Continuing the MLS water vapor record with OMPS LP using neural networks

Michael D. Himes[1,2], Natalya A. Kramarova[2], Krzysztof Wargan[3,2], Sean M. Davis[4], and Glen Jaross[2]

[1]Morgan State University, Baltimore, MD, USA
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]Science Systems Association, Inc., Lanham, MD, USA
[4]NOAA Chemical Sciences Laboratory, Boulder, CO, USA

Correspondence: Michael D. Himes (michael.d.himes@nasa.gov)

Abstract. Stratospheric water vapor (SWV) plays an important role in atmospheric chemistry, dynamics, and radiative forcing. Satellite measurements by the Aura Microwave Limb Sounder (MLS), SciSat-1 Atmospheric Chemistry Experiment ~~(ACE~~Fourier Transform Spectrometer (ACE-FTS), and Stratospheric Aerosol and Gas Experiment III ~~(SAGE III)~~ on the International Space Station (SAGE III/ISS) have provided key constraints on SWV for the past decades. MLS provides the best geographical coverage among these instruments, but it approaches the end of its life cycle in the coming years, which will result in a data desert for satellite-based SWV measurements given that ~~ACE~~ ACE-FTS and SAGE III/ISS only measure at a few dozen geolocations per day. The Ozone Mapping and Profiler Suite Limb Profiler (OMPS LP) is flying aboard the Suomi National Polar-orbiting Partnership (SNPP) and NOAA-21 satellites and is planned for additional platforms in the coming years. While not designed to measure SWV, it shows weak sensitivity to it, particularly in the wake of the Hunga eruption's significant injection of water vapor into the stratosphere. By utilizing the frequent co-locations between OMPS LP and MLS measurements, we developed a neural network-based ~~approach to retrieve~~ retrieval algorithm to estimate SWV from SNPP OMPS LP radiances between 11.5–40.5 km. We find that the LP SWV profiles typically agree with MLS within 5%, and agreement with ~~ACE~~ ACE-FTS and SAGE III/ISS is typically within 10%. We show that the SNPP-trained model is applicable to NOAA-21 OMPS LP without retraining, though minor differences in radiances between the instruments results in a ∼5% bias under most conditions. Our results suggest that OMPS LP can continue the global water vapor record in the lower stratosphere into the 2030s, though continued independent measurements by satellite and balloon-borne instruments will be key to verifying the stability of our approach for quantifying decadal-scale SWV variability.

1 Introduction

Stratospheric water vapor (SWV) influences atmospheric dynamics, chemistry, and radiative forcing (e.g., Ramanathan and Inamdar, 2006; Charlesworth et al., 2023; Niemeier et al., 2023; Fleming et al., 2024). While SWV is typically 3–6 parts per million by volume (ppmv), water vapor concentrations in the upper troposphere can reach up to 1000 ppmv ~~(Read et al., 2022)~~ (e.g., Read et al., 2022). Deep convective systems and tropical upwelling via the Brewer-Dobson circulation can transport tropospheric air into the lower stratosphere, which comprises an important contribution to SWV (e.g., Fueglistaler et al., 2009;

Khaykin et al., 2009; Randel and Jensen, 2013; Dauhut et al., 2016). Rising atmospheric temperature due to climate change increases the amount of water vapor held in tropospheric air, which in turn increases the amount of water vapor transported into the stratosphere (Yue et al., 2019; Nowack et al., 2023). Long-term satellite measurements of SWV provide key constraints on the SWV budget and serve as important sources for data assimilation and validation of reanalysis frameworks (e.g., Davis et al., 2017; Hersbach et al., 2020; Wargan et al., 2023; Knowland et al., 2025).

Presently, satellite retrievals of SWV profiles are performed by the Aura Microwave Limb Sounder ~~(MLS; Livesey et al., 2021)~~ (MLS; Waters et al., 2006; Livesey et al., 2021), SciSat-1 Atmospheric Chemistry Experiment Fourier Transform Spectrometer ~~(ACE-FTS; Boone et al., 2023)~~(ACE-FTS; Bernath et al., 2005; Boone et al., 2023), Stratospheric Aerosol and Gas Experiment III ~~(SAGE III)~~ aboard the International Space Station ~~(Davis et al., 2021; Park et al., 2021)~~(SAGE III/ISS Cisewski et al., 2014), and the Thermosphere Ionosphere Mesosphere Energetics Dynamics (TIMED) satellite's Sounding of the Atmosphere using Broadband Emission Radiometry ~~(SABER; Rong et al., 2019)~~ (SABER; Russell et al., 1999) instruments. These instruments provide well validated H_2O products that ~~agree~~ typically show median differences within $\sim$10% among each other as well as with ground-based and in-situ measurements ~~(e.g., Carleer et al., 2008; Hurst et al., 2014; Rong et al., 2019; Davis et al., 2021; De Los Rí~~ (e.g., Carleer et al., 2008; Hurst et al., 2014; Rong et al., 2019; Davis et al., 2021; Park et al., 2021; De Los Ríos et al., 2024). Since May 2024, the MLS receiver used for the H_2O retrievals now only operates around 6 days per month ~~due to power constraints~~ to preserve measurement lifetime and will continue to do so until the end of the Aura mission, which significantly limits the spatiotemporal coverage of the MLS H_2O product. Following the decommissioning of Aura, ACE-FTS and SAGE III/ISS will continue to provide their H_2O products, but their geographical coverage is limited given they are solar occultation instruments. SABER takes around 1400 scans per day and, depending on time of year, views between 52°S–83°N or 83°S–52°N~~, but~~ ; the local times of the measurements change by up to 12 hours over TIMED's two-month yaw cycle. The Canadian High-altitude Aerosols, Water vapour, and Clouds mission (HAWC; Langille et al., 2025) is planned to launch early next decade, presenting a gap in global geographical coverage of SWV between Aura's decommissioning and HAWC's launch (Salawitch et al., 2025).

The Ozone Mapping and Profiler Suite (OMPS) Limb Profiler ~~(LP)~~ (LP; Dittman et al., 2002; Leitch et al., 2003) is currently flying aboard the Suomi National Polar-orbiting Partnership (SNPP; launched October 28, 2011) and NOAA-21 (launched November 10, 2022) satellites, and it is planned for the Joint Polar Satellite System (JPSS) 4 and 3 satellites, which are estimated to launch in 2027 and 2032, respectively. Using 3 slits, the instrument measures limb-scattered radiances between 290–1000 nm with a 1 km vertical sampling. Each satellite completes 14–15 Sun-synchronous orbits per day. SNPP's OMPS LP takes measurements every $\sim$1° latitude, while NOAA-21's LP takes measurements every $\sim$0.4° latitude, resulting in around 7000 and 17,500 measured radiance profiles per day, respectively. However, OMPS LP is only weakly sensitive to H_2O, which has challenged the application of traditional radiative transfer-based retrieval methods.

Our solution to retrieve water vapor profiles from OMPS LP measurements is deep learning (Goodfellow et al., 2016). Neural networks (NNs) learn to model complex, nonlinear processes in a data-driven manner. Given a set of corresponding inputs and outputs, the NN weights are tuned to approximate the underlying process, without explicit knowledge about it. Given OMPS LP's weak sensitivity to H_2O and the well validated MLS H_2O product, NNs could learn to accurately predict H_2O profiles from OMPS LP measurements at altitudes with sufficient sensitivity by using co-located MLS H_2O profiles as the target outputs. This would result in an MLS-like H_2O product, thereby continuing the MLS global SWV record following the end of the Aura mission.

Here we present an OMPS LP water vapor product between 11.5–40.5 km produced by a NN trained on co-located LP-MLS measurements. In Section 2 we investigate the sensitivity of LP to water vapor under conditions before and after the Hunga ~~eruption~~Tonga–Hunga Ha‘apai (hereafter Hunga) eruption, which injected more than 50 Tg of water into the stratosphere (e.g., Millán et al., 2022; Vömel et al., 2022; Khaykin et al., 2022; APARC, 2025). In Section 3 we discuss our methodology to train the NN as well as validate its predictions using other instruments. We present and discuss our results in Section 4, including the limitations of our approach. Finally, we present our conclusions in Section 5.

2 OMPS LP sensitivity to H_2O

We investigate OMPS LP's H_2O sensitivity using the Gauss-Seidel limb scattering radiative transfer model (RTM) of Loughman et al. (2004). To calculate H_2O cross sections, we use the HITRAN 2020 database (Gordon et al., 2022) via the HITRAN Application Programming Interface (Kochanov et al., 2016). We then convolve those high-resolution cross sections with the OMPS LP bandpasses such that RTM calculations at a given wavelength will be more consistent with what LP would measure for the assumed conditions. Using two selected co-located MLS H_2O profiles from before and after the Hunga eruption, we simulate radiances between 550–1025 nm in 5 nm intervals at altitudes of 3.5–50.5 km in 1 km intervals. For the two RTM simulations, all parameters are kept ~~the same~~ identical except for the H_2O profile ~~to ensure~~ ensuring that any differences in the Jacobians are due solely to ~~differences~~ variations in H_2O.

Based on the Jacobians output by the RTM, we select 12 wavelengths measured by OMPS LP (554, 596, 654, 720, 728, 824, 917, 929, 943, 956, 970, and 983 nm) that show the highest sensitivity to H_2O in their spectral region. Figure 1 shows an example of these Jacobians at 945 nm for the selected H_2O profile after the Hunga eruption. The H_2O enhancement between 20–30 km attributable to the Hunga eruption results in as much as a four times increase in LP's sensitivity to H_2O. However, LP is very weakly sensitive to H_2O, and the sensitivity becomes negligible above 30 km.

3 Deep learning methods

3.1 Data curation

Since there is currently no water vapor product derived from OMPS LP measurements, we instead use the MLS version 5 water vapor product as our ground truth. To prioritize times where SNPP and Aura have closely aligned orbits, we select dates that have

- at least one orbit with 60 consecutive co-locations that are within 30 minutes and within 100 km, and
- at least 250 total co-locations on that day that satisfy the above co-location criteria.

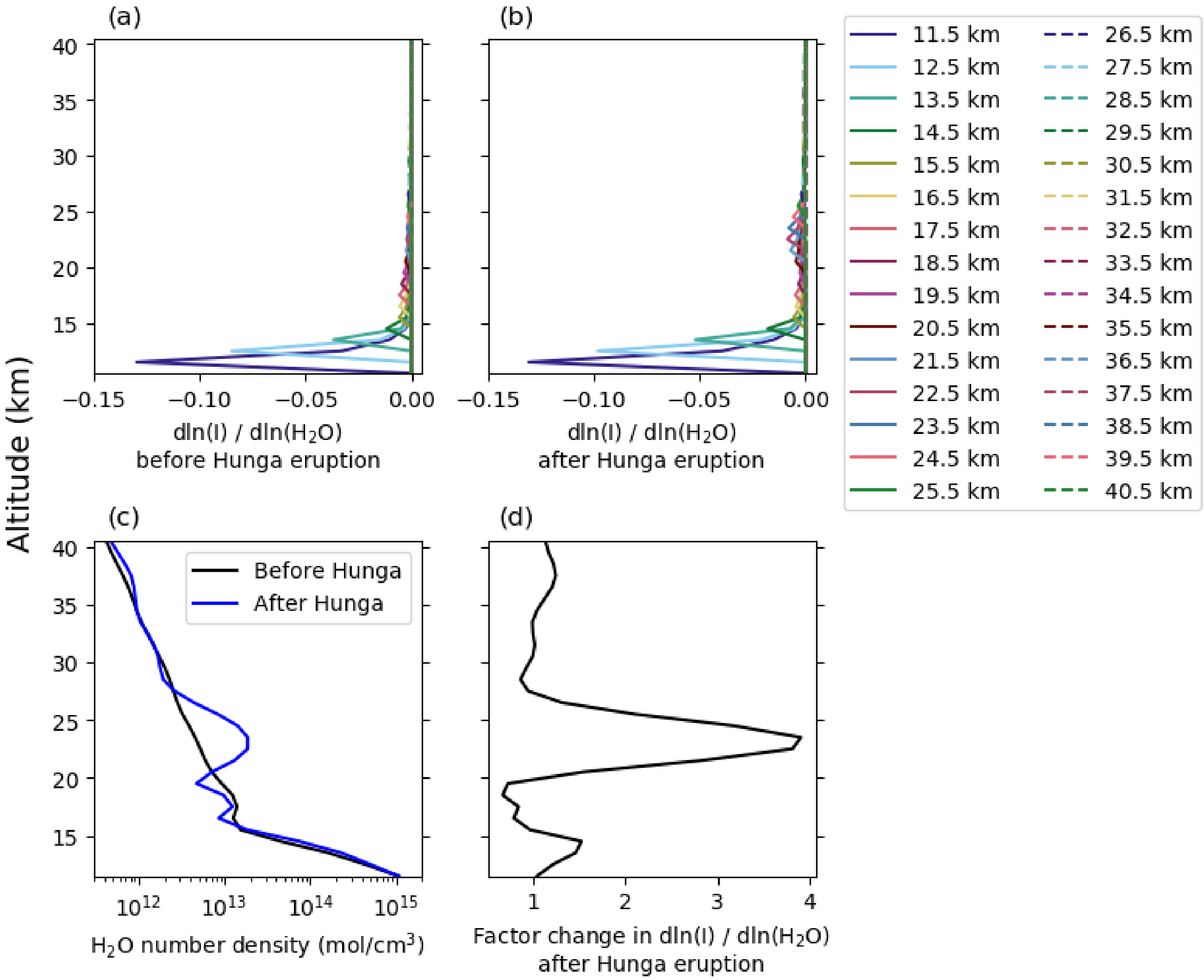

Figure 1. Example of OMPS LP's sensitivity to H_2O at 945 nm in the tropics. Panels **(a)** and **(b)** show the Jacobians for selected MLS H_2O profiles ~~(a)~~ from before and ~~(b)~~ after the Hunga eruption~~peak in the upper troposphere due to the increased water vapor content~~, respectively. Panel **(c)** shows the selected water vapor profiles, which show an enhancement in water vapor around 24 km attributable to the Hunga eruption. Panel **(d)** shows the ratio of (b) to (a). The increased water vapor concentration between 21–27 km due to Hunga results in LP's sensitivity increasing by up to 4× in this altitude range.

These criteria are satisfied every couple days due to their similar equatorial crossing times around 1:30 in the afternoon. On dates between February 2014 and December 2024 that satisfy the above criteria, we co-locate OMPS LP and MLS measurements within 6 hours and 100 km to build a data set of OMPS LP radiances and the corresponding MLS water vapor profiles. This results in 2,074,101 co-locations, with almost half occurring at high latitudes. For context, SNPP OMPS LP ~~collects~~ measures over 2.5 million ~~measurements~~ vertical profiles per year.

