SPARC/IOC/GAW

Assessment of Trends in the Vertical Distribution of Ozone

May 1998

Edited by N. Harris, R. Hudson and C. Phillips

SPARC Report No.1. WMO Ozone Research and Monitoring Project Report No. 43

CONTENTS

Chapter 2. Ozone Data Quality *Lead Authors : James M. Russell III, Herman G. J. Smit*

Chapter 3. Ozone Change as a Function of Altitude *Lead Authors : Richard Stolarski, William Randel*

Appendices

Foreword

This scientific assessment has been carried out jointly by the WCRP project on Stratospheric Processes and their Role in Climate (SPARC) and the International Ozone Commission (IOC), in close co-operation with WMO's Global Atmospheric Watch programme (GAW). The objective of the report is to review critically the measurements and trends of the vertical distribution of ozone. Recently revised data are used where appropriate and the time period covered has been extended into mid 1996. This objective has been achieved, and this report contains a thorough description of the measurements and their associated uncertainties as well as the trends and their associated uncertainties. The executive summary, which follows, first appeared in SPARC Newsletter No. 10 during the spring of 1998.

This is the first assessment of its kind carried out by SPARC and the IOC. It was stimulated by the WMO/UNEP Scientific Assessment of Ozone Depletion : 1994 which found large discrepancies between ozone trends in the lower stratosphere. One of the objectives of this assessment was to ensure a continuity in the international effort necessary to prepare the WMO/UNEP Assessments and the work presented here is being heavily relied on in the production of the 1998 assessment. It is also being used in the preparation of the Special IPCC Report on Aviation and the Global Atmosphere. The SPARC-IOC-GAW assessment also anticipates the need for precise updated observations of the vertical distribution of stratospheric ozone depletion to be used to assess the impact on climate in the Third IPCC Assessment Report on Climate Change due in the year 2000.

The need for such an assessment was recognised by the IOC, SPARC and GAW in 1995 after the publication of the 1994 WMO-UNEP Assessment. The outline of the assessment was determined by an international group of scientists during a workshop held at the Observatoire de Haute Provence in France, in July 1996. The drafts of the chapters were prepared in the following year and an impressive amount of new critical analyses were produced. The draft report was examined by an international panel of reviewers both by mail peer review and at a meeting at Abingdon, United Kingdom in October 1997. This rigorous review greatly improved the report, especially the presentation of the large amount of new material and the consistency between chapters.

The success in producing the present document is the result of the intensive work and enthusiastic co-operation of a large number of scientists world-wide who have worked towards improving the quality of the measurements, validating the trends observed with state of the art techniques, and calculating and validating the trends deduced, testing the models used. This has resulted, for the first time, in a single profile of observed ozone depletion at northern mid-latitudes, combining all the observing systems available. The work of the contributors and reviewers was generously supported by many agencies including WMO-GAW, WCRP, SPARC, DG-XII of the European Commission, NASA, NOAA, UK DETR, CNRS and other national research programmes and institutions.

We take this opportunity to express our gratitude to all the scientists (authors, contributors and reviewers) who helped in the preparation this assessment and to the SPARC and IOC panel members who have been supportive since its inception. Our special gratitude is due to the lead authors of the chapters who worked impressively hard and effectively. Particular thanks must be given to Isabelle Faviez of the Institut Pierre Simon Laplace, France, who helped organise the workshop at the Observatoire de Haute Provence, and Maria Brown of the European Ozone Research Co-ordinating Unit, United Kingdom, for her help in organising the workshops and the final editing of the report. Finally, and by no means least, we sincerely thank our co-editor, Céline Phillips of the SPARC office, for her dedicated hard work at all stages during the preparation of the report, at the meetings and especially in the preparation of the final document.

Neil Harris, European Ozone Research Coordinating Unit , Centre for Atmospheric Science, University of Cambridge, UK
Bob Hudson, Department of Meteorology, University of Maryland, USA Assessment co-chairs.

Summary

1. Introduction

One of the largest uncertainties in determining the effect of CFC's on stratospheric ozone has been the magnitude of the trends in the altitude region between 15 and 20 km. In the 1994 WMO-UNEP ozone assessment, SAGE was reported as giving trends up to -0± 8% per decade at northern mid-latitudes, while the ozonesonde stations gave a trend of -± 3% per decade. In 1996 the SPARC panel on Understanding Ozone Trends and the International Ozone Commission decided to collaborate, under the auspices of the World Climate Research Programme and the World Meteorological Organisation, on a study to carefully re-evaluate the ground-based and satellite data to resolve this discrepancy. The philosophy of the study was similar to that of the International Ozone Trends Panel of 1988 which addressed the total ozone measurements. The published literature was not simply reviewed, but a critical re-analysis and interpretation of the vertical profiles of ozone was performed. One of the principal aims of the report was to determine if there was sufficient confidence in the long-term measurement systems to use them for accurate determination of ozone trends in the stratosphere and troposphere. A major purpose of the study was to validate the quality of the data including quantification of the errors and to determine if there were any limitations in altitude or latitude.

The report is divided into three main chapters. Chapter 1 contains a description of how the various instruments work, and how ozone concentrations are calculated from the raw measurement. Particular attention is paid to the true vertical resolution of each instrument's measurement and to its long term calibration drift as well as to its precision and accuracy. Chapter 2 assesses how well the various measurements agree through a series of rigorous data comparisons. Traditional techniques based on zonal averages and on close matches in time and space are augmented by new techniques which classify the air mass according to its dynamic history. Chapter 3 discusses and uses the various methods available for calculating trends, as well as investigating how well the causes of the trends can be determined by statistical approaches.

2. Findings

2.1 Validity of the Data Sets

Only four measurement techniques were identified that had produced records long enough to assess long term trends, SAGE (I and II), SBUV and SBUV2, Umkehr/Dobson and ozonesondes. The SAGE I and II satellite series extends from February 1979 to June 1996, with a three year interruption beginning November 1981. This series provides altitude coverage from the lower stratosphere to the stratopause. The SBUV-SBUV2 satellites (1978 to present) and the ground-based Umkehr/Dobson instruments (1957 to present) provide data sets for examining trends in the middle and upper stratosphere. Data from the ozonesonde network, started in the early 1960's, and extend to the present. This data set has the potential for providing trends in the lower stratosphere and troposphere. The report also assessed the ozone data quality from measurement systems which have operated over shorter time periods to validate the long term measurement systems above. The purpose of this approach was two-fold, (1) Was there any evidence that the SAGE algorithms produced errors in the measured ozone?, and (2) Was there any evidence for a long-term systematic error in the SAGE data which could affect the derived trends?

The measurement techniques of all the data sets were critically analysed from an instrumental and theoretical perspective. Changes in instrument performance and operation were considered as well as any

errors or uncertainties produced in the algorithms used. Important issues include the correction in the SAGE data for the presence of aerosol (principally important below 20 km and in the aftermath of volcanic eruptions) and the pump correction for ozonesondes (important above 25 km). Tables of uncertainties have been composed which include not only the accuracy and precision of individual measurements, but also, for the first time for measurements of the vertical profile of ozone, estimates of the stability of the various systems over time. These were given as a function of altitude and latitude where appropriate and are being included in the estimates of the uncertainties in the trends given below. The drift uncertainties (2σ) are estimated to be less than 5% per decade for all measurements systems considered, with the exception of the Brewer-Mast ozonesonde in the troposphere (at 5 km).

The inter-comparisons were used to see if the drifts found between instruments with long term records were consistent with these estimates of stability. These inter-comparisons also included data covering shorter periods from the HALOE and MLS instruments on the UARS satellite and ground-based LIDAR and microwave instruments including those in the NDSC.

All SAGE II data used in the report were derived using the version 5.96 algorithm. The known error in the altitude registration of the SAGE I data was corrected according to Wang *et al*. (1996). It was found that the most important screening consideration was to eliminate SAGE II data contaminated by Mt. Pinatubo aerosol absorption. SAGE II ozone retrievals are affected by an inability to remove all the aerosol interference (although the current algorithm is better than previous versions). Based on comparisons with MLS, which is almost unaffected by high aerosol loading, it is recommended that between 1.5 and 2.5 years of data following the Pinatubo eruption be omitted from the SAGE II ozone data at pressures greater than 10 hPa. The detailed recommendations are given in the report as a function of pressure.

The upper altitude limit of SAGE II data for use in detecting trends was determined to be on the order of 50 km based on noise in SAGE II and inter-comparisons with HALOE. The lower altitude limit is less well determined, most likely because of low altitude atmospheric variability and aerosol effects on both the SAGE II measurements and the data used for comparisons. In most instances, the drifts and their associated uncertainties between SAGE II and correlative data start to increase below about 20 km, and they become much more variable thus limiting the lower altitude that can be validated. This is not to say that long-term trends derived from SAGE II are invalid in this range, only that for the measurements systems used in the inter-comparisons, a less definitive statement can be made about trend validity. The smallest long-term drifts that can be verified over the 20 km to 50 km altitude interval are in the range of 0.3%year^{-1} when viewed as a function of latitude.

Between 20 and 40 km, the Dobson/Umkehr measurements constrain SAGE I/II drifts in a narrow latitude band in the northern mid-latitudes to 0.2±0.2%year^{-1} and in the southern mid-latitudes to 0.3 ±0.3%year^{-1} at the 95% confidence level. Globally averaged SAGE I/II trends over this altitude range are constrained by both ground-based and satellite measurements to be valid to a level on the order of 0.2%year^{-1} at the 95% confidence level. It appears that the best agreement in trends occurs for SAGE II comparisons with other satellite data; although the Dobson/Umkehr comparisons provide equally good constraints in the northern and southern mid-latitudes.

The drifts for time series of coincident differences between SAGE II and other measurements for individual stations or latitudes are summarised in Figure 1. Values range from £ 0.3 ± 0.15%year^{-1} to ~0.5 ± 0.7%year^{-1} (sondes, lidar, Umkehr, HALOE) for altitudes between 20 km and 35 km and £ 0.5 ± 0.5%year^{-1} to ~1 ± 1%year^{-1} for altitudes between 35 km and 50 km. Only two systems (sondes and lidar) provide useful trend comparison data for the altitude range between 15 km and 20 km. Trends of

matched pair (i.e. sonde minus SAGE II) differences at individual stations show significant variability, ranging up to 3%year^{-1} at 15 km for Lauder; however, best agreement was obtained when the matched pair differences from the eight sonde stations used in the trend analyses (see section 2.2) were combined into a single time series to calculate the regression slope of the differences. No statistically significant differences were obtained for the combined time series, but the mean difference was about 0.25% ± 0.4%year^{-1} in this lower stratosphere range above ~15 km altitude. This suggests that there is a fair degree of noise in the differences at individual stations due either to atmospheric variability, sampling, or instrumental effects; but it also suggests that SAGE II trends in the lower stratosphere are accurate to the 0.25%year^{-1} level.

While no statistically significant drift was found between SAGE and the ozonesondes, differences in the absolute values of the measured ozone were found. The SAGE II absolute values agree with sondes in the altitude region between 20 km and 28 km to within a few %, but below 20 km SAGE II values start to increase relative to the sondes and reach values which are 15% to 20% larger at 15 km. The low altitude differences are latitudinally dependent which could indicate a problem with using the data for global trend calculations.

The globally averaged drifts of SAGE II versus SBUV, HALOE and MLS range from -0.06% to -0.4% ±~ 0.6%year^{-1}, i.e. SAGE II ozone has become more negative compared to the other instruments. These differences, although statistically insignificant, give a slight indication of a SAGE II drift with time (SAGE II trend is larger). SBUV2 differences with SAGE II are of opposite sign to SBUV, HALOE and MLS, but this is most likely due to algorithm effects brought on by a drifting orbit. Globally averaged analyses of the longest satellite time series - SBUV compared with the composite time series of SAGE I (1979-1981) and SAGE II (1984-89), designated SBUV(*) in Figure 1, shows agreement to (-0.2% to 0.2%) ± 0.2%year^{-1} in the altitude region between 20 km and 50 km.

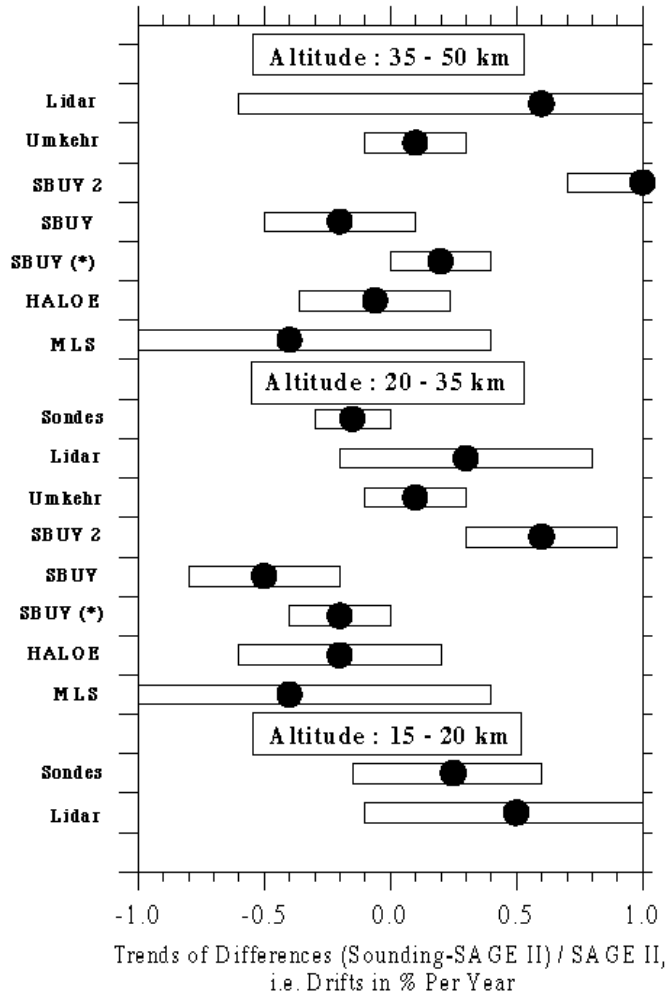


Figure 1 Trends of differences (i.e. drifts) between ozone measurements made by various ozone profiling instruments and SAGE II in % per year [(Sounding - SAGE II)/SAGE II]. Trends with ozonesondes are for the eight northern mid-latitude sounding stations used in Chapter 3 covering 36°N to 56°N. Umkehr differences are averages for eight Northern and Southern Hemisphere stations. Lidar differences are averages for the three stations with the longest records and number of coincidences exceeding 100 (Hohenpeissenberg, OHP and TMF). Trends of satellite differences (SBUV, SBUV2, HALOE and MLS) are presented as global means. The average differences are indicated by the dots and the bars represent the 95% confidence intervals of the drift estimations.

The results of a novel coordinate mapping (CM) study show general agreement with the conventional inter-comparisons, which rely on measurement space and time coincidence. Near 25 km, both the CM and conventional techniques show statistically significant regression slopes of differences in the 40-50°S (1-1.5%year^{-1}) and 60-70ºS (~0.5%year^{-1}) regions. Differences between the two satellite measurements also appear in the CM analysis near 40-50°N in the same altitude range with the CM technique indicating a marginally significant positive drift of SAGE II relative to HALOE of ~1 ± 0.75%year^{-1}. At higher altitudes, CM results show insignificant drifts of less than 0.5%year^{-1} in all regions except the polar summer latitudes. Differences in this region, which cannot be directly compared to conventional results, indicate statistically significant trends of differences, with SAGE ozone trends being larger relative to

HALOE. A Lagrangian approach was tested which used air parcel trajectories to link measurements. This technique was only tested for a limited time period and so was unsuitable for use in assessing long-term stability, but it also shows great promise for applications in the future.Some evidence exists to suggest that SAGE I and SAGE II overlapping measurements are inconsistent (e.g., the comparisons with the ozonesonde measurements at Hohenpeissenberg, Payerne and Uccle), but the results are not statistically significant. It is recalled that SAGE I data have been corrected for a systematic reference height error of approximately 300 m (latitude dependent). The uncertainty in this correction for each latitude is approximately 100 m. Below 20 km altitude a simple upward shift of the SAGE I profiles (as assumed for this report) may be incorrect because of the large Rayleigh scattering contribution to the 0.6 nm extinction at these altitudes. A new inversion of the SAGE I data to correct the altitude registration problem would be preferable. Data can be used for trends with caution below 20 km, but more inter-comparisons are needed to draw firm conclusions.

Comparisons of ozonesondes in the stratosphere with other ozone profiling techniques show consistent results with agreement of about ± (3-5)% at altitudes between the tropopause and 28 km. The precision of the different sonde types is better than ± 3%. Above 28 km the results are not consistent due to instrumental uncertainties (e.g. pump corrections and sensing solution changes) and caution must be used, at least for the non-ECC types of sondes, when applying the data for long-term trend determinations.

There is a dearth of sonde validation studies for the troposphere. Because of the small number of comparisons, only estimates about the reliability of the sonde data records below the tropopause can be made. In general, ECC-sondes provide much more consistent results than the other two types of sondes considered in the report. The precision of the ECC-sonde is better than ± (5-10)% and shows a small positive bias of about 3%. Brewer Mast and KC79-sondes are less precise (± (10-20)%), but there are no indications of any bias larger than ± 5%. Key issues of uncertainty are the background correction and the use of the total ozone normalisation factor.

The main reasons for observed differences between different sonde results from sounding stations using the same type of ozone sonde are believed to be due to differences in the preparation and correction procedures applied at the different launch sites. Although much progress has been made to improve the quality and homogeneity of the ozonesonde data since the last WMO Scientific Assessment of Stratospheric Ozone in 1994, there is still an urgent need to investigate and intercompare the instrumental performance of the different sonde types as well as a need to revise and agree on procedures for preparation and data processing.

2.2 Trend Analyses

The statistical models used in the report were inter-compared using three test data sets. This comparison revealed only minor differences in trends obtained by the models. Somewhat greater differences were found in the uncertainties estimated for the trends and other variables included in the models. Results are most sensitive to the details of the models for time series with significant missing data. Quasi-decadal variations are a ubiquitous feature of ozone observations, in addition to QBO and faster time scale dynamical variability. Inclusion of these terms does not have a strong influence on the calculated trends for long time series. Much of the observed decadal change is approximately in phase with the solar cycle for the observational record, suggesting a solar mechanism. However, current model calculations of the solar effect show some inconsistencies with observations (in terms of magnitude and lower stratospheric response), and this limits confidence in our detailed understanding. It is also likely that a confusion exists between solar and volcanic signals for the recent record. Although these effects have relatively small

impacts on the linear trend estimates, it does limit the ability to interpret decadal variability.

Figure 2 shows the mean trend vs. altitude at northern mid-latitudes obtained for combined measurement systems including estimated uncertainties from both the statistical and instrumental analyses. This averaging of trends from SAGE I/II, ozonesondes, Umkehr and SBUV is possible for the first time in a major assessment because there is now agreement between the systems in the regions where the measurements overlap. A few points are worth noting before discussing the trend findings further:

(a) The trends below 20 km shown in figure 2 are found from the ozonesondes alone. There is now reasonable agreement in the trends at these altitudes between ozonesondes and SAGE I/II (which has occurred mainly as a result of the revised SAGE aerosol correction), but uncertainties in the SAGE I altitude registration below 20 km are considered too large for the SAGE I/II trends to be used in this context,

(b) The ozone losses are statistically significant at all altitudes between 12 and 50 km,

(c) There are two clear maxima in the trends, one around 40 km altitude, the other at about 15 km.

The upper stratosphere (altitudes between about 30 and 50 km) is a region where changes in ozone were originally predicted to occur. This is a region in which the chemistry should be dominated by gas-phase reactions. When the upper stratospheric data are fit to a standard statistical model, negative trends are found throughout the region with statistically significant peak values of -6 to -8% per decade at 40-45 km altitude. There is a factor of two seasonal variation, with a maximum negative trend in winter. There is no significant inter-hemispheric difference in upper stratospheric trends based on SAGE I/II version 5.96 data extended through 1996. There is good agreement between SAGE I/II and Umkehr. The SBUV-SBUV2 combined record shows less negative trends. Less confidence is placed in the SBUV-SBUV2 result due to potential problems with the present version (6.1.2) of the NOAA-11 SBUV2 data.

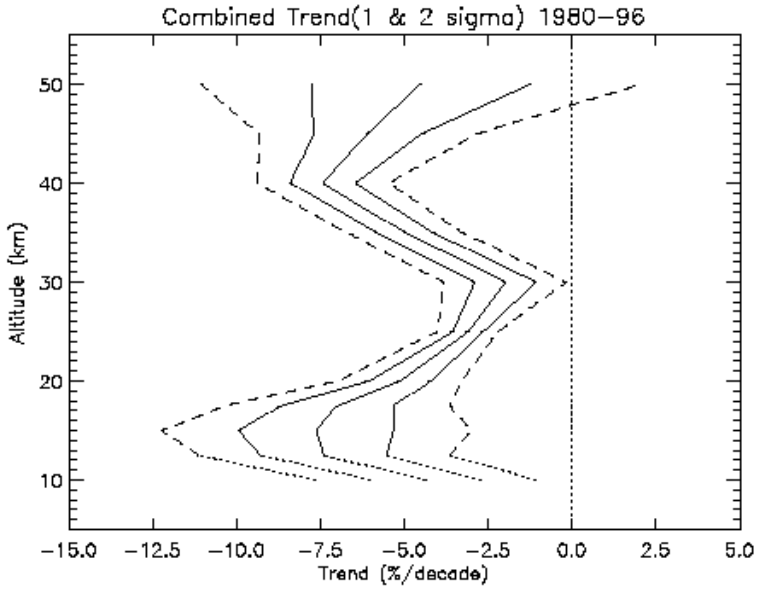


Figure 2. Estimate of the mean trend in the vertical distribution of ozone that has occurred over northern mid-latitudes from 1980-1996 (heavy solid line) calculated using the trends derived from SAGE I/II, ozonesondes, SBUV and Umkehr measurements. Combined uncertainties are shown as 1s (light solid lines) and 2s (dashed lines). The combined trends and uncertainties are extended down to 10 km as shown by the light dotted lines. The results below 15 km are a mixture of stratospheric and tropospheric trends and the exact numbers should be viewed with caution. Combined trends have not been extended lower into the

troposphere because there are concerns regarding the representativeness of any mean trends derived from the small sample of sonde stations.

The lower stratosphere (altitudes below 30 km) is the region where much of the trend which has been deduced from column data is expected to occur. The primary trend instruments in this region are sondes (up to 27 km) and SAGE (20-30 km). For sondes, sampling of the data prior to trend analysis has as much or more effect on derived trends than do the details of the statistical model. The optimal selection criteria for the use of sonde data in trend analyses is a subject of debate. Trends from 8 individual stations in the northern mid-latitudes are negative throughout the lower stratosphere. They range from -3 to -11% per decade at 20 km and are statistically significant at all stations. The trends show little seasonal variability above 20 km. The seasonal variability in the trend in the ozone profile occurs mostly in the altitude range of 10-20 km. The exact time dependence of this seasonal variability is uncertain. European stations show a winter-spring maximum, while Canadian stations show a spring-summer maximum. Details of this seasonal maximum were somewhat different in the two analyses of the sonde data. There is a reasonable agreement between SAGE I/II trends and sonde trends over the altitude region from 15 to 27 km at northern mid-latitudes. The agreement between 15 and 20 km may be fortuitous. SAGE II trends in the 15-20 km region in the tropics are much more negative than those in northern mid-latitudes but there are insufficient sonde records with which to compare these results (or those at southern mid-latitudes).

It is difficult to make generalisations concerning trends in tropospheric ozone. The only data from which to make conclusions is sonde data from a small number of stations. Trends calculated for Canadian stations are negative or near zero for the period from 1970 through 1996 and also for the period from 1980 through 1996. Trends calculated for 3 European stations are strongly positive for the period 1970 through 1996 but are essentially zero at two of these stations when data from 1980 through 1996 are considered. Trends calculated for the Japanese stations show a mixture of positive and insignificant for both time periods.

Trends in the column amount of ozone above 20 km deduced from SAGE I/II are much smaller than the column trends deduced from TOMS. The TOMS/SAGE differences are consistent with the sonde trends below 20 km. There is also a consistent seasonal variation between satellite and sonde data. Both indicate that the primary seasonal variation in mid-latitude ozone trends occurs at altitudes between 10 and 20 km with a clear maximum over northern mid-latitudes during the local winter-spring period and a much smaller seasonal cycle in the southern hemisphere.

References

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

World Meteorological Organisation (WMO), Report of the International Ozone Trends Panel: 1988, Global Ozone and Monit. Network, WMO. Report No. *18*, Geneva, 1990.

WMO/UNEP, *Scientific assessment of ozone depletion: 1994,* Global Ozone and Monitoring Network, WMO. Report No. *37*, Geneva, 1995.

Chapter 1

Characteristics of

Ozone Data Sources Used for Trend Identification

Lead Authors

David Hofmann - Clive Rodgers

Authors

Dirk DeMuer - Lucien Froidevaux - Sophie Godin

Larry Gordley - Ernest Hilsenrath - Rich McPeters - Mike Newchurch

Dave Rusch - Sam Oltmans - Joe Zawodny

1.1. Introduction

In this chapter we discuss the sources of data used to study long-term trends in atmospheric ozone, including the measurement methods used, their characteristics and their error sources, in order to understand the extent to which they should be able to throw light on the actual long-term variation. This includes the characteristics of the instruments and those of the data analysis methods used to retrieve the ozone profile from the original signals obtained by the instruments. A discussion of this kind was included in the report of the 1988 Ozone Trends Panel (WMO 1990), but much has been learned since then, and more data sources have become available, specifically the UARS instruments MLS and HALOE, and the lidars, which now have a significantly longer data record.

The magnitude of the reduction in global total ozone, as reported in the WMO/UNEP 1994 Scientific Assessment of Ozone Depletion, appears to be in reasonable agreement among several available measurement techniques, for example surface-based Dobson and satellite-based SBUV and TOMS. The situation is not as good for measurements of trends in the vertical profile of ozone, especially in the lower stratosphere. The 1994 assessment found considerable differences between the trends derived from ozonesondes and from the SAGE II instrument below about 20 km altitude. Considerably better agreement was obtained above 20 km when comparing the SBUV, SAGE and Umkehr techniques for trends.

For the 1998 Assessments, these discrepancies need to be examined in much more detail than was done in 1994. The analysis presented here is intended to determine the drift uncertainties of the calibration of the available instruments, i.e. whether they are adequate in principle to determine ozone trends to the required accuracy, and to identify problems that need attention for future measurements. The discussion will also provide information as to the suitability of data sources in validating the measurements of other instruments. To this end, we describe how the data are produced for each instrument/technique, including the generation of the raw data and the processing used to obtain the final numbers for analysis. The actual performance of instruments can only be determined by intercomparison. These aspects will be discussed in Chapter 2.

1.1.1. Instrument types and observational techniques considered

Ozone profiling instruments can be generally classified into remote and in-situ devices. Remote instruments such as the satellite sensors (HALOE, MLS, SBUV and SAGE), and ground-based instruments (Dobson-Umkehr, lidar and microwave spectrometers) obtain ozone profile information at a distance by utilising the spectral characteristics of ozone absorption, mainly in the ultraviolet but also in the visible, near infrared and microwave regions of the spectrum. In-situ devices such as balloon-borne ozonesondes and airborne UV absorption (Dasibi-type) instruments sample ambient air drawn into the instrument which is analysed for ozone by chemical or spectral analyses. Both remote and in-situ techniques can be deployed from balloons and aircraft while satellite and ground-based measurements of ozone are obviously limited to remote techniques.

For the analysis of ozone trends, instruments and techniques must have sufficiently long-term records of data with adequate quality and coverage. Ground-based Dobson instruments, observing ozone profiles in the Umkehr mode, have the longest heritage of ozone measurements with substantial records at some stations from 1957. The network of instruments is routinely inter-calibrated and produces ozone profiles in the stratosphere between approximately 20 and 40 km. Vertical resolution is relatively coarse, and spatial coverage is limited by the small number of ground stations. Lidar instruments have fewer ground

stations, but in principle, better vertical resolution and precision. They will be important monitoring instruments in the future, and their length of data record is now sufficiently long that they are considered here. Ozonesondes cover a lower altitude range than other techniques, and are important for monitoring the troposphere and lower stratosphere to about 25 km. Useful measurements are available from a small number of stations since about 1970.

Satellite instruments provide a more complete geographical coverage, but data records are shorter. SBUV (on Nimbus 7) and SBUV/2 on NOAA 11 provide daytime data for the stratosphere since 1979, and SAGE I and II provide solar occultation data (twice per orbit), also for the stratosphere, from 1979. More recently two new instruments, HALOE (infrared solar occultation)and MLS (microwave thermal limb emission) have flown on the Upper Atmosphere Research Satellite, launched in 1991. Its six-year record is not yet sufficient to identify trends satisfactorily, but they already provide useful data sources for validation, and will contribute to future trend assessments.

1.1.2. Instrument and technique advantages

No single existing ozone profiling instrument or technique is capable of measurements at all altitudes, with adequate global and temporal coverage. Thus a combination of instruments and techniques is necessary. There is a range of requirements for data from an instrument or set of instruments to measure ozone trends. These include spatial and temporal coverage and resolution, long-term stability, long-term continuity, length of record and minimal sensitivity to other time varying quantities such as volcanic aerosol.

Horizontal spatial coverage is best obtained by satellite based sensors which measure continuously, such as MLS and SBUV, although SBUV can only make daytime measurements. Occultation instruments such as SAGE and HALOE obtain measurements only at the terminator in a pattern which depends on the spacecraft orbit. Ground-based instruments are limited by the total number of stations that can be provided. The most widely distributed ground-based instrument is the Dobson, for which there are about 90 stations world-wide, although only 19 of those which make Umkehr measurements place their data in the WOUDC.

Vertical spatial coverage varies somewhat. Most of the instruments discussed make measurements in the range from about 25-50 km, but in the critical lower stratosphere from about 15-20 km, only sondes, SAGE and HALOE make useful measurements.

High vertical resolution (better than 1 or 2 km) is obtained by sondes and the occultation instruments SAGE and HALOE and over part of the height range by Lidars. MLS, SBUV and Umkehr all have a rather poorer resolution, around 5±10 km, giving rise to some questions of comparison and interpretation.

All of the radiometric data sources used have some form of internal radiometric calibration. This is particularly simple and reliable in the case of the occultation instruments. The sondes are usually, but not always, referenced to a nearby Dobson total ozone measurement. The precision, accuracy and long-term stability of all of the data sources is discussed in detail in this chapter. Most instruments, namely SAGE, HALOE, SBUV, Lidar and Umkehr are affected to some extent by volcanic aerosol, which has long-term and irregular time variation, and which is eliminated to varying degrees of success. Solar cycle effects may be present in the UV instruments, Umkehr and SBUV, and long-term temperature trends may affect any of the data sources in various ways.

All data sources are subject to problems of long-term continuity of instrumentation. The satellite instruments have a finite lifetime, even if they are reasonably stable during their life, great care is needed

when comparing with successor instruments. The same is true of ground-based instruments, but on a shorter time-scale where the temptation to improve an instrument may induce long-term variations in its record. Even the one-off sondes may be subject to long-term changes in manufacturing techniques and suppliers.

The length of the data record is of primary importance in determining trends. The UARS instruments, MLS and HALOE, have only about 5 or 6 years, whilst the SAGE I/II and SBUV record is 18 years, and the Dobson Umkehr record goes back to 1957 at some stations. The question of the influence of data record, sampling and natural variability on the accuracy with which trends can be measured is not strictly part of the error analysis of an instrument, and will be treated separately in Chapter 3.

1.1.3. Instrument and retrieval analysis approach

The error analysis of the data sources follows the same general approach as the 1988 Ozone Trends Panel (WMO 1990). The retrieved profile is taken to be a smoothed version of the true profile, with smoothing functions given by the averaging kernels (Rodgers, 1990). The width of the averaging kernels describes the vertical resolution of the measurement, and the errors in the smoothed profile estimate are analysed in terms of random and systematic errors.

For each instrument type, sources of error are identified and quantified as far as possible, the results being stated in terms of errors in the retrieved ozone profile. For estimating the accuracy with which long-term trends in ozone may be measured, it is particularly important to identify long-term drifts in the systematic errors which could be interpreted as a long-term trend in ozone. Constant systematic errors are of less importance for trends, and with a long record, random errors will generally be reduced to negligible proportions. However the random and constant systematic errors are still important when considering instruments used for validation of others. The identification of trend error sources is in its infancy, and it can involve many subtle effects that may not normally be considered for single-profile error analysis. Therefore the analysis in this chapter should be regarded as preliminary, and be treated with caution.

1.10. References

Barath, F., et al., The Upper Atmosphere Research Satellite Microwave Limb Sounder Instrument, *J.Geophys. Res.*, 98, 10,751–10,762, 1993.

Basher, R.E., Survey of WMO-sponsored Dobson spectrophotometer intercomparisons, 19, 56 pp., World Meteorological Organization (WMO), Wellington, New Zealand, 1995.

Bass, A. M., and R. J. Paur, The ultraviolet cross-section of ozone, I, Results and temperature dependence in Atmospheric Ozone proceedings of the Quadrennial Ozone Symposium, 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.

Bhartia, P.K., R.D. McPeters, C.L. Mateer, L.E. Flynn, and C.G. Wellemeyer, Algorithm for the estimation of vertical profiles from the backscattered ultraviolet technique, *J. Geophys. Res. 101*, 18,793-18,806, 1996.

Bojkov, R. D., L. Bishop, and V. E. Fioletov, Total ozone trends from quality-controlled ground-based data (1964-1994), *J. Geophys. Res.*, 100, 25867-25876, 1995.

Brewer, A.W. and J. R. Milford, The Oxford-Kew ozone sonde, Proc. R. Soc. London, Ser. A, 256, 470-495, 1960.

Browell, E.V., Differential absorption lidar sensing of ozone, Proc. IEEE, 77, 419-432, 1989.

Bruhl, C., S. R. Drayson, J. M. Russell III, P. J. Crutzen, J.M. McInerney, P.N. Purcell, H. Claude, H. Gernandt, T. J. McGee, I. S. McDermid, and M. R.Gunson, Halogen Occultation Experiment ozone channel validation, *J. Geophys. Res.*, 101, D6, 10,217-10,240, 1996.

Caudill, T.R., D. Flittner, B. Herman, O. Torres, and R. McPeters, Evaluation of the pseudo-spherical approximation for backscattered ultraviolet radiances and ozone retrieval, *J. Geophys. Res.*, *102*, 3881-3890, 1997.

Cebula, R.P., H.W. Park, and D.F. Heath, Characterization of the Nimbus-7 SBUV radiometer for long term monitoring of the stratospheric ozone, *J. Atmos. Oceanic Technol., 5*, 215-227, 1988.

Chu, W. P., M. P. McCormick, J. Lenoble, C. Brogniez, & P. Pruvost, SAGE II Inversion Algorithm, *J. Geophys. Res*,94, pp8339-8351, 1989.

Claude, H., R. Hartmannsgruber, U. K^hler, Measurment of atmospheric ozone profiles using the Brewer/Mast sonde, WMO Global Ozone Research and Monitoring Project, Report No. 17, WMO/TD, No. 179, 1987.

Claude, H., U. Köhler, W. Steinbrecht, New trend analyses of the homogenized ozone records at Hohepeissenberg, Proceedings of the Quadrennial Ozone Symposium, L'Aquila, Italy, September 1996, 1998.

Cunnold, D. M., H. C. Wang, W. P. Chu, and L. Froidevaux, Comparisons between SAGE II and MLS ozone measurements and aliasing of SAGE II ozone trends in the lower stratosphere, *J. Geophys. Res.*, 101,10,061–10,075, 1996.

Dave, J.V., Meaning of successive iteration of the auxiliary equation of radiative transfer, Astrophys. J.,140, 1292-1303, 1964.

Dave, J.V. and C.L. Mateer, A preliminary study on the possibility of estimating total atmospheric ozone from satellite measurements, *J. Atmos. Sci.*, *24*, 414-427, 1967.

Dave, J.V., Investigation of the effect of atmospheric dust on the determination of total ozone from the earth's ultraviolet reflectivity measurements, NTIS Accession No. N77-24690 thru 24692, International Business Machines Corp.,18100 Frederick Pike, Gaithersburg, Md 20760, 1977.

De Backer, H., D. De Muer, E. Schoubs and M. Allaart, A new pump correction profile for Brewer-Mast ozonesondes, Proceedings of the Quadrennial Ozone Symposium, L'Aquila, Italy, September 1996, 1998.

DeLuisi, J.J., Umkehr vertical ozone profile errors caused by the presence of stratospheric aerosols, *J. Geophys. Res.,* 84, 1766-1770, 1979.

DeLuisi, J. J., D. U. Longenecker, C. L. Mateer, and D. J. Wuebbles, An analysis of northern middle-latitude Umkehr measurements corrected for stratospheric aerosols for 1979-1986, *J. Geophys. Res.,* 94, 9837-9846, 1989.

DeLuisi, J. J., I. Petropavlovskikh, C. Mateer, and L. Thomason, An evaluation of the uncertainties in estimated stratospheric aerosol errors to retrieved Umkehr ozone profiles, Proceedings of the XVII Ozone Symposium, L'Aquila, Italy, 1996.

De Muer, D., and H. De Backer, Influence of sulfur dioxide trends on Dobson measurements and on electrochemical ozone soundings, Atmospheric Ozone Conference, Tromsö, Norway, 28-29 June 1993, SPIE Proceedings Series, Vol. 2047, 18-26, 1994a

De Muer, D. and H. De Backer, The discrepancy between stratospheric ozone profiles from balloon soundings and other techniques: A possible explanation, Proceedings of the Quadrennial Ozone Symposium, Charlottesville, VA, June 4-13, 1992, 1994b.

EN-SCI, Instruction manual for the En-Sci Corp. Model 1Z ECC O3 Sondes, En-Sci Corp, Boulder, CO, 1996.

Fegley, A.A, and W.K. Fowler, Radiometric calibration of SBUV/2 instruments: Retrospective Improvements, *Metrologia*, 28, 297-300, 1991.

Fishbein, E. F., et al., Validation of UARS MLS temperature/pressure measurements, *J. Geophys. Res.*, 101, 9983-10,016, 1996.

Flaud, J. M.., C. Camy-Peyret, C.P. Rinsland, M.A.H. Smith, and V. Malathy Devi, *Atlas of Ozone Spectral Parameters from Microwave to Medium Infrared,* Academic Press, San Diego 1990.

Frederick, J.E, R. P. Cebula, and D. F. Heath, Instrument characterization for the detection of long-term changes in stratospheric ozone: An analysis of the SBUV/2 radiometer, *J. Atmos. Oceanic Technol., 3*, 472-480, 1986.

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

Gleason, J.F, R.D. McPeters, Correction to the Nimbus 7 solar backscatter ultraviolet data in the "nonsync" period (February 1987 to June 1990), *J. Geophys. Res*., 100, 16,873-16,877, 1995.

Godin S., Ph.D. thesis, Paris Univ., Paris, 1987.

Godin S., G. Mégie, J. Pelon, Systematic Lidar Measurements of the Stratospheric Ozone vertical Distribution, Geophys. Res. Letters, Vol 16 Nû 16, 1989.

Godin S., G. Mégie, C. David, V. Mitev, D. Haner, Y. Emery, C. Flesia, V. Rizi, G. Visconti, L. Stefanutti, Ozone, aerosols and Polar Stratospheric Clouds Measurements during the EASOE Campaign, Proc. Quad. Ozone Symp., 561, NASA conf. Pub. 3266, 1994.

Godin S., NDSC workshop on stratospheric ozone lidar, Paris, 1996.

Gordley, L. L., B. T. Marshall, and D. A. Chu, LINEPAK: Algorithms for modeling spectral transmittance and radiance, *J. Quant. Spectrosc. Radiat. Transfer*, 52, 563-580, 1994.

Gotz, F. W. P., Zum Strahlungsklima des Spitzbergensommers. Strahlugns-und Ozonmessungen in der Konigsbucht 1929, Gerlands Beltr, 31, 119-154, 1931.

Gotz, F. W. P., A. R. Meetham, and G. M. B. Dobson, The vertical distribution of ozone in the atmosphere, *Proc. Roy. Soc*., A145, 416-446, 1934.

Hahn, J.F., C.T. McElroy, E.W. Hare, W. Steinbrecht, and A.I. Carswell, Intercomparison of Umkehr and differential absorption lidar stratospheric ozone measurements, *J. Geophys. Res*., 100, 25,899-25,911, 1995.

Heath, D. F., A. J. Krueger, H. R. Roeder, B. D. Henderson, The solar backscatter ultraviolet and total ozone mapping spectrometer (SBUV/TOMS) for Nimbus G, *Optical Engineering, Vol. 14*, 323-331, 1975.

Heath, D.F., Z. Wei, W.K. Fowler, and V.W. Nelson, "Comparison of Spectral Radiance Calibrations of SSBUV-2 Satellite Ozone Monitoring Instruments using Integrating Sphere and Flat-Plate Diffuser Technique," *Metrologia, 30*, 259-264, 1993.

Hervig, M.E., J.M. Russell III, L. L. Gordley, S. R. Drayson, K. A. Stone, R.E. Thompson, M. E. Gelman, I. S. McDermid, A. Hauchecorne, P. Keckhut, T. J.McGee, U. N. Singh, and M. R. Gross, Validation of temperature measurements from the Halogen Occultation Experiment, *J. Geophys. Res*., 101, D6,10,277-10,285, 1996.

Hervig, M.E., J. M. Russell III, L. L. Gordley, J. Daniels, S. R. Drayson, and J. H. Park, Aerosol effects and corrections in the Halogen Occultation Experiment, *J. Geophys. Res*., 100, 1067-1079, 1995.

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.

Jarnot, R.F., R. E. Cofield, J. W. Waters, and D. A. Flower, Calibration ofthe Microwave Limb Sounder on the Upper Atmosphere Research Satellite,*J. Geophys. Res.*, 101, 9957-9982, 1996.

Johnson, B.J., Oltmans, S. J., D. J. Hofmann, J. A. Lathrop, Evaluation of ECC ozonesonde performance from recent field and laboratory intercomparisons, Proceedings of the Quadrennial Ozone Symposium, L'Aquila, Italy, September 1996, 1998.

Keating, G.M., M. C. Pitts, and D. F. Young, Ozone reference models for the middle atmosphere (New CIRA), in *Middle Atmosphere Program Handbook for MAP*, edited by G. M. Keating, vol. 31, pp. 1-36, SCOSTEP Secretariat, Univ. of Illinois, Urbana, 1989a.

Keating, G.M., M. C. Pitts, and C. Chen, Improved reference models for middle atmosphere ozone, in *Middle Atmosphere Program Handbook for MAP*, edited by G. M. Keating, vol. 31, pp. 37-49, SCOSTEP Secretariat, Univ. of Illinois, Urbana, 1989b.

Kerr, J.B., I.A. Asbridge, and W.F.J. Evans, Intercomparison of total ozone measured by the Brewer and Dobson spectrophotometers at Toronto, *J. Geophys. Res.*, 93, 11,129-11,140, 1988.

Kerr, J.B., C.T. McElroy, and D.I. Wardle, The Brewer instrument calibration center 1984-1996, Proceedings of the Quadrennial Ozone Symposium, 21-24, 1996.

Klenk, K.F., P.K. Bhartia, A.J. Fleig, V.G. Kaveeshwar, R.D. McPeters, and P.M. Smith, Total ozone determination from the Backscattered Ultraviolet (BUV) experiment, *J. Appl. Met., 21*, 1672-1684, 1982.

Köhler, U. and H. Claude, Homogenized ozone records at Hohenpeissenberg, Proceedings of the Quadrennial Ozone Symposium, LíAquila, Italy, September 1996, 1998.

Komhyr, W.D., Nonreactive gas sampling pump, Rev. Sci. Inst., 38, 981-983, 1967.

Komhyr, W.D., Electrochemical cells for gas analysis, Ann. Geophys., 25, 203, 1969.

Komhyr, W.D. and T. B. Harris, Development of an ECC ozonesonde, NOAA Tech. Rep. ERL 200, APCL 18, , Boulder, CO, 1971.

Komhyr, W.D., Operations handbook-Ozone measurements to 40-km altitude with model 4A electrochemical concentration cell (ECC) ozonesondes (used with 1680 MHz radiosondes), NOAA Tech. Memo. ERL ARL-149, 49p, Air Resources Laboratory, Boulder, CO, 1986.

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.

Komhyr, W.D., C. L. Mateer, and R. D. Hudson, Effective Bass-Paur 1985 ozone absorption coefficients for use with Dobson ozone spectrophotometers, *J. Geophys. Res.*, 98, 20,451-20,456, 1993.

Komhyr, W.D., R. A. Barnes, G. B. Brothers, J. A. Lathrop, and D. A. Opperman, Electrochemical concentration cell ozonesonde performance evaluation during STOIC 1989, J. Geophys. Res., 100, 9231-9244, 1995.

Logan, J. A., Trends in the vertical distribution of ozone: An analysis of ozonesonde data, J. Geophys. Res., 99, 25,553-25,585, 1994.

Marshall, B. T., L. L. Gordley, and D. A. Chu, BANDPAK: Algorithms for modeling broadband transmission and radiance, *J. Quant. Spectrosc. Radiat. Transfer*, 52, 581-599, 1994.

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

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. Dütsch, J. Staehelin, and J. J. DeLuisi, Influence of a priori profiles on trend calculations from Umkehr data, *J. Geophys. Res.,* 101, 16779 - 16787, 1996.

Mateer, C. L., and H. U. Dütsch, Uniform evaluation of Umkehr observations from the world ozone network, NCAR Rep, 105 pp., National Center for Atmospheric Research, Boulder, CO, 1964.

Mateer, C.L., D.F. Heath, and A.J. Krueger, Estimation of total ozone from satellite measurements of backscattered ultraviolet Earth radiance, *J. Atmos. Sci., 28*, 1307-1311, 1971.

McCormick, M.P., J.M. Zawodny, R.E. Veiga, J.C. Larsen, and P.H. Wong, An overview of SAGE I and II ozone measurements, *Planet. Sp. Sci., 37*, 1567-1586, 1989.

McDermid I.S., S. Godin, L.O. Lindquist, Ground based lidar DIAL system for long term measurements of stratospheric ozone, Appl. Opt., vol. 29, 25, 3603-3612, 1990.

McDermid I.S., D. Haner, M. M. Kleiman, T.D. Walsh, M.L. White, Differential absorption lidar systems for tropospheric and stratospheric ozone measurements, Opt. Engineering, 1991, Vol. 30

McElroy, C. T., and J. B. Kerr, Table mountain ozone intercomparison: Brewer ozone spectrophotometer Umkehr observations, *J. Geophys. Res.*, 100, 9293-9300, 1995.

McElroy, C.T., J.F. Hahn, and E. Hare, Determining high-altitude trends in ozone from Brewer Umkehr observations made at Canadian stations, Proceedings of the Quadrennial Ozone Symposium, 25-28, 1996.

McGee T., D. Whiteman, R. Ferrare, J.J., Butler, J.F., Burris, STROZ LITE : Stratospheric Ozone Lidar Trailer Experiment, Opt. Engineering, 1991, Vol. 30

McGee, T.J., M. Gross, R. Ferrare, W.S. Heaps, and U.N. Singh, Raman DIAL measurements of Stratospheric Ozone in the presence of volcanic aerosols, Geophys. Res. Lett., 20, 955-958, 1993.

McGee T.J., P. Newman, M. Gross, U. Singh, S. Godin, A.M. Lacoste, G. Mégie, Correlation of ozone loss with the presence of volcanic aerosols, Geophys. Res. Lett., vol 21, 25, 2801-2804, 1994

McPeters, R.D., Climatology of Nitric Oxide in the upper stratosphere, mesosphere, and thermosphere: 1979 through 1986, *J. Geophys. Res., 94*, 3461-3472, 1989.

McPeters, R.D. et. al., Nimbus-7 Total Ozone Mapping Spectrometer (TOMS) data products user's guide, RP1384, NASA, 1996.

McPeters, R.D., G. J. Labow, B.J. Johnson, A satellite-derived climatology for balloonsonde estimation of total column ozone, J. Geophys. Res., 102, 8875-8885, 1997.

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.

Measures R., M., Laser Remote Sensing, John Wiley, New York, 1984.

Mégie, G., G. Ancellet, J. Pelon, Lidar measurements of ozone vertical profiles, Appl. Opt., 24, 3454-3453, 1985

Mueller, J. I., Flight preparation instructions for the Model 730-8 Ozonesonde, Mast Development Company, Davenport, Iowa, 1976.

Nakane H., N. Sugimoto, S. Hayashida, Y. Sasano, I. Matsui, Five years Lidar Observation of Vertical Profiles of Stratospheric Ozone at NIES, Tsukuba (36N, 140E), 17th ILRC, Sendaï, 1994.

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, C. L. Mateer, and J. Cao, Intercomparison of SAGE with Umkehr[64] and Umkehr[92] ozone profiles and time series: 1979-1991 submitted to *J. Geophys. Res.*, 1997.

Papayannis, A., G. Ancellet, J. Pelon and G. Mégie, Multiwavelength lidar for ozone measurements in the troposphere and lower stratosphere, Appl. Opt., 29, 467-476, 1990.

Parsons, C. L., G. A. Norcross, and R. L. Brooks, Radiosonde pressure sensor performance: Evaluation using tracking radars, J. Atmos. Ocean. Technol., 1, 321-327, 1984.

Paur, R. J., and A. M. Bass, The ultraviolet cross-section of ozone: II. Results and temperature dependence, *Atmospheric ozone*, Zerefos, C.S. and Ghazi, A. (Eds), p. 611, Reidel, Dordrecht, 1985.

Petropavlovskikh, I., J. J. DeLuisi, and P.K. Bhartia, A critical look at the capabilities of the Umkehr and SBUV for ozone information below 20 km, Proceedings of the XVII Ozone Symposium, L'Aquila, Italy, 1996a.

Petropavlovskikh, I., J. J. DeLuisi, B. Herman, R. Loughman, P.K. Bhartia, C.L. Mateer, J. Lenoble, and Y. Belikov, A comparison of radiance calculations by spherical atmosphere radiation transfer codes for modeling the Umkehr effect, Proceedings of the XVII Ozone Symposium, L'Aquila, Italy, 1996b.

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

Rodgers, C. D., Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, Rev. Geophys., 14, 609–624, 1976.

Russell, J.M., III, L.L. Gordley, J.H. Park, S.R. Drayson, A.F. Tuck, J.E. Harries, R.J. Cicerone, P.J. Crutzen, and J.E. Frederick, The Halogen Occultation Experiment, *J. Geophys. Res.*, 98, 10,777-10,797, 1993.

Staehelin, J., H. Schill, B. Högger, P. Viatte, G. Levrat, and A. Gamma, Total ozone observation by sun photometry at Arosa, Switzerland, *Opt. Engin.,* 34, 1977-1986, 1995.

Steele, H.M., R.P. 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, 19665-19681, 1997.

Steinbrecht W. and A.I. Carswell, Evaluation of the effects of Mount Pinatubo aerosol on differential absorption lidar Measurements of stratospheric ozone, J. of Geophys. Res., 1995, 1215-1233.

Strand, O. N., and E. R. Westwater, Statistical estimation of the numerical solution of a Fredholm integral equation of the first kind, *J. Atmos. Comp. Mach.,* 15, 100, 1968.

Swaart D.P.J., A. Apituley, J. Spakman, E. P. Visser, H. B. Bergwerff, RIVMs tropospheric and Stratospheric Ozone lidars for European and Global Monitoring networks, 17th ILRC, Sendaå, Japan, 1994.

Tarasick, D. W., D. I. Wardle, J. B. Kerr, J. J. Bellefleur, and J. Davies, Tropospheric ozone trends over Canada: 1980-1993, Geophys. Res. Lett., 22, 409-412, 1995.

Thomas, R.W.L. and A.C. Holland, Simple relationship between the UV radiation backscattered by the Earth's atmosphere and the vertical ozone profile, *Appl. Opt., 16*, 2581-2583, 1977.

Torres, A. L., ECC ozonesonde performance at high altitude: Pump efficiency, NASA Tech. Memo. 73290, 10pp, 1981.

Twomey, S., On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear systems produced by quadrature, *J. Assoc. Comp. Mach., 10*, 97-101, 1963.

Uchino, O. and Isao Tabata, Mobile lidar for simultaneous measurements of ozone, aerosol, and temperature in the stratosphere, Appl. Opt., Vol 30, Nû 15, 1991.

Uchino,O, M. Maeda, J. Kohno, T. Shibata, C. Nagasawa, M. Hirono, Observation of stratospheric Ozone layer by XeCl Laser radar, App. Phys. Lett., 33, 807-809, 1978.

Wang, H. J., D. M. Cunnold, & X. Bao, A Critical Analysis of Stratospheric Aerosol aand Gas Experiment Ozone Trends, 1996, *J. Geophys. Res.*, 101, pp 12495-12514.

Waters, J. W., Microwave limb sounding, in *Atmospheric Remote Sensing by Microwave Radiometry*, edited by M. A. Janssen, chap. 8, J. Wiley, New York, 1993.

Werner J., K.W. Rothe, H. Walther, Monitoring of the Ozone Stratospheric Layer by Laser Radar, Appl. Phys. B., 32, 113-118, 1983.

WMO, Handbook for Dobson Ozone Data Re-evaluation, World Meteorological Organization Global Ozone Research and Monitoring Project – Report No. 29, Geneva, 1996.

Yarger, D.N., An evaluation of some methods of estimating the vertical atmospheric ozone distribution from the inversion of spectral ultraviolet radiation, *J. Appl. Met., 9*, 921-928, 1970.

1.2. SAGE

1.2.1. SAGE I/II Basic Measurement Description

An initial report and review of the SAGE measurement technique and instrument was presented in the Report of the International Ozone Trends Panel 1988 (WMO, 1988). The bulk of that material remains accurate and only a few changes or pertinent facts and overview material will be presented here.

The SAGE measurements of ozone are made through the solar occultation technique. Twice during a typical orbit, the SAGE instrument enters or exits the Earth's shadow. During these transitions the instrument views the sun through the atmosphere and scans across the solar disk in a plane parallel to the local vertical. For measurements high above the Earth's atmosphere, the sun is unattenuated by the atmosphere and these data can be used to calibrate the instrument responsivity. The high altitude, I_0, scans are ratioed to the other scans to produce atmospheric transmission profiles. The fundamental measurement made by SAGE is a spectral slant path atmospheric transmission profile as a function of altitude. These transmission profiles are then used to infer aerosol extinction and species density profiles of ozone, NO_2, and H_2O as a function of altitude.

1.2.2 SAGE II Measurement spatial and spectral resolution

The solar occultation technique allows for a very narrow field of view to be utilised. Projected on to the limb of the Earth, SAGE II has a vertical field of view that is less than 0.5 km. The sampling and scanning rates are such that adjacent data points overlap by 50% (spaced by ~250 m). Because of the rate at which the sun rises or sets due to the orbital motion of the spacecraft, between 4 and 10 individual scans across the sun may contain observations of a particular altitude. For a 1 km altitude bin there are no fewer than 12, and sometimes as many as 30 separate data samples or transmission measurements. The horizontal extent of the measurements is 200 km along the line of sight by 2.5 km for a 1 km retrieved layer.

SAGE II is a seven channel spectrometer with channels at the following wavelengths: 385, 448, 453, 525, 600, 940, and 1020 nm. With the exception of the 448 and 453 nm pair the spectral resolutions (FWHM) are between 10 and 20 nm. The 448 channel has a bandwidth of 3 nm and the 453 channel a bandwidth of 2 nm.

1.2.3 SAGE II sampling

The location of the measurements is determined by the orbit and is not programmable in flight. SAGE I and II were placed in a mid-inclination, 57o, orbits to provide nearly global coverage (complete seasonal sampling from 55°S to 55°N). Two measurements are made during each orbit, in a pattern similar to that of HALOE (Figure 1.5).

1.2.4 SAGE II instrument error sources

There are three primary instrumental sources of error with SAGE II. The first source of errors is due to the scan mirror reflectivity. The scanning of the field of view is accomplished by moving a scan mirror

slightly less than 2.5o. The relative spectral reflectivity of the mirror must be known so that the I_0 scans at one angular extreme can be transferred to the low altitude observations made at the other extreme. The mirror properties are monitored over a range of angles where the sun is between 100 and 250 km tangent height and found to vary by about 0.1% over the complete range of angular motion. The error in scan mirror reflectivity versus angle is of the order of 5%. The mirror reflectivity vs. angle is assumed to be linear for this small angular range. Early in the mission, the ERB satellite upon which SAGE II orbits was flipped upside down so that the unattenuated sun could be viewed over the full range of mirror motion to confirm the linear model. For the first four years of operation, the mirror properties remained stable. Since 1989 the mirror properties have been changing slowly with time, probably due to the oxidation or removal of the optical coating by atomic oxygen. In the current manner of processing the SAGE II data, mirror calibration data are processed in one year segments and applied retroactively. Data for the current year is processed with last year's (December 31st) value until the end of the current year when the calibration data are processed and the measurements reprocessed.

The second source of instrumental error is the assumption of radiometric stability during an occultation event (<4 minutes). The I_0 scans can only be ratioed to other scans if the instrument sensitivity does not change during the event. In general this is true for SAGE II except for the 448 and 453 channels. It has been found that these channels exhibit a change in responsivity with a decay time of approximately 20 seconds. This does not represent a problem for sunset events because the I_0 scan data can be taken after the initial transient. For sunrise, however, this transient occurs in the atmosphere. Fortunately, the transient is only significant in the 448 and 453 nm channels where it is a 0.25% effect. It is less than 0.05% in the other channels. While the cause is not known, the effect has been named the "thermal shock". The magnitude of the effect as a function of wavelength appears to be related to the amplifier gain (narrowness of the channel spectral bandwidth) and thus leads one to conclude that the cause is probably electronic in nature rather than thermal. Nevertheless, this error source affects the NO_2 retrievals primarily. While there is the potential for errors in the NO_2 retrieval to affect the retrieved ozone indirectly by way of an error in the short wavelength aerosol extinction, there are no observable effects on the ozone profiles. A method for correcting the effect was developed for Version 5.94 of the SAGE II algorithm and is presented in Section 1.2.6.

The last source of instrumental uncertainty relates to the long-term stability of the location and bandwidths of the channels. There is no way of determining the spectral location and bandwidths of the seven SAGE II channels in orbit. Channel locations and bandwidths are set by one of two methods. The SAGE II spectrometer disperses the spectrum and a series of slits and lenses determines the spectral characteristics of most of the broadband channels. The 448 nm channel is too narrow for this method and so the bandwidth is set by an interference filter. Similarly, the 453 and 940 channels employ interference filters since they are in the zero order beam of the grating. It has been determined that the 448 nm channel has been drifting in wavelength and increasing its bandwidth continuously and immediately after launch. To the best of our knowledge, this channel has moved on the order of 1 nm to longer wavelengths in the last 13 years and the bandwidth has increased by 30%. Again, this affects the NO_2 measurements primarily and there is no observable effect on the ozone measurements at the 1% level. A first order correction for this drift in instrumental throughput is included in Version 5.96 of the SAGE II processing algorithm and is discussed later. There is no evidence to suggest that any of the other channels have moved or changed in any way. The changes in the 448 nm channel are thought to be due to the gradual movement, tilting, of the filter in its mount due the relaxation of residual stresses.

1.2.5. Corrections Needed to Link SAGE I and SAGE II

From previous ozone trend evaluation reports, it has been generally accepted that SAGE I's 385 nm filter was actually centred at 382 nm and that this caused an error of approximately 300 m in the reported SAGE I altitudes. The reason that the altitudes were affected by this wavelength error is that the SAGE I ephemerides were somewhat unreliable and the altitudes were therefore constrained by comparisons between the molecular scattering signature in the 385 nm channel and the densities estimated from the United States National Weather Service (NWS) temperature measurements.

More recently Wang *et al*. (1996) have shown evidence that the SAGE I height adjustment should vary with latitude probably because of small uncertainties in the geopotential heights in the upper stratosphere in the NWS measurements in 1979-1981. The first piece of evidence consisted of comparisons between proxies for the estimated heights of the ozone layer during the SAGE I and SAGE II observation periods (Figure 1.1a). An independent evaluation of the SAGE II height error was made by comparing the mean SAGE I and SAGE II (v 5.96) differences above the ozone peak using SBUV observations as a transfer standard (Figure 1.1a). Confidence in the results is provided by the fact that the error bars overlap for the two independent estimates at almost all latitudes.

Ideally, the SAGE I ozone profiles should have been retrieved again using the inferred altitude offsets. However for this report, an ad-hoc procedure has been used: the mean of the offsets shown in Figure 1.1a (see Figure 1.1b) was applied to move the previously retrieved ozone concentration profiles uniformly upward or downward. This procedure is believed to be quite accurate because the Rayleigh contribution to the 600 nm ozone retrieval is small at altitudes above 20 km (Chu *et al*., 1989). Below 20 km altitude however, a new retrieval of the SAGE I ozone profiles is necessary. The current SAGE I results below 20 km altitude are therefore considered to be unreliable. Above 20 km altitude, Figure 1.1b shows that there is an uncertainty of 100-150 m in the offset which should be considered in the evaluation of trends. This corresponds to an uncertainty in the SAGE I ozone values at 1 hPa of approximately 3%. Unfortunately some of the comparisons in Chapter 2 used a SAGE I offset which differed from those shown in Figure 1.1b but in almost all cases it was less than 100 m. This had no effect on the conclusions in that Chapter.

1.2.6. SAGE II Algorithm description

The version of the SAGE II algorithm used in the report is Version 5.96. What follows here is a brief description of the changes to the algorithm which was analysed and reviewed for the 1988 assessment. A lengthy and detailed discussion of the SAGE II algorithm can be found in the Report of the International Ozone Trends Panel 1988 (WMO report no. 18, 1988, pages 148-154) as well as Chu (1989). Since the WMO 1988 assessment, several changes and improvements have been made to the algorithm to account for known deficiencies. In general each distinct improvement is assigned a new version number.

Until the publication of this report Version 5.93 was the current public release of the SAGE II measurements and was basically the product described in the 1988 assessment. Version 5.94 includes a correction for the sunrise NO_2 data. In addition to the changes in the sunrise NO_2 profile there are indirect changes in the 385 nm aerosol extinction profile. The technique for isolating a responsivity change in the data is based on the temporal bias in the measured transmission at a particular altitude. Assuming that any measurement of the transmission at a particular altitude is as good as any other, the presence of a time correlated bias in transmission measured from successive scans is evaluated in the ratio of the 448 to 453 nm channels (this ratio is used to determine the NO_2 slant column). This method produces a measure of the rate of change in responsivity and must be integrated in time to produce a correction. Only the 448/453 nm channel ratio is corrected and the effect of the correction is to change the sunrise NO_2 profile with little if any effect on the retrieved ozone profile. Again, errors in NO_2 may

indirectly affect the ozone profile via errors in the short wavelength extinction. The changes in ozone are found to be less than 1%.

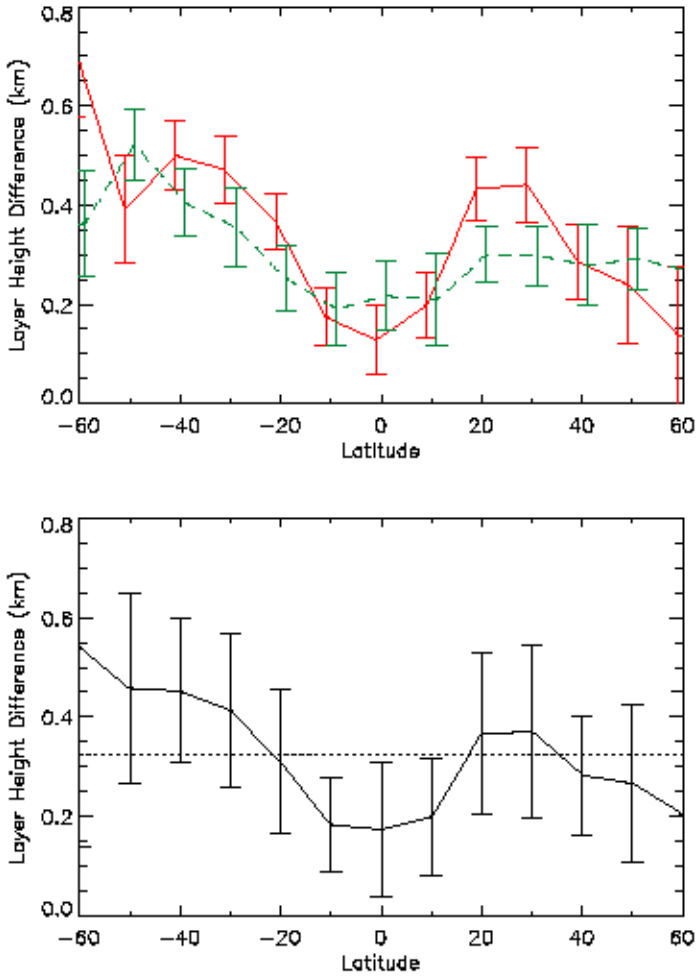


Figure 1.1. The estimated altitude error in the SAGE I observations (SAGE I profiles should be moved upwards by the amounts shown). The upper panel (a) shows the error estimated from comparisons of a proxy for the ozone layer height from SAGE I and SAGE II observations (full line, offset by 1° of latitude southward) and from mean differences between SAGE I and SBUV ozone observations in the upper stratosphere (dashed line, offset by 1° of latitude northward) (see Wang *et al.*, 1996). The lower panel (b) shows the SAGE I height error estimate based on the mean of the two results in the upper panel (a); the standard deviation is based on the differences between the two results - this is particularly large at 60°S.

Version 5.95 contains an improved spectral aerosol extinction model which is designed to reduce the aerosol interference seen in the SAGE II ozone data. This version also includes a more realistic error estimate in the species retrievals attributed to uncertainties in estimating the aerosol extinction. This is discussed in Section 1.2.7. Version 5.96 accounts for a slow wavelength drift in one of the two differential NO_2 channels, the 448 nm channel. The inclusion of this improves the long-term stability of the NO_2 data set over the last 13 years. Like Version 5.94, this change has little if any impact on the retrieved ozone profile.

At this time, the SAGE II V5.96 ozone data can be used from 15 to 50 km for long-term trend studies and from 11 to 60 km for studies which do not strictly require long-term stability of the data set (e.g.

spatial and seasonal investigations). The lower limits of 15 and 11 km must be above the tropopause. While there is a well known 10-15% difference between SAGE and ozonesonde profiles in the 15-20 km region, this difference is very stable with time and does not appear to affect the derived SAGE trends.

SAGE ozone retrievals require knowledge of the Rayleigh scattering contribution to the observed extinction at all wavelengths. Since SAGE I/II do not directly measure atmospheric density they rely on an external source (NCEP) for density information. At 55 km and again at 18 km the Rayleigh extinction is comparable to the ozone extinction. Any systematic or trend errors in density at these altitudes are represented one for one as errors in SAGE ozone.

There is an inconsistency in the use of an oblate Earth model in the SAGE algorithm. While the spacecraft ephemeris calculations use the oblate Earth model, the refraction calculation does not. This has the effect, especially for events taken at large β-angle where the events are 'smeared' over several degrees of latitude, of slightly stretching or compressing the altitude scale. Due to the peculiarities of the SAGE II orbit, the pattern of seasonal sampling, latitude and β-angle repeats from year to year so this bias in altitude is primarily a systematic error. The repetition from year to year is not exact however, and there may be a few isolated latitudes and seasons where the bias may appear as a trend over the SAGE II record.

1.2.7. SAGE II Error and retrieval characterisation

A primary function of the SAGE II retrieval algorithm is the estimation of multi-wavelength aerosol line-of-sight optical depth simultaneously with the determination of line-of-sight ozone density using measurements at 385, 453, 525, 600, and 1020 nm. Since ozone significantly contributes to extinction at 453, 525, and 600 nm and aerosol contributes at all wavelengths, the solution is constrained by asserting that the aerosol contribution at 600 nm (where ozone absorption is the strongest) can be expressed as a linear combination of the aerosol contribution at the other 4 wavelengths. Ultimately, the ozone line-of-sight optical depth at 600 nm can be expressed as

$$\tau(\Omega) = \left(\tau(600) - \sum_i a_i \tau_i \right) / \left(1 - \sum_i a_i \sigma_i \right)$$

where i is the measurement number (1 to 4), the vector $\boldsymbol{a}$ contains the coefficients inter-relating aerosol at 600 nm to that at the other channels, and cs contains the ozone cross sections at those four channels (relative to that at 600 nm). The values used for $\boldsymbol{a}$ can be derived in a number of ways but cannot unambiguously determine the 600 nm aerosol contribution.

Prior to Version 5.95, these coefficients were derived using a fairly tightly constrained linear inversion that performed well for moderate to low aerosol loading. However, in these earlier versions, a positive bias in the retrieval of ozone was observed during periods of elevated aerosol loading including following the eruption of Mt Pinatubo in 1991 and, more subtly, early in the SAGE II observational period (1984-1985). For Version 5.95, coefficients have been derived that perform better at high aerosol loading using a similar, though less tightly constrained, method. Figure 1.2 shows an example of the aerosol extinction estimated with the new coefficients during the post Pinatubo period. In the new approach, the coefficients are allowed to vary with altitude such that the previous "low-loading" coefficients are used at altitudes above 35 km. Unlike previous versions, a component reflecting the ability of this method to estimate aerosol at 600 nm has been included in the ozone error budget. Under most circumstances this does not significantly impact the estimated errors, however, following the eruption of Mt Pinatubo in 1991 errors are considerably larger where the aerosol is itself larger. It should

be noted that this component of the error budget is not necessarily unbiased particularly immediately after the Mt Pinatubo eruption when aerosol size distributions (and extinction spectra) were more complex.

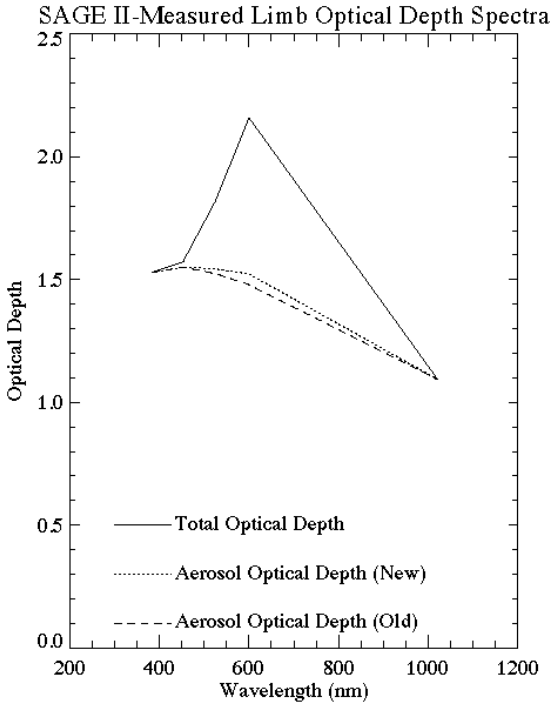


Figure 1.2. This figure shows an example of SAGE II-measured limb or line-of-sight optical depth for an event shortly after the eruption of Mt Pinatubo. The derived aerosol optical depth spectra for this event using the improved (5.96) and old versions (5.931) of the SAGE II processing algorithm are also shown. The new algorithm typically reduces ozone (relative to older versions) in this and other heavily volcanically perturbed periods due to the presence of large aerosol. (Northern mid-latitudes, March 1992, 20.5 km)

There is an apparent altitude bias in the SAGE II (and presumably SAGE I) data in the 15 to 20 km region. This bias, which can be 300 to 500 m, is likely due to a bias in the solar edge detection algorithm in the presence of a strong gradient (curvature actually) in the 1020 nm extinction profile. This bias is probably the cause of the long standing discrepancy between sondes and SAGE in this altitude region. The ozone gradient with height in this region is very steep, changing by 3% in 100 m. The stability of the ozone difference between SAGE and sondes indicates that any long-term trend in the potential altitude bias is less than 200 m over the SAGE II measurement period. The magnitude and temporal variation of this bias will be determined as part of the Version 6.00 development.

1.2.8. SAGE II Outstanding problems and future plans

Version 5.96 still contains some known deficiencies which are under investigation. There remains a small unexpected and unexplained diurnal variation/bias in aerosol extinction and ozone density. The effect is small, a few percent and appears to be constant in time. This may or may not be related to the "January bias" reported by Wang *et al*. (1996). In addition there are at least two known improvements planned for the SAGE II algorithm.

Version 5.97 will merge the water vapour retrieval with that for the other species and include corrections for the mutual interference between H_2O and O_3 in the troposphere. It is anticipated that this will result in a significant improvement in the ability of SAGE II to measure tropospheric ozone.

Version 6.00 utilises an entirely new transmission algorithm to produce spectral transmission profiles with higher vertical resolution (approaching 0.5 km) and significant reductions in uncertainty (scatter or error bars). This new algorithm should remove an altitude bias in the 15 to 20 km region discussed previously.

1.2.9. SAGE Error Analysis Summary

		Ozone error magnitude, 2s	
Source	Region (km)	Systematic (%)	79-96 drift (%/year)
SAGE I ephemeris	15-20	10	0.5
SAGE I Rayleigh subtraction	15-20	-4	-0.2
Aerosol extinctiond	15-30	10a	<0.2
Edge time biasd	15-30	15ab	<0.2b
b-angle and oblateness	<20c	10c	<0.1
448-nm channel filter	30	<1	<0.1
spectroscopy	all	5	<0.1
Scan mirror calibration	>50	5	<0.5
Trends in density (Rayleigh)	>50	?	<0.1

a Some uncertainty on how these errors may be partitioned

b For ozone with error estimates less than 12%

c for certain latitudes and seasons: high b-angle

d may be correlated

Table 1.1.

1.3. HALOE

1.3.1. Basic Measurement Description

The HALogen Occultation Experiment, HALOE, on board the Upper Atmosphere Research Satellite (UARS) performs solar occultation measurements to infer high resolution mixing ratio profiles of trace gases. The technique involves measuring the atmospheric absorption of sunlight as the sun rises and sets relative to the spacecraft. Two radiometric measurement techniques are used to sound the atmosphere. Simple broadband radiometry is used to measure absorption of sunlight by the limb of the atmosphere in channels for CO_2 (2.8 mm, used for pressure registration and temperature/pressure retrieval), H_2O (6.6 mm), NO_2 (6.25 mm) and O_3 (9.6 mm). The measured radiometer voltage during an event (sunrise or sunset) is divided by the exo-atmospheric voltage (observations above the atmospheric limb), giving a direct measurement of the atmospheric transmission for each channel. The retrieval algorithm then models the transmission in each channel to infer mixing ratio profiles.

Broadband gas correlation radiometry is the second measurement technique. Although not used directly for ozone retrievals, these channels are used for inferring aerosol extinction, essential for accurate ozone retrievals. The gas correlation method involves splitting the incoming beam, and passing one branch through a cell containing the same gas to be measured. Signals from separate detectors for the two beams are subtracted and divided by the exo-atmospheric measurement of the beam which bypasses the gas cell. This quantity, referred to as the modulation signal, is modelled in the retrieval algorithm to infer mixing ratio profiles in channels for CH_4 (3.46 mm), HCl (3.4 mm), NO (5.26 mm) and HF (2.45 mm). For a detailed description of the HALOE measurement and retrieval techniques, see Russell *et al.* (1993). The discussion that follows will primarily focus on the O_3 channel. However the CO_2 channel will be discussed as well, since the ability to accurately retrieve O_3 is dependent on the pressure registration and temperature retrievals. The temperature and pressure are retrieved above 35 km, with NCEP temperature fields used below 35 km. Since both CO_2 and O_3 channels are broadband radiometers, we will concentrate on the broadband radiometer technique. We will also only discuss uncertainties that lead to ozone retrieval errors.

The O_3 and CO_2 channels are radiometers. Each outputs a voltage sampled in time, $V(t)$, which is proportional to the measured radiance. We call the exoatmospheric voltage Vexo. The broadband transmission $\tau(t)$ can be estimated by:

$$\tau(t) = V(t)/V\text{exo} \quad (1.3.1)$$

Note that this ratio measurement does not require absolute radiance calibration, a feature that greatly benefits the ability to infer long-term trends.

1.3.2 Measurement Spatial and Spectral Resolution

The angular resolution of the HALOE Instantaneous Field-of-View (IFOV) is 2 arc minutes in elevation by 6 arc minutes in azimuth. The elevation component translates to roughly 1.6 km when projected to the limb tangent point. However the effective vertical resolution, about 2.3 km, is determined by the total system modulation transfer function (MTF), which is the combined effect of optical resolution and electronic signal smoothing in time.

Figure 1.3 is included to help visualise the projection of the IFOV onto the solar disc image. As the sun sets, its refracted image shrinks in apparent elevation and the IFOV, which is positioned at a fixed angular distance from the top edge, covers more area of the sun and moves closer to the centre of the limb darkening curve. This effective changing source function is explicitly modelled using scans of the solar disk viewed above the atmosphere and detailed limb path refraction calculations.

The spectral range of the O_3 channel covers roughly 960-1065 cm^{-1}. The main absorbers in this spectral region are (in order of importance): O_3, H_2O, CO_2, N_2O, CH_4 and aerosol with O_3 contributing 98% or more of the absorption above the troposphere, except for heavy aerosol loading situations where aerosol can dominate the interference. The spectral range of the CO_2 channel covers 3530-3610 cm^{-1}, centred near 2.8 mm. The major absorbers in this spectral region are CO_2, H_2O, and N_2O with CO_2 contributing 98% or more of the absorption. The spectral bandpass filters underwent ground testing to determine the temperature dependence of the spectral response. The filter temperatures are monitored onboard, and the values are used explicitly to model the spectral temperature dependence. Figure 1.4 shows the spectral response of the O_3 channel.

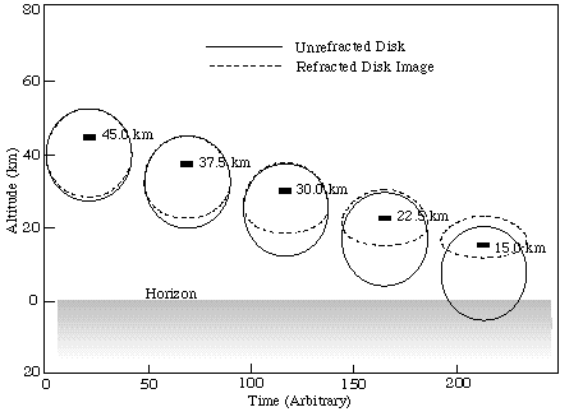


Figure 1.3. Refracted and unrefracted images of the setting sun, with the HALOE IFOV represented by the black rectangle, and tagged with the apparent tangent point altitude. Note that the IFOV is locked (fixed) relative to the top apparent edge of the sun.

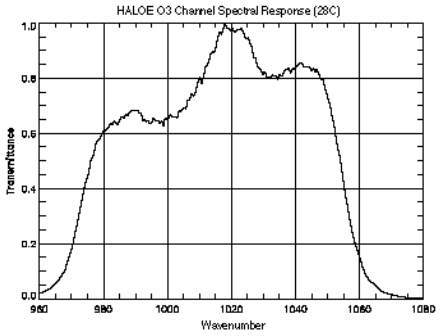


Figure 1.4. The HALOE ozone channel relative spectral response curve at nominal operating temperature of 28°C.

1.3.3. Sampling Strategy

Since HALOE is an occultation experiment, the atmospheric sounding locations depend on spacecraft orbit, rotation of the Earth and the location of the sun. The geometry of the UARS orbit (57° inclination, circular and 600 km) results in 15 sunrises and 15 sunsets daily, with tangent point locations during a day generally confined to small latitude bands which cycle over varying extremes, roughly every 36 days. Figure 1.5 shows the global sunrise and sunset locations for 1995.

The sample rate is 8 Hz which equates to a vertical tangent altitude grid of 0.18 to 0.4 km depending on apparent sun sink rate. This oversamples the effective 2.3 km measurement resolution by a factor of 5 or greater. The data is interpolated to a 0.3 km grid for processing by the retrieval algorithm.

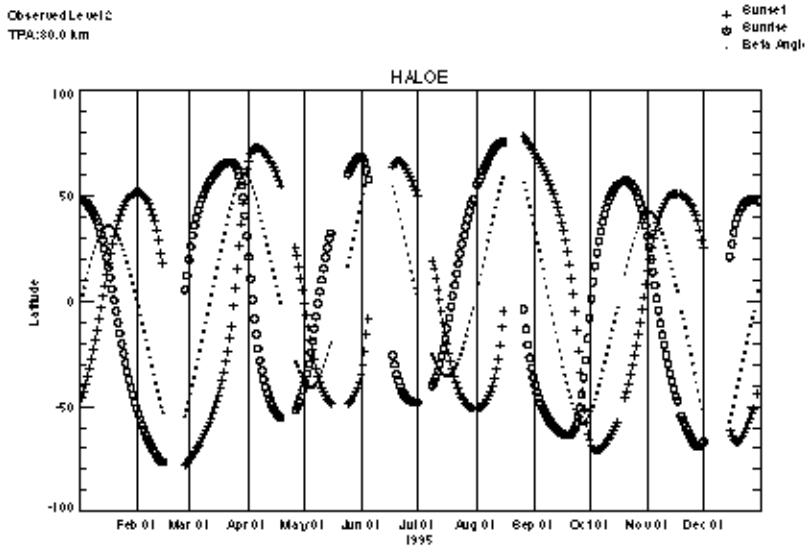


Figure 1.5. The HALOE global sampling pattern for 1995. The **o** represent the daily mean latitude for the sunrises and the **+** represents the daily mean latitude for sunsets. There are 15 sunrise and 15 sunset occulations each day, confined to small latitude bands and separated by approximately 24° in longitude.

1.3.4. Instrument Error Sources

The individual profile error sources for the HALOE instrument have been well characterised and discussed elsewhere (see Russell *et al*. (1993), Bruhl *et al*. (1996), Hervig *et al*. (1996)). In this section, we describe the main sources of instrument error that affect profiles at lower altitudes (below 50 km), as summarised in Table 1.2.

Pointing uncertainty

The uncertainty in IFOV pointing knowledge is often the dominant source of measurement error in the lower stratosphere and below. For solar occultation, the pointing geometry can be precisely determined from spacecraft location and solar ephemeris data, provided the sun is accurately tracked. The HALOE instrument tracks the top edge of the solar disk image and relies on ephemeris data and an atmospheric refraction model to compute tangent point locations. The onboard closed-loop tracking system maintains (at a 64 Hz sample rate) the IFOV at a fixed angular distance below the solar top edge image to a 16.2

arc second precision, stepping every 1/8 second. This provides a pointing precision of 100 meters or better (at the limb) when combined with oversampling and smoothing over the science measurement time constant of about 1 second. The system has proven extremely reliable at altitudes throughout the stratosphere and mesosphere, inducing random retrieval mixing ratio errors of <3% in clear air.

In regions of heavy aerosol or cirrus clouds, conditions can occur in which the algorithm driving the tracker cannot accurately distinguish the top edge of the sun. When the ozone mixing ratio near the troposphere is extremely small (<0.1 ppm) and the overburden large and steep, percent errors due to pointing can be large with absolute magnitude reaching 0.1 to 0.2 ppm near the 50 to 70 hPa level in the tropics. However, the absolute mixing ratio error reduces to well below these values at lower altitudes in clear air. At mid and high latitudes the precision at tropopause height in clear air can approach 0.01 ppm.

	Profile Altitude Region		
Error Source	Mid-Strat. >20 km	Lower Strat. 16 - 20 km	Troposphere <16 km
Pointing Uncertainty	<3%	3 - 10%	10 - 100%
Source Function	<1%	<1%	<1%
Relative Spectral Response	<1%	<1%	<1%
Noise	negligible	negligible	negligible

Table 1.2. The main sources of instrument error that affect HALOE profiles below 20 km, in descending order.

Source Function

HALOE makes exo-atmospheric scans of the sun before sunset and after sunrise events. The scans are measured at high spatial resolution and provide well calibrated source functions. These source functions are also used in the forward signal model to remove the effects of tracking jitter (<100m) which is accurately measured (as discussed above). On occasion, sunspots can be detected in the source functions, although this is an infrequent problem (<5% of events have sunspots large enough to cause significant retrieval error, even early in the mission which was during solar maximum). Because of the orbital geometry, the area of the solar disk viewed above the atmosphere during calibration scans can be slightly different than viewed during occultation. In effect, the solar image appears to rotate, moving into or out of the IFOV. If the sunspot is large enough and positioned near the edge of the IFOV, tracking jitter and/or the apparent solar disk rotation can induce results inconsistent with the modelled source function. The effects are minor in the ozone channel at 10 mm where sunspots are less pronounced, even in times of extreme sunspot activity. The few events that contain significant error due to sunspots are noted in data quality information in the HALOE quality files.

Relative Spectral Response

The spectral response functions of the broadband radiometers are believed to be extremely stable. This is based partly on the analysis of witness filters, which have been found to be stable for over 15 years. Also, peak voltages measured during calibration scans of the sun have been trended since launch to detect any degradation in radiometer response. As of present (over 6 years since launch), the ozone

channel has degraded 1.3% in peak value and the CO_2 channel <2%. There are indications that these effects are non-spectral, in which case they would have no effect on results. However even spectral filter changes, consistent with these response changes, would have negligible effect on results since the mean atmospheric spectral characteristics for these channels would not change appreciably.

Noise

The on-orbit measured instrument noise for the HALOE O_3 channel is of the order of £ 0.00178 Volts. Dividing by the measurement above the atmosphere for the O_3 channel (typically 3.2 Volts) gives an attenuation noise value of 0.00056. In the stratosphere, the attenuation signal (1-t) due to a 2.3 km tangent layer gives S/N values greater than 500, which makes instrument noise values insignificant at stratospheric levels and below.

1.3.5. HALOE Algorithm

The HALOE processing system converts solar observations during occultations into temperature, pressure, mixing ratio and extinction profiles. The algorithm first uses the measurements, including solar scans, to produce band averaged transmission, t(q), profiles as a function of zenith viewing angle, q, for 8 broadband channels. It also infers transmission difference profiles, Dt(q), for the four bandpasses that have additional gas correlation signals. The next step is registration of the profiles with altitude and pressure by modelling the CO_2 channel transmission using temperature and pressure as a function of altitude from an auxiliary data source, NCEP (National Center for Environmental Prediction). These are actually iterative steps since the inferred q depends on calculated refraction effects that alter the apparent location of the solar source image.

Next the retrieval of temperature and pressure above 35 km is performed, using the pressure registration and NCEP temperature at and below 35 km as a lower boundary condition. This is followed by an NO (nitric oxide) mixing ratio retrieval using the NO gas correlation signal divided by the NO broadband signal, Dt/t, which can be shown to be independent of aerosol extinction. Aerosol extinction is then retrieved using the NO mixing ratios and the NO measured broadband t profile. Using Mie theory and index of refraction data for sulphate aerosol, an aerosol extinction profile is modelled for the broadband channels (Hervig 1995). Ozone and water are then retrieved followed by a repeated temperature retrieval that includes aerosol extinction. Next all the above steps, starting with pressure registration using the retrieved temperature instead of NCEP data, are repeated to refine both the pressure and temperature results. Table 1.3. lists the 35 HALOE data processing steps performed for each occultation. The procedure was made sequential for computational efficiency and system simplicity. Simulation determined that this does not compromise accuracy.

Forward Model

The forward models used in the HALOE processing are comprised of subsets of the BANDPAK (Marshall, 1994) and LINEPAK (Gordley, 1994) codes. The gas correlation measurements are simulated by rigorous line-by-line calculations using LINEPAK routines. The broadband measurements are all simulated using BANDPAK routines applying either the CGA (Curtis-Godson Approximation) or PMA (Pseudo Mass Approximation) broadband calculations along with a spectral overlap correction for secondary species, all described in Marshall 1994. The choice of approximation method was determined by accuracy analyses using line-by-line methods.

Steps	Parameter	Forward model
1	t(q) and Dt(q) data	
2, 12	PR (BBR CO_2)	PMA
3, 11, 13, 21	t/p (BBR CO_2)	PMA
4, 14, 22	NO (GCR)	LBL
5, 15, 19, 23	AE_NO (BBR)	PMA
6, 9, 16, 19, 24	AE_CO_2, H_2O, NO_2, O_2 (modelled)	
7, 17, 27	O_3 (BBR)	CGA
8, 18	AE_NO (BBR)	PMA
10, 20, 25	H_2O (BBR)	PMA
26, 29	NO_2 (BBR)	PMA
28	CH_4 (GCR)	LBL
30	HCl (GCR)	LBL
31	HF (GCR)	LBL
32	AE_CH_4 (BBR)	PMA
33	AE_HCl (BBR)	PMA
34	AE_HF (BBR)	PMA
35	Aerosol parameters (size, number density, composition) determined from 4 individually retrieved extinction profiles	

t(q) transmission profiles as a function of zenith angle
AE_*name* Aerosol extinction for *name* bandpass
GCR gas correlation measurement retrieval
BBR broadband measurement retrieval
PMA Pseudo-mass approximation
PR Presssure registration
CGA Curtis-Godson approximation
t/p Temperature and pressure
LBL Line-by-line

Table 1.3. Major HALOE data processing steps.

The HALOE broadband models can be toggled between line-by-line, PMA and CGA. The results in Figure 1.6 are typical of the error in ozone retrievals due to the forward model approximation. In general the retrieved mixing ratio error is in the <3% range above the ozone peak and <5% or <0.2 ppm below the peak, whichever is larger. However, 0.2 ppm error can be a substantial percentage error, but usually only occurs in the tropical lower stratosphere. (See Figure 1.7).

In general, the small mixing ratios in the tropical lower stratosphere are very difficult to retrieve with a

high degree of accuracy. The typically steep vertical gradient below a high mixing-ratio peak presents system accuracy requirements difficult to achieve with the current uncertainty in spectral parameters and broadband approximations. This is compounded by the spectral location of the HALOE ozone channel, which was chosen for overall accuracy between tropopause and mesopause. It results in weighting functions for low altitudes that are marginally adequate for tropical tropopause soundings. However, results at mid and high latitudes in clear air are very good down to just above cloud top.

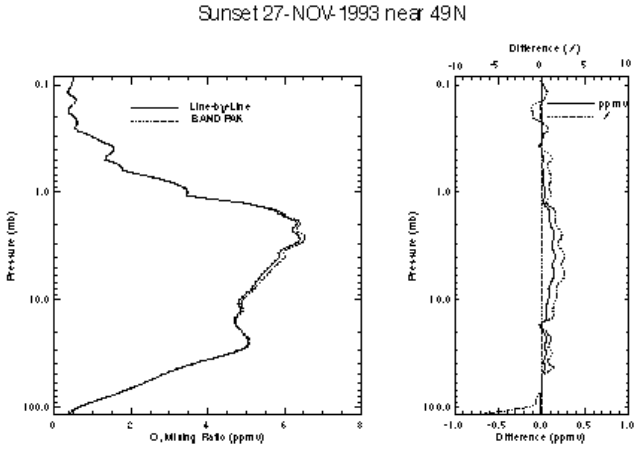


Figure 1.6. Typical mid latitude HALOE ozone retrieval using rigorous line-by-line forward model versus the BANDPAK fast approximation model.

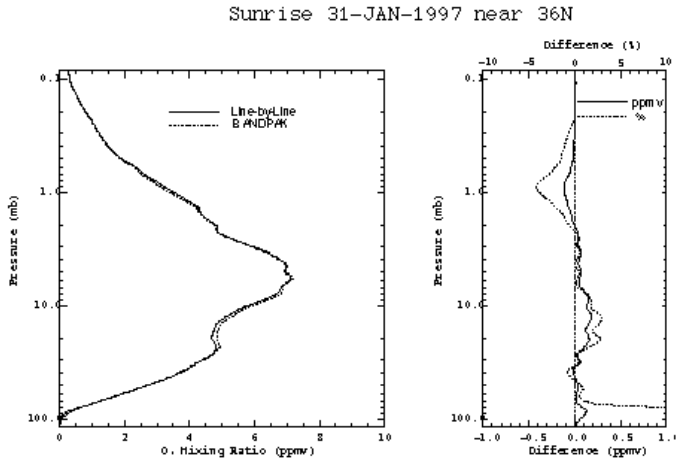


Figure 1.7. Typical low latitude HALOE ozone retrieval using rigorous line-by-line forward model versus the BANDPAK fast approximation model.

The only other significant errors in the forward model are due to spectral parameters. Figure 1.8 shows the effect of a 5% half-width error. The band strength error is now believed by researchers to be about 4% (Flaud, 1990) which results in an equal error in mixing ratio. Under most circumstances a root sum of the squares (rss) of all the forward model error components indicates about a 5% error. At low concentration below large ozone peaks these errors can compound due to weighting function effects

discussed above to give a much larger percentage error. However, repeatability in clear air is excellent at all altitudes and latitudes.

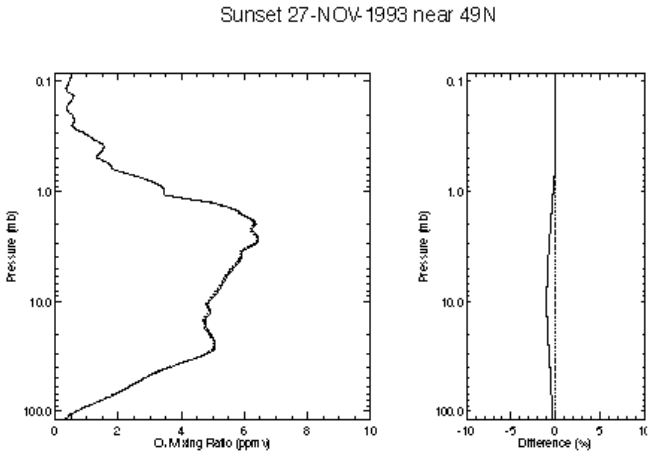


Figure 1.8. Effect of a 5% error in absorption line halfwidth.

The forward model for the CO_2 channel is very accurate (<1% error in simulated extinction) at altitudes above 35 km where the temperature/pressure retrieval is performed. Below 35 km the accuracy degrades and NCEP temperatures are used. The broadband transmission of fundamental vibrational bands has only slight temperature dependence, making the temperature uncertainty a small contribution to the total ozone retrieval error which is dominated by ozone forward model error and pointing error below the ozone peak, and random error and forward ozone model error above the peak.

Inverse Model and Retrieval

The inverse model for the HALOE ozone retrieval is a modified onion peel with no a priori assumption. The only assumption is the mixing ratio at the highest retrieved altitude (>90 km), which is made constant with altitude and equal to the top retrieved value. This is achieved by treating the top of the atmosphere as one extended layer. This top boundary condition has less than 5% effect at 85 km and negligible effect below 80 km. An interleave retrieval strategy is applied which uses all the over sampled data at a resolution consistent with the system resolution, achieving profiles that display features very consistent with that resolution. Each transmission profile, sampled at 0.3 km spacing in apparent tangent altitude, is broken into 7 profiles at 2.1 km spacing. These seven are retrieved with independent 2.1 km layer models, re-interlaced into one 0.3 km profile and smoothed with roughly a 1.5 km smoothing function.

The retrieval process proceeds in an onion peel fashion, alternating between the 7 separately modelled profiles using a modified vector relaxation process that effectively weights (in a maximum likelihood fashion) results from nearby (in altitude) points in the other 6 profiles. Three complete passes are performed to eliminate biases introduced from the first pass, which does not have accurate information from results below the tangent point due to the onion peeling sequence.

The combination of precise solar tracking and an instantaneous field of view of 1.6 km leads to very little error due to resolution limitations, part of what is called null space error (Rodgers 1990). The

averaging kernels for five different altitudes are shown in Figure 1.9. In reality there is one function per 0.3 km altitude grid point. These functions display how the retrieval weights the true mixing ratio profile to obtain a retrieved value at one altitude. Ideally, for an unconstrained onion peel retrieval, these functions would be symmetric triangular functions about the retrieved altitude. However, a slight constraint to retrieved results at adjacent altitudes, coupled with the natural limb viewing contribution functions (Rodgers 1990), gives slightly asymmetric and oscillating averaging kernels. This has a negligible effect on the error budget, which has been verified by simulated retrievals with no constraint.

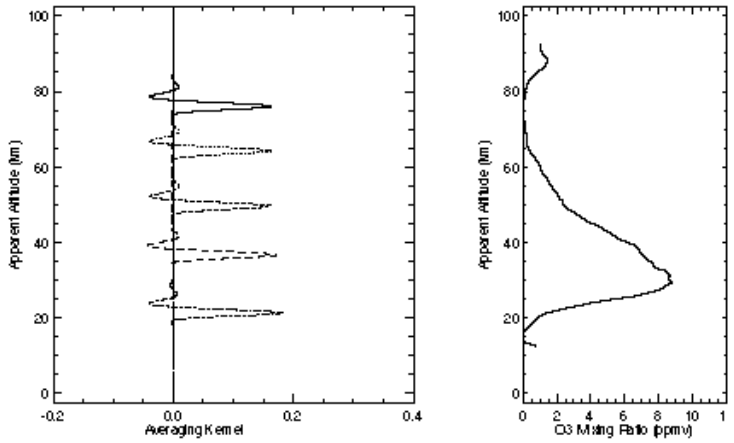


Figure 1.9. Typical HALOE averaging kernels.

1.3.6. Error Mechanisms and Characteristics

Mechanisms causing single event errors are listed in Table 1.4 for a standard mid-latitude profile, along with the order of importance of the five dominant errors for the given altitude range. For a more comprehensive error discussion, consult Bruhl, 1996. As can be seen in the table, the dominant error sources change with altitude. At the lowest altitudes the errors are dominated by pointing and forward model. They can also change with atmospheric state: mixing ratio and temperature profiles, cloud cover, aerosol loading. For clear air and mid-latitude type profiles we list at the bottom of Table 1.4 the current conservative estimates taken from Bruhl, eliminating aerosol and water bias. Water bias is now believed to be insignificant. Aerosol varies with time and must be included only under moderate to extreme conditions. Note that the random error at low altitudes is a small component of the total error.

These traditional single retrieval errors may or may not affect the inference of trends. In trending, random errors become insignificant if large amounts of data are used. Systematic errors such as calibration errors have small effects if stable over time.

Although occultation sensors do not require absolute radiance calibration (only signal ratios are used), as with any spectroscopic technique, stability of spatial and spectral response is essential unless continuously calibrated. In addition the tracking

	Error Mechanism	<18 km	Altitude	25-50 km	>50 km

			18-15 km		
Random	Noise bit and detector noise				1
	Pointing tracker jitter sunspots				2 4
Systematic	Forward model				
	forward model approx.	4	1	2	
	spectral line parameters	5	3	1	
	interfering gas absorbers				
	clouds	2	5		
	aerosol	3	4	3	
	Pointing				
	edge tracker & registration	1	2	5	
	azimuth tracking				3
	sunspots				5
	Response functions				
	spectral			4	
	temporal				
	Typical total error % (mid-lat. profiles)	25	15	9	8-100

Table 1.4. Relative importance of major HALOE error mechanisms for various altitudes (1 being most important).

requirements for occultation sensors, coupled with solar source function characterisation, present unique problems. Finally, the complicated observation pattern, Earth's oblateness and orbital precession can alias seasonal and latitudinal changes, confusing trend estimates if not done carefully.

Below we discuss mechanisms believed to have a potential impact on trends inferred from HALOE data (and in some cases occultation data in general) and methods used to detect, monitor and estimate the magnitude of these mechanisms. The discussion proceeds by error category with Table 1.5 summarising

estimated magnitude.

Aerosol and Cloud Extinction

During the period of large volcanic aerosol loading after the eruption of Pinatubo, the ozone channel transmission at tangent attitudes below 25 km was strongly affected by aerosol extinction. Figure 1.10 shows the effect of aerosol in mid 1992 on the retrieved ozone profile. Figure 1.11 shows the effect in 1997. It was determined, through correlative comparisons that the HALOE processing removes the effects of aerosol with roughly 93% accuracy, or a residual uncertainty equivalent to 7% of the aerosol extinction. This was verified by Steele and Turco (1997). The aerosol error is included in the quality error numbers in the HALOE data set. As can be seen by the two figures, this error can induce a wide range of percentage errors in ozone. From >100% in the lower stratosphere before February 1992 to typically <2% by 1997. Again, the user should check recorded errors in the data set, which are noise plus aerosol effects combined in a root sum squared fashion.

Cloud extinction, which is not spectrally modelled in the current HALOE processing, also can corrupt the broadband signals. Therefore, long-term change in either aerosol or cloud cover could induce trending errors if the affected data are not removed from the analysis. We recommend that trend estimates exclude all data where the aerosol extinction values for the NO channel (5.26 mm) bandpass exceeds 10^{-4} km^{-1}.

Pointing

As described in section 1.3.4, and shown in Table 1.4, pointing error can be a major low altitude contribution to the total error budget. This is primarily an instrument design limitation that becomes significant just above cloud tops. To induce trend errors, this mechanism also requires a long-term change in aerosol or cloud cover. However, the effect of the problem is eliminated by the same data filtering procedure recommended in the previous section.

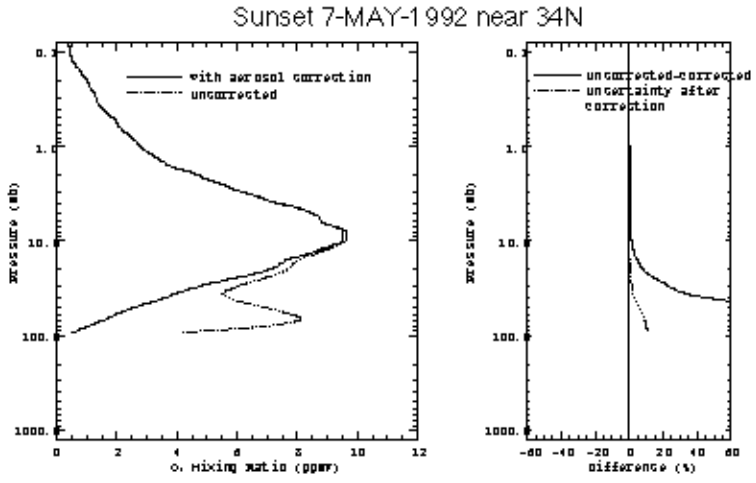


Figure 1.10. Typical HALOE ozone 1992 aerosol correction effects

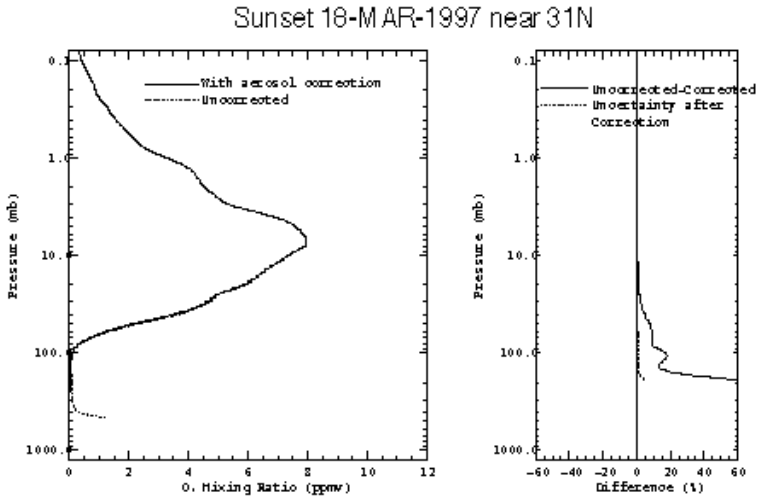


Figure 1.11. Typical HALOE ozone 1997 aerosol correction effects

Spectral Stability

The spectral response of the HALOE instrument is defined primarily by interference filters, with secondary effects coming from dichroic beam splitters and detectors. The stability of these elements has been evaluated using duplicate witness units. The units, now nearly 15 years old, have shown no measurable change in relative or absolute response. The ozone and CO_2 channels use thermistor bolometer detectors which have a long history of reliable use in rugged industrial environments. Duplicate bolometers also show no spectral change.

The spectral filters were manufactured with a lamination process in which small impurities can induce some delamination under extreme (>100 K) thermal cycling. However HALOE experiences less than a 20 K temperature variation in orbit which is not a problem according to the manufacturer. The witness filters were last spectrally scanned in 1996.

Further evidence, from comparisons and solar signal change, and changes of signals from an internal 1000 K blackbody, points to spectral stability.

The only other possible sources of spectral change are the gas cells. However cell content has no direct effect on ozone retrieval. Although the NO gas correlation channel is used to infer aerosol, which is in turn used to correct the ozone channel for aerosol extinction, this procedure is nearly insensitive to gas cell content. Also the cell content for all correlation channels has been proven stable over the mission through an analysis of retrieval correlation with orbital b angle. The overall conclusion is that HALOE has been spectrally stable throughout the mission. Ozone and CO_2 signal changes due to relative spectral response, if they exist, must be well below the 1% level.

