# Water vapor absorption spectroscopy and validation tests of databases in the far-infrared (50-720 cm -1 ). Part 1: Natural water

M. Toureille, A.O. Koroleva, S.N. Mikhailenko, O. Pirali, Alain Campargue

# Water vapor absorption spectroscopy and validation tests of databases in the far-infrared (50-720 $cm^{-1}$). Part 1: Natural water

M. Toureille[1], A.O. Koroleva[1,2], S.N. Mikhailenko[3,4], O. Pirali[5,6], and A. Campargue[1*]

[1] Univ. Grenoble Alpes, CNRS, LIPhy, 38000 Grenoble, France
[2] Institute of Applied Physics, Russian Academy of Sciences, Nizhniy Novgorod, Russia
[3] V.E. Zuev Institute of Atmospheric Optics, SB, Russian Academy of Science, 1, Academician Zuev square, 634055 Tomsk, Russia
[4] Climate and Environmental Physics Laboratory, Ural Federal University, 19, Mira av., 620002 Yekaterinburg, Russia
[5] SOLEIL Synchrotron, L'Orme des Merisiers, Saint-Aubin 91192, Gif-Sur-Yvette, France
[6] Université Paris-Saclay, CNRS, Institut des Sciences Moléculaires d'Orsay, 91405 Orsay, France

Key words: water vapor; far infrared; rotational spectrum; water isotope

* Corresponding author: Alain Campargue (alain.campargue@univ-grenoble-alpes.fr)

**Abstract**

The rotational spectrum of water vapor in natural isotopic abundance has been recorded by high resolution (≈ 0.001 $cm^{-1}$) Fourier transform spectroscopy at the AILES beam line of the SOLEIL synchrotron. The room temperature absorption spectrum has been recorded between 50 and 720 $cm^{-1}$ using five pressure values up to 7 mbar and an absorption pathlength of 151.75 m. Line parameters were retrieved for the five recorded spectra and then combined in a global list of 2867 water lines with line intensity ranging between a few $10^{-26}$ and $10^{-19}$ cm/molecule. 454 of the measured lines are newly observed by absorption spectroscopy. The spectral calibration based on a statistical matching with about 700 accurate reference line positions allows for line center determinations with an accuracy of $5\times10^{-5}$ $cm^{-1}$ for well isolated lines of intermediate intensity.

The large spectral coverage, the achieved position accuracy and sensitivity of the constructed line list make it valuable for validation tests of the current spectroscopic databases. Six water isotopologues ($H_2{}^{18}O$, $H_2{}^{16}O$, $H_2{}^{17}O$, $HD^{18}O$, $HD^{16}O$, and $HD^{17}O$) were found to contribute to the spectrum. The line position comparison to the recent HITRAN2020 spectroscopic database and to the W2020 line lists of $H_2{}^{16}O$, $H_2{}^{17}O$ and $H_2{}^{18}O$, [Furtenbacher et al. J. Phys. Chem. Ref. Data 49 (2020) 043103; https://doi.org/10.1063/5.0030680] shows an overall very good agreement. Nevertheless, a number of significant deviations are observed. Part of them has an amplitude largely exceeding the W2020 claimed error bars. On the basis of the experimental data at disposal for the main isotopologue (1310 transitions), the best agreement is achieved with the positions calculated using the effective Bending–Rotation Hamiltonian [Coudert et al. J Mol Spectrosc 2014;303:36-41. https://doi.org/10.1016/j.jms.2014.07.003].

## 1. Introduction

The present work is devoted to an experimental study of the high resolution absorption spectrum of natural water vapor in the far infrared region (FIR) between 50 and 720 $cm^{-1}$. It follows a recent study of the rotational spectrum of water vapor highly enriched in $^{18}O$ in the same region [1]. In fact, the $^{18}O$ enriched spectrum analyzed in Ref. [1] was a "side-product" of a large measurement campaign performed in 2018 at the AILES beam line of the SOLEIL synchrotron and dedicated to the characterization of the water vapor FIR absorption continuum by Fourier transform spectroscopy (FTS) [2,3]. It turns out that, while we believed that the rotational spectrum involving the lowest vibrational states of water vapor was well characterized, some of the recorded SOLEIL high resolution spectra of $^{18}O$ water vapor allowed for a substantial extension of the previous knowledge in the region, in addition to provide valuable tests of existing databases. This is mainly the result of the unique characteristics of the AILES beamline of the SOLEIL synchrotron light source in the "difficult" FIR spectral region where the performances of laboratory spectrometers are usually not optimum (see the review of various experimental approaches for FIR absorption spectroscopy in the introduction of Ref. [1]). The combination of the synchrotron radiation with a long path absorption cell allows for broad-band high sensitivity FIR recordings with a spectral resolution of about to 0.001 $cm^{-1}$. Note that in the case of the $H_2{}^{18}O$ transitions assigned in [1], the weakest measured lines have intensity on the order of $10^{-25}$ cm/molecule which corresponds to a gain of 4 orders of magnitude of the detectivity threshold compared to previous $H_2{}^{18}O$ absorption measurements.

The importance of the FIR region in the Earth's radiation budget is a major motivation of the Far-infrared-Outgoing-Radiation Understanding and Monitoring (FORUM) mission of European Space Agency (https://www.forum-ee9.eu/). One of the identified goals of this mission is to "fill the observational gap across the far-infrared (from 100 to 667 $cm^{-1}$), never before sounded in its entirety from space". This requires an accurate knowledge of the water spectroscopy in the region and motivated us to apply for beam time at SOLEIL synchrotron for a new measurement campaign dedicated to the high resolution spectroscopy of various water isotopologues (natural, $^{17}O$ and D enriched) using the same experimental setup as for the $^{18}O$ enriched water recordings. The present contribution dedicated to the analysis and discussion of the natural water spectra recorded at five pressures up to 7 mbar, is the first one of the series.

The paper is organized as follows. In Section 2, we recall the experimental details including the spectrum acquisition, line list construction and spectra calibration. The main results are presented in Section 3 which presents an overview of the transitions contributing to the spectra together with a line position comparison with literature. In particular, we will consider *(i)* the HITRAN2020 spectroscopic database [4], *(ii)* the W2020 line lists of $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$ with line positions computed from

empirically determined energy levels [5] and, *(iii)* the calculations by Coudert et al. for the $H_2{}^{16}O$ main isotopologue using an effective Bending–Rotation Hamiltonian [6].

## 2. Experiment

### *2.1. Spectra acquisition*

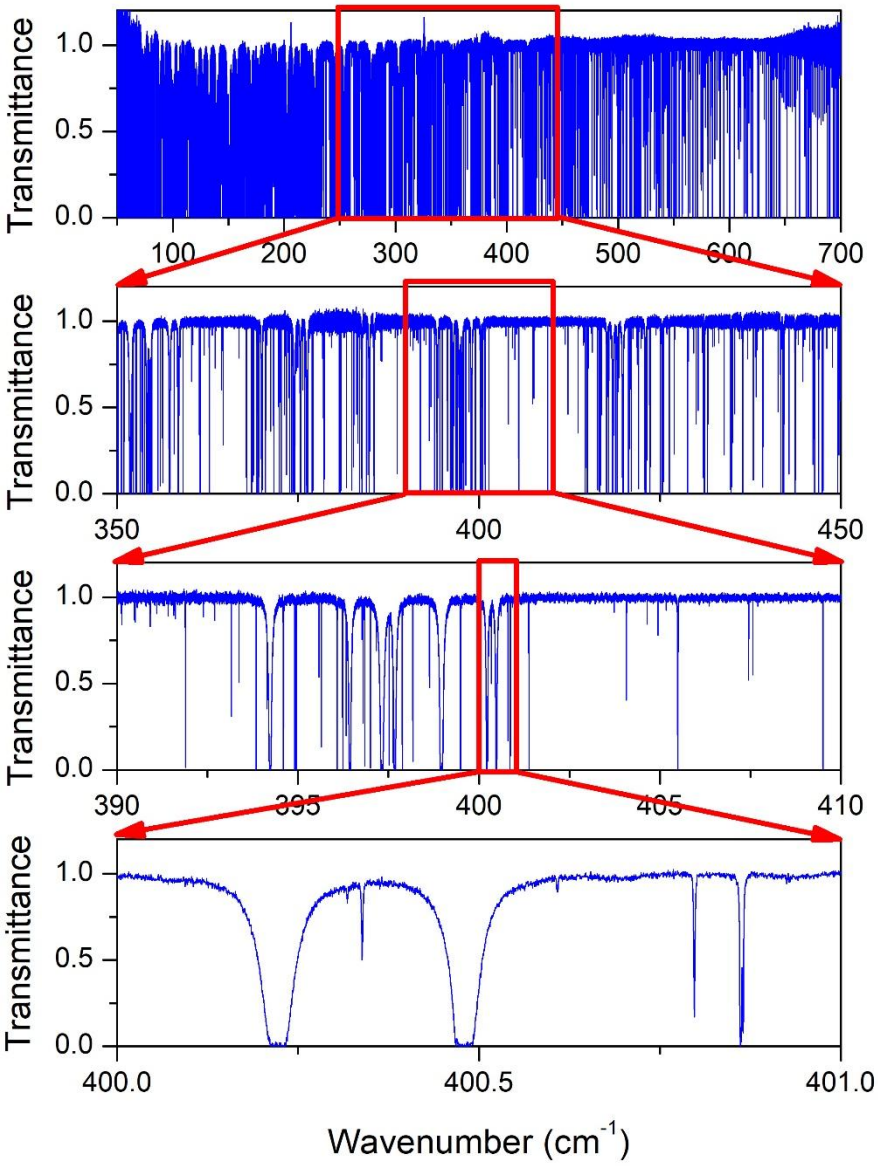


**Fig. 1**
Successive zooms of the FTS spectrum #3 of natural water vapor recorded at SOLEIL synchrotron at room temperature ($P$= 0.91 mbar) between 50 and 720 $cm^{-1}$.

The FTS spectra were acquired on the AILES beam line of SOLEIL synchrotron facility operated in the 500 mA multibunch mode, in September 2021. A Bruker 125 interferometer with a 6 µm mylar-composite beam splitter and a 4 K cooled Si bolometer detector were used for the recordings. The

absorption cell is a multipass cell in White-type configuration. The total absorption path length was set to 151.75±1.5 m corresponding to 60 passes between mirrors separated by 2.52 m and about 0.5 m of space between the 50 µm thick polypropylene films windows. Five spectra were recorded for different pressure values up to 7 mbar, measured by a capacitance gauge (Pfeiffer 10 mbar full range with corresponding accuracy of 0.01 mbar). The de-ionized water used for the recordings was frozen and liquefied several times before injection in the cell in order to minimize the presence of gas impurities. **Table 1** summarizes the experimental conditions and the sequence of the recordings.