We apply the recommended MLS data screening criteria except for the period of January 15 – February 12, 2022, as it has been shown that the recommended screening criteria filter out many valid profiles affected by Hunga (Millán et al., 2024). We limit the MLS water vapor profiles to ≤261 hPa~~, as the 316 hPa pressure level can be affected by the a priori profiles used in the MLS retrievals~~. We log-linearly interpolate the water vapor profiles from the MLS pressure grid to OMPS LP's geometric height (11.5–40.5 km in 1 km steps) using the NASA Global Earth Observing System Forward Processing for Instrument Teams (GEOS FP-IT; Lucchesi, 2015) pressures interpolated in time and space to the LP measurements. We limit the altitude range to 11.5–40.5 km because 10.5 km can exceed 261 hPa and the H_2O sensitivity of OMPS LP becomes ∼0 above 40.5 km.

We also consider a similar methodology ~~but for ACE~~ for ACE-FTS and SAGE III~~data using~~/ISS data with recommended screening criteria applied to both datasets. We utilize co-location criteria of within 1 day, within 2° latitude, and within 1113 km longitude (equal to 10° longitude at the equator), consistent with the criteria used in Davis et al. (2021). These data sets are used to investigate whether it is viable to train exclusively on ~~ACE~~ ACE-FTS or SAGE III/ISS data and whether training on a combination of MLS, ~~ACE~~ACE-FTS, and/or SAGE III/ISS data offers benefits over only training on MLS data.

3.2 Neural network methods

For each co-located measurement, we construct an input–output pair to be used during NN training. The inputs are comprised of

- LP radiances at 554, 596, 654, 720, 728, 824, 917, 929, 943, 956, 970, and 983 nm,
- GEOS FP-IT pressures and temperatures interpolated in time and space to the LP measurements, and
- the solar zenith angle of the LP measurement.

These inputs are formatted as 2-D "images" (wavelength × altitude) with four channels (radiance, pressure, temperature, solar zenith angle), similar to a standard RGBA image. The radiances vary at each point in the 2-D image, the pressures and temperatures vary only with respect to altitude, and the solar zenith angle is constant throughout. For each input image, the corresponding outputs are the co-located MLS H_2O profile that has been interpolated to the LP altitude grid. See Appendix B for a discussion of the relative importance of these inputs.

To address the latitudinal sampling bias inherent in the co-located data set, we first select a subset of the data such that there are roughly the same number of samples in each 5° latitudinal bin, resulting in 1,137,100 input–output pairs. We ensure that extrema for each latitudinal bin are included in this subset. We then split these data into training (used to update NN weights), validation (monitors for overfitting during training), and testing (tests model generality on unseen data after training is complete) sets in a proportion of roughly 75%, 15%, and 10%, respectively. For context, this results in the training and validation sets containing a total of ∼3.7% of all available SNPP OMPS LP data over the 2014–2024 period considered.

To investigate how the training data impacts the algorithm's ability to generalize to unseen data, we run additional experiments where certain years are omitted from the training data set. One test omits 2015–2016, which exhibits typical SWV, while the other test omits 2024, which contains SWV perturbations from Hunga.

For data pre-processing and NN training, we utilize the open-source Python package MARGE (Himes et al., 2022), which uses TensorFlow (Abadi et al., 2016) via the Keras API. We pre-process the data by taking the base-10 logarithm of the OMPS LP radiances, GEOS FP-IT pressures, and MLS H_2O profiles, then scale each input and output parameter to be within the closed interval [-1, 1] based on their training set extrema at each altitude.

To determine a neural network architecture well suited to solving this problem, we perform a Bayesian hyperparameter optimization (Akiba et al., 2019). The selected architecture is similar to the landmark AlexNet architecture (Krizhevsky et al., 2012); for details on our optimization procedure, the selected architecture, and the training details, see Appendix A. Using the chosen architecture, we train an ensemble of 10 neural networks using a mean-squared-error loss function; members are only differentiated by their random initialization. The ensemble ~~'s~~ size is held constant for all retrievals; it was determined by adding ensemble members until the mean and standard deviation of the ensemble's predictions did not significantly change. The ensemble's mean prediction provides the retrieved H_2O profile~~, while the~~. The standard deviation among the ensemble provides an uncertainty estimate ~~. The ensemble size was determined by adding ensemble members until the mean and standard deviation of the ensemble' s predictions did not significantly change.~~ for the NNs' prediction of the MLS H_2O profile. MLS profiles also have some uncertainty; we factor this into the LP uncertainty estimate by combining these sources in quadrature, assuming that they are uncorrelated:

$$\boldsymbol{\sigma}^2_{\mathrm{LP}} = \boldsymbol{\sigma}^2_{\mathrm{NN}} + \boldsymbol{\sigma}^2_{\mathrm{MLS}} \qquad (1)$$

We assume these terms are uncorrelated given that LP's errors at the wavelengths considered are dominated by systematics, and these should be unrelated to any random or systematic errors in the MLS data; this assumption has a negligible effect on the resulting LP uncertainty, as discussed in Section 4.1. Typically, $\boldsymbol{\sigma}_{\mathrm{NN}}$ is small in the stratosphere due to close agreement among the NNs, leading the total LP uncertainties to be dominated by $\boldsymbol{\sigma}_{\mathrm{MLS}}$ in this regime. In the troposphere, where strong scattering limits the NNs' accuracy, the LP uncertainties are dominated by $\boldsymbol{\sigma}_{\mathrm{NN}}$. We estimate the MLS uncertainty by computing 5° monthly zonal means at each pressure level considered for the reported MLS uncertainties over the 2015–2019 period. The chosen time period is a minor consideration, as the MLS data quality and description document reports that other time periods are similar to their reported January–March 2009 period (Livesey et al., 2022). These seasonal uncertainties are then log-linearly interpolated in time and pressure to the LP altitude grid using the GEOS-IT pressure data.

Additionally, we apply these methods to permutations of the co-located MLS, ~~ACE~~ACE-FTS, and/or SAGE III/ISS data sets. When combining more than one data set, we consider two approaches, one where we use the data as is, and another where we bias correct the ~~ACE~~ ACE-FTS and SAGE III/ISS data such that they have a global median difference of 0% at all altitudes with respect to MLS.

3.3 Evaluation and validation

To evaluate the NN's typical accuracy and how well it generalizes to unseen data, we calculate the root mean square error (RMSE) and coefficient of determination (R^2) for the validation and test sets. We validate our LP water vapor product by comparing with satellite measurements, balloon-borne measurements, and an assimilation/reanalysis product.

For satellite measurements, we consider the MLS version 5 ~~(Livesey et al., 2021)~~(Lambert et al., 2020), SAGE III/ISS version 6 (NASA/LARC/SD/ASDC, 2025), and ~~ACE~~ ACE-FTS version 5.3 (Boone et al., 2023) products. For MLS, we consider co-locations within 6 hours, while for SAGE ~~and ACE~~ III/ISS and ACE-FTS we consider co-locations within 24 hours. For all three, we only consider the co-location if it is separated by less than 1000 km and within 2° latitude. When multiple co-locations satisfy these criteria, we use the co-location with the shortest distance. For each of these instruments, we compute both a global median percent difference as well as zonal median percent differences in 5° latitude bins. We found anomalous values in the SAGE ~~and ACE~~ III/ISS and ACE-FTS data sets that differ by a factor of up to 10,000 compared with the layers above and below it, even after applying each product's recommended screening criteria. To screen out the most extreme of these unrealistic values, we apply a very conservative 20σ median rejection routine to the data set of percent differences.

~~Given the reported instrumental drift in version 5 of the MLS water vapor product (Livesey et al., 2021), we~~ We investigate whether our product shows similar properties as the MLS product by performing a multiple linear regression (MLR) on monthly means for a 5°×5° grid, with proxies for a linear trend, seasonal cycle, quasi-biennial oscillation (QBO), and El Niño Southern Oscillation (ENSO), as these terms explain the majority of stratospheric H_2O variability. For the seasonal cycle term, we also include a phase offset term for lag in months, given that it takes time for the effects to propagate upwards through the stratosphere. For the ENSO term, we regress using the sea surface temperature anomaly with a fitted lag in months, as previous work showed this is necessary to maximize correlation (e.g., Garcia et al., 2007; Calvo et al., 2010; Yu et al., 2022). For the QBO term, we use coefficients for the two leading empirical orthogonal functions for the QBO wind time series between January 1956–February 2025.

To compare with balloon-borne measurements, we consider NOAA Frost Point Hygrometer (Hurst et al., 2011) and Cryogenic Frost point Hygrometer (Vömel et al., 2007a, b) soundings from Boulder, USA; Hilo, USA; Lauder, New Zealand; San José, Costa Rica; Lindenburg, Germany; and Biak, Indonesia. For each sounding, we co-locate satellite measurements within 24 hours, 2° latitude, and 1113 km longitude (equal to 10° at the equator). If multiple co-locations meet these criteria, we select the profile that minimizes the distance from the sounding. We calculate the median absolute difference and percent difference over the data set of co-locations between each satellite instrument and the frost point measurements.

We additionally compare with the Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) Stratospheric Composition Reanalysis of Aura MLS (M2-SCREAM) reanalysis product (Wargan et al., 2023), which assimilates MLS products including H_2O, as this guarantees a co-location for all OMPS LP measurements. To assess whether our methodology blindly memorizes the days it sees during training (where co-locations with MLS are frequent) or if it learns to generalize to days it does not see during training (where co-locations with MLS are less frequent), we compute the mean differences and standard deviation of the differences between the LP product and M2-SCREAM for two subsets of 2021 data:

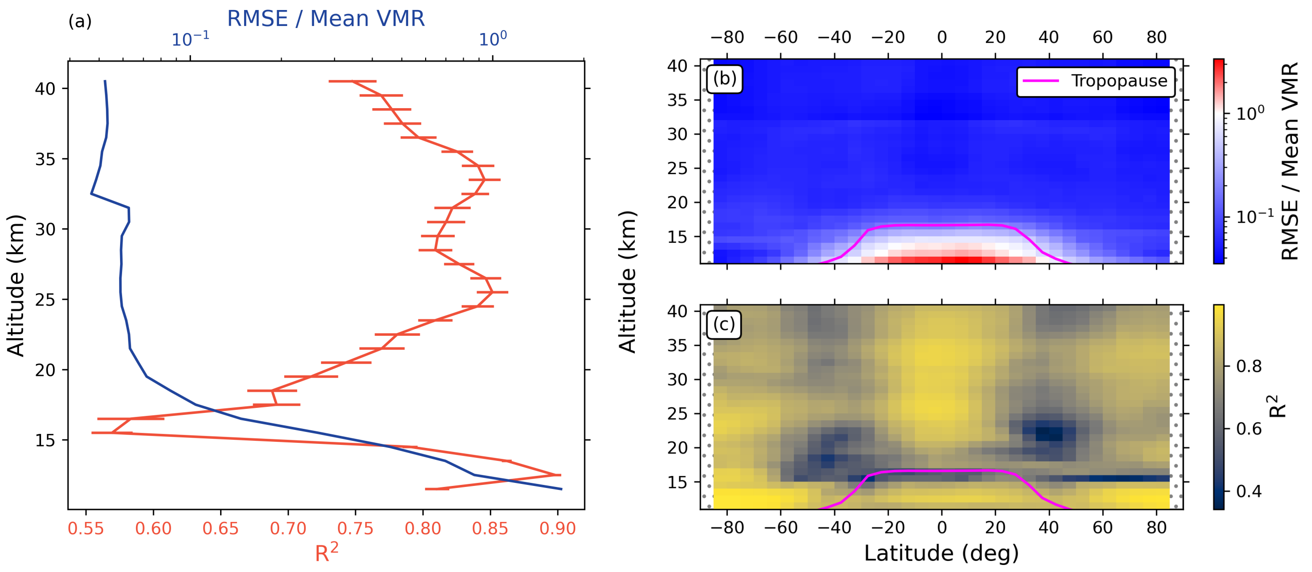

Figure 2. Performance summary for the ensemble of neural networks when applied to the test set. The relative RMSE (RMSE divided by the mean H_2O profile in the training set) is used to better highlight how the errors compare to the average H_2O VMRs at each altitude. Panel **(a)** shows the performance metrics calculated over the full test set. Error bars denote the standard deviation among the NNs' performances. Panels **(b)** and **(c)** show the performance metrics calculated as a function of latitude. Below the tropopause, the RMSE is on the order of or larger than the average H_2O VMR, while in the stratosphere the RMSE is typically <10% of the average H_2O VMR.

one that contains the days where training data were drawn from, and another that contains the days where no data were used during training.

~~Finally, given that OMPS LP is onboard the~~ NOAA-21 OMPS LP has an insufficient period overlapping with MLS measurements (2023–present, with MLS only taking measurements for ∼6 days per month since May 2024), inhibiting the use of NOAA-21 ~~satellite and planned to launch onboard two additional satellites in the coming years, we apply~~ OMPS LP data for our NN training methodology. Given the imminent termination of Aura MLS and that the SNPP satellite will presumably cease operations before the end of NOAA-21 or the subsequent JPSS satellites, we test the application of our SNPP-trained model to NOAA-21 OMPS LP measurements to determine whether our model can generalize to future iterations of the same instrument, as Himes et al. (2025b) found this to be the case for OMPS LP aerosol retrievals.

4 Results and Discussion

4.1 Training on MLS data

Figure 2 shows the mean and standard deviation of the R^2 and RMSE metrics for the NN ensemble when applied to the test set of data not seen during training. R^2 is >0.7 except between 15.5–18.5 km. These altitudes probe the upper troposphere or lower stratosphere depending on latitude, and the LP product shows greater ~~errors~~ median differences when compared to MLS in the troposphere. The 15.5–16.5 km altitudes, which are consistently within the troposphere in the tropics, feature the lowest R^2 values, while 17.5–18.5 km, which are typically not within the troposphere in the tropics, yield an R^2 just below 0.7. When considering only events outside the tropics, R^2 for 15.5–16.5 km increases to $\sim$0.7, while considering only events within the tropics results in R^2 reducing to $\sim$0.55 for these altitudes. In the stratosphere, the RMSE is <10% of the H_2O VMR. Errors increase below 18.5 km as measurements increasingly occur in the troposphere where water vapor VMR increases substantially. The discontinuity in RMSE at 32.5 km is related to the discontinuity in MLS v5 a priori profiles (Millán et al., 2024).

~~We find that omitting specific years during training can be important~~ Table 1 summarizes the OMPS LP H_2O product characteristics. The along-track resolution is $\sim$1° for SNPP OMPS LP and $\sim$0.4° for ~~certain situations~~NOAA-21 OMPS LP. Given that the MLS H_2O product has a vertical resolution of 2.8–3.8 km over the pressure range we consider and that the LP vertical field of view is 1.2–1.5 km, the LP H_2O product's vertical resolution may be between 1.2 and 3.8 km. However, more analysis is needed to better estimate the LP H_2O vertical resolution and how it changes as a function of altitude and under different conditions. At present, we conservatively report a vertical resolution $\sim$3.8 km.

