

Comparison of ER-2 aircraft and POAM III, MLS, and SAGE II satellite measurements during SOLVE using traditional correlative analysis and trajectory hunting technique

M. Y. Danilin,^{1,2} M. K. W. Ko,¹ R. M. Bevilacqua,³ L. V. Lyjak,⁴ L. Froidevaux,⁵ M. L. Santee,⁵ J. M. Zawodny,⁶ K. W. Hoppel,³ E. C. Richard,⁷ J. R. Spackman,⁸ E. M. Weinstock,⁸ R. L. Herman,⁵ K. A. McKinney,⁹ P. O. Wennberg,⁹ F. L. Eisele,^{4,10} and R. M. Stimpfle,⁸ C. J. Scott,¹ J. W. Elkins,⁷ and T. V. Bui¹¹

Received 27 April 2001; revised 7 August 2001; accepted 19 September 2001; published 11 December 2002.

[1] We compared the version 5 Microwave Limb Sounder (MLS), version 3 Polar Ozone and Aerosol Measurement III (POAM III), version 6.0 Stratospheric Aerosol and Gas Experiment II (SAGE II), and NASA ER-2 aircraft measurements made in the Northern Hemisphere in January–February 2000 during the SAGE III Ozone Loss and Validation Experiment (SOLVE). This study addresses one of the key scientific objectives of the SOLVE campaign, namely, to validate multiplatform satellite measurements made in the polar stratosphere during winter. This intercomparison was performed by using a traditional correlative analysis (TCA) and a trajectory hunting technique (THT). TCA compares profiles colocated within a chosen spatial-temporal vicinity. Launching backward and forward trajectories from the points of measurement, the THT identifies air parcels sampled at least twice within a prescribed match criterion during the course of 5 days. We found that the ozone measurements made by these four instruments agree most of the time within $\pm 10\%$ in the stratosphere up to 1400 K (~ 35 km). The water vapor measurements from POAM III and the ER-2 Harvard Lyman α hygrometer and Jet Propulsion Laboratory laser hygrometer agree to within ± 0.5 ppmv (or about $\pm 10\%$) in the lower stratosphere above 380 K. The MLS and ER-2 ClO measurements agree within their error bars for the TCA. The MLS and ER-2 nitric acid measurements near 17- to 20-km altitude agree within their uncertainties most of the time with a hint of a positive offset by MLS according to the TCA. We also applied the Atmospheric and Environmental Research, Inc. box model constrained by the ER-2 measurements for analysis of the ClO and HNO₃ measurements using the THT. We found that: (1) the model values of ClO are smaller by about 0.3–0.4 (0.2) ppbv below (above) 400 K than those by MLS and (2) the HNO₃ comparison shows a positive offset of MLS values by ~ 1 and 1–2 ppbv below 400 K and near 450 K, respectively. Our study shows that, with some limitations (like HNO₃ comparison under polar stratospheric cloud conditions), the THT is a more powerful tool for validation studies than the TCA, making conclusions of the comparison statistically more robust.

INDEX TERMS: 0340 Atmospheric Composition and Structure: Middle atmosphere—composition and chemistry; 0341 Atmospheric Composition and Structure: Middle atmosphere—constituent transport and chemistry (3334); 0394 Atmospheric Composition and Structure: Instruments and techniques; 3360 Meteorology and Atmospheric Dynamics: Remote sensing;

KEYWORDS: measurement comparison, trajectory hunting, correlative analysis, SOLVE

Citation: Danilin, M. Y., et al., Comparison of ER-2 aircraft and POAM III, MLS, and SAGE II satellite measurements during SOLVE using traditional correlative analysis and trajectory hunting technique, *J. Geophys. Res.*, 107, 8315, doi:10.1029/2001JD000781, 2002. [printed 108(D5), 2003]

¹Atmospheric and Environmental Research, Inc., Lexington, Massachusetts, USA.

²Now at The Boeing Company, Seattle, Washington, USA.

³Naval Research Laboratory, Washington, D. C., USA.

⁴National Center for Atmospheric Research, Boulder, Colorado, USA.

⁵Jet Propulsion Laboratory, Pasadena, California, USA.

⁶NASA Langley Research Center, Hampton, Virginia, USA.

⁷NOAA Aeronomy Laboratory, Boulder, Colorado, USA.

⁸Harvard University, Cambridge, Massachusetts, USA.

⁹California Institute of Technology, Pasadena, California, USA.

¹⁰Georgia Institute of Technology, Atlanta, Georgia, USA.

¹¹NASA Ames Research Center, Moffett Field, California, USA.

1. Introduction

[2] The Stratospheric Aerosol and Gas Experiment III (SAGE III) Ozone Loss and Validation Experiment (SOLVE) was a measurement campaign designed to study the processes controlling ozone behavior at high and middle latitudes (see <http://cloud1.arc.nasa.gov/solve/> for details). SOLVE measurements were made mostly in the Arctic region during November 1999 to March 2000 using the NASA DC-8 and ER-2 aircraft together with balloons and ground-based instruments. This mission was also intended to acquire correlative data necessary for validation of the SAGE III measurements, which will be used to assess the behavior of global ozone and its possible trends. However, the launch of the SAGE III instrument was postponed until 2001. Here comparison of the ER-2 measurements during the SOLVE campaign against available satellite instruments (Polar Ozone and Aerosol Measurement III (POAM III), Microwave Limb Sounder (MLS), and SAGE II) is performed. This information can be used to benefit SAGE III validation later.

[3] Validation of any new instrument is a must before its products can be used for scientific studies. Traditional correlative analysis (TCA) compares similar products obtained by new and other well-established platforms which are collocated in time and space as closely as possible. Such an approach is particularly attractive for comparison of the remote sensing and in situ data, since typically in situ measurements are more precise and accurate. Since the sampling volumes of the satellite and in situ measurements are very different (10^3 – 10^4 km³ versus $< 10^{-2}$ km³), some geophysical variability of the matched in situ data may complicate their comparison. Typically, this variability is mitigated by comparing as many matches as possible or assigning only one mean value to in situ measurements for each satellite volume. It is also crucial to make sure that the species compared have a lifetime longer than the temporal mismatch between measurements. To assess the quality of the satellite measurements, they are typically validated with in situ measurements from ozonesondes, balloon, or aircraft (see special issues of the *Journal of Geophysical Research*, (volume 94, issue D6, pp. 8335–8446, 1989; volume 101, issue D6, pp. 9539–10,476, 1996; and volume 102, issue D19, pp. 23,591–23,672, 1997) devoted to validation of SAGE II, UARS, and POAM II data, respectively). However, a limited altitude range and relatively small amount of matches (particularly for the occultation instruments with only 28–30 profiles per day [e.g., Lu *et al.*, 1997a, 1997b]) with in situ measurements could hamper a statistically significant validation of new satellite platforms.

[4] Recently, several new techniques using trajectories have been applied to improve validation of satellite data. Pierce *et al.* [1994] created “synoptic” maps of UARS Halogen Occultation Experiment (HALOE) measurements in order to statistically improve the comparison between HALOE sunrise and sunset measurements. Morris *et al.* [1995] applied trajectory mapping to validate spatially distant UARS MLS and HALOE data. The reverse-domain-filling technique [e.g., Sutton *et al.*, 1994] was used to create uniformly gridded satellite data by initializing trajectories at a regular grid and then assigning them values of the satellite measurements using backward trajectories

and their encounters with satellite observations. Bacmeister *et al.* [1999] used the Lagrangian approach to map ER-2 flights on 2 and 4 November 1994 and CRISTA data after 4 November 1994 to a UT noon on 5 November 1994 for comparison. Lingenfelter *et al.* [1999] compared ER-2 and HALOE ozone measurements in the lower stratosphere using trajectories. Morris *et al.* [2000] showed that trajectory mapping is a very effective tool for comparing one sparse data set with one dense satellite data set (HALOE and MLS O₃), comparing two sparse data sets (HALOE and SAGE II O₃), and estimating instrument precision using MLS H₂O measurements. von der Gathen *et al.* [1995] and Rex *et al.* [1998] applied the match technique, which is similar to the trajectory hunting technique (THT) and finds air parcels sampled twice by ozonesondes in order to calculate ozone loss rates for the matched parcels.

[5] For validation of short-lived species, a box model is required in order to account for photochemical changes along matched trajectories. Pierce *et al.* [1997] applied photochemical calculations along trajectories to compare ER-2 and HALOE measurements and model calculations of radical species against ER-2 measurements. Danilin *et al.* [2000] used the Atmospheric and Environmental Research, Inc. (AER) box model along matched air trajectories in December 1992 and obtained reasonable agreement between the calculated and measured behaviors of ClO, ClONO₂, HNO₃, and aerosol extinction at 780 cm⁻¹ during this episode.

[6] The goal of this study is not to provide a thorough validation study for the MLS, POAM III, and SAGE II instruments during their whole operation periods, but rather to perform a multiplatform data comparison between each other and with the ER-2 measurements for a particular period in January–February 2000. In order to achieve this goal, we will use both the traditional correlative analysis and the trajectory hunting technique, which are described in detail in section 3. The structure of our paper is the following: Section 2 considers the main features of the instruments and the episodes considered. Section 3 describes the methods used. Sections 4, 5, 6, and 7 compare ozone, water vapor, chlorine monoxide, and nitric acid measurements, respectively. Finally, section 8 summarizes the main findings of our study.

2. Period Considered and Characteristics of the Instruments Used

2.1. Period Considered

[7] Figure 1 shows latitude coverage by the MLS (blue), POAM III (red), SAGE II (black), HALOE (yellow), and ER-2 (green) instruments from 10 January to 20 March 2000. The ER-2 aircraft arrived in Kiruna (Sweden, 68°N, 21°E) on 14 January 2000. It made six flights during its first deployment (on 20, 23, 27, and 31 January and 2 and 3 February) and five flights during its second deployment (on 26 February and 5, 7, 11, and 12 March). It returned to the United States on 16 March. To support SOLVE, POAM III changed its routine schedule when it looks every other day either in the Northern Hemisphere or the Southern Hemisphere. The POAM III is an occultation instrument, which provides sunrise measurements in the Northern Hemisphere

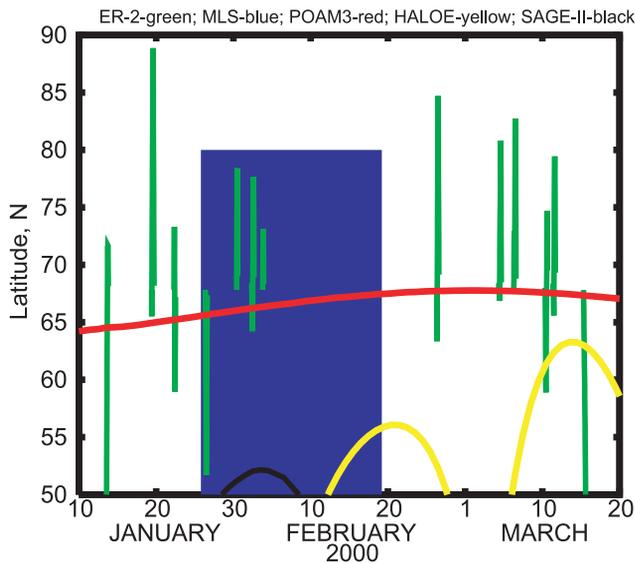


Figure 1. Latitudinal coverage of the MLS (blue), POAM III (red), ER-2 (green), SAGE II (black), and HALOE (yellow) measurements during the 10 January to 20 March 2000 period.

almost every day (except 14 February 2000), with 14 profiles daily most of the time. Additionally, the MLS obtained data from 2 to 13 February 2000 and from 27 to 29 March 2000 (the last period is not shown in Figure 1). To maximize the probability of obtaining useful measurements in the sunlit polar vortex, MLS scans were performed only poleward of 50°N on the “day side” of the orbit, providing several tens of profiles of O_3 , HNO_3 , and ClO daily [Santee *et al.*, 2000b].

[8] Two other occultation instruments (SAGE II and HALOE) operated on their routine basis. SAGE II provided sunrise measurements poleward of 50°N from 29 January 29 to 8 February 2000. HALOE looked poleward of 50°N from 12 to 27 February (sunrise) and from 7 to 24 March (sunset). Since the HALOE latitude overlapped barely with that of MLS on 12–13 February and briefly crossed the path of the ER-2 flights on 11 and 16 March, very few matches are anticipated for the MLS/HALOE and ER-2/HALOE pairs. The poor statistics of the comparison for these pairs precludes us from using the HALOE data in our analysis. We focus our study on the available measurements during the 23 January to 18 February 2000 period, which is characterized by the relatively dense coverage by the satellite and aircraft measurements.

2.2. Instruments Used

[9] Table 1 summarizes the instruments and their principal parameters used in this study.

2.2.1. ER-2

[10] The NASA ER-2 aircraft has 17 instruments aboard, of which the following five are used in this study for comparison with satellite measurements: dual-beam UV absorption ozone photometer [Proffitt *et al.*, 1989], the Harvard Lyman α hygrometer [Weinstock *et al.*, 1994], the Jet Propulsion Laboratory (JPL) laser spectrometer [May, 1998], the Harvard $\text{NO}_2\text{-ClO-ClONO}_2\text{-BrO}$ instrument [Stimpfle *et al.*, 1999], and the California Institute of Technology chemical ionization mass spectrometry (CIMS) instrument (K. A. McKinney *et al.*, manuscript in preparation, 2001). Here we briefly describe the main parameters and principles of measurements of these instruments.

[11] Ozone is measured by the ozone photometer, which uses light from a 254-nm lamp. This light passes through two identical sample chambers, one with ambient air and the other with the same air with completely removed ozone. The absorption cross section of ozone is well known and large at 254 nm. Thus the difference between the detected signals from two chambers allows accurate determination of the ozone concentration in the ambient air. At a measurement frequency of 1 Hz, the minimum detection limit of ozone is about 1.5×10^{10} molecules/cm³ (or 8 ppbv at 20 km) with a total uncertainty of several percent.

[12] The Harvard Lyman α photofragment fluorescence hygrometer measures water vapor aboard the ER-2 aircraft. Ambient flow is ram fed through the nose of the aircraft. To optimize accuracy, the core of the flow is picked off and throttled rapidly through the water vapor detection axis (for a full description of the instrument, see Weinstock *et al.* [1994]). Reported accuracy of the measurements is $\pm 5\%$ with a potential offset of ± 0.1 ppmv [Hintsa *et al.*, 1999]. The 1σ precision of the measurements is typically ± 0.1 ppmv for a 10-s integration time.

