, the differences are ~0.5% yr^{-1} in the 19–24 altitude range (Umkehr layer 4), and they exceed 1% in this layer in 15°N to 15°S (where the accepted record is only 3 years long). We reemphasize that the estimates in Table 1 are based upon the rejection of SAGE lower stratospheric ozone data as described above and the uncertainties refer to trend estimates over the entire 1979–1996 period.

SAGE I data should not have been affected by aerosols because of the low atmospheric aerosol loading in 1979–1981. However, several analyses [e.g., *Newchurch et al.*, 1995], including in particular *Wang et al.* [1996], have demonstrated that SAGE I data possess a reference height error of roughly 300 m. For these (as well as other SPARC) studies, all SAGE I profiles were shifted vertically upward by the latitude-dependent amounts calculated by *Wang et al.* [1996]. This is an ad hoc approach to adjusting for this height error because there should also be an adjustment made for the altered Rayleigh scattering contribution. However, as given by *Wang et al.* [1992], above 21 km altitude, the Rayleigh term is <20% of the ozone contribution at 0.6 μm. Therefore a 300 m height adjustment would produce an effect on ozone of <1% (the effect is larger below 20 km altitude as indicated in Table 1.2.9 of *Hofmann et al.*, 1998). The effects on the trends of the uncertainties in these altitude shifts (~100 m) have been included by repeating the calculations with different altitudinal shifts. They amount to ~0.1% yr^{-1} from 30–38 km altitude and to ~0.15% yr^{-1} above 40 km altitude for the period 1979–1996. Larger trend uncertainties would apply to shorter analysis periods when SAGE I data are used. This source of uncertainty has been included in the trend error bars reported by M. J. New-

Table 1. Estimated Instrumental Systematic Ozone Trend Errors (95% Confidence Limits) for 1979–1996 Above 20 km Altitude for Several Observing Systems

Layer Number	Pressure Level, mb	Approximate Altitude, km	Possible Instrument Calibration Drift, % yr^{-1}					
			SBUV (1979–1989)	SBUV/2 (1989–1994)	Collection of Umkehrs (1979–1996)	SAGE (1979–1996)	HALOE (1992–1996)	MLS (1992–1996)
9	1–2	44–49	0.5	0.5	0.1	0.2	0.1	0.2
8	2–4	39–44	0.5	0.4	0.1	0.2	0.1	0.2
7	4–8	34–39	0.5	0.3	0.1	0.1	0.1	0.2
6	8–16	29–34	0.4	0.6	0.1	0.1	0.1	0.2
5	16–32	24–29	0.3	0.3	0.1	0.2	0.1	0.2
4	32–64	19–24	0.6	0.6	0.3	0.2	0.2	>0.2

The layer numbers in column 1 refer to their Umkehr designation. The only errors included here are those that are not contained in the residual variability after fitting regression lines to the coincident-pair differences between the ozone measurements by different systems. It should be noted that there may be an additional error in layers 6–9 if SAGE values are converted to pressure levels.

church et al. (unpublished data, 1998). Table 1 column 7 also includes this source of uncertainty.

Potential errors not accounted for in the SAGE II ozone trend analyses include a time-dependent scan-mirror calibration error and density/conversion errors. The scan-mirror problem is that the mirror used to scan across the sun has deteriorated slightly and has developed a small angular sensitivity in the 1990s. Above 50 km altitude, the atmospheric extinction in the SAGE channels is small, and the small deterioration in the scan mirror can produce a significant extinction error if the two measurements of the solar irradiance, which are used to define the extinction, are not striking the mirror in exactly the same way. From the variability in the SAGE measurements of the solar irradiance above 65 km altitude, the SAGE II scan mirror changes amounts to a factor of 2 at 65 km altitude, and its effect decreases in proportion to the increase of ozone density. Adjustments have been made for this effect in the SAGE II retrieval algorithm. We believe the remaining uncertainty in the ozone concentrations is <10% at 65 km altitude and, correspondingly, <1% at 50 km altitude. The 5% figure given by *Hofmann et al.* [1998] was intended to represent an average over the 50–65 km region. This suggests a possible ozone trend error at the highest altitude of this study (~47 km) of ~1.7% yr^{-1} for the 1992–1996 period and of less than that for 1985–1996.

SAGE ozone measurements are dependent on National Center for Environmental Prediction (NCEP) measurements of temperature versus pressure for the following reasons: first, to remove the Rayleigh scattering contributions to atmospheric extinction and second, if a comparison is made against ozone mixing ratios as a function of pressure (e.g., SBUV observations), for converting SAGE measurements of ozone concentrations as a function of altitude to mixing ratios as a function of pressure. For the determination of ozone trends, neither a possible bias nor a random error in the NCEP temperature measurements is of much consequence, but an incorrect long-term trend in those measurements is important. *Wang et al.* [1992] have shown that Rayleigh scattering contributes ~1/7 of the extinction at 0.6 μm of 50 km altitude. Moreover, *Cunnold and Wang* [1998] have suggested that the 300–400 m change in the geopotential height in the tropical upper stratosphere from 1987 and 1994 is incorrect. If there was in reality no change whatsoever, an error in the density change of ~5% at a fixed altitude would be present. However, because the Rayleigh scattering contribution to the extinction

measured by SAGE is only a small fraction of the total extinction, the possible error in the SAGE ozone trend from 1979–1996 is roughly 1/7(5%)/17 years or significantly <0.1% yr^{-1}. In addition, in converting SAGE ozone concentrations to pressure levels, an error in the ozone change of ~7% (based on an ozone scale height of 5 km) could occur, and hence an error in the ozone trend from 1979 to 1996 of ~0.4% yr^{-1} in the tropical upper stratosphere is possible (and perhaps as much as 1% yr^{-1} over the period from 1987 to 1994). The errors in this NCEP time series may be the result of the discontinuance of tropical temperature measurements by rocketsondes after the mid-1980s. This interruption did not occur at midlatitudes, and there is no evidence of geopotential-height changes of anywhere near this magnitude at midlatitudes.

Hofmann et al. [1998] list a few other minor sources of possible error in SAGE ozone trends above 20 km altitude. However, all of these errors are <0.1% yr^{-1} (a number of sources of error that only affect SAGE retrievals below 20 km altitude will not be discussed here). The errors in Table 1 (column 7) are those that need to be considered over and above confidence limits placed on the statistical trend or slope estimates derived directly from the analysis of the ozone measurements.

2.2. SBUV Version 6 Measurements

The uncertainty in ozone trends reported by the SBUV instrument arises primarily from the deterioration of the diffuser plate and the accuracy with which this deterioration can be described. The sources and implications of this uncertainty are quantified by *Bhartia et al.* [1995] based on *Herman et al.* [1991], and we assume here that those listed uncertainties are good approximations to the 95% confidence limits. Table 1 (column 4) shows the resulting possible calibration drift for SBUV for the 1979–1989 period. There are also small errors in the trends in SBUV layers (which are similar to Umkehr layers) that arise because the SBUV averaging kernels are >5 km wide. This complication is discussed by *Miller et al.* [1997]; above 30 km altitude the error is small compared to the diffuser plate error, but it is roughly 0.5% yr^{-1} in layer 4 (20–25 km altitude). Table 1 includes the error due to the averaging kernel. Some researchers may argue that layer 4 does not belong in an upper-stratosphere paper. However, it has been included in most of the figures because it is the lowest level for which SBUV and Umkehr observations provide some infor-
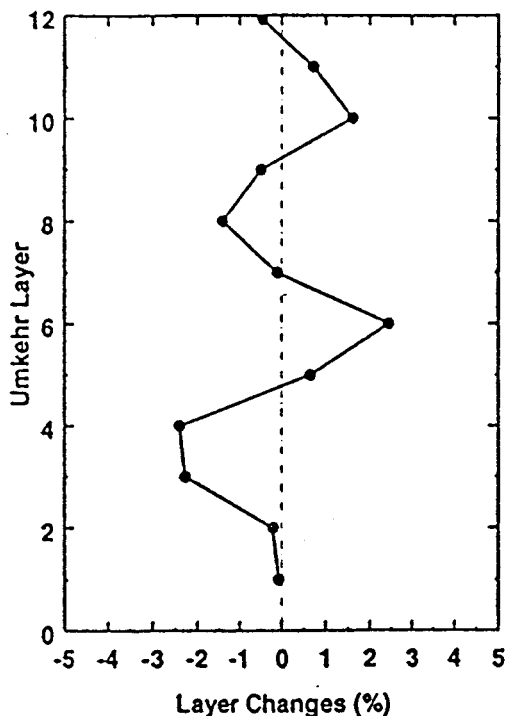
Figure 1. An illustration of how an assumed SBUV/2 calibration error that does not vary with time can produce an apparent time and altitude dependent trend as a result of an orbit that drifts to higher solar zenith angles (from 30° to 60°) [*Hofmann et al.*, 1998, Figure 1.23].

mation on ozone trends, and it would not seem useful to move this one level to the lower stratosphere papers.

In the ensuing, SBUV/2 instruments, a calibration lamp was included to check the diffuser plate calibration. In the NOAA-11 data from 1989 to 1994 used in these trend studies, good calibration was achieved, as demonstrated in annual comparisons against a well-calibrated SBUV instrument on the space shuttle. These instrument and cross-calibrations against SBUV/2 are described by *Hilsenrath et al.* [1995]. Comparisons of albedo measurements indicate a possible drift in SBUV ozone at 1 mbar of 0.5% yr^{-1} from 1989 to 1994 and of 0.3% yr^{-1} at higher pressures. In addition, in contrast to the Nimbus-7 SBUV orbit, the NOAA-11 SBUV/2 orbit drifts relative to the terminator. This results in solar zenith angle (SZA) changes of the observations at each latitude. These SZA changes can combine with interwavelength calibration errors in the SBUV/2 to produce ozone trend errors. Figure 1 shows an example of the type of errors which can occur. It presents model calculations of the effects of a hypothetical set of interwavelength calibration errors (consistent in magnitude with albedo differences discussed by *Hilsenrath et al.* [1995] and observed measurement residuals in SBUV/2 data), combined with the SZA changes from 1989 to 1994 in viewing conditions at the Equator for NOAA-11 SBUV/2. Table 1 (column 5) summarizes the net trend uncertainties for SBUV/2, which are estimated to be similar to those for the SBUV instrument.

2.3. Umkehr Measurements

The most important uncertainties that affect Umkehr ozone trends are instrumental calibration uncertainties and stratospheric aerosol effects. Sources of error in the Dobson instru-

ment include optical alignment, optical wedge calibration, and detector noise [*Hofmann et al.*, 1998]. The Dobson instruments are calibrated against the world standard Dobson instrument 83, which maintains a long-term (>25 year) precision of approximately ±0.5% [*Komhyr et al.*, 1989; *Basher*, 1995]. However, the Dobson instruments are not directly calibrated in the Umkehr mode [*Hofmann et al.*, 1998]. Nevertheless, because of the recalibration efforts made from time to time, this uncertainty should average out, especially in the ensemble average of the trends from many instruments. Its effect should therefore be contained in the standard error of the ensemble of Umkehr trend estimates, and no contribution to the systematic trend error is needed.

Umkehr observations are affected by radiation which is scattered by the stratospheric aerosol layer. Thus large volcanic eruptions have been shown to affect ozone trends, particularly in layer 8 near 42 km altitude [*Mateer and DeLuisi*, 1992]. Three techniques have been used to correct for this effect [*Mateer and DeLuisi*, 1992; *Reinsel et al.*, 1994; *Newchurch and Cunnold*, 1994]. In order to reduce the resulting uncertainties, only observations for which the stratospheric aerosol optical depth is <0.02 have been used. This restriction results in corrections of less than ~6% in the layer 8 ozone amounts, and the corrected ozone amounts differ by <2% among the three correction techniques [*Hofmann et al.*, 1998]. Thus, even in layer 8, the residual error from the potentially incorrect removal of the aerosol effect (following the Mt. Pinatubo eruption) is <0.1% yr^{-1} over the 1979–1996 period. Moreover, even this value may overestimate the systematic error in ozone trends from Umkehr observations because trend results from two of the procedures are included in the analysis of M. J. Newchurch et al. (unpublished data, 1998).

The Umkehr averaging kernels are ~10 km wide and are even wider for altitudes below the ozone maximum. This broad resolution causes a vertical smearing of the ozone trends derived from Umkehr measurements [e.g., *Miller et al.*, 1997] and a reduction in the trends caused by the absence of trends in the a priori ozone profiles that are used in the Umkehr retrievals [*Mateer et al.*, 1996]. On the basis of the results of *Miller et al.* [1997], who calculated the ozone trend errors in simulated Umkehr observations again based on the expected atmospheric ozone changes since 1979, the trend errors in individual layers 4–8 are 0.27% yr^{-1} in layer 4, decrease to essentially zero in layer 6 and increase again to 0.13% yr^{-1} in layer 8.

Table 1 (column 6) shows the estimate of the systematic error in ozone trends obtained from an ensemble of Umkehr instruments. It is based on combining the aerosol removal uncertainty and the smearing error, with the smearing error being the dominant term.