Figure 3 shows a comparison between the NN, MLS, and LP uncertainties for a co-located event in the tropics in January 2019. The NN uncertainty is smaller than the MLS uncertainty by a factor of $\sim$2.5 in the stratosphere, while in the troposphere the NN uncertainty is larger than the MLS uncertainty by a factor of $\sim$3. Consequently, the combined LP uncertainty is dominated by the NN in the troposphere and by MLS in the stratosphere. We show two cases for this combined LP uncertainty, comprising the lower (uncorrelated) and upper (correlated) bounds for the combined NN-MLS uncertainty. They differ by $\sim$0.1 ppm, which is negligible relative to the volume mixing ratios of typical water vapor profiles. The uncertainties reported within the LP H_2O product are calculated assuming no correlation (see Eq. 1), as the random and systematic errors of MLS are unlikely to be correlated with LP's systematic errors at the wavelengths considered.

The results of our experiments that omit certain years from the training set reveal that some years are more important than others for achieving a well-generalized model. When omitting 2015–2016, we find that the model is generally unaffected and still performs well during those years. However, when omitting ~~2024–present~~2024, we find that the model begins producing severely inaccurate predictions by March 2024. This behavior is likely explained by the difference between these considered periods: while 2015–2016 were ordinary years in terms of SWV, the continued ~~presence~~ evolution of elevated SWV from the Hunga eruption into 2024 is atypical and not represented by the data available during ~~training~~2014–2023. By including a small fraction of 2024 data during training, we find that the model continues performing accurately up to the present time of writing this manuscript. As the stratosphere returns to pre-Hunga conditions, we expect that the NNs will continue producing accurate retrievals of H_2O, but continued comparisons with other instruments designed to measure water vapor, such as ~~ACE~~ACE-FTS, will be critical to ensuring that accuracy.

Since the input radiances are at wavelengths affected by aerosols, we analyze the model errors as a function of aerosol extinction reported in the OMPS LP aerosol product. We note that a weak anti-correlation exists between the predicted H_2O VMR and the aerosol extinction at 675 nm in the lower stratosphere, though this is a real phenomenon rather than an artifact

Table 1. OMPS LP H_2O product characteristics

Parameter	Instrument/Altitude	Value
Along-track resolution	SNPP OMPS LP	$\sim$1° latitude
	NOAA-21 OMPS LP	$\sim$0.4° latitude
Vertical resolution		$\sim$3.8 km
Precision, 0°–5°N	11.5 km	11.6 ppmv
	16.5 km	0.324 ppmv
	21.5 km	0.292 ppmv
	26.5 km	0.297 ppmv
	31.5 km	0.311 ppmv
	36.5 km	0.346 ppmv
Precision, 55°–60°N	11.5 km	0.754 ppmv
	16.5 km	0.322 ppmv
	21.5 km	0.316 ppmv
	26.5 km	0.325 ppmv
	31.5 km	0.344 ppmv
	36.5 km	0.377 ppmv

Notes. The reported vertical resolution is a conservative estimate based on the reported vertical resolution for MLS; there is not currently a per-altitude estimate. The reported precision values are the average uncertainties in the specified latitudinal bands.

of our model. Stratospheric aerosol extinction generally peaks immediately above the tropopause and decreases over the few kilometers above it, while stratospheric H_2O VMR typically is at a minimum immediately above the tropopause due to the cold trap and increases over the few kilometers above it. We find that the NNs' percentage error as a function of aerosol extinction at 675 nm is uncorrelated, indicating that the success of our approach is not dependent on aerosol conditions. See Appendix C for additional details.

In general, the ~~error~~ differences with respect to MLS ~~is~~ are independent of the presence of ~~tropospheric clouds~~upper tropospheric clouds identified by the OMPS LP measurements (Chen et al., 2016). However, events affected by polar stratospheric clouds (PSCs) identified by OMPS LP show a median bias around -2% at most altitudes. When considering the error as a function of distance below the PSC, the median bias can exceed -20% at 17 km below the PSC, which only occurs for PSCs at $\geq$28.5 km. However, the standard deviation of these errors can be substantial, where it averages around 33% between 15.5–22.5 km, with a maximum of 66% at 22.5 km. Given this behavior, we recommend that events contaminated by PSCs should not be used for scientific studies; the data product includes quality flags for these events.

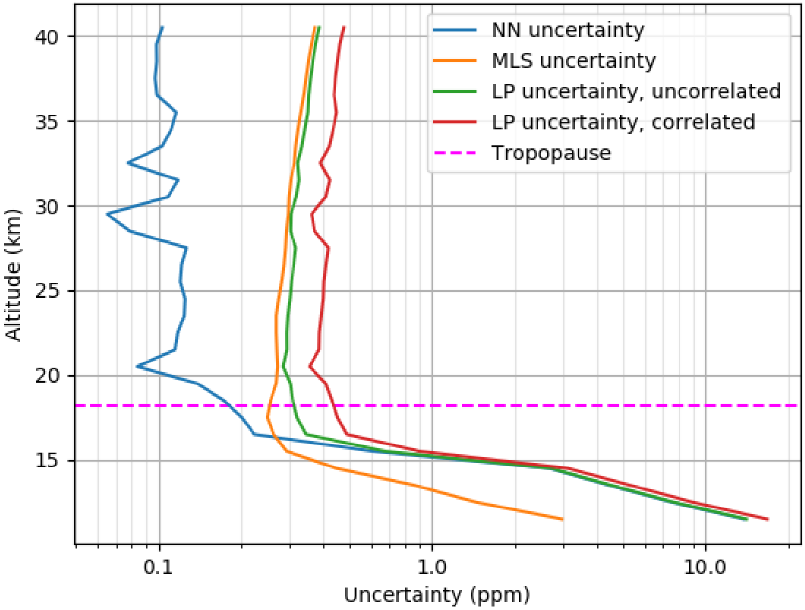

Figure 3. Comparison of uncertainties for a co-located event in the tropics in January 2019. In the troposphere, the NN uncertainty is larger than the corresponding MLS uncertainty by a factor of ∼3, while in the stratosphere the MLS uncertainty is larger than the NN uncertainty by a factor of ∼2.5. The LP uncertainties are a combination of these NN and MLS uncertainties; the uncorrelated and correlated cases differ by ∼0.1 ppm.

4.2 Training on ~~ACE~~ ACE-FTS and SAGE III/ISS data

We also considered training on co-located ~~ACE~~ ACE-FTS and/or SAGE III/ISS data as an alternative to the co-located MLS data set. Results for these experiments were ~~negative~~not satisfactory.

When training exclusively on ~~ACE~~ ACE-FTS or SAGE III/ISS data, we find that the resulting NN models do not properly generalize to unseen data, indicating that these data sets are not sufficient to solve this problem. Part of this may be explained by the typically significant measurement time differences between LP and the other instruments; LP measures in the early afternoon, while ~~ACE and SAGE~~ ACE-FTS and SAGE III/ISS measure at sunrise/sunset, and these times only coincide for select geolocations depending on the time of year. However, the number of co-locations seems to be a more limiting factor. When restricting the MLS data set to similar sizes as the ~~ACE and SAGE~~ ACE-FTS and SAGE III/ISS data sets, we find that the resulting performance is poor. Given the success when using the MLS-LP data set of ∼1 million co-locations but the failure when considering tens of thousands of co-locations, these results emphasize that our methodology relies on a large data set of co-located profiles.

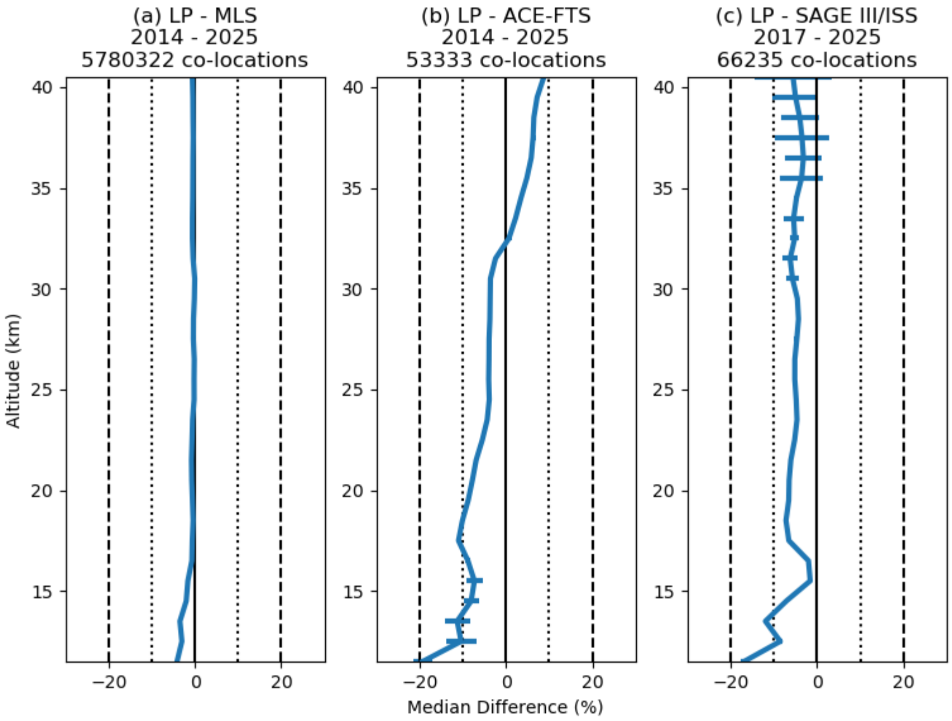

Figure 4. Summary of global median percent differences between the LP stratospheric H_2O profiles and co-located **(a)** MLS, **(b)** ~~ACE~~ACE-FTS, and **(c)** SAGE III/ISS profiles. Horizontal uncertainty bars indicate the standard error of the median.

Additionally, when including ~~ACE~~ ACE-FTS and/or SAGE III/ISS data alongside the MLS data, we find significantly degraded performance, regardless of whether or not the ~~ACE and SAGE~~ ACE-FTS and SAGE III/ISS data were de-biased with respect to MLS. This is likely attributable to the variances between MLS, ~~ACE~~ACE-FTS, and SAGE III/ISS being on the order of the natural variability of water vapor, which inhibits NN learning.

Despite our negative results when training on ~~ACE~~ ACE-FTS and/or SAGE III/ISS data, we cannot rule out that alternative ML approaches not considered here could utilize ~~ACE~~ ACE-FTS and/or SAGE III/ISS data to derive water vapor profiles from LP radiances.

4.3 Comparisons with satellite measurements

Figure 4 summarizes the global median percent differences between LP stratospheric H_2O profiles and co-located MLS, ~~ACE~~ACE-FTS, and SAGE III/ISS profiles. For this comparison, we filter all LP tropospheric measurements by using the nearest co-located tropopause altitude reported in the GEOS FP-IT product. The error bars show the standard error of the median, which is generally negligible except at low altitudes for ~~ACE~~ ACE-FTS and at high altitudes for SAGE III/ISS. Where LP detects a cloud, we exclude any measurements at or below the cloud top, though we find that this criterion does not significantly

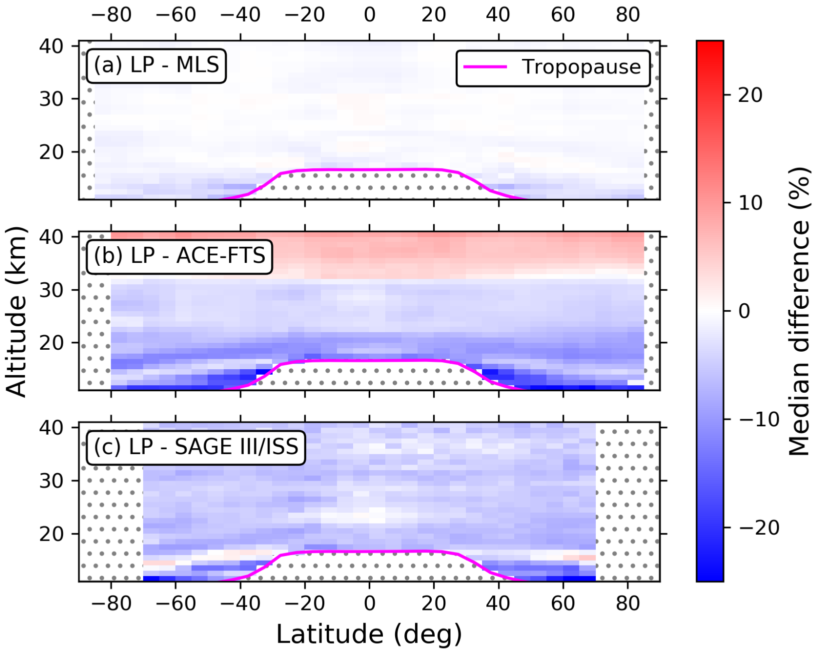

Figure 5. Summary of median percent differences between the LP stratospheric H_2O profiles and co-located **(a)** MLS, **(b)** ~~ACE~~ACE-FTS, and **(c)** SAGE III/ISS profiles in 5° latitudinal bins. The pink line indicates the median tropopause altitude for a given latitudinal bin. Gray stippling indicates where there are no data, whether due to lack of statistical significance (high latitudes) or due to being in the troposphere (below the tropopause).

alter the results. Note that for the MLS comparisons we include all co-located data, including those used during training; this choice does not bias the results, as discussed later in Section 4.5.

Differences with respect to MLS are less than 2% at all altitudes ≥14.5 km, with a maximum difference of 4.1% at 11.5 km. When considering only 2025 data, the most extreme difference is 7.7% at 13.5 km, with a typical difference of ∼5% below 22 km and <2% above 25 km. Given that we trained on MLS data, this close agreement is expected and shows the model has learned a good approximation to retrieve MLS-like water vapor profiles.

When comparing with ~~ACE~~ACE-FTS, we find agreement within 10% except between 11.5–13.5 km, where differences can reach up to 19.3%. In general, the differences increase with altitude from -19.3% at 11.5 km up to 8.8% at 40.5 km. ~~This behavior is generally consistent with earlier studies, such as Davis et al. (2021) which shows a similar pattern of increasing differences between 15–40 km when comparing SAGE and ACE.~~

LP's differences with respect to SAGE III/ISS are generally around 6% or less. Between 11.5–13.5 km, differences can reach up to 16.9%. Given that a similar increase in differences at these altitudes is seen when comparing with both ~~ACE and SAGE~~ ACE-FTS and SAGE III/ISS but not with MLS, this suggests that either ~~ACE and SAGE~~ ACE-FTS and SAGE III/ISS are biased high in this regime, or MLS is biased low in this regime and ~~our~~ the LP product has inherited this bias.

