

## Validation of Polar Ozone and Aerosol Measurement (POAM) III version 4 stratospheric water vapor

J. Lumpe,<sup>1</sup> R. Bevilacqua,<sup>2</sup> C. Randall,<sup>3</sup> G. Nedoluha,<sup>2</sup> K. Hoppel,<sup>2</sup> J. Russell,<sup>4</sup>  
V. L. Harvey,<sup>3</sup> C. Schiller,<sup>5</sup> B. Sen,<sup>6</sup> G. Taha,<sup>7</sup> G. Toon,<sup>6</sup> and H. Vömel<sup>8</sup>

Received 12 October 2005; revised 6 January 2006; accepted 27 January 2006; published 2 June 2006.

[1] The Polar Ozone and Aerosol Measurement (POAM) III solar occultation instrument has been measuring water vapor at high latitudes since April 1998. Retrievals extend from 5 to 50 km, with 5–7% precision throughout the stratosphere and a vertical resolution of 1 (3) km in the lower (upper) stratosphere. Estimated systematic errors in the stratosphere are 10–15%. In this paper, we validate the POAM III version 4 stratospheric water vapor using correlative measurements from satellite, airborne, and balloon-borne platforms. The resulting comparisons show that POAM water vapor is high compared to correlative measurements in the middle to lower stratosphere. The satellite (Halogen Occultation Experiment (HALOE) and Stratospheric Aerosol and Gas Experiment (SAGE) II) comparisons also indicate a sunrise/sunset bias in the POAM data, with sunset (Southern Hemisphere) retrievals larger than sunrise (Northern Hemisphere) retrievals by 5–10%. In the Northern Hemisphere, POAM is approximately 5–10% high compared to all validation data sets between 12 and 35 km. At higher altitudes this difference decreases, such that POAM agrees with HALOE at 40 km and is lower by 10% at 50 km. In the Southern Hemisphere, POAM is 15–25% higher than HALOE below 35 km, with differences decreasing to 10% by 50 km. Similar differences are seen with SAGE II. Despite these systematic differences the POAM water vapor data are self-consistent and show no long-term trends in accuracy or precision. Statistical comparisons of the water vapor variability measured by POAM, HALOE, and SAGE II show very good agreement. The POAM data are therefore valid for scientific studies, and the science community is encouraged to use this unique data set.

**Citation:** Lumpe, J., et al. (2006), Validation of Polar Ozone and Aerosol Measurement (POAM) III version 4 stratospheric water vapor, *J. Geophys. Res.*, 111, D11301, doi:10.1029/2005JD006763.

### 1. Introduction

[2] Water vapor is an extremely important middle atmospheric constituent. It is a greenhouse gas, it is the source of the OH radical, which is important in ozone photochemistry, and it is an ideal tracer of atmospheric motions. Polar stratospheric water vapor is particularly important because of the role of ice bearing Type II polar stratospheric clouds (PSCs) in the dehydration and denitrification of the Antarctic

polar stratosphere, and the potential importance of this phenomenon in initiating ozone loss in the polar regions. This dehydration is a significant sink for middle atmospheric water vapor. Also, because of its tracer properties, water vapor in the polar stratosphere is an important diagnostic of descent in the polar vortex.

[3] The Polar Ozone and Aerosol Measurement (POAM) III solar occultation instrument (hereafter referred to simply as POAM) has been making measurements of stratospheric and upper tropospheric water vapor in the polar regions from April 1998 to the present. For most of this time period, POAM has been the only satellite instrument continuously monitoring water vapor in the polar stratosphere. The Halogen Occultation Experiment (HALOE) and Stratospheric Aerosol and Gas Experiment (SAGE) II instruments have been operational during this time period and also measure water vapor, but the orbits of both instruments are such that the polar regions are only sampled once every 6–7 weeks for several days. In March 2002 the SAGE III instrument also began measuring water vapor from a polar orbit, reaching somewhat higher latitudes than POAM in the Northern Hemisphere (NH) but seldom sampling the Southern Hemisphere (SH) polar vortex. However, at the current

<sup>1</sup>Computational Physics, Inc., Boulder, Colorado, USA.

<sup>2</sup>Naval Research Laboratory, Washington, D. C., USA.

<sup>3</sup>Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, Colorado, USA.

<sup>4</sup>Center for Atmospheric Sciences, Hampton University, Hampton, Virginia, USA.

<sup>5</sup>Institut für Chemie und Dynamik der Geosphäre I, Forschungszentrum Jülich, Jülich, Germany.

<sup>6</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.

<sup>7</sup>Science Systems and Applications, Inc., Lanham, Maryland, USA.

<sup>8</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA.

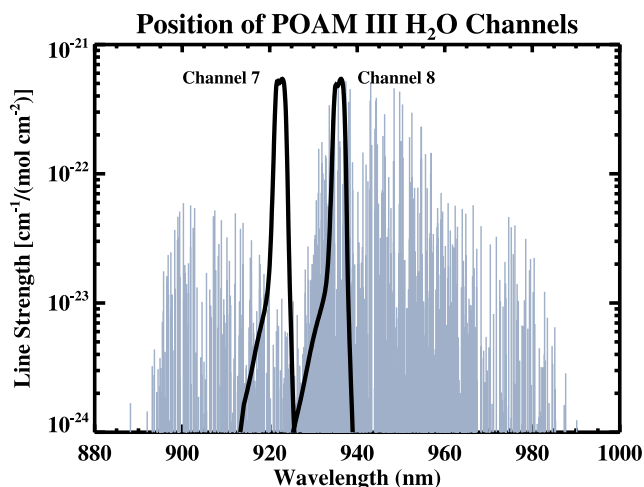
time operational water vapor retrievals from the SAGE III instrument are not yet available. At least two other recent satellite instruments have provided some measurements of polar stratospheric water vapor. These include the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) instrument, which has been measuring water vapor at high latitudes using solar occultation since its launch in August 2003, and the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) instrument on the ENVISAT satellite, which measured water vapor using limb emission, with near-global coverage, from July 2002 through March 2004. Thus, in terms of its temporal and spatial coverage, the 7-year POAM III data record is unique and represents a valuable resource for studying polar stratospheric and upper tropospheric water vapor. The POAM version 3 retrievals were used by *Nedoluha et al.* [2000, 2002a] to document, in unprecedented detail, the dehydration of the Antarctic vortex. In addition, *Benson et al.* [2006] have used POAM version 4 water vapor data with a microphysical model to investigate ice cloud formation and dehydration in the Antarctic vortex.

[4] In this paper, we present a comprehensive validation study of the stratospheric component ( $>10$  km) of the POAM version 4 water vapor retrievals. Version 4, the current POAM operational data set, incorporates a number of changes from version 3 which are briefly summarized in section 2. Since this paper represents the first published validation of the POAM stratospheric water vapor, a detailed discussion of the version 3/version 4 differences is not presented. The version 3 upper tropospheric/lower stratospheric (UT/LS) water vapor measurements were previously validated by *Nedoluha et al.* [2002b], and a separate paper is in preparation which quantifies changes in the version 4 water vapor in this critical region.

[5] The validation analysis presented here relies primarily on extensive statistical comparisons with the HALOE and SAGE II satellite solar occultation instruments. The HALOE version 19 data are available throughout the entire POAM time period, and because they have been extensively validated [see, e.g., *Kley et al.*, 2000] these comparisons form the cornerstone of our analysis. For many years the SAGE II water vapor retrievals had known deficiencies and were in a state of flux. However, this data set has recently stabilized with the release of version 6.2 [Thomason et al., 2004], and therefore it has been included in this analysis. In addition to the satellite data sets, many stratospheric water vapor measurements have been made at high northern latitudes from both airborne and balloon-borne platforms. During the 1999/2000 NH winter POAM participated in the SOLVE/THESEO 2000 campaign, which resulted in a number of coincident measurements with balloons and aircraft. Additional comparisons with balloon-borne instruments are also available from the SOLVE II mission in the winter of 2002/2003, as well as several flights in 2004. These comparisons are presented in this paper as a check on the conclusions drawn from the much more extensive satellite comparisons.

## 2. POAM Instrument and Retrievals

[6] POAM uses the solar occultation technique to measure atmospheric slant path transmission (or optical depth)



**Figure 1.** Filter transmission curves (arbitrarily normalized) for the POAM water vapor “on” channel at 935.9 nm and “off” channel at 922.2 nm. The shaded lines represent the HITRAN 2000 water vapor transition line strengths.

in nine spectral channels from 354 to 1018 nm. Inversion of the optical depth data yield vertical profiles of aerosol extinction, ozone, water vapor, and nitrogen dioxide in the stratosphere and upper troposphere. The instrument and its operations are described by *Lucke et al.* [1999]. POAM was launched on the French SPOT 4 satellite in March 1998 into a polar, Sun-synchronous orbit. In this orbit 14 occultations are obtained per day around a circle of latitude in each hemisphere, with successive observations separated by about  $25^\circ$  longitude. The measurement latitude varies slowly over the course of a year, with an approximate semiannual period that is exactly repeatable from year to year. The latitude range is  $63^\circ$  to  $88^\circ$  in the SH, and  $55^\circ$  to  $71^\circ$  in the NH. All POAM NH measurements are made during spacecraft (s/c) sunrise, but correspond to local sunset. All SH observations are obtained during s/c sunset, and correspond to local sunrise from mid-April through mid-September and local sunset at other times. POAM operated in its nominal auto mode, obtaining 14 events per day in each hemisphere (28 total), for approximately the first year of operations. At that point POAM began to experience a serious anomaly in which it would occasionally stick in azimuth during an event. To reduce wear of the azimuth motor, the POAM nominal observing scheme was changed to alternate daily between sunrise (NH) and sunset (SH) occultations. This of course yields only 14 occultations per day, thus reducing the POAM effective data rate by 50%, but was successful in prolonging the instrument lifetime. The operational mode is programmable from the ground, and has been changed on numerous occasions when a different sampling pattern is desirable (e.g., to provide denser NH coverage during the SOLVE missions).

[7] The POAM water vapor measurement is made using the 940-nm water vapor absorption band. Specifically, POAM utilizes the differential absorption between a primary “on” channel at 935.9 nm and a nearby “off” channel at 922.4 nm. Figure 1 illustrates the location and shape of the spectral band pass for these two channels. Because water vapor absorption in the stratosphere is a weak signal relative

to the background Rayleigh and aerosol scattering extinction, even in the primary H<sub>2</sub>O channel, the use of closely spaced differential channels improves the water vapor sensitivity significantly over a single channel approach. The POAM version 3 retrieval algorithm was described in detail by Lumpe *et al.* [2002a], and the version 4 algorithm does not alter the basic method used to retrieve water vapor density profiles. The fundamental POAM water vapor retrieval product is volume density versus altitude. Conversion to mixing ratio for scientific analysis or comparison with other measurements is accomplished using temperature/pressure profiles derived from the United Kingdom Meteorological Office (Met Office) analysis and interpolated to the POAM occultation location and time.

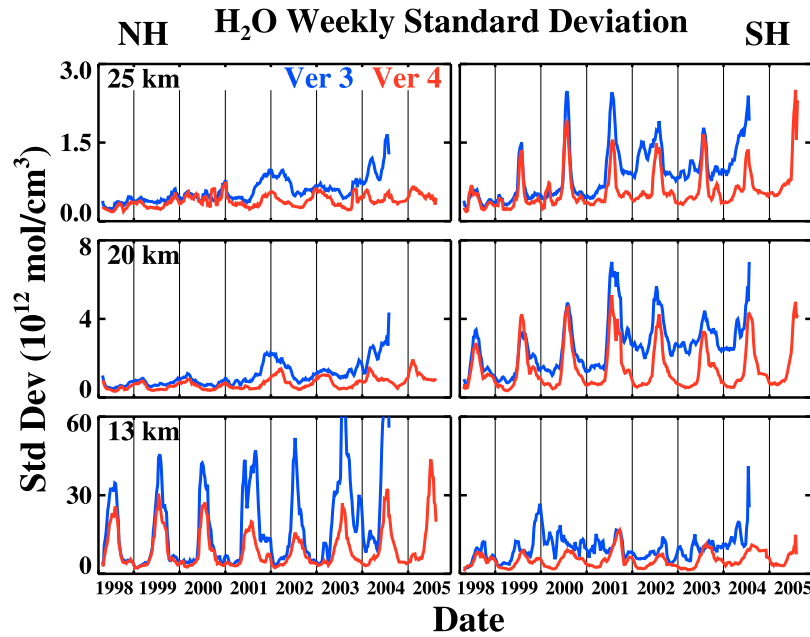
[8] A key feature of the POAM retrieval method is the global nature of the spectral inversion, in which all species (ozone, water vapor, nitrogen dioxide, and aerosols) are retrieved simultaneously. The spectral dependence of the gas species is determined by their respective absorption cross sections, whereas the aerosol optical depth spectrum is constrained by a simple low-order polynomial parameterization. The study by Lumpe *et al.* [2002a] contains a detailed description of the algorithms used to calculate the H<sub>2</sub>O effective cross sections and their implementation in the spectral inversion. The version 4 retrievals use the HITRAN 2000 [Rothman *et al.*, 2003] water vapor line transition parameters. This is a change from version 3, which used the HITRAN 96 line parameters but included the 14.4% correction (line strength increase) in the 940 nm absorption band reported by Giver *et al.* [2000]. These line strength corrections, and other updates to the line widths and temperature coefficients, were incorporated into the HITRAN 2000 database. The spectroscopic change alone results in a slight increase in the retrieved water vapor from version 3 to 4. The magnitude of the increase is generally a few percent above 20 km and goes to zero at lower altitudes. We have evaluated the new HITRAN 2004 data set in the POAM retrievals and found that changes from the HITRAN 2000 results are insignificant within the retrieval error bars.

[9] The primary improvements made in the version 4 algorithm involve changes in the level 1 algorithms that determine the instrument pointing information. During the lifetime of POAM, there has been degradation in the instrument elevation tracking mechanism. This has made it necessary to increase the tracking gain in the elevation servo loop several times during the course of the mission, which, in turn, has resulted in increased jitter in the tracking motion during an occultation scan. Also, nonlinearities in the potentiometer used to monitor the elevation motion have increased over the mission life. These degradations led to an increase over time in the noise and potential biases of the instrument's absolute pointing information, which of course resulted in increased noise and biases in the version 3 retrievals. Most affected were the water vapor and aerosol extinction products, because they represent weak extinction signals and are derived primarily from the long-wavelength channels, which are most sensitive to small pointing errors. There was little degradation evident in the version 3 ozone retrievals. The version 3 validation papers that have been published [e.g., Randall *et al.*, 2003, and references therein] were written earlier in the mission before significant degradation of the elevation pointing system had occurred.

[10] The version 4 algorithm changes were primarily aimed at compensating for these problems. Specific improvements include introducing additional smoothing of the level 0 radiances to decrease the impact of tracking jitter. The smoothing, which is done on the native 18 Hz data grid before binning to the standard POAM 1 km retrieval grid, has a vertical scale of 1 km or less, and therefore does not degrade the effective vertical resolution of the retrievals. The calibration of the potentiometer, which relates the optical head movement to changes in altitude, has also been improved making it less susceptible to errors due to changing potentiometer nonlinearity. Because of the increased problems with the elevation potentiometer, a significant effort was made to reduce the overall dependence on the potentiometer data in the pointing algorithms. As with version 3, the potentiometer data is used to determine instrument pointing only at the lowest altitudes of each scan. At higher altitudes, the pointing is determined by maintaining the expected optical depth ratios between Rayleigh-dominated channels. This method has been improved in version 4, allowing it to be used down to about 26 km (approximately 8 km lower than in version 3).

[11] Another change made in version 4 is the parameterization used to constrain the spectral dependence of the aerosol optical depth ( $\delta$ ). All previous versions assumed a quadratic polynomial in  $\ln \delta(\lambda)$  for the spectral inversion, where  $\lambda$  is the wavelength [Lumpe *et al.*, 2002a]. In the version 4 algorithm we have adopted a simple quadratic fit to  $\delta(\lambda)$  directly. This removes the primary source of nonlinearity in the spectral inversion and leads to a faster and more robust species separation without a significant loss of accuracy. It also eliminates occasional problems with retrieval nonconvergence that occurred in previous versions, and reduces aerosol feedback into the gas retrievals. This is likely because the old, nonlinear parameterization is positive-definite and therefore could not produce a negative aerosol component. While a negative aerosol optical depth is obviously unphysical, there are times (e.g., the cleansed springtime polar vortex) when the aerosol loading is at or below the POAM detection limit and allowing negative aerosol values for individual retrieval profiles produces a more realistic statistical result in the mean.