Time Response

The time response of the measurement system can change the effective resolution due to the sun sink rate. Time response changes can be observed by analysing scans on and off of the sun. No change in response time constant has been observed. A 1% change in resolution can cause up to a 0.05% change in retrieved ozone. We believe resolution has changed less than 2%.

Optical Degradation

Signal characteristics when scanning on and off the sun can also be used to infer off axis scatter and resolution changes. No detectable change has been observed. In addition the gas correlation difference signals give an extremely sensitive response to field-of-view differences (~ 1 ppm) between effective vacuum path and gas cell path signals. There has been no significant change in these signals. Although this is not directly related to the ozone channel, it lends further confidence that the instrument has been optically stable.

Solar Tracking

The tracking system has also shown stable performance. Since the solar source function changes intensity from centre to solar limb (see Figure 1.12), it is essential that the tracking system provide exact field-of-view location on the solar image. Periodically the HALOE instrument makes horizontal scans of the sun to calibrate the tracking system. This ensures that changes in tracking performance are observed and accurately modelled. There has been no observable change in tracking performance. A possible concern is the azimuth tracking stability. As the sun changes shape at low altitudes during an occultation event the ability to centre the FOV changes. As azimuth sensor optics degrade with time, a slight change in centring stability could change results at low altitudes. The flatness of the solar intensity at 9.6 mm greatly reduces this effect relative to the potential liability at shorter wavelengths (again, see Figure 1.12). However, since the NCEP temperatures are used at low altitudes instead of the CO_2 2.8 mm signals, the HALOE ozone trends should not be susceptible to small changes in azimuth tracking knowledge.

One change that has been observed is the response of the diodes used for high resolution tracking in the vertical. The total response has decreased by about 30%. Since the tracking logic uses only relative responses of diodes to detect the solar edge, the tracking performance has shown no detectable change, this stability is expected to continue until response is down by 90% or more, which at the current rate of decrease would take another 15 years.

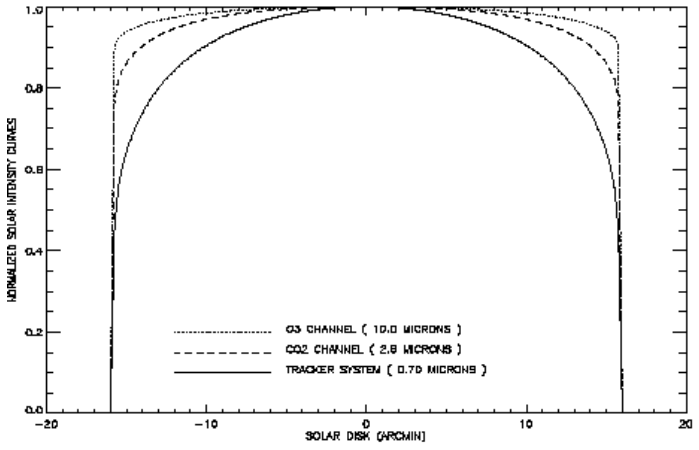


Figure 1.12. Solar intensity curves corresponding to HALOE O_3, CO_2, and tracker system wavelengths.

Orbital Geometry

The UARS orbit precesses with a period of about 70 days. This causes the occultations to occur at various angles (b angles) relative to the orbital plane, inducing changes in solar image sink rate. The eccentricity of the orbit and the Earth's oblateness also combine to cause a time dependent change in the zenith angle versus height relationship during occultation events that occur at large orbital beta angle. These effects are tedious to model (they can cause up to 25% error in mixing ratio if not included) and are most pronounced at specific latitudes (such as southern mid to high) and seasons. The natural variation in pressure height also contributes more error for large b angles. Each year the sequence of orbital beta angle shifts phase by five days and therefore would shift one full precession cycle of 70 days over 14 years. Although great care has been taken to explicitly incorporate these effects in the data analysis, results for very high b angle (>55o) can often differ by up to 3%. This is expected to have a negligible effect on trend estimation. However trends determined with small subsets of data in specific latitude bands can give degraded results, especially if the mean b angle changes systematically with time. Figure 1.5 shows the 1995 sampling and b angle pattern. The five-day shift each year has the potential of aliasing with seasonal and latitudinal variation. Correlative comparisons at a particular site also have the potential to incur a systematic change in space and time coincidence, leading to false comparison trends. In summary, without careful attention to the sampling details, occultation measurement patterns can lead to biases in trends.

Algorithm

The processing algorithm can be a source of faulty trends if time dependent errors occur. For example the HALOE algorithm assumes a 0.47%/year increase in CO_2 mixing ratio. An error in this increase can lead to a nearly equal error in inferred ozone mixing ratios. Another time dependent error that may be quite significant is the aerosol correction. In fact if the data from the first year of the mission is used, the Pinatubo aerosol would likely be the largest contributor to trend error in the lower stratosphere. However, filtering out data through 1992 and other data with large aerosol loading can reduce aerosol-related error to insignificant values. No other mechanisms are known to exist in the processing algorithms that could cause trend errors. The HALOE temperature and pressure retrieval is now independent of auxiliary data above 35 km and therefore not susceptible to errors from those sources. Below 35 km the NCEP temperature is used to build pressure down from the retrieved pressure. The HALOE registration with density at all altitudes is insensitive to NCEP errors, such as height versus pressure relationships.

As discussed above and in section 1.3.4.3, the HALOE instrument has proven very stable over six years of operation. The algorithm uses no a priori other than the upper boundary profile shape, as discussed in section 1.3.5.2. It also uses no time or signal dependent constraint. In addition, high altitude results obtained by retrieving on daily average signal have been found to be nearly identical to results obtained from a daily average retrieval. Considering the extremely noisy character of HALOE retrievals in low S/N at high altitudes, this implies negligible effects due to constraints.

These random characteristics combined with high precision and no requirement for absolute radiance calibration make the data particularly well suited for trending. However the lower stratosphere and upper troposphere trends must be taken from data void of clouds and with little aerosol extinction since the correction for either effect and tracking knowledge can have substantial systematic error components under such conditions.

Table 1.5 lists contributions to trend errors due to instrumental sources only, and does not include errors due to sampling and natural variability. We stress that these are difficult to verify and only best guesses at upper limits. As the mission continues and instrument parameters continue to be monitored, conditions

could change.

Type	<25 km	25-60 km	Comments
Aerosol Extinction	0.15	0.02	Assumes pre-1993 data not used
CO_2 mixing ratio	0.03	0.03	Assumes mixing ratio change over 10 years is known to 0.3% ~ 10 ppm
Solar Tracking	0.05	-	There is no evidence of change, but likely.
Cloud Change	0.03	-	There would have to be a change in cloud cover with time
Spectral Response	0.03	0.03	No evidence
Optical Degradation	0.01	0.01	Assumes 2% FOV widening. No evidence, but likely
RSS	0.07	0.04	Assuming clear air (no aerosol)
	0.16	0.05	Starting with 1993 aerosol

Table 1.5. Sources of ozone trend error in HALOE measurements, %/year.

1.3.7. Outstanding Problems and Future Plans

Most problems affecting HALOE ozone retrievals above the middle stratosphere are believed to have been solved. Little improvement is expected above 25 km. Below 25 km there has been a substantial effort to understand the tracking performance, which is a major contributor to the error budget. Solutions are now understood and will be implemented in the coming year.

A second major obstacle to achieving optimum results into the upper troposphere is the processing algorithm which was designed to address the stratosphere and above. Substantial information for potential low altitude retrieval of water and methane exists in other channels, which is not being used by the current algorithm. An effort to rectify this deficiency will be funded through NASA. This will lead to improved H_2O, CH_4, and cirrus detection, and therefore ozone results. A secondary effort under this same grant will be the improvement of the ozone forward model. The eventual goal, expected within 2 to 3 years, is an improved HALOE ozone, water and methane data set below 20 km into the troposphere.

The current data set can be filtered to remove data corrupted by cloud and pointing deficiencies by using an aerosol extinction threshold. This could result in good statistical information, but would still be substantially inferior to the data set expected from the future retrieval strategy that will optimise information extraction at low altitudes using all channels and measurements.

1.4. Microwave Limb Sounder

1.4.1. Measurement Description

The UARS Microwave Limb Sounder (MLS) measurements are obtained from observations of millimetre-wavelength thermal emission as the instrument field of view is vertically scanned through the atmospheric limb. The MLS instrument is described by Barath *et al*. [1993] (see also Waters, 1993). The instrument has three main subsystems: the antenna, the 3 radiometers (measuring radiances at frequencies near 63, 183, and 205 GHz), and the filter banks. The signal entering the radiometer contains radiances emitted by the limb of the Earth averaged over the antenna field of view (FOV) and each spectral filter, and a small "extraneous radiance" comprised of thermal emission from the instrument (mostly the antenna primary element), scattered radiances from the primary mirror's surface roughness, spill over of the primary past the secondary, and edge diffraction. The thermal component of the extraneous radiance is spectrally flat over the filter bank bandwidth and scan independent, but the scattered and diffracted components may be weakly dependent on scan-angle. Calibration is accomplished by interspersing several space views (at 2.7 K) and calibration target views (at about 300 K) with the (typically 26) limb views during each 65 second scan of the atmospheric limb via a switching mirror located between the antenna and the radiometer(s). The radiances (or brightness temperatures) of the target (inferred from its measured emissivity and temperature) and space, along with pre-launch calibration measurements of the losses along optical paths, are incorporated in a linear model relating limb counts to limb radiances; the radiometric gain is the proportionality constant in this relation (see Jarnot *et al*., 1996). Departures from a linear relationship are estimated to be less than 0.1%.

The MLS instrument uses double-sideband heterodyne radiometers: signals from two sidebands, equally spaced about the local oscillator frequency, are combined (with slightly unequal weights) through the use of whisker diode mixers into a single measurement. The weights, or sideband ratios, are measured for each channel as part of pre-launch instrument calibration. There are fifteen channels across each of the bands obtained in 510 MHz wide filter banks. Channels have similar characteristics on each side of the line centre and are narrower near the centre; signal-to-noise is roughly proportional to the square root of the channel width. The 15 channels dedicated to the 205 GHz ozone line (a line centred at 206.132 GHz, or 1.45 mm wavelength) are referred to as "band 4" . Another ozone line is measured at 184.378 GHz (1.63 mm wavelength), but that radiometer failed in April 1993. The MLS ozone data mentioned in this report refer to the longer-term 205 GHz ozone data. There are other species that emit weakly in the 205 GHz ozone line region, namely SO_2 (with a contribution primarily during the first year after the Mount Pinatubo eruption), HNO_3, and H_2O (from the lowermost stratosphere and upper troposphere). Other contributions include emission from the wing of a N_2O line outside the target band pass, which becomes significant only in the lowermost stratosphere, and very small emission features from HO_2 and a heavy (symmetric $O^{18}OO$ isotope) ozone line (see Froidevaux *et al*., 1996).

1.4.2. Measurement Resolution

The MLS measurements have very high spectral resolution, as alluded to above, with the ozone line being resolved through the use of 15 channels; the centre channel has a width of 2 MHz, and the two broadest wing channels have a width of 128 MHz. The spatial resolution is given by roughly 400 km along the line of sight (perpendicular to the UARS velocity), from radiative transfer considerations; a similar amount of smearing occurs along the tangent track, because of the satellite motion during a 65 s scan (for which an individual ozone profile is retrieved). The vertical resolution is constrained by the

FOV vertical extent at the tangent point (3 km half-power beam-width for the ozone measurements); however, the MLS data (for Version 4 and earlier) are produced on a vertical grid with points spaced each factor of $10^{1/3}$ (or 2.15) change in atmospheric pressure, giving a vertical resolution of ~5.4 km. The limb scan used in normal operations consists of discrete steps between about 90 and 0 km tangent heights, with step spacing for the scan varying between 1 km in the lower stratosphere to 5 km in the mesosphere; individual spectra are measured during 1.8 s dwells between steps.

1.4.3. Sampling

MLS performs routine measurements continuously every day (except when satellite or instrument-related problems occur), both day and night. The measurement latitudinal coverage is from 34° on one side of the equator to 80° on the other. UARS performs a yaw manoeuvre at roughly 36-day intervals (a "UARS month"), when MLS high-latitude coverage switches between north and south. Within each UARS month the UARS orbit plane precesses slowly with respect to the Earth-Sun line. The orbit precession causes the measurements to sweep through essentially all local solar times during the course of a UARS month, becoming 20 min earlier each day at a fixed latitude.

The typical profile spacing along the orbital track is 3 to 4° in latitude, with significantly denser coverage in small bands close to the turn-around points (near 34 or 80° of latitude). The spacing between orbits in longitude is typically about 25°, but since there is overlap arising from ascending and descending portions (for day and night side, typically), the average daily sampling in longitude (for a latitude bin a few degrees wide) is about 12°.

After 2.3 years in orbit (in late December 1993), the antenna-scanning mechanism began to exhibit signs of wear. March 1994 through May 1994, and July 1994 were periods of testing and significantly reduced data gathering; these months have from one third of the days with bad data to almost all bad days (days with no profiles retrieved). August and September of 1994 contain mostly good data, but the months of October 1994 through January 1995 again have very few days of useful atmospheric profile data. Reverse scanning and other modifications to the operations (including short periods of "mechanism rest" every orbit) have been implemented since February 1, 1995. Very little limb data gathering (typically only a few days per month) occurred from February through July 1995, which was a period during which instrument power sharing began for UARS (in May 1995) because of poor solar array performance (the solar array was then parked indefinitely without subsequent rotation).

Since June 1995, MLS has been in a mode of operation characterised by off periods for power savings and on periods during which typically 2 days of full (reverse) scans are obtained followed by one day of limb tracking at altitudes near 18 km. The limb tracking days do not lead to standard catalogued profile data files, although there is some information on the atmosphere from those days. The August 1995 through September 1996 time period generally contains about one half to one third catalogued good days for atmospheric profiles, and this mode of operation is expected to continue. MLS has been operating without "scan slips" since January 1995. However, instrument operations were significantly reduced after June 1997, when one of the 3 UARS batteries failed, further reducing power availability.

1.4.4. Instrument Error Sources

Instrumental sources of error arise from radiance noise (random errors; see Jarnot *et al.*, 1996), and systematic sources, namely possible uncertainties (and possible changes with time) in the antenna FOV

and antenna surface properties, the channel spectral responses (filter shapes), the radiometric gain, the channel sideband ratios, and the extraneous radiance (scattering and emission). Non-instrumental errors can also arise from the forward and inverse numerical models (discussed later). Further discussion of most of the above can be found in Froidevaux *et al*. [1996] and in subsequent parts of this report (e.g. error characterisation section).

1.4.5. Algorithm Description

Forward Model

The forward model used to describe the calibrated radiance signals received by MLS is a first order Taylor series of a local thermodynamic equilibrium (LTE) radiative transfer calculation (with no scattering) averaged over the spatial and spectral responses of the instrument (details are given by Read *et al*., manuscript in preparation, 1997). The independent variable or state vector is a mixing ratio profile gridded every 1/3 in $\log_{10}$ pressure, with the grid points connected by linear segments. The linearisation values assume a climatological mean grouped into eight 20o latitude bins and 10 periods (corresponding to UARS months) per year. The radiative transfer calculation uses spectral filter shapes and a field of view gain pattern collapsed into the elevation dimension determined during pre-launch calibration. The one-dimensional field of view averaging neglects radiance azimuthal variation which would be created by the curvature of the Earth and along track variations in the constituents. Line-of-sight gradients are ignored in the forward (and inverse) model representations so far.

The absorption functions used in the calculation employ a line-by-line plus a background continuum contribution for each species. The total absorption is arrived at by summing all individual species contributions. All "important" species are included in addition to the O_3 emission near 205 GHz, namely N_2O, H_2O, $H_2{}^{18}O$, HNO_3, O_3 v-2, $O^{18}OO$, SO_2, and N_2/O_2 continuum. A Van-Vleck Weiskopf line shape with no interline interference effects (except for O_2 which includes it) is used for each line and the total contribution is obtained from a sum over contributions from all lines and a continuum.

The 205 GHz O_3 radiance is quite linear in concentration and a pre-frequency averaging approximation is used where the absorption coefficient is averaged over the spectral response before performing radiative transfer calculations. This approximation is not used for the stronger and more non-linear 183 GHz O_3 measurement. The derivatives in the Taylor series are computed analytically, fully incorporating the representation basis function shape, which is considerably faster than using finite differences. Given that the linearised forward model is not as accurate as the complete radiative transfer model, the radiance uncertainties have been increased by an "inflator factor" (0.5 % of the radiance for the 205 GHz ozone band).This includes possible errors caused by non-linearity, finiteness of the vertical grid used (1 km grid), and frequency averaging methodology. It does not attempt to include possible errors arising from other factors mentioned above.

Inverse Model

The MLS approach used for retrieving geophysical parameters, such as tangent pressure, temperature, and mixing ratios, from the calibrated radiances (Level 1 data), is based on sequential estimation, as first applied to limb sounding by Rodgers [1976]. This technique combines a priori information (such as climatological profiles with an associated error covariance) with measurement information, to produce a solution for the desired parameters and error covariances. The a priori "state vector" is used to stabilise the solution in regions of poor sensitivity and to provide some continuity. This methodology can be

viewed as a combination of the a priori information and the weighted least squares solution to the linearised equation relating measurements and state vector (see Appendix of Froidevaux *et al*., 1996, for a simple derivation). The calculated radiances and their derivatives are obtained by table look-up (for speed of evaluation) from climatological conditions and then interpolated in the vertical to the tangent pressure. Radiances are interpolated with respect to (logarithm of) tangent pressure using cubic splines and with respect to other variables via the linearised form of the forward model (see above).

In addition to the radiance error "inflator factor" mentioned above in relation to (some of the) forward model uncertainties, a radiance error term is added to account for error contributions from variables which are constrained to certain values (based on a priori or retrievals from other bands). For example, temperature and tangent pressure uncertainties, after their evaluation from band 1 retrievals, have their uncertainties propagated to mixing ratio estimates in this way; also, uncertainties arising from velocities (Doppler shifts) along the FOV direction are treated in this fashion, with no values retrieved, but with the (a priori) uncertainty estimates included in the equations for mixing ratio updates. In terms of the estimated error covariance (whose diagonal elements give the uncertainties in the MLS data files), the retrieval software produces uncertainties which are flagged as negative under conditions of poor measurement sensitivity (see details in Froidevaux *et al*., 1996).

Another important feature of the MLS retrievals used in the production of Version 4 data is the use of an "opacity criterion" which eliminates radiances corresponding to optical depths greater than unity along the line of sight, because of non-linearity concerns (given the non-iterative linear scheme used so far). The UARS MLS Level 3 data are output on the UARS pressure grid ($1000x10^{-n/6}$ hPa values), which is twice as fine as the Version 4 MLS retrieval (Level 2) grid; adjacent Level 2 grid values are averaged to fill in the Level 3 grid on the "odd" UARS surfaces. Finally, it was found from simulations that better retrievals were obtained if radiances arising from tangent pressures larger than 100 hPa were not included, at least for the non-iterative retrievals used to produce Version 4 (and prior) MLS data.

The a priori values used for ozone are based on the ozone climatology from Keating *et al*. [1989a,b], with extrapolations for the poorly represented lower stratosphere (at pressures over 20 hPa). Other details are given in Froidevaux *et al*. [1996]. A priori uncertainties are generally chosen conservatively (i.e., large values are used) in order to minimise possible biases. For ozone a 3 ppmv a priori uncertainty is used in the stratosphere down to 50 hPa, with a 2 ppmv uncertainty at 100 hPa, and significantly tighter constraints in the troposphere (where not enough measurement sensitivity exists). A priori uncertainties decrease smoothly with decreasing pressure in the mesosphere, until about 0.2 hPa, above which an uncertainty of 1 ppmv is assumed. The use of a "global" a priori uncertainty is clearly an oversimplification of the variability of the real atmosphere. The possibility of larger-than-appropriate estimated uncertainties in the retrieved profiles exists as a result of this approach, for regions like the lowermost stratosphere where poorer measurement sensitivity exists; however, this is a safer approach than the use of a priori uncertainties which are too small, since this can lead to undesirable biases with unrealistically small estimated uncertainties. Based on sensitivity tests of the influence of a priori values, it was confirmed that the 205 GHz ozone retrievals are affected in a non-negligible way by the a priori for pressures near 100 hPa and above about 0.46 hPa. The range of recommended pressures for Version 4 (and prior) MLS data has been 46 to 0.46 hPa, with most improvements needed in the lowermost stratosphere (100 hPa in particular), where measurement sensitivity and accuracy are poorer.

Averaging kernels (see the description of error characterisation by Rodgers, 1990; also Marks and Rodgers, 1993) for MLS ozone are well-peaked functions with some degradation in measurement sensitivity occurring in the lower mesosphere and at 100 hPa for the 205 GHz ozone data (see Figure 1.13). Most of the smoothing arising from the MLS measurements is caused by the use of the fairly coarse Level 2 retrieval grid (not by the instrument FOV or radiative transfer smearing).

Click here to download Figure 1.13 by anonymous ftp.

Figure 1.13. Averaging kernels for MLS ozone 205 GHz data (Froidevaux *et al.*, 1996).

MLS Version 4 Data

A few changes were made between the Version 3 ozone data described by Froidevaux *et al*. [1996] and the Version 4 data used in this report. A post-launch estimate of the pointing offset between the 63 and 205 GHz radiometers is now used (this was obtained from MLS antenna scans across the moon, see Jarnot *et al*., 1996); the pointing offset change from Version 3 corresponds to a tangent height change of about 430 m (with larger tangent heights in V4, relative to the 63 GHz FOV). Also, HNO_3 is retrieved in Version 4 from band 4 radiances (in addition to ozone), with some impact on ozone in the lowermost stratosphere. Radiances for tangent pressures of 100 hPa are included in Version 4 by spline interpolation of the radiances to minimise vertical sampling gaps sometimes present in Version 3 (but a criterion to neglect optically thick radiances is still in effect in Version 4). Ozone values are also affected by changes which occurred in band 1 (for temperature and - mostly - tangent pressure, from O_2 spectroscopy changes - see Fishbein *et al*., 1996).

Average differences between V4 and V3 MLS ozone data are displayed in Figure 1.14, for the late 1991 to end of 1996 time period. V4 ozone values are typically a few percent lower than V3 data, in the 22 to 0.46 hPa range, with variations in these differences across month or latitude varying only at the 1 to 2 % level. At 46 hPa, V4 ozone values are typically systematically lower than in Version 3 (by about 15 to 25%), with largest V4-V3 differences in the tropics and smallest (< 10%) differences at high latitudes. At 100 hPa, V4 values are generally larger than V3 values, especially at low latitudes, where the increase is often large (between 0.5 and 1 ppmv), leading to significant changes in latitudinal gradients. Column ozone above 100 hPa is typically 5-10 % smaller in V4 data than in V3 data.

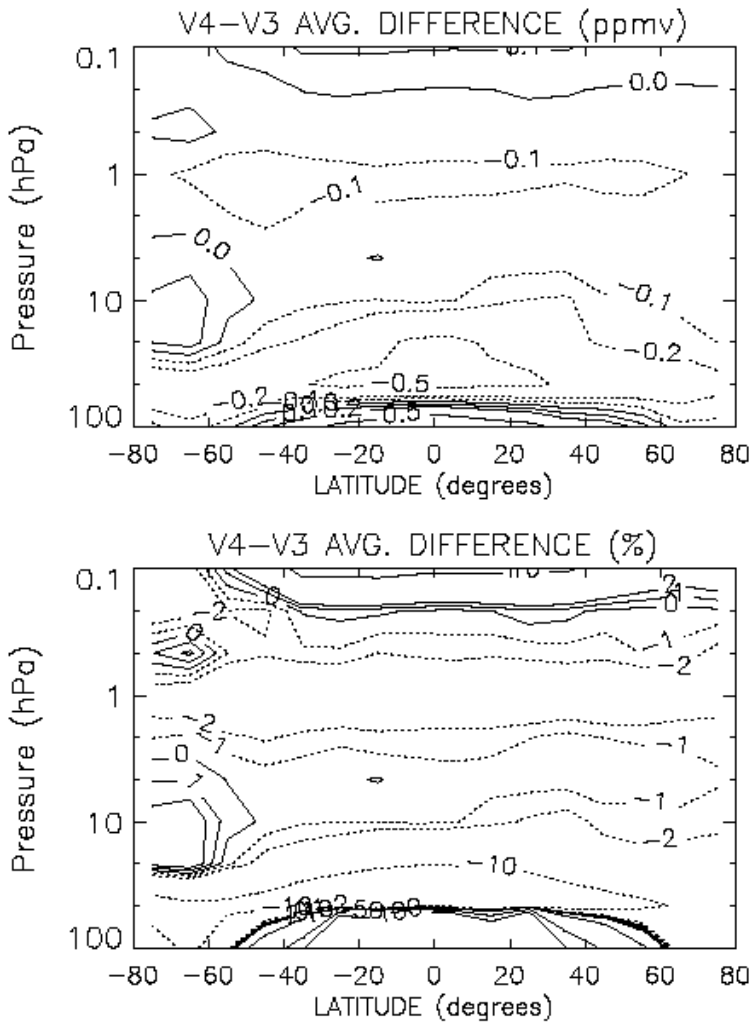


Figure 1.14. Average difference for V4-V3 MLS data versus pressure and latitude, for the time period from late September 1991 to the end of 1996. Top panel shows differences in ppmv, bottom panel in percent.

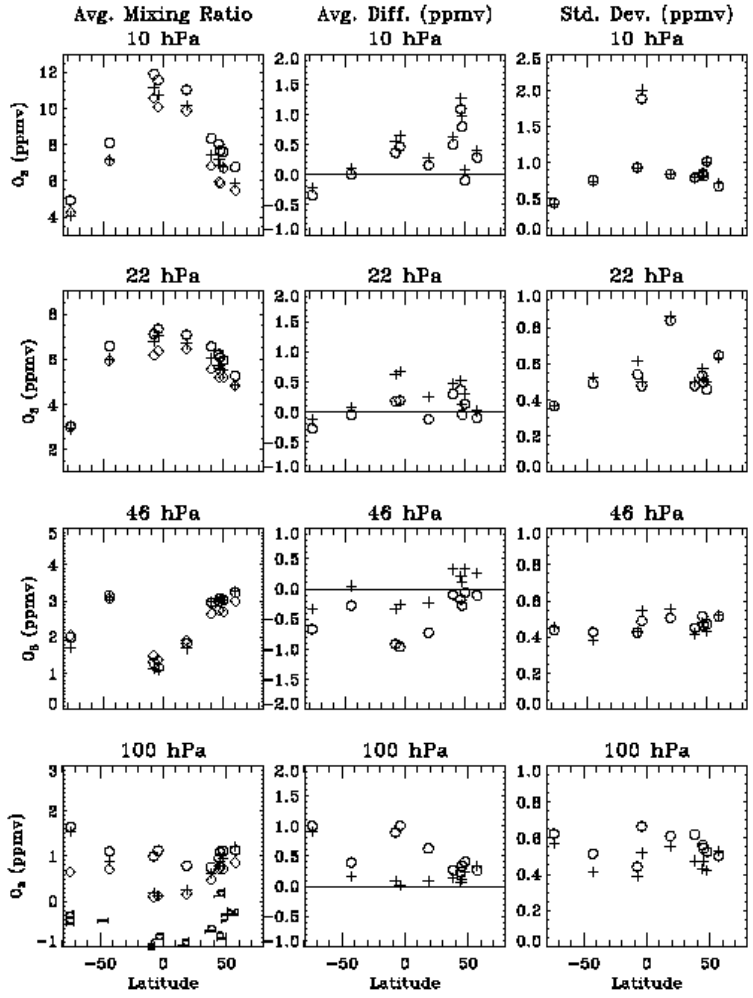


Figure 1.15. MLS data compared to ozonesonde data at various latitudes and pressures. Left panels give average sonde values (diamonds) and MLS version 3 (plus signs) and version 4 (circles) average values. The centre panels give average differences (MLS-sonde), with v3-sonde as plus signs and v4-sonde as circles; the right panels show the standard deviations of the differences (using the same symbols). The site locations from South to North are: McMurdo (78S), Lauder (45S), Ascension Island (8S), Brazzaville (4S), Hilo (19N), Boulder (40N), Hohenpeissenberg (45N), Payerne (46N), Uccle (50N), and Gardermoen (60N). Time periods covered for each case and number of coincident profiles are as follows, in the same order of sites: Aug. 92-Oct. 92 (17 profiles), Oct. 91-Apr. 96 (96 profiles), Oct. 91-Oct. 92 (27 profiles), Oct. 91-Aug. 92 (28 profiles), 1992 data (29 profiles), 1992 data (42 profiles), Oct.91-Apr. 96 (176 profiles), Oct. 91-Dec. 94 (127 profiles), Oct. 91-Apr. 96 (208 profiles), and Feb. 93-Feb. 94 (30 profiles).

There is evidence that the lowermost stratospheric data (from 46 to 100 hPa) at low latitudes in Version 4 are not as reliable (compared to correlative data) as the Version 3 data, although negative biases at 100 hPa existed in Version 3 data. Version 4 values at 100 hPa are sometimes unrealistically large, and the 46 hPa values can be too small (in a compensating way). This is illustrated in Figure 1.15, where ozonesonde data are compared to MLS V4 and V3 retrievals. The main problem occurs at low latitudes in the lowermost stratosphere, with 100 hPa V4 MLS values being significantly too large, and V4 values being too small; the next version of MLS data seems to rectify this kind of bias. At other latitudes, and at other pressures for most latitudes, there are indications that V4 data are slightly improved (versus sonde as well as versus SAGE II data) over V3 data.

To illustrate how Versions 3 and 4 data track each other over the 1991-1996 time period, slope values from linear regression fits of the V4-V3 difference time series (averaged in 10° latitude bins) are displayed in Figure 1.16. Most of the differences in the rates of change are not statistically significant at the 95% confidence level. There is excellent tracking between the two data sets in the 22 to 1 hPa region (typically within a few tenths of a %/year). The largest differences occur where the mean biases between the two versions are largest (and where poorer measurement sensitivity exists), namely in the lowermost stratosphere, primarily at 100 hPa in the tropics.

Given that there are still some bias issues with MLS V4 ozone data in the lowermost stratosphere, particularly between 30°S and 30°N, and that this is also the region where the rates of ozone change have changed the most and are more uncertain (in ppmv/year as well as in %/year), less weight should be given to this data set in that region.

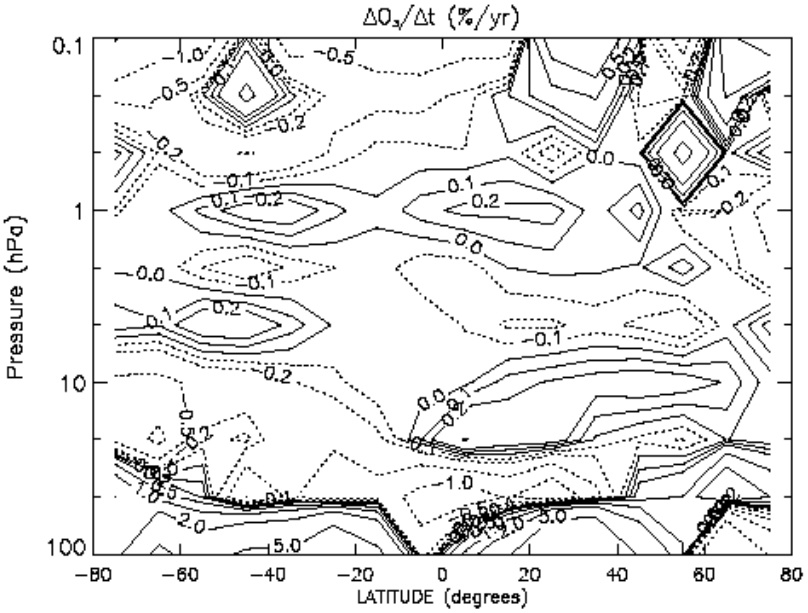


Figure 1.16. Slope (percent/year) of regression line for the MLS V4-V3 difference time series as a function of pressure and latitude (data were binned in 10° bins).

1.4.6. Error and Retrieval Characterisation

The error sources which can impact MLS ozone retrievals can arise for the following reasons: (1) radiance noise (random error); (2) radiance scaling error (includes radiometric, sideband ratio, and line strength errors); (3) error in filter shapes; (4) error in FOV direction (relative to the 63 GHz radiometer FOV used for tangent-point pressure retrievals); (5) error in FOV shape (includes spectral dependence); (6) error in ozone line width; (7) errors from retrieval numerics; (8) error contribution from a priori; (9) error in P_{tan}; (10) error in atmospheric temperature; (11) error in dry air continuum; (12) error in knowledge of H_2O; (13) error in knowledge of HNO_3; (14) error in knowledge of N_2O; (15) error in knowledge of SO_2. The above error sources (except radiance noise) lead to mostly systematic errors, and the MLS ozone sensitivity to these factors has been described by Froidevaux *et al.* [1996]. The main sources of systematic error (especially in the lower stratosphere) arise from possible

errors in tangent pressure, temperature, and the modelled dry air continuum. Precision and accuracy values are plotted and tabulated in Froidevaux *et al*. [1996], based on a consideration of the above error sources, and consistent with conclusions arrived at from inter-comparisons with other data sets; the pressure-dependent (total) accuracy and single-profile precision is given here in Table 1.6.

Pressure	Precision		Accuracy	
hPa	ppmv	%	ppmv	%
0.46	0.37	20	0.3	15
1.0	0.31	10	0.3	10
2.2	0.23	4	0.3	5
4.6	0.20	2	0.3	5
10	0.18	2	0.3	5
22	0.16	3	0.4	5
46	0.22	11	0.4	20
100	0.55	>50		>50

Table 1.6. Estimated Precision and Accuracy of MLS 205 GHz Ozone Profiles.

In terms of potential sources of error for trend analyses of MLS data, time series of engineering data (such as shown in Jarnot *et al*., 1996) suggest that changes in instrument performance (radiometric scaling, spectral response, and extraneous radiance) are significantly less than 1 percent over the 1991 to 1997 time period. A few changes in MLS operations over the course of the mission, such as modifications to the instrument control program or power interruptions (see Fishbein *et al*., 1996), might have had a small impact on instrument behaviour, although nothing is obvious in the retrieved profiles. Since the retrievals take out an offset between modelled and measured radiances, it is improbable that small trends or variations in extraneous radiance (possibly caused by spacecraft surface degradation, as viewed by the antenna spill over lobes) would have a significant impact on the ozone data. Radiometric gains have remained extremely stable over the mission lifetime, but changes in these are removed as part of the calibration. The blackbody target temperature is monitored to calculate target radiance, but the target emissivity could change (in theory); this would have minimal impact because of the compensating effect of the radiation it receives and reflects from the surrounding structure at a similar temperature. A change in antenna surface properties, although unlikely based on past experiences of such aluminium surfaces in space (at microwave wavelengths), would translate into an error in limb radiance (and retrieved ozone). Time series of calibrated radiances from the highest altitudes place limits on such changes; based on such time series, we estimate that the antenna system reflectivity (with a value of 0.989) has changed by less than 0.0001 over 6 years.

This places a bound on possible stratospheric ozone changes in relation to MLS antenna degradation of much less than 1 percent over this time period. Changes in the FOV with time would also be difficult to detect (not easily monitored with high accuracy); however, since the FOV is determined by bulk geometric properties of the reflectors, optics and feed horn, catastrophic events are far more likely sources of change than slow degradation (and they would have been noticed). A frequency shift or drift (in terms of the measurements of the ozone line) could also in theory lead to a spurious ozone trend; this

could be investigated by looking at changes in the characteristics of radiance residuals (or c^2 values) versus time, but this has not yet been done. An unknown change in receiver spectral sideband response would also impact the retrieved ozone, but MLS engineering data on the mixer bias levels provide no indication of any such changes. Errors as a function of time in retrieved temperatures could affect the MLS ozone trends at a small level. Stratospheric aerosol effects at the MLS wavelengths are expected to be negligible (and are not observed in the radiances during the Mount Pinatubo loading of the lower stratosphere); Cunnold *et al*. (1996) have used comparisons between MLS and SAGE II ozone to correct SAGE II data for aerosol effects. Our best current estimates of known possible sources of error and their impact on MLS-derived trends is tabulated below in likely order of importance.

Error Source	Max. Potential Impact on Ozone
Temperature trend error (assume 0.2K/year max.)	<0.2%/year
Antenna degradation	<0.1%/year
Changes in operation mode	Believed negligible
Spectral FOV changes (unknown)	Believed negligible
FOV Spectral changes (unknown)	Believed negligible

Table 1.7. Error sources (in likely order of importance) that could affect MLS-derived trends.

In summary, there are no known mechanisms for significant degradation, and there are indications that excellent stability has been achieved. However, given that it is very difficult (for any instrument) to monitor all possible error sources that might in theory lead to a spurious trend, we should also rely on comparisons between MLS data and other data sources that are believed to be reliable. Based on the comparisons of MLS data versus SAGE II and HALOE data (see Chapter 2), stability at the level of better than 0.1 or 0.2 %/year appears to be achievable (at least above 25 km); this also seems to be borne out by observed stability of MLS temperatures in the tropics, and by comparisons of MLS temperature retrievals versus NCEP data (not shown here). In the lowermost stratosphere however, the error bars are large enough that such stability is not demonstrable in current MLS ozone retrievals; this is a retrieval sensitivity rather than an instrument calibration issue.

1.4.7. Outstanding Problems and Future Plans

There are still some ozone biases in the lowermost stratosphere (see also discussion elsewhere in this report), with MLS Version 4 ozone values at 46 hPa tending to be too small and values at 100 hPa too large, particularly in the tropics. The 100 hPa retrieval problem stems from the small ozone signal and spectral signature (of order 0.3 K/ppmv) at this tangent level and the difficulties involved with modelling continuum and instrumental spectral effects arising at the 0.2 K level; any non-random uncertainty in the calculated radiances, especially in channels furthest away from line centre, can be misinterpreted partly as an ozone signature and will lead to a bias. Relative, rather than absolute radiances are used in the retrievals, since a baseline offset (systematic offset between calculated and observed radiances) is

retrieved as part of the retrievals (see Froidevaux *et al.*, 1996). However, this works best when the wing channels are 'free' of ozone information, such as is the case in the mid- to upper stratosphere. For better separation of these effects down to the lowermost stratospheric levels, increased spectral bandwidth is necessary (and is planned for EOS MLS observations); UARS MLS was not originally designed for focusing on the lowermost stratosphere. The primary issues are nevertheless improvements in the lowermost stratosphere, particularly at 100 hPa in the tropics. The anti-correlation that exists between retrieval levels will be analysed further as well (for example, a combination of adjacent retrieval surfaces might be a better comparison product at some heights).

Future plans (Version 5 data) showing promise for improvements include an iterative retrieval scheme and some improvements in the dry air continuum characterisation, along with retrievals on every UARS pressure surface (rather than on every other one). A more rigorous (non-diagonal) treatment of error covariances from constrained quantities is also planned for Version 5, although this affects the estimated errors but not so much the retrieved ozone values.

1.5. SBUV and SBUV/2

1.5.1. Basic Measurement Description

The Solar Backscatter Ultraviolet instruments, SBUV on Nimbus 7 and SBUV/2 on NOAA-11, are nadir-viewing instruments that infer total column ozone and the ozone vertical profile by measuring sunlight scattered from the atmosphere in the middle ultraviolet. The SBUV instrument was first described by Heath *et al*. [1975]. The follow-on SBUV/2 instruments flown on the NOAA series of spacecraft was described by Frederick *et al*. [1986], and Hilsenrath *et al*. [1995].

The instruments are all of similar design: nadir viewing double grating monochromators of the Ebert-Fastie type. The instruments step through 12 wavelengths in sequence over 32 seconds, while viewing the Earth in the fixed nadir direction with an instantaneous field of view (IFOV) on the ground of approximately 200 by 200 km. To account for the change in the scene-reflectivity due to the motion of the satellite during the course of a scan, a separate co-aligned filter photometer (centred at 343 nm on SBUV; 380 nm on SBUV/2) makes 12 measurements concurrent with the 12 monochromator measurements. The eight shortest monochromator wavelength channels are sensitive to ozone profiles from about 20 hPa to about 1.0 hPa, while the four longest channels detect radiances that reach the Earth's surface and thus are sensitive to ozone column amount.

The instruments are flown in polar orbits to obtain global coverage. Since the SBUV ozone measurements rely on backscattered solar radiation, data are only taken on the day side of each orbit. There are about 14 orbits per day with 26o orbit spacing at the equator. Unfortunately, the NOAA polar orbiting satellites are not exactly sun-synchronous. For NOAA-11 equator crossing times drifted from 1:30 pm (measurements at 30° solar zenith angle at the equator) at the beginning of the data record to 5:00 pm at the end to the data record (measurements at 70° solar zenith angle). As the orbit drifts the terminator moves to lower latitudes and coverage decreases, which can be a problem for trend determination. NOAA-11 SBUV/2 geographic coverage over the data record is shown in Figure 1.17.

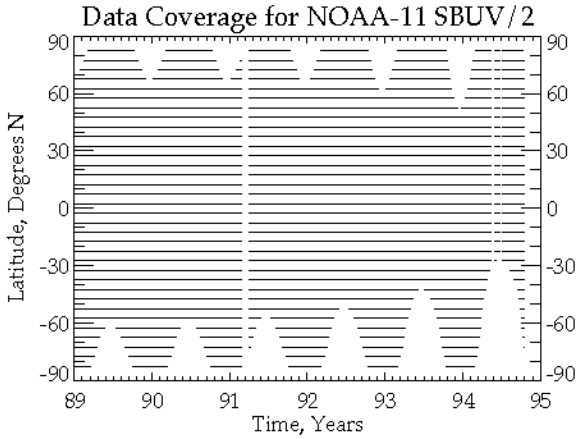


Figure 1.17. The geographic coverage for the NOAA-11 SBUV2 instrument, showing the decreasing coverage at high southern latitudes caused by orbit drift.

Ozone profiles and total column amounts are derived from the ratio of the observed backscattered spectral radiance to the incoming solar spectral irradiance, which is referred to as the backscattered

albedo. The only difference in the radiance and irradiance observations is the instrument diffuser used to make the solar irradiance measurement; the remaining optical components are identical. Therefore, a change in the diffuser reflectivity will result in an apparent trend in ozone. An important improvement of SBUV/2 over the SBUV was the ability to measure changes in solar diffuser reflectivity in flight, using a mercury lamp as a light source (this lamp also provides wavelength calibration). This system did not work for the NOAA-9 SBUV/2 instrument but was redesigned and worked satisfactorily for the NOAA-11 instrument [Hilsenrath *et al.*, 1995].

Spectral Resolution

The spectral resolution for SBUV (and TOMS) monochromators is 1.1 nm, FWHM. The bandwidth of the photometer is about 3 nm. The wavelength channels used for ozone retrievals for Nimbus 7 SBUV were: 256, 273, 283, 288, 292, 298, 302, 306, 312, 318, 331, and 340 nm. The wavelengths for NOAA-11 SBUV/2 were very similar except that Channel 1 was moved from 255.7 nm to 252.0 nm in order to avoid emission in the (04) nitric oxide gamma band that contaminated the SBUV Channel 1 measurement. Channel 1 data were not used in SBUV profile data.

Spatial Resolution

From an 800 km orbit, the instruments' (both Nimbus 7 and NOAA-11) 11.3ox11.3o instantaneous field of view (IFOV) results in an Earth scene size of about 200 km square. As the instrument IFOV moves along the satellite track, at roughly 6 km/sec, the Earth scene moves about 200 km during each 32 second spectral scan. Vertical resolution is determined by the radiative and scattering properties of the atmosphere for a nadir viewing instrument and is roughly 8 km in the upper stratosphere [Bhartia *et al.*, 1996] Vertical resolution is discussed in more detail in the SBUV algorithm section.

1.5.2. Instrument Calibration

Errors can be categorised as instrument calibration errors and algorithm errors. Pre-launch calibration errors are mostly systematic for SBUV measurements. The post-launch calibration can drift over time and directly affect derived ozone trends. For the most part, these can be corrected using on-board calibration systems. Algorithm errors and their dependence over time, such as those due to changing solar zenith angles and orbital drift can be quantified and are discussed in Section 1.5.3.

Pre-launch calibrations

The pre-launch calibrations and instrument characteristics for SBUV/2 have been described by Frederick *et al.* [1986] and Fegley *et al.* [1991]. SBUV and SBUV/2 employ nearly identical standards, traceable to the United States National Institute of Standards and Technology (NIST), and similar procedures for pre-launch calibrations. NIST secondary standards are used for irradiance calibration. Radiance calibration is derived by observing light from the irradiance standard reflected off a diffuse target whose bi-directional reflectance distribution function (BRDF) is accurately known. The lamp irradiance cancels in the albedo calibration. Extensive experiments and analysis have demonstrated that radiometric calibration are precise to better than 1% (2 sigma). Absolute diffuser BRDF values are known to about 3%. Non-linearities, wavelength errors, and other instrument factors have uncertainties that are less than 1%.

Recently, an integrating sphere has been successfully employed to calibrate the radiance sensitivity of BUV instruments (Heath et *al.,* 1993). The integrating sphere has now been used to calibrate the SBUV/2, SSBUV, TOMS, and GOME instruments. These instruments have also been calibrated using

diffuser targets whose BRDFs have been accurately measured. Comparison of integrating sphere and diffuser calibrations has shown agreement of about 1% for SSBUV, SBUV/2, and GOME. For TOMS the agreement ranged from 1-3%.

Out-of-band errors, or response to stray light, is measured pre-flight by pointing the instrument at the sun and comparing the radiance at 200 nm with that at 400 nm. Any relative signals larger than 1 ppm are considered stray light. Some SBUV/2 instruments have exceeded this threshold in the laboratory. For NOAA-11, correlations appear between the shortest profiling wavelengths and surface reflectivity, suggesting stray light problems. This contamination could result in a maximum error of as much as 10% for ozone at 1 hPa under certain lighting conditions.

Post launch calibration: Nimbus 7 SBUV

Long-term SBUV instrument corrections have been discussed elsewhere [Cebula *et al.*, 1988]. Corrections to the data were applied by using a generalised Langley plot method [Bhartia *et al.*, 1995], to further calibrate ozone profiles over the entire version 6 Nimbus 7 data set through 1990. In early 1987 the instrument began to lose synchronisation between its chopper wheel and its electronics, introducing apparent "noise" into the individual measurements. A correction procedure was developed for this non-sync error and the data reprocessed [Gleason *et al.*, 1995]. These data do not appear to be biased but are generally noisier and should be considered inferior to the data preceding the onset of the non-sync errors.

Post launch calibration: NOAA-11 SBUV/2

The NOAA-11 SBUV/2 instrument response was monitored and corrected using on-board calibration systems from late 1988 until October 1994 when the diffuser mechanism failed, terminating the use of the on-board calibration system. Hilsenrath *et al.* [1995] have shown that measured albedo relative to the "day 1" albedo over time is a function of true change in Earth radiance, diffuser degradation, and inter-range ratio (PMT gain). In order to accurately account for solar irradiances changes due to changes in viewing angle on the diffuser plate, the plate's angle-dependent reflectivity must be accurately determined over the life of the instrument.

Photomultiplier Inter-range Ratios

SBUV/2 instruments use three overlapping gain ranges to measure the solar and Earth radiation which varies over six orders of magnitude. Gain range 1 and 2 originate from the photomultiplier anode and can vary with time, while gain range 3 originates from the cathode and should be very stable. This design differs from SBUV where the anode was the source of all output ranges. For SBUV/2, measurements from the Earth and sun are output on different electronic gain ranges at shorter wavelengths and at high solar zenith angles. Therefore the instrument's electronic gain is not always cancelled out in the albedo ratio. The photomultiplier gain must be monitored and carefully evaluated over time for accurate calculations of long-term trends. Change in inter-range ratio (IRR), determined by the ratio of counts in two gain ranges when both ranges have valid data, as a function of wavelength The magnitude of the IRR decreased by about 15% over 5 years. For wavelengths shorter than 297 nm, the IRR's spectral dependence is based on pre-launch data and then extrapolated over time using the temporal dependence at the longer wavelengths. The spectral dependence was assumed to be constant over time. At solar zenith angles greater than 75°E, radiances at wavelengths shorter than 312 nm are on a range different than the corresponding solar irradiance measurement and therefore requires an IRR value for the albedo.

Diffuser Plate Calibration

Relative diffuser reflectivity over time, the most important component of the post launch calibration process for accurate ozone trend determination, is monitored with an on-board calibration lamp system.

Diffuser degradation is monitored at six mercury lines spanning 185 to 405 nm. Figure 1.18 illustrates relative diffuser reflectivity at 254 nm over time in the upper panel. The lower panel illustrates diffuser degradation rate (%/year) for all six spectral lines. The 2 s limits on the fit are shown by the dashed lines. The largest degradation rate for the ozone channels is about 1.5%/year at 252 nm with an uncertainty of 0.2%.

Wavelength calibration is also monitored with the mercury spectral lamp. Absolute wavelength accuracy was determined to about 0.02 nm. Wavelength calibration drift was determined to be less than 0.01 nm over the five years of operation and has no observable effect on the derived ozone values. Beginning in late 1992 the grating drive began to occasionally fail to reach the correct grating position (1 position = 0.07 nm) in 5 channels. A three part correction was implemented to account for these events: 1) ozone cross section, 2) solar irradiance, and 3) multiple scattering and reflectivity. After correction, the effect of grating position errors is much less than 1%.

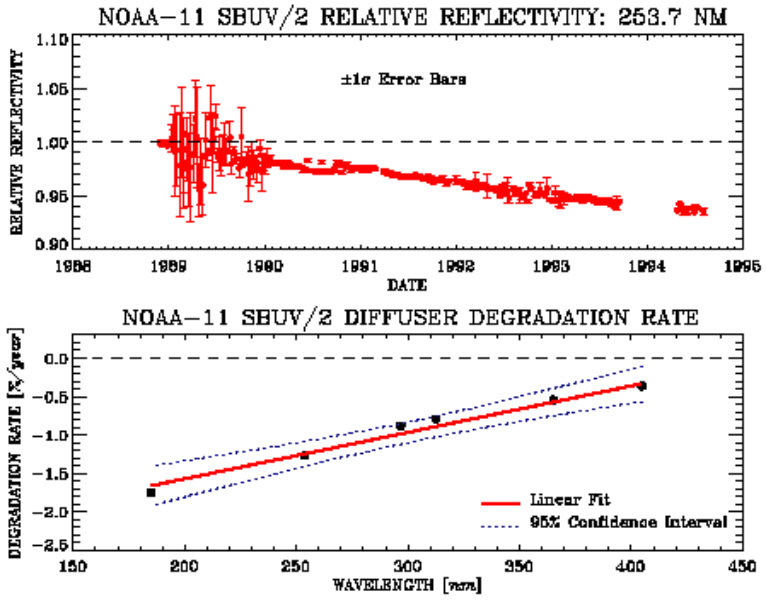


Figure 1.18. The relative diffuser degradation observed in the NOAA-11 SBUV2 instrument at 253.7 nm based on measurements of the on-board mercury lamp (upper panel). The wavelength dependent degradation rate is shown in the lower panel.

1.5.3. SSBUV Calibration Validation for SBUV/2

Periodic Shuttle SBUV (SSBUV) flights were conducted, principally to provide validation of long-term degradation of SBUV/2, but also to provide primary calibration data if the on-board calibration system failed, as was the case for the NOAA-9 instrument [Frederick *et al.*, 1990]. SSBUV absolute calibrations and procedures were similar to those described for SBUV/2. The SSBUV instrument participated in several laboratory intercomparisons to establish calibration biases relative to other instruments measuring solar irradiance. These tests demonstrated that SSBUV absolute irradiance calibration error is less than 1% (2 s) over the ozone wavelengths relative to NIST irradiance standards [*R.. Madden, NIST, private communication,* 1993]. Subsequent comparisons, over the period 1993 to 1995, have shown that SSBUV irradiance calibration is maintained to better than 1% [Hilsenrath *et al.*, 1993].

The on-board calibration system for the NOAA-11 SBUV/2 performed successfully for most of its lifetime. SSBUV's primary role was then to check and set the pre-launch calibration and validate the SBUV/2 performance over time. The validation technique is described by Hilsenrath *et al*. [1995] and is called the albedo normalisation factor (ANF) correction where albedo measurements are compared between two spacecraft. This correction employs coincident (with a one hour window) SSBUV and SBUV/2 albedo measurements. By means of a radiative transfer calculation, small differences in surface reflectivity, solar zenith angle, wavelengths, and ozone are accounted for when the measurements are compared between different spacecraft. The ANF for seven sets of coincident SSBUV and SBUV/2 observations corresponding to seven SSBUV missions are shown in Figure 1.19. Peak-to-peak ANF's (calibration errors) are about 3%, which translates to an ozone error of about 3% and 5% at 10 hPa and 1 hPa, respectively. Drifts in ANFs for the NOAA-11 data period are about -0.5% per year (+1.0% per year for ozone) at wavelengths corresponding to 1.0 hPa ozone and less than ±0.2% per year (±0.3% per year for ozone) at wavelengths corresponding to higher pressures. The residual calibration drift in SBUV/2 detected by SSBUV has large uncertainties because of the few samples (7) provided by the SSBUV missions. However they should be considered along with algorithmic drifts caused by the NOAA-11 drifting orbit when comparing SBUV/2 trends with other instruments.

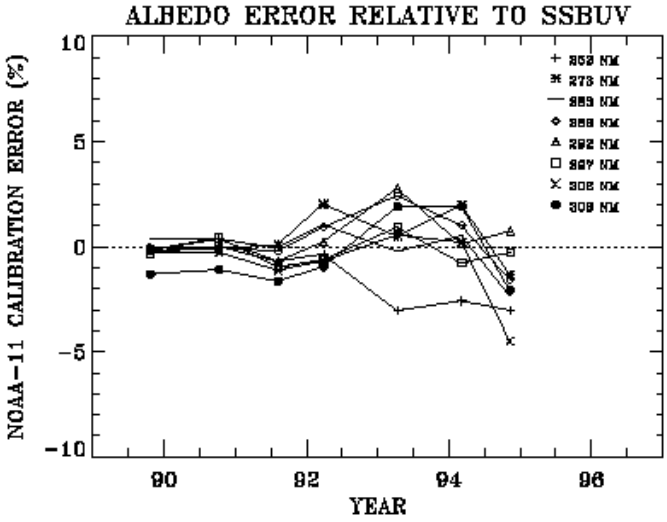


Figure 1.19. The relative error in the NOAA-11 SBUV2 radiance-to-irradiance ratio determined through comparison with 7 flights of the shuttle-based SBUV2 instrument (SSBUV).

1.5.4. SBUV Algorithms

The algorithm developed by NASA's Ozone Processing Team to obtain ozone profiles has been used to process data from the SBUV experiment on the Nimbus 7 satellite, and by follow-on SBUV/2 instruments on NOAA polar satellites. The algorithm described here is the most recent version (V6.0) of such algorithms, now the operational NASA and NOAA algorithm. This description of the algorithm is largely taken from a recent paper by Bhartia *et al*. (1996). The NOAA-11 SBUV/2 data analysed in this report are designated version 6.1.2; the ozone retrieval algorithm is the same as that used for SBUV, but an updated calibration and minor instrument-specific changes have been applied. A similar reprocessing of data from the NOAA-9 SBUV/2 instrument is currently (summer 1997) being done. In the following we shall use the lower case letters "buv" as a generic abbreviation for "backscattered ultraviolet," and the

upper case letters, e.g., SBUV, to refer to specific instruments.

Basis for Retrieval

Figure 1.20 illustrates the physical basis of the buv method for estimating the ozone profile. It shows the contribution to the buv radiances (255.5 nm to 331.2 nm) from different levels in the atmosphere. With increasing wavelength, the photons penetrate more deeply into the atmosphere, and thus sample different regions of the atmosphere. The contribution function is the fractional contribution to the backscattered radiance if one divides the atmosphere in layers of equal Dlog(p). Since the peak of the contribution function varies with a (the ozone cross section) times S (the optical path), one can scan the atmosphere by varying either a or S. The buv technique uses the former method; the ground-based Umkehr technique, which is based on similar principles, uses the latter method. The shortest buv wavelengths form clear contribution function peaks in the stratosphere and are used to infer the ozone vertical profile. The longest 4 buv wavelengths usually get their maximum contribution from the troposphere, so these wavelengths are used to infer total column ozone. At intermediate wavelengths two separate peaks can form and these wavelengths are sensitive to total column ozone as well as to the stratospheric ozone profile. The peak position and hence the information content of the radiances at any given wavelength, changes with solar zenith angle. Thus a wavelength suitable for determining total ozone at small solar zenith angle may become too sensitive to the ozone profile to be used for that purpose at large solar zenith angle.

The Backscattered Radiance Calculation

The problem of estimating buv radiances from the Earth's atmosphere varies considerably in complexity depending upon wavelength. At wavelengths shorter than 290 nm, where strong ozone absorption sharply cuts off the penetration of photons below 30 km (Figure 1.20), the buv radiances can be calculated by solving a relatively simple single-scattering problem in a purely Rayleigh-ozone atmosphere. However, at longer wavelengths the buv photons penetrate deep into the troposphere where they not only suffer multiple scattering in the optically thick Rayleigh atmosphere but are also affected by terrestrial surfaces, aerosols, and clouds. Given the computational complexity of these effects, it is convenient to consider the buv radiances as composed of two parts. The singly-scattered (SS) component, that includes just the primary scattering in an atmosphere consisting only of Rayleigh scattering, and the multiply-scattered and -reflected (MSR) component, that includes all other types of scattering and surface effects. During operational processing, the SS component is estimated accurately by numerical quadrature. The MSR component is estimated separately using a table look-up procedure.

Click here to download Figure 1.20 by anonymous ftp.

Figure 1.20. The altitude dependent contribution to the radiance backscattered from the atmosphere.

Considering only Rayleigh scattering and ozone absorption, the singly-scattered component *I*ss of the buv radiances is given by,

$$I\mathrm{ss}(\lambda) = F\lambda\, b\lambda\, P(\theta)/4\pi \int \exp[-S(p)\,\{a\lambda\, X(p) - b\lambda\, p\}]\, dp$$

where, Fl is the solar flux at wavelength l, bl the effective Rayleigh scattering coefficient per unit pressure, $P(\theta)$ the Rayleigh scattering phase function for a scattering angle q, al the effective ozone

absorption coefficient per unit ozone amount, $S(p)$ the slant path (air mass), and $X(p)$ is the column ozone above the pressure p.

At wavelengths longer than 290 nm, the incoming photons penetrate into the lower atmosphere where they may undergo multiple scattering. In addition, one must consider the effects of surface reflection and Mie scattering by aerosols and clouds. Since these effects are highly variable and cannot be modelled accurately, the buv technique relies on the concept of Lambert-equivalent reflectivity, introduced by Dave (1977). In this concept, one estimates the buv radiances by radiative transfer calculation for an idealised atmosphere that contains ozone and Rayleigh scattering, but no Mie scatterers, and is bounded by an opaque Lambertian surface. The effective reflectivity, R^*, and effective pressure, p^*, of this fictitious surface are estimated by using the measured radiance at the longest buv wavelength (339.8 nm), where the ozone absorption is very weak. The MSR radiances are computed by using the auxiliary equation solutions of the radiative transfer equation (Dave, 1964), which accounts for polarisation and all orders of scattering. However, Earth's sphericity effects are included only in the calculation of primary scattering. Recent calculations (Caudill *et al.*, 1997) indicate that this approximation produces errors of less than 1% in radiances at solar zenith angles up to 88°. For computational efficiency, a table look-up procedure is used to compute the MSR component of I. The table look-up procedure works because the ozone profile dependence of I_{MSR} is weak. The MSR radiation is generated primarily in the troposphere (below 100 hPa).

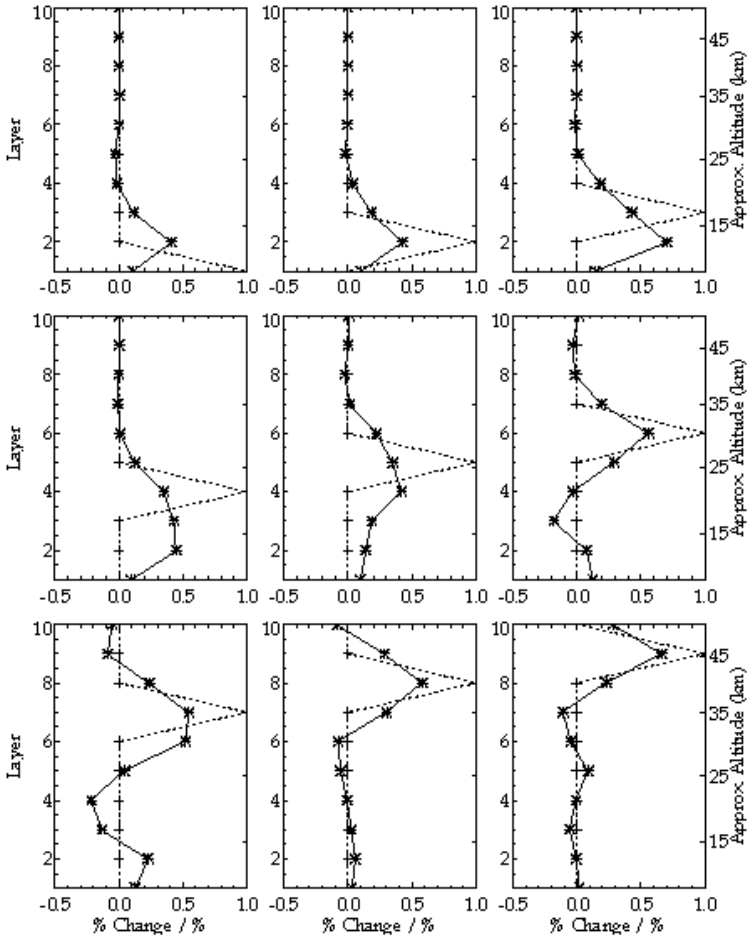


Figure 1.21. Averaging kernels for individual Umkehr layers for the SBUV retrieval algorithm showing the expected sensitivity to perturbations at various altitudes.

The Inversion

The general problem of inverting buv measurements to derive an ozone profile is ill-conditioned. One must apply some sort of constraint to achieve a physically reasonable solution. There are several ways of specifying these constraints. One may, for example, constrain the profile to follow a pre-defined mathematical function with adjustable parameters that are selected to best fit the measurements; or select a profile that has the highest probability of being the correct profile based on a priori statistical information (Rodgers, 1976). The buv profiling algorithm uses a combination of both of these strategies.

The ozone profile is mathematically defined by a cubic interpolating spline (in the Umkehr coordinate system layers log$X$ vs. log$p$ with 12 fixed nodes at p=2-n atm, where n=0, 2, 3...12 (note that n=1 is missing). The algorithm first obtains X at n=0, i.e., the total column ozone, using the longest 4 buv wavelengths. Total ozone is used to estimate I_{MSR} as well as to select first guess ozone values at the lowest 5 nodes from a library of ozone profiles. Next, ozone values at the 4 uppermost nodes are estimated using measurements at the shorter buv wavelengths. A cubic spline curve fitted to these 9 nodes forms the starting profile, which is iteratively adjusted until the retrieval converges to a solution profile. The following sub-sections discuss this procedure in more detail.

Total Column Ozone

Total ozone is derived using the 4 longest wavelength bands of the instrument (Dave and Mateer, 1967; Klenk *et al*., 1982; McPeters *et al*., 1996). Since these references discuss the total ozone algorithm in adequate detail, we provide here just an overview. The buv total ozone technique is based on the observation made earlier that at certain wavelengths the buv photons penetrate most of the ozone column to reach the troposphere, whence they are scattered by the dense atmosphere, clouds, aerosols, and surface. When the peak of the contribution function is in the troposphere, the logarithms of buv radiances decrease almost linearly with ozone column amount, approximately following the Beer's law of absorption. Since this relationship is not perfectly linear but has a small dependence on ozone profile and temperature, one needs to perform accurate radiative transfer calculations using climatological profiles to describe the curves relating the buv radiances to total column ozone under a variety of observing conditions. Total column ozone is derived in a table look-up procedure described in the references cited earlier.

Construction of First Guess Profile

The first guess profile is constructed by fitting a cubic interpolating spline with natural boundary conditions to the log$X$ vs. log$p$ curve. The piece-wise linear curve given by the *Thomas and Holland* (1977) solution, linearly extrapolated, provides the cumulative ozone X at the 4 topmost points, at p=2-n atm, with n=9, 10, 11, 12. The value of X at n=0 is simply the total column ozone. Values of X at n= 2, 3, 4, 5, are obtained using a library of standard ozone profiles constructed by combining data from SAGE (McCormick *et al*., 1989) with ozonesonde data. A profile for a given total ozone amount and latitude of measurement is obtained by bi-linear interpolation (low latitude profiles are extended to the equator and high latitude profiles to the pole). Given the values of X at the 9 nodes, the values at the missing 3 nodes (n=6,7,8) are obtained by spline interpolation.

Retrieval of Final Ozone Profile

The buv algorithm assumes that the MSR component of the total measured radiation contains no retrievable ozone profile information. Since the total ozone information that it does contain has been

obtained previously using the longer buv wavelengths, I_{MSR} is treated simply as a contamination to the measured radiance and is subtracted out. The retrieval scheme, therefore, is based on single-scattering kernels and singly-scattered radiances. Measurements made at the longest 4 wavelengths (3 at high solar zenith angles) that have already been used in deriving R^* and are not used explicitly. Instead, the total ozone derived from them is considered a measurement and is included with the radiance measurements at the other 8 wavelengths (9 at high solar zenith angles, by including the 312.5 nm wavelength). Although this scheme of handling the measured radiances may appear unnecessarily tedious, it was done primarily to save computation time. Back-calculations using the retrieved profiles show that, in most cases, the derived ozone profile explains the measured radiances at all 12 wavelengths to a high degree of precision (~1%).

The final retrieval of an ozone profile is a straightforward implementation of Equation (101) of Rodgers (1976) reproduced below.

$$\mathbf{x}n+1 = \mathbf{x}a + \mathbf{S}x\ K n T\ (\mathbf{K}n\ \mathbf{S}x\ \mathbf{K}n T + \mathbf{S})\text{-}1\ [(\mathbf{y} - \mathbf{y}n) - \mathbf{K}n\ (\mathbf{x}a - \mathbf{x}n)]$$

In this equation, **x**a is the a priori profile. The solution starts with a first-guess profile **x**1, from which the kernels **K**1 and measurements **y**1 are evaluated, and proceeds through n iterations until the retrieved profiles stop changing to within a specified tolerance. Although **x**1 is not required to be the same as **x**a , they are made the same in the buv algorithm. Strictly speaking, since measured radiances are used in constructing **x**a, one does not have a true a priori profile to which the solution is constrained. This violates the tenets of optimal estimation theory upon which this equation is based. However, the principal impact of this violation is to complicate the error analysis of the algorithm.

Ozone profiles derived from the inversion are reported as ozone amounts (in DU) in 12 layers defined above, as well as ozone mixing ratios at 19 standard pressure levels. The latter are obtained from the first derivative of the cubic spline curve. However, as discussed in previous sections, the buv radiances are directly sensitive to the column amount of ozone above a pressure surface. To compute layer ozone amounts or mixing ratios one must difference two numbers. In the upper stratosphere, where the column amount of ozone is changing rapidly, the numerical differentiation process increases the error only modestly. Below the ozone density peak, however, one must difference two large numbers to obtain a small number. This can lead to large error enhancement, particularly in the tropical lower stratosphere, where the amount of ozone in a layer, as a fraction of the ozone column above, is the smallest.

Information Content

A convenient way of examining loss of information due to inversion is to treat the algorithm as a "black box" and study its response to perturbations in the ozone profile. Rodgers (1990) recommends a variant of this technique based on a concept called the averaging kernels. Besides having an elegant physical interpretation, averaging kernels are very convenient for comparing profiles from a low-resolution instrument with those from a high-resolution instrument. Using the averaging kernels and the a priori profile one can immediately tell what profile the algorithm would produce for a given "truth" profile from another sensor.

A form of averaging kernel where the first guess is included in the "black box" is shown in Figure 1.21. Between 1.5 hPa and 20 hPa, the algorithm reproduces the input perturbation reasonably well except for broadening the FWHM to about 10 km. These results are derived using the 252 nm wavelength of SBUV/2. If the shortest SBUV wavelength is not used, the upper range of validity drops below 1.5 hPa. at low solar zenith angles. Above 1.5 hPa and below 20 hPa the response has the wrong shape, with misplaced peak positions. These results are in general agreement with previous conclusions (Rodgers, 1990). The primary reason for the distorted response is that in the region of the atmosphere where the

information from the buv radiances is low, the algorithm relies on a priori information more heavily. Since the Gaussian impulse puts perturbations in the atmosphere that are very different from a priori information, the algorithm finds a compromise answer that poorly reproduces the impulse.

1.5.5. Algorithm Errors

Errors in retrieved ozone profiles come from 3 independent sources: measurement errors, error in calculating radiances from a known ozone profile, called the "forward model errors", and error in inverting the radiances to obtain an ozone profile, called the "inverse model errors". These errors are discussed in detail in Bhartia *et al*. (1996). To estimate the effect of these errors we use high vertical resolution ozone profiles generated by the SAGE instrument. Direct comparison between SAGE II and SBUV derived ozone profiles have been made (McPeters *et al*., 1994; Rusch *et al*., 1994). However, these comparisons fail to provide unambiguous answers about the capability of the buv inversion algorithm itself, for they are affected by differences in the spatial and temporal location, instrumental errors, and forward model errors of both instruments. From a SAGE data set, consisting of 1655 individual ozone profiles matched to SBUV profiles, the buv radiances were computed using the average solar zenith angle seen by the SBUV instrument in the latitude band where SAGE took measurements. The computed buv radiances were then used to derive ozone profiles using the algorithm discussed in this paper.

As expected, the SAGE comparisons show that the buv technique can measure the cumulative ozone amounts above a pressure surface to a high degree of precision. Figure 1.22 shows that, in the absence of measurements errors, layer ozone amounts between 0.7 hPa and 10 hPa are retrieved with a precision of 2-3% with a standard deviation under 2%. Lower layers are not shown since the percentage errors for these layers lose meaning as noted earlier. Cumulative ozone is the direct product of the buv

Click here to download Figure 1.22 by anonymous ftp.

Figure 1.22. An estimate of the ability of the SBUV algorithm to retrieve ozone made using radiances theoretically calculated from an ensemble of 1655 real (SAGE) profiles. The average difference is shown in the left panel, while the standard deviation of the retrieved profiles is shown in the right panel.

inversion - between the surface and 55 km it is derived with a precision of better than 2%. One notes that the errors do not average to zero when large number of profiles are averaged. This is a consequence of the a priori constraint which may introduce systematic errors if the a priori profile is not the ensemble average of the subset of SAGE II data examined. This appears to be the case in the lower layers at southern mid-latitudes.

1.5.6. Solar Zenith Angle Effects

Figure 1.22 showed that ozone perturbations between 1.5 hPa and 20 hPa would be accurately retrieved by the algorithm. Because the altitude at which the scattering contribution peaks is determined by the product of the ozone cross section (wavelength) times the optical path, the maximum altitude accurately retrieved will move upwards as the solar zenith angle increases. If the maximum altitude is 1.5 hPa for zenith angle less than 60o, the profile will be valid up to 1.0 hPa by 75o and to 0.7 hPa by 85°. As a result profile retrievals at high latitudes in winter extend higher in the atmosphere.

Another interpretation is that the same altitude is measured by longer wavelengths as the solar zenith angle increases. This can cause an error in the trends derived from the NOAA-11 SBUV/2 because, as was shown in Figure 1.17, the orbit of the NOAA-11 spacecraft drifted towards the terminator (higher solar zenith angles) over its life. The result is that any wavelength dependent error will directly produce an ozone trend error. Figure 1.23 shows the theoretical change in the profile (trend) caused by an orbit drift from a 30° zenith angle to a 60° zenith angle if a wavelength dependent calibration error is assumed. For the 8 profile wavelengths 252 nm to 306 nm, errors are assumed of -0.5, +1.0, +1.0, 0.0, 0.0, -1.0, +1.0, and 0.0% respectively. These are realistic possible errors suggested by the final residues from the SBUV/2 processing and within the uncertainty of the calibration. They show that a trend error of 1 to 2% over the six year life of SBUV/2 is possible. A similar assumed error in the SBUV calibration would result in an altitude dependent bias, but would not produce a similar trend error because the Nimbus 7 orbit was very stable.

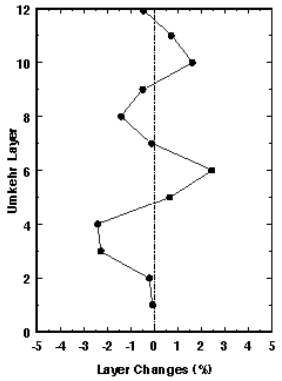


Figure 1.23. An illustration of how an assumed calibration error that does not vary with time can produce an apparent time and altitude dependent trend as a result of an orbit that drifts to higher solar zenith angles (from 30o to 60o).

1.5.7. SBUV Error Summary

Table 1.8. contains estimated errors (90% limits) for SBUV from the algorithm [Bhartia *et al.*, 1996] and from the calibration, and an estimate of the error in maintaining the calibration [Bhartia *et al.*, 1995]. Since many of the components of these are errors that are systematic, they will not be completely eliminated by averaging profiles.

Pressure (hPa)	algorithm %	calibration %	calibration drift %/year
0.5	6	5	0.5
1.0	6	5	0.5
2.0	5	4	0.5

4.0	5	3	0.5
7.0	5	3	0.46
10.0	5	2	0.36
20.0	7	2	0.29
40.0	11	2	0.4
100.0	15	2	0.4

Table 1.8. SBUV Error summary

1.6. LIDAR

1.6.1. Principle of the lidar and the DIAL technique

The lidar (LIght Detection And Ranging) is a remote sensing instrument similar in principle to the radar but operating in the optical range. Depending on the desired measurement, lidar systems use various light-matter interactions such as Rayleigh, Mie and Raman scattering or fluorescence. Measurements of atmospheric ozone, temperature or aerosol are based on the first 3 processes. Generally, a lidar measurement consists in sending into the atmosphere a laser beam ; a small part of this laser radiation is scattered back to the ground, where it is collected by a telescope, detected by a photomultiplier tube and analysed by an electronic acquisition system. Range resolved measurements can be obtained using pulsed lasers. In order to measure the ozone vertical distribution, the Differential Absorption Laser technique (DIAL) is used. This technique requires the simultaneous emission of two laser beams characterised by a different ozone absorption cross-section.

1.6.2. Description of ozone lidar systems

A lidar system includes basically one or several laser sources with optical devices to reduce the divergence of the beam, a telescope which collects the light scattered back by the atmosphere, an optical analysing system with detectors such as photomultipliers to detect the optical signal, and an electronic acquisition system (see Figure 1.24). The analysing systems used to digitise the electronic signal provided by the photomultipliers include photon counting and/or transient analysers. In the case of the DIAL systems characterised by the emission of two laser wavelengths, the optical receiving system comprises spectral analysing optics, such as interference filters or spectrometers.

To monitor atmospheric ozone with the DIAL technique, the choice of the laser wavelengths depends on the altitude range of the measurement (Mégie *et al.*, 1985). The spectral range is chosen first in the ultraviolet where the ozone absorption is more efficient, but the selected wavelengths differ according to whether the measurement is made in the troposphere or in the stratosphere - in the troposphere, the ozone number density is small so the laser wavelengths must correspond to a strong UV absorption, while for stratospheric measurements, the objective is to reach the stratosphere and to detect the high ozone concentrations there (Browell, 1989, Papayannis *et al.*, 1990). Furthermore, in the higher stratosphere, one has to consider the simultaneous decrease of the ozone number density and the atmospheric number density which provides the backscatter radiation. This leads to the need for powerful laser sources in order to reach the high altitude ranges. The absorbed wavelength should not be strongly absorbed in order to reach the stratosphere. Most teams working on this subject use XeCl eximer laser sources, which emit directly in the UV at 308 nm (Uchino *et al.*, 1978) and are very powerful (100 W are commonly reached with the present systems). For the non-absorbed wavelength, different techniques are used, mainly the generation of a wavelength at 353 nm corresponding to the first Stokes radiation by stimulated Raman effect in a cell filled with hydrogen (McDermid *et al.*, 1990, Werner *et al.*, 1983, Swaart *et al.*, 1994), the use of the third harmonic of a Nd:Yag laser (355 nm) (Godin *et al.*, 1989, Uchino *et al.*, 1991), or the use of a XeF laser which provides a wavelength at 351 nm (McGee *et al.*, 1991, Nakane *et al.*, 1994).

At the present time, about 12 stratospheric ozone lidar systems are in operation throughout the world for long-term monitoring. The lidar measurements used in this report are summarised in Table 2.4 of chapter

2. They correspond mainly to the longest ozone lidar time series obtained in the latitude range covered by the SAGE II instrument. Most of these measurements are presently performed in the frame of the NDSC (Network for the Detection of the Stratospheric Changes) which was set up in the beginning of the 1990s for ground based long-term monitoring and satellite validation purposes.

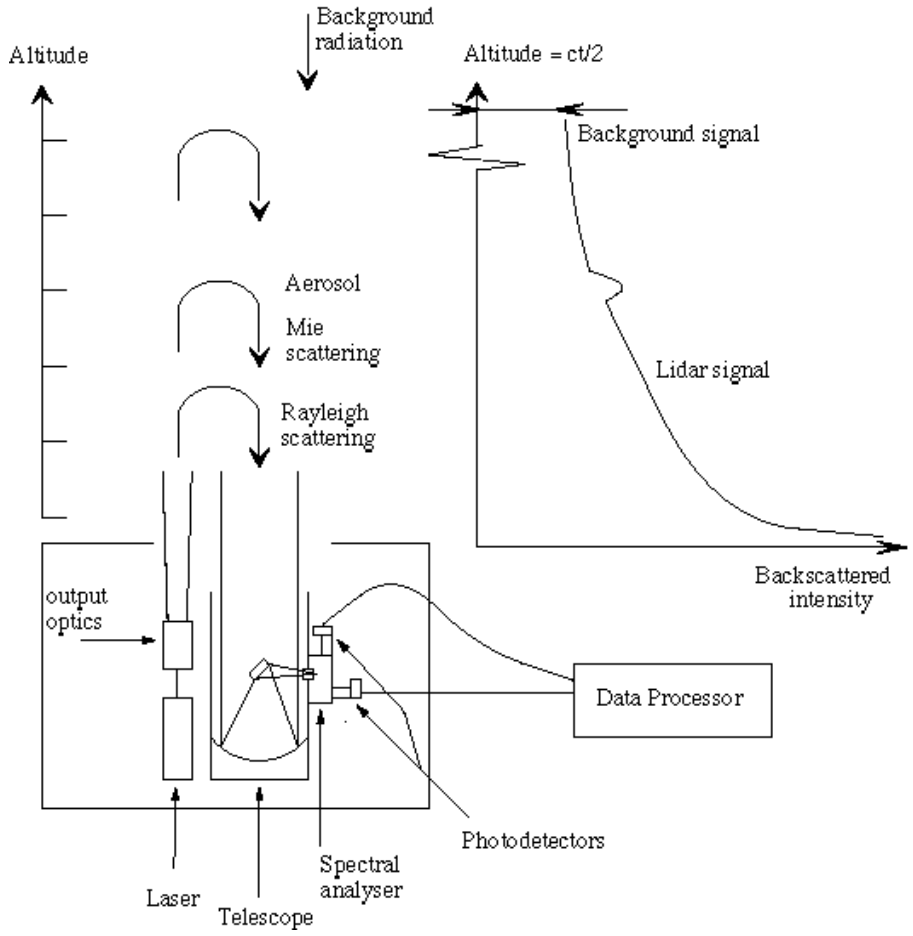


Figure 1.24. Schematic view of the principle of a lidar system

1.6.3. Description of the measurement and sampling strategy

The lidar signals cover a very high dynamic range corresponding to several orders of magnitude, which is not handled by the electronic acquisition systems. This requires the use of simultaneous photon counting and analogue acquisition for the low and the high signals or the separation of the optical signal in two parts (90% and 10%), corresponding to the high and the low altitudes respectively. For reasons of simplicity in terms of electronic acquisition, the latter solution is the most commonly used. Depending on the power and the repetition rate of the laser, an ozone measurement lasts typically several hours, leading to a spatial resolution of the order of 100 km, depending on the atmospheric conditions. The vertical resolution ranges from several hundred meters in the lower range to several kilometres above 40 km. This evaluation is detailed in the following sections. Finally, one main caveat of the lidar measurements is the requirement of clear sky meteorological conditions - laser radiation is rapidly absorbed by clouds and only cirrus can be tolerated for accurate stratospheric measurements.

1.6.4. Ozone retrieval from the lidar signals

Assuming a monochromatic laser impulsion at wavelength l, the received optical power corresponding to the light backscattering at altitude z, is given by (Measure, 1984) :

$$P(z) = K(\lambda)/z^2)\exp[-2\tau(\lambda,z)] \quad (1.6.1)$$

where $K(\lambda)$ is an instrument constant involving the telescope surface area, the emitted power, the optical efficiency of the receiving system and a geometrical factor depending on the alignment of the laser and the telescope axis, $\beta(\lambda,z)$ is the atmospheric backscatter coefficient and $\tau(\lambda,z)$ is the atmospheric optical depth. $\tau(\lambda,z)$ depends on the following parameters :

$$\tau(\lambda,z) = \int_{z_0}^{z}\left[\alpha(\lambda,z') + \sigma_{O_3}(\lambda)n_{O_3}(z') + \sum_i \sigma_i(\lambda)n_i(z')\right]dz' \quad (1.6.2)$$

where $\alpha(\lambda,z)$ is the atmospheric extinction coefficient, $\sigma_{O_3}(\lambda)$ the ozone absorption cross-section, $n_{O_3}(z)$ the ozone number density to be measured and the term $\sum \sigma_i(\lambda)n_i(z)$ corresponds to the extinction by other absorbers.

Applying this formula to the second wavelength and taking into account the background signal, one derives the ozone number density from the received lidar signals:

$$n_{O_3}(z) = \frac{-1}{2\Delta\sigma_{O_3}(z)}\frac{d}{dz}Ln\left(\frac{P(\lambda_1,z) - P_{B1}}{P(\lambda_2,z) - P_{B2}}\right) + \delta n_{O_3}(z) \quad (1.6.3)$$

where *P*bi is the background signal and:

$$\delta n_{O_3}(z) = \frac{-1}{\Delta\sigma_{O_3}(z)}\left[\frac{1}{2}\frac{d}{dz}Ln\left(\frac{P(\lambda_1,z)}{P(\lambda_2,z)}\right) - \Delta\alpha(z) - \sum_i \Delta_i n_i(z)\right] \quad (1.6.4)$$

The laser wavelengths are chosen so that the term $\delta n_{O_3}(z)$ represents less than 10% of the measurement. The derivation of the ozone number density from the laser signals shows thus that the DIAL technique is a self-calibrated technique which doesn't need the evaluation of instrumental constants.

Main steps of the algorithm

The DIAL algorithm follows basically the theoretical derivation of the ozone number density from the lidar signals. The main steps are the following :

- Temporal signal averaging to increase the signal to noise ratio.
- Correction from :
  - background light
  - dead time correction in the case of photon counting acquisition, due to the saturation of the photon counting systems with high intensities.
- Derivation of the ozone number density from the lidar signals corrected from the background light. This requires the differentiation of the lidar signals by numerical derivative filters. Furthermore, as described in the following section, it necessary to degrade the resolution of the measurement in the upper altitude range. A lowpass filter with a varying cut-off frequency with respect to altitude is

thus also applied to the lidar signals.

1.6.5. Accuracy and vertical resolution of the measurement

The precision of a DIAL measurement is defined by the statistical error due to the random character of the detection process which follows basically the Poisson statistics (Measures, 1984). The accuracy of the measurement depends on the approximations made in deriving the ozone number density from the received signals. It depends also on the linearity of the lidar signals.

According to the Poisson statistics, the statistical error on ozone is given by the following formula :

$$\varepsilon_1(z) = \frac{1}{2 N_{b,i}(z) \Delta\sigma_{o_3} \Delta z} \left[\sqrt{\sum_{i,j} \frac{c_j^2 P_{i,j}}{N_i (P_{i,j} - P_{bi})^2}} \right] \quad (1.6.5)$$

where DZ is the initial range resolution of the acquisition system, Pi,j corresponds to the lidar signal at wavelength i from altitude Zj, cj are the coefficients of the low pass derivative filter used to differentiate the signals, Ni is the number of laser shots at wavelength li and Pbi is the background radiation at wavelength li. The final statistical error e s on the measurement is the result of a compromise depending on the experimental system characteristics, the duration of the signal acquisition and the vertical resolution according to the following relation :

$$\varepsilon_1(z) \propto \left(A \Delta z^3 P_0 T_a \right)^{-1/2} \quad (1.6.6)$$

where A is the telescope receiving area, DZ the final range resolution, P_0 the emitted power and Ta the acquisition time. Due to the rapid decrease of the signal to noise ratio in the high stratosphere, it is necessary to degrade the vertical resolution of the measurement in order to limit the statistical error at this altitude range, to reasonable values. Most teams choose low pass filters with varying number of points as a function of the altitude. The DIAL stratospheric ozone lidar profiles are thus generally characterised by a vertical resolution varying from several hundred meters in the lower stratosphere, to several kilometres around 50 km (see Figure 1.25 for an example of ozone lidar vertical resolution profile).

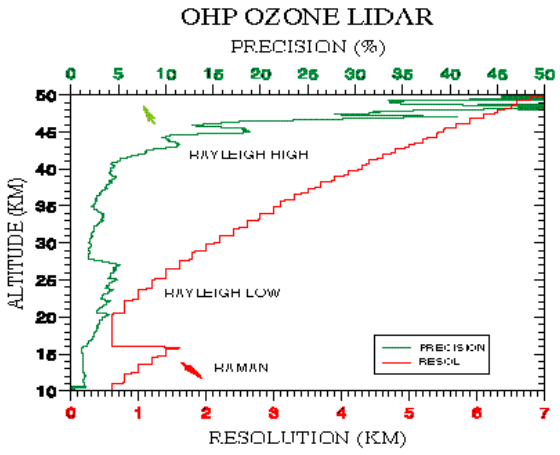


Figure 1.25. Precision and vertical resolution profile of an ozone measurement in the case of the OHP lidar

instrument. Both the precision and the vertical resolution profile depend on the experimental configuration. The precision can vary from one measurement to the other.

The accuracy of the measurement depends on the term $\Delta\beta_a(z)$ (see equation 1.6.4) which corresponds to less than 10% of the value derived directly from the slope of the signals but still has to be corrected using ancillary measurements. It depends also on the accuracy of the ozone absorption cross-sections and on the approximation concerning the monochromaticity of the laser radiations. These error sources are summarised in Table 1.9 which indicates the residual error on the measurement after correction (Godin, 1987).

Error source	Residual
Ozone absorption cross-section	
- Absolute value (Bass&Paur)	2 %
- Temperature sensitivity	< 0.5%
Laser line width	< 0.3%
Rayleigh extinction	<0.6%
Other absorbers	
- SO_2 - normal conditions - after major volcanic eruption	negligible 1%
- NO_2	< 0.3%
Aerosol backscatter and extinction	
- Volcanic conditions : - correction using ancillary size distribution measurements - use of Raman channel	 30% < 5%
- background	< 5%

Table 1.9

For the temperature sensitivity of the ozone absorption cross-sections and the Rayleigh extinction term, the correction requires daily pressure-temperature profiles provided by nearby radiosondes and assimilated satellite data or temperature lidar measurement in the upper altitude range. For the error related to the nitric oxide, an annual average profile is generally used. Since the residual error is directly related to the precision of the parameters and ancillary measurements used in the ozone formulation, it

can be considered as random.

Error due to volcanic aerosol and use of Raman scattering

In the presence of volcanic aerosol, the DIAL ozone profile is locally perturbed at the altitude of the aerosol cloud. The corresponding error is expressed in equation 1.6.7:

$$\varepsilon_a(z) = \frac{1}{\sigma_{O_3}(\lambda_1,z) - \sigma_{O_3}(\lambda_2,z)}\left[\frac{1}{2}\frac{d}{dz}Ln\frac{\beta(\lambda_1,z)}{\beta(\lambda_2,z)} - (\alpha(\lambda_1,z) - \alpha(\lambda_2,z))\right] \quad (1.6.7)$$

It can be separated in two components related to the backscatter and extinction effects respectively. The main contribution to the error is due to the backscatter effect, directly linked to the aerosol since the corresponding molecular contribution, proportional to the atmospheric number density for both wavelengths, cancels out.

ε_a can be determined if one evaluates the spectral behaviour of the aerosol backscatter and extinction coefficients from assumed aerosol size distribution. But the Angström coefficient of the backscatter coefficient is highly non-linear and residual errors can exceed 100% in the case of very strong aerosol loading (Godin *et al.*, 1994, Steinbrecht and Carswell, 1995).

Another more efficient method to handle this problem is to detect the lidar signals corresponding to the first stokes vibrational Raman scattering by molecular nitrogen of the emitted wavelengths (McGee *et al.*, 1993). The advantage of this method is to suppress the aerosol contribution to the backscatter component of the lidar equation but its drawback is a loss of accuracy due to the much smaller Raman scattering efficiency. In order to compensate this effect, one has to degrade the vertical resolution of the corresponding measurement following the compromise expressed in equation 1.6.6 (see Figure 1.26 for an example of ozone Rayleigh and Raman profiles). The residual error related to the aerosol extinction can be evaluated with aerosol models based on experimental measurements. It has been shown experimentally (McGee *et al.*, 1994) to be less than 5% even in the case of heavy aerosol loading.

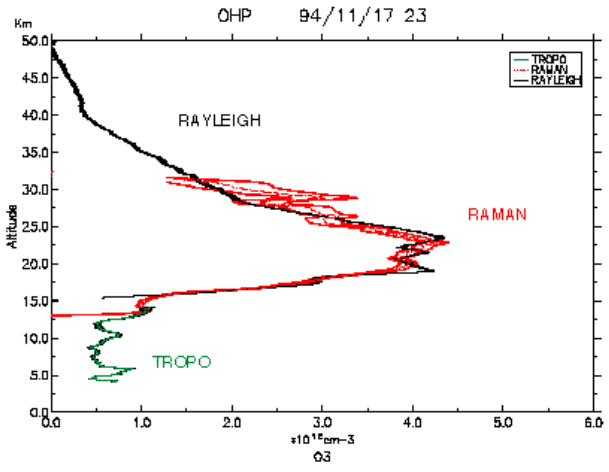


Figure 1.26. Example of Rayleigh and Raman ozone lidar profiles. These measurements are compared to a tropospheric lidar profile obtained on the same site with a different lidar instrument.

Instrumental error sources

As mentioned previously several experimental artefacts can alter the linearity of the lidar signals and thus introduce biases in the retrieved ozone concentration. The main problems encountered with the

DIAL technique are :

- Alignment errors

  In order to obtain a correct ozone measurement, the axis of the laser beams and the receiving telescope must be aligned. Any error in the alignment procedure will effect the slope of the signals and then the ozone retrieved density.

- Signal induced noise in the photomultiplier tubes

  This effect is due to the high intensities originating from the scattering of the laser light in the first few kilometres. It results in a slowly decreasing signal which is superposed on the background light (McDermid *et al.*, 1990). This problem is perceptible at the high altitude ranges, where the signal to noise ratio is low. It can induce high or low biased ozone concentrations depending on which wavelength is the most affected by the signal induced noise. In order to avoid the signal induced noise effect, it is necessary to include mechanical choppers in the experimental set-up, since electronic obscuration of the photomultiplier tubes is not sufficient to tackle this problem.

- Saturation of the acquisition systems

  At low altitude ranges, the photon counting systems cannot handle the high intensities received by the photomultiplier tubes. Most lidar teams use several electronic channels for one wavelength, which allows to preserve the linearity of the signal by switching from one channel to the other. In that respect, the use of Raman channels is very useful, since their low intensity allow to check the linearity of the Rayleigh channels in the lower altitude range.

- Bias introduced by the algorithm in the high altitude ranges

  As mentioned earlier, it is necessary in the DIAL technique to use a low pass filter in order to take into account the rapid decrease of the signal-to-noise ratio in the higher altitude range. The usual method is to use a filter whose number of points varies with altitude, which defines a final vertical range resolution profile. Above 40 km or more, depending on the power of the laser sources, the use of a high number of points introduces a bias in the ozone retrieved quantity. Recent simulations performed in the frame of the NDSC by various lidar teams allowed to quantify this bias for a given configuration. This bias can reach 10% at 45 km (Godin, 1996).

1.6.7. Conclusion

The DIAL method is a powerful technique for the measurement of the ozone vertical distribution in the atmosphere. It is particularly suited for trend studies since it is self-calibrated. The DIAL measurements are characterised by a good precision and good vertical resolution especially in the low stratosphere. Typical values are vertical resolutions ranging from ~0.5 km at 20 km and 4 to 7 km at 50 km and precision ranging from ~1% at 20 km to 10%-50% at 50 km depending on the systems and the weather conditions. Finally, the lidar measurements require clear skies, so this has to be taken into account in the choice of the site for long-term studies.

The various error sources which could have an effect on trend studies are summarised in Table 1.10. As mentioned previously, the error sources related to the term $\Delta\beta_b(z)$ of equation 1.6.4 are corrected in the ozone algorithm so their trend effect should be of second order. More generally, the errors which could affect the trends should be dominated by instrumental artefacts such as the eventual presence of signal

induced noise, detector saturation or alignment problems. But these artefacts are very difficult to quantify since they depend completely on the instrumental configuration and on the operating procedure. This is why the lidar set-ups designed for long-term operation and operating in the frame of the NDSC are required to fulfil several requirements in order to maintain the linearity of the lidar signals in the whole altitude range of the measurement. Among these requirements, one can cite in particular :

- the use of mechanical chopper and high quality photomultiplier tubes for unbiased measurements in the higher altitude range.
- the use of several optical or electronic channels for the analysis of the lidar returns in order to account for the high dynamics of the lidar signals and obtain unbiased measurements in the lower altitude range.

It is also necessary, in order to maintain a long-term homogeneous record, to use the same range resolution profile for the whole record.

Error source	range	Maximum trend effect
Temperature sensitivity of the ozone cross-section	whole profile	0.02%/year (assuming an error on temperature trend of 0.1 K/year)
Rayleigh extinction	whole profile	0.01%/year (assuming an error on the atm. density trend of 0.1%/year
NO_2 absorption	middle stratosphere	negligible
Aerosol extinction	low stratosphere	0.01%/year (assuming an error on the trend in strat. aerosol loading of 0.5%/year
Detector saturation	low stratosphere	System dependent
Alignment	high stratosphere	System dependent
Signal induced noise effect	high stratosphere	System dependent

Table 1-10.

1.7. Umkehr

1.7.1. Umkehr basic measurement description

Götz [1931] discovered that the ratio of zenith sky radiances of two wavelengths in the ultraviolet, one strongly and the other weakly absorbed by ozone, increases with increasing solar zenith angles but suddenly decreases at zenith angles close to 90o. He named this observation the Umkehr effect and realised that such measurements contain information about the vertical distribution of ozone in the stratosphere [Götz *et al*., 1934]. Umkehr observations today are performed with both Dobson [WMO, 1992; Staehelin *et al*., 1995] and Brewer [Kerr *et al*., 1988] spectrophotometers measuring the ratio of diffusely transmitted zenith-sky radiance at a wavelength pair in the ultraviolet, one wavelength strongly, the other weakly absorbed by ozone (e.g., for the C-wavelength pair, 311.5 nm is strongly absorbed and 332.4 nm weakly absorbed). These wavelength pairs are measured in a series of zenith-sky observations with the solar zenith angle changing from 60° to 90° during sunrise or sunset. During the Umkehr observation period, the Dobson or Brewer instrument also independently measures the total ozone column. The actual measurement vector consists of the logarithm of ratios of channel signals (R values) that are converted to radiance using calibrations tables [WMO, 1992; Kerr *et al*., 1996] and reported as N values in N-units for 12 discrete zenith angles between 60° and 90°. These measurements are inverted and archived as layer ozone amounts at the WOUDC.

1.7.2 Measurement spectral and spatial resolution

The Dobson instrument employs a selection of eight wavebands from 305.5 to 339.8 nm [Komhyr *et al*., 1993] while the 8 Brewer wavebands occur between 306.3 and 329.5 nm [McElroy and Kerr, 1995]. Only the C-pair (311.5 and 332.4 nm) measurements are used in the current Dobson/Umkehr inversion [Mateer and DeLuisi, 1992], while the preliminary Brewer/Umkehr inversion uses either 5 or 6 wavebands [McElroy *et al*., 1996]. Both instrument retrievals use the Bass and Paur [1985] ozone absorption cross-sections.

The Umkehr inversions report ozone in 10 Umkehr layers (layers 0 and 1 are combined into layer 1) divided into equal log-pressure vertical intervals starting at the surface (1 atm.=1013 hPa) and extending to layer 10 (9.77x10-4 atm. to the top of the atmosphere). These Umkehr layers are approximately 5 km thick and are centred roughly at the layer number times 5 km in height (e.g. layer 8 is centred at 40 km). Although the standard Umkehr retrievals report 10 layers, because of a combination of the physical atmospheric scattering process, finite instrument spectral resolution, and real atmospheric vertical correlation, the actual retrievals possess, at most, 4 independent pieces of information (i.e., significant eigenvectors [Mateer, 1965, Hahn *et al*., 1995]). The vertical position and extent of this information is contained in the averaging kernels of the measurement technique, which are summarised in section 1.7.6. In general, the Umkehr technique retrieves ozone information between 20 and 40 km and is constrained to retrieve a profile consistent with the total ozone column measurement [Mateer and DeLuisi, 1992].

1.7.3 Sampling strategy

Of the approximately 90 Dobson stations world-wide, 46 Dobson stations have been used recently for total-ozone trend studies [Bojkov *et al*., 1995]. Because the Umkehr observations require much more observing time, only 19 stations are represented in the WOUDC database with Dobson/Umkehr observations. Of those, 15 stations (some of which are automated) have sufficiently long records to be

initially considered in this study. These stations represent both hemispheres but are located predominantly in northern mid-latitudes. While some Dobson/Umkehr records extend back to 1957, the analyses in this report concentrate on the satellite-overlap period 1979-1996.

Because all of the instruments in the Brewer network, which now comprises more than 100 instruments, are automated, all Brewer stations report Umkehr measurements. Several Brewer stations have data records exceeding 10 years, sufficiently long to be used for trend studies [McElroy *et al.,* 1996] and preliminary results compare favourably with other ozone instruments [McElroy and Kerr, 1995; Hahn *et al.,* 1995]. However, the Brewer data require additional scientific effort to invert and validate before they can be used for trend studies. With appropriate inversion and validation, the Brewer records would have made a significant contribution to this report.

1.7.4 Instrument error sources

Sources of error in the Dobson instrument include optical alignment, optical wedge calibration, and detector noise. The Dobson instruments considered here are routinely calibrated for total ozone measurements with the world standard Dobson instrument 83, which maintains a long-term (>25 year) precision of approximately ±0.5% [Komhyr *et al.,* 1989; Basher, 1995]. This calibration of instrument 83 and the secondary standard instruments is maintained by both standard-lamp tests and Langley-plot calculations. While the Dobson instruments are routinely calibrated in the configuration for total ozone column measurements, they are not directly calibrated in the Umkehr mode. During measurements at high solar zenith angles, corresponding to ozone information at high altitudes, the instrument's internal scattered light level approaches the atmospheric signal. Therefore, instrument-to-instrument variability should be greatest in the upper-layer ozone retrievals. The resulting retrieval accuracy is discussed in section 1.7.6.

1.7.5 Umkehr Algorithm Description

The current Umkehr inversion algorithm [Mateer and DeLuisi, 1992] is based on the techniques of Rodgers [1990] and represents the first update of the original (1964) algorithm [Mateer and Dütsch, 1964]. The measured N-values (defined below) are inverted to provide ozone amounts (in Dobson Units, DU) in 10 Umkehr layers through the application of a radiative transfer code that takes into account the primary and multiple atmospheric scattering, atmospheric refraction, and absorption by atmospheric ozone.

The Dobson zenith sky measurements may be written as

$$N(x,z) = 100 \log10\{[I(x,z,L2)/F0(L2)]/[I(x,z,L1)/F0(L1)]\}+ C0$$

where N is the relative logarithmic attenuation for the wavelength pair and is referred to as the N-value. The quantity x refers to the ozone profile, z is the solar zenith angle, $I(x,z,L1)$ is the 311.5 nm radiance and $I(x,z,L2)$ is the 332.4 nm radiance of the C-wavelength pair, $F0$ is the extraterrestrial flux and $C0$ is an instrumental constant. In the forward model, the N-value comprises four components: the primary scattering component, N_p, the multiple scattering component N_{MS}, a refraction component N_R, and an instrumental parameter $C0$.

$$Np = N - NMS - NR - Co$$

By using differences $Yi = Ni - No$, where Ni is the N-value at zenith angle zi and $N0$ the lowest zenith angle, $C0$ and the extraterrestrial flux can be eliminated.

The forward model uses the average of the M.A.P. 40°-50° North and South temperature profile with the temperature-dependent ozone cross sections of Bass and Paur [1985] and Paur and Bass [1985]. For the forward calculation, the atmosphere is divided into 61 layers in the vertical where each layer is 1/4 of the appropriate Umkehr layer. Table 1.11 shows the layer designations used in the Umkehr retrievals along with the pressure levels and approximate altitudes. The inverse model provides ozone content in 10 layers, where layer 10 includes all the ozone above layer 9, and layer 1 includes the retrieved ozone for both layers 0 and 1.

Layer Number	Layer Base Pressure (atm.)	Layer Base Approx. Height (km)
0	1.0	0.0
1	0.5	5.5
2	0.25	10.3
3	0.125	14.7
4	6.25×10^{-2}	19.1
5	3.12×10^{-2}	23.5
6	1.56×10^{-2}	28.0
7	7.81×10^{-3}	32.6
8	3.91×10^{-3}	37.5
9	1.95×10^{-3}	42.6
10	9.77×10^{-4}	47.9
11	4.88×10^{-4}	53.2
12	2.44×10^{-4}	58.3
13	1.22×10^{-4}	63.1
14	6.10×10^{-5}	67.8
15	3.05×10^{-5}	72.2

Table 1.11. Layers used for Umkehr ozone profile retrievals [Mateer and DeLuisi, 1992].

The inverse problem uses the observed N-value corrected for multiple scattering and refraction with the value for the smallest solar zenith angle subtracted. Additionally, the integral of the retrieved ozone is constrained to have the value of the observed total ozone. The retrieval algorithm uses the optimal

estimation method of Strand and Westwater [1968] as formulated by Rodgers [1976]. At the nth iteration, retrieval **x**n+1 is obtained from **x**n by the following expression:

$$\mathbf{x}n+1 = \mathbf{x}_A + [\mathbf{S}x\text{-}1 + \mathbf{K}n\mathrm{T}\mathbf{S}\mathrm{E}\text{-}1\mathbf{K}n]\text{-}1\mathbf{K}n\mathrm{T}\mathbf{S}\mathrm{E}^{-1}[(\mathbf{y}-\mathbf{y}n)-\mathbf{K}n(\mathbf{x}\mathrm{A}-\mathbf{x}n)]$$

where $\mathbf{x}_A$ is the *a priori* ozone profile, $\mathbf{S}_\mathbf{x}$ is the covariance uncertainty matrix for the first guess profile, **S**E is the error covariance matrix for the measurements, **y***n* is the vector of calculated observations for **x***n*, **K***n* is the averaging kernel, and the superscript T represents matrix transposition.

The *a priori* profiles are significantly improved over the 1964 approximations. In layers above 5, these current *a priori* profiles are sinusoidal functions of Julian day and latitude in six latitude bands. The *a priori* amounts below layer 4 are quadratic functions of the measured total ozone amount only derived from ozonesonde climatology. The amounts in layers 4 and 5 result from a cubic fit to upper and lower layers.

1.7.6 Error and Retrieval Characterisation

The contributions to the error are from measurement errors, smoothing errors, forward-model errors, and inverse-model errors. The retrievals are also sensitive to the amount and distribution of atmospheric aerosols. Mateer and DeLuisi [1992] characterise the retrieval errors from various forward model inputs and from the inverse model. For a discussion of aerosol-induced errors see DeLuisi [1979] and Newchurch and Cunnold [1994]. The averaging kernels for an ozone column of 350 DU and a uniform Sx matrix appear in Figure 1.27. Although all Umkehr observations of all qualities are available in the WOUDC database, for comparative studies, one should restrict the population of retrievals to those with small differences between observed and fitted N values (rmsres 1.3 N units) and with convergence criteria dfrms 0.01 in the inversion process. These criteria affect profile comparisons more than trend studies.

