Measurements of atmospheric HDO/H2O in southern California from CLARS-FTS

Zhao-Cheng Zeng[1], Olivia Addington[2], Thomas J Pongetti[3], Robert L Herman[4], Keeyoon Sung[5], Sally Newman[2], Andreas Schneider[6], Tobias Borsdorff[7], Yuk L. Yung[2], and Stanley Sander[4]

[1]University of California, Los Angeles
[2]California Institute of Technology
[3]NASA-JPL
[4]Jet Propulsion Laboratory
[5]Jet Propulsion Lab (NASA)
[6]Earth science group, SRON Netherlands Institute for Space Research, Utrecht, the Netherlands
[7]SRON Netherlands Institute of Space Research

November 26, 2022

Abstract

Atmospheric isotopologues of water vapor (e.g., HDO) are important tracers for understanding Earth's hydrological cycles. Most remote sensing measurements of these isotopologues, however, are column averaged values and sparse in space and time. Measurements targeting the planetary boundary layer (PBL) are much rarer. In this study, we retrieved HDO and H from CLARS-FTS observations (2011-2019). The isotopological abundance δD, which represents the relative difference of the HDO/H_2O ratio to a standard abundance ratio, is also calculated. The averaged δD retrievals are (-156.1±60.0)uncertainty of (6.1+-10.2)an uncertainty of (42.4+-31.6)In LA, the δD shows a seasonal cycle that is primarily driven by the change of atmospheric humidity. The temporal variabilities in δD data between CLARS-FTS and a collocated Total Carbon Column Observing Network (TCCON) observatory are highly correlated. The difference between CLARS and TCCON δD retrievals can primarily be attributed to the difference in their observation geometries. We conclude that the HDO and δD measurements from CLARS-FTS provide high spatial and temporal resolution datasets for further study of hydrological processes in the LA megacity.

Measurements of atmospheric HDO/H_2O in southern California from CLARS-FTS

Zhao-Cheng Zeng[1,2,#], Olivia Addington[2,#], Thomas Pongetti[3], Robert L. Herman[3], Keeyoon Sung[3], Sally Newman[4], Andreas Schneider[5], Tobias Borsdorff[5], Yuk L. Yung[2], and Stanley P. Sander[3]

[1] Joint Institute for Regional Earth System Science & Engineering (JIFRESSE), University of California, Los Angeles, USA
[2] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[3] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[4] Bay Area Air Quality Management District, San Francisco, USA
[5] Earth science group, SRON Netherlands Institute for Space Research, Utrecht, the Netherlands
# equal contribution

Corresponding author: Z.-C. Zeng (zcz@gps.caltech.edu) and S. P. Sander (stanley.p.sander@jpl.nasa.gov)

Key points:

- Measurements of HDO, H_2O, and δD were made over southern California by CLARS-FTS from 2011 to 2019

- The δD values show a significant seasonal cycle that is highly correlated with the change of atmospheric absolute humidity

- δD measurements by CLARS-FTS, TCCON, and TROPOMI are in good agreement

Abstract

Atmospheric isotopologues of water vapor (e.g., HDO) are important tracers for understanding Earth's hydrological cycles. Most remote sensing measurements of these isotopologues, however, are column averaged values and sparse in space and time. Measurements targeting the planetary boundary layer (PBL) are much rarer. In this study, we retrieved HDO and H_2O columns from observations by the California Laboratory for Atmospheric Remote Sensing Fourier Transform Spectrometer (CLARS-FTS), a mountaintop observatory on Mt. Wilson (1.67 km a.s.l.) overlooking the Los Angeles (LA) basin in southern California. CLARS-FTS observations are highly sensitive to the lower atmosphere due to the long light path along the PBL. Retrievals were conducted using spectral windows between 6000-7000 cm^{-1} from CLARS-FTS observations (2011-2019). The isotopological abundance δD, which represents the relative difference of the HDO/H_2O ratio to a standard abundance ratio, is also calculated. The averaged δD retrievals are (−156.1±60.0)‰ with an uncertainty of (6.1±10.2)‰ for LA Basin Survey m and (−344.7±95.0)‰ with an uncertainty of (42.4±31.6)‰ for Spectralon Viewing Observation mode. In LA, the δD shows a seasonal cycle that is primarily driven by the change of atmospheric humidity. The temporal variabilities in δD data between CLARS-FTS and a collocated Total Carbon Column Observing Network (TCCON) observatory are highly correlated. The difference between CLARS and TCCON δD retrievals can primarily be attributed to the difference in their observation geometries. We conclude that the HDO and δD measurements from CLARS-FTS provide high spatial and temporal resolution datasets for further study of hydrological processes in the LA megacity.

1. Introduction

Water vapor is both the most abundant greenhouse gas and an important component of the global hydrological cycle. Uncertainties in global concentrations of water vapor are a major challenge for global climate modeling, but observations of water vapor isotopologues provide additional information in constraining meteorological models and improving understandings of current weather processes and past climate events (**Galewsky et al., 2016**). Given the differences in binding energies for different molecular masses, concentrations of water isotopologues are influenced by phase changes in a process referred to as fractionation, in which heavier isotopologues are more likely to condense compared to lighter isotopologues. Fractionation allows measurements of water isotopologue concentrations to serve as a proxy for observing water vapor transport through the global hydrological cycle.

There are several naturally occurring stable oxygen (^{16}O, ^{17}O, and ^{18}O) and hydrogen (^{1}H and ^{2}H or D) isotopes, all of which can combine to form various stable water isotopologues. The lightest isotopologue ($H_2{}^{16}O$) is the most abundant, but heavier isotopologues (e.g., HDO) are still observed commonly on Earth (**Yoshimura, 2015**). In liquid water, heavier isotopologues have higher binding energies and lower diffusive velocities, making them less likely to evaporate compared to lighter isotopologues (**Craig, 1961; Craig et al., 1965**). Therefore, the resulting water vapor tends to contain a smaller proportion of heavy isotopologues compared to the remaining liquid water, i.e. is more isotopically "depleted" than the liquid water (**Galewsky et al., 2016**). Similarly, when precipitation occurs, heavier isotopologues are more likely to condense, again leaving the remaining gaseous water vapor more isotopically depleted (**Yoshimura, 2015**). Such fractionation processes allow measurements of isotopologue concentrations to serve as a proxy for observing water vapor movement. Generally, a ratio, R, of heavy to light isotopologues, is defined to quantify relative concentrations. Using delta-notation "δ", one can measure the deviation of a given isotopologue concentration (here D) from the standard composition of ocean water:

$$\delta D = \left(\frac{R_D}{R_{D,VSMOW}} - 1\right) * 1000\ [‰] \qquad (1)$$

where, R_D and $R_{D,VSMOW}$ are the ratios of the heavy to light isotopologue (HDO/H_2O) in the sample and standard, respectively. The standard is defined using Vienna Standard Mean Ocean Water (VSMOW). $R_{D,VSMOW} = 3.1152\times10^{-4}$ (**Hagemann et al., 1970**). The units of δD for a given isotopologue are in parts per thousand, or per mil (‰). δD values are generally negative, such that lower values, i.e. more negative, imply greater depletion of the sample in the heavy isotopologue, whereas higher δD values, i.e., closer to zero, imply greater enrichment compared to the sample.

In recent years, important advancements have been made in the field of remote sensing observations of stable water isotopologues, especially HDO. The growth in available global data sets from remote-sensing instruments along with the continual improvements of retrievals has in turn generated interest in data set validation and use of these data sets in modeling. One of the first instruments to demonstrate the potential of remote sensing to study water vapor isotopologues in the stratosphere was the Atmospheric Trace Molecule Spectroscopy (ATMOS) mission on the Space Shuttle (e.g., **Kuang et al., 2003**). Following this, instruments were developed that were sensitive to water vapor isotopologues in the troposphere. Greater spatial coverage of δD was provided by the Interferometric Monitor for Greenhouse gases (IMG) sensor on the Advanced Earth Observing Satellite 1 (ADEOS-1) (**Zhakalov et al., 2004**). Since then, subsequent satellite instruments have improved the temporal and spatial resolution of δD data retrievals and increased the number of data sets (e.g., **Worden et al., 2007; Lacour et al., 2012**). For example, the Aura Tropospheric Emission Spectrometer (TES) and Atmospheric Infrared Sounder (AIRS) use thermal IR radiances to measure the HDO on a global scale (e.g., **Worden et al., 2007; Worden et al., 2019**). The Greenhouse Gas Observing Satellite (GOSAT) launched in 2009, houses a Fourier Transform spectrometer, Thermal and Near Infrared Sensor for Carbon Observations (TANSO-FTS), with the capability of retrieving global HDO and H_2O concentrations (**Frankenberg et al., 2013; Boesch et al., 2013**). The Scanning Imaging Absorption Spectrometer for Atmospheric Chartography instrument (SCIAMACHY) aboard the ESA's environmental research satellite ENVISAT used near infrared (NIR) spectra to retrieve global δD values with high sensitivity in the lower troposphere, where most of the water vapor resides (**Frankenberg et al., 2009; Schneider et al., 2018**). Most recently, The TROPOspheric Monitoring Instrument (TROPOMI) instrument on board Sentinel-5 also uses short-wave infrared spectra to make global total column measurements of HDO and δD, reporting improvements in the signal-to-noise ratio of observations compared to SCIAMACHY (**Schneider et al 2020**).

Ground-based remote sensing instruments, such as global Total Carbon Column Observing Network (TCCON; **Wunch et al., 2011**) which operates in the NIR, have also computed δD values from water vapor isotopologue retrievals (**Rokotyan et al., 2014**). Comparable measurements taken from Network for the Detection of Atmospheric Composition Change (NDACC) global tower network, which are similar to TCCON measurements in viewing geometry but use spectra from middle infrared, are incorporated into project MUSICA (Multi-platform remote Sensing of isotopologues for investigating the Cycle of Atmospheric Water), which includes measurements from ground-based, space-based, and in-situ instruments. MUSICA has performed valuable validation work of water vapor isotopologue measurements (**Schneider et al., 2016**). Specifically, MUSICA applies a bias correction to remote sensing data using vertical isotopologue profiles measured by well-calibrated in-situ instruments with low instrumental uncertainty (**Gonzales et al., 2016**). The MUSICA data product has been used for validation and bias correction of other remote sensing data sets (**Scheepmaker et al., 2015**). Overall, the availability of new data sets has allowed for the improvement of general circulation models (GCM) and prompted many studies aimed at better understanding complicated meteorological processes such as convection, cloud formation, and stratospheric-tropospheric exchange processes, along with the relative contribution of different global sources to atmospheric water vapor (**Yoshimura, 2015; Galewsky et al., 2016**).

However, most measurements of these isotopologues from space-borne and ground-based remote sensing instruments are column averaged values with contributions from all altitudes. For thermal IR based satellites, e.g., TES or AIRS, the measurements have low sensitivity in the PBL. The measurements are also sparse in space and time on an urban scale like the Los Angeles (LA) basin in southern California. Measurements targeting the planetary boundary layer (PBL), the layer that couples the Earth's surface and the atmosphere above, are much rarer. In this study, we study the PBL-targeted measurements of HDO and δD in LA from observations by the California Laboratory for Atmospheric Remote Sensing - Fourier Transform Spectrometer (CLARS-FTS). Compared to conventional remote sensing observation networks, CLARS-FTS observations are highly sensitive to the lower atmosphere due to the long light path along the PBL.

In this paper, we first demonstrate that HDO and δD can be retrieved from CLARS-FTS spectra with sufficiently small fitting error and retrieval uncertainty. The entire spectral record from 2011 to 2019 was processed to provide a novel HDO, H2O, and δD dataset. Secondly, an examination of the temporal variability of XHDO, XH_2O, and δD on annual and interannual timescales is performed using these resulting datasets. Thirdly, we compare CLARS-FTS retrievals to TCCON and TROPOMI retrievals at Caltech to demonstrate their consistency and discrepancy. Finally, we show that the discrepancies between CLARS-FTS and TCCON can be explained by the differences between their observation geometries.

