Frequency comb-referenced cavity ring-down spectroscopy of natural water between 8041 and 8633 cm−1

A.O. Koroleva, S.N. Mikhailenko, S. Kassi, A. Campargue

Frequency comb-referenced Cavity Ring Down Spectroscopy of natural water between 8041 and 8633 cm^{-1}

A.O. Koroleva [1,2], S.N. Mikhailenko [3,4] S. Kassi [1], A. Campargue [1,*]

[1] *Univ. Grenoble Alpes, CNRS, LIPhy, Grenoble, France*
[2] *Institute of Applied Physics of RAS, Nizhniy Novgorod, Russia*
[3] *Laboratory of Theoretical Spectroscopy, 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*

Running head: CRDS spectrum of water vapor near 1.2 µm

Number of pages: 33
Number of figures: 10
Number of Tables: 5

Keywords: Frequency comb; cavity ring down spectroscopy; Water; H_2O; Rovibrational assignments; HITRAN

∗ Corresponding author. *E-mail address:* Alain.Campargue@univ-grenoble-alpes.fr (A. Campargue)

Abstract

The 1.25 µm atmospheric transparency window is of importance for a number of atmospheric applications. As a continuation of our previous works on the improvement of water vapour line parameters in the near infrared, the room temperature absorption spectrum of water vapour in natural isotopic abundance is recorded with unprecedented sensitivity between 8041 and 8633 cm^{-1} using comb referenced cavity ring-down spectroscopy. The line positions and intensities of more than 5400 lines were retrieved. Their intensities range between 3.6×10^{-30} and 1.5×10^{-22} cm/molecule. The high sensitivity and low noise level of the recordings ($\alpha_{min}\approx 10^{-11}$ cm^{-1}) allow for measuring more than 1600 new lines and determine their positions with an accuracy of about 10^{-4} cm^{-1} in the case of isolated features. The rovibrational assignments were performed using known experimental energy levels and calculated spectra based on variational calculations by Schwenke and Partridge. The final line list is assigned to more than 5400 transitions of the first six water isotopologues ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, $HD^{16}O$, $HD^{18}O$ and $HD^{17}O$). The measured line positions allow to determine the energy of 79 new levels of $H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, and $HD^{16}O$, and to correct 141 previously reported term values. Although a good agreement is generally observed, the comparison to the water vapor line lists provided by the HITRAN2020 spectroscopic database and to the W2020 transition frequencies reveals a number of discrepancies both for line positions and line intensities. The lack of traceability of some HITRAN line parameters and some biases in the derivation procedure of the W2020 energy levels are confirmed in the studied range. Validation tests of the theoretical values of the line intensities against measured values show both band-by-band variations of the deviations on the order of a few % and line-by-line fluctuations within a given band.

1. Introduction

The present work is devoted to a detailed analysis of the highly sensitive water vapor absorption spectrum between 8041.46 and 8633.41 cm^{-1} by comb-referenced cavity ring down spectroscopy (CR-CRDS). The investigated spectral interval is included in the overview displayed in **Fig. 1**. It corresponds to the high energy edge of the 1.25 µm atmospheric window. The present study extends to higher energy the spectral range of our series of CRDS studies which now cover continuously the 5690-8633 cm^{-1} range [1-7]. Note that the 7911 – 8337 cm^{-1} interval, largely overlapping with the presently studied region, was previously investigated by CRDS in 2015 [6]. In the present work, a higher accuracy on the line parameters is expected from the referencing of the CRDS ring downs to a self-referenced frequency-comb (line position accuracy at a 1×10^{-4} cm^{-1} level or better are achieved for unblended lines). This improvement compared to the 2015 measurements results from *(i)* the absolute frequency calibration of the spectra recorded in this work while reference line positions had to be used to calibrate the spectra of Ref. [6] and *(ii)* a better line profile determination as an absolute frequency value is associated to each ring-down event.

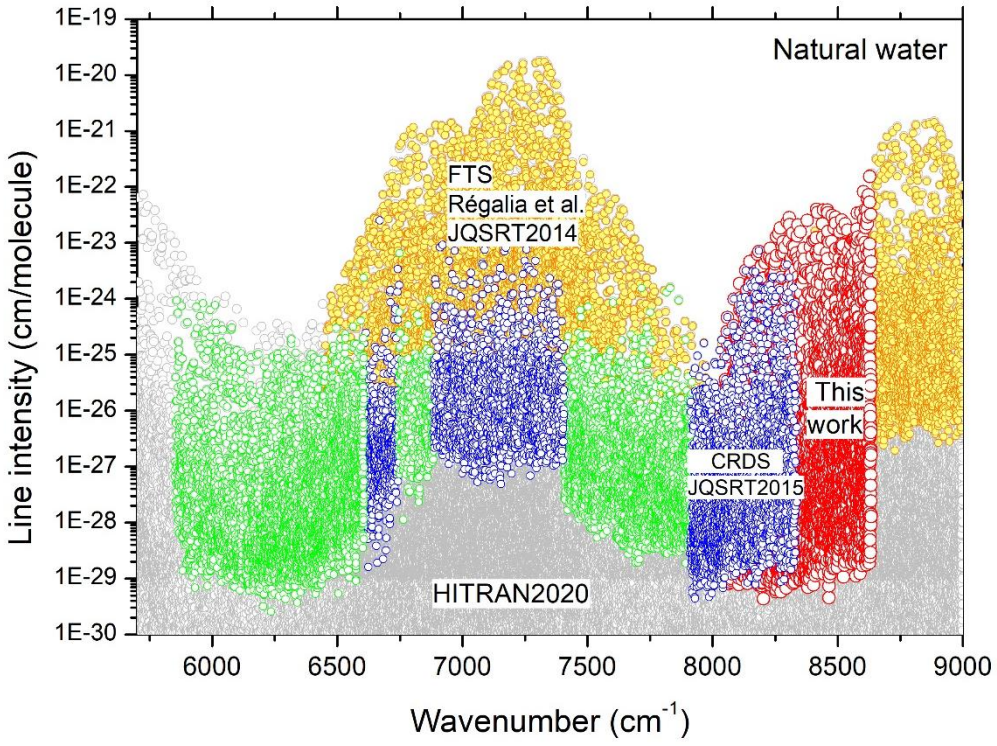

Fig. 1.
Overview of the water vapor spectrum between 5700 and 9000 cm^{-1}. Experimental observations are superimposed to the HITRAN2020 line list [9] (grey circles). The present study (red open circles) extends to higher energy our series of previous CRDS studies (alternate green and blue circles) [1-7]. Note in particular, the CRDS study in the 7911 – 8337 cm^{-1} interval [6], largely overlapping the present study. The line list elaborated by Régalia et al. from a series of FTS spectra [8] is also plotted for comparison (yellow dots).