[13] Water vapor is also measured by the JPL laser hygrometer [May, 1998]. This is a single-channel, near-infrared tunable diode laser spectrometer operating near 1.37- μm wavelength with a multipass optical cell in the Herriott configuration. To ensure no contamination, the open-path optical cell is mounted external to the boundary layer of the right-wing superpod of the ER-2. This instrument is calibrated in the laboratory with a known water vapor standard over the range of pressures experienced during flight and compared with a chilled mirror frost point hygrometer. For SOLVE data, the 1σ precision is typically better than ± 0.05 ppmv over a 1.3-s integration time. The

Table 1. Instruments Analyzed in This Study and Their Characteristics (Measured Parameters With Their Uncertainty $\pm 1\sigma$, Vertical Resolution Δz , and Solar Zenith Angle (SZA) of Measurements)

Instrument	Parameters Measured	Δz , km	SZA, deg
MLS	$\text{O}_3(\pm 0.3 \text{ ppmv})$, $\text{HNO}_3(\pm 1 - 2 \text{ ppbv})$, $\text{ClO}(\pm 0.3 - 0.6 \text{ ppbv})$	4, 6 ^d	any
POAM III	$\text{O}_3(\pm 0.5 \text{ ppmv})$, $\text{H}_2\text{O}(\pm 1 \text{ ppmv})$, $\text{NO}_2(\pm 0.5 \text{ ppbv})$	1	90
SAGE II	$\text{O}_3(\pm 1\%)$, $\text{NO}_2(\pm 10\%)$, $\text{H}_2\text{O}(\pm 0.5 \text{ ppmv})$	0.5	90
ER-2 ^a	$\text{O}_3(\pm 0.05 \text{ ppmv})$, $\text{H}_2\text{O}(\pm 5\%b, \pm 5\%c)$, $\text{HNO}_3(\pm 0.75 \text{ ppbv})$, $\text{ClO}(\pm 17\%)$	<0.1	any

^aPartial list of the parameters measured by ER-2 is shown; POAM III and SAGE II also measure aerosol extinctions.

^bHarvard Lyman α hygrometer.

^cJPL laser hygrometer.

^dValues of 4 km for O_3 and ClO and 6 km for HNO_3 .

reported accuracy is $\pm 5\%$ at pressures < 100 hPa and $\pm 8\%$ at pressures between 100 and 200 hPa [May, 1998].

[14] Chlorine monoxide is measured by the Harvard NO_2 -ClO-CIONO₂-BrO instrument, which comprises two separate instruments: a thermal dissociation, resonance fluorescence instrument for detection of halogen radical and reservoir species [Stimpfle *et al.*, 1999] and a laser-induced fluorescence instrument for the detection of NO_2 [Perkins *et al.*, 2001]. Within the halogen system, ClO is measured by reaction with injected NO to form Cl atoms, followed by resonance fluorescence detection of Cl atoms at 118.9 nm. The instrument sensitivity is calibrated in the laboratory with known Cl atom densities and normalized to the Rayleigh scattering signal. The ClO measurements are acquired with a 35-s temporal resolution with an uncertainty (1σ) and detection limit of $\pm 17\%$ and 3 parts per trillion by volume (pptv), respectively.

[15] HNO_3 is measured by CIMS. The instrument inlet selectively samples either gas or particles in the ambient air by using a modified virtual impactor technique. Typically, aerosol and gas phases are sampled alternately during flight for periods of 3 min each. The sample then passes through a flow tube at ~ 290 K to evaporate condensed HNO_3 from aerosols. HNO_3 is ionized by chemical reaction with the precursor ion, CF_3O^- : $\text{CF}_3\text{O}^- + \text{HNO}_3 \rightarrow \text{HF-NO}_3^- + \text{CF}_2\text{O}$. Ions are then directed through an aperture into the vacuum system where, following mass selection, they are detected by a channel electron multiplier. A constant, known amount of isotopically labeled HNO_3 is added to the sample, allowing continuous, simultaneous calibration and measurement. For the data used here (gas phase measurements from the January–February ER-2 deployment) the precision is about ± 0.75 ppbv (1σ) for a 7-s integration period. The accuracy is the greater of $\pm 25\%$ and ± 1 ppbv.

2.2.2. MLS

[16] The microwave limb sounding technique and the MLS instrument are described in detail by Waters [1993] and Barath *et al.* [1993], respectively. The MLS instrument began acquiring millimeter-wavelength emission measurements in late September 1991. MLS measurements were made daily or near daily until 1995, after which data coverage became progressively sparser; in July 1999 the MLS instrument was placed in standby mode and was not powered up again until February–March 2000, when the 205-GHz radiometer was operated to provide measurements of ClO, O_3 , and HNO_3 in the Northern Hemisphere in conjunction with the SOLVE campaign. Santee *et al.* [2000b] present these measurements and describe the ways in which the operational strategy employed when these data were collected differed from the previous mode. The year 2000 data may have small shifts (typically of the order of a few percent) in comparison with normal version 5 (v.5) operational retrievals (with the 63-GHz radiometer providing tangent pressure and temperature information), but the 2000 data are still considered useful for the type of analyses performed here. In this study we analyze the ClO, O_3 , and HNO_3 measurements obtained during the February 2000 observing period.

[17] A special issue of the *Journal Geophysical Research* (volume 101, issue D6, pp. 9539–10,476, 1996) was devoted to the UARS validation and discussed the MLS version 3 (v.3) results in detail. Subsequently, version 4 data

were released; additional information on version 4 ozone data was provided in by the *World Meteorological Organization (WMO)* [1998]. Version 5 MLS data have recently become available. In v.5, quantities are retrieved on every UARS surface (six surfaces per decade in pressure, as opposed to three in previous MLS data sets), although the true vertical resolution of the data has not doubled. Overall, the data quality and the vertical range of reliability has been improved in v.5, especially for ozone in the lower stratosphere. Estimated vertical resolutions and total uncertainties of the individual measurements of ozone, nitric acid, and chlorine monoxide are about 4 km, 6 km, and 4 km and about 0.4 ppmv, 2 ppbv, and 0.4 ppbv, respectively, at the altitudes of interest in this study. A detailed description of the MLS v.5 data processing algorithm and validation of the various v.5 data products will be given elsewhere [Livesey *et al.*, 2002]. Information about the quality of the v.5 data is also available from the MLS web site (<http://mls.jpl.nasa.gov>).

2.2.3. POAM III

[18] The POAM III instrument is a nine-channel visible/near-infrared photometer which measures O_3 , H_2O , NO_2 , and aerosol extinction using the solar occultation technique. It was launched on 23 March 1998 into a Sun-synchronous 98.7° inclination orbit with a period of 101.5 min and is still in operation. Occultation measurements provide 14 sunrise and 14 sunset profiles daily with a vertical resolution of about 1–2 km (depending on species and altitude) and a longitudinal distance of 25.7° between successive measurements. Because of the orbit inclination, the measurements cover the 55°N – 71°N and 62° – 88°S latitude bands providing sunrise and sunset profiles in the Northern and Southern Hemispheres, respectively. The center wavelengths of the POAM III channels are located at 354, 439.6, 442.2, 603, 761.3, 779, 922.4, 935.9, and 1018 nm. A detailed description of POAM III and its early validation results are presented by Lucke *et al.* [1999] and at the POAM web site (<http://wvms.nrl.navy.mil/POAM>). Also, the special POAM II validation section of the *Journal of Geophysical Research* (volume 102, issue D19, pp. 23,591–23,672, 1997) could be useful for the POAM III data analysis, since these two instruments are similar.

2.2.4. SAGE II

[19] The SAGE II instrument was launched into a 57° inclination orbit aboard the Earth Radiation Budget Satellite on 5 October 1984 [Russell and McCormick, 1989] and is still operational. SAGE II employs the solar occultation technique and measures O_3 , NO_2 , H_2O , and aerosol extinctions with a vertical resolution of better than 1 km. This instrument measures attenuated solar radiation at seven wavelengths centered near 0.385, 0.448, 0.453, 0.525, 0.60, 0.94, and 1.02 μm . Details of the SAGE II retrieval algorithm and validation of the SAGE II measurements are given elsewhere [e.g., Chu *et al.*, 1989] (see also *Journal of Geophysical Research*, volume 94, issue D6, pp. 8335–8446, 1989; <http://www-sage2.larc.nasa.gov>). We use the latest version 6.0 of the SAGE II data in this study.

[20] While papers detailing the v.6.0 algorithm and validation of the data products are in preparation, an overview of the changes implemented in version 6.0 can be found at <http://www-sage2.larc.nasa.gov/cd-rom>. The primary motivation for the version 6.0 development was to understand

and correct the longstanding bias in the ozone profiles between SAGE II and ozonesondes in the 15- to 20-km region [WMO, 1998]. Version 6.0 development was also focused on improving the overall quality of the SAGE II ozone and aerosol products. The primary improvements were to the transmission algorithm, in the way of improved altitude registration, and to the methodology of calculating and propagating the known sources of error. Compared with the previous version, v.5.96, this new version of ozone has finer vertical resolution, considerably smaller error estimates, reduced aerosol artifacts, and a greater vertical range extending into the midtroposphere.

3. Analysis Techniques

3.1. Traditional Correlative Analysis

[21] TCA is a conventional method used for validation of atmospheric measurements during the last several decades. TCA finds nearly coincident profiles measured by the different platforms that satisfy a prescribed match criterion. TCA also assumes that the variability of the compared parameters, due to photochemistry or mixing, is small within a spatial-temporal distance of the match criterion used. The averaged profiles measured by different instruments are compared and their differences (in absolute units or in percent) are analyzed in order to find averaged offsets between different instruments. This approach is used, for example, to compare UARS and SAGE II measurements with radiosonde or ground-based measurements (*Journal of Geophysical Research*, volume 94, issue D6, pp. 8335–8446, 1989; volume 101, issue D6, pp. 9539–10,476, 1996; WMO [1998]). The main drawback of the TCA is the small number of matches. In order to improve the statistics of the comparison, more relaxed match criteria can be used. However, care is needed in order to avoid compromising the comparison because of using an overly relaxed criterion. For example, ozone becomes a relatively short-lived gas above 35 km, and its diurnal variability can spoil a comparison of the nadir versus occultation measurements if a temporal match criterion of several hours is used.

[22] Additionally, comparison of zonally averaged data can be made. In this case, meteorological variability and small-scale effects are mostly removed and the statistical significance of such a comparison is improved. However, the very large temporal and spatial difference between the measurements used in this approach is a shortcoming. For comparison of short-lived species (like NO₂ or N₂O₅), the solar zenith angle of measurements is important in order to avoid diurnal modulation of the species of interest. Despite all these caveats and limitations, the TCA remains the most trusted and widely used method of multiplatform data validation, and special validation campaigns are usually planned for a new platform (like EOS Aura).

3.2. Trajectory Hunting Technique

[23] THT identifies air parcels sampled at least twice by the same or different platforms and compares measurements along the matched trajectories [Danilin et al., 2000]. There are four stages in applying the THT for validation studies. At the first stage, backward and forward trajectories are calculated from the locations of measurements of interest. In this study, only 5-day backward and forward trajectories are

used for all pairs. We did not use longer trajectories because of the potentially larger uncertainties in their calculations [e.g., Schoeberl and Sparling, 1994; Morris et al., 1995]. Also, in our previous study [Danilin et al., 2000] it was shown that during a 5-day period the uncertainties of the trajectory calculations are reasonably small. This statement is supported in sections 4.1–4.2, where the 5-day and 1-day trajectories showed quite similar results.

[24] We calculated 670 backward and forward trajectories that originated from the ER-2 points at $p < 110$ hPa (i.e., above ~ 15 km) for the ER-2 flights on 23, 27, and 31 January and 2 and 3 February 2000. Since the frequency of measurements varies for the different instruments aboard ER-2, for the sake of convenience we use the ER-2 merged files provided by R. J. Salawitch with 10-s averaging. However, the trajectories were calculated from the ER-2 points distanced with an interval of 2 min along the flight track. Comparing SAGE II and MLS ozone data, we calculated 4044 backward and forward trajectories between 15 and 35 km with a 1-km step originated from the SAGE II profiles made during the 1–15 February 2000 period. For comparison of POAM III and MLS ozone measurements, 5130 backward and forward trajectories are calculated between 9- and 35-km altitude with a 1-km vertical step originated from the POAM III profiles measured during the 1–18 February period. We did not calculate trajectories above 35 km (~ 1400 K), since ozone has a short photochemical lifetime (less than 1 day) above this level and could not be considered as a passive tracer. Since MLS does not provide reliable measurements for $p > 100$ hPa, we also did not consider trajectories below this level. In the remaining altitude range (i.e., from 16 km to 35 km), the photochemical lifetime of ozone is of the order of weeks and months (see the discussion in section 4 about slow photochemical ozone loss in February 2000). This lifetime is much longer than the duration of the trajectories used; thus ozone may be considered as a passive tracer.

[25] Molecular mixing with ambient air in the lower and middle stratosphere is quite slow (with a typical timescale of several tens of days [e.g., Prather and Jaffe, 1990]); thus it will not noticeably affect our THT results, which are separated on average by about 2 days. However, mixing with ambient air becomes rapid in the mesosphere and upper stratosphere [Shepherd et al., 2000], thus limiting the effectiveness of the THT there.

[26] The diabatic trajectories used in this study were computed by using three-dimensional (3-D) winds derived from the temperature and geopotential height fields from the U.K. Meteorological Office (UKMO) assimilation scheme [Swinbank and O'Neill, 1994] on a global latitude-longitude grid of 2.5° by 3.75°. We used UKMO temperature and geopotential height fields instead of UKMO horizontal wind in order to calculate vertical velocities in a self-consistent way, since their UKMO values were of concern in some previous studies [Massie et al., 2000]. The zonal mean wind is calculated by using the gradient zonal wind approximation. The eddy components of the zonal and meridional winds are obtained by using an approximation that is consistent with the zonal mean gradient wind [Hitchman et al., 1987]. The mean meridional wind and the vertical wind (mean and eddy components) were calculated by using the thermodynamic and continuity equations in the

Table 2. Times and Locations of the Matches Between the Instruments Shown According to the TCA Satisfying the Match Criterion of (2 hours, 2°, 2°)

Pair	Date	Time, UT, Location
MLS/POAM III	2 February	1137, 66.36°N, 50.55°E / 1203, 66.31°N, 48.74°E
MLS/POAM III	3 February	2252, 65.34°N, 123.99°W / 2333, 66.44°N, 122.74°W
MLS/SAGE II ^a	7 February	1950, 52.38°N, 85.19°W / 1323, 50.76°N, 87.91°W
MLS/SAGE II ^a	7 February	2144, 52.33°N, 109.52°W / 1459, 50.71°N, 112.04°W
MLS/SAGE II ^a	8 February	1015, 51.92°N, 55.88°E / 0350, 50.24°N, 54.90°E
MLS/SAGE II ^a	8 February	1328, 51.82°N, 7.22°E / 0702, 50.12°N, 6.63°E
MLS/SAGE II ^a	9 February	1017, 51.14°N, 51.00°E / 0353, 49.16°N, 52.81°E
MLS/ER-2 ^b	2 February	1137, 66.36°N, 50.55°E / 1100, 64.50°N, 48.74°E
MLS/ER-2 ^b	2 February	1138, 63.08°N, 54.59°E / 1114, 64.72°N, 54.65°E
POAM III/ER-2 ^b	23 January	1337, 65.34°N, 17.97°E / 1203, 65.38°N, 19.00°E
POAM III/ER-2 ^b	2 February	1203, 66.31°N, 48.74°E / 1055, 64.41°N, 49.31°E

^aThe match criterion of (8 hours, 2°, 3°) is used for the MLS/SAGE II pair.