2. Data and methods

2.1 CLARS-FTS

(a) (b)

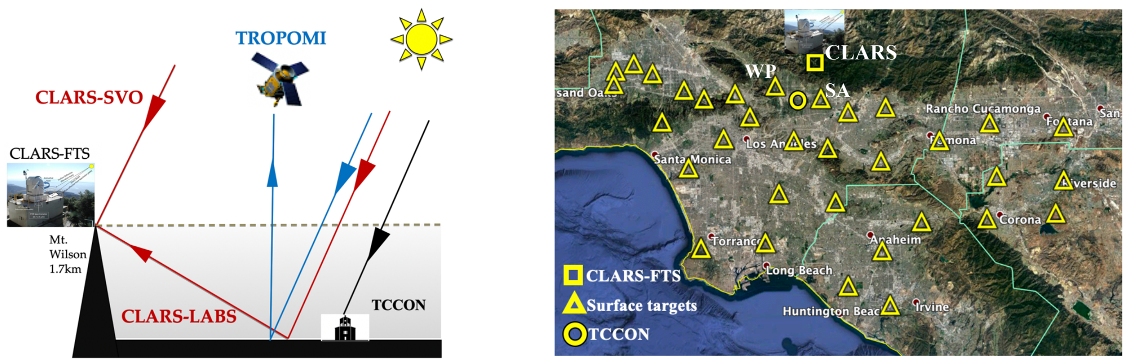


Figure 1. (a) Schematic figure showing the observations of CLARS-FTS, TROPOMI, and TCCON in the Los Angeles Basin; (b) The spatial distribution of CLARS-FTS surface reflection targets. The location of TCCON at Caltech, and the surface targets of West Pasadena (WP) and Santa Anita (SA) are also indicated.

The CLARS-FTS instrument is located on Mt. Wilson at 1.67 km a.s.l. and makes daily observations of solar spectra reflected from 33 different surface targets distributed around the LA basin (**Figure 1(b)**). CLARS-FTS is a first-of-its-kind mountaintop observation system to monitor urban emissions by collecting surface reflected light from a top-down perspective. CLARS-FTS has two observing modes: Los Angeles Basin Survey (LABS) and Spectralon Viewing Observation (SVO) (**Figure 1(a)**). In the SVO mode, the spectrometer is pointed toward a Spectralon plate directly below the instrument. It receives the reflected sunlight to retrieve the column abundance of the atmosphere above the CLARS level. The SVO observation enables CLARS-FTS to measure background concentration during the day. In the LABS mode, the instrument is pointed towards one of the 33 surface reflection targets. The LABS observation mode has a longer light path in the PBL relative to satellite measurements and therefore higher sensitivity to urban emissions. The observation time for each surface target is about 3 minutes, which means high temporal resolution retrievals can be achieved. The spatial coverage spanned by the collection of reflection points provides a mapping capability over the entire LA basin. CLARS-FTS can perform one basin-wide scan in approximately 90 minutes and cycles through the entire measurement cycle around 5-8 times a day. Given its location above the top of the PBL, CLARS-FTS is a unique instrument which is both ground-based but employs a retrieval geometry similar to that of a geo-stationary satellite. The result of this instrument configuration, coupled with the frequency with which measurements are taken, yields a record of spectral observations possessing high spatial and temporal resolution, compared to other remote sensing instruments. Furthermore, as CLARS-FTS began taking daily measurements in 2011, the spectral record represents the longest available data record of atmospheric gases for the entire LA basin. A detailed description of the FTS and the surface reflection targets can be found in **Fu et al. (2014)** and **Wong et al. (2015)**.

2.2 Retrievals of HDO, H_2O, and δD

CLARS-FTS operates in the NIR from 4000 cm^{-1} to 13500 cm^{-1} with a spectral resolution of 0.06 cm^{-1}. Recorded solar spectra are converted to slant-column densities (SCD), or total numbers of absorbing molecules per unit area along a Sun-Earth-instrument optical path, using a modified GFIT algorithm developed at JPL (**Fu et al. 2014**). The GFIT algorithm, within the GGG 2014 Software Suite, is employed by other ground-based remote sensing instruments, including TCCON FTS, for the retrieval of greenhouse gases (**Wunch et al., 2011**). GFIT provides a recommended list of spectral windows for HDO and H_2O retrieval, along with associated input parameters. The broad spectral interval for HDO windows cover multiple HDO features, which leads to better and more consistent retrievals. **Figure A1** in **Appendix A** shows a comparison of the nine spectral window candidates by their contribution from HDO absorption to the overall gas absorption and the spectral fitting error for each spectral window using a set of ~4000 observations from 4 distinct days in 2013. We rejected the six spectral windows between 4000 cm^{-1} to 6000 cm^{-1} because of their large fitting error and used the three spectra windows between 6000 cm^{-1} and 7000 cm^{-1} that are more robust, as shown in **Table 1**. For H_2O, we selected the five spectral windows between 6200 cm^{-1} and 6500 cm^{-1} from the TCCON list (**Wunch et al., 2015**) that are close to the HDO windows, as shown in Table 1. These spectral windows have been fully tested for TCCON observations. In this study, we further evaluate these spectral windows for CLARS observations by their fitting errors and retrieval uncertainties. Examples of spectral fit, including their fitting residuals and contributions from target gases as well as interfering gases, are shown in **Figure 2** for HDO windows and in **Appendix B** for H_2O windows.

Table 1. Spectral Windows for HDO and H_2O with Associated Parameters

Gas	**Center (cm^{-1})**	**Width (cm^{-1})**	**Gases to fit**	**Continuum Basis Functions**
	6330.05	45.50	HDO, H_2O, CO_2	2-order polynomial for continuum
HDO	6377.40	50.20	HDO, H_2O, CO_2	
	6458.10	41.40	HDO, H_2O, CO_2	
	6255.95	3.60	H_2O, CO_2 , HDO	
	6301.35	7.90	H_2O, CO_2 , HDO	2-order polynomial for continuum
H_2O	6392.45	3.10	H_2O, HDO	
	6401.15	1.15	H_2O, HDO, CO_2	
	6469.60	3.50	H_2O, CO_2, HDO	

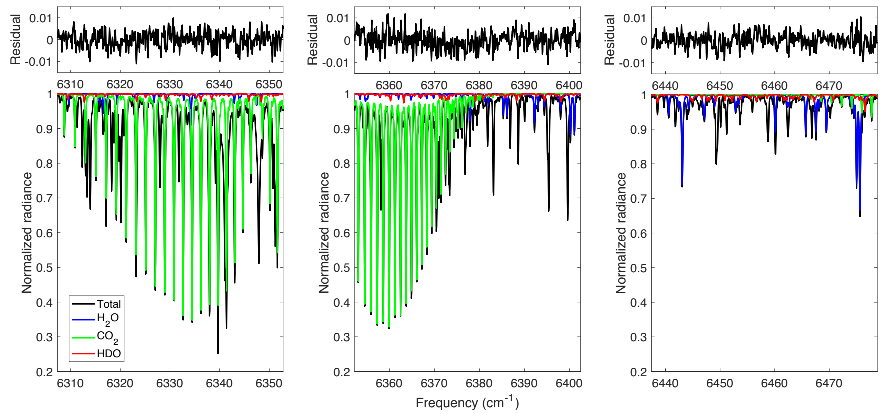


Figure 2. Examples of CLARS-FTS spectral windows of (left) 6330.05 cm^{-1}, (center) 6377.40 cm^{-1}, and (right) 6458.10 cm^{-1} for retrieving HDO in this study. These samples of normalized spectra are taken from a mid-day observation on 7/14/2013 over the West Pasadena surface target. The lower panel shows the full spectral fit with total contribution in black, contribution from HDO in red, from H_2O in blue, and from CO_2 in green. The upper label shows the residuals of the spectral fits, defined as the difference in total measured and total calculated radiance. Similar examples of spectral fit for H_2O are shown in the Appendix **Figure B1**.

Using these optimized spectral windows, we separately retrieved HDO and H_2O SCDs using the CLARS GFIT algorithm. For each spectral observation, CLARS GFIT outputs four numerical results with which to calculate a SCD for HDO and H_2O: an original vertical column density (OVC) in unit of molecules/cm^2, an air mass value (AM), a volume mixing ratio (VMR) scale factor (VSF), and the error in the VSF (VSF error). The air mass value represents the number of vertical columns the light travels through in its slant column path. The first three results are multiplied together to determine the SCD:

$$HDO_{SCD} = AM \times OVC \times VSF \tag{2}$$

The same calculation is conducted to produce H_2O SCD. The uncertainty in the SCD is determined using the same formula, but using VSF error in place of VSF, according to conventional rules of error propagation. Dry-air column averaged mixing ratios of HDO (XHDO) and H_2O (XH_2O) are computed from the retrieved SCDs by normalizing the SCD measurements to the dry-air total column, which can be derived from the measured SCD for O_2 and the dry-air O_2 mole fraction:

$$XHDO = 0.2095 \times \frac{SCD_{HDO}}{SCD_{O2}} \tag{3}$$

Using this method improves mixing ratio measurements since any existing systematic errors in retrievals of both HDO or H_2O and O_2 SCDs will be minimized in computing the ratio (**Fu et al., 2014**). In this study, O_2 is retrieved using a spectral window centered on 7885 cm^{-1}, whose retrieval results yield very low spectral fitting residual and VSF error values (**Zeng et al., 2020**). XH_2O is retrieved using the spectral windows in the same wavelength range as XHDO. This method reduces overall uncertainty in the final δD result by avoiding the possible complication of wavelength dependent noise such as that from the aerosol scattering effect (further discussion in **Section 4.1**). Ratioing of XHDO and XH_2O values to some extent cancels out possible errors existing in both XHDO and XH_2O retrievals. Finally, XHDO and XH_2O

measurements can be used to compute δD (**Equation 1**) for each CLARS-FTS observation, yielding a data set extending from 2011-2019 and spanning the LA basin.

2.3 Calibration of XH_2O, XHDO, and δD from CLARS-GFIT algorithm

Since CLARS-FTS retrievals employ the same GFIT algorithm as TCCON, the calibration developed by TCCON can also be applied to CLARS-FTS retrievals. As shown in **Wunch et al. (2015)**, XH_2O retrievals from TCCON observations have been compared against radiosonde measurements, resulting in a bias correction factor of 1/1.0183 being applied to TCCON XH_2O values. For XHDO retrievals, **Schneider et al. (2020)** derived a correction factor by scaling the TCCON XHDO to match the calculated δD between TCCON and the MUSICA dataset, whose δD profiles have been validated against aircraft measurements (**Schneider et al., 2016**). As a result, an error-weighted average correction factor of 1/1.0778 for XHDO was derived based on multiple TCCON sites. These two scale factors are used to calibrate XH_2O and XHDO retrievals, respectively, from CLARS-FTS, and the resulting calibrated δD is re-generated.

2.4 HDO and δD observations from TCCON and TROPOMI

2.4.1 TCCON

The TCCON FTS measures direct solar spectra in the NIR and retrieves the column-averaged abundances of many atmospheric gases, including H_2O and HDO, using the GFIT algorithm. XH_2O reported in the TCCON data product is the mean of retrievals from fifteen spectral windows and for XHDO it is from six spectral windows. A detailed introduction of the configuration of TCCON FTS, the characteristics of the observed spectra, and the calibration of the GFIT retrievals using aircraft measured profiles can be found in **Wunch et al. (2011)**. As discussed in **Section 2.3**, for TCCON XH_2O, a correction factor of 1/1.0183 is applied, and for XHDO, a correction factor 1/1.0778 is applied. The XH_2O, XHDO, and δD data were collected from 2012 to 2019 by the TCCON-Caltech site (**Figure 1(a)**; **Wennberg et al., 2015**) on the Caltech campus.