We have included in **Fig. 1**, the line list of the most sensitive study by Fourier transform spectroscopy (FTS) available in the literature. It was performed by Régalia et al, in 2014, with absorption pathlengths up to 1203 m [8]. The sensitivity of the CRDS technique allows lowering the detection limit by two to three orders of magnitude (The weakest lines measured by CRDS have an intensity smaller than 10^{-29} cm/molecule). As a result, in the present work, an important number of transitions are newly observed above

8337 cm^{-1}. These transitions are included in the usual spectroscopic databases (HITRAN [9], GEISA [10]) with calculated values of their line positions and line intensities. The present set of measurements will thus allow for valuable validation tests of the existing databases (see Section 4). In the next Section 2, the experimental set-up and the line list construction are presented. The rovibrational analysis leading to the derivation of a number of new energy levels will be detailed in Section 3.

2. Experimental set up

The room temperature absorption spectrum of natural water vapor was recorded by high sensitivity frequency comb referenced cavity ring-down spectroscopy [11], in flow regime at a pressure of 1.0 Torr. The weak flow was set through a needle valve connecting the cell to a turbo pump group. The pressure was continuously measured by a capacitance gauge (MKS Baratron, 10 Torr, 0.25% accuracy of the reading) and actively regulated to 1.0 Torr using a computer based Proportional/Integral controller.

An external cavity diode laser (ECDL) was used as light source to cover the 8040 – 8630 cm^{-1} range. The spectrum quality is illustrated by three successive zooms presented in **Fig. 2.** The noise equivalent absorption is lower than 10^{-11} cm^{-1}.

Following Refs. [12,13], a self-referenced frequency comb (Model FC 1500-250 WG from Menlo Systems) was used for the frequency calibration of the CRDS spectra. As detailed in Ref. [11], an accurate frequency is determined "on the fly" and attached to each ring down event. The frequency determination requires *(i)* the frequency measurement of the beat note between a fraction of the ECDL light and a tooth of the frequency comb and *(ii)* the tooth number deduced from the frequency value provided by a commercial Fizeau type wavemeter (HighFinesse WSU7-IR, 5 MHz resolution, 20 MHz accuracy over 10 hours). As illustrated in Ref. [12], the CR-CRDS line centers have in routine an absolute accuracy at the 1×10^{-4} cm^{-1} level or better for unblended lines.

The line parameters were retrieved from the spectra by using a homemade interactive least squares multi-lines fitting written in LabVIEW. A Voigt profile with the width of the Gaussian component fixed to the calculated Doppler broadening was adopted for each line (see an example of spectrum reproduction in **Fig. 3**). Note that, in order to improve the retrieved line parameters, in a number of spectral intervals corresponding to strong overlapping of several transitions, we included in the fit lines with line parameters constrained to their literature values (empirical line positions and variational line intensity). The parameters of these 1213 "frozen" lines cannot be reliably retrieved but their inclusion in the fit is expected to help to increase the accuracy of the fitted parameters of the nearby overlapping lines. The "frozen" lines are given for information in the experimental list (tag "F" in the first column) but are excluded from the following discussion and comparisons.