**Table 1**

Experimental conditions of the five FTS spectra of natural water under analysis. The temperature was 295.5 K.

| | Sample | Pressure | Resolution, $cm^{-1}$ | Nb of scans |
|---|---|---|---|---|
| **#1** | Empty cell | ≈ 1.7 µbar | 0.001 | 320 |
| **#2** | Baseline<br>**Sample** | Pumping on the cell<br>**≈ 0.080 mbar** | 0.05<br>0.001 | 200<br>400 |
| **#3** | Baseline<br>**Sample** | Pumping on the cell<br>**≈ 0.91 mbar** | 0.05<br>0.001 | 200<br>480 |
| **#4** | Baseline<br>**Sample** | Pumping on the cell<br>**≈ 7 mbar** | 0.05<br>0.002 | 200<br>280 |
| **#5** | Empty cell | ≈ 0.15 µbar | 0.001 | 220 |

The first and last spectra were recorded at very low pressure, pumping on the cell in order to measure the strongest lines (intensity up to $10^{-18}$ cm/molecule) and to check the stability of the frequency scale (see below). The other pressure values were set to about 0.085, 0.91 and 7.0 mbar. Except for the 7 mbar spectrum for which the pressure broadening allows for using a spectral resolution of 0.002 $cm^{-1}$, the other spectra were recorded at the maximum spectral resolution of 0.00102 $cm^{-1}$ (defined as 0.9/MOPD where MOPD= 882 cm is the maximum optical path difference). No apodization of the interferogram was used (boxcar option of the Bruker software). The number of co-added spectra ranges between 200 and 480 (200 spectra corresponds to about 10 hours' acquisition at 0.001 $cm^{-1}$ spectral resolution or 5 hours at 0.002 $cm^{-1}$ resolution). The baseline fluctuations were corrected by division by a lower resolution (0.05 $cm^{-1}$) spectrum acquired prior to each high resolution recording. The temperature of 295.5(3) K was monitored by a pair of platinum sensors mounted on the cell external surface. An overview of the spectrum recorded at 0.91 mbar is displayed on **Fig. 1**, which includes successive zooms. The absorption coefficient was determined as $\alpha_{total} = 1/L \ln(I_0(\nu)/I(\nu))$, where $L$= 151.75 m, and $I(\nu)$ and $I_0(\nu)$ correspond to the spectrum with the cell filled with water vapor and evacuated, respectively.

The observed line profiles result from different contributions. At 1 mbar, the pressure broadening (about $4\times10^{-4}$ $cm^{-1}$ HWHM [4]) is equivalent to the width of the apparatus function (about $3.5\times10^{-4}$ $cm^{-1}$ HWHM) while the Doppler broadening (proportional to the transition frequency) is on the order of

$1.5\times10^{-4}$ $cm^{-1}$ HWHM near 100 $cm^{-1}$. The present study being mainly focused on line positions, each of the five transmittance spectra was fitted independently assuming the standard Voigt line profile as line shape (with adjusted Gaussian and Lorentzian widths) and no particular care was taken for the treatment of the apparatus function. The line parameters retrieval was performed using a homemade multiline fitting program in LabVIEW and C++. The HITRAN2020 line list was taken as starting point of the fit. For each spectrum, we selected a ten of isolated lines of intermediate intensity and adjusted their position, area, Gaussian and Lorentzian widths. The obtained Gaussian widths are larger than the Doppler broadening as they include an important contribution of the apparatus function. The average value of the fitted Gaussian widths was adopted as default value for all the fitted lines of a given spectrum. **Fig. 2** illustrates the adopted procedure in a small spectral interval showing a very strong doublet (intensity larger than $10^{-19}$ cm/molecule) near 303 $cm^{-1}$ for the spectra #1 (very low pressure) and #3 (0.91 mbar). During the fit of the 0.91 mbar spectrum, the doublet was omitted and line parameters of narrow lines were obtained from a local fit, even when these lines were located on the wing of the strong doublet. The line center of the two components of the doublet could be determined from a fit of the very low pressure spectrum (#1). The (obs. – calc.) residuals included in **Fig. 2** are at the noise level (~ 1%). This value corresponds to a noise equivalent absorption on the order of $7\times10^{-7}$ $cm^{-1}$ and a detectivity threshold of about $10^{-25}$ cm/molecule for the line intensities measured in the 0.91 mbar spectrum.

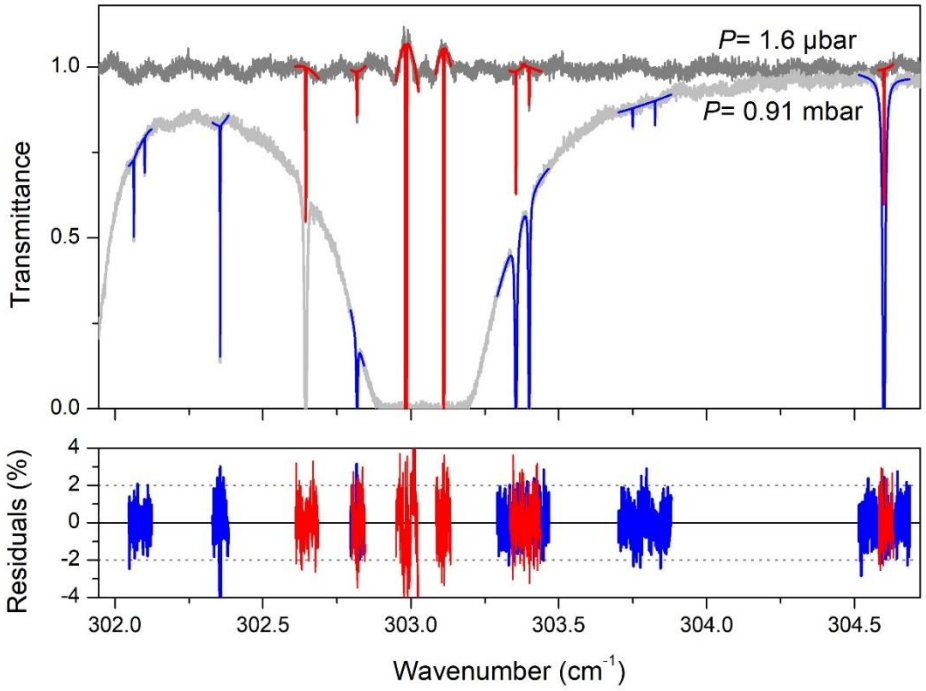


**Fig. 2**
Line parameter retrieval from the FTS spectra of water vapor near 303 $cm^{-1}$. The line profile fit was performed in narrow spectral intervals around the lines which are not too saturated.

*Upper panel:* Recorded spectra #1 and #3 at ~ 1.6 µbar and 0.91 mbar, respectively, (gray and light gray, respectively) with corresponding best fit spectra (red and blue, respectively).
*Lower panel:* Corresponding (exp. – fit) residuals in %.

Overall, line parameters of 804, 1778, 1987, 1367 and 306 absorption features were retrieved for the spectra #1 to #5, respectively. In order to obtain a unique global line list, for each line, we selected the pressure condition of the best parameter determination of the considered line (intermediate transmittance value, limited overlapping with nearby lines) and kept the corresponding values for the global list. For a large fraction of lines, two spectra could be selected and the average line position and line intensity were adopted for the global list. In the list provided as Supplementary Material, the spectrum number (1 to 5, see **Table 1**) is attached to each line indicating the spectrum or spectra used for the retrieval. On the overview of the global list presented in the upper panel of **Fig. 3**, different colors are used according to the source of the line parameters.

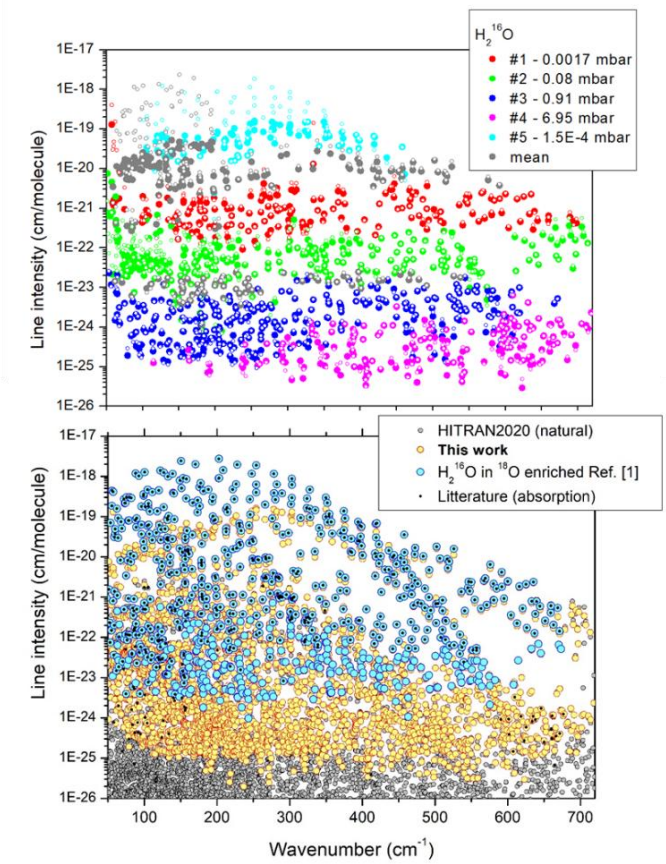


**Fig. 3.**
Overview of different line lists for water vapor between 40 and 720 $cm^{-1}$.
*Upper panel:* global line list constructed in this work from five spectra of natural water recorded at SOLEIL synchrotron at different pressures. Different colors are used for the spectrum chosen as source of the line parameters. Grey symbols mark the lines for which the position and intensity correspond to values averaged from two spectra. Small open circles correspond to the HITRAN list.
*Lower panel:* comparison of our global list (yellow dots) to *(i)* previous measurements available in the literature by absorption spectroscopy [7-20] (black dots), *(ii)* $H_2{}^{16}O$ transitions identified in the spectrum of $^{18}O$ enriched water

recorded at SOLEIL synchrotron in [1] (blue dots), *(iii)* present recordings for natural water (yellow dots) and *(iv)* the HITRAN2020 line list [4] (grey dots).

During the data treatment, a number of lines were found in addition to those included in the HITRAN list of water vapor. They were identified as due to three impurity species: $CO_2$, $NH_3$ and HF. The $CO_2$ lines are located in the high energy part of the recorded region above 640 $cm^{-1}$ and belong to the $\nu_2$ bending band. The relative abundance of $CO_2$ was found to decrease sharply (from 5000 to 5 ppm) with the water vapor pressure indicating that it is due to a small leak or desorption. The HF and ammonia lines are rotational lines. About six HF lines are observed between 100 and 350 $cm^{-1}$ while ammonia lines are observed below 200 $cm^{-1}$. Ammonia with a relative concentration of a few ppm is believed to be present as an impurity in the used water sample. HF is probably due to degassing from the O-ring seals of the cell.

After removal of the impurity lines and combination of the five datasets, the obtained global list includes 2867 water lines with intensities above $2\times10^{-26}$ cm/molecule (see overview in **Fig. 3**). Note that in the case of the spectra #1 and #5, the pressure was too small to be measured with the used pressure gauge. Intensity comparison to the HITRAN values allowed to estimate the corresponding pressure values to about 1.7 µbar and 0.15 µbar, respectively. These pressure values were used in the calculation of the line intensities derived from these two spectra. As a result of the large pressure range of the recordings (from 0.15 µbar to 7 mbar), the obtained experimental list includes line intensities spanning six orders of magnitude.

*2.2. Frequency calibration*

Before combining the five line lists into a single global list, we checked that no significant variation of the frequency calibration of the different spectra was observable: the position differences compared to HITRAN were found independent on the spectrum (see left panel of **Fig. 4**). The absolute frequency calibration was performed considering line positions provided in the HITRAN database with an uncertainty better than $10^{-5}$ $cm^{-1}$. A total of about 700 isolated lines with good signal-to-noise ratio were selected in the different recorded spectra and matched to the HITRAN reference lines. Note that the 7 mbar spectrum for which pressure shifts on the line positions might be significant was excluded. According to Refs. [21-27], the self-pressure shifts for rotational lines are between -0.058 and +0.045 $cm^{-1}$/atm. So, it means that at the maximum pressure of 7 mbar in our measurements, the line shifts can reach $\pm4\times10^{-4}$ $cm^{-1}$, a value larger than the precision of the line center determination in the 7 mbar spectrum. The differences between the experimental line centers and the HITRAN values were globally linearly fitted. Excluding a ten of outliers, an *rms* deviation of $2.42\times10^{-5}$ $cm^{-1}$ was obtained for the linear fit, thus larger than the accuracy of the position of the reference lines. The obtained empirical correction of the frequencies $(+8.4(6)\times10^{-5}-6.30(7)\times10^{-7}\sigma$, where $\sigma$ is the measured wavenumber in $cm^{-1}$) was applied to all the values of the global list. A value of $5\times10^{-5}$ $cm^{-1}$ seems to be a reasonable estimate for the resulting

uncertainty on the experimental position values of the "good" lines (isolated unsaturated lines measured in the low pressure spectra). In the global line list provided as Supplementary Material, the fit error on the line position determination is included. For a significant fraction of the lines, the fit uncertainty (thus excluding the frequency calibration error) was found smaller than $5\times10^{-5}$ cm$^{-1}$ and is thus believed to largely underestimate the real uncertainty on the line position. For all these lines, we replaced the fit uncertainty by a value of $5\times10^{-5}$ cm$^{-1}$. Note that the uncertainty of the weakest or highly blended lines can reach a value of $1\times10^{-3}$ cm$^{-1}$ in the worst cases.