2.4.2 TROPOMI

The TROPOMI instrument on board Sentinel-5 uses short-wave infrared spectra at 4225 cm^{-1} (i.e., 2.3 μm) to retrieval global total column measurements of HDO, H_2O, and δD (**Schneider et al., 2020**). The measurements have a daily coverage (overpass at around 1:30 pm local time) and a spatial resolution of up to 3.5 km×7 km at nadir. The retrievals are filtered by strict criteria to exclude measurements contaminated by clouds and aerosols. A detailed description of the instruments, retrieval algorithm, and data screening can be found in **Schneider et al. (2020)**. From a comparison with collocated TCCON measurements, the TROPOMI retrievals have a mean bias of (−1.1±7.3)% for XHDO and (−14±17)‰ for δD. The δD dataset available from late 2017 to 2019 in the LA region is used in this study. To allow more measurements for comparison, we relaxed the cloud filter slightly from 1% to 5% in the filters. Averaged δD before and after relaxing the filter are the same (about -236‰), indicating no bias caused. In total, there are 635 valid observations over the study area shown in **Figure 1(b)**.

3. Results

3.1 Retrieval and spectra fitting errors from CLARS-FTS

Based on preliminary investigation of the quality of spectral fittings and the retrieval uncertainties in HDO and H_2O mixing ratios, the entire spectral record of CLARS-FTS from 2011 to 2019, was reprocessed using the spectral windows with central wavenumbers in the range 6000-7000 cm^{-1}. The CLARS-GFIT parameters used are the same as those shown **Table 1**. Before additional analysis was performed on HDO and H_2O observations, the data were passed through a series of filters, as summarized in **Table 2**. Data with poor spectral fitting, identified as instances with large solar zenith angles (SZA), low signal-to-noise (SNR) ratios, and large root-mean-square-error (RMSE) values from the spectral fitting, are removed. Additionally, the ratio between retrieved and geometric O_2 slant column densities (O_2 ratio) are used to remove retrievals affected by cloud and aerosol scattering. The geometric O_2 SCD is calculated assuming no scattering occurs, along with additional assumptions outlined in **Fu et al. (2014)**. Because oxygen is well-mixed in the atmosphere, deviations in the retrieved O_2 SCD from the geometric O_2 SCD implies variations in the light path due to clouds and/or aerosols (**Zeng et al., 2018; Zeng et al., 2020**), and can therefore be used to identify observations that represent especially cloudy or hazy days. Along with the above criteria, retrievals with high uncertainty values, defined VSF error values (one of the outputs from CLARS-GFIT) as more than two times the standard deviation from the mean VSF error, are also removed. This filter helps to remove δD results which would necessarily have very high uncertainties, since VSF error is propagated through calculation of SCDs, mixing ratios, and eventually δD values. Note that the VSF error is the uncertainty of the profile scaling factor while the fitting RMSE is the error of residuals from the spectral fit.

Table 2. Data filters for HDO and H_2O retrievals from CLARS-GFIT

Filters	Selection Criterion
Low clouds and/or aerosols	LABS O_2 ratio between 0.9 and 1.04
High clouds	SVO O_2 ratio between 1.0 and 1.08
Large SZA	SZA less than 70 degrees
Low SNR	SNR larger than 100
Poor spectral fit	Spectral fitting RMSE less than 1 standard deviation above mean
High VSF Error	VSF error less than 2 standard deviation above mean

Figure 3 shows relevant statistics for the entire data set once it has been filtered according to the above criteria. The figure shows the fraction VSF error in retrievals, a quantification of retrieval uncertainty, for HDO, where the data is separated into histograms according to observation modes (LABS and SVO) and spectral windows (6330.05 cm^{-1}, 6377.40 cm^{-1}, and 6458.10 cm^{-1}). For LABS retrievals, one can observe that the fraction VSF error in both HDO and H_2O yield similar distributions for the three spectral windows. In other words, retrievals from the three spectral windows have similar retrieval uncertainties for LABS measurements. In addition, the majority of retrievals have VSF error values less than 10% for HDO. However, this consistency is not observed for SVO measurements, where the distributions for the two higher wavenumber spectral windows (6377.40 cm^{-1} and 6458.10 cm^{-1}) have very long tails. The reason maybe at these two windows are contributions from interference due to other gases, mainly CO_2 and H_2O as shown in **Figure 2**, are much stronger than HDO for the

portion of the atmosphere above CLARS. As a result, retrieved VSFs for HDO are associated with large uncertainty even as the fitting errors are small, as shown in the following **Figure 4**. For H_2O as shown in **Figure B2**, most retrievals, except for the 6255.95 cm^{-1} spectral window, have VSF error less than 10% for LABS and 20% for SVO. The 6255.95 cm^{-1} spectral window, however, shows larger VSF error, especially for SVO. We therefore conclude that this window is not robust for H_2O retrievals and are not included in the following analysis.

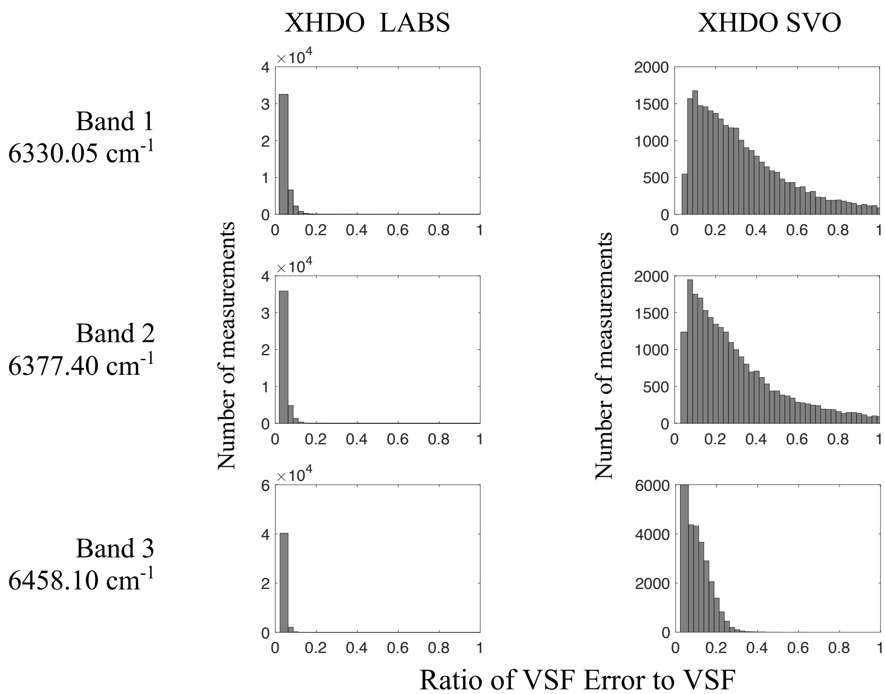


Figure 3. Retrieval error, in ratio of VSF error to VSF, for the entire filtered XHDO datasets from 2011 to 2019. Both VSF and VSF error values are calculated from CLARS-GFIT. The data are separated into histograms according to observation modes (LABS and SVO) and three spectral windows (6330.05 cm^{-1}, 6377.40 cm^{-1}, and 6458.10 cm^{-1}). The retrieval errors for H_2O data are shown in the Appendix **Figure B2**.

Figure 4 shows the histograms of RMSE from spectral fitting for the entire HDO data set, again with SVO and LABS observations separated. In the LABS histograms, we do see slightly different distributions when comparing the two lower frequency (6330.05 cm^{-1} and 6377.40 cm^{-1}) spectral windows to the 6458.10 cm^{-1} spectral window, which tends to have greater counts of higher RMSE. For SVO, the fitting errors are in general less than those for LABS, because SVO mode is measuring the background concentration above PBL with small perturbations such as impacts from aerosol and surface scattering. However, the 6330.05 cm^{-1} and 6377.40 cm^{-1} retrievals have significantly larger fitting errors compared to the 6458 cm^{-1} spectral window. These discrepancies are consistent with what is shown in **Figure 3** for the retrieval uncertainty. Therefore, we conclude that the 6330.05 cm^{-1} and 6377.40 cm^{-1} retrievals for SVO mode are not robust for HDO retrievals. For the following analysis, we used the 6458.10 cm^{-1} spectral window only for SVO HDO retrievals. For H_2O spectral windows as shown in **Figure B3**, all windows show consistent RMSE for both LABS and SVO.

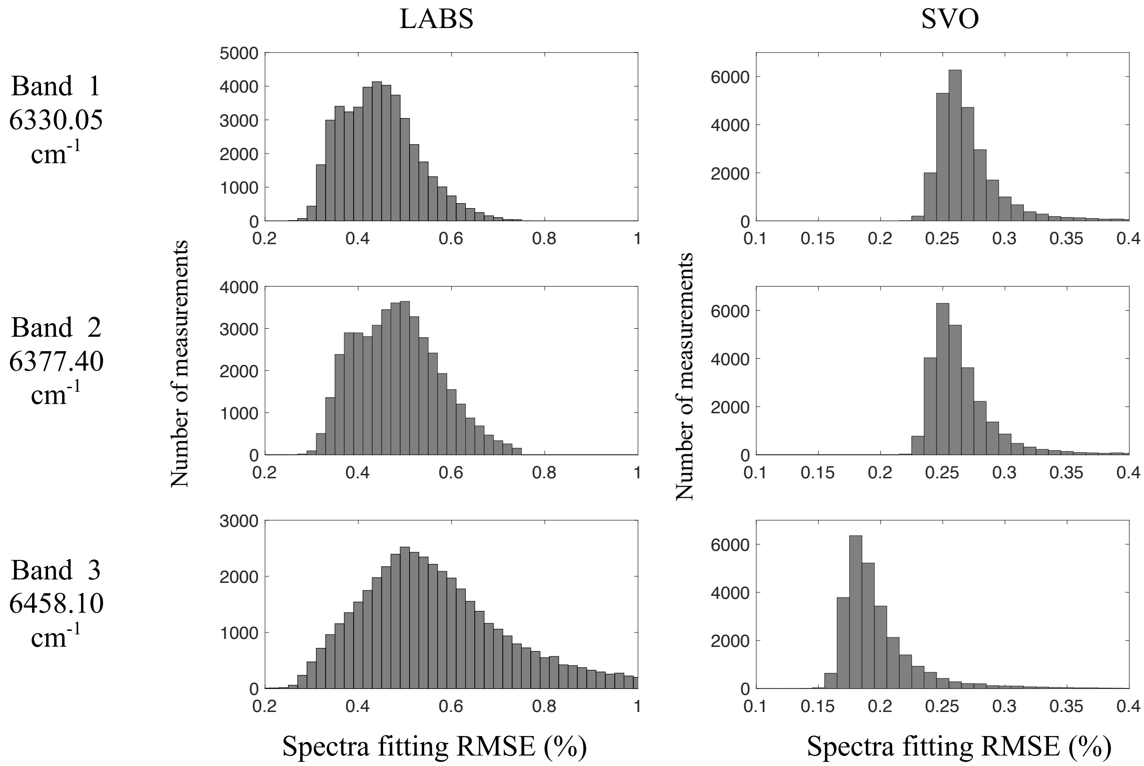


Figure 4. Histograms of RMS error from spectral fitting of HDO windows for the entire data set from 2011 to 2019. The data are separated into histograms according to observation mode (LABS and SVO) and three spectral windows (6330.05 cm^{-1}, 6377.40 cm^{-1}, and 6458.10 cm^{-1}). Similar figures for H_2O windows are shown in the Appendix **Figure B3**.