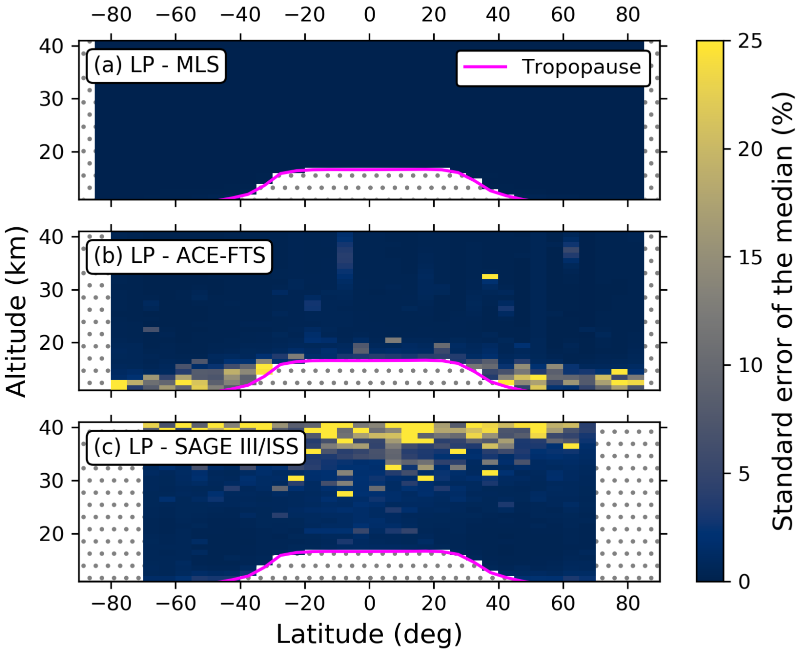

Figure 6. Like Figure 5, but for the standard error of the median.

Figure 5 and Figure 6 show the median percent differences and standard error of the median, respectively, in 5° latitudinal bins for the comparisons with MLS, ~~ACE~~ACE-FTS, and SAGE III/ISS. The results are generally consistent with those shown in Figure 4. The notable exceptions occur at the lowest altitudes. For the comparisons with MLS, differences can reach up to 8% between 11.5 - 13.5 km just outside of the tropics. For ~~ACE~~ACE-FTS, the deviations at these altitudes can exceed 21%, but notably the standard error of the median is typically $\sim$11% in this region, suggesting that this low bias may not be as substantial as it appears. However, comparisons with SAGE III/ISS also show this low bias at these altitudes, where the standard error is negligible. Together, this suggests that the LP product has a slight systematic low bias at 11.5–13.5 km just outside the tropics, but at latitudes $\geq$45°, the agreement with MLS suggests the biases when comparing with ~~ACE and SAGE~~ ACE-FTS and SAGE III/ISS are related to statistical differences between those products and MLS.

Figure 7 shows the “tape recorder” of alternating positive and negative anomalies in H_2O VMR, primarily attributable to seasonal changes in H_2O. For this plot, we subtract the pre-Hunga mean profile from each daily zonal mean to produce the daily anomaly. Before 2025, the OMPS LP and MLS tape recorders show excellent agreement throughout the stratosphere, with OMPS LP correctly capturing the increase in H_2O due to the Hunga eruption. ~~Beginning~~ For $\sim$3 months in early 2025, OMPS LP shows a positive bias >~~1~~ 1.0 ppm above 30 km that is not seen in the corresponding MLS data (Figure 7c). This bias is likely due to the weak H_2O sensitivity at these altitudes, which ~~inhibit~~ inhibits the NNs’ ability to reliably infer the H_2O VMR at these altitudes. Since the NNs correctly infer the H_2O at these altitudes before 2025, it suggests that the pre-2025 data were successfully predicted based on the shape of profiles at the lower altitudes that have sensitivity. With an absence of

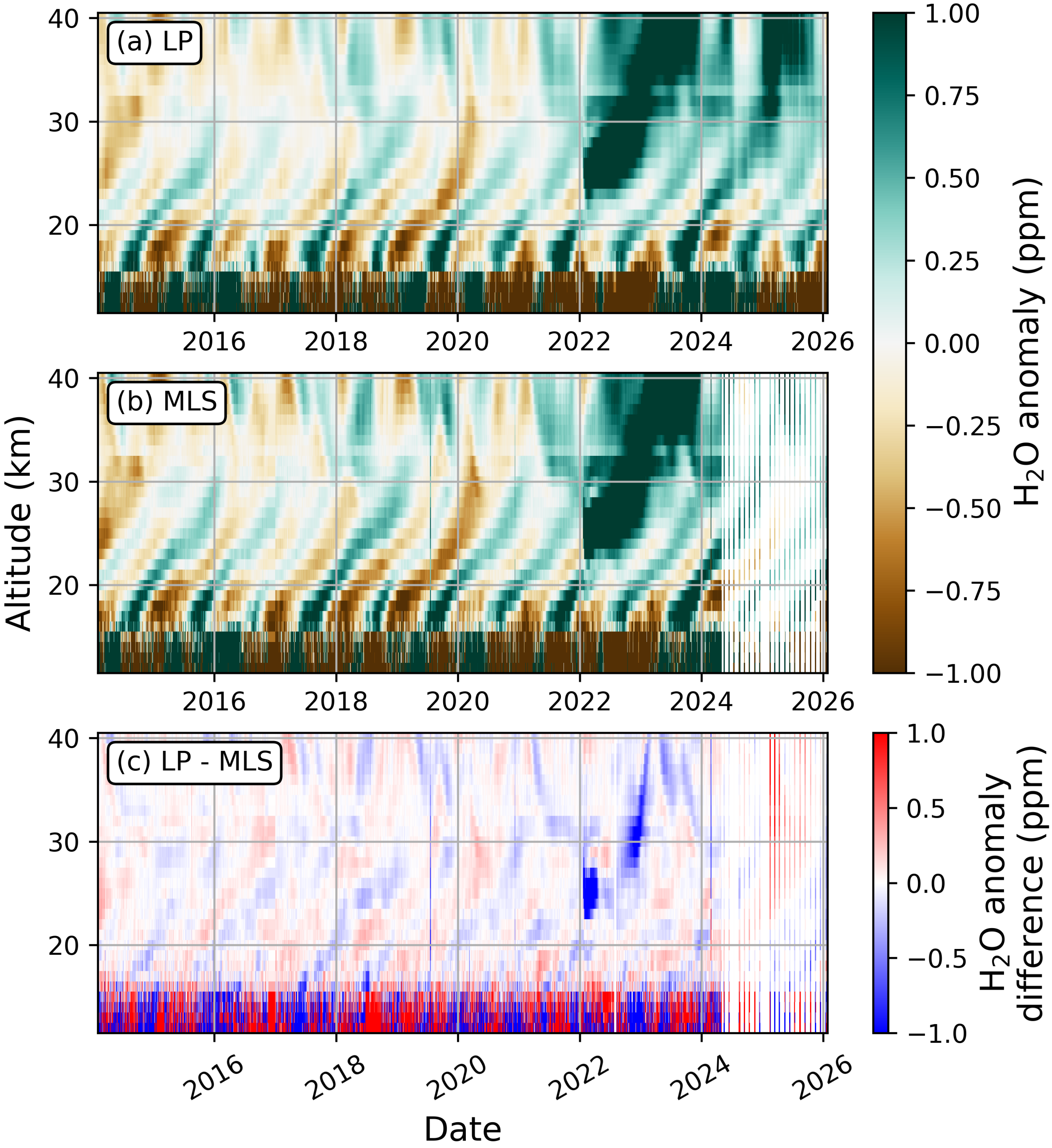

Figure 7. Parts-per-million anomaly in H_2O VMRs for the daily zonal means within 2.5° of the Equator for the **(a)** OMPS LP and **(b)** MLS water vapor products. Panel **(c)** shows the difference between the LP and MLS H_2O anomalies. The anomaly is determined by subtracting the pre-Hunga mean profile from each daily zonal mean profile. Beginning in May 2024, the MLS data become more sparse due to only taking measurements around 6 days each month.

2025 data in ~~training, they guess based on similar~~ the training set, the NN model relies on profiles from the training ~~set, which are evidently~~ data, primarily those influenced by ~~the~~ elevated H_2O ~~from Hunga~~ associated with the Hunga eruption. In 2025, the NNs perform reasonably well below 30 km, ~~indicating~~ suggesting that there is sufficient sensitivity for the determined approximation to remain accurate when applied to unseen data, though there is a slight overestimation ($\sim$0.25 ppm) in mid-2025 between 25–30 km. We therefore advise that users exercise caution when using the OMPS LP H_2O product above 30 km ~~.~~ in the tropics. Additionally, as the excess SWV from Hunga continues to evolve, it is possible that those conditions will be different enough from the training data that the NNs' predictions become inaccurate; we will continue to monitor their performance to ensure the LP NN-based algorithm continues to provide reasonable H_2O profiles.

4.4 Multiple linear regression analysis

In general, the results of our MLR analysis show similar behavior between the LP and MLS products, but the fitted coefficients for the LP data tend to be less than the corresponding MLS coefficients. The main exception to this is the seasonal phase offset, where both products closely agree. Figure 8 shows an example of the fitted coefficients for the seasonal amplitude at 14.5 km. The South Asian monsoon stands out clearly in both panels, though the LP product's fitted amplitudes for this region are around 1 ppm less than the corresponding fits for MLS.

Regarding water vapor trends, the LP product generally shows ~~greater trends in the troposphere and~~ weaker trends in the stratosphere when compared with MLS. In some locations (particularly south and southeast Asia, central Africa, and central America), the LP product shows greater trends in the upper troposphere. Where both products show a trend of increasing H_2O, LP tends to show a lesser trend than MLS (Figure 9). ~~As MLS v5 is known to still contain some statistically significant drifts in the lower stratosphere when compared with balloon measurements (Livesey et al., 2021), our results suggest that the NN methodology reduces these drifts.~~

4.5 Comparisons with M2-SCREAM

Figure 10 compares the OMPS LP and M2-SCREAM H_2O products for the year 2021. Note that M2-SCREAM assimilates MLS v4.2, which is biased high for water vapor, while LP is trained on MLS v5, resulting in a persistent $\sim$0.5 ppmv bias between the products. Days in which data were (Fig. 10a) or were not (Fig. 10b) included during training are shown separately but look almost identical, highlighting that the NNs' predictions are equally accurate whether or not they saw data from that day during training. Additionally, the standard deviation of the differences between OMPS LP and M2-SCREAM are consistently less than the standard deviation among OMPS LP or M2-SCREAM profiles, indicating that the OMPS LP product is more accurate than the natural variability of H_2O. Overall, our results suggest that the NN predictions are in good agreement with M2-SCREAM for data not seen during training.

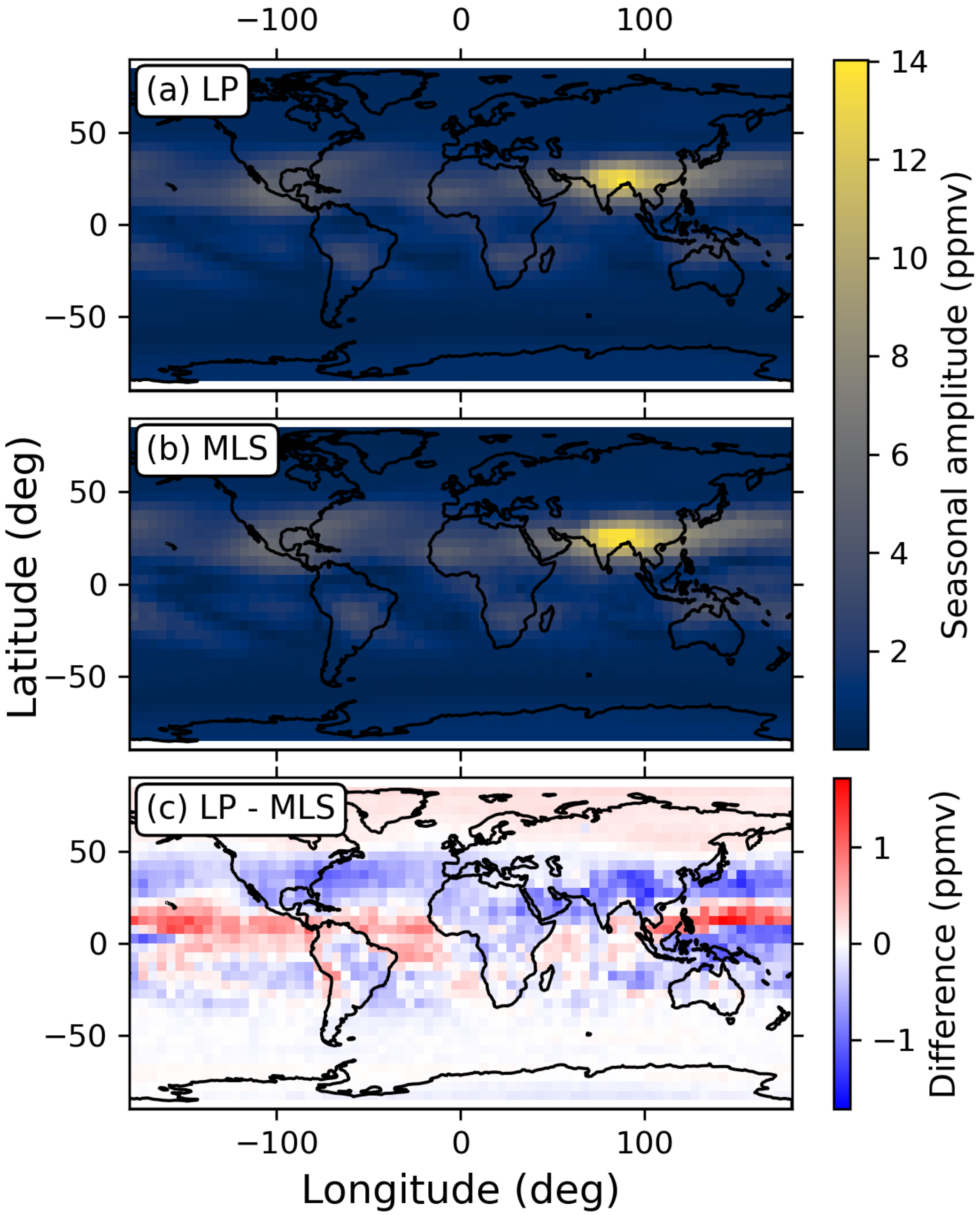

Figure 8. Results for the seasonal amplitude at 14.5 km fitted via MLR for the **(a)** LP and **(b)** MLS water vapor products. The large amplitude over South Asia is attributable to the annual monsoon's strong seasonal impact on water vapor. Panel **(c)** shows the difference between the fitted seasonal amplitude for the LP and MLS products.

