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 (2s) 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.
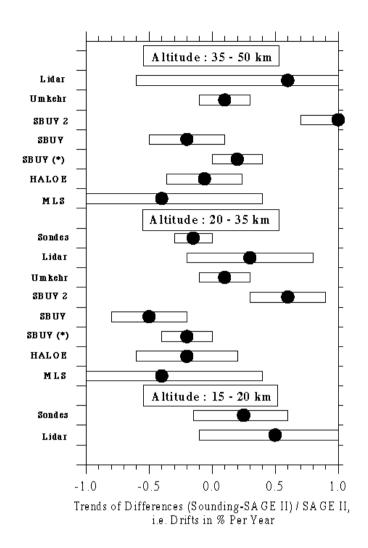
Figure 1 Trends of differences (i.e. drifts) between ozone measurements made by various ozone profiling instruments and SAGE II in % per year [(Sounding - SAGE II)/SAGE II]. Trends with ozonesondes are for the eight northern mid-latitude sounding stations used in Chapter 3 covering 36°N to 56°N. Umkehr 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 60-70°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 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.
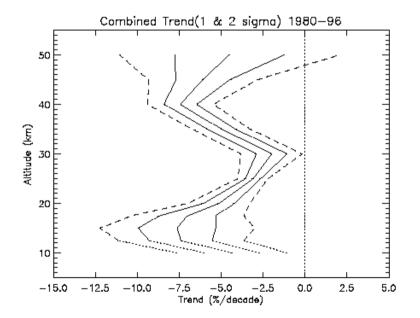

Figure 2. Estimate of the mean trend in the vertical distribution of ozone that has occurred over northern mid-latitudes from 1980-1996 (heavy solid line) calculated using the trends derived from SAGE I/II, ozonesondes, SBUV and Umkehr measurements. Combined uncertainties are shown as 1s (light solid 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.