[12] These improvements have resulted in a significant improvement in the precision and long-term consistency of the water vapor data set. The precision of the version 3 water vapor retrievals degraded over the life of the mission as a direct result of the increased pointing uncertainties. This has been greatly reduced in the version 4 retrievals, as Figure 2 demonstrates. This figure shows the retrieval standard deviation in one-week time bins at selected altitudes from the version 3 and 4 data sets. In both hemispheres the standard deviation is a maximum in the winter, reflecting a large increase in geophysical variability. As discussed by Lumpe *et al.* [2002a] the summertime retrieval variability provides a reasonable upper limit to the actual retrieval precision (see also Figure 4). It is clear from Figure 2 that in version 3 the retrieval precision degraded dramatically in recent years, primarily at 25 km and below but also to a lesser extent above 25 km (not shown). The variability in version 4 is greatly reduced, and for the most part shows no statistically significant long-term trend in the minimum summertime periods. However, some degradation



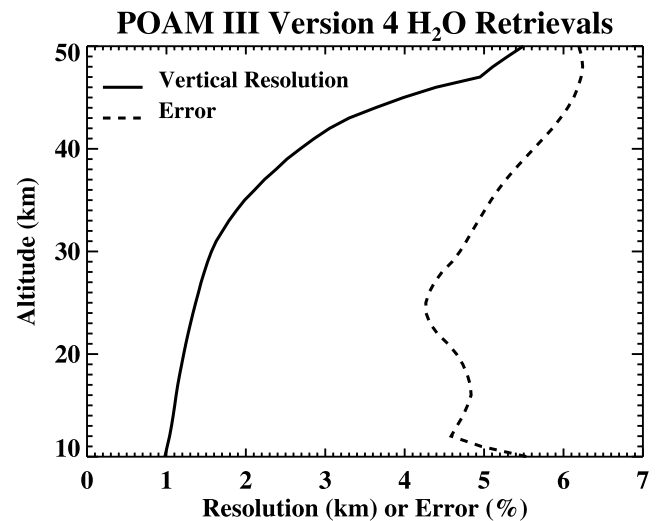
**Figure 2.** Comparison of version 3 (blue) and version 4 (red) retrieval standard deviation at selected altitudes during the POAM mission. The left and right panels show results from the Northern and Southern Hemispheres, respectively.

is still evident in the recent SH data at 25 km and above. This is probably a residual effect of the increased instrument tracking jitter that is not completely removed from the data. We continue to monitor instrument degradation, primarily by studying noise levels in the elevation potentiometer data and the retrieval variability, as shown in Figure 2. If these increase significantly in the future it might be necessary to increase the level of smoothing in the radiance data.

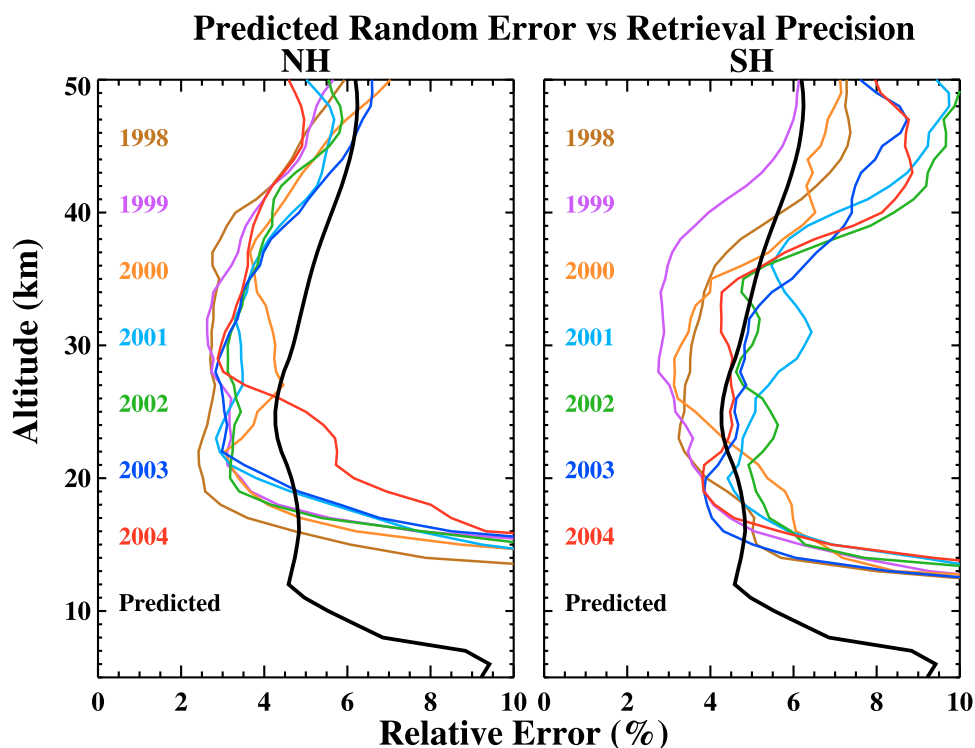
[13] A comprehensive retrieval characterization and error analysis for the version 3 water vapor retrievals was presented by Lumpe *et al.* [2002a], and this analysis has been repeated for the version 4 algorithm. The water vapor retrieval altitude range is 5 km (or cloud top) to 50 km; however, in this paper, we restrict our attention primarily to altitudes  $>10$  km. Figure 3 shows the estimated water vapor retrieval total random errors ( $1-\sigma$ ) and the full-width-half-maximum (FWHM) of the retrieval averaging kernels, which we use as a measure of the retrieval vertical resolution. The vertical resolution of the retrievals is about 1 km in the lower stratosphere, and begins to degrade with increasing altitude above 30 km to about 3 km at 40 km. The resolution analysis includes the full, realistic instrument model simulation and the increased smoothing added in version 4. The water vapor precision is generally  $<5\%$  ( $1-\sigma$ ) from 12 to 35 km, and better than 7% down to 14 km. This represents the total random error predicted from the formal error analysis, and is dominated by measurement noise and random altitude registration uncertainties. Two other important, but sporadic, sources of error in the water vapor retrievals arise from incomplete removal of sunspot artifacts and aerosol feedback errors. These are not included in the random error profile shown in Figure 3, but are discussed in detail below.

[14] It is instructive to compare the prediction of the formal error analysis with the actual version 4 retrieval

variability, which should provide a reasonable upper limit to the true retrieval precision during geophysically quiescent times. Measurement variability is defined for this purpose by calculating the standard deviation from the mean of the retrievals in one-week time bins (the same quantity plotted in Figure 2). Only summertime data were used in each hemisphere, and to be consistent with the calculated precision, all data flagged for sunspot and aerosol artifacts were removed (see discussion below). Figure 4 compares the minimum retrieval standard deviation obtained in all POAM years with the calculated precision. The comparison is quite



**Figure 3.** Estimated total random error and vertical resolution of the POAM version 4 water vapor retrievals. These results are obtained from the analysis described by Lumpe *et al.* [2002a], updated to the version 4 algorithm.



**Figure 4.** Comparison of calculated random error (black curve) to retrieval precision determined from the POAM version 4 water vapor data (colored curves). The random error profile is identical to the dashed curve in Figure 3. For each year the retrieval precision corresponds to the minimum summertime standard deviation calculated from one-week time bins. Left and right panels correspond to Northern and Southern Hemispheres, respectively.

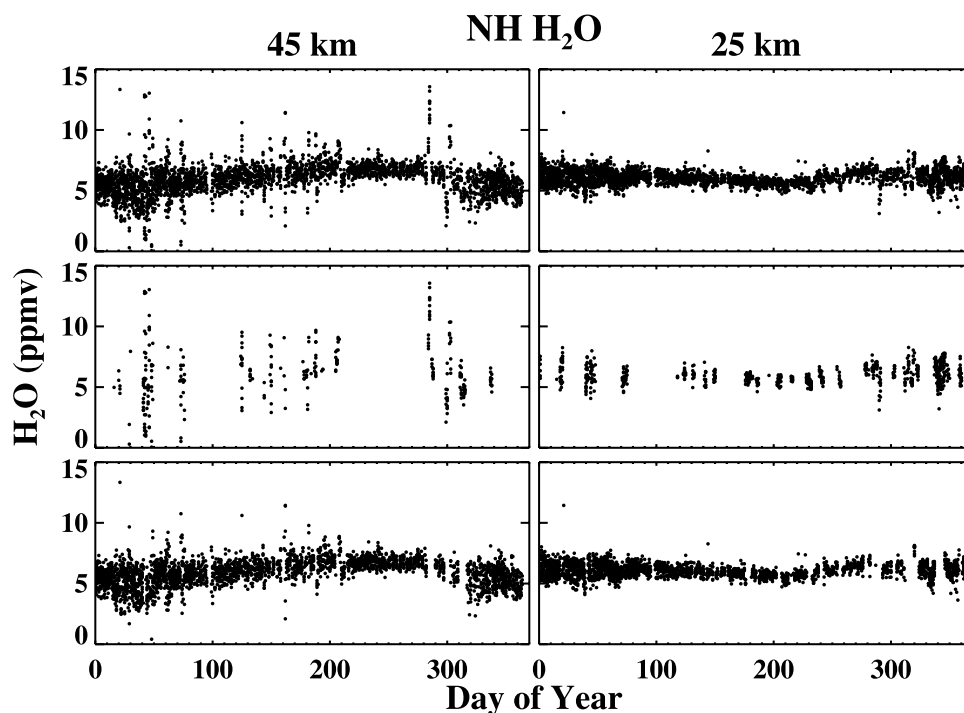
good in both hemispheres above approximately 15 km. At lower altitudes the measurement variability increases sharply, reflecting real atmospheric variability near the tropopause region that we do not expect to be reproduced by the error analysis. This increase in variability occurs at higher altitudes in the NH because the POAM summertime measurements there occur at lower latitudes than in the SH, and therefore the tropopause is higher. Above 15 km the NH measurement precision is actually better than the predicted error throughout most of the stratosphere whereas the SH shows somewhat higher variability, which on average is well represented by the calculated errors. We believe this likely results from larger actual pointing errors in the SH measurements than in the NH measurements. The random pointing error assumed in the error analysis calculations is thus probably more consistent with what actually occurs in the SH, but overestimates the NH errors. One notable exception to the agreement in the SH is the upper stratosphere, where the variability is significantly larger than the predicted random error. This could be a consequence of the observed hemispheric asymmetries in temperature and wave activity in the summertime mesosphere and upper stratosphere [Siskind *et al.*, 2003]. Another interesting hemispheric difference apparent in these results is the greater interannual variability of the retrieval precision in the SH compared to the NH above 15 km.

[15] Water vapor products from two previous POAM retrieval versions have been presented in the literature. Version 2 data was used in the Antarctic dehydration study

by Nedoluha *et al.* [2000] and in the SPARC upper tropospheric and stratospheric water vapor assessment report [Kley *et al.*, 2000]. The Arctic/Antarctic dehydration studies of Nedoluha *et al.* [2002a] and the UT/LS validation study of Nedoluha *et al.* [2002b] both used the version 3 data. Since this study represents the first comprehensive validation of the POAM stratospheric water vapor we will not present a detailed discussion of the evolution of the water vapor data set. The results presented by Nedoluha *et al.* [2000, 2002a] describe the temporal evolution and morphology of the polar stratospheric dehydration observed by POAM. Differences between version 4 water vapor data and previous versions do not affect the conclusions reached in those studies. The SPARC report [Kley *et al.*, 2000] used HALOE as a transfer standard in evaluating POAM and other stratospheric water vapor data sets, and those conclusions are superseded by the newer POAM/HALOE comparisons presented in this study. Finally, as discussed previously, the POAM UT/LS water vapor validation presented by Nedoluha *et al.* [2002b] is being extended using version 4 data and will be published separately.

## 2.1. Sunspot Errors and Screening

[16] As discussed by Lumpe *et al.* [2002a], sunspot errors in the POAM retrievals are caused mainly by small pointing inaccuracies, which result in uncertainties in the tracking position on the solar disk during a scan. This in turn creates a normalization error when dividing the measured radiances by the unattenuated radiance to calculate transmission. This



**Figure 5.** This figure illustrates the sunspot retrieval screen. The top panel shows all POAM water vapor retrievals obtained in 1999 in the Northern Hemisphere at 45 km and 25 km (left and right panels, respectively). The middle panel shows the points that are flagged because of potential sunspot artifact contamination. The lower panel shows the resultant screened data set.

error can be particularly significant in the presence of sunspots because of the resultant large inhomogeneity in the radiance across the solar disk. POAM performs a Sun scan above the atmosphere during each occultation event, which provides the normalization factors required for calculating transmission and allows the presence of sunspots to be detected. For each sunspot detected a sunspot index value (defined as the fractional decrease in radiance caused by the sunspot) is computed as a function of angular position on the solar disk. In the retrieval process these sunspot index values are then mapped onto the retrieval altitude grid. The sunspot index is a function of altitude because the instrument tracks the Sun's center of brightness, which moves up on the disk of the Sun as the instrument scans lower in the atmosphere. Therefore sunspots near the Sun center affect retrievals at high altitudes, while sunspots occurring at high solar latitudes affect retrievals at lower altitudes.

[17] Lumpe *et al.* [2002a] developed an empirical analysis to characterize the retrieval errors resulting from sunspot artifacts, as a function of the sunspot index. This analysis has been repeated for the version 4 retrievals and the resulting sunspot errors are included in the error bars archived in the version 4 data set. For water vapor, these errors increase as a function of altitude and can exceed 100% at 40 km for the largest sunspots. The difficulty in characterizing the sunspot retrieval errors statistically is that their distribution tends to be non-Gaussian. That is, errors at the several-sigma level tend to occur much more often than expected from a Gaussian distribution. Therefore, in addition to including the sunspot component in the total error

bars, we flag points with potentially large sunspot contamination, based on the statistical error value, and omit the flagged points in validation or scientific studies. The value of the sunspot error used to flag contaminated data was determined empirically using the same technique described for NO<sub>2</sub> by Randall *et al.* [2002]. This choice is necessarily somewhat arbitrary and represents our best estimate of a trade off between avoiding false negatives (contaminated data not flagged) while minimizing false positives (flagged data which is not contaminated). The data quality flag archived with the version 4 water vapor data alerts the reader to all points that are potentially contaminated by sunspots.

[18] The effectiveness of the sunspot data flagging procedure is illustrated in Figure 5. The top panels of this figure show all of the POAM NH water vapor retrievals obtained in 1999 at 45 and 25 km (left and right panels, respectively). The middle panels show those points with a sunspot error exceeding the screen value (the flagged points), and the lower panels show the remaining unscreened points. First, it is clear that the sunspot problem is much worse at 45 than 25 km. In fact, at 25 km the benefit of the data screening is very modest. However, at 45 km the sunspot artifacts tend to be larger, and the screening is more important. This is because the magnitude of the sunspot error decreases with increasing optical depth. Figure 5 shows that the screening procedure does a good job of removing obvious outliers. However, many of the flagged points are not evidently contaminated, and there are a few points (less than 2% out of more than 20,000 retrievals examined at 45 km) that appear contaminated but are not flagged. On average, over



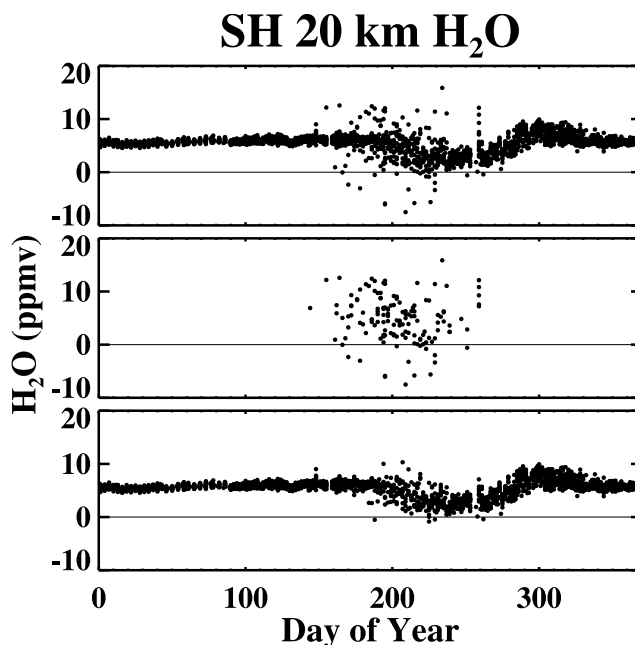
the entire data set, approximately 15 to 20% of all POAM water vapor data above 15 km is flagged for potential sunspot contamination. We recommend that users of the POAM data utilize the flag values provided but, because of the small possibility of false negatives, users should be cautious with any data that has an enhanced total error value. Above 25 km and outside the winter vortex, such an enhancement can only be caused by a nonzero sunspot error value.

## 2.2. Aerosol Errors and Screening

[19] Water vapor retrievals made in the presence of high aerosol loading are also subject to enhanced errors because of inaccuracies in the characterization and removal of the aerosol extinction. Aerosol contamination is generally only a problem in the presence of PSCs, and is therefore much more important in the SH than in the NH because of the much larger number of PSCs observed in the South. As with sunspots, an empirical error analysis was developed for the aerosol contamination (based on the measured aerosol extinction at 1  $\mu\text{m}$ ), and this error component is included in the  $\text{H}_2\text{O}$  total error bar. Analogous to the sunspot discussion above, we also use this error source to flag data points with potentially large aerosol contamination.