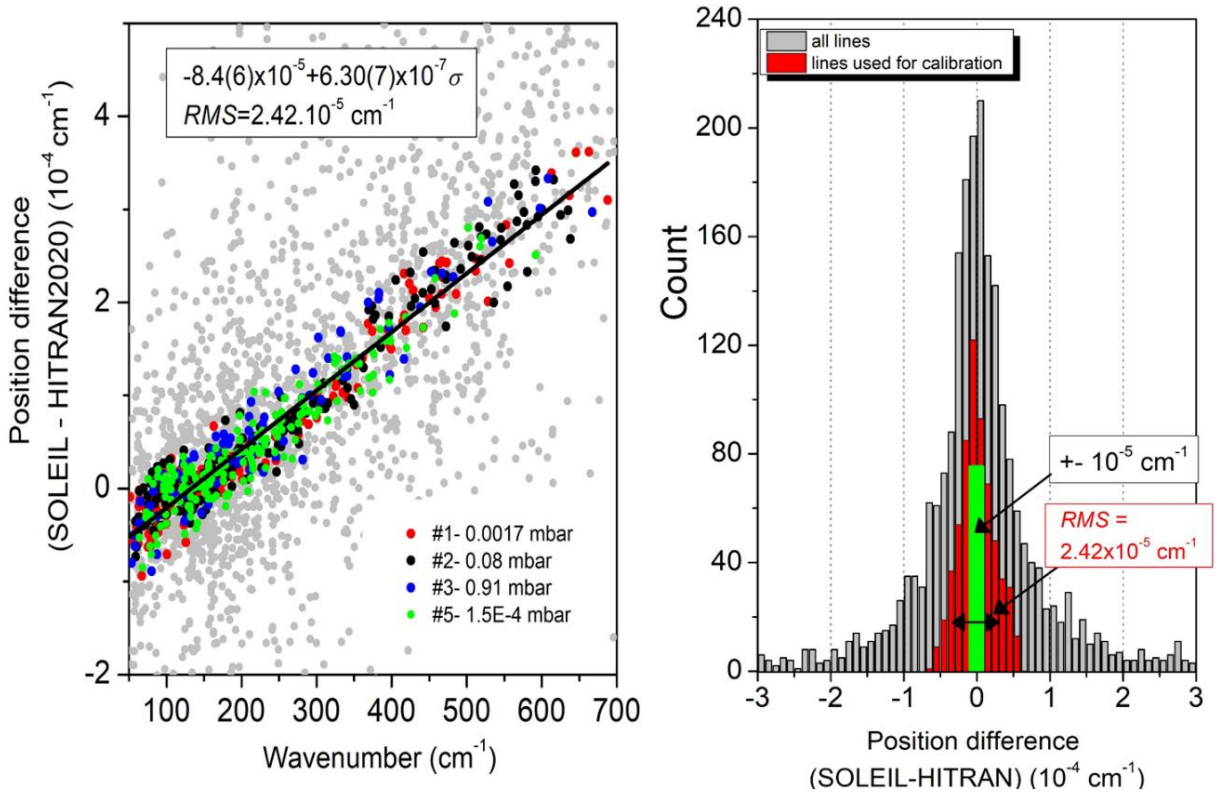


**Fig. 4**
Frequency calibration of the FTS spectra of natural water recorded at SOLEIL synchrotron.
*Left panel:* The colored symbols correspond to the differences between the line centers retrieved from the different recorded spectra and reference line positions provided with an uncertainty better than $10^{-5}$ cm$^{-1}$ in the HITRAN database [4]. The 7 mbar spectrum for which pressure shifts on the line positions might be significant is excluded. The global dataset was fitted by a linear function providing the frequency correction to be applied to the measurements. Grey symbols correspond to the comparison to all the HITRAN line positions (*rms* deviation of $2.42\times10^{-5}$ cm$^{-1}$),
*Right panel:* Histograms of the (meas. – HITRAN) position differences for the lines used for the frequency calibration (red) and for the whole set of measurements. The green bars correspond to the HITRAN uncertainty of the used reference lines ($\pm10^{-5}$ cm$^{-1}$).

The histogram of the position differences is presented in **Fig. 4** for the lines used for the frequency calibration and for the whole set of measurements. A significant fraction of the measured lines shows position differences larger than our experimental error bar. In the following, we will examine these lines in order to determine the reasons for these discrepancies.

As concerns line intensities, we do not claim for a high accuracy of the reported values. The limited number of points describing the line profile and the impact of the apparatus function which was roughly taken into account in the line parameter retrieval limit the accuracy of the retrieved values. The fit uncertainty included in the global line list is only indicative. Comparisons to HITRAN intensity values indicate that, excluding the two lowest pressure spectra (#1 and #5), most of the line intensities agree with HITRAN values within ±10%. The comparison of the fitted intensities to HITRAN values shows deviations largely exceeding the fit uncertainty. Relying on HITRAN values, we conclude that the fit error of our experimental intensities is probably underestimated for a small fraction of the measurements. Our accuracy was nevertheless sufficient to evidence a clear HDO enrichment in the spectra, probably due to a previous contamination by deuterated species adsorbed on the cell walls. Compared to the natural abundance, the enrichment was larger by about a factor of two for the first recorded spectra and then, as a consequence of the successive fillings and evacuations of the cell, gradually decreased to reach practically the natural value for the last recording. In the case of the most intense lines which were retrieved from spectra #1 and #5 (intensity up to $10^{-18}$ cm/molecule), saturation effects are considerable and the fitted values of the intensities are strongly underestimated by a factor up to 10 in the worst cases (see **Fig. 3**).

## 3. Analysis and comparison with literature

The HITRAN2020 rovibrational assignments of practically all observed water transitions have been validated and transferred to our list. The HITRAN rovibrational assignment of five high $J$ transitions in the pure band of $H_2{}^{17}O$ (from Lodi and Tennyson [28]) is incomplete in the HITRAN list [4]. The complete assignment is provided in our list. Overall, the 2867 measured water lines were assigned to 3001 transitions of six isotopologues ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, $HD^{16}O$, $HD^{18}O$, and $HD^{17}O$). **Table 2** presents some global characteristics of the assigned transitions, including the maximum values of the $J$ and $K_a$ rotation numbers and the numbers of transitions firstly observed compared to previous absorption studies available in the literature. Overall 454 transitions are newly observed by absorption. The list of these newly observed lines is provided as a second Supplementary Material. Part of these lines (306) was previously observed by emission spectroscopy (mostly in [19] and [29] – see below). The line position comparison to the most relevant emission study is included the Supplementary Material.

**Table 2**

Statistics of the rovibrational assignments of the five spectra of natural water recorded at SOLEIL at various pressures up to 7 mbar.

| Molecule | $NT$ [a] | $NT_{new}$ [b] | $J$ | $K_a$ | Range, $cm^{-1}$ |
|---|---|---|---|---|---|
| $H_2{}^{16}O$ | 1310 | 96 (385) | 22 | 14 | 51.43 – 718.43 |
| $H_2{}^{18}O$ | 564 | 3 (3) | 18 | 12 | 53.57 – 702.59 |
| $H_2{}^{17}O$ | 407 | 3 (3) | 17 | 10 | 53.51 – 694.38 |
| $HD^{16}O$ | 674 | 46 (63) | 19 | 12 | 50.27 – 591.88 |
| $HD^{18}O$ | 42 | | 10 | 7 | 78.51 – 310.10 |
| $HD^{17}O$ | 4 | | 6 | 4 | 168.17 – 255.19 |

| | | | | |
|---|---|---|---|---|
| Total | 3001 | 148 (454) | 22 14 | 50.27 – 718.43 |

Notes

[a] Number of assigned transitions

[b] Number of newly observed transitions. The first number considers only previous studies by absorption. The number between parenthesis considers all previous studies (absorption and emission).

### *3.1. Overview of previous experimental studies in the region*

The review of previous works on the rotation and vibration-rotation transitions of water in our region includes studies by absorption [7-20] and by emission spectroscopy in flames or discharges [6,19,29-34]. Overall, 686 $H_2{}^{16}O$ transitions were recorded in absorption [7-20] for the five lowest states – (000), (010), (020), (001), and (100) – up to maximum value of the rotation number $J_{max} = 17$. About 240 transitions were newly measured by absorption in Ref. [1], leaving 375 new observations for the present study. More than 6000 transitions involving 14 vibration states were assigned in emission spectra [6,19,29-33] in our region. The maximum value of the rotation number of the ground state (000) is $J_{max} = 41$ [33]. The overall position accuracy of emission studies is generally worse than the present accuracy. The most precise set of line positions reported by emission spectroscopy [6,19,29] have an accuracy $4.0\times10^{-4}$ cm$^{-1}$ at best.

In the spectrum of water vapor highly enriched in $^{18}O$ recorded at SOLEIL synchrotron with the same setup as the one used for the present recordings, 37 additional $H_2{}^{16}O$ transitions up to $J_{max} = 19$ were reported [1]. Compared to this study, the average value of the position differences is $3.44\times10^{-5}$ cm$^{-1}$ with an *rms* deviation of $13.37\times10^{-5}$ cm$^{-1}$ for 604 transitions in common in the two studies. Taken into account that the $H_2{}^{16}O$ relative abundance was only 5% in the spectrum analyzed in Ref. [1], the achieved agreement is very satisfactory and illustrates the consistency of the independent frequency calibration of the two spectra.

The observations relative to the $H_2{}^{18}O$ minor isotopologue are more limited. Overall, 525 (mainly pure rotation) transitions were reported in seven absorption [7,9,13,34-37] and one emission [38] studies. Only six rotation transitions for the (010), (020) and (100) states were measured in these previous studies [9,34]. The most extended $H_2{}^{18}O$ line list (more than 730 transitions) was obtained from the SOLEIL spectrum analyzed in Ref. [1]. It includes more than 300 transitions of the (010) – (010) and (020) – (020) bands. The comparison of our $H_2{}^{18}O$ line positions with the observations shows very good agreement to highly precise data of Matsushima et al. [37]. The *rms* deviation is $6.19\times10^{-5}$ cm$^{-1}$ for 98 transitions observed in both measurements.

For $H_2{}^{17}O$, 362 pure rotation transitions were reported by absorption in Refs. [7,13,35-37]. The list of $H_2{}^{17}O$ rotation transitions was extended up to 685 in Ref. [1], including 114 transitions of the (010) – (010) band. Now, we are reporting 407 transitions (see **Table 2**). A very good agreement is obtained with the highly precise positions of Matsushima et al. [37]. The *rms* deviation is $4.02\times10^{-5}$ cm$^{-1}$ for 103

transitions observed in both measurements. The comparison with our previous measurement [1] gives an *rms* deviation of $2.93\times10^{-4}$ cm$^{-1}$ for 345 transitions with position discrepancies greater than 0.001 cm$^{-1}$ for weak component of unresolved pure rotational doublet $17_{1\ 17} - 16_{0\ 16}$ at 319.6949 cm$^{-1}$ and strongly blended components of the doublet $6_{6\,1} - 5_{5\,0}$ and $6_{6\,0} - 5_{5\,1}$ at 301.142 cm$^{-1}$.