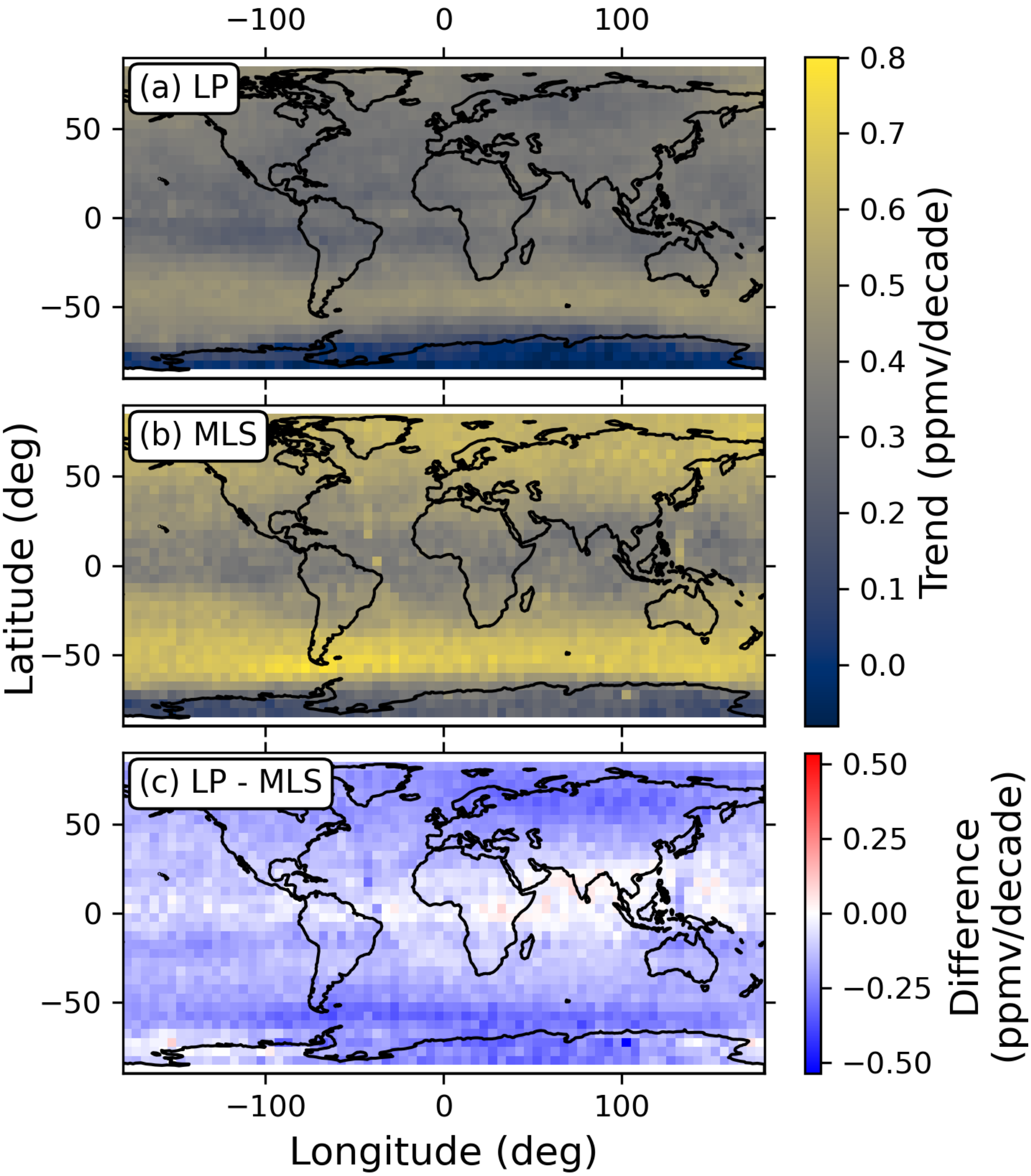

Figure 9. Results for the linear trend in H_2O at 18.5 km fitted via MLR for the **(a)** LP and **(b)** MLS water vapor products. In general, LP shows a weaker trend of increasing H_2O than MLS.

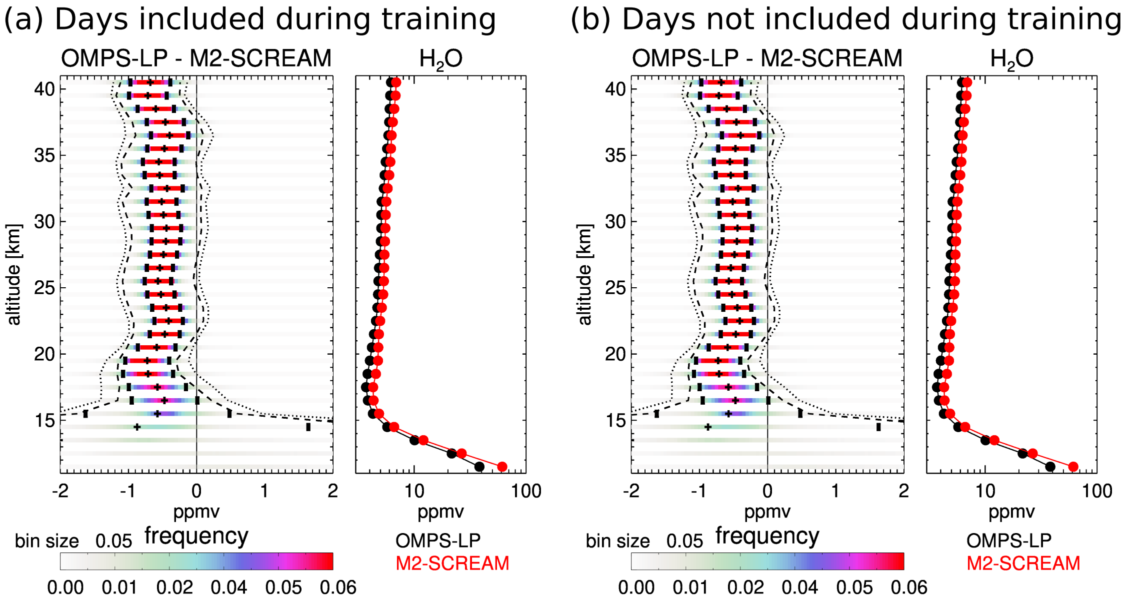

Figure 10. Comparison between OMPS LP and M2-SCREAM for days in 2021 where **(a)** some data were included during training and **(b)** no data were included during training. Note that for days in which data were included during training, only a small percentage (1–4%) of data on those days were used in the training data set. For each panel, the left subpanel shows the probability density function of the differences between OMPS LP and M2-SCREAM as horizontal colored bars, one standard deviation of the differences as the vertical marks, the mean difference as plus signs, the mean difference ± the standard deviation of OMPS LP H_2O as dashed lines, and the mean difference ± the standard deviation of M2-SCREAM as dotted lines. The right subpanel shows the mean H_2O profile for each product. Panels **(a)** and **(b)** look nearly identical, indicating that the model is retrieving H_2O from the information content embedded in LP radiances rather than blindly memorizing the training data. The ∼0.5 ppmv offset between the products is due to differences in the MLS version used by each product.

4.6 Comparisons with balloon-borne measurements

Figure 11 shows the median differences between satellite instruments (MLS, SAGE III/ISS, ACE-FTS, ~~ACE,~~ and OMPS LP) and the frost point hygrometer soundings from the six stations considered (see Section 3.3). The LP product agrees with the frost point measurements within 0.3 ppmv and within 10% between 16.5–27.5 km; this is in close agreement with the MLS results, which is expected given that we trained on MLS data. The only notable difference in agreement between MLS and LP is that the LP product shows a slightly reduced bias between 16.5–21.5 km. Like in Davis et al. (2021), we find that the satellite instruments show a dry bias in the upper troposphere compared to the frost point measurements, which may be due to spatiotemporal variability between the co-located measurements and/or reduced data quality in this regime.

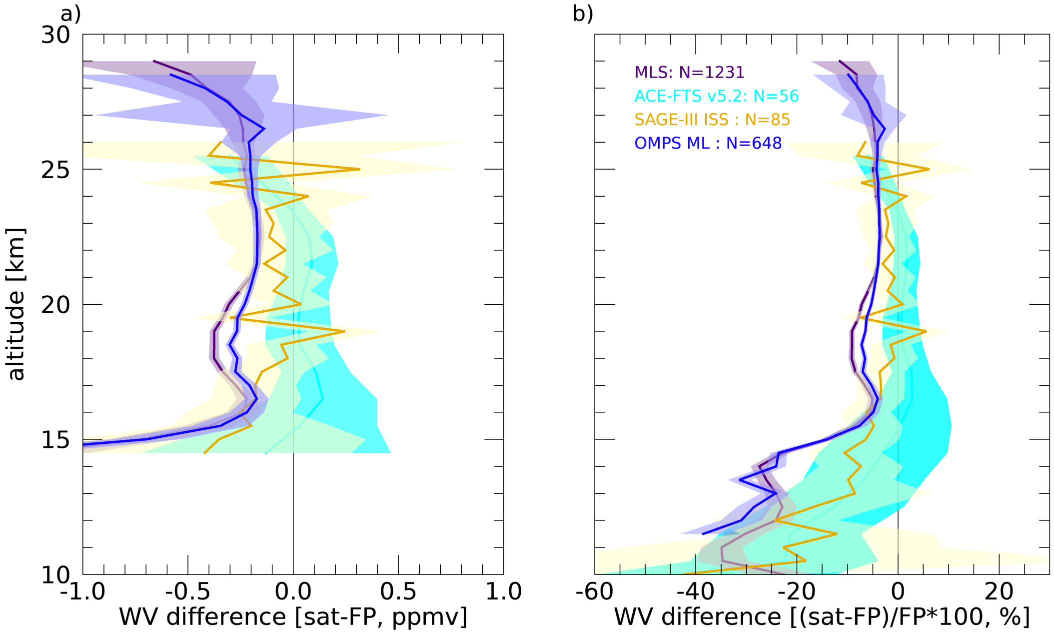

Figure 11. Comparisons between satellite instruments and frost point balloon measurements. Panel **(a)** shows the median difference in ppmv limited to altitudes where the mean water vapor is <10 ppmv, and panel **(b)** shows the median percent difference. In both panels, the shaded regions show the median ± 2 standard errors of the mean.

4.7 Application to NOAA-21 OMPS LP

Paralleling the SNPP comparisons in Section 4.3, Figures 12 and 13 show respectively the global median percent differences and the 5° zonal median percent differences between NOAA-21 OMPS LP and MLS, ~~ACE~~ACE-FTS, and SAGE III/ISS. For global comparisons, NOAA-21 OMPS LP shows a persistent ∼5% offset with respect to the corresponding SNPP comparisons at all altitudes. A similar low bias is seen in the NOAA-21 OMPS LP aerosol data, suggesting that this bias is attributable to differences in radiances between OMPS LP on SNPP and NOAA-21. Given our methodology, it is unclear whether the NNs have learned to implicitly account for a bias in the SNPP radiances or if the problem is related to the calibration of NOAA-21 radiances. However, this bias is not strictly a -5% shift for all conditions; the zonal comparisons show that the tropics exhibit a positive bias not seen in the corresponding SNPP comparisons. Further investigation is necessary to understand the cause of these biases. If the origin of these biases is not able to be determined, they could be addressed via a soft calibration approach.

5 Conclusions

We presented a water vapor retrieval product derived from SNPP OMPS LP measurements via a neural network (NN) trained on co-located MLS version 5 water vapor profiles. In general, the LP H_2O product is consistent with other water vapor products

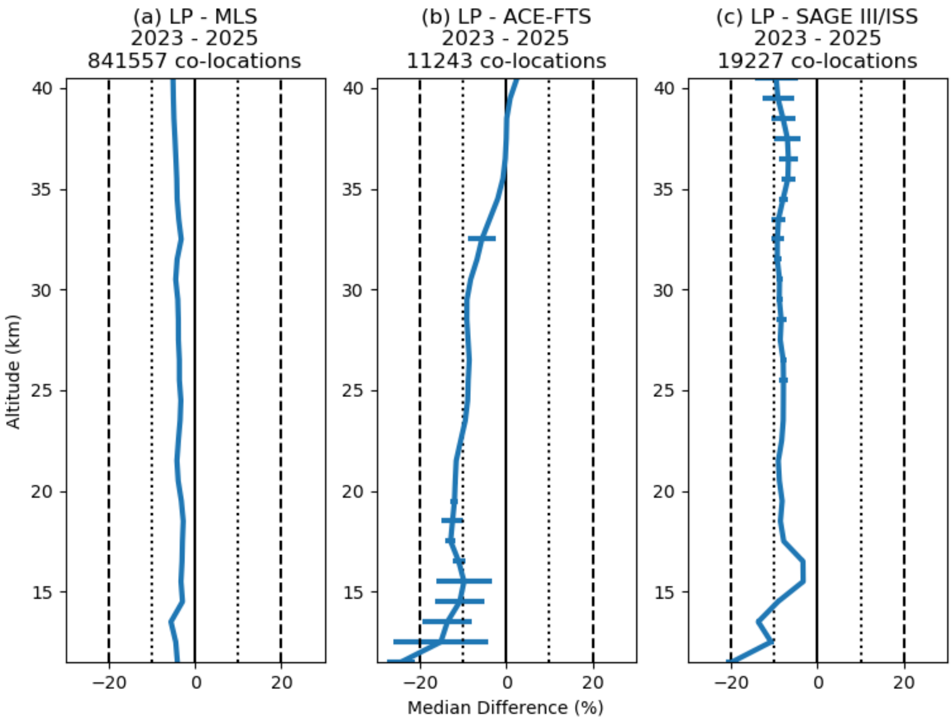

Figure 12. Like Figure 4, but for NOAA-21 OMPS LP.

considered here. We find that our method typically agrees with MLS within 5% at all altitudes considered. The results of our multiple linear regression analysis show good correspondence between LP and MLS for seasonal water vapor variations, including for the south Asian monsoon. LP's tape recorder in the tropics also shows close agreement with MLS, capturing both the alternating positive and negative seasonal anomalies as well as the large water vapor injection from the Hunga eruption. Agreement with SAGE III/ISS version 6 and ~~ACE~~ ACE-FTS version 5.3 water vapor profiles is typically within 10% above 15 km and within 20% below 15 km. When compared with frost point balloon measurements, OMPS LP generally agrees within 10% in the stratosphere, closely mirroring comparisons between those frost point measurements and MLS. Comparisons with the M2-SCREAM reanalysis product show similar behavior between days included in training and days omitted from training, indicating that our method is retrieving H_2O from the LP radiances rather than memorizing the training data. Overall, we find that the LP product generally performs comparably to MLS over the 11.5–40.5 km altitude range considered, enabling the continuation of the MLS water vapor record for these altitudes. The exception is in the tropics above 30 km, where a period in 2025 shows larger errors relative to MLS than those seen during the 2014–2024 training period; we advise that users exercise caution when using LP H_2O data in this regime.