[20] The effectiveness of the aerosol screening is illustrated in Figure 6, which concentrates on the 2003 POAM retrievals at 20 km in the SH. In the top panel all the data are plotted, and clearly show an increase in the retrieval scatter during the austral winter. In the middle panel of Figure 6 we plot those points that have been flagged for possible aerosol contamination. The points affected are mainly those in which a PSC was present, and almost all of the outliers are contained in this subset. The bottom panel shows the resulting water vapor time series with the flagged data removed. As with the sunspot screen, the noise in the screened data set is considerably reduced while the essential geophysical features (e.g., the rapid dehydration in July) are retained. Nevertheless, there is still some residual increase in retrieval variability in the core of the winter vortex, probably due to a combination of enhanced natural variability due to transport, and aerosol extinction values which are elevated compared to normal background levels, but still below the threshold value. Thus we conclude that even after screening for aerosol contamination, the presence of enhanced aerosol extinction in the SH winter lower stratosphere is a source of additional noise in the individual POAM water vapor retrievals.

[21] The fraction of POAM water vapor retrievals flagged for potential aerosol contamination in the version 4 data set peaks in the SH at approximately 8% between 15 and 25 km. In the NH the problem is essentially nonexistent because far fewer PSCs are observed there. The exception is the 1999/00 NH winter (the SOLVE/THESEO 2000 winter), which was very cold and resulted in many optically thick PCS being observed by POAM [Bevilacqua *et al.*, 2002]. This is discussed in more detail in section 4, where comparisons with aircraft and balloon measurements obtained during the SOLVE/THESEO campaign are presented. A small fraction of points are also screened in the NH below 13 km as a result of high cirrus contamination. It is important to note that the 8% value quoted for the fraction of aerosol-screened retrievals in the SH corresponds to the



**Figure 6.** Same as Figure 5, but for the aerosol screening. The top panel shows all POAM retrievals obtained at 20 km in the Southern Hemisphere in 2003. The middle panel shows all points flagged because of potential aerosol contamination, and the bottom panel shows the resultant screened data set.

entire data set, which is somewhat misleading because all this screening (aside from cirrus contamination) occurs during the winter season. In some years in the SH, during the winter months of July through September as much as 50% of the data are screened for potential aerosol contamination at 20 km.

## 3. Comparisons With Satellite Measurements

[22] The POAM measurement coverage overlaps several times a year in both hemispheres with the HALOE and SAGE II solar occultation measurements, primarily in spring and summer. Over the 7-year POAM mission this overlap has produced an extensive set of coincident measurements that are available for validation purposes. The HALOE and SAGE II water vapor retrievals have similar vertical resolution and precision to the POAM data and thus are uniquely suitable for POAM validation. The HALOE water vapor data set in particular is mature and has been extensively validated. The SAGE II water vapor data set is less mature, as discussed below, but given the number of coincident measurements with POAM it is still valuable.

[23] The HALOE instrument has been measuring middle atmospheric water vapor from the UARS satellite since 1991. HALOE is an IR solar occultation instrument that measures  $\text{H}_2\text{O}$  at 6.6  $\mu\text{m}$  from 15 to 85 km, with approximately 2.3 km vertical resolution [Russell *et al.*, 1993]. For this study we use the version 19 HALOE retrievals. This data set has been compared extensively with in situ and remote measurements made from balloon and aircraft, and with other satellite-based measurements. These comparisons are summarized by Harries *et al.* [1996] and the SPARC

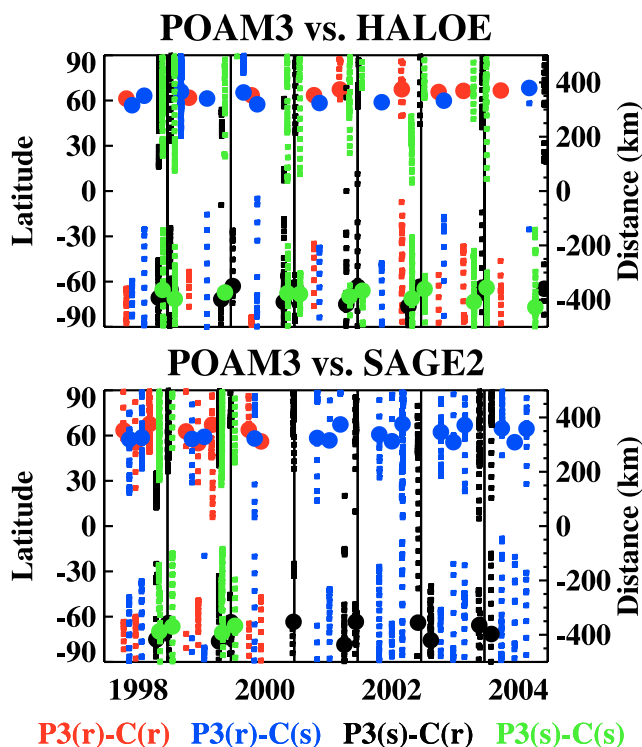
water vapor report [Kley *et al.*, 2000]. The situation is not completely straightforward, but the preponderance of the comparisons suggest that the HALOE version 19 water vapor retrievals are biased low by about 5% in the stratosphere, with no clear altitude dependence. Also, several comparisons suggest that the HALOE version 19 retrievals do not resolve the hygropause as well as would be expected based the vertical resolution estimates. This sometimes leads to large apparent HALOE negative biases compared with other measurements just below the hygropause, but does not affect the HALOE data above 20 km. This issue is discussed extensively by Kley *et al.* [2000].

[24] The SAGE II instrument has been in operation on the ERBS satellite since 1984. Like POAM, it measures solar extinction in the UV to near-IR spectral region and thus utilizes the same 940-nm absorption bands to retrieve water vapor. Unlike POAM, which has differential water vapor channels, SAGE II measures water vapor with a single channel at  $\sim 945$  nm. As mentioned previously, for many years the SAGE II water vapor retrievals suffered from known deficiencies and were not suitable for use as a validation source. Recently, however, a new version 6.2 retrieval has been released which appears to fix many of these problems, and this version has been used in the following study. As discussed by Thomason *et al.* [2004] it was determined that a previously uncharacterized post-launch shift in the spectral response of the SAGE II water vapor channel was responsible for most of the systematic error in the water vapor retrievals.

[25] The version 6.2 retrievals incorporated a modified spectral band pass for this channel, with a position and width determined empirically by minimizing the mean bias between SAGE II retrievals and a 10-year (1992–2002) climatology constructed from the HALOE version 19 water vapor. The details of this procedure are described by Thomason *et al.* [2004]. For our purposes, the relevant point is that the SAGE II version 6.2 water vapor cannot be used independently of the HALOE version 19 data to evaluate the absolute value of the POAM retrievals. However, it is still useful to compare relative variations between POAM and SAGE II. Taha *et al.* [2004] have presented a validation analysis of the SAGE II version 6.2 water vapor, concluding that between 15 and 40 km SAGE II agrees to within 10% with HALOE, POAM (version 3), ILAS and the CMDL frost point hygrometer, and within 15–20% with MkIV. Above 40 km the SAGE II data become very noisy and exhibit an increasing positive bias.

### 3.1. Method

[26] Statistical analyses of comparisons between POAM retrievals and measurements from HALOE and SAGE II were performed following procedures described previously [e.g., Randall *et al.*, 2003]. The criteria used to define a coincident measurement were 500 km spatial separation and a time difference of 12 hours. Coincidences were also required to occur within 5 degrees in equivalent latitude. Equivalent latitude is defined as that latitude that would enclose the same area between it and the pole as a given potential vorticity contour. In the wintertime comparisons this additional requirement insures that coincident measurements occur in similar vortex conditions, and thus minimizes differences due to real atmospheric variability.



**Figure 7.** Coincidence latitudes and separation distances for all measurement pairs between POAM and (top) HALOE or (bottom) SAGE II. The “r” and “s” notations refer to sunrise and sunset spacecraft occultations, respectively, and “C” denotes the correlative instrument (HALOE or SAGE II). Small symbols denote separation distance (right axis), and larger circles denote average latitudes in each coincidence time period (see Tables 1 and 2). Distances are negative when POAM was at a lower latitude than the correlative instrument.

Comparisons were performed separately for satellite sunrise (“r”) and sunset (“s”) occultations, since occultation instruments are known to often suffer from systematic errors related to the satellite occultation type. Note that HALOE and SAGE II spacecraft occultations occur as both sunrises and sunsets in each hemisphere. Figure 7 and Tables 1 and 2 summarize the coincidence statistics for the satellite-satellite comparisons.

[27] The POAM-HALOE and POAM-SAGE II coincidences occur primarily during the spring and summer months in each hemisphere. Water vapor exhibits significant seasonal variability at the POAM measurement latitudes throughout the stratosphere, as shown in Figure 8 [see also Nedoluha *et al.*, 2002a]. During the summer, the water vapor profile generally increases gradually with increasing altitude in the stratosphere, primarily because of increasing methane oxidation. In the fall and winter, the higher-mixing-ratio air in the upper stratosphere descends to the middle stratosphere, while low-mixing-ratio air from the mesosphere (not shown) descends to the upper stratosphere, resulting in a broad mixing ratio maximum centered near 30 km. During the springtime, lower mixing ratios in the upper stratosphere are replenished with air moving poleward. A local minimum (hygropause) is reached in the lower stratosphere near 15 km that, in the NH, is of lowest



**Table 1.** Statistics of POAM/HALOE Coincidences<sup>a</sup>

Dates	Number	Pocc/Hocc	Latitude	Distance, km	$\Delta$ Time (Min/Max), hours
9–10 May 1998	7	SR/SR	59.7	410	−7.1 (−7.6/−6.8)
11–12 June 1998	8	SR/SS	55.4	383	−0.2 (−0.8/0.2)
20–21 Aug. 1998	10	SR/SS	61.9	337	−0.2 (−0.7/0.0)
9–14 Nov. 1998	50	SS/SR	−70.6	322	−2.9 (−4.8/−1.0)
30 Nov. to 8 Dec. 1998	53	SS/SS	−65.1	327	−0.7 (−3.2/1.2)
15–21 Jan. 1999	33	SS/SR	−64.5	346	−4.8 (−5.9/−3.5)
9–14 Feb. 1999	35	SS/SS	−71.2	344	−0.2 (−1.7/1.2)
17–24 March 1999	59	SR/SS	67.0	358	−0.0 (−0.7/0.6)
5 May 1999	9	SR/SR	60.6	338	−7.8 (−8.2/−7.4)
16–17 Aug. 1999	7	SR/SS	60.6	296	−0.2 (−0.5/0.3)
5–10 Nov. 1999	14	SS/SR	−71.9	308	−3.0 (−4.9/−1.2)
25–30 Nov. 1999	12	SS/SS	−66.3	352	−0.9 (−3.1/0.3)
11–12 Jan. 2000	9	SS/SR	−63.5	246	−5.0 (−5.8/−4.4)
13–16 March 2000	14	SR/SS	67.3	466	−0.0 (−0.3/0.2)
30 April 2000	5	SR/SR	61.4	475	−7.8 (−8.1/−7.6)
31 May to 2 June 2000	8	SR/SS	56.5	308	−0.3 (−0.6/0.1)
31 Oct. to 3 Nov. 2000	29	SS/SR	−73.7	280	−3.6 (−5.0/−1.4)
20–26 Nov. 2000	33	SS/SS	−67.4	301	−0.1 (−2.6/1.3)
4–9 Jan. 2001	23	SS/SR	−63.1	397	−4.6 (−5.7/−3.4)
31 Jan. to 4 Feb. 2001	32	SS/SS	−68.1	337	−0.3 (−1.5/0.8)
24 April 2001	7	SR/SR	62.4	294	−8.1 (−8.4/−7.9)
26 May 2001	7	SR/SS	57.2	304	−0.3 (−0.6/−0.0)
17–23 Sept. 2001	6	SR/SR	69.0	421	−4.3 (−11.1/1.1)
26–28 Oct. 2001	18	SS/SR	−75.6	370	−4.2 (−5.5/−2.8)
13–17 Nov. 2001	20	SS/SS	−69.3	347	−0.5 (−2.6/0.8)
31 Dec. 2001 to 6 Jan. 2002	20	SS/SR	−62.6	314	−4.2 (−5.5/−1.6)
26–30 Jan. 2002	10	SS/SS	−66.5	416	−0.1 (−1.4/0.9)
19–20 May 2002	9	SR/SS	57.9	356	−0.0 (−0.6/0.2)
9–17 Sept. 2002	28	SR/SR	67.3	285	−10. (−11./−9.6)
20–21 Oct. 2002	8	SS/SR	−78.0	437	−4.8 (−5.3/−4.2)
6–1 Nov. 2002	46	SS/SS	−71.3	250	−0.0 (−1.9/1.4)
25–29 Dec. 2002	8	SS/SR	−62.6	385	−4.1 (−4.8/−1.9)
20–23 Jan. 2003	25	SS/SS	−64.9	379	−0.0 (−1.0/0.7)
12 April 2003	5	SR/SR	64.2	393	−9.0 (−9.3/−8.8)
12 May 2003	8	SR/SS	59.3	217	−0.2 (−0.7/0.0)
2–4 Sept. 2003	13	SR/SR	65.1	330	−9.5 (−10./−9.1)
30 Oct. to 3 Nov. 2003	29	SS/SS	−74.1	370	0.4 (−0.9/1.6)
17–21 Dec. 2003	29	SS/SR	−63.2	332	−3.4 (−4.9/−1.8)
8–16 Jan. 2004	28	SS/SS	−63.8	337	−0.8 (−3.8/1.0)
3–5 April 2004	10	SR/SR	65.3	373	−10. (−10./−9.5)
14–16 Sept. 2004	3	SR/SS	68.3	285	0.0 (−0.4/0.6)
19–21 Oct. 2004	17	SS/SS	−78.0	341	0.9 (−0.6/2.0)
8–14 Dec. 2004	20	SS/SR	−64.1	334	−3.2 (−5.0/−1.4)

<sup>a</sup>Columns show the inclusive dates of different coincidence time periods, number of coincident pairs, spacecraft occultation type for POAM (Pocc) and HALOE (Hocc) (each coincidence time period includes only a single type of occultation per instrument), average latitude of the coincidences, average separation distance, and average time separation as well as the minimum and maximum time separation during that time period. Differences are calculated as POAM minus HALOE (e.g., negative  $\Delta$  time means that the POAM measurement preceded the HALOE measurement). SS, sunset; SR, sunrise.

magnitude during the spring and summer. This minimum is due to air that has recently passed through the relatively dry tropical tropopause. The hygropause is usually several kilometers above the tropopause, since high-mixing-ratio air is transported isentropically from the tropics to high latitudes during summer and fall without passing through the tropical tropopause [Nedoluha *et al.*, 2002b; Prados *et al.*, 2003]. In the SH, dehydration due to the descent of ice particles in PSCs causes a minimum to occur over a broad range of altitudes during the late winter to spring [e.g., Nedoluha *et al.*, 2000]. As seen from Figure 8, NH coincidences tend to occur near times of minimum mixing ratios near 15 km and just before times of maximum mixing ratios in the upper stratosphere and near the tropopause. Coincidences in the SH tend to occur shortly after the main period of lower stratosphere dehydration.