^bMean values of ER-2 time and location during the match are shown.

same manner as that of *Smith and Lyjak* [1985]. The net diabatic heating rates used in the thermodynamic equation were calculated as described by *Gille and Lyjak* [1986] using UKMO temperatures and monthly mean H₂O and O₃ climatology. The 3-D wind at the locations and times required by the trajectory calculation was obtained by linearly interpolating the daily wind fields in space and time. Our trajectory code is a commonly used and widely accepted source of trajectory calculations.

[27] At the second stage of the THT, we check whether each trajectory launched from the locations of the first platform measurements passes within a prescribed temporal-spatial distance from the other platform measurements. If it does, the hunting for this trajectory is successful and it is a subject for further analysis; if not, we drop this trajectory from our study. It is better to launch trajectories from the relatively sparse measurements (ER-2 or POAM III) and to hunt for the more frequent measurements (like MLS) rather than vice versa, since this approach requires less computer time and memory while providing the same results.

[28] The third stage is devoted to the interpolation in the vertical coordinate of the matched measurements. This interpolation is required only for the targeted measurements (i.e., MLS), since the values at the initial points of the trajectories are known. For the results shown below, we used a linear interpolation in the log-pressure scale. However, our sensitivity analysis shows that a choice of the vertical coordinate (e.g., pressure versus log-pressure or θ versus pressure) does not noticeably affect the results. Once the matched points are known, the same interpolation code is applied for all species sampled by both platforms.

[29] At the final, fourth stage, grouping and statistical analysis of the matched measurements are performed. We define “grouping” as a procedure that bins all matched data as a function of any vertical coordinate (potential temperature, pressure, or altitude). For example, below we bin all satellite/satellite matched measurements with a step of 50 K and 100 K from 350 K to 1000 K and above 1000 K, respectively. We increased the grouping step above 1000 K in order to get comparable statistics with the data below 1000 K, since a vertical step in kilometers per 100 K in potential temperature is smaller in the middle than in the lower stratosphere. For the ER-2/MLS and ER-2/POAM III pairs we apply a constant vertical step of 20 K.

[30] This brief description of the THT shows that TCA matches are only a subset of the THT matches, when a

match is obtained near initial points of trajectories. Thus the THT is a statistically more powerful tool than the TCA for validating atmospheric measurements. THT is also a more cost-efficient way to carry out validation campaigns, since it allows for obtaining as much useful information as possible from independent measurements in the background atmosphere in addition to specially deployed platforms. In our paper we show comparisons for the same pair of instruments using both the TCA and THT in order to increase confidence in the results shown.

[31] For all but the MLS/SAGE II results shown below, we use the match criterion of ($\Delta\text{time} \leq 2$ hours, $\Delta\text{latitude} \leq 2^\circ$, $\Delta\text{longitude} \leq 2^\circ$). For the Kiruna latitude, the spatial difference of 2° in latitude and longitude is translated into 237 km. For the MLS/SAGE II ozone measurements, we apply the match criterion of ($\Delta\text{time} \leq 8$ hours, $\Delta\text{latitude} \leq 2^\circ$, $\Delta\text{longitude} \leq 3^\circ$), since for a shorter Δtime no matches are found for the period considered using the TCA. Below, for the sake of simplicity, we write the match criterion, for example, as (2 hours, 2°, 2°), omitting the words “ $\Delta\text{time} \leq$,” “ $\Delta\text{latitude} \leq$,” and “ $\Delta\text{longitude} \leq$.” In order to facilitate analysis of the results presented, we will show in some cases the difference between two instrument measurements in both absolute (ppmv or ppbv) and relative (percent) units.

4. Comparison of the Ozone Measurements

[32] The figures comparing MLS with satellite (POAM III and SAGE II) and ER-2 measurements are shown with a vertical resolution of ~ 1 –1.5 km or ~ 0.5 km, respectively, while the vertical resolution of the v.5 MLS data is about 4–6 km. This caveat should be kept in mind. We decided to show a finer vertical resolution linearly interpolating the MLS data in the log (pressure) coordinate to the data by POAM III, SAGE II, and ER-2. The results shown below look quite similar, and conclusions of our study are not affected when we use a vertical step of ~ 4 km after degrading the POAM III and SAGE II vertical resolution to the MLS vertical grid. For the ER-2/MLS pair the MLS vertical resolution provides only one level for comparison.

4.1. MLS Versus POAM III

[33] We found only two MLS/POAM III profiles satisfying the chosen match criterion of (2 hours, 2°, 2°), which are shown in Table 2. We averaged the matched MLS and

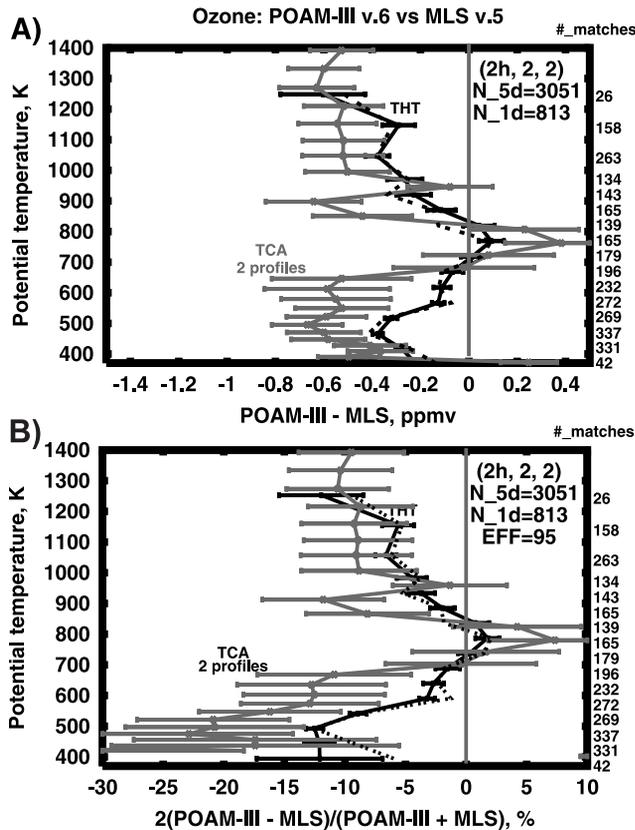


Figure 2. Vertical profiles of the difference between POAM III and MLS ozone measurements (a) in ppmv and (b) in percent (panel B) from the TCA (grey curves). Black curves show results for THT with 5-day (solid curves) and 1-day (dashed curves) trajectories. Numbers of matches for the THT with 5-day trajectories are shown near the right vertical axis. Total numbers of matches are 3051 and 813 for 5- and 1-day trajectories, respectively. Match criterion of (2 hours, 2° , 2°) is used. The error bars shown represent only the random (not systematic) error determined in Appendix A and correspond to $\pm 1\sigma$ (or 67% confidence level).

POAM III profiles separately and then show their difference by grey lines in Figure 2. Launching 5-day trajectories from the POAM III points, we found a total of 3051 matches satisfying the same match criterion. The black line in Figure 2

shows the difference between the POAM III and MLS ozone measurements, when all these matches are grouped with the 50 K and 100 K steps below and above 1000 K, respectively.

[34] Both methods show that POAM III ozone values are smaller than those of MLS almost everywhere below 1400 K. This difference ranges from -0.4 to $+0.4$ ppmv (or from -12% to 2%) for the THT and from -0.7 to $+0.4$ ppmv (or from -30% to 7%) for the TCA. Results for the TCA could be suspicious because of the small statistics (i.e., only two individual profiles). Indeed, the rapid changes near 800–900 K are caused by a zigzag structure in one POAM III profile, which propagated almost unmitigated in Figure 2. However, the THT results have about 100–300 individual matches per level shown and should not suffer from possible large discrepancies in the individual ozone measurements. Additionally, the THT results are very similar for hunting using 1-day backward and forward trajectories (shown by the black dashed lines in Figure 2), thus supporting the credibility of this technique.

[35] The effect of photochemical changes along the matched trajectories is unlikely to be an issue for this comparison, since the number of matches for forward (1605 matches) and backward (1446 matches) trajectories is about the same (see Table 3 for statistics of the THT). This factor indeed should be considered if we are in the region of intensive photochemical ozone destruction. In this case, POAM III ozone values will be reduced (increased) by the photochemistry for the forward (backward) trajectories originating from the points of POAM III measurements. However, detailed analysis of the photochemical ozone loss rate shows a modest ozone depletion in the beginning of February due to the lack of sunlight [Hoppel *et al.*, 2002]. Also, our box model calculations for the ER-2/MLS ozone comparison show only a small photochemical ozone loss in the lower stratosphere (see sections 4.3 and 6 for details). All these arguments convince us that the average difference between the POAM III and MLS ozone measurements is real.

4.2. MLS Versus SAGE II

[36] Table 2 shows that because of the geometry of the SAGE II and MLS measurements, there are no matches within a 6-hour interval. To get at least a minimum statistic for the MLS/SAGE II comparison using the TCA, we relaxed the match criterion from (2 hours, 2° , 2°) to (8 hours, 2° , 3°). For this criterion, we found five matched profiles.

Table 3. Statistical Data for the THT With the Match Criterion of (2 hours, 2° , 2°) Showing the Total Number of Relevant Trajectories N_{tra} and Matches N_t^m , Number of Matches for the Backward (N_b^m) and Forward (N_f^m) Trajectories, Number of Matches per Trajectory Launched (N_t^m/N_{tra}), and Mean Temporal Distance Between the Matches ($\overline{\Delta t}$)

Pair	Species	N_{tra}	N_t^m	N_b^m	N_f^m	N_t^m/N_{tra}	$\overline{\Delta t}$, days
MLS/POAM III	O ₃	5130	3051	1446	1605	0.59	2.4
MLS/SAGE II ^a	O ₃	4044	3397	907	2490	0.84	2.2
MLS/ER-2	O ₃	387	525	75	450	1.36	1.8
MLS/ER-2	ClO	387	525	75	450	1.36	1.8
MLS/ER-2	HNO ₃	387	127	21	105	0.33	1.9
ER-2/POAM III	O ₃	670	405	148	257	0.60	2.3
ER-2/POAM III	H ₂ O ^b	670	402	147	255	0.60	2.3
ER-2/POAM III	H ₂ O ^c	670	354	132	222	0.53	2.3

^aMatch criterion of (8 hours, 2° , 3°) was applied for this pair.

^bHarvard Lyman α hygrometer.

^cJPL laser hygrometer.

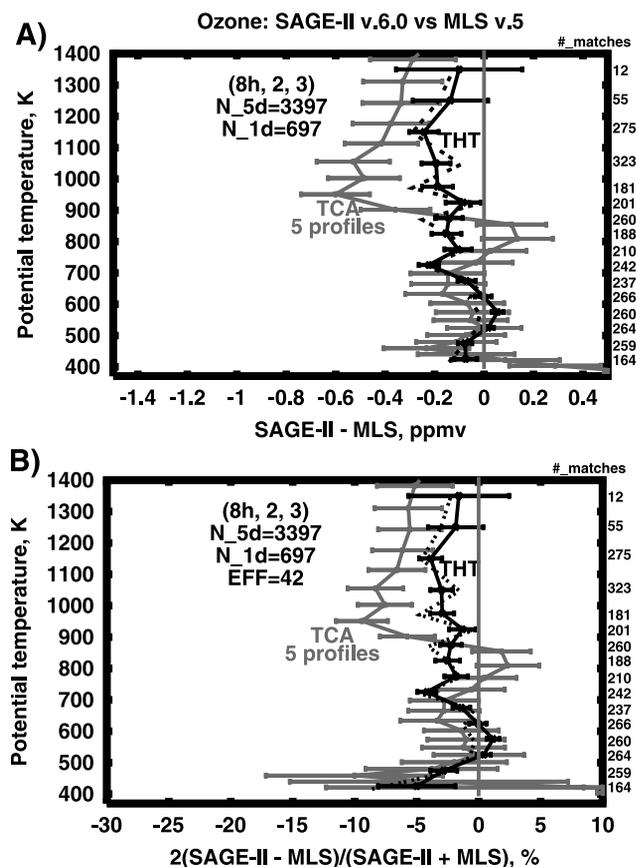


Figure 3. The same as Figure 2, but for the SAGE II and MLS ozone measurements and a match criterion of (8 hours, 2°, 3°).

Again, we averaged them separately for each instrument and depict the difference between their averaged values by grey lines in Figure 3. The difference obtained by the THT is shown by black lines, and the number of the matches found is depicted near the right vertical axis. We found 3397 matches for the 4044 trajectories launched.