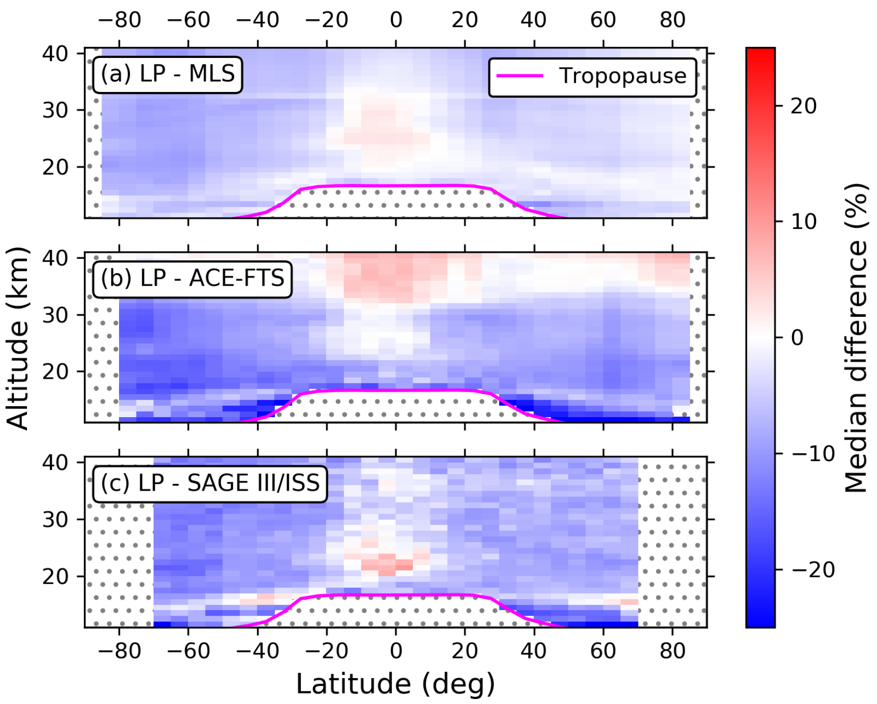

Figure 13. Like Figure 5, but for NOAA-21 OMPS LP.

When applying the same methodology but using SAGE III/ISS and/or ~~ACE~~ ACE-FTS data for training, we find significantly reduced performance. We similarly find poor performance when limiting the MLS-LP data set to the same size as the SAGE ~~and ACE~~ III/ISS and ACE-FTS data sets, which suggests that the success of our approach relies on a large training data set of co-located profiles. When including SAGE III/ISS and/or ~~ACE~~ ACE-FTS data alongside MLS data, we also find poor performance, regardless of whether or not the SAGE III~~and ACE~~ /ISS and ACE-FTS data are bias corrected to have a 0% median difference with MLS at all altitudes. This suggests that the variances between the three satellite data products inhibit NN learning.

Despite insufficient co-located data to train a well generalized model specific for NOAA-21 OMPS LP, we find that the SNPP-trained NN is applicable to NOAA-21 OMPS LP measurements without retraining. For NOAA-21 data, we find a persistent negative bias of ∼5% under most conditions when compared with the corresponding SNPP results; this pattern is also seen in comparisons between the SNPP and NOAA-21 OMPS LP aerosol products, suggesting that it is due to differences in the radiances rather than poor generalization of the SNPP-trained NN. However, the source of this bias is unclear at the time of writing; future work should explore approaches to identify the origin of this bias, characterize it, and correct it, if possible. Assuming that the NN model continues performing well over the coming years, our results suggest that this SNPP-trained model will be applicable to OMPS LP onboard JPSS-4 and 3, which are planned to launch in 2027 and 2032, respectively, thereby extending the MLS water vapor record into the 2030s, albeit at a reduced altitude range. Continued satellite and balloon-borne measurements from instruments with physics-based stratospheric H_2O products, such as ~~ACE~~ACE-FTS, SAGE III/ISS, and frost point hygrometer soundings, will be integral to ensuring that our NN-based retrievals continue to perform well in the coming years.

Code and data availability. The MARGE software is available on GitHub at https://github.com/exosports/MARGE (Himes, 2022). All data and results related to the MARGE software for this work are publicly available under the Reproducible Research Software License at https://doi.org/10.5281/zenodo.17237404 (Himes et al., 2025a). The SNPP OMPS LP version 1.0 H_2O data product is available at https://doi.org/10.5067/C1BD8BLEBH04 (Himes, 2025a). The NOAA-21 OMPS LP version 1.0 H_2O data product is available at https://doi.org/10.5067/XNK38X2VQGZ0 (Himes, 2025b). The SNPP OMPS LP version 2.6 L1G data product is available at https://doi.org/10.5067/YVE3FSNJ59RQ (Jaross, 2023).

Appendix A: Neural network optimization, architecture, and training

To optimize the neural network architecture for this problem, we performed a Bayesian hyperparameter optimization over the number and types of layers, number of nodes per layer, and activation functions. We considered fully connected NNs, convolutional NNs, and architectures that utilize both fully connected and convolutional layers. In addition to standard fully connected layers, we also considered Concrete Dropout layers (Gal et al., 2017), which include a trainable parameter for the layer's dropout rate. For convolutional architectures, we considered architectures with and without pooling layers.

The selected architecture is similar to the landmark AlexNet architecture (Krizhevsky et al., 2012), with hidden layers and activation functions consisting of Conv2D(32)–ReLU–MaxPool2D–Conv2D(64)–ReLU–MaxPool2D–CD(256)–ReLU–CD(256)–ReLU, where Conv2D(m) indicates a two-dimensional convolutional layer with m feature maps using a kernel size of 5, ReLU indicates the rectified linear unit activation function, MaxPool2D indicates a two-dimensional pooling layer that selects the maximum value within a 2×2 window, and CD(n) indicates a Concrete Dropout layer with n nodes. This is followed by a fully-connected output layer of 30 nodes, corresponding to the H_2O VMR at the 30 altitudes spanning 11.5–40.5 km. Note that other architectures performed similarly to the selected architecture; we found that the training data set played a more significant role in model performance.

We optimized the learning rate policy according to the method described by Himes et al. (2025b) and trained each NN using the mean-squared-error loss over the validation set until early stopping engaged after a patience of 60 epochs. On average, models trained for 638 epochs, which required an average of almost 7 hours to train using an Nvidia V100 graphics processing unit.

On our processing system, running our retrieval algorithm using the central processing unit requires around 12 and 16 seconds to process one SNPP and ~~N21~~ NOAA-21 orbit, respectively; specialized graphics processing units would reduce this runtime.

Appendix B: Importance of the neural network inputs

The NNs presented in the manuscript utilize a combination of LP radiances and solar zenith angles as well as GEOS-IT data for temperatures and pressures. To roughly assess the relative importance of these inputs, we conducted experiments where NNs were trained on (1) only the meteorological GEOS-IT data, (2) only the LP radiances, and (3) using pressure/temperature climatology in place of GEOS-IT data.

Figure B1 shows the NN ensemble's RMSE and R^2 values for the model presented in the manuscript alongside those performance metrics for NNs trained on (1) only the GEOS-IT meteorological data and (2) only LP radiances. The reduced performance of the NN trained on only GEOS-IT data indicates that the LP radiances enable the determination of a better solution for retrieving SWV. This performance difference can also be seen in Figure B2, which shows the tape recorder plot for the NN trained only on the GEOS-IT data. Notably, the GEOS-IT-only NN model features positive anomaly artifacts (especially in mid 2018), misses the SWV enhancement in January 2022 attributable to the Hunga eruption, and consistently shows larger deviations from MLS in 2025 than those seen in Figure 7. In contrast, the NNs trained on only LP radiances and on LP + climatological meteorological data result in almost the same performance as the NN ensemble presented in the main text of the manuscript. Given these results, we conclude that the LP radiance data are the primary source of information content for this problem.

Appendix C: Model insensitivity to aerosol extinction

The LP wavelengths used in this study are affected by aerosols. However, as discussed in Section 4.1, the NN model errors are generally uncorrelated with aerosol extinction. Figure C1 shows a 2-D histogram of LP-MLS differences vs. aerosol extinction at 675 nm. The vast majority of the probability density is found clustered around a difference of ~0%. At the largest extinctions (>0.01), there appears to be a slight negative bias, but this is generally related to the negative bias associated with PSCs discussed in Section 4.1. Figure C2 shows a time series of mean orbital LP-MLS differences, with three major eruptions over this time period marked for reference. The LP-MLS differences do not feature any notable deviations in the time following these eruptions, suggesting that variations in aerosol loading do not negatively impact the model performance.

Author contributions. MDH curated the data set used for the NNs; trained the NNs; developed the software that produces the data product; performed comparisons with MLS, SAGE III/ISS, and ACE-FTS; applied MLR to separate trend, seasonal, QBO, and ENSO contributions; and wrote the initial draft of the manuscript. NAK oversaw the project, provided guidance on the validation efforts, and provided advice on figures. KW performed the comparisons between the LP H_2O product and M2-SCREAM. SMD performed the comparisons between satellite instruments and balloon-borne measurements. GJ suggested the initial idea for the project and provided advice based on previous efforts to retrieve H_2O from OMPS LP. All co-authors advised ways to investigate and improve the performance of the method, reviewed the manuscript, and provided advice on the text and figures.

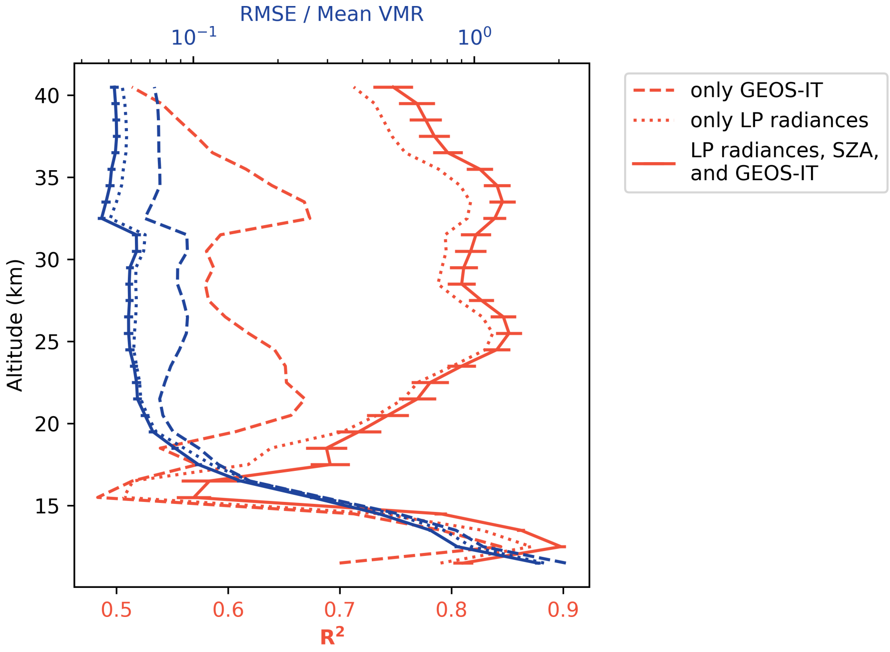

Figure B1. Like Figure 2, but also including results for NNs trained on (1) only GEOS-IT meteorology and (2) only LP radiances. The performance gap between the model trained on only GEOS-IT data and those trained with LP radiances indicates that the LP radiances offer additional information that enables a more accurate model.

Competing interests. At least one of the authors is a member of the editorial board of *Atmospheric Measurement Techniques*.

Acknowledgements. The authors would like to thank the OMPS LP characterization team for producing the Level-1 gridded data used in this work. We are grateful for Alyn Lambert, William Read, Luis Millan, Nathaniel Livesey, and all members of the Aura MLS water vapor team for their work on the valuable data product that made this work possible. We also thank the science teams responsible for the ~~ACE, SAGE~~ ACE-FTS, SAGE III/ISS, and frost point hygrometer data used for validation in this work. Additionally, we thank James Johnson and Daniel Kahn for feedback on the data product formatting and assisting with releasing the data product. We appreciate the thoughtful comments provided by the anonymous reviewers as they have improved the quality of this manuscript. We also recognize contributors to NumPy, SciPy, Matplotlib, TensorFlow, Keras, Optuna, Dask, the Python Programming Language, and the free and open-source community.

Financial support. This work was supported by the NASA Goddard Space Flight Center Earth Science Division's Strategic Science program.

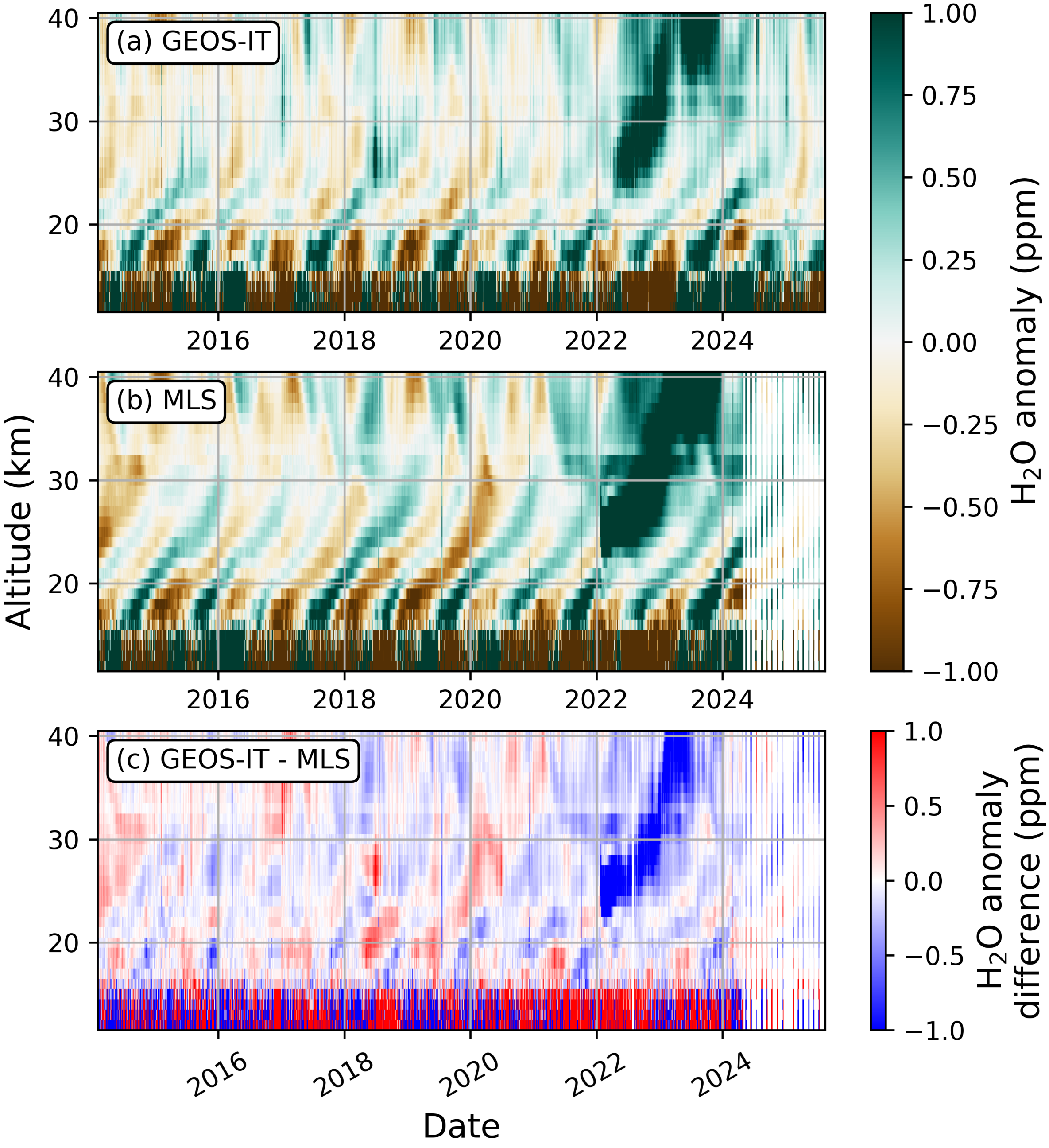

Figure B2. Like Figure 7, but for the NN trained on only the meteorological GEOS-IT data. The absence of LP radiance and solar zenith angle data results in reduced performance.

