We thank the reviewer for providing additional thoughtful comments on the manuscript. Our response to each comment is provided below.

> *1. You partially addressed my previous concern about whether this method is truly retrieving water vapor. Your ablation tests are useful and support the importance of the radiance inputs in the machine-learning model. However, these tests are not equivalent to a direct channel importance or sensitivity analysis. The revision still does not provide a quantitative per-channel importance assessment, permutation test, or another physically interpretable attribution analysis. In particular, I do not see a direct response to my previous comments about the physical detectability of water vapor in the 930 -950 nm region, given that SAGE III/ISS retrievals in this spectral range are noisy even with much finer spectral resolution and solar occultation geometry, whereas OMPS-LP uses limb-scattered radiation with much coarser spectral resolution.*

Evidence of water vapor is visible in the OMPS LP measurements and can be confirmed via RTM simulations, particularly in the SH and tropics. Figure R3.1 shows altitude-normalized radiances at 16.5 km from RTM simulations with and without H_2O based on the viewing conditions and atmospheric state of a measurement by OMPS LP. At LP's resolution, there is an H_2O absorption feature centered around the 942 nm channel, as well as weaker features around 824 and 720 nm. Spectra at other altitudes look similar. Figure R3.2 shows altitude-normalized radiances at various altitudes measured by OMPS LP corresponding to the observational conditions in R3.1. A decrease in radiances is observed in the 930 – 970 nm range, including in the lower stratosphere, consistent with water vapor absorption. This is generally true for measurements throughout the southern latitudes and tropics.

In the northern latitudes, the water vapor spectral pattern becomes less clear. The impact of aerosols becomes more significant due to forward scattering viewing conditions of the LP instrument. The interaction between aerosol scattering and H_2O causes the H_2O signal to weaken under these conditions, leading to no notable changes or even a small increase in radiances, rather than decrease as seen in the tropics and southern latitudes. Figure R3.3 shows RTM simulations for the viewing conditions and atmospheric state of a measurement by OMPS LP in the northern mid-latitudes. The simulation with H_2O shows no water vapor absorption but rather shows a small increase in radiances. It is important to note that the RTM contains a number of simplifying assumptions that are not representative of reality (e.g., it assumes a simple reflective surface at the bottom of the atmosphere and ignores water vapor absorption within clouds) and therefore may be missing some important contributions in this regime necessary to model real measurements.

Figures R3.4 and R3.5 show northern mid-latitude cases measured by OMPS LP. Some cases, as in Figure R3.4, only show evidence of aerosol, with no clear H_2O signal; this mostly occurs at latitudes associated with the reduction in R^2 seen in Figure 2c around 40° N. This appears to be a transition region where the H_2O spectral features change from a decrease in radiances to an increase in radiances, inhibiting the ML model's performance. Others, as in Figure R3.5, show a flat or negative spectral slope over the wavelength range associated with H_2O, rather than monotonically increasing as would be expected under conditions with a strong aerosol signal and a negligible H_2O signal, which we believe may be indicative of H_2O.

Based on the analysis of simulated and measured radiances, we believe that the ML model derives water vapor from the spectral dependence of measured radiances and is able to distinguish different conditions by being exposed to a large range of cases during the training. Given the ML model's agreement with MLS, ACE-FTS, and SAGE III/ISS when applied to LP measurements in 2025 and 2026, the model can infer H_2O at northern high latitudes, because otherwise a lack of information content would manifest as reduced performance in NH regions compared to other regions, which are not seen in our validation results. Our test run indicated no changes in the model performance when we added SSA as additional parameter (note that SSA is consistently different between southern and northern hemispheres for the LP viewing geometry), which also confirms that the model rather relies on the spectral signature of measured radiances to infer the information about the water vapor concentration under different conditions.

We have added a new appendix (Appendix A, since it is referenced first in the text) to discuss this in detail, which includes the figures below as well as additional figures that highlight the importance of altitude normalization for this problem.