3.2 HDO, H_2O, and δD retrieval from CLARS-FTS

Figure 5 shows the diurnal observations of XHDO, XH_2O, and δD values on the sampled set of retrievals from four days in 2013. Over these four days, CLARS LABS was targeting two closer surface reflection targets: West Pasadena and Santa Anita. The LABS XHDO results from three spectral windows are differentiated by color and the SVO XHDO results are plotted in black. XH_2O results from four spectral windows are shown. To derive δD shown in **Figure 5**, weighted averaged XH_2O are calculated first over all the spectral windows. Equation (1) is then applied to obtain the δD values for each of the three XHDO retrievals. From **Figure 5**, we observe relatively good agreement in the retrievals from the three spectral windows, which also suggests these retrieval results are robust. Furthermore, uncertainties in retrieval results, shown with error bars, are relatively small for all spectral windows on each of the four days. The time series indicates that XHDO and XH_2O values exhibit both diurnal and seasonal variability for the LABS observations. For all four days, XHDO and XH_2O increase consistently until mid-day and then decrease in the late afternoon into the evening, as is reasonable based on diurnal temperature patterns. In July, we note that the XHDO ranges between 1−4 ppm and XH_2O ranges between 5,000 and 10,000 ppm whereas in September, the concentrations of both species are smaller. For SVO, both XHDO and XH_2O are lower than for LABS. The results of δD are also shown for LABS and SVO. The δD shows a smaller diurnal variability compared to XHDO and XH_2O. The SVO δD is smaller than the LABS data. This is consistent with the vertical distribution of δD, in which *δ*D generally decreases (more depleted) with elevation (**Galewsky**

et al., 2014). The seasonal change of δD is associated with the changes in specific humidity, which is illustrated in **Section 3.5**.

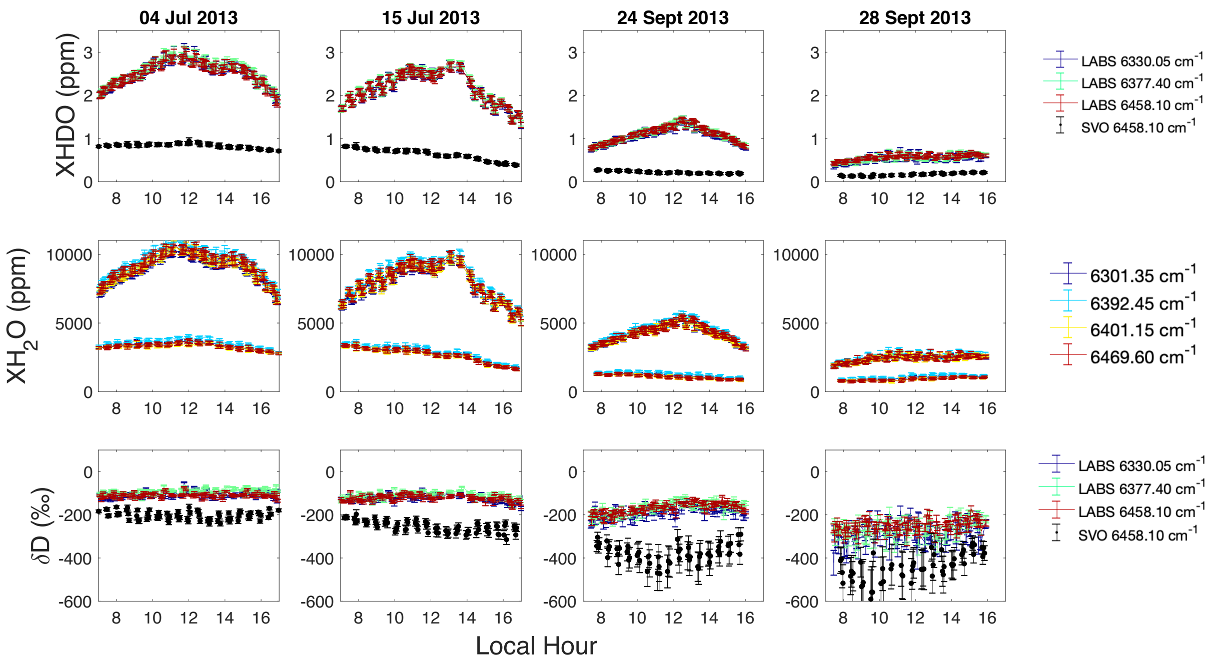


Figure 5. Diurnal observations of XHDO, XH_2O, and δD values on the sampled set of retrievals from four days in 2013, where LABS results from the different spectral windows are differentiated by color. These data are combined observations from West Pasadena and Santa Anita surface targets, which were two closer ones targeted over these four days by CLARS-FTS. . For XHDO and δD, the SVO data are from band 6458.10 cm^{-1} only, while for XH2O, the SVO data (lower time series) are from all available bands as LABS. Some data are not visible because of overlapping. Error bars show uncertainty values in individual observations. To derive δD, weighted averaged XH_2O are calculated first over all the spectral windows. Equation (1) is then applied to obtain the δD values for each of the three XHDO retrievals.

HDO and H_2O SCDs, XHDO and XH_2O, and δD values are computed separately from each spectral window. **Figure 6** shows the correlation results of LABS XHDO and δD from one spectral window versus another. The three columns show the three correlations that can be done using the set of three spectral windows. The three rows are HDO and δD correlation results. As one can see from **Figure 6**, there is very good agreement between mixing ratio results among the three spectral windows. Visually, each of the correlations appears as a straight line with a slope of approximately 1. This is verified quantitatively by the high correlation coefficients indicated for each correlation plot. Furthermore, the δD values from different spectral windows are also significantly correlated. The reason the correlation for δD becomes weaker is because relative difference in δD between spectral windows is amplified when ratioing against XH_2O following equation (1), especially when XH_2O is small. The point-by-point absolute difference is about (25.7±24.1) ‰ on average. The window-to-window XH_2O retrievals (correlation plots not shown here) are also highly consistent, as can be observed from examples in **Figure 5**. These results suggest that averaging the LABS retrieval results from the three spectral windows will not bias the overall determination of mixing ratio, and in turn will reduce overall uncertainty in the δD values.

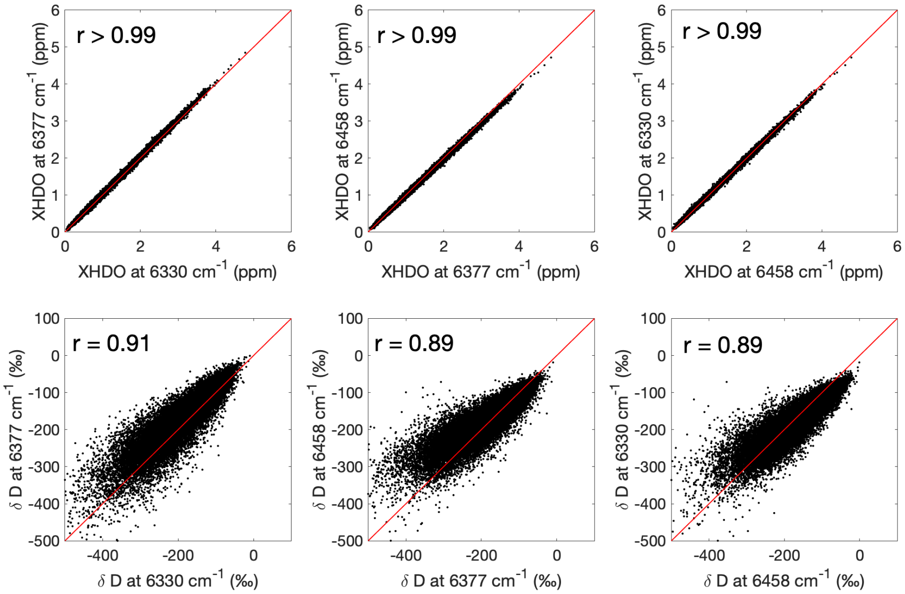


Figure 6. Correlations of HDO and δD retrievals from all the LABS observations for one spectral window versus another. In order of column, the correlations are: 6377 cm^{-1} vs 6330 cm^{-1}, 6458 cm^{-1} vs 6330 cm^{-1}, 6458 cm^{-1} vs 6337 cm^{-1}, respectively. The two rows show XHDO correlations and δD correlations, respectively. The 1:1 line in red and the correlation coefficients (r) are also indicated for δD. For XHDO, the correlation coefficients all larger than 0.99.

3.3 Seasonal cycles of HDO, H_2O, and δD

Given the LABS observations that retrieval uncertainties are not necessarily identical for all spectral windows (**Figure 5**), a weighted average rather than a simple average is used to compute a final mixing ratio value for each observation, given by (using XHDO as an example):

$$\overline{\mathrm{XHDO}} = \frac{\sum_{i=1}^{i=3} \mathrm{XHDO}_i \times \mathrm{w_i}}{\sum_{i=1}^{i=3} \mathrm{w_i}} \text{ where } \mathrm{w_i} = \frac{1}{\sigma_i^2} \qquad (4)$$

The weight for the mixing ratio value associated with each spectral window is defined as the reciprocal of the retrieval uncertainty, which in turn is defined using conventional error propagation of SCD uncertainties for the water vapor isotopologue and O_2 according to **Equation 3**. The same calculations (**Equation 4**) are made for XH_2O and δD for both LABS and SVO observations. **Figure 7** shows the time series of weighted means of XHDO, XH_2O, and δD. For SVO XHDO, only the 6458.10 cm^{-1} spectral window results are shown, as explained in **Section 3.1**. The averaged δD retrievals are (−156.1±60.0)‰ for LABS and (−344.7±95.0)‰ for SVO. The δD retrieval uncertainties are (6.1±10.2)‰ for LABS and (42.4±31.6) ‰ for SVO. From **Figure 7**, we can observe first that XHDO and XH_2O tend to be smaller for SVO observations compared with LABS observations. This is reasonable given the fact that XH_2O decreases with altitude due to decreasing temperatures and pressures. More evaporation also implies decreasing concentrations of the heavy isotopologue, which is again consistent with the plot in that XHDO and δD are smaller for SVO observations. On average, the difference between monthly averaged LABS and SVO measurements is (204.9±35.2)‰. We also find that the SVO δD values are more variable compared to the LABS measurements. This may be because of the much thinner atmosphere above PBL and therefore they more sensitive to any perturbations to the water vapor abundance.

From the first two panels, we see that XHDO and XH_2O values reach minimum values in the winter and maximum values in the summer months. This is again reasonable based on interannual patterns of temperature. From the last panel, we see that δD also reaches a minimum in the winter, which is interpreted as the time of greatest depletion in HDO. The δD value increases to around zero in the summer, implying when observed fractionation levels in atmospheric water vapor are approximately equivalent to that of the standard VSMOW. On average, the δD for the peak months from July to September is −112.4‰ and bottom months from December to February is −208.9‰. According to Rayleigh distillation (**Rayleigh et al. 1902**), there is preferential condensation of the heavier isotopologues (HDO here) compared with H_2O when the water vapor mixing ratio is lower. As a result, seasonal variations in δD are driven by the variations in humidity to the first order. This pattern fits with the seasonal temperature change for Los Angeles, higher temperature and absolute humidity (water content) from spring to summer and lower temperature and absolute humidity in the wintertime. The seasonal differences may also be due to varied contributions from different sources. For instance, there is likely more surface evaporation with higher δD when it is hotter during summer seasons. Therefore, lower absolute humidity in winter is associated with greater depletion in the heavy isotopologue, as is expected. The correlation between absolute humidity and δD is further explored in **Section 3.5**. Overall, CLARS-FTS provides continuous and robust estimations of XH_2O, XHDO, and δD for the LA basin.