[28] Both relative  $[(P - C)/\text{Avg}, \%]$  and absolute  $[(P - C), \text{ppmv}]$  differences were calculated at each altitude for

each coincidence pair. Here P refers to the POAM measurement, C refers to the correlative measurement, and Avg refers to the average of the two. Before calculating the differences, correlative data were interpolated to a 1-km vertical grid to match the vertical scale of the POAM data. For any given coincidence pair, P and/or C could have had missing data due to failed retrievals; these are generally found at the high and/or low ends of the altitude range. In addition, data were flagged as “bad” and thus not included in the analysis, if the stated errors on the measurements exceeded 100%. POAM data were further screened for sunspot or aerosol artifacts as described in section 2. HALOE data are screened for cloud artifacts before being posted to the distribution web site (<http://haloedata.larc.nasa.gov/home/index.php>), so no cloud-contaminated data were included in the comparisons. SAGE II measurements were screened according to Taha *et al.* [2004], whereby all measurements for which the

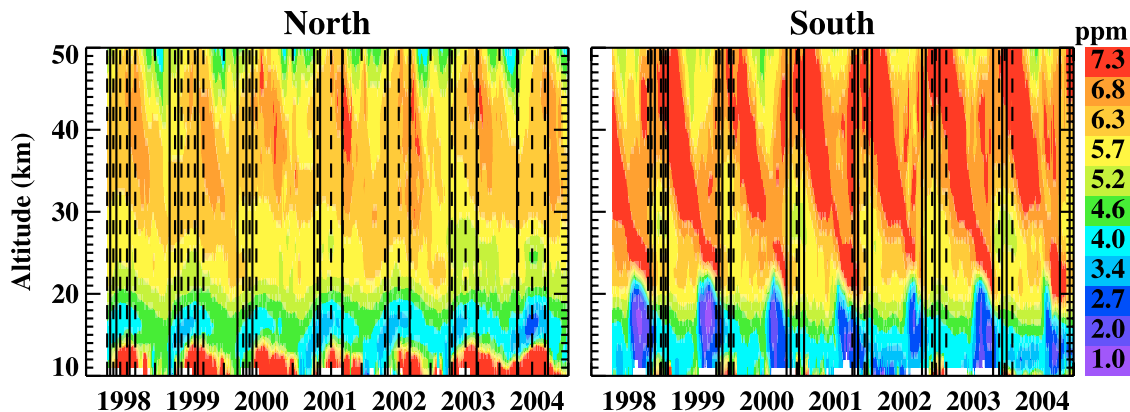
**Table 2.** Same as Table 1, but for POAM/SAGE II Coincidences

Dates	Number	Pocc/Hocc	Latitude	Distance, km	$\Delta$ Time (Min/Max), hours
22–24 April 1998	8	SR/SR	62.5	393	–8.5 (–9.3/–7.8)
25–27 May 1998	19	SR/SS	57.2	312	–0.1 (–0.8/0.5)
30 June to 2 July 1998	11	SR/SR	54.7	350	–6.9 (–7.6/–6.5)
5–7 Aug. 1998	23	SR/SS	58.6	326	–0.0 (–0.7/0.5)
16–23 Sept. 1998	24	SR/SR	69.0	374	–6.4 (–11./11.)
26 Oct. to 1 Nov. 1998	49	SS/SR	–74.9	180	–2.0 (–5.1/0.3)
12–20 Nov. 1998	69	SS/SS	–69.7	302	–1.6 (–4.2/1.3)
1–9 Jan. 1999	43	SS/SR	–63.1	365	–1.9 (–4.6/0.7)
29–31 Jan. 1999	22	SS/SS	–67.0	304	0.0 (–0.9/0.9)
17–19 April 1999	5	SR/SR	63.4	412	–8.6 (–9.1/–8.3)
21–23 May 1999	8	SR/SS	57.7	321	–0.3 (–0.7/0.1)
26–28 June 1999	9	SR/SR	54.5	317	–6.9 (–7.5/–6.3)
1–2 Aug. 1999	5	SR/SS	57.4	418	–0.2 (–0.5/–0.0)
12–18 Sept. 1999	20	SR/SR	67.6	268	–10. (–11./–10.)
23–27 Oct. 1999	55	SS/SR	–76.3	330	–2.9 (–5.5/–0.5)
8–16 Nov. 1999	71	SS/SS	–70.6	306	–0.9 (–4.2/1.3)
27 Dec. 1999 to 5 Jan. 2000	26	SS/SR	–62.7	402	–2.3 (–5.0/0.4)
24–27 Jan. 2000	15	SS/SS	–65.8	338	–0.3 (–1.4/0.9)
12–15 April 2000	9	SR/SR	64.0	424	–9.3 (–9.9/–8.9)
15–19 May 2000	15	SR/SS	58.5	319	–0.0 (–0.8/0.5)
22–23 June 2000	7	SR/SR	54.8	359	–6.5 (–7.1/–6.1)
21–30 Dec. 2000	21	SS/SR	–62.7	366	–1.7 (–4.9/0.0)
12–14 May 2001	13	SR/SS	59.2	344	–0.0 (–0.6/0.4)
22–24 July 2001	15	SR/SS	56.2	337	0.0 (–0.5/0.5)
23 Sept. 2001	2	SR/SS	69.5	489	–0.0 (–0.1/0.0)
15–18 Oct. 2001	30	SS/SR	–79.5	357	–4.7 (–6.1/–3.2)
16–25 Dec. 2001	40	SS/SR	–62.9	354	–2.3 (–5.0/0.1)
4–7 May 2002	23	SR/SS	60.2	331	–0.1 (–0.7/0.5)
16–19 July 2002	9	SR/SS	55.3	409	–0.0 (–0.6/0.5)
12–1 Sept. 2002	63	SR/SS	68.1	296	0.0 (–0.7/0.7)
9–17 Dec. 2002	26	SS/SR	–63.4	254	–1.6 (–5.0/0.4)
24–28 Feb. 2003	28	SS/SR	–76.6	359	–6.9 (–8.1/–5.8)
23–26 April 2003	16	SR/SS	62.2	336	–0.0 (–0.7/0.6)
5–7 July 2003	12	SR/SS	54.8	386	–0.1 (–0.6/0.4)
8–12 Sept. 2003	29	SR/SS	66.8	305	–0.0 (–0.7/0.6)
29 Nov. to 7 Dec. 2003	43	SS/SR	–65.5	313	–1.9 (–4.8/0.9)
8–14 Feb. 2004	38	SS/SR	–70.6	351	–5.2 (–6.7/–3.6)
7–13 April 2004	33	SR/SS	64.5	360	–0.0 (–0.7/0.6)
2–24 June 2004	13	SR/SS	54.7	297	–0.1 (–0.7/0.4)
29–31 Aug. 2004	18	SR/SS	64.5	312	–0.0 (–0.7/0.6)

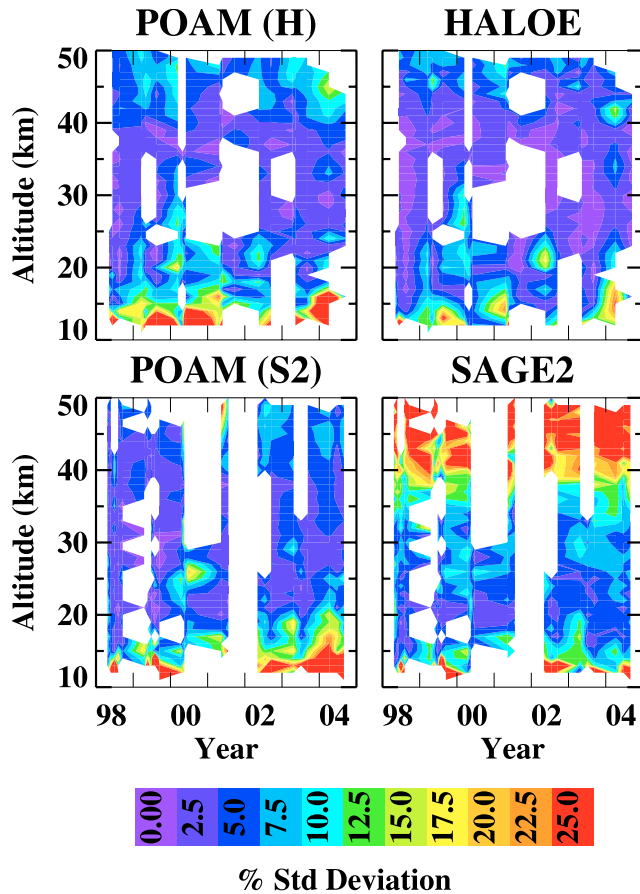
1020-nm aerosol extinction was greater than  $2 \times 10^{-4} \text{ km}^{-1}$  were removed. The error analysis and screening process results in varying numbers of coincident comparisons at each altitude.

### 3.2. Satellite Comparison Results

[29] Figures 9 and 10 compare the variability measured by each of the instruments during the coincident time periods. As discussed in section 2, this variability is determined by a combination of finite retrieval precision



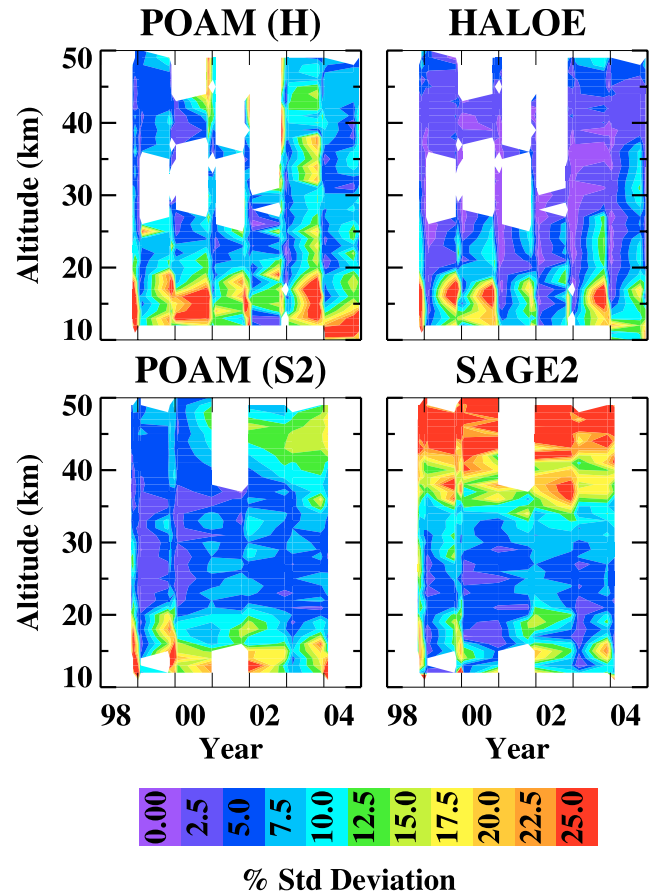
**Figure 8.** Ten-day running averages of POAM water vapor mixing ratios in the (left) NH and (right) SH. Solid (dashed) vertical black lines denote times of HALOE (SAGE II) coincidences as in Tables 1 and 2.



**Figure 9.** Time series of the variability in mixing ratio profiles measured by (left) POAM and (right) HALOE or SAGE II during each of the NH coincidence time periods listed in Tables 1 and 2. The variable plotted is the 1- $\sigma$  standard deviation (percent, relative to the average mixing ratio) of the distribution of mixing ratios measured during each coincident time period.

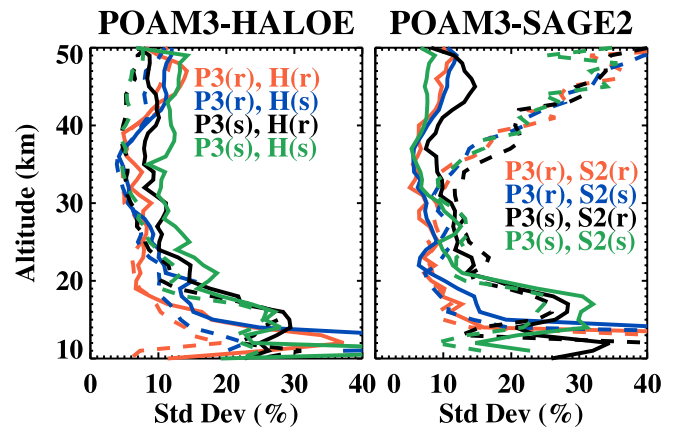
and real geophysical variability. If the instruments exhibit similar precision, the variability measured by each should be similar. The standard deviation of all of the measurements from a given instrument during each coincidence time period was calculated as a function of altitude, and is shown as time series in Figure 9 (NH) and 10 (SH). Data gaps, indicated by white portions of the contour, are caused by either a lack of coincidences or data from one or both instruments being screened as described above. At most altitudes, the variability is roughly constant in time, within the limits of the overall spread in any given year, suggesting that there are no significant trends in the instrument precision. One exception to this is the POAM data above 40 km in the SH SAGE II comparisons, where the variability appears to increase in time. Variability is generally largest at the high and low ends of the altitude range depicted. POAM variability is similar to HALOE in the NH, but significantly larger in the SH throughout most of the altitude range. In contrast, compared to SAGE II, POAM variability is similar below about 35 km in both hemispheres, but significantly lower above 35 km.

[30] Profiles of the variability averaged over all coincidences with both HALOE and SAGE II are shown in

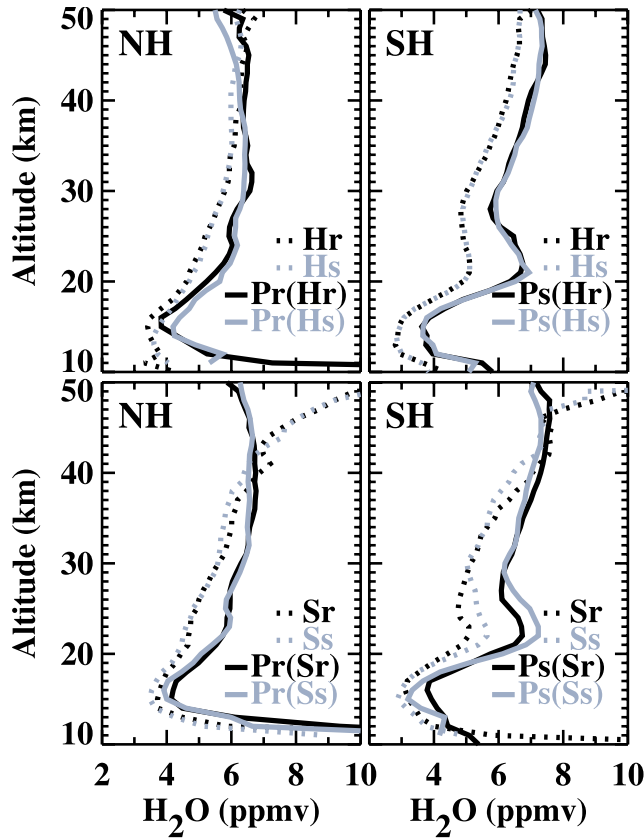


**Figure 10.** Same as Figure 9, but for the SH.

Figure 11. Above about 20 km, standard deviations are around 10% as measured by POAM and HALOE, with up to 5% larger variability in the POAM measurements. Since the coincidence time periods include times when the polar



**Figure 11.** Profiles of the standard deviation in water vapor mixing ratios (percent, relative to the average mixing ratio) of the distribution of coincident measurements over the entire time period (1998–2004) for (left) HALOE and (right) SAGE II coincidences. Solid lines give the standard deviation measured by POAM; dashed lines give the standard deviation measured by HALOE or SAGE II. Colors signify the various combinations of spacecraft occultations (see also Figure 7).



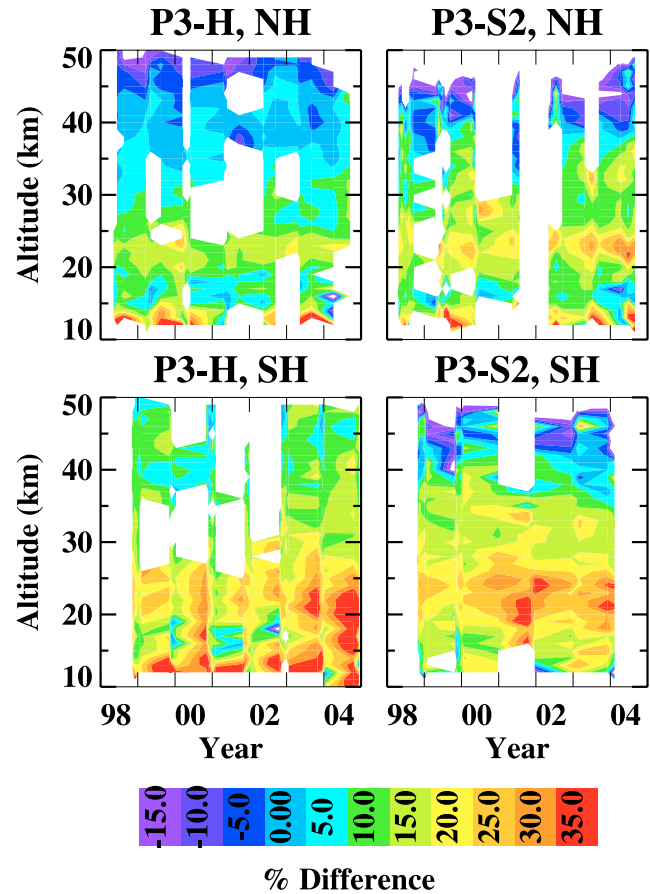
**Figure 12.** Average water vapor profiles for all coincidences in the (left) NH and (right) SH between POAM III and HALOE (top panels) or SAGE II (bottom panels). Solid lines represent POAM III, and dashed lines represent either HALOE or SAGE II. The “r” and “s” designations in the labels refer to spacecraft sunrise and sunset occultations, respectively, with “P” denoting POAM, “H” denoting HALOE, and “S” denoting SAGE II.

vortex was present (e.g., March in the NH; October and November in the SH), these standard deviations include significant geophysical variability in addition to random measurement error. As shown in Figure 4, the measured upper limit to the precision of POAM water vapor measurements is on the order of 3–6% in the middle stratosphere, suggesting that roughly half of the variability shown in Figure 11 is due to geophysical variations. The hemispheric asymmetry shown in Figure 4, in which POAM precision is slightly worse for sunset measurements, is also apparent in Figure 11 below 40 km. Variability in the SAGE II and POAM measurements is comparable below 30 km, but above this altitude variability in the SAGE II measurements increases significantly. This is broadly consistent with *Taha et al.* [2004], who showed increasing noise in the SAGE II data above 40 km. All three instruments show increasing variability below 20 km. This is primarily because of increasing geophysical variability at these altitudes, as steep gradients in water vapor are encountered near the hygropause.