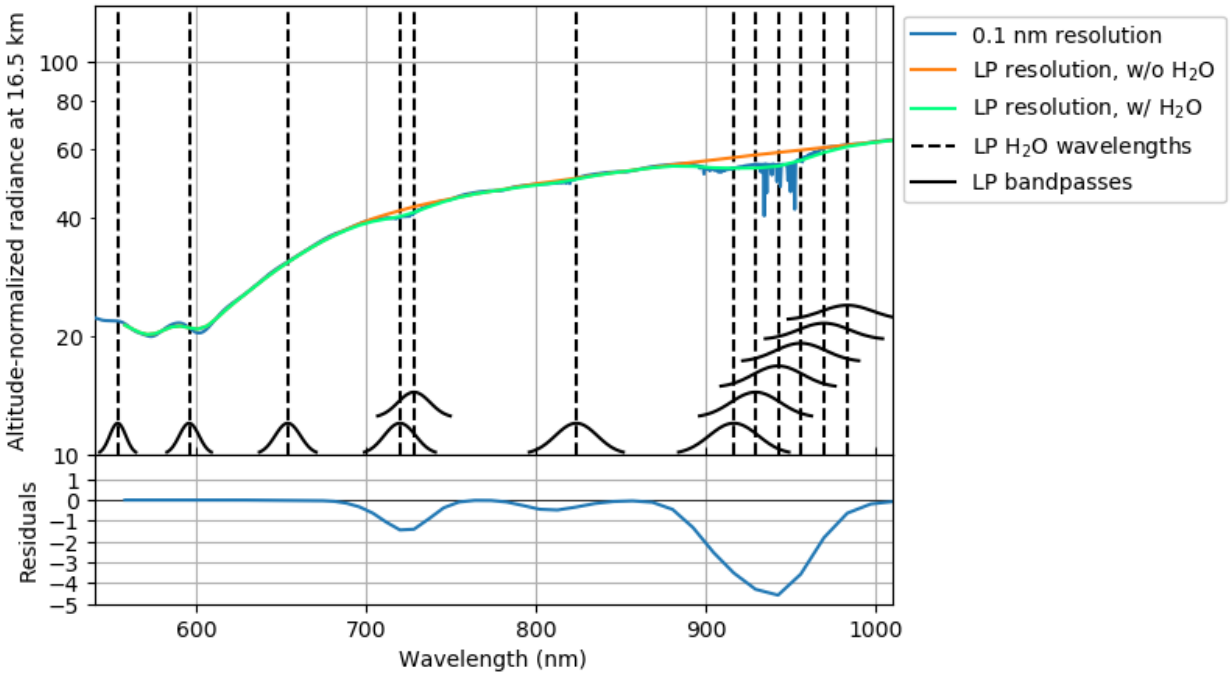

Figure R3.1. Radiative transfer simulations with and without H_2O based on a measurement by OMPS LP on 3 May 2019 at 5.4° N latitude, 106.7° E longitude, where the solar zenith angle is 23.2° and the single scattering angle is 103.6°. Temperature and pressure profiles are from GEOS IT; O_3, surface reflectivity, and the aerosol extinction profile at 675 nm (extrapolated to other wavelengths assuming a gamma-function size distribution) are from the OMPS LP ozone retrieval algorithm; and the H_2O profile is from a co-located MLS measurement. Radiances are altitude normalized at 40.5 km. Black dashed vertical lines mark LP channels used in the ML model for the H_2O retrievals. The LP bandpasses are shown as black curves, which are vertically shifted to avoid overlaps. Original RTM simulations with water vapor absorption are performed at a 0.1 nm spectral resolution (blue curve). Simulated radiances convolved with LP bandpasses are shown in green. Convolved radiances simulated without accounting for water vapor absorption are shown in orange. The lower panel shows the differences between convolved radiances simulated with and without water vapor (differences between orange and green lines), highlighting spectral regions where water vapor absorption is expected.