2.4. HALOE Version 18 Measurements

HALOE and MLS measurements began more recently (at the end of 1991) than those from the other instruments discussed in this report. The experience with these instruments has therefore been less extensive than with the other instruments. The only known problem which can affect the ozone trends in the v18 HALOE retrievals is the contamination by the high aerosol concentrations that followed the Mt. Pinatubo eruption. The correction for aerosols in the HALOE retrievals for May 1992 at 34°N amount to ~5% of the ozone at 30 km altitude, 22% at 25 km altitude, and almost 150% at 20 km altitude [*Hofmann et al.*, 1998, Figure 1.11]. Comparisons against nonaerosol affected data sets have indicated that the

HALOE correction procedure is successful in removing 93% of the aerosol contamination [*Steele and Turco*, 1997; *Hofmann et al.*, 1998]. On the basis of the decay of those aerosols which affect infrared wavelengths, a factor of ~2 decay per 200 days [*Hervig et al.*, 1996], the ozone trend uncertainty for the period of 1993–1996 can be calculated. The procedure used here is to estimate the residual aerosol-produced error in the HALOE ozone retrievals, equal to 100% − 93% = 7% of the aerosol correction, on January 1, 1993. Thus, for example, at 25 km altitude, the residual error is 7% of 22%/2 = 0.8% where the factor of 2 results from the decay of the aerosols in the period of ~200 days from May 1992 to January 1, 1993. This error can be extrapolated forward in time based on the decay rate, and the time series of errors can then be analyzed by a standard trend-fitting procedure to estimate the error in the HALOE ozone trends from 1993 to 1996. The estimated trend error is <0.1% yr^{-1} for altitudes above 23 km and ~0.4% yr^{-1} at 20 km altitude. Trend comparisons below ~25 km altitude should therefore be restricted to a starting date of no earlier than January 1, 1993 [also, see *Bhatt et al.*, 1999]. For comparisons against SAGE trends, this restriction is not a problem because SAGE measurements are more severely impacted by aerosols, and the filtered periods for SAGE following Pinatubo tend to be longer than those for HALOE.

Hofmann et al. [1998] suggest other possible sources of ozone trend error in the HALOE measurements. None of these sources amounts to >0.05% yr^{-1}, and there is no evidence for the presence of any of those errors. Nevertheless, we conservatively assign a small uncertainty of 0.1% yr^{-1} to HALOE ozone trends above 25 km together with an uncertainty of 0.2% yr^{-1} in the 19–24 km layer due to aerosols provided that HALOE measurements in this layer are not used before 1993.

2.5. MLS Version 4 Measurements

There are no known mechanisms for significant degradation in the MLS 205 GHz ozone measurements, although the temporal coverage has degraded over the years because of a combination of satellite problems and MLS antenna scanning problems (after the first 2 years of operation). Potential antenna surface degradation seems to be a small source of error, based on analyses of radiances from the highest altitudes of the MLS scans; these analyses (and the constant nature of the radiances) lead to an estimated upper limit for the change in antenna reflectivity (which has a value of 0.989) of 0.0001 over 6 years. This in turn would lead to much less than a 1% change in ozone over this time period. Potentially, the largest source of error arises from an imperfect knowledge of temperature trends, which could translate into an imperfect ozone trend. On the basis of comparisons (not shown here) between MLS-derived temperature trends in the tropics versus NCEP data in the 20–1 mbar range, a possible average temperature error is ~0.2 K/year (assuming that NCEP data were themselves to be perfectly accurate). The MLS data are therefore deemed "intrinsically" capable of very high stability, and the total contribution to ozone trend errors from possible error sources (at least for the middle and upper stratosphere) is believed to be ~0.2% yr^{-1} (see also the discussion by *Hofmann et al.* [1998]). Systematic issues in the lowermost stratosphere should make one more cautious about deducing trends there; the MLS version 5 data will provide a significant improvement in the systematic differences (versus SAGE II data).

3. Slopes Derived From Comparisons Between Instruments

As previously indicated, the comparisons between the time series of ozone observations will focus on the SAGE data, and the time series analyzed will in most cases consist of other instrument minus SAGE differences between coincident measurements expressed as percentages of the SAGE values. A particular emphasis of this report is on the question of whether the comparisons show any evidence of temporal changes of unknown origin in SAGE calibration. The presentation will include not only slopes of the differences in individual latitude-altitude bins but also latitudinally averaged slopes and sometimes globally averaged values. The procedure used to provide these averages is to take a mean over the slopes from the N latitudinal bins, and the standard error of the mean is then

$$\left[\left(\sigma^2 + \sum_{i=1}^{N} \sigma_i^2 \bigg/ N \right) \bigg/ (N-1) \right]^{1/2} \qquad (1)$$

were σ^2 is the variance of the N slopes and σ_i^2 is the variance of the slope estimate for the ith latitudinal bin. However the variances σ^2 and σ_i^2 have been increased to allow for the autocorrelation of the slope-fit residuals, and these residuals are represented by an autoregressive first order, AR(1), process, for which the correlation decreases as r^n, where r is the correlation for a 1 month delay and n is the number of months of the delay for the calculated autocorrelation. In these calculations it is found that the AR(1) term is typically significant for the σ_i^2 calculation only in Umkehr layers 5 and 6. However, the latitudinal correlation between the slope estimates is almost always large (and the altitudinal correlation is not), and the σ^2 therefore has been increased substantially. In the following, presentation of slopes in the differences between SAGE and other individual measurements of ozone the results will be presented in 10° latitude bins and for Umkehr layers. However, the latitudinally averaged results at each level will in most cases be deferred until the discussion at the end of the section where they can be simultaneously compared for all the instruments.

All comparisons (except those otherwise noted) are based on differences between coincident measurements of ozone. Coincidence criteria are generally ±1 day, ±2° latitude, and ±12.5° longitude. For comparisons against SAGE or HALOE measurements, these coincidence criteria will typically provide ~1 day of sunrises and 1 day of sunsets per month (~30 coincidence profiles). However, ozone differences are then averaged over 10° latitude bins; this means that the monthly averages in any bin are based on >60 coincident profiles.

3.1. Comparisons Against SBUV

For comparisons against SBUV, in addition to transforming the SAGE profiles to mixing ratio profiles on pressure surfaces, we have summed the ozone measurements over layers ~5 km thick in order to correspond to the layers used in the SBUV retrievals. The differences between layer-mean values for SBUV and SAGE I/II values are fairly systematically equal to a few percent in Umkehr layers 6–9 (~30–50 km, see Table 1). In layers 4 and 5 (~20–30 km altitude), there are differences of ±10% which are mostly related to limitations in the a priori profiles used in the SBUV retrievals [*Wang et al.*, 1996].

Slopes in the differences between SBUV and SAGE for the SAGE I period alone are not useful because this period was only 2.5 years long. SBUV and SBUV/2 provided overlapping
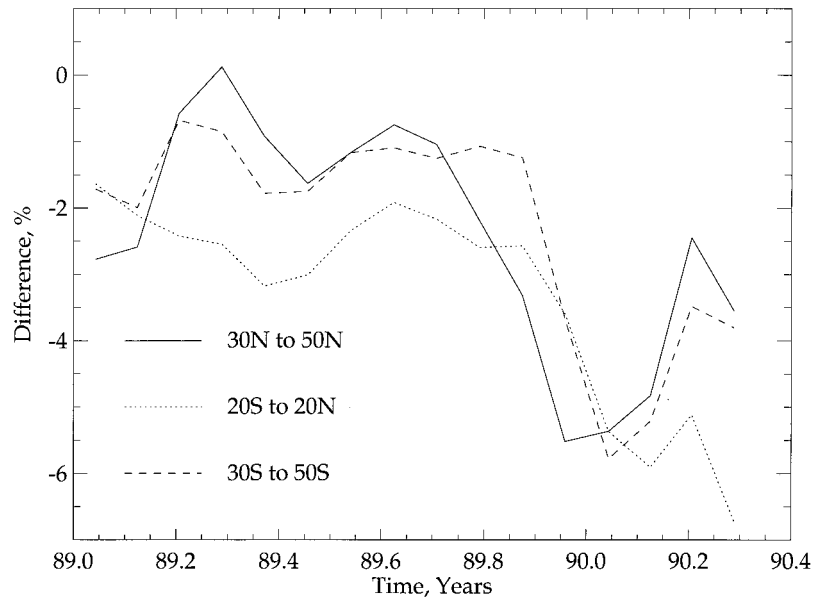
Figure 2. Time series of zonal/monthly mean differences (percent) between SBUV minus SBUV/2 for Umkehr layer 8. Solid line represents 30°–50°N, the dashed line represents 30°–50°S, and the dotted line is 20°–20°N [*Russell et al.*, 1998, Figure 2.2].

measurements in 1989 and part of 1990. Figure 2 shows a time series of the differences between ~2 and 4 mbar (Umkehr layer 8) during the overlap period. At other levels, the differences are of similar magnitude, but they are of opposite sign below ~25 mbar. Because of these obvious calibration differences and the uncertainties in SBUV/2 described in section 2, we chose to analyze the SBUV and the SBUV/2 time series separately. Moreover, it is evident from Figure 2 that something changed in one of the instruments at the end of 1989. There is evidence from D-pair total ozone checks of the internal calibration that the SBUV instrument was deteriorating at that time. SBUV comparisons are therefore terminated at the end of 1989. SBUV type observations are somewhat influenced by the presence of a massive stratospheric aerosol layer such as that which occurred shortly after the Mt. Pinatubo eruption [see *Torres and Bhartia*, 1995]. As a result of these considerations, the trend analysis of SBUV/2 data does not use ozone data from mid-1991 to mid-1992.

Figure 3 shows the slopes determined from SBUV and SAGE I/II differences over the period from 1979 to 1989. The slopes are seen to be essentially zero in layers 5 and 7, and there are fairly latitudinally consistent slopes of up to 0.5% yr^{-1}, in layers 6, 8, and 9. The smaller slopes in the tropics in layers 8 and 9 may be related to the tropical temperature/geopotential height uncertainties introduced in converting SAGE data to the mixing ratio versus pressure scale [*McPeters et al.*, 1994; *Cunnold and Wang*, 1997]. If the NCEP geopotential heights in the tropics in 1989 are too low, SAGE ozone values on the real pressure surfaces will be too high, and therefore slopes in SBUV-SAGE values will be slightly underestimated (see section 2.1).

There are no significant differences with latitude (at the 95% confidence level) in the slopes in Figure 3. Combining the latitudinal and altitudinal values into a single overall mean slope yields a slope of 0.1 ± 0.1% yr^{-1} (95% confidence limit) in these six layers (i.e., from ~19 to 49 km altitude). It is possible to interpret the significant differences between the

latitudinally averaged positive slopes in layers 8 and 9 and the negative slopes in layer 6 as evidence of small errors in the SBUV trends which are within the uncertainties given in Table 1. These errors can occur because different SBUV wavelengths are used to measure ozone in these layers, and, as was illustrated in the SBUV/2 comparisons (although we emphasize that SBUV did not have a drifting orbit), SBUV-type errors can have an effect of producing offsetting differences in these layers.

Figure 4 shows the comparisons of SAGE II data against SBUV data from 1984 to 1989. Similar results for approximately the same period have previously been reported by *McPeters et al.* [1994]. For this shorter period, there are clearly larger slopes in the differences. Again in layers 8 and 9 the SAGE data may be yielding excessive tropical trends on NCEP pressure levels. This period displays less overall consistency with latitude than for the 1979–1989 period (e.g., layers 4 and 9). However, excluding the low values at 0° latitude in layers 8 and 9, the larger slope variability with latitude is associated primarily with the expected factor of approximately three increase in trend uncertainties produced by a factor-of-two reduction in the period of comparison. There are some significant differences in the comparisons against SBUV for the two periods (Figures 3 and 4), which will be addressed in the summary discussion section 3.6. The overall mean of the slope difference slopes in layers 4–9 in Figure 4 is −0.1 ± 0.3% yr^{-1}.