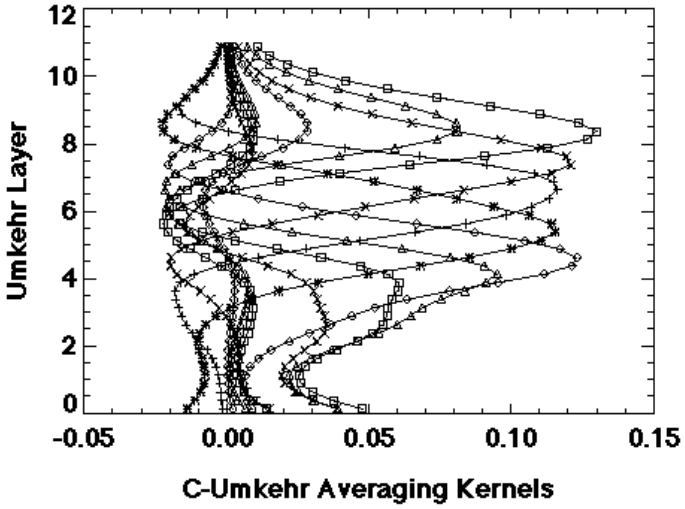


Figure 1.27. Averaging kernels from the 1992 C-Umkehr inversion algorithm for an ozone profile with 350 DU at 45oN using the uniform Sx covariance matrix. Ordinate labels represent the bottom of the designated layer. Symbols corresponding to retrieval layers are: layer 10, diamonds peaking at layer 8; layer 9, triangles peaking at layer 8; layer 8, squares peaking at layer 8; layer 7, crosses peaking at layer 7; layer 6, plusses peaking at layer 6; layer 5, asterisks peaking at layer 5; layer 4, diamonds peaking at

layer 4; layer 3, triangles peaking at the top of layer 3; layer 2, squares peaking in layer 3, and layer 1, crosses peaking in layer 2.

Measurement errors

For a sufficiently large number of retrievals, one can calculate the error covariance matrix, **S**M given the measurement error covariance, **S**E, which adopts a variance of 1 (N-unit)2 for all zenith angles and 3 DU for the total ozone measurement. The result of this calculation indicates that above layer 9 there is no retrievable information. In the middle layers, measurement errors contribute 3-6% uncertainties to the profiles. In the lowest layers (3,2, and 1), retrieval errors increase.

Smoothing errors

Smoothing errors include profile structure that is either not seen or poorly seen in the Umkehr measurement. The smoothing error for layers 4 through 8 is approximately 10%. For layers below layer 4, it increases to 15% and for layers above 8 to 25%.

Forward model errors

Forward model errors include quadrature errors, parameter errors, and others such as the lack of inclusion of the effects of SO2 absorption, the failure to include multiple scattering in the partial derivative calculation, and the effect of temperature profile errors. Errors due to quadrature vary from 0.0 to 1.1% as a function of layer. The relevant parameters are the ozone absorption and Rayleigh scattering coefficients. Ozone errors due to reasonable uncertainties in these parameters range from 0.2% to 2.6% all functions of altitude. Errors due to heavy SO2 pollution (10 DU) result in ozone errors of approximately 3% in layers 4 through 8. Errors due to multiple scattering neglect in the partial derivative are generally less than 0.5% in layers 4-8 [Mateer and DeLuisi, 1992]; although in some cases the error maybe somewhat larger. Errors due to the assumption of a single temperature profile are less than ±3% and are strong functions of layer with a minimum less than 1% in layer 5 and increasing both above and below that layer.

Inverse model errors

Three areas that contribute to the inverse model errors are errors in the SE matrix, errors in the Sx matrix, and errors in the *a priori* profile. The first two jointly contribute about 1-4% uncertainty to the retrievals for levels between 4 and 8. The uncertainty is larger at levels above 8 and below 4.

Aerosol errors

The 1992 algorithm continues to be sensitive to atmospheric aerosols, particularly stratospheric aerosols. This aerosol interference is an optical effect caused by the scattering of light by the aerosols in the Junge aerosol layer (~20 km) on radiation that was initially scattered into the nadir by the atmosphere at much higher altitudes (e.g., 40 km for layer 8) [DeLuisi, 1979]. Ozone errors due to total aerosol optical depths of 0.016, with 0.012 of that atmospheric total in the stratosphere, are largest in layer 9. In layer 8, the error is approximately 4% with smaller errors in the other layers. For this reason, in this report we use only Umkehr with stratospheric aerosol optical depths less than 0.02 where the correction is less than approximately 6% and the uncertainty in that correction is on the order 1% in ozone amounts. This restriction results in omitting approximately one year of observations after both the eruptions of El Chichon and Mount Pinatubo but ensures minimal residual error due to aerosol interference.

A priori influence on trends

The determination of trends in layers below 5 is affected by a potential trend in the *a priori* profiles [Mateer *et al*., 1996]. Using differences in retrievals between synthetic and actual profiles the authors conclude that a time-dependent *a priori* profile would influence the calculated trend in layers 1-3 and somewhat in layer 4. Trends in layers above 4, however, are not affected by a potential trend in the *a priori* profile. The problem is moderated by using a time-dependent *a priori* ozone profile, combining layers 1, 2, and 3 into a single layer, and layers 8, 9, and 10 into a single layer. This finding is confirmed by the study of Petropavlovskikh *et al*. [1996a]. In this report, we take a somewhat more conservative approach and combine layers 1-4 into a single layer, but also calculate layers 4-8 separately. We then combine layers 8, 9, and 10 into a single 8+ layer. Alternative approaches could combine layers 1-3, 1-5, or 9-10.

The effect of appropriate errors on ozone trend calculations is summarised in Table 1.12. In layers 8 and 8+, the rss estimated error due to all sources on trend calculations is 0.23±0.21%/year diminishing to 0.2±0.2 %/year in layers 6, 5, and 4. The largest instrumental sources of uncertainty regarding trend estimation are the Umkehr calibration uncertainty and the aerosol optical interference correction.

Error Source	Ozone Trend Error %/ year	Uncertainty %/year
Calibration		
Total ozone	0.02	0.04
Umkehr	0.2	0.2
Smoothing Error		
All sources	< 0.01	< 0.01
Forward Model		
Quadrature	0.0	0.0
Spectroscopic parameters	0.0	0.0
SO2 absorption	0.01	0.01
Multiple scattering	0.0	0.0
Temperature (per K)	0.01	0.01
Inverse Model		
SE matrix	< 0.01	< 0.01
Sx matrix	< 0.01	< 0.01
a priori profiles		
layer4	< 0.01	0.0
layers 5-8	0.0	0.0
Aerosol Optical Interference		
layers 8, 8+	0.01	0.03
layers 7, 4-	0.07	0.02

layer 6	0.03	0.01
layers 5, 4	0.0	0.0
RSS Total		
layers 8, 8+	0.23	0.21
layers 7,4-	0.21	0.2
layers 6, 5, 4	0.20	0.2

Table 1.12. Error sources in trends calculated from Dobson/Umkehr ozone profile retrievals.

1.7.7 Outstanding Problems and Future Plans

Correcting Umkehr observations for the high aerosol loads within 1 year after a major eruption requires sophisticated analysis. For example, DeLuisi *et al*. [1996] performed an extensive error analysis of the aerosol correction to the Umkehr ozone retrievals that included effects of size distribution errors, vertical profile errors, and errors using climatological profiles of ozone and aerosols. The result of this work clearly showed that the errors possess a non-linear dependence on the aerosol optical thickness. A linear multivariate regression analysis was conducted to examine the relationship between the aerosol load in one of the Umkehr layers to the retrieved ozone errors in all layers 1 through 10. The magnitude and sign of the contributions (absolute error) vary according to the amount of aerosol in the layer and the altitude location of the layer. The contribution of the aerosol in the layer with the maximum load (usually layer 4) produces a strong effect in all layers. For long-term trend studies, however, one can accurately correct the Umkehr observations in the linear part of the curve that occurs during smaller aerosol loadings and omit the enhanced aerosol periods (approximately 1 year after major eruptions).

Some questions remain concerning the accuracy of the Umkehr forward model, especially with regard to the treatment of multiple scattering. Petropavlovskikh *et al*. [1996b] compared various radiative transfer codes to evaluate the accuracy of the Dave scalar code used in the past to model the aerosol effect on the Umkehr retrieved ozone profile. The differences between the codes was found to be a function of the solar zenith angle of the calculation. They found that the calculated ozone errors are sensitive to the radiative transfer code adopted. The authors conclude that there are non-negligible differences between the results from the different codes and further study is warranted. Recent comparisons of Umkehr to SAGE profiles [Newchurch *et al*., 1997] indicates a discrepancy of approximately 5% between the Umkehr ozone columns and the SAGE columns (SAGE higher than Umkehr). This bias exhibits an altitude structure increasing from 0% in layer 4 to 15% in layer 8.

Under the auspices of the European Commission, the Umkehr data record from most stations world-wide is being reviewed to improve the retrieved profiles in several areas including the following: 1) instrumental calibration and compatibility with total ozone column measurements, 2) sensitivity to the temperature dependence of the ozone absorption cross sections, 3) *a priori* sensitivities, 4) aerosol corrections using satellite and ground-based aerosol measurements, and 5) comparisons to other ozone measurements. This work is an ambitious review and potential modification of essentially all important aspects of the Umkehr observations and ozone profile retrieval. It is currently in its early stages and has not affected the Umkehr database for the results reported here.

In summary, Dobson instruments produce long-term, well-calibrated measurements of ozone profiles in coarse vertical resolution in the stratosphere. The current retrieval uses the maximum likelihood formulation of Rodgers [1976], *a priori* ozone profiles as a function of latitude and Julian day, and the temperature-dependent ozone absorption coefficients of Bass and Paur, [1985]. It is a significant improvement over the previous inversion algorithm. The aerosol effect on the ozone retrievals can be accounted for at the expense of approximately one year of observations following major eruptions. The Brewer network comprises well-calibrated instruments some of which now have significant records that should be analysed for ozone trends after adequate scientific attention is given to the inversion algorithm and data validation.

1.8. Ozonesondes

Ozonesondes are small, light weight, balloon-borne instruments capable of making measurements from the surface to near 35 km in altitude. All of the ozonesonde data discussed in this report is from instrumentation using the reaction of ozone with potassium iodide in an aqueous solution as the method of detection. Three basic types of ozonesondes were used to obtain data for this assessment. These three sonde types are, however, used by a number of individual sounding programs. In practice there are important differences in the way the individual programs use an instrument of a given type. In this section of the chapter we will characterise each sonde type and will also provide relevant information from the individual programs that has a bearing on the final data product. The three sonde types considered here are the Brewer-Mast (BM), the electrochemical concentration cell (ECC), and the Japanese ozonesonde (KC).

1.8.1 Basic Measurement Description

Principle of Operation

The ozonesonde consists of a pump and ozone sensing cell coupled to a standard meteorological radiosonde through an electronics interface. The information from the ozonesonde is telemetered to the ground through the radiosonde transmitter. The parameters normally measured are the ozone concentration, ambient air pressure, temperature, humidity, and, in some cases, the wind direction and speed. Each sounding is made with an individual disposable instrument. The aqueous chemical sensing element is based on the redox reaction:

$$2KI + O_3 + H_2O \rightarrow I_2 + O_2 + 2KOH \quad 1.8.1$$

The method of detecting the free iodine, I_2, is an important characteristic of the three types of electrochemical cells considered in this report. Ambient air is continuously forced into the sensing cell by a battery driven sampling pump. An electrical current is generated proportional to the mass flow rate of ozone through the cell. By knowing the volume flow rate and temperature, the electrical current can be converted to an ozone concentration under the assumption that the ozone reaction with potassium iodide is quantitatively known.

Brewer-Mast Ozonesonde

The Brewer-Mast sonde evolved from the Oxford-Kew ozonesonde developed by Brewer and Milford (1960). The Brewer-Milford type ozone sensor consists of a single electrochemical cell with a silver anode and platinum cathode immersed in an alkaline potassium iodide solution. A polarising potential of 0.42 V is applied between the electrodes such that no current will flow unless free iodine is present. In operation, ozone in the sampled ambient air is forced through the sensing solution in the electrochemical cell (bubbler) to produce free iodine according to redox reaction (1.8.1). Transported by the stirring action of the bubbler, the iodine in contact with the platinum cathode is converted back to iodide ions by the uptake of two electrons per molecule of iodine.

$$I_2 + 2e - Pt \rightarrow 2I^- \quad \text{[cathode reaction]} \quad 1.8.2$$

Correspondingly at the anode two electrons are released through the ionisation of two silver atoms

$$2Ag \rightarrow 2Ag^+ + 2e \quad \text{[anode reaction]} \quad 1.8.3$$

to form nearly insoluble silver iodide, which prevents iodine from entering the solution. In principle, each ozone molecule entering the sensor causes a current of two electrons to flow through the external circuit. The reaction chamber (bubbler) is made of Plexiglas and contains a cylindrical platinum mesh cathode (~6 cm^2) and a thin silver wire as anode. The bubbler is filled with 2 ml of neutrally buffered aqueous solution of potassium iodide (0.1%). The electrically driven gas sampling pump is mounted at the side of the bubbler and forces about 220 sccm/min. of ambient air through the bubbler. The sonde is protected by a Styrofoam flight box. The Brewer-Mast sondes were manufactured by the Mast Keystone Corporation (Reno, Nevada, USA) and its predecessor the Mast Development Corporation.

ECC Ozonesonde

The ECC ozonesonde was developed by Komhyr (1969, 1971). The ECC ozone sensor is an electrochemical cell consisting of two half cells, made of Teflon, which serve as cathode and anode chamber, respectively. Both half cells contain a platinum mesh serving as electrodes. They are immersed in KI solution of different concentrations. The two chambers are linked together by an ion bridge in order to provide an ion pathway and to prevent mixing of the cathode and anode electrolytes.

The ECC does not require an external electrical potential. This is in contrast to the Brewer-Milford type of electrochemical ozone sensor (Brewer and Milford, 1960). The ECC gets its driving electromotive force from the difference in the concentration of the KI solution in the cathode and anode chamber, 0.06 Mol/l (1%KI) and "8.0 Mol/l (saturated KI) respectively. A non-reactive gas sampling pump (Komhyr, 1967), made of Teflon, forces ozone (flow ~220 sccm/min.) in ambient air through the cathode cell with the lower concentration of KI solution causing an increase of free iodine (I_2) according to the redox reaction (1.8.1). At the surface of the Pt cathode, I_2 will be converted to I- through the uptake of two electrons (1.8.2), while at the anode surface, I- is converted to I_2 through the release of two electrons, such that the overall cell reaction is

$$3I^- + I_2 \rightarrow I_3^- + 2I^- \quad 1.8.4$$

Thus one ozone molecule causes two electrons to flow in the external circuit. This electrical current is directly related to the uptake rate of ozone in the cathode chamber. The sonde is flown in a polystyrene protective box. Five different models of the ECC type have been used: the SPC 3A, 4A, 5A, 6A models manufactured by Science Pump Corporation (Camden, New Jersey, USA) and the 1Z model manufactured by EN-SCI Corporation (Boulder, Colorado, USA).

Japanese Ozonesonde: KC-68 and RSII-KC79

This ozonesonde type is based on a modified version of the carbon-iodine ozone sensor (Komhyr, 1969). The ozone sensor is an electrochemical cell containing a platinum gauze as cathode and an activated carbon anode immersed in an aqueous neutral potassium iodide/potassium bromide solution. Ozone in ambient air is forced through the sensing solution generating free iodine molecules (I_2) according to redox reaction (1.8.1). At the Pt cathode the iodine is reconverted into iodide/bromide by the uptake of two electrons, while correspondingly at the activated carbon anode under the release of two electrons the following reaction takes place:

$$C + 2OH^- \rightarrow CO + H_2O + 2e \quad 1.8.5$$

One ozone molecule produces an electrical current of two electrons in the external circuit. The gas sampling pump and the electrochemical cell are made of methacrylate resin. The pump flow rate is about 400 sccm/min. The sonde is flown in a Styrofoam flight box. The RSII-KC70 is manufactured by Meisei

Electric Company, Ibaraki, Japan.

1.8.2 Observational Program and Sampling Strategy

Ozonesonde observations are carried out by a number of different operating agencies with varying sampling protocols, covering a variety of observational periods. Most of the locations have observations up through the 1996 period covered by the assessment and have done soundings for 10 years or more. Information on location, period of observation, operating agency, ozonesonde type, radiosonde type, frequency of observations, and altitude resolution is summarised in Table 1.13. The stations are grouped geographically and by operating agency. Several factors that deal with the processing algorithm discussed in the next section are also included in the table.

At four Canadian sites with long records, the type of ozonesonde was changed in 1980 from the BM to the ECC. At present the discontinuity (primarily in the troposphere) introduced by this change has not been satisfactorily characterised from an instrumental (as opposed to statistical) point of view. The difference in the troposphere in the sensitivity to ozone between the BM and ECC sondes is discussed further in Chapter 2 in conjunction with intercomparison results.

The standard practice is a program of weekly balloon launches. Since soundings are occasionally missed or the instrument fails, the actual number of profiles is somewhat less. During portions of the record at most of the sites there are only scattered soundings because of a reduced or intermittent schedule of launches. The three European sites (Uccle, Hohenpeissenberg, and Payerne) have carried out a more ambitious sounding program with 2-3 launches per week over a period of 20 years or more. All but a few of the sites have major gaps in their record, which range from a few months to more than a year.

Table 1.13 Station and Sounding Program Characteristics

					Frequency						
		Sonde		Operating	(time) of	Altitude	Radiosonde/	Normalization	Pump	Background	Box/Pump
Latitude	Longitude	Type	Period of Operation	Agency	Sounding	Resolution	Interface	To Total	Correction	Signal	Temperature
Resolute, Canada											
74.72° N	94.98° W	BM	Jan. 1966-Nov. 1979	AES/Canada	1/week (12Z)	Mandatory and significant levels	VIZ	Yes (avg. corr. when none)	Standard	None	Constant
		ECC	Dec. 1979-present	AES/Canada	1/week (12Z)		VIZ 1979-93 Vaisala 1994-	Yes (data on B-P scale)	Komhyr (1986) Komhyr (1995)	Declining	Measured box
Churchill, Canada											
58.75° N	94.07° W	BM	Oct. 1973-Sept. 1979	AES/Canada	1/week (12Z)	Mandatory and significant levels	VIZ	Yes (avg. corr. when none)	Standard	None	Constant
		ECC	Sept. 1979-present	AES/Canada	1/week (12Z)		VIZ 1979-93 Vaisala 1994-	Yes (data on B-P scale)	Komhyr (1986) Komhyr (1995)	Declining	Measured box
Edmonton, Canada											
53.55°	114.10°	BM	Oct. 1970-	AES/Canada	1/week	Mandatory	VIZ	Yes (avg. corr.	Standard	None	Constant

N	W		Aug. 1979		(12Z)	and significant levels		when none)			
		ECC	May 1978-present	AES/Canada	1/week (12Z)		VIZ 1979-93 Vaisala 1994-	Yes (data on B-P scale)	Komhyr (1986) Komhyr (1995)	Declining	Measured box
Goose, Canada											
53.32° N	60.30° W	BM	June 1969-Aug. 1980	AES/Canada	1/week (12Z)	Mandatory and significant levels	VIZ	Yes (avg. corr. when none)	Standard	None	Constant
		ECC	Sept. 1980-present	AES/Canada	1/week (12Z)		VIZ 1979-93 Vaisala 1994-	Yes (data on B-P scale)	Komhyr (1986) Komhyr (1995)	Declining	Measured box
Uccle, Belgium											
50.80° N	4.35° E	BM	Nov. 1966-present Occasional breaks Little data Feb. 1983-Jan 1985	Belgium Meteorological Institute	3/week (10Z)	Mandatory and significant levels	VIZ to 1990-Vaisala 1990-	Yes (all data on B-P scale	Standard to 1989 From 1989 based on measured	None	Average from measured. Measured from 1990
Hohenpeissenberg, Germany											
47.80° N	11.02° E	BM	Nov. 1966-present	German Weather Service	2/week Nov.-April 3/week	Mandatory and significant levels	VIZ with hypsometer Vaisala from 1994	Yes (after 1968)	Standard	None	Constant
Payerne, Switzerland											
46.82° N	6.95° E	BM	Nov. 1966-present	Swiss Meteorological Institute	3/week (11Z)	Fixed pressure levels (~.25 km)	VIZ to 1980 Swiss (2 types)	Yes (to Arosa total)	Standard	None	Constant
Haute Provence (OHP), France											
43.93° N	5.70° E	ECC	1989-present		1/week		Vaisala/TMZX	Yes		Constant	Measured
Boulder, USA											
40.03° N	105.07° W	ECC	March 1979-present. Missing Jan. 1990-June 1991	NOAA/CMDL	1/month to 1985 1/week 1985-	0.25 km	Vaisala/TMAX	Yes (data on B-P scale)	Average (Johnson, 1997) 1991 meas.	Constant	Measured box to 1990. Measured pump 1991-
Hilo, USA											
19.72° N	155.07° W	ECC	Sept. 1982-present	NOAA/CMDL	1/week	0.25 km	Vaisala/TMAX	Yes (all data on B-P scale)	Average (Johnson, 1997) 1991 meas.	Declining	Measured box to June 1991 Measured pump July 1991-
Sapporo, Japan											
43.05° N	141.33° E	KC	Dec. 1968-present	Japan Meteorological Agency			Japanese	Yes	Standard	Declining	Measured box
Tsukuba (Tateno), Japan											
36.05° N	140.10° E	KC	Jan. 1968-present	Japan Meteorological Agency			Japanese	Yes	Standard	Declining	Measured box
Kagoshima, Japan											

31.55° N	130.55° E	KC	May 1968-present	Japan Meteorological Agency			Japanese	Yes	Standard	Declining	Measured box
Wallops Island, USA											
37.93° N	75.48° W	ECC	July 1969-present	NASA/WFF	2/month	Mandatory and significant levels	VIZ to 1995 VIZ digital 1995-	WOUDC data renormalized to B-P scale	Measured (Torres, 1981)	Constant	Measured box
Lauder, NZ											
45.03° S	169.68° E	ECC	Aug. 1986-present	NIWA	1/week Aug-Nov 2/week	0.10 km	VIZ to 1990 Vaisala/TMAX 1990-	No	(Johnson, 1997) average	Declining to 1990	Measured box

1.8.3 Calculation of the ozone profile

All three of the ozonesonde types use the following equation to convert the sensor output (current) into the partial pressure of ozone (P_{O_3}) in millipascals:

$$P_{O_3} = 4.308 \times 10^{-4} (i - i_b) t T \quad 1.8.6$$

where *t* is the time in seconds to pump 100 ml of air at a temperature *T* (in K), *i* is the output current (in microamperes) and *ib* is the background current (the output with ozone free air flowing through the sensor). This equation is valid if no ozone is destroyed in the air sampling system, each ozone molecule produces two iodine molecules, and there are no interfering reactions in the solution from minor constituents other than ozone in the air sample.

1.8.4 Factors influencing the determination of trends

Three primary types of factors may influence the determination of an ozone profile and each of these types can lead to changes in instrument performance that may vary with time and potentially introduce a bias into the measurement time series. These three types include algorithmic factors, operating procedures and instrument manufacture, and environmental conditions. In some cases these factors may fall into more than one category, but are treated here under a single factor.

1.8.4.1. Algorithmic Factors

These factors deal primarily with the computation of the ozone profile from the measured sensor output and include: 1) The determination of the correction for loss of pumping efficiency at reduced atmospheric pressure; 2) Subtraction of the "background" signal (reading for zero ozone) from each reading; 3) Application of the pump temperature; 4) Response time of the sensor; and 5) "Normalisation" of the integrated profile amount to an independently measured total column ozone amount.

Pump Efficiency Correction

The standard procedure is to use an average pump efficiency correction for each type of instrument. This correction is usually determined at the time of development of the instrument and not periodically checked. Recent tests on the BM sonde (DeBacker *et al.*, 1998) show that the correction is significantly larger above 20 km than previously determined. The pump efficiency has also probably deteriorated over time as a result of changes in manufacturing practices (DeBacker *et al.*, 1998). If the pump efficiency changes with time but the correction remains unchanged (the usual case to this point) the profile may be affected not only at the altitudes where this correction is significant (>25 km), but also at other altitudes as well because of the procedure of correction to the independently measured total ozone amount. At

Uccle corrections have been applied for likely changes in pump performance.

For the ECC sonde the correction for the loss of air pump efficiency is done according to the data in Komhyr (1986) or Komhyr *et al*. (1995). The latter is adopted by the Vaisala (a large provider of ECC sondes) manual and gives about 3% more ozone at 5 hPa and 1% more at 10 hPa than Komhyr (1986). Recent determinations of the ECC pump efficiency correction suggest that previous measurements underestimated the actual efficiency loss (Johnson *et al*., 1998). It does not appear that there have been large changes over time in the performance of the ECC pumps. In 1995 the Canadian stations changed the correction, but this was a small change that likely has not affected the record.

The Japanese ozonesonde has used the same pump and average corrections over the period of measurements. The pump efficiency correction measurements for the various sondes are summarised in Table 1.14.

Pressure hPa	BM standard	BM measured at Uccle	Japanese	ECC Torres 1981	ECC Komhyr 1986	ECC Komhyr *et al*. 1995	ECC NOAA / CMDL Johnson *et al*. 1998	ECC measured at Uccle
100	1.01	1.06	1.02	1.000	1.011	1.007	1.027	1.04
50	1.03	1.085	1.04	1.013	1.015	1.018	1.046	1.06
30	1.065	1.125	1.07	1.025	1.022	1.029	1.063	1.065
20	1.09	1.17	1.11	1.035	1.032	1.041	1.081	1.075
10	1.17	1.27	1.25	1.063	1.055	1.066	1.133	1.10
7	1.23		1.40	1.085	1.072	1.087	1.183	
5			1.66	1.105	1.092	1.125	1.240	

Table 1.14. Correction for Efficiency Loss in the Ozonesonde Pump

Background Current Measurements

Measurement of the ozonesonde signal when no ozone is present is not routinely done for BM sondes, where this signal is assumed to be zero except at Payerne, where in recent years this is checked but is not applied as part of the processing.

For the ECC and Japanese sondes the background current is measured for each sounding prior to the balloon launch. At all stations except the NOAA/Climate Modelling and Diagnostics Laboratory (CMDL) stations and Lauder, this offset gradually declines with the decreasing pressure and becomes insignificant in the stratosphere. The method of removal of the background signal is of some importance in the troposphere especially in the tropics where this signal can be a significant fraction of the measured ozone signal. The application of this offset primarily affects trends when the procedure changes. The one station where this is likely to have occurred is Hilo, Hawaii, where a correction procedure has been implemented.

Pump/Box Temperature Measurement

In order to correct for changes in the air volume flow due to temperature changes (eq. 1.8.6), the temperature at the pump or in the instrument enclosure (box) should be measured. The practice for the BM sondes is to assume that the pump temperature is constant. Only at Uccle is it measured, and the temperature is found to decline by 15-20°C over the course of the flight, which leads to a 5% error in ozone computation for the highest altitudes.

In the ECC and Japanese sondes either the box or pump temperature is measured. There is normally about a 2-5°C difference (pump warmer in most configurations) at the top of the flight. Several of the ECC stations have changed from a box to a pump temperature measurement, but the change is likely not significant.

Sensor Response Time

For the ECC sonde (the other types are similar) the time for a 90% response to a step change is approximately 50 s. This translates into a height resolution of about 150 m and a slight displacement of the profile in the vertical relative to a measurement made with a faster responding sensor such as a UV photometer. The usual practice is to make no correction for the lag in response. At Uccle, where both ascent and descent profiles are measured, a deconvolution that accounts for the response is used so that the ascent data can be compared with the descent data, which is obtained from the rapidly descending instrument after balloon burst.

Normalisation to Measured Total Ozone

Most of the stations making ozonesonde measurements over a long period of time adjust the integrated ozone amount from the profile to an independently measured column amount. This procedure requires a homogeneous total ozone data set and a consistent method for estimating the ozone amount above the profile ceiling level (which is usually between 15-7 hPa). The total ozone scale for the Dobson spectrophotometer was adjusted in 1992 to account for a re-determination of the ozone cross-section. This requires a reduction of ozone amounts by 2.6% (data scaled by 0.9743). Reprocessed data for Hohenpeissenberg, Uccle, the NOAA/CMDL sites, and Japanese stations have applied this correction. The Canadian stations, which use the Brewer spectrometer for total ozone, were already using the Bass-Paur scale for the ECC series. The BM data were adjusted to the Bass-Paur scale for this report. For Wallops Island data, a homogeneous total ozone data set was applied to an archived data set, which was not normalised. At Lauder the data were not normalised.

The standard procedure for computing ozone above the profile top (the residual ozone) is to assume a constant mixing ratio equal to the mixing ratio at the top of the profile. Because the profile is the most uncertain near the top, due to uncertainties in the pumping efficiency, another procedure is to use a climatology based on SBUV profiles to get the residual (McPeters *et al.*, 1997). This was done for the NOAA/CMDL and Wallops Island data.

Before a profile is normalised to the total ozone, it is important to apply all altitude dependent corrections. Since a single factor is used in correcting the integrated profile to the total ozone an uncorrected error in one part of the profile will be translated into errors in other altitudes. This is illustrated in the trend computation at Uccle for various corrections applied to the profile (fig. 1.28). Even though the SO_2 correction (discussed later) and pump corrections mainly apply in the lower and upper portions of the profile respectively, they affect the entire profile. At Uccle both of these factors

have changed with time but have been accounted for in the final data set.

The re-evaluated data set of ozone soundings at Uccle from 1969 to 1996 was used for a sensitivity study of instrumental effects on calculated trends. Two effects were explored, the potential effect of the pump correction, and that of SO_2 interference. The results are shown in Figure 1.28. The calculated trends are based on a least squares analysis of de-seasonalised data. With the standard algorithm and corrections for calculation of ozone profiles for Brewer Mast sondes described by Claude *et al*. (1987), the ozone trend is -1.4%/year in the middle troposphere.

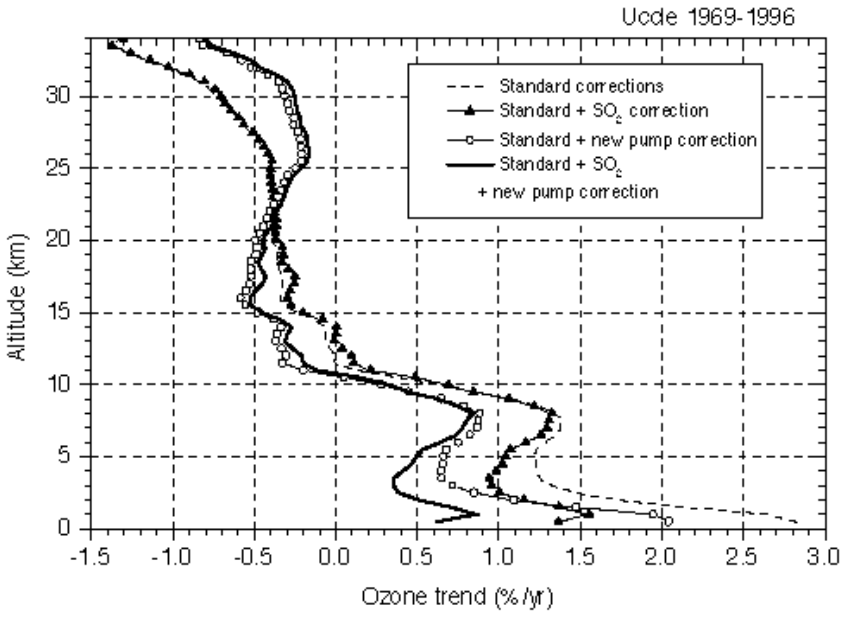


Figure 1.28. Trends at Uccle computed after having applied different corrections to the data.

Before re-evaluation of the Uccle time series a jump was observed in 1989, resulting in too high ozone values in the troposphere and too low values above the ozone maximum (~22 km). After application of the pump efficiency profile that was measured in the laboratory (which shows much higher pump corrections than adopted in the standard profile) this jump was largely removed. Since each ozone profile is normalised with respect to the measured ozone column, the effect of a new pump correction profile propagates to the entire profile, although it is largest at the upper altitudes. Consequently the calculated ozone trends are affected for the entire profile, with trends of -0.8%/year at 33 km, -0.5 %/year in the lower stratosphere, and 0.6-0.8%/year in the middle troposphere.

The SO_2 levels in the lower troposphere above the urban area of Brussels decreased rapidly in the 1970s and to a lesser extent in the 1980s. After applying a correction for the SO_2 interference as described by De Muer and De Backer (1994a), the ozone trends in the boundary layer are reduced by about half. Tropospheric trends with both corrections applied are 0.35-0.85%/year. Changes in boundary layer ozone also propagate to trends at other altitudes because of the normalisation procedure, but changes in the upper tropospheric and stratospheric trends are small. The corrections applied to the data here have little effect on the trends calculated near the mean level of the ozone maximum. Trends should be most

reliable around that level.

The issue of SO_2 interference is likely largest for Uccle of all the sonde stations, since the measurements are made in an urban area. SO_2 may have affected boundary layer values for ozone for Hohenpeissenberg, Wallops Island, and Sapporo, as discussed in Logan (1994), but probably to a lesser extent. No attempt has been made to account for potential SO_2 effects at other stations.

At some stations, primarily at high latitude, a much smaller number of independent total ozone measurements are available. In these cases an average correction is applied (the BM sondes) or the profile is left un-normalised (ECC and Japanese sondes). The profile also cannot be normalised to the total ozone if the balloon does not reach sufficient altitudes so that the integrated profile from the sonde encompasses most of the ozone layer. This cut-off altitude is between 15-20 hPa for most stations. The use of the climatology for computing the residual would probably allow a greater number of the profiles to be successfully normalised.

The comparison between the integrated sonde total ozone and the measured total is also a factor used to determine the quality of the measured profile. Profiles with correction factors outside prescribed limits are usually rejected. For ECC and Japanese sondes these limits are often set at 0.80 £ C.F. £ 1.20 and for BM sondes 0.90 £ C.F. £ 1.5. The variation of this correction factor for various locations and types of sondes is shown in Figure 1.29. Periods when the correction factor is flat indicate that an average correction was applied (primarily the Canadian sites). The correction factor should not be the only criterion used for judging of profile quality, though it is often the only one available to the user of archived data. By including the profile in the archived data set, the originator of the data has usually judged that the profile is of acceptable quality.

Click here to download Figures 1.29 (a), (b) and (c) by anonymous ftp.

1.8.4.2. Procedural and Environmental Factors Influencing Trends

Several factors which may influence trends are considered within this group. They include: (1) changes in operating procedures, (2) changes in instrument type, (3) changes in ozonesonde construction, (4) radiosonde changes and error in radiosonde measurements, and (5) boundary layer pollution.

Changes in Operating Procedures

For BM sondes the procedures described by Claude *et al*. (1987) are generally followed with various modifications. These procedures were not implemented uniformly in time at all three BM sites, however. All of the BM stations have improved their pre-flight conditioning procedures including better cleaning of the sondes and longer exposure to sensing solutions. The primary influence of such changes over time is likely to be in the troposphere. At Uccle the effect of improved techniques was evident and a correction has been applied to the data.

Procedures for preparing and operating ECC sondes are described by Komhyr (1986) and these are generally incorporated into the Vaisala system manual. Many of the deviations from these procedures are related to the algorithmic factors described in the previous section. Recently, changes to the sensing solution have been introduced at Lauder and Wallops Island. These have occurred recently enough that they will have little influence on the trend determination for this report.

A change in launching time at sites with a large diurnal variation may influence the trend at the lowest altitudes. At Edmonton and Payerne such an effect has been found. At Payerne a correction has been applied based on nearby continuous surface measurements. At many sites the frequency of measurements

has varied greatly over the period of observation with most stations having significant gaps in the record (see Figure 1.29).

Changes in Instrument Type

Of the records being considered for this report, the only instrument type change was in 1979 from the BM to ECC sonde at four Canadian stations. The BM sondes were not prepared according to the procedures of Claude *et al*. (1987) but used an earlier, less rigorous procedure with no special cleaning or calibration. Because of the shift in tropospheric ozone amounts at the time of the switch from the BM to the ECC sonde, the combined tropospheric record has not been deemed suitable for trend determinations by the program investigators (Tarasick *et al*., 1995).

Changes in Ozonesonde Construction

Modifications have been made in the manufacture of all three types of sondes. Most of these seem to be relatively minor in terms of the effect on the data record. A temporary deterioration in quality of the BM sonde was noted with the relocation of the factory in 1989. There are strong indications that pump performance of the BM sondes has also deteriorated with time based on recent pump efficiency measurements made on older and newer sondes at Uccle (DeMuer, private communication, 1997). At Uccle this change has been accounted for in the post-flight processing. At Hohenpeissenberg and Payerne some modifications in the operating procedures seem to have compensated for the change in instrument quality.

The ECC sonde has gone through several model changes. Most ECC records begin with the 3A sonde. The 4A sonde, introduced in the early 1980s, had a redesigned pump that was more efficient at higher altitudes. Differing curves for the pump efficiency correction for the 4A sonde dependent on the amount of sensing solution (affecting the head pressure) were recommended. The 5A sonde is identical to the 4A except it is designed to be used with a digital electronic interface. Recently a 6A sonde has been introduced with a thermistor for measuring the pump temperature mounted in the pump. In 1994 a second manufacturer began supplying sondes. These are nearly identical in construction to sondes from the competing manufacturer.

The Japanese sonde was modified in 1979 to include the radiosonde and ozonesonde in a single integrated package. The construction of the sensor may also have changed at this time from a double- to single-cell design. No apparent changes in sonde performances have been noted with these changes.

Radiosonde Changes and Errors

The radiosonde to which the ozonesonde is coupled is the source of error in the ozone profile determination through errors in the pressure and temperature determination, which affect the computation of ozone mixing ratio and altitude. Bias errors may appear when there is a switch from one radiosonde type to another. Such a switch has occurred at Uccle, Payerne, Hohenpeissenberg, the NOAA/CMDL stations, the Canadian stations, and Lauder. DeMuer and DeBacker (1994b) compared both the VIZ (without hypsometer) and Vaisala radiosonde to radar measured altitudes and found that at 30 km the Vaisala sonde gave altitudes ~300 m greater and the VIZ sonde ~800 m less than determined by radar. At 30 km the error in ozone partial pressure is about 7% for altitude deviations of 500 m. Switching from one radiosonde to another can thus introduce significant errors.

At several sites (NOAA/CMDL and Hohenpeissenberg) the VIZ radiosonde was used with a hypsometer that accurately measures the pressure at higher altitudes (Parsons *et al*., 1984). Only at the Japanese stations have the radiosonde sensors remained the same over the record. At Uccle and Payerne corrections have been made for the change in radiosonde type and errors in pressure measurement.

Boundary Layer Pollution

Sulphur dioxide is a negative interferent in electrochemical sondes of all types. Near urbanised areas and in the case of Hilo, Hawaii, the presence of an erupting volcano, SO_2 can have a significant impact on boundary layer ozone measurements. It usually takes a period of time (several minutes) for the accumulation of SO_2 to be purged from the sensor, which could affect measurements 1-2 km above the boundary layer itself. In the Uccle record, the changing SO_2 concentrations with time are accounted for to some extent (see figure 1.28), although complete SO_2 measurements are not available. At Hilo the data below ~650 hPa are probably not very reliable during periods of volcanic activity and air flow toward Hilo. At some places strong NO sources near the launch site may artificially depress the ozone amount which would make such measurements unrepresentative of regional conditions. In recent years the control of a number of pollution sources may artificially produce an increasing trend in ozone near the surface if the changing pollution conditions are not accounted for.

Estimates of the contribution to the trend error for a number of the factors discussed previously are summarised in Table 1.15 for each ozonesonde type at three altitudes. These are considered likely estimates for the ensemble of stations using a particular type of ozonesonde. The uncertainty could be much larger for a particular factor at an individual station.

Trend Uncertainty Factor		ECC	Brewer-Mast	Japanese
Normalisation to total ozone	5 km	0.07	0.07	0.07
	20 km	0.05	0.05	0.05
	30 km	0.07	0.07	0.07
Pump efficiency correction	5 km	0.02	0.20	0.04
	20 km	0.01	0.04	0.02
	30 km	0.03	0.16	0.06
Ozone zero signal	5 km	0.02	0.01	0.02
(background current)	20 km	0.00	0.00	0.00
	30 km	0.01	0.00	0.01
Pump temperature	5 km	0.01	0.01	0.01
	20 km	0.00	0.00	0.00
	30 km	0.02	0.01	0.01
Boundary layer/SO_2	5 km	0.02	0.05	0.01

pollution				
	20 km	0.00	0.01	0.01
	30 km	0.00	0.04	0.02
Pressure/altitude measurements	5 km	0.01	0.05	0.01
(radiosonde)	20 km	0.01	0.01	0.01
	30 km	0.04	0.04	0.02
Unknown procedural changes	5 km	0.03	0.03	0.03
	20 km	0.01	0.01	0.01
	30 km	0.02	0.03	0.03
Sum of the Uncertainties	5 km	0.18	0.42	0.20
	20 km	0.08	0.11	0.09
	30 km	0.19	0.31	0.20

Table 1.15**.** Measurement Uncertainties Influencing Trends in Ozone Measured by Ozonesondes (%/year), 1970-1996.

1.8.5. Characteristics of Individual Station Records

A number of stations in Europe and North America have records of sufficient length and quality that they should be the preferred choices for trend analysis. These are: Hohenpeissenberg (Germany), Payerne (Switzerland), Uccle (Belgium), the stations run by NASA and NOAA in the United States, and the Canadian stations from the beginning of the ECC record (1980). A significant number of stations (e.g., Haute Provence) have started making soundings too recently for reliable trends to be deduced. A number of others (e.g., Aspendale, Lindenberg) have switched sonde types within the past 5-10 years. There are well-known problems with attempts to derive trends from combined ECC and Brewer-Mast data sets [Tarasick *et al*., 1995]. Many stations have recently revised their total ozone records and these have been applied to the sonde time series where total ozone normalisation is currently practised. Even if the total ozone normalisation (correction) factor is not used to correct the sonde profile, it should be calculated, as it provides an excellent test for suspect data. Trends or discontinuities in normalisation factor time series are also cause for concern (see Figure 1.29).

The records discussed below have all been subjected to careful examination and re-evaluation by the scientists responsible for ensuring the quality of the data for each station. Correction factors have been examined for trends and discontinuities, and in most cases the total ozone record has also been re-evaluated. Corrections and changes and their effects (if known) are described, and known problems are noted.

Hohenpeissenberg

Regular soundings with Brewer-Mast ozonesondes were started at Hohenpeissenberg in November 1966. VIZ Hypsometer radiosondes (types 1072, 1192, 1292 and 1393) were used, with occasional use of other types, until the fall of 1994, when the station switched to the Vaisala RS80-30 NES. Major differences have been detected in altitude calculations due to differences in pressure measurements between the sondes. Quantifying and correcting of this break in the time series is in progress. The pre-flight preparation procedures of the sondes have been consistent and are in general as described in Claude *et al.*, [1987]. Some modifications of these procedures have been introduced fairly recently, including the use of potassium cyanide for cleaning the silver electrode and filling the sonde with KI solution at least 24 hours before launch. No systematic effects of these changes on sonde performance have been noted. Data reduction methods and pump corrections are consistent through the record. The recently re-evaluated Dobson total ozone time series (Köhler and Claude, 1996) has been applied to the ozonesonde record, with the use of Bass-Paur absorption coefficients. At the same time the entire sonde record was re-examined, and suspect data points, especially above 17 hPa, were removed (Claude *et al.*, 1996). The resulting time series shows less scatter in total ozone normalisation factors, but very similar trends.

Uccle

Regular soundings have been made since 1969 with Brewer-Mast ozonesondes. The pump temperature has been measured since 1990 when Vaisala (instead of VIZ) radiosondes were put into use. For soundings before 1990 a mean pump temperature profile (derived from the 90-91 data) is used. The altitude scale of profiles obtained before 1990 has been corrected for the mean difference between the VIZ and Vaisala radiosonde altitude measurements using wind-finding radar data for reference altitude measurements (De Muer and De Backer, 1994b). A correction for SO_2 levels in the lower troposphere over Brussels, based on ground-level SO_2 measurements, has been applied. This affects tropospheric values especially during the 1980s (De Muer and De Backer, 1994a). The change to the use of an electrostatic discharge ozoniser instead of a UV lamp for pre-conditioning of the sondes, beginning in October 1981, caused an apparent jump in tropospheric ozone values. This effect was corrected in the earlier data (De Backer, 1994) based on a comparison of ascent and descent profiles (after deconvolution of the profiles to take the time response of the sensor into account) during a 2-year period both before and after October 1981.

Special attention has been paid to the best possible application of the pump efficiency correction to the profile. Since 24 April 1996 the measured pump correction profile for each individual sonde is applied whenever it is available. Since 22 May 1992 when the absolute calibration of the efficiency of the sonde pumps started, the pump correction profile that is applied is such that the overall correction factor is the same as the correction factor measured in the laboratory. From 28 April 1989 to 21 May 1992, which was a period with worse ozonesonde pump characteristics, an empirical relationship between the overall correction factor and the pump correction profile deduced from the soundings after 21 May 1992 is used.

Homogenised Dobson values on the Bass-Paur scale were used to normalise all the ozone profiles. For the set of data used in this study, only the profiles with an overall correction factor between 0.92 and 1.50 were retained.

Payerne

Measurements started in November 1966 in Kirchberg, near Zurich. Since August 1968 measurements have been performed at Payerne, 150 km to the west. Launch times have changed several times, with measurements since January 1982 made at 12:00 UT. Important changes in pre-flight procedures occurred in 1977, 1983, and 1993. Before 1976 sondes were prepared at Zurich and then transported to Payerne for launch. In 1983 the improved pre-flight protocol of the group at Hohenpeissenberg (Claude

et al., 1987) was adopted. Since 1993 several changes have been introduced, including complete disassembly and meticulous cleaning of new sondes, followed by repeated calibration with fresh sensing solution on each occasion. Since January 1983 a four-point calibration with an UV photometer has been routinely performed. This method allows the determination of the background current (although this information has not yet been used), as well as giving some information on the quality of the sonde. Since 1984 pump flow has been measured by an electronic flow meter, rather than a bubble apparatus. Some adjustments in order to homogenise the measured flows were applied. Before November 1980 the VIZ radiosonde was used. This was changed to the old mechanical Swiss radiosonde CH, then to the new electronic Swiss Radiosonde SRS-400 in April 1990. The latter change caused a marked discontinuity in the time series that has been corrected by applying a regression model on 23 standard levels. A constant temperature of 300K is assumed for the box with the VIZ sonde, and 280K afterward. The standard pump efficiency correction is applied throughout the data record. A correction of 1.04 (the "Dütsch factor") was applied before November 1980, as well. This has no effect on normalised profiles. All flights are normalised, either to the Dobson at Arosa (200 km east), or where this is unavailable the TOMS measurement. An estimated (average) correction factor is used before 1978 or after 1993 when no measured total ozone is available. Calculation of the residual between 30 and 10 hPa is via the SBUV climatology between 30 and 17 hPa, and assuming constant mixing ratio between 17 and 10 hPa. Soundings ending below 30 hPa are rejected.

NOAA/CMDL

The Climate Monitoring and Diagnostics Laboratory (CMDL) has used the ECC type ozonesonde throughout its measurement program. The three primary time series, which began in the early to mid-1980s, are from Boulder, Colorado; Hilo, Hawaii; and South Pole, Antarctica. There have been several operational changes over the course of the program. The first of these was the conversion from the analogue system to a digital system, which took place in 1991. The analogue system was based on the VIZ radiosonde with hypsometers for measuring pressure at pressures less than about 50 hPa. This system also measured the "box temperature" as a surrogate for the pump temperature required to determine the volume of air pumped into the sensor. With the advent of the digital system the Vaisala RS-80 radiosonde was mated with the ozone sensor through the TMAX-C interface board. With the conversion to the digital sonde the measurement of the instrument temperature was switched to a place on the pump body. In 1994 the supplier of the ozonesonde pump/sensor was switched from the Science Pump Corp. to EnSci Corp. Although the design is very similar there may have been some slight differences in performance characteristics. Starting in 1991 individual pump efficiencies were measured. Soundings have been reprocessed using uniform procedures throughout the time series. These procedures include: (1) Use of pump efficiency corrections prior to 1991 that are an average of the measured efficiencies made since 1991. (This has included the development of a unique oil bubble flow-meter technique for checking earlier pump efficiency measurement methods.) (2) Subtraction of a constant background current throughout the entire flight profile. (3) A correction for the over sensitivity of the 1% sensing solution to ozone exposure (required by the use of the larger pumping efficiency corrections measured. This correction is based on measurements using solutions of varying composition in comparison to a calibrated photometer). (4) Normalisation to Dobson total column ozone amounts that have been reprocessed so that they are on the Bass-Paur scale. In addition the residual ozone has been calculated using the SBUV climatology rather than the extrapolation using a constant mixing ratio. Using these procedures a relatively homogeneous data set has been produced. Corrections that have not been made are for possible inconsistencies in the pressure and temperature measurements associated with the change in radiosonde. The changes introduced by the change in radiosonde appears to be relatively small.

NASA/Wallops Island

This station began soundings in May 1970. Measurement density varies from one or more observations

per week to two per month, dependent on available resources. A 1.5% KI solution was used through the data of 1994. Beginning with the use of the digital VIZ radiosonde in 1995 (prior sondes were analogue) the station switched to use of 1% solution. Standard operational procedures have otherwise remained the same since the late 1970's. Individual pump efficiency calibrations are made for each sonde in a pressure chamber. The background current (zero offset) correction is independent of pressure. Background current is measured about one-half hour prior to instrument release before the cathode cell has been exposed to ozone. Comparison to a Dasibi UV monitor is made prior to launch, as a quality check. Dobson normalisation factors are also calculated (using constant mixing ratio to calculate the residual), again as a quality check. Neither measurement is used to correct profile data submitted to the World Ozone and Ultraviolet Radiation Data Center (WOUDC), Toronto, Canada. For this study the total ozone record on the Bass-Paur scale was applied to the data to account for changes from the application of the correction factor in the earlier portion of the archived record to the more recent procedure of not applying this correction.

Canadian Network

Prior to 1980 the Canadian network flew sondes of the Brewer-Mast type. These were prepared in the fashion described in Mueller (1976); that is, sondes were conditioned with a high ozone source for about 30 minutes, then filled with solution and launched; no special cleaning or calibration was performed. Sondes were filled with solution about one hour before launch. Correction factors for this period are fairly high (~1.3) and there is considerable scatter. In addition, there is a clear bias in tropospheric measurements that will produce spurious trends if these data are combined with the ECC data of the later period [Tarasick *et al*., 1995]. These data have not been re-evaluated and are not reliable for trend assessment in their current state.

Since 1980 onwards the Canadian network has flown ECC sondes. Two others have been added more recently. Sondes are flown weekly at all stations, with additional flights at times of special interest. ECC 3A sondes were used in Canada until about 1983, when stations changed to the 4A. Since 1994/1995 the digital ECC type 5A and 6A, manufactured by Science Pump Corp., and the EN-SCI 1Z have been used in Canada. Differences between these types and the 4A appear to be only in the electronics. The sonde is coupled to a Vaisala RS80 radiosonde via an RSA11 interface. Prior to 1993 an analogue system based on the VIZ radiosonde was used. Hypsometers for measuring pressure were used only on a few soundings, on an experimental basis. The possible effect of the change of radiosonde has not yet been examined. Operational constraints have at times required some variation in launch times. This appears to have had a significant effect in the lower troposphere at Edmonton and Alert [Tarasick *et al*., 1995]. In the mid-1980s at Edmonton mean values increased by about 42% at the surface when the standard launch time was changed from morning to afternoon.

Sonde preparation procedures have been consistent throughout the Canadian program. Solutions have been prepared in accordance with Komhyr (1986), i.e., 2.5 ml of 1% KI solution in the cathode cell and 1.5 ml of anode solution. Concentrations have not been changed at any time in the ECC record, and preparation has been at one central location. Sondes are typically charged with solution one week to one month before launch, and solutions are topped up weekly, where necessary, with solution. Sondes are charged with fresh solution on the day of launch. Sondes are tested before launch with an uncalibrated ozone source and rejected if they show excessive background current, but complete laboratory characterisation of individual sondes has not been standard procedure in Canada. Standard pump corrections recommended by the manufacturer (which necessarily differ between 3A and 4A models) have been used for all sondes. The correction curve supplied in 1983 has been used for all 4A flights; this was changed slightly in Komhyr (1986) and as the Vaisala software uses the newer correction curve this has caused a change with the 5A sondes. Pump flow rate at surface pressure is measured in the pre-flight procedure and used in the data reduction. A pressure-dependent correction for background current

is used throughout the Canadian record. At several stations there are significant trends in the recorded background current, which could affect derived trends in the upper troposphere, if the background current were in fact constant with altitude.

Where a total ozone measurement is available, profiles are scaled linearly to the total ozone measurement. The residual ozone above 17 hPa is approximated by the average of all measurements above that height, using the assumption of constant mixing ratio. No correction is applied for flights that fail to reach 17 hPa. All correction factors have been adjusted to the recently re-evaluated time series of Brewer total ozone data for those stations. Zenith sky total ozone measurements were not used for the ECC portion of the record for this study. At the arctic stations total ozone measurements are not available in the winter, which may lead to a seasonal bias in the corrected data. Overall, the correction factors at Canadian stations for the ECC record are very close to 1.0, and show insignificant trends.

Japanese stations

Three stations in Japan have been operating since late 1968 (in addition to the Japanese Antarctic station, Syowa). A fourth was added in 1989. Since 1991 soundings are weekly; before that time they were bimonthly, with no soundings in the summer. The Japanese sonde is a galvanic cell like the Brewer-Mast, with a carbon anode. Before 1979 a model with two solution chambers was used. This change has changed response time to ozone changes. Sensing solution is 0.04% KI with 34% KBr and a buffer. This has not changed. Temperature is measured near the inner surface of the solution cell. Background current is measured twice: 15 minutes before the launch, and also during the pre-flight calibration, at about 20 minutes after exposure of the cell to ozone. Background current is subtracted with decreasing atmospheric pressure. Pump efficiency under low pressure is calculated by from an empirical formula. Values at various pressures are tabulated in Table 1.15. Normalisation to the Dobson total column ozone is done assuming constant mixing ratio after balloon burst. The total ozone records have been re-evaluated for Bass-Paur absorption coefficients. There have been some changes in radiosonde, and a radiation correction has been applied to temperature sensors since 1979 (when the new Japanese ozonesonde was introduced). This may have caused an altitude shift in the measurements.

NIWA/Lauder

The National Institute of Water and Atmospheric Research (NIWA) has been performing ECC ozonesonde soundings since August 1986. Ozonesondes flown have included the 4A, 5A and 1Z series ozonesondes manufactured by the EN-SCI Corporation (1Z) and the Science Pump Corporation (4A and 5A). The 4A sondes were flown from 1986 to October 1989 with telemetry, pressure, temperature and humidity information being provided by a 1680 MHz VIZ radiosonde. From 1989 the instrument has been coupled to a Vaisala Incorporated, 403 MHz RS-80-15 meteorological radiosonde. Coupling is achieved through TMAX-HMOS electronic interface circuitry. A modified version of Terry Deshler's (University of Wyoming) ANLOZ program is used to process the data. Measured parameters telemetered to the ground receiving station are ozone, sonde pump or box temperature, air pressure, air temperature and humidity. In general, 5A sondes have been flown between August 1989 to June 1994, and 1Z sondes from July 1994, although there are occasional exceptions to this. Frequency of flights is generally one per week. From September to about the middle of December the frequency of flights is generally two per week. This can vary for intensive campaigns.

From 3 August 1986 to 31 July 1996 the KI cathode solution concentration was 1%. Based on recommendations by Komhyr (private communication), a 0.5% KI cathode solution concentration has been used from 1 August 1996. This change has produced better comparisons with quasi-simultaneous or simultaneous Dobson total column ozone measurements (Bass-Paur coefficients) and ozone lidar measurements. Three dual ozonesonde soundings consisting of ozonesondes using 1% and 0.5% KI

cathode solutions have been made at Lauder (October 1996, March, and June 1997). With these results as well as field and laboratory comparisons carried out by other groups we intend to eventually be able to homogenise these two data sets.

Two pump efficiency corrections are employed at Lauder. For 4A and 5A sondes, the pump efficiency corrections used are based on measurements made at the University of Wyoming in 1988 by Terry Deshler's group. For 1Z sondes the corrections in EN-SCI (1996) are employed. Third order polynomial fits to the quoted pump efficiency correction factors are used to interpolate values for other pressures. The two corrections are quite different, and reflect, to a certain extent, the different methods that are used to determine pump efficiency corrections. Results do suggest, however, that there is a systematic bias between 5A and 1Z series sonde pump performance in the mid-stratosphere.

The background current measurement is determined after all preparations have been done and the sonde is about to be launched. In most cases this background has been obtained shortly after exposing the sensors to ozone, but in more recent times the pre-launch procedures have been performed up to several hours before the actual launch - the background current has then, subsequently, stabilised at a lower value. For 5A and 1Z sondes the background current is assumed to remain constant through the flight. For 4A sondes a pressure scaled background current has been applied to the data.

Normalisation to total column ozone is not done, but a normalisation factor for all flights is determined based on an assimilated total column value from sources such as Dobson (Bass-Paur scale), Nimbus-7 and Meteor-3 TOMS V.6 and V.7, and ground-based scanning spectrometers. For sondes containing a 1% solution, ozone data is scaled by 0.9743 to match the Bass-Paur coefficient correction to Dobson data. This is not done for the soundings using a 0.5% solution. A residual ozone calculation is made by using a 40 to 50oS zonal SBUV climatology (McPeters *et al.,* 1997). This calculation is applied from the top of the flight to 1 hPa.

For 4A sondes a thermistor measuring box temperature was positioned near the base of the ozonesonde. For 5A sondes the thermistor was glued to the base or top of the Teflon block near the pump. For 1Z sondes the sondes have a thermistor positioned in the Teflon block near the piston. Experiments performed at Lauder suggest that the 5A sonde 'pump temperature' is between 2-3K cooler than the 1Z sonde 'pump temperature' throughout the sounding, while the 4A sonde 'box temperature' measures about 1K cooler than the 1Z sonde in the troposphere, increasing to about 4K in the mid-stratosphere. Temperature changes of this magnitude probably do not have a big impact of the profile. In recent years work has been done to standardise the package configuration and general procedures. This was done specifically to generate consistent internal temperature profiles which give an indicated temperature between 10-17oC by the top of the profile.

No correction for direct radiative heating of the radiosonde temperature sensor has been made to the data. From 28 March 1996 a shadow band around the radiosonde temperature sensor has been used for day time flights. Corrections to pressure have been made to all backscatter sonde flights (incorporating ozonesondes) since 1992 – these packages are flown approximately monthly – these corrections were made by measuring the pressure differences between radiosonde and test unit at 10 or 15 hPa. Corrections to other sonde flights have been made since 7 August 1996, based on comparison with an accurate pressure sensor at the station and assuming a constant difference through the flight.

Other Stations

A number of other stations have records whose overall quality remains unassessed, but which represent the only sources of ozonesonde profile data for a significant portion of the globe. These include the Australian and Indian stations, as well as Natal, Brazil. In some cases the data record is sparse or

intermittent. Because of their geographic locations, these data sets must be considered in a global trends assessment, but they should be regarded with less confidence than the data records described above.

1.8.6. Conclusions

- Procedures for operating ozonesondes, producing an ozone profile from the measured sensor signal, and correcting data for known changes in ozonesonde characteristics are sufficiently different between programs that results from one program cannot routinely be applied to another location. This can include assignment of error or measurement uncertainties with the potential to influence trends.
- Correction of an ozone profile by use of an independent measure of the total ozone (normalisation) should only be done after making all known altitude-dependent corrections to the profile.
- The correction factor derived from normalisation of the integrated profile to the total ozone should not be the sole determiner of sonde quality. In particular, there is no justification for using a single criterion for all ozonesondes even for a particular type of ozonesonde.
- Normalisation of the profile using total ozone may provide stability to the ozone time series but also produces some risks that regions with smaller ozone amounts (the troposphere and above 30 km) will be adversely affected if altitude-dependent corrections are not adequately known.
- ECC pump efficiencies have probably not changed significantly with time. On the other hand, BM pump efficiency has likely deteriorated with time (see Chapter 2 where evidence is presented based on intercomparison studies).
- Recent studies suggest that the effect of changes in ECC sensing solution concentration on ozone measurements is complex and requires further study before being implemented. These changes have little impact on the data sets discussed here because they have only been implemented in a few places relatively recently (<1 year).
- Because normalisation of ECC sondes gives average correction factors near 1.0, the skewing of the profile shape is likely to be minor and the trend stability of time series of normalised profiles should be good from the lower troposphere to 30 km.
- Data from BM sondes used at the Canadian stations prior to 1980 should be used with great caution since operating procedures for preparing these sondes did not meet standards employed by the three existing European BM stations and are thus uncharacterised in their performance.
- Re-evaluation of the data record at Payerne is not complete.

1.9. Summary and Conclusions

In this chapter the sources of data used in this document have been reviewed to provide information on the measurement techniques, the data analysis methods used, and their characteristics and sources of error, both for individual measurements and for the extent to which the measurement of long-term trends are affected.

Ground-based Dobson instruments, observing ozone profiles in the Umkehr mode, have the longest heritage of ozone measurements with substantial records at some stations from 1957. The network of instruments is well calibrated and produces reliable ozone profiles in the stratosphere between approximately 20-40 km. Vertical resolution is relatively coarse, but adequate for analysis of ozone trends. Spatial coverage is limited by the relatively small number of ground stations but could be improved substantially by including the Brewer instrument network, an independent, well-calibrated system, which also makes Umkehr observations, and has been in operation since the mid 1980s.

While chemical ozonesondes perform well below about 25 km, above this altitude, uncertainties related to gas sampling techniques become a limiting factor not present in the UV-absorption in-situ instruments. At altitudes beyond the capability of aircraft (about 20 km) and balloons (about 40 km) remote techniques are clearly necessary. The lidar and microwave remote ozone profiling techniques, operated from the surface, provide adequate accuracy in the 25-50 km range and allow for continuous measurements, not generally possible with balloon and aircraft measurements. The coverage by microwave instruments is not yet adequate for them to be considered in this report.

In terms of global and temporal coverage, satellite-based instruments are obviously superior to surface-based ones. The various satellite instruments have different individual characteristics with respect to long-term calibration stability, global coverage, vertical resolution or sensitivity to contamination by stratospheric aerosol.

The solar occultation instruments, SAGE and HALOE, which utilise limb extinction at several wavelengths during sunrise and sunset events, have good long-term stability as they are able to calibrate before (sunset) or after (sunrise) each vertical profile measurement by viewing the sun directly with negligible intervening atmosphere. Vertical resolution is good, being on the order of 1 km. Spatial coverage is relatively poor because of the requirement of a sunrise or sunset.

Limb viewing in the visible/near infrared is subject to contamination by volcanic aerosols. Since a major volcanic eruption, which perturbs the stratospheric aerosol distribution, can be anticipated on average at least every ten years, this presents a limitation for maintaining long-term records with present instruments. Future instruments, with increased wavelength resolution, should allow an improvement in aerosol extinction corrections.

The nadir-viewing buv-type instruments have good global coverage for ozone profiles above about 20 km, but are subject to calibration uncertainties, the possibility of long-term drift, and have poor vertical resolution. Aerosol contamination is present, but is less of a problem for these instruments.

The MLS instrument, operating in the microwave wavelength range, is the least susceptible to aerosol contamination. Its global coverage is limited to viewing from 34o in one hemisphere to 80o in the other during a each orbital precession period (about 36 days) of the UARS spacecraft, although this is a platform, not an instrument, limitation. Its vertical resolution (potentially 3 km, but currently 5.4 km) and its calibration stability are intermediate to the other two satellite techniques.

The accuracy of estimates of ozone trends will depend on drift in systematic measurement errors in, as well as on the statistical fitting error due to sampling limitations in the presence of natural variability. Table 1.16 summarises the drift uncertainty (1s) in the calibration of each data source, but must be read in conjunction with the detailed description and caveats given in the individual sections. The analysis of drift uncertainty is still at an early stage, as it can involve many subtle effects that may not normally be considered for single-profile error analysis. This table should be regarded as preliminary, and to be treated with caution.

Data Source	Altitude, km					
	5	15- 0	20-25	25-30	30-35	35-40
SAGE I	-	0.3	0.3	0.3	0.07	0.05
SAGE II	-	0.15	0.15	0.15	0.07	0.05
HALOE[1]	-	0.16	0.16	0.05	0.05	0.05
MLS	-	> 0.3	0.2	0.1	0.1	0.1
SBUV	-	-	0.25	0.2	0.25	0.3
Lidar[2]	-	0.025	0.025	0.025	0.02	0.02
Umkehr[3]	-	-	0.2	0.2	0.2	0.2
Ozonesondes						
ECC	0.08	0.08	0.08	0.19	-	-
Brewer-Mast	0.42	0.11	0.11	0.31	-	-
Japanese	0.20	0.09	0.09	0.20	-	-

1 For aerosol conditions starting 1993. For low aerosol, estimated drift error is 0.07 below 25 km and 0.04 above.

2 Not including factors which are unique to individual instrument designs.

3 These estimates are all £ 0.2.

Table 1.16. Summary of the drift uncertainty (1s) in the calibration of the various data sources, measured in %/year. This table must be read in conjunction with the detailed description and caveats given in the individual instrument sections.

Chapter 2

Data Quality

Lead Authors

James M. Russell III - Herman G. J. Smit

Co-Authors

Roger J. Atkinson - Brian J. Connor - Derek M. Cunnold - Lawrence E. Flynn

Lucien Froidevaux - Sophie Godin - Stacey Hollandsworth - James B. Kerr

Richard D. McPeters - A. J. Miller - Gary A. Morris - Mike Newchurch

David W. Rusch - Larry W. Thomason - Ray H. J. Wang

Contributors

Gerard Ancellet - Matthias Beekmann - Greg Bodeker - Rumen Bojkov

Ellen J. Brinksma - Hans Claude - Hugo De Backer - Pierre Jeannet - Bryan Johnson

N. Kämpfer - Gordon J. Labow - Ulf Köhle - I. Stuart McDermid - Thomas McGee

Hideaki Nakane - Samuel Oltmans - Francis J. Schmidlin

Johannes Staehelin - Wolfgang Steinbrecht - Daan P.J. Swart - David Tarasick

Osamu Uchino - Pierre Viatte - Joseph M. Zawodny

2.1. Introduction

No uniform data source exists to allow careful review of the current state of knowledge of long-term changes in the vertical ozone distribution. The data available for trend studies comes from different ozone measuring platforms using different sensing techniques with different spatial and temporal coverage. Therefore, in preparation for this assessment the current data quality status for the different measuring systems was evaluated by a group of experts at a workshop held at the Observatoire de Haute Provence in July 1996 (Harris, 1996). This study resulted in the definition of an inventory of the most reliable data sets with time records longer than 5-10 years (see Table 2-1).

A. Middle/Upper stratosphere above 25 km:

Satellites:	
SAGE I	Feb., 1979 - Nov. 1981
SAGE II	Oct. 1984 - present
SBUV + SBUV2	1978 - present
HALOE	Oct. 1991 - present
MLS	Oct. 1991 - present
Ground-based:	
Microwave	1989 - present
Umkehr/Dobson	1957 - present
Lidar	1990 - present

B. Lower stratosphere below 25 km:

Satellites:	
SAGE-I	Feb. 1979 - Nov. 1981
SAGE-II	Oct. 1984 - present
HALOE	Oct. 1991 - present
MLS	Oct. 1991 - present

Ground-based:	
Microwave	1989 - present
Umkehr/Dobson	1957 - present
Balloon borne sondes	1965 - present
Lidar	1985 - present

C. Troposphere:

Ground based:	
Balloon borne sondes	1965 - present
Lidar	1990 - present

Table 2-1**:** Inventory of the most "reliable" data sets with time records longer than 5-10 years.

It is obvious that only four measurement systems have records long enough to assess long term ozone trends: SAGE (I and II), SBUV and SBUV2, Umkehr/Dobson and ozone sondes. These measurements cover the lower stratosphere to the stratopause. The SAGE I and II satellite series extends from February 1979 to the present with a 3 year interruption from November 1981 to November 1984. The SBUV/SBUV2 satellites and ground-based Umkehr/Dobson instruments spanning the period from 1957 to the present provide a data set for examining middle and upper stratospheric trends. The ozonesonde network starting in the early 1960's extends to the end of 1996 and has generated a data set for potentially examining lower stratospheric/tropospheric ozone trends. There are also other ozone data records available over shorter time periods which can be used to validate the long-term measuring systems including the HALOE and MLS instruments operating on the UARS satellite, ground-based lidar, and ground-based microwave instruments. We report here on studies done to evaluate the validity of the long-term trends obtained by these observing systems.

A major objective of this chapter is to determine if there is sufficient confidence in the long-term measurement systems to use them for accurate determination of ozone trends in the stratosphere and troposphere. The purpose of this study is therefore to validate the quality of these data sets including quantification of errors and to determine if there are any limitations in altitude or latitude regarding use of the data for deriving global trends. Important data quality questions addressed in the study include:

- What are the lower and upper altitude limits of the different data sets to be used in validation or trend studies?
- What is the smallest long-term ozone change that can be verified and over what altitude range?
- Are there any evidences of instrumental drifts of the measuring systems with time?

- What special screening methods should be used for the different data sets?
- Should SAGE sunrise, sunset or both data sets be used for trend determinations?
- How should data offsets of SAGE I compared to SAGE II be handled?

Answers to these questions will provide the basis for recommendations needed to derive the ozone trends reported in Chapter 3.

The general philosophy of this study was first to do internal data consistency studies for the long-term ozone measuring platforms through intercomparisons and data analysis. Next the long-term ozone records were compared with data collected over shorter time periods (e.g. HALOE, MLS, Lidar and Microwave) to look for any possible seasonal or latitudinal biases, which would indicate systematic errors in the instruments or which point to inherent instrumental drifts.

Above 25 km SAGE I and II, SBUV, SBUV2, and the Umkehr/Dobson-network provide the major data sources to derive global ozone trends in the middle and upper stratosphere. The good spatial coverage of the SBUV and SBUV2 data are used to examine the effect of the poorer spatial coverage of SAGE on the derived trends. In addition, the HALOE and MLS data were compared to SAGE II over shorter time intervals to determine if there are any seasonal or latitudinal dependencies in SAGE II trends. Umkehr/Dobson, Lidar, and Microwave data were used to verify the accuracy of individual profiles measured by the SAGE I and II satellite instruments and to further assess the validity of trends.

An important issue in this study is the derivation of global ozone trends between the tropopause and 25 km which is the region with the largest ozone depletion. SBUV and SBUV2 and Umkehr/Dobson are not sensitive enough in this region which leaves SAGE I and II as the only global data set with temporal and spatial coverage sufficient to meet this objective. However, the coverage of SAGE I and II is still limited to between 60°S and 60°N. Further, the validity of SAGE measurements for Z £ 25 km needs to be carefully scrutinised due to possible errors in the SAGE data caused for example by high aerosol loading resulting from volcanic eruptions (e.g., Mt. Pinatubo) or ice crystal absorption due to high cirrus clouds (see also Cunnold *et al.*, 1996a). Ozonesondes provide data to derive long-term trends in the lower stratosphere, but with very poor spatial coverage at a limited number of stations, mainly located on the continents in the northern hemisphere (NH). Ozonesondes alone therefore cannot provide global trends. However they provide valuable data sets for intercomparison with SAGE results to investigate any instrumental biases and drifts in indicated trends. In the troposphere ozonesondes provide the only data for use in deriving an ozone trend. The availability of other data to validate the sonde measurements is very limited and can only be done over very short time intervals. Only a few dedicated intercomparison campaigns have been performed in the past. However, over the last 5-10 years several studies have been performed to validate sonde data by comparison with other techniques like Lidar or UV-photometric measurements from aboard aircraft.

During the pre-satellite period before 1978/1979 ozone sondes and Umkehr/Dobson measurements provide the only data sets for study of ozone change. Therefore, it is important to characterise the quality of the sondes and Umkehr time series during this period (i.e. 1965-1979). There are indications that in the tropopause region up to 17-20 km ozonesondes are probably more reliable then satellites. We have attempted in this study to use the newest, latest, revised versions of the different ozone data sets which were made available before the end of June 1997 and which were stored on or linked to the dedicated SPARC temporary data base.

Coincidence criteria used for all comparisons shown in this report are ± 2° latitude, ± 12.5° longitude, and 2 days time unless otherwise noted. We recognise that these are rather loose criteria. It was found for

example when validating the Nimbus 7 LIMS data, that even 3 hours could make an important difference in the temperature agreement; but because SAGE II uses the solar occultation experiment approach, coincidence opportunities are limited requiring adoption of the criteria used. When comparisons are done against SAGE II, SAGE II is always the reference used in calculating percentages with the formulation:

$$D = 100 * (\text{Correlative } O_3 - \text{SAGE II } O_3) / \text{SAGE II } O_3 \, \% \quad (2.1)$$

where correlative refers to balloon sondes, lidar, and other techniques. Also unless otherwise noted, the error bars shown throughout this chapter are the 2 standard error mean bars (or 2 sem) calculated by dividing twice the standard deviation by the square root of the number of samples used in the comparison.

2.2. General Issues

2.2.1. Strategy for Intercomparison Analyses

There are now several stratospheric ozone profile data sets which overlap the long-term data sets for a period of several years. We used intercomparisons among these various data sets to partially test the validity of trends determined from the longer data records. Some limited comparisons are also included from more restricted campaigns (spatially and temporally), to present additional information about possible systematic errors.

Given two data sets covering a few years of data, we do not try here to deduce a linear trend after allowing for modelled (fitted) variations such as seasonal and solar cycle effects (see WMO Report of the International Ozone Trends Panel: 1988). This is because we mainly want to check how the data sets track each other over time, without trying rigorously to take into account the variations in terms of real "trends." Also, certain phenomena like solar and QBO effects cannot be removed reliably, based on a few years of data alone. Chapter 3 deals with the trend issues for the long-term datasets.

One component of the data quality analysis will rest upon simple linear regression (least-squares) fits for the time series, or difference time series for coincident data sets, to assess possible drifts with time in the measurement systems. For global datasets such as those from SAGE II and UARS, comparisons are given for rates of change from coincident measurement sets as well as from zonally-binned datasets. Ground-based measurements, such as lidars or sondes, are also used to check rates of change by comparison with coincident satellite data sets (e.g. versus SAGE II). Temporal sampling issues can be analysed by comparing these values against comparisons of the rates of change inferred from each "full" dataset. The standard error in these slopes are provided but it must be kept in mind that these are generally underestimates of the true error because the analyses ignore the temporal autocorrelations in the data (WMO Report of the International Ozone Trends Panel: 1988). Also, zero error is generally assumed for the time-varying component of systematic error, which is an idealised assumption that could be a root cause for differences between datasets. Trends of differences (i.e. drifts) between data sets are used to address this issue.

Significant changes in tracking or consistency between datasets with altitude (or latitude), can provide clues about possible reasons for differences. In particular, newer data sets such as UARS data should be checked against other existing datasets in the upper stratosphere, since there is good documented consistency between SAGE, SBUV, and Umkehr trends in that region (e.g. WMO Scientific Assessment of Stratospheric Ozone: 1994). This necessary (but not sufficient) check (for SAGE II data, in particular), provides added confidence in the mid- to lower stratosphere intercomparisons. Another useful analysis is to compare datasets under different conditions of aerosol loading (e.g., use a time interval shortly after the eruption of Mt Pinatubo, as well as one from a more quiescent time); this issue comes up because there is evidence that SAGE ozone retrievals can be contaminated by severe aerosol loading (see Cunnold *et al*., 1996a). This chapter will give recommendations for screening the SAGE data so that improved data quality exists for trend determination, even if this comes at the cost of somewhat more limited time coverage.

2.2.2 Effects of Viewing Geometry and Vertical Resolution

SAGE II and other limb-viewing satellite instruments (including HALOE and MLS) have nominal instantaneous vertical fields of view (FOV) ranging from 0.5 to about 5 km. On the other hand, both ozonesondes and ozone lidar have considerably higher vertical resolution. Since ozone profiles may

exhibit strong vertical structure, comparisons of ozone measurements by different instruments should reflect those differences. In the lower stratosphere, ozone decreases rapidly below the ozone peak (~21 km). Since the IFOV vertically averages the measured slant path transmission (or emission), considerable biases in the ozone profile derived from limb-viewing instruments are possible (depending on the slope of the gradient and the field of view of the instrument). For example, Figure 2-1 shows the impact of the IFOV on the retrieval of a species with steep gradients both above and below a peak (much like ozone) based on convolving the FOV with the natural weighting function but without measurement noise and algorithm smoothing. For larger IFOVs, there are subtle problems around the peak but a significant bias is obvious below the peak. To properly interpret differences between measurement systems, comparisons of ozone (both spaceborne and in situ) should reproduce the vertical smoothing scheme inherent in SAGE II as well as HALOE and MLS (including both geometry and IFOV). It should be noted that, simple vertical smoothing of retrieved profiles does not fully correct for the field-of-view effect (though such smoothing may be adequate for narrow fields of view) but it is the approach which has been adopted for most of the comparisons performed in this chapter.

Click here to download Figure 2.1 by anonymous ftp.

Figure 2-1 The changes in an inferred mixing ratio profile (absolute: right panel; relative: left panel) as a function of the vertical field-of-view (FOV) for a limb-viewing instrument. FOV averaging was performed on the line-of-sight optical depth and then profiles were retrieved with an onion-peeling type algorithm.

2.2.3 Altitude to Pressure Conversion Issues

Ozone measurements made by SAGE or Lidar are recorded at geometric altitudes but many other ozone measurements (e.g. SBUV, MLS and ozonesondes) have pressure as their natural reference. A comparison of ozone trends and inferred solar cycle effects shows that there is better consistency between SAGE and the other satellite measurements and long term self-consistency in the SAGE measurements if the trends and solar cycle effects are reported on altitude levels (Cunnold and Wang, 1997). This better consistency is probably related to the presence of long term trends in tropical temperatures in the upper stratosphere in the NWS data that are being used to convert SAGE altitudes to pressures (Cunnold and Wang, 1997). These trends exceed those predicted by models and result in geopotential height changes of approximately 300 m between 1988 and 1994 in the tropics at 1 hPa. This produces a trend difference at that location over the 1988 to 1994 period between SAGE trends on altitude surfaces versus those on pressure surfaces of approximately 1%year^{-1}. At ~30 km altitude and above in the mid-latitudes the effect of temperature uncertainties on this scale conversion is likely to impact ozone trends by ~0.2%year^{-1} over as short a period as 6 years. It is therefore recommended that ozone trends be reported in the natural vertical coordinate of the individual measurement technique. This would avoid the use of temperatures with trends which have their own inherent uncertainties. If this is not done, direct comparisons between measurements would then require an assumption that temperature changes were insignificant. A comparison of SAGE ozone trends on altitude and pressure levels would however probably bound the effects of temperature uncertainties.

There are also some altitude to pressure conversion issues to be considered in use of SBUV2 data. The SBUV2 retrieval algorithm assumes a constant pressure gradient/mass density relationship which ignores the inverse dependence of gravitational force on distance squared. Equation (2.2) provides a simple relationship which can be used to correct for this effect where P_c is the corrected pressure, P_i is the initial pressure, Z is altitude, and R_e is the radius of the Earth.

$$P_c(Z) = P_i(Z)\left[1-(2Z+14)/R_e\right] \quad (2.2)$$

2.2.4 Space and Time Sampling Questions

In comparing measurements from different instruments the comparison process itself introduces uncertainty because of space and time sampling questions. There are really three questions for each comparison: are the measurements at the same location, are the measurements at the same time, and are the instruments measuring the same thing (i.e. high vertical resolution point measurement versus a low vertical resolution measurement over a broad area)? The answers vary considerably for different measurement systems, limiting the quality of comparisons that can be done.

Space/time sampling ranges from sporadic point measurements (balloon sondes, Umkehr) to full global coverage on a daily basis (TOMS). Lidar, balloon sondes, and Umkehr data sets comprise point measurements made at a few locations around the globe and must be regarded as both spatially and temporally sparse. Lidar measurements are generally made at night, require clear conditions, and the data records are only a few years long (since 1988 at Table Mountain). Umkehr measurements are made near sunrise and near sunset and also require clear conditions. These measurements have been made for many years and are of interest for extending the ozone record back in time, but even in the early 1980's the record is sporadic and difficult to use to systematically evaluate ozone profile changes. Balloon sonde measurements have the advantage that they can be made day or night and under cloudy conditions, but at many sites balloons are flown only monthly, or weekly at best. Ground-based microwave measurements also have the advantage that they can be made day and night, but again the available data records are short.

The advantage of satellite measurements is that coverage is much better, but there are still limitations depending on the platform and on the measurement technique. Moreover, a space-based measurement is usually an average over some IFOV which covers a finite area on the ground, an issue when doing a comparison with a point measurement. Many of the comparisons in this report are relative to SAGE II. Because SAGE uses the solar occultation technique, it makes only two profile measurements per orbit, about 30 measurements per day. SAGE II is in a precessing orbit such that the sunrise and sunset latitudes shift from high northern latitudes to high southern latitudes approximately monthly. Coverage is fairly good between 60°N and 60°S but is sporadic in the polar regions. SAGE has good altitude resolution, ~1km, but it samples a horizontal volume of about 300 km along the line of sight at the tangent point. HALOE, on the Upper Atmosphere Research Satellite (UARS), is also an occultation experiment and has similar spatial and temporal coverage. The MLS instrument, also on UARS, measures limb emission and thus has continuous coverage along the orbital track; but it has similar horizontal averaging characteristics as SAGE and HALOE. The view geometry gives coverage from about 30 degree latitude in one hemisphere to about 80 degree latitude in the other hemisphere, with the prime hemisphere changing every 34 days as the orbit precesses. The SBUV and SBUV/2 instruments measure ozone using backscattered ultraviolet sunlight. Since both Nimbus 7 and NOAA 11 were in polar orbits and the instruments are nadir viewing, they produce coverage from 80°S to 80°N on a daily basis, but only along the orbital track. No measurements are possible at high latitudes in winter (polar night). The IFOV on the Earth surface is a 200 km square.

Analysis of the TOMS total ozone data shows that at spatial separations between measurements greater than about 5^{o} great arc distance, the correlation begins to drop. This is most evident at mid to high latitudes in winter and early spring, when the ozone field is highly variable. The effect depends on altitude - ozone in the upper stratosphere is much less variable and a match within $<20^{o}$ will be highly correlated. As the time separation increases, the effect is equivalent to an increasing spatial separation. An air parcel measured by a balloon sonde today may have been a thousand kilometres away two days

prior. Two approaches to improving the number and quality of matches will be discussed in section 2.2.5 - trajectory mapping and potential temperature-potential vorticity interpolation.

Various data sets are compared in this report, with SAGE as the prime data set acting as a transfer standard. If, for example, a single SAGE profile is compared with an individual balloon sonde profile at a site like Hohenpeissenberg, the number of matches will be quite small. In general, comparisons will be more significant statistically if monthly zonal means are compared. In the 30°-50°N zone there are enough balloon sonde stations (or Umkehr stations) for a good comparison. Station-to-station biases can be a problem if the data records from the different stations are non-uniform - i.e. if a station with a data record only in the late 80's has a bias relative to a station that has data mostly in the early 80's and both are included in the long term comparison.

The comparison of the lidar stations is possibly the most difficult because of the small number of sites and the short data record length. Any altitude dependent biases can be determined, limited by the number of matches that go into the comparison, but the time dependence will be very uncertain because of the short length of record.

The satellite-to-satellite comparisons will generally have many more matches, though the UARS orbit was such that HALOE still has only limited direct matches with SAGE II. Here again good comparisons of zonal means can be made. Some of the best comparisons can be made of SBUV2 with SAGE II, because there will be an SBUV-IFOV that can be matched with each SAGE-IFOV every day. A limiting factor here is the differing vertical resolution - 1 km for SAGE versus about 8 km for the SBUV experiment.

Given the space and time sampling of the various instruments, we conclude that a good job can be done on determining the altitude dependent biases, but that the ability to determine time dependence of differences will be limited because of the shortness and/or non-uniformity of many of the data records

2.2.5 The Use of Dynamical Techniques in Ozone Data Validation.

The most frequently used methods for ozone data validation in the past have involved intercomparison either of individual pairs of "coincident" measurements or of zonally-averaged data. These techniques generally provide an effective means of validating the observations from instruments achieving essentially global coverage on a daily basis, but this is not the case for instruments which sample the atmosphere more sparsely (in either space or time), such as the satellite-borne solar occultation instruments and the ground-based instruments. For these latter instruments, the "coincidence" approach relies upon extensive data validation campaigns that provide a relatively small subset of "coincident" measurements, where the definition of "coincidence" often implies separations of hundreds of kilometres in space and/or many hours in time. This approach most often produces a statistically unsatisfying comparison of only a small subset of the observations, which must then be assumed representative of the larger data set. Alternatively, zonal averaging of the individual data sets and subsequent intercomparison of zonal mean analyses generally provides statistically significant results, but effectively eliminates the possibility of examining zonal variations. In both cases, normal meteorological variability adds considerably to the statistical uncertainty associated with the validation results.

Recently, two alternative approaches to data validation have been developed: coordinate mapping (or constituent reconstruction) and trajectory mapping. Each approach takes advantage of a number of quasi-conserved properties of the atmospheric flow, both to enable relaxation of the data coincidence constraints, and to reduce the effects of the short period and large amplitude dynamical fluctuations which are ubiquitous in the atmosphere.

In the lower stratosphere where the greatest uncertainty remains with respect to both ozone measurement accuracy and long-term trends, ozone acts as a relatively long-lived tracer in the absence of heterogeneous chemical processes. Ozone mixing ratio tends therefore to be well conserved in a lagrangian or "along the flow" sense. Furthermore the potential temperature (PT) of an air parcel tends also to be well-conserved along the flow, as does an air parcel's potential vorticity. Coordinate mapping and trajectory mapping exploit two or more of these three quasi-conserved properties of the stratospheric flow to enhance our ability to effectively intercompare ozone data sets.

Trajectory mapping, by employing O_3 and PT conservation assumptions, allows a stricter definition of "coincidence" to be used than in the traditional technique, increases the number of coincident data pairs available for comparison, and at the same time enables a decrease in statistical uncertainty to be achieved. Coordinate mapping employs all three of the above conservation assumptions, to enable virtually the entire data set from each instrument to be used in the intercomparison, while greatly reducing the contribution to uncertainty attributable to atmospheric variability. The details of each of these two approaches and their application to an intercomparison between SAGE II and HALOE data are provided in Section 2.4.6 below.

2.3. SAGE Data Use Considerations

2.3.1. Aerosol Effects on SAGE II Ozone Data Quality

Based on comparisons between MLS and SAGE II ozone data, Cunnold *et al*. (1996a) demonstrated that, following the eruption of Mt. Pinatubo, SAGE II ozone values were being systematically over-estimated because of a failure to correctly account for the removal of the aerosol contribution at 600 nm. Steele and Turco (1997) further demonstrated that the aerosol contribution at 600 nm in the SAGE II observations was being incorrectly simulated especially in the absence of major volcanic eruptions. The SAGE II retrieval team reached a similar conclusion at approximately the same time and modifications were made to the algorithm by incorporating a more realistic simulation of typical aerosol size distributions and the wavelength dependence of the aerosol extinction contribution. SAGE II version 5.96 resulted from these changes (together with other changes which had only minor effects on ozone but which increased sunrise/sunset ozone differences somewhat in the upper stratosphere). The effect of the algorithm modification was primarily a significant change in ozone retrievals below 20 km altitude (see Figure 2.2). The changes were fairly systematic except in 1991-1994 (e.g. see Figure 2.2c).

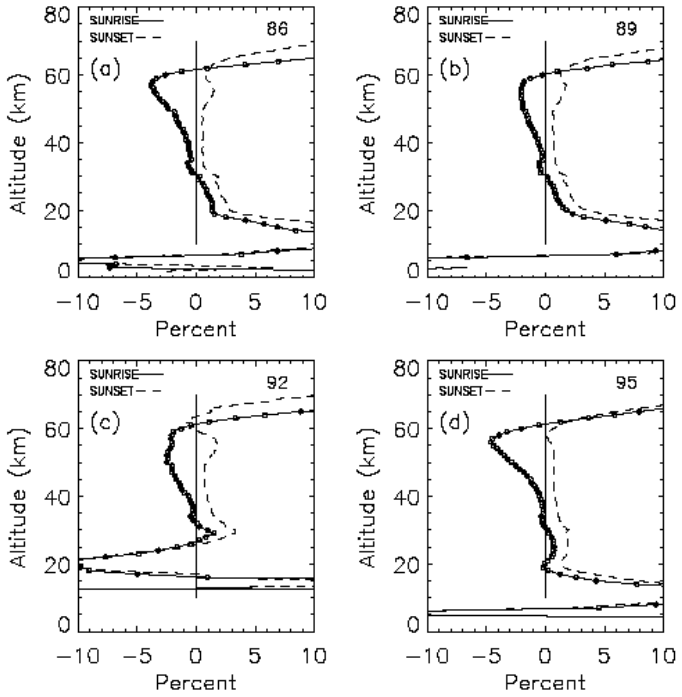


Figure 2.2 The ratios of all SAGE II retrieved ozone concentrations during (a) 1986, (b) 1989, (c) 1992 and (d) 1995 between algorithm versions 5.96 and 5.93 for sunrise (full lines) and sunsets (dashed lines) and expressed as (V.96 - V.93)/V.93. Note the separation of the two lines above 40 km altitude, the significant differences in the retrievals below 20 km altitude, and the unique behaviour of the differences below 30 km altitude in 1992 (caused by a new treatment of aerosol effects).

Figure 2.3 shows a time series of the SAGE II (v 5.96) results at 45°S and 18.5 km altitude and Figure 2.4a shows the differences between SAGE II and MLS ozone values at 68 hPa at mid-latitudes as a function of time. Both figures indicate that the SAGE II ozone measurements were anomalously variable in 1992 and 1993 and that the SAGE II values were larger relative to the MLS values than in other years. It is concluded that the improvements associated with the version 5.96 SAGE II retrieval algorithm have not totally removed the aerosol contamination of the ozone signal during times when the

aerosol extinction was large.

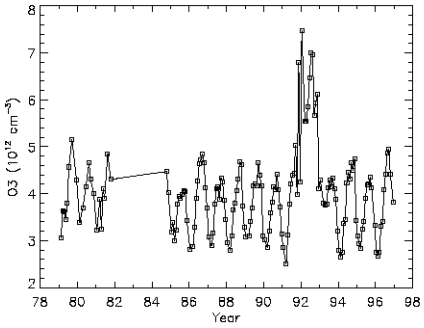


Figure 2.3 A time series of zonal monthly mean SAGE ozone measurements at 18.5 km altitude, 45°S. Note the anomalous ozone values from mid-1991 until mid-1993 which were caused by aerosol interference effects on the SAGE retrieval algorithm.

In the trend studies it is therefore recommended that periods with high aerosol loading be filtered out of the SAGE II ozone time series. Several filters have been considered for this purpose. These include the elimination of specific time periods at different locations (see Table 2.2), eliminating values based on the aerosol extinction value, or eliminating values based on the magnitude of the SAGE II error bars for ozone. Figure 2.4(b) and (c) shows that error bars greater than 20% were typical when SAGE II ozone values were high relative to MLS values, whereas error bars less than approximately 12% were typical of background conditions after 1994 when SAGE-MLS differences were relatively systematic at the 10% level. This is because, in SAGE version 5.96 retrievals, an uncertainty arising from the inability to fully separate ozone and aerosol contributions has been added to the SAGE error bars for ozone. This uncertainty is approximately proportional to the aerosol loading. It is evident from the SAGE/MLS differences in Figure 2.4a that SAGE error bars still do not completely reflect the aerosol interference effect since the SAGE/MLS % differences exceed the SAGE error bars when the error bar greater than 25%. Nevertheless, from Figure 2.4a and 2.4b, it is seen that the SAGE-MLS difference stabilised when the SAGE error bar decreased to approximately 12%. A SAGE II error bar of less than 12% has been used to filter data in a number of the comparisons presented in this Chapter. However at the 100 hPa level, SAGE II error bars are often larger than 12% throughout the entire time series because of the small and highly variable ozone concentrations there. The temporal filter shown in Table 2.2 is preferred because it eliminates the few remaining measurements of unknown quality in 1991-1993. In addition, in order to eliminate a few spurious data points, all SAGE II ozone values with an error bar exceeding 50% have been eliminated.

Click here to download Figure 2.4 by anonymous ftp.

Figure 2.4 (a) Comparisons between SAGE II and coincident MLS ozone measurements binned in 35°N-55°N and 35°S-55°S and by month at 68 hPa. Note the large SAGE II values prior to 1993. (b) Average, reported SAGE II error bars as a function of time for the data in (a). (c) The percentage differences between SAGE II and MLS ozone observations, for data in (a), plotted as a function of the SAGE II mean error bars.

Pressure	55°N - 35°N	35°N - 35°S	35°S - 55°S
14 hPa	92.0	92.0	92.0

21 hPa	92.0	92.5	92.0
32 hPa	92.5	93.0	92.5
46 hPa	93.0	93.5	93.0
68 hPa	93.5	94.0	93.5
100 hPa	94.0	94.0	94.0

Table 2.2: Time filtering criteria for SAGE data. It is recommended that SAGE II ozone data from the time of the Mt Pinatubo eruption in June 1991 until the times indicated in the table be eliminated from trend calculations.

2.3.2. SAGE I Data Validation Studies

2.3.2.1. SAGE I/SBUV Comparisons

Figure 2.5 shows mean percentage differences between coincident SAGE I and SBUV observations where the SAGE I observations have been integrated over the SBUV (or equivalently Umkehr) layers using the NWS pressure/ temperature/geopotential height profiles which were supplied with each SAGE profile. Recalling that each SAGE I profile has received an altitude adjustment as a function of latitude, it is likely that the anomalous differences in layers 6-9 at 70°S reflect systematic differences between SAGE I and SBUV observations. This might in part be due to a small SBUV error at large solar zenith angle. At other latitudes, the differences between SAGE I and SBUV in layers 6-9 fairly systematically equal a few percent with the altitudinal and latitudinal structure of the differences being consistent with the structure of the differences between SAGE II and SBUV observations (see Figure 2.6). There were almost no sunrise observations for SAGE I (only approximately 3 months) and the SAGE II sunrise/sunset differences are addressed in Section 2.4.8. In layers 4 and 5 SAGE/SBUV differences vary strongly with latitude. This reflects the lack of vertical resolution of the SBUV observations below approximately 20 hPa and inadequacies of the SBUV a priori profiles.

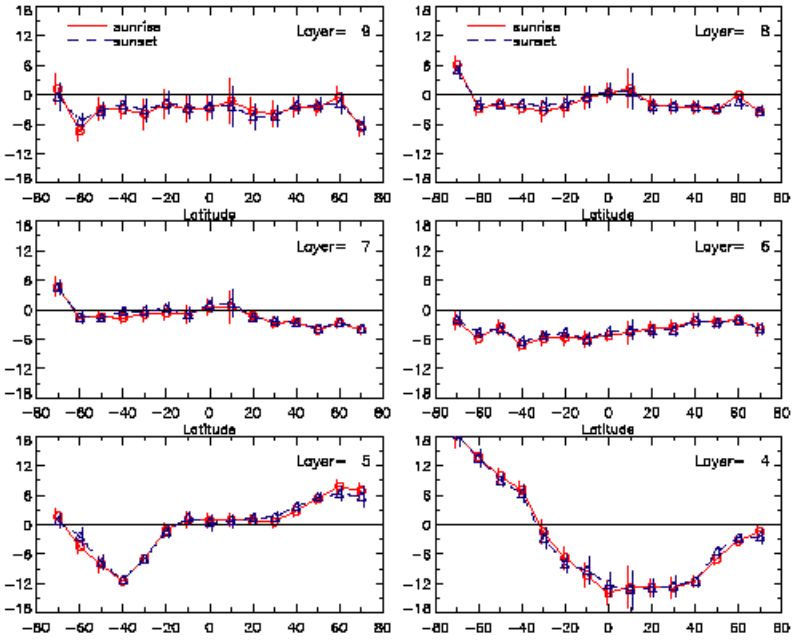


Figure 2.5 The differences between the altitude adjusted SAGE I observations and coincident SBUV measurements expressed as percentage of the SBUV values. SAGE sunsets and sunrises have been treated separately and the data have been placed in 10° wide latitude bins. The SAGE data have been placed on pressure levels using the NWS temperature profiles and then summed over Umkehr layers for compatibility with the reported SBUV values. The error bars are twice the standard errors of the differences.

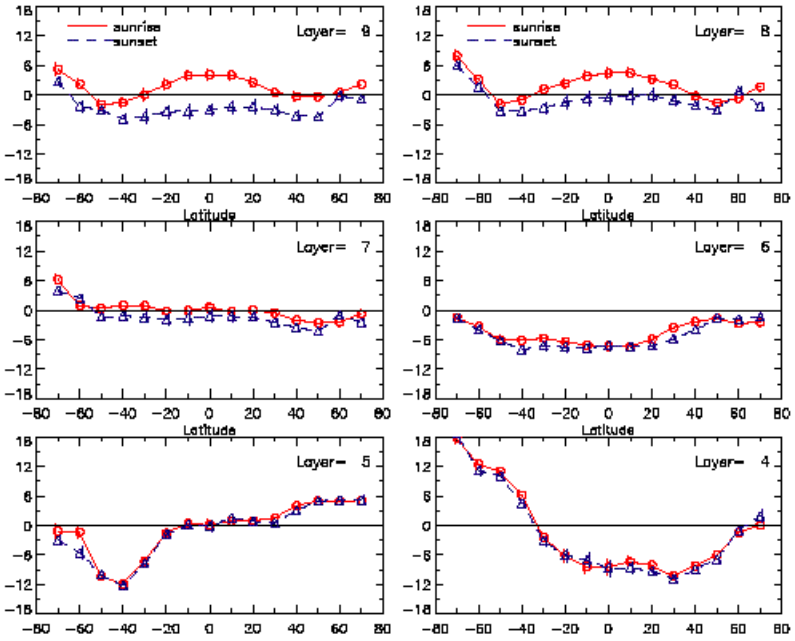


Figure 2.6 As for Figure 2.5, but for the SAGE II measurements from October 1984 to June 1990.

A second comparison between SAGE I and coincident SBUV profiles over almost 3 years of SAGE I operation was obtained by calculating the amplitude of the 12 month ozone periodicities. These annual amplitudes, expressed as a percentage of the layer mean values at each latitude, are shown in Figure 2.7 (a) and (b). The similarity in structure of the two plots is excellent, considering that the uncertainties in the estimates are approximately 3%. It is also noteworthy that SAGE II observations show a similar structure in the annual cycle (Figure 2.7 (c)). Both the vertical structure in the plots and the hemispheric differences are similar. On the basis of these comparisons, we suggest that there are no obvious deficiencies in the SAGE I observations in the middle and upper stratosphere.

SAGE I utilised fewer wavelengths than SAGE II and it is therefore more difficult to remove the effects of aerosols from the 600 nm channel than for SAGE II. Fortunately the stratospheric aerosol loading during the SAGE I period was unusually low. The SAGE II measurements are a little more accurate than those for SAGE I which could also impact the retrievals at very high and very low altitudes. Of most concern, however, for the SAGE I results below 20 km altitude is that there have been no new retrievals of the SAGE I profiles using the empirically-derived SAGE I reference height errors. It is therefore believed, as noted earlier, that SAGE I data below 20 km possesses significant uncertainties.

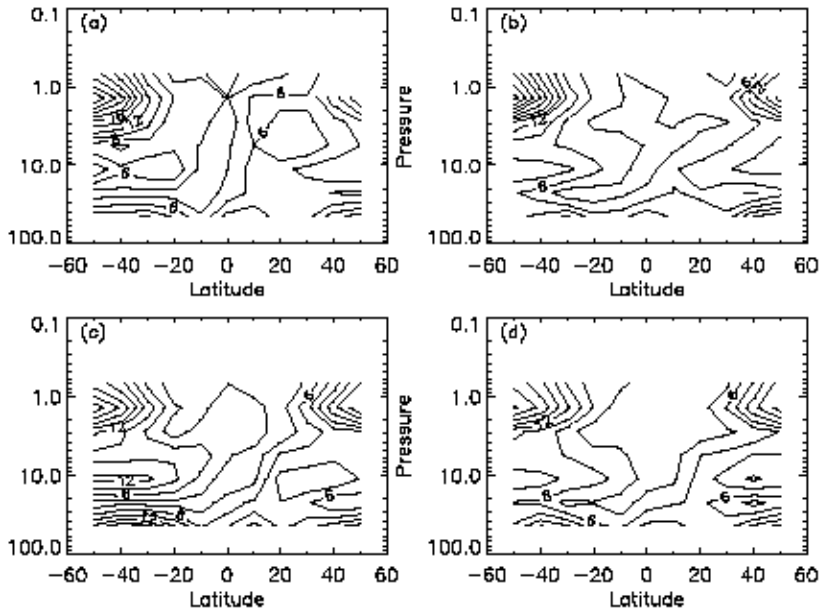


Figure 2.7 A comparison of the Umkehr layer mean amplitudes of the annual cycle in ozone (expressed as percentages of the annual mean values) from (a) SAGE I, (b) coincident SBUV, (c) SAGE II and (d) coincident SBUV2 observations.

2.3.2.2 SAGE I/Sonde Comparisons

We compare SAGE I, SAGE II, and ozonesonde measurements in this subsection relying on the assumption that the sondes can be used as a transfer standard between SAGE I and SAGE II. Good agreement would imply that SAGE I ozone measurements can be used in combination with SAGE II

results in the calculation of ozone trends. Comparisons were performed only at mid to high latitudes because of the vanishingly small number of SAGE I/sonde coincidences at lower latitudes. SAGE I ozone densities less than $5x10^{11}$ or greater than $1x10^{13}$ cm^{-3} were eliminated in doing the comparisons. Also, since the SAGE I data set is just three years in length, no attempt was made to compare SAGE I/sonde time dependent differences. Instead, the percent differences with three sonde stations (Hohenpeissenberg, Payerne, and Uccle) are presented as a function of altitude for both SAGE I and SAGE II.

Comparisons of SAGE I (asterisks), SAGE II (diamonds), and ozonesonde data are shown in Figure 2.8. The data sets were combined because of the small number of coincidences between the sondes and SAGE I. The comparisons above 27 km are not deemed significant due to potential ozonesonde errors. In the 27 to 25 km region, the SAGE I/sonde and SAGE II/sonde differences are outside the 2 sem of the differences, a fact that must be taken into account when combining the data sets. Below 25 km, the uncertainties overlap at all altitudes except 16.5 km, even though substantial differences exist in the comparisons at certain altitudes. For example, at 17.5 km, even though the bars overlap, the SAGE I/sonde difference is 14% whereas the SAGE II/sonde difference is about 7%; at 16.5 km, the differences are 5 and 10%. At 15.5 km the SAGE II difference is about 9% and the SAGE I difference is about 2%. Because the differences are quite large at many altitudes, caution should be exercised in using the combined SAGE I/SAGE II data sets in calculating long-term ozone trends at mid- to high latitudes.

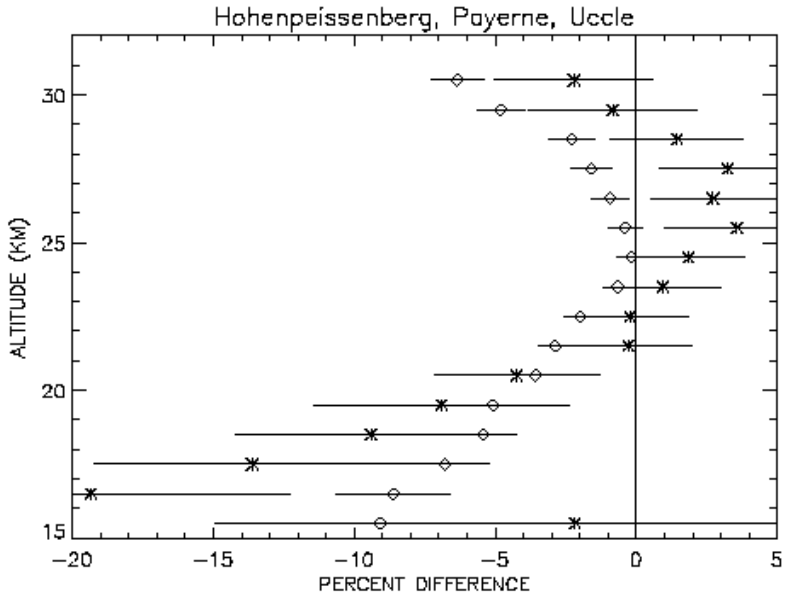


Figure 2.8 Percent difference in measured ozone between sonde and SAGE I (asterisks) and sonde and SAGE II (diamonds) as a function of altitude for the combined coincidences at Hohenpeissenberg, Uccle, and Payerne.