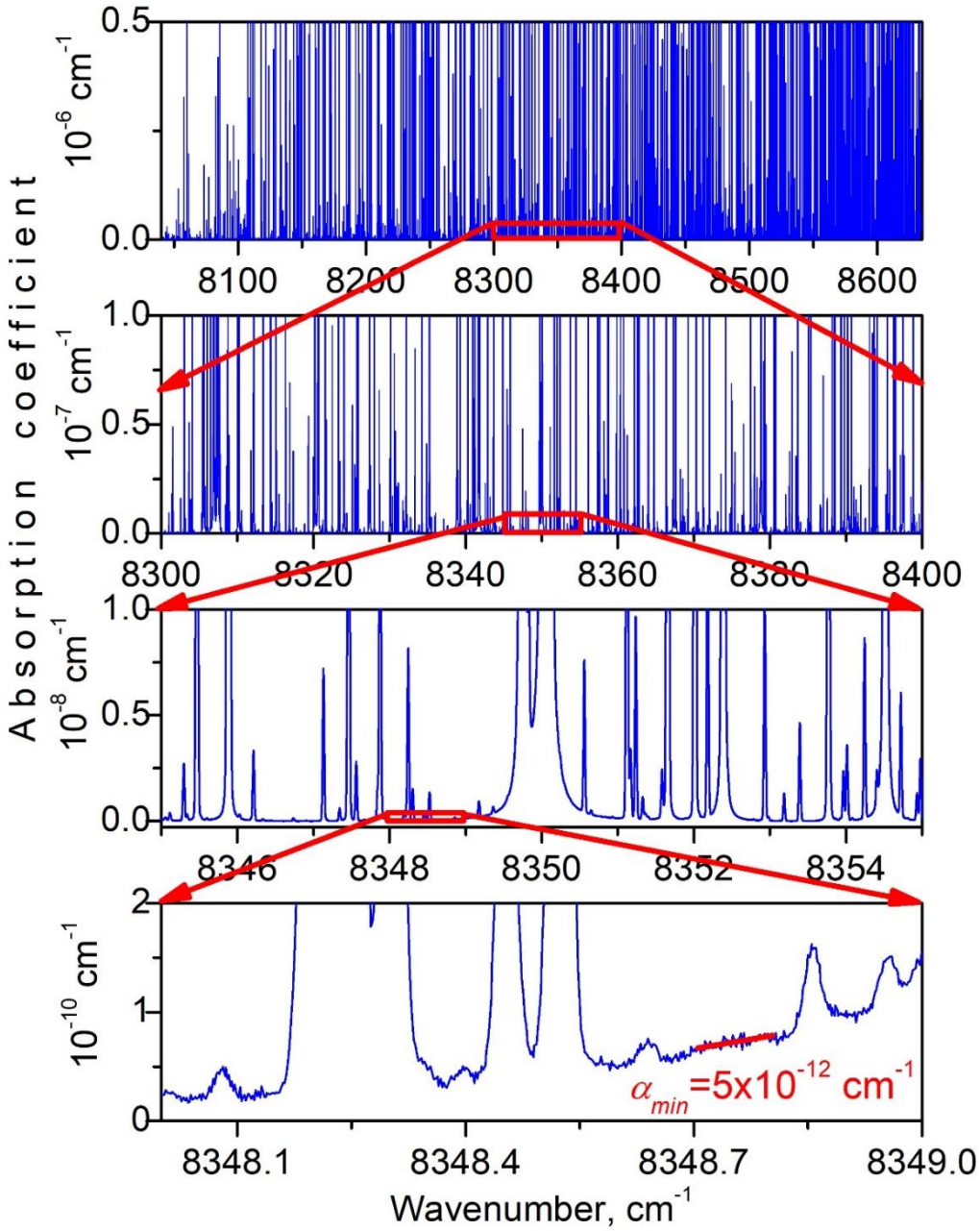

Fig. 2.
Room temperature CRDS spectrum of natural water vapor around 8348 cm^{-1}. The sample pressure was about 1.0 Torr. The successive enlargements illustrate the high dynamics of the recordings and the noise level on the order of α_{min} ~ 5×10^{-12} cm^{-1}.

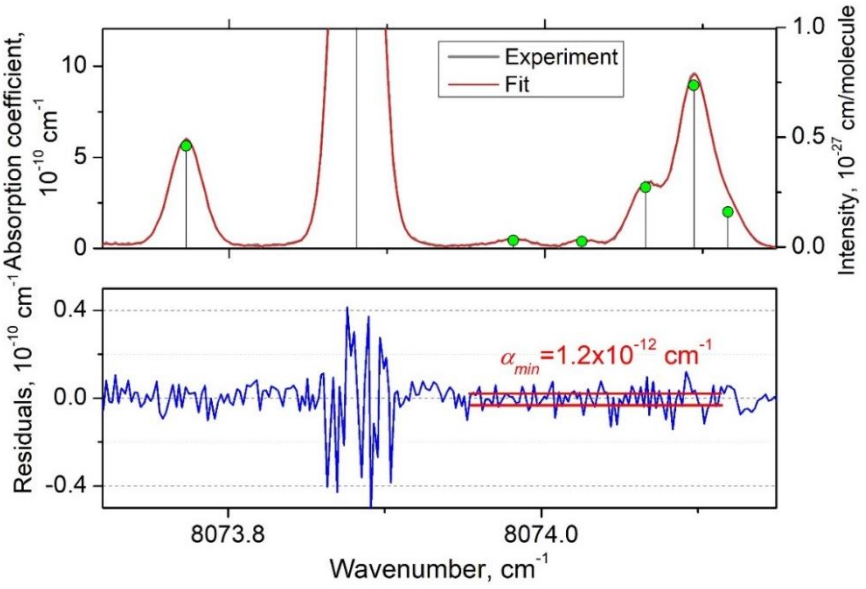

Fig. 3.
Example of spectrum reproduction of the CRDS spectrum of water vapor near 8074 cm^{-1}. The obtained line list is superimposed as a stick spectrum (green dots). The two weak lines around 8074 cm^{-1} have an intensity of about 2×10^{-29} cm/molecule.

Overall, line parameters of 5447 absorption features were retrieved from the recorded spectrum between 8041.46 and 8633.41 cm^{-1}. 184 and 55 of them are due to NH_3 and CO_2 molecules, respectively. They were identified by position and intensity comparison with HITRAN values. As detailed in the next section, 5191 lines were assigned to 5430 transitions of six water isotopologues (red open circles in **Fig. 1**). Their intensities range between 3.6×10^{-30} and 1.5×10^{-22} cm/molecule. Detailed information about the assignment procedure is given in the next section and summarized in **Table 1**. Twenty weak lines were left unassigned at the end of the assignment process.

Table 1.
Statistical overview of the water transitions previously reported and observed in the present study between 8041.4 and 8633.5 cm^{-1}.

Molecule	This study				Published data				
	N_{TW} [a]	Region, cm^{-1}	J_{max}	$K_{a\,max}$	$N_{Lit.}$ [b]	Region, cm^{-1}	J_{max}	$K_{a\,max}$	N_{new} [c]
$H_2{}^{16}O$	3237	8041.4 – 8633.4	19	10	2508	8041.4 – 8633.4	20	10	980 [d]
$H_2{}^{18}O$	911	8043.9 – 8632.5	15	8	869	8043.9 – 8632.5	14	7	208
$H_2{}^{17}O$	540	8042.5 – 8632.5	15	7	282	8042.6 – 8632.3	13	6	291
$HD^{16}O$	707	8041.6 – 8631.1	15	9	785	8041.6 – 8633.5	16	8	170
$HD^{18}O$	32	8443.0 – 8633.2	8	5	382	8042.7 – 8633.2	13	8	
$HD^{17}O$	3	8538.3 – 8595.3	3	2	30	8469.3 – 8629.8	7	4	

Notes
[a] N_{TW} – number of transitions assigned in the present work
[b] $N_{Lit.}$ – number of transitions reported from previous absorption studies
[c] N_{new} – number of new transitions reported in this work. Note that N_{new} is not equal to $N_{TW} - N_{Lit.}$ because due to saturation effect of some strong lines or spectral gaps in our CRDS spectrum, some previously reported transitions ($N_{Lit.}$) could not be measured in the current study.
[d] This number does not consider the (less accurate) measurements by emission spectroscopy [14].