In the case of $HD^{16}O$, 510 absorption and 1422 emission transitions were reported in Refs. [7,9,10,39,40]. In addition, 133 previously unobserved transitions were reported in Ref. [1]. 674 $HD^{16}O$ transitions were measured in the present work, 63 of which being firstly detected in absorption. The best agreement of the line position comparison is for Refs. [1,9,10] with corresponding *rms* values are $1.49\times10^{-4}$, $2.78\times10^{-4}$ and $1.74\times10^{-4}$ cm$^{-1}$ for 488, 204 and 267 transitions respectively. The comparison to the less accurate data of Refs. [7,40] leads to *rms* values of $3.22\times10^{-3}$ and $3.66\times10^{-3}$ cm$^{-1}$ for 60 and 55 transitions, respectively. The maximum deviation reaches a value of 0.0138 cm$^{-1}$ for the absorption study of Ref. [7] and 0.0169 cm$^{-1}$ for the emission study of Ref. [40].

For $HD^{18}O$, the first measurements were due to Johns [9] who reported 148 pure rotation transitions between 78 and 220 cm$^{-1}$. In our recent study [1], we extended the observations to 1126 transitions in the range 44.86 – 645.74 cm$^{-1}$ All 42 transitions reported in the present work were observed in Ref. [1]. Note that some of the present line positions are nevertheless believed to be more precise because the corresponding lines were saturated in Ref. [1]. The line position comparison with data of Johns [9] gives the *rms* deviation of $3.66\times10^{-4}$ cm$^{-1}$ for 26 transitions.

*3.2. Validation tests of the HITRAN2020, W2020 and 14CoMaPi line positions*

Let us now compare our results to the most relevant spectroscopic databases in our region.

Following the approach developed by a task group (TG) of the International Union of Pure and Applied Chemistry (IUPAC-TG) [41-44], improved sets of empirical energy levels have been recently released for $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$. The W2020 empirical energy levels of Ref. [5] were derived from an exhaustive collection and review of measured transitions in all spectral regions [5,45]. The procedure and code xMARVEL [46,47] were applied to the constructed catalog of measured absorption and emission line positions. In particular, all the above reviewed sources of measurements in the rotational region of interest were incorporated in the transition datasets. In the present work, no new energy levels could be determined from the recorded spectra *i.e.* all the energy levels involved in the transitions presently measured have been determined empirically (tag "M" in the W2020 lists [5]). An important advantage of the Ritz principle which is the basis of the MARVEL procedure is that it allows to obtain line positions with experimental accuracy even for transitions which were not previously measured. Direct validation tests when experimental spectra become available are nevertheless suitable (see *e.g.* Refs. [48-50]. The transitions newly observed in the present SOLEIL spectra are particularly valuable in that context. Let us underline that the rotational region under analysis is of importance as a large fraction of the involved

energy levels are lower state of transitions located in the whole frequency range of the water absorption spectrum. Errors or inaccuracies evidenced in the rotational region may impact line positions calculated from empirical energy values in the entire water spectrum.

As mentioned above, according to the HITRAN2020 source list, most of the HITRAN position values should coincide to the W2020 empirical positions of Ref. [5]. In fact, as already discussed in Ref. [50] and illustrated below, this is not the case. The origin of this situation remains unclear and might possibly lead to an update of the HITRAN2020 list. Thus separated comparison has to be performed to the HITRAN and W2020 line lists.

The bending-rotation Hamiltonian developed by Coudert et al. [6] (14CoMiPa, hereafter) provides an alternative source of calculated line positions in our region for the main isotopologue. Due to the failure of the approach based on a rotational Hamiltonian of a semi-rigid molecule in the case of water, a bending rotation Hamiltonian has been developed to account for the $H_2{}^{16}O$ spectrum up to the second triad and to $J$ = 30 [6]. The effective Hamiltonian parameters were determined from a weighted least-squares fitting of a large set of 24,461 literature data including rotational energy levels, microwave, far infrared, and infrared measured transition frequencies.

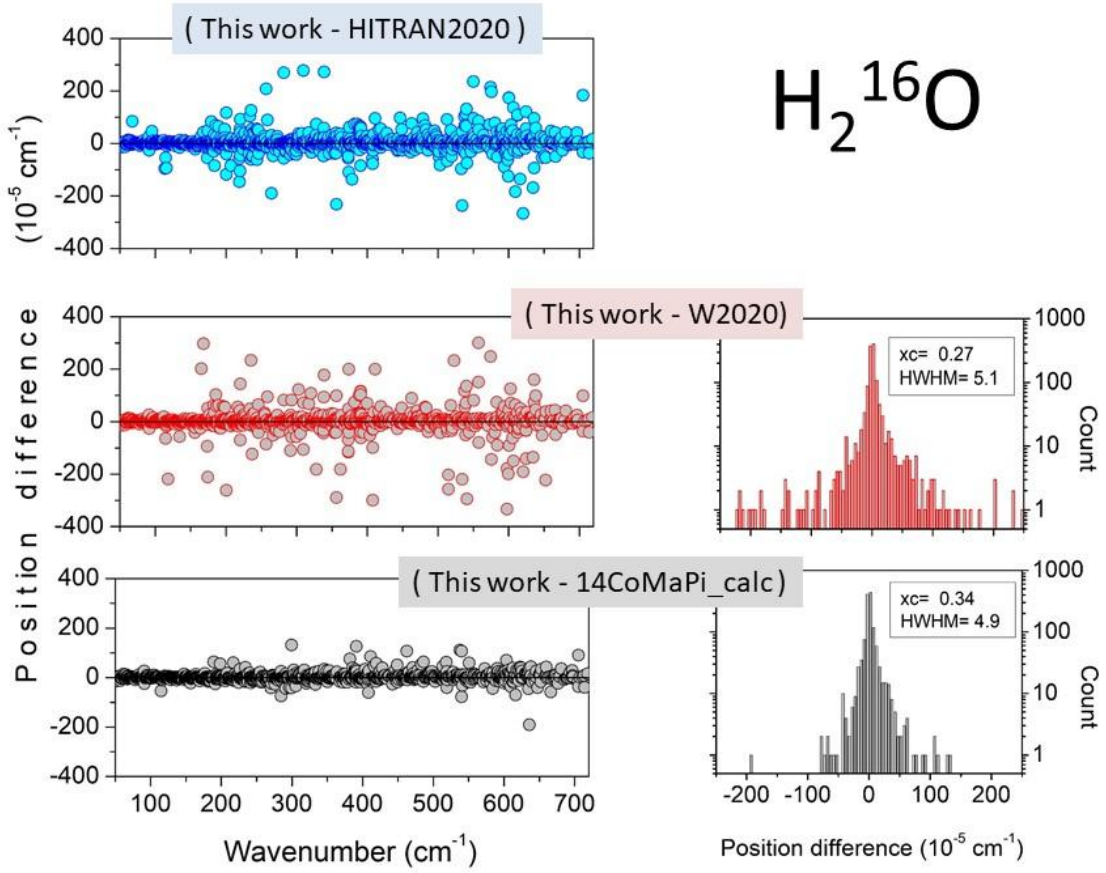


**Fig. 5**
Position comparison of the $H_2{}^{16}O$ lines in the 50-720 $cm^{-1}$ range.

*Left panels:* Differences between the experimental values presently determined from SOLEIL spectra to the HITRAN2020 data, to the W2020 empirical values and the values calculated using a rotation-bending Hamiltonian [6] (14CoMaPi),
*Right panels:* Histograms of the deviations between the experiment and the W2020 and 14CoMaPi positions. The center (xc) and HWHM obtained from a Gaussian fit of the histograms are given on the plot (in $10^{-5}$ $cm^{-1}$ units).

We present in **Fig. 5** an overview comparison of our line positions of the main isotopologue to the values included in the HITRAN database, to the W2020 values and to the calculated values of Coudert et al. [6]. The corresponding comparison table is provided as a third Supplementary Material. **Fig. 5** illustrates that the overall agreement between the experiment and the three databases is excellent. For instance, the histograms of the W2020 and 14CoMaPi deviations from experiment show a similar Gaussian distribution centered at about of $3\times10^{-6}$ $cm^{-1}$ for 1310 transitions with a HWHM of $5\times10^{-5}$ $cm^{-1}$, corresponding to our claimed position uncertainty of the "good" lines. The most noticeable differences between the three comparisons is the presence of some relatively large deviations for the W2020 and HITRAN2020 data while these "outliers" are mostly absent for 14CoMiPa. At this point, it is worth mentioning that the comparison applies to the same set of 1310 transitions as all our $H_2{}^{16}O$ observations have a counterpart in the three databases (except for a couple of doublets missing in the HITRAN database – see below).

The direct comparison to the spectra presented in **Figs. 6** and **7** gives insights on the significance of these outliers. **Fig. 6** shows six examples for which clear deviations are observed for the W2020 line positions of Ref. [5] while HITRAN position (given with W2020 source!!) agree with the measurements. The W2020 position uncertainties (displayed on the figure) are significantly smaller than the observed deviations. For instance, for the lines displayed at 338.7656 and 557.5464 $cm^{-1}$, the (Meas. - W2020) position differences are $1.77\times10^{-3}$ and $1.51\times10^{-3}$ $cm^{-1}$, respectively, 74 and 101 times larger than the claimed W2020 uncertainties ($2.4\times10^{-5}$ and $1.5\times10^{-5}$ $cm^{-1}$, respectively [5]).

In **Fig. 7**, we present six examples of $H_2{}^{16}O$ transitions where neither the W2020 position nor the HITRAN position shows a satisfactory agreement with experiment. All these lines are supposed to have a W2020 source in the HITRAN2020 database.

The $H_2{}^{16}O$ calculated stick spectrum of 14CoMaPi [6] is included in the different panels of **Figs. 6** and **7**. In all the displayed examples, the 14CoMaPi line positions agree with experiment. This observation nicely illustrates the fact that the physics included in the bending-rotation Hamiltonian of Ref. [33] brings constraints which improves the accuracy of the calculations beyond the experimental accuracy of the input data. Although the accuracy of the previous frequency measurements of the transitions observed in the SOLEIL spectra was not optimum, the fit of the effective Hamiltonian parameters benefitted from a large set of spectroscopic data extending beyond our spectral range. It is important to note that all the experimental sources used by Coudert et al. [6] in 2014 are publicly available and were used for the

derivation of the W2020 energy levels. The absence of "outliers" in the 14CoMaPi list suggests that for our set of observations in the rotational range, the effective Hamiltonian approach is more efficient to avoid "outliers" than the xMARVEL approach [46,47] implemented to derive the W2020 energy levels [5,45].