2.4. SAGE II Intercomparison Analysis

2.4.1. Ozonesondes

2.4.1.1. Density Comparisons

Information on sites used for this study is given in Table 2.3. As before, SAGE II ozone densities less than $5x10^{11}$ or greater than $1x10^{13}$ cm^{-3} were not used in doing the comparisons and the time filtering criteria presented in Table 2.2 were adopted. The SAGE II/sonde differences for 10 stations are shown in Figure 2.9 a to j. In general, the differences are small (3 to 5%) near 25 km but they increase with decreasing altitude reaching values from -5 to nearly -30% depending upon the site.

STATION	Latitude (deg.)	Longitude (deg.)	SAGE II	Sonde
Boulder	40°N	105°W	'84 to '95	ECC
Hohenpeissenberg	48°N	11°E	'84 to '95	BM
Uccle	51°N	4°E	'84 to '95	BM
Hilo	20°N	155°W	'84 to '95	ECC
Payerne	47°N	7°E	'84 to '96	BM
Lauder	45°S	170°E	'86 to '96	ECC
Tateno	36°N	140°E	'85 to '96	CI
Sapporo	43°N	141°E	'84 to '97	CI
Edmonton	54°N	114°W	'85 to '95	ECC
Goose Bay	53°N	60°W	'85 to '96	ECC

Table 2.3: Ozonesonde Stations to compare with SAGE. CI is Carbon-Iodine; ECC is Electrochemical Concentration Cell; BM is Brewer-Mast.

2.4.1.2. Latitude Variations of the Differences

Figure 2.10 shows the SAGE II/sonde differences for selected altitudes at each station plotted as a function of the integral of the SAGE II aerosol extinction. The integrated aerosol extinction is calculated

by integrating the SAGE II altitude dependent extinction data (version 5.96) collected at the times of the coincidences in altitude and averaging over the time period of the coincidences. The linear fits to the data in Figure 2.10 show strong negative slopes at all altitudes which are significant at the one sigma level. The reasons for this correlation are unknown, but could be either aerosol algorithm correction effects or altitude registration errors.

Possible Algorithm Errors:

Accurate retrieval of ozone densities at altitudes below about 19.5 km (the exact altitude depends on the aerosol loading and its variation with altitude) is dependent on the accuracy achieved in removing aerosol extinction from the 0.6 m m channel used by SAGE II for ozone sensing. The aerosol extinction at 0.6 m m is extrapolated using Mie theory from measured aerosol extinctions in the other SAGE II channels. Ozone errors will depend in a non-linear way on errors in aerosol extinction.

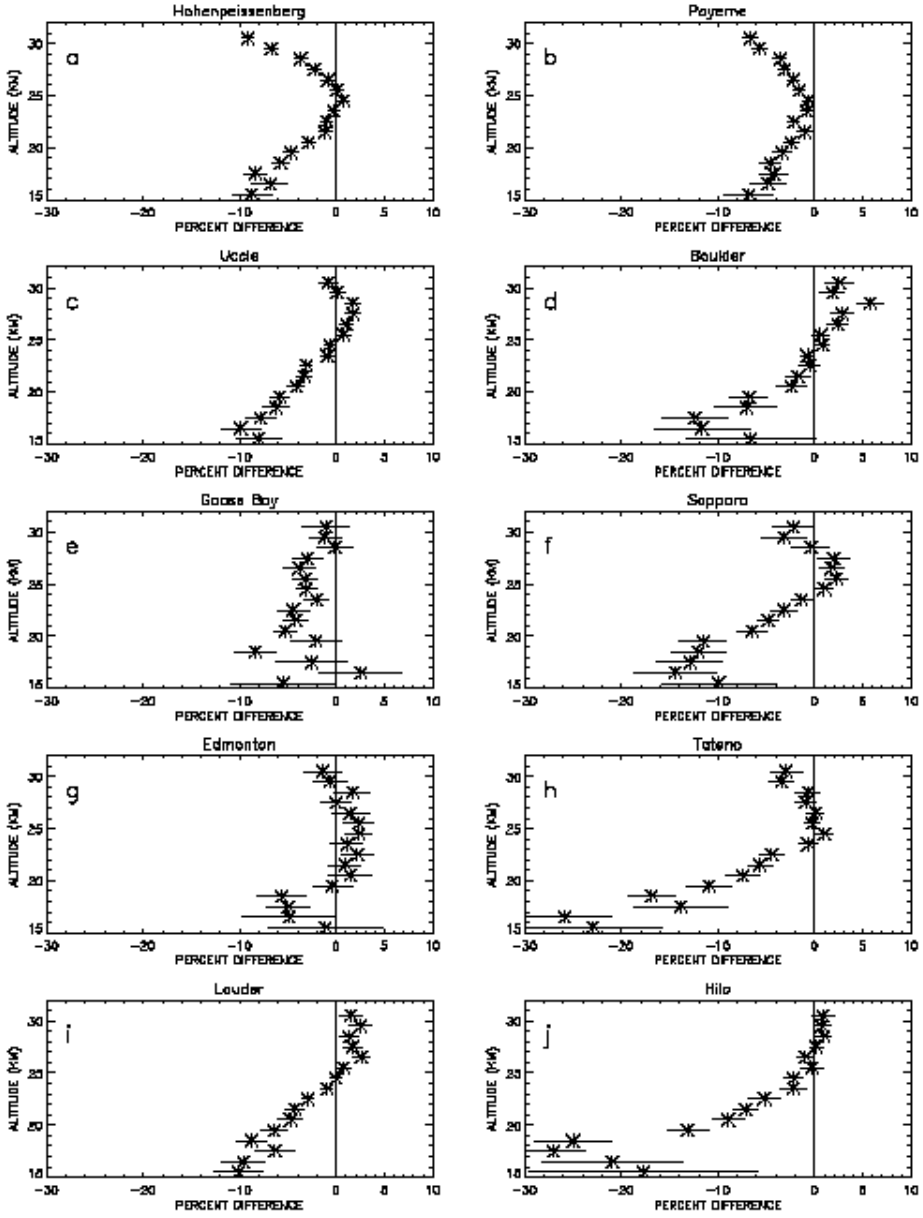


Figure 2.9. Percent differences between sonde and SAGE II ozone at 10 sites. Information on the site locations is given in Table 2.3.

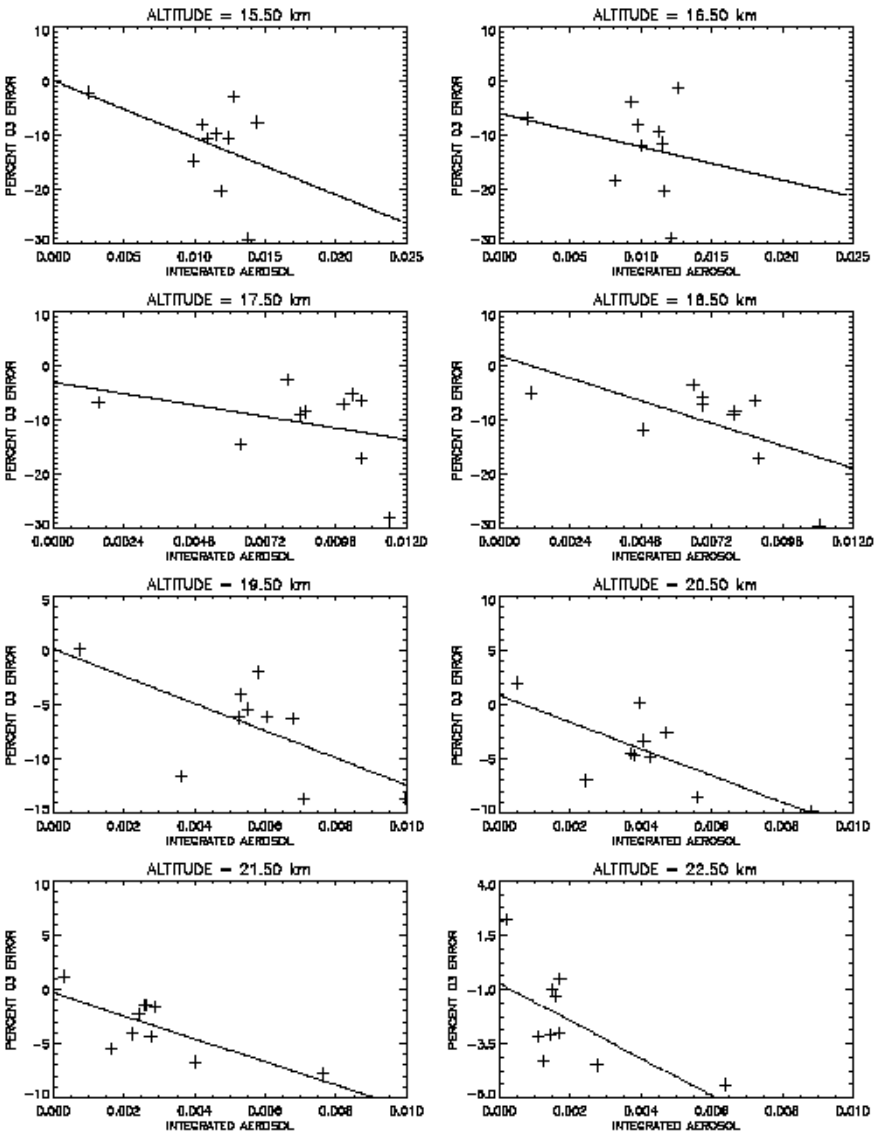


Figure 2.10. The percent differences between sondes and SAGE II as a function of the integrated aerosol extinction at various altitudes.

Possible Altitude Registration Errors:

An alternative explanation for the latitudinal difference (see Chapter 1) postulates the existence of altitude registration biases in the SAGE II data in the 15 to 20 km region. The biases, which may be as large as 0.5 km, are most likely caused by errors in the SAGE II solar edge detection algorithm in the presence of a strong gradient in the 1.02 m m extinction profile, which is the channel most sensitive to aerosol extinction. An altitude registration error would give an incorrect removal of Rayleigh scattering extinction from the 0.6 m m channel and therefore produce errors in the ozone determination.

2.4.1.3 Trends of Sonde Minus SAGE II Differences

Trends of sonde minus SAGE II differences as a function of time are calculated by a simple linear regression. Figure 2.11 shows an example difference time series for 20.5 km at Hohenpeissenberg along with the fitted line. The positive slope means that the sonde ozone values are increasing with time relative

to the SAGE II data. The difference vector has an auto-correlation coefficient of less than 0.2 and thus the individual data points are independent. The slope of the fitted line in Figure 2.11 is 0.68 ± 0.38%year^{-1} where the uncertainty is the 2 sigma error on the slope.

The altitude dependent trends in the SAGE II/sonde differences are shown in Figure 2.12 for the eight stations used later in Chapter 3 to compare SAGE II and sonde trends. These stations cover the latitude range from 36^{o}N to 56^{o}N. The trends of differences at individual stations vary widely in comparison to SAGE II. In the 20 to 25 km region the trends of differences vary from -1.7%year^{-1} at Edmonton to +1%year^{-1} at Sapporo. At lower altitudes, the variations are larger, ranging from values greater than -2%year^{-1} At Sapporo to about 3%year^{-1} at Boulder (although not significant at the 2 sigma level).

Another way to compare the data is to combine the matched pair differences from all eight stations into one data set, construct a time series at each altitude, and calculate the trend of the differences. These results are shown in Figure 2.13. Analysing the data in this way gives small but non-significant trends of differences (£ 0.3% ± ~0.3 to 0.5%year^{-1}) at all altitudes except 15.5 km where the value is about 0.83± 0.6%year^{-1}, 2s . It was decided to combine the matched pair differences from the various stations in this way because of the small number of coincidences at the lower latitude stations in the set. An alternate method would be to average the ozone change differences from the eight stations. This would give larger overall differences and would weight the stations with fewer coincidences the same as those with a large number of coincidences. Because the data used in Figure 2.13 are matched points differences at each station, spatial effects are not important.

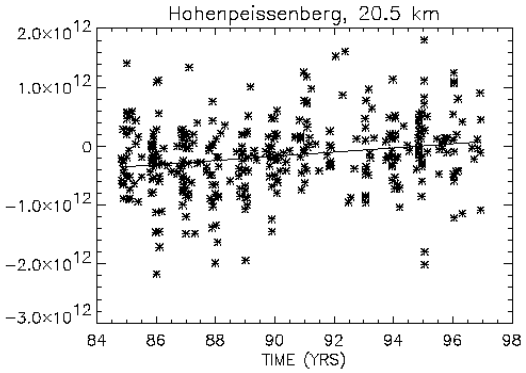


Figure 2.11 Differences in sonde minus SAGE II ozone as a function of time for Hohenpeissenberg at 20.5 km.

2.4.1.4. Summary and Conclusions

SAGE II/sonde ozone comparisons (see Figure 2.9) reveal small differences in the 25 km range, but differences increase with decreasing altitude, reaching maximum values of nearly 30% at two of the ten sonde sites. The sites with the largest differences are generally those where the low altitude ozone densities are smallest. There are at least two explanations for the altitude dependence of the differences between sonde and SAGE ozone measurements shown in Figure 2.10: algorithm issues or altitude registration errors. The sonde/SAGE II comparisons themselves do not shed light on the causal mechanism for the differences.

Trends of sonde minus SAGE II differences at individual sonde sites show large variability and care should be taken in using SAGE II ozone data to calculate ozone trends in narrow latitude ranges. On the other hand, using the combined data set covering the latitude range from 36 to 56° N and above 15.5 km, the trend of the differences are £ 0.3%/year (none of which are significant to the 2 sigma level except at 19.5 km). At 15.5 km, the largest difference occurs at 15.5 km where it is about 0.8%year^{-1} and is statistically significant.

2.4.2. Lidar

Extensive intercomparisons between stratospheric ozone lidars and the SAGE II instrument are presented in this section. Table 2.4 summarises the location of the lidar sites, the lengths of the lidar time series, and the institutions of the eight lidar teams participating in this study.

2.4.2.1. Difference Time Series

Time series of differences between lidar and SAGE II measurements were calculated for each site at 10 altitude levels : 15, 17.5, 20, 22.5, 25, 30, 35, 40, 45 and 50 km. The differences are formulated as (Lidar-SAGE)/SAGE and are given in %. No vertical scale or units conversion was necessary since the SAGE II and the ozone lidars provide ozone vertical profiles in the same units. The SAGE II data correspond to algorithm version 5.96. The matching criteria for selecting the SAGE II measurements around each site was ±2.5° in latitude, ±12° in longitude and ±24 hours or ±48 hours coincidence. The GSFC measurements were performed at various locations since this instrument is a mobile system designed for intercomparisons with other ozone or temperature lidars involved in the NDSC (Network for the Detection of Stratospheric Change); so the differences for the GSFC lidar correspond to various sites; GSFC, TMF, OHP, Hawaï and Lauder. Depending on the experimental systems, the whole altitude domain is not covered by all the lidar systems. Also, the altitude range covered can change over the length of the lidar data record due to improvements which allowed the measurements to be extended in altitude (e.g. implementation of Raman channels as discussed in Chapter 1). The altitude range can vary from one lidar site to another, depending on the meteorological conditions or the power of the laser sources. Both the SAGE II and the lidar measurements were perturbed in the lower stratosphere by the volcanic aerosol cloud emanating from the Pinatubo eruption. SAGE II data were screened to accept only data with errors lower than 12% in the low stratosphere. Most of the lidar systems, which did not include Raman channels could provide valid measurements only above the volcanic cloud. The upper altitude of the cloud reached 30 km or more in the fall of 1991 and decreased steadily with a time scale of about 8 months afterwards. Due to decreases in signal-to-noise ratio of the lidar signals with altitude, the resolution of the lidar measurements is lower in the upper stratosphere (generally above 40 km) requiring that the resolution of the SAGE data be degraded to match the lidar resolution in this altitude range.

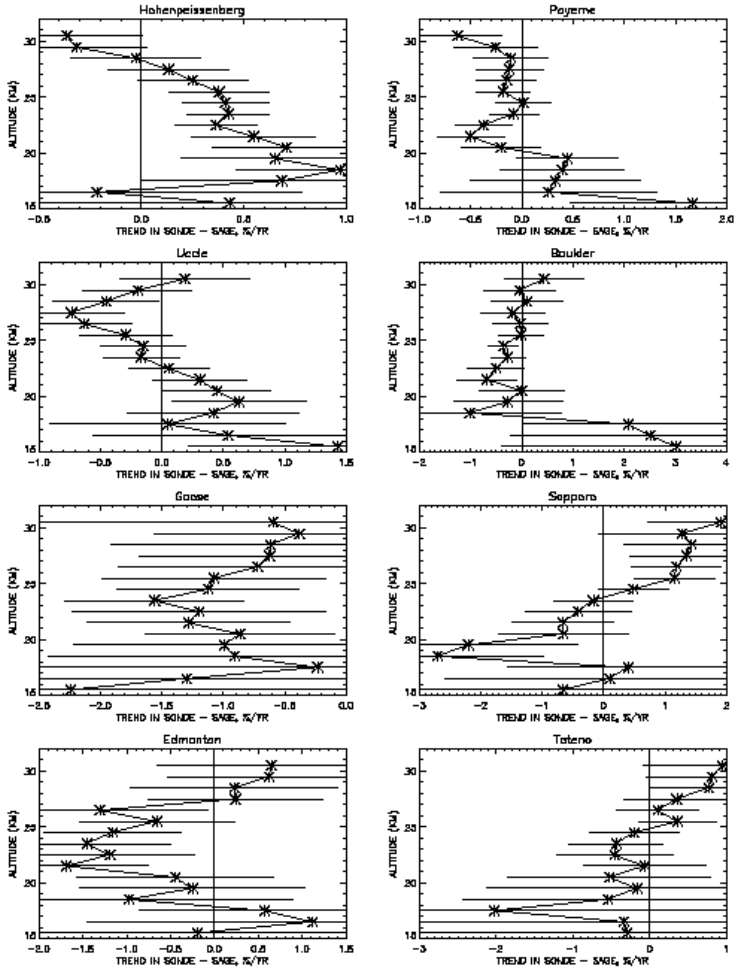


Figure 2.12. Trends in the sonde-SAGE II ozone differences as a function of altitude for eight sonde stations between 35° N and 53° N. The bars are 2 Standard Error.

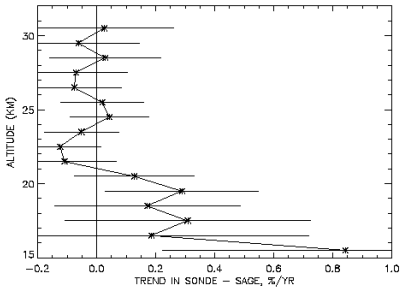


Figure 2.13 The altitude dependence of the trend in the sonde - SAGE II differences for the combined coincidences of the eight sonde sites shown in Figure 2.12.

Institution	Site	Location	Time series
RIVM	Lauder	45.05°S, 169.68°E	1992 - 1996
DWD	Hohenpeissenberg	47.48°N, 11.06°W	1987 - 1996
SA-CNRS	OHP	43.94°N, 5.71°E	1986 - 1996

NIES	Tsukuba	36.05°N, 140.13°E	1991 - 1996
MRI	Tsukuba	36.05°N, 140.13°E	1988 -1991
NASA GSFC	mobile instrument		1989 - 1996
JPL	TMF	34.40°N, 117.70°W	1988 - 1996
JPL	MLO	19.54°N, 155.58°W	1992 - 1996

Table 2.4: Lidar stations

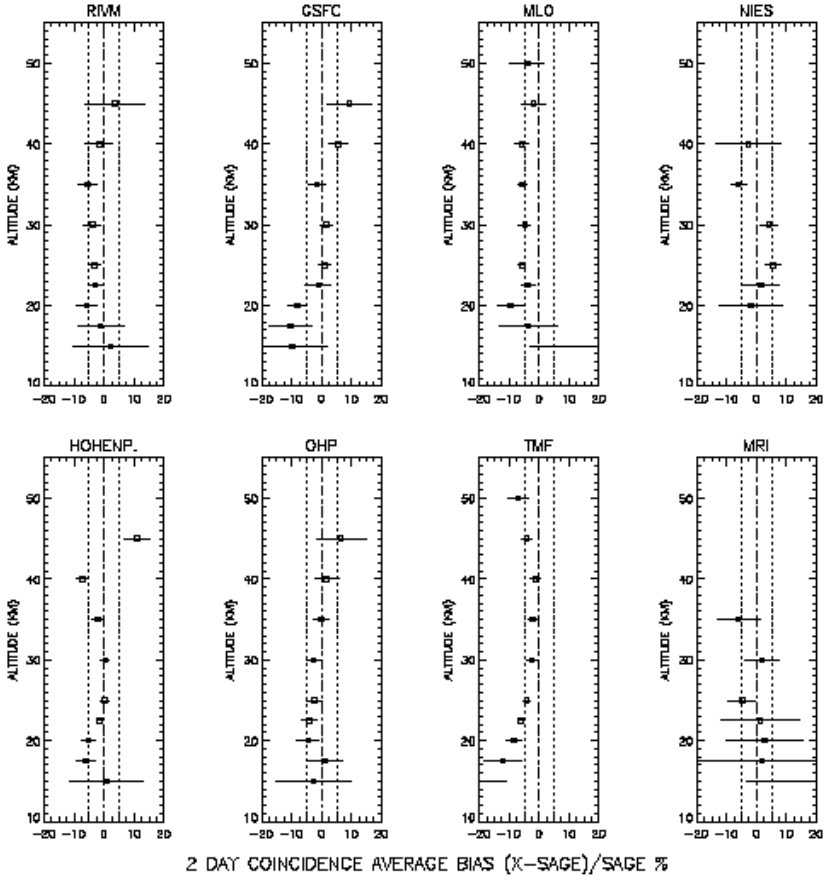


Figure 2.14. Average difference between the lidars and SAGE II as a function of the altitude for the ±2 day coincidence criterion. The error bars correspond to two sigma standard error.

2.4.2.2 Absolute Differences as a Function of Altitude

As in Section 2.4.1, first the results of the absolute differences are presented to an indication of the confidence level that can be placed on trend values. The vertical profile of the average differences for the two day coincidence criterion is displayed in Figure 2.14, together with the two sigma standard error and the number of coincidences. The results for one day coincidences occasionally show some small

differences in the average values and the corresponding error bars but these differences do not change the global picture of the results for the various data sets. The number of coincidences (see Table 2.5) is highly variable from one station to the other, depending on the length of the time series and the frequency of the measurements. The maximum number of coincidences for each station ranges from 10 (MRI) to 190 (DWD) for the two day coincidence criterion. Four stations provide more than a maximum of 50 events: DWD, OHP, TMF and GSFC.

The best agreement between ozone lidars and SAGE II is found between 22.5 km and 35 km. In this altitude range, the average difference is slightly negative and sometime significant at the two sigma standard error confidence level, but it generally does not exceed ±5%; the largest differences are observed in the MLO time series (tropics), where the negative bias stays around -5% from 25 km to 40 km. In the lower stratosphere, the variability is much larger at 15 km and some biases exceeding 20% in absolute value are found (MRI, TMF and MLO) but the lidars and SAGE II generally agree within the two sigma level. Around 20 km, the average difference is generally negative and significant with values ranging from -4.5 ±3.7% to -9.5 ± 4.1%, except for the Japanese stations where it is smaller. These results, like the sonde comparisons discussed in Section 2.4.1, raise a note of caution about using SAGE data for trends in the lower stratosphere below about 20 km. Above 40 km, the average difference increases and can be positive (DWD, GSFC) or negative (TMF, MLO). The 50 km altitude is hardly reached by the lidar measurements, except by the TMF and MLO instruments.

alt. km	RIVM	DWD	OHP	NIES	MRI	GSFC	TMF	MLO
15.0	39	71	41		3	46	13	22
17.5	44	131	51		4	53	77	35
20.0	44	147	57	18	5	59	109	40
22.5	44	152	62	29	5	61	118	45
25.0	44	158	89	29	10	63	135	45
30.0	44	186	131	36	10	66	176	46
35.0	44	189	147	35	10	69	187	46
40.0	44	180	131	7		67	185	46
45.0	40	146	56			60	180	46
50.0	13					48	134	40

Table 2.5: Number of lidar - SAGE II coincidences for the various stations

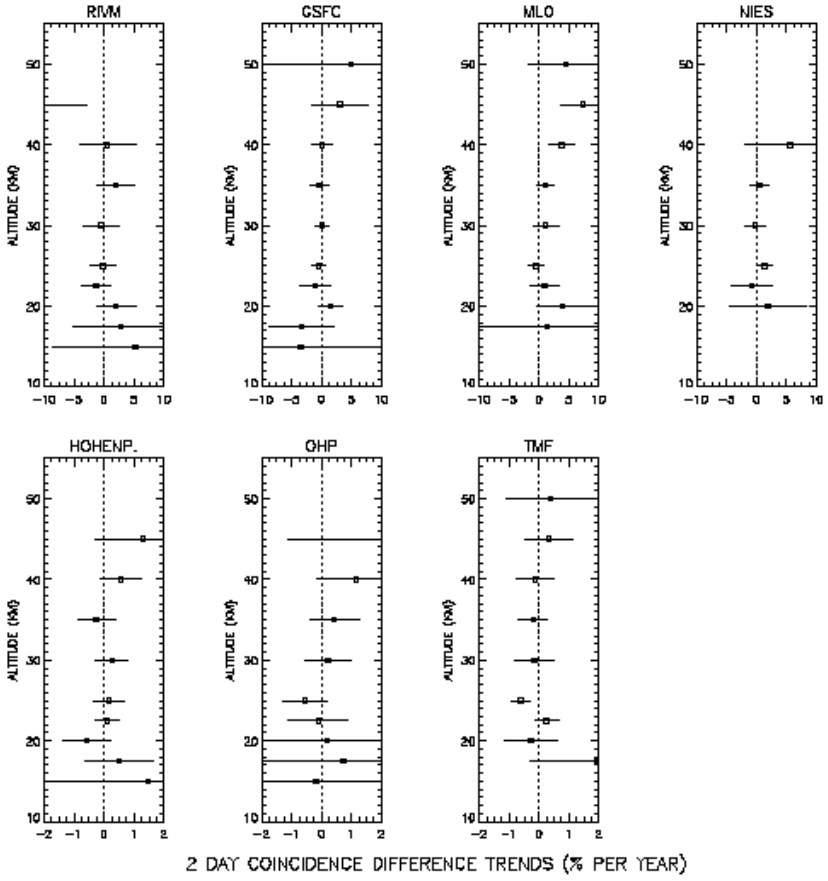


Figure 2.15. Trend of the average difference time series (± 2 days coincidence). The error bars correspond to the two standard error value of the trend.

2.4.2.3. Trends of Lidar - SAGE Differences

Trends of the difference time series are displayed in Figure 2.15 as a function of altitude, for each station and for the two day coincidence criterion. The error bars correspond to the two sigma standard error of the trend. Again, the results are not significantly different from the one day coincidence, so only the two day coincidence results will be discussed here. Likewise, the analysis focuses on the stations and altitude levels characterised by a number of coincidences large enough to provide reliable results on trends. Note that the scale on Figure 2.15 has been increased to ±2% in the case of Hohenpeissenberg, OHP and TMF.

In most cases, the lidar-SAGE difference time series show generally no trend significant at the two standard error confidence level. Due to the higher variability of ozone at altitudes below 20 km, the standard deviations are higher in this altitude range, and in general, the trends of the differences are larger than 1%year^{-1}. The only exceptions are for the OHP and Hohenpeissenberg time series, which show non-significant trends of < 1%year^{-1}. The trends are also higher above 40 km, due to decreases in the lidar measurement signal-to-noise. In these regions the trends generally exceed 1%year^{-1}. Over the 22.5 km to 35 km interval, the trend does not exceed 0.5 ±1%year^{-1} on average. Values less than 0.3%year^{-1} with standard errors not exceeding 0.4% are found in the case of the longest records.

2.4.2.4. Comparative Study of Lidars and SAGE II Changes Over the Entire Time Period

As another approach, difference of trends comparisons were done at each station determined by calculating the trend for each time series separately (SAGE II and lidar) without trying to match observations in time. The length of the data record used was defined by the shortest time series (the lidar in all the cases). The changes considered here correspond to a simple linear fit to the data points without removing the seasonal variations. The SAGE II data were selected according to the same spatial criteria as in the previous paragraphs (±2.5° in latitude and ±12° in longitude). Since the seasonal variation was not removed from the data, care was taken to calculate the changes over exactly the same period of time. However, due to the sampling of the SAGE II measurements at each site, some differences can arise, especially in the case of short time series. For some of the longer time series (DWD and OHP), significant differences in the trends are found around 20 km. In the case of the DWD lidar, significant differences are seen at 20 and 22.5 km, where the change is negative for the lidar while for SAGE II, it is positive at 20 km and slightly negative at 22.5 km. At 20 km, the difference reaches -1.6 ±1%year^{-1}. Concerning the OHP time series, a significant difference of -1.4 ±1.1%year^{-1} is observed at 20 km, where both instruments show a negative trend (larger in absolute value in the case of the lidar). Since the previous section showed that the SAGE II trends could be verified to about 0.5%year^{-1} by these two lidars, such differences should be checked with a more elaborate trend model.

2.4.2.5. Summary and Conclusions

The comparison between 8 lidars and SAGE II shows general agreement in absolute ozone density values for both types of measurements especially in the middle stratosphere (25-35 km). Poorer agreement is obtained in the lower stratosphere probably due to atmospheric variability and in the upper stratosphere (above 35 or 40 km, depending on the instruments) due to the lower signal-to-noise of the lidar measurement. Using coincidence time series corresponding to the 48 hours criterion and based on more than 40 events, the non significant average differences range from (-1.6 to 1.7)± 3% (two sigma standard error) in the middle stratosphere. The significant biases found in this altitude range are always negative (with the SAGE values being high with respect to the lidar data) and range from (-2.4 to -5.6) ±1.5%. Differences are largest on average in the southern hemisphere mid-latitude and northern hemisphere tropical stations. In the lower stratosphere, SAGE shows a positive bias with respect to the lidar especially at 20 km where it ranges from (-4.5 to -9.5) ±4%. At 15 km, no significant bias is generally observed but large differences exist and the variability is much larger.

Concerning possible drifts of SAGE II with respect to the lidars determined using matched coincidence events, the trends in lidar minus SAGE II differences for the longer coincidence times series (DWD, OHP, TMF) show generally non significant slopes in the 22.5 - 35 km altitude range varying from (0.1 to 0.5) ±0.4%year^{-1}. The shorter coincidence time series generally indicate trends within 1.5 ±1.5%year^{-1}. Below 20 km, the trends of the longer coincidence times series are within 1% to 2%year^{-1} and the two sigma uncertainty increases to 1% per year on average.

In summary, based on the SAGE II - lidar comparisons, the SAGE II trends from 20 to 35 km can be validated to about the 0.5%year^{-1} level. Below 20 km, although there is, in general, no statistically significant drift between SAGE II and the lidars the standard errors of both the absolute differences and the trends of differences are larger. Atmospheric variability is high in this range and the number of long-term records is smaller. According to the longest record, the SAGE II trends could be validated only to about 1%year^{-1} down to 17 km.

2.4.3. Umkehr

2.4.3.1. Introduction

The Umkehr data used for the analysis in this chapter came from the World Ozone and Ultraviolet Radiation Data Center (WOUDC) where they were all inverted in March 1997 from N-values using the Umkehr inversion algorithm [Mateer and DeLuisi, 1992] employing the uniform Sx error covariance matrix. This study further restricted the population of retrievals to those with small differences between observed and fitted N values (rmsres £ 1.3 N units) and with convergence criteria feps £ 0.15 and dfrms £ 0.01 in the inversion process. For such Umkehrs, the difference between observed and retrieved total column ozone is less than 1 DU in two-thirds or more of the retrievals. All Umkehr values are reported on the Bass and Paur [1985] scale. Table 2.6 lists the 19 Umkehr stations, grouped in broad latitude intervals, considered in this analysis, along with the record length of station observations within the 1979-1996 satellite period. The table also lists the fractional number of months represented within that period and the fraction of the Umkehrs available from the WOUDC database that passed the stringent convergence-criteria tests. Calculations performed in the Chapter 3 trends analyses indicate that the more restrictive convergence criteria do not significantly affect the computed trends from these Umkehr stations. Nevertheless, to insure the highest quality coincidence-comparison results, we applied the more stringent criteria to the Umkehr/SAGE comparisons reported in this chapter. The SAGE I data are altitude corrected according to Wang *et al*. [1996]. The SAGE II data are version 5.96. For all SAGE data, we omit ozone values with error estimates greater than 50% and also omit observations with an anomalous flag set to true; however, we do not screen any data based on aerosol amounts, which may affect the SAGE results below 20 km (layer 4).

Because aerosols in the lower-stratospheric aerosol layer affect the UV radiation scattered down from the zenith sky during the Umkehr observations, these aerosols cause retrievals of ozone in the upper stratosphere (primarily in layers above layer 6, that is > 30 km, i.e., well above the aerosol-layer altitude) to be underestimated during periods of high aerosol loading [DeLuisi, 1979]. Two different Umkehr aerosol corrections are applied separately through the analysis; the theoretically derived correction is described in Mateer and DeLuisi [1992], while the empirical correction is described in Newchurch and Cunnold [1994] and Newchurch *et al*. [1995, 1997]. At each station, for the population of measurements coincident with SAGE observations, we calculate the layer-mean ozone differences ±2 sem for individual Umkehr layers 4 through 8, layers 2+3+4 combined, layers 8+9+10 combined, and the columns for layers

WMO No.	Station	Country	latitude North	longitude East	Record length from 79-96	% months present	% stringent-convergence acceptance	Comment
	Northern mid-latitude							
190	Naha	Japan	26.20	127.68	2/79-12/96	83	64	level shift
10	New Delhi	India	28.65	77.22	3/79-12/95	61	50	compared
152	Cairo	Egypt	30.08	31.28	2/79-11/95	79	47	compared
7	Kagoshima	Japan	31.60	130.60	2/79-	80	55	level shift

					12/96			
14	Tateno	Japan	36.05	140.10	2/79-5/94	85	57	compared
252	Seoul	Korea	37.57	126.95	2/86-12/96	47	40	level shift
67	Boulder	USA	40.03	-105.25	2/79-12/95	93	64	compared
12	Sapporo	Japan	43.05	141.33	2/79-12/96	88	61	level shift
40	Haute Provence	France	43.93	5.70	9/83-12/95	69	60	compared
68	Belsk	Poland	51.84	20.79	2/79-11/96	66	55	compared
	Tropics							
214	Singapore	Singapore	1.33	103.88	4/79-12/96	70	47	few coincidences
187	Poona	India	18.53	73.85	3/79-12/95	55	52	few coincidences
31	Mauna Loa	USA	19.53	-155.58	3/84-12/95	65	60	unknown problem
245	Aswan	Egypt	23.97	32.45	4/85-12/96	62	56	level shifts
	Southern mid-latitudes							
265	Irene	South Africa	-25.25	28.18	7/90-12/95	26	36	few coincidences
27	Brisbane	Australia	-27.42	153.12	2/79-12/91	65	58	few coincidences
159	Perth	Australia	-31.92	115.95	2/79-12/95	76	67	compared
256	Lauder	New Zealand	-45.03	169.68	2/87-12/95	49	71	compared
	Antarctica							
101	Syowa	Japan	-69.00	39.58	9/88-11/96	21	57	few coincidences

Table 2.6: Information concerning Umkehr stations considered in this study including station number, name, country, latitude, longitude, record length within the SAGE I/II (1979-1996) study period, fraction of month observed during the study period, fraction of observations that passed the more stringent inversion convergence criteria, and comment regarding whether the station was compared to SAGE I/II or reason that it was not compared.

4-10, 3-10, and 2.10. These layers and layer combinations derive from the Umkehr retrieval averaging

kernels, which indicate the altitude and extent of the information content of this method [Mateer and DeLuisi, 1992]. We also calculate the linear regressions separately for the Umkehr, SAGE, and differences of coincident pairs of Umkehr-SAGE time series for all layers and combinations of layers expressed in both percent per year and Dobson Units [DU] per year ±1s confidence interval on the linear regression slope. The layer-mean ozone differences and the linear regression slopes are also separately plotted as a function of layer. All of these quantities are calculated separately for two time-difference criteria, 24 and 48 hours, and for the two aerosol corrections. We then average, within broad latitudinal ranges, the layer-mean ozone differences, layer regression slopes, and the regression slopes of the coincident differences for stations considered suitable for comparative analysis.

2.4.3.2. SAGE and Umkehr Time Series

As an example of the results at an individual station, Figure 2.16 shows the time series of Umkehr observations coincident with SAGE I and SAGE II in all specified layers and layer combinations plus the total ozone column indicated as layers 1+…+10. The regression lines are simple, linear fits to the data; they are not trend calculations in that they do not take into account any external influences (solar, dynamical, etc.) nor seasonality nor autocorrelation. These particular data were corrected for optical aerosol interference using the Mateer and DeLuisi factors. Using the Newchurch *et al*. aerosol corrections makes no significant (1s) difference in the regression slopes in any layer or combination of layers; however, the two corrections differ by approximately 2% in the average column ozone amounts. The close correspondence of the two independent aerosol correction methods enhances our confidence in the applicability of both methods for trend calculations. However, because the Mateer and DeLuisi factors conserve total column ozone, we recommend using those factors. These small differences between the aerosol correction methods are characteristic of all stations. Figure 2.17 shows quantities analogous to Figure 2.16 Umkehr measurements, but for the SAGE I and II observations. There is no aerosol correction applied to the SAGE data in these analyses. Because SAGE rarely measures ozone in layer 1, the layer 1-10 total-column time series is not shown. The data missing in both the Umkehr and SAGE time series result from missing SAGE observations and, therefore, missing coincidences (although the Umkehr data are complete in all layers). The persistent data gap between late 1981 and late 1984 results from the missing SAGE observations between the end of SAGE I and the beginning of SAGE II. Figure 2.18 shows the differences (in DU) between the Umkehr and SAGE at individual coincidences plus the linear regression through those differences. The uncertainties for all slopes in Figures 2.16 through 2.18 are 1-s confidence intervals. Regression slopes calculated for the SAGE II period only (not shown) and projected through the SAGE I period fit the SAGE I period well, supporting the conclusion that the SAGE I and SAGE II instruments are well calibrated with respect to each other. Because Figure 2.18 results from differences of coincident pairs of observations, secular geophysical variations are intrinsically accounted for and resulting regression slopes significantly different from zero indicate real differences between the two measuring systems.

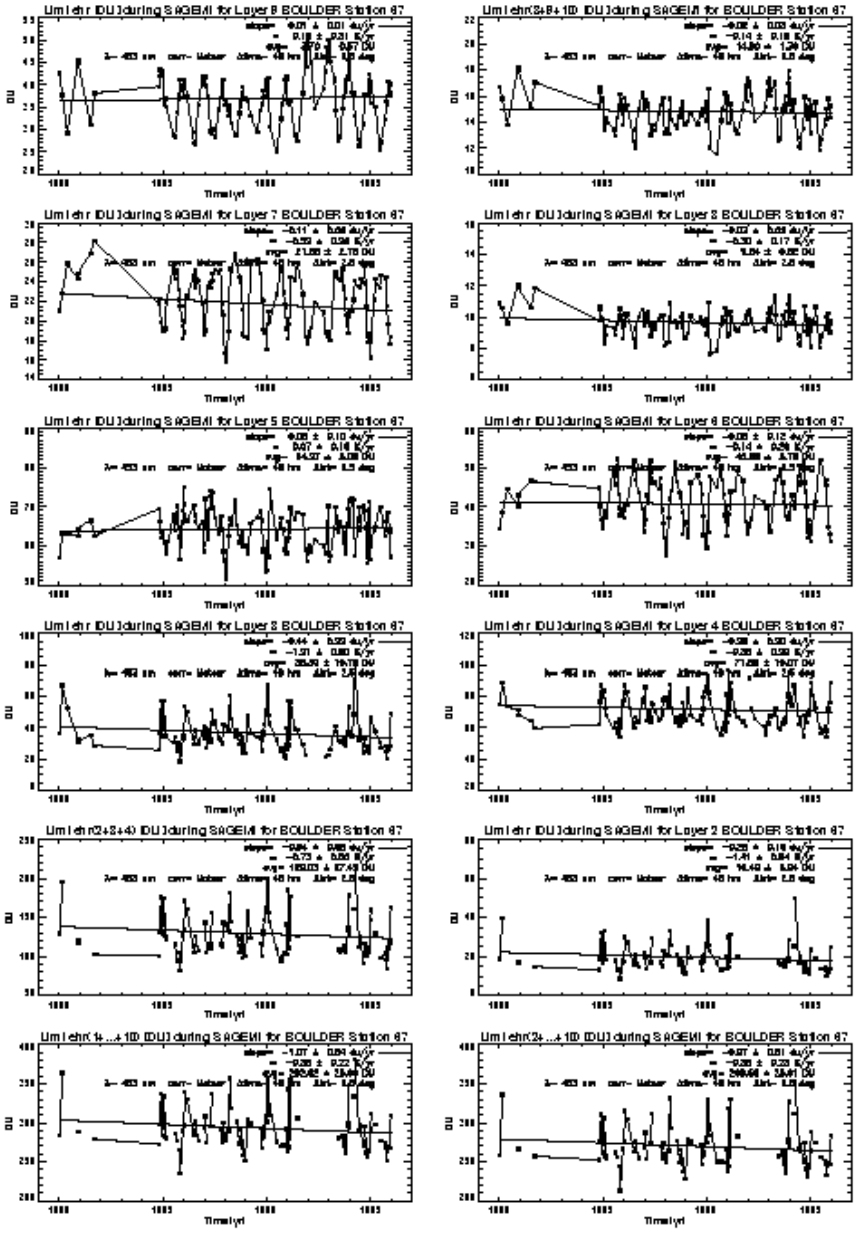


Figure 2.16 Time series of layer-ozone retrievals from Umkehr observations at the Boulder station coincident with SAGE I/II measurements. Panels display individual layers 2 through 9 and combinations layers 1 through 10, 2 through 10, 2+3+4, and 8+9+10 as indicated. The slope of the linear regression +/- 1 s.d. in DU year^{-1} and %year^{-1} (as a fraction of the indicated average layer-ozone amount) along with the maximum allowed time separation (48 hr) and latitude separation (2.5 degrees) are indicated. Umkehr ozone amounts were corrected for aerosol effects using the Mateer and DeLuisi [1992] factors.

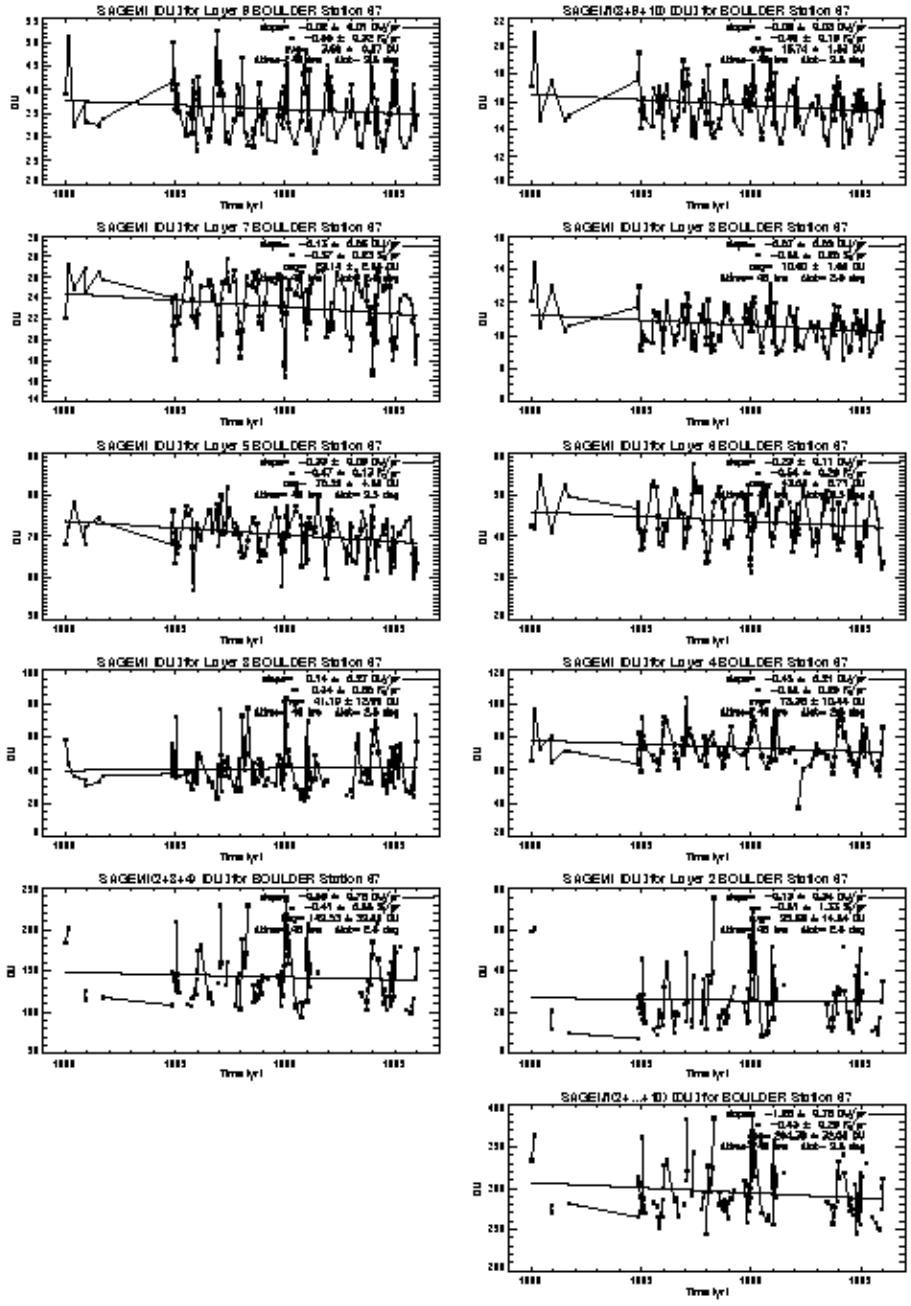


Figure 2.17 As in Figure 2.16, but for the SAGE I and II observations. Because SAGE rarely observes ozone in layer 1, the panel representing combined layers 1 through 10 is omitted. The data gap from 1982-1985 occurs between the end of SAGE I and the beginning of SAGE II.

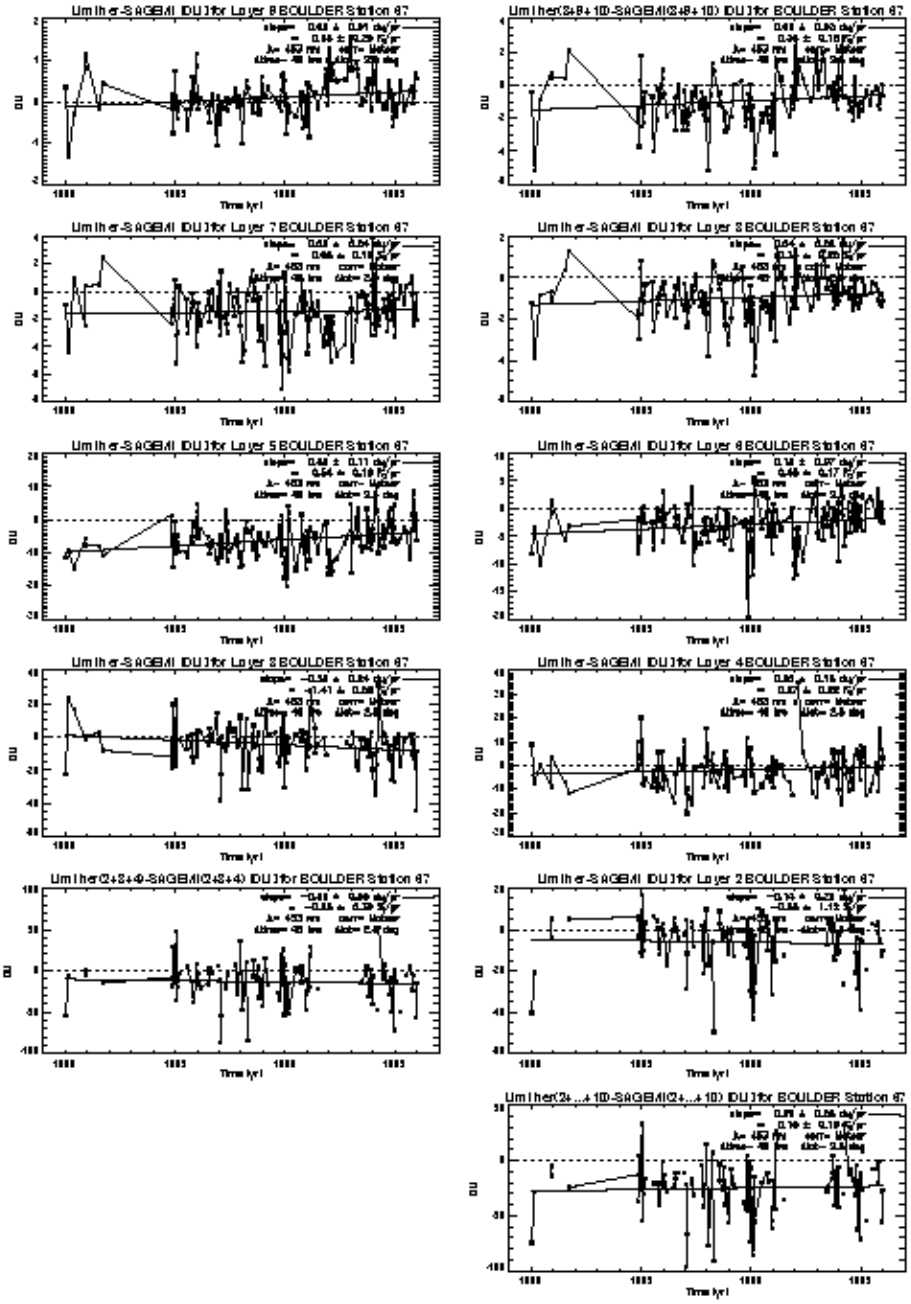


Figure 2.18 As in Figure 2.16, but for the differences (Umkehr - SAGE) in DU of the individual coincident pairs.

2.4.3.3. Layer-mean ozone and regression slopes

Figure 2.19 presents the layer-mean ozone differences ±2 standard errors of the mean (2 sem) during the SAGE I/II period for the 48-hour time intervals and the Mateer and DeLuisi aerosol correction method. The layer averages of 2 southern-hemisphere stations, Lauder and Perth, appear in panel (a), while the layer averages of 6 northern-hemisphere stations, New Delhi, Cairo, Tateno, Boulder, Haute Provence, and Belsk, appear in panel (b). Neither a more restrictive 24-hour time requirement nor an alternate aerosol correction method significantly affects these results. In the northern-hemisphere results, one can see significant SAGE-high biases in all layers, except layer 4, resulting in column biases of 7±1% for the

Mateer and DeLuisi aerosol correction (5±1% for the Newchurch aerosol correction). The SAGE-high biases in layers 2 and 3 result from the ozonesonde climatology that forms the Umkehr *a priori* profiles. These *a priori* profiles dominate the Umkehr retrieval in the lower layers. This bias is also consistent with the direct SAGE-ozonesonde comparisons in this report. The column biases are essentially the same as the column biases between Umkehr and SAGE version 5.93 reported by Newchurch *et al*. [1997], although the vertical structure varies somewhat from station to station, especially in the upper layers. Therefore, the average vertical profile of layer-mean biases depends, somewhat, on which stations are included in the average. SAGE II version 5.93 is the publicly available version at the time of this report. Version 5.96 will be made available by the time of the publication of this report. Differences between allowed time separations between 24 and 48 hours are not significant. The combination layers 2+3+4 results are plotted at level 3.5 and the layer 8+9+10 results appear at level 8.5 both in larger, disconnected symbols. The southern-hemisphere results possess very similar structure and magnitude, but larger variances.

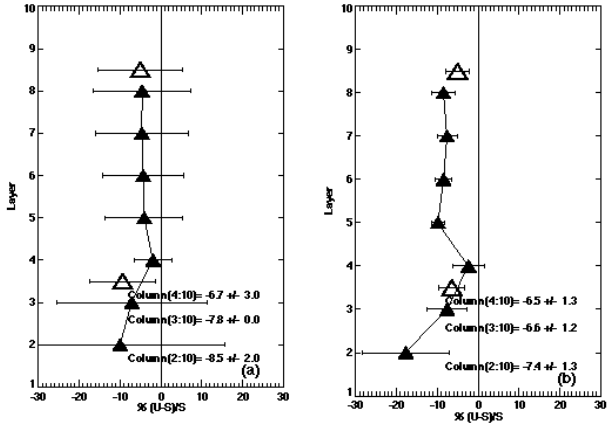


Figure 2.19. Layer-mean ozone differences expressed as Umkehr-SAGE I/II in percent ±2 sem of SAGE I/II layer-mean amounts for individual layers 2 through 8 (solid triangles), combined layers 2+3+4 (open triangles) plotted at layer 3.5, and combined layers 8+9+10 (open triangles) plotted at layer 8.5 for stations Perth (32° S) and Lauder (45° S) in the southern hemisphere in panel (a) and stations New Delhi (29° N), Cairo (30° N), Tateno (36° N), Boulder (40° N), Haute Provence (44° N), and Belsk (52° N) in the northern hemisphere in panel (b). Average differences ±2 standard errors of the mean for column amounts 2.10, 3-10, and 4-10 are also indicated. Coincident criteria are 2.5° latitude and 48 hours. All Umkehr observations are individually corrected for aerosol interference using the Mateer and DeLuisi correction factors.

Figure 2.20 displays the regression slopes of the Boulder station (as a typical example) for both 24-hour coincidences (squares) and 48-hour coincidences (circles) plotted against layer. Panel (a) presents the regression slopes of the Umkehr observations, panel (b) presents the slopes of the SAGE I/II observations, and panel (c) presents the slopes of the coincident-pair differences (Umkehr-SAGE). The large, disconnected symbols at level 2.5 represent the regression slopes of the layer 2.10 column amounts; the symbols at level 3.5 represent the layer 2+3+4 amount slopes; and the symbols at level 8.5 represent the layer 8+9+10 amount slopes. Smaller connected symbols represent individual layers. Results for both allowed time separations plotted together with 1-s uncertainties reinforce the similarity of these two time criteria. The Boulder Umkehr slopes are between 0.0 and 0.7%year^{-1}, depending on layer and SAGE I/II slopes are nearly constant at 0.5%year^{-1} above layer 3. The slope uncertainties are consistent between Umkehr and SAGE measurements and also between time coincident criteria.

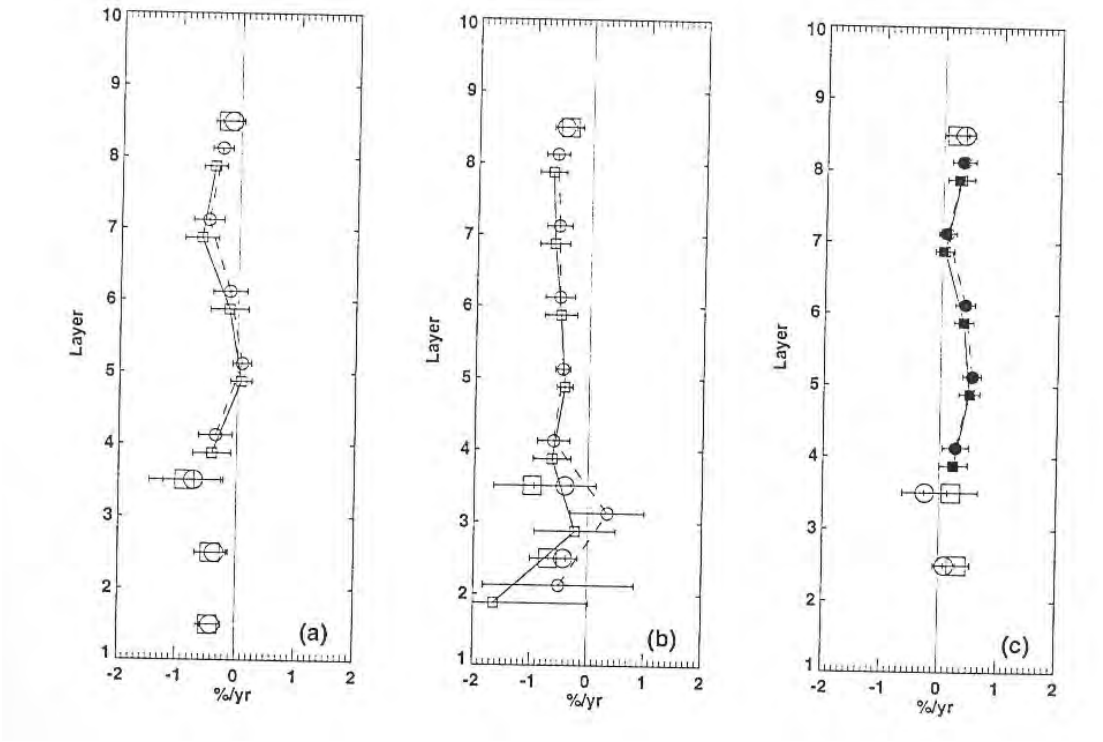


Figure 2.20 Regression slopes (%/year) of the layer-ozone time series of coincident observations ±1s confidence interval about the regression slope for the Boulder station (40° N, 255° E) . Panel (a) shows the slopes of the Umkehr observations; panel (b) shows the slopes of the SAGE I/II observations; and panel (c) shows the slopes of the coincident-pair differences (Umkehr-SAGE) within 2.5° latitude and 24 hours (squares) or 48 hours (circles). The disconnected open symbols represent layers 1 through 10 plotted at layer 1.5 (Umkehr only); layers 2 through 10 plotted at layer 2.5; layers 2+3+4 plotted at layer 3.5; and layers 8+9+10 plotted at layer 8.5. The connected symbols represent individual layers 2 through 8 for SAGE and 4 through 8 for Umkehr. All Umkehr observations are individually corrected for aerosol interference by using the Mateer and DeLuisi aerosol correction factors.

The results of these analyses indicated that certain stations should not be used for comparison to the SAGE regression slopes either because the coincidence density was too low or because discontinuities in the time series would produce anomalous discrepancies without intervention analysis. Indications for the 19 stations considered here are listed in Table 2.6. Because we chose not to remove the level shifts with interventions, these stations are not included in the latitudinal averages of layer-mean ozone nor averages of the regression slopes.

The northern mid-latitude (NML) average regression slopes and associated 95% confidence intervals appear in Figure 2.21 for Umkehr (panel a), SAGE I/II (panel b), and the coincident-pair differences Umkehr-SAGE (panel c). Panels (a) and (d) show the average slopes ±2 standard errors of the mean (i.e., an estimate of the variance between stations) of the Umkehr observations; panels (b) and (e) show the average slopes of the mean of the SAGE I/II observations; and panels (c) and (f) show the variance-weighted average slopes ±95% confidence intervals of the coincident-pair differences (Umkehr-SAGE within 2.5° latitude and 48 hours). The variance-weighted 95% confidence interval (also shown) is calculated as $2\,[(\Sigma s_i^{-2})^{-1}]^{1/2}$, where s_i is the standard deviation of the regression slope of an individual station time series. The largest discrepancy between Umkehr and SAGE regression slopes (panel c) occurs in layers 8 and 8^+ where the slope of the coincident-pair differences is 0.3±0.2%/year, significantly different from zero at the 95% confidence level. In all other layers and layer combinations, the absolute slope of the differences is less than or equal to 0.2±0.2%/yr. Similarly, the slopes expressed in $DUyear^{-1}$ present a consistent picture between Umkehr and SAGE observations, on average. The

analogous figures for the southern mid-latitude (SML) averages appear in Figures 2.21 panels d, e, and f. The southern mid-latitude averages show marginally significant discrepancies between Umkehr and SAGE only in layers 7 and 8^{+}. The vertical structure of the SAGE/Umkehr differences in the variance-weighted SML averages is similar to the NML results. Primarily due to the smaller sample size, the uncertainties in the SML are somewhat larger (~0.3%/year) than the uncertainties in the NML averages.

The 95% confidence-interval uncertainty in the total-column proxy (layers 2 through 10) regression slopes is on the order of 0.2%/year in the NML where the Umkehr/SAGE slope differences are not significantly different from zero and on the order of 0.3%/year in the SML where the slope differences are also zero. This close correspondence between SAGE and Umkehr total columns and ozone profiles suggests a fundamental agreement over time between these two measurements systems with some variation in the vertical distribution of the regression slopes at the 95% confidence level of 0.2%/yr.

2.4.3.4. Summary and Conclusions

In summary, we find significant layer ozone-amount biases (Umkehr lower than SAGE) increasing from 2±5% (2 standard errors of the mean) in layer 4 to 8±3% in layers 5 through 8 resulting in significant column (layers 2.10) biases of 7±1% (Umkehr lower than SAGE I/II). These layer-ozone biases differ by approximately 2% as a function of the 2 independent aerosol correction schemes employed. Comparisons of regression slopes through the time series of the coincident pairs of observations in layers 2+3+4, layers 8+9+10, the sum of layers 2 through 10, and layers 4 through 8 individually show no significant SAGE/Umkehr secular-drift differences greater than $0.3\%\text{year}^{-1}$ (most layers less than or equal to $0.2\%\text{year}^{-1}$) on average in NML. This result of close correspondence between Umkehr and SAGE regression slopes of times series of coincident observations holds equally well for both Umkehr aerosol corrections where we find no significant residual aerosol effect contaminating the regression trends. This result is consistent with the Chapter 1 estimate of the uncertainty in the aerosol correction being about $0.1\%\text{year}^{-1}$ in the corrected ozone amount. The 95% confidence interval in determining significant differences in the regression slopes (secular drifts) between SAGE and Umkehr observations is about $0.2\%\text{year}^{-1}$ in the northern mid-latitudes and 0.3%/year in the southern mid-latitudes. The 95% confidence uncertainty in the total column proxy (layers 2 through 10) regression slopes is on the order of $0.2\%\text{year}^{-1}$ in the northern mid-latitudes where the Umkehr/SAGE slope differences are marginally equal to zero and on the order of $0.3\%\text{year}^{-1}$ in the southern mid latitudes where the slope differences are equal to zero in the mean.

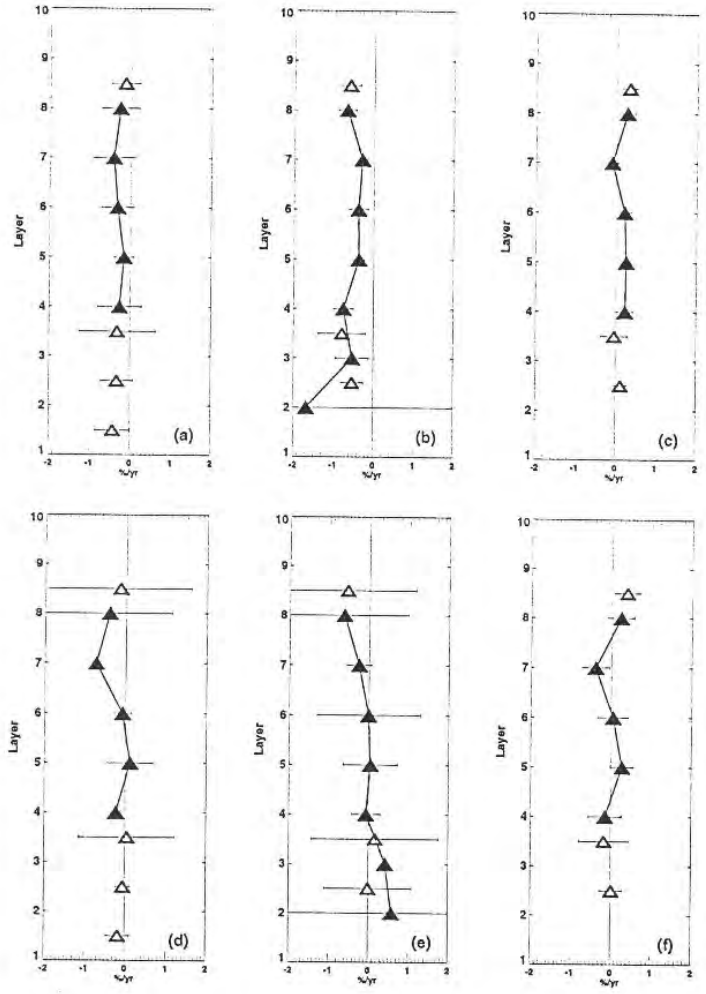


Figure 2.21 Hemispheric average regression slopes (%year^{-1}) of the layer-ozone times series of coincident observations for the northern hemisphere in the top panels and for the southern hemisphere in the bottom panels. These averages result from the same stations used in Figure 2.19. All errors bars shown are ±2 standard errors of the mean (i.e., an estimate of the variance between stations). Panels (a) and (d) show the average slopes of the Umkehr observations; panels (b) and (e) show the average slopes of the SAGE I/II observations; and panels (c) and (f) show the variance-weighted average slopes ±95% confidence intervals associated with the variance-weighted means of the coincident-pair differences (Umkehr-SAGE within 2.5° latitude and 48 hours). The variance-weighted 95% confidence interval is calculated as 2 $[(S\, s_i^{-2})^{-1}]^{1/2}$, where s_i is the standard deviation of the regression slope of an individual station time series. The disconnected open triangles represent layers 1 through 10 plotted at layer 1.5; layers 2 through 10 plotted at layer 2.5; layers 2+3+4 plotted at layer 3.5; and layers 8+9+10 plotted at layer 8.5. The solid triangles represent individual layers. All Umkehr observations are individually corrected for aerosol interference by using the Mateer and DeLuisi aerosol correction factors.

The combination of Umkehr errors, SAGE II uncertainties, and atmospheric variability allows statistically significant information, at the 95% confidence level, to be obtained regarding the correspondence of Umkehr and SAGE I/II layer-ozone trends to within approximately 0.2%/year.

2.4.4. SBUV2 and SBUV

It is desirable to create a composite data set from SBUV and SBUV2 measurements so that a comparison can be made against the SAGE measured ozone trends in the upper stratosphere. There is a 17 month overlap between the two SBUV data sets beginning in January 1989. The percent differences between the two sets as a function of time for Umkehr layer 8 are illustrated in Figure 2.22 for three latitude bands.

There are consistent offsets throughout most of 1989 but near the end of that year they begin to drift apart. Similar patterns are observed in other layers. Comparisons with Nimbus 7 TOMS total ozone measurements suggest that changes in the SBUV data are responsible for the change in behaviour. It is therefore recommended that the data for 1989 alone be used to determine an adjustment for the SBUV to SBUV2 transition of approximately 2% in layer 8 and by -4 to 4% in the other layers. Furthermore the 1990 SBUV data should be discarded.

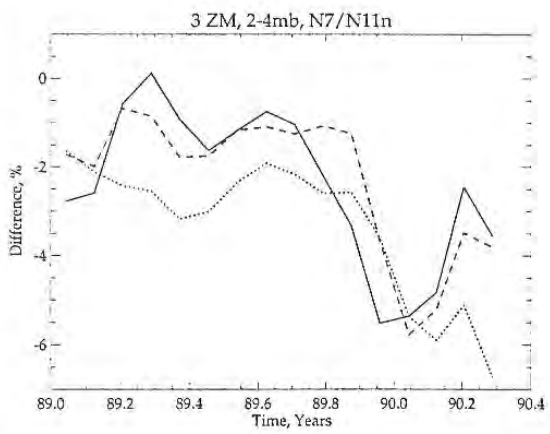


Figure 2.22 Time series of zonal/monthly mean differences (percent) between SBUV and SBUV2 for Umkehr layer 8. Solid line is for 30° N -50° N, the dashed line is for 30° S -50° S and the dotted line is 20° N to 20° N.

The mean differences between SAGE II and the coincident, adjusted SBUV2 from 1989 to 1994 are shown in Figure 2.23. The differences are just a few percent different from those between SAGE and SBUV measurements (see Figures 2.5 and 2.6 in section 2.3.2.1), but they show the same features, particularly in the latitudinal structure seen in layers 4 and 5. Of concern however is that the linear slopes of SBUV2 data from 1989 to 1994 are mostly larger than the linear slopes in SAGE data (shown in Figure 2.24). Although these slope differences are typically not significant individually at the 2 standard error level, there are obvious systematic differences in the global averages at levels 5 (+0.8 ±0.2%/year^{-1}), 7 (+0.7 ±0.2%year^{-1}), 8 (+1.3 ± 0.2%year^{-1}) and 9 (+0.6 ±0.3%year^{-1}).

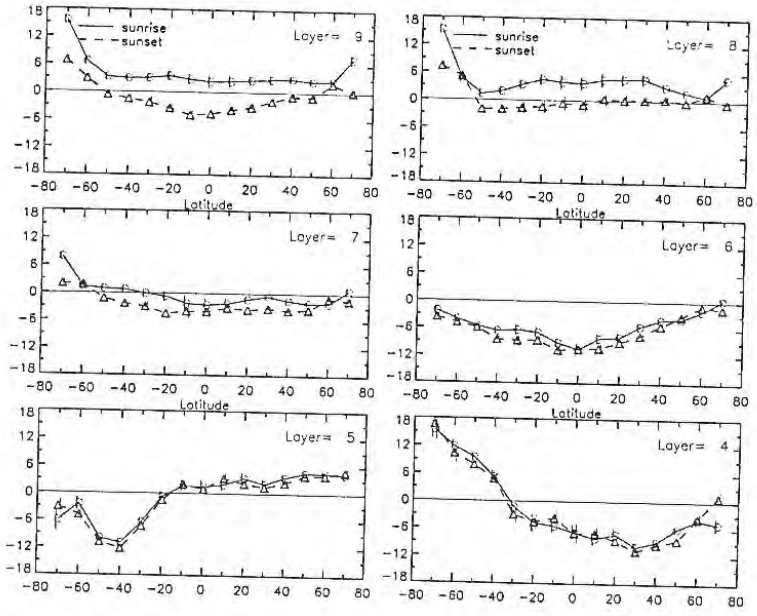


Figure 2.23 Mean percentage differences between SAGE II and coincident SBUV2 ozone values (SBUV2-SAGE) as a function of Umkehr layer and latitudinally binned in regions 10 wide. The error bars are twice the standard errors of the differences and SAGE sunrise and sunset measurements have been analysed separately.

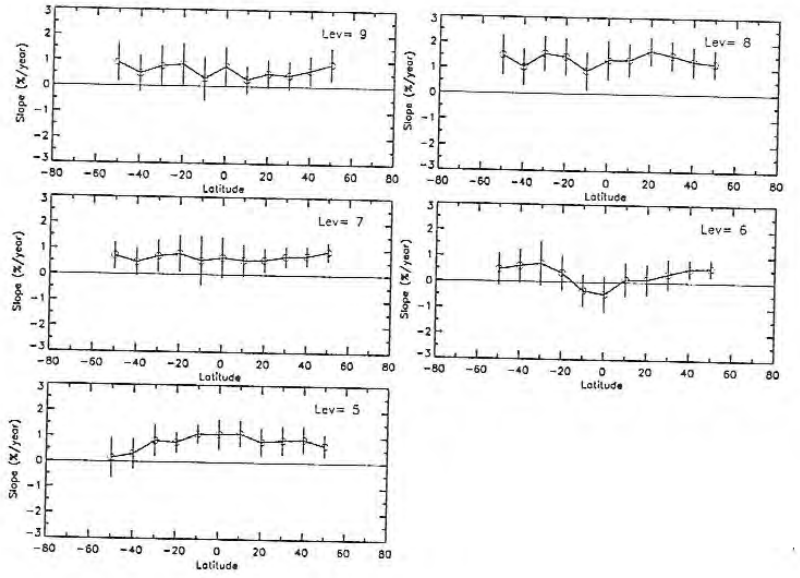


Figure 2.24 Mean slopes of SBUV2-SAGE coincident differences in Umkehr layers and in 10° latitude bins over the period of 1989-1994. Slopes are expressed as percent of the SAGE values per year. SAGE sunrise and sunset observations have been combined into a single time series.

These comparisons are based on the conversion of SAGE values to pressure surfaces which can cause changes in SAGE trends, whereas the other satellite instruments, such as SBUV2, intrinsically report ozone values on pressure surfaces. During the 1989 to 1994 period, geopotential height changes in the upper stratosphere (layers 7 to 9) could have produced anomalous trends of up to 1%/year in the tropics (see Section 2.2.2). However the effect would have been in the direction of increasing the SAGE ozone trends. It is concluded therefore that there must be other reasons for the larger slopes for SBUV2 than for SAGE.

It is to be noted that SBUV comparisons (see Chapter 1) show that SBUV2 may have had a calibration drift of approximately 0.5%year^{-1} at 1 hPa. Also the NOAA-11 satellite had changing equator-crossing times over time, and thus SBUV2 did not have a repeating pattern of solar zenith angles (SZA) from year to year as was the case for SBUV. These changing SZAs could have combined with interchannel calibration errors and modelling imprecision (e.g., in multiple scattering and temperature profiles) to create false changes in the profiles (see Chapter 1).

Despite these slope differences, the good overall quality of the two data sets may be seen in a comparison of the amplitudes of the annual cycles in the SAGE II and coincident SBUV2 measurements. These results are shown in Figure 2.7 (c) and (d) in section 2.3.2.1. As in the SAGE I/SBUV comparison the differences in the seasonal cycles are at the level of the uncertainties in the estimates (approximately 3%). There is remarkable similarity in the principal features of the two results but there is also possible evidence of a loss of vertical resolution in the SBUV2 observations at altitudes below 10 hPa. For example, the smaller maximum at 12 hPa in the southern hemisphere in the SBUV2 observations may be noted; however the SBUV observations for 1985-1989 (not shown) show a maximum similar to the SAGE II results in Figure 2.7 (c) in section 2.3.2.1.

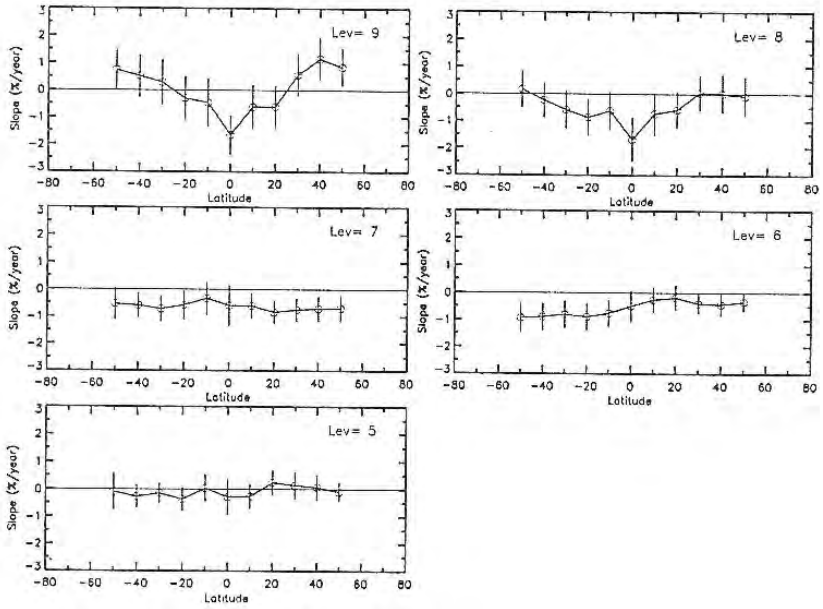


Figure 2.25 Mean slopes of SBUV-SAGE differences based on coincident SBUV measurements from 1984 to 1990. SAGE values have been summed over Umkehr layers and the differences have been accumulated in monthly bins which are 10 degree wide in latitude. Slopes are expressed as percentages of the SAGE mean values per year. SAGE sunrise and sunset differences have been combined.

Differences in the slopes calculated from SAGE II and coincident SBUV measurements from the end of 1984 to mid-1990 are shown in Figure 2.25. The differences are less systematic than for the SBUV2 comparisons, and in contrast to those results, they show that SAGE II exhibits more positive slopes than SBUV. The latitudinally averaged slope differences in layers 8-9 are approximately -0.2 ±0.4%year^{-1} (where the second term is the 2 standard error of the differences). The globally-averaged slope differences average -0.4 ±0.3%year^{-1} in layers 5 to 9. Again some of the tropical slope differences between SAGE II and SBUV in layers 8 and 9 in particular might be introduced by anomalous NWS temperature trends in the tropical upper stratosphere (and in contrast to the SBUV2 comparisons the possible errors are in the correct direction to explain the differences).

SBUV2 ozone profile retrievals are often given as 12 layer amounts. This does not mean that the instrument is capable of providing independent information for each of these layers. Since the SBUV2 ozone profile retrieval system is complex, it is difficult to determine whether trend results from other sources are consistent with SBUV2 results. However a study (Miller *et al.*, 1997), based on perturbing SBUV profiles with simulated retrievals of ozone changes measured by SAGE and the sondes, shows that SBUV2 should provide accurate estimates of the ozone trends in layers 6-9 and in the sum of layers 1-5. Instead of layer 9 alone, it may also be better to use the sum of layers 9 to 12.

The eruptions of El Chichon in April 1982 and Pinatubo in June 1991 impacted the ability of SBUV and SBUV2 to make accurate measurements of ozone. The aerosol load injected into the lower stratosphere increased UV scattering and enhanced the backscatter signal. It is recommended that one year following both the Pinatubo and El Chichon eruptions SBUV data should be deleted from any time series for trend analysis. A more detailed characterisation of the relationship between the size and profile of these errors and the aerosol loading is presented in Torres and Bhartia (1995).

A comparison of the linear slopes in SAGE I and SBUV data is not meaningful because the period of operation of SAGE I was less than 3 years. A more useful slope comparison is obtained from the combined SAGE I/SAGE II data versus the coincident SBUV measurements. The slopes in the monthly

mean differences are shown in Figure 2.26. SAGE data possess a more positive slope in layers 5 and 6 and a less positive slope in layers 8 and 9. The most significant difference is +0.4 ±0.2%year^{-1} in layer 9. However because of the offsetting differences in layers 5 to 7, it is concluded that overall there is no evidence of a drift in the SAGE (and overall SBUV) calibration of greater than 0.2%year^{-1}, based on SBUV comparisons over the period 1979-1990. This is perhaps the longest period over which comparisons can be made between coincident ozone measurements obtained by essentially two separate instruments. It is likely to provide the strongest constraint possible on the absence of a drift in the SAGE ozone calibration.

In summary, SBUV comparisons for the 1984 to 1990 pre-Mt Pinatubo period validate the globally averaged SAGE II trends to within 0.4 ±0.3%year^{-1} with SAGE II showing a larger trend and they validate the globally averaged data for the longer record from 1979 to 1990 (i.e. including SAGE I data in the time series), to the 0.2%year^{-1} level for layers 5 to 9. Comparison of trends with SBUV2 show that SAGE II trends are smaller than SBUV2 by 0.8±0.3%year^{-1} and it is believed that the problem is caused by a solar zenith angle correction problem with the SBUV2 data.

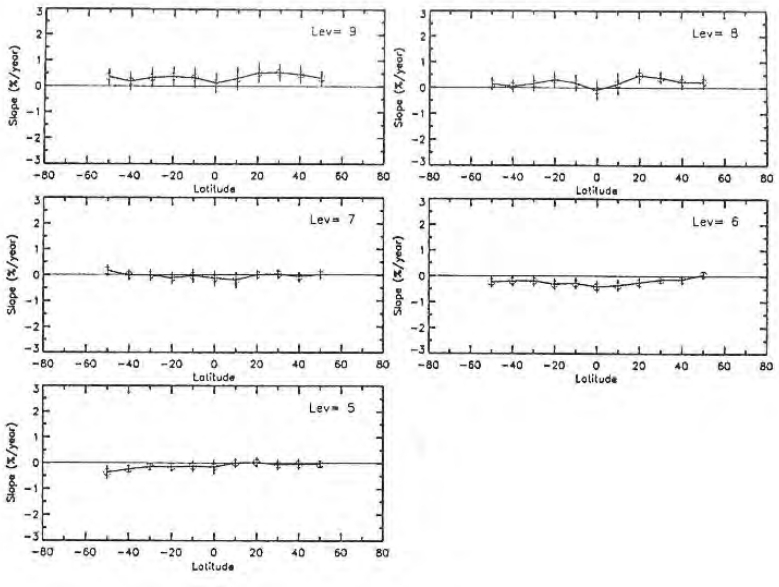


Figure 2.26 Linear slopes of the differences (SBUV-SAGE) between SAGE and coincident SBUV ozone observations from 1979 to 1990 expressed as a percentage of the SAGE Umkehr layer. Mean values were binned by month and placed in 10 degree wide latitude bins. Error bars are twice the standard errors of the differences in slopes.

2.4.5. HALOE

2.4.5.1. General Considerations

The SAGE II and HALOE (Russell *et al.*, 1993) experiments are in identical 600 km, 57^{o} inclination circular orbits. Because the orbital insertion times for these two experiments are different however, actual coincidences do not occur very often especially in the tropics because the ground tracks of the measurements sweep very rapidly through the region. In the range from 40° to 50° north or south, there are many more measurement coincidence opportunities because the measurement track is "turning around" in latitude and there is a tendency for the measurement location to "dwell" for some number of

days before beginning another "rapid" pole-to-pole traversal. Three kinds of intercomparisons were performed; (1) absolute SAGE versus HALOE profile comparisons made at "coincident" points, done primarily to form an idea of the overall quality of the two data sets; (2) comparisons of annual changes in ozone obtained from linear fits to each of the two data sets; and (3) comparisons of changes in SAGE and HALOE differences with time obtained from spatially and temporally matched pairs of measurements from the two experiments. In addition to linear fits to the unmodified data sets, annual and semi-annual oscillations (AO and SAO) were also removed while doing the linear fits to assess possible improvements in agreement between the data sets.

2.4.5.2. Absolute Profile Differences

Statistical comparisons of SAGE II and HALOE profiles in the tropics and mid-latitudes were performed for the July/August periods of 1992 and 1996 respectively. These years cover times of peak Mt. Pinatubo aerosol loading and near background levels. The coincidence criteria used to select the SAGE II/HALOE pairs were 1 day, 2^o latitude, and 12.5^o longitude. SAGE II was used as the reference in computing percentage differences. The first point to note is that between 25 and 50 km, the two satellite results agree to within 5% or less. At lower altitudes, the differences are larger. In 1992, the SAGE II data do not extend below about 22 km due to effects of aerosol interference. Differences with HALOE at that level are large (~20%). In 1996 (when the aerosol loading is reduced) the SAGE II data extend down to lower altitudes and differences as large as 20% are not reached until about 18 km. All the reasons for differences in the two measurements below about 20 km are not known, but it is clear that aerosol effects in one or both experiments are a major cause for the discrepancies. This suggests that it will be difficult to draw meaningful conclusions about the validity of SAGE II long-term changes in the region below 20 km using SAGE II/HALOE comparisons. The SAGE II and HALOE differences are similar in sign and altitude dependence to the SAGE II versus ozonesonde differences shown in Section 2.4.1.

2.4.5.3. Long-Term Changes

The measurements from each experiment were prepared for long-term change comparisons by binning the data into 10^o latitude intervals. Additionally for SAGE II, the data were filtered at altitudes below 25 km to remove any points with ozone errors greater than 12% to reduce possible aerosol effects on calculated long-term changes. The data in each latitude bin were then linearly interpolated to a 5 km altitude grid covering the altitude range from 15 km to 50 km. The ozone parameter analysed was number density. A least squares, linear fit was made to the entire SAGE II and HALOE data sets for the 1991-1996 and 1993-1996 time periods. These periods were selected to examine long-term change measurements during times when the Mount Pinatubo aerosol cloud was a maximum and when the effect was reduced. At each altitude, the linear change in percent per year was calculated using the middle point on the fitted line as the reference value. In the first set of comparisons, no attempt was made to remove natural cycles; we next removed the AO and SAO effects to assess how this would affect the results. In doing this, daily O_3 values for a particular latitude bin were averaged longitudinally to get zonal means and fits were done to determine coefficients of mean, AO, SAO, and linear terms. There was a slight improvement (~0.5%$year^{-1}$) in agreement with the natural cycles removed, so these data were used for the conclusions we draw.

Examples of the SAGE II and HALOE time series for the 30°-40°S bin are shown in Figure 2.27 for 20 km, 30 km, and 40 km altitude levels. Both time series show annual and semi-annual cycles at certain altitudes and the amplitudes of the variations are close in magnitude. At 20 km, clear differences between SAGE II and HALOE measurements exist in 1992, the year of peak aerosol loading. These general difference features exist for all latitude bins. Figure 2.28 shows the SAGE II and HALOE slopes as a function of latitude for each altitude level in 5 km increments for the 1991-1996 period with the AO

and SAO removed. The error bars on the plots are the 2 standard error mean (sem) slope uncertainties. Note that down to 30 km, the differences in indicated annual changes are £ 0.5% at all latitudes except 55°S where they are ~1%. This is also true for 25 km at many latitudes, but there are two latitude bands, i.e. 15°N and 55°S, where the differences are 1.5% to 2%year^{-1}. At 20 km, the differences are on the order of 2 to 4%year^{-1}. A similar plot for the 1993-1996 period showed only small differences except at altitudes at and below ~25 km where there were slight improvements. The global mean differences in indicated annual changes for the 60°S to 60°N range are given in Table 2.7. The results show that from 50 km down to 25 km, while there are no statistically significant differences, the values are less than 0.2%year^{-1} at all altitudes except 30 km where it is 0.4% ± 0.25%year^{-1}. At 20 km, the globally averaged differences are barely significant at the 1.4% ± 1.2%year^{-1} level. Below 20 km, the differences are large and no useful results are obtained in the tropics because of poor sampling.

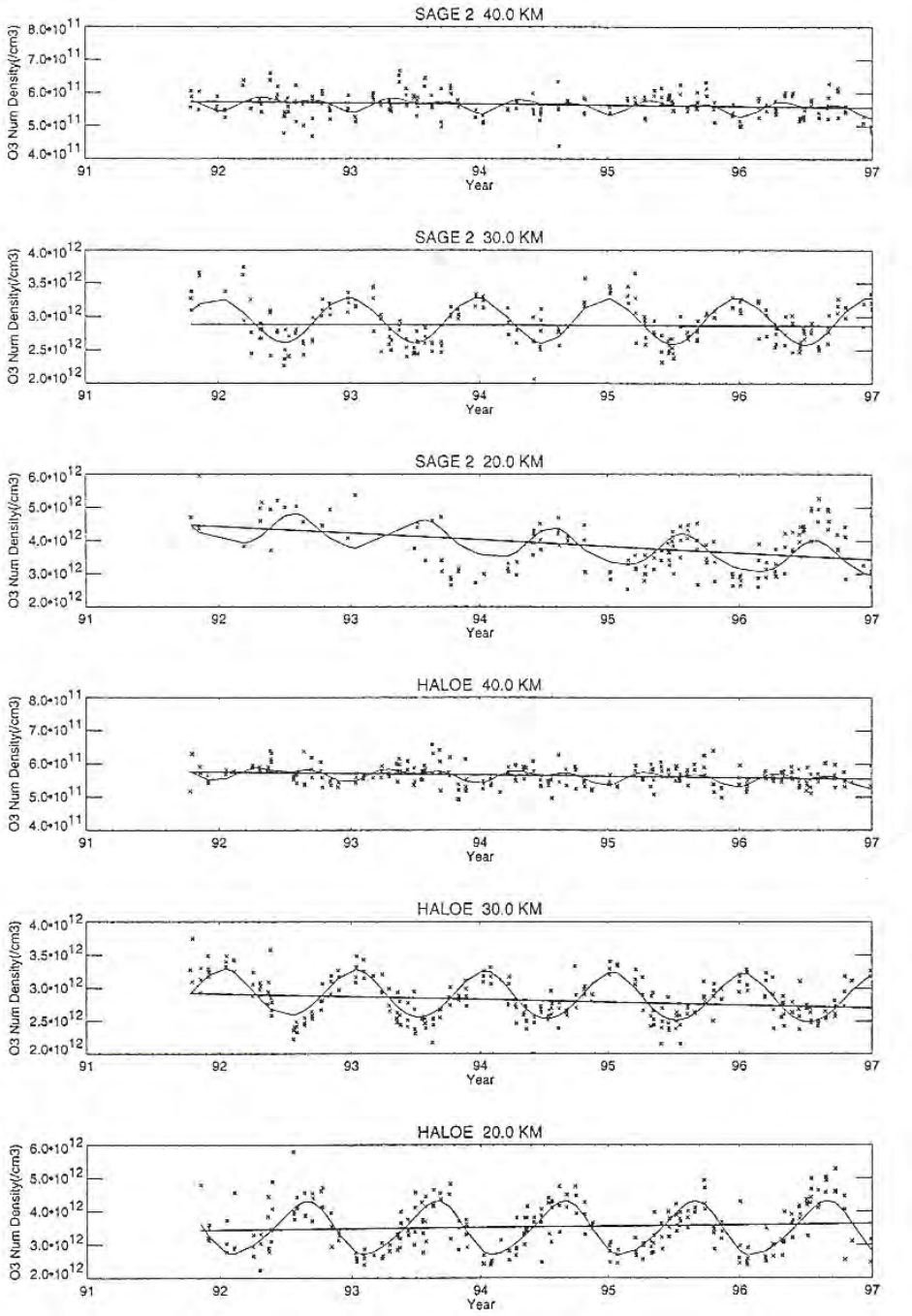


Figure 2.27 SAGE II and HALOE time series for the UARS period in the 30°S-40°S region for the 20 km,

30 km, and 40 km altitude levels. Annual, semi-annual, and linear fits are shown.

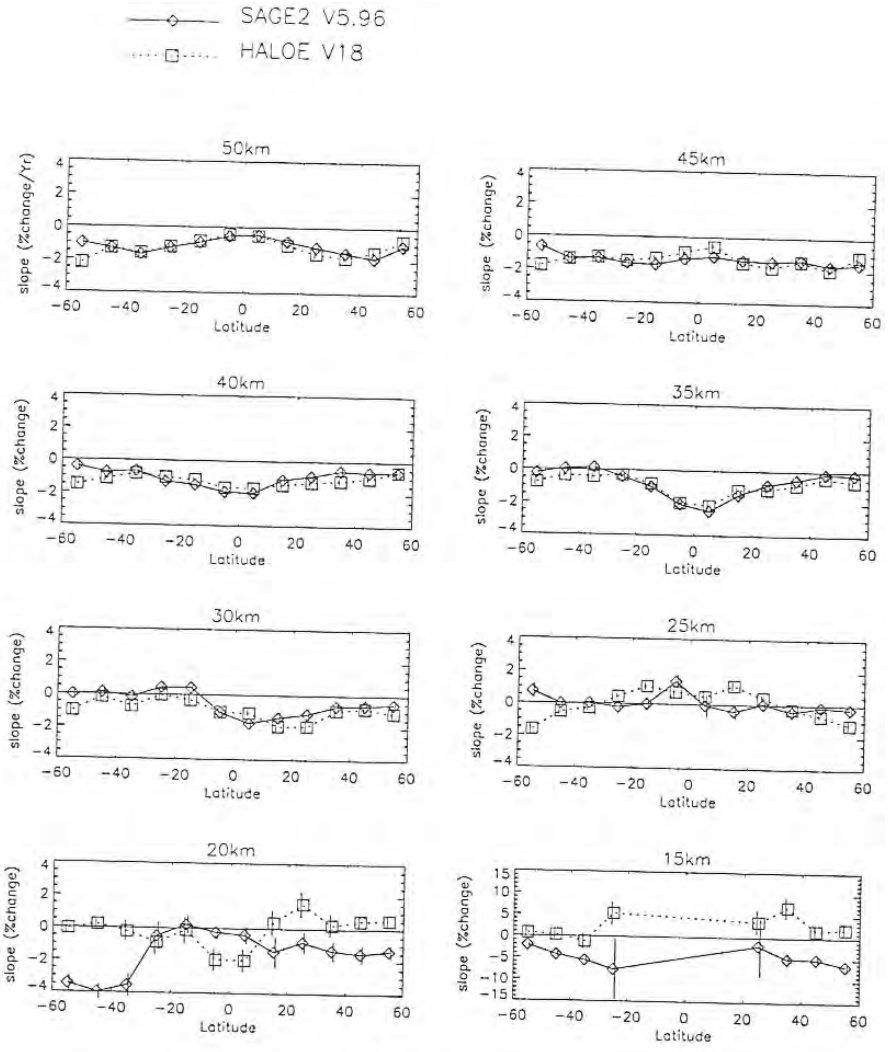


Figure 2.28 SAGE II and HALOE linear trends versus latitude for the 15 km to 50 km altitude range. Annual and semi-annual oscillations were removed while performing the linear fit.

Altitude (km)	Mean Difference	2 Standard Error Mean
50	0.09	0.22
45	0.01	0.27
40	-0.18	0.26
35	-0.15	0.20
30	-0.40	0.25
25	-0.08	0.62
20	1.38	1.16

Table 2.7: Global Mean (60°S-60°N) Difference (HALOE -SAGE II) in Indicated Annual Ozone Change in units of %year^{-1}.

2.4.5.4. Changes in SAGE II/HALOE Differences

The next comparison performed was to calculate differences in SAGE II and HALOE data "co-located" in time and space. The main purpose of this comparison was to examine the data to see if there was any time dependence of the differences, either seasonal or longer term, which would imply an instrumental or experiment effect that may be giving rise to false trends. In each latitude bin, profiles from the two experiments were matched using coincidence criteria mentioned above. Each data set was interpolated to a 1 km altitude grid and percent differences were calculated using the average of SAGE II and HALOE as a reference.

The percent difference values were then interpolated to the standard 5 km altitude grid used in this report. One problem with this approach is that the number of data points used for calculating long-term changes dramatically reduces when a matching criterion is applied; so much so in the tropics, that meaningful comparisons cannot be made. The latitude range with the most coincidences is 30°-40°N or S. A marginal number of coincidences exist at 20 km and there are not enough at 15 km to use in this study. There appears to be a slight seasonal dependence of the differences but without other corroborating data (which has not been borne out in this study), no conclusions can be drawn. Over the 25 km to 50 km range, the only latitude ranges where there are consistent statistically significant slopes to the difference time series are 40^0-50°S (~1.5%± 1%year^{-1}) and 60°-70°S (0.5± 0.5%year^{-1}). Below 25 km, the differences are not statistically significant. In each latitude bin, the change is negative, i.e. SAGE II appears to be getting progressively larger with time than HALOE or vice versa. When viewed globally, i.e. when the same analysis is done on a single time series of HALOE-SAGE II differences using the entire 60°N-60°S range, the results are very similar to those shown in Table 2.7.

2.4.5.5. Summary and Conclusions

Comparisons of annual changes determined from simultaneous annual, semi-annual, and linear fits to the full SAGE II and HALOE data sets over the 1991-1996 time period show agreement in the 50 km to 30 km interval to better than ± 0.5%year^{-1} at all latitudes except 55°S where they are ~ 1%year^{-1}. Agreement to £ ± 0.5%year^{-1} is also obtained for 25 km at all latitudes except 15°N and 55°S where they are 1.5% to 2%year^{-1}. If the AO and SAO effects are not included in the fit, the agreement degrades by about 0.5%year^{-1}. This suggests that sampling differences between the two experiments may be falsely indicating long-term change. The globally averaged mean differences in annual change from 50 km down to 25 km are 0.2± 0.2% -0.6%year$^{?1}$ (-0.4 ± 0.3%year$^{?1}$ at 30 km). Using matched SAGE II/HALOE pairs, plotting percent differences versus time, and then doing a linear fit through the difference points shows no statistically significant SAGE II drifts with time relative to HALOE over the 50 km to 25 km interval except at 40°-50°S (~1.5± 1%year^{-1})and 60°-70°S (~0.5± 0.5%year^{-1}). No statistically significant drifts are observed at any other latitudes. On a global basis, the averaged "trends" of the HALOE-SAGE II difference time series at various altitudes show similar results as those obtained using non-coincident data. Because of the large differences in annual change observed below 25 km between SAGE II and HALOE and without knowing the cause of the differences, no conclusions can be drawn regarding the validity of SAGE II data for trend calculations in this region.

2.4.6. Intercomparison of SAGE II and HALOE Retrievals Using Dynamical Mapping Techniques

2.4.6.1. Introduction

The concept of employing dynamical techniques for the intercomparison of sparse data sets was introduced in Section 2.2.5. Here the Coordinate Mapping (CM) technique is used to compare the UARS-period mean ozone values and simple linear trends in SAGE and HALOE observations. Section 2.4.6.2 describes the technique and discusses the advantages and disadvantages of its use as a data intercomparison tool. In Section 2.4.6.3 the effectiveness of the CM technique is verified through comparison with the more conventional zonal mean approach for a test period in March 1995, during which SAGE and HALOE made observations at similar latitudes. Results of the CM analysis for the UARS period are reported in Section 2.4.6.4. Finally, Section 2.4.6.5 introduces the Trajectory Mapping (TM) technique, a second dynamical approach to data intercomparisons which has promising potential for future work.

2.4.6.2. Coordinate Mapping

The general concept of employing quasi-conservative coordinate mapping as a means of increasing the utility of available stratospheric trace gas data dates back to McIntyre (1980), who referred to it as a "modified Lagrangian mean" approach to stratospheric analysis. The method, formally developed by Schoeberl and Lait (1993), has been utilized in recent years as a powerful tool with numerous applications (e.g., Lait *et al*. 1990; Randel and Wu, 1995; Atkinson and Plumb, 1997), including the intercomparison of sparse data sets at high latitudes, where the ozone field tends to be most variable (Redaelli *et al*., 1994). The technique involves a transformation of trace gas data from 3-D physical space (latitude-longitude-pressure) into 2-D potential vorticity (PV) ? potential temperature (PT) coordinates, a flow-following "dynamical" reference frame. The 2-D character of this reference frame makes it analogous to a zonal mean analysis. However, because the dynamical coordinates move with the meteorology when the flow is adiabatic, it is possible in this reference frame to reduce the contribution to "local" variance in ozone concentrations due to short-period meteorological variability. Also, since the instantaneous large-scale stratospheric wind field is typically closely aligned along isopleths of isentropic PV, ozone mixing tends to be very rapid along a PV contour as compared to mixing across it. The isentropic ozone mixing ratio contours therefore closely mimic those of the PV [Leovy *et al*., 1985; Butchart and Remsberg, 1986; Manney *et al*., 1995], so that over suitably short periods (compared to photochemical and diabatic time-scales), "longitudinal" and temporal variations in this dynamical reference frame will be small. Therefore, comparing sparse data sets in this coordinate system should reduce biases introduced by the different time and spatial sampling of the individual instruments. The reliability of the CM results is dependent upon the quality of the meteorological analyses, and the extent to which the PV, PT, and ozone mixing ratio are conserved along the flow. The validity of these assumptions varies as a function of latitude, altitude, and season, but these assumptions generally hold true in the lower stratosphere. Overall, the technique is a computationally inexpensive and efficient means of comparing sparse data sets.

2.4.6.3. Technique Verification

To gauge the reduction in variance possible through use of dynamical techniques, we calculate the standard deviation of March 1995 SAGE observations in 5 degree latitude bins on altitude surfaces (traditional approach) and 5 degree equivalent latitude (Butchart and Remsberg, 1986; Lary *et al*., 1995)

bins on potential temperature surfaces. Equivalent latitude is a form of normalised PV closely associated with geographical latitude. In regions where the flow is zonally symmetric, or the dynamical time-scale is much longer than chemical or radiative time-scales, equivalent latitude reduces to geographical latitude. The standard deviations of the observations showed general patterns across broad latitude bands, and were thus averaged into broader bands. Figure 2.29 shows the results for (a) 30-50°S, (b) 50-65°S, (c) 30°-50°N and (d) 50°-70°N. In regions where the photochemical and dynamical time scales are long, such as at high latitudes in both hemispheres (panels b and d) and at northern mid-latitudes (panel c) in the lower stratosphere, CM shows a reduction in variance as compared to the traditional approach. In the tropics (not shown) and southern mid-latitudes (panel a), where the flow is less disturbed or photochemical time scales are short (i.e. in the tropical upper stratosphere), the CM results are consistent with traditional results. The period March 19-21, 1995 was selected to assess the effectiveness of the CM technique. During this 3-day period, HALOE and SAGE were observing at the same latitudes in both hemispheres. As a result, the traditional zonal mean approach should have produced its most reliable results. First, zonal means of SAGE and HALOE profiles meeting standard coincidence criteria were created, similar to the analysis in section 2.4.5.2. Then SAGE and HALOE observations for March 6-31 were mapped into dynamical coordinates using the CM technique. Using the National Centers for Environmental Prediction (NCEP) global analyses interpolated to the satellite measurement sites, we interpolated HALOE and SAGE ozone data for this period onto PT surfaces and separated the data into bins of 2 PV units (1 PVU = 3×10^{-6} K m^2 kg^{-1} s^{-1}) on each level. Measurements in each bin were averaged to define a composite field in PV ? PT space for each instrument, using data collected for periods of ± 10 days (21 days total). Composite fields were constructed on March 16 and 21, and in-between days were linearly interpolated. SAGE data with corresponding quality flags in excess of 12% of the observation values were excluded. To compare with the traditional results, the HALOE and SAGE measurements were then mapped back into physical space. Using the 12Z March 20 NCEP dynamical data and the PV-PT composite fields, we reconstructed synoptic SAGE and HALOE measurements at the resolution of the dynamical data (2 degrees latitude by 5 degrees longitude). Finally, we calculated the zonal mean of the CM coincident measurements.

Figures 2.30a shows the percent difference between SAGE and HALOE standard zonal mean ozone profiles in the northern (64°-66°N) and southern (29°-35°S) Hemispheres. The error bars represent the 2s standard error in the zonal mean calculation. The agreement between the two data sets is within 5% in the northern hemisphere, with no statistically significant differences. Larger differences exist in the southern hemisphere below 22 km, with HALOE ozone generally lower than SAGE ozone. These differences may be due to the presence of aerosols. Figure 2.30b shows the percent difference between zonal mean CM HALOE and CM SAGE ozone measurements as a function of height at 65°N and 30°-34°S. The agreement between these zonal mean difference profiles is within 5% of the traditional results in both hemispheres. The CM technique indicates greater statistical significance in the northern hemisphere lower stratosphere, where the technique should be performing well (see Figure 2.29d). This demonstrates the validity of the technique as compared to the traditional approach during a time period when the traditional approach is performing its best. This is true in both hemispheres, despite the reduced variability in the southern mid-latitudes (see Figure 2.29a). Further, small errors introduced by meteorological uncertainties, for example, should be random, and thus not contribute to trend uncertainties. The CM technique gives results consistent with traditional comparisons, but, unlike the traditional approach, CM can comment on the agreement between SAGE and HALOE at latitudes and during time periods outside of those when SAGE and HALOE are coincidentally observing the same latitudes.

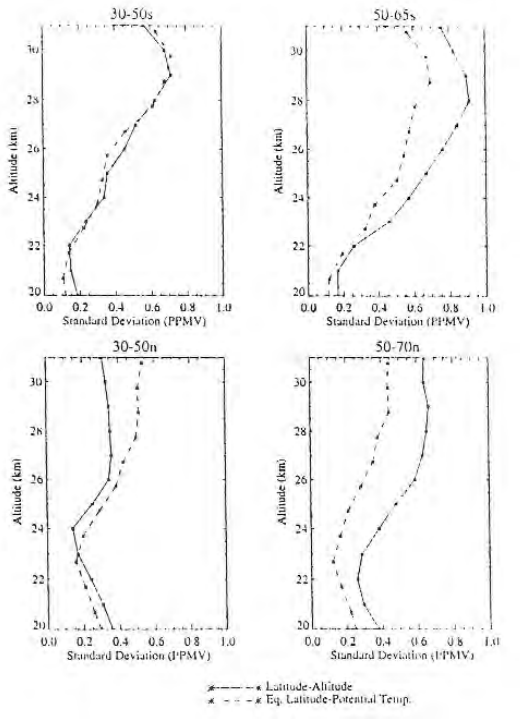


Figure 2.29 Standard deviation of March 1995 SAGE observations in 5 degree latitude bins on altitude surfaces (traditional approach) and in 5 degree equivalent latitude bands on potential temperature surfaces (Lagrangian-mean approach). Plotted are average standard deviations at (a) 30°S-50°S; (b) 50°S-65°S; (c) 30°N-50°N; and (d) 50°N-70°N in each coordinate system.