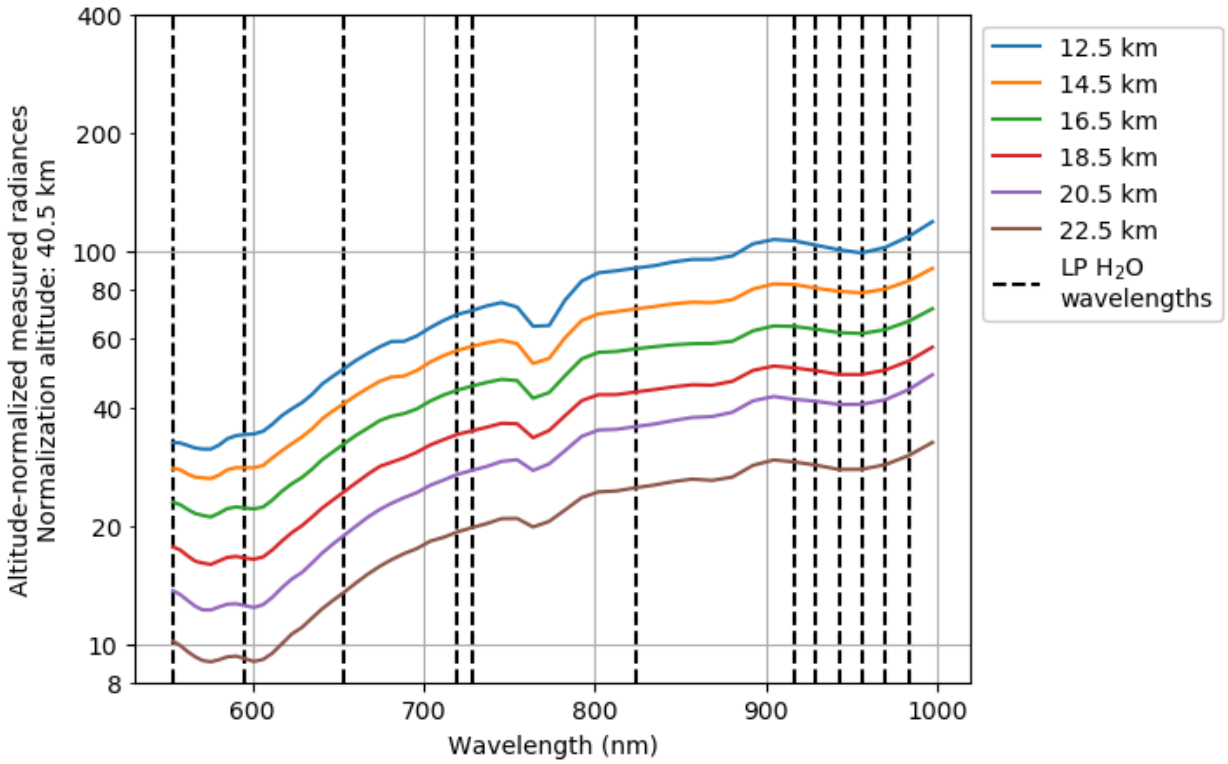

Figure R3.2. Altitude-normalized radiances measured by OMPS LP for the event modeled in Figure R3.1. At these altitudes, including in the lower stratosphere, a small decrease in radiances is observed around 930 – 970 nm, which we believe is attributed to water vapor absorption. However, the measured radiances do not show reduction around 720 and 824 nm as predicted by the RTM simulations.

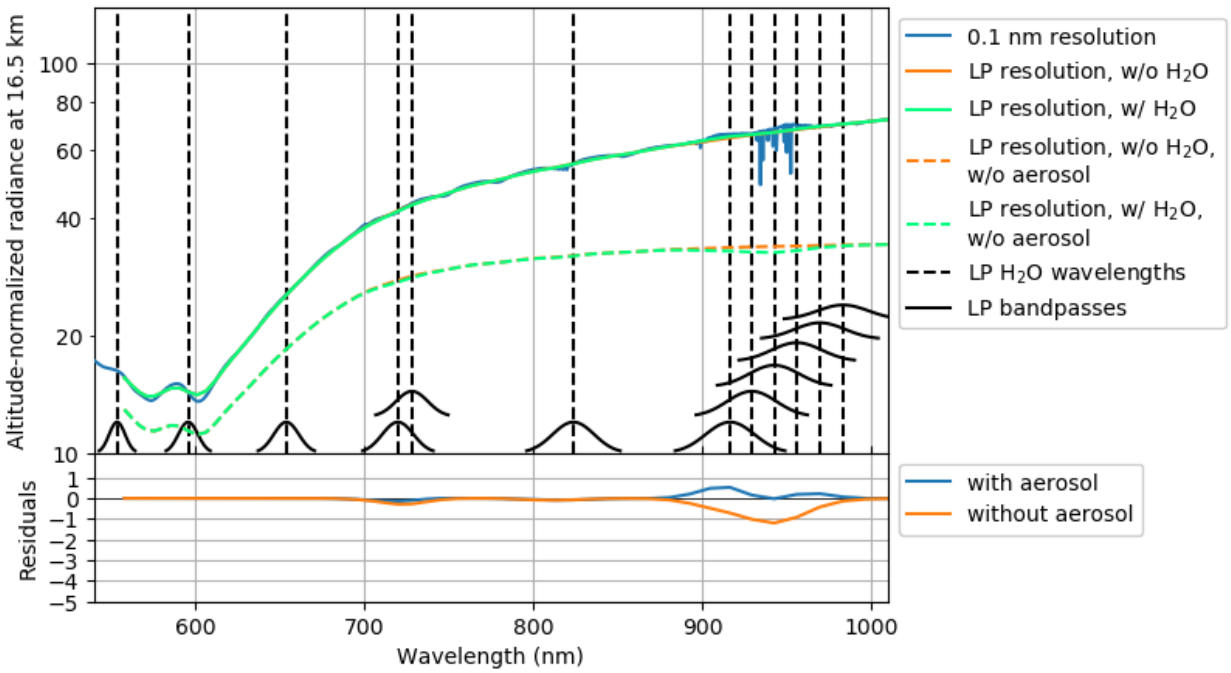

Figure R3.3. Like Figure R3.1, but for a measurement at 61.6° N latitude, 87.0° E longitude, solar zenith angle of 46.2°, and a single scattering angle of 49.5°. Simulations without aerosol are shown as dashed orange and green lines. Unlike in Figure R3.1, the lower panel here shows almost no differences between the simulations with and without H_2O in 930 – 970 nm spectral range, and even indicate a small increase in radiances simulated with H_2O compared with the simulation without H_2O. When omitting aerosol, radiances decrease due to H_2O absorption, confirming the increase in radiances for these viewing conditions is due to aerosol.