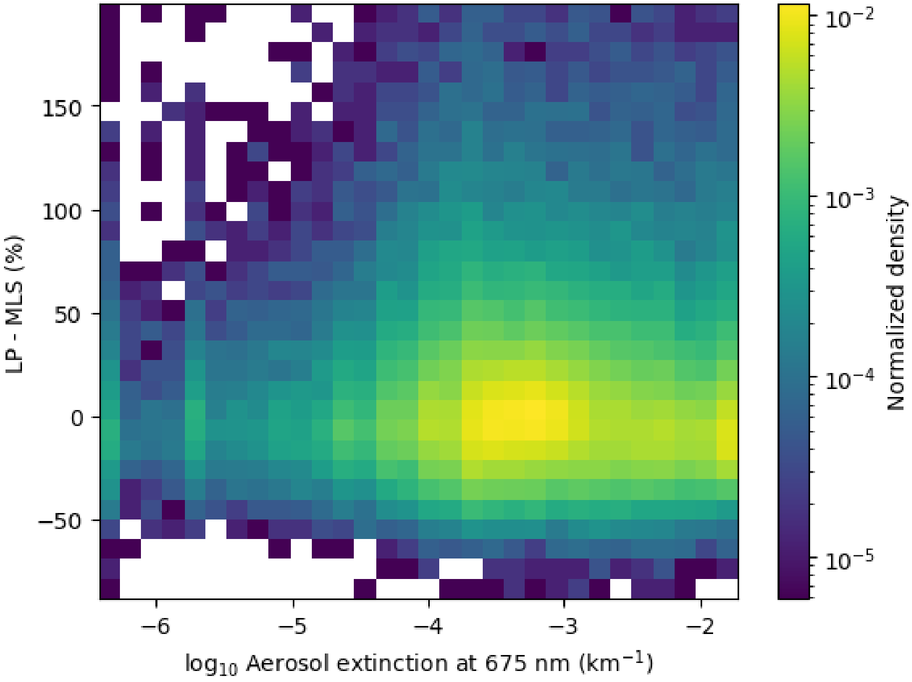

Figure C1. Histogram of the LP-MLS percent differences vs. OMPS LP aerosol extinction at 675 nm at 11.5 km. Errors are clustered around 0%, indicating aerosols do not significantly impact model performance.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al.: Tensorflow: a system for large-scale machine learning., in: OSDI, vol. 16, pp. 265–283, 2016.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M.: Optuna: A Next-Generation Hyperparameter Optimization Framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD ’19, p. 2623–2631, Association for Computing Machinery, New York, NY, USA, ISBN 9781450362016, https://doi.org/10.1145/3292500.3330701, 2019.

APARC: The Hunga Volcanic Eruption Atmospheric Impacts Report, Yunqian Zhu, Graham Mann, Paul A. Newman, and William Randel (Eds.), APARC Report No. 11, WCRP-10/2025, https://doi.org/10.34734/FZJ-2025-05237, 2025.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, Geophysical Research Letters, 32, https://doi.org/https://doi.org/10.1029/2005GL022386, 2005.

Boone, C., Bernath, P., and Lecours, M.: Version 5 retrievals for ACE-FTS and ACE-imagers, Journal of Quantitative Spectroscopy and Radiative Transfer, 310, 108 749, https://doi.org/https://doi.org/10.1016/j.jqsrt.2023.108749, 2023.

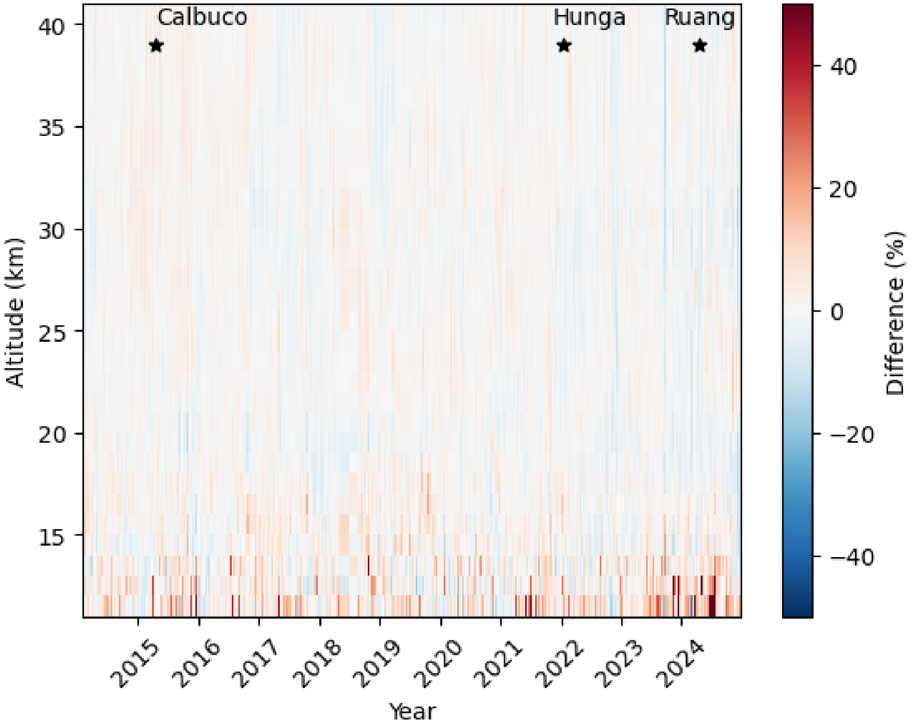

Figure C2. Time series plot of the average percentage difference between LP-MLS co-locations per LP orbit. Three major eruptions are denoted on the plot for context. Errors appear unrelated to those major eruptions.

Calvo, N., Garcia, R. R., Randel, W. J., and Marsh, D. R.: Dynamical Mechanism for the Increase in Tropical Upwelling in the Lowermost Tropical Stratosphere during Warm ENSO Events, Journal of the Atmospheric Sciences, 67, 2331 – 2340, https://doi.org/10.1175/2010JAS3433.1, 2010.

Carleer, M. R., Boone, C. D., Walker, K. A., Bernath, P. F., Strong, K., Sica, R. J., Randall, C. E., Vömel, H., Kar, J., Höpfner, M., Milz, M., von Clarmann, T., Kivi, R., Valverde-Canossa, J., Sioris, C. E., Izawa, M. R. M., Dupuy, E., McElroy, C. T., Drummond, J. R., Nowlan, C. R., Zou, J., Nichitiu, F., Lossow, S., Urban, J., Murtagh, D., and Dufour, D. G.: Validation of water vapour profiles from the Atmospheric Chemistry Experiment (ACE), Atmospheric Chemistry and Physics Discussions, 8, 4499–4559, https://doi.org/10.5194/acpd-8-4499-2008, 2008.

Charlesworth, E., Plöger, F., Birner, T., Baikhadzhaev, R., Abalos, M., Abraham, N. L., Akiyoshi, H., Bekki, S., Dennison, F., Jöckel, P., Keeble, J., Kinnison, D., Morgenstern, O., Plummer, D., Rozanov, E., Strode, S., Zeng, G., Egorova, T., and Riese, M.: Stratospheric water vapor affecting atmospheric circulation, Nature Communications, 14, 3925, https://doi.org/10.1038/s41467-023-39559-2, 2023.

Chen, Z., DeLand, M., and Bhartia, P. K.: A new algorithm for detecting cloud height using OMPS/LP measurements, Atmospheric Measurement Techniques, 9, 1239–1246, https://doi.org/10.5194/amt-9-1239-2016, 2016.

Cisewski, M., Zawodny, J., Gasbarre, J., Eckman, R., Topiwala, N., Rodriguez-Alvarez, O., Cheek, D., and Hall, S.: The Stratospheric Aerosol and Gas Experiment (SAGE III) on the International Space Station (ISS) Mission, in: Sensors, Systems, and Next-Generation

Satellites XVIII, edited by Meynart, R., Neeck, S. P., and Shimoda, H., vol. 9241 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 924107, https://doi.org/10.1117/12.2073131, 2014.

Dauhut, T., Chaboureau, J.-P., Escobar, J., and Mascart, P.: Giga-LES of Hector the Convector and Its Two Tallest Updrafts up to the Stratosphere, Journal of the Atmospheric Sciences, 73, 5041 – 5060, https://doi.org/10.1175/JAS-D-16-0083.1, 2016.

Davis, S. M., Hegglin, M. I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G. L., Nash, E. R., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP, Atmospheric Chemistry and Physics, 17, 12 743–12 778, https://doi.org/10.5194/acp-17-12743-2017, 2017.

Davis, S. M., Damadeo, R., Flittner, D., Rosenlof, K. H., Park, M., Randel, W. J., Hall, E. G., Huber, D., Hurst, D. F., Jordan, A. F., Kizer, S., Millan, L. F., Selkirk, H., Taha, G., Walker, K. A., and Vömel, H.: Validation of SAGE III/ISS Solar Water Vapor Data With Correlative Satellite and Balloon-Borne Measurements, Journal of Geophysical Research: Atmospheres, 126, e2020JD033 803, https://doi.org/https://doi.org/10.1029/2020JD033803, e2020JD033803 2020JD033803, 2021.

De Los Ríos, K., Ordoñez, P., Stiller, G. P., Raspollini, P., Gai, M., Walker, K. A., Peña Ortiz, C., and Acosta, L.: Comparison of the H_2O, HDO and δD stratospheric climatologies between the MIPAS-ESA V8, MIPAS-IMK V5 and ACE-FTS V4.1/4.2 satellite datasets, Atmospheric Measurement Techniques, 17, 3401–3418, https://doi.org/10.5194/amt-17-3401-2024, 2024.

Dittman, M. G., Leitch, J. W., Chrisp, M., Rodriguez, J. V., Sparks, A. L., McComas, B. K., Zaun, N. H., Frazier, D., Dixon, T., Philbrick, R. H., and Wasinger, D.: Limb broadband imaging spectrometer for the NPOESS Ozone Mapping and Profiler Suite (OMPS), in: Earth Observing Systems VII, edited by Barnes, W. L., vol. 4814, pp. 120 – 130, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.453751, 2002.

Fleming, E. L., Newman, P. A., Liang, Q., and Oman, L. D.: Stratospheric Temperature and Ozone Impacts of the Hunga Tonga-Hunga Ha'apai Water Vapor Injection, Journal of Geophysical Research (Atmospheres), 129, e2023JD039298, https://doi.org/10.1029/2023JD039298, 2024.

Fueglistaler, S., Dessler, A. E., Dunkerton, T. J., Folkins, I., Fu, Q., and Mote, P. W.: Tropical tropopause layer, Reviews of Geophysics, 47, https://doi.org/https://doi.org/10.1029/2008RG000267, 2009.

Gal, Y., Hron, J., and Kendall, A.: Concrete dropout, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, p. 3584–3593, Curran Associates Inc., Red Hook, NY, USA, ISBN 9781510860964, 2017.

Garcia, R. R., Marsh, D. R., Kinnison, D. E., Boville, B. A., and Sassi, F.: Simulation of secular trends in the middle atmosphere, 1950–2003, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2006JD007485, 2007.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, 2016.

Gordon, I., Rothman, L., Hargreaves, R., Hashemi, R., Karlovets, E., Skinner, F., Conway, E., Hill, C., Kochanov, R., Tan, Y., Wcisło, P., Finenko, A., Nelson, K., Bernath, P., Birk, M., Boudon, V., Campargue, A., Chance, K., Coustenis, A., Drouin, B., Flaud, J., Gamache, R., Hodges, J., Jacquemart, D., Mlawer, E., Nikitin, A., Perevalov, V., Rotger, M., Tennyson, J., Toon, G., Tran, H., Tyuterev, V., Adkins, E., Baker, A., Barbe, A., Canè, E., Császár, A., Dudaryonok, A., Egorov, O., Fleisher, A., Fleurbaey, H., Foltynowicz, A., Furtenbacher, T., Harrison, J., Hartmann, J., Horneman, V., Huang, X., Karman, T., Karns, J., Kassi, S., Kleiner, I., Kofman, V., Kwabia–Tchana, F., Lavrentieva, N., Lee, T., Long, D., Lukashevskaya, A., Lyulin, O., Makhnev, V., Matt, W., Massie, S., Melosso, M., Mikhailenko, S., Mondelain, D., Müller, H., Naumenko, O., Perrin, A., Polyansky, O., Raddaoui, E., Raston, P., Reed, Z., Rey, M., Richard, C., Tóbiás, R., Sadiek, I., Schwenke, D., Starikova, E., Sung, K., Tamassia, F., Tashkun, S., Vander Auwera, J., Vasilenko, I., Vigasin, A., Villanueva, G., Vispoel, B., Wagner, G., Yachmenev, A., and Yurchenko, S.: The HITRAN2020 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 277, 107 949, https://doi.org/https://doi.org/10.1016/j.jqsrt.2021.107949, 2022.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Himes, M. D.: MARGE, GitHub, https://github.com/exosports/MARGE (last access: 15 August 2025), 2022.

Himes, M. D.: OMPS-NPP LP L2 Water Vapor (H2O) Vertical Profile swath orbital 3slit V1.0, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) [data set], https://doi.org/10.5067/C1BD8BLEBH04, 2025a.

Himes, M. D.: OMPS-N21 LP L2 Water Vapor (H2O) Vertical Profile swath orbital 3slit V1.0, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) [data set], https://doi.org/10.5067/XNK38X2VQGZ0, 2025b.