Figure 5 shows the SBUV/2/SAGE II slopes of the differences over the period from 1989 to 1994. Here what is most significant is that the mean slopes at all the levels are positive; the positive slopes are found at almost every latitude, and there are no significant latitudinal differences in the slopes at any level. The similarity between the latitudinal structure in the slopes in layer 4 in Figures 4 and 5 might be related to deficiencies in SBUV(/2) a priori profiles that were referred to earlier. Overall the SBUV/2 trends are larger than SAGE trends by 0.7 ± 0.3% yr^{-1} (and this is associated with actual increases in ozone from the SBUV/2 observations over this
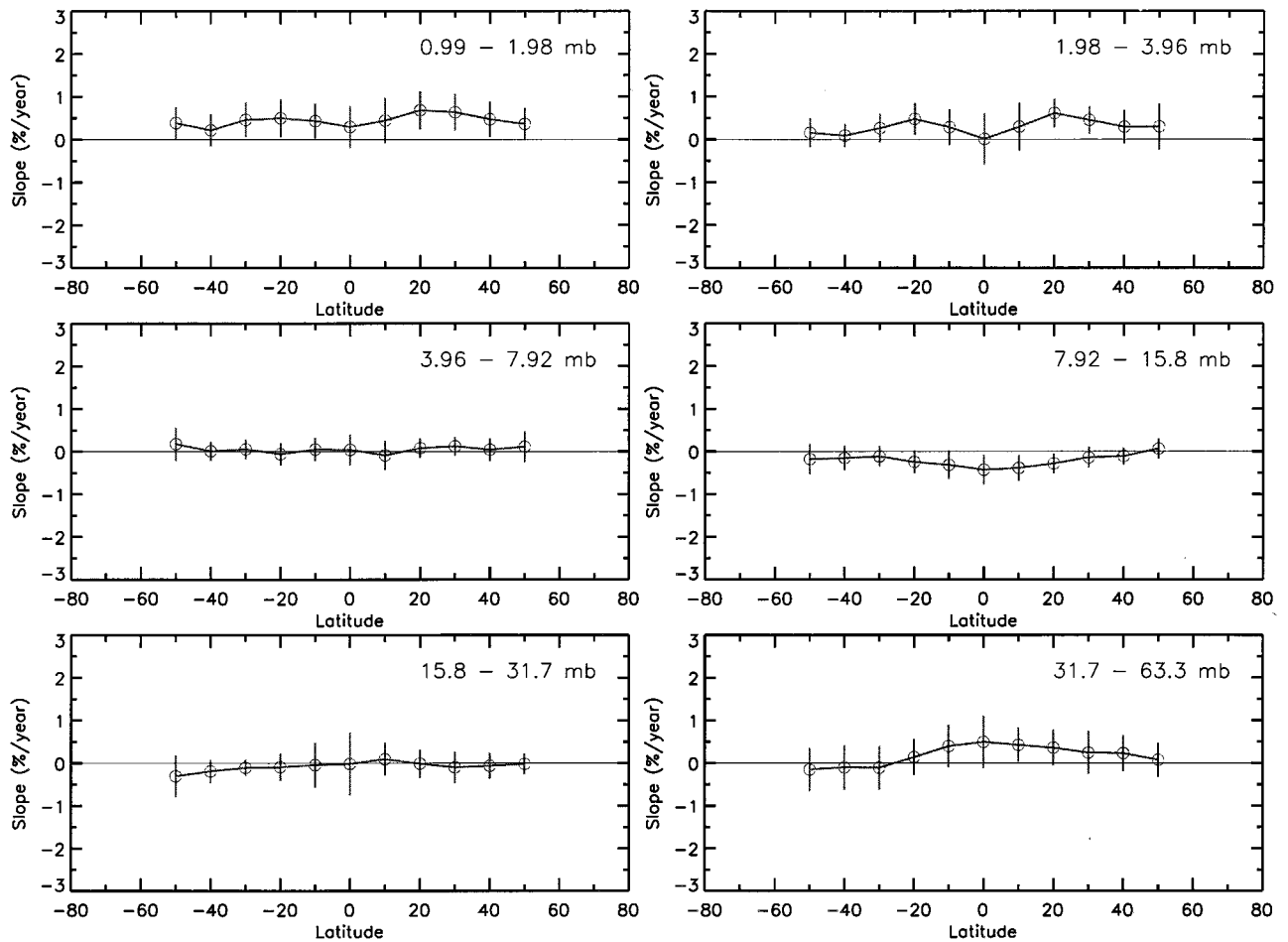
Figure 3. Linear slopes of the differences (SBUV-SAGE) between SAGE and coincident SBUV ozone observations from 1979 through 1989 expressed as a percentage of the SAGE Umkehr layer content. Mean values are binned by month and placed in 10° wide latitude bins. Error bars are twice the standard errors of the differences in slopes (as given by *Russell et al*. [1998, Figure 2.26] except that the error bars have been increased to account for temporal autocorrelation).

period). This result is probably explained by the calibration uncertainties identified in the comparisons between SBUV/2 and SSBUV, which were discussed in section 2.2. The latitudinal means of the slopes in layers 4–9 are 0.1 ± 0.8, 0.8 ± 0.6, 0.4 ± 0.6, 0.7 ± 0.3, 1.4 ± 0.4, and $0.6 \pm 0.3\%$ yr^{-1}; the error bars are larger than those by *Russell et al*. [1998] because the autocorrelation of the residuals have been incorporated here. The large slope in layer 8 and the small slope in layer 6 may be related to the drifting local time, and hence the changing solar zenith angle of the SBUV/2 observations, superimposed on the change in SBUV/2 calibration. This was illustrated in Figure 1 in a simulation for the SBUV/2 observing conditions. In this figure, the extremes of the changes in layers 6 and 8 are obvious, (although they have a sign opposite to that needed to explain the Figure 5 results). The conclusion is that the inferred (from SSBUV) drift in SBUV/2 calibration, and the drifting orbit, makes the current SBUV/2 retrievals questionable for inferring long-term changes in ozone. Independently, A. J. Miller (private communication, 1998) has found evidence of an upward drift in the SBUV/2 measurements in comparisons with measurements from ground-based LIDAR and microwave instruments. These SBUV/2 comparisons are not used

therefore in the assessment of possible drifts in SAGE calibration.

3.2. Comparisons Against HALOE and MLS Measurements

Between 1.5 and 43 mbar the mean ozone concentrations measured by MLS, SAGE II, and HALOE are only ~5% different with the MLS values being larger than the other two (see Figure 6). However, at 1 mbar, MLS values are equal to, or slightly less than, the SAGE II and HALOE values. MLS V04 ozone data are of uncertain quality at pressures >46.4 mbar and also at 46.4 mbar in the tropics [*Froidevaux et al*., 1996].

Slopes have been calculated from MLS/SAGE II differences and from MLS/HALOE differences between coincident measurements in UARS layers. Figure 7 shows the results for the period from 1991 to 1996 in UARS layers, which have been chosen to have centers at approximately the same midpressures as the Umkehr layers (the results at 5.62 mbar were obtained by averaging two UARS layers). Because of the post-Pinatubo SAGE filtering, the MLS/SAGE II time series is shorter at 46.4 mbar than at higher altitudes. For direct com-
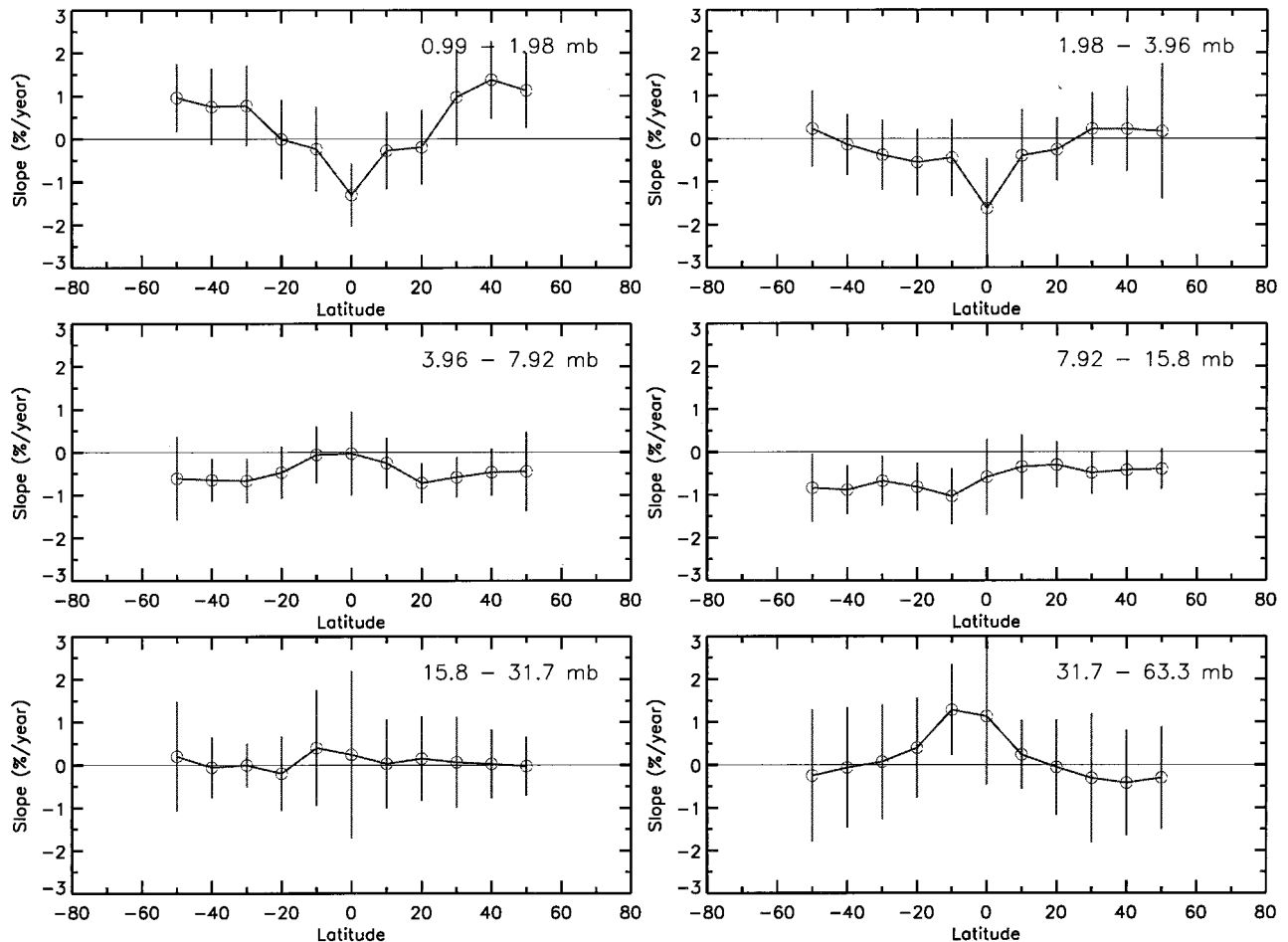
Figure 4. Mean slopes of SBUV-SAGE differences based on coincident SBUV measurements from 1984 through 1989. SAGE values are summed over Umkehr layers, and the differences are accumulated in monthly bins that are 10° wide in latitude. Slopes are expressed as percentages of the SAGE mean values per year. SAGE sunrise and sunset differences are combined. Error bars are twice the standard errors of the slope estimates (as given by *Russell et al.* [1998, Figure 2.25] except that the error bars have been increased to account for temporal autocorrelation).

parability and because HALOE measurements were also affected by the Mt. Pinatubo aerosols [*Hervig et al.*, 1996], although to a lesser extent than SAGE, the MLS/HALOE comparisons were chosen to have the same starting times as the MLS/SAGE comparisons. It should be noted that these comparisons are less continuous than those against SBUV in part because of the 36-day gaps in the MLS coverage at middle and high latitudes but more particularly because of the MLS instrumental and satellite operating problems after 1994 which has resulted in only half the coincident comparisons compared to the years prior to 1994. Because of this sampling problem, which makes an auto-correlation analysis of the residuals less useful, no adjustment to the σ_i^2 (in equation 1) for an AR(1) term has been made for the MLS, HALOE and SAGE comparisons.

The slope results show that there are significant differences between the SAGE and MLS trends in coincident ozone measurements at some latitudes and levels (e.g., 21.5 mbar). However, this statement is equally as true for the MLS-HALOE comparisons over the same period. This is illustrated most dramatically at 21.5 mbar. Further investigation reveals that at 21.5 mbar the MLS coincident samples for HALOE and