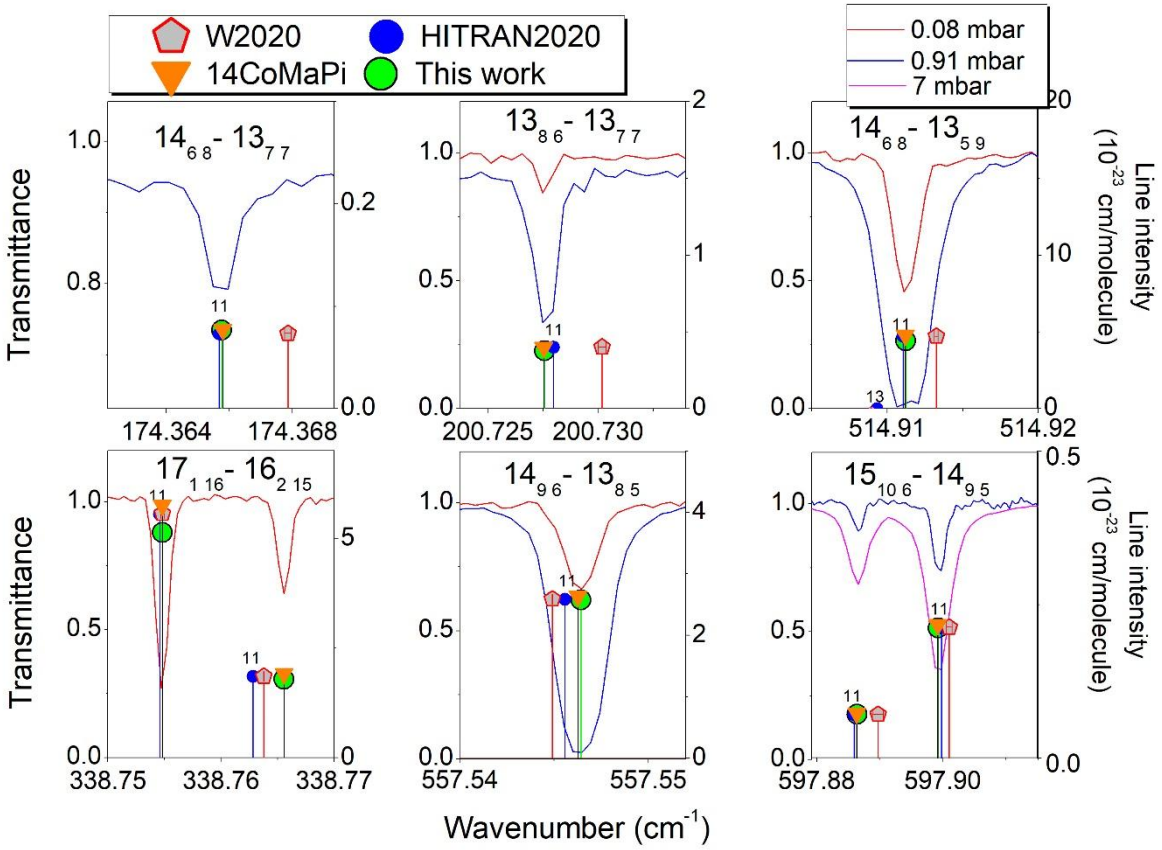


**Fig. 6**
Comparison of FTS spectra of natural water vapor recorded at SOLEIL and corresponding global line list constructed in this work (green circles) to the W2020 list of $H_2{}^{16}O$ [5] (red pentagons). The HITRAN2020 stick spectrum of natural water (blue circles) is superimposed together with the effective Hamiltonian calculated positions of Ref. [6] (orange triangles - 14CoMaPi). All the displayed examples correspond to inaccuracies of the HITRAN2020 positions of the $H_2{}^{16}O$ main isotopologue (isotopologue code "11"). All the problematic lines correspond pure rotational transitions. Their rotational assignments are given on each panel.

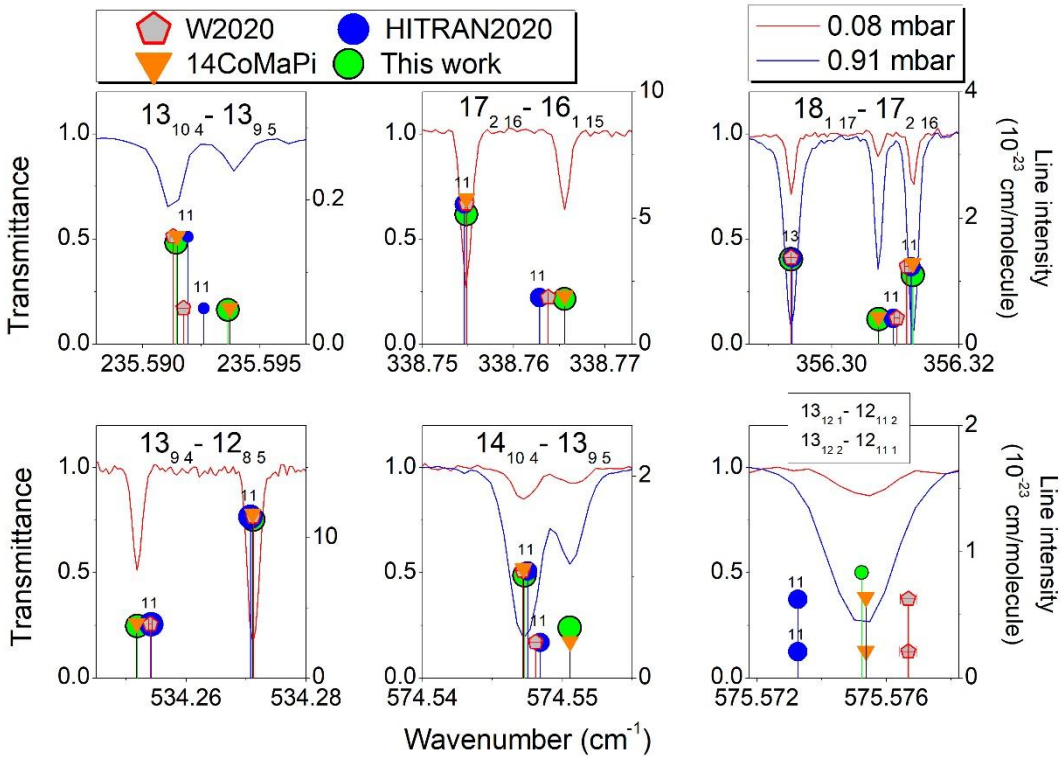


**Fig. 7**
Examples of inaccurate line positions in both the W2020 and HITRAN2020 lists.
Comparison of FTS spectra of natural water vapor recorded at SOLEIL and corresponding global line list constructed in this work (green circles) to the W2020 list of $H_2{}^{16}O$ [5] (grey pentagons), to the HITRAN2020 stick spectrum of natural water (blue circles) and to the $H_2{}^{16}O$ calculated positions of Ref. [33] (orange triangles - 14CoMaPi). All the displayed examples correspond to inaccuracies of both the W2020 and HITRAN2020 positions of the $H_2{}^{16}O$ main isotopologue (isotopologue code "11"). All the problematic lines correspond pure rotational transitions. Their rotational assignments are given on each panel.

In principle, the W2020 empirical energy levels and transition frequencies also benefitted from experimental sources in other spectral regions which should improve their accuracies but the constraints brought by the large networks connecting the energy levels are different compared to those brought by the effective Hamiltonian of Ref. [6]. As a result, some individual energy levels can be strongly impacted by an experimental source reported an inaccurate measurement and the resulting W2020 transitions frequency is inaccurate. In three previous contributions by cavity ring-down spectroscopy (CRDS), we presented validation tests of the W2020 positions. In the 8040-8630 $cm^{-1}$ region and in the region of the A-band of $O_2$ near 760 nm [48-50], a significant number of substantial deviations of the W2020 positions from the measurements were evidenced. Part of the W2020 positions were found less accurate than some original sources included in the W2020 transition database and W2020 uncertainties were found unreliable. In particular, at high energy, a significant fraction of the W2020 position were found to deviate from observations by amounts exceeding their claimed uncertainty by factors >10-100 [49,50]. They are due to the fact that the authors of Ref. [5] used not only high-precision data to determine the energies and

some less accurate data are 'spoiling' higher quality data sources. For instance, a large number of transitions from emission spectra [30-33,51-58], including very old data [59-61] were incorporated in the xMARVEL procedure [5,45]. As a result, the few high-precision data did not have a decisive influence on the determination of the corresponding W2020 energies. For example, the very accurate line positions [1] of the pure rotation transitions $17_{2\,16} - 16_{1\,15}$ (338.76545 $cm^{-1}$, [1]) and $18_{1\,17} - 17_{2\,16}$ (356.30704 $cm^{-1}$ [1]) were included in the input data file of Refs. [5,46] but the resulting W2020 values reported with error bars of $2.4\times10^{-5}$ and $4.6\times10^{-4}$ $cm^{-1}$ deviate by $-1.67\times10^{-3}$ $cm^{-1}$ and $+3.21\times10^{-3}$ $cm^{-1}$, respectively. The position values of Ref. [1] are confirmed by the present measurements (differences of $1.0\times10^{-4}$ $cm^{-1}$ and $3.1\times10^{-4}$ $cm^{-1}$, respectively).

In a fourth Supplementary Material, we provide a table listing the lines of the global line list whose measured position differs from the W2020 value by more than $9\times10^{-4}$ $cm^{-1}$. A sample of the table is presented in **Fig. 8**. The ratio $R = \frac{|\nu_{exp} - \nu_{W2020}|}{Unc_{W2020}}$ which compares the absolute deviation of the W2020 position from the measured value to the claimed W2020 position uncertainty has been added. Over the 60 listed transitions, all but two have an $R$ value larger than unity and 21 correspond to $R$ >10 (with maximum value $R$ = 459 for pure rotational transition $15_{5\,11} - 14_{4\,10}$ at 388.70559 $cm^{-1}$), illustrating the limited reliability of the W2020 uncertainties for the considered set of lines. The (exp-14CoMaPi_calc) position differences are included in the table. For 45 of the 60 transitions listed, the deviation from the 14CoMaPi calculated value is less than $2\times10^{-4}$ $cm^{-1}$. Only three (exp-14CoMaPi_calc) deviations exceeding $9\times10^{-4}$ $cm^{-1}$ are noted. They correspond to very weak lines (line intensity smaller than $1.6\times10^{-25}$ cm/molecule) and have probably an experimental origin.

As concerns the HITRAN database, except for two $J'$= 21 doublets, the source given for the $H_2{}^{16}O$ line positions is the W2020 line list [5]. Even if the position values are not exactly the original values published by Furtenbacher et al. [5] (see for instance the upper panels of **Fig. 6**), the presence of outliers in the HITRAN list has the W2020 origin discussed above. For the ($21_{0\,21} - 20_{1\,20}$ and $21_{1\,21} - 20_{0\,20}$) and ($21_{1\,20} - 20_{2\,19}$ and $21_{2\,20} - 20_{1\,19}$]) doublets, the HITRAN database reproduces the values of Lanquetin et al. [62] (thus obtained following the same approach as adopted in 14CoMaPi). The comparison to experiment indicates that this choice was judicious as for instance, the measured position of the first doublet is closer to the value Lanquetin et al. [62] than to the W2020 value ($6.4\times10^{-4}$ $cm^{-1}$ and $1.1\times10^{-3}$ $cm^{-1}$ position differences, respectively).