[31] Figure 12 compares the water vapor profiles measured by POAM III and either HALOE or SAGE II, averaged over all coincidences in each hemisphere. The

figure shows HALOE and SAGE II sunrise and sunset occultations separately, but note that the results are nearly indistinguishable, confirming the lack of a sunrise/sunset bias in these instruments. All of the comparisons indicate a POAM high bias with respect to the other occultation instruments throughout most of the altitude range; they also indicate a high bias in the SAGE II water vapor above 40 km [see *Taha et al.*, 2004]. In general, however, mixing ratios are fairly constant above 15 km in all measurements, but rapidly increase below 15 km when the hygropause is encountered. That the HALOE NH measurements do not show this feature can most likely be attributed to the lower vertical resolution in this instrument. The SH profiles all show a local maximum near 20–23 km. As indicated in Tables 1 and 2, many of the SH coincidences were acquired during November. Figure 8 shows that at this time, the water vapor mixing ratio maximum had descended to these altitudes, and water vapor had not yet been replenished in the upper stratosphere.

[32] Figure 13 shows the time series of average relative differences between POAM and the correlative measurements for each time period listed in Tables 1 and 2. In both hemispheres, POAM is on average higher than either HALOE or SAGE II at most altitudes in all coincidence time periods, with maximum differences below 15 km, and



**Figure 13.** Time series of the relative differences (percent) between POAM and (left) HALOE or (right) SAGE II for all coincidence time periods in the NH (top panels) and SH (bottom panels).

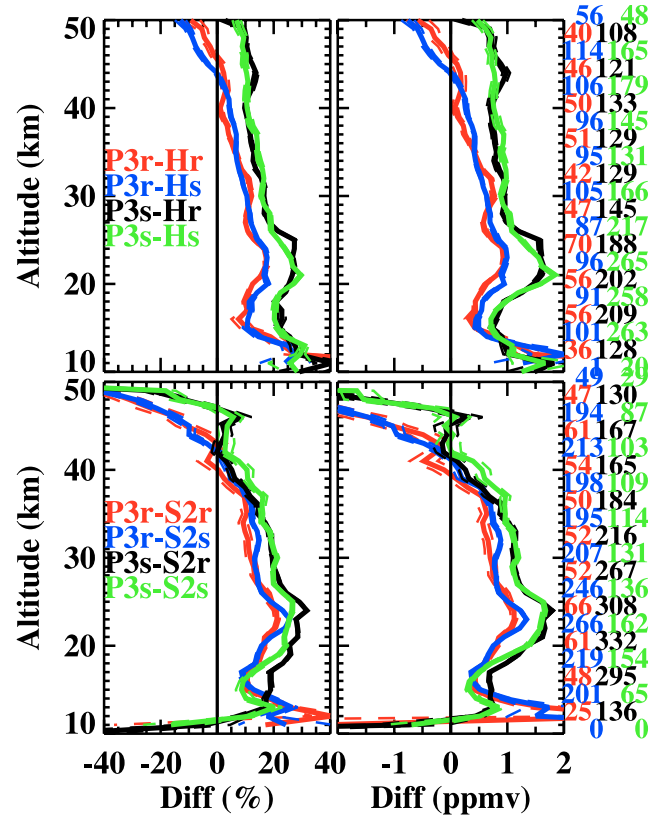


a relative maximum near 20–25 km. Differences are smaller in the NH, and often become negative (POAM smaller than the correlative instrument) above about 40 km. Given that the SAGE II spectral band pass has been calibrated by using the HALOE climatology, the consistency of much of the systematic difference between POAM and these two instruments is not surprising. Although results vary from time period to time period, there is no indication of a significant, systematic trend in the comparisons. A possible exception to this is the SH near 20 km, where differences with HALOE appear to increase in 2003 and 2004, but the comparison statistics are not adequate to quantify this. No such trend is apparent in the 2003 SAGE II differences, and unfortunately there are no SH POAM/SAGE II coincidences after February 2004 to compare to the HALOE results. Overall differences are discussed more quantitatively below.

[33] The differences shown in Figure 13 are summarized in Figure 14, in both relative (percent) and absolute (ppmv) units. There is a clear sunrise/sunset (NH/SH) bias in the POAM data, such that POAM sunrise measurements are about 5–10% lower than the sunset measurements. Below 40 km the POAM retrieved water vapor is higher than both HALOE and SAGE II in both hemispheres. Differences are near 0 around 40 km, and increase with decreasing altitude to maximum values of  $\sim 15\%$  (25%) in the altitude range from 20 to 25 km in NH (SH). These maximum relative differences correspond to absolute differences of about 1–1.5 ppmv. Differences are also large below 13 km, but measurement variability and lack of data precludes definitive conclusions regarding instrument biases. By 50 km, POAM is lower than HALOE by about 15% in the NH, and higher by 5–10% in the SH. *Taha et al.* [2004] showed that SAGE II water vapor was biased high above 40 km, differing by more than 30% with respect to coincident HALOE measurements at 50 km. This is consistent with the large negative POAM/SAGE II differences near 50 km that are evident in Figure 14.

#### 4. Comparisons With Aircraft and Balloon Data

[34] In this section we present comparisons of POAM water vapor with measurements made by a number of balloon- and aircraft-borne instruments using both in situ and remote sensing techniques. These coincident measurements were primarily obtained during the SOLVE/THESEO 2000 and SOLVE II Arctic campaigns. POAM was an operational component of both of these missions, which resulted in a number of comparison opportunities with other dedicated water vapor instruments. These include the following: the JPL Tunable Diode Laser (TDL) spectrometer and the Harvard Lyman- $\alpha$  hygrometer both aboard the NASA ER-2 aircraft, the balloon-borne JPL MkIV interferometer, the Forschungszentrum Jülich Fast In situ Stratospheric Lyman- $\alpha$  hygrometer (FISH), and the NOAA CMDL frost point hygrometer. For the most part, each of these measurements provided only a few POAM validation opportunities compared to the large number of satellite coincidences shown in section 3. However, these comparisons are important because the balloon and aircraft measurements generally have high accuracy and vertical resolution, particularly in the critical lower stratosphere. Also, given that SAGE II does not provide an independent constraint on



**Figure 14.** (left) Relative and (right) absolute differences between POAM and HALOE (top panels) or SAGE II (bottom panels) water vapor for all coincidences as defined in Figure 7. Colors denote the different combinations of spacecraft occultation event types, as in previous figures. Dashed lines, which often overlap the solid lines, denote the uncertainty in the mean differences. The uncertainty here is defined as  $1\text{-}\sigma$  standard deviation of the distribution of differences at each altitude, divided by the square root of the number of comparisons, which are given to the right of the figure.

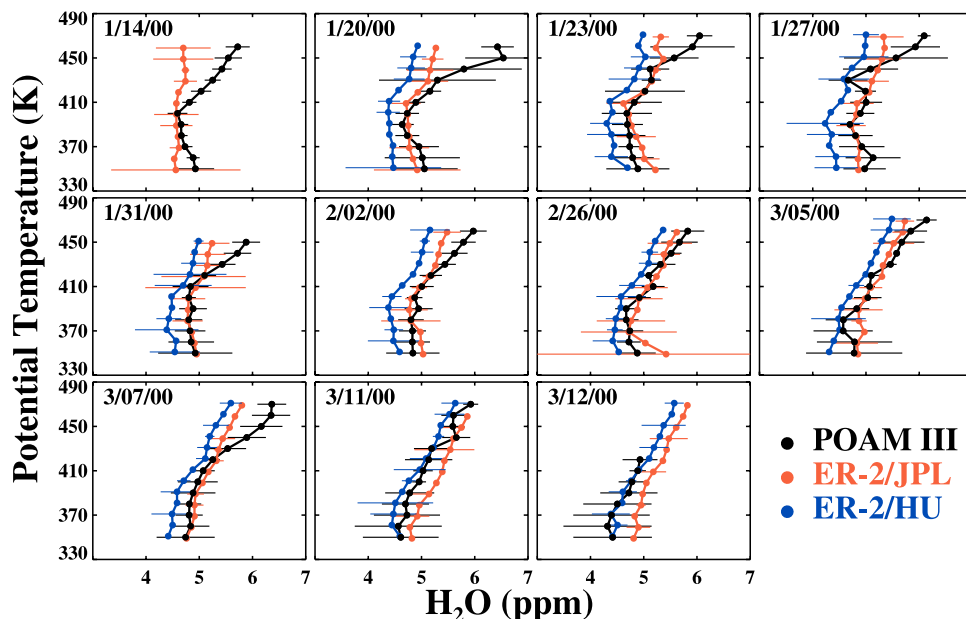
POAM retrieval biases, these comparisons are crucial in helping to interpret the observed POAM/HALOE differences. In the following sections we analyze the comparisons with each of these instruments separately, then summarize with some conclusions about the combined aircraft/balloon validation and its relation to the satellite comparisons. As with the satellite comparisons, the POAM data were screened by eliminating any points having a nonzero flag for sunspot or aerosol contamination.

##### 4.1. ER-2 Comparisons

[35] The Harvard Lyman- $\alpha$  hygrometer and the JPL TDL instrument were both included as part of the complement of instruments aboard the NASA ER-2 for the SOLVE/THESEO mission. Both of these instruments have been extensively validated, and have participated in several NASA aircraft campaigns. A brief summary of the expected precision and accuracy of the instruments is given below.

[36] A detailed description of the Harvard Lyman- $\alpha$  hygrometer is given by *Weinstock et al.* [1994]. From laboratory calibrations and in-flight calibration checks the

### POAM III and ER-2 Vortex-averaged H<sub>2</sub>O



**Figure 15.** Vortex average profiles obtained from POAM and the ER-2 Harvard Lyman- $\alpha$  hygrometer and the JPL TDL instruments on the indicated flight days during the SOLVE/THESEO 2000 mission. The ER-2 profiles were calculated using all measurements obtained inside the inner edge of the vortex during each flight, and the POAM profile is the average of all individual POAM measurements obtained inside the inner edge of the vortex within  $\pm 1$  day of the ER-2 flights. The error bars represent the statistical uncertainty in the mean profiles.

instrument error is estimated to be  $\pm 5\%$  and the precision is  $\pm 0.1$  ppmv [Hinsta *et al.*, 1999]. Comparisons of the Harvard instrument with several other measurements are presented in the SPARC water vapor report [Kley *et al.*, 2000]. Comparisons with the balloon-borne NOAA CMDL frost point hygrometer were made during three flights from Crows Landing, California, as part of the SPADE mission. The Harvard measurements were consistently higher than CMDL by about 10% at 14 km, increasing to near 40% at 17 km, and decreasing above to 25% at 20 km. Comparisons between the Harvard instrument and the NOAA Lyman- $\alpha$  hygrometer, both flying on the ER-2 also during SPADE, show that the Harvard instrument is consistently higher than the NOAA instrument by 10 to 15%. Comparisons between the Harvard instrument and HALOE showed Harvard measurements higher by 10–20%.

[37] The ER-2 version of the JPL TDL spectrometer is described by May [1998]. The instrument measurement precision is estimated to be 1–2% at stratospheric water vapor levels, and the estimated accuracy is 5% for pressures <100 hPa, 8% for pressures 100–200 hPa, and 10% at higher pressures [Kley *et al.*, 2000]. The TDL and the Harvard Lyman- $\alpha$  instrument were deployed together on the ER-2 for the POLARIS mission. Measurements from these two instruments obtained during POLARIS are compared by Hinsta *et al.* [1999] and Kley *et al.* [2000]. These analyses show average agreement to within 1%, with the error generally <1% at ER-2 cruise altitudes (19–21 km), and an increasing tendency for the TDL measurements to be higher than the Harvard measurements at lower altitudes (to a maximum difference of 5%). Comparisons between the

TDL and the NOAA Lyman- $\alpha$  show a consistent high bias in the JPL TDL measurements of 10–15%, with considerable altitude structure in the difference profile [Kley *et al.*, 2000].

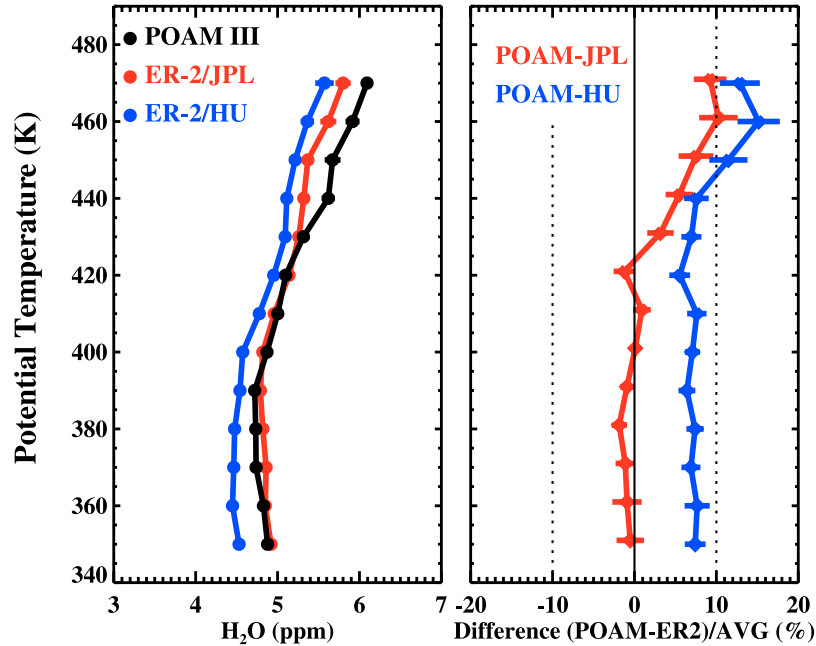
[38] Excluding transit flights, a total of 11 ER-2 science flights were obtained in two deployments during SOLVE/THESEO 2000. Although the number of direct coincidences in space and time between the ER-2 and POAM were relatively few, both POAM and the ER-2 obtained good vortex surveys during each of the ER-2 flight days. In this paper, we compare the POAM and ER-2 water vapor measurements utilizing the same techniques used by Lumpe *et al.* [2002b] in the comparison of POAM/ER-2 ozone. This analysis uses three independent methods to compare the data sets: comparison of vortex average profiles, comparisons using a trajectory-matching technique, and finally comparison of direct ER-2/POAM measurement coincidences. The details of the techniques used to make these three types of comparisons were described in detail by Lumpe *et al.* [2002b].

#### 4.1.1. Vortex Average Comparisons

[39] The results of the vortex average comparisons are illustrated in Figures 15 and 16. Figure 15 shows the vortex-averaged ER-2 TDL and Harvard profiles along with the averaged POAM profile for each flight date. The average POAM profile was formed by using all POAM measurements obtained within  $\pm 1$  day of the ER-2 flight within the inner vortex edge obtained using the objective vortex placement technique described by Nash *et al.* [1996] and the Met Office analysis. Since these are not direct coincidences, it is much more appropriate to do the comparisons



### POAM III and ER-2 In-vortex: Grand Average



**Figure 16.** Grand average POAM and ER-2 Harvard Lyman- $\alpha$  hygrometer and JPL TDL vortex profiles obtained over the entire SOLVE/THESEO mission.

in potential temperature rather than altitude. Also note that the comparisons are done on a 10 K potential temperature grid to retain the full vertical resolution of the ER-2 measurements. However, this results in considerable oversampling of the POAM retrievals, which have a vertical resolution in the comparison region of about 1 km (a 10 K potential temperature increment is about 0.5 km). Figure 15 shows for the most part excellent agreement between the JPL and the Harvard profile shapes, but with a consistent offset between them.

[40] In the lower part of the measurement range, the POAM profile tends to agree well with the JPL measurements throughout most of the mission, but is in better agreement with the Harvard profiles in the last two flights on 11 and 12 March. However, POAM is generally higher than both ER-2 measurements in the upper part of the profile. Notice also that in many cases there is very good agreement in the details of the profile shape between the ER-2 and POAM measurements. For example, the inflection point in the water vapor profile, which occurs near 420 K in the earlier flights and descends as the season progresses, is generally very well reproduced in all three data sets.

[41] Figure 16 shows the grand average profiles and relative differences from all 11 flights. The bias between the JPL and Harvard instruments decreases monotonically from roughly 8% at the lowest potential temperatures to 3–5% in the upper part of the profile. This is a somewhat larger discrepancy than that obtained in the POLARIS campaign. At 420 K and below, POAM agrees exceptionally well on average with the JPL measurements (differences less than 2%) and is approximately 8% higher than Harvard. However, above 420 K an increasing high POAM

bias is observed, reaching a maximum of 8% (JPL) to 12% (Harvard) by 470 K.