SAGE agree well up to the end of 1994, but after that the MLS samples show larger differences than do the SAGE and HALOE time series themselves. The 95% confidence limits on the slopes are similar to those for the shorter period SBUV-SAGE comparisons, as might be expected because of the similar length of the periods, but there is some tendency for the MLS-SAGE residuals to be larger than the MLS-HALOE residuals at the four upper levels, and vice versa at 46.4 mbar. The latitudinal structure of the slopes do not seem to be correlated between SAGE and HALOE nor are there any obvious correlations with the SBUV-SAGE slopes. At 46.4 mbar the slopes, particularly for MLS-HALOE differences, are large and variable (note the scale change in Figure 7). This leads to uncertainties in the latitudinal mean slopes of ~2% yr^{-1} (95% confidence limit). Thus it is concluded that at this level, comparisons between SAGE, HALOE, and MLS provide no effective constraint on calibration uncertainties. This is due to known tropical deficiencies in the MLS data at this level and to the reduced period for which comparisons are possible because of the Mt. Pinatubo aerosol effects on SAGE (and HALOE) data in conjunction with substantially reduced operation of the MLS instrument after October 1994. For the upper
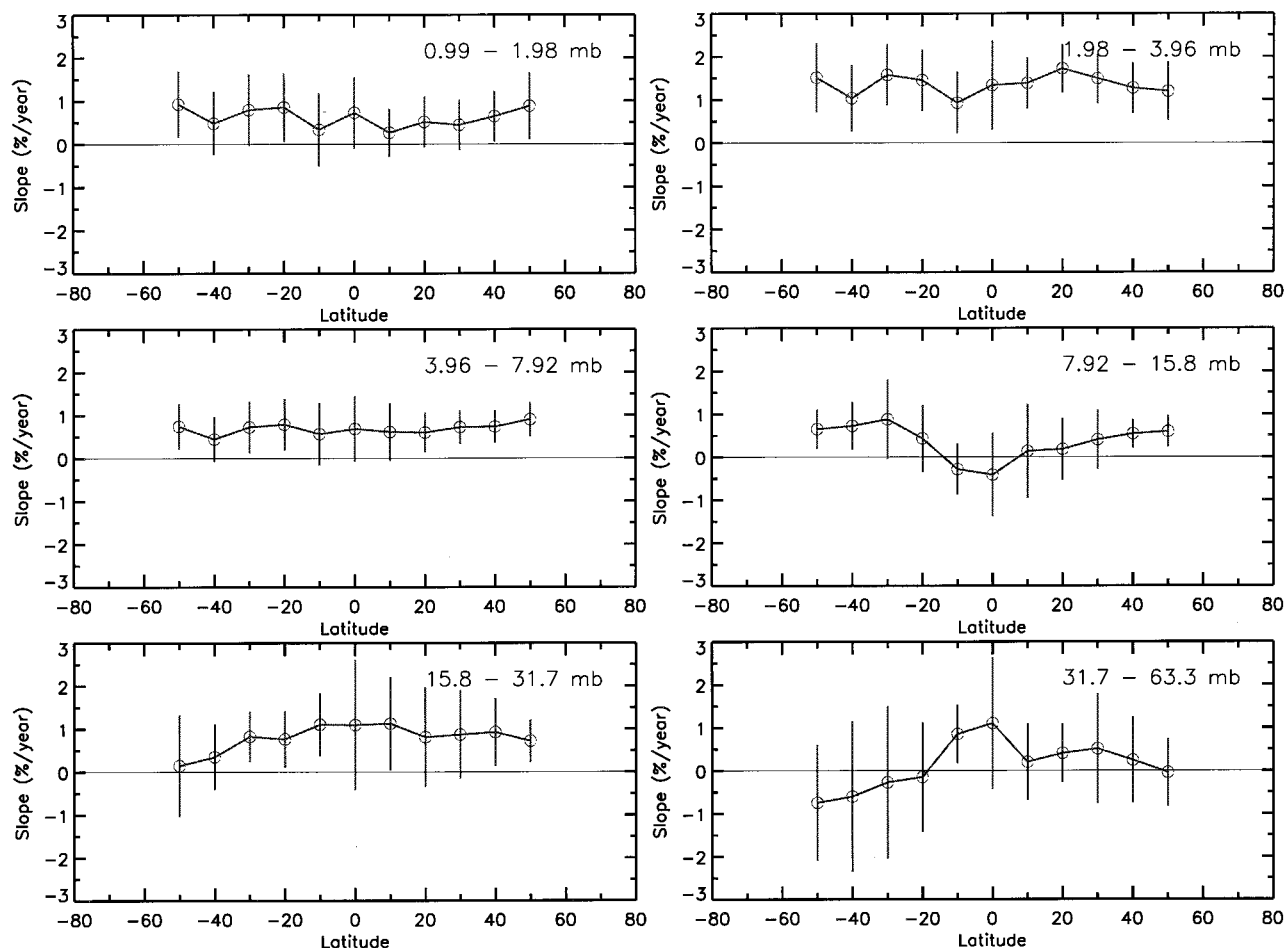
Figure 5. Mean slopes of SBUV/2-SAGE coincident differences in Umkehr layers and in 10° latitude bins over the period of 1989–1994. Slopes are expressed as percent of the SAGE values per year. SAGE sunrise and sunset observations are combined into a single time series. Error bars are twice the standard errors of the slope estimates (as given by *Russell et al.* [1998, Figure 2.24] except the error bars have been increased to account for the temporal autocorrelation).

five levels, the latitudinal-altitudinal mean is $-0.5 \pm 0.3\%$ yr^{-1} for MLS-SAGE coincidences and is $-0.7 \pm 0.5\%$ yr^{-1} for HALOE-SAGE data with MLS used as the basis for comparison.

The calculations just described examined HALOE/SAGE differences using MLS observations as a common basis for comparison. The resulting indirect approach to the comparison of HALOE and SAGE II trends was necessary because there have been very few coincident measurements by HALOE and SAGE. However, for completeness we include here a comparison of linear trends fitted to the SAGE and HALOE observations separately over similar time periods (Figure 8). In these calculations, seasonal cycle terms were included in the fitting procedure. The HALOE trends are more negative than the SAGE trends at 21.6 mbar, but at the other levels, excluding 46.4 mbar where the error bars are large, the slopes are quite similar. The trend difference, HALOE-SAGE, averaged latitudinally and over the upper five levels is $-0.2 \pm 0.3\%$ yr^{-1}; if the anomalous point at 1.46 mbar, 50°S is omitted, the mean is -0.1% yr^{-1}.

It may be noted that there are larger than expected differences between the HALOE-SAGE slopes in the two approaches described. The differences in a few latitude-altitude

bins have therefore been examined in more detail. As already indicated, the number of months for which data coincident with MLS exists is not large (~36 of the possible 72 months). For calculating the HALOE-SAGE slopes, it does not seem to matter whether trends are inferred separately from all the SAGE and HALOE data or from the subsets of the measurements taken at the same times when MLS measurements were also being made. However, when the MLS data are used in the comparisons, the HALOE-SAGE slope changes are found to arise primarily from differences between the MLS trends over the separate HALOE and SAGE sampled coincident measurements time periods, particularly after 1994 data and particularly at 21.5 mbar. It is also worth noting, that when SAGE and HALOE time series are analyzed separately, the amplitudes and phases of the annual and semi-annual terms are somewhat different in the lower stratosphere. The conclusion therefore is that there are sampling limitations which are affecting the slope estimates and that these do not seem to be adequately accounted for in the error bars. This might be associated with the absence of the AR(1) term in the analysis of the time series residuals caused by the irregular sampling.

The differences between the HALOE-SAGE-MLS comparisons presented here and those reported by *Russell et al.* [1998]
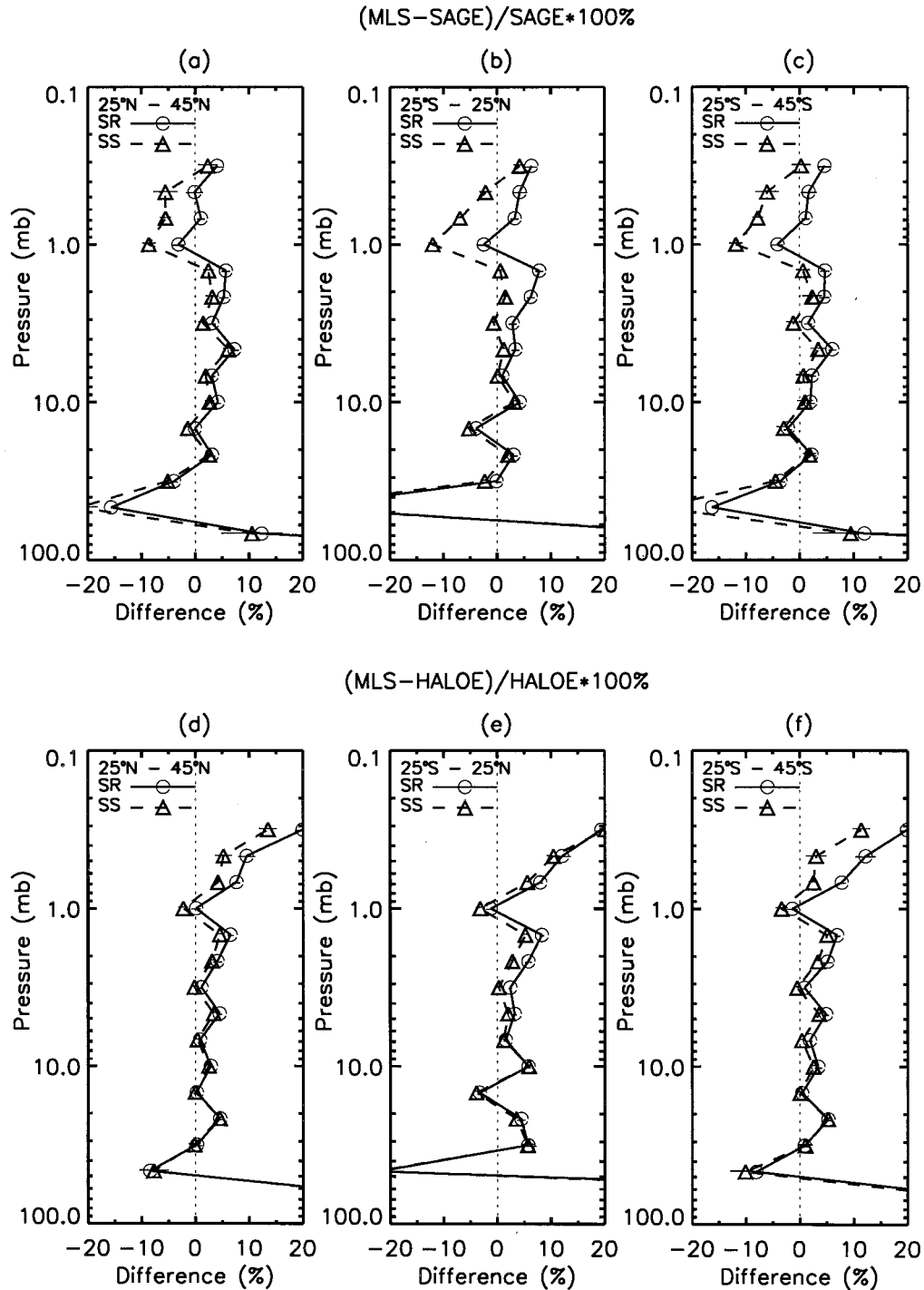
Figure 6. (a)–(c) MLS-SAGE II differences (d)–(f) MLS-HALOE differences for coincident nighttime ozone measurements expressed as percentages of the MLS values. SAGE and HALOE sunrises and sunsets are treated separately, and the data were placed in three separate bins that extended from 25° to 45°S, 25°S to 25°N and 25° to 45°N. The SAGE data are placed on pressure levels using the NCEP temperature profiles and then summed over UARS layers which have boundaries at $10^{3-1/6}$ mbar. The error bars are twice the standard errors of the differences. The period of comparison is October 1991 to December 1996, but the SAGE data are filtered as recommended to remove the Pinatubo aerosol interference effects [*Russell et al.*, 1998, Figure 2.35].

also should be mentioned. These occur because in this work later starting times have been used for the SAGE and HALOE time series at the lower two levels; this has eliminated some of the probably aerosol-contaminated SAGE data in 1992 and 1993 (as well as eliminating some possibly contaminated HALOE data [see *Bhatt et al.*, 1999]). There are also differences in the error bars; these arise from the inclusion of the AR(1) correlation term in the latitudinal averages and from
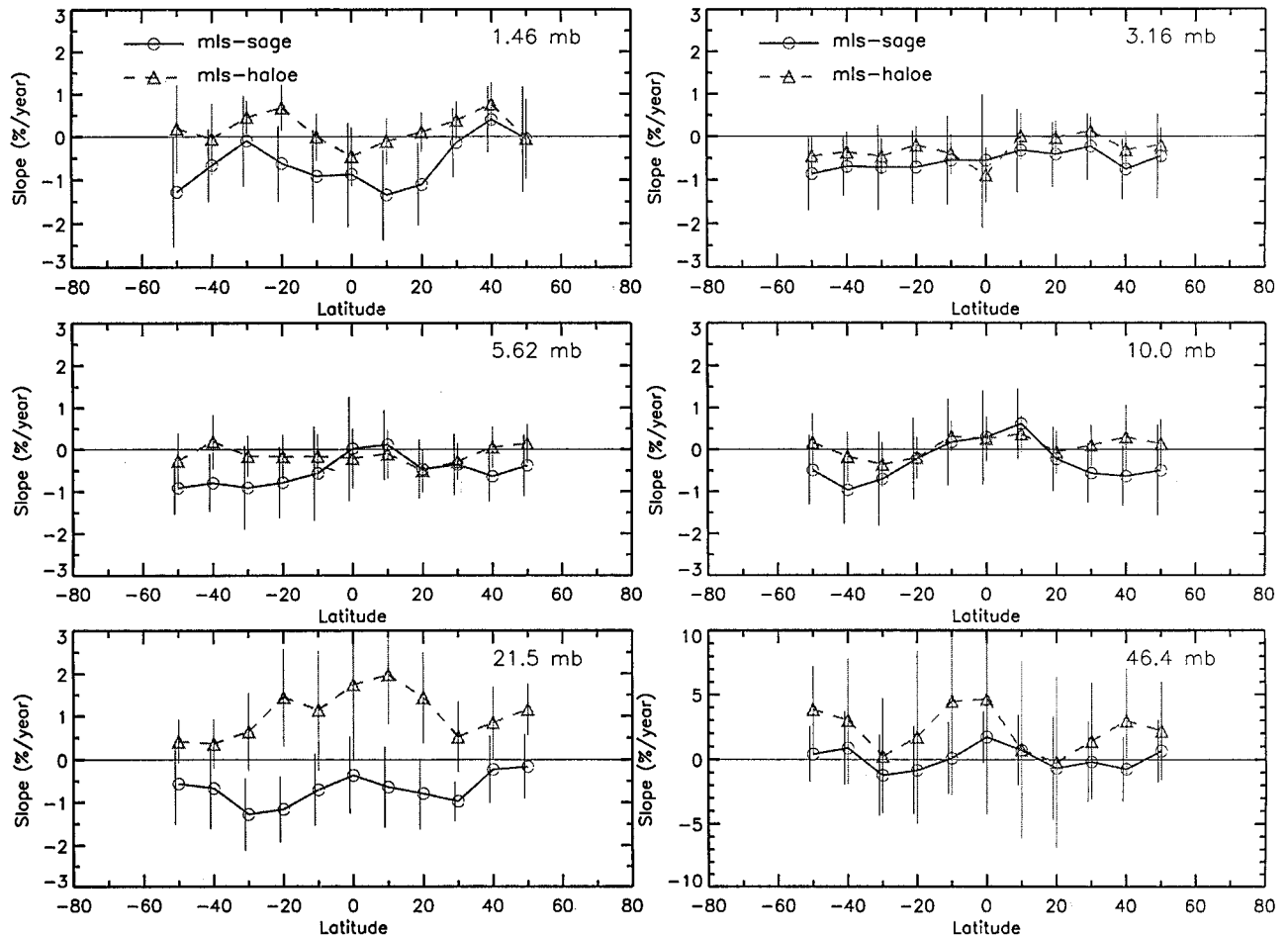
Figure 7. The slopes in the differences between coincident MLS ozone observations from 1991 to 1996 and SAGE II (line with circles) and HALOE (dashed line with triangles) in selected UARS layers in 10° latitude bins. The SAGE II data are filtered at 21.5 and 46.4 mbar by the recommended procedure following the Pinatubo eruption. HALOE comparisons have been started at these same times. The error bars are twice the standard errors of the slope estimates. Note the change of scale in the 46.4 mbar results.