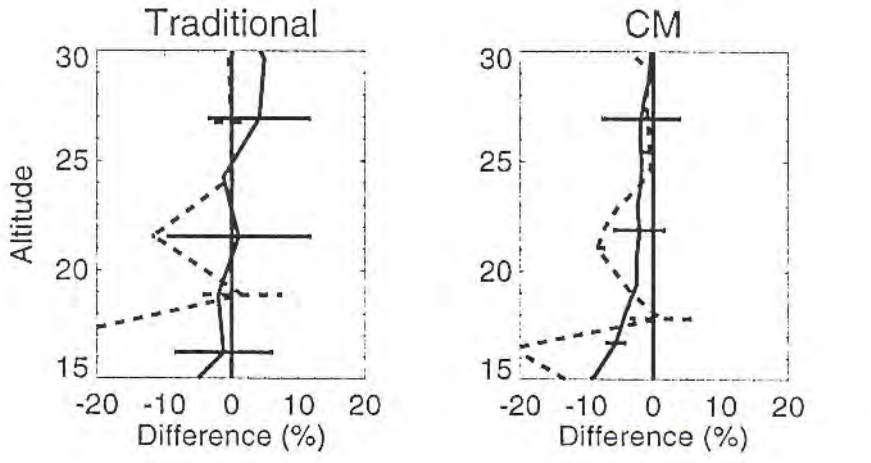


Figure 2.30 March 19-21, 1995 (HALOE-SAGE)/SAGE ozone percent difference profiles at 64°N-66°N (solid line) and 29°S-35°S (dashed line) for (a) Traditional results and (b) Coordinate Mapping results. Traditional results are coincident-matched profile differences calculated on pressure surfaces. CM results are percentage differences between CM HALOE and CM SAGE calculated on potential temperature surfaces for 65°N (solid line) and 30°S-34°S (dashed line).

2.4.6.4. Coordinate Mapping-Based Intercomparison of SAGE II with HALOE

2.4.6.4.1. Analysis Technique

Each SAGE and HALOE profile from 20 September 1991 to 31 December 1996 was first interpolated to

24 isentropic levels from 300K and 1200K. Daily NCEP stratospheric analyses were then used to transform the data at each level from geographic into equivalent latitude. Finally, a 2-D gaussian interpolation scheme was used to produce 5 degree equivalent latitude by 1 month ozone fields for each instrument. The individual measurements were weighted by the distance from the grid point and by the reported measurement error. This weighting was used to eliminate spurious extrapolated values from consideration. At each isentropic level and equivalent latitude, UARS period mean and seasonal differences and simple least-squares trends in the differences were calculated. The latter quantity should be indicative of UARS period mean drift between SAGE and HALOE.

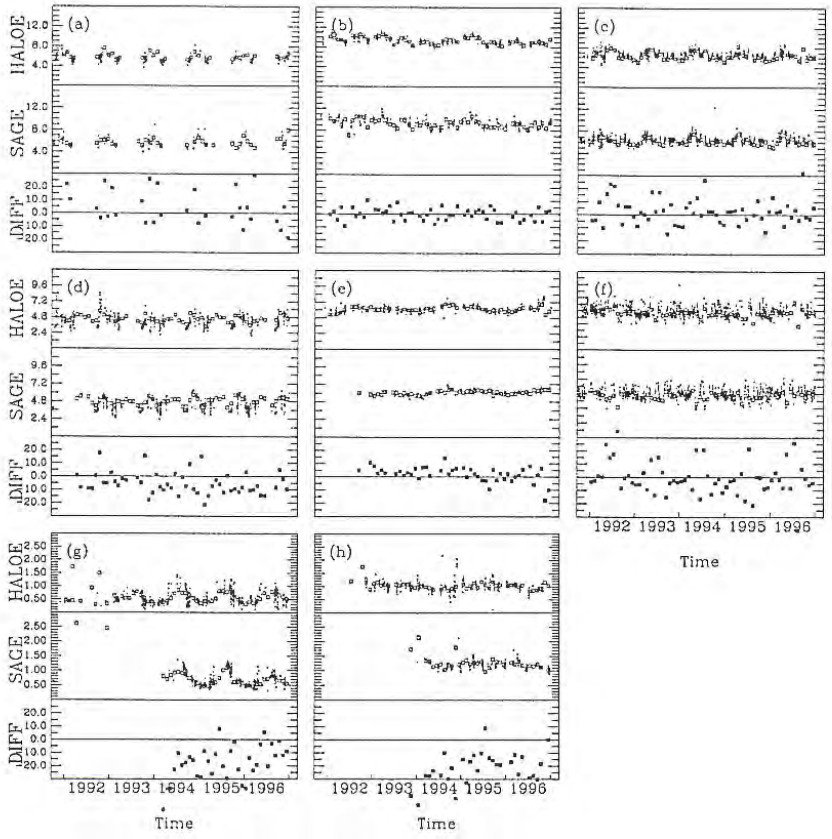


Figure 2.31 UARS period time series of raw and analysed ozone mixing ratio for HALOE, SAGE, and their percentage differences at a range of equivalent latitudes and potential temperatures. (a) 76°S at 1100K, (b) 4°S at 1000K, (c) 60°N at 1000K, (d) 64°S at 650K, (e) Equator at 650K, (f) 40°N at 700K, (g) 60°S at 400K, and (h) 4°S at 475K. Crosses denote raw observations made within 2 degrees equivalent latitude of the central latitude, and open squares the monthly gridded values. Filled squares are the monthly percentage differences.

2.4.6.4.2 Results

Figure 2.31 shows a selection of UARS-period time series of the data from SAGE and HALOE, and their percentage differences, for selected values of potential temperature and equivalent latitude. This figure demonstrates the degree of effectiveness of the analysis technique in describing the "modified Lagrangian mean" ozone evolution observed by each instrument during the period. The seasonal data gaps evident in Figures 2.31a and d, for the high southern latitudes, suggest that any results for this region are likely to be seasonally biased. Figures 2.31e, g, and h reveal the very limited record available for the calculation of trends in the lower stratosphere.

Figure 2.32 shows the period mean percentage difference (HALOE-SAGE) between SAGE and HALOE observations. In general, SAGE sees considerably more ozone than HALOE throughout the lower stratosphere, more so with decreasing altitude (see Figures 2.31g and h). This is consistent with the sonde results reported in section 2.4.1. Above about 600K (~25 km) SAGE and HALOE typically agree to within a few percent, though larger differences are present in the mid- to high latitude upper stratosphere and the tropical middle stratosphere (around 700K) where HALOE ozone exceeds SAGE by 2%. In the mid-latitude middle stratosphere, SAGE exceeds HALOE by 2.5%. An examination of the seasonal variations in the differences (not shown) indicates that the upper stratospheric differences are largely a summer feature (see Figure 2.31a and c). Figure 2.31e suggests that HALOE may see a stronger QBO cycle in ozone than SAGE in the tropical middle stratosphere below the ozone maximum.

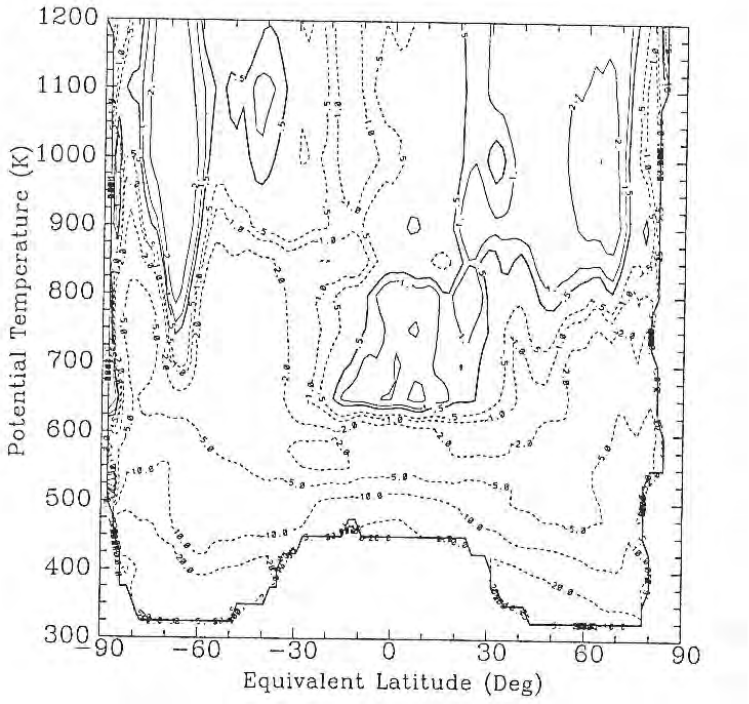


Figure 2.32 UARS period mean, modified Lagrangian mean cross-section (equivalent latitude versus potential temperature) of percentage difference between HALOE and SAGE ozone mixing ratio (i.e. 100*(HALOE-SAGE)/SAGE).

Simple least squares linear trends were calculated for each instrument and for the differences between the two instruments. The relative drift between instruments is shown as a function of equivalent latitude at 4 PT levels in Figure 2.33. At 650K (~25km) HALOE observations at most latitudes in the middle stratosphere tend to have drifted downward relative to SAGE, by up to a percent or so per year. This is where, according to Figure 2.32 period mean HALOE ozone is lower than SAGE, suggesting the gap between the two instruments has widened during the UARS period. The CM analysis shows three regions of statistically significant drift between the instruments at this level. At southern high latitudes, the results are in good agreement with traditional comparisons (section 2.4.5) showing negative HALOE drift relative to SAGE of 1-2%/year from 45^{o}-60^{o}S. The CM results are also consistent with a HALOE negative drift of 0.5%/year at 60^{o}-70^{o}S (section 2.4.5), keeping in mind that the drift in this latitude range maps to the average drift in the equivalent latitude range of ~60^{o}-90^{o}S. In this case the individual instrument analyses (not shown) show a statistically significant decrease in HALOE ozone, while SAGE shows an insignificant increase over the same period. The positive HALOE relative drift in the southern

subtropics of ~ 2± 1%/year is consistent with results shown in Figure 2.28 in Section 2.4.5.3. However, the significant negative HALOE drift at the equator is most likely a consequence of the two extraneous data points at the end of the HALOE time series, as seen in Figure 2.31e. In the 40°-50°N equivalent latitude range, the CM analysis indicates a negative drift of HALOE with respect to SAGE of ~1± 0.75%/year. In this region the trend in SAGE observations is near zero, while the HALOE data show an ozone decrease on the order of 1-1.5%/year. A similar tendency can be seen in the trend comparisons in Figure 2.28 in Section 2.4.5.3.

Above 800K (~30 km) the CM analysis suggests that the mean drift between SAGE and HALOE is generally small (less than 0.5%/year) and not significant, except at polar latitudes. In the northern polar regions at 800K and 1000K (~ 36 km), summertime HALOE observations have drifted downward significantly with respect to SAGE (~ 3-4% ± 3%/year) with SAGE during the period showing ozone increases not detected by HALOE. A similar feature is present in the southern hemisphere at 1200K (~ 40 km) as also seen in Figure 2.31a. The two instruments agree on a significant downward drift in tropical upper stratospheric ozone (above 750K) of up to 2%/year during the period (see Figure 2.31b). This is consistent with the findings using traditional techniques shown in Figure 2.28 at 35 and 40 km (see Section 2.4.5.3). In the lower stratosphere (not shown), the short record length considered makes it meaningless to interpret more than the period mean differences. Of note in Figure 2.31g, though, and to a lesser extent in Figure 2.31h, is an indication that the drifts derived for these locations may be unduly influenced by aerosol-affected observations from the Pinatubo period which have escaped rejection by the data quality filter.

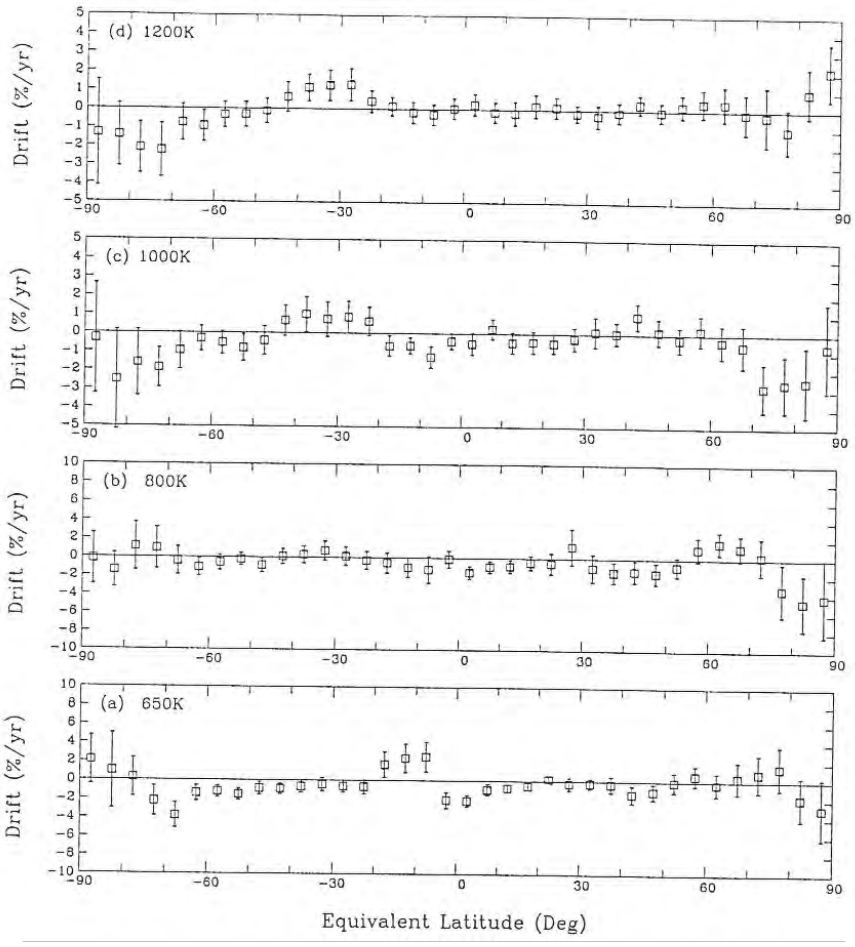


Figure 2.33 UARS period drift between monthly analysed SAGE and HALOE ozone (%/year) as a function of equivalent latitude at (a) 1200K (~ 40km); (b) 1000K (~ 36km); (c) 800K (~ 30 km); (d) 650K (~ 25 km). Error bars are 95% confidence limits.

2.4.6.5. Trajectory Mapping

Trajectory mapping (TM) is another dynamical technique useful for the intercomparison of sparse data sets. Whereas the CM approach has its equivalent in the intercomparison of monthly zonal mean data, the TM technique is the "dynamical" equivalent of the traditional coincidence approach. Like CM, it takes advantage of quasi-conserved quantities following the motion, namely the ozone mixing ratio and potential temperature. Morris *et al*. (1995) noted that the latter of these is reasonably well conserved for about 7 to 10 days.

To create a synoptic trajectory map from satellite data, air parcels are initialised in the model at the time and location of each satellite observation. These air parcels are then isentropically advected forward or backward in time by the model using analysed wind fields. Because trajectory maps include observations from a number of days, they yield substantially enhanced coverage compared to asynoptic maps produced from a single day of measurements. Furthermore, by correctly accounting for dynamical changes in the atmosphere between observations, TM, like CM, provides better representations of the measured constituent fields than asynoptic schemes. Trajectory mapping can be straightforwardly applied to data intercomparison, as shown in Pierce *et al*. (1994a and b), Morris (1994), Morris *et al*. (1995) and Morris *et al*. (1997). By creating trajectory maps of observations from one instrument at the time of the observations of a second instrument, individual pairs of observations (which may have been made at widely different locations and times) can closely coincide and be directly compared. Frequently, these pairs would have failed the coincidence criteria used in the traditional approach and would therefore have been eliminated from consideration. The main disadvantages of the TM technique are the need for accurate global wind velocity fields (like the CM approach) and the high computational cost (unlike CM) which makes the technique impractical for intercomparisons over decadal time-scales.

The use of TM techniques in satellite data validation is summarised in Morris *et al*. (1995) and Morris *et al*. (submitted 1997; Web: ftp:/hyperion.gsfc.nasa.gov/pub/papers/morris/jgr.paper4). Morris *et al*. (1997) presents results using the TM analysis for the same March 1995 test period as above. Figure 2.34 shows a scatter plot of (a) the coincident ozone measurements retrieved by HALOE and (b) 3-day, trajectory-mapped HALOE averages versus SAGE observations at the 800 K level. Using a 400 km and 12 hours coincidence criteria gives 120 coincidences (a stricter time criterion of 10 hours results in fewer than 10 coincidences), or ~20% of the total measurements made by HALOE and SAGE during this time period. Furthermore, the coincident approach can only comment on the relationship between the two instruments in the restricted latitude bands of concurrent observation (e.g., 64°?66°N, 29°?35°S). Using the TM technique, the number of coincidences increases from 120 to 202. More coincidences occur with longer trajectory calculations, but this improvement is balanced by the increased uncertainty with longer trajectories. At 800K, the RMS difference is 10.0% for the traditional coincidences compared to 7.1% for the TM pairs. In this study, trajectory advection periods of up to 7 days duration produced RMS differences in measurements which were smaller than those found using the traditional coincident comparison technique at all altitudes in the 64°-66°N range. The reason for this behaviour can probably be attributed to the method's ability to reduce spatial and temporal variability.

2.4.6.6. Discussion and Implications

Both CM and TM techniques are effective methods for comparing "sparse" data sets. They increase the number and spatial extent of observations from each instrument that can be intercompared. This analysis and results from Morris *et al*. (1997) suggests the TM approach is more accurate than the CM technique, but it is also much more computationally expensive. As a result, TM is probably better used in closely examining particular pairs of observations, whereas the computational efficiency of the CM approach

makes it ideally suited for intercomparing data sets over extended periods. The results obtained for the SAGE/HALOE intercomparison using the CM technique generally support those obtained using the classical "zonal mean" approach, but tend to be more significant. The CM intercomparison has also provided insight into instrument performance in regions of the atmosphere where data sparsity inhibits intercomparison by the classical coincidence techniques, and where "difference in trend" analyses likely suffer uncertainties due to the differential instrument sampling in space and time.

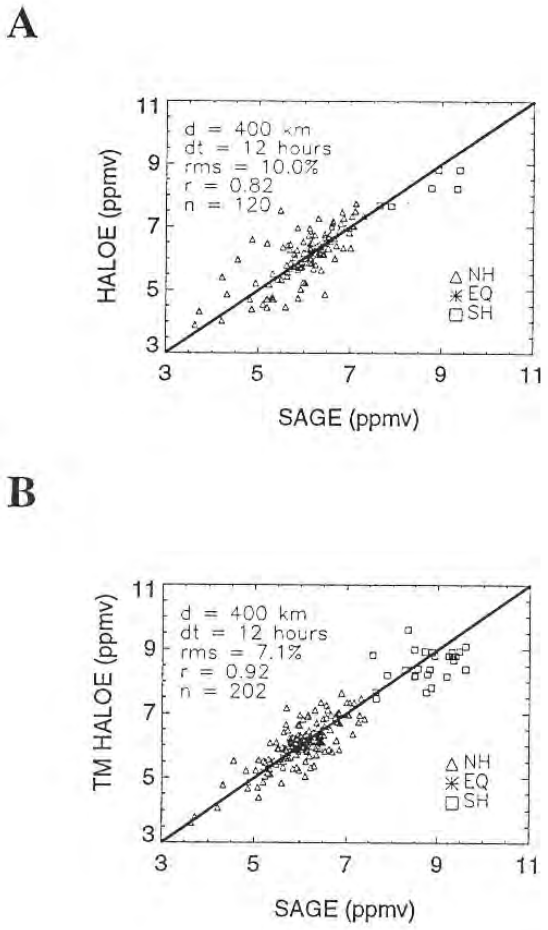


Figure 2.34 (a) HALOE vs. SAGE ozone using traditional coincidence criteria of 400 km maximum separation and 12 hour maximum time difference for the period March 12 - April 5, 1995 at the 800 K potential temperature surface (~10 hPa or 30 km). (b) HALOE vs. SAGE ozone on the 800 K (~30 km) potential temperature surface using the trajectory mapping approach. Trajectory calculations are limited to 3 days duration. Advected HALOE measurements within 400 km of the SAGE observations are considered coincident.

Two important qualifications should be made regarding the CM drift results. First, they are based on just five years of data, at most: whether or not the analysed drifts between the two instruments, for example, are representative of instrument behaviour over longer time scales, has not been addressed here. Second, erroneous temperature drifts in either the NCEP stratospheric analyses or the HALOE temperature retrievals could have a significant impact on this analysis. At these altitudes, both SAGE and HALOE ozone retrievals are based on first guess temperature and pressure profiles interpolated from the NCEP analyses, whereas HALOE retrieves its own temperature and pressure profiles above 30 km but uses NCEP data below. The first step in the CM-based analysis was the interpolation of all ozone profiles to isentropic levels. For this, the temperature profiles provided by the respective instrument team with each ozone profile were used (NCEP for SAGE, HALOE for HALOE), so any overall drift between HALOE

and NCEP temperatures would induce an isentropic ozone trend wherever there is a vertical gradient in ozone. Perhaps as important, erroneous trends in the NCEP analyses could directly impact the SAGE retrievals through the pressure dependence of Rayleigh scattering in the SAGE ozone retrieval. On the other hand, real temperature trends depicted by both NCEP and HALOE should not influence the results presented here.

The CM analysis reveals ~ 0.5% differences in drifts above 30 km in all except the polar regions. At ~25 km the results show significant negative drift of HALOE with respect to SAGE at southern mid-high latitudes and to a lesser extent in northern mid-latitudes. The results also indicate a positive drift of HALOE relative to SAGE in the southern sub-tropics, which may be a manifestation of QBO differences in the data. These results generally agree with the conventional comparisons described in Section 2.4.5. The conclusion that no confidence can be placed in differences in differences in trends below about 22 km also agrees with the conventional analysis results. The analysis for the lower stratosphere suggests care should be exercised in interpreting trends based on data that include aerosol-affected measurements. Given the documented sensitivity of SAGE ozone retrievals to stratospheric aerosol, this analysis suggests, in support of the classical analyses, that it may be advisable to time-filter, rather than quality-filter SAGE data.

2.4.7. SAGE II Comparisons Against MLS and HALOE

Comparisons have been made between SAGE II and coincident MLS measurements by averaging SAGE ozone measurements over UARS pressure levels. Occasional spikes found in the MLS data were found to be correlated with high values of the MLS ozone error bars. These spikes have been removed from the comparisons. Mean differences between SAGE II and MLS (version 4) are shown in Figure 2.35(a) to (c). MLS observations are retrieved with a vertical resolution of two UARS layers (~5 km). Comparisons between SAGE II and MLS will therefore emphasise 4.6, 2.2 and 1.0 hPa (and corresponding levels which differ by a factor of ten in pressure). MLS observations may be seen to yield a few percent larger ozone concentrations from 2.2 to 22 hPa. At 1.0 hPa MLS measurements are smaller than SAGE II values. This has been noted for some time (e.g. Cunnold *et al.,* 1996b) but perhaps more remarkable is the apparent difference between the SAGE sunrise and sunset measurements of approximately 10% at 1 hPa. This difference is slightly larger than for SAGE version 5.93 results (see Figure 2.2 in Section 2.3.1). Fortunately at levels other than 1 hPa, the SAGE II sunrise/sunset differences are less than 5%, including in Umkehr layer 9 (see Figures 2.23 in Section 2.4.4. and 2.6 in Section 2.3.3.1 for example).

At 46 hPa SAGE II ozone values are larger than MLS values. This suggests that MLS values are too small at this level, particularly in the tropics; evidence for this has been previously discussed by Froidevaux et al (1996). However ozonesonde comparisons also suggest that SAGE II values are approximately 5% to 10% too large between 15 and 20 km altitudes (see Section 2.4.1); this may be contributing to the SAGE/MLS differences. At the 100 hPa level it is clear, and from other comparisons also, that the MLS values are substantially too large; there are known difficulties with the MLS retrievals at this level. Accordingly, trends at 100 hPa are not included in the comparisons of this section.

The relationship between SAGE II and HALOE ozone measurements can be examined by using MLS as a transfer standard (since there are only a smaller number of coincidences between SAGE II and HALOE measurements). Comparisons between HALOE (version 18) and MLS measurements (Figure 2.35(d) to (f)) show that HALOE values are also a few percent less than MLS values from 2.2 to 22 hPa. They confirm that MLS values are relatively smaller at 1 hPa (by approximately 5%) and they show a similar structure to the MLS/SAGE differences at 46 and 100 hPa, thereby also suggesting deficiencies in the

MLS profiles at these two levels.

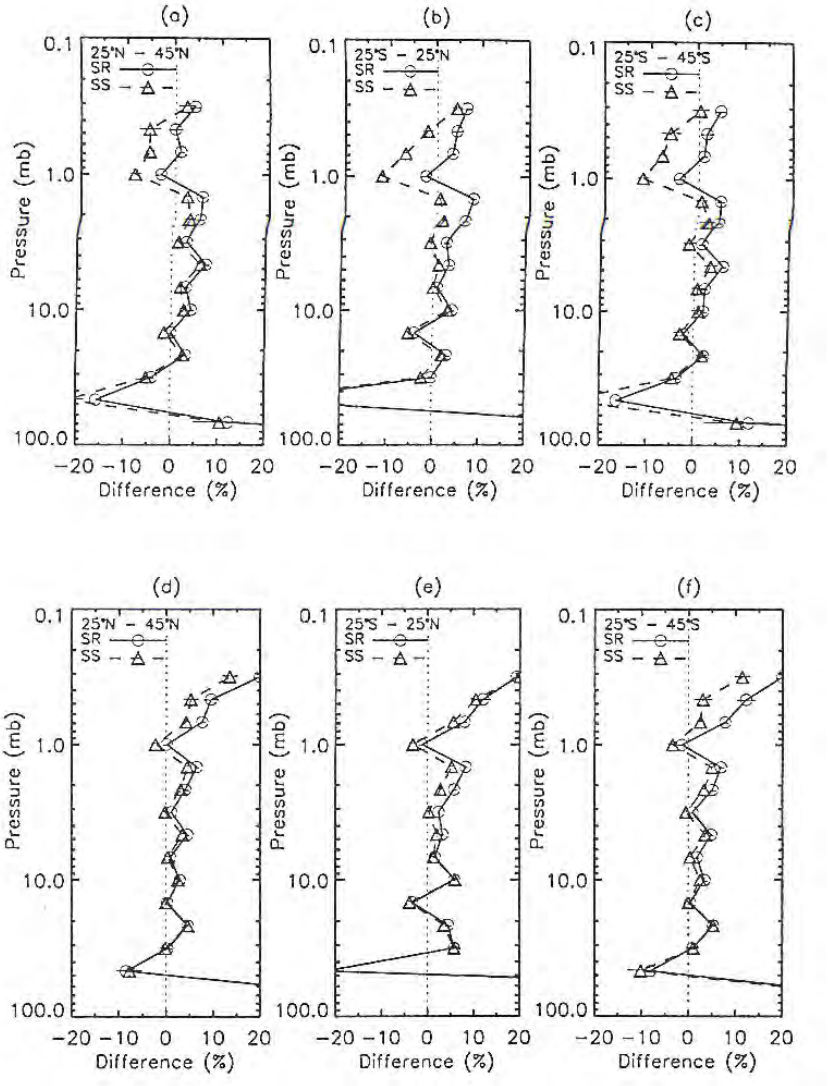


Figure 2.35 Differences (MLS-SAGE II) between (a), (b) and (c) SAGE II and (d), (e) and (f) HALOE (version 18) and coincident MLS (version 4) night time ozone measurements expressed as percentages of the MLS values. SAGE and HALOE sunrises and sunsets have been treated separately and the data were placed in three separate bins which extended from 25° S-45° S, 25° S-25° N and 25° N -45° N. The SAGE data were placed on pressure levels using the NWS temperature profiles and then summed over UARS layers which have boundaries at $10^{3-I/6}$ hPa. The error bars are twice the standard errors of the differences. The period of comparison was October 1991 to December 1996 but the SAGE data has been filtered as recommended to remove the Pinatubo aerosol interference effects.

Since the 15-20 km altitude range is so important for this report, it is highly desirable to compare SAGE and HALOE at 68 and 100 hPa. Unfortunately the MLS observations are not sufficiently accurate at these levels that MLS data can be used as a transfer standard. Therefore a direct comparison of the mean SAGE and HALOE values over similar periods is shown in Figure 2.36. Differences are between ±2% at altitudes above 32 hPa but SAGE values are larger than HALOE values at altitudes below this particularly in the tropics. At 46 hPa similar differences are derived by using MLS as a transfer standard. At mid-latitudes below 20 km altitude (50 hPa) a hemispheric asymmetry of the differences (of approximately 5%) is also shown. This asymmetry is also present if pre-Pinatubo SAGE II data are projected forward in time using a time series fit to SAGE I/SAGE II trends. The low altitude differences

might be related to altitude registration problems for SAGE II below 20 km altitude but the latitudinal asymmetry suggests that the effects may not be systematic and that SAGE II trends below 20 km altitude might be affected.

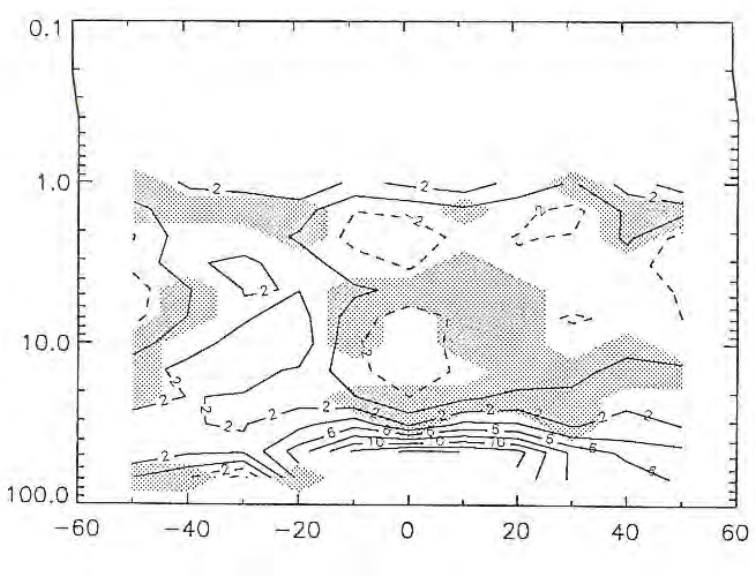


Figure 2.36 Mean differences between SAGE and HALOE measurements in percent determined from linear fits to SAGE and HALOE monthly means in UARS layers (but including annual and semi-annual cycles). The time series extended from October 1991 to December 1996 but started later at altitudes below 30 km where SAGE II data were filtered, as recommended, in order to remove the effects of aerosol interference.

Figure 2.37 shows a comparison of MLS and HALOE annual cycles in ozone. These may be compared against the SAGE II values which were compared earlier against the SBUV2 values (Figure 2.7 in section 2.3.2.1). In the lower stratosphere there are differences which reach 6% where the annual cycles are large and are changing rapidly with height. This is additional evidence that there are significant differences between the various measurements below 20 km altitude. Elsewhere however the results provide confirmation of the excellent quality of all three measurement systems.

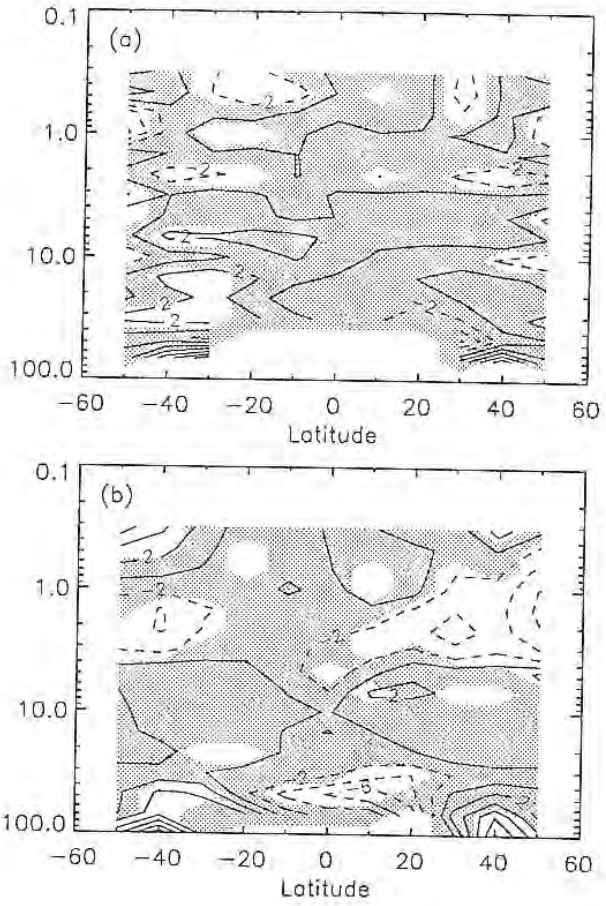


Figure 2.37 A comparison between the amplitudes of the annual cycles in zonal mean ozone expressed as percentages of the annual mean values from (a) HALOE and (b) coincident MLS observations. The HALOE measurements have been summed over individual UARS layers and data from October 1991 to December 1996 were used.

Linear slope differences from coincident SAGE II and MLS observations are shown in Figure 2.38. The overlap period is less than 6 years (and is only half this at altitudes below 40 hPa because of the prescribed filtering of SAGE II data). SAGE II slopes equal or exceed MLS slopes in the latitudinal mean at all levels from 1 to 46 hPa but there is no recognisable pattern in the differences and only at 22 hPa is the difference barely significant at the 95% confidence level. The mean of the slope differences between 1 and 46 hPa is -0.4 ±0.8%/year. As before, SAGE ozone slopes may be related to HALOE ozone slopes based on a comparison of slopes in coincident HALOE and MLS measurements. This comparison is also shown in 2.38. In this case, differences at 46 hPa seem to be related to an MLS problem during the first year after the Pinatubo eruption during which MLS ozone values, in the tropics particularly, exhibited a strong upward trend at 46 hPa which was not captured by the ozonesondes (Froidevaux *et al.*, 1996) (the SAGE/MLS comparisons did not contain this period at these levels because of Pinatubo effects on SAGE II). The only significant MLS/HALOE slope difference at the 95% confidence level between 1 and 22 hPa occurs at 22 hPa where the HALOE slope is 0.9%year^{-1} less than the MLS slope. The differences in the 1 to 22 hPa region are latitude and altitude dependent and average -0.1 ±0.8%year^{-1}.

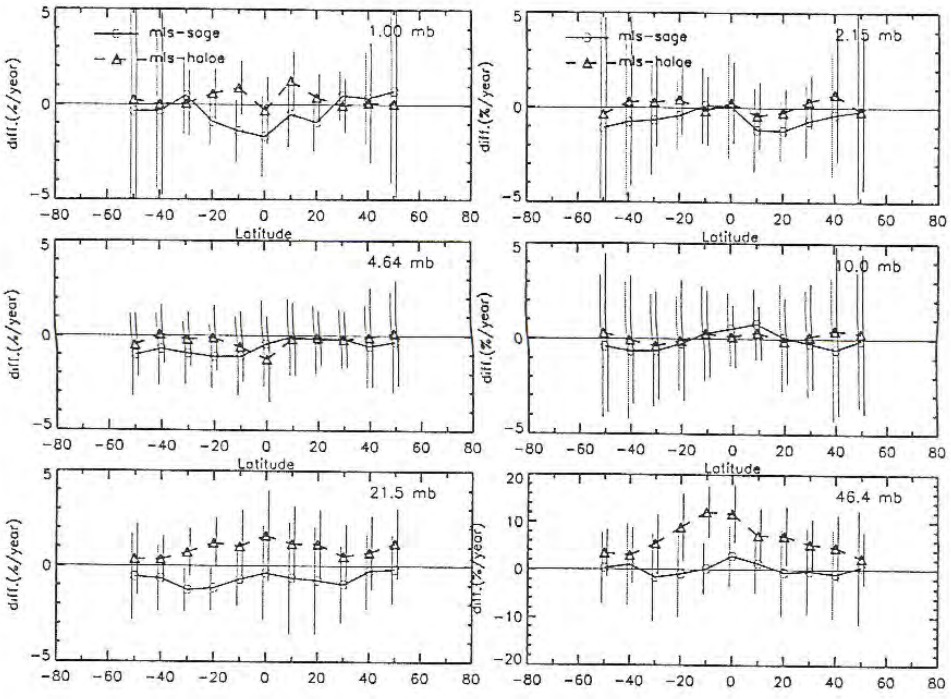


Figure 2.38 The slopes in the differences between coincident night-time MLS ozone observations from 1991 to 1996 and SAGE II (full line with circles) and HALOE (dashed line with triangles). The SAGE II data has been summed over UARS layers and filtered by the recommended procedure following the Pinatubo eruption. The error bars are twice the standard error of the slope estimates.

It is concluded that HALOE, MLS and SAGE II do not drift relative to each other by more than 0.5%year^{-1}, which is within the measurement uncertainties. However all three sets of measurements yield significantly different slopes at 22 hPa with SAGE II yielding the most positive slope and HALOE showing 1.6%year^{-1}. These differences are almost independent of latitude.

The SAGE II sunrise/sunset differences shown in Figure 2.35 (a)-(c) combined with the absence of sunrise/sunset differences in SAGE I observations suggest that there are long-term trend differences in the upper stratosphere between the sunrise and sunset results in the SAGE I/II record. These trend differences are shown in Figure 2.39. Similar differences between the two SAGE II alone trends are calculated if SAGE I observations are not used; furthermore comparisons against the SBUV measurements suggest that only approximately 0.1%year^{-1} of the differences can be attributed to sampling differences. The SAGE II comparisons against HALOE and SBUV measurements do not suggest a preference for sunrise or sunset observations. Furthermore Wang *et al*. (1996) noted the occasional presence of large differences of opposite sign between SAGE II sunset and sunrise observations at the same latitude when comparing against SBUV measurements (e.g. in January 1987). Thus SAGE trends are reported from the combination of sunrise and sunset observations and one-half the differences shown in Figure 2.39 are included in the reported (2) error bars.

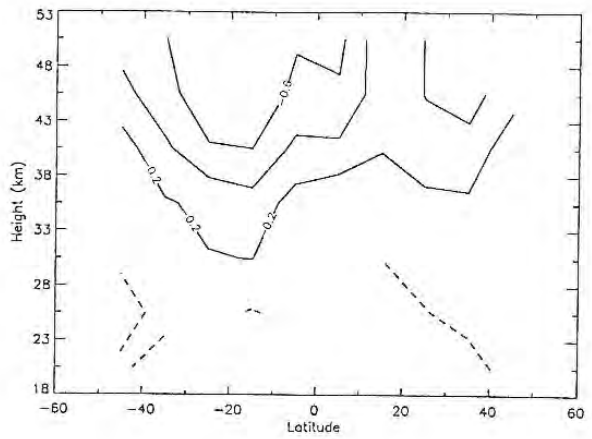


Figure 2.39 The difference between SAGE I/II sunrise and sunset trends from 1979 to 1996 (percent/year). The dashed lines are contours of zero percent/year and the latitude bins were 10 degrees wide.

2.4.8. Short Term Intercomparisons

Although intercomparisons which extend over relative short periods are not useful for evaluating the accuracy of long term ozone profile trends, they can be very valuable for assessing the relative accuracy and precision of the different instruments as a function of altitude. This section reports on results of a series of short term intercomparisons made during the last decade. First, SAGE II data are compared with observations obtained during the Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment deployed on the Space-Lab-3 mission in April/May 1985 and the ATLAS-3 Space Shuttle mission in November 1994. Also, within the frame of the Network for the Detection of Stratospheric Change (NDSC) dedicated short term campaigns were conducted to compare multiple instruments that measure the stratospheric ozone vertical distribution. The results of three NDSC intercomparison campaigns are evaluated: STOIC (Stratospheric Ozone Intercomparison Campaign), OPAL (Ozone Profiler Assessment at Lauder) in April 1995 and MLO3 (Mauna Loa Ozone). Finally, time series of microwave observations made at Table Mountain and at Lauder were compared with SAGE II.

ATMOS/SAGE II comparisons:

For the most recent flight of ATMOS, that on ATLAS-3 [Gunson *et al.,* 1996], there were 16 SAGE II sunset coincidences between 45°-54°N and 32 sunrise coincidences between 65°-74°S. In this particular comparison, both sensors recorded sunset occultations in the northern hemisphere and sunrise occultations in the southern hemisphere. The average ATMOS latitudes were 47.7°N and 71.4°S for sunset and sunrise occultations, respectively. The corresponding average SAGE latitudes were 51.5°N and 68.3°S, respectively. Sunset coincident days cover Nov. 4-6 for ATMOS and Nov. 3-7 for SAGE. Sunrise coincident days cover Nov. 7-11 for ATMOS and Nov. 6-12 for SAGE. The coincidence criteria are 5 degrees latitude, 1000 km distance, and 24 hours. Figure 2.40 shows the average ATMOS-SAGE percent difference for the 1994 flight for the sunset profiles in the upper left panel and for all the sunrise profiles in the lower left panel. The results of both sunrise and sunset comparison are similar. SAGE II and ATMOS agree at 40 km, but at higher altitudes, the ATMOS measurements are higher by about 10% at 50 km. Below 40 km, the ATMOS measurements are increasingly lower than SAGE II, reaching a maximum difference of 30% at 15 km. Shifting the altitude registration of ATMOS down approximately 1 km with respect to SAGE, or equivalently shifting the SAGE altitude registration up 1 km, would reconcile the altitude pattern of the coincident pairs.

To compare SAGE II to ATMOS nine years earlier in 1985, we computed zonal average ozone profiles for both sensors. The ATMOS Space-Lab-3 occultations between 29 April and 1 May comprised 11

sunset profiles between 26°N and 35°N at an average latitude of 30°N. In the same latitude band, 35 SAGE II profiles between 26 April and 28 April constituted the SAGE II zonal average. The average ATMOS-SAGE sunset profile percent differences (±2 RSS mean errors) for the 1985 flight are shown in upper right panel of Figure 2.40 This altitude pattern of differences is, within error bars, exactly equal to that for the ATMOS/SAGE II comparison made 9 years later. For the sunrise measurements only 3 ATMOS profiles are available, at an average latitude of 48°S (see lower right panel of Figure 2.40). The SAGE II zonal profile was derived from 75 SAGE II sunrise observations between 43°S and 53°S (i.e. 48° ±5°) made between 26 April and 1 May, 1985. The profile differences are shown in the lower right panel. While the structure of the comparison is similar above 40 km, below that attitude, it appears that the 3 ATMOS profiles are sampling different air masses than the SAGE II observations. Because of the small ATMOS sample size, this comparison is less meaningful than the other ATMOS/SAGE II comparisons which have a more adequate sample size.

It is concluded that, while there appears to be a small altitude offset between ATMOS and SAGE II, the SAGE II sensor did not experience a significant drift in ozone observations over the nine years between the two comparisons. This comparison places an upper limit on the possible SAGE drift of ±1%/year with 95% confidence.

STOIC:

In July/August 1989 STOIC (Margitan *et al.,* 1995) was conducted at the Table Mountain Facility (2300 m altitude, 24.4°N, 117.7°W), run by the Jet Propulsion Laboratory (JPL) in California, USA. Various ozone profiling instruments were compared: UV-Lidar from JPL (McDermid *et al.,* 1995), UV-Lidar from NASA/GSFC (McGee *et al.,* 1995), Microwave Radiometer from NASA/LaRC (Connor *et al.,* 1995), ROCOZ-A from NASA/WWF (Barnes *et al.*, 1995), ECC-Sondes from NASA/WWF and NOAA/CMDL (Komhyr *et al.,* 1995), Brewer/Umkehr by AES, Canada (McElroy *et al.,* 1995) and Dobson/Umkehr (Komhyr *et al.,* 1995). In general, the comparison showed good agreement (on average within about ±5%) among the different ozone measuring techniques, especially in the 20-40 km altitude range. Although atmospheric variability was rather low during the campaign, individual daily comparisons show a spread of about 10%.

By 30-35 km altitude the ECC-sonde performance began to fall off, which is consistent with results from previous comparisons like BOIC in 1983/1984 (Hilsenrath *et al.,* 1986). Above 40 km the performance of the two lidars began to deteriorate. The GSFC is limited to a slightly lower altitude than the JPL instrument. At altitudes below 40 km the agreement among the lidars, microwave , and sondes is generally excellent. Daily comparisons of the lidars and ECC-sondes show the ability to discern significant vertical structure, especially in the lower stratosphere. Comparison of the Brewer Umkehr and Dobson Umkehr profiles with the STOIC reference profile, which is the average of all measured vertical profiles during STOIC, showed agreement within 15% over the 20- to 45 km region and 5% between 30 and 40 km.

see next page

Click here to download Figure 2.40 by anonymous ftp.

Figure 2.40 The average ±2 standard errors of the SAGE II/ATMOS-ATLAS-3 coincident pairs for sunsets (northern hemisphere) in panel (a) and sunrises (high southern latitudes) in panel (c). The differences between SAGE II and Space-Lab-3 ATMOS sunset zonal averages ±2 RSS of the zonal-average standard errors appear in panel (b). Panel (d) represent the 48° S sunrise zonal averages for ATMOS SL-3 (only 3 profiles) and for SAGE II (75 profiles). The SL-3 ATMOS measurements occurred in 1985, 9 years before the ATLAS-3 measurements.

Only a limited number of matching SAGE II profiles were obtained during STOIC so there is only a

limited basis for evaluating SAGE II. Compared to the other techniques, SAGE II ozone between 20 and 30 km was high by 3-4%, there was excellent agreement between 30 and 40 km, and above 40 km SAGE ozone was slightly lower.

OPAL:

The OPAL intercomparison campaign (McDermid *et al.,* 1996) was carried out at the National Institute of Water and Atmosphere (NIWA) facility at Lauder, New Zealand (45.05°S, 169.68°E) in April 1995. Participating instruments were the NIWA/RIVM-Lidar (Swart *et al.,* 1994), the NASA/NIWA-Microwave Radiometer (Connor *et al.,* 1995), NIWA/ECC-sondes and the mobile NASA/GSFC Lidar (McGee *et al.,* 1995). Between 20 and 45 km altitude both lidar systems and microwave agreed to within 5%. Between 15 and 30-35 km the ECC-sondes agreed with both lidars and microwave to within ±5%. Below 17 km the RIVM/NIWA-Lidar ozone was up to 15% higher than that from the NASA/GSFC-Lidar (Brinksma *et al.,* 1996). There were only four overpasses of SAGE II during OPAL, so the intercomparison of SAGE II with the other ozone observing techniques is limited. Between 20 and 40 km SAGE II showed agreement within 5-10%, while below 20 km SAGE II overestimated ozone by more than 10% compared to the other instruments.

MLO3:

In August 1995 the MLO3 comparison was conducted at the Mauna Loa Observatory, Hawaii (19.5°N, 155.6°W, 3400 m altitude). The instruments compared in MLO3 included two UV LIDAR systems, one from JPL (McDermid *et al.,* 1995) and one from GSFC (McGee *et al.,* 1995), ECC sondes (NOAA/CMDL), Microwave Radiometer from NASA/LaRC (Connor *et al.*, 1995), and two Dobson instruments. Satellite based ozone profiles were provided by UARS/MLS (11 overpasses), SAGE II (1 overpass) and SBUV/2 (10 overpasses).

Initial results show that much better consistency among instruments is being achieved, usually to within the instrument uncertainties. The lidars now appear capable of accurate measurement of the ozone profile from 15 km to 50 km. On average, the JPL and GSFC lidars agreed to within ±3% between 23 and 43 km. Above 43 km the GSFC lidar became noisy, while the JPL lidar performed well to 50 km. Below 23 km the JPL lidar had an increasingly large bias caused by an instrument saturation problem that has since been corrected. The GSFC lidar agreed with the ECC-sonde to within ±5% in the 16 to 25 km region.

The single SAGE II profile agreed with the lidar profiles within ±5%, between 20 and 37 km, but above that altitude SAGE II was 5-10% higher than the JPL lidar profile. SAGE II performed well in the lower stratosphere, agreeing with the ECC sonde profile within ±5% in the 16 to 25 km region. The ground-based microwave instrument agreed with the lidars within ±5% from its lower limit of 21 km up to 43 km. The maximum discrepancy relative to the JPL lidar was -10% at 48 km. The ECC sonde agreed with lidar, SAGE II, and microwave to within ±5% at altitudes up to 25 km, but at higher altitudes an apparent pump correction error led to an increasing discrepancy, reaching +10% above 30 km.

In this campaign instruments that measure ozone density versus altitude were compared with ones that measure ozone mixing ratio versus pressure. The pressure-altitude conversion adds uncertainty to the comparison, particularly for instruments that do not measure a temperature profile and must rely on NCEP data for the conversion. Lidar and SAGE II comparisons with MLS, SBUV/2, and Umkehr are affected by this problem. MLS agreed with the JPL lidar within ±10% between 43 hPa (the MLS lower limit) and 1 hPa. Umkehr generally agreed with lidar within ±10%, but the errors were systematic with altitude, going from -5% at 30 hPa, to agreement at 10 hPa, to -10% at 3 hPa, to -5% at 1 hPa.

Groundbased Microwave/SAGE II Comparisons:

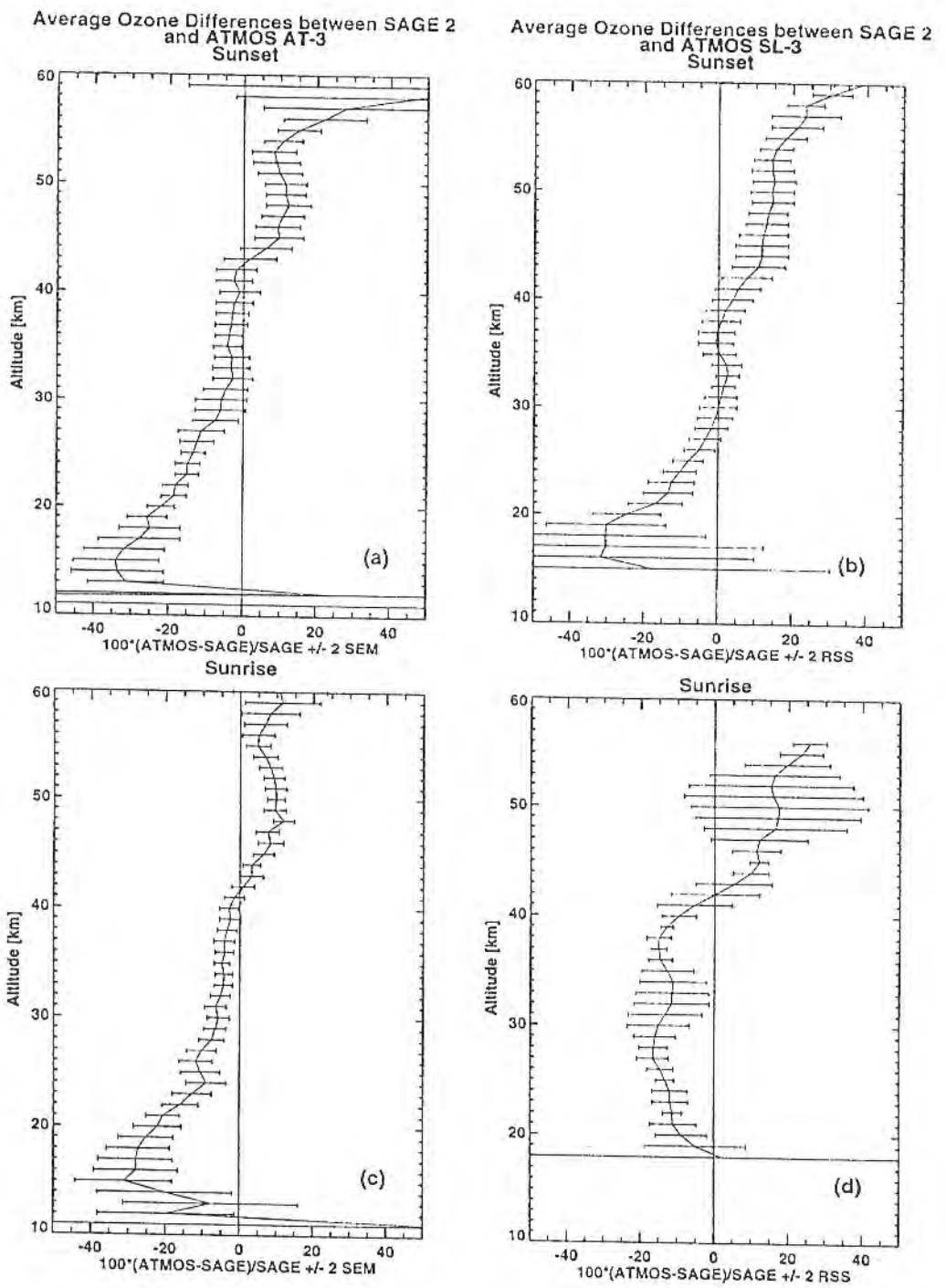


Figure 2.40 The average ±2 standard errors of the SAGE II/ATMOS-ATLAS-3 coincident pairs for sunsets (northern hemisphere) in panel (a) and sunrises (high southern latitudes) in panel (c). The differences between SAGE II and Space-Lab-3 ATMOS sunset zonal averages ±2 RSS of the zonal-average standard errors appear in panel (b). Panel (d) represent the 48° S sunrise zonal averages for ATMOS SL-3 (only 3 profiles) and for SAGE II (75 profiles). The SL-3 ATMOS measurements occurred in 1985, 9 years before the ATLAS-3 measurements.

limited basis for evaluating SAGE II. Compared to the other techniques, SAGE II ozone between 20 and 30 km was high by 3-4%, there was excellent agreement between 30 and 40 km, and above 40 km SAGE ozone was slightly lower.

OPAL:

The OPAL intercomparison campaign (McDermid *et al.,* 1996) was carried out at the National Institute of Water and Atmosphere (NIWA) facility at Lauder, New Zealand (45.05°S, 169.68°E) in April 1995. Participating instruments were the NIWA/RIVM-Lidar (Swart *et al.,* 1994), the NASA/NIWA-Microwave Radiometer (Connor *et al.,* 1995), NIWA/ECC-sondes and the mobile NASA/GSFC Lidar (McGee *et al.,* 1995). Between 20 and 45 km altitude both lidar systems and microwave agreed to within 5%. Between 15 and 30-35 km the ECC-sondes agreed with both lidars and microwave to within ±5%. Below 17 km the RIVM/NIWA-Lidar ozone was up to 15% higher than that from the NASA/GSFC-Lidar (Brinksma *et al.,* 1996). There were only four overpasses of SAGE II during OPAL, so the intercomparison of SAGE II with the other ozone observing techniques is limited. Between 20 and 40 km SAGE II showed agreement within 5-10%, while below 20 km SAGE II overestimated ozone by more than 10% compared to the other instruments.

MLO3:

In August 1995 the MLO3 comparison was conducted at the Mauna Loa Observatory, Hawaii (19.5°N, 155.6°W, 3400 m altitude). The instruments compared in MLO3 included two UV LIDAR systems, one from JPL (McDermid *et al.,* 1995) and one from GSFC (McGee *et al.,* 1995), ECC sondes (NOAA/CMDL), Microwave Radiometer from NASA/LaRC (Connor *et al.*, 1995), and two Dobson instruments. Satellite based ozone profiles were provided by UARS/MLS (11 overpasses), SAGE II (1 overpass) and SBUV/2 (10 overpasses).

Initial results show that much better consistency among instruments is being achieved, usually to within the instrument uncertainties. The lidars now appear capable of accurate measurement of the ozone profile from 15 km to 50 km. On average, the JPL and GSFC lidars agreed to within ±3% between 23 and 43 km. Above 43 km the GSFC lidar became noisy, while the JPL lidar performed well to 50 km. Below 23 km the JPL lidar had an increasingly large bias caused by an instrument saturation problem that has since been corrected. The GSFC lidar agreed with the ECC-sonde to within ±5% in the 16 to 25 km region.

The single SAGE II profile agreed with the lidar profiles within ±5%, between 20 and 37 km, but above that altitude SAGE II was 5-10% higher than the JPL lidar profile. SAGE II performed well in the lower stratosphere, agreeing with the ECC sonde profile within ±5% in the 16 to 25 km region. The ground-based microwave instrument agreed with the lidars within ±5% from its lower limit of 21 km up to 43 km. The maximum discrepancy relative to the JPL lidar was -10% at 48 km. The ECC sonde agreed with lidar, SAGE II, and microwave to within ±5% at altitudes up to 25 km, but at higher altitudes an apparent pump correction error led to an increasing discrepancy, reaching +10% above 30 km.

In this campaign instruments that measure ozone density versus altitude were compared with ones that measure ozone mixing ratio versus pressure. The pressure-altitude conversion adds uncertainty to the comparison, particularly for instruments that do not measure a temperature profile and must rely on NCEP data for the conversion. Lidar and SAGE II comparisons with MLS, SBUV/2, and Umkehr are affected by this problem. MLS agreed with the JPL lidar within ±10% between 43 hPa (the MLS lower limit) and 1 hPa. Umkehr generally agreed with lidar within ±10%, but the errors were systematic with altitude, going from -5% at 30 hPa, to agreement at 10 hPa, to -10% at 3 hPa, to -5% at 1 hPa.

Groundbased Microwave/SAGE II Comparisons:

Records of time series of microwave observations made at the Table Mountain Facility (August 1989-June 1992) and subsequently at Lauder (1993-1996) were compared with coinciding SAGE-II measurements. The microwave radiometer instrument, described by Parrish *et al*. (1992), has a precision of 4-5%, an accuracy of 5-7%, and a vertical resolution of 8-13 km in the range 56-1 hPa (Tsou *et al*., 1995 and Connor *et al*., 1995).

Figure 2.41 (left panel) shows the average and standard deviation of the difference between microwave and SAGE II during the 2 years of operation at Table Mountain. The average difference, based on 37 coincidences, is typically < 3% from 20-47 km. Examination of subsets of these data show no significant dependence on season or on the type of SAGE II event (sunrise or sunset). Good agreement between the data sets is achieved despite the very different vertical resolution of the instruments. Tsou *et al*. showed empirically that accounting for the vertical resolution correctly, by degrading the SAGE resolution to that of the microwave, has only a very small effect on the mean difference.

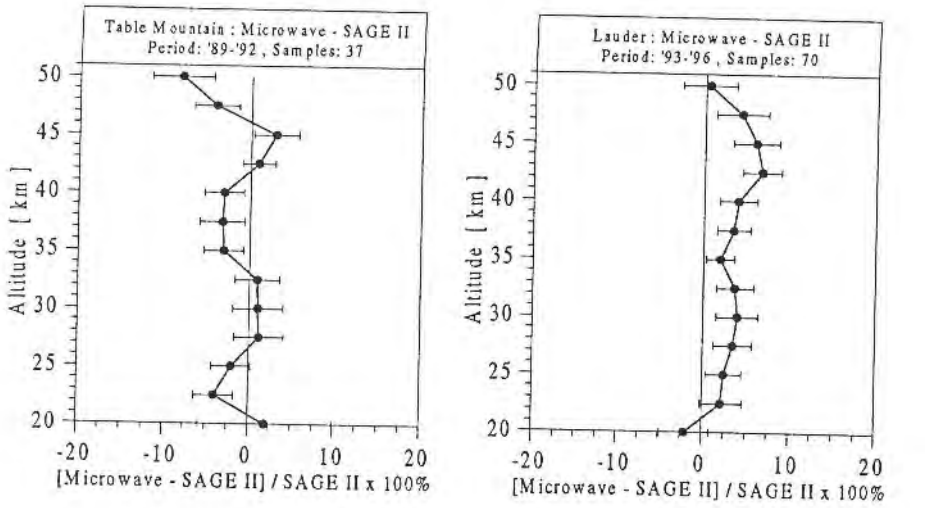


Figure 2.41 Mean of the relative differences between Microwave and SAGE II at Table Mountain from 1989 until 1992 (left panel) and at Lauder from 1993 until 1996 (right panel). Error bars are 2-sigma standard error.

The results of the comparison at Lauder are displayed in the right panel of Figure 2.41 and are based on 70 microwave/SAGE II coincidences. The mean difference is in the range +0-6% at altitudes between 20-50 km. Unlike the Table Mountain results, the sunrise and sunset comparisons are significantly different. The sunrise difference is about 5% larger at all altitudes than the sunset difference, suggesting that the sunrise measurements would be systematically smaller than the sunset measurements if they were sampling the same air. However, it is known that the microwave data at Lauder are affected by a bias of 1-3% during the observational period.

It is concluded that the ground-based microwave data shows good agreement with SAGE II and is a promising instrument for the future to validate and assess ozone trends in the stratosphere.

2.5. Ozonesonde Analyses

2.5.1. Introduction

In the previous sections 2.3.3.2 and 2.4.1, time series of ozonesonde measurements obtained by various sounding stations were compared with SAGE I and SAGE II respectively to investigate differences or drifts between the different long-term time series of ozone measurements in the lower/middle stratosphere. The emphasis of this section is on comparison of ozonesondes with other non-satellite ozone measuring systems to quantify the precision and accuracy of the three different types of ozone sondes, i.e. Brewer-Mast (BM), Electrochemical Concentration Cell (ECC), and the Carbon Iodine (KC79), that are in routine operation at the various sounding stations (see chapter 1).

Since 1970 dedicated short term campaigns to intercompare the different types of ozone sondes used in operational networks have been carried out at different times during the course of the ozone sonde record. These intercomparisons are intended to assess the performance of and to quantify any systematic differences between the various types of sondes. Most of the campaigns have been conducted in the field where at least two ozone sondes were carried to altitudes of about 30 to 35 km by the same balloon. The WMO Jülich Ozone Sonde Intercomparison Experiment (JOSIE) held in 1996 is the first intercomparison campaign that compared all operational types of ozonesondes in a controlled environmental chamber capable of simulating real flight conditions (Smit *et al.*, 1997). The controlled environment plus the fact that the ozonesonde measurements could be compared to an accurate UV-Photometer as a reference allow the opportunity to carry out experiments to address questions arising from the previous field comparisons. However, short term intercomparisons are more or less "snap shots" and may not necessarily reflect the performance of ozonesondes under operational field conditions. Comparison studies of time series of ozonesonde data with other simultaneously operating ozone monitoring devices like lidar or microwave are more valuable to assess the data quality of the ozonesonde measurements in regular operation.

In the free troposphere ozone soundings provide the only time series of measurements to derive long term ozone trends. Due to the much lower concentrations of ozone in the troposphere compared to the stratosphere the performance of the sondes and their typical instrumental/operational factors determining precision and accuracy are different in the two regions of the atmosphere. In this section special attention will be paid to how much confidence exists in the data quality of the different ozonesonde records available for ozone trend assessment in the troposphere.

2.5.2. Laboratory Studies: Jülich Ozone Sonde Intercomparison Experiment (JOSIE)

In February/March 1996 JOSIE was conducted in the environmental simulation chamber at the Research Centre Jülich (Germany) to assess the performance of the major types of ozonesondes (ECC , Brewer-Mast , KC79) used within the Global Atmosphere Watch (GAW) and Global Ozone Network (GLONET). Eight ozone sounding laboratories participated in the intercomparison effort. Four ozonesondes were "flown" simultaneously and compared to a UV-photometer as a reference, while pressure, temperature and ozone concentrations were regulated to simulate different types of vertical soundings with an ascent velocity of about 5 m/s up to 35 km altitude (Smit *et al.*, 1994).

Four different types of ozonesondes were "flown" simultaneously covering six simulated ascents (four mid-latitude and two tropical profile types). Quantitative results of the sonde comparisons with the UV-photometer are shown for each participating sounding laboratory in Figures 2.42a to 2.42h. The comparison is presented as relative deviation of the sonde readings from the UV-photometer (typically

within ±10-20%). However, some systematic deviations were observed with increasing magnitude in the middle stratosphere for some participants (CMDL [2], JMA [8]). Apart from systematic deviations, there are precision errors of varying size for the different participating laboratories. The best precision (±5%) was achieved by CMDL [2], KFA [3], AES [6] and CNRS [8] (excepted one outlier profile) which were operating the ECC-type ozonesondes. Conversely, the non-ECC types of ozonesondes, operated by MOHP [1], IMD [4], ASP [5] and JMA [8], exhibit a somewhat lower precision of about ±10-15%. The relative precision is best for all sondes in the middle stratosphere where ozone is at a maximum. The non-ECC type sondes show a considerably lower precision in the troposphere and stratosphere compared to the ECC sondes.

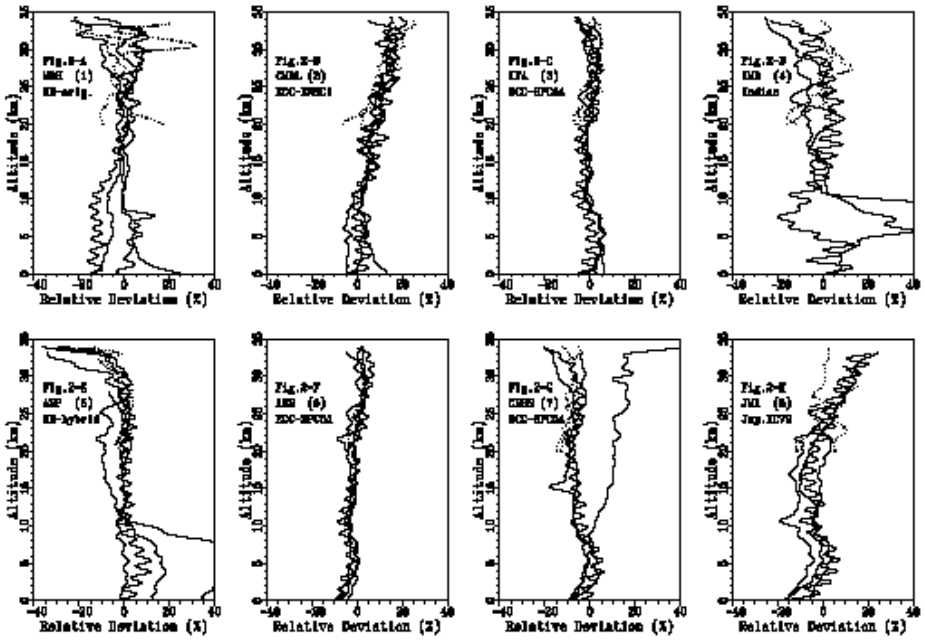


Figure 2.42 JOSIE : Relative deviations of the individual sonde readings from the UV-photometer for each participating laboratory obtained from the six simulation runs of the mid latitudinal (solid lines) and the tropical profiles (dotted). Tropospheric part of tropical profiles below 20 km are excluded. ECC-SPC5A and ECC-SPC6 are manufactured by Science Pump Corporation and ECC-ENSCI by ENSCI-Corporation. BM-orig. is the original model of the Mast-Company, BM-hybrid is a hybrid of an ozone sensor from Mast-Company coupled with a teflon pump from ENSCI-Corporation. The Indian sonde is made from a Brewer-Mast like ozone sensor and a teflon pump manufactured by the Indian Meteorological Department. Jap.KC79 is the Japanese KC79 sonde.

Table 2.8 shows the total ozone column normalisation factor, which is the ratio of the integrated ozone profile measured by the UV-photometer and each sonde, for each sonde "flown" plus the average factor and standard deviations for each participating laboratory. Although all individual normalisation factors range between 0.9 and 1.1, the variability (standard deviation) of the ECC-sonde types with values of ±0.02 for three of the four ECC-stations are significantly smaller than the variability of the other types of sondes with values of ±(0.07-0.10). In the case of the CNRS [7], the standard deviation of 0.07 is mainly caused by the relative low normalisation factor obtained from the first simulation run in campaign II (see Table 2.8 and Figure 2.42G).

Sonde Type / Participant [No]	Total Ozone Normalisation Factor individual simulation runs						Average	Standard Deviation
BM (original) / MOHP [1]	0.95	1.03	1.01	1.13	0.94	0.94	1.00	0.07
ECC (ENSCI) / CMDL [2]	0.93	0.92	0.90	0.88	0.93	0.92	0.91	0.02
ECC (SPC6A) / KFA [3]	1.01	0.98	0.97	0.98	1.00	1.01	1.00	0.02
Indian / IMD [4]	1.06	(*)	0.97	1.05	0.90	1.06	1.01	0.08
BM-Hybrid / ASP [5]	1.03	1.01	0.91	1.06	1.22	1.05	1.05	0.10
ECC (SPC5A/ENSCI) / AES [6]	0.96	0.97	0.94	(*)	0.97	0.94	0.96	0.02
ECC (SPC5A) / CNRS [7]	0.90	1.09	1.03	1.09	1.07	1.07	1.04	0.07
KC79 / JMA [8]	0.86	0.96	0.84	0.89	1.09	0.85	0.91	0.10

Table 2.8 : JOSIE: Summary of total ozone column normalisation factors. Factors marked by (*) are non valid sonde data. Total ozone normalisation were only applied by the following stations: MOHP [1] , IMD [4] , ASP [5] , AES [6] and JMA [8].

Even if the total ozone normalisation (correction) factor is not used to correct the sonde profile it provides an excellent screening test for suspect soundings. However, the normalisation factor is not a guarantee that the profile is correct. Trends or discontinuities in normalisation factor time series can also be a cause for concern (see Chapter 3).

The results of JOSIE presented here, show that in general the sondes agree well with the ozone reference. Most sondes track the simulated ozone profile quite well, even under extreme low ozone concentrations in the tropical troposphere. The JOSIE results show that the ECC-type sondes perform better than other sonde types with regard to precision and accuracy. The observed differences are mostly due to differences in the preparation and correction procedures applied by the different laboratories. The performance of the non-ECC type sondes is quite different as are the observed deviations from the UV-photometer measurements. The results of JOSIE also demonstrate that there is an urgent need for the homogenisation of operating procedures. Further, the larger variability among ozonesonde measurements in the middle stratosphere is due mainly to greater uncertainties in the pump efficiency and the amount of sensing solution that has evaporated.

The increasing variability above 25 km altitude, particularly observed in the case of the non-ECC sonde types, is mostly caused by the decaying pump efficiency and its increasing uncertainty at lower pressures. At the present the pump flow rate is typically corrected at lower pressures by using an average pump efficiency measurement curve (see chapter 1). More recent measurements indicate that the pump efficiency is smaller than the average by about 10-12% at 5 hPa for the ECC-Sonde (Johnson *et al.*, 1997) and by even larger amounts for the Brewer Mast sonde (Steinbrecht *et al.*, 1997; De Backer *et al.*, 1997). It should be noted that significant differences (more than 5-10% at 5 hPa) are observed between the various experimental methods used to determine the pumpflow efficiency at lower pressures.

An additional source of uncertainty is that by the time an ozonesonde reaches the middle stratosphere it has been operating for nearly 90 minutes and at this stage of the flight the uncertainties in the sensor cell characteristics are greater. First, a certain percentage of the sensing solution has evaporated at a rate dependent on the temperature of the cell and ambient pressure encountered during flight. For the ECC-sondes this means that due to evaporation, the concentration of the sensing solution increases, which can have an enhancing effect on the sensitivity of the ECC-sensor and thus on the measured ozone (Barnes *et al*., 1985; Komhyr, 1969).

There are some experimental indications for ECC-sondes, that if the preparation and correction procedures prescribed by Komhyr (1986) are used, this sensitivity enhancement effect can compensate for the too low conventional pumpflow correction in the middle stratosphere. During JOSIE, this may be the case for the ECC-sondes operated by KFA [3], AES [6] and CNRS [7] which were using the same initial concentrations of sensing solutions and the same pumpflow corrections (Komhyr, 1986) and did not show any systematic deviations from the UV-photometer in the middle stratosphere. This in contrast to the overestimation above 25 km altitude by the ECC-sondes operated by CMDL [2] which is probably caused by the larger pumpflow correction (based on individual pumpflow efficiency measurements in the laboratory) applied in combination with the sensitivity enhancement of the ECC-sensor due to evaporation. However, the process is not understood in detail up to now and more investigations are necessary to study this particularly in relation to the initial KI-concentration, the actual temperature of the sensing solution and the pumpflow correction efficiency which also has an important influence on the sonde performance.

The uncertainty in sonde measurements of tropospheric ozone is relatively high since the signal-to-noise is low. The impact of instrumental errors is larger when measuring the much lower values of tropospheric ozone. One instrumental error is uncertainty in the sensor background current which varies in magnitude from one sonde to another as well as from one sonde type to another. Correction for the background current can have a significant impact on the measured tropospheric values in regions where the ozone concentration is low, i.e. near the tropopause. For ECC-sondes the conventional method of correction prescribed by Komhyr (1986) assumes the background current is dependent on the oxygen partial pressure and decreases with altitude. Several laboratory studies (Thornton *et al*., 1982; Smit *et al*., 1994) do not show any oxygen dependence on the background current and the accuracy of the ECC-sonde is significantly improved by using a constant background current correction throughout the entire vertical profile.

Recent laboratory studies of ECC-sondes have also shown that the background signal can be correlated to past exposure to ozone. As a consequence, the background current should be measured before exposure to ozone in the preparation procedure as was shown in a field study by Reid *et al*. (1996). No background correction is made to the BM-sonde records, but prior to flight the BM-sonde readings are electronically compensated for the background current. Nothing is known about the fate of the background of the BM-sonde during flight. Any changes in the magnitude of the background current over the sonde record will directly affect the trends derived for the free troposphere. It is important that more research be dedicated to the size and impact of the changes in the background current of different sonde types.

Another source of uncertainty is the influence of local air pollution which can have detrimental effect on the sonde performance. Most known is the negative interfering effect by SO_2, which lowers the ozone readings and can have a memory effect if excess of SO_2 is accumulated in the sensing solution of the sensor (Schenkel *et al*., 1982). At Uccle the SO_2 contamination was recognised as a serious problem in the lower tropospheric part of the sounding profile and a procedure to correct for the SO_2 interference has been developed (DeMuer *et al*., 1993). It may be possible that in the early 1970's when the SO_2

concentrations were much higher than in the 1980's and 1990's, the records of other stations like Hohenpeissenberg, Tateno and Sapporo are also affected by SO_2 interferences (Logan, 1994).

2.5.3. Dedicated Short Term Ozonesonde Intercomparison Campaigns

During the course of developing the ozonesonde record, several dedicated, short term intercomparison campaigns in the field were conducted. A summary of the campaigns, which started in 1970, is given in Table 2.9.

Name	Location	Date	Brewer-Mast	ECC	KC68/79, Japan	B/M-India	UV-Photometer	Lidar	Microwave	Chemolum.	Satellite	ROCOZ	Reference
WMO-I	Hohenpeissenberg, Germany	Jan-Feb, 1970	x	x	x	x							Attmanspacher et al., 1970
WMO-II	Hohenpeissenberg, Germany	Apr, 1978	x	x	x	x							Attmanspacher et al., 1981
BOIC	Palestine, USA	Jan, 1983 & Mar, 1984	x	x			x						Hilsenrath et al., 1986
MAP/ GLOBUS	Gap, France	Sep, 1983	x	x			x			x			Aimedieu et al., 1987
STOIC	Table Mountain, USA	Jul-Aug, 1989		x				x	x		x	x	Komhyr et al., 1995-A/-B
OHP-I	Haute Provence, France	Sep, 1989	x	x				x					Beekman et al., 1994
OHP-II	Haute Provence, France	Mar-Apr, 1991	x	x			x	x					Beekman et al., 1995
WMO-III	Vanscoy, Canada	May, 1991	x	x	x	x							Kerr et al., 1997
MLO-3	Hilo, Hawaii	Aug, 1995		x				x	x		x		
JOSIE	Juelich, Germany	Feb, 1996	x	x	x	x	x						Smit et al., 1997
SONDEX	Payerne, Switzerland	May, 1996	x	x				x					
SCUV	Gap, France	Jun, 1996		x									

Table 2.9 : Summary of ozone sonde intercomparisons between 1970 and 1996. Dates and locations of the intercomparisons are shown and the types of profiling instruments with a reference are indicated.

The earlier campaigns included only ozone sondes so comparison of sonde results to a reference profile measured by a separate "standard" technique is not possible. This type of comparison allows the determination of relative errors between the different sonde types as a function of altitude but not absolute systematic errors. However, all comparisons normalised the profiles to a common ground-based total ozone measurement to minimise systematic differences. Later intercomparisons (BOIC, MAP/GLOBUS, STOIC, OHP I & II, MLO3, JOSIE, GAP) used a reference profile measured by other techniques and results of these comparisons yield better estimates of absolute errors for the sonde measurement as a function of altitude. A significant amount of information has been learned from these intercomparisons and a review of results from the campaigns has provided information regarding the confidence with which the sonde data may be used for long term trend analysis. The campaigns offer the opportunity for operators to compare flight preparation procedures and to evaluate the effects of any differences.

Results of the intercomparisons indicate that sonde measurements made in the lower stratosphere (12 to

27 km) are quite reliable for use in long-term trend analyses. The systematic differences between different sonde types have been less than 5% and the random variability from one sonde to another for all sonde types has been less than 5%. This relatively low variability is a result of the normalisation of the profile to ground-based total ozone measurements (see Chapter 1). The normalisation is mainly weighted to the ozone in the lower stratosphere which contains most of the column ozone.

The intercomparison projects have shown that the variability among ozonesondes generally increases again in the middle stratospheric region (27 km to balloon burst altitude). For example, in the BOIC campaign 3 ECC ozonesondes (from different groups) and 1 Brewer-Mast sonde showed differences of about 10% as the balloon approached 30 km altitude. Results of the WMO III Vanscoy intercomparison included two additional types of sondes, the Indian and Japanese KC79 sondes. The variability in comparing all 4 types of sondes was even higher at about 10-15%, reaching 20% at 5 hPa (~35 km). Radiosonde pressure variability adds to the uncertainty in ozone profiles in the middle stratosphere, which was observed in the Vanscoy (WMO III) intercomparison. The magnitude of variations and errors in ozonesonde measurements in the middle stratosphere have been fairly consistent in showing larger variability in this region from past intercomparison projects and individual tests, thus indicating that trend analysis may be less reliable at these altitudes. However, recent results from the JOSIE campaign have shown very good precision and accuracy extending into the middle stratosphere, especially for the ECC type sondes.

The issue of tropospheric ozone measurement accuracy is addressed fully in the next section, but there are a number of results from the intercomparison exercises, which can have a bearing on trend determinations in the troposphere. Results of the intercomparisons show systematic differences between sonde types which are typically 10 to 15% but can be as much as 25% for some campaigns. Most campaigns (e.g. WMO I, WMO II, BOIC, OHP I, OHP II) indicate that the ECC sonde measures about 15% to 30% more than the Brewer-Mast in the troposphere. However, results of WMO III suggest that the Brewer-Mast measures about 15% more than the ECC in the troposphere. This marked change over time , with the BM-sonde lower than the ECC in 1970 and larger in 1991 may indicate that any determination of the long term ozone trend in the troposphere is subject to considerable uncertainty, particularly when different ozonesonde types are been used for different parts of the record.

Like the JOSIE-campaign findings, the field intercomparisons have shown that in the troposphere the sonde measurements are less precise (typically about ± 15%, 2-sigma) than in the stratosphere. In general the Brewer-Mast type sondes are less precise than the ECC sondes. One outcome of the intercomparisons is the discovery that occasionally sondes measure erratically (errors >± 50%) in the troposphere with no apparent explanation. The tropospheric error would not be apparent if these profile measurements were considered on their own.

It should be noted that the short term campaign results may not necessarily reflect the performance of ozone sondes under operational field conditions. Particularly for the non-ECC-type of sondes, like the older established BM-sounding stations (Hohenpeissenberg, Payerne and Uccle), the in-flight performance characteristics appear to be strongly coupled to the operational procedures followed at the different stations, but also to the location of the launch site. In the next sections the performance of the different sonde types is evaluated through comparison of time series with other ozone profiling techniques in regular operation at the sounding station site.

2.5.4. Ozonesonde Comparison Studies in the Troposphere

Identification of trends in tropospheric ozone from the ozonesonde record demands that the measurement

accuracy of the sondes remains constant over a long time. As already mentioned, from the WMO intercomparison campaigns it appears that large relative changes in accuracy between BM and ECC sondes may have occurred between 1970 and 1991. However, as will be shown in this section, there is also strong evidence that the ECC and BM records (at least at Hohenpeissenberg) are more reliable than results from the WMO intercomparisons would imply.

The ozone profiling capabilities of the ECC-sonde during regular operations were evaluated in a study by Ancellet and Beekmann (1997) through comparison with routine lidar measurements made at the Observatoire de Haute Provence in the 1990-1995 period (Ancellet *et al.*, 1989). The seasonal means of mid tropospheric ozone (4.5-5.5 km altitude) obtained by the sonde and lidar are summarised in Table 2.10. ECC values are not corrected by total ozone normalisation. Both data sets show excellent agreement in the annual mean (54.3 ppb for ECC, 53.8 ppb for Lidar) and in seasonal variations in the mid troposphere. Differences in particular periods are generally in the range of 2 ppb (5%). For 15 simultaneous ECC versus lidar profiles the mean of the differences observed between 4 and 7 km was 2.5± 1.8 ppbv (4± 3%) which is in good agreement with differences between the seasonal means (see Table 2.10). This comparative study indicates that tropospheric ozonesonde profiles measured with ECC-sondes should not be corrected for the total ozone column. Supporting evidence exists from simulation experiments performed at the ozonesonde calibration facility in Jülich (Smit *et al.*, 1994). Although there is a dearth of comparative studies to evaluate the performance of ozonesondes in the troposphere, this present study under regular field operations together with the studies under controlled laboratory conditions show that there is sufficient confidence in the performance of the ECC-sondes to use them for tropospheric trend assessment studies.

Instrument	ECC / OHP		Lidar / OHP	
Period	1991-95		1990-95	
Time	mean	n	mean	n
Nov.-Feb.	49.0 ± 0.8	66	47.2 ±1.2	96
March	53.5 ± 2.7	18	53.8 ± 2.7	23
Apr.-Jun.	65.2 ±1.5	52	63.2 ±2.3	56
Jul.-Sep.	63.0 ± 1.5	54	60.7 ± 1.8	74
October	53.1 ± 2.1	52	55.2 ±3.0	14
all	54.3 ± 0.6	207	53.8 ± 1.2	263

Table 2.10 : Seasonal mean ozone values and the standard deviation (in ppbv) of the mean in the mid-troposphere (4.5 -5.5 km height) at the Observatoire de Haute Provence; n denotes the number of profiles.

As mentioned before, one major uncertainty in BM profiles is the pump flow correction at lower

pressures (<100 hPa). Individual BM pumpflow calibrations in the laboratory [De Backer *et al.*, 1997, Steinbrecht *et al.*, 1997] have shown that at lower pressures the pumpflow correction should be significantly larger than the current correction recommended by the WMO [Claude *et al.*, 1986]. Although the correction itself is only applied to the upper part of the profile, the additional act of normalising to the total column measurement means that erratic pump corrections will also show up in the tropospheric part of the profile.

The effect of the pump correction factor on the tropospheric performance of BM-sondes in combination with the use of the total ozone normalisation was investigated at Uccle using both ECC and BM sondes on the same balloon. For each BM-sonde flown, the pump efficiency was individually measured in the laboratory before launch. The results from 20 dual sounding are shown in Figure 2.43. When the BM sondes are corrected with the standard WMO pump correction, the BM-sondes measure about 10% more ozone in the troposphere than the ECC-sondes and 3-5% less ozone above 20 km altitude. However, when using the individual measured pump efficiency corrections, the differences between BM and ECC soundings are within ± 2% and show no significant altitude dependence. It should be noted that the rms difference of 8-10% between the ECC and BM measurements in the troposphere is consistent with the results obtained during JOSIE (see section 2.5.2). This study shows that the standard pump correction recommended is inadequate for the BM-sonde and because of the total ozone normalisation, it can have a significant effect on tropospheric ozone values.

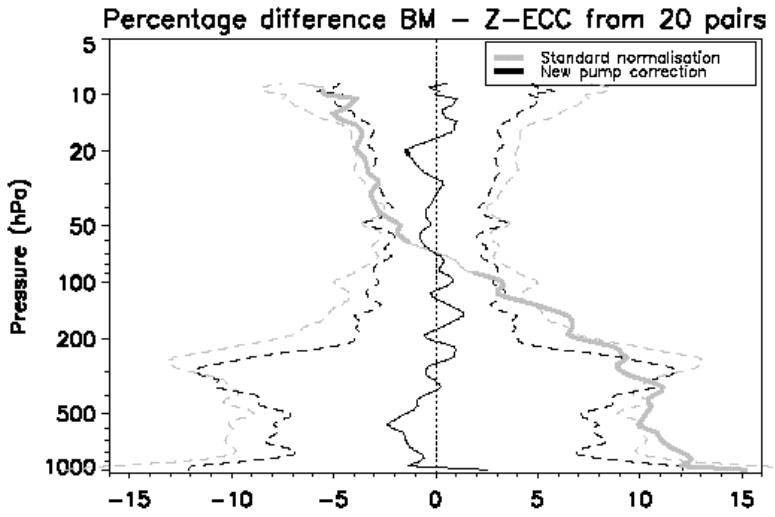


Figure 2.43 Mean percentage differences between ozone profiles obtained from measurements with BM and ECC sondes with the standard (grey) and the new pump (black) corrections in Uccle. The domains where the differences are statistically significant at the 95% confidence limit, according to a 2-tailed student-t test are indicated by thick lines. The dashed lines give the 1 sigma levels of the differences.

This issue can be relevant for trends if the recently measured pumpflow corrections, which are larger than the standard correction recommended by WMO, has been caused by a decrease in the BM-pump efficiency (at lower pressures) during the sonde record. However, at present this is not clear and the observed differences may be caused by differences between the different experimental methods used to measure the pump efficiencies at low pressures. Unfortunately, the standard pump flow corrections recommended for both the BM-and ECC-sondes are based on rather old pump flow efficiency measurements which were performed more than 25 years ago and they never have been officially re-evaluated in the literature. Therefore, it is rather difficult to trace any long term changes in the pumpflow

efficiencies at lower pressures during the sonde record.

The previous study at Uccle also suggests the need to reconsider the concept of applying the total ozone normalisation factor to the tropospheric part of the ozone profile measured by the BM-sonde. Results from comparison studies (Thouret *et al*., 1997; and Jeker, private communication) of BM-soundings at Hohenpeissenberg and Payerne with ozone measurements made from aboard in-service aircraft suggest that it is indeed more proper not to use the total ozone normalisation factor in the tropospheric part of the sounding. Similar findings were obtained from the BM/ECC-comparison flights during SONDEX 96 (Schmidlin, private communication). However, this issue is still in an explorative stage and more research is required.

To evaluate the long term performance of the BM ozone soundings at Hohenpeissenberg the sonde record obtained at 700-800 hPa was compared with night-time measurements made at the Zugspitze mountain station at 2970 m altitude (700 hPa). Night-time measurements at this site are generally indicative of free tropospheric conditions and they have been made with a UV-photometer or a chemiluminescent instrument since 1978. The results of the comparison is shown in Figure 2.44. Although there is an offset between the measurements (800 hPa data from the sondes agrees better with the Zugspitze data than the 700 hPa measurements), there is no major anomaly in the long term variations measured by the sonde and at the Zugspitze. At both stations the strongest increase in ozone mixing ratio was not observed until the beginning of the eighties. Since then no significant trend has been observed. These observations strongly suggest that the trend in tropospheric ozone derived from the Hohenpeissenberg data is reliable to about 0.3% per year and it is obviously not affected by uncertainties induced via the pump flow correction in combination with total ozone normalisation. However, the offset between Hohenpeissenberg 700 hPa and Zugspitze data (comparable altitude) might be due to regional variations in the vertical ozone distribution, but it could also be caused by the total ozone normalisation leading to an 8-10% shift in the tropospheric data or from other non identified offsets.

A similar comparative study has been done for the 700 hPa level of the Payerne BM-ozone soundings which were compared with the data sets of 2 high alpine surface monitoring stations (Jungfraujoch and Zugspitze) to identify the anomalies observed between 1990 and 1993 (see Figure 2.45). Since 1992, very good agreement between the annual ozone concentrations at Zugspitze and Jungfraujoch has been found, and the deviations relative to the sounding data remained approximately constant with time before 1990 and after 1993. The small temporal drift in the difference "sounding-Zugspitze" has been further reduced by applying the results of the pre-flight laboratory calibration of the ozone sondes (see chapter 1). However, it should be noted that the 1982/1983-period the Payerne sonde record show inconsistencies in the troposphere which cannot be explained and are probably not of atmospheric origin (Staehelin, private communication). Therefore, the ozone record over this period has to be taken with caution to derive tropospheric ozone trends.

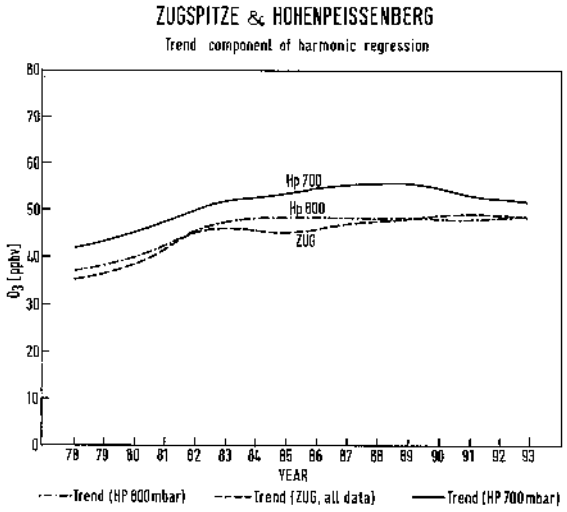


Figure 2.44 Ozone trend component based on monthly means of Hohenpeissenberg Brewer Mast data at 700 and 800 hPa (approx. 3000 and 2000m a.s.l.) and continuous measurements at the Zugspitze (2960 m a.s.l.) since 1978 *[courtesy of K.E. Scheel].*

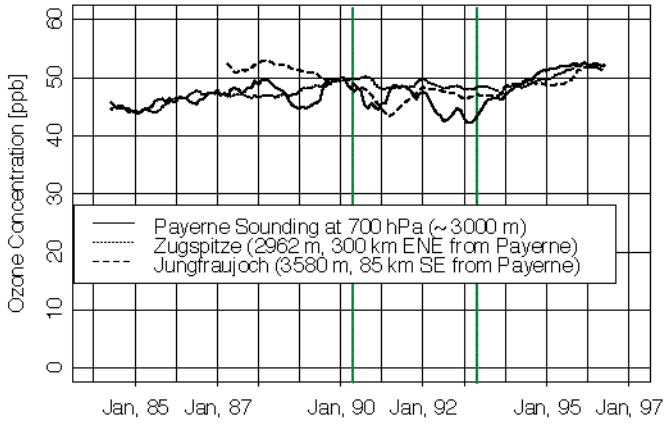


Figure 2.45 Time series as 12-months moving averages of the monthly means of ozone [ppb] of the Payerne sounding (700 hPa, ~ 3000 m asl) and of the surface monitoring stations of Jungfraujoch and Zugspitze between 1984 and 1996. *[Zugspitze-data by courtesy of Dr H. E. Scheel, Jungfraujoch-data by EMPA, Switzerland].*

2.5.5. Ozonesonde Comparison Studies in the Stratosphere

The results of the short term intercomparisons have in general shown that the sondes are giving consistent results in the lower/middle stratosphere between 12 and 27 km provided the profiles are normalised to a ground based total ozone measurement. The different sonde types have consistently agreed to within 5% for this height range. The results are not as conclusive for altitudes above about 27 km. In this section the conclusions drawn from the "snap shot" intercomparisons are evaluated through comparison of individual sonde records with an independent ozone profiling technique done under regular operation at the sounding site.

At a few sounding stations over the last 5-10 years, other vertical ozone profiling techniques like lidar and/or microwave instruments, have been used to monitor the vertical ozone distribution. These data

have been used in this assessment to validate the different types of sondes in regular operation. The sounding sites, sonde type, "reference" instrument and period of comparison are listed in Table 2.11. All installed instruments are dedicated to stratospheric profiling as part of the Network for the Detection of Stratospheric Change (NDSC). The lidar systems listed have been also used in this assessment to validate SAGE II data (section 2.4.2), and the microwave radiometer has been used in a similar way at Lauder (section 2.4.8).

Comparison Specifications					[(Sonde-Ref)/Ref] x 100%			
Sounding Site	Sonde Type	Comparison Instrument	Observation Period	Number Pairs	15-20 km	20-25 km	25-30 km	30-33 km
Lauder	ECC	Lidar	11/94-12/96	» 80	-2± 2	2± 2	4± 2	3± 2
Lauder	ECC	Microwave	11/92-12/95	» 130		-2± 2	-1± 2	-1± 3
Haute Provence	ECC	Lidar	11/91-12/96	» 55	2± 4	3± 2	2± 3	-3± 3
GSFC Mobile Station	ECC	Lidar	10/88-08/95	» 90	6± 5	4± 2	3± 1	1± 2
Hohenpeissenberg	BM	Lidar	11/87-11/95	(*)	4± 3	3± 2	-3± 3	-10± 5
Payerne	BM	Microwave	11/94-09/96	» 300	-3± 3	-2± 2	-5± 5	
Haute Provence	BM	Lidar	07/86-12/90	5		5± 5	-2± 8	
Tsukuba	KC79	Lidar	11/88-05/91	» 45		-2± 4	3± 3	6± 4

Table 2.11: Summary of ozonesonde comparisons with other ozone profiling techniques in the stratosphere. Pairs marked with asterisks (*) are based on differences between monthly means of sonde and lidar data

Time series of the differences between the sonde record and the co-existing "reference" instrument were calculated for each site at altitude levels between 15 and 35 km with an altitude step of about 2.5 km. The time coincidence criterion between sondes and microwave measurements was ± 24 hours. Differences between the monthly means of the sonde and the lidar were obtained only for the BM-Sonde/Lidar comparison at Hohenpeissenberg. The vertical profiles of the mean relative differences (± 2 standard error) between the different sondes compared with the co-existing "reference" are shown in Figure 2.46 for ECC sonde comparisons and in Figure 2.47 for the non-ECC sonde comparisons. In addition, the major results of the comparisons are summarised in Table 2.11.

The ECC sonde (Figure 2.46) shows agreement with lidar or microwave data to better than ± 2.3% for individual comparisons throughout the lower/middle stratosphere up to the balloon burst altitude at 33-35 km. Averaged over the three comparisons with lidar the differences are ~2± 3% at 15-20 km, 2± 2% at 20-25 km, 2± 2% at 25-30 km, and -1± 3% at 30-33 km. It is obvious that the ECC-sonde tends to slightly overestimate the ozone compared to the lidar, even in the altitude region above 25 km where the conventional pump correction is too low. This is probably caused by the compensating effect of the ECC-sonde sensitivity increase due to an enhancement of the concentration of the sensing solution by evaporation during the course of the sounding (see section 2.5.2). This effect even slightly dominates up to an altitude of 30 km. The small negative differences obtained from the comparison with the microwave at Lauder can be also caused by the fact that the microwave instrument slightly overestimated ozone during the observational period (see microwave SAGE II comparison in section 2.4.8).

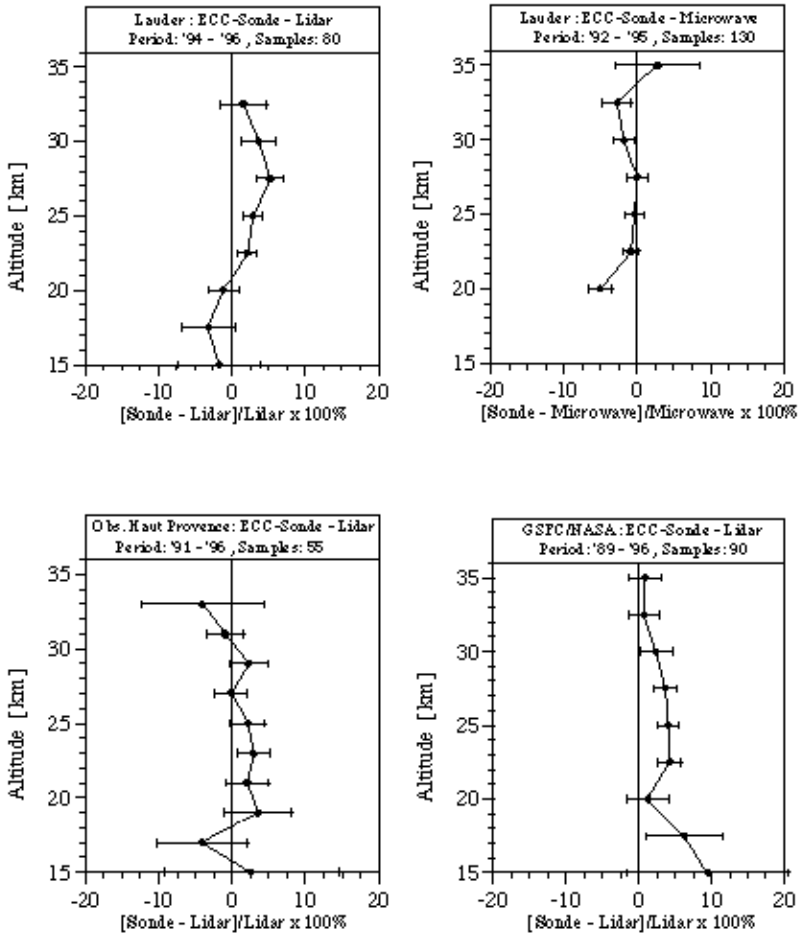


Figure 2.46 Average differences between ECC sondes and lidar or microwave radiometer as a function of altitude at a ± 24 hours coincidence criterion. The error bars are ± two sigma standard error.

The long term comparison between BM sonde and lidar data at Hohenpeissenberg is shown in Figure 2.47 (upper left panel). Between 15 and 28 km altitude the sonde readings are about 3± 3% larger than the lidar readings. Above 28 km altitude the BM-sonde readings are significantly decreasing with altitude compared to the lidar (-6% at 35 km). This is due to the deteriorating pump efficiency of the sondes, which is not properly corrected by the current WMO procedure [Steinbrecht *et al.*, 1997]. However, at present there is no evidence that this problem can cause any trend effects in the sonde record at Hohenpeissenberg.

BM sondes at Payerne (see Figure 2.47, upper right panel) underestimate ozone in the lower stratosphere by about -3± 3% compared to microwave measurements. Above 28 km the sonde readings are too low due to improper pump corrections of the sonde data, similar to the results of the BM-soundings obtained at Hohenpeissenberg.

Comparison of Japanese KC79 sondes with lidar at Tsukuba (see upper right panel of Figure 2.47 (Fujimoto *et al.*, 1997)) show that the sondes tend to underestimate ozone between 18 and 23 km altitude, but overestimate ozone by about 2 to 4% in the region between 23-30 km and more than 5% above 30 km.

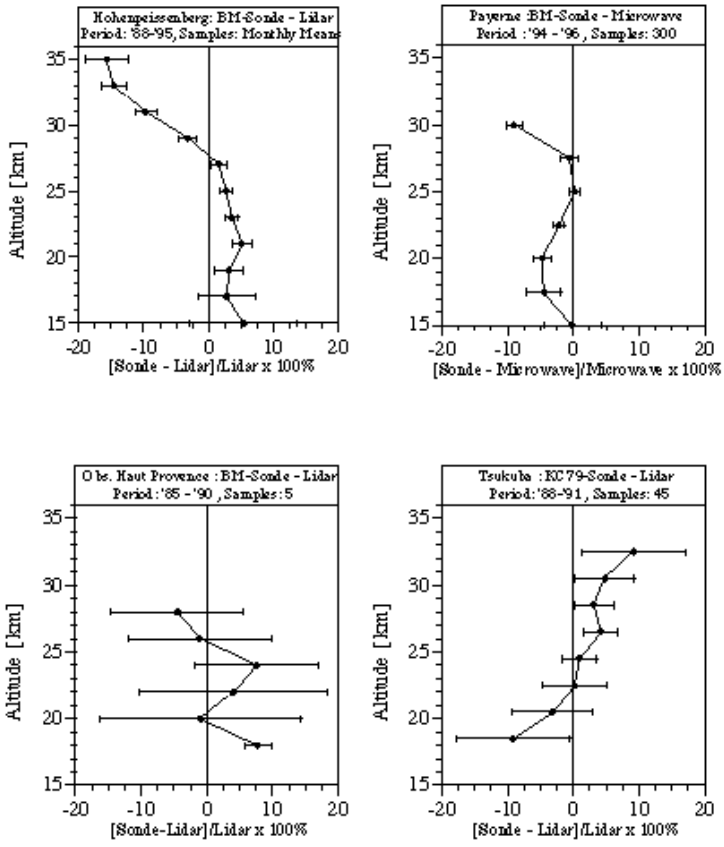


Figure 2.47 Average differences between non ECC sondes (Brewer Mast and KC79) and lidar or microwave radiometer as a function of altitude at a ± 24 hours coincidence criterion. The error bars are ± two sigma standard error.

2.5.6. Summary and Conclusions

In general all intercomparison studies have indicated that in the lower to middle stratosphere between the tropopause and ~ 28 km the three different sonde types show consistent results provided the individual measured sonde profiles have been normalised to ground based total ozone column measurements at the launch site. In this altitude range the precision of the various sonde types is within ± 3% , while any systematic bias compared to other ozone sensing techniques are smaller than ± 5%. For altitudes above 28 km the results are not so conclusive and the measurement behaviour of the sondes are different and cannot be generally characterised. The Brewer Mast sonde used by the established (long term record) stations (Hohenpeissenberg, Payerne, Uccle) show systematic under-estimations of ozone which increase with altitude (-15% at 30 km). For the ECC-sondes, there is some evidence suggesting that

measurements agree with each other and to reference techniques to within ± 5%. The Japanese KC68/79 tends to overestimate ozone above 30 km. The data quality of sondes above 28 km is strongly influenced by the performance of the air sampling pump and its decaying efficiency at lower pressures. However, most of the intercomparison studies show that the performance of the ECC sondes between 28 and 35 km is still good and tends even to overestimate the ozone compared to lidar measurements. There are some experimental indications that for ECC-sondes there is a compensating effect due to evaporation of the KI-sensing solution which will cause an increase in concentration and may result in higher measured ozone. However, in general the sonde data above about 28 km are less reliable and should be used with caution for long-term trend evaluations, at least for the non-ECC-sonde types.

For the troposphere where the ozone concentrations are much smaller the results are not consistent at all. It is noted that validation of data from different types of sondes used in this study is only based on a very sparse number of experimental comparisons. Also, because ozone values are much lower in the troposphere than in the stratosphere, the impact of instrumental errors and instrumental variability is greater. Intercomparison campaigns done between 1970 and 1990 have shown systematic differences between sonde types typically varying from 10 to 15%. Recent campaigns (WMO-III, JOSIE, etc.) have shown that the Brewer Mast and KC79-sondes are less precise than the ECC-types and that the ECC-sondes are much more consistent than the other two types of sondes. The precision of the ECC-sonde is better than ± (5-10)% and shows a small positive bias of about 3%. The Brewer Mast and KC79 sonde showed precisions in the range of ± (10-20)% , but there are no indications of any bias larger than ± 5%.

There is some experimental evidence suggesting that the total ozone normalisation should not be used for ECC-sonde tropospheric ozone profile measurements while for the other sonde types it is less clear. Because this issue is still in an exploratory stage it is recommended that the normalisation be applied for any tropospheric data used in this trend assessment report.

Recent intercomparisons have shown that the observed differences between different sounding stations using the same sonde type are for a major part due to the differences in the preparation and correction procedures applied at the different launch sites. Fortunately some of the instrumental factors with the potential to influence the observed ozone trends regard post-flight data processing, and the data may be re-evaluated when the influences of these instrumental factors and their uncertainties are better understood. Although much progress have been made to improve the quality and homogeneity of the ozonesonde data since the last WMO Scientific Assessment of Stratospheric Ozone : 1994, there is still an urgent need to investigate and intercompare the instrumental performance of the different sonde types as well as to revise and achieve agreement on procedures for preparation and data processing.