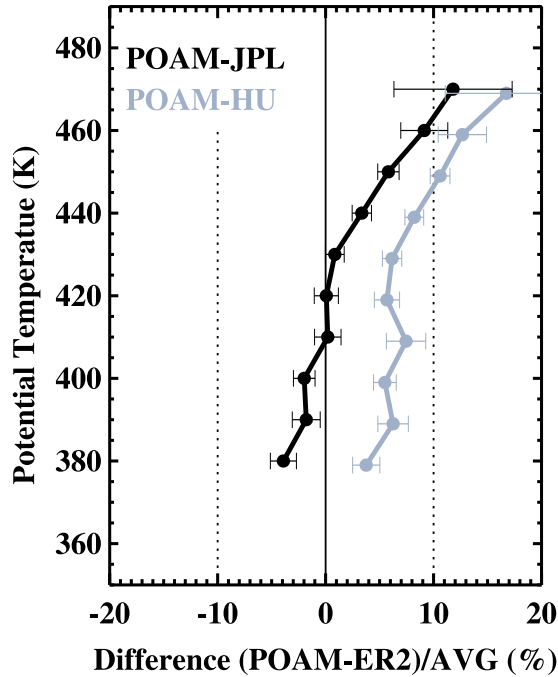
#### 4.1.2. Trajectory-Matching Comparisons

[42] We have also compared the POAM and ER-2 water vapor measurements using a trajectory-hunting approach first used by *Danilin et al.* [2002]. This approach uses an analysis identical to that used for the POAM/ER-2 ozone measurement comparisons described by *Lumpe et al.* [2002b]. That is, forward and backward trajectories were launched from each POAM measurement location at 10 K potential temperature increments using the Met Office wind analysis. The coincidence criteria used to find correlative measurements was 300 km in horizontal separation, 5 K in potential temperature, and 1.2 hours in time. With these criteria, 249 matches were found between 380 and 470 K and the average results are shown in Figure 17. The error bars denote the standard error of the mean estimate. These results are in excellent agreement with the mean vortex average comparisons shown in Figure 15. Again we see that below 430 K POAM agrees well with JPL and is biased high by 5–7% compared to Harvard, whereas at higher altitudes POAM becomes systematically high compared to both ER-2 measurements. This high-altitude difference is somewhat larger than that suggested by the vortex average results, and the altitude gradient in the difference profiles appears somewhat larger.

#### 4.1.3. Direct Coincidence Comparisons

[43] Finally, we present comparisons of direct coincidences between POAM and the ER-2 measurements. In order to search for direct coincidences, we used coincidence criteria of  $\pm 3^\circ$  latitude,  $\pm 10^\circ$  longitude, and  $\pm 3$  hours in time. Most of the coincidences actually occurred in the vicinity of Kiruna during ER-2 takeoff or landing, thus covering a range of altitudes. Coincidences were obtained for all ER-2

### In-Vortex Grand Average

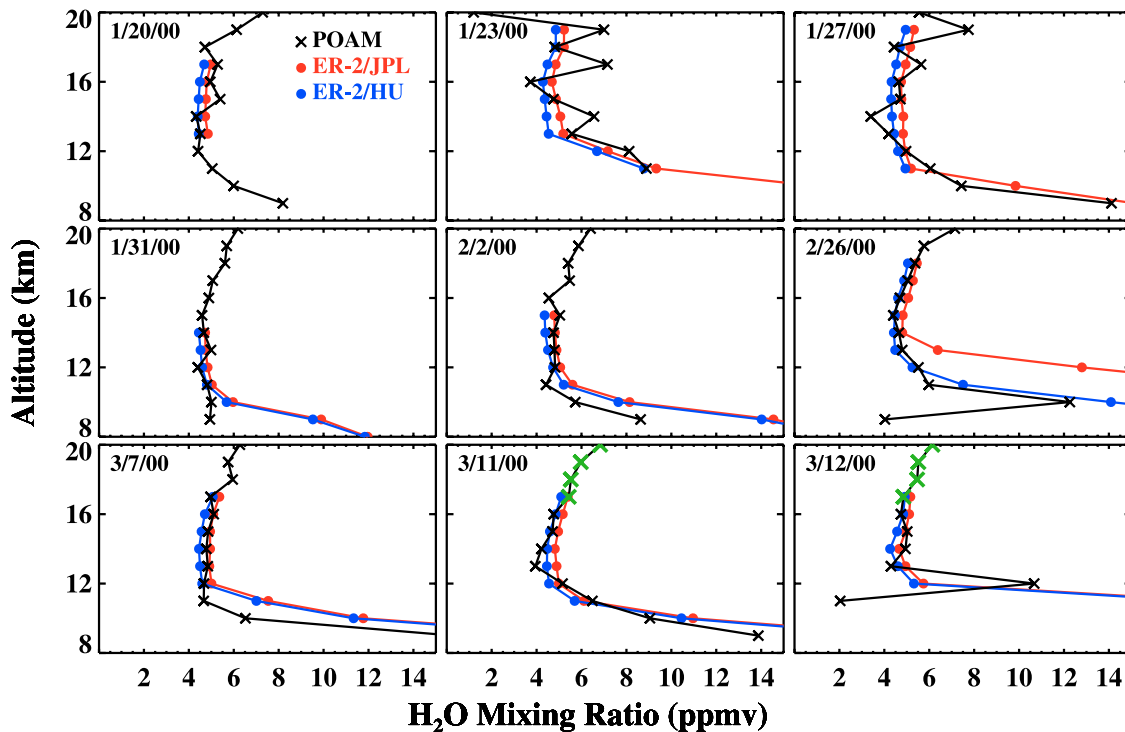


**Figure 17.** Average POAM and ER-2 JPL and Harvard differences obtained over the entire SOLVE/THESEO 2000 mission using the trajectory-matching technique described in the text.

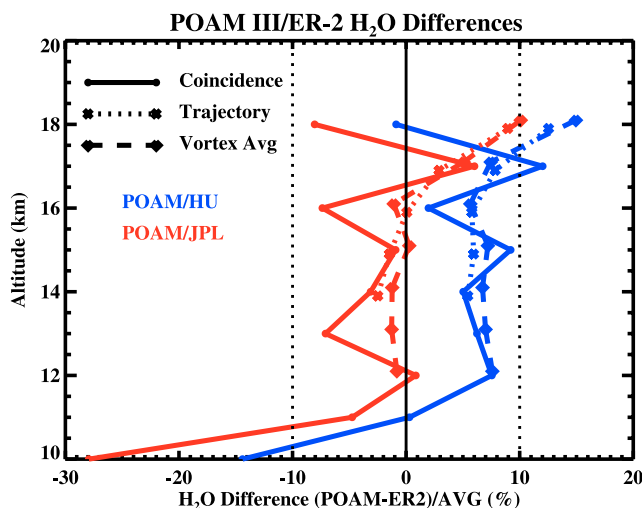
flights during SOLVE/THESEO II except 14 January and 5 March 2000 and the individual comparisons are shown in Figure 18. Note that all of the POAM measurements included in this plot are individual profiles, not averages as were used in all previous comparisons given in this paper. The increased noise in some of the individual profiles is probably primarily due to enhancements in the aerosol optical depth which are below the level at which they are flagged by the aerosol screen. For example, on 23 and 27 January POAM measured high-extinction PSC layers at altitudes just above 20 km, and the structure below 20 km is probably ringing because of the PSC. Likewise, in the 26 February and 12 March measurements there are tropopause cloud layers at 10 and 12 km, respectively, which could be affecting the profiles at the higher altitudes. In all these cases, the aerosol feedback errors are below the adopted threshold values discussed in section 2. This illustrates the imperfection of the aerosol screen and supports our recommendation that users must exercise caution when using individual POAM version 4 water vapor profiles. On the other hand, the 11 and 12 March POAM measurements flagged for sunspot errors above 16 km have no obvious retrieval artifacts and appear to be fine. Again, this illustrates the difficulty in quantifying these two error sources.

[44] Figure 19 shows the average difference profiles from the coincidence comparisons along with the average results from the trajectory-hunting and vortex average comparisons. These two comparisons were placed on an approximate altitude scale using an average potential temperature/altitude profile. From 12 to 17 km there is good agreement

### POAM III and ER-2 Coincident H<sub>2</sub>O Profiles



**Figure 18.** Comparison of POAM (black), JPL (red), and Harvard (blue) H<sub>2</sub>O profiles measured from direct POAM/ER-2 coincidences during the SOLVE/THESEO 2000 mission. In this case all the plotted POAM profiles are from individual POAM retrievals. POAM data points denoted by green symbols on 11 and 12 March are flagged for sunspot contamination.



**Figure 19.** POAM/ER-2 JPL TDL and Harvard Lyman- $\alpha$  hygrometer grand average relative measurement differences obtained using the vortex average technique, the trajectory-matching technique, and direct coincidences.

among the comparison techniques, and we have good confidence that the biases between POAM and the ER-2 measurements are well characterized over this altitude range. It is interesting to note, however, that the coincidence difference profile does not agree as well with the two other methods for the POAM/JPL comparisons as for POAM/Harvard. In both cases the direct coincidence result has more structure, which is reasonable since it is not as heavily averaged. However, for POAM/JPL the direct coincidence analysis seems to indicate a small ( $\sim 5\%$ ) low POAM bias, whereas the vortex average and trajectory analyses show almost no POAM/JPL bias, as discussed previously. At 18 km the coincidence profile average deviates largely from the other two comparisons. However, there are only three individual coincidence profiles that extend as high as 18 km, and the average at 18 km is dominated by the 23 and 27 January events where POAM clearly has large artifacts at that altitude. Thus the coincidence average difference at this altitude is probably not representative of true POAM/ER-2 biases. Below 12 km we only have comparisons from the direct coincidence measurements and these show a general tendency for POAM to be lower than the ER-2 measurements. Figure 18 shows that these altitudes are dominated by small altitude mismatches in the position of the hygropause rather than true POAM/ER-2 biases at these altitudes. These can be caused by small altitude registration errors in POAM, or may also reflect differences in the apparent position of the hygropause over the POAM measurement volume compared to that of the ER-2. Nedoluha *et al.* [2002b] performed a detailed comparison of POAM and MOZAIC in situ measurements in the 9–12 km altitude region and that study did not show any evidence of a POAM low bias. This analysis is currently being reevaluated using the POAM version 4 retrievals.

#### 4.2. MkIV Comparisons

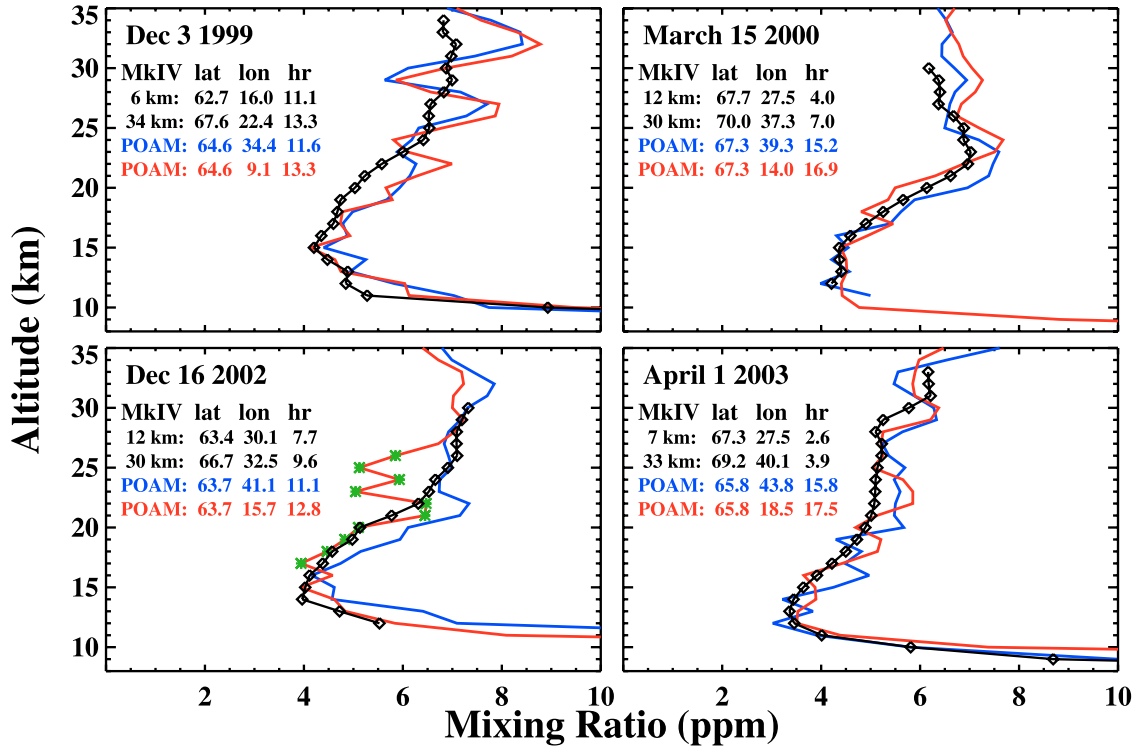
[45] The balloon-borne MkIV solar occultation spectrometer made two flights during each of the SOLVE/THESEO

2000 and SOLVE II missions. This instrument, which is very similar to the JPL ATMOS instrument, is described by Toon [1991], and additional descriptive material is given by Kley *et al.* [2000] and Lumpe *et al.* [2002b]. MkIV measurements have been compared to both the Harvard Lyman- $\alpha$  instrument and the JPL TDL instrument during the POLARIS campaign [Kley *et al.*, 2000]. The results show a great deal of scatter, but the mean MkIV/Harvard difference was about 5% (MkIV lower than Harvard) with little altitude dependence. However, the JPL TDL comparisons show a large altitude dependence with mean MkIV/JPL TDL differences ranging from about 15% at 12 km down to 1 to 2% at 19 km (again with MkIV low).

[46] In comparing the POAM and MkIV water vapor measurements we have used exactly the same methodology employed by Lumpe *et al.* [2002b] to compare POAM and MkIV ozone during SOLVE/THESEO 2000. The comparisons are summarized in Figures 20 and 21. Figure 20 shows the MkIV water vapor profiles measured on each day along with the two nearest POAM profiles on that day. Because the MkIV tangent points are spread out over several great circle degrees, it makes sense to compare the balloon profile to the two POAM measurements bracketing each flight. The time and location of the POAM occultations, along with the corresponding information from the minimum and maximum altitudes of the MkIV occultation, are given in each panel of Figure 20. On the December 1999 flight the MkIV occultation measurement was made at local sunset, whereas on all other days MkIV measured local sunrise occultations. POAM of course always measures at local sunset in the NH. Because the day is so short in December at this latitude the absolute time difference between sunrise and sunset is very small; however for the March/April flights there is a 10 to 12 hour time difference between the MkIV and POAM measurements. For the final MkIV flight on 1 April 2003, POAM was actually not operational on that day so we have used the nearest spatial coincidences from the POAM occultations on 31 March. This has little effect on the absolute time difference between the measurements, which is 10 to 12 hours either way.

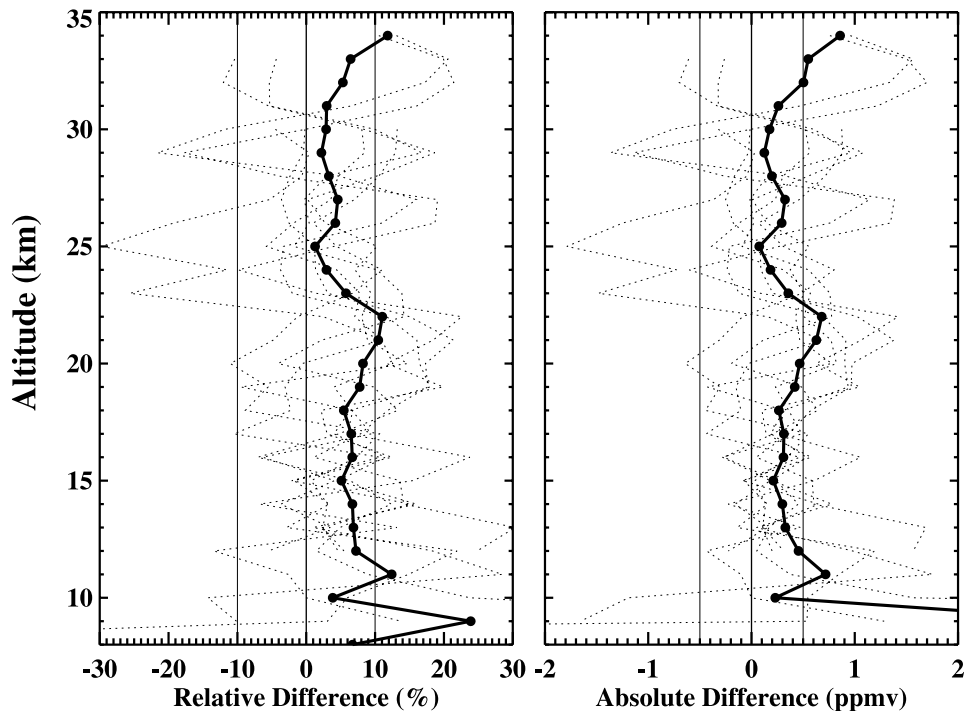
[47] The coincidences from the SOLVE/THESEO 2000 campaign are discussed in detail by Lumpe *et al.* [2002b]. The 3 December 1999 balloon launch provides a much cleaner opportunity for POAM validation than the 15 March 2000 launch for a number of reasons. On the 3 December flight the two POAM measurements were nearly equidistant from the MkIV measurement for the entire profile. However, on 15 March the balloon drifted east a large distance during the occultation measurement, bringing it much closer to one of the POAM measurements at the lowest measurement altitudes. Also, the 3 December MkIV measurement and the two closest spatially coincident POAM measurements were all made deep in the vortex, whereas on 15 March the balloon observation was made closer to the vortex edge. The POAM 39.3°E measurement on 15 March was made at similar equivalent latitude to the MkIV, but the 14°E POAM measurement was made outside the vortex middle edge in the lower stratosphere. A similar situation occurred during the 16 December 2002 flight, where the MkIV measurement and the POAM measurement at 41°E both occurred inside the vortex, but the POAM measurement at 15.7°E was made near the vortex edge. Finally, on the 31 March 2003 flight

## Comparison of POAM III and MkIV H<sub>2</sub>O Profiles



**Figure 20.** POAM/MkIV coincident measurements obtained during the SOLVE/THESEO 2000 and SOLVE II missions. For each MkIV balloon flight the two closest individual POAM measurements are plotted. Green POAM points on 16 December 2002 have been flagged for sunspot artifacts.