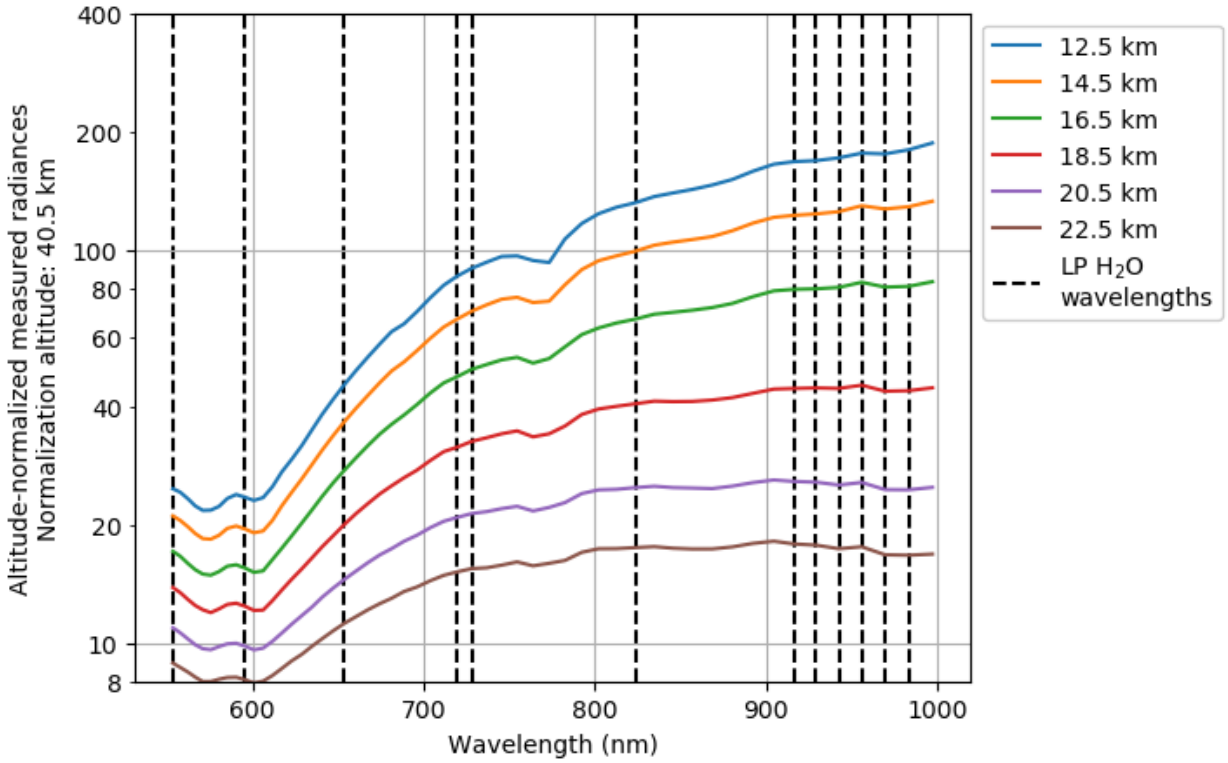

Figure R3.4. Like Figure R3.2, but for the event simulated in Figure R3.3. Unlike in Figure R3.2, there are no obvious spectral features attributable to H_2O in the 930 – 970 nm spectral band.