some mistakes in the error bar estimates given by *Russell et al.* [1998]. It should be noted that the error bars for all the 5 year computations of slopes in the differences (SBUV, HALOE, MLS, SAGE) are now similar at individual levels.

A way to improve the precision of the HALOE-SAGE slope estimates is to use potential vorticities on isentropic surfaces to define coincidences between the two sets of measurements. By limiting the analysis to regions where ozone and potential vorticity change slowly (i.e., the lower stratosphere below ~25 km altitude), it is possible to obtain larger numbers of coincidences to produce more precise estimates of slopes of the differences. However, this technique requires some justification, and because it works best in the lower stratosphere, it is not discussed further in this paper (but see *Russell et al.* [1998]).

Figure 9 summarizes the latitudinal averages of the MLS or HALOE-SAGE slopes. The UARS layer slopes are discussed here, and in later figures, as if they also applied to Umkehr layers. As already stated, the layers are centered at approximately the same pressures and tests have indicated that there are only small changes in the slopes when they are recalculated over 5 km thick Umkehr layers instead of over the 2.5 km thick UARS layers. In spite of the unexpected differences between

the two ways of obtaining the HALOE-SAGE slopes when the error bars are considered, there is some consistency between the three sets of results. Ignoring layer 4 where the error bars are too large to provide useful results, these comparisons suggest that ozone trends in this period of ~5 years (1992–1996) are more negative than SAGE is estimating, although the results are less negative in direct comparisons between SAGE and HALOE than if MLS data are used. The overall mean slopes of the differences for layers 5–9, based on the three estimates times five layers, is $-0.4 \pm 0.4\%$ yr^{-1}.

3.3. Umkehr Comparisons

Umkehr observations were obtained from the World Ozone and Ultraviolet Radiation Data Center (WOUDC). These were inverted to produce ozone profiles on the *Bass and Paur* [1985] scale, using the *Mateer and DeLuisi* [1992] algorithm (using the uniform covariance matrix). The ozone comparison was restricted to those Umkehr profiles for which the observed and simulated radiance ratios at the two wavelengths used to measure ozone were in excellent agreement (an rms residual of less than ~1.5%); this results in agreement between the derived ozone column and that measured by the Dobson instrument of better than 1 DU in most cases. Table 2 lists the
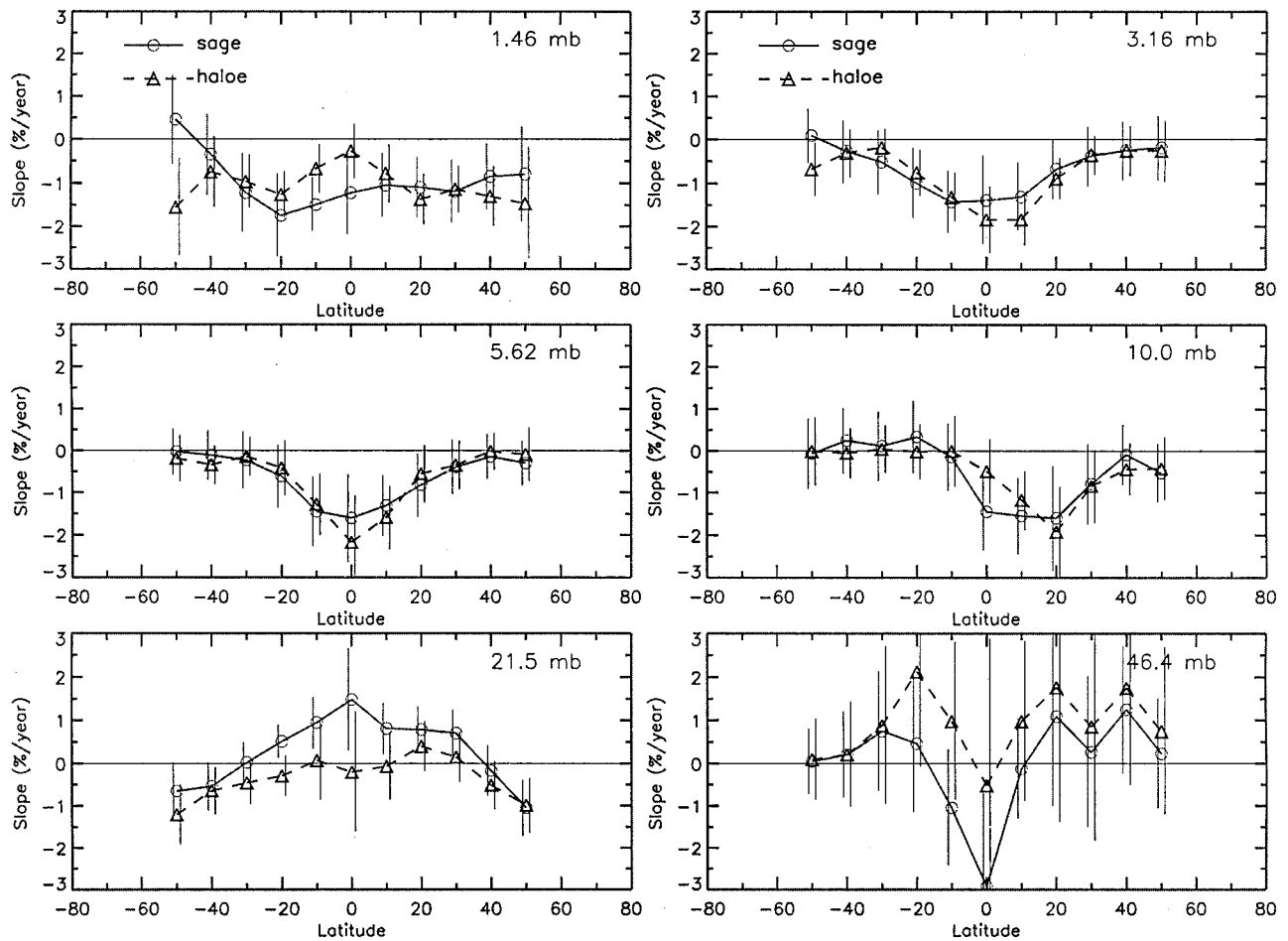
Figure 8. Differences between HALOE and SAGE II linear trends over the same periods as in Figure 7 and at the same locations. Annual and semiannual oscillations are removed during the linear fitting procedure. Note that this calculation, in contrast to all the other plots, was not based on coincident measurements.

Umkehr stations used in this analysis; a number of other records were rejected as being too short or because of fairly obvious discontinuities.

The comparisons [see, also, *Russell et al.*, 1998] were made for coincidence intervals of both 24 and 48 hours. Since no significant ozone trend differences were found in the two sets of results, the 48 hour results are presented because of their smaller standard errors in the monthly means.

Calculations were made with both the *Newchurch et al.* [1995, 1998b] and the *Mateer and DeLuisi* [1992] aerosol correction factors, but because the Mateer and DeLuisi factors also conserve total ozone, the latter factors are used in the results reported here. In the mean, above layer 4, Umkehr ozone values are 5–10% smaller than SAGE values [*Newchurch et al.*, 1998b; *Russell et al.*, 1998]. Because of this bias, SAGE measures larger ozone column above layer 2 by ~7%. The origin of these differences is not understood [see *Newchurch et al.*, 1998b].

Figure 10 shows the seven-station mean of Umkehr-SAGE differences and the 95% confidence limits based on the standard errors of those means. Results are shown for individual Umkehr layers 4–8, for layers 2–4 combined and for layers 8–10 combined. There is good agreement between the Umkehr and the SAGE trends with the mean slope of the
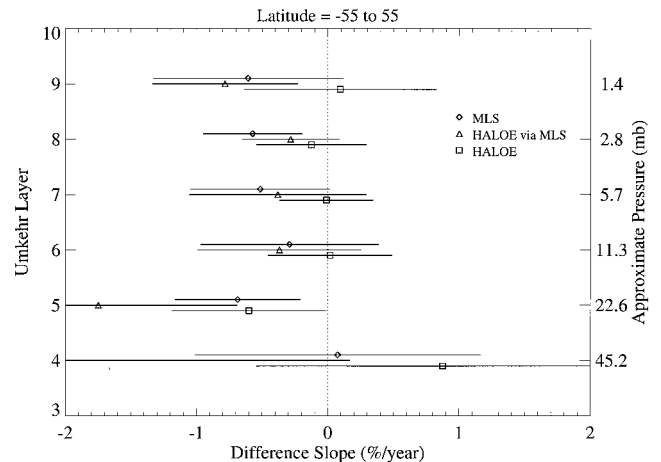

Figure 9. A summary of the latitudinally averaged slopes and 95% confidence limits for roughly the 1992–1996 period. The diamonds are for MLS-SAGE, the triangles are for HALOE-SAGE based on MLS data, and the squares are HALOE-SAGE trend differences. Results are displayed at Umkehr layer midpoints based on using UARS layers centered at approximately the same central pressures; the results at 5.62 mbar are obtained by averaging the results in two UARS layers.

Table 2. Information Concerning Umkehr Stations Considered in This Study

WMO Number	Station	Country	Lat. North	Long. East	Record Length from 1979–1996	Months Present, %	Stringent-convergence acceptance, %
10	New Delhi	India	28.65	77.22	March 1979 to Dec. 1995	61	50
152	Cairo	Egypt	30.08	31.28	Feb. 1979 to Nov. 1995	79	47
14	Tateno	Japan	36.05	140.10	Feb. 1979 to May 1994	85	57
67	Boulder	United States	40.03	−105.25	Feb. 1979 to Dec. 1995	93	64
40	Haute Provence	France	43.93	5.70	Sept. 1983 to Dec. 1995	69	60
68	Belsk	Poland	51.84	20.79	Feb. 1979 to Nov. 1996	66	55
159	Perth	Australia	−31.92	115.95	Feb. 1979 to Dec. 1995	76	67

The table also uses fraction of months observed during the study period and the fraction of observations that passed the more stringent inversion convergence criteria. This table contains a subset of the data by *Russell et al.* [1998, Table 2.6]. Lat., latitude; Long., longitude.

differences in layers 4–8 being $0.2 \pm 0.2\%$ yr^{-1} over the period 1979–1996. There was no AR(1) term included in the Umkehr analysis, but its absence is almost certainly not important because the variance of the slope estimates is dominated by the slope differences between the individual Umkehr site estimates.

Because there are many more coincidences between SBUV and Umkehr measurements than between SAGE and Umkehr measurements, it seems useful to also check the time series consisting of SBUV and Umkehr coincidences. Of course the time period covered in this case is only from 1979 to 1989. Figure 10 shows the calculated mean slopes and their two standard errors based on the seven Umkehr sites listed in Table 2. For layers 4–8 the mean slope is larger in the Umkehr observations than in SBUV observations by $0.1 \pm 0.2\%$ yr^{-1}. These two results are consistent with the excellent agreement already shown in the more direct comparisons between SBUV and SAGE. However, it should be emphasized that the

Umkehr sites used in these analyses are primarily located in midlatitudes of the northern hemisphere.