3. Rovibrational assignments

As in our previous studies (see, for example, Refs. [6,8,15]) the vibration-rotation assignments of the water lines were performed using variational vibration-rotation (VR) line lists computed by S.A. Tashkun

[16] based on the results of Schwenke and Partridge [17,18] (hereafter "SP line lists") and empirical values of VR levels known from previous studies. Here, we are using a set of VR energies previously used for the construction of an accurate empirical water line list in the 5690 – 8340 cm^{-1} region [19].

Most of the lines could be straightforwardly assigned by comparison to the literature and/or using line positions calculated using empirical VR energies (so called "trivial assignment"). This is in particular the case of the $HD^{18}O$ and $HD^{17}O$ isotopologues which are present at very low abundance in our natural water sample (6.23×10^{-7} and 1.16×10^{-7}, respectively). All lines associated to these two molecules (35 transitions of the $\nu_2+2\nu_3$ band, in total) were previously observed by FTS of deuterated water highly enriched in ^{18}O [15].

Table 2. Term values of newly determined energies of four water isotopologues

$V_1V_2V_3$	J	K_a	K_c	Energy	dE	NT
				$H_2{}^{16}O$		
012	15	1	15	11267.59539	76	
040	13	9	4	10491.32266	217	
050	11	6	6	10469.34626	80	
050	12	7	5	11095.00221	305	
051	3	3	0	11763.21886	426	
070	2	0	2	10155.14662	71	
070	4	0	4	10314.95101	58	
111	13	7	6	11709.19591	186	
111	14	7	7	12075.24608	308	
120	14	10	5	11005.62060	540	
130	10	6	4	10458.46562	54	3
130	12	7	6	11271.80777	120	
130	12	7	5	11271.87771	371	
130	13	5	9	11073.64853	369	
130	14	4	10	11271.89153	104	2
130	15	4	11	11633.90887	46	
210	9	9	1	11034.93160	423	
210	9	9	0	11034.93160	423	
210	11	7	5	11114.88878	198	
210	12	2	10	10683.92477	342	
210	12	4	8	10937.90318	212	
210	14	0	14	10743.67461	97	
210	14	2	12	11244.99079	218	
210	15	1	14	11309.67474	274	
				$H_2{}^{17}O$		
031	8	6	2	10032.79258	108	
031	9	2	7	9610.15869	280	
031	9	6	4	10250.76622	338	
031	9	6	3	10250.86402	771	
031	10	3	8	9887.88534	527	
111	11	3	8	10584.43224	255	
111	11	5	6	10768.69671	443	
111	15	1	15	11047.63569	241	
130	7	3	5	9167.12046	51	
130	9	3	6	9625.44662	365	
130	9	4	5	9758.80525	226	
130	10	2	9	9596.49369	221	
210	7	3	5	9557.02032	236	2
210	9	4	6	10084.37692	124	

$V_1V_2V_3$	J	K_a	K_c	Energy	dE	NT
				$H_2{}^{18}O$		
002	14	2	12	9873.26122	330	
012	9	5	5	10429.63069	210	
012	10	5	6	10669.45636	138	
031	12	5	7	10808.76445	200	
050	4	4	0	8445.45262	169	
060	9	1	8	10106.07540	201	
111	11	4	7	10651.83616	82	
111	11	6	5	10900.36303	352	
111	14	2	12	11264.23450	163	
111	15	2	14	11321.67216	268	
130	8	4	4	9523.28133	216	
130	8	5	3	9724.54545	174	2
130	8	6	3	9958.65018	371	
130	9	3	7	9548.31109	303	
130	9	4	6	9734.02573	194	3
130	9	5	4	9940.45713	93	
130	13	3	10	10764.20530	155	
140	1	1	0	9760.92728	315	
140	2	1	2	9794.92167	498	
140	2	2	1	9911.44404	245	
140	4	2	3	10074.51093	391	
210	10	2	8	10147.62425	240	
				$HD^{16}O$		
012	10	9	2	10604.62588	1085	
012	10	9	1	10604.62619	1085	
012	14	3	12	10342.65890	149	
012	15	1	15	10191.11483	278	
012	15	1	14	10388.94771	538	
012	15	2	14	10389.07700	498	
012	15	2	13	10557.95975	423	
031	8	7	2	9303.89502	135	2
031	8	7	1	9303.89510	135	2
031	10	5	6	9156.80832	515	
031	11	4	7	9164.40425	419	
220	6	4	2	8724.78613	480	
220	6	6	1	9088.02358	78	
220	6	6	0	9088.02355	79	
220	7	4	4	8830.65087	390	
220	7	5	3	8996.39853	251	
220	7	5	2	8996.39920	249	
300	10	8	3	9650.68217	615	
300	10	8	2	9650.68200	619	

Notes

$V_1V_2V_3$ – vibration quantum numbers; J K_a K_c – rotation quantum numbers; Energy/cm^{-1} – empirical term values; dE – estimated uncertainty of the term value in 10^{-5} cm^{-1} units; NT – number of line positions used for the energy determination if it is larger than 1.

Overall, less than 300 lines corresponding to new or inaccurate upper state energy levels required a non-trivial identification. The term values of the 79 newly determined levels are listed in **Table 2.** In a number of cases, it was possible to use several lines to determine a given upper level. The number (*NT*) of used line positions is given in the last column of the table. One hundred and forty energy levels were found to deviate by more than 5×10^{-3} cm^{-1} compared to the literature. They will be given and discussed in the following sections. The experimental line list is provided as Supplementary Material. It includes for each line, the line position and line intensity with their uncertainty as provided by the fit, the empirical value of the lower state energy, together with the comparison with the line parameters provided by the HITRAN database [9] or included in the W2020 list [20] (see next section), when available.