The pure rotation $H_2{}^{16}O$ doublet, $22_{1\,22} - 21_{0\,21}$ and $22_{0\,22} - 21_{1\,21}$, assigned to the line measured at 407.74639 $cm^{-1}$ (intensity of $1.313\times10^{-25}$ cm/molecule) is the only observed line absent in the HITRAN2020 list. The W2020 positions and intensities [5] agree satisfactorily with the measurements (407.74939 $cm^{-1}$ and $1.04\times10^{-25}$ cm/molecule, respectively).

| | | | | | HITRAN2020 | | | | | | | | | | W2020 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| spc | Nu_exp | dN | S_exp | dS | Nu | S | assignment | | | | | | | | d1 | Ratio_1 | Nu | t | dN1 | S | d2 | Ratio_2 | R | d3 |
| 3 | 118.04454 | 5 | 1.869E-24 | 6 | 118.044535 | 1.997E-24 | 000 | 14 | 6 | 8 | 000 | 14 | 5 | 9 | 1 | 0.936 | 118.04673 | M | 4 | 2.002E-24 | -219 | 0.934 | 54.8 | -9 |
| 3 | 165.07328 | 5 | 1.435E-24 | 3 | 165.073002 | 1.598E-24 | 000 | 14 | 7 | 7 | 000 | 14 | 6 | 8 | 28 | 0.898 | 165.07126 | M | 35 | 1.601E-24 | 202 | 0.896 | 5.8 | 1 |
| 3 | 168.09872 | 11 | 2.589E-25 | 11 | 168.098288 | 2.399E-25 | 000 | 15 | 5 | 11 | 000 | 14 | 6 | 8 | 43 | 1.079 | 168.09575 | M | 16 | 2.405E-25 | 297 | 1.077 | 18.6 | -8 |
| 3 | 173.65217 | 8 | 8.392E-25 | 8 | 173.652719 | 8.024E-25 | 000 | 15 | 6 | 10 | 000 | 15 | 5 | 11 | -55 | 1.046 | 173.65311 | M | 20 | 8.044E-25 | -94 | 1.043 | 4.7 | 1 |
| 3 | 174.36575 | 5 | 7.695E-25 | 4 | 174.365673 | 7.317E-25 | 000 | 14 | 6 | 8 | 000 | 13 | 7 | 7 | 8 | 1.052 | 174.36787 | M | 11 | 7.338E-25 | -212 | 1.049 | 19.3 | -3 |
| 3 | 185.67453 | 5 | 1.807E-24 | 3 | 185.673836 | 1.911E-24 | 000 | 15 | 5 | 11 | 000 | 15 | 4 | 12 | 69 | 0.946 | 185.67350 | M | 14 | 1.916E-24 | 103 | 0.943 | 7.4 | -2 |
| 3 | 200.72759 | 5 | 3.778E-24 | 5 | 200.727999 | 4.003E-24 | 000 | 13 | 8 | 6 | 000 | 13 | 7 | 7 | -41 | 0.944 | 200.73021 | M | 17 | 4.013E-24 | -262 | 0.941 | 15.4 | 3 |
| 3 | 220.52085 | 5 | 1.150E-24 | 3 | 220.519932 | 1.041E-24 | 000 | 14 | 9 | 6 | 000 | 14 | 8 | 7 | 92 | 1.105 | 220.51942 | M | 6 | 1.043E-24 | 143 | 1.103 | 23.8 | 4 |
| 3 | 235.59390 | 5 | 4.761E-25 | 7 | 235.592641 | 4.969E-25 | 000 | 13 | 10 | 4 | 000 | 13 | 9 | 5 | 126 | 0.958 | 235.59157 | M | 33 | 4.983E-25 | 233 | 0.955 | 7.1 | -9 |
| 3 | 236.92234 | 7 | 3.869E-25 | 6 | 236.922838 | 3.507E-25 | 000 | 14 | 10 | 5 | 000 | 14 | 9 | 6 | -50 | 1.103 | 236.92358 | M | 35 | 3.516E-25 | -124 | 1.100 | 3.5 | 0 |
| 3 | 258.66921 | 5 | 3.136E-24 | 3 | 258.668316 | 3.176E-24 | 000 | 16 | 2 | 14 | 000 | 16 | 1 | 15 | 89 | 0.987 | 258.66827 | M | 14 | 3.184E-24 | 94 | 0.985 | 6.7 | 5 |
| 2 | 291.44779 | 5 | 2.479E-23 | 3 | 291.448007 | 2.408E-23 | 000 | 16 | 2 | 15 | 000 | 16 | 1 | 16 | -22 | 1.029 | 291.44889 | M | 8 | 2.415E-23 | -110 | 1.027 | 13.8 | 10 |
| 1 | 302.64445 | 5 | 1.114E-21 | 3 | 302.644081 | 9.033E-22 | 000 | 16 | 1 | 16 | 000 | 15 | 0 | 15 | 37 | 0.925 | 302.64323 | M | 3 | 9.059E-22 | 122 | 0.922 | 40.7 | -16 |
| 3 | 309.78911 | 5 | 5.228E-24 | 3 | 309.789058 | 5.544E-24 | 000 | 17 | 1 | 16 | 000 | 17 | 0 | 17 | 5 | 0.943 | 309.79017 | M | 8 | 5.560E-24 | -106 | 0.940 | 13.3 | 14 |

**Fig. 8**

Beginning of a Table provided as a Supplementary Material listing the absorption lines extracted from the global line list of water vapor between 50-720 $cm^{-1}$ whose measured position differs from the W2020 position by more than 0.0009 $cm^{-1}$. The (exp-HITRAN2020), (exp-W2020) and (exp-14CoMaPi_calc) position differences are given in $10^{-5}$ $cm^{-1}$ units in columns, d1, d2 and d3, respectively. The ratio R= |d2|/dN1compares the absolute deviation of the W2020 position to the W2020 position uncertainty (dN1).

Let us now consider the minor isotopologues

The W2020 transition frequencies have been published only for the non-deuterated species ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$) [5]. In the case of the deuterated species – $HD^{16}O$, $HD^{18}O$, and $HD^{17}O$ –, the HITRAN database reproduces the line positions calculated by Kyuberis et al. [63] from updated IUPAC-TG energy levels [42]. The general agreement is satisfactory for the line positions of deuterated species. The *rms* deviations of our line positions from those of HITRAN [4] are $3.74\times10^{-4}$, $6.62\times10^{-4}$ and $4.54\times10^{-4}$ cm$^{-1}$ for 674, 42 and 4 transitions of $HD^{16}O$, $HD^{18}O$ and $HD^{17}O$ respectively. Nevertheless, the line position differences exceed 0.001 cm$^{-1}$ for twelve $HD^{16}O$ and one $HD^{18}O$ transitions. The maximum deviations are for high $J,K_a$ transitions of $HD^{16}O$ (up to 0.0043 cm$^{-1}$) and for the $2_{2\,0} - 1_{1\,1}$ $HD^{18}O$ transition (0.0042 cm$^{-1}$). Note, the *rms* deviation for 41 transitions of $HD^{18}O$ (excluding the $2_{2\,0} - 1_{1\,1}$ transition) is $1.26\times10^{-4}$ cm$^{-1}$.

For the $H_2{}^{18}O$ species, again we noted some differences between HITRAN2020 and W2020 line positions while the HITRAN source of all the considered $H_2{}^{18}O$ positions is supposed to be the W2020 line list [5]. For example, the HITRAN position of the $13_{11\,3} - 12_{10\,2}$ transition is 558.84113 cm$^{-1}$ differs by more than $10^{-3}$ from its W2020 value (558.84022 cm$^{-1}$ [5]). Our measured value of 558.84025 cm$^{-1}$ coincides to the W2020 value. The *rms* deviation of the present SOLEIL line positions from those of HITRAN2020 is $1.57\times10^{-4}$ cm$^{-1}$ for all 565 assigned $H_2{}^{18}O$ transitions

In the case of the $H_2{}^{17}O$ isotopologue, for an unknown reason, the W2020 source was rejected for 83 transitions: five and 78 were taken from Lodi and Tennyson [28] and updated values determined by a task group (TG) of the International Union of Pure and Applied Chemistry (IUPAC-TG) [41], respectively. Five examples of deviations between the HITRAN $H_2{}^{17}O$ positions and the SOLEIL spectra are presented in **Fig. 9** (one additional panel concerns $HD^{16}O$). Interestingly, the original W2020 line positions of Ref. [5] of all the problematic $H_2{}^{17}O$ lines agree very well with the experimental spectrum. Note that the HITRAN positions of the lines near 319.71 and 320.12 cm$^{-1}$ are due to Lodi and Tennyson [28] while the W2020 source is given for the three other $H_2{}^{17}O$ lines. For the 78 positions with IUPAC-TG origin, the average position difference compared to our measurements is $1.09\times10^{-4}$ cm$^{-1}$ with a large *rms* deviation of $2.87\times10^{-3}$ cm$^{-1}$. These values are reduced to $5.06\times10^{-5}$ cm$^{-1}$ and $1.79\times10^{-4}$ cm$^{-1}$ for the corresponding W2020 values, respectively, indicating that overall, the corresponding W2020 position values of Ref. [5] should have been preferred.

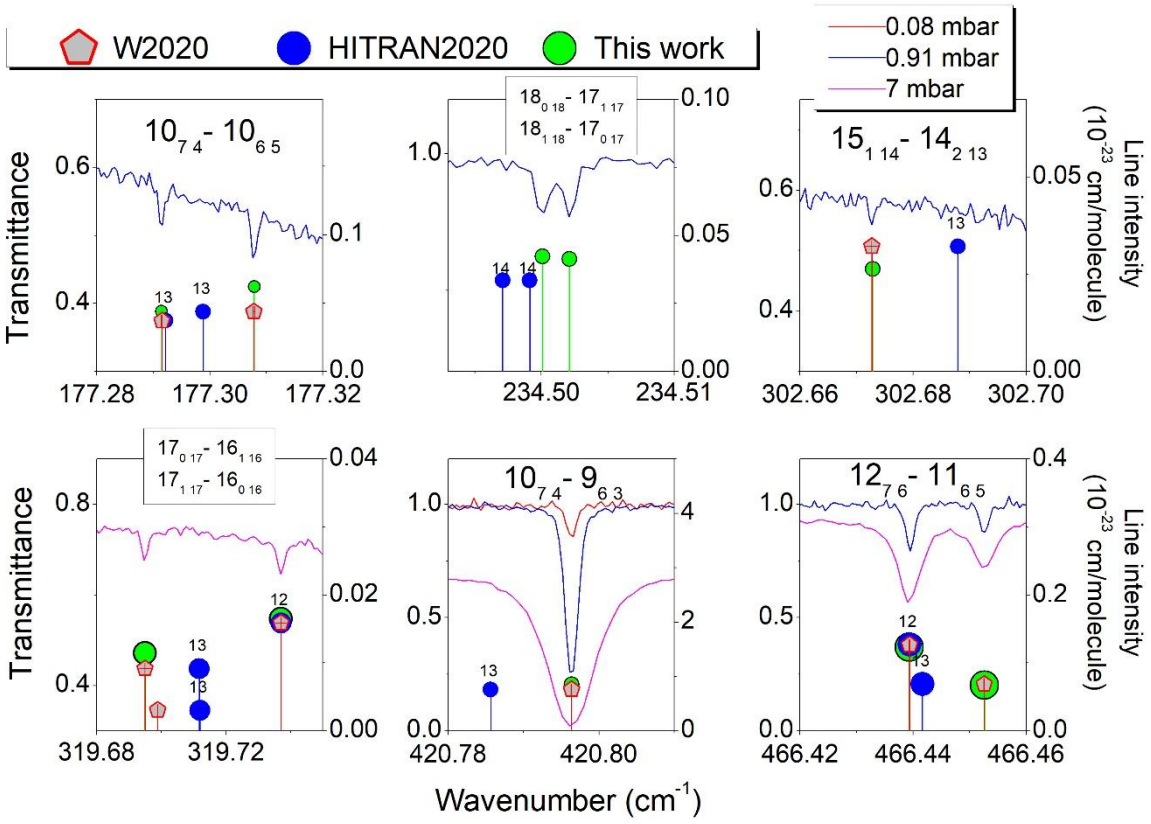


**Fig. 9**

Examples of inaccurate line positions in the HITRAN database.

Comparison of FTS spectra of natural water vapor recorded at SOLEIL and corresponding global line list constructed in this work (green circles) to the HITRAN2020 list of natural water [4] (blue circles). All but the $HD^{16}O$ doublet (isotopologue code "14") near 234.50 $cm^{-1}$ are relative to the $H_2{}^{17}O$ minor isotopologue (isotopologue code "13"). The W2020 line list of $H_2{}^{17}O$ [5] is superimposed (grey pentagons) with corresponding error bars. In all the displayed examples, the W2020 line positions are in good agreement with experiment. All the problematic lines correspond pure rotational transitions. Their rotational assignments are given on each panel.