[37] The obvious feature of Figure 3 is the good agreement between the v.5 MLS and the v.6.0 SAGE II measurements. The difference never exceeded the range of $\pm 10\%$ ($\pm 5\%$), being within $\pm 5\%$ ($\pm 3\%$) at most levels for the TCA (THT). The largest difference of 7–8% is obtained near 1000 K (~ 28 km). It is likely that this difference arises from the small number of statistics obtained for the TCA, when the effects of possible large differences for individual profiles is weakly mitigated for the averaged profiles. Indeed, we saw a large difference ($\text{MLS} - \text{SAGE II} > 1$ ppmv) at 26–32 km for the first of five matched ozone profiles listed in Table 2, which was propagated in Figure 3. Based on a PV field analysis, we conclude that this difference may be caused by sampling different air masses near the polar vortex edge. On the other hand, the credibility of the THT results is increased by their insensitivity to the duration of the trajectories used (since black solid and dashed lines in Figure 3 are very similar). The statistically significant difference in the range of $[-2\%:-4\%]$ and $[-5\%:-9\%]$ is obtained in the middle stratosphere (900 K to 1200 K) for the THT and TCA, respectively. This difference is smaller than 5% of the MLS positive offset for the v.3 MLS and v.5.93 SAGE II compar-

ison above 30 hPa reported by *Cunnold et al.* [1996] and *Froidevaux et al.* [1996], thus showing an improvement obtained in the latest versions of these data sets. The very good agreement between MLS and SAGE II throughout the stratosphere shows no evidence of the large biases seen in earlier versions of the SAGE II algorithms. Our comparisons of the ozone measurements for the MLS/SAGE II and MLS/POAM III pairs are consistent with the results of *Manney et al.* [2001], showing that the agreement between v.5 MLS and six other instruments (including v.6.0 SAGE II and v.6 POAM II) is better than 0.25 ppmv in the stratosphere. Summarizing, such a good agreement between MLS, SAGE II, and POAM III validates the use of the latest versions of MLS, POAM III, and SAGE II ozone measurements for scientific purposes.

4.3. MLS Versus ER-2

[38] There are two ER-2/MLS matches satisfying the match criterion of (2 hours, 2°, 2°) obtained during the 2 February 2000 flight using the TCA. During this flight, ER-2 flew at its cruise altitude between the two consecutive MLS profiles taken at 1137 and 1138 UT (see also Table 2). The ER-2 stayed 14 min within the chosen vicinity from each of the coincident MLS profiles. Figure 4 depicts the

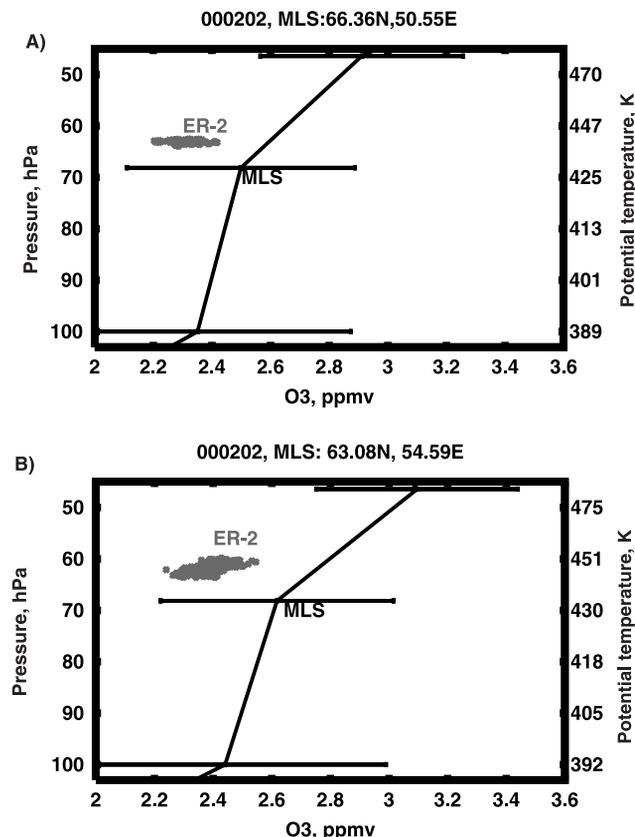


Figure 4. Two matches (a and b) between ER-2 and MLS ozone measurements for the ER-2 flight on 2 February 2000 according to TCA with the match criterion of (2 hours, 2°, 2°). The error bars show the uncertainties of the individual MLS profiles. Approximate potential temperature is shown near the right vertical axis. Locations of the matched MLS profiles are shown as well.

ER-2 and MLS ozone values obtained during these matches. The MLS ozone values are provided at 100, 68, and 46 hPa, while ER-2 sampled at 60–65 hPa. The ER-2 and MLS ozone measurements agree within their uncertainties for these two episodes. However, a possible positive offset in the MLS data of 0.2–0.3 ppmv is suspected. When the match criterion was relaxed to (3 hours, 2°, 5°), six matches were found for the 2 February 2000 ER-2 flight (not shown). For all these matches, the ER-2 and MLS ozone measurements agree within their uncertainties with even closer agreement than that shown in Figure 4. In such comparisons, one needs to keep in mind the huge difference in the individual volume sampled ($400 \times 400 \times 4 \text{ km}^3$ for MLS and much less than 1 km^3 for ER-2, which does not sample much of the altitude region sensed by MLS). This fact and the small number of matches make it difficult to assess any systematic difference (outside the error bars) between the two sets of measurements, especially for the TCA method.

[39] The THT allows better quantification of the difference between the MLS and ER-2 ozone measurements, since more matches are found covering a wider vertical range. We found 525 matches for the 387 trajectories launched from the ER-2 points during its flights on 31 January and 2 and 3 February 2000. Other ER-2 flights are irrelevant for comparison with MLS, since they occurred more than 5 days before the start or after the end of the MLS measurements in February. Figure 5 shows that ER-2 ozone measurements are smaller than those from MLS by 0.2–0.3 ppmv in the range 360–460 K. These results are statistically more robust than those of Figure 4. Part of the possible small discrepancy between MLS and ER-2 data may be linked to small offsets obtained for the single-radiometer mode retrievals used after mid-1997, in comparison with the multiradiometer retrievals used prior to this with the 63-GHz data available for tangent pressure and temperature retrievals. Indeed, tests performed by the MLS team indicate that the single-radiometer retrievals tend to overestimate the results from the full-blown retrievals by 0.1 to 0.2 ppmv at high northern latitudes. Our box model calculations initialized with the ER-2 measurements showed a modest averaged photochemical ozone depletion of 0.005 ppmv during the averaged temporal distance of 1.8 days between the matched ER-2 and MLS measurements (see section 6 for details). Thus the possible photochemical changes along the matched air parcels did not affect the results of the ER-2/MLS ozone comparison and can be ignored.

4.4. POAM III Versus ER-2

[40] Using the TCA, we found two ER-2/POAM III matches on 2 February and 23 January, when ER-2 was cruising and descending to Kiruna, respectively. These matches are shown in Figure 6, which depict 10-s-averaged ER-2 data and part of the corresponding coincident POAM III profile. The ER-2 was 10 min and 52 min within (2 hours, 2°, 2°) from the coincident POAM III profile. While the ER-2 and POAM III ozone measurements show good agreement within their uncertainties, one can suspect a possible small positive bias in the POAM III measurements. In order to check this assumption, we relaxed the match criterion to (3 hours, 2°, 5°) and found seven matches (not shown). However, they showed a similar behavior of the POAM III and ER-2 measurements shown in Figure 6.

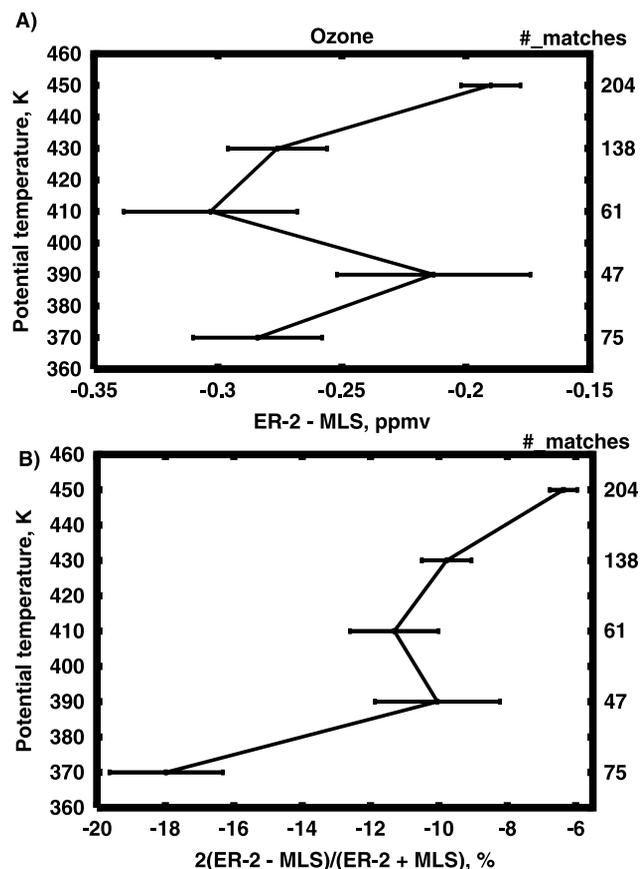


Figure 5. Vertical profiles of the difference between ER-2 and MLS ozone measurements (a) in ppmv and (b) in percent using the THT with the match criterion of (2 hours, 2°, 2°). The error bars show only the random errors determined in Appendix A. Numbers of matches are shown near the right vertical axis.

Applying the THT to the POAM III/ER-2 pair, we found 405 matches, which are shown in Figure 7 after grouping them with a vertical step of 20 K. The agreement between ER-2 and POAM III is good (better than 8%). Possible variability of the ER-2 ozone measurements (which is about ~ 0.1 – 0.2 ppmv at ER-2 cruise altitudes as shown in Figures 4 and 6) within a single POAM III sampling volume is strongly mitigated by a factor of $N^{1/2}$ owing to the large number of matches (N) in THT, thus making the POAM III/ER-2 comparison meaningful.

[41] Lumpe *et al.* [2002] also compared ozone measurements made by ER-2 and POAM III during SOLVE. In general, their results are very similar to ours, confirming good agreement between the ER-2 and POAM III ozone measurements. For example, they showed an agreement between ER-2 and POAM III ozone measurements within $\pm 10\%$ in the range 350–470 K using both the vortex-averaged and trajectory-matching techniques.

5. Comparison of the ER-2 and POAM III H₂O Measurements

[42] POAM III measures H₂O using the 935.9- and 922.4-nm channels. Previous comparisons of v.1.4 POAM III and

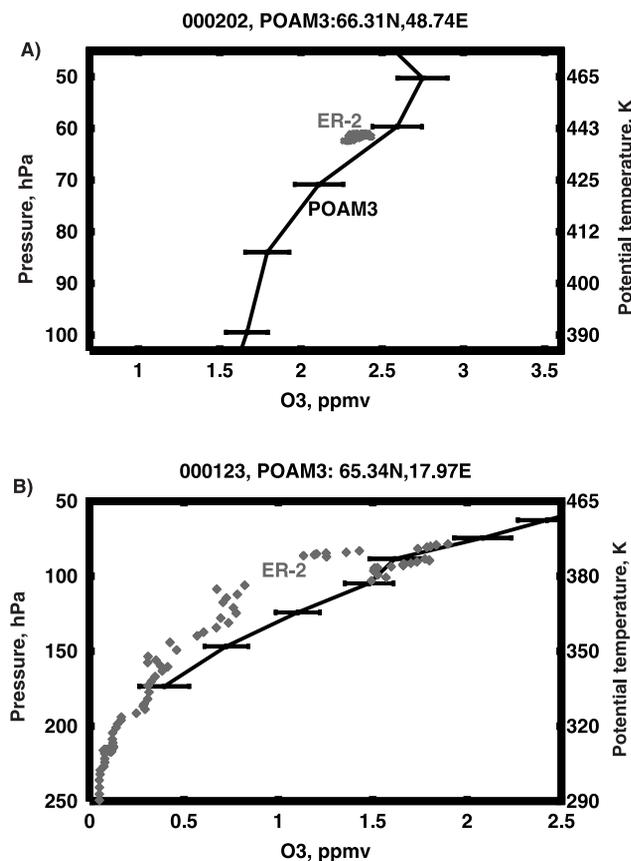


Figure 6. The same as Figure 4, but for the POAM III and ER-2 ozone measurements on 2 February and 23 January 2000.

v.19 HALOE measurements showed a high bias of about 15% in the stratosphere in the POAM III measurements [Lucke et al., 1999]. Recently, Bevilacqua et al. (unpublished manuscript, 2001) compared v.3 POAM III and v.19 HALOE H₂O measurements and found that the POAM III measurements are higher by <10% in the range 20–40 km in the Northern Hemisphere. The ER-2 water vapor measurements have been made during many aircraft campaigns and involved in the recent *Stratospheric Processes and Their Role in Climate (SPARC)* [2001] H₂O assessment.

[43] Figure 8 shows the ER-2/POAM III comparison for water vapor. Unlike the ozone case, there are two instruments aboard ER-2 which measure H₂O (see section 2 for their details). During the 2 February 2000 flight, both ER-2 instruments showed very compact water vapor values of about 5 ppmv at 62 hPa, almost overlapping each other. POAM III measured slightly higher water content, which could be considered consistent with the ER-2 measurements within uncertainties of the measurements. During the descent on 23 January 2000, both the Harvard and JPL instruments showed an increase of H₂O which is also hinted at in the POAM III measurements. However, individual water vapor measurements by POAM III can be noisy in the lower stratosphere/upper troposphere, especially if sunspots and clouds are present (see Bevilacqua et al. (unpublished manuscript, 2001) and Nedoluha et al. [2000] for details). For example, the POAM III H₂O profile shown in the

bottom panel of Figure 8 is strongly affected by a polar stratospheric (PSC) presence at this location according to the sharply increased aerosol extinction detected by the POAM III aerosol channels.

[44] Water vapor behaves as an inert tracer in the stratosphere, unless temperature drops below the ice frost point causing removal of H₂O from the gas phase. However, based on the analysis of temperature along the trajectories used, such cold temperatures did not occur during the period considered here. Thus neglect of possible microphysical changes is easily justified for the use of THT for H₂O. We found 402 and 354 matches for Harvard and JPL hygrometers, respectively. The number of matches is smaller for the JPL instrument because of some gaps in its data during three ER-2 flights (000131, 000202, and 000203). Figure 9 shows the difference between the H₂O measurements by the ER-2 instruments and POAM III, which changes from +8% (–3%) at 370 K to –8% (–12%) at 470 K for the JPL (Harvard) hygrometer. Most of the time at cruise altitude, the difference between the Harvard and JPL hygrometers lies within the reported uncertainties [Hints et al., 1999]. However, there is a pressure-dependent systematic bias between the two instruments that is greatest at 100- to 200-hPa pressure, which has been attributed to the JPL laser hygrometer [SPARC, 2001]. As is shown in Figure 9, this bias is about 0.4 ppmv at 390 K reducing to ~0.2 ppmv at 490 K (JPL values are always larger).