## Comparison of POAM III and MkIV H<sub>2</sub>O Profiles



**Figure 21.** (left) Relative and (right) absolute differences for all individual POAM/MkIV coincidence events (dotted lines) and the mean for all coincidences (solid lines).

both MkIV and POAM were measuring outside the vortex, albeit close to the vortex outer edge and thus possibly in perturbed conditions.

[48] Although some scatter in the individual POAM profiles is clearly evident in these plots, the comparisons overall are very good. In particular, the excellent agreement between POAM and MkIV in the resolution of the hygro-pause in the December 1999 and March 2003 comparisons suggests good altitude registration in the POAM measurements. The structure in the POAM profiles above 25 km on 3 December 1999 is characteristic of a sunspot artifact that was not removed in the data screening. Closer study of the POAM data shows that in fact sunspots were detected in these events but at levels below the threshold at which we flag the data. This is therefore a good example of a situation where a significant retrieval artifact was not successfully flagged in the error analysis. The 16 December 2002 POAM profile at 15°E, on the other hand, also has a sunspot artifact, which in this case has been correctly flagged (as denoted by the green symbols in the plot, between 17 and 26 km).

[49] Figure 21 shows both relative and absolute POAM/MkIV differences for all 8 coincidences, as well as the mean difference profiles. On the whole, these comparisons are remarkably consistent and indicate a POAM high bias with respect to MkIV of 5 to 10% from 10 to 22 km, decreasing to 5% or less between 23 and 33 km. There is little evidence of significant altitude dependence in the differences, although it is worthwhile to reiterate that only the two events from December 1999 and March 2003 contribute to the comparisons below 10 km, so the statistics are poor and the scatter is relatively large in this region.

#### 4.3. Comparisons With NOAA Frost Point Hygrometer

[50] The NOAA CMDL balloon-borne frost point hygrometer has provided the most extensive series of stratospheric water vapor measurements yet obtained. The instrument was originally developed at the Naval Research Laboratory [Mastenbrook, 1968], and used to make measurements over Washington, D. C., from 1964 to 1980. The instrument program was then taken over by the NOAA Climate Modeling and Diagnostics Laboratory (CMDL) and with only minor modifications since 1980, has been flown routinely at Boulder, CO from 1980 to the present. In addition, over the past 20 years, the instrument has participated in many field campaigns. A thorough discussion of the current instrument and its operation is given by Vömel *et al.* [1995].

[51] The quoted accuracy of the NOAA CMDL frost point hygrometer measurements is about 10% [Vömel *et al.*, 1995]. Over the years the NOAA CMDL frost point hygrometer measurements have been compared with many other water vapor measurements, and these comparisons are summarized by Kley *et al.* [2000]. Comparisons with HALOE have been made during 25 balloon flights at Boulder, Colorado (14 in HALOE s/c sunrise, and 11 in s/c sunset). For the HALOE s/c sunrise comparisons, the results show no discernable bias between the two instruments between 21 and 29 km. The same is true in the s/c sunset comparisons up to 25 km, but above this altitude there is an indication that HALOE may be 10–15% higher

than the frost point measurement. Below 20 km in both the s/c sunrise and sunset comparisons the frost point measurements become increasingly high compared to HALOE, reaching about 20% at 15 km. Much of this difference appears to result from the problems with resolving the water vapor structure in the vicinity of the hygro-pause discussed previously. It has recently been pointed out that there is evidence of a change in the agreement between HALOE and the CMDL Boulder record in the late 1990s [Randel *et al.*, 2004]. While the agreement looks very good up to 1997, after that the Boulder record seems to show slightly higher water vapor than HALOE. Nevertheless, the Boulder water vapor record is consistent with the drop in 2001 observed by POAM and HALOE, as discussed by Randel *et al.* [2004]. In addition to these comparisons, we have discussed previously the frost point comparisons with the airborne Harvard Lyman- $\alpha$  hygrometer, which show a 15–30% measurement difference (the Harvard instrument consistently higher than the NOAA instrument).

[52] We have analyzed a total of 17 coincidences between POAM and the CMDL frost point instrument, including nine flights made from Kiruna (67.9 N, 21.1 E) and eight from Sodankyla (67.4 N, 26.7 E). Six of the Kiruna flights occurred during the 1999/2000 SOLVE/THESEO winter and the remaining three occurred during the 2002/2003 SOLVE II winter. Of the Sodankyla coincidences, three coincided with the SOLVE II mission, while the remaining five occurred in January/February of 2004. For each CMDL balloon flight we compare the POAM occultation closest to the balloon location on that day (no other time coincidence criteria are used). All coincidences occurred within the 500 km separation used as the coincidence criterion for the satellite comparisons in section 3, with a mean separation of 325 km. The temporal coincidences are also quite good: the time difference for most coincidences is between two and three hours, and never exceeds six hours.

[53] Figure 22 summarizes the results of these 17 comparisons. Since the number of coincidences is significant we have not shown the individual water vapor profiles, but rather the individual and mean difference profiles. The results indicate good agreement between the two instruments for these coincidences, with no evidence of either a systematic difference between the two balloon stations or any time dependence in the POAM/CMDL measurement differences.

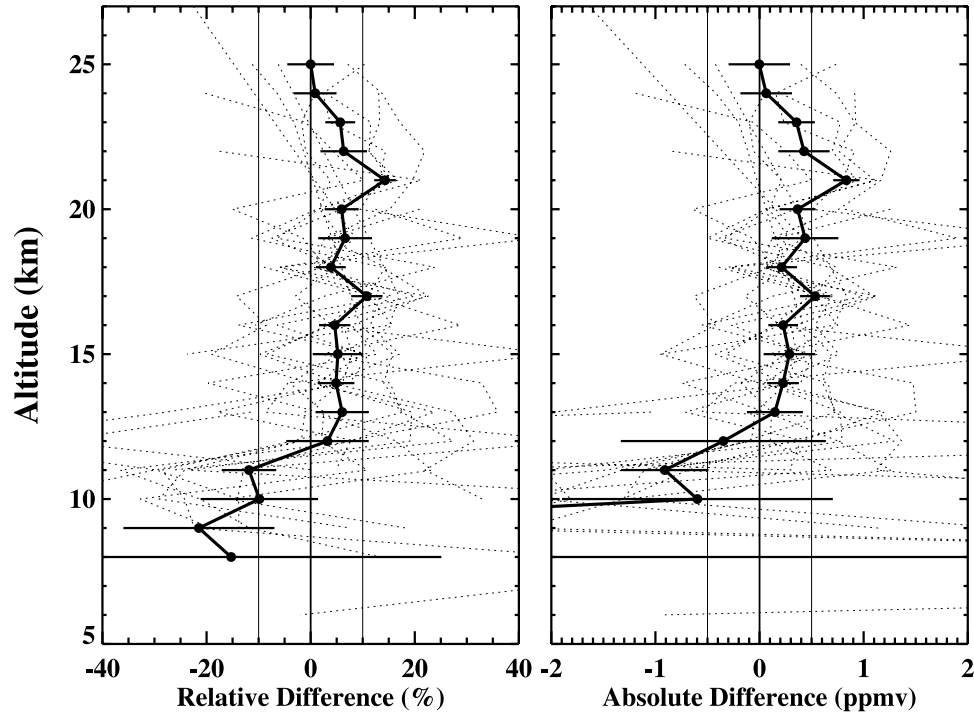
[54] Differences are within 10% at all altitudes between 12 and 25 km, and generally within 5%. In this altitude range the comparisons indicate a very constant and systematic POAM high bias of approximately 5% relative to CMDL. This is very similar to the POAM/MkIV comparisons shown in section 4.2. Below 12 km POAM becomes systematically lower than the frost point measurement, reaching a maximum difference of 15–20% at 8 km. However, the statistics are poor here and the large negative differences at 10 km and below are driven by a single Sodankyla coincidence with poor agreement in the hygro-pause region.

#### 4.4. FISH Comparisons

[55] The Fast In situ Stratospheric Hygrometer (FISH) is a Lyman- $\alpha$  photofragment fluorescence instrument developed at the Forschungszentrum Jülich (Germany). The

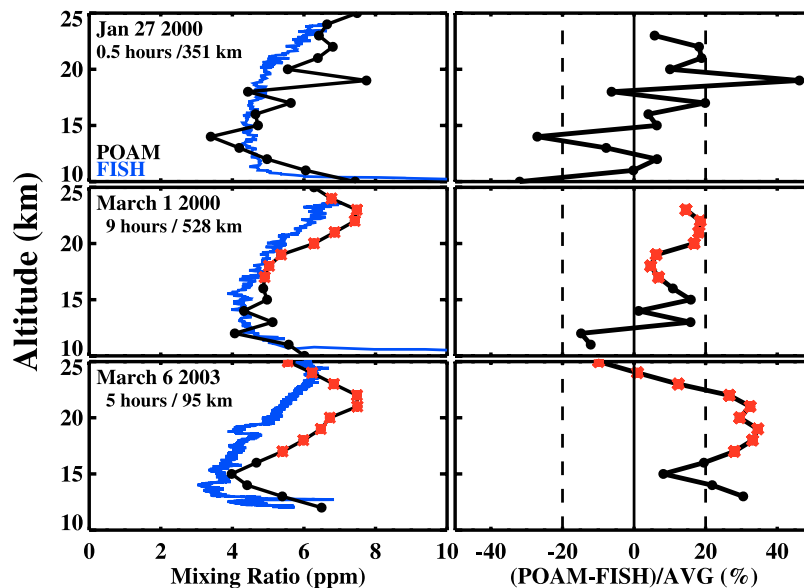


### Combined POAM III/CMDL Coincidences



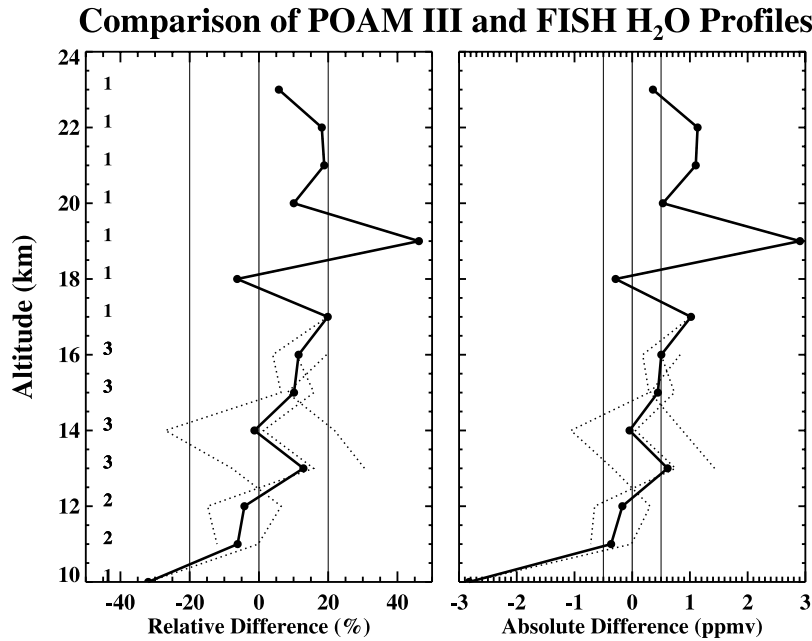
**Figure 22.** A summary of the 17 POAM/NOAA CMDL frost point hygrometer coincident measurements from December 1999 to February 2004 in Kiruna and Sodankyla. The left (right) panel shows the relative (absolute) error profiles for all individual coincidences (dotted lines) as well as the mean difference profile (solid lines). Error bars correspond to the standard deviation of the comparisons.

### POAM III & FISH Coincident Measurements



**Figure 23.** POAM/FISH water vapor measurement comparisons for the three coincidences obtained during SOLVE/THESEO 2000 and SOLVE II. The left panels show the individual water vapor profiles, with POAM plotted in black and FISH plotted in blue. Right panels show the corresponding measurement differences for each coincidence. Red POAM points indicate data flagged for sunspot contamination.





**Figure 24.** (left) Relative and (right) absolute differences for all individual POAM/FISH coincidence events (dotted lines) and the mean for all coincidences (solid lines). Numbers on left axis indicate the number of points contributing to the mean at each altitude.

instrument and the calibration procedure are described in detail by Zöger *et al.* [1999] and Kley *et al.* [2000]. The quoted accuracy of the FISH instrument is 5%, and the precision at typical lower stratospheric mixing ratios is 0.15–0.2 ppmv. The FISH water vapor mixing ratios were compared to several in situ hygrometers and spaceborne sensors of Kley *et al.* [2000], e.g., to those made by the Laboratoire de Meteorologie Dynamique (LMD) during a single pair of balloon flights in France in May 1999. The measurements generally agreed to within 0.2 ppmv with maximum deviations above the tropopause of 8%. There is no clear indication of a consistent bias between the two measurements, although the pair of measurements suggests (or at least is consistent with) a very small (no more than a few percent) low bias in FISH relative to LMD. Very good agreement was also found with the NOAA frost point hygrometer launched from Northern Sweden on 27 January 2000 [Schiller *et al.*, 2002], confirming the indirect comparison between both hygrometers using HALOE measurements as a transfer standard [Kley *et al.*, 2000].

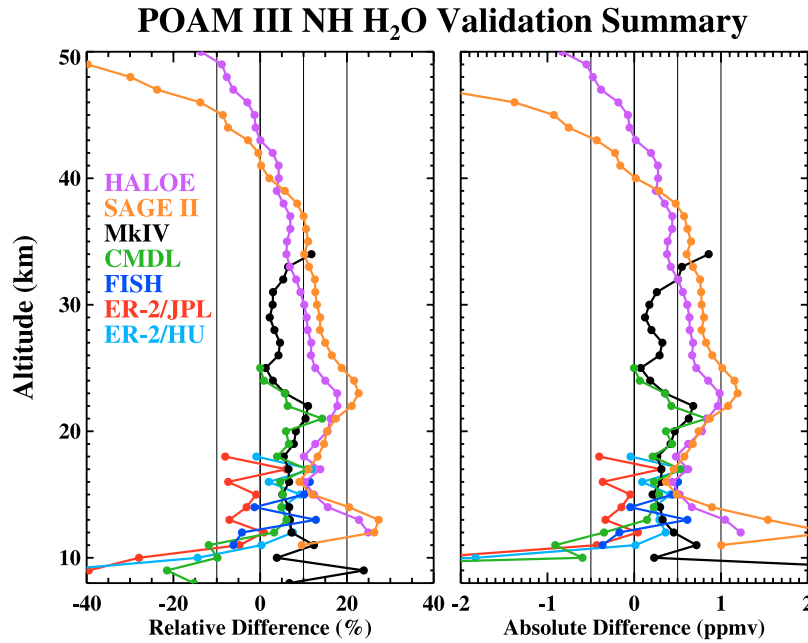
[56] Three POAM coincidences occurred with FISH balloon flights launched from Kiruna, the first two during SOLVE/THESEO 2000 and the third during SOLVE II. For each flight we have compared the FISH data with the POAM measurement made nearest to the balloon location on that day. The FISH measurements, which are reported as mixing ratio as a function of pressure, are first converted to a geometric altitude grid using a colocated Met Office pressure profile. Since the FISH measurements are made at much higher vertical resolution than POAM, the data are then averaged into 1-km altitude bins and interpolated to the native POAM 1-km grid for the comparisons.