4. Line positions - Comparison with literature

4.1. Overview comparison

The literature review indicates that previous analyses of high resolution absorption spectra of water vapor in the 8041 – 8633 cm^{-1} range were reported in Refs. [6,8,21-25] for the main isotopologue ($H_2{}^{16}O$) and in Refs. [6,8,15,24,26], Refs. [6,8,24,26], Refs. [6,8,15,24,27,28], and Refs. [12,29,30] for the $H_2{}^{18}O$, $H_2{}^{17}O$, $HD^{16}O$, and $HD^{18}O$ minor isotopologues, respectively. $HD^{17}O$ transitions were reported in Ref. [15]. Additional vibration-rotation transitions of the $H_2{}^{16}O$ molecule have been published by Zobov et al. [14] from an analysis of high temperature emission spectra. A statistical comparison of our observations to previous data is included in **Table 1**. Overall, more than 1600 transitions of the four most abundant isotopologues were observed for the first time (they include the about 300 transitions used to derive new or corrected energy levels, discussed just above).

In the following, we limit mainly the position and intensity comparisons to the following line lists: the FTS study by Régalia et al. [8] (hereafter, FTS2014), our previous CRDS study in the 7911 – 8337 cm^{-1} interval (hereafter, CRDS2015) [6], the HITRAN2020 list [9] and the W2020 lists [20] (the latter being available only for non-deuterated species, $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$). In the case of the deuterated species, the energy levels and transition wavenumbers recommended by an IUPAC task group (IUPAC-TG, hereafter) will be considered for $HD^{16}O$, $HD^{17}O$, and $HD^{18}O$ [31].

The overview of the deviations of the positions provided by these four sources from the present measurements is plotted in **Fig. 4**. This figure deserves several comments.

An overall reasonable agreement is observed but all the four considered references [6,8,9,20] seem to have their positions systematically larger than our values (TW). The histograms of the $(\nu_{TW} - \nu_{Ref.})$ position differences between our values and those of the FTS2014, CRDS2015 and W2020 lists were plotted and fitted with a Gaussian. The fitted value of the center of the Gaussian was found to be -8.2×10^{-4}, -7.4×10^{-4}, and -3.4×10^{-4} cm^{-1}, respectively. These values are largely above the claimed uncertainty of our frequency calibration performed with the help of a self-referenced frequency comb. In order to understand the origin of these apparent systematic shifts, we examine the frequency

calibration of the other sources. In our spectral region, the frequency calibration of the FTS2014 spectrum [8] was made difficult by the lack of accurate reference lines which obliged Régalia et al. to calibrate their spectra on « a set of accurate line positions of the main isotopologue with positions calculated with well-known experimental energy levels ». The accuracy of the resulting calibration of the FTS2014 spectra is not reported in Ref. [8]. The frequency calibration of the CRDS2015 spectra required also accurate reference lines to refine frequency values provided by a wavemeter [6]. Although it is not explicitly indicated in Ref. [6], we used the most accurate line positions available at that time, namely the FTS2014 position values to calibrate the CRDS spectra. This is confirmed by the very similar average deviation of the FTS2014 and CRDS2015 line positions compared to the present values (see **Fig. 4**).

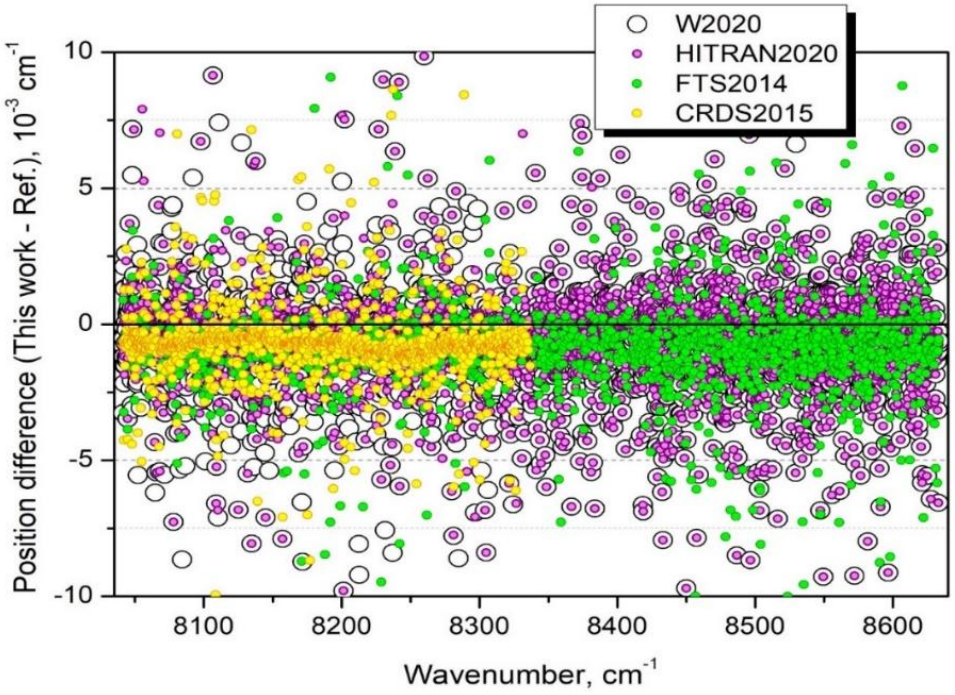

Fig. 4.
Position differences of the water vapor lines measured by CRDS between 8040 and 8620 cm^{-1} compared to various literature sources: FTS line list elaborated by Régalia et al. (FTS2014 – green dots) [8], pour previous CRDS line list in the 7911 – 8337 cm^{-1} interval (CRDS2015 – yellow dots) [6], HITRAN2020 values [9] and W2020 line list [20].

As concerns the systematic difference with HITRAN2020 position values, it is difficult to draw a conclusion as the HITRAN2020 list uses a variety of sources with experimental or theoretical origins of different quality (see below). In particular, many of the HITRAN2020 positions coincide with W2020 empirical line positions which explains the close average deviations of the two datasets compared to our values.