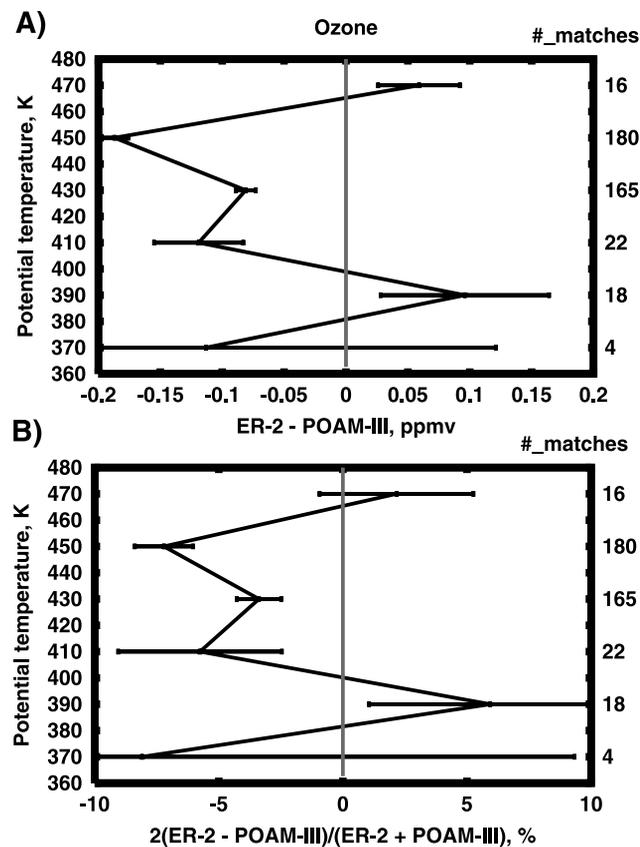


Figure 7. The same as Figure 5, but for the POAM III and ER-2 O₃ measurements.

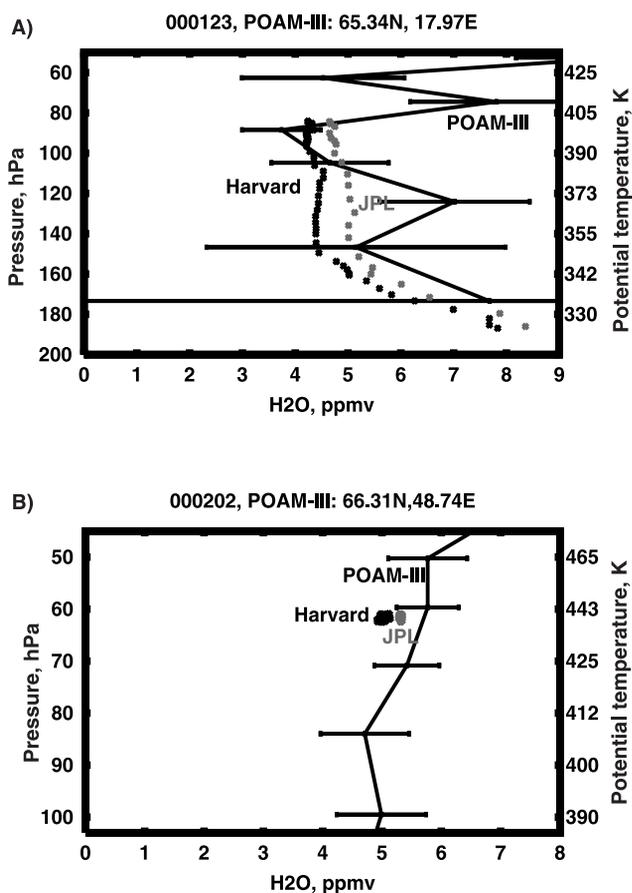


Figure 8. The same as Figure 6, but for the H₂O measurements. JPL and Harvard University instrument measurements are shown by grey and black symbols, respectively.

[45] The SOLVE POAM III/ER-2 H₂O comparison in our paper shows some inconsistency with the previous POAM III/HALOE and ER-2/HALOE comparisons [SPARC, 2001]. For example, our study and that by Bevilacqua et al. (manuscript in preparation, 2002) show that v.3 POAM III water vapor measurements are higher by 5–10% and 10–15% than the Harvard hygrometer and v.19 HALOE data, respectively, at 15- to 20-km altitude. This implies that the Harvard data should be higher by 5–10% than the v.19 HALOE measurements. However, the SPARC [2001] study shows that the Harvard hygrometer data are 20% higher than v.19 HALOE measurements. This inconsistency is discussed elsewhere (Bevilacqua et al., unpublished manuscript, 2001), and its further analysis is outside the scope of this paper.

6. Comparison of the ER-2 and MLS CIO Measurements

6.1. Analysis Using the TCA

[46] Previously, version 3 MLS CIO measurements were validated against balloon, aircraft, and ground-based measurements [Waters et al., 1996]. It was found that the v.3 MLS CIO measurements agree with available correlative measurements within their combined uncertainties. Subsequently, v.4 MLS CIO data were compared with the 204-

GHz measurements from the millimeter-wave atmospheric sounder (MAS) on the space shuttle [Feist et al., 2000]. The agreement was well within the combined error bars over a pressure range of 0.4–40 hPa. In v.5 MLS, use of a finer vertical grid in the retrievals allows better definition of the peak and results in generally smoother profiles (although the true vertical resolution of the v.5 CIO measurements, 4–5 km at the altitudes shown here, is coarser than the retrieval grid). A positive bias of 0.1 ppbv in the lower stratosphere is known to exist in the v.5 CIO data (based on averages of nighttime data over the first full year of the mission) [Livesey et al., 2002]. Here we compare the v.5 MLS data with SOLVE measurements.

[47] Figure 10 compares CIO measurements made by the Harvard NO₂-CIO-CIONO₂-BrO instrument aboard the ER-2 (grey dots) and MLS (black lines). For the two ER-2/MLS matches shown, the measurements by these instruments agree within their uncertainties. During these matches, chlorine was activated, with values between 1 and 2 ppbv in the range 70–45 hPa. Such a small number of matches can be understood from Figure 1, which shows that only two ER-2 flights (on 2 and 3 February) are available for direct comparison with MLS measurements. Unfortunately, the 3 February flight was very short and remote from the MLS points by at least 10 hours, making comparison of the measurements of short-lived CIO impossible for this flight using the TCA. To increase the number of coincident ER-2 and MLS CIO measurements, we relaxed the match criterion to (2 hours, 2°, 5°). In this case, we found six matches (not shown here). In all these matches, ER-2 and MLS agreed within their error bars. However, ER-2 values of CIO were smaller by at least 0.1 ppbv, consistent with the known high bias in the MLS CIO data at these altitudes.

[48] The agreement between remote sensing values of CIO in a box 400 × 400 × 4 km³ and very accurate in situ ER-2 measurements is good. It is clear that lower values of MLS CIO in v.5 compared with their values from the previous versions mitigate a concern about their positive bias raised previously by modelers [e.g., Chipperfield et al., 1996; Lutman et al., 1997] and mentioned by Waters et al.

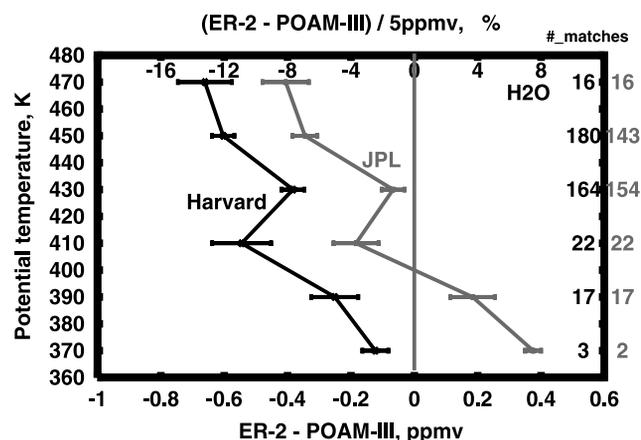


Figure 9. The same as Figure 7, but for the POAM III and ER-2 H₂O measurements. Black (grey) lines and symbols correspond to the Harvard (JPL) measurements. The horizontal axis in Figure 9a approximates the difference in percent.

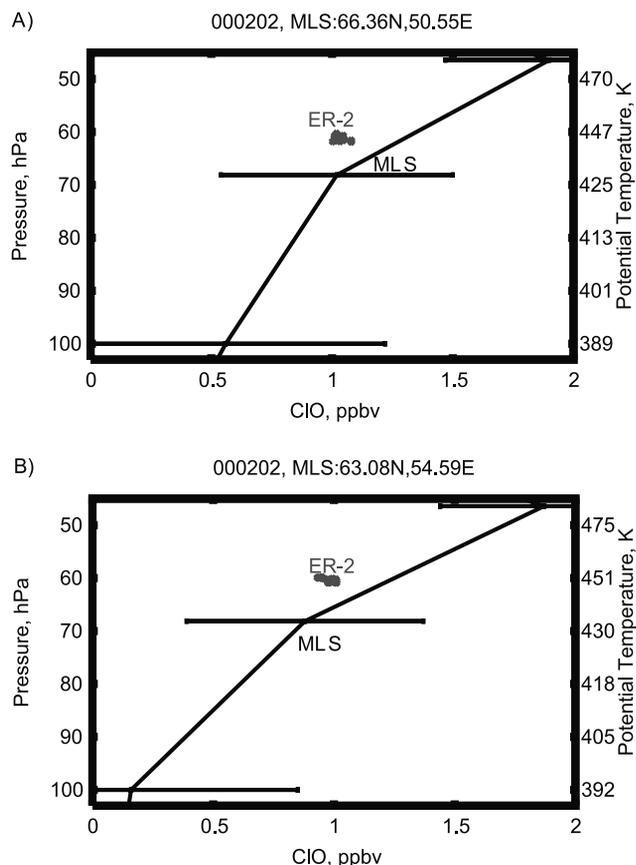


Figure 10. The same as Figure 4, but for the ER-2 and MLS CIO measurements.

[1996]. Since chlorine monoxide is a key ozone depleting species, accurate measurements of CIO and their agreement with models are crucial for our understanding of the changes in the global ozone layer.

6.2. Analysis Using the THT

[49] Because of the short photochemical lifetime of CIO in the stratosphere, the THT requires a photochemical box model for the matched parcels. We use the AER box model [Danilin *et al.*, 2000] constrained by the ER-2 measurements for analysis of the CIO measurements. The same model runs are used for analysis of the HNO₃ measurements presented in section 7.2. For the forward trajectories originated from the ER-2 points, our model is initialized by using the ER-2 O₃, NO, NO₂, HNO₃, NO_y, CIO, Cl₂O₂, HCl, ClONO₂, H₂O (JPL hygrometer values), CH₄, and N₂O values. Initial concentrations of other chlorine and nitrogen species, which are not measured by the ER-2 (namely, HOCl, Cl₂, OClO, HONO, N₂O₅, and HNO₄), are determined from the ER-2 measurements of NO_y and Cl_y using the partitioning among these species according to the AER 2-D model at 67°N in February [Weissenstein *et al.*, 1998]. However, the initial concentrations of these six species are small, contributing only several percent to the NO_y and Cl_y amounts, and are not important for our comparison of the CIO and HNO₃ measurements. In order to justify the last statement, we performed a sensitivity model run with zero initial concentrations of these six

species and obtained the same results for CIO and HNO₃ as shown below. Some details of nitrogen species partitioning between NO_y and HNO₃ in gas and condensed phases do not affect our model CIO results and are given in section 7.2. Initial total inorganic bromine Br_y is determined by using the N₂O-Br_y correlation based on the analysis of Wamsley *et al.* [1998] and is about 18 pptv at the ER-2 cruise altitude. The initial partitioning among bromine species is determined according to the AER 2-D model. The initial aerosol surface area density (SAD) is taken according to the Focused Cavity Aerosol Spectrometer (FCAS) [Jonsson *et al.*, 1995] measurements. We discuss later why the choice of the SAD measured by FCAS or Multiple-Angle Aerosol Spectrometer Probe (MASP) [Baumgardner *et al.*, 1996] has little impact on our results. Our model has an option to use either nitric acid trihydrate (NAT) [Hanson and Mauersberger, 1988] or supercooled ternary solution (STS) [Tabazadeh *et al.*, 1994] PSC schemes. Details of our model treatment of PSCs are given elsewhere [Danilin *et al.*, 2000]. Briefly, in the NAT scheme, the PSC SAD appears at $T < T_{\text{NAT}}$ in addition to the sulfate aerosol SAD (which was constant in this scheme) and is proportional to the amount of HNO₃ condensed. In the STS scheme, the aerosol SAD increases depending on the amount of HNO₃ and H₂O condensed at low temperature. Also, the reaction probabilities are different on the NAT and STS surfaces and are taken according to Sander *et al.* [2000]. One also should keep in mind that the amount of HNO₃ condensed in the NAT and STS scheme is quite different, especially in the vicinity of T_{NAT} . Below we applied both schemes to analyze CIO and HNO₃ measurements.

[50] It was suggested that the UKMO assimilated temperature has a warm bias in the polar lower stratosphere during winter [e.g., Manney *et al.*, 1996]. Indeed, we found that the UKMO temperature was higher on average by 1.2 K compared with the ER-2 measurements in the initial points of the 450 forward trajectories. To offset this warm bias, the forward trajectory temperatures were forced to match ER-2 temperatures in the initial points. One should keep in mind that despite this procedure, the assimilated temperature in other points of trajectories could still deviate from its “real” values. This procedure is standard in model analysis of ER-2 measurements during polar winters [e.g., Kawa *et al.*, 1997], allowing for considerable reduction of uncertainties of this crucial parameter (± 0.3 K for ER-2 versus several K for UKMO).