2.6. SBUV and UMKEHR Analyses

2.6.1. Introduction

In section 2.4.3 the comparison of Umkehr data with SAGE II data was described. Here we compare data from 8 of the 13 stations used there with data from the Nimbus 7 SBUV instrument. SBUV (the solar backscatter ultraviolet instrument) operated from November of 1978 until June of 1990 when the instrument failed. Details of the SBUV calibration and algorithm are described in Chapter 1. The comparison of SBUV with Umkehr is of interest because the two techniques are very similar. The altitude in the atmosphere at which the scattering contribution function peaks depends on the product of the ozone cross section times the optical path. SBUV measures an ozone profile by scanning through wavelength (varying cross section) while Umkehr does so by measuring over a range of solar zenith angles (varying optical path). The information content and vertical resolution of the two techniques are similar. SBUV is considered to have good information in Umkehr layers 9 through 6 (and possibly 5) while Umkehr is considered to have good information in layers 8 through 5 (and possibly 4). SBUV is much less sensitive to volcanic aerosol effects than Umkehr because it views from above, and most of the aerosol is below most of the ozone.

The SBUV data used in this comparison were produced with the version 6 algorithm described in Chapter 1. Until 1983 SBUV was on a three-day-on one-day-off schedule because of spacecraft power limitations; after that it operated full time. In February 1987 the instrument began to experience a high rate of chopper synchronisation errors, which caused an apparently random error of ~3% in the radiance measurements. A correction was applied [Gleason *et al.*, 1995] and the post-1987 data do not appear to have a bias. Recent comparisons with NOAA-11 SBUV/2 data suggest increasing calibration error in the last year of data (the instrument failed in June of 1990). Since SBUV is a nadir-only measurement on a polar orbiting spacecraft, on a given day the SBUV orbital track could be as much as 13^{o} away in longitude from the Umkehr site. We use SBUV data interpolated to each Umkehr station using a distance-weighted average of observations within 2^{o} of latitude and 15^{o} of longitude.

The Umkehr data used were described in section 2.4.3. They were corrected for volcanic effects using the Mateer aerosol correction described in that section. Both Umkehr and SBUV processing used the Bass and Paur ozone cross section data and should be consistent in that respect. When both sunrise and sunset Umkehr observations are available, the two observations are averaged rather than being treated as independent observations. Data from 8 stations (Table 2-12) with what appear to be consistent data records were used in this comparison. Data for individual Umkehr layers 4 through 8, for combined layers 1+2+3+4 and for layers 8+9+10 combined are compared.

2.6.2. SBUV and Umkehr time series

To illustrate some of the limitations of analysing the SBUV versus Umkehr comparisons, the time series for one station, Boulder, Colorado, is shown in Figure 2-48. The matched time series of Umkehr observations (solid squares) and SBUV observations (open triangles) is shown in the upper plot, while the percent difference (SBUV-Umkehr)/Umkehr*100 is shown in the lower plot. Data for only one layer - Umkehr layer 8 (approximately 38 to 43 km) - are shown here. This station was chosen to illustrate several common problems. First, the Umkehr data record is usually not homogeneous, with few observations before about 1985. As another example, there were no data in the WOUDC data archive for Lauder, New Zealand until 1987. Second, the data records at some stations appear to have shifts. In 1979-1982 period the average Boulder difference from SBUV in Umkehr layer 8 averaged near zero,

while for the record in the post-El Chichon period (1984-1991) the difference averaged almost 10%, producing an apparent relative trend at this altitude of +6.6% per decade. If only data after 1984 are analysed, the trend would be near zero, but with a large uncertainty because of the shortness of the record. If such a shift relative to SBUV were seen at every station, this would be clear evidence that the calibration shift is in the SBUV instrument. Since the shifts occur at different times for different stations, they likely reflect changes in the ground-based instruments.

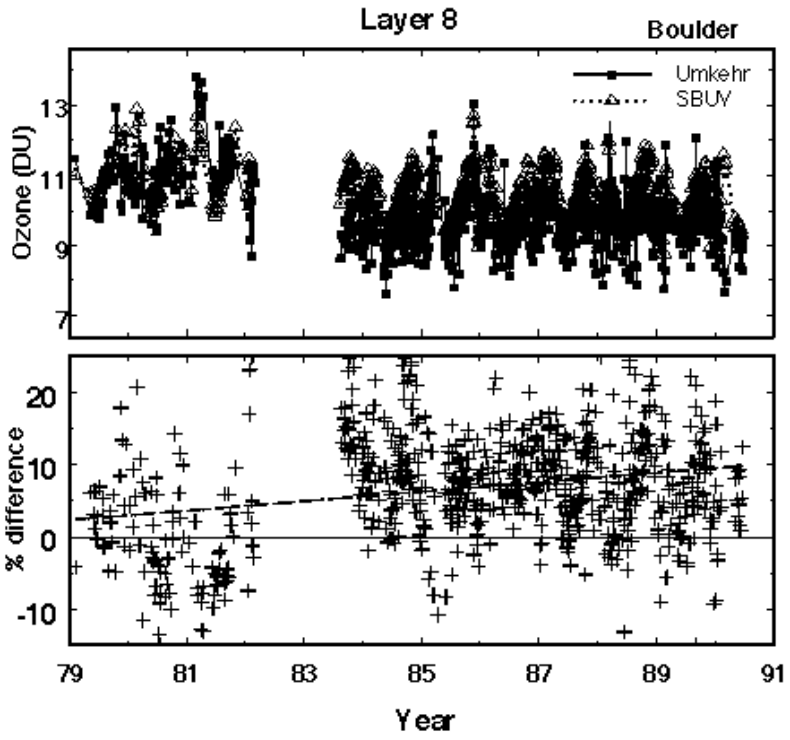


Figure 2-48 A sample comparison of one station - Boulder - for one layer in the atmosphere (Umkehr layer 8 covers approximately 38 km to 43 km). The ozone time series for Umkehr and SBUV are shown in the upper plot, while the percent difference (SBUV-Umkehr) is shown in the lower plot.

2.6.3. Regression Analysis Results

The regression analysis used here is quite simple - a linear fit is made to the time series of the SBUV-Umkehr difference. A difference plot removes most of the seasonal, QBO, and solar cycle variation, so a simple linear fit should be adequate. Based on previous experience [Hollandsworth *et al.*, 1995] a more sophisticated multiple regression fit does a better job of determining the uncertainties of the trend, but the trend itself can be obtained accurately from such a fit. The results of this analysis for each station and for each layer are given in Table 2-12. The percent bias for each layer is given in the upper half of the table, while the relative trend is given in the lower half.

	% bias (SBUV-Umkehr)/Umkehr						
Umkehr Layer	1-2-3-4	4	5	6	7	8	8-9-10
10 New Delhi	-0.8%	-7.1%	9.8%	-0.9%	8.7%	16.1%	7.6%

14 Tateno	-0.0%	-10.0%	9.3%	3.4%	5.4%	9.1%	1.9%
40 Haute Provence	-3.9%	-9.3%	11.8%	4.9%	-0.6%	4.1%	-2.5%
67 Boulder	2.8%	-8.3%	10.0%	3.0%	1.4%	6.4%	0.1%
68 Belsk	-2.0%	-0.9%	17.9%	8.4%	-0.7%	1.2%	-3.7%
152 Cairo	-3.7%	-11.5%	9.7%	3.7%	6.4%	10.2%	2.6%
159 Perth	2.1%	1.7%	-0.8%	-0.2%	10.0%	12.1%	3.9%
256 Lauder	10.4%	11.2%	-9.4%	-0.3%	3.7%	3.7%	-2.3%
all stations	0.5%	-5.5%	6.9%	2.6%	4.2%	7.7%	0.7%

	relative trend in % per decade (SBUV-Umkehr)/Umkehr						
10 New Delhi	1.4%	6.4%	-1.2%	-0.9%	4.8%	0.8%	-2.6%
14 Tateno	-0.0%	0.2%	1.1%	2.3%	4.0%	1.5%	-0.4%
40 Haute Provence	-3.5%	-1.1%	1.9%	-0.9%	-5.3%	-1.0%	2.3%
67 Boulder	-1.1%	0.5%	-2.4%	-3.6%	1.7%	6.6%	4.7%
68 Belsk	2.1%	-0.1%	-4.3%	0.5%	-2.7%	-10.0%	-8.3%
152 Cairo	-1.5%	0.6%	-1.6%	1.5%	-3.1%	-7.1%	-3.9%
159 Perth	-4.0%	-0.6%	0.9%	-8.8%	-14.1%	-3.6%	-1.7%
256 Lauder	15.5%	10.3%	-7.7%	-2.0%	0.3%	5.1%	9.9%
all stations	0.1%	5.4%	-6.8%	-2.2%	-0.6%	-1.1%	-1.2%

Table 2-12: Average bias and trend of SBUV relative to Umkehr

The left half of Figure 2-49 is a plot of the average of the station bias between SBUV and Umkehr for the 8 stations, for layers 4 through 8, and for combined layers 1-4 and 8-10. There is a tendency for SBUV to be higher than Umkehr by 5-7% in layers 5 through 8, lower by 5% in layer 4, and near agreement in the low combined layers 1-2-3-4, and in the high combined layers 8-9-10. The total column comparison shows SBUV about 3% higher than Umkehr.

The bias between SBUV and Umkehr is less interesting than the relative drift between the two measurement systems. One would like to see agreement between the systems since this would give confidence to the long term trends derived from each system. Unfortunately, there are significant differences. As shown in the right half of Figure 2-49, SBUV and Umkehr on average agree well, within 0.3% per year, in all layers, but with a fairly large variance among the stations. The Lauder time series is really too short to be compared independently and contributes much of the variance in the lower levels.

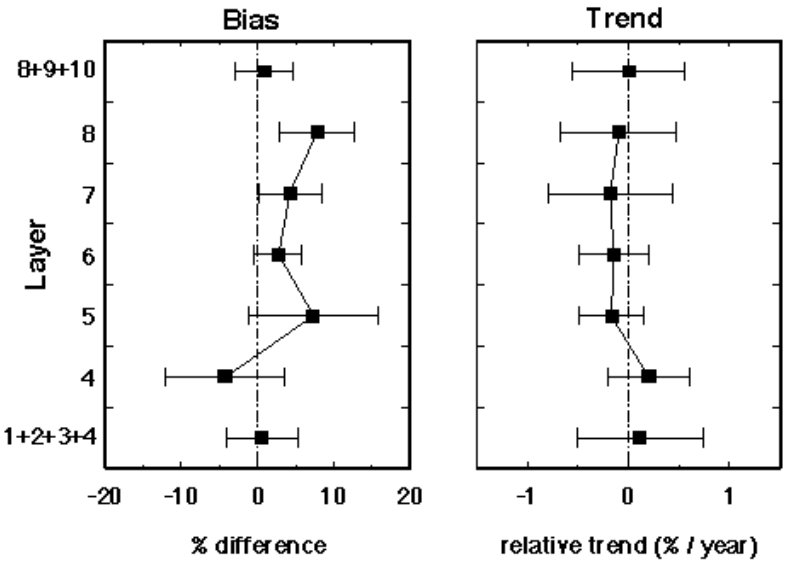


Figure 2-49 The altitude dependent bias (left plot) and trend (right plot) of SBUV relative to Umkehr. Note that combined layers are shown for Umkehr layers 1-4 and for layers 8-10.

2.6.4. Summary and Conclusions

While the SBUV/Umkehr comparisons show fairly good agreement, there are some problems with the comparison that should be kept in mind. First, it is important to screen the Umkehr stations to those that have stable instruments. Of the original 13 stations discussed in section 2.4.3, 5 stations had to be rejected because of obvious problems in the data record. There is also the problem that there was little Umkehr data in the early 1980's. One can argue that if station errors are random and enough stations are averaged, the result can be fairly stable. An optimistic conclusion would be that SBUV and Umkehr agree very well - to within 0.3%year^{-1}, but the station-to-station variance is 0.3-0.7% per year. Umkehr should be much more usable for trend evaluation in the period after 1985 when the number of observations is much greater and the instrument calibrations are presumably better as a result of the program to inter-compare Dobson and Brewer instruments.

2.7. Summary and Recommendations

2.7.1. Important Questions Addressed

Most of the questions raised in the introduction to this chapter have been addressed. The upper altitude limit of SAGE II data for use in determining trends is on the order of 50 km based on noise in SAGE II and intercomparisons with HALOE. The lower limit is less determined, most likely because of low altitude atmospheric variability and aerosol effects on both in SAGE II measurements and the data used for comparisons. In most instances, the differences in trends between SAGE II and correlative data start to increase below about 20 km and they become much more variable thus limiting the lower altitude at which derived trends can be validated. This is not to say that SAGE trends are invalid in this range, only that for the measurements systems used in the intercomparisons, a less definitive statement can be made about trend validity. The smallest long-term changes that can be verified over the 20 km to 50 km altitude interval are in the range of 0.3%year^{-1} when viewed as a function of latitude. Between 20 km and 40 km, the Dobson/Umkehr measurements constrain SAGE I/II drifts in a narrow latitude band in the northern mid-latitudes to 0.2±0.2%year^{-1} and in the Southern mid-latitudes to 0.3±0.3%year^{-1} at the 95% confidence level. Globally averaged SAGE I/II trends over this altitude range are constrained by both ground-based and satellite measurements to be valid to a level on the order of 0.2%year^{-1} at the 95% confidence level. It appears that the best agreement in trends occurs for SAGE II comparisons with other satellite data; although the Dobson/Umkehr comparisons provide equally good constraints in the northern and southern mid-latitudes. There is some evidence, even though it is statistically insignificant, that there is a slight drift in SAGE II of about 0.4%year^{-1} when compared to MLS and HALOE. The most important screening consideration is to eliminate SAGE II data in the lower stratosphere contaminated by Mt. Pinatubo aerosol absorption. Table 2-2 (see Section 2.3.1) provides a time and altitude screening guide. There are some obvious sunrise/sunset differences in the SAGE II data that are important only above approximately 45 km as discussed below. Finally, the question of SAGE I/SAGE II data offset was addressed and it was determined that the main offset is due to altitude registration between the two sets of measurements. The correction developed by Wang *et al*.(1996) is recommended as a satisfactory correction and was used in this study.

2.7.2. Conclusions and Recommendations

SAGE II absolute values agree with sondes in the altitude region between 20 km and 28 km to within a few % but below 20 km, SAGE II values start to increase relative to the sondes and reach values which are 15% to 20% larger at 15 km. The low altitude differences are latitudinally dependent which could indicate a problem with using the data for trend calculations.

The regression slopes for time series of coincident differences between SAGE I/II and other measurements for individual stations or latitudes is summarised quantitatively in Table 2-13 and graphically in Figure 2-50. Values range from £ 0.3± 0.15%year^{-1} to ~0.5± 0.7%year^{-1} (sondes, lidar, Umkehr, HALOE) for altitudes between 20 km and 35 km and £ 0.5± 0.5%year^{-1} to ~1 ± 1%year^{-1} for altitudes between 35 km and 50 km. Only two systems (sondes and lidar) provide useful trend comparison data for the altitude range between 15 km and 20 km. Trends of matched pair (i.e. sonde minus SAGE II) differences at individual stations show significant variability, ranging up to 3%year^{-1} at 15 km for Lauder; however, best agreement was obtained when the matched pair differences from the eight sonde stations used in Chapter 3 were combined into a single time series to calculate the regression slope of the differences. No statistically significant differences were obtained for the combined time

series, but the mean difference was ~0.25% ± 0.3-0.4%year^{-1} in this lower stratosphere range above ~15 km altitude. This suggests that there is a fair degree of noise in the differences at individual stations due either to atmospheric variability, sampling, or instrumental effects; but it also suggests that SAGE II trends in the lower stratosphere are accurate to the 0.25%year^{-1} level.

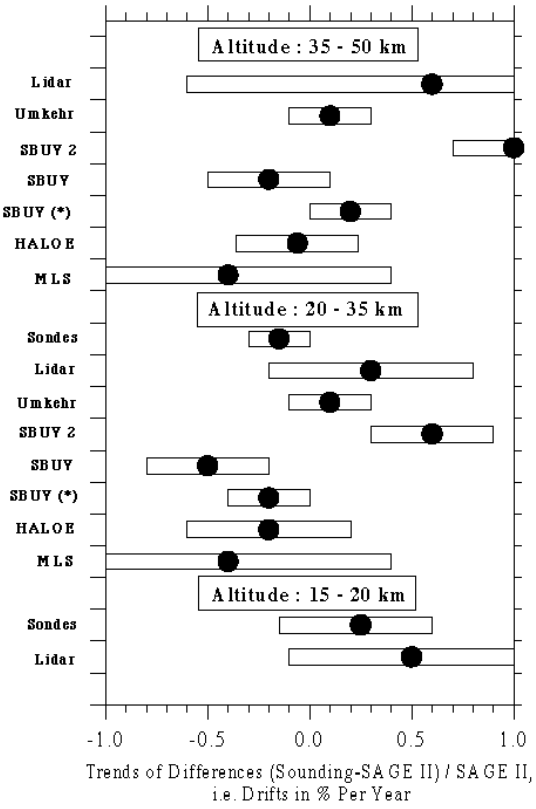


Figure 2.50 Trends of differences (i.e. drifts) between ozone measurements made by various ozone profiling instruments and SAGE II in % per year [(Sounding - SAGE II)/SAGE II]. Trends with ozonesondes are for the eight northern mid-latitude sounding stations used in Chapter 3 covering 36°N to 56°N. Umkehr differences are averages for eight Northern and Southern Hemisphere stations. Lidar differences are averages for the three stations with the longest records and number of coincidences exceeding 100 (Hohenpeissenberg, OHP and TMF). Trends of satellite differences (SBUV, SBUV2, HALOE and MLS) are presented as global means. The average differences are indicated by the dots and the bars represent the 95% confidence intervals of the drift estimations.

While no statistically significant drift was found between SAGE and the ozonesondes, differences in the absolute values of the measured ozone were found. The SAGE II absolute values agree with sondes in the altitude region between 20 km and 28 km to within a few %, but below 20 km SAGE II values start to increase relative to the sondes and reach values which are 15% to 20% larger at 15 km. The low altitude differences are latitudinally dependent which could indicate a problem with using the data for global trend calculations.

Method	15-20 km	> 20-35 km	> 35-50 km
Ozonesondes (1984-1996) for NML	0.7 to 3-4 0.25± 0.35(1)	± 0.7 for Z£ 28 km -0.15± .0.15	
Lidar (various time periods)	(1 to 2)± 1	0.1±0.4 to 0.5±0.7	(0.2 to 2)±0.7
Umkehr for NML (1984-1996)		-0.1±0.2 to 0.2±0.2	-0.1±0.2 to 0.3±0.2
Umkehr for NML (1984-1996) Column Amount		Layers 2 to 10 0.1±0.2	
SBUV 2 (1989-1994) *Global Mean*		1±1.5 *Layers 5-7 0.6 ±0.3*	1±1.5 *Layer 8, 9 1.0 ±0.3*
SBUV (1984-1989) *Global Mean*		(0.5 to 1) ±0.6 *Layers 5-7 -0.5 ±0.3*	(0.5 to 1.5) ±0.6 *Layer 8 ,9 -0.2 ±0.3*
SBUV - SAGE I&II (1979-1989)		-0.2±0.2 for Z>30 km	0.2±0.2
HALOE (1991-1996) *Global Mean*		(-0.4 to +0.4) ±0.5 , [except 15°N, and 55°S: ~(1-2)±0.5] *-0.2 ±0.4* *(-0.4 ± 0.3 @ 30 km)*	(-0.5 to + 0.5) ±0.5, [except 55°S: ~(1-2)±0.5] *-0.06 ±0.3*
MLS (1991-1996) *Global Mean* **HALOE** *(Global Mean using MLS-Transfer)*		*-0.4 ± 0.8* *-0.3 ± 0.8*	*-0.4 ± 0.8* *-0.4 ± 0.8*

(1) Below 16 km slopes of the differences become much larger and reach 0.85%year^{-1} at 15.5 km
Average of global means: Z>20-35 km: -0.3± 0.5%year^{-1}, Z>35-50 km: -0.2 ± 0.5%year^{-1}

Table 2-13: Trends of differences (Sounding-SAGE II)/SAGE II, i.e. possible drifts (%year^{-1}), Uncertainties are 95% confidence intervals

Examination of latitudinally averaged trends of differences (i.e. global trends) shows negative drifts of -0.06% to -0.4% ± ~0.6%year^{-1} for SBUV, HALOE, and MLS comparisons (SAGE II trend larger than other measurements). These differences, although statistically insignificant, give a slight indication of a SAGE II drift with time (SAGE II trend is larger). SBUV2 differences with SAGE II are of opposite sign to SBUV, HALOE and MLS, but this is most likely due to algorithm effects brought on by a drifting orbit as discussed in Chapter 1. Globally averaged analyses of the longest satellite time series - SBUV compared with the composite time series of SAGE I/II (1979 to 1989) - shows agreement to (-0.2% to 0.2%) ± 0.2%year^{-1} in the altitude region between 20 km and 50 km.

While some evidence exists to suggest that SAGE I and SAGE II overlapping measurements are inconsistent (e.g. Hohenpeissenberg, Payerne and Uccle comparisons), the results are not statistically significant. It is recalled that SAGE I data have been corrected for a systematic reference height error of approximately 300 m (latitude dependent). The uncertainty in this correction for each latitude is approximately 100m. Below 20 km altitude a simple upward shift of the SAGE I profiles (as has been assumed for this report) may be incorrect because of the large Rayleigh scattering contribution to the 0.6 nm extinction at these altitudes. An inversion of the SAGE I data to correct the altitude registration problem would be preferable. Data can be used for trends with caution below 20 km, but more intercomparisons are needed to draw firm conclusions. This is consistent with the Chapter 1 statement on this issue.

SAGE II data show a difference of ~10% between sunrise and sunset ozone values at 1 hPa (SS higher). This is unphysical and is much greater than the difference between HALOE sunrise/sunset values of approximately 2%. There is no reason to prefer SAGE II sunrise measurement over sunset (or vice versa). We recommend combining SR and SS for doing trend studies, but because the reason for the SR/SS differences is unknown, the error bars for derived trends will be increased. This effect is important only for altitudes above 45 km.

The results of the coordinate mapping (CM) study showed general agreement with the conventional analyses, which rely on measurement space and time coincidence. Near 25 km, both the conventional and CM technique shows statistically significant regression slopes of differences in the 40-50°S (1-1.5%year^{-1}) and 60-70°S (~0.5%year^{-1}) regions. Differences between the two satellite measurements also appear in the CM analysis near 40-50°N in the same altitude range with the CM technique indicating a marginally significant positive drift of SAGE relative to HALOE of ~1 ± 0.75%year^{-1}. At higher altitudes, CM results show insignificant drifts of less than 0.5%year^{-1} in all regions except the polar summer latitudes. Differences in this region, which cannot be directly compared to conventional results, indicate statistically significant trends of differences, with SAGE ozone trends being larger relative to HALOE.

Comparisons of ozonesondes in the stratosphere with other ozone profiling techniques show consistent results with agreement of about ± (3-5)% at altitudes between the tropopause and 28 km. The precision of the different sonde types is better than ± 3%. Above 28 km the results are not consistent due to instrumental uncertainties (e.g. pump corrections and sensing solution changes) and caution must be used, at least for the non-ECC types of sondes, when applying the data for long-term trend determinations.

A dearth of sonde validation studies exist for the troposphere and because of the small number of comparisons, only estimates about the reliability of the sonde data records below the tropopause can be made. In general, ECC-sondes provide much more consistent results than the other two types of sondes. The precision of the ECC-sonde is better than ± (5-10)% and shows a small positive bias of about 3%. Brewer Mast and KC79-sondes are less precise (± (10-20)%), but there are no indications of any bias larger than ± 5%. Key issues of uncertainty are the background correction and the use of the total ozone normalisation factor.

The main reasons for observed differences between different sonde results from sounding stations using the same type of ozone sonde are believed to be due to differences in the preparation and correction procedures applied at the different launch sites. Although much progress has been made to improve the quality and homogeneity of the ozonesonde data since the last WMO Scientific Assessment of Stratospheric Ozone in 1994, there is still an urgent need to investigate and intercompare the instrumental performance of the different sonde types as well as a need to revise and agree on procedures for preparation and data processing.

2.8. References

Aimedieu, P., W.A. Matthews, W. Attmannspacher, R. Hartmannsgruber, J. Cisneros, W. Komhyr, D.E. Robbins, Comparison of in-situ stratospheric ozone measurements obtained during the MAP/Globus 1983 campaign, Planet. Space Sci., 35, 563-585, 1987.

Ancellet, G., A. Papayannis, J. Pelon, and G. Mégie, DIAL Tropospheric Ozone Measurement using a ND:Yag Laser and the Raman shifting Technique, J. Atm. Ocean. Tech. 6, 832-839.

Ancellet, G., and M. Beekmann, 1997, Evidence for changes in the ozone concentration in the free troposphere over Southern France from 1976 - 1995, Atmos. Env. 31, 2835-2851.

Atkinson, R. J., and R. A. Plumb, Three-dimensional ozone transport during the ozone hole breakup in Dec. 1987, J. Geophys. Res., 102, 1451-1466, 1997.

Attmannspacher, W. and H. Dütsch, International Ozone Sonde Intercomparaison at the Observatory of Hohenpeissenberg, Berichte des Deutschen Wetterdienstes, 120, 1970.

Attmannspacher, W. and H. Dütsch, 2nd International Ozone Sonde Intercomparaison at the Observatory of Hohenpeissenberg, Berichte des Deutschen Wetterdienstes, 157, 1981.

Barnes, R.A., A.R. Bandy,, and A.L. Torres, Electrochemical concentration cell ozonesonde accuracy and precision, J. Geophys. Res., 90, 7881-7887, 1985.

Barnes, R.A., C.L. Parsons, and A.P. Grothouse, ROCOZ-A ozone measurements during Stratospheric Ozone Intercomparison Campaign (STOIC), J. Geophys. Res., 100, 9209-9224, 1995.

Bass, A. M., and R. J. Paur, The ultraviolet cross-section of ozone, I, Results and temperature dependence, In: Atmospheric Ozone proceedings of the Quadrennial Ozone Symposium, edited by C. S. Zerefos and A. Ghazi, 606-610, D. Reidel, Norwell, Mass., 1985.

Beekmann, M., G. Ancellet, G. Megie, H.G.J. Smit, and D. Kley, Intercomparison campaign for vertical ozone profiles including electrochemical sondes of ECC and Brewer-Mast type and a ground based UV-differential absorption lidar, J. Atmos. Chem., 19, 259-288, 1994.

Beekmann, M., G. Ancellet, D. Martin, C. Abonnel, G. Duverneuil, F. Eidelimen, P. Bessemoulin, N. Fritz and E. Gizard, Intercomparison of tropospheric ozone profiles obtained by electrochemical sondes, a ground based lidar and airborne UV-photometer, Atmos. Env., 29, 1027-1042, 1995.

Brinksma, E.J., D.P.J. Swart, J.B. Bergwerff, Y.J. Meijer, and F.T. Ormel., RIVM stratospheric ozone Lidar at NDSC-station Lauder: Routine measurements and validation during the OPAL-campaign, In: Advances in Atmospheric Remote Sensing with Lidar, Ansmann, Neuber, Rairoux and Wabndinger (Eds.), Springer Berlin, 529-532, 1996.

Butchart, N., and E. Remsberg, The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface, J. Atmos. Sci., 43, 1319-1339, 1986.

Claude, H., Hartmannsgruber R., and Köhler U., Measurements of atmospheric ozone profiles using the Brewer-Mast sonde. WMO Global Research and Monitoring Project- Report No. 17, 1-61, 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, 16757-16770, 1994.

Connor, B.J., A. Parrish, J.J. Tsou, and M.P. McCormick, Error analysis for the ground-based microwave measurements during STOIC, J. Geophys. Res. 100, 9283-9291, 1995.

Cunnold, D.M., H. Wang, W. Chu, and L. Froidevaux, Comparisons between SAGE II and MLS ozone measurements and aliasing of SAGE II ozone trends in the lower stratosphere, J. Geophys. Res., 101, 10,061-10,075, 1996-A.

Cunnold, D.M., L. Froidevaux, J. Russell, B Connor, and A. Roche, An overview of UARS ozone validation based primarily on intercomparisons among UARS and SAGE II measurements, J. Geophys. Res., 101, 10,335-10,350, 1996-B.

Cunnold, D.M. and H.J. Wang, The effects of temperature uncertainties on the interpretation of ozone trends in the upper stratosphere, Proceedings of the XVIII Quadrennial Ozone Symposium, Eds. R. Bojkov , and G. Visconti, L'Aquila, Italy, September 1996, 1998.

De Backer H., D. De Muer, E. Schoubs and M. Allaart, A new pump correction profile for Brewer-Mast ozonesondes, Proceedings of the XVIII Quadrennial Ozone Symposium, Eds. R. Bojkov , and G. Visconti, L'Aquila, Italy, September 1996, in 1998.

DeLuisi, J.J., Umkehr vertical ozone profile errors caused by the presence of stratospheric aerosols, J. Geophys. Res., 84, 1766-1770, 1979.

De Muer, D. and H. De Backer, Influence of sulfur dioxide trends on Dobson measurements and on electrochemical ozone soundings, in Atmospheric Ozone, Edited by T. Hendriksen, Proc. SPIE 2047, 18-26, 1993.

Froidevaux, L., W.G. Read, T.A. Lungu, R.E. Cofield, E.F. Fishbein, D.A. Flower, R.F. Jarnot, B.P. Ridenoure, Z. Shippony, J.W. Waters, J.J. Margitan, I.S. McDermid, R.A. Stachnik, G.E. Peckham, G. Braathen, T. Deshler, J. Fishman, D.J. Hofmann, and S.J. Oltmans, Validation of UARS Microwave Limb Sounder ozone measurements, J. Geophys. Res., 101, 10,017-10,060, 1996.

Froment, G., 1981, Surveillance de l'ozone stratosphérique, Note n°11 de l'etablissement des Etudes et de Recherches Météorologiques, ed. Meteo France , Toulouse , France.

Fujimoto, T. and O. Uchino, Estimation of the error caused by smoothing on DIAL measurements of stratospheric ozone, J. Meteor. Soc. Japan, 72, 605-611, 1994.

Fujimoto, T., O. Uchino and T. Nagai, Four year intercomparison of MRI ozone lidar, ozonesonde and SAGE II data, Proceedings of the XVIII Quadrennial Ozone Symposium, Eds. R. Bojkov, and G. Visconti, L'Aquila, Italy, September 1996, 1998.

Gleason, J. F. and R. D. McPeters, Corrections to the Nimbus 7 solar backscatter ultraviolet data in the "nonsync" period (February 1987 to June 1990), J. Geophys. Res., 100, 16873-16877, 1995.

Gunson, M.R., M.M. Abbas, M.C. Abrams, M. Allen, L.R. Brown, T.L. Brown, A.Y. Chang, A. Goldman, F.W. Irion, L.L. Lowes, E. Mahieu, G.L. Manney, H.A. Michelsen, M.J. Newchurch, C.P. Rinsland, R.J. Salawitch, G.P. Stiller, G.C. Toon, Y.L. Yung, and R. Zander, The Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment: Deployment on the ATLAS Space Shuttle missions, Geophys. Res. Lett., 23, 2333-2336, 1996.

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.

Harris, N. *et al*., Report of the SPARC/IOC-workshop at the Observatoire de Haute Provence (France) from July 8-11, SPARC Newsletter 7, 1996

Hilsenrath, E., W. Attmannspacher, A. Bass, W. Evans, R. Hagemeyer, R.A. Barnes, W. Komhyr, K. Mauersberger, J. Mentall, M. Proffitt, D. Robbins, S. Taylor, A. Torres, and E. Weinstock. Results from the balloon ozone intercomparison campaign (BOIC), J. Geophys. Res., 91, 13137-13152, 1986.

Johnson, B. J., S. J. Oltmans, D. J. Hofmann, J. A. Lathrop, Evaluation of ECC ozonesonde performance from recent field and laboratory intercomparisons, Proceedings of the XVIII Quadrennial Ozone Symposium, Eds. R. Bojkov, and G. Visconti, L'Aquila, Italy, September 1996, 1998.

Kerr, J.B., H. Fast, C.T. McElroy, S.J. Oltmans, J.A. Lathrop, E. Kyro, A. Paukkunen, H. Claude, U. Kohler, C.R. Sreedharan, T. Takao and Y Tsukagoshi, The 1991 WMO international ozonesonde intercomparison at Vanscoy, Canada, Atmosphere Ocean, 4, 685-716, 1994.

Komhyr, W. D., Electrochemical concentration cells for gas analysis, Ann. Geophys., 25, 203-210, 1969.

Komhyr, W.D. [1986] Operations Handbook-Ozone Measurements to 40-km Altitude with Model 4A Electrochemical Concentration Cell (ECC) Ozonesondes (Used with 1680-MHz Radiosondes). Technical Memorandum ERL ARL-149, National Oceanic and Atmospheric Administration, Boulder, Colorado.

Komhyr, W.D., R.A. Barnes, G.B. Brothers, J.A. Lathrop and D.P. Opperman, Electrochemical concentration cell ozonesonde performance evaluation during STOIC 1989, J. Geophys. Res., 100, 9231-9244, 1995-A.

Komhyr, W.D., B.J. Connor, I.S. McDermid, T.J. McGee, A.D. Parrish, J.J. Margitan, Comparison of STOIC 1989 ground-based lidar, microwave spectrometer, and Dobson spectrophotometer Umkehr ozone profiles with ozone profiles from balloon-borne electrochemical concentration cell ozonesondes, J. Geophys. Res., 100, 9273-9282, 1995-B.

Lait, L. R., M. R. Schoeberl, P. A. Newman, M. H. Proffitt, M. Loewenstein, J. R. Podolske, S. E. Strahan, K. R. Chan, B. Gary, J. J. Margitan, E. Browell, M. P. McCormick, and A. Torres, Reconstruction of O3 and N2O fields from ER-2, DC-8, and balloon observations, Geophys. Res. Lett., 17,521-17,524,1990.

Lary, D.J., M.P. Chipperfield, W.A. Norton, and L.P. Riishojgaard, Three-dimensional tracer initialization and general diagnostics using PV latitude-potential-temperature coordinates, Quart. J. Roy. Meteor. Soc., 121, 187-210, 1995.

Leovy, C. B., C.R. Sun, M.H. Hitchman, E.E. Remsberg, J.M. Russell III, L.L. Gordley, J.C. Gille, and L.V. Lyjak, Transport of ozone in the middle stratosphere: Evidence for planetary wave breaking, J. Atmos. Sci., 42, 230-244, 1985.

Manney, G.L., L. Froidevaux, J.W. Waters, and R.W. Zurek, Evolution of microwave limb sounder ozone and the polar vortex during winter, J. Geophys. Res., 100, 2953-2972, 1995.

Margitan, J.J., R.A. Barnes, G.B. Brothers, J. Butler, J. Burris, B.J. Connor, R.A. Ferrare, J.B. Kerr, W.D. Komhyr, M.P. McCormick, I.S. McDermid, C.T. McElroy, T.J. McGee, A.J. Miller, M. Owens, A.D. Parrish, C.L. Parsons, A.L. Torres, J.J. Tsou, T.D. Walsh, and D. Whiteman, Stratospheric Ozone Intercomparison Campaign (STOIC) 1989: Overview, J. Geophys. Res., 100, 9193-9208, 1995.

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

McDermid, I.S., S.M. Godin, and T.D. Walsh, Results from the Jet Propulsion Laboratory Stratospheric Ozone LIDAR during STOIC 1989, J. Geophys. Res., 100, 9263-9972, 1995.

McDermid, I.S., T.J. McGee, D.P.J. Swart, *et al*, NDSC Lidar Intercomparisons and validation: OPAL and MLO3 campaigns in 1995, In: Advances in Atmospheric Remote Sensing with Lidar, Ansmann, Neuber, Rairoux and Wabndinger (Eds.), Springer Berlin, 525-528, 1996.

McElroy, C.T., and J.B. Kerr, Table Mountain ozone intercomparison: Brewer ozone spectrophotometer Umkehr observations, J. Geophys. Res., 100, 9293-9300, 1995

McGee, T.J., R.A. Ferrare, D.N. Whitman, J.J. Butler, J.F. Burris, and M.A. Owens, LIDAR Measurements of Stratospheric Ozone during the STOIC Campaign, J. Geophys. Res., 100, 9255-9262, 1995.

McIntyre, M. E., Towards a Lagrangian-mean description of stratospheric circulations and chemical transports, Phil. Trans. R. Soc. Lond., 296, 129-148,1980.

Miller, A.J. *et al.,* Information content of Umkehr and SBUV(2) satellite ozone measurements for trends, J. Geophys. Res., in press, 1997.

Morris, G.A., A Demonstration and Evaluation of Trajectory Mapping, Ph.D. Thesis, Rice University, 1994.

Morris, G.A., M.R. Schoeberl, L.C. Sparling, P.A. Newman, L.R. Lait, L. Elson, J. Waters, R.A. Suttie, A. Roche, J. Kumer, and J.M. Russell, III, Trajectory mapping and applications to data from the Upper Atmosphere Research Satellite, J. Geophys. Res., 100, 16,491-16,505, 1995.

Morris, G.A., *et al*., Satellite data validation using trajectory mapping, Part I: Accuracy assessments, J. Geophys. Res., to be submitted, 1997.

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, 14029-14042, 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 submitted to J. Geophys. Res., 1997.

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, 1992.

Pierce, R. B., W. L. Grose, J. M. Russell III, and A. F. Tuck, 1994: Evolution of Southern Hemisphere Spring Air Masses Observed by HALOE, Geophys. Res. Lett., 21, 213-216, 1994-A.

Pierce, R. B., W. L. Grose, J. M. Russell III, A. F. Tuck, R. Swinbank, and A. O'Neill, 1994: Spring Dehydration in the Antarctic Vortex Observed by HALOE, J. Atm. Sci., 51, 2931-2941, 1994-B.

Randel, W.J. and F. Wu, TOMS total ozone trends in potential vorticity coordinates, Geophys. Res. Lett., 22, 683-686, 1995.

Redaelli, G., L. R. Lait, M. Schoeberl, P. A. Newman, G. Visconti, A. D'Altorio, F. Masci, V. Rizi., L. Froidevaux, J. W. Waters, and A. J. Miller, UARS MLS 03 soundings compared with lidar measurements using the conservative coordinates reconstruction technique, Geophys. Res. Lett., 21, 1535-1538, 1994.

Reid S.J., G. Vaughan, A.R. Marsh, and H.G.J. Smit, Intercomparison of ozone measurements by ECC sondes and BENDIX chemiluminescent analyser, J. Atm. Chem., 25, 215-226, 1996.

Russell, J. M. III, L. L. Gordley, J. H. Park, S. R. Drayson, D. H. Hesketh, R. J. Cicerone, A. F. Tuck, J. E. Frederick, J. E. Harries, and P. J. Crutzen: The Halogen Occultation Experiment, J. Geophys. Res., 98, 10,777-10,797, June 20, 1993.

Schenkel, A. and Broder B., Interference of some trace gases with ozone measurements by the KI-method, Atmos. Environ. 16, 2187-2190, 1982.

Schoeberl, M. R., and L. R. Lait, Conservative-coordinate transformations for atmospheric measurements, The use of EOS for studies of atmospheric physics, Proceedings of the International School of Physics Enrico Fermi, Italian Phys. Soc., 419--431, 1993.

Smit, H.G.J., Sträter, W., Kley, D., Proffitt, M.H., The evaluation of ECC-ozone sondes under quasi flight conditions in the Environmental simulation chamber at Jülich, A contribution to subproject TOR , In: Transport and Transformation of Pollutants in the Troposphere, Proceedings of EUROTRAC Symposium '94, Garmisch-Partenkirchen, 11-15 April 1994 , P. Borrell, P.M. Borrell and W. Seiler Editors, SPB Academic Publishing, The Hague, The Netherland, pp. 349-353, 1994

Smit H.G.J., W. Sträter, M. Helten, D. Kley, D. Ciupa, H.J. Claude, U. Köhler, B. Hoegger, G. Levrat, B. Johnson, S.J. Oltmans, J.B. Kerr, D.W. Tarasick, J. Davies, M. Shitamichi, S.K. Srivastav, C. Vialle, and G. Velghe, JOSIE: The 1996 WMO International intercomparison of ozonesondes under quasi flight conditions in the environmental simulation chamber at Jülich, Proceedings of the XVIII Quadrennial Ozone Symposium, Eds. R. Bojkov, and G. Visconti, L'Aquila, Italy, September 1996, 1998.

Steele, H.M. and R.P. Turco, The measurement and prediction of aerosol extinction: Implications for SAGE II ozone concentrations and trends, J. Geophys. Res., in press, 1997.

Steinbrecht W., R. Schwartz, and H. Claude, New pump correction for the Brewer Mast Sonde: Experimental determination and application to instrument intercomparisons, Submitted to: J. Atmos. & Ocean. Technol., 1997.

Swart, Daan P.J., Spakman, Jan, Bergwerff, Hans B., 1994: RIVM's Stratospheric Ozone Lidar for NDSC Station Lauder: System Description and First Results. Abstracts of Papers of the 17th International Laser Radar Conference, Sendai, Japan, 405-408

Thornton, D.C., and N. Niazy, Sources of background current in the ECC-ozone sonde: Implication for total ozone measurements, J. Geophys. Res., 87, 8943-8950, 1982.

Thouret, V., A. Marenco, P. Nedelec, C. Grouhel, and J.A. Logan, Comparisons of ozone vertical profiles from the MOZAIC airborne program and ozonesonde stations : A study at Hohenpeissenberg and Wallops Island station, Submitted to J. Geophys. Res., September 1997.

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

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

WMO Report of the International Ozone Trends Panel: 1988, World Meteorological Organization Global Ozone Research and Monitoring Project, Report No. 18, 1988

WMO Scientific Assessment of Stratospheric Ozone: 1994, World Meteorological Organization Global Ozone Research and Monitoring Project, Report No. 37, 1994

Chapter 3

Ozone Change as a Function of Altitude

Lead Authors

Richard Stolarski - William Randel

Co-Authors

Lane Bishop - Derek Cunnold - Dirk DeMuer

Sophie Godin - Dave Hofmann - Stacey Hollandsworth

Lon Hood - Jennifer Logan - Jim Miller - Mike Newchurch

Sam Oltmans - Osamu Uchino - Ray Wang

3.1. Introduction

The SPARC/IOC Ozone Trends Panel has been asked to provide an in-depth review of our knowledge of trends in the concentration profiles of ozone. This report will serve as an input to the next WMO/UNEP Scientific Assessment of Ozone Depletion. The present chapter assesses our understanding of the long-term ozone changes that have taken place since 1970 with an emphasis on the profile. As a result of the studies carried out under Chapter 2, data sets with improved long-term calibration and internal consistency are available for this analysis.

This report builds on the chapter on ozone trends of the 1994 WMO/UNEP report (WMO, 1995 and Harris *et al*., 1997). The emphasis of that report was trends in the total column amount of ozone, although it also contained significant material on trends in the altitude profile. This report will focus on trends in the profile. We will consider trends in the column amount only to check for consistency with column-integrated trends.

The 1994 report concluded that the trends in the ozone profile were generally consistent with the trends in total ozone, but that significant uncertainties remained. One of these was that the trends in the lower stratosphere deduced from SAGE I/II data were significantly larger than those observed by sondes. The introduction of a latitude-dependent correction for the altitude registration of the SAGE I data by Wang *et al*. (1996) has improved this situation. This correction has also made the upper stratospheric trends deduced from SAGE I/II and those deduced from the SBUV-SBUV/2 instruments more consistent.

Upper stratospheric ozone is only a small portion of the total column. However, trends in this region are important because they should reflect a straightforward response of ozone to the increases in chlorine amount. The chemistry of this region is gas phase and relatively well known, but confirmation through observational results is needed. This report uses SAGE I data as corrected by Wang *et al*. (1996). The analysis uses version 5.96 SAGE II data as described in Chapter 1. The SBUV/2 data from NOAA-11 used in these calculations is a reprocessed version using an updated calibration as described in Chapter 2.

Chapters 1 and 2 have included an extensive reanalysis of the data from a number of ozone sonde stations. These have included corrections for pump efficiency, solution strength, and other factors. These corrections potentially have a proportionally larger effect on the results of trend analyses in the troposphere. In the troposphere, the sondes remain the only viable measurements with long-term records. An important issue for trends in tropospheric ozone is whether trends have occurred throughout the period since 1970 or have been larger in the 1970s than in the 1980s and 1990s. To examine this question, trends from both 1970 and from 1980 are considered in this chapter.

In the present report we employ statistical analyses to test for significant occurrences of linear trends. This approximates the expected behaviour of ozone due to changes in the stratospheric chlorine amount. The time periods used are either 1970 to the present or 1979-80 to the present. Because of the implementation of the Montreal Protocol, chlorine amounts have reached their maximum in the troposphere and begun to decline. The amount of chlorine in the stratosphere should peak within a year or two and the ozone decline should then begin to reverse itself. This is the last report in which we will be able to address the primary question by searching for a linear trend.

3.2. Description of data sets

Chapter 1 of this report describes the instruments used for profile measurements in detail. The data from several instruments are inter-compared in Chapter 2. In this section we give a brief description of the data sets used, emphasising those issues most important to the trend analyses. The four data sets used are ozone sondes, Umkehr, SAGE, and SBUV. TOMS data is used to check for consistency of the altitude profile of trends with those in total column ozone.

3.2.1. Ozone Sondes

The SPARC/IOC committee re-evaluated the data at several stations as shown in Table 3.1. Data for the Canadian stations are those available from the World Ozone and Ultraviolet Radiation Data Center (WOUDC) in July 1997. This activity did not include re-evaluation of the Canadian data. There are concerns about the quality of the Brewer Mast data for these stations, as discussed in Chapter 1. Most of the sonde data used here for trend analysis are scaled to ozone column measurements on the Bass-Paur scale (Bass and Paur, 1985; Paur and Bass, 1985) by use of a correction factor. The Canadian Brewer-Mast sonde data are multiplied by 0.9743 to put them on the Bass-Paur scale. The Lauder data has no correction factor applied for total ozone. The data for Wallops Island from the WOUDC were reprocessed to scale each profile to re-evaluated ozone column data (Oltmans *et al.*, 1997). The data for Payerne are undergoing further analysis, and an interim data set was provided for the analyses here.

a) Data re-evaluated for this report

No.	Station	Lat.	Long.	Type	Period
53	Uccle	51	4	BM	1/69 - 12/96
99	Hohenpeissenberg	48	11	BM	11/66 - 12/96
156	Payerne	47	7	BM	11/66 - 12/96
12	Sapporo	43	141	KC	12/68 - 12/96
67	Boulder	40	-105	ECC	3/79 - 12/96
14	Tateno	36	140	KC	11/68 - 12/96
7	Kagoshima	32	131	KC	1/69 - 12/96
109	Hilo	20	-155	ECC	9/82 - 12/96
256	Lauder	-45	170	ECC	8/86 - 12/96

b) Data available at the WOUDC

No.	Station	Lat.	Long.	Type	Period
24	Resolute	75	-95	BM	1/66 - 11/79
				ECC	12/79 - 2/96
77	Churchill	59	-94	BM	10/73 - 8/79

				ECC	9/79 - 12/96
21	Edmonton	53	-114	BM	10/72 - 8/79
				ECC	9/79 - 12/96
76	Goose Bay	53	-60	BM	6/69 - 8/80
				ECC	9/80 - 12/96
107	Wallops Is.	38	-76	ECC	5/70 - 5/95

Table 3.1. Sonde stations, sonde types and time period of measurement. (BM = Brewer Mast, KC = Carbon-iodine Cell, ECC = Electrochemical concentration cell)

The sonde data were provided in a variety of formats. To facilitate trend analysis, these were processed into a common format. This format gives the column of ozone in Dobson units for 33 equally spaced layers of log pressure ranging from 1000 to 6.3 hPa (30 layers up to 10 hPa). The vertical resolution, ~1 km, was chosen to be similar to that of SAGE data. The WOUDC data and the data reanalysed for this report were put into this format.

Two groups conducted trend analyses of the sonde data for this report, Logan and Megretskaia from Harvard University, and Tiao, Choi, and Zhang from the University of Chicago. These groups treated the sonde data differently using selection criteria which depended on the correction factor, CF (the factor scaling the sonde data to the total ozone column). The Harvard group used the same criteria for the CFs as Logan (1994), 0.9-1.35 for BM sondes except for Hohenpeissenberg, where 0.9-1.2 was used, and 0.8-1.2 for ECC and KC sondes. They analysed the trends for 33 layers and for 11 layers obtained by summing 3 consecutive layers; only the latter results are shown here. The Chicago group used the CF criteria of 0.9-1.2 for BM and 0.9-1.15 for ECC and KC sondes as in their earlier work (Miller *et al.*, 1995). They required a total ozone column reading for the day of the sounding, and a balloon burst occurring above 16 hPa. They aggregate the 33 layers into 15 Umkehr layers that they have used in their previous analyses. Furthermore they removed the scaling to the total ozone column from the sonde data by dividing each profile by the correction factor. If there is a trend in the CF this can result in different trends being derived for the ozone profile (e.g., Logan, 1985, 1994; Miller *et al.*, 1995). The fractions of soundings that met various criteria for each method (BM and ECC) are given in Table 3.2.

The stricter requirements used by the Chicago group result in a significantly larger fraction of the soundings being omitted, particularly for the BM stations, as shown in Table 3.2. The second column gives the fraction of soundings retained in the Harvard analysis, and the last column the fraction retained in the Chicago analysis. Hohenpeissenberg is the exception, since both groups use the same CF criteria, 0.9-1.2. In the worst cases, the Canadian BM data and Payerne, only 25-44% of the soundings meet the Chicago group's criteria compared to 70-90% that meet the Harvard group's criteria. For all of the BM stations, a larger fraction of soundings are omitted from the earlier part of the record than from the later part. This results from the downward trend in the correction factors (see Chapter 1). For the Canadian ECC soundings, Wallops Island, and the Japanese stations, the Chicago group retain about 45-65% of the soundings, while the Harvard group retain about 85-95%. Trends in the correction factor using the Chicago data selection criteria are given in Table 3.3. There are small but significant trends in the CF at several stations, mostly in the range -1 to -4%/decade. Trends were computed for 1970 (or the beginning of the record if after 1970) to 1996 by both groups and for 1980-96 by the Harvard group.

There is a significant difference in tropospheric ozone values measured by BM and ECC sondes that must be accounted for when deriving trends for the Canadian stations. Intercomparisons in the 1970s and early 1980s showed that ECC sondes measured more ozone in the troposphere than BM sondes by about 15 to 20% (with a range of 7-38%) (Logan, 1994), as discussed in Chapter 1. Tiao *et al*. (1986) used an intervention term in their statistical trend model at the time when BM sondes were replaced with ECC sondes, and this approach was adopted here by both groups. The magnitude of the intervention term is similar to the differences between EEC and BM sondes found in the intercomparisons, and varies among stations (Tiao *et al*., 1986). For trends starting in 1980 (September 1980 for Goose Bay), the Canadian data are obtained exclusively with ECC sondes, so no intervention term is necessary.

a) Brewer Mast data

Station	Years	0.9-1.35	0.9-1.35	0.9-1.2	0.9-1.2	0.9-1.2
			16 hPa		16 hPa	16 hPa
						Tot O_3
Uccle	70-96	0.87	0.82	0.58	0.54	0.54
Hohenpeissenberg	70-96	-	-	0.93	0.91	0.91
Payerne	70-96	0.83	0.76	0.47	0.44	0.44
Resolute	70-79	0.91	0.51	0.75	0.39	-
Churchill	73-79	0.69	0.41	0.44	0.25	0.25
Edmonton	72-79	0.78	0.52	0.53	0.34	0.34
Goose Bay	70-80	0.81	0.58	0.42	0.32	0.32

b) ECC type data

Station	Years	0.8-1.2	0.8-1.2	0.9-1.15	0.9-1.15	0.9-1.15
			16 hPa		16 hPa	16 hPa
						Tot O_3
Resolute	80-96	0.96	0.65	0.90	0.59	-
Churchill	80-96	0.97	0.75	0.89	0.67	0.31
Edmonton	80-96	0.95	0.78	0.84	0.68	0.63

Goose Bay	80-96	0.95	0.70	0.87	0.63	0.50
Sapporo	70-96	0.88	0.80	0.67	0.60	0.60
Tateno	70-96	0.90	0.81	0.65	0.59	0.59
Kagoshima	70-96	0.85	0.76	0.53	0.46	0.46
Wallops	70-96	0.98	0.86	0.88	0.77	0.66
Boulder	80-96	0.94	0.91	0.91	0.88	0.88
Hilo	82-96	0.89	0.87	0.87	0.85	0.85
Lauder	86-96	0.92	0.82	0.92	0.84	0.84

Table 3.2. Fraction of soundings that met various criteria. The second column of the table gives the fraction of soundings that met the CF criteria required for the Harvard analysis (except for Hohenpeissenberg); the third column gives the fraction of the soundings that also reach 16 hPa. The fourth column gives the fraction that meet the Chicago CF criteria; the fifth column gives the fraction of these soundings that also reach 16 hPa; and the sixth column gives the fraction for which there is also a measurement of the total ozone column amount. For Boulder, Hilo, and the Japanese stations, no CF is given if there is no total ozone column measurements so the soundings fail the CF criteria. For the BM Canadian stations, a default CF is given which fails the Chicago CF criteria except for Resolute. For the ECC soundings, the default CF is 1.0, which meets the CF criteria; Resolute is not required to have a total ozone column measurement in winter.

The models used for trend analysis are described in the Appendix (section 3.9). The Harvard and Chicago models are similar in that they fit 12 monthly means and 4 seasonal trends, and allow for the dependence of ozone on the QBO and solar flux. The major differences between the models are the inclusion of autoregression, the assumption of zero trend before 1970 in the Chicago model and the fact that the Chicago group remove outliers from the model fit. Eight of the stations have data before 1970 (Table 3.1). Vertical profiles of trends are best compared as relative trends (e.g., percent per decade), since each group used different layers; unfortunately percentage trends were computed relative to a different reference. The Harvard trends are given relative to the mean of the time series for which the trend is calculated. The Chicago trends relative to the seasonal intercept in 1970 (or beginning of series if later) adjusted for solar effects and intervention if used. Column trends in the troposphere (1000-250 hPa) and stratosphere (250-16 hPa) are shown, so that the results can be compared in the same units (DU) for the same columns.

Station	**Years**	**Trend (%/decade)**	**Years**	**Trend (%/decade)**
Uccle	70-96	-2.6±0.6	80-96	-3.7±1.3
Payerne	70-96	-1.9±0.5	80-96	-3.0±0.8
Hohenpeissenberg	70-96	-1.3±0.5	80-96	Not Significant
Sapporo	70-96	Not Significant	80-96	Not Significant

Tateno	70-96	-1.3±0.9	80-96	-3.3±1.7
Kagoshima	70-96	Not Significant	80-96	Not Significant
Wallops Is.	70-96	-2.0±1.1	80-96	-2.5±1.8
Boulder			80-96	-3.7±1.2
Hilo			82-96	-3.7±1.3
Lauder			86-96	-1.5±1.5
Resolute	70-79	-2.3±2.5	80-96	Not Significant
Churchill	70-79	Not Significant	80-96	-2.3±2.4
Edmonton	70-79	5.1±6.0	80-96	-2.7±1.4
Goose Bay	70-79	Not Significant	80-96	2.0±1.7

Table 3.3. Trend in correction factor (%/decade). The trend in the correction factor was calculated using a least squares fit to monthly mean values; two standard errors are given. Most of the trends listed as not significant are smaller than 1%/decade; values are given for trends that are significant or are close to significant. The Chicago data selection criteria were used, i.e., the sonde reached 16 hPa, there was an ozone column measurement (except for Resolute), and the CF was within the range 0.9-1.2 (BM) and 0.9-1.15 (ECC). Trends are given separately for the two types of sonde for the Canadian stations, and for the two analysis periods at the other stations.

3.2.2. Umkehr

The Dobson/Umkehr data used in this study resulted from a concerted effort by the World Ozone and Ultraviolet Radiation Data Center (WOUDC) to insure that all Umkehr data were in the database, were on the Bass-Paur scale, and that the most current calibrations were used. A significant number of previously unavailable Dobson/Umkehr observations entered the database as a result of this effort and were made available for this report. Because of limited availability, no Brewer/Umkehr data were used in this analysis. However, Umkehr records from the Brewer network would provide a significant enhancement to the Dobson/Umkehr records and are highly desirable for future trend analyses.

The Umkehr data used in this analysis were current in the WOUDC in October 1997 where they were inverted from N-values using the Umkehr[92] inversion algorithm (Mateer and DeLuisi, 1992) employing the uniform S_x error covariance matrix. Previously, some Umkehr records had been inverted at individual stations (not at WOUDC) introducing some uncertainty in the uniformity of the inversion algorithm. All the Umkehr records in this report resulted from a single inversion algorithm with uniform characteristics.

The Chapter 2 regression analysis of the time series of SAGE-Umkehr differences for 15 stations indicated that not all of those records were reliable for ozone trend estimation. Difficulties with record length, missing data periods, and level offsets reduced the number of stations used for trend analysis to two Southern Hemisphere stations, Perth and Lauder; and six northern hemisphere stations, New Delhi, Cairo, Tateno, Boulder, Belsk, and Haute Provence. The Arosa data were not available from WOUDC at the time of this study and, therefore, were not included in the Chapter 2 analysis. However, one of the research groups doing trend analyses for this Chapter used data obtained directly from the Arosa station for analysis.

In concert with Chapter 2 recommendations, we analysed Umkehr trends in layers 1+2+3+4 (termed layer 4^-), individual layers 4 through 8, and layers 8+9+10 (termed layer 8^+). We also report trends in total ozone for the selected stations. All analysed Dobson/Umkehr records were corrected for aerosol effects either before the trend analysis using the correction factors of Mateer and DeLuisi (1992) or with an explanatory variable in the statistical trend model. In both approaches, periods of very high stratospheric aerosol loading (~1 year) following the eruptions of El Chichon (April 1982) and Mt Pinatubo (June 1991) were omitted. While several Dobson/Umkehr stations report observations prior to 1977, all analyses in this chapter consider either the period 1979-1996 (Newchurch and Yang) or 1977-1996 (Reinsel) for trend evaluation.

3.2.3. SAGE

SAGE I ozone measurements were almost all made at sunset. Stratospheric aerosol loading in 1979-1981 was exceptionally low and therefore the impact of aerosols on the ozone measurements is expected to have been small or non-existent. An important uncertainty in the SAGE I measurement concerns the reference altitude for each profile. The ad hoc correction for this error (Wang *et al.*, 1996) is an attempt to remove this systematic error source above an altitude of approximately 20 km. Ozone trends will be reported separately for the SAGE I/II time series and for the SAGE II time series alone.

The SAGE II data used in this report are from retrieval version 5.96 that was released in February 1997. There was no comparable recent re-retrieval of the SAGE I data. SAGE II v5.96 possesses a non-physical separation of sunrise and sunset ozone values of approximately 10% above 45 km altitude. This separation is larger than in v5.93. Chapter 2 studies indicate that this separation is relatively systematic and should not affect the ozone trend results. Removal of aerosol influences from the ozone retrievals has been improved in v5.96 but SAGE ozone values below an altitude of approximately 22 km were still contaminated for approximately 2 years after the Pinatubo eruption. The SAGE II error bars reflect this contamination and only SAGE data with less than 12% error bars are used in the trend analyses. At higher altitudes, the SAGE II data set is more complete and extends from October 1984 to December 1996. All SAGE ozone trends are reported on altitude levels (see Chapter 2).

3.2.4. TOMS

The TOMS total ozone record from the Nimbus 7 instrument extended from November 1979 until May 1993. The Meteor 3 instrument was launched in 1991 and lasted until the end of 1994. These data were combined into a single time series following McPeters *et al.*, 1996. There is then a gap in the TOMS record until the Earth Probe (EP) instrument that was launched in August, 1996. A final calibration has yet to be established for the EP TOMS data. Therefore total ozone trends are calculated from November 1978 through October 1994.

3.2.5. SBUV

The Version 6.0 NIMBUS-7/SBUV data and the Version 6.1.2 NOAA-11/SBUV2 data are used in this report. Both data sets are processed with the Version 6 BUV algorithm described in Chapter 1 and Bhartia *et al.* (1996). The instruments make a series of measurements approximately every 2 degrees of latitude over the sunlit portions of 14 orbits separated by 27 degrees of longitude. These are used in a

profile retrieval algorithm and reported as ozone profiles in DU for 12 Umkehr layers (~5 km thick each) although, as noted in Chapters 1 and 2, the actual vertical resolution is poorer, particularly in the lower stratosphere.

The SBUV data are available for November 1978 to May 1990. The current Nimbus 7 SBUV data include an updated calibration to correct for time-dependent instrument changes. Bhartia *et al*. (1996) have estimated the drift errors in the SBUV profile to be ± 5%/decade at 1 hPa and ± 2%/decade at 10 hPa. The data after February 1987 are affected by an out-of-synchronisation condition and have been corrected by a "scene-stabilisation" method described in Gleason and McPeters (1995). As recommended in Chapter 2, SBUV data for 1990 are not used in either the SBUV to SBUV2 adjustment or the trend calculations.

The SBUV2 data are available for January 1989 to October 1994 with decreasing coverage of the southern hemisphere over the instrument lifetime due to precession of the NOAA-11 equator crossing times. The changing orbit results in a loss of coverage as far north as 20° S in July of 1994. The NOAA-11 SBUV2 data have recently been reprocessed using updated calibrations and instrument behaviour characterisations and an algorithm change to correct for grating position errors in the latter part of the record to produce the Version 6.1.2 data set. The NOAA 11 absolute calibration was adjusted to match the Shuttle SBUV (SSBUV) calibration (Hilsenrath *et al*., 1995), while the time-dependent calibration was maintained using an on-board calibration lamp system and verified though comparisons with SSBUV measurements (Ahmad *et al*., 1994). However, known errors in the data set remain. The Version 6.1.2 algorithm is summarised in Chapters 1 and 2.

The NIMBUS 7 version 6 SBUV data through December 1989 and the NOAA-11 version 6.1.2 SBUV2 data from January 1989 through October 1994 are used to create a combined data set. The 1990 N7 SBUV data are disregarded. The SBUV and SBUV2 data are joined by adjusting the SBUV data by the average difference between the SBUV and SBUV2 data over 1989. These empirical adjustment factors are functions of both latitude and layer and range from -15% to 12% for the layer from 32 hPa to the ground (layers 4-0) with smaller adjustments in the upper layers. The adjustments in layer 8 (2-4 hPa), for example, range from 0.6 to 3.8%. The adjusted SBUV and SBUV2 data are then averaged during the overlap period to create a consistent time series. The SBUV2 daily average, 5° zonal means are filtered by average solar zenith angle and average latitude to eliminate data taken at extreme angles as a result of the NOAA-11 drifting orbit. On any given day, if the average solar zenith angle for a latitude zone is greater than 80° or the average latitude of the measurements in the bin is greater than 1° off the centre latitude of the bin, the daily average value for that zone is ignored. A 5° zonal average requires ~70% of the measurements to be present, and both 5° averages are required to create a daily 10° zonal mean. Ten daily, zonal average values are required to create the monthly average value for that bin. These filters are consistent with Chapter 2 recommendations. Only "Error Code 0" data (Fleig *et al*., 1990) have been used. These are data from the ascending part of the orbit for which no error flags occurred. The data have also been screened for measurement contamination from volcanic aerosols (Torres and Bhartia, 1995) by deleting all data equatorward of 40^{o} for one year following the April 1982 El Chichon and June 1991 Mt Pinatubo eruptions.

3.3. Comparison of Statistical Models

3.3.1. Objectives

A wide variety of statistical models have been used to derive trends in stratospheric ozone and to determine the effects on ozone of other variables such as the solar cycle and the quasi-biennial oscillation (QBO). The 1989 Scientific Assessment of Ozone Depletion (WMO, 1990, Chapter 2) contained brief intercomparison results of a few of these statistical models. It was found that variations in the statistical model or in the ancillary variables used (solar, QBO, nuclear effect) had relatively minor effects on the calculated ozone trend, at least for total ozone. However, when different research groups derive different trends, questions continue to arise as to how much of the difference in the trends or standard errors is due to differences in data used, and how much is due to the differences in the statistical model calculations.

The purpose of this section is to compare the statistical trend calculations of participating research groups on common sets of ozone data in order to determine the effect on ozone trends of statistical model differences.

The following sections describe the three test data series used for the intercomparisons; illustrate the statistical model issues using the example of one particular statistical model; and discuss the differences in results of the contributing participants.

3.3.2. Sample Data Sets

Three sample ozone data sets were chosen to represent some of the issues involved in fitting a statistical model. The data are plotted in Figure 3.1 a,b,c.

- Test data 1: Nimbus 7 TOMS (version 7) 40-50°N monthly zonal mean from 11/1978 through 04/1993 (contributed by Stacey Hollandsworth).
- Test data 2: Uccle ozone sondes monthly mean at 13 km from 01/1969-12/1996 (contributed by Dirk DeMuer).
- Test data 3: SAGE I and SAGE II monthly mean at 40.5 km, 35-45°N from 02/1979-12/1996. These SAGE I data are not altitude-corrected. (contributed by Ray Wang).

Test data 1 (TOMS) is a "clean" time series, with no missing values. Test data 2 (Uccle sondes) has stretches of missing values, including one of fifteen months in the middle of the series. Test data 3 (SAGE) has serious missing-data issues, especially for the early years of the series, including a break of almost three years between the SAGE I and SAGE II data; also there are no June values until 1990, causing additional issues for monthly trend models.

Ancillary time series of related variables are also frequently included in trend models, either from interest in the effect on ozone of the ancillary variables themselves, or in hopes of improving the ozone trend estimate by removing the effect of the other variables. The two effects most commonly included are the 11-year solar cycle and the quasi-biennial oscillation (QBO) in equatorial stratospheric winds. The latter is frequently lagged by several months with lag depending on latitude and altitude. Different researchers use different proxies for these effects. For solar effects, the 10.7 cm (2800 MHz) solar radio flux or sunspot number are typical choices, while 50 hPa or 30 hPa equatorial winds are frequently chosen for the QBO. For reference and verification purposes, the 10.7 cm flux and a QBO series (average 50 hPa wind at three stations) are available.

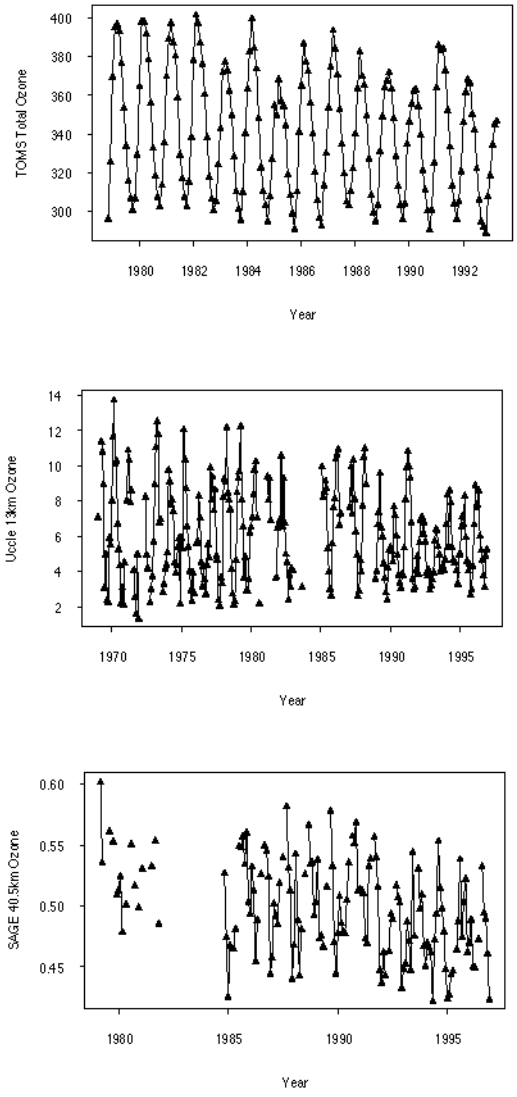


Figure 3.1 a) Test Data 1 - TOMS monthly average total ozone 11/78-4/93, 40°N-50°N, in Dobson Units. b) Test Data 2 - Uccle 13 km sonde monthly average ozone 01/69- 12/96, in mPa. c) Test Data 3 - SAGE I/II 40.5 km monthly average ozone, 03/79 -12/96, 35°N-45°N, in 10^{12} cm^{-3}.

3.3.3. Example Statistical Model

This section summarises one form of statistical model as a context for discussion of the statistical issues in the intercomparison of models. For some additional discussion of the terms and statistical issues, see Bojkov *et al*. (1990).

Let y_t represent monthly ozone values for one of the test series; in some cases y_t is missing for some months, and this is addressed in the notes below. The statistical model for y_t is of the form:

$$y_t = \text{(Monthly mean)} + \text{(Monthly trend)} + \text{(Solar effect)} + \text{(QBO effect)} + \text{Noise}$$

or more precisely,

$$y_t = \sum_{i=1}^{12} \mu I_{i,t} + \sum_{i=1}^{12} \beta I_{i,t} R_t + \gamma_1 Z_{1,t} + \gamma_2 Z_{2,t} + N_t$$

with the definitions:

m_i	Ozone mean in month i, i = 1…12
$I_{i;t}$	Indicator series for month *i* of the year; i.e., 1 if the month corresponds to month *i* of the year, and 0 otherwise
b_i	Trend in Dobson units/year in month *i* of the year
R_t	Linear ramp function measuring years from the first month of the series; equal to $(t-t_o)/12$. For series beginning before 1970, it is often taken to be a ramp function equal to zero for $t < t_o$, where t_o corresponds to 12/69, and then $(t-t_o)/12$ for t ? t_o.
$Z_{1;t}$	Solar 10.7 cm flux series, with γ_1 the associated coefficient
$Z_{2;t}$	QBO series lagged some appropriate number of months (latitude and altitude dependent), with γ_2 the associated coefficient
N_t	Residual noise series

Although this is the underlying model used by most researchers, a variety of statistical model issues are handled differently by different researchers, or even by the same researcher depending upon data features, for example if the proportion of missing data is very high. The following notes illustrate some of the frequent differences, but are not intended to be exhaustive, especially on minor issues.

3.3.4. Seasonal issues

Ozone trends are sometimes reported (usually graphically) on a monthly basis, but it has become common to tabulate the trends on an aggregated seasonal basis. Common seasons are Dec...Feb, Mar...May, Jun...Aug, Sep...Nov plus year round, but others have been used (for example Dec...Mar, Apr, May...Jul, Sep...Nov). To calculate seasonal trends from a monthly model, one averages the monthly trends over the months in each season. If percentage trends are desired, some groups calculate these from the average ozone trends in each season divided by the average ozone intercept (m_i) for the months in that season; other groups divide by the mean value for that season over the period of the time series (also, occasionally averages of monthly percentage trends have been calculated).

Proper calculation of the standard errors of seasonal trends when a monthly model is used requires use of the covariance matrix of the trend estimates. This is particularly important when an autocorrelated noise model is used, and if disregarded, may lead to serious underestimates of the standard errors of seasonal trends because the average of, for example, Dec-Jan-Feb trend estimates, does not represent the average of three independent quantities.

Sometimes the seasonal model is fit with only four seasons by explicitly working with quarterly data

(using Q1=Dec...Feb, etc.). A variation on this when data is sparse is to fit 12 monthly means, but only quarterly trends. In this case the standard errors are handled automatically, except for the year round trend. This approach has been used for deriving trends from sonde data (Miller *et al.*, 1995).

Sometimes too, the seasonal mean and/or the seasonal trends will be treated as the sum of sine and cosine harmonics, with the higher order terms dropped if not statistically significant. This is almost necessary if one models daily or weekly data, and has also been used to reduce the number of terms in the model when the solar or QBO terms are treated on a seasonal basis.

Seasonal trends can also be calculated by fitting multiple regression models separately for each month or for each season, using time and ancillary variables as predictors. This automatically handles seasonal weighting, but allows the solar and QBO coefficients to be different in each month or season. If done by season (quarterly ozone values), this approach automatically handles the autocorrelation issue for the standard error for the seasonal trend. It does not provide the necessary covariance matrix for the standard error of a year round trend. It may cause problems when data is consistently missing for a specific month.

3.3.5. Autocorrelation

For many ozone series, especially total, the noise term N_t is autocorrelated, i.e. N_t plotted against N_{t-1} will show a substantial correlation. This is due to dynamic or other effects not explicitly accounted for in the statistical model. This has usually been accounted for by allowing N_t to be an autoregressive series of order 1, i.e., AR(1) model, (although occasionally higher order autoregressive series have been used):

$$N_t = fN_{t-1} + e_t$$

where e_t is an uncorrelated series.

Properly accounting for an autocorrelated noise series normally changes the trend estimates by little, but leads to larger standard errors of the trend estimates than ordinary least squares. This is particularly important for seasonal and year round trend estimates produced by averaging monthly estimates, as discussed above.

Computations for fitting regression models with autocorrelated errors usually follow a conditional least squares approach (conditional on the first observation in the case of an AR(1) model) or, preferably, unconditional least squares or maximum likelihood (see, for example, Box *et al.*, 1994). These calculations are programmed into some statistical packages, although not all can well handle internal missing data values. Some researchers use their own FORTRAN routines to perform maximum likelihood calculations.

Some researchers use ordinary least squares estimates for trend calculations, but use a bootstrap (resampling) technique to calculate the standard errors (e.g. Efron and Tibshirani, 1993). The use of seasonal segments in such resampling techniques allows for proper handling of autocorrelation and seasonal weighting in the regressions. This has been found to lead to similar standard errors as the explicit autoregressive model.

The jack-knife technique (Miller, 1974) is also used in cases where data are missing in the time series. In this approach, consecutive blocks of data are dropped from the time series instead of resampling the data. The variability of the parameters with each run gives an overall error estimate for each model term.

3.3.6. Trend term

For data series beginning after 1970, the trend term, R_t, is taken to be a linear function of time, with the zero point normally the first available data point. For series which extend back into the 1960's or earlier, a "hockey stick" trend has commonly been used: a ramp function which is zero up to (usually) 12/69, and then linear with time thereafter. This was originally designed to approximate expected ozone depletion from emissions of CFCs. In the 1994 Ozone Assessment (WMO, 1995), the data before 1980 was allowed to have a trend independent of that after 1980. This "double jointed hockey stick" was used to estimate differences in trends in the 1970's vs. the 1980's. It is also possible to use the predicted effect of chlorine driven depletion from photochemical models, although this has the disadvantage of changing as the chemical models are improved.

A few researchers continue to use 12/69 as the zero point for the trend term R_t which begin later (e.g., TOMS in 11/78). This does not affect estimated trends in ozone, but does affect the intercepts, m_i, and therefore also affects percentage trends for those researchers who use the intercepts, rather than the monthly or seasonal mean, as the base for percentage trends. These differences are usually small unless the trends themselves are large.

3.3.7. Seasonal weighting

Ozone is more variable in the winter than in the summer, and application of monthly or seasonal weights in the regression analysis accounts for this. This will have the effect of increasing the standard errors of the ozone trend estimates for the high variability months, while decreasing them for the low variability months, compared to an ordinary unweighted analysis. Initial weights are usually chosen inversely proportional to individual monthly ozone variances, and then updated from the residuals of an initial regression fit. This can be iterated to convergence if desired, although using only one or two iterations leads to essentially the same results.

When auto-regressive models are fit, the monthly or seasonal weights can be applied to the noise N_t series or the uncorrelated e_t series with nearly identical results, and the choice is reasonably made based on computational convenience.

3.3.8. Intercomparisons

Participating researchers were asked to perform trend calculations on the three test data sets and report the results for the intercomparison, together with notes on their models. Trend results were requested by season (Dec-Jan-Feb, Mar-Apr-May, Jun-Jul-Aug, and Sep-Oct-Nov) as well as Year Round. Trends were requested in natural units per year (e.g., DU/year for total ozone) and %/decade (with respect to the mean of the slope period, or noted otherwise).

Ten researchers (or research groups) responded with at least some test analyses:

Bishop:	Allied Signal

Fioletov:	Atmospheric Environmental Services, Canada
Hollandsworth and Flynn:	NASA/GSFC
Logan and Megretskaia:	Harvard Univ.
McCormack and Hood:	Univ. Arizona
Newchurch and Yang:	Univ. Alabama/Huntsville
Randel:	NCAR
Reinsel:	Univ. Wisconsin/Madison (same model as Tiao *et al.*/ Univ. of Chicago)
Wang and Cunnold:	Hampton Univ. /Georgia Tech
Ziemke and Chandra:	NASA/GSFC

Figures 3.2, 3.3, and 3.4 show the trend results in natural units together with one standard error uncertainties.

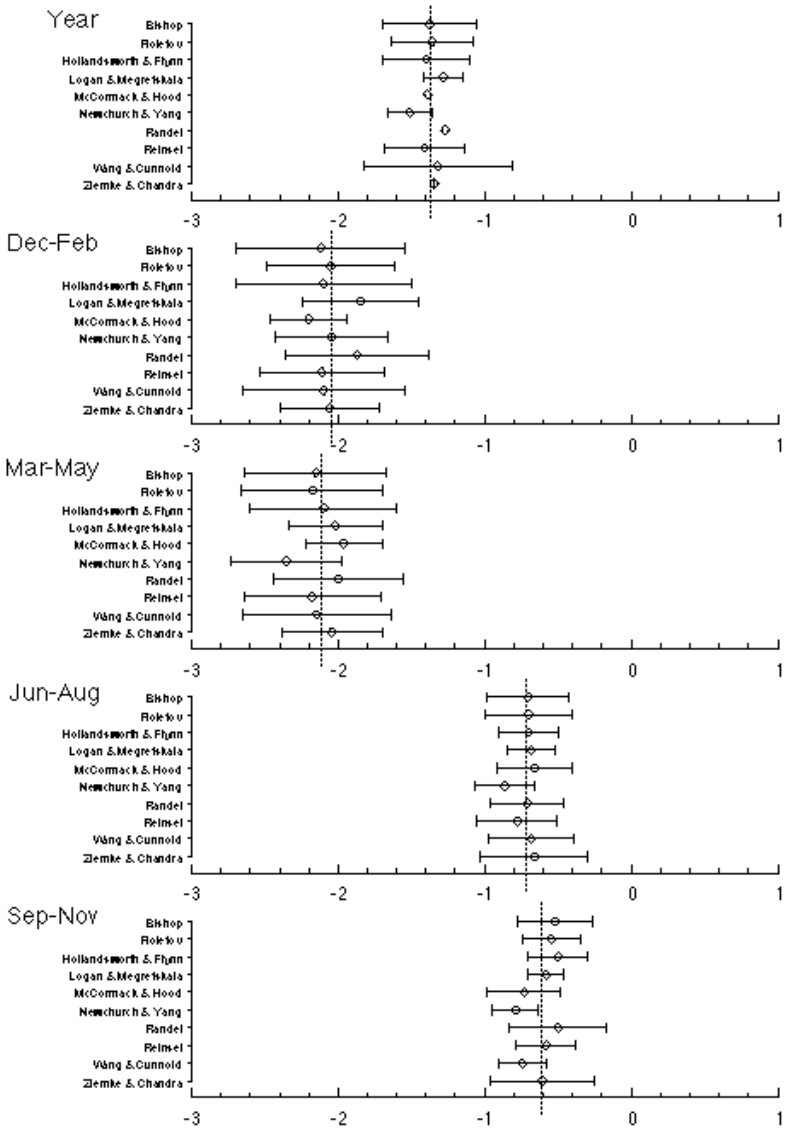


Figure 3.2 Ozone trend estimates by researcher for Test Set 1, TOMS monthly total ozone 11/78-4/93, 40°N-50°N. Trends are in DU/year. Uncertainty intervals are one standard error. Year round trends (average over all months) were calculated where possible if not given directly by the researcher, but in some cases not enough information was given to calculate the corresponding standard error. Vertical dotted line is the mean of all researchers' trend estimates.

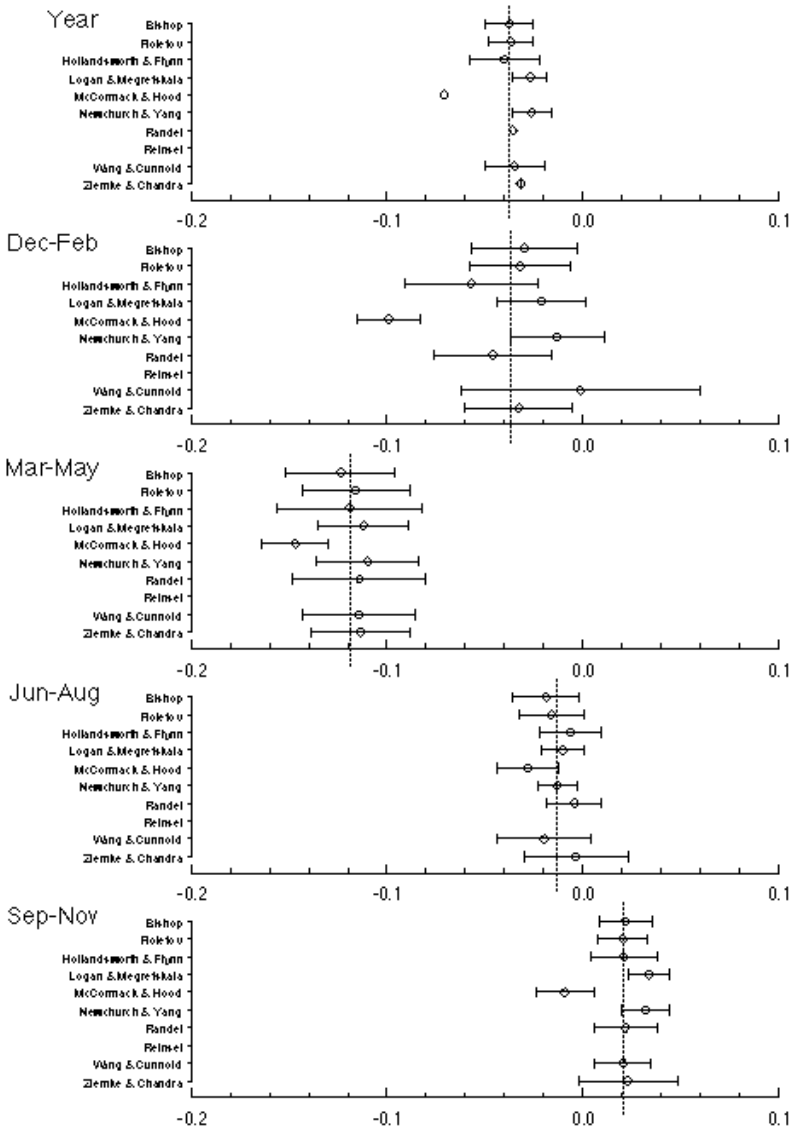


Figure 3.3 Ozone trend estimates by researcher for Test Set 2, Uccle sonde 13 km ozone 01/69-12/96. Trends are in mPa/year. Uncertainty intervals are one standard error. Year round trends (average over all months) were calculated where possible if not given directly by the researcher, but in some cases not enough information was given to calculate the corresponding standard error. Vertical dotted line is the mean of all researchers' trend estimates.

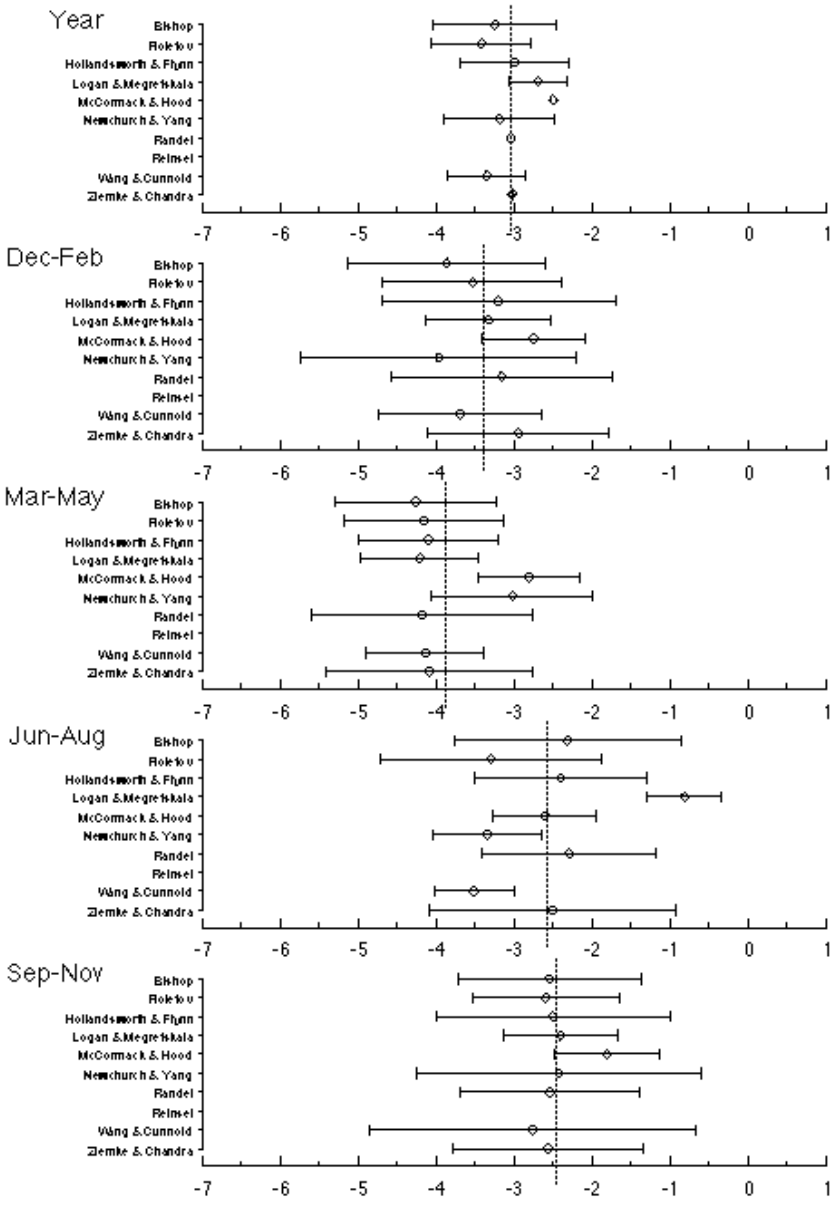


Figure 3.4 Ozone trend estimates by researcher for Test Set 3, SAGE I and II 40.5 km ozone 11/78-4/93, 35°N-45°N. Trends are in 10^9 cm^{-3}/year. Uncertainty intervals are one standard error. Year round trends (average over all months) were calculated where possible if not given directly by the researcher, but in some cases not enough information was given to calculate the corresponding standard error. Vertical dotted line is the mean of all researchers' trend estimates.

3.3.9. TOMS test data set

The ten sets of trend estimates are very similar in the case of the TOMS test data set, although standard errors vary noticeably, particularly for the case of the year round trend (average of all months or seasons). In all but one case, the mean estimate over all researchers is contained within the one standard error estimate of each researcher, and in every case within two standard errors. The TOMS test data is different from the others in that there are no missing data, and it is likely that this feature has led to the greater degree of consistency than is found in the Uccle and SAGE results. The primary differences among researchers' results on the TOMS data set are found in the standard errors:

· Researchers with seasonal weighting obtain estimates with larger standard errors in the winter and spring, and smaller standard errors in the summer and fall.

· Researchers using autoregressive models quote larger standard errors than those who don't, particularly for the year round trend (when available). See discussion on the example model earlier in this section.

3.3.10. Uccle test data set

Most, but not all, researchers obtained consistent results. McCormack and Hood calculated trend estimates substantially lower (more negative, or less positive) than the others for unknown reasons. This was especially noticeable in the winter and fall seasons. Wang and Cunnold reported a standard error for the Dec-Jan-Feb trend estimate twice as large as the other researchers ? although this is the most variable season for ozone, several other researchers also include seasonal weighting in their model. Their estimate was also much larger, though the one standard error uncertainty interval still included the mean of all researchers, since the interval was so large. Other than these two discrepancies, the primary differences were as discussed for TOMS above, in the seasonality and size of the standard errors.

3.3.11. SAGE test data set

For seasons other than the summer (Jun-Jul-Aug), most researchers obtained reasonable consistency for both trend and standard errors (other than the size and seasonality variations due to autoregression and seasonal weighting model differences). Possible exceptions were McCormack/Hood and Newchurch/Yang in Mar-Apr-May, whose estimates were larger than the others.

The SAGE data pose a particular problem in regard to missing data, with the very irregular presence of data for SAGE I, followed by a break, and then specific patterns of missing data for SAGE II. As an example, there are no values for June until 1990, and August values are missing from 1986 through 1993 inclusive. A researcher who calculates monthly trends and averages these into seasonal trends will obtain a different summer trend than one who computes seasonal trends directly, either through a seasonal model using monthly data or, especially, one based on seasonal mean ozone where one must either estimate the missing June and August values, or discard the May-Jun-Jul season ozone altogether if any one month is missing.

The summer estimates (Jun-Jul-Aug) for SAGE I/II pose special problems as discussed above. There is a peculiar pattern of missing data in these months, and the researchers' decisions as to how to handle this appear to have led to substantial differences in trends and standard errors. The average estimate was about $-2.5¥10^9$/year, with all researchers between -2 and -3.5 except for Logan and Megretskaia at -0.8. Logan and Megretskaia also quoted an unusually small standard error of 0.5. Even for the group of researchers who obtained fairly similar estimates for the summer, the quoted standard errors varied by a factor of three. When fitting their monthly model, Newchurch and Yang obtained a very large negative trend estimate in the summer season with an extremely large standard error, and they suggested use of their seasonal model (based on seasonal ozone averages) to maintain stability; their seasonal results are those shown in the figure. They believe the summer instability in the monthly model is attributable to use of a separate AR(1) coefficient in each month ? their method of calculation leads to the loss of data for both neighbouring months whenever either is missing, and this results in little data for fitting the summer AR(1) coefficients and trends.

3.3.12. Effect of neglecting solar, QBO and other terms

One important question is how the inclusion or neglect of the solar, QBO and other terms in the statistical models influences the derived trend and standard error estimates (i.e. how sensitive are the trend results to details of other model terms?). In order to examine this, the test time series were analysed using models with only a linear trend component. Comparison to the 'full' model trend results showed relatively small (~10%) changes in values of the trends. For the TOMS test data the trend standard errors were increased somewhat (by up to 20%) by neglect of the other components. On the other hand, the trend standard errors for the Uccle and SAGE test data were slightly reduced by use of the simpler model. Detailed changes in standard error would likely be sensitive to location (such as at the equator, where the QBO component is relatively important). However, the overall conclusion is that the trend results are relatively insensitive to inclusion of other terms in the statistical models. This is probably due to the fact that the time series are sufficiently long compared to the 11-year solar cycle and ~2 year QBO periodicities. A similar insensitivity of trend results is found concerning the inclusion or neglect of data during the El Chichon and/or Mount Pinatubo time periods (for data through 1996, as used throughout much of this report). The episodic influence of volcanic events will only have a strong effect on trends if they occur at either extreme end of the time series.

3.3.13. Conclusions and implications

· On the test data with no missing monthly values (TOMS), all researchers obtained similar results for trends. There were variations in standard errors, however, which were large enough to give some concern, as they affect the statistical significance of the trends. For example, long-term total ozone trends near the equator border on statistical significance (WMO, 1995) - lack of proper calculation of standard errors in such a situation may result in non-significant trends being declared statistically significant, or vice versa.

· For the test data sets with missing values, particularly the SAGE set with large numbers of missing values in a strong pattern form, some researchers' results deviated substantially from the average, both with respect to trend estimate and, especially, standard error. Some researchers feel that in such situations, it is better to fit a simpler model to maintain stability; for example, by fitting seasonal trends directly or by use of reduced numbers of harmonic terms for the seasonal trends and possibly also the seasonal cycle.

· Based on these intercomparisons, it is important that researchers provide good documentation for the features of their statistical model. Particularly, when any patterns of missing data exist which have strong time dependent features, the methods of handling the missing data should be discussed in detail.

3.4. Natural and forced variations in ozone

Global ozone observations display both a long-term decline and considerable interannual variability. Sources of long-term variability include changes in solar output, temporal changes in the abundance of stratospheric aerosols, and variability in the stratospheric circulation leading to changes in ozone transport on seasonal, interannual and decadal time scales. In regards to such variability, there are two questions of relevance for trend estimates: 1) how do the presence and lack of detailed knowledge of such signals affect the estimation of trends?, and 2) how are uncertainty estimates affected? In order to quantify these influences, it is important to characterise as much of the observed variance as possible in the regression models used to estimate trends.

Several components of interannual variability have been identified in ozone observations and used in prior statistical analyses. In particular, variations associated with the stratospheric quasi-biennial oscillation (QBO) and 11-year solar cycle have been studied extensively (see references below). The QBO makes a relatively large contribution to interannual variance, and the fact that over six complete QBO cycles have been observed in global satellite data allows accurate empirical fitting of this signal. The relatively short oscillation period of the QBO means there is little effect on long-term trend estimates, although uncertainty estimates (particularly in the tropics) are significantly improved by including a QBO term in the regressions. In the case of the solar cycle less than two complete oscillations have been observed during the satellite record, and there is not detailed agreement between the satellite observations and idealised model calculations (as discussed below). The proper inclusion of the solar effect can have a large impact on trends only for relatively short time series (<10 years); there is less of an influence for the ~18 year (or longer) time series of focus here. However, detailed understanding of the decadal variation in ozone, including a departure from linear trends and minimisation of trend uncertainty, requires accurate knowledge of the solar effect. An important complicating factor arises in the recent observational record in that there is a near temporal coincidence between the declining phases of the solar cycles in 1982-84 and 1992-94 with the volcanic eruptions of El Chichon (1982) and Mount Pinatubo (1991), such that there is a possible aliasing of these decadal signals (Solomon *et al.*, 1996). Additionally, there is a suggestion that decadal variations in stratospheric temperature and dynamical fields may influence ozone transport and trends (Hood and Zaff, 1995; Peters and Entzian, 1996; McCormack and Hood, 1997; Hood *et al.*, 1997). The influence of dynamical changes on ozone trends is a topic of current research and debate. We discuss below the current understanding of these various signals, including outstanding uncertainties and the implications for trend calculations.

3.4.1. Solar Cycle

Changes in solar ultraviolet spectral irradiance directly modify the production rate of ozone in the upper stratosphere (e.g. Brasseur, 1993), and hence it is reasonable to expect a solar cycle variation in ozone amount. Analyses of ground-based records extending over three to six decades indicate the existence of a decadal time scale variation of total ozone that is approximately in phase with the solar cycle (e.g., Angell, 1989; Miller *et al.*, 1996; Zerefos *et al.*, 1997). The global satellite ozone records since 1979 show evidence for a decadal oscillation of total ozone with maximum amplitude (~2%) at low latitudes (Hood and McCormack, 1992; Chandra and McPeters, 1994; Hood, 1997).

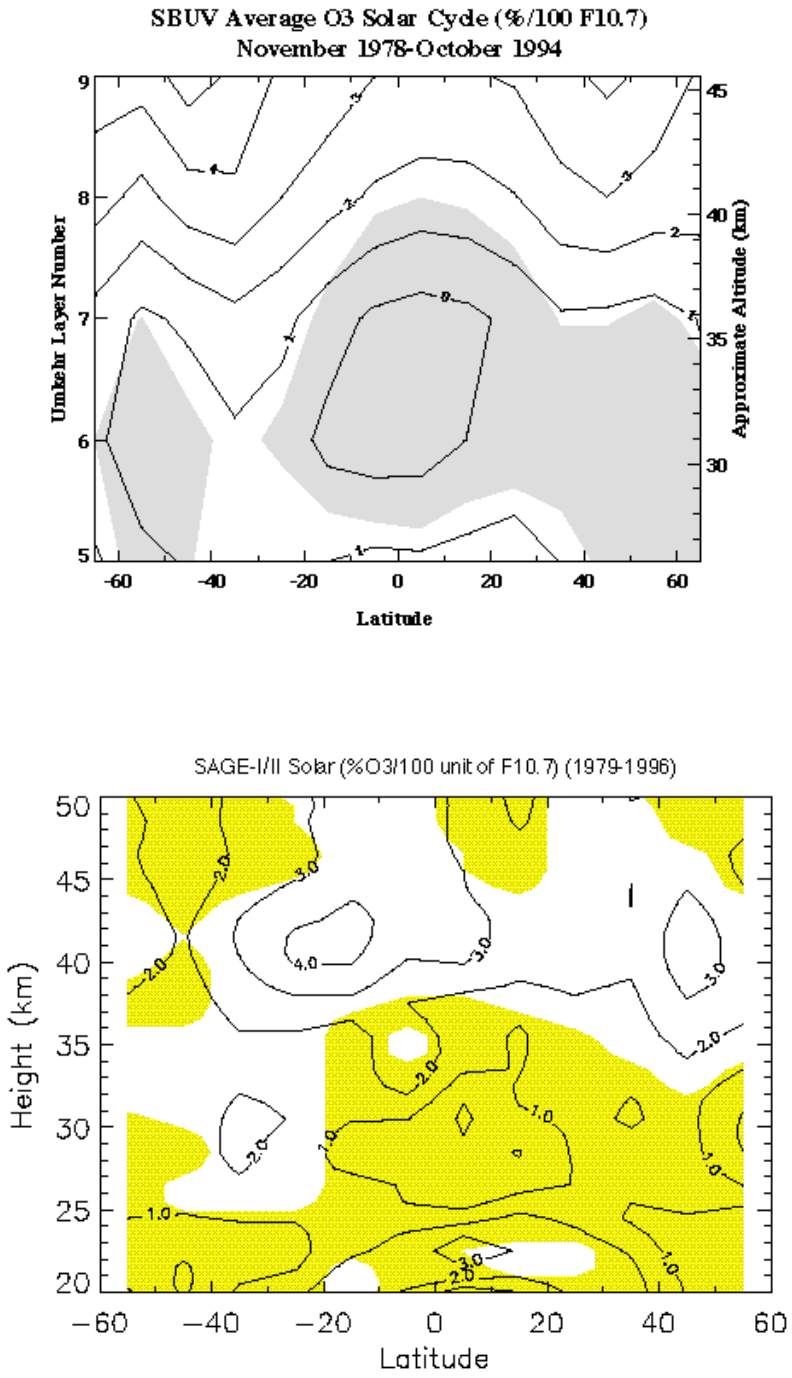


Figure 3.5 Meridional cross sections of the solar signal derived from a) SBUV2 data and b) SAGE I/II data using regression analysis. Contour interval is 1%/100 units of F10.7.

During at least the last three solar cycles, the decadal oscillation has been approximately in phase with proxies for solar ultraviolet flux, and statistical representation is typically based on correlation with the 10.7 cm solar radio flux (F10.7) or the Mg II core-to-wing ratio (Heath and Schlesinger, 1986; DeLand and Cebula, 1993).

The dependence on altitude of the ozone solar cycle signal has been investigated using SBUV/SBUV2 data (Hood *et al.*, 1993; Chandra and McPeters, 1994; McCormack and Hood, 1996) and using SAGE I/II data (Wang *et al.*, 1996). Figure 3.5a,b shows meridional cross sections of the solar signal derived from SBUV/SBUV2 and SAGE I/II data using regression analysis (from the standard model fits

discussed above). These plots show percentage ozone changes per 100 units of 10.7 cm solar flux; 130 units is the approximate difference between solar maximum and solar minimum. These data show reasonably consistent patterns in the upper stratosphere, with an in-phase solar cycle variation of 2-4%. Similar sized solar variations in the upper stratosphere are derived from a somewhat longer record (1968-onwards) of Umkehr data (Miller *et al*., 1995). The SAGE I/II data (with resolution below 25 km) furthermore show a maximum over 23-28 km centred in the tropics (£30°N-S). The vertical integral of these SAGE I/II data results in a column ozone variation of approximately 2%, and this is in reasonable agreement with the TOMS and SBUV/SBUV2 column ozone solar signal derived over the same time period (1979-1996). Note the SBUV/SBUV2 profile data do not extend into the lower stratosphere, and hence this lower stratospheric tropical decadal variation (~ 85% of the column signal) is not resolved (as discussed in Hood, 1997).

Figure 3.6 shows the solar cycle variation derived from the model results of Jackman *et al*. (1996), who include detailed treatment of the solar spectral characteristics derived from UARS data. This model result shows a broad latitudinal maximum in the upper stratosphere with maximum amplitude near 1.5%; this is similar in altitude dependence but smaller in magnitude (by about 50%) than the satellite results shown in Figure 3.5. The model does not show a maximum in the tropical lower stratosphere (i.e. the maximum seen in SAGE I/II data in Figure 3.5b). Consistent with this, the model column variation is approximately 1%, somewhat less than the 2% derived from the relatively short record of TOMS. Very similar model estimates of the solar cycle variation have been obtained by Brasseur (1993) and Haigh (1994).

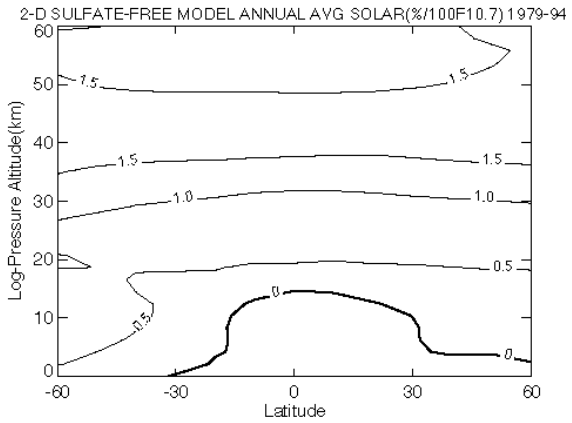


Figure 3.6 Solar cycle variation derived from a 2D photochemical-dynamical model including solar cycle variation of the ultraviolet solar flux responsible for ozone production [Jackman *et al*. (1996)].

In summary, the decadal variation in upper stratospheric ozone observed in satellite data is similar in spatial pattern to model calculations of the solar cycle, but the observed changes are approximately a factor of two larger than model results. In addition, the satellite data show an in-phase variation in the lower stratosphere not seen in model calculations. Consistent with this, the decadal oscillation in column ozone derived from data is somewhat larger than model predictions. This source of observed-model discrepancy is unresolved at present; it may be due to ozone transport effects that are unresolved in the models, or to uncertainties associated with the relatively short observational record. This uncertain knowledge of solar effects is one limiting factor in understanding long-term ozone variability.

3.4.2. Aerosol Effects

Record low ozone levels observed after the eruption of Mount Pinatubo in 1991 (Gleason *et al.*, 1993) led to a renewed appreciation of the effects of volcanoes on stratospheric ozone. Hofmann and Solomon (1989) proposed such an explanation for relatively small (~2%) ozone losses after El Chichon (in 1982), based on heterogeneous chemistry occurring on the volcanic sulphate aerosols. Observed losses after Pinatubo were substantially larger, of order 5-10% in NH middle-high latitudes (e.g. Gleason *et al.*, 1993; Hofmann *et al.*, 1994; Bojkov *et al.*, 1993; Kerr *et al.*, 1993; Grant *et al.*, 1994; Randel *et al.*, 1995; Coffey, 1996). Numerous modelling studies have shown ozone depletion with similar magnitude and space-time structure (maximising in NH winter 1991-92 and 1992-93) (e.g. Tie *et al.*, 1994; Bekki *et al.*, 1994; Rosenfield *et al.*, 1997). This consistency in modelling the 'fingerprint' of Pinatubo on ozone supports confidence in physical understanding of the heterogeneous chemical mechanisms involved.

Recent studies have shown that including realistic time series of sulphate aerosol surface area density in 2-D model calculations of total ozone improves the agreement with observations in the northern hemisphere middle latitudes (Solomon *et al.*, 1996) and in global ozone (Jackman *et al.*, 1996). Thus a primary source of ozone variability appears to be related to the amount of stratospheric sulphate aerosols, which is reasonably well constrained by observations (Thomason *et al.*, 1997). The episodic nature of the volcanic events allows their influence to be isolated, and the effects on trend calculations is minimal unless they occur near one end of the time series (the volcanic time periods are often simply omitted from the trend calculations).

As noted by Solomon *et al.* (1996), the occurrence of two major volcanic eruptions nine years apart during each of the last two declining phases in solar activity could lead to some confusion in separating volcanic and solar effects on ozone. This will not strongly influence trend estimates for the long time records, but may have implications for isolation of the solar cycle in the short observational record. This is demonstrated using model output in Figure 3.7, where the model results of Jackman *et al.* (1996) including sulphate aerosol effects are used to calculate the apparent solar cycle variation over 1979-1994; this result should be compared to the 'true' model solar cycle in Figure 3.6 (derived from a 'no-sulphate' model run). This comparison shows a large projection of the aerosol effects in the lower stratosphere onto the solar cycle during this time period; the derived solar signal in the lower stratosphere is a completely spurious result in the aerosol simulation. This result suggests caution should be applied to detailed interpretation of apparent solar cycle variations in the relatively short satellite data record since 1979 (especially for lower stratospheric profile and column ozone).