The W2020 line positions [20] have the advantage to be provided with individual error bars which should, in principle, help to disentangle the situation. Although our previous studies have demonstrated that the W2020 error bars should be used with caution [32,33] because they are frequently strongly underestimated, we selected for comparison the most accurate W2020 line

positions *i.e.* those which are given with an uncertainty smaller than 5×10^{-4} cm^{-1}. About 1400 lines of $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$ were found in common (see Supplementary Material where the W2020 position uncertainty is reproduced). The center of the corresponding histogram of the position deviations was fitted at -1.7×10^{-4} cm^{-1}, instead of -3.4×10^{-4} cm^{-1} if the whole W2020 dataset is used. This improvement of the level of agreement by a factor of 2 may be interpreted as a confirmation of the accuracy of the calibration of our spectra. Although the FTS2014 and CRDS2015 line positions are part of the large number of sources used to derive the W2020 empirical energy levels, the impact of the small calibration error of these two sources on the W2020 positions is probably reduced by the involvement of other sources in the W2020 transition database used to derive the W2020 energy levels [20].

As further checks of the accuracy of the frequency calibration of our spectra, we consider a few individual lines with line positions accurately measured in the literature. We first compared the water position values to those measured in low pressure CO spectra where water was present as an impurity [34]. The CO spectra were recorded by comb-reference CRDS with the same setup used in this work. Over a set of 34 lines, excluding two outliers, the average of the water line position differences was found to be of -2.9×10^{-5} cm^{-1} with a standard deviation of 1.4×10^{-4} cm^{-1}. This excellent agreement obtained for water lines which are very weak in the CO spectrum and strong ($S > 10^{-24}$ cm/molecule) in the pure water spectrum gives confidence in the absence of bias in the calibration of the two considered spectra. The same conclusion was drawn from validation tests performed using lower state combination difference relations between the present measurements and $H_2{}^{16}O$ line positions measured at lower frequency by frequency-comb spectroscopy of water vapor enriched in ^{17}O [12,13].

Finally, for an independent check, we consider the set of 70 positions of isolated lines measured by Sironneau and Hodges between 7710 and 7920 cm^{-1}, reported with an average combined uncertainty of approximately 3 MHz (10^{-4} cm^{-1}) [35]. From these measured positions, it was possible to determine accurate upper state energy levels shared by 26 transitions observed in our region. These transitions are all weak ($10^{-29} < S < 5\times10^{-27}$ cm/molecule). If we limit the comparison to the four lines with intensity larger than $S > 10^{-27}$ cm/molecule, we obtain an average position differences of 2.2×10^{-5} cm^{-1} with a standard deviation of 7.5×10^{-5} cm^{-1} (if all but one transitions are compared, the average and standard errors are 2.5×10^{-5} cm^{-1} and 7.5×10^{-4} cm^{-1}, respectively). This comparison to an independent dataset confirms the reliability of the frequency calibration of our spectra and that the systematic shifts observed compared to FTS2014, CRDS2015, HITRAN2020 and W2020 are not due to the present data.

The position deviations displayed on **Fig. 4** show a significantly smaller dispersion for CRDS2015 compared to the three other sources (roughly by a factor of 4). The visual impression given by **Fig. 4** is valid when CRDS data (below 8337 cm^{-1}) are compared to HITRAN2020 and W2020 because these lists are mostly complete and thus the number of plotted position differences is similar. In the case of the less sensitive FTS2014 data, the number of lines in common is reduced but

the average scattering of the position differences is similar to that of the HITRAN2020 and W2020 lists. A number of large discrepancies (or position outliers) are noted for all the sources but appear to be more frequent for HITRAN2020 and W2020. In the following we will examine in details some of these outliers.

4.2. Comparison with FTS2014 [8] (Régalia et al. JQSRT 136 (2014) 119-136)

The overview of the FTS2014 list is included in **Fig. 1**. In our region, the FTS2014 list includes 2008 lines attributed to 2225 transitions of four water isotopologues ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, and $HD^{16}O$) with minimum intensity values of a few 10^{-27} cm/molecule. Although a number of absorption features could be fitted as multiplets while they are reported as single lines in FTS2014, all the FTS2014 assignments are confirmed in the present analysis. The comparison shows a good agreement for the line positions. Nevertheless, as discussed above (**Fig. 4)**, there is a systematic difference between two sets of the line positions. The average deviation for 1910 line positions is -9.67×10^{-4} cm^{-1} with a largest difference of about 0.018 cm^{-1} for the $H_2{}^{18}O$ $\nu_1+\nu_2+\nu_3$ $3_{2\,2}$ – $4_{4\,1}$ line at 8502.54803 cm^{-1}. The root mean square ratio for 1907 line intensities is 0.975.

4.3. Comparison with CRDS2015 [6] (Campargue et al. JQSRT 157 (2015) 135-152)

In the 8041 - 8337 cm^{-1} interval in common, the CRDS2015 spectrum and the present study count 2163 and 2176 lines assigned to 2284 and 2262 water transitions, respectively. In spite of the very good consistency of the two datasets, the position differences of 72 transitions were found to exceed 0.005 cm^{-1}. This is mainly related to the presence of lines of ammonia present as an impurity in both spectra. In the CRDS2015 spectrum, the overlapping with ammonia lines limited the accuracy of the line parameters of some weak water lines [6]. A direct comparison of the two spectra shows that the amount of ammonia in the new spectrum is about 20 times smaller than in CRDS2015 spectrum. However, a small number of significant differences in the line positions are due to other reasons. In order to give to interested readers an idea of the origin of the largest deviations, we detail below the few transitions showing position differences larger than 0.015 cm^{-1}.

First, let us mention that the term value of the $H_2{}^{16}O$ 002 $17_{3\,15}$ level (10892.9135 cm^{-1} in Table 2 of Ref. [6]) appears to be incorrect. This value was derived in Ref. [6] from a single transition $2\nu_3$ $17_{3\,15} - 16_{0\,16}$ at 8231.9671 cm^{-1}. This line is absent in the (more sensitive) new spectrum under analysis. It is thus probably due to an impurity and should be removed from the water list of Ref. [6].