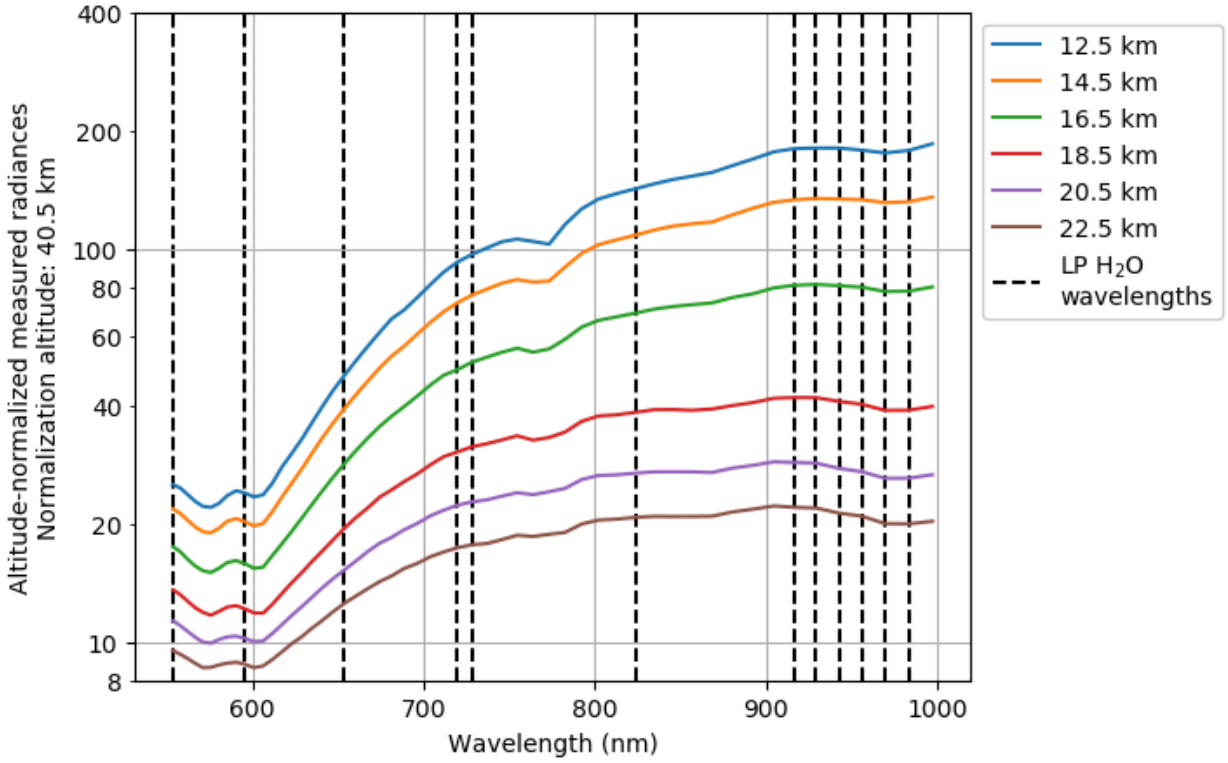

Figure R3.5. Like Figure R3.2, but for a measurement at 72.7° N latitude, 73.1° E longitude, solar zenith angle of 57.3°, and single scattering angle of 38.7°. In contrast to Figure R3.4, a decrease in radiances is observed at the longest wavelengths considered, possibly due to H_2O absorption.

2. The resampling in Line 116 addresses latitudinal sampling bias, but my concern is different: I am asking about the distribution of the training data across water-vapor concentration regimes and profile states, not only across latitude. Please discuss whether the training set is sufficiently representative across low, typical, and enhanced water vapor conditions, and whether rare or evolving regimes are underrepresented.

The training data set is comprised of a random sampling of measurements between 2014-2024 on "golden" days where LP and MLS have frequent co-locations that are close in time and space, which is then sub-sampled to ensure uniform coverage across latitudes. No special consideration is given to the water vapor concentration regimes or profile states when sampling, and consequently, rare regimes are inherently underrepresented. In earlier tests, we found that excluding data from 2022 – present resulted in reduced model performance during this period, as the H_2O perturbations from Hunga – both the initial water vapor injection, as well as its evolution – were significantly different than what the ML model saw during training. Including ~1% of the data from 2022 – 2024 was sufficient to ensure the model learned how to interpret the LP radiances during this time period as well as in 2025 – present.

Still, there are rare events that were not seen during training, and we have found that it does not prohibit the ML model from identifying those regimes. Figure R3.6 shows a water vapor profile from 8 January 2020 over the Pacific Ocean that was not seen during training plotted over the distribution of water vapor VMRs encountered during training. This case features a water vapor injection in the lower stratosphere due to the intense wildfires in Australia during December 2019 – January 2020, which was then transported eastward by winds. Despite that the VMRs at the altitudes of this enhancement were rarely seen during training (<0.045% probability for all altitudes containing the enhanced water vapor, most of which are from profiles that do not match the shape of this LP-retrieved profile), the model identifies the enhancement. The LP and MLS VMRs of this enhancement differ by up to a factor of 2, but it is unclear whether this is due to LP underestimating the VMR, MLS overestimating the VMR, a combination of both, or differences in measured air masses.