Himes, M. D., Harrington, J., Cobb, A. D., Güneş Baydin, A., Soboczenski, F., O'Beirne, M. D., Zorzan, S., Wright, D. C., Scheffer, Z., Domagal-Goldman, S. D., and Arney, G. N.: Accurate Machine-learning Atmospheric Retrieval via a Neural-network Surrogate Model for Radiative Transfer, Planetary Science Journal, 3, 91, https://doi.org/10.3847/PSJ/abe3fd, 2022.

Himes, M. D., Kramarova, N. A., Wargan, K., Davis, S. M., and Jaross, G.: Research Compendium for Himes et al. (2025): "Continuing the MLS water vapor record with OMPS LP", Zenodo [data set], https://doi.org/10.5281/zenodo.17237404, 2025a.

Himes, M. D., Taha, G., Kahn, D., Zhu, T., and Kramarova, N. A.: Using neural networks for near-real-time aerosol retrievals from OMPS Limb Profiler measurements, Atmospheric Measurement Techniques, 18, 2523–2536, https://doi.org/10.5194/amt-18-2523-2025, 2025b.

Hurst, D. F., Oltmans, S. J., Vömel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A. F.: Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/https://doi.org/10.1029/2010JD015065, 2011.

Hurst, D. F., Lambert, A., Read, W. G., Davis, S. M., Rosenlof, K. H., Hall, E. G., Jordan, A. F., and Oltmans, S. J.: Validation of Aura Microwave Limb Sounder stratospheric water vapor measurements by the NOAA frost point hygrometer, Journal of Geophysical Research: Atmospheres, 119, 1612–1625, https://doi.org/https://doi.org/10.1002/2013JD020757, 2014.

Jaross, G.: OMPS-NPP L1G LP Radiance EV Wavelength-Altitude Grid swath orbital 3slit V2.6, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) [data set], https://doi.org/10.5067/YVE3FSNJ59RQ, 2023.

Khaykin, S., Pommereau, J.-P., Korshunov, L., Yushkov, V., Nielsen, J., Larsen, N., Christensen, T., Garnier, A., Lukyanov, A., and Williams, E.: Hydration of the lower stratosphere by ice crystal geysers over land convective systems, Atmospheric Chemistry and Physics, 9, 2275–2287, https://doi.org/10.5194/acp-9-2275-2009, 2009.

Khaykin, S., Podglajen, A., Ploeger, F., Grooß, J.-U., Tence, F., Bekki, S., Khlopenkov, K., Bedka, K., Rieger, L., Baron, A., Godin-Beekmann, S., Legras, B., Sellitto, P., Sakai, T., Barnes, J., Uchino, O., Morino, I., Nagai, T., Wing, R., Baumgarten, G., Gerding, M., Duflot, V., Payen, G., Jumelet, J., Querel, R., Liley, B., Bourassa, A., Clouser, B., Feofilov, A., Hauchecorne, A., and Ravetta, F.: Global perturbation of stratospheric water and aerosol burden by Hunga eruption, Communications Earth and Environment, 3, 316, https://doi.org/10.1038/s43247-022-00652-x, 2022.

Knowland, K. E., Wales, P. A., Wargan, K., Weir, B., Pawson, S., Damadeo, R., and Flittner, D.: Stratospheric Water Vapor Beyond NASA's Aura MLS: Assimilating SAGE III/ISS Profiles for a Continued Climate Record, Geophysical Research Letters, 52, e2024GL112 610, https://doi.org/https://doi.org/10.1029/2024GL112610, e2024GL112610 2024GL112610, 2025.

Kochanov, R. V., Gordon, I. E., Rothman, L. S., Wcisło, P., Hill, C., and Wilzewski, J. S.: HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data, Journal of Quantitative Spectroscopy and Radiative Transfer, 177, 15–30, https://doi.org/10.1016/j.jqsrt.2016.03.005, 2016.

Krizhevsky, A., Sutskever, I., and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105, 2012.

Lambert, A., Read, W., and Livesey, N.: MLS/Aura Level 2 water vapor (H2O) mixing ratio V005, Goddard Earth Sciences Data and Information Services Center (GES DISC) [data set], https://doi.org/10.5067/Aura/MLS/DATA2508, 2020.

Langille, J., Rieger, L. A., Blanchard, Y., Blanchet, J.-P., Bourassa, A., Degenstein, D., Huang, Y., Strong, K., Walker, K., Zawada, D., Braun, S., Cole, J., Mariani, Z., Mclinden, C., Paquin-Ricard, D., Sioris, C., Qu, Z., Wolde, M., Wang, X., Al-Abadleh, H. A., Ariya, P., Beltrami, H., Chang, R., Fletcher, C., Goldblatt, C., Grenier, P., Gyakum, J., Kushner, P., Luca, A. D., MacDougall, A. H., O'Neill, N., Pausata, F., Sica, R., Tan, I., Thériault, J. M., Tegtmeier, S., Toohey, M., Ward, W., and Wiacek, A.: The High-altitude Aerosols, Water vapour and Clouds mission: concept, scientific objectives and data products, Bulletin of the American Meteorological Society, pp. BAMS–D–23–0309.1, https://doi.org/10.1175/BAMS-D-23-0309.1, 2025.

Leitch, J. W., Rodriguez, J. V., Dittman, M. G., Frazier, D., McComas, B. K., Philbrick, R. H., Wasinger, D., Valle, T. J., Dixon, T., Dooley, D., Munzer, R., and Larsen, J. C.: Limb scatter ozone profiling sensor for the NPOESS ozone mapping and profiler suite (OMPS), in: Optical Remote Sensing of the Atmosphere and Clouds III, edited by Huang, H.-L., Lu, D., and Sasano, Y., vol. 4891 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 13–21, https://doi.org/10.1117/12.467527, 2003.

Livesey, N. J., Read, W. G., Froidevaux, L., Lambert, A., Santee, M. L., Schwartz, M. J., Millán, L. F., Jarnot, R. F., Wagner, P. A., Hurst, D. F., Walker, K. A., Sheese, P. E., and Nedoluha, G. E.: Investigation and amelioration of long-term instrumental drifts in water vapor and nitrous oxide measurements from the Aura Microwave Limb Sounder (MLS) and their implications for studies of variability and trends, Atmospheric Chemistry and Physics, 21, 15 409–15 430, https://doi.org/10.5194/acp-21-15409-2021, 2021.

Livesey, N. J., Read, W. G., Wagner, P. A., F. L., Santee, M. L., Schwartz, M. J., Lambert, A., Millan Valle, L. F., Pumphrey, H. C., Manney, G. L., Fuller, R. A., Jarnot, R. F., Knosp, B. W., and Lay, R. R.: Earth Observing System Aura Microwave Limb Sounder Version 5.0x Level 2 and 3 data quality and description document, Tech. Rep. JPL D-105336 Rev. B, Jet Propulsion Laboratory, https://mls.jpl.nasa.gov/data/v5-0_data_quality_document.pdf, 2022.

Loughman, R. P., Griffioen, E., Oikarinen, L., Postylyakov, O. V., Rozanov, A., Flittner, D. E., and Rault, D. F.: Comparison of radiative transfer models for limb-viewing scattered sunlight measurements, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2003JD003854, 2004.

Lucchesi, R.: File Specification for GEOS-5 FP-IT, Tech. Rep. GMAO Office Note No. 2, NASA Global Modeling and Assimilation Office, https://gmao.gsfc.nasa.gov/pubs/docs/Lucchesi865.pdf, 2015.

Millán, L., Santee, M. L., Lambert, A., Livesey, N. J., Werner, F., Schwartz, M. J., Pumphrey, H. C., Manney, G. L., Wang, Y., Su, H., Wu, L., Read, W. G., and Froidevaux, L.: The Hunga Tonga-Hunga Ha'apai Hydration of the Stratosphere, Geophysical Research Letters, 49, e2022GL099 381, https://doi.org/10.1029/2022GL099381, e2022GL099381 2022GL099381, 2022.

Millán, L., Read, W. G., Santee, M. L., Lambert, A., Manney, G. L., Neu, J. L., Pitts, M. C., Werner, F., Livesey, N. J., and Schwartz, M. J.: The Evolution of the Hunga Hydration in a Moistening Stratosphere, Geophysical Research Letters, 51, e2024GL110 841, https://doi.org/https://doi.org/10.1029/2024GL110841, e2024GL110841 2024GL110841, 2024.

NASA/LARC/SD/ASDC: SAGE III/ISS L2 Solar Event Species Profiles (HDF5) V006, https://doi.org/10.5067/ISS/SAGEIII/SOLAR_HDF5_L2-V6.0, 2025.

Niemeier, U., Wallis, S., Timmreck, C., van Pham, T., and von Savigny, C.: How the Hunga Tonga—Hunga Ha'apai Water Vapor Cloud Impacts Its Transport Through the Stratosphere: Dynamical and Radiative Effects, Geophysical Research Letters, 50, e2023GL106 482, https://doi.org/https://doi.org/10.1029/2023GL106482, e2023GL106482 2023GL106482, 2023.

Nowack, P., Ceppi, P., Davis, S. M., Chiodo, G., Ball, W., Diallo, M. A., Hassler, B., Jia, Y., Keeble, J., and Joshi, M.: Response of stratospheric water vapour to warming constrained by satellite observations, Nature Geoscience, 16, 577–583, 2023.

Park, M., Randel, W. J., Damadeo, R. P., Flittner, D. E., Davis, S. M., Rosenlof, K. H., Livesey, N., Lambert, A., and Read, W.: Near-Global Variability of Stratospheric Water Vapor Observed by SAGE III/ISS, Journal of Geophysical Research: Atmospheres, 126, e2020JD034 274, https://doi.org/https://doi.org/10.1029/2020JD034274, e2020JD034274 2020JD034274, 2021.

Ramanathan, V. and Inamdar, A.: The radiative forcing due to clouds and water vapor, p. 119–151, Cambridge University Press, 2006.

Randel, W. J. and Jensen, E. J.: Physical processes in the tropical tropopause layer and their roles in a changing climate, Nature Geoscience, 6, 169–176, 2013.

Read, W. G., Stiller, G., Lossow, S., Kiefer, M., Khosrawi, F., Hurst, D., Vömel, H., Rosenlof, K., Dinelli, B. M., Raspollini, P., Nedoluha, G. E., Gille, J. C., Kasai, Y., Eriksson, P., Sioris, C. E., Walker, K. A., Weigel, K., Burrows, J. P., and Rozanov, A.: The SPARC Water Vapor Assessment II: assessment of satellite measurements of upper tropospheric humidity, Atmospheric Measurement Techniques, 15, 3377–3400, https://doi.org/10.5194/amt-15-3377-2022, 2022.

Rong, P., Russell, J. M., Marshall, B. T., Gordley, L. L., Mlynczak, M. G., and Walker, K. A.: Validation of water vapor measured by SABER on the TIMED satellite, Journal of Atmospheric and Solar-Terrestrial Physics, 194, 105 099, https://doi.org/https://doi.org/10.1016/j.jastp.2019.105099, 2019.

Russell, III, J. M., Mlynczak, M. G., Gordley, L. L., Jr., J. J. T., and Esplin, R. W.: Overview of the SABER experiment and preliminary calibration results, in: Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, edited by Larar, A. M., vol. 3756, pp. 277 – 288, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.366382, 1999.

Salawitch, R. J., Smith, J. B., Selkirk, H., Wargan, K., Chipperfield, M. P., Hossaini, R., Levelt, P. F., Livesey, N. J., McBride, L. A., Millán, L. F., Moyer, E., Santee, M. L., Schoeberl, M. R., Solomon, S., Stone, K., and Worden, H. M.: The Imminent Data Desert: The Future of Stratospheric Monitoring in a Rapidly Changing World, Bulletin of the American Meteorological Society, 106, E540 – E563, https://doi.org/10.1175/BAMS-D-23-0281.1, 2025.

Vömel, H., Barnes, J. E., Forno, R. N., Fujiwara, M., Hasebe, F., Iwasaki, S., Kivi, R., Komala, N., Kyrö, E., Leblanc, T., Morel, B., Ogino, S.-Y., Read, W. G., Ryan, S. C., Saraspriya, S., Selkirk, H., Shiotani, M., Valverde Canossa, J., and Whiteman, D. N.: Validation of Aura Microwave Limb Sounder water vapor by balloon-borne Cryogenic Frost point Hygrometer measurements, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2007JD008698, 2007a.

Vömel, H., David, D. E., and Smith, K.: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/https://doi.org/10.1029/2006JD007224, 2007b.

Vömel, H., Evan, S., and Tully, M.: Water vapor injection into the stratosphere by Hunga Tonga-Hunga Ha'apai, Science, 377, 1444–1447, https://doi.org/10.1126/science.abq2299, 2022.

Wargan, K., Weir, B., Manney, G. L., Cohn, S. E., Knowland, K. E., Wales, P. A., and Livesey, N. J.: M2-SCREAM: A Stratospheric Composition Reanalysis of Aura MLS Data With MERRA-2 Transport, Earth and Space Science, 10, e2022EA002 632, https://doi.org/https://doi.org/10.1029/2022EA002632, e2022EA002632 2022EA002632, 2023.

Waters, J., Froidevaux, L., Harwood, R., Jarnot, R., Pickett, H., Read, W., Siegel, P., Cofield, R., Filipiak, M., Flower, D., Holden, J., Lau, G., Livesey, N., Manney, G., Pumphrey, H., Santee, M., Wu, D., Cuddy, D., Lay, R., Loo, M., Perun, V., Schwartz, M., Stek, P., Thurstans, R., Boyles, M., Chandra, K., Chavez, M., Chen, G.-S., Chudasama, B., Dodge, R., Fuller, R., Girard, M., Jiang, J., Jiang, Y., Knosp, B., LaBelle, R., Lam, J., Lee, K., Miller, D., Oswald, J., Patel, N., Pukala, D., Quintero, O., Scaff, D., Van Snyder, W., Tope, M., Wagner, P., and Walch, M.: The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1075–1092, https://doi.org/10.1109/TGRS.2006.873771, 2006.

Yu, W., Garcia, R., Yue, J., Russell III, J., and Mlynczak, M.: Variability of Water Vapor in the Tropical Middle Atmosphere Observed From Satellites and Interpreted Using SD-WACCM Simulations, Journal of Geophysical Research: Atmospheres, 127, e2022JD036 714, https://doi.org/https://doi.org/10.1029/2022JD036714, e2022JD036714 2022JD036714, 2022.

Yue, J., Russell III, J., Gan, Q., Wang, T., Rong, P., Garcia, R., and Mlynczak, M.: Increasing Water Vapor in the Stratosphere and Mesosphere After 2002, Geophysical Research Letters, 46, 13 452–13 460, https://doi.org/https://doi.org/10.1029/2019GL084973, 2019.