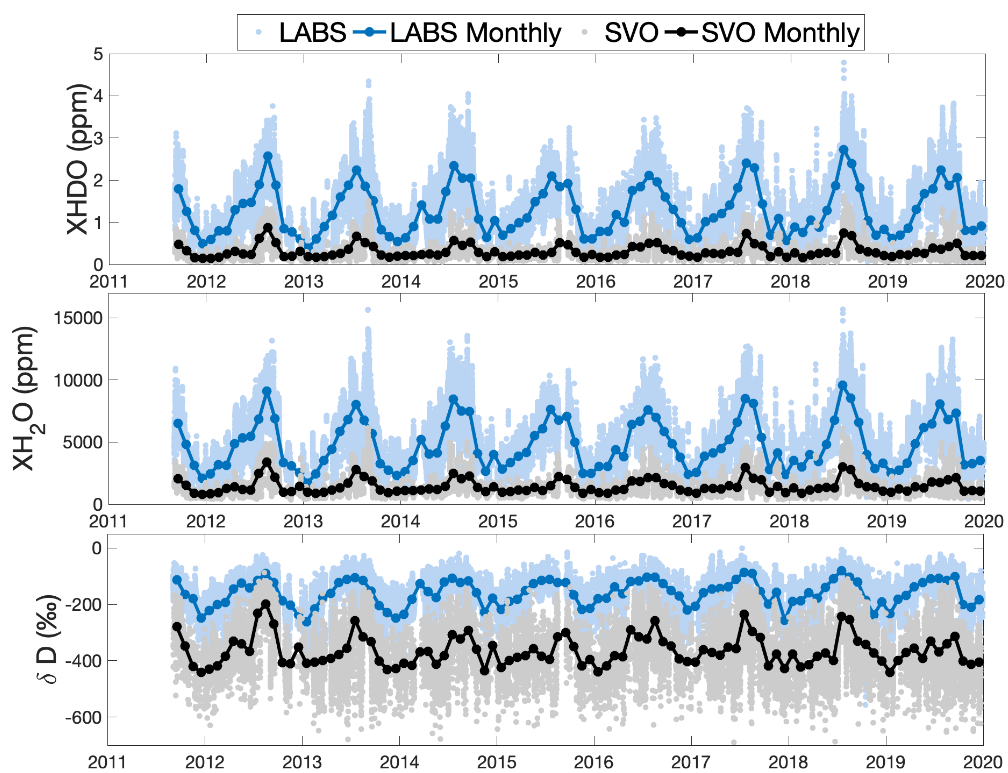


Figure 7. Time series of XHDO, XH_2O, and δD observed by CLARS-FTS from Sept. 2011 to Dec. 2019. LABS measurements are the weighted means from retrievals of the three spectral windows (6330.05 cm^{-1}, 6377.40 cm^{-1}, and 6458.10 cm^{-1}). The weights are calculated by their retrieval uncertainty. The SVO measurements are from retrievals using 6458.10 cm^{-1} spectral window only. The monthly means of these measurements are also shown.

3.4 Comparison of δD with TCCON and TROPOMI

Measurements of δD from CLARS, TCCON, and TROPOMI recorded at the same time and location are expected to differ due to many differences which include observation geometry, vertical profile sensitivity, radiative transfer modeling, and data filtering. However, these measurements should show similar temporal patterns due to the seasonal water cycles that drive the concentration of water isotopologues. As shown in **Figure 8(a)**, CLARS-FTS and TCCON δD measurements show high consistency in seasonal cycles with peaks in summers and troughs in winters. The seasonal amplitudes from monthly variations are 138.2‰ and 106.9‰ for CLARS-FTS and TCCON, respectively. On a daily basis, both data sets show high correlation (correlation coefficient (r) = 0.92) for daily averaged values. However, there is a systematic offset between the two measurements, with monthly averaged CLARS-FTS data higher than TCCON data by (40.4±18.0)‰ on average. Also, data for summer months (59.2‰ on average) have a higher average difference than winter months (30.5‰ on average). This systematic difference is expected because CLARS-FTS, compared to TCCON, measures an extra reflected light path in the PBL, where the heavy water isotopologue is more abundant than high altitudes above the PBL. As a result, CLARS-FTS measures higher δD values compared to TCCON. Further discussion to reconcile this difference is described in **Section 4.2**. Since TROPOMI and TCCON have similar measurement geometry, we see a similar comparison result between CLARS-FTS and TROPOMI (**Figure 8(b)**). Unfortunately, very few TROPOMI observations (in total 635) are available for comparison during the two-year period of overlapping data. However, their correlation is still high (r=0.89) for the daily averaged values. The offset between these two daily averaged datasets is about −61.7‰, with TROPOMI data being more negative than CLARS-FTS data.

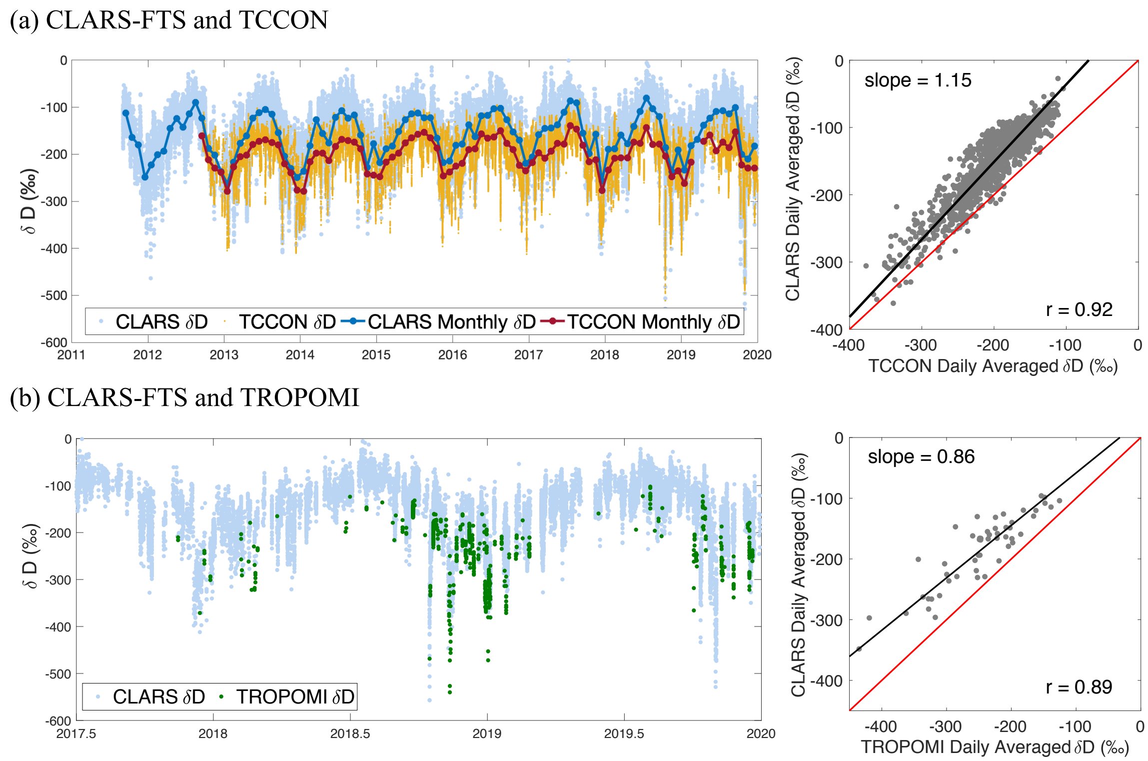


Figure 8. (a) left: comparisons of time series of δD measurements between CLARS-FTS and TCCON; right: scatter plot of their δD measurements; (b) left: comparisons of time series of δD measurements between CLARS-FTS and TROPOMI; right: scatter plot of their δD measurements.

3.5 Correlation between δD and humidity

According to Rayleigh distillation theory (**Rayleigh, 1902**), when water vapor abundance is lower, heavier isotopologues, such as HDO, are more likely to condense compared with H_2O. As a result, heavier isotopologues are more depleted (lower δD) in the air when humidity is lower. Therefore, the seasonal variations in δD are found to be primarily driven by variations in humidity. However, small departures from this humidity–δD correlation as predicted by Rayleigh distillation could provide new insights into secondary processes of the hydrological cycle related to evaporation and condensation. In this section, we will focus on the characterization of the humidity–δD correlations from measurements and leave further comparison with theoretical calculations to future studies.

In **Figure 9(a)**, we show the comparison of time series of XH_2O and δD from CLARS-FTS. The seasonal cycles show highly consistent seasonal patterns. Even on a daily time scale, both datasets closely track one another, as shown in **Figure 9(b)**. For example, the many low value anomalies during winter months shown in XH_2O are also shown in δD. Such a strong correlation can be seen in the scatter plots in **Figure 9(c-e)** from all CLARS-FTS, TCCON-Caltech, and TROPOMI data. This nonlinear correlation from remote sensing data has been reported by many other studies (e.g., **Worden et al., 2007; Noone, 2012; Schneider et al., 2020**). The progressive decrease of δD with a decreasing water vapor mixing ratio clearly demonstrates the preferential condensation of HDO compared with H_2O. Another possible process that may explain the δD and humidity correlation is the airmass mixing model, as described in **Noone (2012)**, that represents an exchange between two reservoirs with different H_2O mixing ratio and δD. For example, the mixing of the atmosphere with evaporation from the surface. A thorough exploration of these theoretical models to explain the measurements is beyond the scope of this paper and will be our future works.

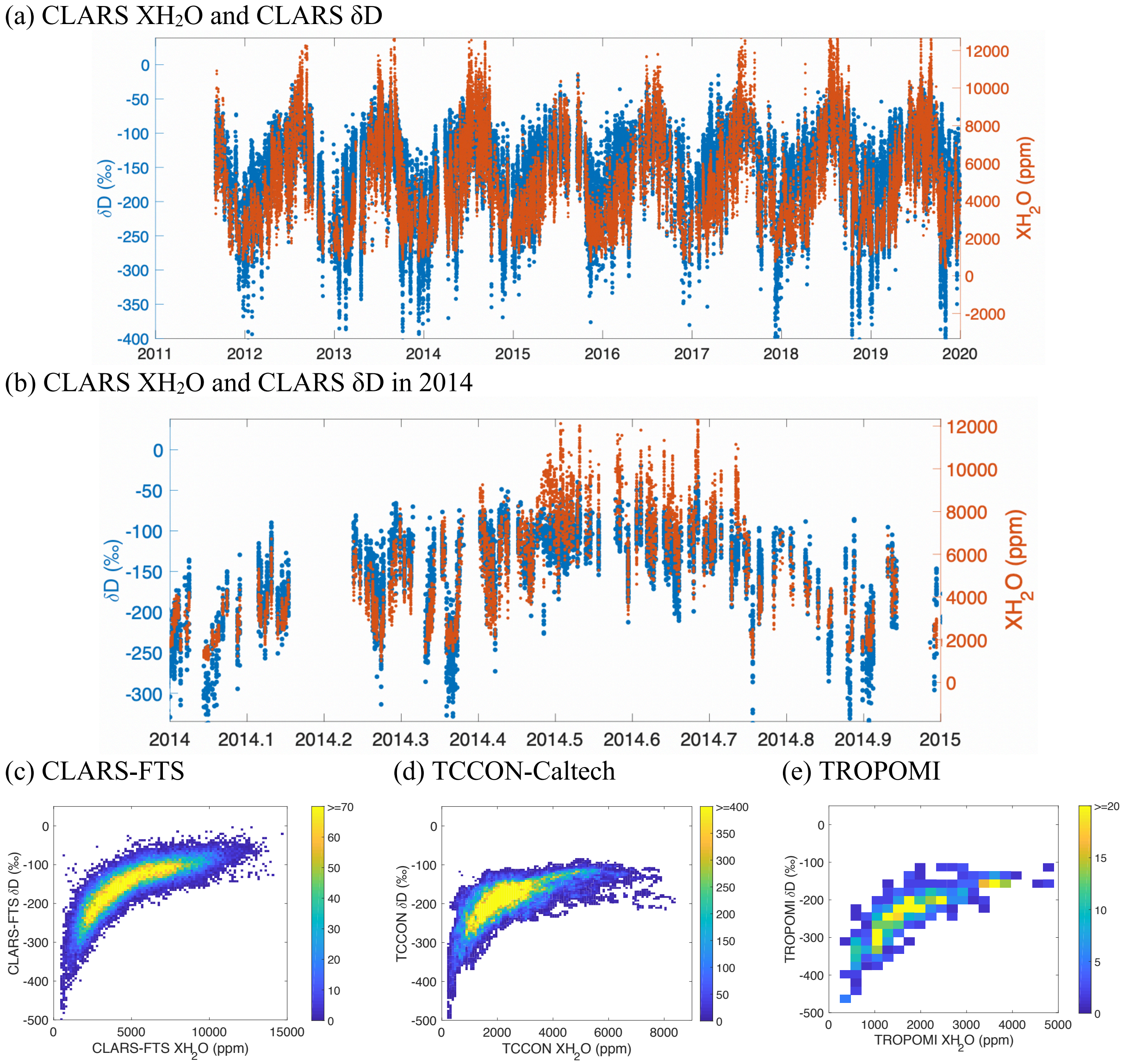


Figure 9. (a) The time series of δD and XH_2O from CLARS-FTS. For illustration purposes, some of the high XH_2O values above 12000 ppm are not shown; (b) The same as (a) but zooming into the year 2014; (c-e) Density plots between XH_2O and δD from CLARS-FTS, TCCON-Caltech, and TROPOMI data, respectively.