Attempts to statistically model the volcanic influence on ozone in regression analysis is a relatively novel idea and a topic of current research. One simplified, first-order statistical approach toward separating volcanic and solar effects involves including an explanatory variable linked to the atmospheric aerosol content in multiple regression models (McCormack *et al.*, 1997). This is only an approximate approach since, for example, it neglects the multiplicative effect of increasing chlorine loading during the 1980's (Tie and Brasseur, 1995), or the important effects of temperature variations.

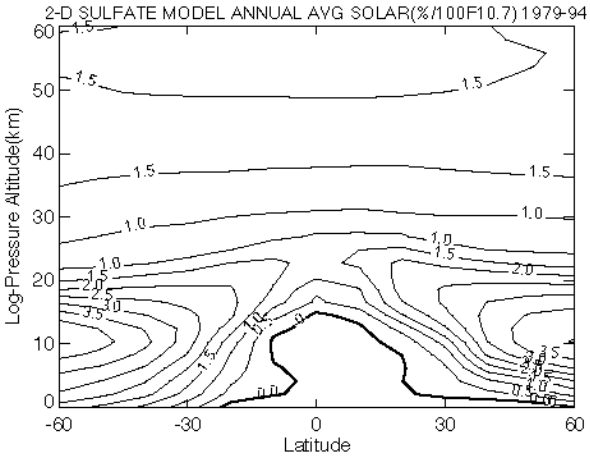


Figure 3.7. Apparent solar cycle variation derived from a 2D photochemical-dynamical model which included solar cycle variation of ultraviolet solar flux and variation of the aerosol concentration due to volcanoes such as El Chichon and Mount Pinatubo [C.H. Jackman, private communication].

3.4.3. Quasi-biennial Oscillation

Column Ozone

Interannual variability in tropical ozone is dominated by an approximate 2-year cycle, which is closely linked with the quasi-biennial oscillation (QBO) in zonal wind and temperature in the tropical stratosphere. A strong QBO component in ozone is also observed at extra-tropical latitudes in both hemispheres. Analyses of long time records of global satellite data from TOMS have clearly documented characteristics of the global QBO in column ozone (Bowman, 1989; Lait *et al*., 1989; Chandra and Stolarski, 1991; Randel and Cobb, 1994; Tung and Yang, 1994, Yang and Tung, 1994). The main results of these analyses show: 1) column ozone variations near the equator (£10^{o} latitude), approximately in phase with equatorial zonal winds near 30 hPa, and 2) extra-tropical anomalies over approximately 15^{o}-60^{o} latitude in each hemisphere, approximately out of phase with the tropical signal. The amplitude of the column ozone QBO anomalies is of order 2-4% of the mean ozone amount.

An intriguing and as yet poorly understood aspect of the extra-tropical QBO is that it is seasonally synchronised, such that the QBO influence is observed only during winter-spring of each respective hemisphere. Similar results have been derived from ground-based measurements (e.g. Hamilton, 1989; Yang and Tung, 1994) and other satellite-derived column ozone data (Hasebe, 1983; Hollandsworth *et al*., 1995).

Besides the QBO signals in the tropics and mid-latitudes, analysis of TOMS data and other long-term records suggest a further region of QBO influence in the winter polar vortices (Oltmans and London, 1982; Garcia and Solomon, 1987; Bowman, 1989; Lait *et al*., 1989; Randel and Cobb, 1994). The polar ozone QBO is approximately in phase with mid-latitudes (and out of phase with the tropics), and also seasonally synchronised (maximum in spring). Observational evidence for the polar ozone QBO is less statistically significant than that in the tropics or mid-latitudes, at least partly because of the high level of "natural" interannual variability in springtime vortex structure.

Statistical modelling and isolation of the column ozone QBO is straightforward, using a regression

analysis tied to the equatorial zonal winds. It is important to include seasonally-dependent coefficients in such regressions to accommodate the seasonally synchronised nature of the mid-latitude QBO. The details of such analyses can depend on the specific altitude of the reference time series (because the QBO propagates downward in time), and some studies include a latitude-dependent time lag. Randel *et al*. (1995) have used a linear combination of zonal winds over 10-70 hPa as a basis time series for column ozone QBO regressions, and Ziemke *et al*. (1997) have shown that this provides optimal fits and maximum reduction in variance (and that no time lag is necessary).

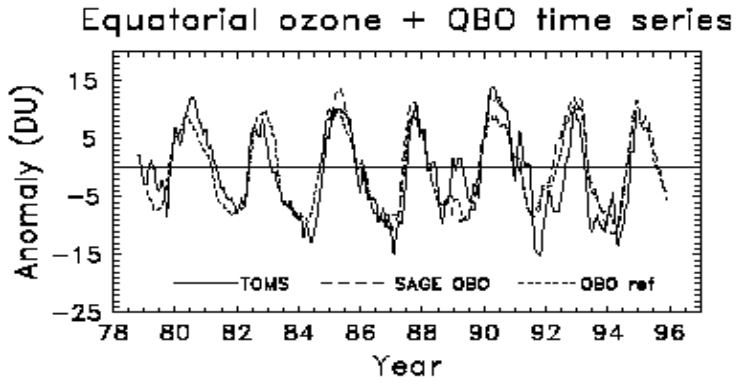
Figure 3.8a shows interannual anomalies in TOMS column ozone at the equator, together with this optimal QBO reference time series. The TOMS anomalies at 28° N are shown in Figure 3.8b, together with the QBO fit derived by regression analysis (note the large anomalies during NH winter, approximately out of phase with the tropical QBO variations in Figure 3.8a).

Ozone Profile

The vertical structure of the ozone QBO has been analysed using SAGE II data by Zawodny and McCormick (1991), Hasebe (1994) and by Randel and Wu (1996). Their results show that there is a two-cell vertical structure to the tropical and mid-latitude QBO, peaking in the lower stratosphere (20-27 km) and middle stratosphere (30-38 km), with a node near 28 km. This is shown in Figure 3.9a,b, for anomalies over the equator and at 28° N. The two maxima are approximately in quadrature in the tropics, but are in phase at mid-latitudes. Both the lower and middle stratosphere components contribute important fractions to the column ozone amounts (approximately 2/3 and 1/3 fractions, respectively). The vertical integrals of these SAGE II anomalies are included in the time series shown in Figure 3.8, showing good agreement with the QBO component in the TOMS data.

SAGE II measurements of nitrogen dioxide furthermore show a strong QBO signal in the middle stratosphere (Zawodny and McCormick, 1991; Randel and Wu, 1996), and Chipperfield *et al*. (1994) demonstrated from modelling results that the two-cell vertical structure in ozone was due to transport (in the lower stratosphere) and reactive-nitrogen associated chemical effects (in the middle stratosphere), respectively.

Accurate statistical modelling of the profile ozone QBO may be accomplished using regression analyses based on two (orthogonal) reference QBO time series (Randel and Wu, 1996), which allows for the continuous vertical propagation observed in the profile anomalies. Alternatively, the fitting can be accomplished by using a single reference time series with variable lag at each latitude and height.



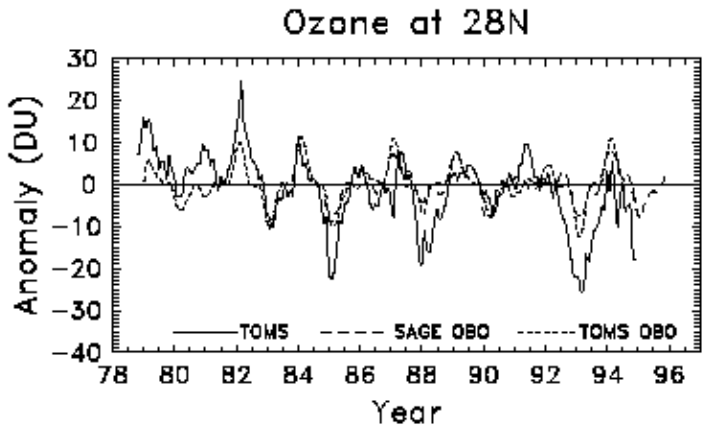


Figure 3.8 Time series of deseasonalized anomalies in TOMS column ozone at the equator and 28°N. Light dashed lines in (a) show the QBO reference time series, derived from an optimal linear combination of zonal winds over 70-10 mb (Randel et al., 1995). Light dashed lines in (b) show the statistical QBO fit at 28°N, derived using seasonally-varying regression. Heavy dashed lines in both panels show column ozone anomalies above 20 km derived from the SAGE II data shown in figure 3.9.

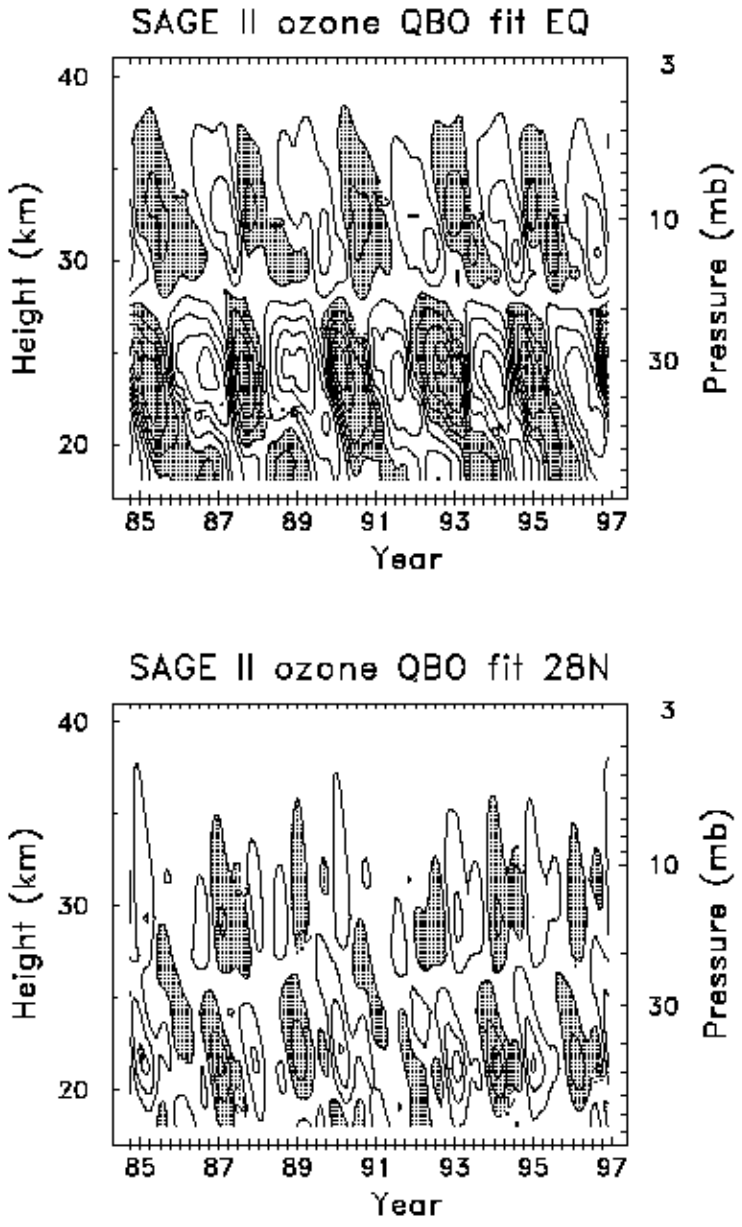


Figure 3.9 Altitude-time sections at the equator (a) and 28°N (b) of SAGE II ozone density anomalies (DU/km) associated with the QBO. Contour values are ± pm 0.2, 0.6, 1.0, ... DU/km with positive values shaded. Note the separate maxima in the lower and middle stratosphere; these contribute about 2/3 and 1/3 of the column amounts, respectively. The vertical integral of these anomalies are included in Figure 3.8.

3.4.4. Dynamical Proxies

Ozone is strongly influenced by meteorological variability on both short and inter-annual time scales. Accurate characterisation of such dynamical variations therefore allows increased precision in trend estimates (or at least a reduction in trend uncertainties). An example of "anomalous" meteorological variability is the El Nino-Southern Oscillation (ENSO), which is known to affect column ozone amounts (Shiotani, 1992; Randel and Cobb, 1994). However, the effects of ENSO are mainly longitudinally localised, and the influence on zonal mean ozone is typically less than 1%, so that little reduction in overall variance is gained by including an ENSO proxy in regression models for zonal mean ozone.

With regard to "natural" meteorological variability, Ziemke *et al*. (1997) have investigated the use of several meteorological fields for use as dynamical proxies for column ozone, including temperature, geopotential height, relative and potential vorticity, and other fields. The strongest correlation (and reduction in regression variance) was obtained using lower stratospheric temperature data (from Microwave Sounding Unit (MSU) Channel 4). An important point to note is that lower stratospheric temperatures also exhibit negative trends (Miller *et al*., 1992; Spencer and Christy, 1993), with spatial structure similar to that of observed ozone trends (Randel and Cobb, 1994). Modelling studies suggest that much of the observed temperature trend can be explained as a radiative response to the observed ozone changes (Ramaswamy *et al*., 1996). In addition, a portion of this temperature trend may be of dynamical origin (Hood *et al*., 1997). In either case, using the full time series of temperature data (including a long-term trend) as a dynamical proxy can result in a change (underestimate) in the derived ozone trends. A practical alternative is to de-trend the temperature data prior to use in the regression model, and use the de-trended data to model the high-frequency components of the ozone variance. The overall effect of including such a dynamical proxy is a small reduction in the derived trend uncertainty estimates for total ozone (Ziemke *et al*., 1997).

Summary

Decadal variations are a ubiquitous feature of ozone observations, in addition to QBO and faster time scale dynamical variability. These latter terms have little influence on calculation or interpretation of trends. Much of the observed decadal changes are approximately in phase with the solar cycle for the observational record, suggesting a solar mechanism. However, current model calculations of the solar effect show some inconsistencies with observations (in terms of magnitude and lower stratospheric response), and this limits confidence in our detailed understanding. There is also likely a confusion of solar and volcanic signals for the recent record. Although these effects have relatively small impacts on trend estimates, it does limit our ability to interpret decadal variability.

3.5. Trend Results

3.5.1. Upper stratosphere

Previous studies of stratospheric ozone trends (WMO, 1995; DeLuisi *et al*., 1994; McPeters *et al*., 1994; Reinsel *et al*., 1994; Miller *et al*., 1995, 1996) are summarised in Harris *et al*. (1997). They found that ozone amounts from approximately 1979-1991 had been declining at the rate of 5-10%/decade in northern mid-latitudes in the altitude region of maximum active chlorine (35-45 km). Those trend estimates were reasonably consistent between three measurement systems Dobson/Umkehr, SAGE I/II, and SBUV. The altitude-latitude structure of the measured ozone trends is also consistent with both the theoretical understanding of chlorine-catalysed ozone destruction (Solomon and Garcia, 1984; Kaye and Rood, 1989) and the measured latitudinal distribution of ClO (Aellig *et al*., 1996; Waters *et al*., 1996). Subsequent analyses of SAGE I/II trends through the same period employing an altitude correction for the SAGE I observations (Wang *et al*., 1996) reconciled differences between SAGE I/II and SBUV trends that had been present in the tropical lower stratosphere. Subsequent analysis of combined SBUV and SBUV2 [SBUV/SBUV2] trends (Hollandsworth *et al*., 1995), extended through 1994, did not substantially change that general agreement. These upper-stratospheric trends exhibited latitudinal and seasonal variations such that the trends were more negative in the winter and spring seasons at high latitudes. The trends in the tropical latitudes are less negative throughout the stratosphere and exhibit little seasonal or altitudinal variation. At somewhat lower altitudes (10-20 hPa ~30 km) these three systems, in addition to ozonesonde observations, concurred in finding no significant ozone loss at any latitude over the 1979-1991 period.

The deduced trends in ozone concentration are in general agreement with theoretical predictions (e.g., Chandra *et al*., 1995; Jackman *et al*., 1996) that implicate halogen-induced ozone destruction. The effect of the Montreal Protocol and its amendments is now apparent in measured tropospheric halogen amounts (Montzka *et al*., 1996), and similar declines in upper-stratospheric halogen levels are anticipated. Theoretical models predict a reversal in the ozone decline to occur near the end of this decade as a result of this decline in halogen amounts (Jackman *et al*., 1996).

One of the objectives of this report is to calculate the linear trends from measurements of stratospheric ozone through 1996. Because of the anticipated reversal in declining ozone amounts, subsequent studies will likely focus on detecting positive deviations from the negative linear trends established by these analyses. In the context of the Chapter 2 assessment of appropriate measurement systems for calculating stratospheric ozone trends, we consider times series of Dobson/Umkehr, SAGE I/II, and SBUV/SBUV2 observations from 1979 to 1996 [1994 for SBUV/SBUV2]. These data have been carefully screened such that only highly reliable measurements are used in the trend calculations.

3.5.1.1. Time series of Umkehr, SAGE I/II, and SBUV/SBUV2

We will begin by focusing on northern mid-latitudes and an altitude of 40 km (Umkehr layer 8). Before calculating trends, we first illustrate the time series from 1979 to 1996 [1994 for SBUV/SBUV2] of ozone observations from three independent sensors, Dobson/Umkehr (Figure 3.10a), SAGE I/II (Figure 3.11a), and SBUV/SBUV2 (Figure 3.12a).

Umkehr

The Umkehr observations in Figure 3.10a show the Boulder station as an example of the northern mid-latitudes. These monthly ozone averages, reported in natural Umkehr units of Dobson Units (DU) in the

layer, are corrected for aerosol interference. The missing data periods result from eliminating observations during periods of very high stratospheric aerosol loading. The ozone values calculated by the statistical model also appear in the top panel and cover the missing data periods. The model residuals (i.e., the measurements minus the statistical model) for the Boulder Dobson/Umkehr record appear in the bottom panel as an example of the remaining statistical variance after removing the mean, trend, solar, and QBO terms.

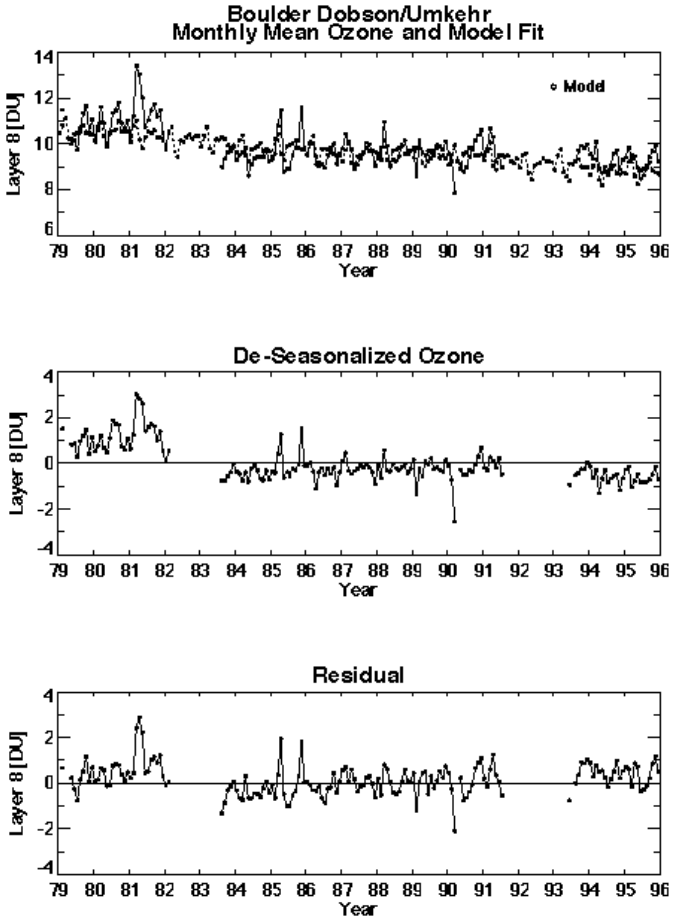


Figure 3.10 a) Monthly time series of layer 8 (40 km) ozone amounts [DU] from the Boulder Dobson/Umkehr station corrected for aerosol effects with periods of high stratospheric aerosol optical depth omitted. Also shown is the statistical model fit to this data. b) Deseasonalized data from Boulder. c) The residual between the data at Boulder and the model fit.

SAGE I and SAGE II

The SAGE I/II observations in Figure 3.11a are monthly averages of version 5.96 ozone retrievals in a 5-km thick layer between 38-43 km between latitudes 40°-50°N. These averages, which include both sunrise and sunset occultation measurements, are reported in the natural SAGE units of ozone number density versus geometric altitude.

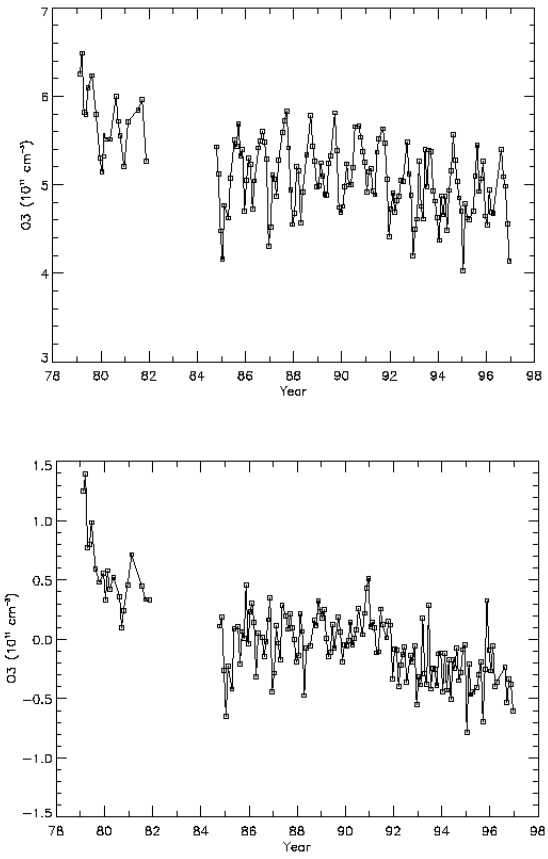


Figure 3.11 a) Monthly time series of layer 8 (38-43 km) ozone concentrations (10^{11}cm^{-3}) from SAGE I and SAGE II sunrise and sunset observations between 40°N-50°N. b) Deseasonalized data from a).

SBUV and SBUV2

Figure 3.12a shows the monthly-mean SBUV/SBUV2 ozone data in layer 8 from 1979 to 1994 averaged from 40°-50° North. These data are reported in natural SBUV units of DU versus pressure in Umkehr layers. The SBUV2 daily average 5° zonal means are filtered to eliminate data taken at extreme angles as a result of the NOAA-11 drifting orbit, consistent with Chapter 2 recommendations. Then 10^{o} monthly zonal means are created from this data.

The corresponding, de-seasonalised time series for these three sensors appear in Figures 3.10b, 3.11b, and 3.12b for Umkehr, SAGE I/II, and SBUV/SBUV2, respectively. These measurements are all analysed and presented in their natural units to avoid uncertainties introduced by conversion errors. Taken together, a number of consistencies emerge. First, a significant ozone decrease between the SAGE I and SAGE II time periods (1982-1984) is apparent, with decreases of similar magnitude (~10%), in both the Umkehr and SBUV/SBUV2 observations. Second, the relatively constant level between 1984 and 1992 (with some suggestion of a slight increase between 1987-1990 in the SAGE series) and then a significant decrease after 1992 occurs in Umkehr and SAGE data but not in SBUV/SBUV2. Third, the amplitude of the annual variation is similar for all three sensors. The correspondence between these three independent ozone time series suggests that we should have considerable confidence in the trends computed from these data and should expect them to return similar results.

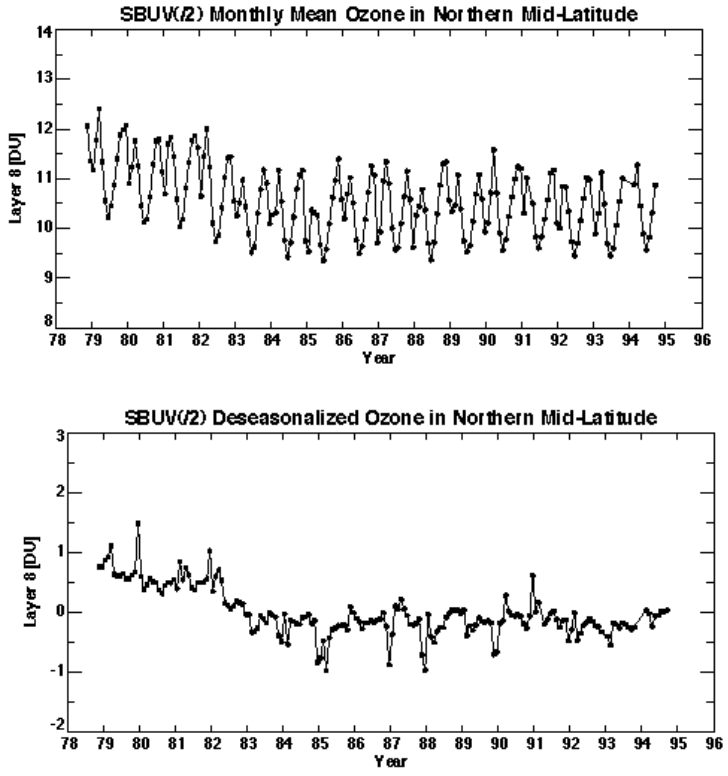


Figure 3.12 a) Monthly time series of layer 8 ozone amounts [DU] 40°N-50°N from the Nimbus-7 SBUV and NOAA-11 SBUV/2 data joined by adjusting the SBUV data by the average difference between the SBUV and SBUV/2 data over 1989. b) Deseasonalized data from a).

3.5.1.2. Trend Analyses of Umkehr Observations

The well-known aerosol interference in the Umkehr observations is an optical effect and not an in situ ozone-aerosol interaction. The following three methods have historically been employed to identify the magnitude of and correct for this aerosol interference: 1) theoretical radiative transfer calculations (Mateer and DeLuisi, 1992), 2) statistical calculations (i.e., time-series regression models employing exogenous aerosol records) (DeLuisi *et al*., 1994; Reinsel *et al*., 1994), and 3) comparisons to other ozone measurements (Newchurch and Cunnold, 1994).

In this report, to correct for the optical interference, we employ two methods. Method (1), employed by Newchurch and Yang, uses the coefficients of Mateer and DeLuisi (1992) with aerosol data from coincident SAGE II aerosol extinction measurements (Newchurch and Cunnold, 1994; Newchurch *et al*., 1995) along with Garmisch lidar backscatter measurements (H. Jaeger, private communication) for the period prior to 1984. The lidar backscatter measurements are converted to optical depths at 320 nm by regression against SAGE II coincident measurements over 1984 to 1995. The resulting continuous aerosol time series is then lagged appropriately to correct Umkehr data at various latitudes. Additionally, Umkehr data during high stratospheric aerosol optical depth periods (> 0.025 corresponding to ~1 year after El Chichon and ~1 year after Mount Pinatubo) are omitted from the analyses.

Method (2), employed by Reinsel, uses an empirical statistical model approach in which transformed stratospheric optical thickness (transmission) data are used as an explanatory variable for the Umkehr measurements. The stratospheric aerosol optical thickness data are constructed from SAGE II satellite

information for the period 1985-1996. They are formed first by linear interpolation of SAGE II aerosol data available at wavelengths 1020, 525 and 453 nm to 694.3 nm, and then by linear combination of the integrated-from-tropopause and integrated-from-20-km data sets, to approximate aerosol optical thickness values integrated from 15 km (at 694.3 nm). The resulting data were appended to the previous existing optical thickness data based on composite lidar (and SAGE II) aerosol information through December 1984 (Reinsel *et al.*, 1994). This procedure was performed separately for each of three ten-degree latitude zones, 30°-40°N, 40°-50°N, 50°-60°N. For the two southern hemisphere stations, only SAGE II aerosol data since 1985 were needed because the available Umkehr data from Lauder and Perth do not extend prior to that time.

Based on the close correspondence between the results of the three historically used correction methods for the aerosol conditions considered here, we conclude that the corrected ozone data possess less than ~2% residual error due to the aerosol interference in the worst case analysed here (immediately after the omitted periods following El Chichon and Mount Pinatubo eruptions).

The two independent Umkehr analyses both report trends for the following eight Umkehr stations grouped into three latitude zones: 30°-45°S including Perth and Lauder; 29°-36°N including New Delhi, Cairo, and Tateno; and 40°-52°N including Boulder, Haute Provence, and Belsk. The Reinsel analysis also calculated trends at the Arosa station. At each of these stations, both groups computed trends for total-column ozone, for aggregate Umkehr layers 1+2+3+4 (labelled 4^-), individual layers 4, 5, 6, 7, and 8, and aggregate layers 8+9+10 (labelled 8^+).

Miller *et al*. (1997) tested the ability of the Umkehr method to retrieve trends by superimposing a prescribed trend on an artificial data set. They then used the Umkehr technique to retrieve the data. They found that the Umkehr technique underestimated their prescribed trend in layer 8 by approximately 30% (-5%/decade computed vs. -7.5%/decade prescribed). Similarly, the Umkehr technique significantly underestimated the actual trend in layer 4 (-1%/decade computed vs. -4%/decade prescribed). In light of their trend sensitivity results combined with the solar-response simulation results of the same study, they recommend using layers 5^- and 8^+ as a conservative solution.

Mateer *et al*. (1996) focused on Umkehr ozone trends compared to ozonesonde trends and recommended using layers 1+2+3 combined, 4-7 individually, and 8^+. They noted that individual layer 4 represents trends less well than individual layers 5, 6, or 7, and the partitioning of the trend between the lower 3 or 4 layers is a strong function of the total-column ozone trend and of the a priori profile partitioning. They did not address the question of individual layer 8 retrievals because the ozonesondes extend no higher than layer 7.

These results indicate that the fidelity of trend results in individual layers 4 or 8 may be questioned. However, the computations in this report indicate that the trends in layer 8 are very similar to the trends in layer 8^+. Likewise, individual layer 4 trends are not significantly different from layer 4^- trends.

Four differences between the two Umkehr analyses in this report are 1) slightly different starting times: 1977 for Reinsel, 1979 for Newchurch and Yang; 2) the use of a level offset term in the Tateno analysis of Reinsel, but not in the analysis of Newchurch and Yang; 3) somewhat less stringent convergence criteria for the data used in the Reinsel analysis; and 4) different treatments for the aerosol optical interference. Both of these groups have used reasonable, if somewhat different, approaches to estimating ozone trends in Umkehr data. Because of the similar trend results of both groups (in fact, of all groups) from the test-sets described in section 3.3, we can be confident that different trend results arise from different analyses of the data and not from differences in numerical or statistical methods. That is to say if both groups had taken the same analysis approach, they would have computed the same trends. The

fact that the trend estimates from these two groups are, in most cases, not significantly different from one another enhances our confidence that average computed trends represent the true atmospheric condition and that the statistical confidence intervals are reasonable estimates of the uncertainties associated with these trend calculations.

Figure 3.13 shows the trend results for both Umkehr analyses as variance-weighted averages and associated 95% confidence intervals in three latitude regions. The southern mid-latitude trends (in the left panel) from both groups for stations Lauder and Perth are well within the 95% confidence intervals of each other for all Umkehr layers and for the total ozone column. Note that the data records for the two southern hemisphere Umkehr stations are relatively short (data since 8/84 and 2/87 for Perth and Lauder, respectively), so that the trend estimation results for the southern hemisphere are much less reliable than for the northern hemisphere zones. The northern mid-latitude (29°-36°N) trends from stations New Delhi, Cairo, and Tateno (centre panel) also indicate similar estimates from both analyses. The largest differences in this latitude zone occurs in layers 6 and 7 where the two trend estimates are significantly different at the 95% confidence level. These differences are due to a combination of the inclusion of a level shift at Tateno in the Reinsel analysis, but not in Newchurch and Yang, and a difference in the population resulting from different data-selection criteria. The northern mid-latitude (40°-52°N) trends (right panel) from stations Boulder, Haute Provence, and Belsk are similar in all layers except the highest layers 8 and 8^+ where the Newchurch and Yang results (-0.7 to -0.8%/year) are more negative than the Reinsel results (-0.4 to -0.5%/year). Most of these differences probably occur because of differences in the beginning time of the time series analysed at Boulder and Belsk (1977 for Reinsel and 1979 for Newchurch and Yang). Also there is a difference due to different aerosol corrections. The Haute Provence trends are similar for both analyses most likely because the beginning of record is the same for both analyses (1984). Including the Arosa results in the Reinsel average does not appreciably change the average values because the Arosa trends are nearly equal to the 3-station mean trends in all layers. Both groups assume that trend estimates for individual stations are uncorrelated and the error bars reflect statistical uncertainty only.

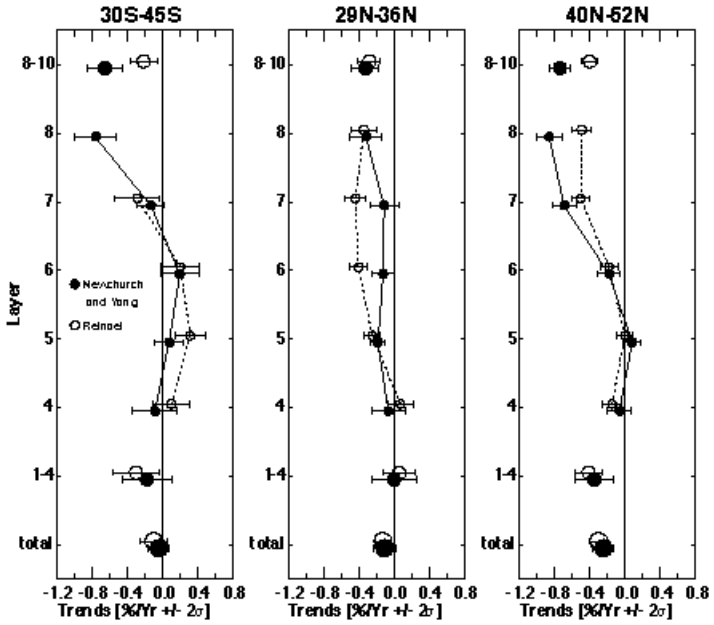


Figure 3.13 Annual ozone trends as a function of altitude deduced from Umkehr data for the following three broad latitude regions: left panel (30°S-45°S) including Perth and Lauder; center panel (29°N-36°N) including New Delhi, Cairo, and Tateno; and right panel (40°N-52°N) including Boulder, Haute Provence, and Belsk. Trends are

variance weighted averages within latitude band with 95% confidence intervals. Results are reported for combined layers 1-4, 8-10, and individual layers 4-8 as indicated for two research groups: Reinsel (open circles) and Newchurch and Yang (closed circles).

3.5.1.3. Altitude profiles of annual trends from Umkehr, SAGE I/II, and SBUV/SBUV2

Applying the statistical models to Dobson/Umkehr, SBUV/SBUV2, and SAGE I/II, time series from 1979 (November 1978 for SBUV/SBUV2) to 1996 in various layers and combinations of layers, while accounting for the potential QBO and solar effects, produces the ozone annual trends in Figure 3.14, as percent per year relative to the mean of the period. For Umkehr observations we compute trends for layers 1-4 combined into a single layer, 4^-, plotted as a vertical bar from ~1 to 22 km with a triangle, for layers 4-8 individually, and for all layers above seven, 8^+, plotted as a vertical bar from ~37 to 54 km. The Umkehr trends in Figure 3.14 are simple averages of the two group analyses presented in Figure 3.13 with root-mean-square 2s error bars. Each individual group average, however, is a variance weighted mean. SAGE I/II trends (diamonds) are computed in natural concentration-versus-altitude coordinates for individual 1-km layers from 15 to 50. However, these are not plotted because they form a smooth curve though the trends computed for 5-km thick layers corresponding to the Umkehr layers.

Click here to download Figure 3.14 by anonymous ftp.

The SAGE I/II results in this chapter include error estimates attributable to the SAGE I altitude correction (for figures including SAGE I data), the sunrise/sunset trend differences (important above layer 7 and discussed below) and the statistical errors. Trends from SBUV/SBUV2 (circles) are calculated from weekly averages (3 daily values required to create weekly average) and are reported for layers 1-4 combined and for layers 5-9 individually.

In the upper stratosphere (i.e., layer 8), trend estimates range from -0.5 ±0.2%/year for SBUV/SBUV2 to -0.8 ±0.2%/year for SAGE I/II. The Dobson/Umkehr trends for layers 8 and 8^+ are both intermediate the SAGE I/II and SBUV/SBUV2 trends and have error bars smaller by a factor of 2. The 2s confidence intervals overlap for the three sensors suggesting that the average layer 8 trend of approximately -0.6 ±0.2%/year is a reasonable estimate of the true atmospheric ozone decline at 40 km.The similarity in the vertical structure of the trends from several instruments provides additional confidence that we are measuring the true atmospheric trend. The largest decline occurs near layer 8 and diminishes to a minimum trend in layer 5 (~25 km) where the trend estimates range from 0.0 ±0.1%/year by Dobson/Umkehr to -0.2±0.2 %/year by SAGE I/II with SBUV/SBUV2 intermediate. These results suggest a trend in layer 5 of -0.1 ±0.1%/year, not significantly different from zero. In layer 9, the SAGE I/II trend is -0.7 ±0.4%/year while the SBUV/SBUV2 estimate is -0.4 ±0.2%/year. The large SAGE I/II error bars at 45-50 km altitude result from uncertainties due to unresolved trend differences between sunrise and sunset observations (described below).

Trends for layers below 5 become more negative in all estimates. The Umkehr trend estimates for layers 4 and 4^- are -0.1 ±0.1%/year and -0.4 ±0.2%/year, respectively. The SAGE I/II estimate for layer 4 is -0.5 ±0.4%/year, which is more negative than that deduced from Umkehr measurements. The same is true for the trend in the combined layers 1-4 where the SBUV/SBUV2 trend is -0.7 ±0.3%/year The

additional uncertainty in the SAGE I altitude registration at tangent heights below 20-25 km increases the uncertainty in the SAGE I/II trends in layers 3 and 4. The trend of -0.4%/year in Umkehr layer 4^- is more consistent with the Dobson total ozone trends of - 0.1 to -0.3%/year calculated by Bojkov *et al.*, (1995) as one would expect from the stringent constraint for Umkehr layers 1-10 to equal the Dobson total column. These results clearly indicate declining ozone in the range of -0.3 ±0.2%/year in layer 4 and of -0.5 ±0.2%/year in stratospheric ozone below approximately 20 km.

Trends calculated from all sensors over the shorter time period of 1979-1989 (corresponding to the lifetime of the Nimbus 7 SBUV) appear in Figure 3.15. The general altitude structure is much the same as calculated for the longer time period. The layer-8 average trend is more negative (-0.8 ±0.3%/year), due to more negative trends from all three sensors. Likewise, layer 9 trends are somewhat more negative. The other layers, however, are not appreciably different. The confidence intervals are generally larger because of the shorter time period.

3.5.1.4. Meridional cross sections of annual trends from SAGE I/II and SBUV/SBUV2

The latitudinal coverage for the SAGE I/II observations extends from 55°S to 55°N. Calculating trends over the period 1979-1996, which comprises the entire SAGE I/II measurement set, results in the altitude-latitude contour plot in Figure 3.16. These trends are based on concentration changes in 1 km altitude layers (the natural coordinates for SAGE). They are referenced to the concentrations at the mid-point of the time series (1987). Typical uncertainties on the annual trends are 0.2 to 0.3%/year (95% confidence limits). A vertical section at 40°-50°N for SAGE I/II is shown in Figure 3.14. In this global view, the maximum ozone trends remain at 40 km (layer 8),

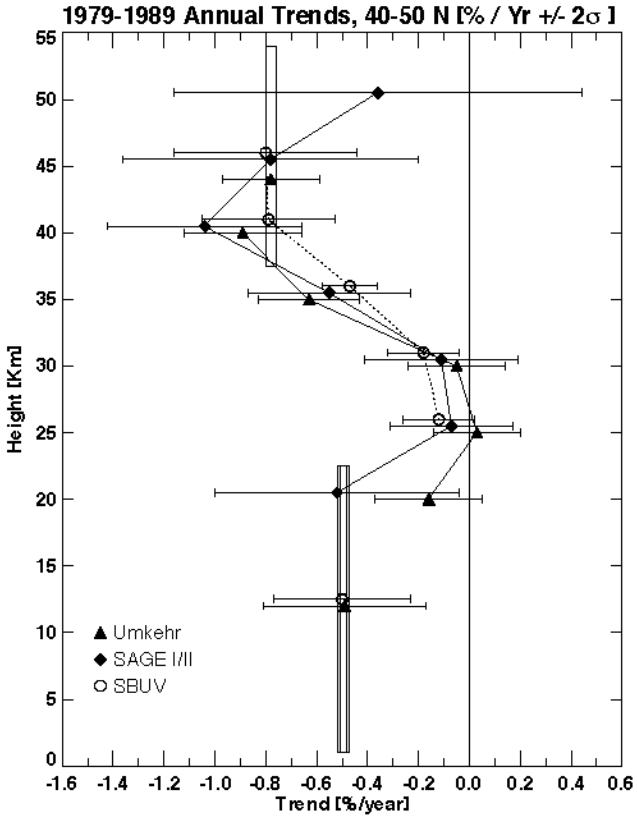


Figure 3.15 Trends as in Figure 3.14 calculated for all sensors over the shorter time period of Nimbus-7 SBUV

(1979-1989).

but are larger in the extra-tropics than in the tropics. The minimum trend at 25 km (layer 5) occurs at all latitudes (with slightly positive trends in the tropics). In general, the SAGE I/II trends are significantly different from zero, at the 95% confidence level (approximately ±0.2%/year), at all latitudes outside of the tropics for altitudes above 30 km (layer 6). The trends exhibit remarkable hemispheric symmetry except below 30 km for latitudes poleward of 40°.

The altitude structure of the trends in layers 4-8 from two southern hemisphere Dobson/Umkehr stations at 32°S and 45°S, Perth and Lauder, respectively (Figure 3.13, left panel) agrees with a section through the 32°-45°S area of the SAGE I/II contour plot; however, the SAGE I/II trends in all layers, but especially in layer 6, are somewhat more negative than the Umkehr estimates. SAGE trends computed with SAGE II data only (1984-1996**,** Figure 3.17), which corresponds more closely to the Perth and Lauder time periods, correspond better in the southern hemisphere and equally well in the northern hemisphere. In the northern sub-tropics, three Umkehr stations are located between 19°N and 24°N, Poona, Mauna Loa, and Aswan, however, because of unresolved data continuity problems at each of these stations they are not used for trend calculations at this time. Three Umkehr stations between 29°N and 36°N, New Delhi, Cairo, and Tateno, exhibit, in Figure 3.13, (centre panel), average trends 0.2 to 0.4%/year less negative than the SAGE I/II estimates in the upper layers, but agree well in the lower layers. The correspondence between the northern mid-latitude Dobson/Umkehr (Figure 3.13, right panel), SAGE I/II, and SBUV/SBUV2 was discussed in the previous section.

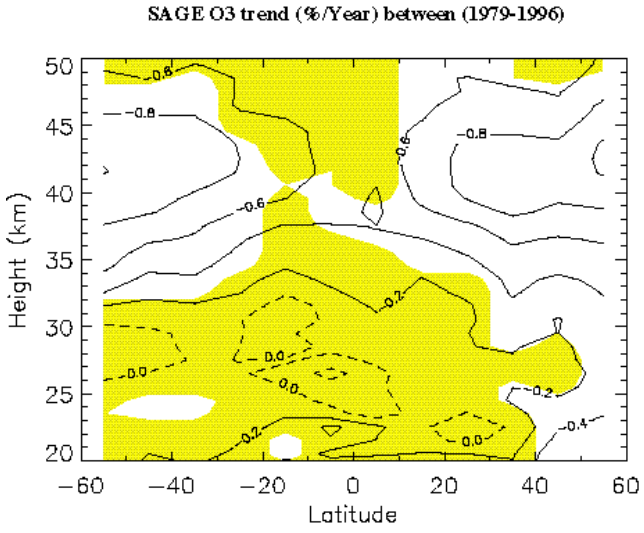


Figure 3.16 Annual ozone trends calculated from SAGE I/II observations from 1979 to 1996 expressed in %/year of the mid-point of the time series (1987). Results are contoured from calculations done in 5° latitude bands and 1 km altitude intervals. Contours differ by 0.2%/year with the dashed contours indicating zero or positive trend. The shaded area indicates where the trends do not differ from zero within 95% confidence limits. The estimate of uncertainty contains terms due to the SAGE I reference height correction and the SAGE II sunrise/sunset trend differences.

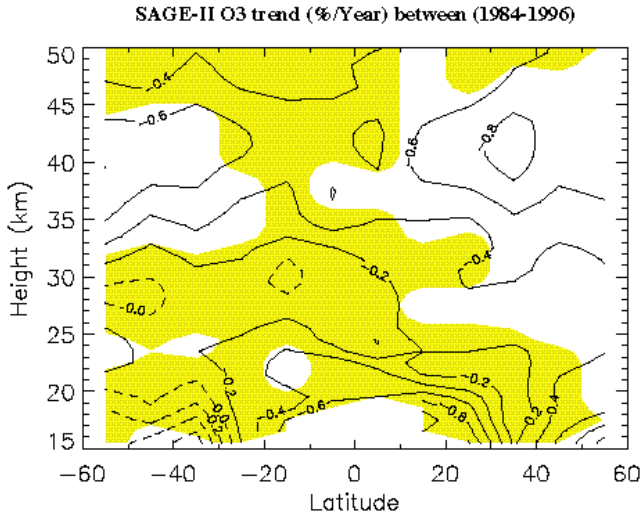


Figure 3.17 Annual ozone trends as in Figure 3.16 for SAGE II only for the time period 1984 to 1996.

The SBUV/SBUV2 trends over November 1978 to October 1994 in percent per year relative to the mean ozone amount from the combined time series at each latitude and layer appear in Figure 3.18. The dark shading indicates the regions in which the derived trend is not significantly different from zero at the 2s level calculated from statistical error bars only; inclusion of the instrument error gives no statistically significant trends between 35°S and 45°N at the 2s level. The lighter shading indicates regions in which the length of the time series during the winter season is compromised as a result of the NOAA-11 orbit drift and where the true annual trend over this time period cannot be calculated. In the northern hemisphere, the data loss is not as extensive as in the southern hemisphere; therefore, the northern high-latitude trends are much more representative of the true annual trends over this time period. On the other hand, data loss in the southern hemisphere high latitudes begins as early as 1990, such that the trends plotted here are actually an average of the spring, summer and autumn trends through 1994 and the winter trend through 1990. The largest ozone losses occur during winter in the profile data through 1990, leading to a substantial bias in the mid to high-latitude trend results in the southern hemisphere over this time period. Comparing the SAGE I/II trends in Figure 3.16 to the SBUV/SBUV2 trends in Figure 3.18 indicates similar small trends of 0.0 to -0.3%/year in layers 5 and 6. In the upper stratosphere, the altitudinal and latitudinal structures of the SAGE I/II and SBUV/SBUV2 trends are similar; however, the SAGE I/II trends are substantially more negative at all latitudes.

In the upper stratosphere (layers 7-9), trends calculated from SBUV alone (1978-1989), as seen in Figure 3.19, exhibit altitude/latitude structure similar to SAGE I/II trends calculated over the same period (Figure 3.20) but the SBUV values remain less negative than the SAGE I/II values. The SBUV trends in layers 8 and above are approximately 0.2%/year less negative than SAGE I/II in both hemispheres. Note that the SBUV latitudinal coverage (±65°) is somewhat greater than the SAGE I/II coverage (±55°). In the lower stratosphere, on the other hand, while the SBUV and SAGE I/II trend structures remain similar, the SBUV trends are more negative than the SAGE I/II estimates by approximately 0.3%/year The SBUV trends are negative or zero throughout the stratosphere, while the SAGE I/II trends are actually positive (e.g., a maximum of +0.4%/year in the tropics) in layer 5. These SAGE I/II trends, however, are statistically insignificant at the 95% confidence level.

Focusing on SAGE trends over three overlapping periods of the 17-year record, we see that trends in the first 11 years (Figure 3.20, from SAGE I/II combined 1979-1989) are more negative in the upper stratosphere and less negative in the middle stratosphere than trends in the last 12 years (Figure 3.17, SAGE II only 1984-1996). The less negative trends seen in the upper stratosphere during the later period

result primarily from the influence of relatively small ozone declines in the middle years ~1984-1992 followed by greater declines after 1992 (cf. Figure 3.11). Middle stratospheric trends have become more severe. In fact, the middle stratospheric trends (20-25 km) changed from 0.2%/year *increases* in the first period to 0.2%/year *decreases* in the last period. Trends over the 17 year period (Figure 3.16) are intermediate to the estimates for the first 11 years and for the last 12 years.

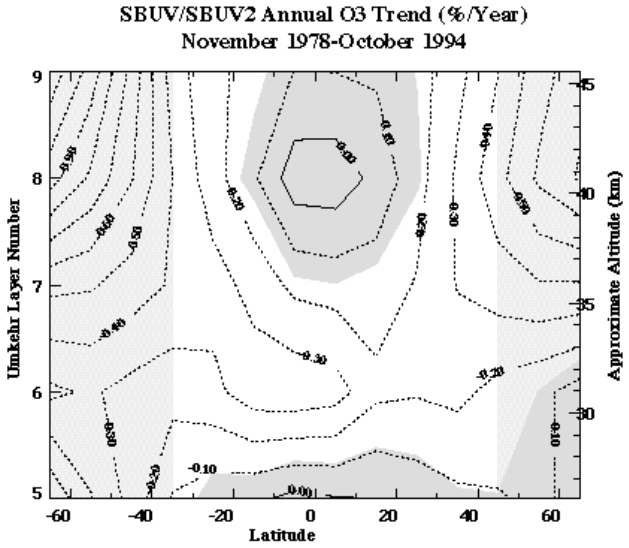


Figure 3.18 Annual SBUV/SBUV2 trend calculated from November 1978 through October 1994 as a function of latitude and Umkehr layer. The trend is in percent per year relative to the mean ozone amount from the combined time series at each latitude and layer. The dark shading indicates the regions in which the derived trend is not significantly different from zero at the 2s level calculated from statistical error bars only; inclusion of the instrument error renders no statistically significant trends between 35°S and 45°N. The lighter shading indicates regions in which the length of the time series during the winter season is compromised as a result of the NOAA-11 orbit drift.

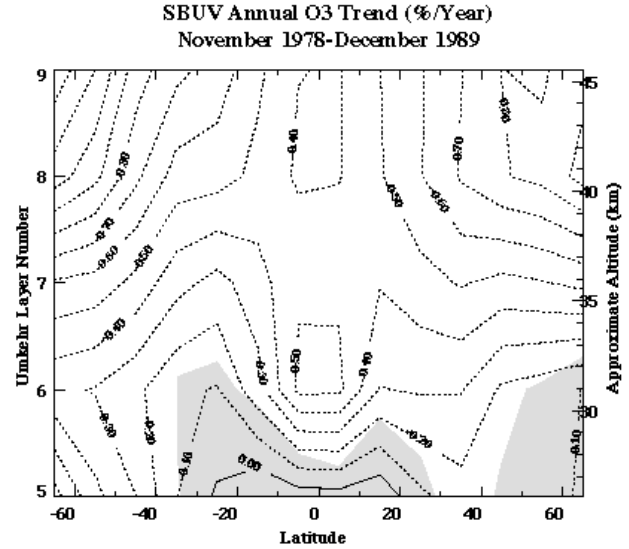


Figure 3.19 The annual SBUV trend calculated from November, 1978 through December, 1989 as a function of latitude and Umkehr layer. The trend is in percent per year relative to the mean ozone amount over this period at

each latitude and layer. The contour interval is 0.1 percent per year. The shading indicates the regions in which the derived trend is not significantly different from zero at the 2s level. The trend uncertainty is calculated from the statistical errors only.

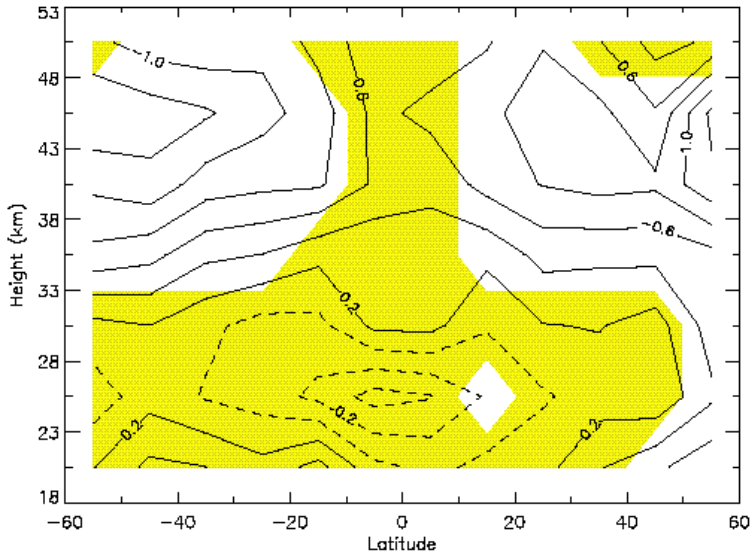


Figure 3.20 Altitude/latitude contours of ozone trends (%/year) calculated from SAGE I/II sunrise and sunset observations (1979-1990) corresponding approximately to the SBUV-only time period shown in Figure 3.19. Shading indicates 95% confidence intervals that include uncertainties due to SAGE I altitude correction, SAGE II sunrise/sunset differences, and statistical uncertainties.

3.5.1.5. Seasonal variation of the SAGE I/II and SBUV/SBUV2 trends

Investigating the trends in layer 8 as a function of latitude and month, we find that both SAGE I/II (Figure 3.21) and SBUV/SBUV2 (Figure 3.22) indicate roughly a factor of 3 seasonal variation in the mid-latitude ozone trends with the most negative trends occurring in the winter (between November and February in the north, June and July in the south). Figure 3.23 shows the SAGE I/II trends as a function of altitude and latitude for individual 3 month periods. The winter maximum trend appears clearly at 40-45 km, with the trend begin strongest during southern hemisphere winter. This seasonal pattern seen in the satellite data is consistent with results from the Dobson/Umkehr analyses, although the seasonal variation of the Umkehr trends is somewhat smaller. SAGE I/II results indicate more seasonal variation in the tropics (0.0 to -0.6%year) than does SBUV/SBUV2 (0.0 to +0.2%/year) and also significantly more negative trends as noted earlier in the annual mean results. The heavy black line surrounding polar night in Figure 3.22 designates the regions with no SBUV2 data because of the drifting NOAA-11 orbit; the thin black line designates the regions with no SBUV2 data after 1992 due to continuing drift of the orbit. In 1994, there are no data in the southern hemisphere outside of the tropics after about April. The structure and magnitudes of the seasonal variation in the southern hemisphere are similar for both SAGE I/II and SBUV/SBUV2, although the area of low trends propagating out of the austral winter extends further equatorward and later in the austral springtime in the SAGE I/II estimates. Note that the SBUV/SBUV2 latitudinal coverage is somewhat larger than that of SAGE I/II as reflected in the latitude-scale differences.

In the high latitudes of the northern hemisphere, the most negative trend in the SAGE I/II data occurs in the late winter while the most negative trend in SBUV/SBUV2 data occurs 2 months earlier in the early winter. Analysing only the NIMBUS-7 SBUV data (i.e., over the shorter period of 1978-1989) in the trend calculation (Figure 3.24) increases the magnitude of the tropical trends markedly to ~-0.3 to -0.6%/year (also seen in Figure 3.26) but has less effect on the seasonal structure at higher latitudes.

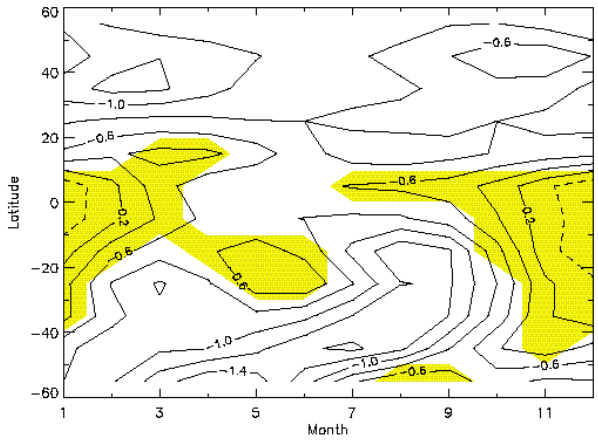


Figure 3.21 Seasonal variation of ozone trends in layer 8 (%/year) calculated from SAGE I/II (1979-1996) for latitudes 55°S to 55°N. Shading indicates 95% confidence intervals that include uncertainties due to the SAGE I altitude correction, SAGE II sunrise/sunset differences, and statistical uncertainties.

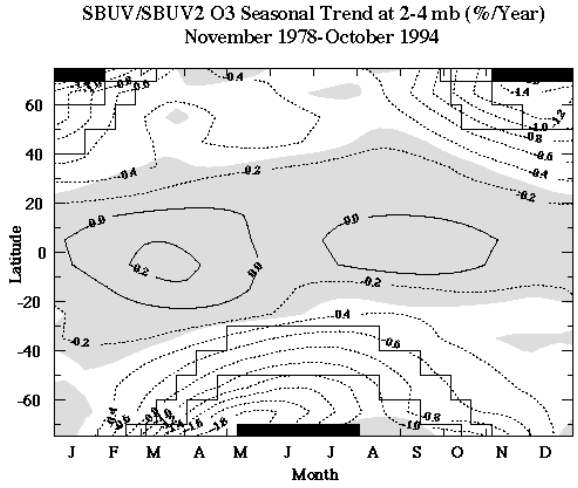


Figure 3.22 Seasonal variation of ozone trends in layer 8 (%/year) calculated from SBUV (version 6) and SBUV2(version 6.1.2) (1978-1994). The shading indicates 95% confidence intervals from statistical errors only. The solid black shading indicates regions of no data due to polar night. The heavy black line surrounding polar night designates the regions with no SBUV2 data because of the drifting NOAA-11 orbit; the thin black line designates the regions with no SBUV2 data after 1992 due to continuing drift of the orbit. In 1994 there are no data in the southern hemisphere outside of the tropics after about April.

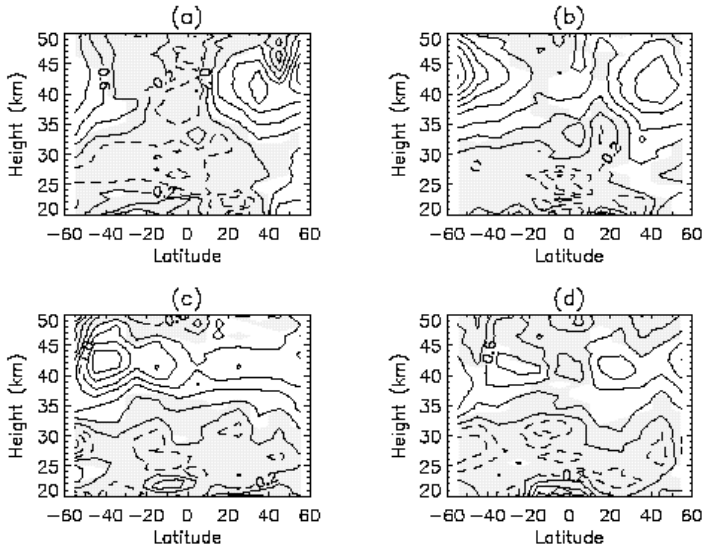


Figure 3.23 SAGE I/II trends as a function of altitude and latitude as in Figure 3.16, but for individual 3 month periods consisting of a) December, January, February; b) March, April, May; c) June, July, August; and d) September, October, November. The statistical model used was of standard form except that the trend term used two seasonal harmonics.

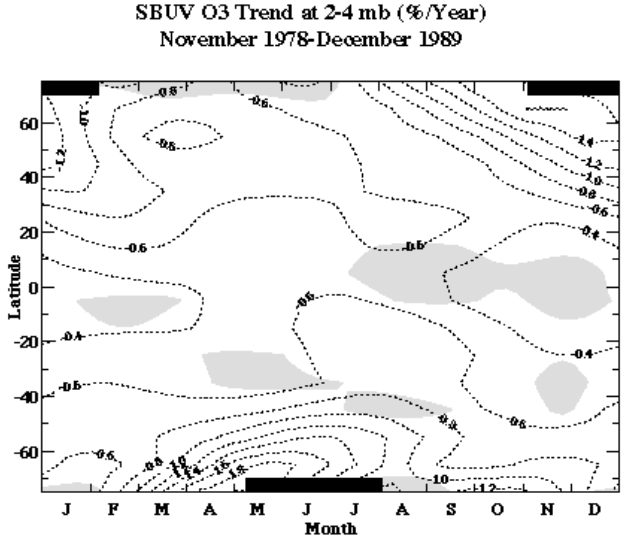


Figure 3.24 Seasonal variation of ozone trends in layer 8 (%/year) calculated from SBUV (version 6) for November 1978 to December 1989 for latitudes 75°S to 75°N. The trend is relative to the seasonal mean (i.e. January mean for the January trend, etc..) ozone amount from the combined layer 8 time series at each 10-degree latitude band. The shading indicates trends that are not significant at the 95% confidence level from statistical errors only. The solid black shading indicates regions of no data due to polar night.

3.5.1.6. Hemispheric symmetry of trend at 40 km (layer 8) from SAGE and SBUV/SBUV2

In order to test the inter-hemispheric symmetry in the ozone trend, we focus on layer 8 and inspect the ozone trends as a function of latitude. Figure 3.25 shows the layer 8 SAGE I/II trends as a function of latitude for the first 11 years (1979-1990, diamonds), the last 12 years (1984-1996, triangles), and all 17

years (1979-1996, solid line). In none of the three periods is an inter-hemispheric difference in the trends evident. The analogous results for SBUV/SBUV2 and only N7 SBUV appear in Figure 3.26. Note that data within ±40°N should not be affected directly by the NOAA-11 drifting orbit. Although minor inter-hemispheric differences exist in the mean results, these differences are statistically insignificant. In layer 9, the N7 SBUV trend asymmetry is ~0.5%/year, but is still not statistically significant due to large inter-annual variability at high latitudes.

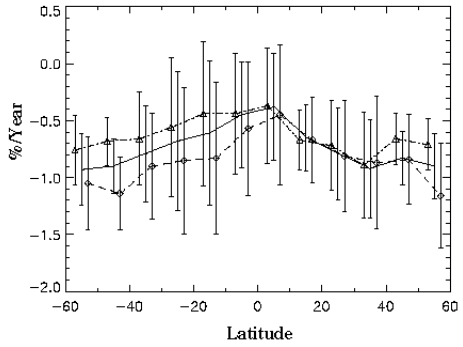


Figure 3.25 SAGE I/II ozone trends at 40 km (layer 8) for latitudes 55°S to 55°N for the following three time periods: The first 11 years (1979-1990, diamonds and dotted line), the last 12 years (1984-1996, triangles and dot-dash line), and all 17 years (1979-1996, no symbols, solid line). Error bars represent 95% confidence intervals that include uncertainties in the SAGE I altitude correction, SAGE II sunrise/sunset differences, and statistical variations.

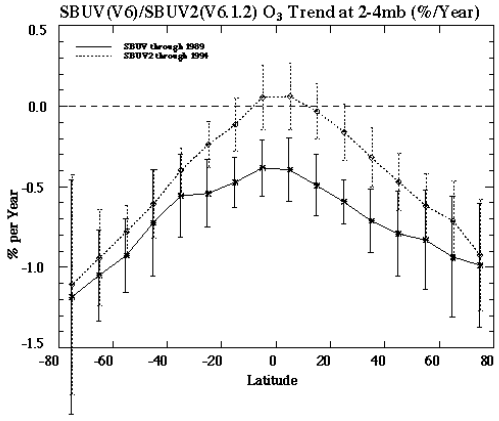


Figure 3.26 SBUV/SBUV2 2-4 hPa (layer 8) ozone trends for latitudes 75°S to 75°N for the period 1979-1994 (dotted line) and for 1979-1989 (solid line). Error bars represent 95% statistical confidence intervals.

3.5.1.7. Sunrise/Sunset differences in SAGE trends

Trends computed exclusively from either sunrise or sunset SAGE observations result in different values above 35 km (layer 7) as seen in Figure 3.27, which shows the difference between sunrise and sunset (sunrise-sunset) trends as functions of altitude and latitude over the period 1979-1996. In layer 8 (40 km), for example the sunrise trends are approximately 0.2-0.4%/year more negative than the sunset trends with some latitudinal variation. In layer 9, differences reach 0.6%/year. Calculating sunrise-sunset differences over only the SAGE II period results in somewhat different latitudinal dependence of the

trend differences, but supports the general conclusion that layer 8 indicates 0.3%/year differences and layer 9 experiences 0.6%/year sr/ss differences at some latitudes.

There is some indication that the version 5.96 SAGE II data used for these analyses may exhibit a larger sunrise/sunset discrepancy than the previous version 5.93 SAGE II data; although the version 5.93 trends (Wang *et al.*, 1996) are approximately the same as the version 5.96 trends. There is insufficient independent evidence to prefer one set of occultation measurements to the other (e.g., sunset to sunrise). Additionally, because SAGE I returned sunrise measurements only during the first three months, limiting the SAGE time series to only sunrise or only sunset observations introduces an undesirable sampling bias. We note 1) that the SAGE sunrise-sunset differences are insignificant below layer 8, 2) that the differences in layer 8 are of the same order as the 2s statistical trend uncertainties, and 3) the differences in layer 9 are significantly larger than the trend uncertainties.

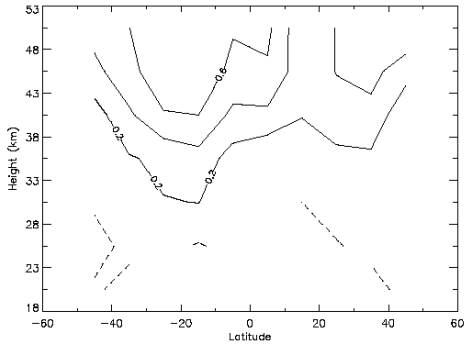


Figure 3.27 Difference between the SAGE I/II ozone trends for sunrise and sunset measurements (%/year) as a function of altitude and latitude over the period 1979-1996.

3.5.2. Lower stratosphere

3.5.2.1. Sondes

Time series

Monthly mean values of ozone near 90 hPa are shown in Figure 3.28a for three stations, using the Harvard selection criteria. The right hand panels show the time series of monthly anomalies (de-seasonalised monthly means). Figure 3.28b shows the time series using the Chicago selection criteria. Measurements are made 2-3 times a week at the European stations and weekly in Canada; at the Japanese stations measurements were made weekly at the beginning and end of the time series, with infrequent data and few summer measurements in between. The frequency of measurements is reflected in the variability in the monthly values (Figure 3.28a). The effect of the stricter Chicago criteria is to introduce more gaps in the time series, and to increase the variance in the monthly anomalies. Time series up to ~1994 for all sonde stations (WOUDC data) are shown in Logan (1994).

An example of the fit of the statistical model to the sonde data is shown for a level near 20 km in Figure 3.29. The top panel shows the monthly mean values, the middle panel shows the fit to the data after removing the seasonal cycle determined by the statistical model, and the bottom panel shows the residuals of the fit for the model.

Uccle, Hohenpeissenberg, and Payerne are the closest together of the sonde locations. There do not appear to be any major biases in ozone values near 90 hPa, except for the early 1970s where values at Hohenpeissenberg and Payerne tend to exceed those at Uccle (Figure 3.30).

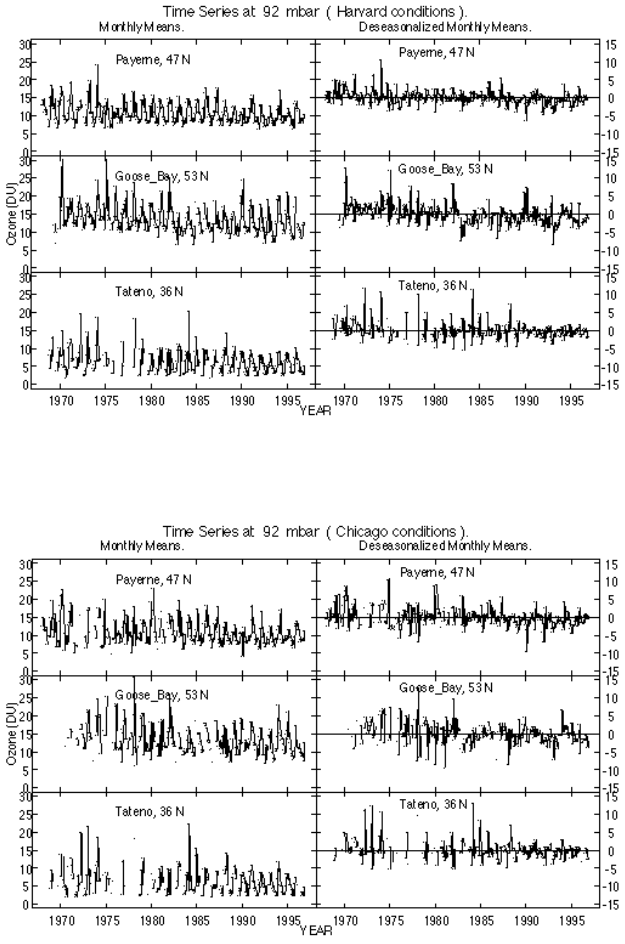


Figure 3.28 Time series of monthly mean values for ozone in DU and deseasonalized monthly means for selected stations. The data selection criteria used by a) the Harvard group and b) the Chicago group, were applied (see text). Values are shown for one of the 33 levels near 90 mbar, and the same relative scale is used for both sets of means.

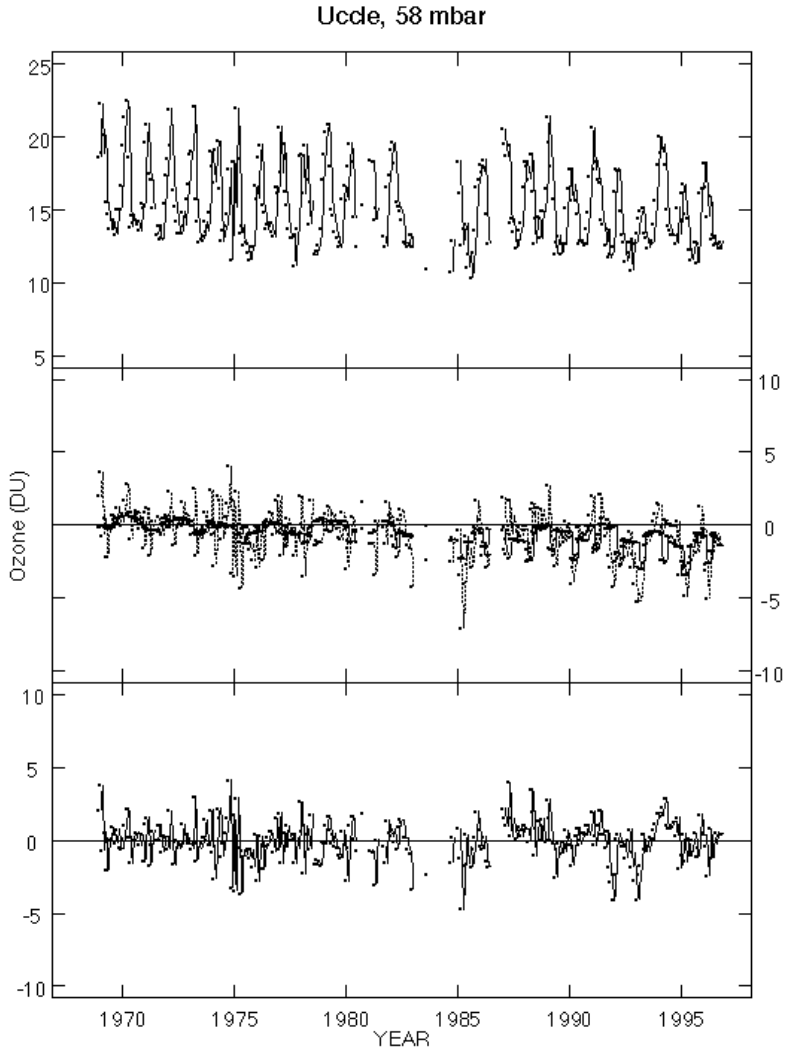


Figure 3.29 Trend analysis of Uccle near 58 mb (~20 km), Harvard results. The top panel shows the monthly mean time series. The middle panel shows the times series with the seasonal component of the model fit removed (dotted line), and the model fit to this (solid line). The lower panel shows the residuals of the complete model fit to the time series.

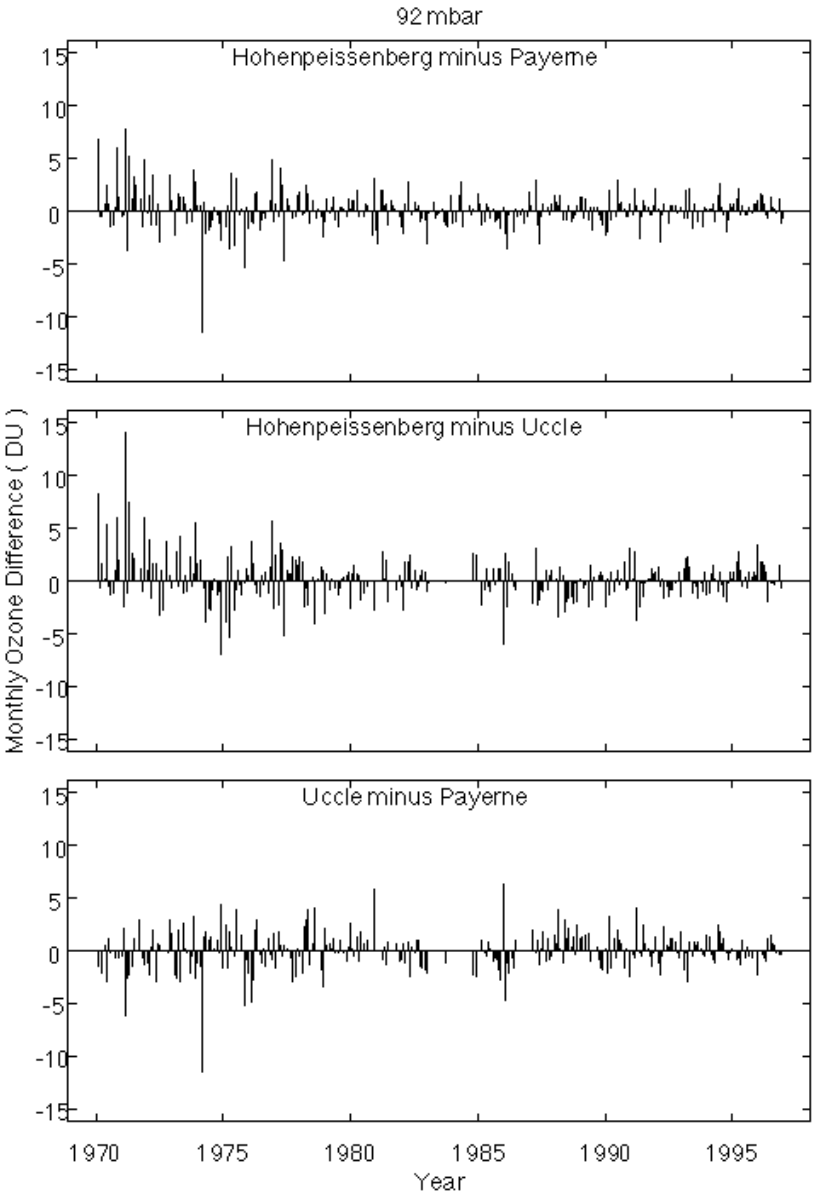


Figure 3.30 Difference of monthly mean values for the three European stations near 90 hPa. The correction factors used by the Harvard group were applied (see text).

Trends for 1970-96

Seasonal trends in the vertical distribution of ozone are shown in Figure 3.31a,b,c for the 11 sonde stations for 1970-96, in percent per decade; the solid lines are those calculated by the Harvard group and the dashed lines those from Chicago. Annual trends are shown in Figure 3.32. The sonde data are expected to be reliable for trend determination up to 27 km, according to the recommendations in Chapter 1. Results are shown for pressure levels centred below 13 hPa, ~29 km, so results for the top level should be considered less reliable than for the other levels.

The seasonal trends determined by the two groups usually agree within ± 3%/decade for most stations, and almost all trends agree within their standard errors in the stratosphere. Agreement is worst for Churchill, with differences of ± 10%/decade in summer and winter; this station suffers from serious data loss with the Chicago data selection criteria. The annual mean trends (Figure 3.32a) agree within 2%/decade in the stratosphere, with the exception of results from Uccle and Payerne.

The annual trends are shown in absolute units (DU/km/decade) in Figure 3.32b. This comparison avoids the problem of the percentage trends being referenced to the mean at different times; pressure was converted to altitude using the U.S. Standard Atmosphere. This figure, like Figure 3.32a, shows that the two groups obtain similar results for the ozone trends except for Payerne and Uccle above 50 hPa. The dashed line shows the absolute trends computed with the Harvard statistical model, but with the data treatment of the Chicago group; their CF criteria, the requirement that the balloon reaches 16 hPa and that there is a measurement of the ozone column, and with the data divided by the correction factor. The results from the two groups are similar at most altitudes at most stations when they use the same data treatment, and the discrepancies at Uccle and Payerne are removed.

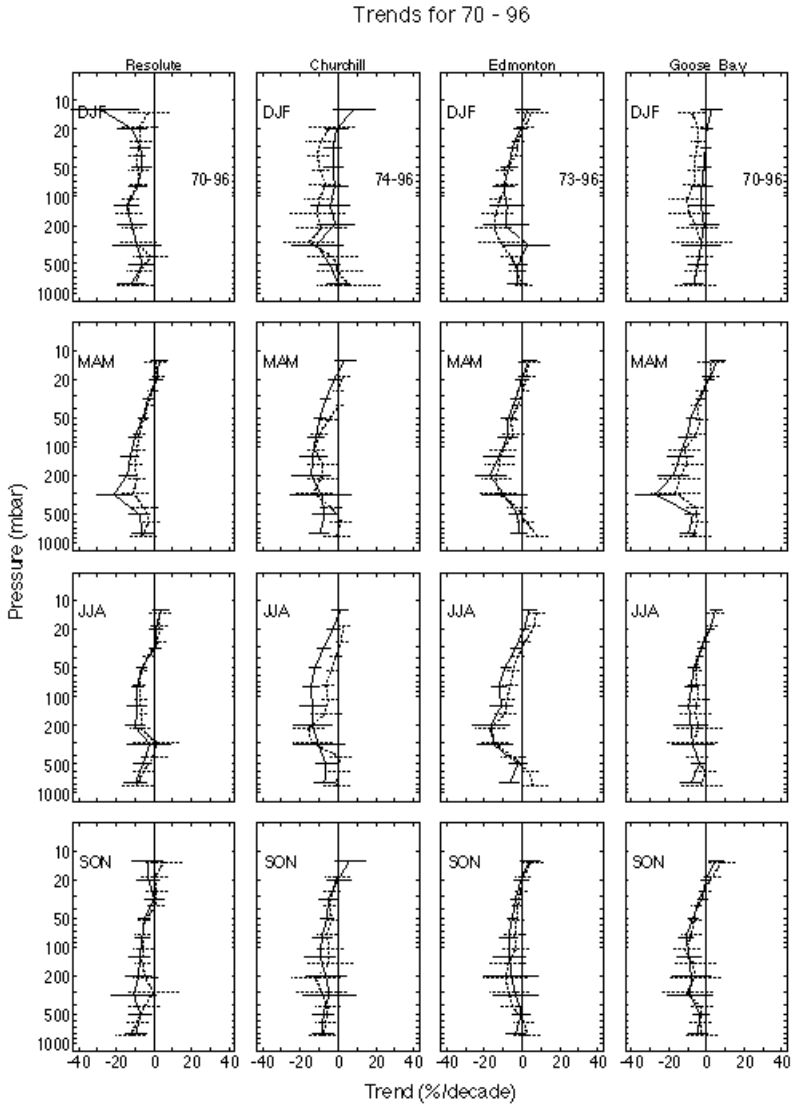


Figure 3.31 a) Seasonal trends in the vertical distribution of ozone for 1970-96 at Resolute (75°N, 95°W), Churchill (59°N, 94°W), Edmonton (53°N, 114°W), and Goose Bay (53°N, 60°W). The Harvard results are shown by the solid line and the Chicago results by the dashed line. Two standard errors for the statistical fit are shown.

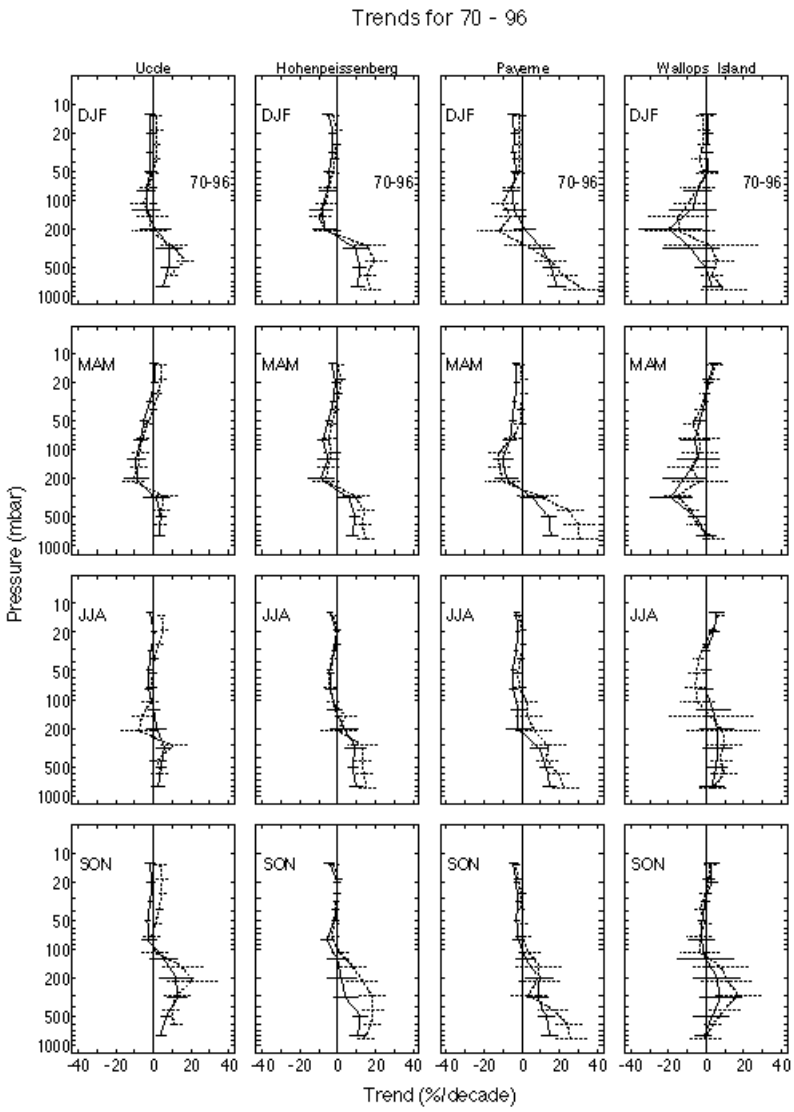


Figure 3.31 b) Seasonal trends in the vertical distribution of ozone for 1970-96 at Uccle (52°N, 4°E), Hohenpeissenberg (48°N, 11°E), Payerne (47°N, 7°E), and Wallops Island (38°N, 75°W). The Harvard results are shown by the solid line and the Chicago results by the dashed line. Two standard errors for the statistical fit are shown.

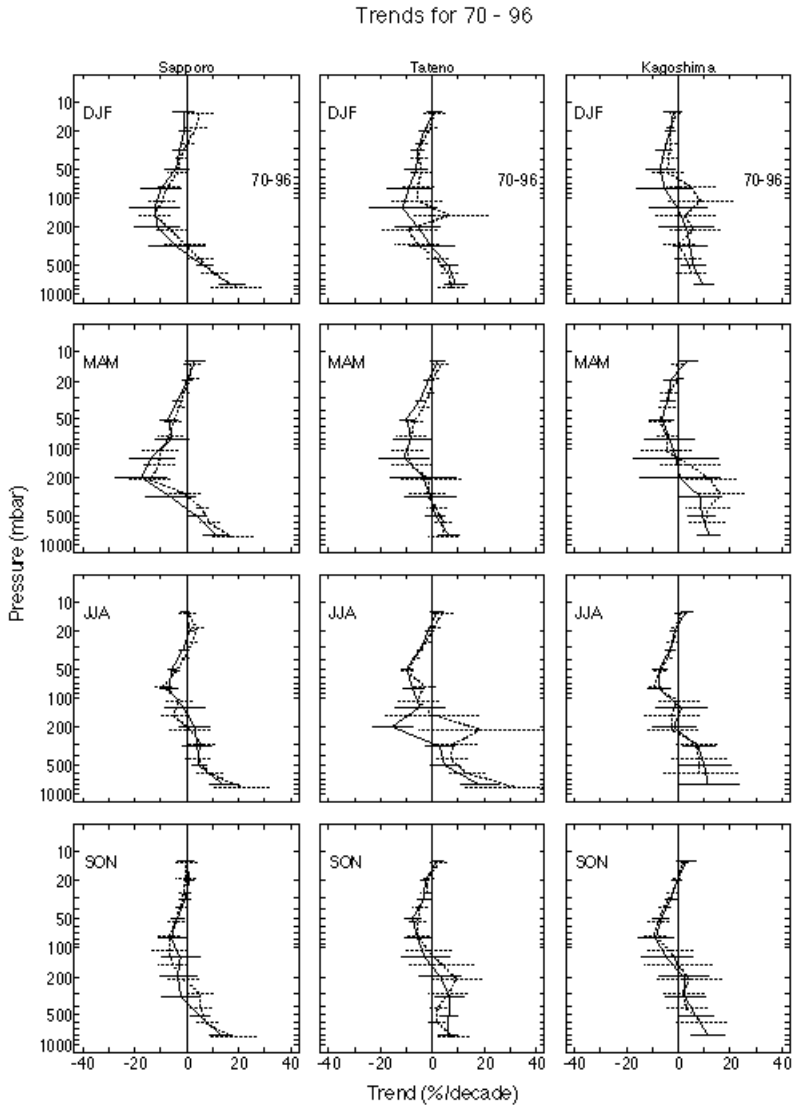


Figure 3.31 c) Seasonal trends in the vertical distribution of ozone for 1970-96 at Sapporo (43°N, 141°E), Tateno (36°N, 140°E), and Kagoshima (32°N, 140°E). The Harvard results are shown by the solid line and the Chicago results by the dashed line. Two standard errors for the statistical fit are shown.

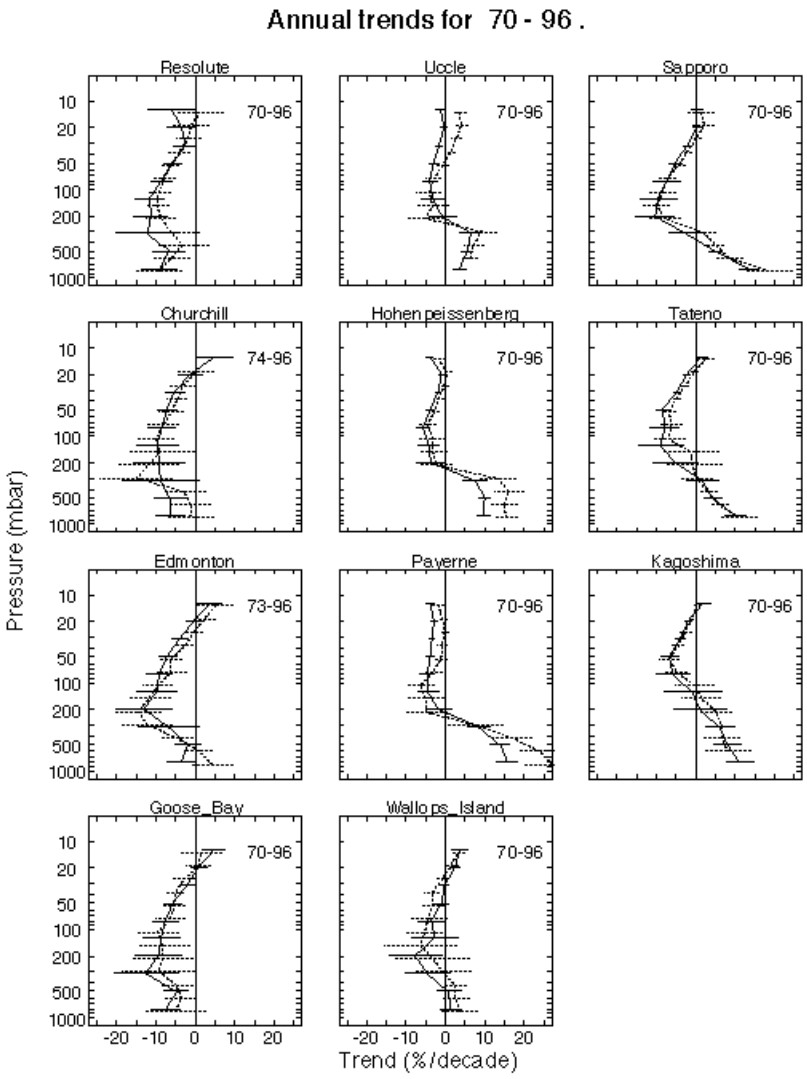


Figure 3.32 a) Annual trends in the vertical distribution of ozone in %/year for 1970-96. The Harvard results are shown by the solid line and the Chicago results by the dashed line. Two standard errors are shown.

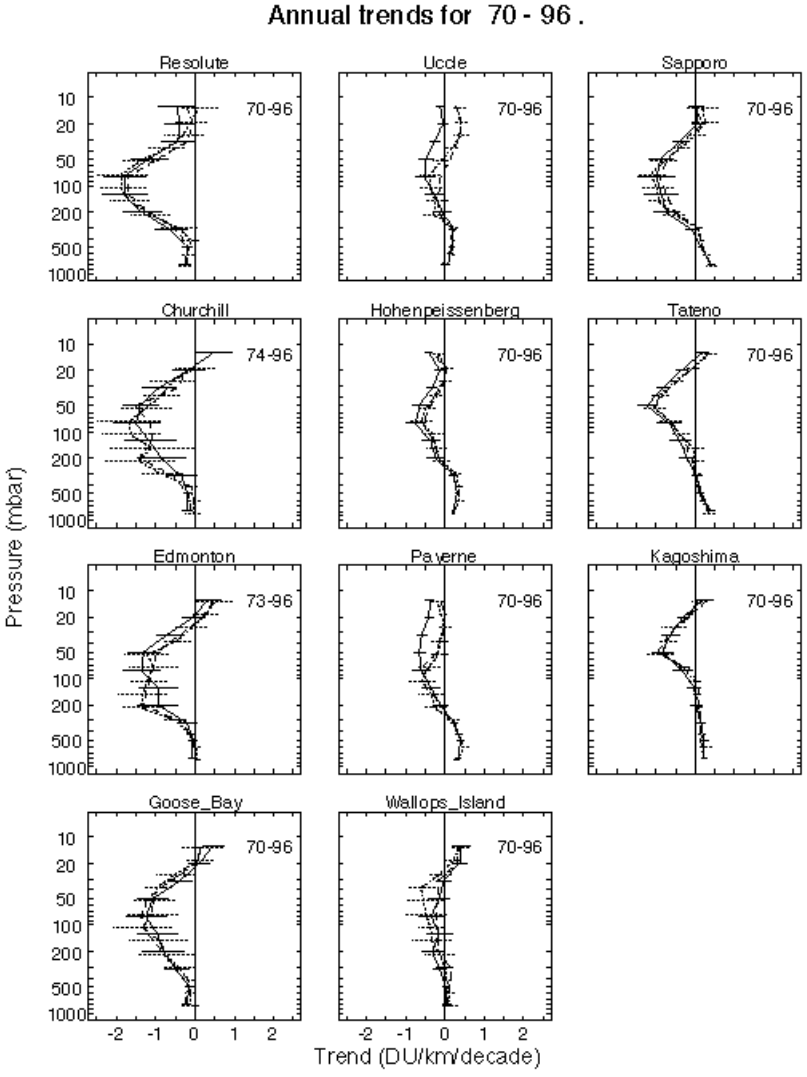


Figure 3.32 b) Annual trends in the vertical distribution of ozone for 1970-96 in DU/km/decade. The Harvard results are shown by the solid line and the Chicago results by the dotted line. Two standard errors are shown. The dashed line shows trends derived with the Harvard model run with the Chicago data treatment (CF criteria, balloon reaching 16 mbar, an ozone column measurement, and the data divided by the CF).

In order to separate the effects of data selection and of dividing by the correction factor, trends were calculated with the Harvard model using the Chicago data selection criteria, with and without removing the scaling to the ozone column. Results are shown in Figure 3.33 (in percent per decade) where the difference between the solid line and the dashed line is caused by data selection criteria, and the difference between the dashed line and the dotted line is caused by dividing by the correction factor.

For stations with no trend in the correction factor, Sapporo and Kagoshima, the different trends in Figure 3.33 are due primarily to the data selection criteria, while for Hohenpeissenberg, where the selection criteria were almost the same, the different trends are due primarily to the trend in the correction factors (see Table 3.2). The effect on trends of dividing by the correction factor is largest (2-3%/decade) for Uccle, Payerne, and Wallops Island, the stations with the largest trend in the correction factors (Table 3.3); for Uccle, the CFs show a distinct jump in 1989 when a new pump correction factor was applied (see Chapter 1). For the remaining stations, the effect of dividing by the correction factor is <2%/decade. Changing the data selection criteria sometimes changes the shape of the trend vertical

profile, while dividing by the CF generally shifts the profile to larger or smaller trends depending on the magnitude of the trend in the CF.

Figure 3.34 shows the Harvard and Chicago results as the column trend (in DU per decade) for 250 to 16 hPa; this comparison also avoids the problems of the absolute trends being computed on different layer thicknesses and percentage trends being referenced to the mean at different times. The Harvard standard model generally gives more negative trends for stratospheric ozone than the Chicago model; however, when the Harvard model is run with the data treatment of the Chicago group, there is much less difference compared to the Chicago results (as shown also in Figure 3.32b). The results in Figure 3.32b and 3.34 imply that the treatment of data prior to trend analysis generally has more effect on derived trends than the details of the statistical model. The inclusion of autocorrelation in the Chicago model gives standard errors that appear only slightly larger than those in the Harvard model (see Figure 3.32a for Hohenpeissenberg, where the data selection criteria were almost the same). Once the Chicago data treatment is used in the Harvard model, the remaining causes for differences are the assumption of no trend prior to 1970, the inclusion of autocorrelation in the Chicago model, and the removal of outliers in the Chicago model.

Trends in ozone are largest in the lower stratosphere, with maximum annual trends of -3 to -10%/decade (Figure 3.32). There are statistically significant decreases in ozone at all the stations analysed here. The decrease in ozone is located between about 30 hPa and the tropopause. These results are similar to those shown in Logan (1994), Miller *et al*. (1995), WMO (1995), and Harris *et al*. (1997) for analyses of WOUDC data that ended a few years earlier. Bojkov and Fioletov (1997) analysed the sonde data at WOUDC relative to the tropopause height, and found significant decreases only 1-2 km above the tropopause. It is important to note that Bojkov and Fioletov (1997) found that the use of altitude relative to tropopause height as a vertical coordinate significantly reduced the variability in the data set near the tropopause. Analysis with respect to the tropopause was not done in the present study, and should be done in the near future with the reanalysed data sets of this study.

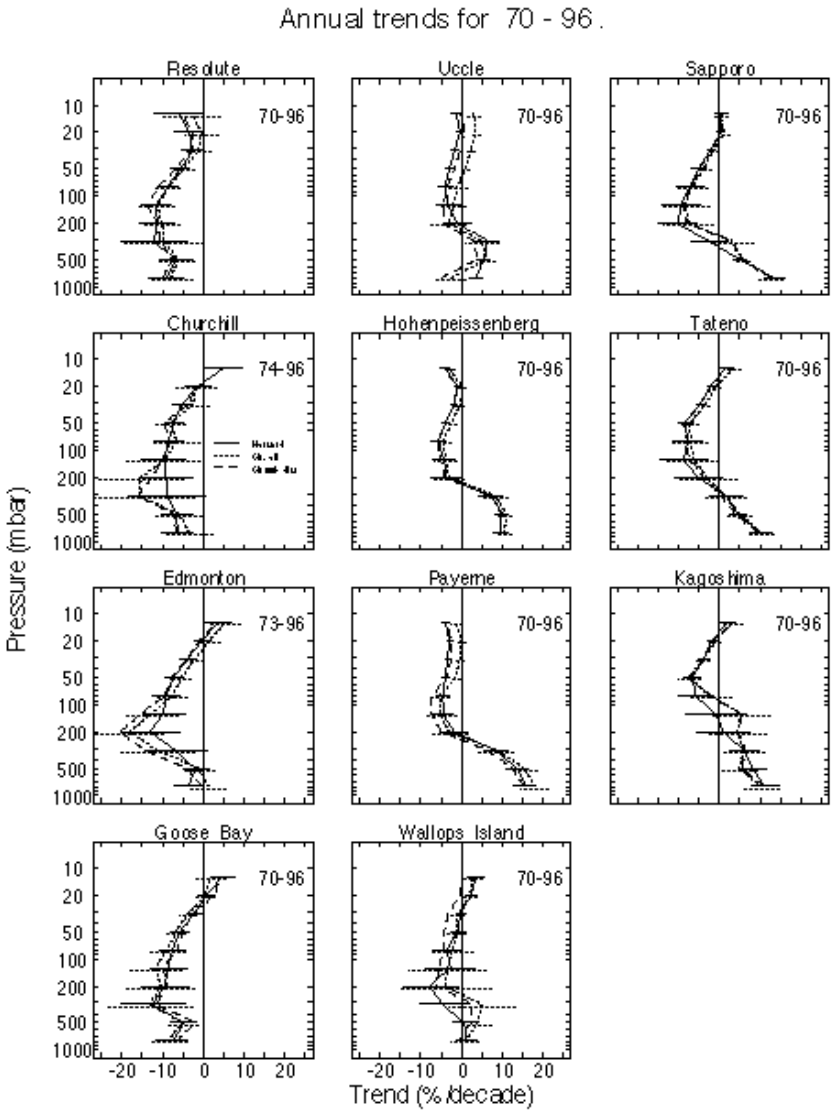


Figure 3.33 Sensitivity of annual trends in the vertical distribution of ozone to data treatment prior to trend analysis. All trends were derived with the Harvard model. The solid line shows results for the Harvard data selection criteria; the dotted line shows results with the Chicago data treatment (their CF criteria, the balloon reaching 16 mbar, and an ozone column measurement, and with the data divided by the CF); the dashed line shows results with the Chicago data selection criteria, but without dividing by the CF.

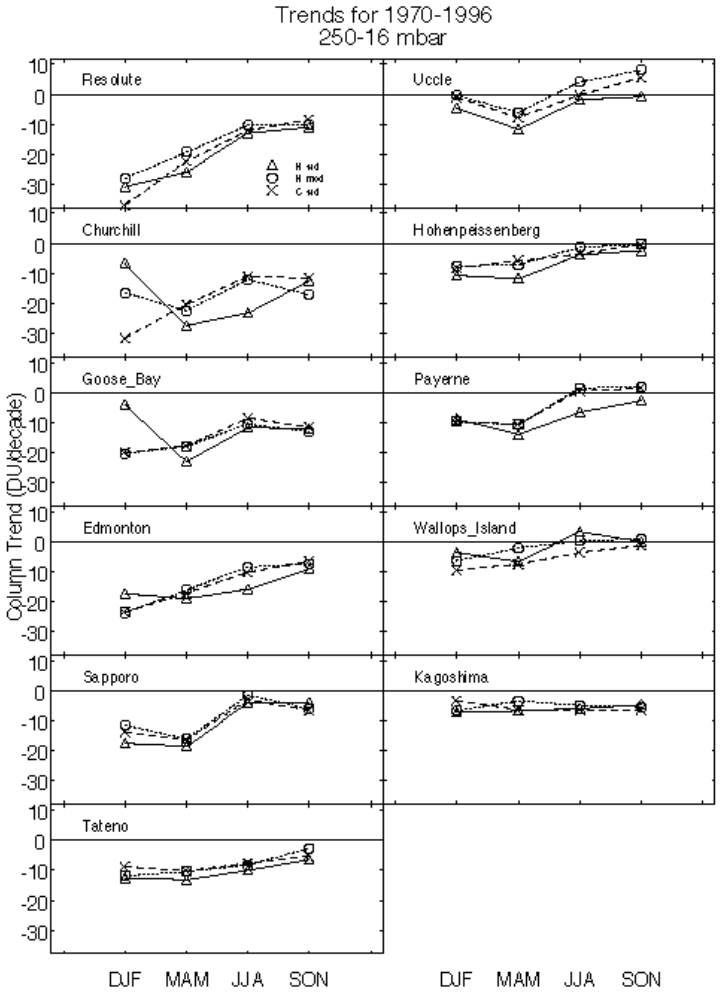


Figure 3.34 Column trend in ozone in DU/decade from 250 to 16 mbar for 1970-96. The triangles are results from the Harvard group, the crosses those from the Chicago group, and the circles are results for the Harvard model, with three conditions of the Chicago group used in data selection, their CF criteria, the balloon reaching 16 mbar, and an ozone column measurement, and with the data divided by the CF.

The annual trends for sonde stations located between 36°N and 59°N are compared in Figure 3.35. The various stations form a reasonably compact band of trends from 125 to 30 hPa, with Wallops Island (Harvard results) and Uccle (Chicago results) showing somewhat less negative trends above 70 hPa than the other stations; the Uccle trend derived by the Chicago group has been influenced by the step function in the correction factors in 1989 (see Chapter 1). The mean trend in ozone for the sonde stations located from 36°N to 59°N is shown in Figure 3.36. Both groups find a mean trend of -7%/decade near 100 hPa. From 80 to 15 hPa, the mean Chicago trend is about 1%/decade less negative than the mean Harvard trend. Referencing the trend to 1970 (Chicago) rather than the mean over 1970-96 (Harvard) may give less negative trends. Both negative and positive trends are seen near 13 hPa, but these results are less reliable than those at 20 hPa and below. The error shown for each layer is the standard error of the annual trends. This error was larger for layers than the root mean square error of the annual trend errors for the individual stations.

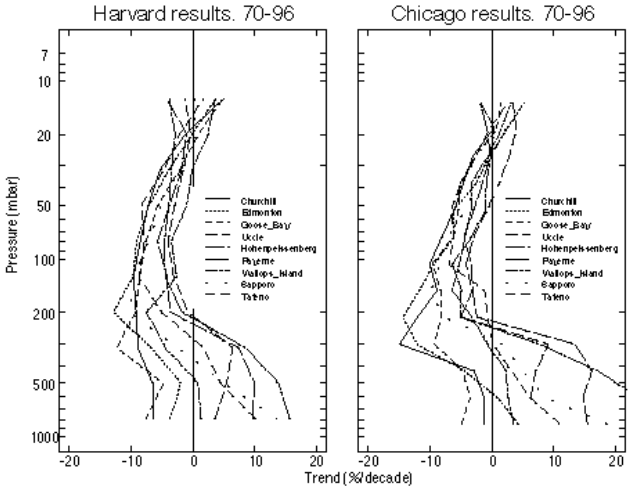


Figure 3.35 Annual trends for individual sonde stations located between 59°N and 36°N, superimposed; a) Harvard results and b) Chicago results.

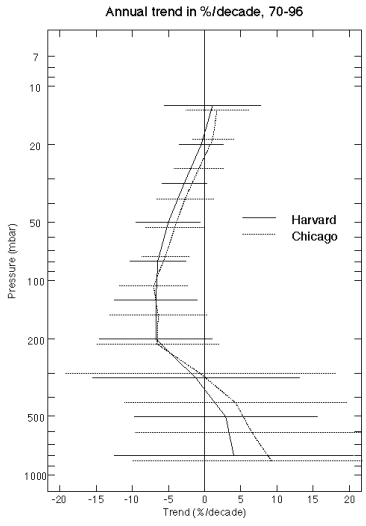


Figure 3.36 Mean annual trend for the sonde stations located between 36°N and 59°N. The solid line shows the Harvard results, the dashed line the Chicago results. Two standard errors are shown; these were calculated as the standard error of the nine trend values at each pressure level.