The largest difference, $\nu_{TW} - \nu_{CRDS2015}$ = 0.0605 cm^{-1}, corresponds to the $2\nu_3$ $17_{5\,12} - 16_{4\,13}$ transition of $H_2{}^{16}O$. In Ref. [6], this transition was attributed as second assignment of the line at 8065.4741 cm^{-1} assigned to the $5\nu_2$ $7_{5\,2} - 8_{4\,5}$ transition of $H_2{}^{18}O$. As indicated in the Supplementary Materials of Ref. [19], this assignment of the $2\nu_3$ $17_{5\,12} - 16_{4\,13}$ transition relied on the IUPAC energy value of the 002 $17_{5\,12}$ upper level, 11504.759237(1905) cm^{-1} [36]. In fact, the $2\nu_3$ $17_{5\,12} - 16_{4\,13}$ transition of $H_2{}^{16}O$ should be attributed to the line at 8065.5346 cm^{-1}. Note that this line is present in the CRDS2015 spectrum but was missed in the line list of Ref. [6]. Consequently, the upper energy of

the 002 $17_{5\,12}$ level should be corrected by +0.0605 cm^{-1} compared to the value given in Table 3 of Ref. [6].

The next large difference, (ν_{TW} - $\nu_{CRDS2015}$) = -0.04745 cm^{-1}, concerns the $H_2{}^{16}O$ 050 $11_{6\,5}$ upper level. In Ref. [6], the $5\nu_2$ $11_{6\,5} - 12_{5\,8}$ assignment was attributed to the line at 8194.1965 cm^{-1}, the first assignment being to the $H_2{}^{18}O$ $\nu_1+3\nu_2$ $4_{2\,3} - 5_{1\,4}$ transition. This assignment of the $5\nu_2$ $11_{6\,5} - 12_{5\,8}$ transition relied on the IUPAC energy value of the 050 $11_{6\,5}$ upper level, 10469.566775(3062) cm^{-1} [36]. Now we attribute the $5\nu_2$ $11_{6\,5} - 12_{5\,8}$ transition to the line at 8194.14905 cm^{-1} observed in the present work. This assignment is confirmed by ground state combination difference relations (GSCD) with the $5\nu_2$ $11_{6\,5} - 11_{5\,6}$ transition at 8470.52401 cm^{-1}.

The $H_2{}^{17}O$ $\nu_1+3\nu_2$ $4_{3\,2} - 5_{2\,3}$ transition, (ν_{TW} - $\nu_{CRDS2015}$) = -0.03312 cm^{-1}, should be attributed to the line at 8294.67838 cm^{-1} (this study) and not to the line at 8294.7115 cm^{-1} (Ref. [6]). This assignment is confirmed by GSCD relation with the $\nu_1+3\nu_2$ $4_{3\,2} - 4_{2\,3}$ transition at 8441.03301 cm^{-1}. Note the line at 8294.6811 cm^{-1} was left unassigned in Ref. [6].

The lines of the $H_2{}^{16}O$ $\nu_1+2\nu_2$ $13_{10\,4} - 12_{7\,5}$ transition in both spectra (Ref. [6] and this study) are very weak noisy lines but we prefer the new value of the line position (8058.48831 cm^{-1} instead of 8058.4674 cm^{-1} [6]).

In Ref. [6], the $H_2{}^{16}O$ $\nu_1+\nu_3$ $18_{3\,16} - 17_{1\,17}$ transition was attributed to the line of an unresolved doublet at 8068.2458 cm^{-1}. In the present spectrum, this line was fitted as two components at 8068.24613 cm^{-1} ($18_{2\,16} - 17_{0\,17}$) and 8068.22529 cm^{-1} ($18_{3\,16} - 17_{1\,17}$). The upper energy of the 101 $18_{3\,16}$ level obtained from this latter position is 11049.58373 cm^{-1}. This term value is confirmed by its agreement with the IUPAC-TG value of 11049.5814 cm^{-1} [36] and by the line position of the $\nu_1+\nu_3$ $18_{3\,16} - 17_{3\,15}$ transition predicted at 7482.32953 cm^{-1}, in excellent agreement with its experimental value of Ref. [3] (7482.3297 cm^{-1}). We thus conclude that the 101 $18_{3\,16}$ level should have not be corrected in Ref. [6] (11049.6061 cm^{-1} in Table 3 of this reference).

From the present spectra, a more precise line center was derived for the very weak $H_2{}^{16}O$ $\nu_1+3\nu_2$ $14_{5\,10} - 15_{4\,11}$ transition at 8157.37971 cm^{-1} (S = 5.3×10^{-29} cm/molecule) instead of 8157.3604 cm^{-1} (S = 6.5×10^{-29} cm/molecule) in Ref. [6]. The corresponding 130 $14_{5\,10}$ upper level is thus corrected by 0.01931 cm^{-1}.

A similar situation is found for the line corresponding to the $H_2{}^{16}O$ $4\nu_2+\nu_3-\nu_2$ $10_{5\,5} - 11_{5\,6}$ transition at 8231.9271 cm^{-1} which is blended by impurity lines in the CRDS2015 spectra. The present line center determination at 8231.91083 cm^{-1} is confirmed by GSCD relation with the line position 8450.77756 cm^{-1} of the $4\nu_2+\nu_3-\nu_2$ $10_{5\,5} - 11_{3\,8}$ transition.

The line at 8249.5876 cm^{-1} assigned to the $H_2{}^{16}O$ $\nu_2+2\nu_3$ $13_{2\,12} - 14_{3\,11}$ transition in Ref. [6] is not observed in the present spectra and is believed to be due to an impurity. In the present work, the $H_2{}^{16}O$ $\nu_2+2\nu_3$ $13_{2\,12} - 14_{3\,11}$ transition is assigned to an extremely weak line at 8249.56889 cm^{-1}. (S = 7.4×10^{-30} cm/molecule). This assignment is confirmed by GSCD with the $\nu_2+2\nu_3$ $13_{2\,12} - 13_{1\,13}$

transition observed by Régalia et al. at 9182.3256 cm^{-1} [8]. The corresponding 012 $13_{2\,12}$ upper level is thus corrected by -0.018 cm^{-1}.