[51] We performed model calculations for the 450 forward trajectories that originated from the locations of ER-2 measurements. We did not make model runs for the backward trajectories from the ER-2 locations, since only MLS O₃, CIO, and HNO₃ values are available in the initial MLS points (i.e., where model calculations start), thus introducing additional uncertainties of the model initialization due to unknown initial partitioning of the chlorine and nitrogen species. Also, we notice some inconsistency between the MLS HNO₃ and ER-2 NO_y measurements, since the MLS HNO₃ values often (in 14 of 21 cases) were larger than the ER-2 NO_y for forward matches, thus making model initialization difficult. Owing to the timing of the ER-2 and MLS measurements (see Figure 1), the forward trajectories provide the dominant part of all ER-2/MLS matches (450 of 525, or $\sim 86\%$). Indeed, since MLS started its measurements

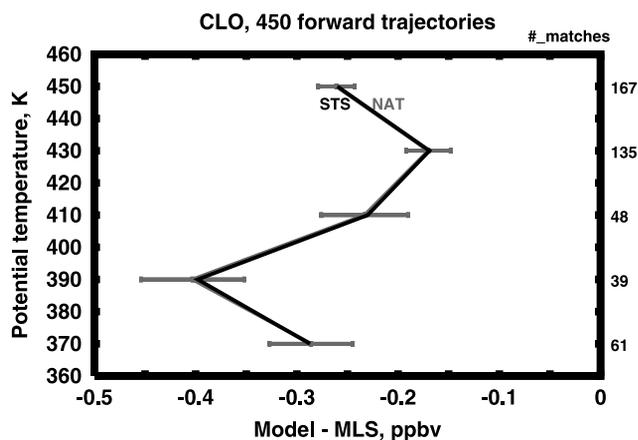


Figure 11. Difference between the model-calculated and MLS-measured CIO for the forward trajectories from the ER-2 locations using the NAT (grey line) and STS (black line) PSC schemes. The AER box model calculations are constrained by the ER-2 measurements as mentioned in the text.

on 2 February, only about 0.5 and 2 days of the MLS data are match-available for the backward trajectories that originated from the ER-2 locations during the 2 and 3 February flights, respectively. On the other hand, 5 and 3 days of the MLS measurements are match-available for the forward trajectories during these two flights and the 31 January flight, respectively.

[52] The grey line in Figure 11 shows the vertical profile of the difference between the MLS-measured and model-calculated CIO values using the NAT PSC scheme. This difference has maximum values of about -0.4 ppbv near 390 K and smaller values of about -0.2 ppbv above 400 K. For the STS scheme, the results (black line) are almost the same as for the NAT scheme. The main reason for this is the very low initial values of ClONO_2 (<0.05 ppbv) and HOCl (a few tens of pptv). Thus despite the difference in the reaction probabilities and SADs for the NAT and STS schemes, the $\text{ClONO}_2 + \text{HCl} \rightarrow \text{Cl}_2 + \text{HNO}_3$ and $\text{HOCl} + \text{HCl} \rightarrow \text{Cl}_2 + \text{H}_2\text{O}$ heterogeneous reactions are shut down, precluding any additional formation of active chlorine.

[53] It is important to notice that the error bars in Figure 11 show the standard error of the mean differences (as in all previous figures) and do not account for systematic error of the MLS CIO measurements. If this error is also taken into account, the error bars in Figure 11 should be increased threefold below 400 K and twofold above this level. Model initialization and calculations also introduce additional uncertainties. However, thorough analysis of this very important issue is outside the scope of this paper, requires numerous sensitivity calculations (as in the work of *Consideine et al.* [1999] or *Cohen et al.* [2000]), and deserves a separate study.

7. Comparison of the ER-2 and MLS HNO_3 Measurements

7.1. Analysis Using the TCA

[54] Designed primarily to measure stratospheric abundances of CIO, O_3 , and H_2O , UARS MLS also measures

HNO_3 . Previous v.4 MLS measurements of HNO_3 are discussed in detail by *Santee et al.* [1999]. Recently, *Santee et al.* [2000a] presented a first preliminary validation of the v.5 MLS HNO_3 data with Atmospheric Trace Molecular Spectroscopy (ATMOS) and UARS Cryogenic Limb Array Etalon Spectrometer (CLAES) HNO_3 measurements. The v.5 MLS HNO_3 values agree well with both ATMOS (v.3.1) and CLAES (v.9) data at most altitudes in the tropics. At midlatitudes, the agreement between these platforms is generally good, except for the range 46–22 hPa, where the MLS data are higher by up to 35% and 50% compared with ATMOS and CLAES, respectively. Under conditions when HNO_3 is enhanced inside the winter polar vortices in regions of limited PSC activity, MLS values can exceed those of the infrared measurements by 15–60%. Detailed validation of the v.5 MLS HNO_3 data is presented by *Livesey et al.* [2002]. The CIMS is a new instrument aboard ER-2 and operated for the first time during the SOLVE campaign. CIMS validation studies are still under way (K. A. McKinney et al., manuscript in preparation, 2001).

[55] Figure 12 shows the comparison of the matched ER-2 and MLS measurements on 2 February 2000. Because the CIMS instrument measures gaseous HNO_3 during approximately half of the flight time (during the other half it measures condensed-phase HNO_3) and because of the longer integration time for the HNO_3 data, fewer measurements of nitric acid are reported for a given ER-2 flight compared with measurements for ozone. Consequently, fewer ER-2 points are shown in Figure 12 than in Figure 4. In order to increase the number of matches, we relaxed the match criterion to (3 hours, 2° , 10°). If the match criterion of (2 hours, 2° , 2°) is used, only the two matches in the top row of Figure 12 remain. It is difficult to quantify the difference based on the TCA results shown in Figure 12 because of the poor statistics available, large variability of the ER-2 data, and a narrow vertical range of 52–63 hPa covered by ER-2 during its 2 February flight. Usually, the MLS and CIMS ER-2 HNO_3 measurements agree within their uncertainties. On average, perhaps, one can say that the ER-2 values tend to be smaller than the MLS data (especially in Figures 12a and 12e). However, Figure 12f and to a lesser extent Figure 12b show the cases when ER-2 values are larger than those by MLS.

7.2. Analysis Using the THT

[56] To improve the statistical significance of the MLS/ER-2 nitric acid comparison, we applied the THT with a match criterion of (2 hours, 2° , 2°). We found a total of 148 matches (127 and 21 matches for the forward and backward trajectories, respectively), less than for the ozone and CIO cases because of fewer available HNO_3 measurements by ER-2. The results of this analysis are shown by the red, blue, and green lines in Figure 13a for all, forward, and backward trajectories, respectively.

[57] The results shown in Figure 13a assume that gaseous nitric acid is a passive tracer along the matched trajectories. For a good passive tracer, the difference between the forward and backward trajectories should be small, indicating no changes in the tracer concentration with time. However, the large difference (up to 4 ppbv at 430 K) between results for the forward and backward trajectories in

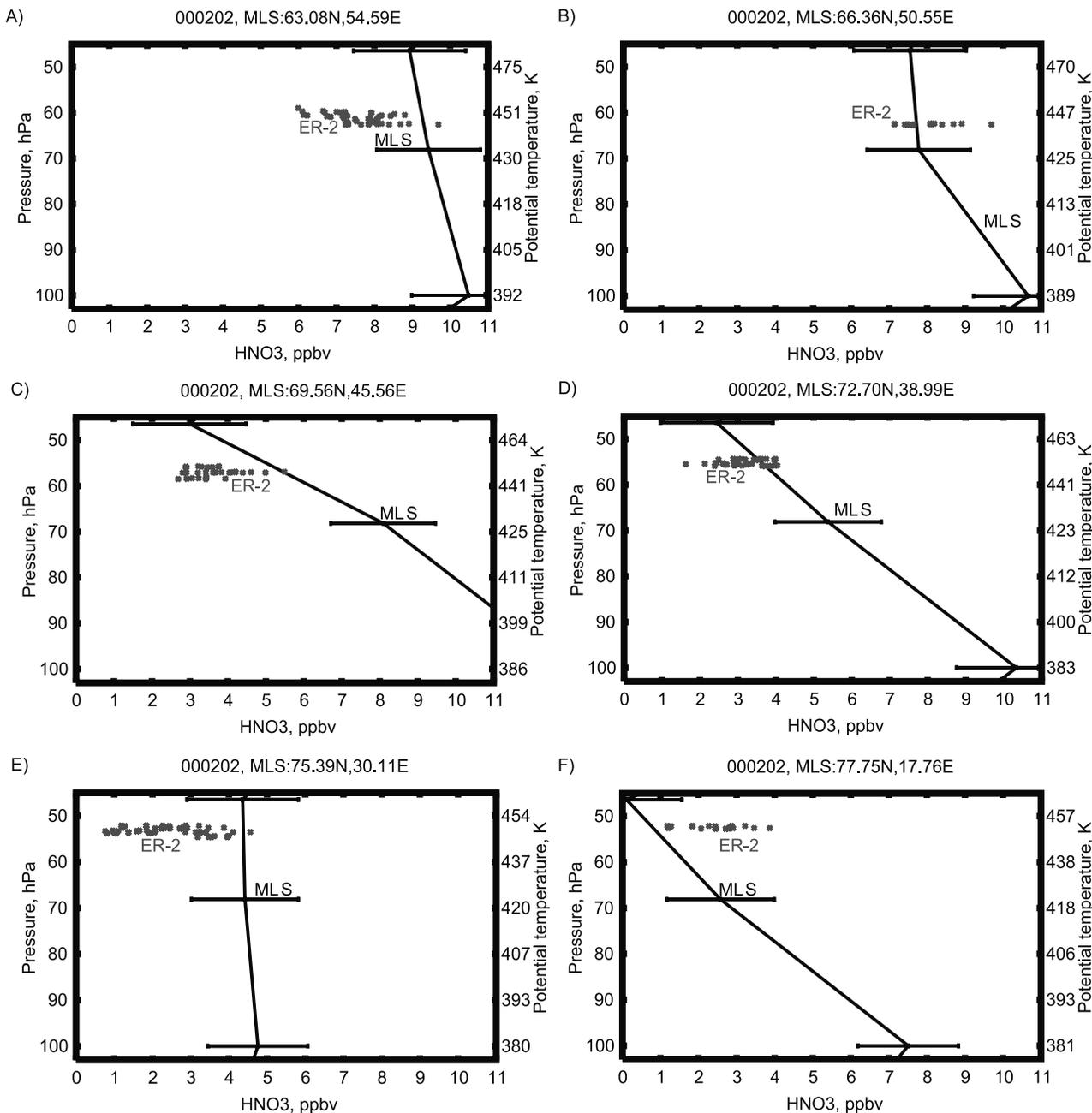


Figure 12. The same as Figure 4, but for the six matches (a-f) of the ER-2 and MLS HNO_3 measurements satisfying the match criterion of (3 hours, 2° , 10°).

Figure 13a clearly signals that nitric acid is not a passive tracer above 400 K. We investigated whether the changes in the HNO_3 concentrations are caused by the following reasons: (1) photochemistry, (2) heterogeneous reactions, (3) irreversible denitrification of the HNO_3 -containing PSC particles, and (4) temporary removal of nitric acid from gas to condensed phases and back via condensation to and evaporation from the PSC particles, respectively.

[58] The photochemical changes of nitric acid are small in our study (less than few tens of pptv according to our model calculations below). However, this statement could be easily confirmed without any detailed model calculations using the fact that, on average, the matched air parcels were illuminated about 6 hours. Assuming a HNO_3 photolysis rate of

the order of $10^{-7} - 10^{-6} \text{ s}^{-1}$, one gets that about 0.2–2% (or $\sim 0.01-0.1$ ppbv) of HNO_3 can be destroyed. This value is consistent with our model calculations and much smaller than the ER-2/MLS differences shown in Figures 12–13.

[59] The effects of heterogeneous reactions on sulfate aerosol and PSCs are also small for HNO_3 . The main reason for this is very small initial values of nitrogen (NO , NO_2 , N_2O_5 and ClONO_2) species during this period, which could be eventually converted to HNO_3 via heterogeneous reactions. All these species have concentrations smaller than several tens of pptv, which could produce at most 0.1 ppbv of additional HNO_3 . The last value is an order of magnitude smaller than the differences we saw in Figure 13a. We performed our model runs using the FCAS aerosol SAD.

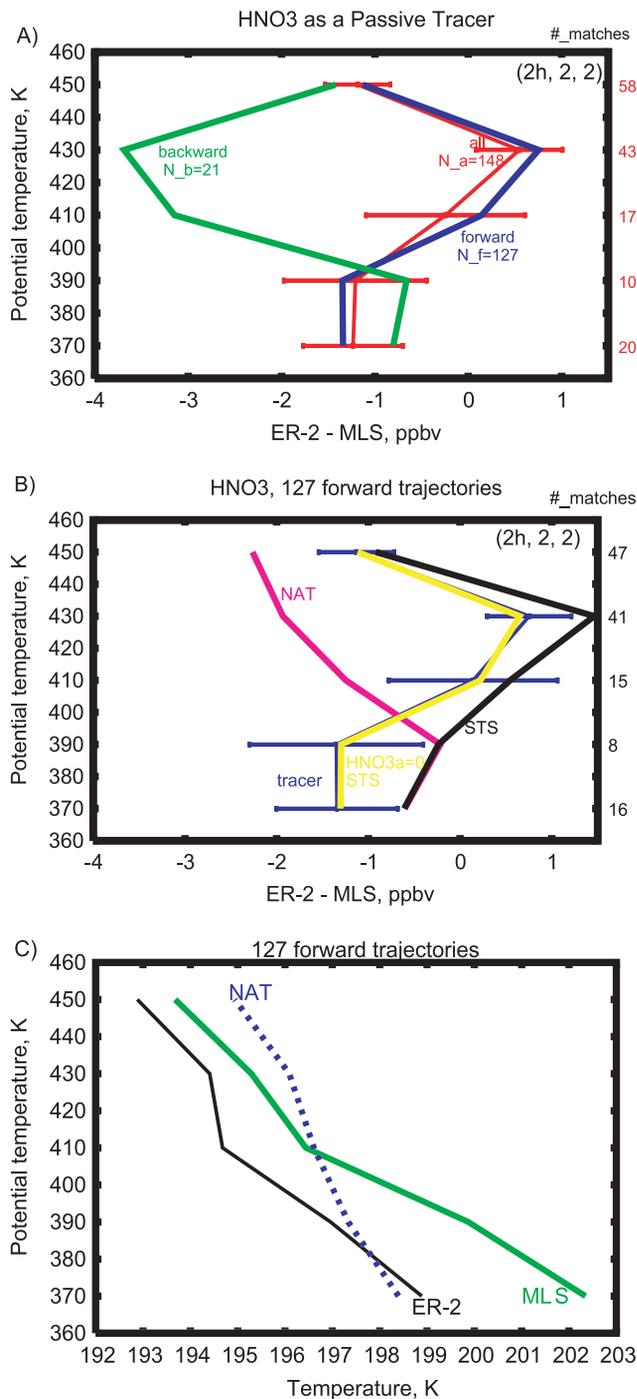


Figure 13. (a) Difference between the MLS and ER-2 gaseous HNO_3 measurements determined by all (red curve with error bars), forward (blue curve), and backward (green curve) trajectories assuming that HNO_3 is a passive tracer. (b) The same as Figure 13a, but for the forward trajectories only. The pink and black lines show the model results with the NAT and STS PSC schemes, respectively. The yellow line shows model results with the STS scheme and zero initial content of HNO_3 in aerosol. The blue line with error bars is the same as that in Figure 13a. (c) Profiles of the averaged temperature in the initial (labeled ER-2) and final (labeled MLS) points of the matched forward trajectories and the NAT threshold (labeled NAT).