[57] The measured water vapor profiles and corresponding relative differences for all three coincidences are shown in

Figure 23. The spatial and temporal separation between the two measurements is given in the top panel for each flight. Note that the January 2000 FISH measurement was made near local sunset, and thus is close in time to the POAM occultation, whereas in March of 2000 and 2003 the FISH measurements were made in the morning and therefore have larger time differences compared to POAM.

[58] The noise level in the January 2000 POAM measurement is significant ( $\pm 20\%$ ), resulting from aerosol retrieval artifacts due to a PSC that was detected at 20 km. This PSC obviously affected the water vapor retrieval, yet was not large enough to trigger the aerosol flag, so these data points have been used in the validation. Unfortunately, in both of the other two coincidences large sunspots were present in the POAM scans, resulting in significant systematic biases in the water vapor retrievals above 16 km. These events are unusual for two reasons: sunspot artifacts usually do not cause a systematic bias in the POAM retrievals, and are generally not important at altitudes as low as 17 km. Nevertheless, these points were appropriately flagged by the error analysis and are denoted by the red symbols in the plots.

[59] The individual POAM/FISH difference profiles for the three events are plotted in the right panels of Figure 23 and the mean relative and absolute error profiles are shown in Figure 24. Because of the sunspot contamination in the March 2000 and March 2003 POAM measurements, only the single coincidence in January 2000 yields a difference profile over the complete altitude range. The numbers along the left side of the vertical axis in Figure 24 indicate the number of points that went into calculating the average at each altitude. Obviously, above 16 km only the single profile from the January 2000 coincidence contributes to the difference. Even though the statistics are not nearly as



**Figure 25.** Summary of the POAM water vapor validation comparisons obtained in the NH. Each colored profile corresponds to the mean (left) relative or (right) absolute difference for all NH comparisons with a given validation data set.

good for the POAM/FISH comparisons as for the other instruments, the results are consistent. Above 12 km, POAM appears to have a more or less altitude-independent high bias of  $\sim 10\%$  relative to FISH. At 12 km and below this difference approaches zero, but again only one or two coincidences are available and therefore the statistical significance is low.

## 5. Summary and Conclusions

[60] The results of this validation study tell a generally consistent story regarding the precision and accuracy of the POAM version 4 water vapor measurements. Figure 25 summarizes the results of all the comparisons obtained in the NH. These profiles are identical to the mean difference profiles for the various validation data sets shown in the previous sections. For the POAM/ER-2 differences we have used only the results of the direct coincidence comparisons, since they span the largest altitude range, while the SAGE II and HALOE differences are determined from the mean of the two NH (P3r) curves for each instrument in Figure 14. Figure 25 clearly shows that the NOAA/CMDL frost point hygrometer, FISH, ER-2 Harvard Lyman- $\alpha$  and MkIV comparisons are remarkably consistent in suggesting that the POAM NH water vapor retrievals are biased high by approximately 5–10% between 12 and 34 km. The satellite comparisons show differences that are about 5–10% larger in this altitude range. As discussed previously, however, HALOE is probably biased approximately 5% low, and the SAGE II measurements, with regard to average biases, are not independent of the HALOE data. Thus most of the measurements appear to be consistent with a POAM high bias of 5–10% in the NH stratosphere below 35 km. Above 35 km we have only the satellite comparisons for guidance.

Here the POAM/HALOE difference decreases monotonically with increasing altitude, changing sign above 40 km such that POAM is lower than HALOE by  $\sim 10\%$  at 50 km. The SAGE II comparisons show the same overall trend, but the SAGE II data above 40 km are increasingly noisy and are known to have a significant high bias, which is apparent in these results.

[61] Of the NH correlative data sets considered in this study, the JPL TDL instrument on board the ER-2 is the only one that does not suggest a POAM high bias in the stratosphere. The POAM/JPL direct coincidence differences plotted in Figure 25 actually implies a POAM low bias of  $\sim 5\%$ . However, as discussed in section 4.1.3, this result is not entirely consistent with the other two methods used for the POAM/ER-2 validation, which yield essentially zero bias between POAM and JPL. Either way, there is a 5–10% discrepancy in the POAM comparisons with the two ER-2 instruments. This is somewhat larger than the Harvard/JPL differences obtained in previous direct comparisons between these two instruments, however these differences have changed over time and are still under investigation.

[62] There is currently no accepted explanation for the apparent high POAM bias in the low and middle stratosphere. Errors due to water vapor spectroscopy and radiative transfer modeling are not considered likely candidates. The version 4 algorithm has been run using all recent HITRAN line-parameter compilations as input and the resulting retrieval differences are less than the observed POAM/HALOE biases. Likewise, the POAM line-by-line radiative transfer code has been validated against other comparable codes to the level of  $\sim 1\%$  [Lumpe et al., 2002a].

[63] One possible explanation for the POAM high bias could be an uncharacterized shift in the position or width of

the optical band pass in the POAM water vapor channels. This is of course the same type of instrument model error thought to be responsible for the SAGE II systematic water vapor biases, for which an empirical correction was implemented in the version 6.2 retrievals. However, there is currently no independent evidence to suggest that these filters were either incorrectly characterized on the ground, or have changed during flight. We will continue to investigate possible causes of the apparent POAM high bias, but we feel that for the version 4 data set this bias has been well characterized by the analysis presented here.

[64] The other significant conclusion drawn from the satellite comparisons in this study is that the POAM water vapor appears to have a *s/c* sunrise/sunset bias of 5–10%, with sunset (SH) retrievals larger than sunrise (NH). This bias also existed in the POAM version 2 and 3 data sets, and has been noted and discussed in previous comparisons [Kley *et al.*, 2000; Taha *et al.*, 2004]. It is interesting to note that the POAM/SAGE II aerosol extinction comparisons presented by Randall *et al.* [2001] show biases which may be consistent with a small (175 m) hemispheric-dependent pointing error. However, retrieval simulations suggest that the POAM version 4 water vapor retrievals would be largely insensitive to pointing errors of this magnitude. Thus it is not likely that the hemispheric-dependent pointing errors suggested by the aerosol comparisons (should they actually be present) are responsible for the observed sunrise/sunset biases in the water vapor retrievals. It is more probable that the bias is the result of small shifts in the band pass of the POAM channel 8 optical filter caused by heating (and thus temperature) differences in the sunrise and sunset occultation events. This possibility is being investigated using the POAM instrument model. The one observation that does not support this hypothesis is that the POAM NO<sub>2</sub> retrievals, which are also quite sensitive to the assumed shape of the filter band pass, do not exhibit a significant sunrise/sunset bias. However, the POAM filters may not necessarily all have exactly the same response or sensitivity to heating and temperature gradients.

[65] Despite these small systematic errors the POAM version 4 water vapor data is self-consistent and robust and we believe it constitutes a unique and valuable data set that is valid for scientific use. The validation analysis spans the entire POAM mission and does not indicate any significant temporal trends in the accuracy or precision of the version 4 water vapor retrievals. The measured retrieval precision has been shown to be well characterized by the version 4 error analysis, and the measured water vapor variability is very consistent with the other satellite data sets used in this study.

[66] **Acknowledgments.** This work was supported by the NASA Scientific Data Purchase program through contract NAS13-02045 and the NASA Solar Occultation Satellite Science Team (SOSST) program through contract NNH04CC14. The POAM version 4 data are archived at the NASA Langley Research Center's Atmospheric Sciences Data Center. We would like to thank E. Weinstock and J. R. Spackman for providing the Harvard Lyman- $\alpha$  hygrometer data and helping in its interpretation. We also acknowledge R. L. Herman for providing the JPL TDL data and useful discussions on the ER-2 data comparisons.

## References

- Benson, C. M., K. Drdla, G. E. Nedoluha, E. P. Shettle, K. W. Hoppel, and R. M. Bevilacqua (2006), Microphysical modeling of southern polar dehydration during the 1998 winter and comparison with POAM III observations, *J. Geophys. Res.*, **111**, D07201, doi:10.1029/2005JD006506.
- Bevilacqua, R. M., M. D. Fromm, J. M. Alfred, J. H. Hornstein, G. E. Nedoluha, K. W. Hoppel, J. D. Lumpe, C. E. Randall, and E. P. Shettle (2002), Observations and analysis of PSCs detected by POAM III during the 1999/2000 Northern Hemisphere winter, *J. Geophys. Res.*, **107**(D20), 8281, doi:10.1029/2001JD000477.
- Danilin, M. Y., et al. (2002), Trajectory hunting as an effective technique to validate multiplatform measurements: Analysis of the MLS, HALOE, SAGE-II, ILAS, and POAM-II data in October–November 1996, *J. Geophys. Res.*, **107**(D20), 4420, doi:10.1029/2001JD002012.
- Giver, L. P., C. Chackerian, and P. Varanasi (2000), Visible and near-infrared H<sub>2</sub>O line intensity corrections for HITRAN-96, *J. Quant. Spectrosc. Radiat. Transfer*, **66**, 101–105.
- Harries, J. E., J. M. Russell III, A. F. Tuck, L. L. Gordley, P. Purcell, K. Stone, R. M. Bevilacqua, M. Gunson, G. Nedoluha, and W. A. Traub (1996), Validation of measurements of water vapor from the Halogen Occultation Experiment (HALOE), *J. Geophys. Res.*, **101**(D6), 10,205–10,216.
- Hints, E. J., E. M. Weinstock, J. G. Anderson, R. D. May, and D. F. Hurst (1999), On the accuracy of in situ water vapor measurements in the troposphere and lower stratosphere with the Harvard Lyman- $\alpha$  hygrometer, *J. Geophys. Res.*, **104**(D7), 8183–8190.
- Kley, D., et al. (2000), SPARC assessment of upper tropospheric and stratospheric water vapor, *WCRP 113, WMO/TD-1043, SPARC Rep. 2*, World Clim. Res. Program, Geneva.
- Lucke, R. L., et al. (1999), The Polar Ozone and Aerosol Measurement (POAM III) instrument and early validation results, *J. Geophys. Res.*, **104**(D15), 18,785–18,799.
- Lumpe, J. D., R. M. Bevilacqua, K. W. Hoppel, and C. E. Randall (2002a), POAM III retrieval algorithm and error analysis, *J. Geophys. Res.*, **107**(D21), 4575, doi:10.1029/2002JD002137.
- Lumpe, J. D., et al. (2002b), Comparison of POAM III ozone measurements with correlative aircraft and balloon data during SOLVE, *J. Geophys. Res.*, **107**, 8316, doi:10.1029/2001JD000472. [printed 108(D5), 2003]
- Mastenbrook, H. J. (1968), Water vapor distribution in the stratosphere and the high troposphere, *J. Atmos. Sci.*, **5**, 299–311.
- May, R. D. (1998), Open-path, near-infrared tunable diode laser spectrometer for atmospheric measurements of H<sub>2</sub>O, *J. Geophys. Res.*, **103**(D15), 19,161–19,172.
- Nash, E. R., P. A. Newman, J. E. Rosenfield, and M. R. Schoeberl (1996), An objective determination of the polar vortex using Ertel's potential vorticity, *J. Geophys. Res.*, **101**(D5), 9471–9478.
- Nedoluha, G. E., R. M. Bevilacqua, K. W. Hoppel, M. Daehler, E. P. Shettle, J. H. Hornstein, M. D. Fromm, J. D. Lumpe, and J. E. Rosenfield (2000), POAM III measurements of dehydration in the Antarctic lower stratosphere, *Geophys. Res. Lett.*, **27**(12), 1683–1686.
- Nedoluha, G. E., R. M. Bevilacqua, and K. W. Hoppel (2002a), POAM III measurements of dehydration in the Antarctic and comparisons with the Arctic, *J. Geophys. Res.*, **107**(D20), 8290, doi:10.1029/2001JD001184.
- Nedoluha, G. E., R. M. Bevilacqua, K. W. Hoppel, J. D. Lumpe, and H. Smit (2002b), Polar Ozone and Aerosol Measurement III measurements of water vapor in the upper troposphere and lowermost stratosphere, *J. Geophys. Res.*, **107**(D10), 4103, doi:10.1029/2001JD000793.
- Prados, A. I., G. E. Nedoluha, R. M. Bevilacqua, D. R. Allen, K. W. Hoppel, and A. Marengo (2003), POAM III ozone in the upper troposphere and lowermost stratosphere: Seasonal variability and comparisons to aircraft observations, *J. Geophys. Res.*, **108**(D7), 4218, doi:10.1029/2002JD002819.
- Randall, C. E., R. M. Bevilacqua, J. D. Lumpe, and K. W. Hoppel (2001), Validation of POAM III aerosols: Comparison to SAGE II and HALOE, *J. Geophys. Res.*, **106**(D21), 27,525–27,536.
- Randall, C. E., et al. (2002), Validation of POAM III NO<sub>2</sub> measurements, *J. Geophys. Res.*, **107**(D20), 4432, doi:10.1029/2001JD001520.
- Randall, C. E., et al. (2003), Validation of POAM III ozone: Comparisons with ozonesonde and satellite data, *J. Geophys. Res.*, **108**(D12), 4367, doi:10.1029/2002JD002944.
- Randel, W. J., F. Wu, S. J. Oltmans, K. Rosenlof, and G. E. Nedoluha (2004), Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures, *J. Atmos. Sci.*, **61**, 2133–2148.
- Rothman, L. S., et al. (2003), The HITRAN molecular spectroscopic database: Edition of 2000 including updates through 2001, *J. Quant. Spectrosc. Radiat. Transfer*, **82**, 5–44.
- Russell, J. M., L. L. Gordley, J. H. Park, S. R. Drayson, A. F. Tuck, J. E. Harries, R. J. Cicerone, P. J. Crutzen, and J. E. Frederick (1993), The Halogen Occultation Experiment, *J. Geophys. Res.*, **98**(D6), 10,777–10,797.

Benson, C. M., K. Drdla, G. E. Nedoluha, E. P. Shettle, K. W. Hoppel, and R. M. Bevilacqua (2006), Microphysical modeling of southern polar de-

- Schiller, C., et al. (2002), Dehydration in the Arctic stratosphere during the SOLVE/THESEO-2000 campaigns, *J. Geophys. Res.*, *107*(D20), 8293, doi:10.1029/2001JD000463.
- Siskind, D. E., S. D. Eckermann, J. P. McCormack, M. J. Alexander, and J. T. Bacmeister (2003), Hemispheric differences in the temperature of the summertime stratosphere and mesosphere, *J. Geophys. Res.*, *108*(D2), 4051, doi:10.1029/2002JD002095.
- Taha, G., L. W. Thomason, and S. P. Burton (2004), Comparison of Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 water vapor with balloon-borne and space-based instruments, *J. Geophys. Res.*, *109*, D18313, doi:10.1029/2004JD004859.
- Thomason, L. W., S. P. Burton, N. Iyer, J. M. Zawodny, and J. Anderson (2004), A revised water vapor product for the Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 data set, *J. Geophys. Res.*, *109*, D06312, doi:10.1029/2003JD004465.
- Toon, G. C. (1991), The JPL MkIV interferometer, *Opt. Photonic News*, *2*, 19–21.
- Vömel, H. S., J. Oltmans, D. J. Hofmann, T. Deshler, and J. M. Rosen (1995), The evolution of the dehydration in the Antarctic stratospheric vortex, *J. Geophys. Res.*, *100*(D7), 13,919–13,926.
- Weinstock, E. M., E. J. Hintsa, A. E. Dessler, J. F. Oliver, N. L. Hazen, J. N. Demusz, N. T. Allen, L. B. Lapson, and J. G. Anderson (1994), New fast response photofragment fluorescence hygrometer for use on the NASA ER-2 and the Perseus remotely piloted aircraft, *Rev. Sci. Instrum.*, *65*, 3544–3554.
- Zöger, M., et al. (1999), Fast in situ stratospheric hygrometers: A new family of balloon-borne and airborne Lyman- $\alpha$  photofragment fluorescence hygrometers, *J. Geophys. Res.*, *104*(D1), 1807–1816.
- R. Bevilacqua, K. Hoppel, and G. Nedoluha, Naval Research Laboratory, 4555 Overlook Avenue SW, Washington, DC 20375, USA.
- V. L. Harvey and C. Randall, Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, 1234 Innovation Drive, Boulder, CO 80303-7814, USA.
- J. Lumpe, Computational Physics, Inc., 1650 38th Street, Suite 105W, Boulder, CO 80301, USA. (lumpe@cpi.com)
- J. Russell, Center for Atmospheric Sciences, Hampton University, Hampton, VA 23668, USA.
- C. Schiller, Institut für Chemie und Dynamik der Geosphäre I, Forschungszentrum Jülich, D-52425 Jülich, Germany.
- B. Sen and G. Toon, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA.
- G. Taha, Science Systems and Applications, Inc., 10210 Greenbelt Road, Suite 600, Lanham, MD 20706, USA.
- H. Vömel, Cooperative Institute for Research in Environmental Sciences, University of Colorado, 216 UCB, Room 318, Boulder, CO 80309-0216, USA.