## 4. Concluding remarks

A series of five high quality FTS spectra of natural water at room temperature has been recorded in the far infrared region using the SOLEIL synchrotron radiation source combined with a cell providing a 151.75-m absorption pathlength. The spectral coverage (50-720 $cm^{-1}$), the high position accuracy (on the order of and $5\times10^{-5}$ $cm^{-1}$ for isolated lines of intermediate intensity), and the high sensitivity of the recorded spectra make them valuable for stringent validation tests of spectroscopic databases. An experimental list of more than 3000 transitions of six water isotopologues ($H_2{}^{16}O$, $H_2{}^{17}O$, $H_2{}^{18}O$, $HD^{18}O$, $HD^{16}O$, $HD^{17}O$) were constructed from the five spectra recorded at different pressures up to 7 mbar. The measured line intensities span more than five orders of magnitude. More than 450 lines are newly detected by absorption spectroscopy. The HITRAN2020 [4] and W2020 [5] line lists and the effective Hamiltonian predictions of the $H_2{}^{16}O$ spectrum by Coudert et al. [6] (14CoMaPi) were compared to the experimental results.

The comparison shows an overall very satisfactory agreement but, even scarce, inaccuracies are of importance as the observed transitions involve rotational levels of the lowest vibrational states – (000), (010) and (020) – which are the lower states of most of the water vapor transitions. Energy levels corrections will thus propagate to calculated line positions in all the energy ranges of the water spectrum. The main results of the comparison can be summarized as follows,

*(i)* The 14CoMaPi calculated positions of the main isotopologue agree with the observations within the experimental accuracy,

*(ii)* Although the agreement with the W2020 and HITRAN2020 positions is in general very good, a number of transition deviates from the observations significantly. These situations are less numerous in the presently studied rotational range but confirm the conclusion of the validation tests of the W2020 positions in the near infrared region [49] and in the region of the oxygen A-band near 760 nm [48,50]. The evidence of such general problem underlines the necessity of a careful validation tests before implementing the W2020 empirical line positions in reference spectroscopic databases used for atmospheric applications,

*(iv)* the uncertainty values attached to the W2020 empirical positions and energy levels can be strongly underestimated. A few examples show deviations exceeding the W2020 uncertainty by factors larger than 10 (up to 459) (see **Fig. 8**),

*(v)* according to the HITRAN2020 database, the W2020 empirical positions of Ref. [5] were adopted as the main source of line positions in the region. In fact, the HITRAN2020 and the W2020 original line positions do not coincide. As a result, we found series of examples where the position agreement of the recorded spectra was better for HITRAN2020 list than for W2020 list and *vice versa* (see **Figs. 6 and 7**),

*(vi)* overall, the W2020 line positions of $H_2{}^{17}O$ show a better agreement with measurements than some HITRAN values and should have been systematically preferred.

Although relatively limited (no new energy levels were determined), the present study has allowed to illustrate the advantage of the state-of-the-art effective Hamiltonian approach for the first variational levels of the main isotopologue. Note that only 1310 of the 3001 reported line positions belong to the main isotopologue. Transitions of the minor isotopologues could be used for testing the rotation-bending Hamiltonian developed some years ago for the $H_2{}^{18}O$ and $HD^{16}O$ species [64,65]. During the measurement campaign of the natural water spectra at SOLEIL, FTS spectra of water samples enriched in $^{17}O$ or D (and both) were recorded. The analysis of these spectra will make available a large amount of new information for validation tests of transition frequencies and energy levels of the minor isotopologues (in particular for doubly-substituted species). The obtained results will be reported in separated contributions.

**Acknowledgements**

This work became possible due to the Project No20210051 supported by SOLEIL Synchrotron Team. SNM activity was supported in the frame of the Russian Science Foundation, grant no. 18-11-00024-П. The support of the CNRS (France) in the frame of International Research Project SAMIA is acknowledged.

## References


1. Mikhailenko SN, Béguier S, Odintsova TA, Tretyakov MYu, Pirali O, Campargue A. The far-infrared spectrum of $^{18}O$ enriched water vapour (40 – 700 $cm^{-1}$). J Quant Spectrosc Radiat Transf 2020;253:107105. https://doi.org/10.1016/j.jqsrt.2020.107105
2. Odintsova TA, Tretyakov Myu, Zibarova AO, Pirali O, Roy P, Campargue A. Far-infrared self-continuum absorption of $H_2{}^{16}O$ and $H_2{}^{18}O$ (15-500 $cm^{-1}$). J Quant Spectrosc Radiat Transf 2019;227:190-200. https://doi.org/10.1016/j.jqsrt.2019.02.012
3. Odintsova TA, Tretyakov MYu, Simonova A, Ptashnik I, Pirali O, Campargue A. Measurement and temperature dependence of the water vapor self-continuum in the 70–700 $cm^{-1}$ range. J Mol Structure 2020;1210:128046. https://doi.org/10.1016/j.molstruc.2020.128046
4. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, et al. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf 2022;277:107949. https://doi.org/10.1016/j.jqsrt.2021.107949
5. Furtenbacher T, Tobias R, Tennyson J, Polyansky OL, Kyuberis AA, Ovsyannikov RI, et al. W2020: A database of validated rovibrational experimental transitions and empirical energy levels of water isotopologues. II. $H_2{}^{17}O$ and $H_2{}^{18}O$ with an update to $H_2{}^{16}O$. J Phys Chem Ref Data 2020;49:043103. https://doi.org/10.1063/5.0030680
6. Coudert LH, Martin-Drumel M-A, Pirali O. Analysis of the high-resolution water spectrum up to the second triad. J Mol Spectrosc 2014;303:36-41. https://doi.org/10.1016/j.jms.2014.07.003
7. Kauppinen J, Kärkkäinen T, Kyrö E. High-resolution spectrum of water vapor between 30 and 720 $cm^{-1}$. J Mol Spectrosc1978;71:15-45. https://doi.org/10.1016/0022-2852(78)90073-5
8. Kauppinen J, Jolma K, Hornaman VM. New wave-number calibration tables for $H_2O$, $CO_2$, and OCS lines between 500 and 900 $cm^{-1}$. Appl Opt 1982;21:3332-6. https://doi.org/10.1364/AO.21.003332
9. Johns JWC. High-resolution far infrared (20-350 $cm^{-1}$) spectra of several species of $H_2O$. J Opt Soc Am B 1985;2:1340-54. https://doi.org/10.1364/JOSAB.2.001340
10. Paso R, Horneman VM. High-resolution rotational absorption spectra of $H_2{}^{16}O$, $HD^{16}O$, and $D_2{}^{16}O$ between 110 and 500 $cm^{-1}$. J Opt Soc Am B 1995;12:1813-38. https://doi.org/10.1364/JOSAB.12.001813
11. Matsushima F, Odashima H, Iwasaki T, Tsunekawa S, Takagi K. Frequency measurement of pure rotational transition of $H_2O$ from 0.5 to 5 THz. J Mol Structure 1995;352:371-8. https://doi.org/10.1016/0022-2860(94)08531-L
12. De Natale P, Lorini L, Inguscio M, Nolt IG, Park JH, Di Lonardo G, et al. Accurate frequency measurements for $H_2O$ and $^{16}O_3$ in the 119-$cm^{-1}$ OH atmospheric window. Appl Opt 1997;36:8526-32. https://doi.org/10.1364/AO.36.008526
13. Toth RA. Water vapor measurements between 590 and 2582 $cm^{-1}$: line positions and strengths. J Mol Spectrosc 1998;190:379-96. https://doi.org/10.1006/jmsp.1998.7611
14. Chen P, Pearson JC, Pickett HM, Matsuura S, Blake GA. Submillimeter-wave measurements and analysis of the ground and $\nu_2$=1 states of water. Astrophys J Suppl Series 2000;128:371-85. https://doi.org/10.1086/313377
15. Horneman VM, Anttila R, Alanko S, Pietilä J. Transferring calibration from $CO_2$ laser lines to far infrared water lines with the aid of the $\nu_2$ band of OCS and the $\nu_2$, $\nu_1$-$\nu_2$, and $\nu_1$+$\nu_2$ bands of $^{13}CS_2$: Molecular constants of $^{13}CS_2$. J Mol Spectrosc 2005;234:238-54. https://doi.org/10.1016/j.jms.2005.09.011
16. Matsushima F, Tomatsu N, Nagai T, Moriwaki Y, Takagi K. Frequency measurement of pure rotational transitions in the $\nu_2$=1 state of $H_2O$. J Mol Spectrosc 2006;235:190-5. https://doi.org/10.1016/j.jms.2005.11.003
17. Cazzolli G, Puzzarini C, Buffa G, Tarrini O. Pressure-broadening of water lines in the THz frequency region: Improvements and confirmations for spectroscopic databases. Part II. J Quant Spectrosc Radiat Transf 2009;110:609-18. https://doi.org/10.1016/j.jqsrt.2008.12.001

18. Drouin BJ, Yu SS, Pearson JC, Gupta H. Terahertz spectroscopy for space applications: 2.5 – 2.7 THz spectra of HD, $H_2O$ and $NH_3$. J Mol Structure 2011;1006:2-12. https://doi.org/10.1016/j.molstruc.2011.05.062
19. Yu SS, Pearson JC, Drouin BJ, Martin-Drumel M-A, Pirali O, Vervloet M, et al. Measurement and analysis of new terahertz and far-infrared spectra of high temperature water. J Mol Spectrosc 2012;279:16-25. https://doi.org/10.1016/j.jms.2012.07.011
20. Yu SS, Pearson JC, Drouin BJ. Terahertz spectroscopy of water in its second triad. J Mol Spectrosc 2013;288:7-10. https://doi.org/10.1016/j.jms.2013.03.011
21. Markov VN, Temperature dependence of self-induced pressure broadening and shift of the $6_{43} - 5_{50}$ line of the water molecule. J Mol Spectrosc 1994;164:233-8. https://doi.org/10.1006/jmsp.1994.1069
22. Markov VN, Krupnov AF. Measurements of the pressure shift of the $1_{10} - 1_{01}$ water line at 556 GHz produced by mixtures gases. J Mol Spectrosc 1995;172:211-4. https://doi.org/10.1006/jmsp.1995.1168
23. Toth RA, Brown LR, Plymate C. Self-broadened widths and frequency shifts of water vapor lines between 590 and 2400 $cm^{-1}$. J Quant Spectrosc Radiat Transf 1998;59:529-62. https://doi.org/10.1006/jmsp.1995.1168
24. Zou Q, Varanasi P. Laboratory measurement of the spectroscopic line parameters of water vapor in the 610-2100 and 3000-4050 $cm^{-1}$ regions at lower-tropospheric temperatures. J Quant Spectrosc Radiat Transf 2003;82:45-98. https://doi.org/10.1006/jmsp.1995.1168
25. Podobedov VB, Plusquellic DF, Fraser GT. THz laser study of self-pressure and temperature broadening and shifts of water vapor lines for pressures up to 1.4 kPa. J Quant Spectrosc Radiat Transf 2004;87:377-85. https://doi.org/10.1016/j.jqsrt.2004.03.001
26. Golubiatnikov GYu. Shifting and broadening parameters of the water vapour 183-GHz line ($3_{13} - 2_{20}$) by $H_2O$, $O_2$, $N_2$, $CO_2$, $H_2$, He, Ne, Ar, and Kr at room temperature. J Mol Spectrosc 2005;230:196-8. https://doi.org/10.1016/j.jms.2004.10.011
27. Koshelev MA, Tretyakov MYu, Golubiatnikov GYu, Parshin VV, Markov VN, Koval IA. Broadening and shifting of the 321-, 325- and 380-GHz lines of water vapor by pressure of atmospheric gases. J Mol Spectrosc 2007;241:101-8. https://doi.org/10.1016/j.jms.2006.11.005
28. Lodi L, Tennyson J. Line lists for $H_2{}^{18}O$ and $H_2{}^{17}O$ based on empirical line positions and *ab initio* intensities. J Quant Spectrosc Radiat Transf 2012;113:850-8. https://doi.org/10.1016/j.jqsrt.2012.02.023
29. Coudert LH, Pirali O, Vervloet M, Lanquetin R, Camy-Peyret C. The eight first vibrational states of the water molecule: Measurements and analysis. J Mol Spectrosc 2004;228:471-98. https://doi.org/10.1016/j.jqsrt.2012.02.023
30. Polyansky OL, Busler JR, Guo B, Zhang K, Bernath PF. The emission spectrum of hot water in the region between 370 and 930 $cm^{-1}$. J Mol Spectrosc 1996;176:305-15. https://doi.org/10.1006/jmsp.1996.0091
31. Polyansky OL, Tennyson J, Bernath PF. The spectrum of hot water: rotational transitions and difference bands in the (020), (100), and (001) vibrational states. J Mol Spectrosc 1997;186:213-21. https://doi.org/10.1006/jmsp.1997.7443
32. Polyansky OL, Zobov NF, Viti S, Tennyson J, Bernath PF, Wallace L. High-temperature rotational transitions of water in sunspot and laboratory spectra. J Mol Spectrosc 1997;186:422-47. https://doi.org/10.1006/jmsp.1997.7449
33. Coheur P-F, Bernath PF, Carleer M, Colin R, Polyansky OL, Zobov NF, et al. A 3000 K laboratory emission spectrum of water. J Chem Phys 2005;122:074307. https://doi.org/10.1063/1.1847571
34. Benedict WS, Pollack MA, Tomlinson III WJ. The water-vapor laser. IEEE J Quant Electronics 1969;QE-5:108-24. Doi:10.1109/JQE.1969.1075731
35. Winther F. The rotational spectrum of water between 650 and 50 $cm^{-1}$ $H_2{}^{18}O$ and $H_2{}^{17}O$ in natural abundance. J Mol Spectrosc 1977;65:405-19. https://doi.org/10.1016/0022-2852(77)90280-6
36. Kauppinen J, Kyrö E. High resolution pure rotational spectrum of water vapor enriched by $H_2{}^{17}O$ and $H_2{}^{18}O$. J Mol Spectrosc 1980;84:405-23. https://doi.org/10.1016/0022-2852(80)90032-6