While usually the MASP SAD values are larger than those by FCAS, we do not anticipate that our comparisons of the HNO_3 ER-2/MLS measurements will be affected because of the small values of NO , NO_2 , N_2O_5 , and ClONO_2 .

[60] In this study we assume that there is no irreversible denitrification caused by sedimentation of HNO_3 -containing particles between the matched points. This assumption is supported by the following facts. First, temperature along the matched air trajectories was always above the ice frost point, thus precluding formation of large ice particles. Second, *Tabazadeh et al.* [2000] showed that irreversible denitrification could happen in air parcels that stayed for ~ 1 week at temperatures below T_{NAT} and above T_{ice} . In our study, the averaged temporal difference between ER-2 and MLS points is 1.9 days, too short for the nucleation, growth, and sedimentation of the new solid HNO_3 -containing particles to occur along the trajectory, even assuming that all this time temperature was below T_{NAT} . However, the recent discovery of the large 10- to 20- μm particles (so-called “rocks”) [Fahey et al., 2001], which could come from higher layers and reduce (increase) the HNO_3 amount by sedimentation (evaporation), complicates the justification of the above assumption. Indeed, some of these particles were obtained by the forward channel of the NO_y instrument during some periods of the 31 January and 3 February flights [Fahey et al., 2001]. Because the nucleation mechanism for the “rocks” is still unknown, and because modeling of their growth and sedimentation would require a three-dimensional model of the stratosphere, a realistic evaluation of their effect is beyond the scope of this paper. We therefore do not assess possible effects of the “rocks” on our ER-2/MLS HNO_3 comparison and acknowledge that their possible impact could introduce additional error.

[61] Nitric acid may experience rapid changes caused by condensation to and evaporation from PSC particles in the polar lower stratosphere during winter. Qualitatively, the difference for the forward and backward trajectories in Figure 13a could be understood if one takes into account temperatures in the ER-2 and MLS matched points, which is illustrated for the forward trajectories in Figure 13c. The ER-2 flew under temperatures below the NAT threshold most of the time during the three flights considered. For such conditions, HNO_3 may be removed from the gas to solid or liquid phase. On the other hand, on average, temperatures in the MLS points were warmer than those in the ER-2 points. As a result, if the gas and condensed phases are assumed to be in equilibrium, less HNO_3 may be condensed onto PSCs in the MLS locations. Thus for the forward trajectories originating from the ER-2 points, one would anticipate a release of the HNO_3 condensed into the gas phase as the parcels move toward the MLS points, thus making the apparent values of the ER-2 HNO_3 , and therefore the ER-2 – MLS difference, increase. Based on this analysis, one also can say that the large difference between the THT HNO_3 results for forward and backward trajectories indicates possible PSC events captured during the period considered.

[62] To account for the microphysical and, less important, photochemical changes of HNO_3 along the matched forward trajectories, we use the AER box model with NAT or STS PSC schemes [Danilin et al., 2000]. Some information of the model initialization relevant for analysis of the ClO

measurements was given in section 6.2. Here we provide some details of the model initialization that are important for comparison of the HNO_3 measurements. Ideally, the initial concentrations of all nitrogen species should be derived from the ER-2 measurements. However, some nitrogen species (like N_2O_5 or HNO_4) are not measured or are not available simultaneously (like HNO_3 in gas and condensed phases). We use the following approach in order to initialize our model runs. If the difference between the NO_y measured by the backward inlet of the NO_y instrument [Fahey *et al.*, 2001] and the sum of all other nitrogen species is positive, we assigned the balance to nitric acid condensed in aerosol (HNO_3^a). We also checked that the nonzero HNO_3^a values are consistent with low temperature in the ER-2 locations. Indeed, T_{ER2} was less than T_{NAT} in most ($\sim 90\%$) such cases. For the remaining ($\sim 10\%$) cases, we kept the nonzero HNO_3^a value despite relatively warm temperature in these points. Also, the ER-2 measurements of nitrogen species were not always internally consistent and the CIMS HNO_3 values were larger than the ER-2 NO_y data in 15 of 127 matches. The reason of this inconsistency is discussed elsewhere (R. K. McKinney *et al.*, manuscript in preparation, 2001). In such cases, we initialized our model using the ER-2 measurements of HNO_3 , NO_2 , NO , and ClONO_2 , assuming concentrations of all other nitrogen species equal to zero. Even in such cases, we avoid any scaling of initial HNO_3 that may hamper our further comparison. These details are important for the analysis below.

[63] Figure 13b shows that the THT analysis of the MLS and ER-2 HNO_3 measurements is a very challenging task even for the forward trajectories, when we have the best currently available constraints by the ER-2 data on the model initialization and a relatively large amount of matches compared with the backward trajectories case. The blue line with error bars is reproduced from Figure 13a, while the black and pink lines show the model results when the photochemical and PSC-related changes in HNO_3 are accounted for the matched parcels using the STS and NAT schemes, respectively. The obvious feature of this figure is a large difference between the model results with different PSC schemes.

[64] Again, qualitatively this difference could be understood in terms of the averaged temperature in initial and final points of the matched trajectories. Below the 400 K level, the results for the STS and NAT calculations are the same, since the averaged temperature in the final points was above the NAT threshold. Above the 400 K level, the averaged temperature in the final points drops below T_{NAT} , thus causing a big difference in the amount of HNO_3 condensed onto PSC particles in NAT and STS schemes. It is well known that more HNO_3 condenses in the NAT than STS scheme in the temperature range of $[T_{\text{NAT}} - 3\text{K}; T_{\text{NAT}}]$ [Tabazadeh *et al.*, 1994; Carslaw *et al.*, 1994]. This fact is also widely used in order to discriminate between different PSC schemes using satellite HNO_3 measurements [e.g., Santee *et al.*, 1998]. In our study, we see the maximum difference of ~ 3 ppbv between the NAT and STS schemes at the 430 K level.

[65] The yellow line in Figure 13b shows the results of the model runs with the STS PSC scheme and zero initial HNO_3^a . We perform this calculation in order to study the sensitivity of the ER-2/MLS HNO_3 comparison to the initial

aerosol HNO_3 , which was not measured by ER-2 for the matches shown. This scenario is believed to be unrealistic, since ER-2 measurements indicate that condensed-phase HNO_3 is present, and temperatures at the ER-2 locations were below T_{NAT} . However, this case provides a basis for assessing the effects of gas phase and heterogeneous chemistry on gas phase HNO_3 along the trajectories. It is obvious that the yellow line is very close to the blue line and far from the black line. Thus the difference between results for the model STS run (black line) and the tracer scenario is caused by the assumed initial HNO_3^a , which averaged values were equal to 0.66, 1.03, 0.28, 0.78, and 0.28 ppbv at 370, 390, 410, 430, and 450 K, respectively, for our model runs. The barely noticeable difference of 0.04 ppbv between the yellow and blue lines at 370 and 390 K is caused by photochemistry, since the effects of HNO_3 condensation onto PSC particles are removed. Indeed, the initial HNO_3 content in aerosol is equal to zero by the assumption for this scenario, while the final HNO_3^a values are also equal to zero because of the high temperature in the MLS points (see Figure 13c).

[66] Results shown in Figure 13 indicate that the THT suffers for the ER-2/MLS HNO_3 measurement comparison in the polar lower stratosphere during winter because of the rapid HNO_3 condensation to or evaporation from PSC particles. Obtaining an accurate quantitative estimate of the difference between the MLS and ER-2 measurements using the THT is further complicated by its high sensitivity to the PSC scheme used. On the basis of the presented results, we can conclude that the MLS HNO_3 measurements are larger by 1–2 ppbv at 450 K and by ~ 1 ppbv at 370–390 K than those from the ER-2. It is difficult to quantify the difference in the 400–440 K range because of the high sensitivity of the HNO_3 results to the PSC schemes (NAT versus STS).

[67] A big difference in the vertical and horizontal resolution between MLS and ER-2 measurements is another important caveat, which should be kept in mind. These factors limit the usefulness of the THT in providing robust conclusions for the ER-2/MLS HNO_3 comparison in the polar lower stratosphere during winter. However, we anticipate that the THT analysis should be more successful for the HNO_3 measurements in any other regions of the stratosphere that are free from PSC activity.

8. Summary and Conclusions

[68] The SOLVE campaign provided a unique opportunity to validate remote sensing satellite measurements (POAM III, MLS, and SAGE II) with the in situ ER-2 measurements in the polar winter stratosphere. We compared these platform measurements during the January–February 2000 period. We also decided that uncertainties introduced by model initialization and calculations are outside the scope of this paper and deserve a separate study. Analyzing the latest versions of the MLS (v.5), POAM III (v.3), SAGE II (v.6.0), and ER-2 measurements, we conclude the following:

1. The ozone measurements by MLS and SAGE II are in excellent agreement (better than 5% and mostly within 2%), with MLS values slightly larger than those from SAGE II. The POAM III ozone measurements are up to 12% smaller than those from MLS. The MLS ozone values are larger by

6–11% than those from the ER-2 in the 400–480 K vertical range, but huge sampling volume differences make it difficult to provide significant confidence in offset determinations between MLS and ER-2 measurements. The POAM III and ER-2 ozone measurements show better than 8% agreement in the 380–500 K vertical range with some evidence of a very small (about 5%) POAM III positive offset (also mentioned by Lumpe *et al.* [2002]). These differences between the various techniques typically fall within the expected combined accuracies of the different data sets and other assumptions of the methods used, which we estimate to be of the order of 10%. Such small differences between the ozone measurements made by these platforms are very encouraging for scientific applications of these ozone data sets.

2. The POAM III water vapor measurements are in agreement with both the Harvard Lyman α and JPL laser hygrometers within ± 0.5 ppmv (or about $\pm 10\%$) with a hint of a higher offset as large as several tenths of 1 ppmv in the POAM III data in the 360–480 K range.

3. According to the TCA, the MLS and ER-2 ClO measurements agree within their error bars, and MLS shows a positive offset consistent with the 0.1 ppbv bias known to be present in MLS v.5 ClO data in the lower stratosphere. Model calculations constrained by the ER-2 measurements show that the MLS measurements are higher by 0.2 (0.4) ppbv above (below) 430 K. However, the uncertainties introduced by the model calculations are not evaluated in this study and can be large. Also, the sampling volumes for the MLS and ER-2 measurements are very different.

4. The MLS and CIMS ER-2 HNO₃ measurements are consistent within their uncertainties most of the time with some hint of a positive offset by MLS according to the TCA. PSC processing and the high sensitivity of the HNO₃ to the choice of PSC scheme complicate quantification of their difference in the 400–440 K range using the THT. However, at 450 K and below 400 K the positive offset of MLS is 1–2 ppbv and about 1 ppbv, respectively, according to the THT. The statistical significance of these values is not easy to assess with high confidence, however, given the model uncertainties related to the PSC microphysics discussed in section 7 and the different volume sampling by MLS and ER-2.

[69] We estimated the random error when comparing different pairs of instruments and did not account for any systematic errors or possible biases. Assessing the systematic errors of the instruments requires very detailed validation studies, which are outside the scope of this paper. It is possible that our results will be used in a broader framework of these studies, especially those which are currently under way (e.g., for v.5 MLS and v.6.0 SAGE II). In general, our analysis shows that our results agree with the preliminary values of the combined systematic errors of the instruments.

[70] Table 3 summarizes some important statistical data characterizing the THT for the period studied. For example, we found more than 3000 matches for the MLS/POAM III and MLS/SAGE II pairs using the 5-day trajectories compared with 2 and 5 profiles (or 50 and 125 matches assuming 25 matches per profile), respectively, satisfying the same match criteria as in the THT. Table 3 confirms the fact that more matches are obtained for more frequent measurements and more relaxed match criteria.

[71] Our study shows that the trajectory hunting technique is an effective tool in validation of multiplatform measurements that provides more statistically robust conclusions than the TCA, owing to the much larger number of matches obtained. As a first step toward assessing the uncertainties caused by the trajectory calculations, we compared results for the 5- and 1-day trajectories. The similarity between solid and dashed black lines in Figures 2–3 demonstrates that the THT provides stable results. In the future, we would like to investigate how successful the THT could be in conjunction with a photochemical box model for analysis of short-lived species and to estimate model uncertainties for such comparisons. The methodology presented here could be used for validation studies of future space-borne platforms (like SAGE III or EOS Aura).

Appendix A: Error Analysis

[72] The standard deviation of the differences between any two platform measurements is defined according to (1):

$$SD = \sqrt{\sum_{i=1}^N (\Delta_i - \bar{\Delta})^2 / (N - 1)}; \quad (1)$$

here N is the number of matches, Δ_i is the difference for the i -th pair (e.g., $\Delta_i = O_{3i}^{\text{MLS}} - O_{3i}^{\text{POAM III}}$), and $\bar{\Delta} = \sum_{i=1}^N \Delta_i / N$ is the mean difference between the two instrument measurements. The values of Δ_i could be expressed in absolute (ppmv or ppbv) or relative (percent) units. The standard deviation characterizes the spread of the distribution near the mean value (i.e., $\bar{\Delta}$) and is a measure of the combined random error of both instruments. Thus it does not account for the systematic errors. The error bars shown in all figures presenting the THT results are determined according to (2):

$$ERR = SD / \sqrt{N}, \quad (2)$$

showing that the error bars become smaller for larger N . The values of ERR represent the standard error of the mean differences and correspond to the 67% confidence level. The values of ERR should be doubled for the 95% confidence.

[73] **Acknowledgments.** We thank R.J. Salawitch for providing the ER-2 merged files and D. W. Fahey, C. R. Webster, K. K. Perkins, J. C. Wilson, and D. G. Baumgardner for their instrument data used in our model initialization. We thank A. Tabazadeh for her STS code. Comments of two anonymous reviewers are appreciated. Work at AER, Inc., was supported by the UARS Guest Investigator Program (NAS5-98131) and NASA Atmospheric Chemistry and Model Analysis Program (NAS5-97039 and NAS1-00138). M.Y.D. acknowledges the travel support from the NASA Atmospheric Effects of Aviation Project for his involvement in SOLVE. L.V.L. is partially supported by the UARS Guest Investigator Program (S10109-X). Work at the Jet Propulsion Laboratory, California Institute of Technology, was done under contract with NASA. The ER-2 instrument coauthors acknowledge support from the NASA Upper Atmosphere Research Program. NCAR is supported by the National Science Foundation.