We have added text to the end of Section 4.3 as well as Figure 8 (matches Figure R3.6) to discuss this:

The NNs are largely trained on normal conditions. While Hunga is an anomaly, its impact has lasted years and can be seen globally, and thus the NNs are exposed to various Hunga-affected conditions. Since our data selection process does not consider H_2O concentration or profile states, rare short-lived events are severely undersampled

relative to nominal conditions. Despite this, we find that the NN models can still retrieve reasonably accurate profiles for rare cases not seen during training. Figure 8 shows the H_2O profiles retrieved by LP and MLS for an event over the Pacific Ocean on 8 January 2020. This event exhibits a water vapor enhancement in the lower stratosphere due to the intense wildfires in Australia during December 2019 - January 2020. Pyrocumulonimbus clouds carried moist, tropospheric air into the lower stratosphere, where it was then transported eastward by winds. Despite that these H_2O concentrations were seen in <0.045% of the training data set, the NNs are able to identify this enhancement. The LP and MLS VMRs of this enhancement differ by up to a factor of two, but it is unclear whether this is due to LP underestimating the VMR, MLS overestimating the VMR, a combination of both, or differences in measured air masses.

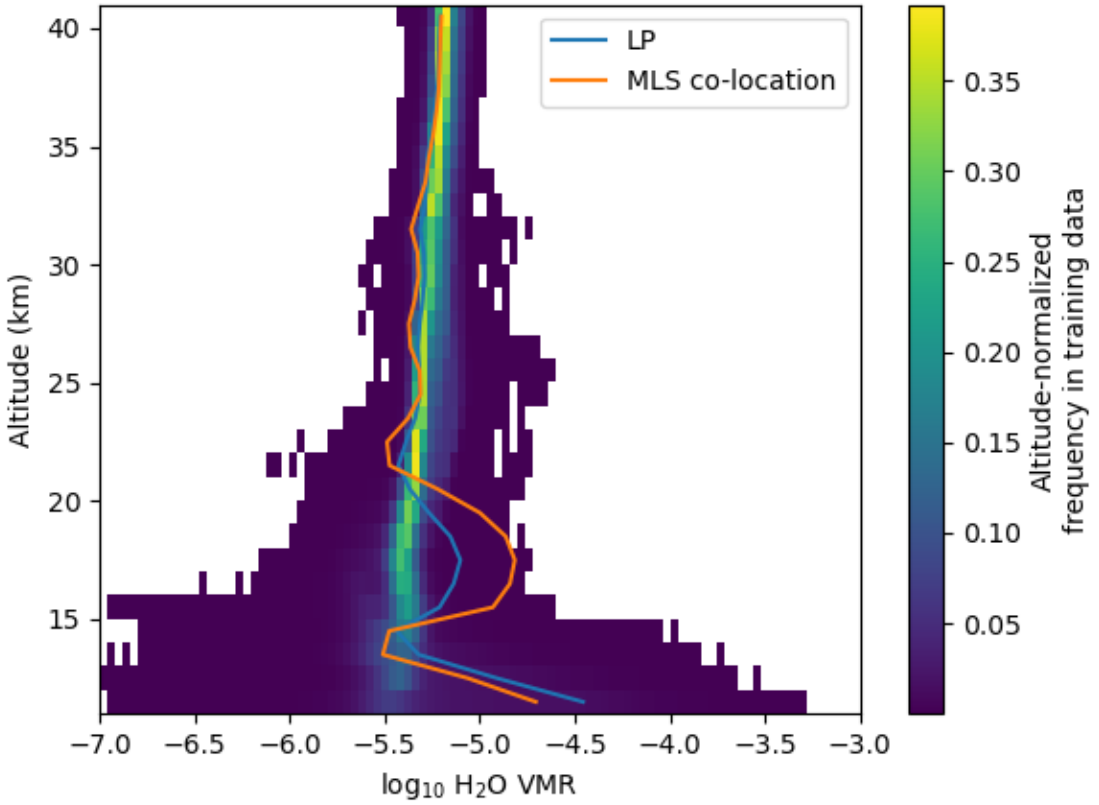

Figure R3.6. LP and MLS water vapor profiles from 8 January 2020, which feature a lower stratospheric enhancement caused by the Australian wildfires, are plotted over a 2D histogram of the water vapor profiles in the training data set. Each altitude is normalized such that the probabilities sum to 1. White regions indicate no data. Despite the low frequency of the VMRs of this enhancement in the training data set (<0.045%), the ML model successfully identifies this rare situation.

3. Figure 2c shows a localized mid-latitude vertical structure that differs from neighboring latitudes. Please explain whether this feature is related to dynamics, solar-zenith-angle dependence, training-data distribution, or a retrieval artifact.