4. Discussion

4.1 Uncertainty in CLARS-FTS δD retrievals due to the aerosol scattering effect

Since aerosol scattering is not incorporated into the GFIT algorithm, we have used a set of tight filters (**Table 2**) to screen the data that may be affected by clouds and aerosols. For δD retrievals, the effects of aerosol scattering and surface pressure variations largely cancel out since δD is derived from the ratio of HDO and H_2O columns using spectral windows in the same wavelength range. This was demonstrated by **Boesch et al. (2013)** from a series of retrieval sensitivity tests. Here we conducted a sensitivity study for CLARS-FTS retrievals by tightening the aerosol filters and evaluating the impact on the seasonal cycles of the observed monthly averaged δD. As shown in **Figure 10**, we can observe a high consistency between retrievals with different O_2 ratio filters, as defined in **Section 3.1**. Since the O_2 ratio is an effective indicator of the aerosol scattering effect, this result indicates that the temporal variabilities in δD retrievals are only slightly affected by the impacts of aerosols. This also confirms the conclusions from **Boesch et al. (2013)**.

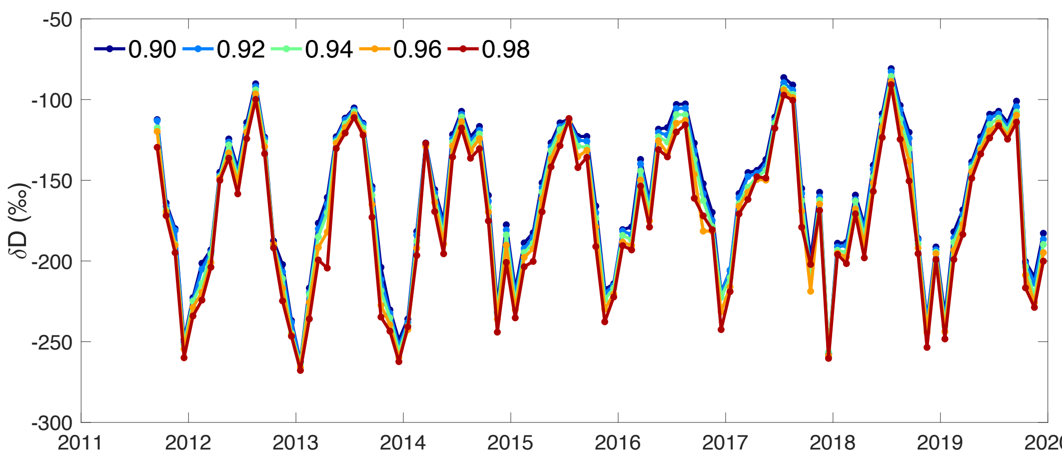


Figure 10. Time series of δD measurements from CLARS-FTS under different threshold for the O_2 ratio filters (0.90, 0.94, 0.94, 0.96, and 0.98) for screening retrievals affected by aerosol scattering effect.

4.2 Reconciling the difference between CLARS and TCCON

As shown in **Figure 8(a)**, a systematic offset is observed in δD retrievals between CLARS and TCCON. Here we specifically investigate the effects of differences in observation geometries and averaging kernels on the measured offset. The averaging kernel defines the sensitivity of the retrievals to the true column. Since different observing systems have different averaging kernels, their impacts on the retrievals should be evaluated for interpreting the retrieval discrepancies.

Our strategy to investigate the causes of the difference between CLARS and TCCON is to first construct a set of "true" H_2O, HDO and δD profiles. These "true" profiles are created based on a priori GFIT profiles, shown in **Figure 11**(a) and (b), which are then scaled in a way that they will generate similar column δD retrievals with TCCON and partial column δD retrievals with free tropospheric measurements from CLARS-SVO. We then apply the observation operators and averaging kernels (**Figure 11(c)**) from CLARS-FTS and TCCON, separately, on the "truth" profiles to simulate what (synthetic) δD values can be measured from both instruments. Finally, we assess the differences of the synthetic δD values from TCCON and CLARS and compare with the differences from real measurements.

Examples of monthly averaged a priori profiles selected from different seasons for H_2O and δD are shown in **Figure 11(a)** and **(b)**. The a priori δD profiles are generated by GFIT with assumed fractionation parameters for the troposphere and stratosphere. After the a priori profiles are scaled to match the TCCON total δD, we found they significantly overestimate (not shown here) the free troposphere partial δD column when compared with CLARS-SVO. This indicates that the free troposphere above PBL has been relatively overestimated. We therefore adjusted the HDO profiles by scaling all levels above CLARS (1.6 km) by 0.9 and the levels below CLARS by 1.1 in order to match the CLARS and TCCON observations within their uncertainties. Using these scaled profiles, we applied the CLARS-LABS observation operator and averaging kernel to generate synthetic δD values. All synthetic δD values for TCCON, CLARS-SVO, and CLARS-LABS are shown in **Figure 12**. We can see that the observed time series of δD values can be very well reproduced from the synthetic profiles. This consistency suggests that the discrepancies between CLARS and TCCON are primarily driven by the difference in the observation geometries and averaging kernels. We conducted a further experiment using the observation operator only and assuming TCCON and CLARS have the same averaging kernel. The results (not shown here) show a very small difference, which suggests that the averaging kernel has a smaller contribution than the observation operator. Moreover, the fact that in **Figure 8** and **Figure 12** summer months have higher difference (between CLARS-LABS and TCCON) than winter months is because of the geometries. Winter months have larger solar zenith angle and therefore larger air mass in the incident light path compared to reflected light path. As a result, the contribution from the reflected path becomes smaller and so CLARS-FTS and TCCON are getting closer estimates.

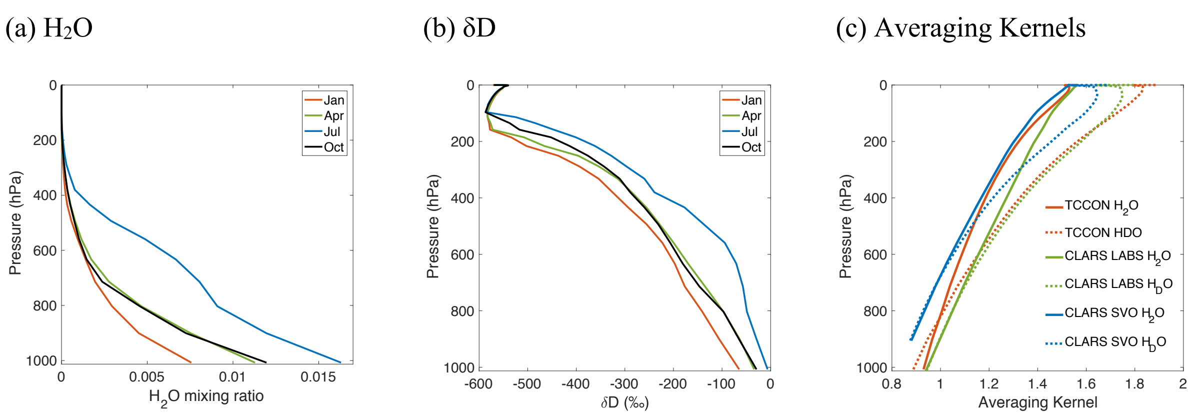


Figure 11. Examples of monthly averaged a priori profiles from GFIT selected from different seasons for (a) H_2O and (b) δD. The original a priori profiles are available on a daily basis. (c) Examples of column averaging kernels from TCCON, CLARS-SVO, and CLARS-LABS observations. These column averaging kernels are averaged profiles from all available retrieval windows. For CLARS, they are from the three HDO windows (6330.05 cm^{-1}, 6377.40 cm^{-1}, and 6458.10 cm^{-1}) and four H_2O windows (6301.35 cm^{-1}, 6392.45 cm^{-1}, 6401.15 cm^{-1}, and 6469.60 cm^{-1}). For TCCON, they are from the spectral windows as listed in Table 3 of **Wunch et al. (2015)**, which includes 15 spectral windows for H_2O and 6 spectral windows for HDO.

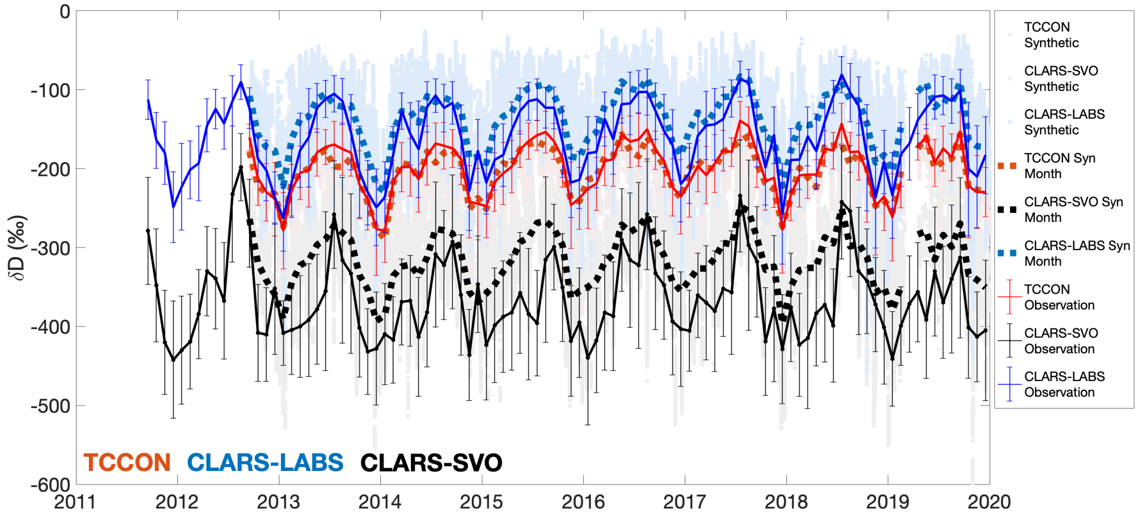


Figure 12. The reconstructed δD time series for TCCON and CLARS-FTS (including SVO and LABS modes) from applying the observation operators, which is associated with the observation geometries, and the averaging kernels to the reconstructed "truth" H_2O and HDO profiles. For comparison, the observed δD time series are also shown. The "truth" profiles are constructed from the a priori H_2O and HDO profiles generated from the GFIT program. The profiles are first scaled to match the TCCON δD retrievals. Then a scale factor (0.9) is further applied to all levels above CLARS (1.6 km) and another scale factor (1.1) to levels below CLARS in order to match the CLARS and TCCON δD observations within their uncertainties.