3.4. SAGE II Sunrise/Sunset Comparisons

Comparisons between SAGE II and MLS and SBUV coincidentally measured mean values have indicated differences between SAGE II sunrise and sunset measurements which are significant above ~40 km altitude and which can exceed 5% above 45 km altitude [*Wang et al.*, 1996; *Cunnold et al.*, 1996a, b]. These differences widen slightly in v5.96 retrievals (see Figure 6). SAGE I measurements, however, do not exhibit sunrise/sunset differences; this was determined from SAGE I comparisons against SBUV measurements. The absence of such differences might be related to sampling limitations resulting from the battery failure on SAGE I which limited sunrise measurements to only the first 3 months of the mission. It also may be due to the retrieval algorithm differences for SAGE I and SAGE II associated with the larger number of
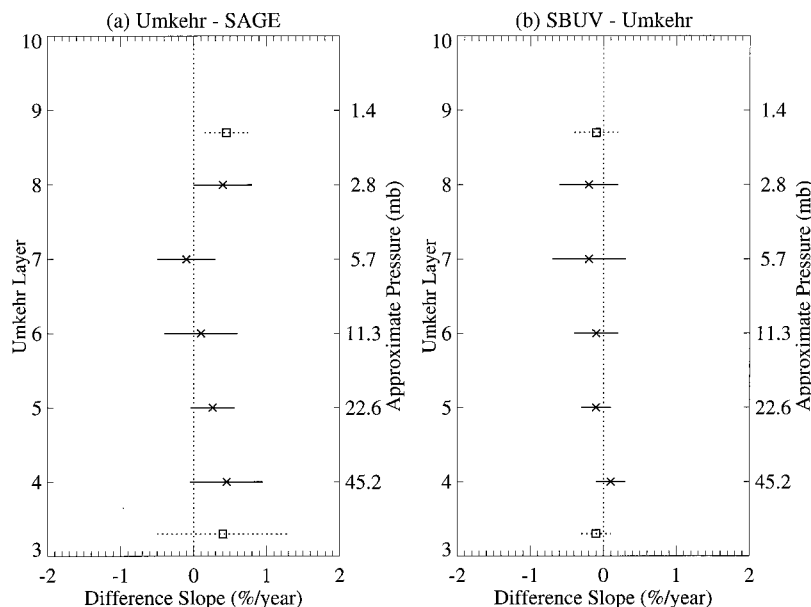

Figure 10. (a) Average regression slopes (% yr^{-1}) of the layer-ozone time series of the differences of coincident observations (Umkehr-SAGE) between 1979 and 1996 from seven Umkehr sites, six of which are located in the Northern Hemisphere. The disconnected open triangles represent layers 2 + 3 + 4 plotted at layer 3.3 and layers 8 + 9 + 10 plotted at layer 8.7. The crosses represent individual layers. All were individually corrected for aerosol interference using the *Mateer and DeLuisi* [1992] aerosol correction factors. (b) Global average regression slopes (i.e., relative drift in % yr^{-1}) of SBUV-Umkehr coincident measurements (1979 through 1989) for the same seven Umkehr sites as were used in Figure 10a.
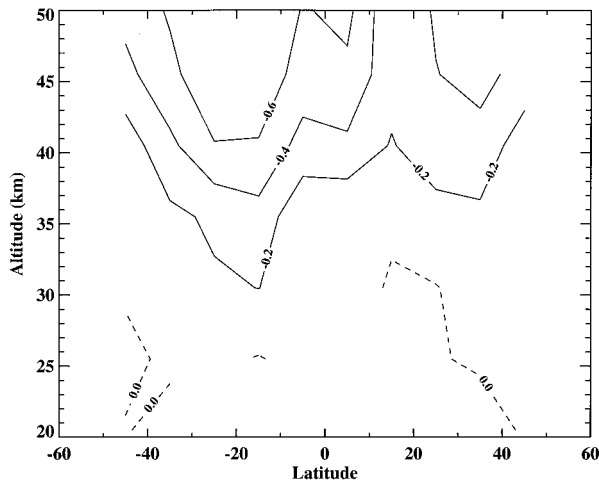
Figure 11. The difference between SAGE I/II sunrise and sunset trends from 1979 to 1996 (% yr^{-1}) expressed as a percentage of the sunset mean values. The dashed lines are contours of zero % yr^{-1} and the latitude bins were 10° wide [*Russell et al.*, 1998, Figure 2.39].

I/II. However, more generally, sunrise/sunset differences occurring in the SAGE II data indicate that there will be differences in the SAGE I/II trend results when SAGE II sunrise or sunset measurements are used separately. Figure 11 shows trend differences based on an empirical model fit (which included seasonal, QBO, and solar cycle terms) to the SAGE I/II data for 1979–1996 using the SAGE II sunrise and sunset data in separate time series (but all the SAGE I data in both the time series). When the SAGE II data alone are used (i.e., 1984–1996), differences also occur between the sunrise and sunset trends. This is illustrated in Figure 12 where it is evident that after about the end of 1987, an increased separation between the two time series occurs. The difference between the sunrise and sunset trends for SAGE II alone are about one-half those shown in Figure 11.

The next question might then be whether the SAGE II sunset measurements are more correct than the sunrise measurements. Intuitively, this might be the case because the SAGE team has historically had difficulty in deriving NO_2 from the sunrise measurements [e.g., *Cunnold et al.*, 1991]. However, it appears that SAGE sunrise ozone values are in better agreement in the mean with MLS, HALOE, and SBUV measurements at 1 mbar (where the sunrise/sunset separation is largest, e.g., Figure 6). Therefore, although the tropical upper stratosphere SAGE sunset trends are of smaller magnitude and are in better agreement with SBUV trends than the sunrise trends, there is no compelling reason to exclusively favor
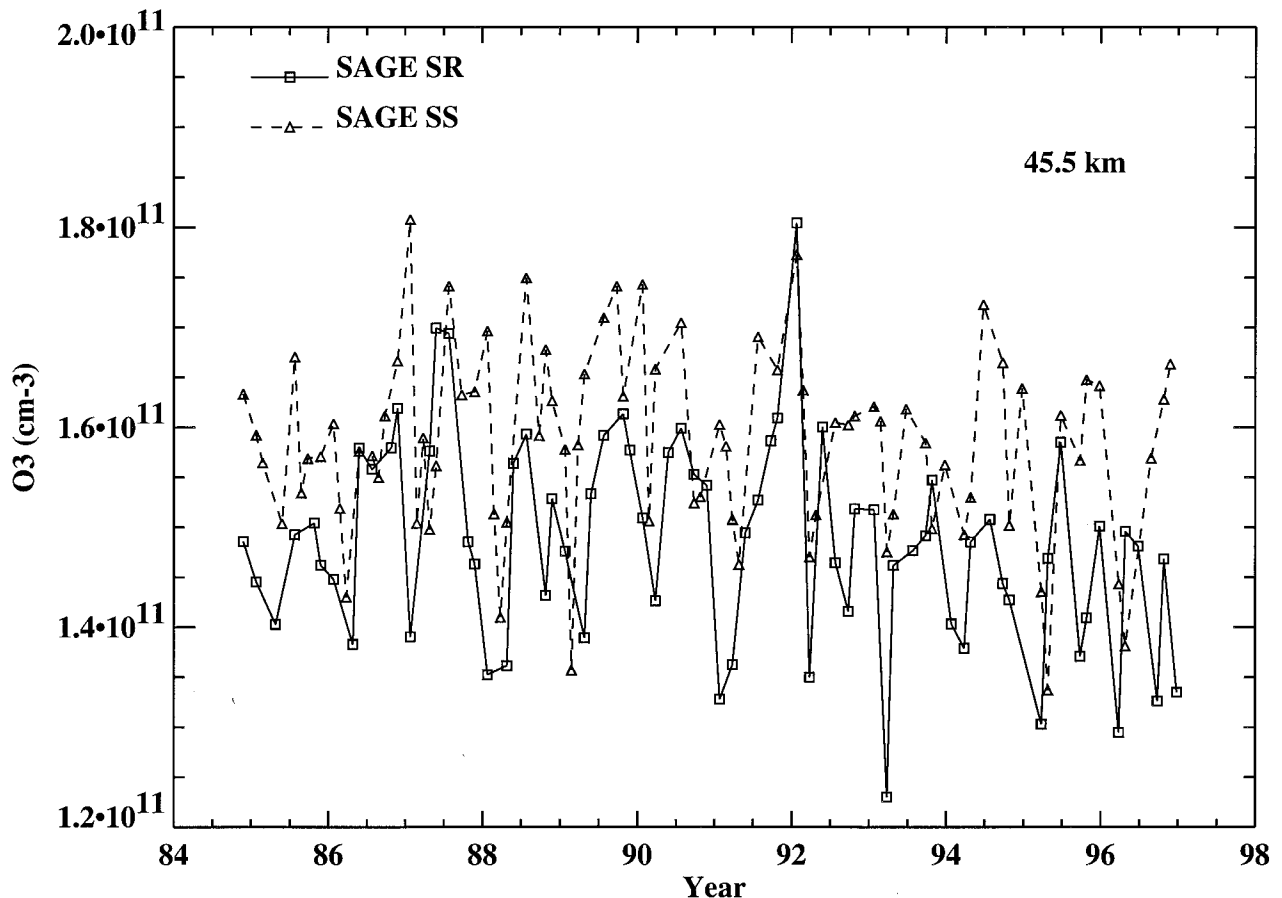
wavelengths used for SAGE II. All analyses reported in this paper and in the other related papers (M. J. Newchurch et al., unpublished data, 1998; *Cunnold et al.*, this issue; *Randel et al.*, 1999] used a combined sunrise/sunset time series for SAGE

Figure 12. Time series of the separate SAGE II monthly zonal means of ozone concentrations for sunrise and sunset observations at 45.5 km, 5°S.
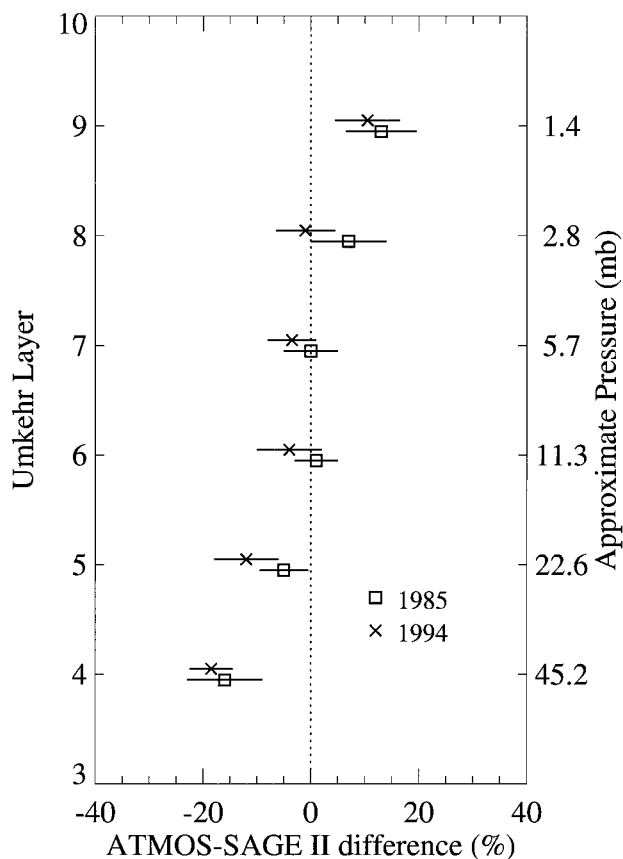
Figure 13. Differences between ATMOS and SAGE II coincident sunset measurements of zonal mean ozone at midlatitudes of the Northern Hemisphere and 95% confidence limits on the mean differences. The squares indicate the results for the Space-Lab-3 ATMOS flight in 1985 and the crosses are for the ATLAS-3 mission in 1994.

either sunrise or sunset trends at this time. In order to allow for the possibility of a systematic problem in either the sunrise or the sunset measurements, the 95% confidence limits on SAGE ozone trends by M. J. Newchurch et al. (unpublished data, 1998) have been increased using $\pm 1/2$ the trend differences shown in Figure 11. This increase in the uncertainties assumes that the range of trends between the sunrise and sunset SAGE observations represents the extreme of the 95% confidence interval.

3.5. ATMOS/SAGE II Comparisons

Figure 13 shows mean differences between ATMOS ATLAS-3 measurements [*Gunson et al.*, 1996] and SAGE II for 16 sunset coincidences between 45° and 54°N in 1994. The retrieved profile differences have been averaged over 5 km thick layers and then plotted at Umkehr levels using the conversion indicated in Table 1 (column 3). Also shown in Figure 13 are the mean differences between 11 sunset ATMOS Space-Lab 3 measurements between 26° and 35°N at the end of April 1985 and 35 SAGE II profiles in the same latitude band measured 2 days earlier. The ATMOS/SAGE II differences suggest that in both 1985 and 1994, the profiles measured by one of the instruments are vertically offset by ~1 km compared to those measured by the other instrument. However, a straightforward interpretation of these comparisons made 9 years apart is that

ATMOS measures a negative ozone drift with respect to SAGE II from layer 4 to layer 9 of -0.3 ± 0.9, -0.8 ± 0.8, -0.6 ± 0.8, -0.4 ± 0.7, -0.9 ± 1.0, and $-0.3 \pm 1.0\%\ yr^{-1}$.