This feature in the mid-latitudes is mostly a numerical effect due to low variability in these regions. The formula for R^2 is

$$R^2 = 1 - (\frac{RSS}{TSS})$$

where RSS is the residual sum of squares, and TSS is the total sum of squares (proportional to the total variance). The NN's mean squared error loss function combined with the data normalization procedure optimizes the model such that the relative difference between LP and MLS is roughly uniform across all cases considered. Since RSS is relatively constant (as demonstrated by the RMSE in the stratosphere in Figure 2b), changes in TSS (driven by the variance in MLS water vapor for a given combination of latitude and altitude) will adjust the R^2 value higher or lower even if agreement between LP and MLS remains the same. This is confirmed by considering the natural variance in MLS data, shown in Figure R3.7, and comparing the daily zonal means of LP and MLS in regions with low and high R^2, shown in Figure R3.8. The regions in Figure R3.7 that have the lowest variability are also the regions where the LP model shows the lowest R^2. The selected latitudinal zones in Figure R3.8 include the local minima in the mid-latitudes as well as regions with much higher R^2. Despite having lower R^2 values, the mid-latitude regions are not systematically worse; they tend to exhibit more consistent percent errors than the selected regions with higher R^2.

However, if the numerical effect of the R^2 calculation were solely responsible, then it would be expected that the southern mid-latitudes would show the lowest R^2 values, given that R3.7 shows that the southern mid-latitudes exhibit the lowest variability in H_2O. Figure 2c shows that the northern mid-latitudes exhibit the lowest R^2 values. This is due to a combination of this R^2 numerical effect as well as a reduction in H_2O spectral features as discussed in #1. As shown in Figures R3.1 and R3.3, the 930 – 970 nm spectral region changes from a reduction in radiances to an increase in radiances at certain solar zenith and single scattering angles due to aerosol. The transition between these regimes will have negligible H_2O spectral features, thereby reducing model accuracy. These effects combine and result in the R^2 minimum occurring in the northern mid-latitudes.

Additionally, we have conducted an experiment that rules out model confusion between similar solar zenith angle values near 40° S and 40° N, confirming that the ML model properly distinguishes these regimes, and therefore the reduction in H_2O feature strength in the northern mid-latitudes is not a factor in the reduction of R^2 in the southern mid-latitudes.

We have added the following text as the second paragraph in Section 4.1:

> *Figure 2c exhibits local minima in R^2 in the mid-latitudes. This reduction in R^2 is a numerical effect associated with how the NN models are trained combined with the variability of MLS water vapor. The equation for R^2 involves the ratio of the residual sum of squares (RSS) to the total sum of squares (TSS). The combination of the MSE loss function and data normalization procedure optimizes the NN models to make predictions such that the errors are relatively evenly distributed (homoscedasticity). TSS, on the other hand, is proportional to the variance of the MLS water vapor values. When MLS water vapor variance is reduced, the R^2 must also reduce if the agreement between LP and MLS remains constant. These mid-latitude regions are local minima in MLS water vapor variance, and thus they are also local minima in R^2 given the expected homoscedasticity of the NN models. However, the absolute minimum in water vapor variance occurs in the SH mid-latitudes, while the absolute minimum in R^2 occurs in the NH mid-latitudes. In the NH, there is an additional effect due to reduced water vapor sensitivity at certain combinations of altitudes, SZA, single scattering angle, aerosol, and H_2O. Based on RTM simulations, the spectral features of water vapor at 930 -- 970 nm shift from reductions in radiances in the SH and tropics to increases in radiances in the NH. At the transition between these regimes, water vapor sensitivity is at its minimum, corresponding to the minimum in R^2 in the NH mid-latitudes. For more details on this, see Appendix A.*