5. Conclusions

In this study, we retrieved XHDO, XH_2O, and δD using the high resolution NIR observations from CLARS-FTS that demonstrates high sensitivity to the PBL atmosphere in the LA basin. Diurnal observations of XHDO, XH2O, and δD with high temporal resolution (as high as 3 minutes) are generated over 33 surface reflection targets covering the LA basin from 2011 to 2019. The temporal variabilities in δD data between CLARS-FTS and a collocated TCCON observatory and TROPOMI observations are highly correlated. CLARS-FTS observes higher values due to its longer path along the PBL where HDO is more abundant. The difference between CLARS and TCCON or TROPOMI δD retrievals can be attributed to the difference primarily in the observation geometries and secondarily in the averaging kernels. From CLARS measurements, the XH_2O and XHDO time series show strong seasonal cycles that are associated with the seasonal variation of temperature that drives the evaporation. The δD shows low values in winter (more depletion of HDO) and high values in summer (less depletion of HDO), mainly driven by the change of atmospheric humidity. HDO and δD from CLARS-FTS provide high spatial and temporal resolution datasets for further study of hydrological processes in southern California.

The data set resulting from this work possesses a large amount of potential for future study. Immediate next steps include examining the spatial variability in the δD data. Analysis of spatial variability can be performed by mapping δD values according to surface reflection targets location associated with the observation. In doing so, one may be able to identify regions of the LA basin which show similar δD patterns throughout the year and therefore are potentially influenced by similar water vapor sources. If such regions can be identified, additional analysis of temporal patterns can be performed in order to attempt to understand the relative importance of various water vapor sources in that region. The time series can also be examined for correlation with other meteorological patterns, such as precipitation and wind patterns, in order to further examine climatological trends or weather anomalies in LA. Finally, there is a large amount of potential to use this data set in modeling studies regarding meteorological processes, such as cloud formation, and atmospheric water vapor sources for LA.

Acknowledgements

We would like to thank Geoffrey Toon (JPL) for proofreading and suggestions on the first draft of this manuscript, and Paul Wennberg (Caltech) and Coleen Roehl (Caltech) for providing the TCCON data. OA would like to thank the National Science Foundation COSMOS Scholarship, the Maximizing Student Potential in STEM program at the Jet Propulsion Laboratory, and Occidental College Undergraduate Research Center for helping to fund the summer experience. The CLARS project receives support from the California Air Resources Board. However, the results presented in this study do not necessarily present the views of the funding agency. Part of the CLARS-FTS data are available from the NASA Megacities Project at https://megacities.jpl.nasa.gov. The TROPOMI HDO data set from this study is available for download at ftp://ftp.sron.nl/open-access-data-2/TROPOMI/tropomi/hdo/9_1/ (last access: 12 December 2020). TCCON data are available from the TCCON Data Archive: https://doi.org/10.14291/tccon.ggg2014.pasadena01.r1/1182415. Part of the research described in this article was performed at the Jet Propulsion Laboratory, California Institute of Technology and NASA under contracts with the National Aeronautics and Space Administration.

References

Boesch, H., Deutscher, N. M., Warneke, T., Byckling, K., Cogan, A. J., Griffith, D. W. T., et al. (2013), HDO/H2O ratio retrievals from GOSAT. *Atmospheric Measurement Techniques*, 6, 599–612, doi:10.5194/amt-6-599-2013

Craig H., Gordon L. I. & Horibe Y. (1963), Isotopic exchange effects in the evaporation of water. *Journal of Geophysical Research*, 68, 5079–5087.

Craig, H. (1961), Isotopic variations in meteoric waters. *Science*, *133*(3465), pp.1702-1703.

Frankenberg, C., Yoshimura, K., Warneke, T., Aben, I., Butz, A., Deutscher, N., et al. (2009), Dynamic processes governing the isotopic composition of water vapor as observed from space and ground. *Science*, 325, 1374–1377. https://doi.org/10.1126/science.1173791

Frankenberg, C., Wunch, D., Toon, G., Risi, C., Scheepmaker, R., Lee, J.-E., et al. (2013), Water vapor isotopologue retrievals from high-resolution GOSAT short-wave infrared spectra. *Atmospheric Measurement Techniques*, 6, 263–274, https://doi.org/10.5194/amt-6-263-2013

Fu, D., Pongetti, T. J., Blavier, J.-F. L., Crawford, T. J., Manatt, K. S., Toon, G. C., et al. (2014), Near-infrared remote sensing of Los Angeles trace gas distributions from a mountaintop site. *Atmospheric Measurement Techniques*, 7, 713–729. https://doi.org/10.5194/amt-7-713-2014

Galewsky, J., Steen-Larsen, H. C., Field, R. D. Worden, J., Risi, C., & Schneider, M. (2016), Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle. *Reviews of Geophysics*, 54, 809–865, https://doi.org/10.1002/2015RG000512

González, Y., Schneider, M., Dyroff, C., Rodríguez, S., Christner, E., García, O. E., et al. (2016), Detecting moisture transport pathways to the subtropical North Atlantic free troposphere using paired H_2O−δD in situ measurements, *Atmospheric Chemistry and Physics*, 16, 4251–4269, https://doi.org/10.5194/acp-16-4251-2016

Hagemann, R., G. Nief, & E. Roth (1970), Absolute isotopic scale for deuterium analysis of natural waters. Absolute D/H ratio for SMOW. *Tellus*, 22 (6), 712–715, 1970. doi:10.3402/tellusa.v22i6

Kuang, Z., Toon, G.C., Wennberg, P.O. & Yung, Y.L. (2003), Measured HDO/H2O ratios across the tropical tropopause. *Geophysical Research Letters*, *30*(7).

Lacour, J.-L., Risi, C., Clarisse, L., Bony, S., Hurtmans, D., Clerbaux, C., et al. (2012), Mid-tropospheric δD observations from IASI/MetOp at high spatial and temporal resolution. *Atmospheric Chemistry and Physics*, 12, 10817– 10832. https://doi.org/10.5194/acp-12-10817-2012

Noone, D. (2012), Pairing Measurements of the Water Vapor Isotope Ratio with Humidity to Deduce Atmospheric Moistening and Dehydration in the Tropical Midtroposphere. *Journal of Climate*, 25, 4476–4494, https://doi.org/10.1175/JCLI-D-11-00582.1

Rayleigh, J. W. S. (1902), On the distillation of binary mixtures, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 4, 521–537, https://doi.org/10.1080/14786440209462876

Rokotyan, N. V., Zakharov, V. I., Gribanov, K. G., Schneider, M., Bréon, F.-M., Jouzel, J., et al. (2014), A posteriori calculation of $\delta^{18}O$ and δD in atmospheric water vapour from ground-based near-infrared FTIR retrievals of $H^{16}O$, $H^{18}O$, and $HD^{16}O$. *Atmospheric Measurement Techniques*, 7, 2567–2580, https://doi.org/10.5194/amt-7-2567-2014

Scheepmaker, R. A., Frankenberg, C., Deutscher, N. M., Schneider, M., Barthlott, S., Blumenstock, T., et al. (2015), Validation of SCIAMACHY HDO/H2O measurements using the TCCON and NDACC-

MUSICA networks. *Atmospheric Measurement Techniques*, 7, 11799–11851. https://doi.org/10.5194/amt-8-1799-2015

Schneider, M., Wiegele, A., Barthlot, S., Gonzálex, Christner, E., Dryoff, C., et al. (2016), Accomplishments of the MUSICA project to provide accurate, long-term, global and high-resolution observations of tropospheric {H_2O,δD} pairs – a review. *Atmospheric Measurement Techniques*, 9, 2845–2875. https://doi.org/10.5194/amt-9-2845-2016

Schneider, A., Borsdoff, T., aan de Brugh, J., Aemisegger, F., Feist, D. G., Kivi, R., et al. (2020), First data set of H_2O/HDO columns from the Tropospheric Monitoring Instrument (TROPOMI). *Atmospheric Measurement Techniques*, 13, 85–100. https://doi.org/10.5194/amt-13-85-2020.

Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., & Allen, N. T. (2015). *TCCON data from Caltech (US), Release GGG2014.R1* (Version GGG2014.R1) [Data set]. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2014.PASADENA01.R1/1182415

Wong, K. W., Fu, D., Pongetti, T. J., Newman, S., Kort, E. A., Duren, R., et al. (2015), Mapping CH4: CO2 ratios in Los Angeles with CLARS-FTS from Mount Wilson, California, *Atmospheric Chemistry and Physics*, 15, 241–252. https://doi.org/10.5194/acp-15-241-2015

Wong, C. K., Pongetti, T. J., Oda, T., Rao, P., Gurney, K. R., Newman, S., et al. (2016), Monthly trends in methane emissions in Los Angeles from 2011 to 2015 inferred by CLARS-FTS observations. *Atmospheric Chemistry and Physics*, 16, 13121-13130. https://doi.org/10.5194/acp-16-13121-2016.

Worden, J., Noone, D., & Bowman, K. (2007), Importance of rain evaporation and continental convection in the tropical water cycle. *Nature*, 445, 528–532.

Worden, J. R., Kulawik, S. S., Fu, D., Payne, V. H., Lipton, A. E., Polonsky, I., et al. (2019), Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor, *Atmospheric Measurement Techniques*, 12, 2331–2339, https://doi.org/10.5194/amt-12-2331-2019

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, X., Feist, D. G., and Wennberg, P. O. (2015), The Total Carbon Column Ob- serving Network's GGG2014 Data Version, Tech. rep., TCCON, https://doi.org/10.14291/tccon.ggg2014.documentation.R0/1221662

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., et al. (2011), The Total Carbon Column Observing Network, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240

Yoshimura, Kei (2015), Stable Water Isotopes in Climatology, Meteorology, and Hydrology: A Review. *Journal of the Meteorological Society of Japan*, 93 (5), 513-533. https://doi.org/10.2151/jmsj.2015-036.

Zakharov, V. I., Imasu, R., Gribanov, K. G., Hoffmann, G., & Jouzel, J. (2004), Latitudinal distribution of the deuterium to hydrogen ratio in the atmosphere water vapor retrieved from IMG/ADEOS data. *Geophysical Research Letters*, 31, L12104, https://doi.org/10.1029/2004GL019433.