37. Matsushima F, Nagase H, Nakauchi T, Odashima H, Takagi K. Frequency measurement of pure rotational transitions of $H_2{}^{17}O$ and $H_2{}^{18}O$ from 0.5 to 5 THz. J Mol Spectrosc 1999;193:217-23. https://doi.org/10.1006/jmsp.1998.7736
38. Mikhailenko SN, Tyuterev VlG, Mellau G. (000) and (010) states of $H_2{}^{18}O$: analysis of rotational transitions in hot emission spectrum in the 400 – 850 $cm^{-1}$ region. J Mol Spectrosc 2003;217:195-211. https://doi.org/10.1016/S0022-2852(02)00018-8
39. Toth RA. HDO and $D_2O$ low pressure, long path spectra in the 600 – 3100 $cm^{-1}$ region. I. HDO line positions and strengths. J Mol Spectrosc 1999;195:73-97. https://doi.org/10.1006/jmsp.1999.7814
40. Janca A, Tereszchuk K, Bernath PF, Zobov NF, Shirin SV, Polyansky OL, Tennyson J. Emission spectrum of hot HDO below 4000 $cm^{-1}$. J Mol Spectrosc 2003;219:132-5. https://doi.org/10.1016/S0022-2852(03)00015-8
41. Tennyson J, Bernath PF, Brown LR, Campargue A, Carleer MR, Császár AG, et al. IUPAC critical evaluation of the rotational-vibrational spectra of water vapor. Part I — Energy levels and transition wavenumbers for $H_2{}^{17}O$ and $H_2{}^{18}O$. J Quant Spectrosc Radiat Transf 2009;110:573-96. https://doi.org/10.1016/j.jqsrt.2009.02.014
42. Tennyson J, Bernath PF, Brown LR, Campargue A, Császár AG, Daumont L, et al. IUPAC critical evaluation of the rotational-vibrational spectra of water vapor. Part II: Energy levels and transition wavenumbers for $HD^{16}O$, $HD^{17}O$, and $HD^{18}O$. J Quant Spectrosc Radiat Transf 2010;111:2160-84. https://doi.org/10.1016/j.jqsrt.2010.06.012
43. Tennyson J, Bernath PF, Brown LR, Campargue A, Császár AG, Daumont L, et al. IUPAC critical evaluation of the rotational-vibrational spectra of water vapor, Part III. Energy levels and transition wavenumbers for $H_2{}^{16}O$. J Quant Spectrosc Radiat Transf 2013;117:29-58. https://doi.org/10.1016/j.jqsrt.2012.10.002
44. Tennyson J, Bernath PF, Brown LR, Campargue A, Császár AG, Daumont L, et al. IUPAC critical evaluation of the rotational-vibrational spectra of water vapor. Part IV. Energy levels and transition wavenumbers for $D_2{}^{16}O$, $D_2{}^{17}O$, and $D_2{}^{18}O$. J Quant Spectrosc Radiat Transf 2014;142:93-108. https://doi.org/10.1016/j.jqsrt.2014.03.019
45. Furtenbacher T, Tobias R, Tennyson J, Polyansky OL, Császár AG. W2020: A database of validated rovibrational experimental transitions and empirical energy levels for $H_2{}^{16}O$. J Phys Chem Ref Data 2020;49:033101. https://doi.org/10.1063/5.0008253
46. Tobias R, Furtenbacher T, Császár AG. Cycle bases to the rescue. J Quant Spectrosc Radiat Transf 2017;203:557-64. https://doi.org/10.1016/j.jqsrt.2017.03.031
47. Tobias R, Furtenbacher T, Tennyson J, Császár AG. Accurate empirical rovibrational energies and transitions of $H_2{}^{16}O$. Phys Chem Chem Phys 2019;21:3473-95. https://doi.org/10.1039/C8CP05169K
48. Vasilchenko S, Mikhailenko SN, Campargue A. Water vapor absorption in the region of the oxygen A-band near 760 nm. J Quant Spectrosc Radiat Transf 2021;275:107847. https://doi.org/10.1016/j.jqsrt.2021.107847
49. Campargue A, Mikhailenko SN, Kassi S, Vasilchenko S. Validation tests of the W2020 energy levels of water vapor. J Quant Spectrosc Radiat Transf 2021;276:107914. https://doi.org/10.1016/j.jqsrt.2021.107914
50. Vasilchenko S, Mikhailenko SN, Campargue A. Cavity ring down spectroscopy of water vapor near 750 nm: a test of the HITRAN2020 and W2020 line lists. Mol Phys 2022;120:e2051762. https://doi.org/10.1080/00268976.2022.2051762
51. Esplin MP, Watson RB, Hoke ML, Rothman LS. High-temperature spectrum of $H_2O$ in the 720 – 1400 $cm^{-1}$ region. J Quant Spectrosc Radiat Transf 1998;60:711-39. https://doi.org/10.1016/S0022-4073(98)00079-X
52. Zobov NF, Polyansky OL, Tennyson J, Lotoski JA, Calarusso P, Zhang KQ, Bernath P.F. Hot bands of water up to $6\nu_2$-$5\nu_2$ in the 933 – 2500 $cm^{-1}$ region. J Mol Spectrosc 1999;193:118-36. https://doi.org/10.1006/jmsp.1998.7732

53. Zobov NF, Polyansky OL, Tennyson J, Shirin SV, Nassar R, Hirao T, et al. Using laboratory spectroscopy to identify lines in the *K*- and *L*-band spectrum of water in a sunspot. Astrophys J 2000;530:994-8. https://doi.org/10.1086/308419
54. Tereszchuk K, Bernath PF, Zobov NF, Shirin SV, Polyansky OL, Libeskind NI, et al. Laboratory spectroscopy of hot water near 2 microns and sunspot spectroscopy in the *H*-band region. Astrophys J 2002;577:496-500. https://doi.org/10.1086/342167
55. Zobov NF, Shirin SV, Polyansky OL, Barber RJ, Tennyson J, Coheur P-F, et al. Spectrum of hot water in the 2000 – 4750 $cm^{-1}$ frequency range. J Mol Spectrosc 2006;237:115-22. https://doi.org/10.1016/j.jms.2006.03.001
56. Zobov NF, Shirin SV, Ovsyannikov RI, Polyansky OL, Barber RJ, Tennyson J, et al. Spectrum of hot water in the 4750–13 000 $cm^{-1}$ wavenumber range (0.769–2.1 μm). Month Not R Astron Soc 2008;387:1093-8. https://doi.org/10.1111/j.1365-2966.2008.13234.x
57. Rutkowski L, Foltynowicz A, Schmidt FM, Johansson AC, Khodabakhsh A, Kyuberis AA, et al. An experimental water line list at 1950 K in the 6250 – 6670 $cm^{-1}$ region. J Quant Spectrosc Radiat Transf 2018;205:213-9. https://doi.org/10.1016/j.jqsrt.2017.10.016
58. Czinki E, Furtenbacher T, Császár AG, Eckhardt AK, Mellau GCh. The 1943 K emission spectrum of $H_2{}^{16}O$ between 6600 and 7050 $cm^{-1}$. J Quant Spectrosc Radiat Transf 2018;206:46-54. https://doi.org/10.1016/j.jqsrt.2017.10.028
59. Flaud J-M, Camy-Peyret C, Maillard JP. Higher ro-vibrational levels of $H_2O$ deduced from high resolution oxygen-hydrogen flame spectra between 2800 – 6200 $cm^{-1}$. Mol Phys 1976;32:499-521. https://doi.org/10.1080/00268977600103251
60. Pine AS, Coulombe MJ, Camy-Peyret C, Flaud J-M. Atlas of the high-temperature water vapor spectrum in the 3000 – 4000 $cm^{-1}$ region. J Phys Chem Ref Data 1983;12:413-65. https://doi.org/10.1063/1.555689
61. Dana V, Mandin J-Y, Camy-Peyret C, Flaud J-M, Rothman LS. Rotational and vibrational dependences of collisional linewidths in the $nv_2$-$(n-1)v_2$ hot bands of $H_2O$ from Fourier-transform flame spectra. Appl Opt 1992;31:1179-94. https://doi.org/10.1364/AO.31.001179
62. Lanquetin R, Coudert LH, Camy-Peyret C. High-lying rotational levels of water: an analysis of the energy levels of the five first vibrational states. J Mol Spectrosc 2001;206:83-103. https://doi.org/10.1006/jmsp.2001.8300
63. Kyuberis AA, Zobov NF, Naumenko OV, Voronin BA, Polyansky OL, Lodi L, et al. Room temperature line lists for deuterated water. J Quant Spectrosc Radiat Transf 2017;203:175-85. https://doi.org/10.1016/j.jqsrt.2017.06.026
64. Coudert LH, Chelin P. Line position and line intensity analyses of the high-resolution spectrum of $H_2{}^{18}O$ up to the first triad and $J = 17$. J Mol Spectrosc 2016;326:130-5. https://doi.org/10.1016/j.jms.2016.01.012
65. Coudert LH, The bending potential energy function of HDO obtained from high-resolution data. J Mol Spectrosc 2016;330:112-9. https://doi.org/10.1016/j.jms.2016.07.008