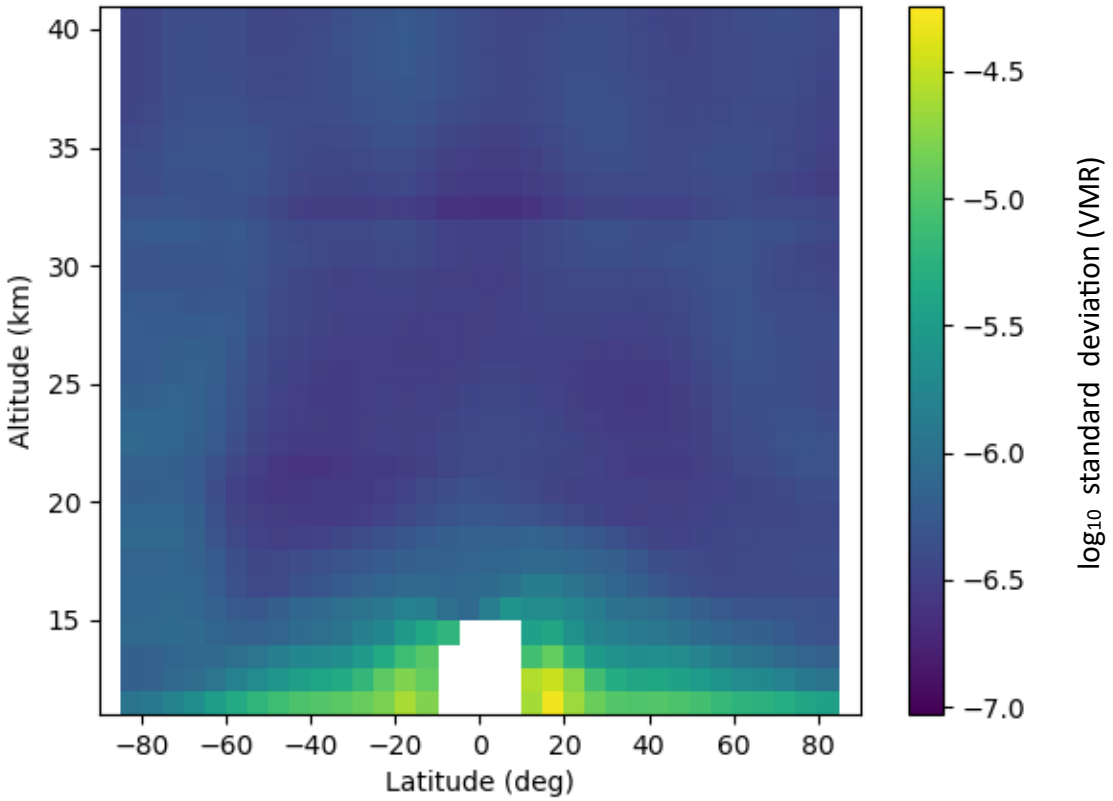

Figure R3.7. Zonal standard deviation of MLS between 2014 – 2021 for all co-locations with LP. The lowest variability occurs in the southern mid-latitudes. The northern mid-latitudes also show reduced variability relative to other regions. An artifact in MLS version 5 results in the band of reduced variability around 32 – 33 km.

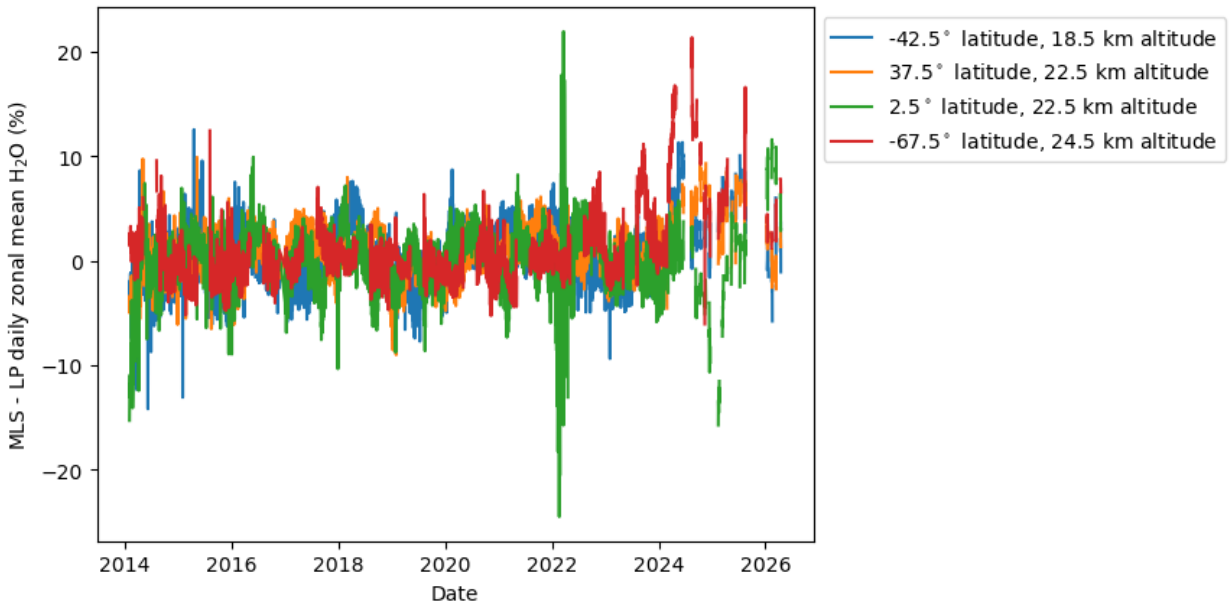

Figure R3.8. Time series of percentage differences in daily zonal means between MLS and LP for selected combinations of latitude and altitude. The blue and orange lines correspond to the southern and northern mid-latitudes where R^2 is at a minimum, while the green and red lines correspond to the tropics and high southern latitudes where R^2 is around its maximum. Despite the mid-latitudes' R^2 values suggesting reduced agreement between LP and MLS, the differences are consistent with other regions that show higher R^2 values, confirming that their reduced R^2 values are largely a numerical effect associated with low variability (Figure R3.7), rather than dynamics or a model artifact.