Zeng, Z.-C., Zhang, Q., Natraj, V., Margolis, J., Shia, R.-L., Newman, S., et al. (2017), Aerosol Scattering Effects on Water Vapor Retrievals over the Los Angeles Basin. *Atmospheric Chemistry and Physicsics*, 17, 2495–2508. https://doi.org/10.5194/acp-17-2495-2017

Zeng, Z. C., Natraj, V., Xu, F., Pongetti, T. J., Shia, R. L., Kort, E. A., et al. (2018), Constraining Aerosol Vertical Profile in the Boundary Layer Using Hyperspectral Measurements of Oxygen Absorption. *Geophysical Research Letters,* 45 (19), 10-772-10-778. https://doi.org/10.1029/2018GL079286

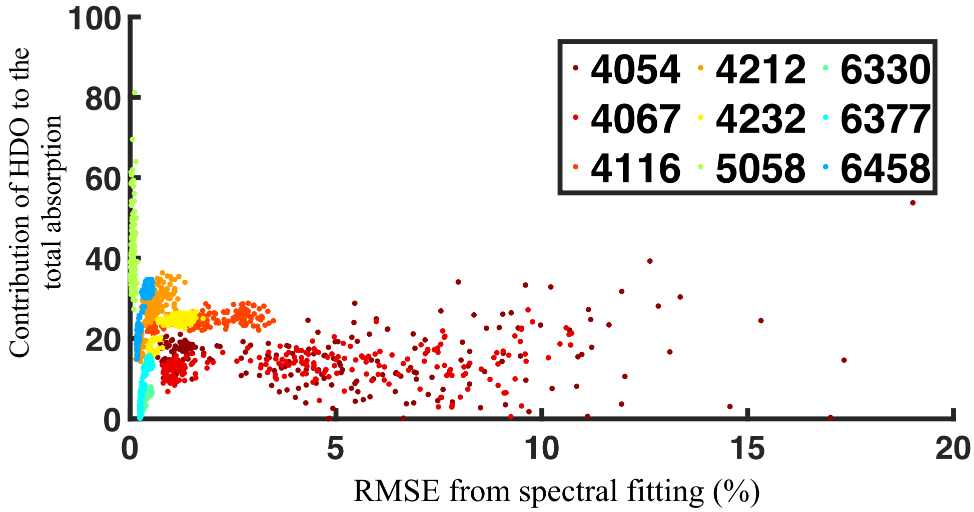
Zeng, Z. C., Xu, F., Natraj, V., Pongetti, T. J., Shia, R. L., Zhang, Q, et al. (2020), Remote Sensing of Angular Scattering Effect of Aerosols in a North America Megacity. Remote Sensing of Environment, 242, 11760. https://doi.org/10.1016/j.rse.2020.111760

Appendix A: Comparison of Spectral Windows

We tested 9 spectral windows (**Table A1**) using CLARS-FTS spectral observations on a sample of 4-6 days in which the instrument took measurements for two surface reflection point locations throughout the course of the day. The results from each spectral window were compared according to overall quality of spectral fittings and resulting retrieval uncertainties. In addition to recommended TCCON windows, a micro-window around 5058 cm^{-1}, was tested based on a survey of successful spectral windows from prior TCCON studies (**Rokotyan** et al., 2014). **Figure A1** shows a plot of the RMSE from the spectral fitting for each spectral window versus the contribution from HDO to the overall absorption using a set of ~4000 observations from 4 distinct days in 2013. From this plot, one can see that the lower frequency spectral windows, which are plotted red/orange/yellow colors, have higher spectral fitting error, implying a worse quality fit for these spectral windows. Comparing VSF errors for this set observations, the lower spectral windows also tended to yield observations with higher overall uncertainties. Therefore, the spectral windows in the 4000-5000 cm^{-1} were eliminated as possible candidates. The green points, referring to the 5058 cm^{-1} micro-window, indicate lower spectral fitting error and also relatively high contribution due to HDO. However, examining the spectra for this micro-window indicated that it contains only one main absorption line, compared to the 6000-7000 cm^{-1} spectral windows, which contain many more. Therefore, the results using 6000-7000 cm^{-1} were considered more robust. From this examination of spectral fit quality from the 9 tested spectral windows, the three spectral windows with central wavenumbers in the 6000-7000 cm^{-1} range were selected as primary candidates for further investigation.

Table A1. Tested Spectral Windows with Associated Parameters

Center (cm^{-1})	Width (cm^{-1})	Gases to fit	continuum basis functions
4054.60	3.30	HDO, H_2O, CH_4	2-order polynomial for continuum; frequency shift, solar lines
4067.60	8.80	HDO, H_2O, CH_4	
4116.10	8.00	HDO, H_2O, CH_4	
4212.45	1.90	HDO, H_2O, CH_4	
4232.50	11.00	HDO, H_2O, CH_4, CO	
5058.95	1.60	HDO, H_2O, CO_2	
6330.05	45.50	HDO, H_2O, CO_2	
6377.40	50.20	HDO, H_2O, CO_2	
6458.10	41.40	HDO, H_2O, CO_2	



743 **Figure A1**: A plot of RMSE from the spectral fitting versus the contribution from HDO to the
744 overall absorption for a sample for ~4000 observations from 4 days in 2013. Results from
745 various spectral windows are indicated by color moving from red to blue as the central
746 wavenumber of the spectral window increases.

Appendix B. H_2O spectral windows and their fitting residuals and retrieval errors

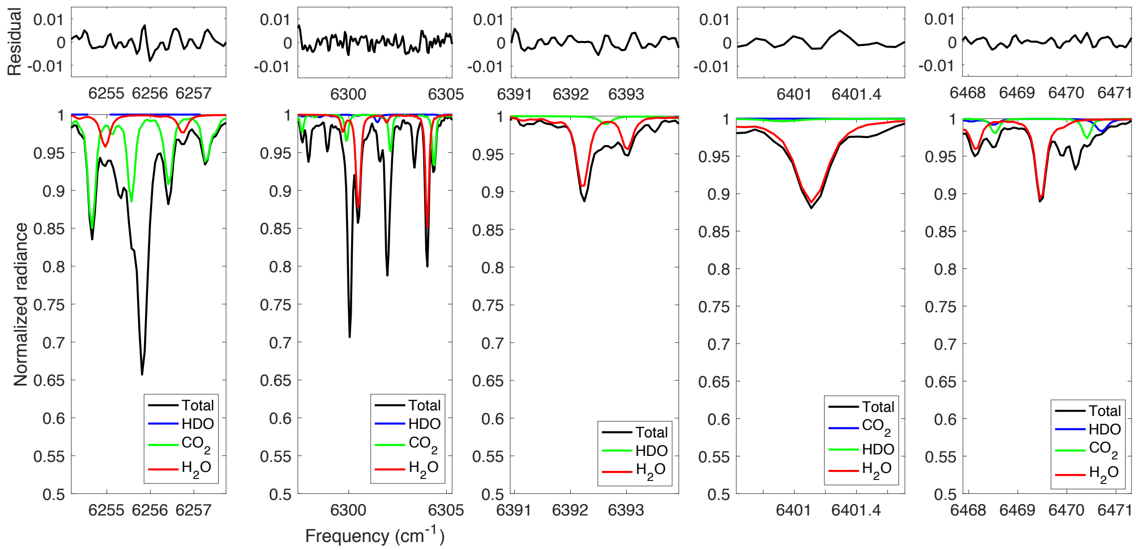


Figure B1. Examples of CLARS-FTS spectra windows of (from left to right) 6255.95 cm^{-1}, 6301.35 cm^{-1}, 6392.45 cm^{-1}, 6401.15 cm^{-1}, and 6469.60 cm^{-1} for retrieving H_2O in this study. These samples of normalized spectra are taken from a mid-day observation on 7/14/2013 over the West Pasadena surface target. The lower panel shows the full spectral fit with total contribution, contribution from H_2O, and from other interfering gases. The upper label shows the residuals of the spectral fits, defined as the difference in total measured and total calculated radiance.

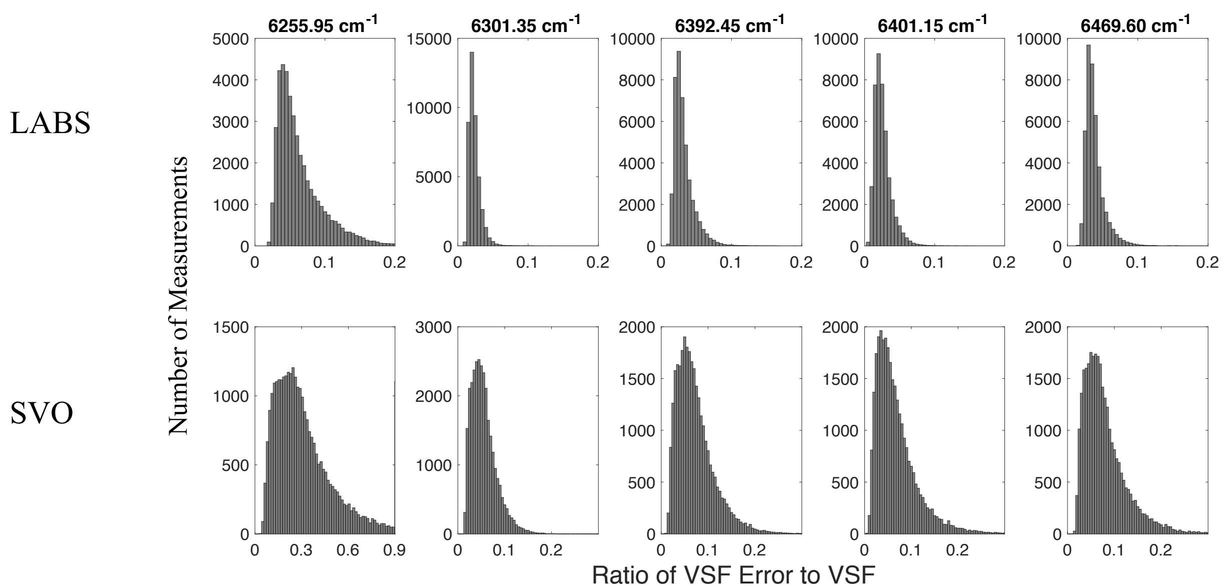


Figure B2. Retrieval error, in ratio of VSF error to VSF, for the entire filtered XH_2O datasets from 2011 to 2019. Both VSF and VSF error values are calculated from CLARS-GFIT. The data are separated into histograms according to observation modes (LABS and SVO) and five spectral windows (6255.95 cm^{-1}, 6301.35 cm^{-1}, 6392.45 cm^{-1}, 6401.15 cm^{-1}, and 6469.60 cm^{-1}). Note that the SVO at 6255.95 cm^{-1} has a different x-axis range.

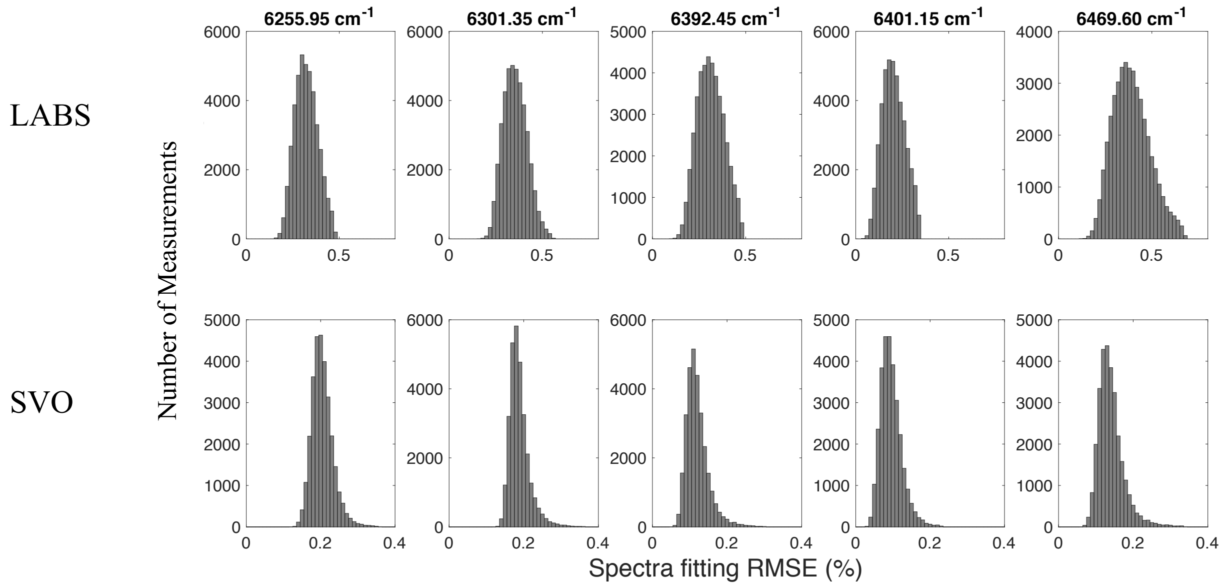


Figure B3. Histograms of RMS error from spectral fitting for the entire data set from 2011 to 2019. The data are separated into histograms according to observation mode (LABS and SVO) and five spectral windows (6255.95 cm^{-1}, 6301.35 cm^{-1}, 6392.45 cm^{-1}, 6401.15 cm^{-1}, and 6469.60 cm^{-1}).