3.6. Summary of Other Sensor/SAGE Slope Differences

Figure 14 summarizes the slope results of sections 3.1–3.3. The latitudinal variations shown in Figures 3, 4, 7, and 8 have been averaged into just three latitudinal regions: southern hemisphere midlatitudes, the tropics and northern hemisphere midlatitudes. It is evident from Figure 14 that there is remarkable consistency between the regression slopes in the two hemispheres, and even the tropical values only exhibit a few noteworthy differences: SBUV (5 year) in layers 8 and 9, HALOE slopes in layer 9, and MLS and HALOE slopes in layer 6. There are also some latitude-dependent differences for SBUV in layer 4. Of these differences, only the SBUV (5 year) differences in layer 9 between the tropics and midlatitudes are significant at the 95% confidence level. There are, however, significant differences between the slopes in the comparisons against SBUV for the two periods, particularly in layers 6–8. After correction for calibration drift, *Bhartia et al.* [1995] deduced a residual SBUV drift uncertainty of ~5%. This is the source of SBUV uncertainty of ~0.5% yr^{-1} for the 1979–1989 period given in Table 1. However, most of the SBUV instrumental change occurred after 1984 [*Bhartia et al.*, 1995]. Therefore a much larger trend uncertainty is associated with the 1985–1989 period (by a factor of 2). This uncertainty has been treated as a source of systematic error in the SBUV trend, and it is not included in the Figure 14 error bars. This suggests that the 1979–1989 estimates are considerably more accurate than those for the 1985–1989 period.

Because of the fairly small variations with latitude, the results can be most effectively summarized by presenting just the vertical structure of the latitudinally averaged slopes. These are shown in Figure 15a. Here the MLS and HALOE results which were shown in Figure 9a have been combined into a single estimate because they were obtained from measurements over the same period of time; the MLS/HALOE results in layer 4 have been omitted because of their exceptionally large error bars resulting from the reduced data set at this level. The SBUV (10 year) estimates, which are the most precise, typically lie in between the other estimates with MLS/HALOE suggesting that SAGE is underestimating the negative trends from 1992 to 1996 and Umkehr suggesting that SAGE is slightly overestimating the negative trends from 1979 to 1996. The MLS/HALOE result in Umkehr layer 5 is clearly an anomalous result which is related to substantial MLS/HALOE differences at 21.5 mbar and to MLS sampling issues; if the HALOE/SAGE comparison via MLS is neglected, the MLS/HALOE-SAGE slope is $-0.6 \pm 0.4\%\ yr^{-1}$.

The results are combined into a single vertical profile of the slopes of the differences in Figure 15b. The SBUV (10 year) results are also shown in the figure for comparison because of their good precision. The results then show that SAGE may be slightly overestimating the negative slopes in layers 4, 8, and 9 and possibly giving a less negative slope in layers 5–7. Overall the results in Figure 15b suggest that SAGE is giving trends with an accuracy of between 0.6 and $-0.4\%\ yr^{-1}$ with 95% confidence. Moreover, it is most likely that the overall accuracy of the SAGE (combined sunrise/sunset) trends is between approximately -0.2 and $0.2\%\ yr^{-1}$.

The combination of the SBUV (10 year) results in layers 8 and 9, suggesting that SAGE is overestimating the downward
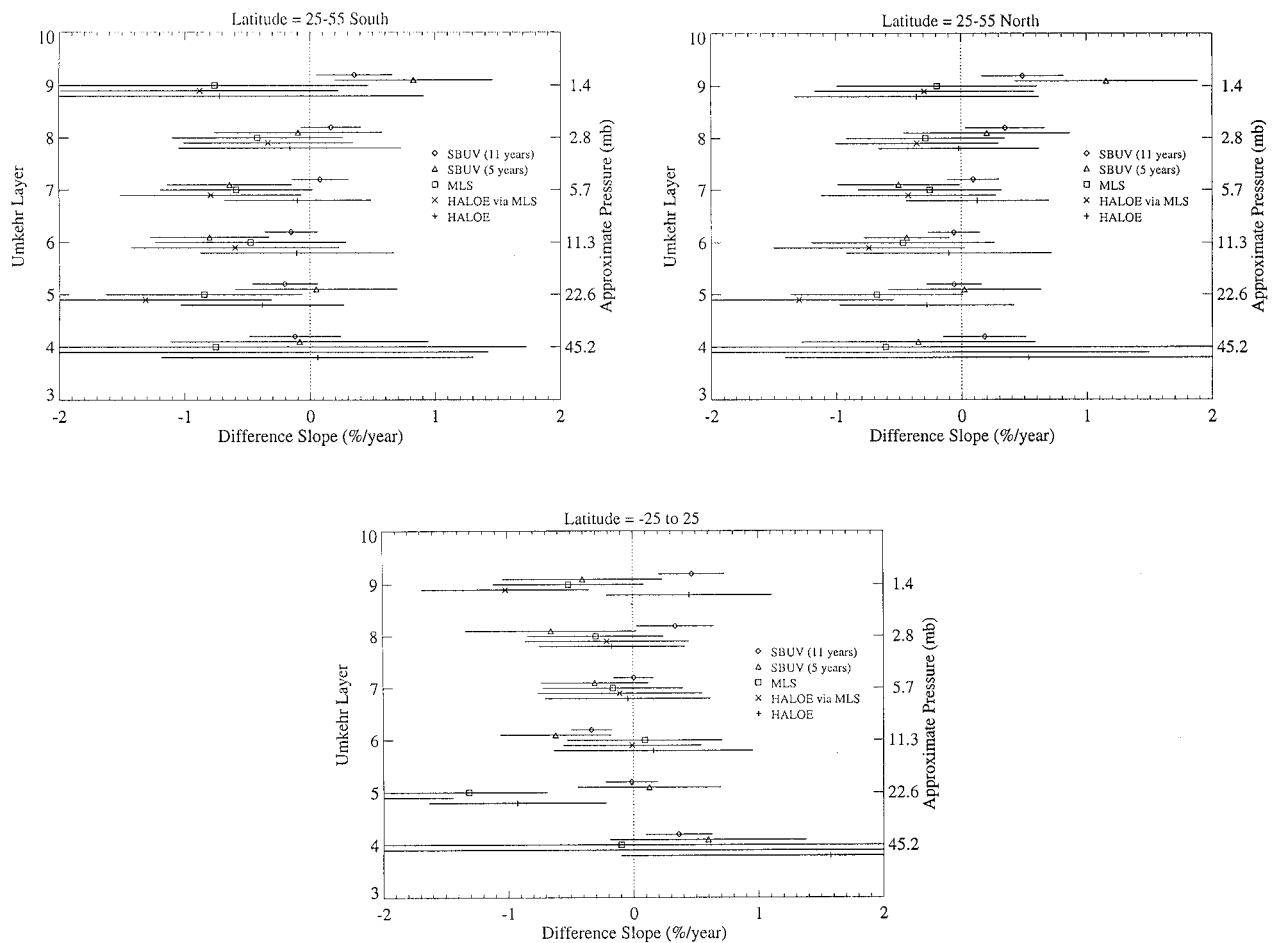
Figure 14. A summary of the difference slopes relative to SAGE averaged over three latitude bands: 25° to 55°S, 25°S to 25°N and 25° to 55°N. Mean slopes and 95% confidence limits are shown for SBUV (1979–1989; diamonds), SBUV (1985–1989; triangles), MLS (1992/1993–1996; squares), HALOE via MLS (crosses), and HALOE (1992/1993–1996) (pluses).

trend there (by ∼0.4% yr^{-1}), with the significant differences between the SAGE sunset- and sunrise-derived trends, and the SBUV (5 year) tropics versus midlatitude slope differences in layers 8 and 9 suggest a need for further study of the region above 40 km altitude. For example, the SAGE II sunrise/sunset ozone trend differences need to be understood.

4. Conclusions

For ozone trend calculations, which are reported by *Newchurch et al.* [1998a], SAGE I/II provides the longest, most spatially extensive set of stratospheric ozone data. To assess the accuracy of the trends derived from SAGE I/II measurements from 1979 to 1996, we have (1) estimated the uncertainty in ozone trends based on instrument considerations and (2) compared SAGE I and II with coincident SBUV, SBUV/2, Umkehr, and MLS measurements and calculated the regression slopes in the coincident-observation-pair time series. We have also analyzed linear trends in SAGE and HALOE ozone measurements over periods extending from 1992 to 1996 at 1 mbar and from 1994 to 1996 in the tropics at 46.4 mbar.

From instrumental considerations, we estimate the systematic uncertainty in trends calculated from SAGE I/II observations from 1979 to 1996 to be 0.1% yr^{-1} between 4 and 16

mbar (∼30–40 km) and 0.2% yr^{-1} in the altitude range from 16 to 24 mbar (∼20–30 km altitude) and from 1 to 4 mbar (∼40–50 km). However, differences exist between the trends derived from the time series consisting of SAGE I/II sunrise measurements and time series consisting of SAGE I/II sunset measurements. These trend differences are largest at 20°S above 40 km altitude and are typically larger than the estimated systematic instrumental error in the SAGE trends above 35 km altitude. These differences must be accounted for in the error bars on SAGE ozone trends in the upper stratosphere. We therefore recommended that either (1) the trends from SAGE sunrise and sunset measurements be calculated separately or (2) the size of the confidence interval about the trend calculated from the grouped sunrise and sunset observations be increased by ±1/2 the difference between the individual sunset and sunrise trend values. The effect of choice 2 would be to increase the 95% error bar on the SAGE trends in layers 8 and 9 by approximately ±0.2% yr^{-1}. We estimate the instrumental uncertainties of both an ensemble of Dobson Umkehr stations and HALOE observations to be 0.1% yr^{-1} above layer 4, of the MLS instrument to be no more than 0.2% yr^{-1} in the same height range, and of the SBUV instrument, to be ∼0.5% yr^{-1} over the 1979–1989 period.
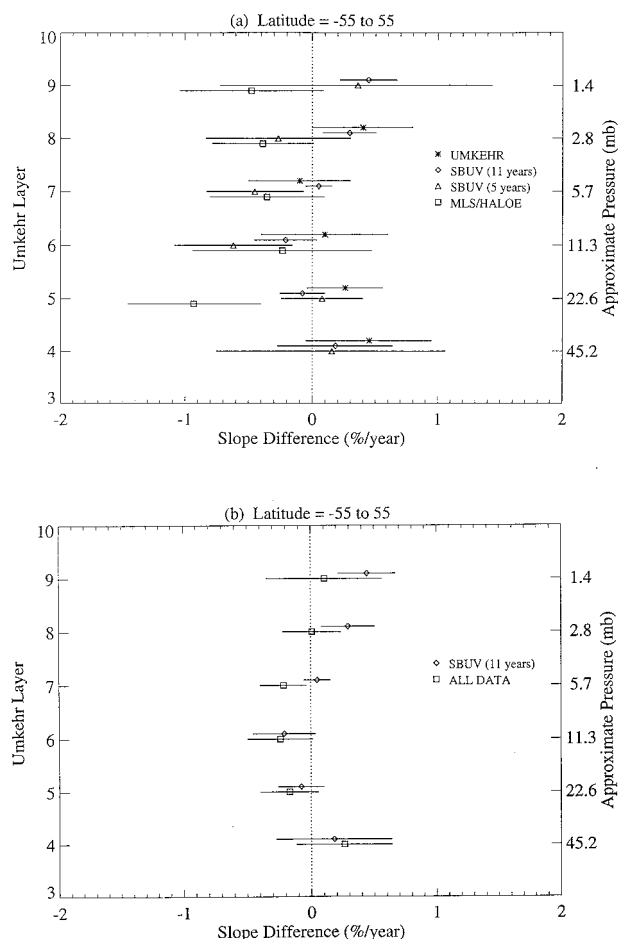
Figure 15. (a) Latitudinally averaged slopes (% yr^{-1}) and 95% confidence limits for Umkehr-SAGE (asterisks), SBUV-SAGE (1979–1989; diamonds), SBUV-SAGE (1985–1989; triangles), and from MLS/HALOE-SAGE (squares) ozone measurements. (b) The combined (all sensors-SAGE) latitudinally averaged slopes obtained using the Figure 14a results (squares) and 95% confidence limits. The SBUV-SAGE (1979–1989; diamonds) results are also shown for comparison because of their inherently smaller error bars compared with the other slope results.

Extensive comparisons between SAGE and the measurements of ozone by the other instruments exhibit good consistency in the slopes of the differences from one latitude to another. There is especially good agreement between the northern and southern hemisphere midlatitude results. The latitudinal means of the mostly northern hemisphere Umkehr-SAGE slopes in layers 4–8 for 1979–1996 are in the range −0.1–0.4% yr^{-1} and the SBUV-SAGE latitudinal mean slopes for 1979–1989 in layers 4–9 are in the range −0.2–0.4% yr^{-1}. MLS/HALOE-SAGE mean slopes, in contrast, for the 1992–1996 period for Umkehr layers 5–9 vary from −0.9–0.0% yr^{-1}. However, mean trend differences between HALOE and SAGE only differ from ±0.1% yr^{-1} in layer 5, and in that layer the magnitude of the mean MLS-HALOE slope is larger than either the MLS or HALOE-SAGE slopes.