The ensemble of seasonal trends for 36°-59°N from the Harvard model is shown in Figure 3.37. The closest agreement of trends is found in spring near the ozone maximum, where it is easiest to measure ozone. There is a seasonal variation in the trends that appears to depend on region, as discussed previously by Logan (1994) and Bojkov and Fioletov (1997). Decreases in ozone are largest in spring and winter at the European stations, and in spring and summer at the Canadian stations, as shown in Figure 3.38. There is no significant decrease in ozone below 90 hPa at the European stations in summer and autumn, while the decrease persists in spring to 200 hPa. The decreases in ozone are larger for the Canadian stations than for the European stations, and persist to 200 hPa and below in all seasons. Given the concerns about the Brewer Mast data for the Canadian stations (Chapter 1) these results are subject to some uncertainty. The Japanese stations Sapporo and Tateno also show largest decreases in winter and spring.

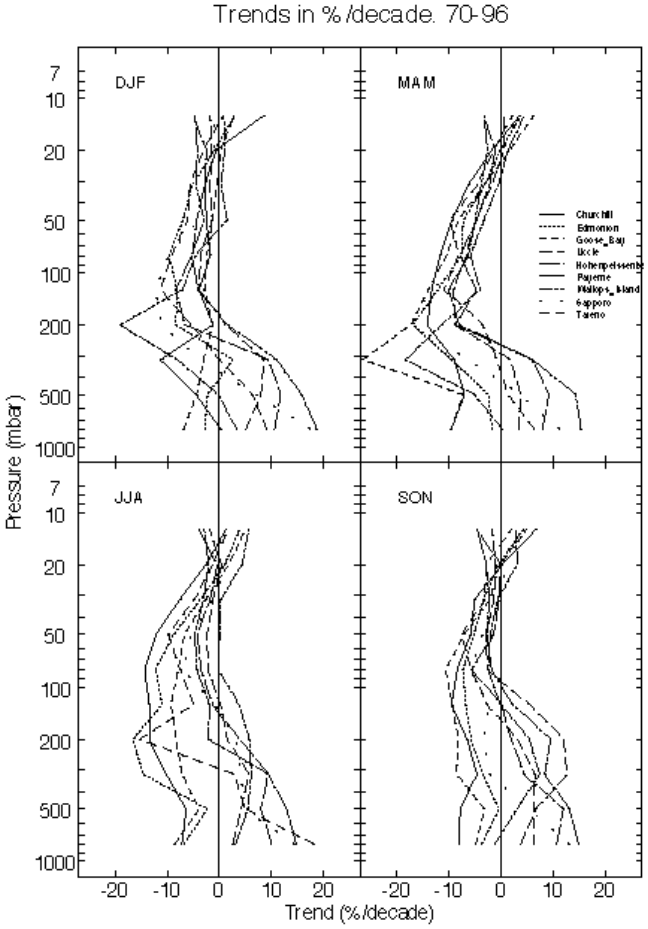


Figure 3.37 Superimposed seasonal mean trends for the sonde stations located between 59°N and 36°N. Harvard results.

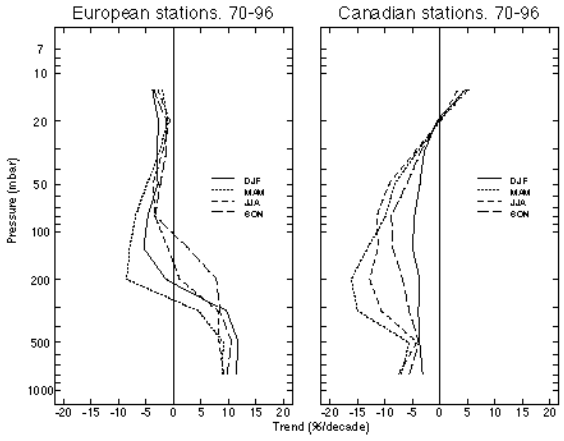


Figure 3.38 Seasonal mean profiles of ozone trend for three European stations, 48-51°N (left) and three Canadian stations, 53-59°N (right). Harvard results.

The composite seasonal behaviour of the eight stations located between 36 and 53°N is shown in Figure 3.39. Both the Harvard and Chicago groups find strong seasonality in the trends from about 300 to 90 hPa, but the seasonal patterns are somewhat different. Both groups find the largest losses in ozone in spring and the smallest losses in autumn in this region. For the Chicago results, the losses in winter are almost as large as in spring, while for the Harvard results, the winter losses are smaller than those in spring.

Which trend results for 1970-96 are likely to be more reliable? Rather than make a judgement, we offer some comments. The major difference between the Harvard and Chicago groups is the choice of data selection prior to trend analysis. The Harvard group use less strict criteria, with the goal of maximising the amount of data of reasonable quality available for analysis. They do not require the soundings to reach a certain height because 95-98% of the soundings at the re-evaluated stations reach 20 hPa. For these stations an ozone column measurement is given whenever a correction factor is given. The Chicago group's criteria are designed to maximise the quality of data used in trend analysis, but the end result is the loss of 70% of the Canadian Brewer-Mast data, and loss of 40-55% of the Uccle, Payerne, Canadian ECC, and Japanese data as shown in Table 3.2. The main cause of data loss is the stricter correction factor criteria, although the other requirements cause significant data loss for the Canadian stations. The effect is that gaps are introduced in the time series, as shown in Figure 3.28. The optimal selection criteria for use of sonde data in trend analysis is clearly a subject of debate. However, both groups find similar results for mean stratospheric trends from 1970 to 1996, even with these differences in analysis.

Another important issue brought out by the differences in the approaches used by the Chicago group and the Harvard group is the scaling to the total column ozone (the correction factors, CFs). Although there are some difficulties with using a uniform scaling for all altitudes, Chapter 1 of this report recommends that the sonde data should be scaled to the ozone column for derivation of the most reliable stratospheric data. The Chicago group selected their data treatment prior to this recommendation and they thus present trends in the uncorrected measurements. For most stations this makes little difference. However, some stations have a trend or a jump in the correction factors and this can make a difference of up to 2-3%/decade as discussed previously by Logan (1985, 1994) and Miller *et al*. (1995). Given the recommendation of Chapter 1 the trends using correction factors are likely to be more reliable than those without.

Tiao *et al*. (1990) emphasise the need for accounting for autocorrelation in the data. For the sondes, this appears to have a minor effect on errors derived for trends (e.g., results for Hohenpeissenberg in Figure 3.32), while the treatment of the data prior to deriving trends has a larger impact on trends and associated errors.

Re-evaluation could clearly improve the Canadian data set. About 20% of the soundings fail to reach 20 hPa and many ozone column values are missing (mostly for ECC soundings). Reprocessing of the data using the upper stratospheric profiles from SBUV data to derive ozone columns after 1980 could solve the latter problem (McPeters *et al*., 1996).

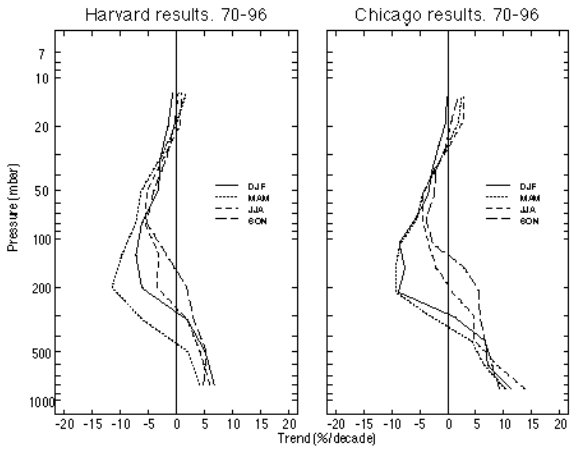


Figure 3.39 Seasonal mean profiles of ozone trend for 8 sonde stations located between 59°N and 36°N. a) Harvard results and b) Chicago results.

Comparison of trends for 1980-96 with 1970-96

Annual trends for individual stations from 1980 to 1996 are shown in Figure 3.40, while the mean trend for 36°-59°N is shown in Figure 3.41 (solid line). In each figure the results are compared to the trend for 1970-96 (dashed line). For about half the stations, the decrease in ozone in the lower stratosphere is larger for the period 1980-96 than for the period 1970-96. The mean trend for the 1980-96 period is more negative by 1-2.5%/decade, with a maximum trend in the lower stratosphere of -10%/decade. The errors in the seasonal trends are larger for the shorter period, because of the short length of the time series. For the Japanese stations the summer trends have extremely large errors because of the paucity of data in much of the 1980s and these contribute to the large errors in the annual trends. The errors in the mean trend for 1980-96 are larger than the errors in the mean trend for 1970-96, because of these factors.

Figure 3.40 includes results for three stations that started operation after the 1970s. Boulder (1980-96) shows significant decreases in ozone from 80 to 20 hPa, while Hilo (1982-96) shows a significant decrease only at 80 hPa; it shows a significant increase at 15 hPa. The data from Lauder (1986-96) do not show any significant trends in the lower stratosphere, but show an increase at 30 hPa. The re-evaluated data were not scaled to the correction factor, and as shown in Table 3.3, there is a small but significant trend in the correction factor.

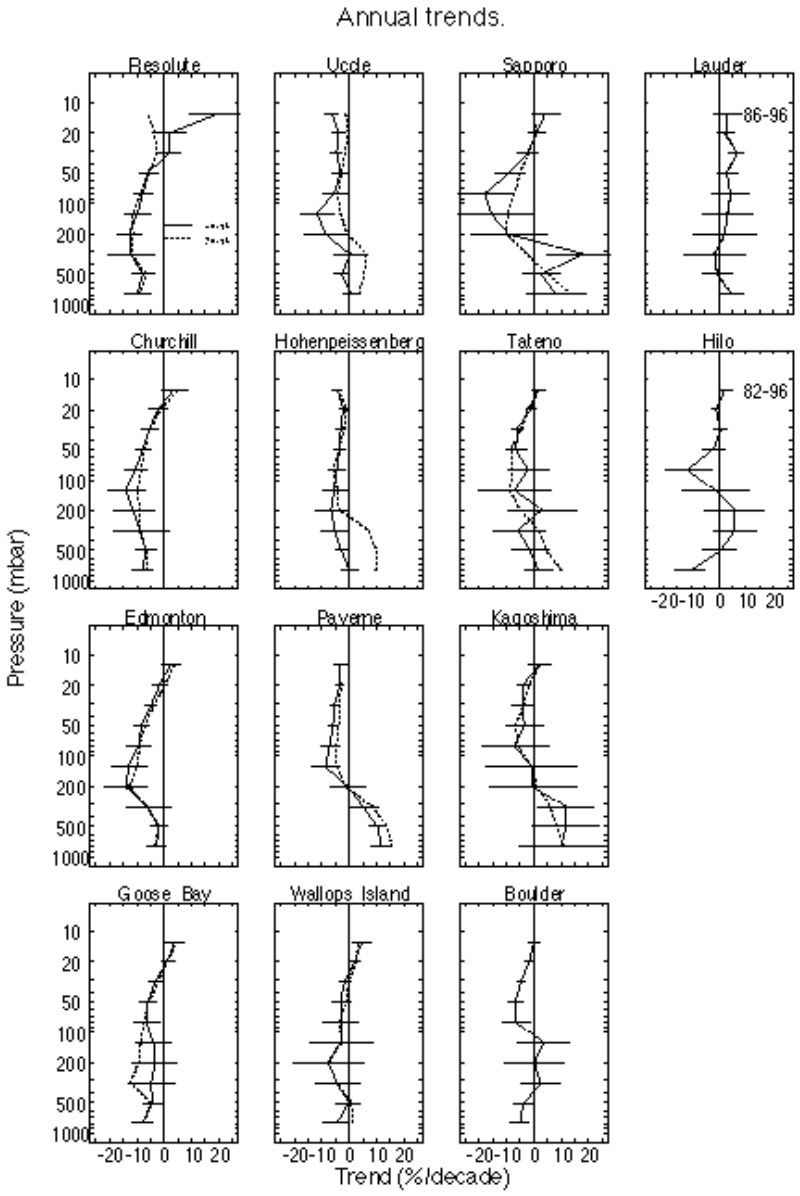


Figure 3.40 Annual trends for 1980-96 (solid lines) compared to trends for 1970-96 (dashed lines) where available. The Lauder trends are for 1986-96, and the Hilo trends for 1982-96. Harvard results.

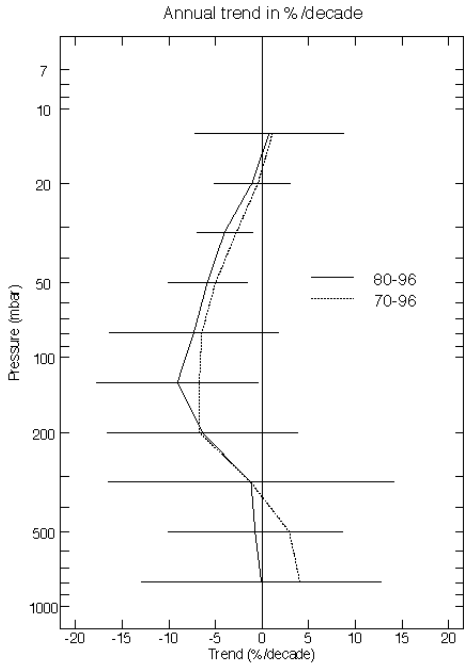


Figure 3.41 Mean annual trend for the sonde stations located between 36°N and 59°N for 1980-96 (solid line) compared to the mean trend for 1970-96 (dotted line). Harvard results. Two standard errors are shown; these were calculated as the standard error of the nine trend values at each pressure level.

Differences in trends derived from WOUDC data and the data re-evaluated for this report

The Chicago group have applied the same trend model to the data re-evaluated for this report and to the data archived at WOUDC prior to re-evaluation for Hohenpeissenberg, Payerne, and the three Japanese stations. Results from the two data sets for each station are fairly similar with the exception of results for Payerne in the troposphere, discussed below. The WOUDC data give slightly more negative trends than the re-evaluated data for the lower stratosphere at Hohenpeissenberg in all seasons. For the Japanese stations and Payerne the results are similar for both data sets from about 80 to 20 hPa.

3.5.2.2. SAGE ozone trends compared with sondes

Figure 3.42a shows time series of monthly mean SAGE ozone values in Dobson units between 20 and 21 km altitude and in the latitude band between 45°N and 50°N. Sunrise and sunset events have been combined in this figure because in the lower stratosphere the differences between sunrise and sunset retrievals are small (much less than those in the upper stratosphere in v5.96). The periods of missing data between the SAGE I and SAGE II observations in 1982-84 and the Pinatubo affected period in 1992-93 are obvious. There is a strong annual cycle in ozone at this location, with highest values in late winter - early spring and smallest values six months later. Figure 3.42b shows the same data after deseasonalisation. These data are relatively compact, except for some wintertime downward spikes associated with years when ozone failed to reach typical high values. These are associated with QBO modulations in mid-latitudes; note the similar patterns at 28°N in Figure 3.9. There is an clear downward trend in the deseasonalised SAGE ozone values. The solid line in Figure 3.42b indicates the statistical model fit to the data and Figure 3.42c shows the residual from this fit.

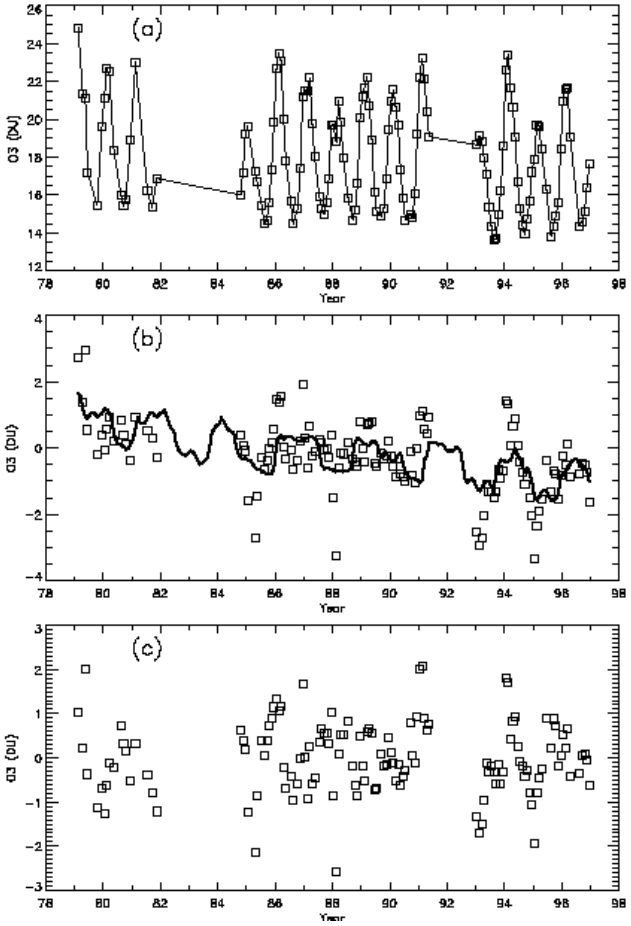


Figure 3.42 a) SAGE I/II monthly mean ozone measurements between 45°N and 50°N integrated from 20 to 21 km altitude in Dobson units. b) The data after deseasonalization. The black line shows the deseasonalized fitted trend model. c) The residuals of the measurements from the trend model.

In Figure 3.16 a meridional cross section of ozone trends in %/year was shown for the combined SAGE I/II data set from 1979-1996. Note these trends are calculated to a lower level of 20 km, as the combined SAGE I/II are not reliable for trend estimates below this altitude (see Chapters 1 and 2). The SAGE I/II trends in the lower stratosphere are significantly different from zero only over latitudes of 40°N to 55°N, where values of order -0.4%/year are found. Figure 3.43a compares the annual mean SAGE I/II trends over 45°N-50°N with averaged sonde results for stations in the latitude band 40°N-53°N. Note that the SAGE data has been extended down to 15 km altitude despite the large uncertainty below 20 km for SAGE I. The agreement with the sondes is good above 20 km. Below 20 km the SAGE I/II trends are different from those derived from sondes but the difference is within the estimated statistical uncertainties (2s). There is little seasonal variation in the SAGE I/II trends in the lower stratosphere (above 20 km), consistent with the lack of seasonality observed in the sondes above 70 hPa (see e.g. Figure 3.39).

Figure 3.43b compares the annual mean SAGE II trends over 45-50°N with averaged sonde results for stations in the latitude band 40°N-53°N for the time period 1984-1996. The agreement is excellent over the entire altitude range from 15 to 27 km. This helps to confirm the use of SAGE II data down to an altitude of 15 km for determination of trends.

Trends between 15 and 20 km derived from the SAGE II data are not, in general, statistically significant (Figure 3.17). The derived trends are negative throughout the northern hemisphere and are strongly negative in the tropics (but not statistically significant because there are large uncertainties in this relatively short data record). In the southern hemisphere mid-latitudes the SAGE II data show a positive trend which is also insignificant, however it is worth noting that the Lauder (45°S) sondes (Figure 3.41) show positive annual trends for the time period 1986-96.

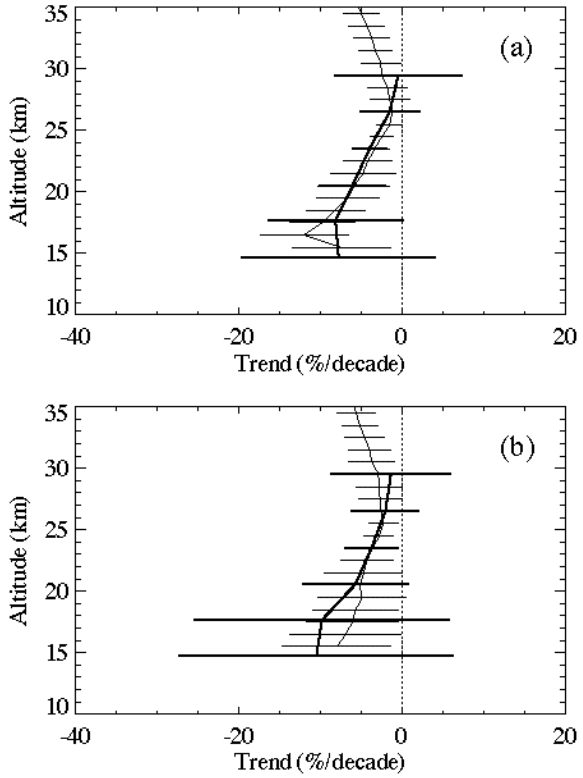


Figure 3.43 Average ozonesonde trends for Uccle, Hohenpeissenberg, Payerne, Edmonton, Goose Bay, Sapporo, and Boulder (40°N-53°N) for the time periods a) 1980-1996 and b) 1984-1996 in %/decade (darker lines). These are compared to a) trends for SAGE I/II and b) trends for SAGE II alone for 45-50°N (lighter lines). The SAGE trends were analyzed at 1 km intervals and the ozonesonde trends have been expressed as a function of altitude using the 1976 standard, mid-latitude atmosphere. 2s error bars are given.

3.5.3. Troposphere

3.5.3.1. Time series

Monthly mean values of ozone near 500 hPa are shown in Figure 3.44 for selected stations; the right hand panels show the monthly anomalies. The data from Goose Bay are for BM sondes prior to August 1980, and ECC sondes thereafter. An intervention term was used in both statistical models to account for the jump, which intercomparisons from the 1970s and early 1980s suggest to be about 15-20%. Difference time series for the three European stations are shown in Figure 3.45 for 500 hPa. There are systematic biases for ozone values in the mid-troposphere in the data reanalysed for this report, as were found in the WOUDC data for Hohenpeissenberg and Payerne (Logan, 1994). Even with the re-evaluated data, ozone values at Hohenpeissenberg are systematically higher than those at Payerne from about 1978 to 1990, while there is less bias in the 1970s. Values at Uccle are higher than those at Payerne

and at Hohenpeissenberg up to about 1986. These biases result in different trends for the three stations. The data used here for Payerne are provisional, as discussed in Chapter 1, and are likely to be revised further. There are particular concerns about the consistency of the tropospheric data for Payerne in the 1980s.

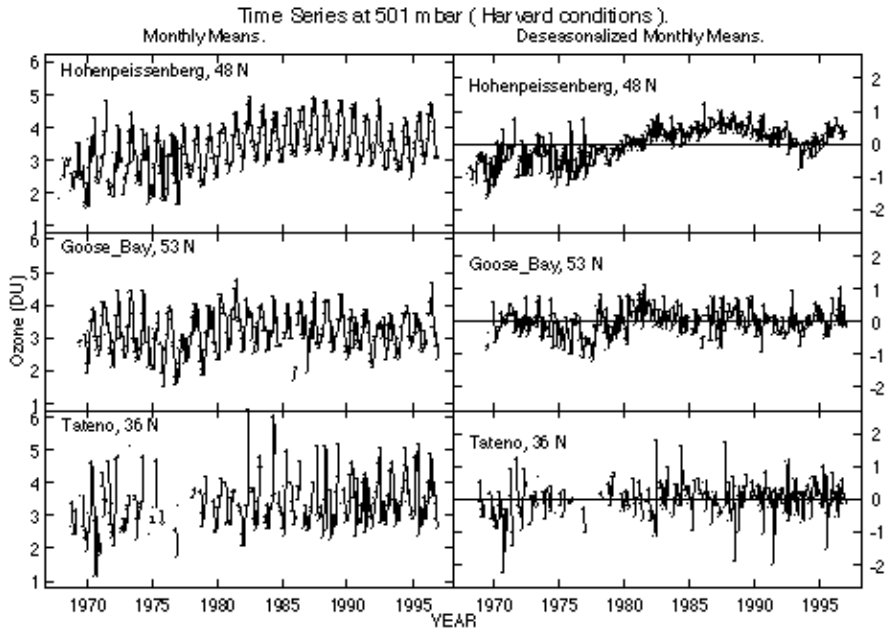


Figure 3.44 Time series of monthly mean values for ozone in DU and deseasonalized monthly means for selected stations. The selection criteria used by the Harvard group were applied (see text). Values are shown for a level near 500 mbar. The same relative scale in used for both the monthly means and the deseasonalized monthly means.

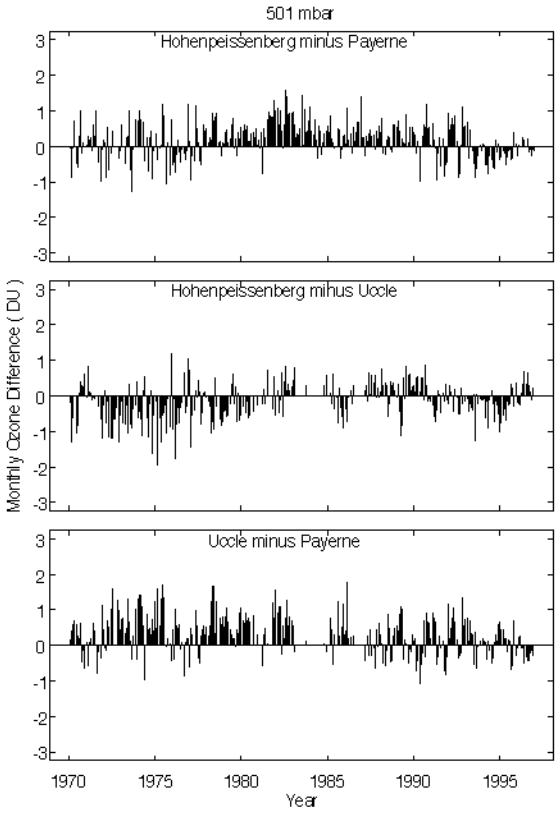


Figure 3.45 Difference of monthly mean values near 500 mbar for the three European stations. The selection criteria used by the Harvard group were applied (see text).

3.5.3.2. Trends for 1970-96

The most obvious feature of the tropospheric trends shown in Figures 3.31a,b,c and 3.32 is that there are significant spatial variations in the magnitude and sign of the trends, with decreases, or no significant trend at the Canadian stations and increases at the European and Japanese stations. The increases over Europe and at Kagoshima are significant up to 300 hPa, while the increases at Sapporo and Tateno are significant only at 500 hPa and below. There is not a significant seasonal variation in the tropospheric trends (e.g., Figures 3.31, 3.38, 3.39). Compared to the Harvard results, the Chicago trends are less negative for the Canadian stations, more positive for the European stations and Wallops Island, and about the same for the Japanese stations. The Harvard group find increases of 5-15%/decade for the European stations, while the Chicago group find increases of 6-25%/decade. Part of this difference is caused by referencing the trend to the beginning rather than the middle of the record. For example, at the lowest layer at Hohenpeissenberg, the Chicago trend is 16%/decade, while the Harvard trend is 10%/decade; however, the Chicago trend would be only 12%/decade, if referenced to the same value as the Harvard trend. Similarly, the Chicago trend in the lowest layer at Wallops Island would be reduced from about 3%/decade to 1.2%/decade (similar to the Harvard trend) if referenced to the same value as the Harvard trend. The results from the two groups for Payerne and Hohenpeissenberg for the absolute trend (Figure 3.32b) and for the tropospheric column trend are in closer agreement than the results for the percentage profile trends (Figure 3.46, solid and dashed lines), another indication that the reference point makes an important difference to the results in %/decade. The Harvard and Chicago results for the column trends for 1000-250 hPa are similar for most stations, and using the Chicago data treatment with the Harvard model generally improves the agreement.

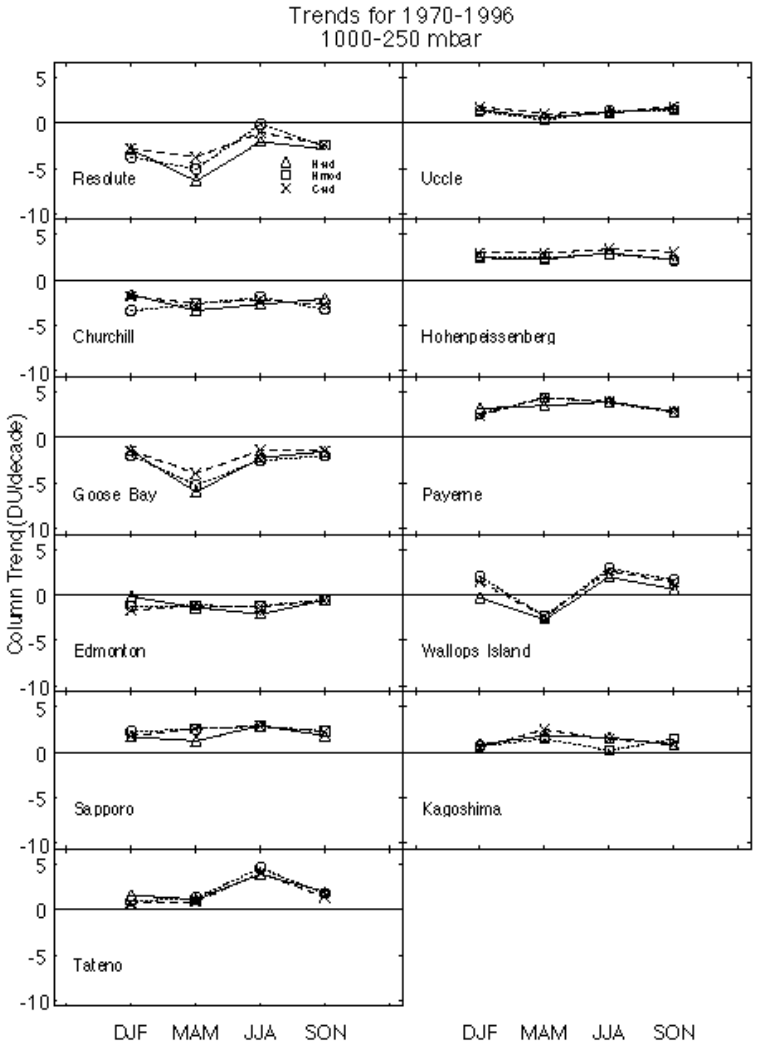


Figure 3.46 Column trend in ozone in DU/decade from the lowest layer to 250 mbar. The triangles are results from the Harvard group, the crosses those from the Chicago group, and the circles are results for the Harvard model, with three conditions of the Chicago group used in data selection, their CF criteria, the balloon reaching 16 mbar, and an ozone column measurement, and with the data divided by the CF.

The ozone increases in Europe and Japan from the Harvard group for 1970-96 are somewhat less than those reported for 1970-91 for the WOUDC data (Akimoto *et al*., 1993; Logan, 1994). This is caused by the relatively flat values of ozone in the last few years (see Figure 3.44). Bojkov and Fioletov (1997) find the increase at Hohenpeissenberg to be significant only 1 km below the tropopause. By comparing data for the first 5 years from the Canadian stations with data for 1987-91, Logan (1994) concluded that there was no evidence for a long-term increase in ozone at the Canadian stations, given the different responses of BM and ECC sondes. This result contradicted the earlier analysis of Wang *et al*. (1993) who reported an increase of 10%/decade for all the Canadian stations. Using an intervention in the statistical model to treat the change in sonde type, the Harvard group finds long-term decreases for all the Canadian stations (-2 to -9%/decade), while the Chicago group finds similar decreases or no significant trend at the same stations. Oltmans *et al*. (1998) report an increase of 15%/decade for Hohenpeissenberg for 1968-95, using a least squares fit to annual mean values. They also find no trend at Wallops Island and results very similar to those in Figure 3.32 for Tateno. They analysed only these three stations, selected for the

consistency of their record, and used the data at WOUDC.

3.5.3.3. Trends for 1980-96

There is a major change in many of the trends for the period 1980-96 compared to 1970-96 (Figure 3.40). For the European stations, only Payerne shows an increase, of ~10%/decade, while Uccle shows no change in ozone, and Hohenpeissenberg has a slight negative trend in the middle troposphere. There are concerns about the consistency of the tropospheric data for Payerne in the 1980s. Tateno also shows no change in ozone, while Sapporo and Kagoshima have increases of 5-15%/decade, not all of which are significant. There are many gaps in the data record for these two stations in the early 1980s, particularly for summer. These gaps give rise to large errors in the summer trends, which contribute to the large errors in the annual trends. The Canadian stations show decreases of -2 to -8%/decade, and these are more reliable than the results for 1970-96 since ECC sondes were used for the whole record. Previous analyses of the Canadian ECC data also showed decreases (Logan, 1994; Tarasick *et al*., 1995; Oltmans *et al*., 1998). Oltmans *et al*. (1998) find no significant trend in ozone for Hohenpeissenberg, Boulder, Wallops Island, Tateno, and Hilo for 1979-95 in the middle troposphere, in agreement with the results in Figure 3.40. The increases in ozone apparent in the late 1960s and 1970s appear to have levelled off at several of the sonde stations, and also at remote surface sites, as discussed by Logan (1994) and by Oltmans *et al*. (1998). The mean trend for the stations from 59°-36°N is zero, 4%/decade less than the mean trend for 1970-96 (Figure 3.41). Mean trends for both periods are insignificant. The concept of a mean trend is less appropriate for the troposphere than for the stratosphere. The locations of the sonde stations in remote regions of Canada, and more polluted regions of Europe and Asia, may lead to different regional influences on tropospheric ozone from trends in emissions of NOx and from changes in stratospheric input of ozone (e.g., Logan, 1994).

3.5.3.4. Differences in trends derived from WOUDC data and reanalysed data

The Chicago group finds similar tropospheric trends using the reanalysed data and WOUDC data for Hohenpeissenberg, Sapporo, Tateno and Kagoshima. The results for the two data sets are dramatically different for Payerne, with much larger tropospheric increases derived for the WOUDC data than for the data used here (Figure 3.47). The history of problems with the Payerne data is documented in Chapter 1, and the WOUDC data contained erroneously high values for ozone in early 1990s. The data used here are considered more reliable, but further revisions are expected in the Payerne data. Miller *et al*. (1995) used the WOUDC data and reported anomalously high trends.

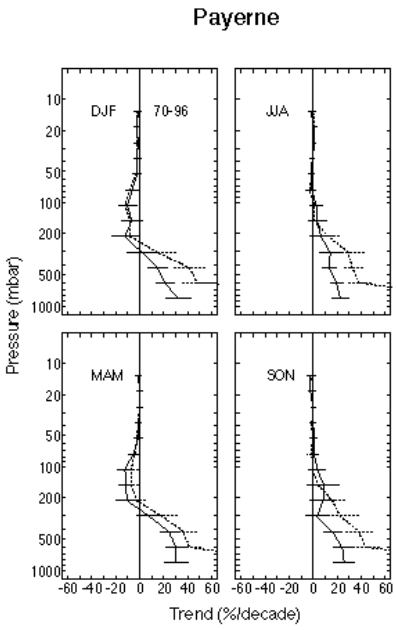


Figure 3.47 Comparison of seasonal trends for (i) sonde data re-evaluated for this study (solid lines) and (ii) for sonde data archived at WOUDC (dotted lines), for Payerne. The same data selection criteria and trend model was applied to both sets of data. Chicago results.

3.5.4. Integral and column ozone

In this section the column ozone trends derived from TOMS are compared to the integrated ozone profile trends derived from SAGE I/II and ozonesonde data, in order to determine consistency between these independent results. TOMS data are the combined version 7 Nimbus 7 - Meteor 3 data, spanning January 1979-December 1994 (as reported in McPeters *et al.*, 1996). SAGE I/II trends are calculated over January 1979-December 1996 (with no data available during January 1982-September 1984), and ozonesonde results span January 1980-December 1996. We consider near-global comparisons between TOMS and SAGE I/II, and northern hemisphere mid-latitude results for TOMS, SAGE I/II and ozonesondes in subsequent sections.

3.5.4.1. Comparison of TOMS and SAGE I/II trends

SAGE I/II data are reliable for trend analyses over the altitude range 20-55 km, so the column integral of SAGE data is calculated over these levels. For reference, the seasonal climatology of column ozone over 20-55 km from SAGE data is shown in Figure 3.48b, together with the TOMS climatology (Figure 3.48a), and their difference (i.e. the inferred column ozone below 20 km, Figure 3.48c). Note the lack of SAGE sampling in polar regions in Figure 3.48b, so that direct comparisons are restricted to ~60°N-S. In the extra-tropics, a large fraction of the ozone column lies below 20 km. The column above 20 km shows relatively weak seasonal or latitudinal variations, whereas large latitudinal gradients and strong seasonality is seen in the lower stratosphere (below 20 km).

Click here to download Figure 3.48 by anonymous ftp.

Figure 3.49a shows the TOMS column ozone trends as a function of latitude and season. This is similar to the results shown in McPeters *et al.*, (1996), except the units here are DU/decade to facilitate the comparisons below. Instrumental uncertainty is not included in the significance estimates in Figure 3.49. The TOMS data show negative trends in northern hemisphere mid-latitudes during late winter-spring; the

March-April values of ~ -24 DU/decade in Figure 3.49a correspond to percentage changes of ~ -6%/decade. There are smaller magnitude mid-latitude trends in the southern hemisphere (~ -10 DU/decade). No significant TOMS trends are found in the tropics.

Trends in column ozone over 20-55 km derived from SAGE I/II data are shown in Figure 3.49b as a function of latitude and season. Significant negative trends of order -4 to -8 DU/decade are observed in northern hemisphere mid-latitudes throughout most of the year (with a maximum in April-May near 50°N), and in southern hemisphere mid-latitudes during all seasons except summer. Significant trends are not observed near the equator. Comparison with the much larger TOMS trends in northern hemisphere spring mid-latitudes suggests that a majority of these column trends occur at altitudes below 20 km.

Figure 3.49c shows trends in the difference between the TOMS column ozone data and the SAGE I/II data integrated from 20 to 55 km. This corresponds to trends in the column below 20 km (i.e. the difference between Figures 3.49a,b). Negative trends are observed polewards of 40° in each hemisphere; there are relative maxima in northern hemisphere spring (discussed above) and in southern hemisphere summer (although the former are not statistically significant). Small positive trends are seen in the tropics (~ 30°N-S); because of the small background (seasonal) values in the tropics below 20 km (see Figure 3.48c), these latter trends equate to a relatively high percentage (~5-10%/decade).

Click here to download Figure 3.49 by anonymous ftp.

3.5.4.2. Comparison of ozonesonde, SAGE I/II, and TOMS trends

The majority of long-term ozonesonde records are in northern hemisphere mid-latitudes, so we focus here on comparisons with ozonesonde trends over the latitude range 40°-53°N. The ozonesonde results are an average of seven stations (Boulder (40°N), Sapporo (43°N), Payerne (47°N), Hohenpeissenberg (48°N), Uccle (51°N), Edmonton (53°N), and Goose Bay (53°N), while for the satellite data we use the zonal means over 40°-52°N. Figure 3.50a,b,c,d shows vertical profiles of the ozonesonde and SAGE I/II trends for each season, expressed in units of DU/km/decade for each of four seasons. The uncertainties for the seasonal trends in Figure 3.50 are relatively large (much larger than the annual means in Figure 3.43); the sonde uncertainties in DJF are larger than those in other seasons because of large geographical variability between the different stations. Also shown in Figure 3.50 are the column integrals, separated according to 0-20 km and 20-30 km levels for the sondes, sums above 20 km for SAGE, and the total column trends from TOMS. Comparison of the sondes and SAGE I/II trends shows reasonable agreement over altitudes of 20-27 km with SAGE slightly less negative (as shown also in Figure 3.43). The column integrals of the sondes show good agreement with TOMS trends, in particular showing a maximum in the winter (DJF) and spring (MAM). A substantial fraction of the ozonesonde trends occur below the 20 km level for all months except autumn. In spring the trends below 20 km are particularly large, accounting for approximately two-thirds of the total column. Both the magnitude and seasonality in the ozonesonde trends below 20 km are consistent with the differences between TOMS and SAGE I/II shown in Figure 3.49 above (at least for the latitude band 40°-53°N).

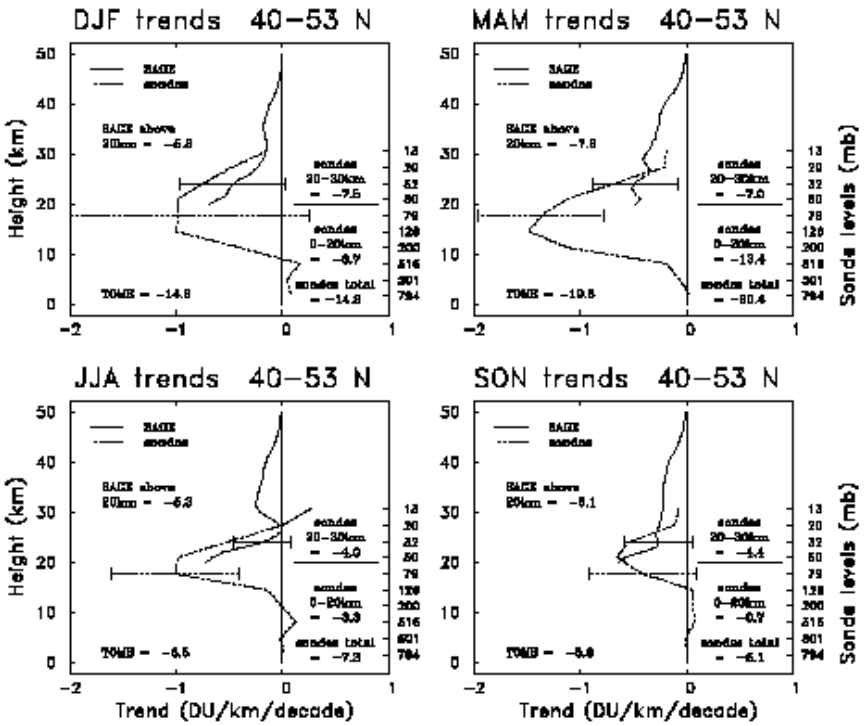


Figure 3.50 Altitude profiles of ozone trends over 40°N-53°N for; a) December, January, February; b) March, April, May; c) June, July, August; d) September, October, November, showing SAGE I/II and ozonesonde results. The column integrals for the SAGE data (above 20 km) and ozonesonde data (over 0-20 and 20-30 km separately) are indicated, along with TOMS total column trend results.

3.5.4.3. Comparison of sonde, Dobson, and TOMS trends at the sonde locations

The sonde trends in DU were integrated from the lowest layer to the layer with its top boundary near 16 hPa, omitting trends derived from the less reliable data obtained near the top of the soundings. These are compared to trends derived from the Dobson (or Brewer) column data that were obtained on the same day as the sonde measurements. (The column data were sometimes unavailable for the Canadian ECC sondes). The trends in the column data were derived using the Harvard trend model used for sonde data, omitting measurements on days that did not meet the CF criteria for sondes used by the Harvard group. Figure 3.51 compares the column trends and the integrated sonde trends, the latter for both the Harvard and Chicago results. The Harvard results are in somewhat better agreement with the column trends for the European stations, while the Chicago results underestimate the column loss more than the Harvard results. For the Canadian and Japanese stations, there is no systematic bias between the Harvard and Chicago results with respect to the Dobson trends. If there is ozone loss above 16 hPa, the integrated sonde trends should be less negative than the Dobson trends, which is sometimes but not always the case.

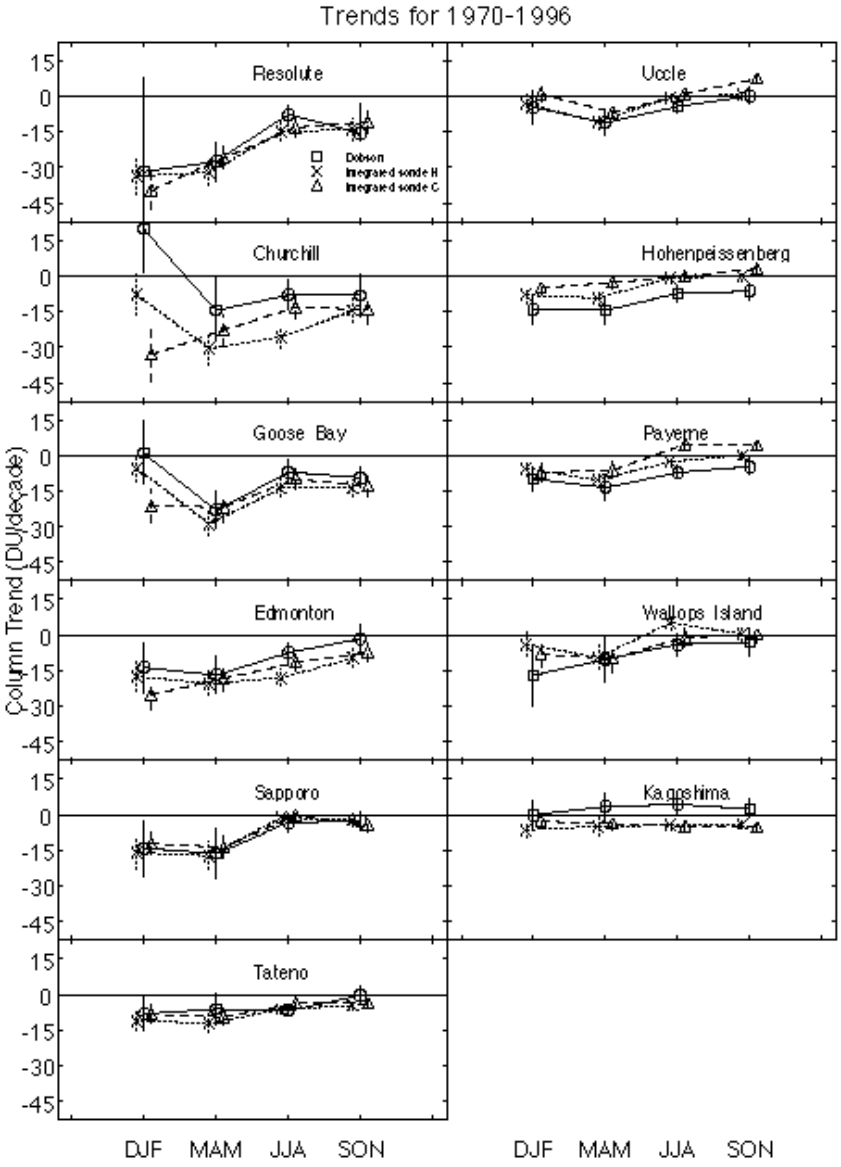


Figure 3.51 Comparison of column trends for ozone for 1970-96. The circles show the trend in the Dobson overhead ozone column measured on the same days as the sondes used in the Harvard analysis, computed with the Harvard trend model. The crosses show the integrated sonde trend up to 16 mbar, computed with the Harvard trend model, and the triangles show the integrated sonde trend up to 16 mbar computed by the Chicago group. Two standard errors are shown. The errors for the sonde data are only approximate, as they do not account for any correlation between ozone at one layer and the next, and the Chicago errors are smaller in part because they are for 12 layers rather than 9. The three results for each season are offset for clarity.

Figure 3.52 shows the column trends derived from TOMS by Hollandsworth for November 1978 to October 1994, compared to the integrated sonde trends and Dobson column trends for 1980-1996. The three trends agree within their standard errors, but the agreement is best for the European stations, Boulder and Tateno. The agreement is worst for Churchill and Resolute in the winter, when Dobson data are sparse. The TOMS data confirm that the percent decrease in ozone (not shown) is largest in spring and summer at the Canadian stations (except Goose Bay with largest losses in autumn) and in winter and spring at the European stations. Figure 3.52 is not an ideal comparison of the ground-based and TOMS ozone columns, since it uses the ground based column data only on the days when there was an ozone sounding.

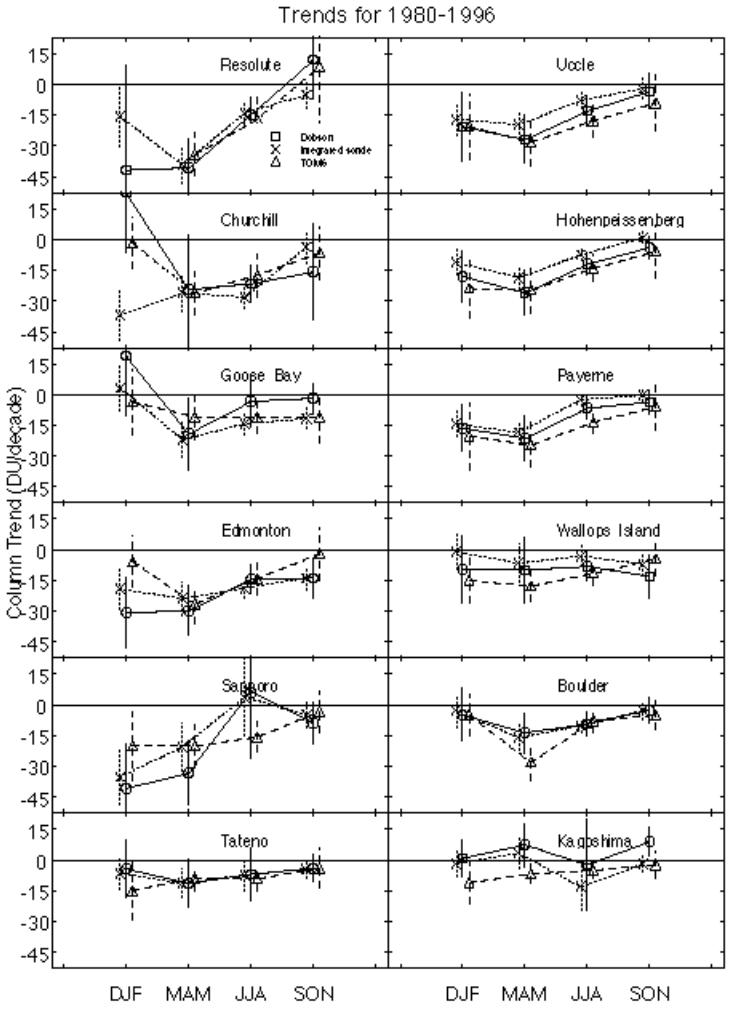


Figure 3.52 Comparison of column trends for ozone for 1980-96. Circles and crosses show the column and integrated sonde trends defined as in Figure 3.51. The triangles show the trend in TOMS data for November 1978 to October 1994, computed by Hollandsworth. The three results for each season are offset for clarity.

3.5.4.4. Summary

A large fraction of the ozone column in middle-high latitudes (and most of the seasonal variation) occurs below 20 km. Trends in the column ozone above 20 km derived from SAGE I/II data show negative trends over mid-latitudes of both hemispheres, but the values are substantially less than TOMS in high latitudes (particularly for northern hemisphere spring and southern hemisphere summer). Trends derived from ozonesonde data over 40°N-53°N show that large fractions of the column trends occur below 20 km. Both the magnitude and seasonal variation of the ozonesonde trends are consistent with TOMS data, and the ozonesonde results below 20 km are in agreement with the observed TOMS - SAGE I/II trend differences.

3.6. Combined trend at northern mid-latitudes and uncertainty estimate

In this chapter we have calculated trends as a function of altitude from 4 different measurement systems, SAGE, Umkehr, SBUV, and sondes. Each trend calculation was derived with an associated estimate of the statistical uncertainty. An additional uncertainty is that due to the potential drift of the instrument system over time. Such uncertainties are not included in the standard uncertainty derived from the statistical time-series models. For each of the systems, an attempt was made in Chapter 1 of this report to estimate these instrument-drift uncertainties. This section takes those estimates and combines them with the estimate for the statistical sampling uncertainty from the time-series models. The result is an estimate of overall uncertainty for each of the instrument systems.

	Sapporo			Edmonton			Uccle			Payerne		
Alt	Drift	Stat	Tot	Drift	Stat	Tot	Drift	Stat	Tot	Drift	Stat	Tot
2.5	0.92	4.98	5.07	0.72	1.56	1.71	2.30	1.56	2.78	2.30	1.66	2.83
5.0	0.92	3.45	3.57	0.72	1.62	1.77	2.30	1.38	2.68	2.30	1.60	2.80
7.5	0.86	5.77	5.84	0.69	3.54	3.60	2.03	2.51	3.23	2.03	2.58	3.28
10.0	0.80	6.67	6.72	0.66	3.98	4.03	1.75	3.39	3.82	1.75	3.03	3.50
12.5	0.74	7.17	7.21	0.63	3.69	3.75	1.48	3.62	3.91	1.48	2.99	3.33
15.0	0.68	7.48	7.51	0.59	3.14	3.20	1.21	3.08	3.31	1.21	2.50	2.78
17.5	0.62	5.54	5.57	0.56	2.29	2.36	0.93	2.32	2.50	0.93	1.87	2.09
20.0	0.56	3.65	3.69	0.53	1.61	1.70	0.66	1.67	1.80	0.66	1.24	1.40
25.0	0.78	1.74	1.91	0.72	1.09	1.30	1.23	0.99	1.58	1.23	0.65	1.39

	Boulder			Goose Bay			Hohenpeissenberg		
Alt	Drift	Stat	Tot	Drift	Stat	Tot	Drift	Stat	Tot
2.5	0.72	1.72	1.87	0.72	1.89	2.02	2.30	1.32	2.65
5.0	0.72	1.87	2.01	0.72	1.66	1.81	2.30	1.15	2.57
7.5	0.69	3.13	3.21	0.69	3.64	3.71	2.03	2.22	3.01
10.0	0.66	4.50	4.55	0.66	4.10	4.16	1.75	2.79	3.30
12.5	0.63	5.12	5.16	0.63	3.77	3.82	1.48	2.76	3.13
15.0	0.59	4.51	4.54	0.59	3.12	3.17	1.21	2.19	2.50
17.5	0.56	2.72	2.78	0.56	2.39	2.45	0.93	1.75	1.98
20.0	0.53	1.79	1.87	0.53	1.76	1.84	0.66	1.25	1.42
25.0	0.72	0.66	0.98	0.72	1.11	1.33	1.23	0.68	1.41

Table 3.4. Estimated uncorrected-drift uncertainties (all 1s) for each station along with statistical uncertainties of trend analysis and combined uncertainties. The drift uncertainties were determined from estimates in Chapter 1 for each type of sonde. The totals were determined as a root sum of squares of the drift and statistical uncertainties.

For ozonesondes, 7 stations at northern mid-latitudes were used. These are indicated in Table 3.4 along with the estimate for the uncorrected-drift uncertainty and the statistical uncertainty in the trend fit. These were combined as a root sum of squares (rss) to give an estimate for the overall trend uncertainty of each sonde station. Combination by rss is a reasonable way to combine these errors but it is not the only way. The trends and uncertainties for these 7 stations were then combined by using weighted means with the weighting factor being the reciprocal of the square of the total standard error. These results are given in Table 3.5 along with the standard error of the mean calculated from the 7 trend values at each altitude. At altitudes of 20 km and below, the standard error of the mean is larger than the overall uncertainty estimated from the station characteristics. It was decided to use the larger of the two values as a first cut at the uncertainty in the trend estimate from the sondes. However, the estimates of individual instrument uncertainties are still used to determine weighting to get the mean trend.

Alt	Sappo	Edmon	Uccle	Bould	Goose	Hohen	Payer	Comb	SE Mean
2.5	5.07	1.71	2.78	1.87	2.02	2.65	2.83	0.87	2.48
5.0	3.57	1.77	2.68	2.01	1.81	2.57	2.80	0.85	1.99
7.5	5.84	3.60	3.23	3.21	3.71	3.01	3.28	1.32	2.72
10.0	6.72	4.03	3.82	4.55	4.16	3.30	3.50	1.52	1.84
12.5	7.21	3.75	3.91	5.16	3.82	3.13	3.33	1.48	1.96
15.0	7.51	3.20	3.31	4.54	3.17	2.50	2.78	1.25	2.28
17.5	5.57	2.36	2.50	2.78	2.45	1.98	2.09	0.93	1.68
20.0	3.69	1.70	1.80	1.87	1.84	1.42	1.40	0.66	1.07
25.0	1.91	1.30	1.58	0.98	1.33	1.41	1.39	0.51	0.47

Table 3.5. Combined drift and statistical uncertainties for each sonde station as a function of altitude (1s). The weighted combined uncertainty is also shown along with the standard error of the mean for the trend at each station. Because the standard error of the mean was larger for most altitudes, the total uncertainty was taken as the larger of the combined uncertainties and the standard error of the mean. At all altitudes except 25 km this means that the standard error of the mean was used to estimate the combined sonde uncertainty.

The fact that the standard error of the mean for the 7 sonde stations is larger than the estimated combined statistical uncertainty highlights an issue that is most important for the interpretation of tropospheric ozone trends. The trends obtained in various geographical regions are sufficiently different from one another that they are likely measures of geophysical changes in the regions which are the result of

different combinations of cause. We will go through the formality of creating a combined trend in the troposphere, but advise caution on interpreting this trend as a true zonal mean. Specifically, the sonde mean is obtained here by weighting each station inversely by its total variance, including estimates for instrument drift. These uncertainty estimates in the troposphere are largest for the Brewer-Mast ozonesondes in use at the European stations. They are thus accorded less weight in this analysis, while the Canadian ECC sondes are accorded more weight. This weights the mean more towards the negative trends found at the Canadian stations since 1980 and yields a somewhat different result than obtained above in section 3.5. There is a further issue in forming the mean in the troposphere when the results seem to indicate that the processes leading to change are different in different regions. We cannot be sure that the distribution of the locations of the sonde stations is correctly representing the area of the globe which has undergone change due to a particular set of causes. The station distribution may overestimate or underestimate the importance of the negative tropospheric trends found at the Canadian stations to a zonal mean effect.

	Sonde		SAGE		SBUV		Umkehr		Mean	
Alt(km)	Trend	Unc	Trend	Unc	Trend	Unc	Trend	Unc	Trend	Unc
2.5	-1.87	2.48	-	-	-	-	-	-	-1.87	2.48
5.0	-1.55	1.99	-	-	-	-	-	-	-1.55	1.99
7.5	-0.23	2.72	-	-	-	-	-	-	-0.23	2.62
10.0	-3.33	1.84	-	-	-	-	-	-	-3.33	1.84
12.5	-6.40	1.96	-	-	-	-	-	-	-6.40	1.96
15.0	-7.34	2.28	-	-	-	-	-	-	-7.34	2.28
17.5	-6.81	1.68	-	-	-	-	-	-	-6.81	1.68
20.0	-5.99	1.07	-4.90	2.84	-	-	-1.10	2.34	-5.12	0.92
25.0	-3.67	0.51	-1.90	1.39	-0.90	1.71	0.30	2.13	-3.12	0.45
30.0	-	-	-2.20	1.24	-1.70	1.77	-1.90	2.04	-2.01	0.91
35.0	-	-	-5.00	1.03	-3.50	2.41	-6.00	2.04	-4.99	0.86
40.0	-	-	-8.30	1.23	-4.70	2.64	-6.60	2.05	-7.42	0.98
45.0	-	-	-7.20	2.02	-4.00	2.75	-	-	-6.08	1.63
50.0	-	-	-4.50	3.28	-	-	-	-	-4.50	3.28

Table 3.6. Trends and uncertainties in %/decade for each measurement system for northern mid-latitudes. All uncertainties are given as 1s. Derivation of mean trend and uncertainty is described in text.

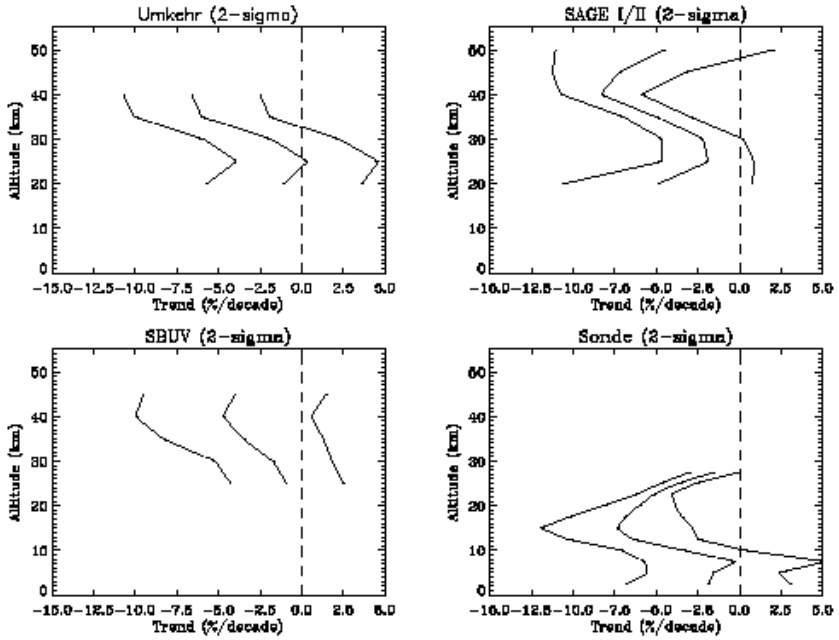


Figure 3.53 Estimates of mean trend and combined statistical and instrument drift uncertainty estimates at northern mid-latitudes for a) Umkehr, b) SAGE I/II, c) SBUV/SBUV2, and d) sondes. Estimates were made at 2.5 km intervals from 2.5 to 20 km and 5 km intervals from 20 to 50 km. Uncertainties shown are 1s. Note that the sonde results were obtained as variance-weighted means. In the troposphere there is an additional unquantified uncertainty about the representativeness of the small number of stations. See text and Tables 3.4-3.6 for details.

The statistical uncertainties and instrument-drift uncertainties for the other measurement systems were also combined by rss. The result for each of the systems for northern mid-latitudes is given in Table 3.6 and shown in the 4 panels of Figure 3.53. This figure shows the mean trend and one standard error calculated for these systems as a function of altitude and the mean trend plus or minus one standard error. All estimates were interpolated to 2.5 km altitude intervals starting at 2.5 km and going through 20 km and 5 km intervals from 20 km through 50 km. Sonde results are shown down to 2.5 km but caution should be exercised in interpreting the meaning of any trends in the troposphere.

Once the estimates for trend and uncertainty were obtained for each system, these were combined into one trend versus altitude plot as shown in Figure 3.54. At altitudes where only one system has measurements, the trend and uncertainty estimates from this system were used. At altitudes where more than one system has measurements, these were combined as a weighted mean and weighted uncertainty using the square of each system's standard error.

The trends shown in Figure 3.54 for the upper stratosphere are dominated by those determined from the SAGE instruments because these have the smallest estimated uncertainty. The result of this estimation of the trend at 40 km is -7.4%/decade ±2.0%/decade, a highly significant trend. The trend at 50 km is estimated from the SAGE instruments only and is not significant at the 2s level because an uncertainty has been estimated at this altitude for SAGE which includes the observed sunrise-sunset differences in ozone from SAGE II.

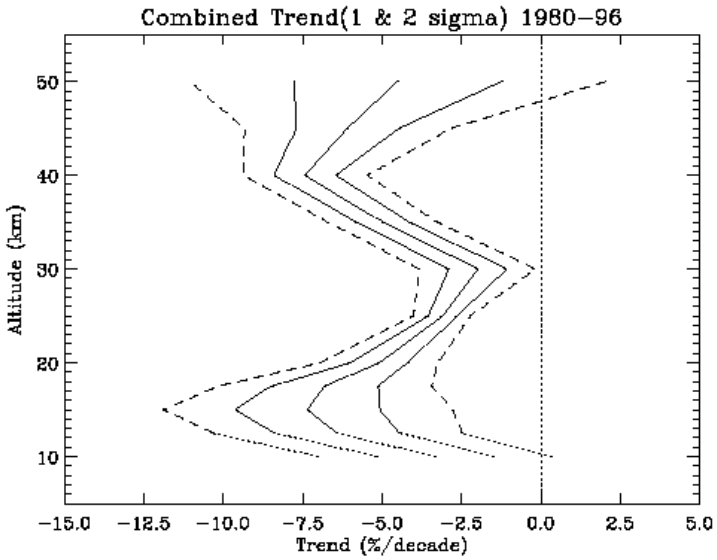


Figure 3.54 Estimate of mean trend using all 4 measurement systems at northern mid-latitudes (heavy solid line). Combined uncertainties are also shown as 1s (light solid line) and 2s (dashed line). Combined trends have not been shown in the troposphere because the small sample of sonde stations have an additional unquantified uncertainty concerning their representativeness of mean trends.

Below 20 km the trend estimate comes from the sondes alone. Data at 12.5 km is generally in the stratosphere while that at 10 km is generally in the troposphere, except during winter. Trends have not been shown in this figure below 10 km because of the issues concerning the unquantified uncertainty in their representativeness in the troposphere. At 20 km the trend is from sondes, SAGE, and Umkehr. The Umkehr trend at 20 km is actually that for layer 4 which is centred a little above 20 km. This trend is quite different from that derived from the sondes or SAGE. The uncertainty estimate is large enough that, after weighting, it only reduces the estimated trend from about -6% to about -5%. It could be argued to not use Umkehr at this low level, but the overall conclusion would not change significantly. The estimated SAGE I/II uncertainties at 20 km are quite large. They are dominated by the uncertainty estimated for correcting the altitude offset of SAGE I. This will need more study if we are to be able to give trends at other latitudes in the lowest stratosphere with reasonably small uncertainties.

It must be emphasised that this is a first attempt to put together an overall trend versus altitude with instrumental and statistical uncertainties. A number of assumptions had to be made to estimate the instrument drift uncertainties. More assumptions had to be made to determine how to combine the various uncertainties. This was done only for the northern mid-latitudes where the data is most plentiful. It is hoped that this exercise will stimulate others to examine the input to these calculations and the assumptions made so that they can improve upon this first attempt.

3.7. Summary

- Statistical models used in the analysis were compared by using 3 test data sets. This comparison revealed only minor differences in trends obtained by the models. Somewhat greater differences were found in the uncertainties estimated for the trends and other variables included in the models. Results are most sensitive to the details of the model for time series with significant missing data.

- Decadal variations are a ubiquitous feature of ozone observations, in addition to QBO and faster time scale dynamical variability. Inclusion of these terms does not have strong influence on the calculated trends for long time series. Much of the observed decadal changes are approximately in phase with the solar cycle for the observational record, suggesting a solar mechanism. There is a strong potential confusion of solar and volcanic signals for the recent record. Although these effects have relatively small impacts on trend estimates, it does limit our ability to interpret decadal variability.

- The upper stratosphere (altitudes between about 30 and 50 km) is a region where changes in ozone were originally predicted to occur. At these altitudes the chemistry should be dominated by gas-phase reactions. When the upper stratospheric data are fit to a standard statistical model, negative trends are found throughout the region with statistically significant peak values of -6 to -8%/decade at 40-45 km altitude. There is good agreement between SAGE I/II and Umkehr. The SBUV/SBUV2 combined record shows less negative trends. We place less confidence in the SBUV/SBUV2 result due to potential problems with the present version (6.1.2) of the NOAA-11 SBUV2 data. There is a factor of two seasonal variation, with a maximum negative trend in winter. There is no significant inter-hemispheric difference in upper stratospheric trends based on SAGE I/II version 5.96 data extended through 1996.

- The lower stratosphere (altitudes between about 10 and 30 km) is the region where we expect much of the trend which has been deduced from column data to occur. The primary trend instruments in this region are sondes (up to 27 km) and SAGE (20-30 km). For sondes, sampling of the data prior to trend analysis has as much or more effect on derived trends than do the details of the statistical model. The optimal selection criteria for the use of sonde data in trend analyses is a subject of debate.

- Trends from 8 individual ozonesonde stations in the northern mid-latitudes are negative throughout the lower stratosphere. For 1970 through 1996 they range from -3 to -10%/decade at 17 km and are statistically significant at all stations. The trends show little seasonal variability above 20 km. The seasonal variability of the trend in the ozone profile occurs mostly in the altitude range of 10-20 km. The exact time dependence of this seasonal variability is uncertain. European stations show a winter-spring maximum, while Canadian stations show a spring-summer maximum. Details of this seasonal maximum were somewhat different in the two analyses of the sonde data.

- There is a reasonable agreement between SAGE I/II trends and sonde trends over the altitude region from 15 to 27 km at northern mid-latitudes (for the time period 1979-1996). This is a significant improvement compared to previous comparisons mainly due to the inclusion of the latitude-dependent altitude correction to the SAGE I data. The agreement in the derived trends from SAGE II only and the sondes is excellent (for the time period 1984-1996). SAGE II trends in the 15-20 km region in the tropics are much more negative than those in northern mid-latitudes but there are insufficient sonde records with which to compare these results.

- It is difficult to make generalisations concerning trends in tropospheric ozone. The only data from

which to make conclusions are sonde data from a small number of stations. Trends calculated for Canadian stations are negative or near zero for the period from 1970 through 1996 and also for the period from 1980 through 1996. Trends calculated for 3 European stations are strongly positive for the period 1970 through 1996 but are essentially zero at two of these stations when data from 1980 through 1996 are considered. Trends calculated for the Japanese stations are strongly positive for the period from 1970 through 1996 but are either positive or not significantly different from zero over the period from 1980 through 1996.

- Trends in the column amount of ozone above 20 km deduced from SAGE I/II are much less than the column trends deduced from TOMS. The TOMS/SAGE differences are consistent with the sonde trends below 20 km. There is also a consistent seasonal variation for satellite and sonde data. Both indicate that the primary seasonal variation in mid-latitude ozone trends occurs between 10 and 20 km altitude, with maximum during northern hemisphere winter and spring.

- A first attempt has been made to combine the trends and uncertainties estimated from all available measurement systems. This has been done for northern mid-latitude measurements. The result is a statistically significant (at the 2s level) negative trend at all altitudes between 10 and 45 km. The combined trend has two local extremes, -7.4 ±2.0%/decade at 40 km and -7.6 ±4.6%/decade at 15 km altitude. A minimum trend of -2.0 ±1.8%/decade was deduced at 30 km altitude.

Acknowledgement

We thank the many contributors to this document who are not listed as Panel Members. In particular we thank the Chicago group, Tiao, Choi, and Zhang, and Inna Megretskaia of Harvard for the extensive calculations that they performed on trends using sonde data. We thank Gordon Labow for his work in preparing data sets to make the calculation of trends much easier. Finally, we thank the members of the Chapter 1 and 2 panels for the extensive work that resulted in improved data sets for analysis and improved ability to estimate uncertainties in the trends.

3.8. References

Aellig, C.P. et al., Stratospheric chlorine partitioning: constraints from shuttle-borne measurements of [HCl],[ClNO3], and [ClO], Geophys. Res. Lett. 17, 2361-2364, 1996.

Ahmad, Z., et al., "Accuracy of Total Ozone Retrieval from NOAA SBUV/2 Measurements: Impact of Instrument Performance," J. Geophys. Res., 99, 22975-22984, 1994.

Akimoto, H., H. Nakane, and YU. Matsumoto, The chemistry of oxidant generation: troposheric ozone increase in Japan. accepted for publication in, Chemistry of the Atmosphere, The Impact on Global Change, ed. J. W. Birks, J. G. Calvert, R. E. Sievers, pub. American Chemical Society, Washington, DC, 1993.

Angell, J.K., On the relation between atmospheric ozone and sunspot number. J. Climate, 2, 1404-1416, 1989.

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

Bekki, S., and J.A. Pyle, A two-dimensional modeling study of the volcanic eruption of Mount Pinatubo. J. Geophys. Res., 99, 18861-18869, 1994.

Bhartia, P.K., et al., Algorithm for the Estimation of Vertical Ozone Profiles from the Backscattered Ultraviolet Technique, J. Geophys. Res., 101, 18793-18806, 1996.

Bojkov, R. D. and V. E. Fioletov, Change of the lower stratospheric ozone over Europe and Canada, J. Geophys. Res., 102, 1337-1347, 1997.

Bojkov, R., L. Bishop, W. J. Hill, G. C. Reinsel, and G. C. Tiao, A Statistical Trend Analysis of Revised Dobson Total Ozone Data Over the Northern Hemisphere, J. Geophys. Res., 95, 9785-9807, 1990.

Bojkov, R. D., C. S. Zerefos, D. S. Balis, I. C. Ziomas, and A. F. Bias, Record low total ozone during northern winters of 1992 and 1993, Geophys. Res. Lett. 20, 1351-1354, 1993.

Bojkov, R. D., L. Bishop, and V. E. Fioletov, Total ozone trends from quality-controlled ground-based data (1964-1994), J. Geophys. Res., 100, 25867-25876, 1995.

Bowman, K.P., Global patterns of the quasi-biennial oscillation in ozone, J. Atmos. Sci., 46, 3328-3343, 1989.

Box, George E. P., Gwilym M. Jenkins and Gregory. C. Reinsel, Time Series Analysis: Forecasting and Control, Third Edition, Prentice Hall, Englewood Cliffs NJ, 1994.

Brasseur, G.P., The response of the middle atmosphere to long-term and short-term solar variability: A two-dimensional model, J. Geophys. Res., 98, 23079-23090, 1993.

Chandra, S. and R.D. McPeters, The solar cycle variation of ozone in the stratosphere inferred from Nimbus 7 and NOAA 11 satellites, J. Geophys. Res., 99, 20665-20671, 1994.

Chandra, S., and R. S. Stolarski, Recent trends in stratospheric total ozone: Implications of dynamical and ElChichon perturbations, Geophys. Res. Lett., 18, 2277-2280, 1991.

Chandra, S., C. H. Jackman, and E. L. Fleming, Recent trends in ozone in the upper stratosphere: Implications for chlorine chemistry, Geophys. Res. Lett., 22, 843-846, 1995.

Chipperfield, M.P., L.J. Gray, J.S. Kinnersley and J. Zawodny, A two dimensional model study of the QBO signal in SAGE II NO2 and O3, Geophys. Res. Lett., 21, 589-592, 1994.

Coffey, M.P., Observations of the impact of volcanic activity on stratospheric chemistry, J. Geophys. Res., 101, 6767-6780, 1996.

Deland, M., and R.P. Cebula, The composite Mg II solar activity index for solar cycles 21 and 22, J. Geophys. Res., 98, 12809-12823, 1993.

DeLuisi, J. J., C. L. Mateer, D. Theisen, P. K. Bhartia, D. Longenecker, and B. Chu, Northern middle-latitude ozone profile

features and trends observed by SBUV and Umkehr, 1979-1990, J. Geophys. Res., 99, 18,901-18,908, 1994.

Efron, B. and R.J. Tibshimani, An Introduction to the Bootstrap, Chapman and Hall, New York, 436 pp., 1993.

Fleig, A.J., et al., Nimbus 7 Solar Backscatter Ultraviolet (SBUV) Ozone Products User's Guide, NASA Publication 1234, 1990.

Garcia, R.R., and S. Solomon, A possible relationship between interannual variability in Antarctic ozone and the quasi-biennial oscillation, Geophys. Res. Lett., 14, 848-851, 1987.

Gleason, J.F., and R.D. McPeters, Corrections to Nimbus-7 solar backscatter ultraviolet data in the "nonsync" period (February 1987 to June 1990), J. Geophys. Res., 100, 16873-16877, 1995.

Gleason, J.F. et al., Record low global ozone in 1992, Science, 260, 523-526, 1993.

Grant, W.B., et al., Aerosol induced changes in tropical stratospheric ozone following the eruption of Mount Pinatubo, J. Geophys. Res., 99, 8197-8211, 1994.

Haigh, J.D., The role of stratopsheric ozone in modulating the solar radiative forcing of climate, Nature, 370, 544-546, 1994.

Hamilton, K.P., Interhemispheric asymmetry and annual synchronization of the ozone quasi-biennial oscillation, J. Atmos. Sci., 46, 1019-1025, 1989.

Harris, N. R. P., G. Ancellet, L. Bishop, D. J. Hoffman, J. B. Kerr, R. D. McPeters, M. Prendez, W. J. Randel, J. Staehelin, B. H. Subbaraya, A. Volz-Thomas, J. Zawodny, and C. S. Zerefos, Trends in stratospheric and free tropospheric ozone, J. Geophys. Res., 102, 1571-1590, 1997.

Hasebe, F., Interannual variations of global ozone revealed from Nimbus 4 BUV and ground based observations, J. Geophys. Res., 88, 6819-6834, 1983.

Hasebe, F., Quasi-biennial oscillations of ozone and diabatic circulation in the equatorial stratosphere, J. Atmos. Sci., 51, 729-745, 1994.

Heath, D., and B. Schlesinger, The Mg 280 nm doublet as a monitor of changes in solar ultraviolet irradiance, J. Geophys. Res., 91, 8672-8682, 1986.

Hilsenrath, E., et al., 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.J., and S. Solomon, Ozone destruction through heterogeneous chemistry following the eruption of El Chichon, J. Geophys. Res., 94, 5029-5041, 1989.

Hofmann, D.J., et al., Ozone loss in the lower stratosphere over the United States in 1992-1993: Evidence for heterogeneous chemistry on the Pinatubo aerosol, Geophys. Res. Lett., 21, 65-68, 1994.

Hollandsworth, S. M., R. D. McPeters, L. E. Flynn, W. Planet, A. J. Miller, and S. 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., The solar cycle variation of total ozone: Dynamical forcing in the lower stratosphere, J. Geophys. Res., 102, 1355-1370, 1997.

Hood, L.L., and J.P. McCormack, Components of interannual ozone change based on Nimbus 7 TOMS data, Geophys. Res. Lett., 19, 2309-2312, 1992.

Hood, L.L., and D. Zaff, Lower stratospheric stationary waves and the longitude dependance of ozone trends in winter, J. Geophys. Res., 100, 25791-25802, 1995.

Hood, L.L., et al., Altitude Dependence of Stratospheric Ozone Trends Based on Nimbus 7 SBUV Data, Geophys. Res. Lett., 20, 2667-2670, 1993.

Hood, L.L., J.P. McCormack and K. Labitzke, An inverstigation of dynamical contributions to midlatitude ozone trends in winter, J. Geophys. Res., 102, 13079-13093, 1997.

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.

Kaye, J. A. and R. B. Rood, Chemistry and transport in a three-dimensional stratospheric model: chlorine species during a simulated stratospheric warming, J. Geophys. Res., 94, 1057-1083, 1989.

Kerr, J.B., D.I. Wardle and D.E. Tarasick, Record low ozone values over Canada in early 1993, Geophys. Res. Lett., 20, 1979-1982, 1993.

Lait, L.R., M.R. Schoeberl and P.A. Newman, Quasi-biennial modulation of the Antarctic ozone depletion, J. Geophys. Res., 94, 11559-11571, 1989.

Logan, J.A. Tropospheric ozone: seasonal behavior, trends and anthropogenic influence, J. Geophys. Res., 90, 10,463-10,482, 1985.

Logan, J.A., Trends in the vertical distribution of ozone: An analysis of ozonesonde data, J. Geophys. Res., 99, 25553-25585, 1994.

McMormack, J.P. and L.L. Hood, Apparent solar cycle variations of upper stratospheric ozone and temperature: Latitude and seasonal dependencies, J. Geophys. Res., 101, 20,933-20,944, 1996.

McCormack, J.P., and L.L. Hood, Modeling the spatial distribution of total ozone in northern hemisphere winter: 1979-1991, J. Geophys. Res., 102, 13711-13717, 1997.

McCormack, J.P., L.L. Hood, R. Nagatani, A.J. Miller, W.G. Planet and R.D. McPeters, 1997: Approximate separation of volcanic and 11-year signals in the SBUV-SBUV/2 total ozone record over the 1979-1995 period, Geophys. Res. Lett., 1997, in press.

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

McPeters, R.D., S.M. Hollandsworth, L.E. Flynn, J.R. Herman and C.J. Seftor, Long-term ozone trends derived from the 16-year combined Nimbus 7 - Meteor 3 TOMS Version 7 record, Geophys. Res. Lett., 23, 3699-3702, 1996.

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. D(tsch, J. Staehelin, and J. J. DeLuisi, Influence of a priori profiles on trend calculations from Umkehr data, J. Geophys. Res., 101, 16779 - 16787, 1996.

Miller, A.J., R.M. Nagatani, G.C. Tiao, X.F. Niu, G.C. Reinsel, D. Wubbles and K. Grant, Comparisons of observed ozone and temperature trends in the lower stratosphere. Geophys. Res. Lett., 19, 929-932, 1992.

Miller, A. J., G. C. Tiao, G.C. Reinsel, D. Wuebbles, L. Bishop, J. Kerr, R.M. Nagatani, J.J. deLuisi, and C. L. Mateer, Comparisons of observed ozone trends in the stratosphere though examination of Umkehr and balloon ozonesonde data, J. Geophys. Res., 100, 11,209-11218, 1995.

Miller, A.J., S.M. Hollandsworth, L.E. Flynn, G.C. Tiao, G.C. Reinsel, L. Bishop, R.D. McPeters, W.G. Planet, J.J. DeLuisi, C.L. Mateer, D. Wuebbles, J. Kerr, and R.M. Nagatani, Comparisons of observed ozone trends and solar effects in the stratosphere through examination of ground-based Umkehr and combined solar backscattered ultraviolet (SBUV) and SBUV 2 satellite data, J. Geophys. Res., 101, 9017-9021, 1996.

Miller, A.J., L.E. Flynn, S.M. Hollandsworth, J.J. DeLuisi, I.V. Petropavlovskikh, G.C. Tiao, G.C. Reinsel, D.J. Wuebbles, J. Kerr, R.M. Nagatani, L. Bishop, and C.H. Jackman, Information content of Umkehr and solar backscattered ultraviolet (SBUV) 2 satellite data for ozone trends and solar responses in the stratosphere, J. Geophys. Res., 102, 19,257-19263, 1997.

Miller, R. G., The jackknife- a review, Biometrika, 61, 1, 1974

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 tropspheric abundance of halogen from halocarbons: implications for stratopheric 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, 14029-14042, 1995.

Oltmans, S.J., and J. London, The quasi-biennial oscillation in atmospheric ozone, J. Geophys. Res., 87, 1567-1586, 1982.

Oltmans, S. J. and 17 others, Trends of ozone in the troposphere, Geophys. Res. Lett., 25, 139-142, 1998.

Paur, R. J., and A. M. Bass, The ultraviolet cross-section of ozone: II. Results and temperature dependence, Atmospheric ozone, Zerefos, C.S. and Ghazi, A. (Eds), p. 611, Reidel, Dordrecht, 1985.

Peters, D., and G. Entzian, January ozone anomaly over the North Atlantic - European region: Longitude dependant decadal change in total ozone during 1979-1992, Meteorol. Z.N.F., 5, 41-45, 1996.

Ramaswamy, V., M.D. Schwarzkopf and W.J. Randel, Fingerprint of ozone depletion in the spatial and temporal pattern of recent lower stratospheric cooling, Nature, 382, 616-618, 1996.

Randel, W.J. and J.B. Cobb, Coherent variations of monthly mean column ozone and lower stratospheric temperature, J. Geophys. Res., 99, 5433-5447, 1994.

Randel, W.J., and F. Wu, Isolation of the ozone QBO in SAGE II data by singular value decomposition, J. Atmos. Sci., 53, 2546-2559, 1996.

Randel, W.J., F. Wu, J.M. Russell III, J. W. Waters and L. Froidevaux, Ozone and temperature changes in the stratosphere following the eruption of Mt. Pinatubo, J. Geophys. Res., 100, 16753-16764, 1995.

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.

Rosenfield, J.E., D.B. Considine, P.E. Meade, J.T. Bacmeister, C.H. Jackman and M.R. Schoeberl, Stratospheric effects of Mount Pinatubo aerosol studied with a coupled two-dimensional model, J. Geophys. Res., 102, 3649-3670, 1997.

Shiotani, M., Annual, quasi-biennial and El Nino Southern Oscillation (ENSO) timescale variations in equatorial total ozone, J. Geophys. Res., 97, 7625-7633, 1992.

Solomon, S. and R. R. Garcia, On the distribution of long-lived tracers and chlorine species in the middle atmosphere, J. Geophys. Res., 89, 11633-11644, 1984.

Solomon, S., R.W. Portman, 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.

Spencer, R.W., and J.R. Christy, Precision lower stratospheric monitoring with the MSU: Technique, validation and results 1979-1991, J. Climate, 6, 1194-1204, 1993.

Tarasick, D. W., D. I. Wardle, J. B. Kerr, J. J. Bellefleur, and J. Davis, Tropospheric ozone trends over Canada: 1980-1993, Geophys. Res. Lett., 22, 409-412, 1995.

Thomason, L.W., L.R. Poole and T. Deshler, A global climatology of stratospheric aerosol surface area density deduced from Stratospheric Aerosol and Gas Experiment II measurements: 1984-1994, J. Geophys. Res., 102, 8967-8976, 1997.

Tiao, G.C., G.C. Reinsel, J.H. Pedrick, G.M. Allenby, C.L. Mateer, A.J. Miller, and J.J. DeLuisi, A statistical analysis of ozonesonde data, J. Geophys. Res., 91, 13,121-13,136, 1986.

Tie, X. X., G. P. Brasseur, B. Breiglib and C. Granier, Two dimensional simulation of Pinatubo aerosol and its effect on stratospheric chemistry, J. Geophys. Res., 99, 20545-20562, 1994.

Tie,X. X. and G. Brasseur, The response of stratospheric ozone to volcanic eruptions: sensitivity to atmospheric chlorine loading, Geophys. Res. Lett., 22, 3035-3038, 1995.

Torres, O., and P.K. Bhartia, Effect of Stratospheric Aerosol on Ozone Profile from BUV Measurements, Geophys. Res. Lett., 22, 235-238, 1995.

Tung, K.K., and H. Yang, Global QBO in circulation and ozone. Part I: Reexamination of observational evidence, J. Atmos. Sci., 51, 2699-2707, 1994.

Wang, W-C., Y-C. Zhuang, and R. D. Bojkov, Climate implications of observed changes in ozone vertical distributions at middle and high latitudes of the northern hemisphere, Geophys. Res. Lett., 20, 1567-1570, 1993.

Wang, H.J., D.M. Cunnold, and X. Bao, A critical analysis of stratospheric aerosol and gas experiment ozone trends, J. Geophys. Res., 101, 12495-12514, 1996.

Waters, J.W., et al., Validation of UARS Microwave Limb Sounder ClO measurements, J. Geophys. Res., 101, 10091-10127, 1996.

WMO, Scientific Assessment of Stratospheric Ozone: 1989, World Meteorological Organization Global Ozone Research and Monitoring Project - Report No. 20, Geneva, 1990.

WMO, Scientific Assessment of Ozone Depletion: 1994, World Meteorological Organization Global Ozone Research and Monitoring Project - Report No. 37, Geneva, 1995.

Yang, H. and K.K. Tung, Statistical significance and pattern of extratropical QBO in column ozone, Geophys. Res. Lett., 21, 2235-2238, 1994.

Zawodny, J.M., and M.P. McCormick, Stratospheric Aerosol and Gas Experiment II measurements of the quasi-biennial oscillation in ozone and nitrogen dioxide, J. Geophys. Res., 96, 9371-9377, 1991.

Zerefos, C.S., K. Tourpali, R.D. Bojkof, D.S. Balis, B. Rognerund and I.S. Isaksen, Solar acticity - total ozone relationships: observations and model studies with heterogeneous chemistry, J. Geophys. Res., in press, 1997.

Ziemke, J. R., S. Chandra, R.D. McPeters and P.A. Newman, Dynamical proxies of column ozone with applications to global trend models, J. Geophys. Res., 102, 6117-6129, 1997.

3.9. Appendix

Description of statistical models

A varying amount of detail was made available regarding the statistical models used by the researchers in section 3.3 on the test data sets, and in section 3.5 on trend results. This section encapsulates the model information available. In some cases the particular model fitted impacts the trend results, as discussed in sections 3.3 and 3.5.

The models used for the primary text discussion on trends (section 3.5) are described more completely below. Following that, all model information is summarised in bullet form.

Logan and Megretskaia (Harvard model, sondes)

The model includes monthly means, four seasonal trends, a lagged QBO and a solar dependence. The monthly means are weighted by the inverse of the interannual monthly variance. The variances are computed with two iterations after fitting the model, to remove the variation due to the trend. The model does not account for autoregression. An intervention term is included in the model for the four Canadian stations at all pressure levels, at the date of the change from BM to ECC sondes; a similar term is included at Payerne for the tropospheric levels in April 1977 to account for the change in the time of soundings from ~1600 to 0930 (Logan, 1985). The QBO time series used is the 30 hPa winds for Singapore and the 10.7 cm solar radio flux at Ottawa/Penticton. Lags were determined for the QBO as described in Logan (1994). Percentage trends are calculated relative to the seasonal mean for the time series being analysed.

Tiao, Choi and Zhang (Chicago model, sondes)

The model includes monthly means, four seasonal trends, a lagged QBO, a solar dependence, and an intervention term as above for the Canadian stations and for Payerne. The residual noise is modelled as a first order autoregressive model, with different variances in different seasons. The QBO time series is the average of 50 hPa winds at Singapore, Balboa, and Ascension Island. The model assumes zero trend before Jan. 1, 1970. Outliers are removed; these are points whose residuals are more than three standard deviations away from the model fit. Percentage trends are calculated relative to the seasonal intercept in 1970 (or beginning of series if later) adjusted for solar effects, and intervention if used.

Hollandsworth and Flynn (SBUV and TOMS total)

The basic regression model for SBUV trend analysis in this report is of the form:

$$O_3(t) = m(t) + A(t)X(t) + B(t)X_{qbo}(t-l) + C(t)X_{solar}(t) + N(t)$$

This model is fitted to weekly ozone data in section 3.5, but monthly data for the trend comparisons on the test data sets in section 3.3. Here $m(t)$ is the seasonal cycle, represented by a constant and 4 harmonics (9 terms); $A(t)$ is the seasonally-varying trend coefficient. The trend proxy $X(t)$ is time (time=0 at 1/1/78) and the coefficient varies with 2 seasonal harmonics (5 terms); $B(t)$ is the seasonally-varying QBO coefficient. The QBO proxy is the 30-hPa monthly-mean zonal-mean Singapore wind, lagged in time with respect to the ozone values by up to a quarter cycle (backward or forward) to produce the minimum residuals. The QBO lags are a function of latitude and altitude. The coefficient varies with one seasonal harmonic (3 terms); C(t) is the seasonally-varying 11-year solar cycle coefficient. The solar cycle proxy is the 5-week smoothed 10.7 cm solar radio flux. The coefficient varies with one seasonal harmonic (3 terms). The full model has 20 terms. There is no seasonal weighting, and autoregressive

noise is taken into account in the standard errors via a jack-knife refitting method.

Newchurch and Yang (Umkehr)

The trend model used to estimate the annual ozone trends from the Umkehr data is a uniform (non-seasonal) trend model of the form:

$$O_3(t) = m + cs_{12} + cs_6 + cs_4 + cs_3 + b.t + g_1 Z_1(t) + g_2 Z_2(t-k) + N(t)$$

where:

μ = ozone monthly mean
$\beta \cdot t$ = trend (no monthly variation)
$cs_{12} = c_{12}\cos(2\pi t/12) + s_{12}\sin(2\pi t/12)$
$cs_6 = c_6\cos(2\pi t/6) + s_6\sin(2\pi t/6)$
$cs_4 = c_4\cos(2\pi t/4) + s_4\sin(2\pi t/4)$
$cs_3 = c_3\cos(2\pi t/3) + s_3\sin(2\pi t/3)$
$\gamma_1 Z_{1,t}$ = 10.7 cm solar flux effect
$\gamma_2 Z_2(t-k)$ = Effect of Singapore wind lagged k months
$N(t)$ = residual noise series

The model is an AR(1) autoregressive one, fit using multiple regression, and with the AR(1) parameter varying by month.

For the test data sets, however, Newchurch and Yang fit a monthly trend model for the TOMS and Uccle data for comparison to other researchers' results. For reasons discussed in the text, they reported results from a seasonal trend model (based on fitting seasonal ozone averages) for the SAGE test data. In both cases, the solar and QBO effects were omitted for the test data sets.

Wang and Cunnold (SAGE trends)

The SAGE trend estimates result from the statistical model of Wang *et al*. (1996) with two significant differences. First, the model used here includes a seasonal modulation term where the Wang *et al*. model assumed none. Second, this model uses an exogenous series (Singapore winds) for the QBO coefficient, where the Wang *et al*. model used a fixed 28-month period. The linear regression model for the ozone trend is then:

$$O_3(t) = A.t + Seasonal + C.QBO + D.Solar + N(t)$$

where:

$O_3(t)$ = monthly mean ozone
$A = A_1 + A_2\cos(wt) + A_3\sin(wt) + A_4\cos(2wt) + A_5\sin(2wt)$
Seasonal $= B_1 + B_2\cos(wt) + B_3\sin(wt) + B_4\cos(2wt) + B_5\sin(2wt)$
QBO = lagged Singapore wind at 30mb (C is a constant)
Solar = MgII index (except 10.7 cm flux for Uccle test set; D is a constant)
$N(t)$ = AR(1) autoregressive residual

Note: the QBO phase lag is determined by the maximum correlation between the 30 hPa Singapore zonal wind and the residual N(t) (fitted without QBO term)

Model summary

For easy comparison, here are the main features of the trend models in bullet form (ordered alphabetically by first researcher):

Bishop:

- seasonal cycle: 12 monthly means (intercepts at 1970 or beginning of series if later)
- seasonal trend: 12 monthly trends
- solar: 10.7 cm flux (1 term)
- QBO: average 50 hPa winds at Singapore, Ascension and Balboa (1 term), lag determined as that most statistically significant
- autoregressive AR(1), full maximum likelihood estimation
- monthly weights based on variance of N_t residuals
- covariance matrix of monthly trends used to calculate standard errors of seasonal and year round trends

Fioletov:

- seasonal cycle: 12 monthly means
- seasonal trend: 12 monthly trends
- solar: 10.7 cm flux (1 term)
- QBO: normalised equatorial wind at 30 and 50 hPa (2 terms)
- autoregressive AR(1), conditional least squares
- monthly weights based on variance of N_t residuals
- covariance matrix of monthly trends used to calculate standard errors of seasonal and year round trends

Hollandsworth and Flynn:

- fit to weekly ozone (except test data sets, which were fit to monthly ozone)
- seasonal cycle: constant and 4 harmonics (9 terms)
- seasonal trend: constant and 2 harmonics (5 terms)
- solar: 10.7 cm flux, effect varies by season (3 terms), 5-week smooth
- QBO: lagged smoothed Singapore 30 hPa wind, effect varies by season (3 terms)
- not autoregressive, but comparable errors determined by statistical jack-knife methods
- not seasonally weighted

Logan and Megretskaia (Harvard model):

- seasonal cycle: 12 monthly means
- seasonal trend: 4 seasonal trends
- solar: 10.7 cm flux (1 term)
- QBO: lagged 30 hPa winds at Singapore
- not autoregressive
- monthly weights based on variance of N_t residuals
- intervention terms for sonde change at Canadian stations
- covariance matrix of seasonal trends used to calculate the standard error of year round trend

McCormack and Hood:

- form of model not known, but appears not to be seasonally weighted

Newchurch and Yang:

- Umkehr trends
- monthly model
- seasonal cycle: constant and 4 harmonics (9 terms)
- uniform, non-seasonal, trend (1 term)
- solar: 10.7 cm flux
- QBO: lagged Singapore 30 hPa wind
- stratospheric optical depth series at 320 nm
- AR(1), with parameter varying by month
- TOMS and Uccle test data sets
- monthly model
- seasonal cycle: 12 monthly means (intercepts)
- seasonal trends: 12 monthly trends
- solar, QBO, optical depth, not included
- AR(1), with parameter varying by month
- SAGE test data set
- model based on seasonal ozone averages
- seasonal cycle: 4 seasonal means (intercepts)
- seasonal trends: 4 seasonal trends
- solar, QBO, optical depth, not included
- AR(1), with parameter varying by season

Randel:

- seasonal cycle: subtraction of 12 monthly means to form deseasonalised anomalies
- seasonal trend: constant and 2 harmonics (5 terms)
- solar: 10.7 cm flux (constant and 2 harmonics)
- QBO: linear combination of equatorial winds over 70-10 hPa (constant and 2 harmonics)
- standard errors calculated using bootstrap resampling technique (effects of autocorrelation are accounted for)

Reinsel (only TOMS test data was submitted):

- same as Bishop, except:
- weights are applied to e_i, not N_i, series, and only 5 degrees of freedom for weights are used (similar months are grouped)

Tiao, Choi and Zhang (Chicago model):

- seasonal cycle: 12 monthly means (intercepts)
- seasonal trend: 4 seasonal trends
- hockey stick model beginning at 1970
- solar: 10.7 cm flux (1 term)
- QBO: lagged average 50 hPa winds at Singapore, Ascension and Balboa (1 term)
- AR(1) autoregressive
- weights based on variance of ϵ_t residuals
- intervention terms for sonde change at Canadian stations
- covariance matrix of seasonal trends used to calculate the standard error of year round trend
- outliers are removed

Wang and Cunnold:

- seasonal cycle: constant and 2 harmonics (5 terms)
- seasonal trend: constant and 2 harmonics (5 terms)
- solar: MgII index for TOMS and SAGE, 10.7 cm flux for Uccle
- QBO: lagged Singapore wind at 30 hPa
- autoregressive AR(1)

Ziemke and Chandra:

- seasonal cycle: constant and 4 harmonics (9 terms)
- seasonal trend: constant and 3 harmonics (7 terms)
- solar: 10.7 cm flux, constant and 2 harmonics (5 terms)
- QBO: 30 hPa Singapore wind, constant and 2 harmonics (5 terms)
- not autoregressive
- standard error analysis not clear

Appendix I

Co-chairs, Authors, Contributors and Reviewers

Co-chairs

Neil Harris	European Ozone Research Cooordinating Unit	UK
Robert Hudson	University of Maryland	US

Chapter 1

Lead Authors

David Hofmann	NOAA Climate Monitoring and Diagnostics Laboratory	US
Clive Rodgers	University of Oxford	UK

Authors

Dirk DeMuer	Royal Meteorological Institute (KMI)	Belgium
Lucien Froidevaux	NASA Jet Propulsion Laboratory	USA
Sophie Godin	Service d'Aeronomie, CNRS	France
Larry Gordley	GATS inc., VA	US
Ernest Hilsenrath	NASA Goddard Space Flight Center	US
Rich McPeters	NASA Goddard Space Flight Center	US
Mike Newchurch	University of Alabama	US
Dave Rusch	University of Colorado	US
Sam Oltmans	NOAA Climate Monitoring and Diagnostics Laboratory	US
Joe Zawodny	NASA Langley Research Center	US

Chapter 2

Lead Authors

James M. Russell III	Hampton University	US
Herman G. J. Smit	Forschungszentrum Jülich	Germany

Co-Authors

Roger J. Atkinson	Meteorology CRC, Bureau of Meteorology	Australia
Brian J. Connor	National Institute of Water and Atmospheric	New

	Research	Zealand
Derek M. Cunnold	Georgia Institute of Technology	US
Lawrence E. Flynn	NOAA National Environmental Satellite, Data and Information Services	US
Lucien Froidevaux	NASA Jet Propulsion Laboratory / Caltech	US
Sophie Godin	Service d'Aeronomie, CNRS	France
Stacey Hollandsworth	Space Applications Corporation	US
James B. Kerr	Atmospheric Environment Service	Canada
Richard D. McPeters	NASA Goddard Space Flight Center	US
A. J. Miller	NOAA National Centers for Environmental Prediction	US
Gary A. Morris	University of Maryland, Baltimore	US
Mike Newchurch	University of Alabama Huntsville	US
David W. Rusch	University of Colorado	US
Larry W. Thomason	NASA Langley Research Center	US
Ray H. J. Wang	Georgia Institute of Technology	US

Contributors

Gerard Ancellet	Service d'Aeronomie, CNRS	France
Matthias Beekmann	Service d'Aeronomie, CNRS	France
Greg Bodeker	National Institute of Water and Atmospheric Research	New Zealand
Rumen Bojkov	World Meteorological Organization	Switzerland
Ellen J. Brinksma	Vrije Universiteit, Amsterdam	Netherlands
Hans Claude	Deutscher Wetterdienst	Germany
Hugo De Backer	Institut Royal Meteorologique de Belgique	Belgium
Pierre Jeannet	Schweizerische Meteorologische Anstalt	Switzerland
Bryan Johnson	NOAA Climate Monitoring and Diagnostics Laboratory	US
N. Kämpfer	Universität Bern	Switzerland
Gordon J. Labow	NASA Goddard Space Flight Center	US
Ulf Köhle	Deutscher Wetterdienst	Germany
I. Stuart McDermid	NASA Jet Propulsion Laboratory / Caltech	US
Thomas McGee	NASA Goddard Space Flight Center	US
Hideaki Nakane	National Institute for Environmental Studies	Japan
Samuel Oltmans	NOAA Climate Monitoring and Diagnostics Lab.	US
Francis J. Schmidlin	NASA Goddard Space Flight Center	US
Johannes Staehelin	Eidgenossische Technische Hochschule, Zurich	Switzerland

Wolfgang Steinbrecht	Deutscher Wetterdienst	Germany
Daan P.J. Swart	Rijksinstituut voor Volksgezondheid en Milieu	Netherlands
David Tarasick	Atmospheric Environment Service	Canada
Osamu Uchino	Japan Meteorological Agency	Japan
Pierre Viatte	Schweizerische Meteorologische Anstalt	Switzerland
Joseph M. Zawodny	NASA Langley Research Center	US

Chapter 3

Lead Authors

Richard Stolarski	NASA Goddard Space Flight Center	US
William Randel	National Center for Atmospheric Research	US

Co-Authors

Lane Bishop	Allied Signal, Inc.	US
Derek Cunnold	Georgia Institute of Technology	US
Dirk DeMuer	Institut Royal Météorologique de Belgique	Belgium
Sophie Godin	Service d'Aéronomie, CNRS	France
Dave Hofmann	NOAA Climate Monitoring and Diagnostics Laboratory	US
Stacey Hollandsworth	Applied Research Corporation	US
Lon Hood	University Of Arizona	US
Jennifer Logan	Harvard University	US
Jim Miller	NOAA / National Center for Environmental Prediction	US
Mike Newchurch	University of Alabama	US
Sam Oltmans	NOAA Climate Monitoring and Diagnostics Laboratory	US
Osamu Uchino	Meterological Research Institute	Japan
Ray Wang	Georgia Institute of Technology	US

Reviewers

Roger Atkinson	Atmosphere Watch Section	Australia
John Burrows	University of Bremen	Germany
Arthur Downey	Atmosphere Watch Section	Australia
Herb Fischer	Forschungszentrum Karlsruhe	Germany
John Gille	National Center for Atmospheric Research	US
Rod Jones	University of Cambridge	UK
Brian Kerridge	Rutherford Appleton Laboratory	UK
Richard	National Institute of Water and Atmospheric	New

McKenzie	Research	Zealand
Howard Roscoe	British Antarctic Survey	UK
Yasuhiro Sasano	National Institute for Environmental Studies	Japan.
Susan Solomon	NOAA	US
Geraint Vaughan	University of Wales	UK

Appendix II

Acronyms and Abbreviations

a.s.l. above sea level

AES Atmospheric Environment Service, Canada

ANF Albedo Normalisation Factor

AO annual oscillation

ASP Aerological Station, Payerne, Switzerland

ATMOS Atmospheric Trace Molecule Spectroscopy

BM Brewer-Mast ozonesondes

CF correction factor

CM coordinate mapping

CMDL Climate Monitoring and Diagnostics Laboratory, NOAA, USA

CNRS Centre National de la Recherche Scientifique, France

DIAL Differential Absorption Laser Technique

DU Dobson units

DWD Deutscher Wetterdienst

ECC Electrochemical Concentration Cell ozonesondes

ERB Earth Radiation Budget Satellite

FOV field of view

GAW Global Atmosphere Watch

GLONET Global Ozone Network

GSFC Goddard Space Flight Center

HALOE HALogen Occultation Experiment

IFOV instantaneous field of view

IMD Indian Meteorology Department

IOC International Ozone Commission

IRR inter-range ratio

JMA Japanese Meteorological Agency

JOSIE Jülich Ozone Sonde Intercomparison Experiment

JPL Jet Propulsion Laboratory

KFA Research Centre, Jülich, Germany

Lidar LIght Detection And Ranging

MLO3 Mauna Loa Ozone

MLS Microwave Limb Sounder

MOHP Meteorological Observatory Hohenpeissenberg, Germany

MRI Meteorological Research Institute, Japan

MSU Microwave Sounding Unit

NASA National Aeronautics and Space Administration, USA

NCEP National Center for Environmental Prediction (formerly NMC) USA

NDSC Network for Detection of Stratospheric Change

NIES National Institute for Environmental Studies, Japan

NIST National Institute of Standards and Technology (formerly NBS) USA

NIWA National Institute of Water and Atmosphere, New Zealand

NOAA National Oceanic and Atmospheric Administration, United States

NWS National Weather Service, USA

OHP Observatoire de Haute Provence

OPAL Ozone Profiler Assessment at Lauder

QBO Quasi-Biennial Oscillation

RIVM Dutch Institute of Public Health and the Environment

rmsres root mean square residual

rss root sum of the squares

SA Service D'Aéronomie, CNRS

SAGE Stratospheric Aerosol and Gas Experiment

SAO semi-annual oscillation

SBUV Solar Backscatter Ultraviolet Instrument

sem standard error of the mean

SPARC Stratospheric Processes and their Role in Climate

SSBUV Space Shuttle SBUV

STOIC Stratospheric Ozone Intercomparison Campaign

TM trajectory mapping

TMF Table Mountain Facilities

TOMS Total Ozone Monitoring Spectrometer

UARS Upper Atmosphere Research Satellite

WOUDC World Ozone and Ultraviolet Radiation Data Centre