References

- Bacmeister, J. T., V. Kuell, D. Offermann, M. Riese, and J. W. Elkins, Intercomparison of satellite and aircraft observations of long-lived tracers using trajectory mapping, *J. Geophys. Res.*, *104*, 16,379–16,390, 1999.

- Barath, F., et al., The UARS Microwave Limb Sounder instrument, *J. Geophys. Res.*, *98*, 10,751–10,762, 1993.
- Baumgardner, D. G., et al., Refractive indices of aerosols in the upper troposphere and lower stratosphere, *Geophys. Res. Lett.*, *23*, 749–752, 1996.
- Bevilacqua, R. M., Introduction to the special section: Polar Ozone and Aerosol Measurement (POAM II), *J. Geophys. Res.*, *102*, 23,591–23,592, 1997.
- Carslaw, K. S., et al., Stratospheric aerosol growth and HNO₃ gas phase depletion from coupled HNO₃ and wake uptake by liquid particles, *Geophys. Res. Lett.*, *21*, 2479–2482, 1994.
- Chipperfield, M. P., et al., Analysis of UARS data in southern polar vortex in September 1992 using a chemical transport model, *J. Geophys. Res.*, *101*, 18,861–18,881, 1996.
- Chu, W. P., M. P. McCormick, J. Lenoble, C. Brogniez, and P. Pruvost, SAGE II inversion algorithm, *J. Geophys. Res.*, *94*, 8339–8351, 1989.
- Cohen, R. C., et al., Quantitative constraints on the atmospheric chemistry of nitrogen oxides: An analysis along chemical coordinates, *J. Geophys. Res.*, *105*, 24,283–24,304, 2000.
- Considine, D. B., et al., A Monte Carlo uncertainty analysis of ozone trend predictions in a two-dimensional model, *J. Geophys. Res.*, *104*, 1749–1765, 1999.
- Cunnold, D. M., H. 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*, 10,061–10,076, 1996.
- Danilin, M. Y., M. L. Santee, J. M. Rodriguez, M. K. W. Ko, J. L. Mergenthaler, J. B. Kumer, A. Tabazadeh, and N. Livesey, Trajectory hunting: Analysis of rapid chlorine activation in December 1992 as seen by UARS, *J. Geophys. Res.*, *105*, 4003–4018, 2000.
- Fahey, D. W., et al., The detection of large HNO₃-containing particles in the winter Arctic stratosphere, *Science*, *291*, 1026–1031, 2001.
- Feist, D. G., et al., Validation of stratospheric ClO measurements from MAS, *J. Geophys. Res.*, *105*, 9053–9062, 2000.
- Froidevaux, L., et al., Validation of UARS MLS ozone measurements, *J. Geophys. Res.*, *101*, 10,017–10,060, 1996.
- Gille, J. C., and L. V. Lyjak, Radiative heating and cooling rates in the middle atmosphere, *J. Atmos. Sci.*, *43*, 2215–2229, 1986.
- Glaccum, W., et al., The Polar Ozone and Aerosol Measurement (POAM) instrument, *J. Geophys. Res.*, *101*, 14,479–14,488, 1996.
- Hanson, D. R., and K. Mauersberger, Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere, *Geophys. Res. Lett.*, *15*, 855–858, 1988.
- Hints, E. J., E. M. Weinstock, J. G. Anderson, R. D. May, and D. F. Hurst, On the accuracy of in situ water vapor measurements in the troposphere and lower stratosphere with the Harvard Lyman α hygrometer, *J. Geophys. Res.*, *104*, 8183–8189, 1999.
- Hitchman, M. H., C. B. Leovy, J. C. Gille, and P. L. Bailey, Quasi-stationary asymmetric circulations in the equatorial lower mesosphere, *J. Atmos. Sci.*, *44*, 2219–2236, 1987.
- Hoppel, K., R. M. Bevilacqua, G. Nedoluha, C. Deniel, F. Lefevre, I. D. Lumpe, M. D. Fromm, C. E. Randall, J. E. Rosenfield, and M. Rex, POAM III observations of Arctic ozone loss for the 1999/2000 winter, *J. Geophys. Res.*, *107*(D20), 8262, doi:10.1029/2001JD000476, 2002.
- Jonsson, H. H., et al., Performance of the focused cavity aerosol spectrometer for measurements in the stratosphere of particle size of 0.06–2 μ m diameter size, *J. Atmos. Oceanic Technol.*, *12*, 115–129, 1995.
- Kawa, S. R., et al., Activation of chlorine in sulfate aerosol as inferred from aircraft observations, *J. Geophys. Res.*, *102*, 3921–3933, 1997.
- Lingenfelter, G. S., W. L. Grose, E. E. Remsberg, T. D. Fairlie, and R. B. Pierce, Comparison of satellite and in situ ozone measurements in the lower stratosphere, *J. Geophys. Res.*, *104*, 13,971–13,979, 1999.
- Livesey, N. J., W. G. Read, L. Froidevaux, J. W. Waters, M. L. Santee, H. C. Pumphrey, D. L. Wu, Z. Shippony, and R. F. Jarnot, UARS Microwave Limb Sounder version 5 data set: Theory, characterization, and validation, *J. Geophys. Res.*, *107*, 10.1029/2002JD002273, in press, 2002.
- Lu, J., V. A. Mohnen, G. K. Yue, and H. Jäger, Intercomparison of multi-platform stratospheric aerosol and ozone observations, *J. Geophys. Res.*, *102*, 16,127–16,136, 1997a.
- Lu, J., V. A. Mohnen, G. K. Yue, R. J. Atkinson, and W. A. Matthews, Intercomparison of stratospheric ozone profiles obtained by SAGE II, HALOE, and ozonesondes, *J. Geophys. Res.*, *102*, 16,137–16,144, 1997b.
- Lucke, R. L., et al., The POAM III instrument and early validation results, *J. Geophys. Res.*, *104*, 18,785–18,799, 1999.
- Lumpe, J. D., K. W. Hoppel, R. M. Bevilacqua, E. Browell, T. McGee, E. C. Richard, G. Toon, B. Sen, K. Pfeilsticker, and H. Boesch, Comparison of POAM III ozone measurements with correlative aircraft and balloon data during SOLVE, *J. Geophys. Res.*, *107*, doi:10.1029/2001JD000472, in press, 2002.
- Lutman, E. R., et al., Three-dimensional studies of the 1991/1992 northern winter using domain-filling trajectories with chemistry, *J. Geophys. Res.*, *102*, 1479–1488, 1997.
- Manney, G. L., et al., Comparison of UKMO and U.S. NMC stratospheric analyses during northern and southern winter, *J. Geophys. Res.*, *101*, 10,311–10,334, 1996.
- Manney, G. L., et al., Comparison of satellite ozone observations in coincident air masses in early November 1994, *J. Geophys. Res.*, *106*, 9923–9944, 2001.
- Massie, S. T., et al., Chlorine activation during the early 1995–1996 Arctic winter, *J. Geophys. Res.*, *105*, 7111–7131, 2000.
- May, R. D., Open-path, near-infrared tunable diode laser spectrometer for atmospheric measurements of H₂O, *J. Geophys. Res.*, *103*, 19,161–19,172, 1998.
- Morris, G. M., et al., Trajectory mapping and applications to data from the UARS, *J. Geophys. Res.*, *100*, 16,497–16,506, 1995.
- Morris, G. M., J. F. Gleason, J. Ziemke, and M. R. Schoeberl, Trajectory mapping: A tool for validation of trace gas observation, *J. Geophys. Res.*, *105*, 17,875–17,894, 2000.
- Nedoluha, G. E., et al., POAM III measurements of dehydration in the Antarctic lower stratosphere, *Geophys. Res. Lett.*, *27*, 1683–1686, 2000.
- Perkins, K. K., et al., The NO_x-HNO₃ system in the lower stratosphere: Insights from in situ measurements and implications of the J_{HNO₃}[OH] relationship, *J. Phys. Chem.*, *105*, 1521–1534, 2001.
- Pierce, R. B., W. L. Grose, and J. M. Russell III, Evolution of Southern Hemisphere air masses observed by HALOE, *Geophys. Res. Lett.*, *21*, 213–216, 1994.
- Pierce, R. B., et al., Photochemical calculations along air mass trajectories during ASHOE/MAESA, *J. Geophys. Res.*, *102*, 13,153–13,168, 1997.
- Prather, M. J., and A. H. Jaffe, Global impact of the Antarctic ozone hole: Chemical propagation, *J. Geophys. Res.*, *95*, 3473–3492, 1990.
- Proffitt, M. H., et al., In situ ozone measurements within the 1987 Antarctic ozone hole from a high-altitude ER-2 aircraft, *J. Geophys. Res.*, *94*, 16,547–16,555, 1989.
- Rex, M., et al., In situ measurements of stratospheric ozone depletion rates in the Arctic winter 1991/1992: A Lagrangian approach, *J. Geophys. Res.*, *103*, 5843–5853, 1998.
- Russell, P. B., and M. P. McCormick, SAGE II aerosol data validation and initial data use, *J. Geophys. Res.*, *94*, 8335–8338, 1989.
- Sander, S. P., et al., Chemical kinetics and photochemical data for use in stratospheric modeling, evaluation 13, *JPL Publ. 00-3*, Jet Propul. Lab., Pasadena, Calif., 2000.
- Santee, M. L., et al., UARS MLS HNO₃ observations: Implications for Antarctic PSCs, *J. Geophys. Res.*, *103*, 13,285–13,313, 1998.
- Santee, M. L., et al., Six years of UARS MLS HNO₃ observations: Seasonal, interhemispheric, and interannual variations in the lower stratosphere, *J. Geophys. Res.*, *104*, 8225–8246, 1999.
- Santee, M. L., et al., HNO₃ measurements from MLS on the UARS and EOS CHEM satellites, paper presented at Am. Meteorol. Soc. Meet., Long Beach, Calif., Jan. 2000a.
- Santee, M. L., et al., UARS MLS observations of chemical processing in the 2000 Arctic late winter, *Geophys. Res. Lett.*, *27*, 3213–3216, 2000b.
- Schoeberl, M. R., and L. Sparling, Trajectory modeling, in *Proceedings of the International School of Physics: Diagnostic Tools in Atmospheric Physics*, edited by G. Fiocco and G. Visconti, pp. 289–305, Elsevier Sci., New York, 1994.
- Shepherd, T. G., J. N. Koshyk, and K. Ngan, On the nature of large-scale mixing in the stratosphere and mesosphere, *J. Geophys. Res.*, *105*, 12,433–12,446, 2000.
- Smith, A. K., and L. V. Lyjak, An observational estimate of gravity wave drag from the momentum balance in the middle atmosphere, *J. Geophys. Res.*, *90*, 2233–2241, 1985.
- Stimpfle, R. M., et al., The coupling of ClONO₂, ClO, and NO₂ in the lower stratosphere from in situ observations using the NASA ER-2 aircraft, *J. Geophys. Res.*, *104*, 26,705–26,714, 1999.
- Stratospheric Processes and Their Role in Climate (SPARC), SPARC water vapor assessment, *Rep. 2*, edited by D. Kley, J. M. Russell III, and C. Phillips, World Meteorol. Org., Geneva, 2001.
- Sutton, R. T., H. Maclean, R. Swinbank, A. O'Neill, and F. W. Taylor, High-resolution stratospheric tracer fields estimated from satellite observations using Lagrangian trajectory calculations, *J. Atmos. Sci.*, *51*, 2995–3005, 1994.
- Swinbank, R., and A. O'Neill, A stratosphere-troposphere data assimilation system, *Mon. Weather Rev.*, *122*, 686–702, 1994.
- Tabazadeh, A., R. P. Turco, and M. Z. Jacobson, A model for studying the composition and chemical effects of stratospheric aerosol, *J. Geophys. Res.*, *99*, 12,897–12,914, 1994.
- Tabazadeh, A., M. L. Santee, M. Y. Danilin, H. C. Pumphrey, P. A. Newman, P. J. Hamill, and J. L. Mergenthaler, Quantifying denitrification and its effect on ozone recovery, *Science*, *288*, 1407–1411, 2000.

- von der Gathen, P., et al., Observational evidence for chemical ozone depletion over the Arctic in winter 1991–92, *Nature*, 375, 131–134, 1995.
- Wamsley, P. R., et al., Distribution of halon 1211 in the upper troposphere and lower stratosphere and the 1994 total bromine budget, *J. Geophys. Res.*, 103, 1513–1526, 1998.
- Waters, J. W., Microwave limb sounding, in *Atmospheric Remote Sensing by Microwave Radiometry*, edited by M. A. Janssen, pp. 383–496, John Wiley, New York, 1993.
- Waters, J. W., et al., Validation of UARS MLS ClO measurements, *J. Geophys. Res.*, 101, 10,091–10,127, 1996.
- Weinstock, E. M., et al., A new fast response photofragment fluorescence hygrometer for use on the NASA ER-2 and Perseus remotely piloted aircraft, *Rev. Sci. Instrum.*, 65, 3544–3554, 1994.
- Weisenstein, D. K., et al., The effects of sulfur emissions from HSCT aircraft: A 2-D model intercomparison, *J. Geophys. Res.*, 103, 1527–1548, 1998.
- World Meteorological Organization (WMO), SPARC Assessment of trends in the vertical distribution of ozone, *Rep.* 43, 289 pp., Geneva, 1998.
- T. V. Bui, NASA Ames Research Center, Mail Stop 245-5, Moffett Field, CA 94035, USA.
- M. Y. Danilin, The Boeing Company, MS 0R-RC, P. O. Box 3707, Seattle, WA 98124-2207, USA. (danilin@h2o.ca.boeing.com)
- F. L. Eisele and L. V. Lyjak, ACD, NCAR, P.O. Box 3000, Boulder, CO 80303, USA.
- J. W. Elkins and E. C. Richard, NOAA Aeronomy Laboratory, 325 Broadway, Boulder, CO 80303, USA.
- L. Froidevaux, R. L. Herman, and M. L. Santee, NASA JPL, 4800 Oak Grove Drive, Pasadena, CA 91109, USA.
- M. K. W. Ko and C. J. Scott, AER, Inc., 131 Hartwell Avenue, Lexington, MA 02421, USA.
- K. A. McKinney and P. O. Wennberg, Division of Geology and Planetary Sciences, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA.
- J. R. Spackman, Department of Earth and Planetary Sciences, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA.
- R. M. Stimpfle and E. M. Weinstock, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA.
- J. M. Zawodny NASA Langley Research Center, Hampton, VA 23681, USA.
-
- R. M. Bevilacqua and K. W. Hoppel, Naval Research Laboratory, 4555 Overlook Avenue, SW, Washington, D. C., 20375, USA.