The results from the three comparison periods have been combined with the SBUV-SAGE II slopes for 1985–1989 to yield a single set of layer means and standard errors. These standard errors included allowances for the uncertainties in the

means from each instrument and also allowance for the auto-correlation of the time series residuals (for SBUV-SAGE) and for the autocorrelation of the variation of the slopes with latitude (in all cases). The overall layer means with 95% confidence limits then covered the range of −0.4 to +0.6% yr^{-1} and the most likely slopes varied from −0.2 to +0.2% yr^{-1} in Umkehr layers 4–9. The values are slightly negative in layers 5–7 and slightly positive in layers 4, 8, and 9. Comparisons between SAGE and ATMOS measurements in 1985 and 1994 fail to provide any evidence of nonzero drifts between the two instruments and provide no evidence to contradict the above conclusion.

Mean SBUV/2-SAGE slopes, however, were positive in all layers and ranged from 0.1 ± 0.8% yr^{-1} in layer 4 to 1.4 ± 0.4% yr^{-1} in layer 8. Positive slopes at all latitudes and altitudes and their vertical structure suggest a calibration problem in the SBUV/2 data probably associated with the changing solar zenith angles of the NOAA-11 measurements.

The accuracy of the 11 year SBUV-SAGE slopes is ~3 times better than for the 5 year slope estimates. There is some incompatibility between the two sets of SBUV results also which suggests an additional source of uncertainty in the SBUV 5 year comparisons; this difference might be associated with the deterioration of the SBUV instrument after 1984. SBUV-SAGE mean slopes of 0.4% yr^{-1} in layer 9 and 0.3% yr^{-1} in layer 8, where Umkehr also yields 0.4% yr^{-1}, combined with the SAGE II sunrise/sunset trend differences above 40 km altitude, indicate that in layers 8 and 9 SAGE I/II trends for 1979–1996 are accurate to within ±0.4% yr^{-1}. In layers 4–7 the SAGE I/II trends are indicated to have an accuracy of ±0.2% yr^{-1}.

References

Basher, R. E., Survey of WMO-sponsored Dobson spectrophotometer intercomparisons, *Rep. 19*, 56 pp., World Meteorol. Organ., Wellington, N. Z., 1995.

Bass, A. M., and R. J. Paur, The ultraviolet cross-section of ozone, I, Results and temperature dependence, in *Atmospheric Ozone: Proceeding of the Quadrennial Ozone Symposium Held in Halkidiki, Greece*, edited by C. S. Zerefos and A. Ghazi, pp. 606–610, D. Reidel, Norwell, Mass., 1985.

Bhartia, P. K., S. Taylor, R. D. McPeters, and C. Wellemeyer, Application of the Langley plot method to the calibration of the solar backscatter ultraviolet instrument on the Nimbus 7 satellite, *J. Geophys. Res.*, *100*, 2997–3004, 1995.

Bhatt, P. B., E. E. Remsberg, L. L. Gordley, J. M. McInerney, V. G. Brackett, and J. M. Russell III, An evaluation of the quality of Halogen Occultation Experiment ozone profiles in the lower stratosphere, *J. Geophys. Res.*, *104*, 9217–9231, 1999.

Cunnold D. M., and H. J. Wang, The effects of temperature uncertainties on the interpretation of ozone trends in the upper stratosphere, *Atmospheric Ozone: Proceedings of the XVIII Quadrennial Ozone Symposium*, edited by R. Bojkov and G. Visconti, Edigrafital S.p.A., S. Atto, Italy, 1998.

Cunnold, D. M., et al., Validation of SAGE II NO$_2$ measurements, *J. Geophys. Res.*, *96*, 12,913–12,925, 1991.

Cunnold, D. M., H. Wang, W. P. Chu, and L. Froidevaux, Comparisons between Stratospheric Aerosol and Gas Experiment II and Microwave Limb Sounder ozone measurements and aliasing of

SAGE II ozone trends in the lower stratosphere, *J. Geophys. Res.*, *101*, 10,061–10,076, 1996a.

Cunnold, D. M., L. Froidevaux, J. M. Russell, B. Connor, and A. Roche, Overview of UARS ozone validation based primarily on comparisons among UARS and Stratospheric Aerosol and Gas Experiment II measurements, *J. Geophys. Res.*, *101*, 10,335–10,350, 1996b.

Cunnold, D. M., R. Weiss, R. G. Prinn, D. Hartley, P. G. Simmonds, P. J. Fraser, B. Miller, F. N. Alyea, and L. Porter, GAGE/AGAGE measurements indicating reductions in global emissions of CCl_3F and CCl_2F_2 in 1992–1994, *J. Geophys. Res.*, *102*, 1259–1269, 1997.

Cunnold, D. M., H. J. Wang, L. Thomason, J. Zawodny, J. A. Logan, and I. A. Megretskaia, SAGE (version 5.96) ozone trends in the lower stratosphere, *J. Geophys. Res.*, this issue.

Froidevaux, L., et al., Validation of UARS Microwave Limb Sounder ozone measurements, *J. Geophys. Res.*, *101*, 10,017–10,060, 1996.

Gunson, M. R., et al., The Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment: Deployment on the ATLAS Space Shuttle Missions, *Geophys. Res. Lett.*, *23*, 2333–2336, 1996.

Harris, N., R. Hudson, and C. Phillips, *Assessment of Trends in the Vertical Distribution of Ozone*, SPARC Rep. 1, 289 pp., World Meteorol. Org., Geneva, Switzerland, 1998.

Herman, J. R., R. D. Hudson, R. McPeters, R. Stolarski, Z. Ahmad, X. Gu, S. Taylor, and C. Wellemeyer, A new self-calibration method applied to TOMS/SBUV backscattered ultraviolet data to determine long-term global ozone change, *J. Geophys. Res.*, *96*, 7531–7545, 1991.

Hervig, M. E., J. M. Russell III, L. L. Gordley, J. H. Park, S. R. Drayson, and T. Deshler, Validation of aerosol measurements from the Halogen Occultation Experiment, *J. Geophys. Res.*, *101*, 10,267–10,276, 1996.

Hilsenrath, E., R. P. Cebula, M. T. Deland, K. Laamann, S. Taylor, C. Wellemeyer, and P. K. Bhartia, Calibration of the NOAA-11 Solar Backscatter Ultraviolet (SBUV/2) ozone data set from 1989 to 1993 using in-flight calibration data and SSBUV, *J. Geophys. Res.*, *100*, 1351–1366, 1995.

Hofmann, D., et al., Characteristics of ozone data sources used for trend identification, in *Assessment of Trends in the Vertical Distribution of Ozone*, SPARC Rep. 1, pp. 3–96, World Meteorol. Org., Geneva, Switzerland, 1998.

Hollandsworth, S. M., R. D. McPeters, L. E. Flynn, W. Planet, A. J. Miller, and C. Chandra, Ozone trends deduced from combined Nimbus 7 SBUV and NOAA 11 SBUV/2 data, *Geophys. Res. Lett.*, *22*, 905–908, 1995.

Hood, L. L., R. D. McPeters, J. P. McCormick, L. E. Flynn, S. M. Hollandsworth, and J. F. Gleason, Altitude dependence of stratospheric ozone trends based on Nimbus 7 SBUV data, *Geophys. Res. Lett.*, *20*, 2667–2670, 1993.

Jackman, C. H., E. L. Fleming, S. Chandra, D. B. Considine, and J. E. Rosenfield, Past, present, and future modeled ozone trends with comparisons to observed trends, *J. Geophys. Res.*, *101*, 28,753–28,767, 1996.

Komhyr, W. D., R. D. Grass, and R. K. Leonard, Dobson spectrophotometer 83: A standard for total ozone measurements, 1962–1987, *J. Geophys. Res.*, *94*, 9847–9861, 1989.

Logan, J. A., et al., Trends in the vertical distribution of ozone: A comparison of two analyses of ozonesonde data, *J. Geophys. Res.*, *104*, 26,373–26,400, 1999.

Mateer, C. L., and J. J. DeLuisi, A new Umkehr inversion algorithm, *J. Atmos. Terr. Phys.*, *54*, 537–556, 1992.

Mateer, C. L., H. U. Dutsch, J. Staehelin, and J. J. DeLuisi, Influence of a priori profiles on trend calculations from Umkehr data, *J. Geophys. Res.*, *101*, 16,779–16,787, 1996.

McCormick, M. P., P. R. E. Veiga, and W. P. Chu, Stratospheric ozone profile and total ozone trends derived from the SAGE I and SAGE II data, *Geophys. Res. Lett.*, *19*, 269–272, 1992.

McPeters, R. D., T. Miles, L. E. Flynn, C. G. Wellemeyer, and J. Zawodny, Comparison of SBUV and SAGE II ozone profiles: Implications for ozone trend, *J. Geophys. Res.*, *99*, 20,513–20,524, 1994.

Miller, A. J., et al., Information content of Umkehr and Solar Back-scattered Ultraviolet (SBUV) 2 satellite data for ozone trends and solar responses in the stratosphere, *J. Geophys. Res.*, *102*, 19,257–19,263, 1997.

Montzka, S. A., J. H. Butler, R. C. Myers, T. M. Thompson, T. H. Swanson, A. D. Clarke, L. T. Lock, and J. W. Elkins, Decline in the tropospheric abundance of halogen from halocarbons: Implications for stratospheric ozone depletion, *Science*, *272*, 1318–1322, 1996.

Newchurch, M. J., and D. M. Cunnold, Aerosol effect on Umkehr ozone profiles using Stratospheric Aerosol and Gas Experiment II measurements, *J. Geophys. Res.*, *99*, 1383–1388, 1994.

Newchurch, M. J., D. M. Cunnold, and H. J. Wang, Stratospheric Aerosol and Gas Experiment II: Umkehr ozone profile comparisons, *J. Geophys. Res.*, *100*, 14,029–14,042, 1995.

Newchurch, M. J., D. M. Cunnold, C. L. Mateer, and J. Cao, Intercomparison of SAGE with Umkehr (64) and Umkehr (92) ozone profiles and time series: 1979–1991, *J. Geophys. Res.*, *103*, 31,277–31,292, 1998.

Randel, W., R. Stolarski, D. Cunnold, J. A. Logan, and M. J. Newchurch, Trends in the vertical distribution of ozone, *Science*, *285*, 1689–1692, 1998.

Reinsel, G. C., G. C. Tiao, D. J. Wuebbles, J. B. Kerr, A. J. Miller, R. M. Nagatani, L. Bishop, and L. H. Ying, Seasonal trend analysis of published ground-based and TOMS total ozone data through 1991, *J. Geophys. Res.*, *99*, 5449–5464, 1994.

Russell, J. M., III, et al., Ozone data quality, in *Assessment of Trends in the Vertical Distribution of Ozone*, SPARC Rep. 1, pp. 97–188, World Meteorol. Org., Geneva, Switzerland, 1998.

Solomon, S., R. W. Portmann, R. R. Garcia, L. W. Thomason, L. R. Poole, and M. P. McCormick, The role of aerosol variations in anthropogenic ozone depletion at northern midlatitudes, *J. Geophys. Res.*, *101*, 6713–6727, 1996.

Steele, H. M., and R. T. Turco, Separation of aerosol and gas components in the Halogen Occultation Experiment and the Stratospheric Aerosol and Gas Experiment (SAGE) II extinction measurements: Implications for SAGE II ozone concentrations and trends, *J. Geophys. Res.*, *102*, 19,665–19,682, 1992.

Torres, O., and P. K. Bhartia, Effects of stratospheric aerosol on ozone profile from SBUV measurements, *Geophys. Res. Lett.*, *22*, 235–238, 1995.

Wang, H. J., D. M. Cunnold, and X. Bao, A critical analysis of SAGE ozone trends, *J. Geophys. Res.*, *101*, 12,495–12,514, 1996.

Wang, P.-H., M. P. McCormick, W. P. Chu, J. Lenoble, R. M. Nagatani, M. L. Chanin, R. A. Barnes, F. Schmidlin, and M. Rowland, SAGE II stratospheric density and temperature retrieval experiment, *J. Geophys. Res.*, *97*, 843–864, 1992.

World Meteorological Organization, Scientific Assessment of Stratospheric Ozone: 1994, World Meteorol. Org. *Rep. 37*, Geneva, Switzerland, 1994.

D. M. Cunnold and H. J. Wang, School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 221 Bobby Dodd Way, Atlanta, GA 30332. (cunnold@eas.gatech.edu)

L. E. Flynn, Office of Research and Applications, National Environmental Satellite Data and Information Service, National Oceanic and Atmospheric Administration, Washington, D.C. 20233.

L. Froidevaux, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109.

R. McPeters, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771.

M. J. Newchurch, Atmospheric Sciences Department, University of Alabama, Huntsville, Huntsville, AL 35807.

J. M. Russell, Center for Atmospheric Science, Hampton University, Hampton, VA 23668.

J. M. Zawodny, NASA Langley Research Center, Hampton, VA 23681.