The present determination of the line position of the $HD^{16}O$ $3\nu_1$ $8_{3\,6} - 7_{1\,7}$ transition at 8179.97385 cm^{-1} instead of 8179.9901 cm^{-1} (Ref. [6]) is more correct. This position is much more consistent with line positions of other transitions reaching the 300 $8_{3\,6}$ upper state present in both spectra.

In Ref. [6], the $H_2{}^{18}O$ $3\nu_2+\nu_3$ $8_{5\,4} - 9_{5\,5}$ transition was attributed to the line of two unresolved transitions at 8325.5676 cm^{-1}. In the present spectrum, the corresponding line could be fitted as two components at 8325.56436 cm^{-1} ($H_2{}^{16}O$ $\nu_2+2\nu_3$ $9_{3\,7} - 10_{6\,4}$) and 8325.58345 cm^{-1} ($H_2{}^{18}O$ $3\nu_2+\nu_3$ $8_{5\,4} - 9_{5\,5}$), leading to a 0.016 cm^{-1} correction on the position of the latter transition.

4.4. Comparison with HITRAN2020 database

In our region, the water line list of the HITRAN2020 database [9] includes 11,650 transitions of seven isotopologues ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, $HD^{16}O$, $HD^{18}O$, $HD^{17}O$, and $D_2{}^{16}O$) between 8041.45 and 8633.41 cm^{-1} (see **Fig. 1**). The HITRAN intensity cut off is 10^{-30} cm/molecule, except for the $D_2{}^{16}O$ species for which the intensity cut off was fixed to 10^{-32} cm/molecule, for an unknown reason. Note that 1124 transitions of $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$ species do not have complete vibration-rotation assignment.

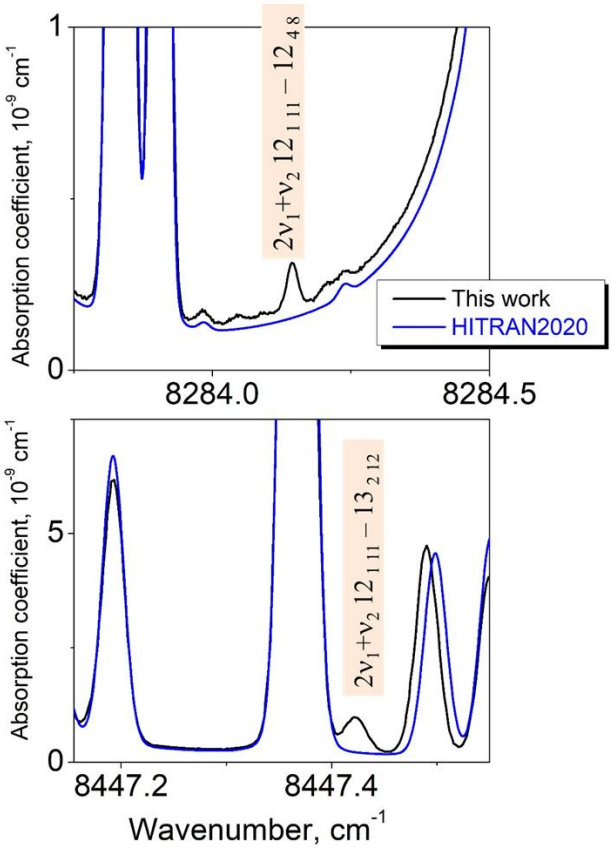

Fig. 5.
Two of the four $H_2{}^{16}O$ lines of the $2\nu_1+\nu_2$ band missing in the HITRAN2020 database.

First of all, among our 5429 assigned transitions, 49 are missing in the HITRAN line list. 45 of these 49 transitions belong to minor isotopologues and have their intensity near the HITRAN intensity cut-off: $H_2{}^{18}O$ (6 transitions), $H_2{}^{17}O$ (10 transitions) and $HD^{16}O$ (29 transitions). Nevertheless, the four $H_2{}^{16}O$ lines missing have larger intensities between 7.90×10^{-29} and 2.26×10^{-27} cm/molecule. They are $2\nu_1+\nu_2\ 12_{1\,11} - 12_{4\,8}$ (8284.14410, 1.04×10^{-28}), $2\nu_1+\nu_2\ 12_{1\,11} - 13_{2\,12}$ (8447.42231, 5.58×10^{-28}), $2\nu_1+\nu_2\ 10_{0\,10} - 11_{1\,11}$ (8503.04576, 7.90×10^{-29}), and $2\nu_1+\nu_2\ 12_{1\,11} - 12_{2\,10}$ (8529.59296, 2.26×10^{-27}). Two of them are shown in **Fig. 5**.

According to HITRAN reference code, 5608 of the 7141 of the line positions of the main isotopologue, $H_2{}^{16}O$, were taken from the W2020 line list [20]. This dataset is supplemented with positions from Refs. [19,37,38] (more than 1500) and several lines from Refs. [8,36]. Note none of the 502 transitions of Ref. [37] has full vibration-rotation assignment. Their line positions are not empirical but calculated using a semi-empirical potential energy surface. The main set of electric-dipole transitions is supplemented by 47 very weak electric quadrupole transitions with line positions calculated from W2020 empirical energy levels [20] and calculated line intensities [39].

As mentioned above, for 5608 transitions, the source of the line positions is the W2020 line list. However, a careful analysis of these data shows that for 40 transitions, the HITRAN positions differ from the original W2020 values [20]. The position difference reaches a value of 0.00569 cm^{-1}. The vibration-rotation identification of the transitions does not match for 25 transitions. In addition, the positions of seven transitions cannot be calculated from W2020 energies since the corresponding upper levels ($060\ 8_{8\,1}$, $070\ 6_{0\,6}$ and $080\ 2_{0\,2}$) are not in the list of energies of Ref. [20].

According to HITRAN's reference code, 1421 line positions of the $H_2{}^{18}O$ isotopologue are empirical values from the W2020 line list [20] and 180 are calculated values from Bubukina et al. [37] given with incomplete vibration-rotation assignment. The line positions of 210 transitions given with W2020 source do not coincide with those from the W2020 line list [20].

Unlike the first two species, only twelve $H_2{}^{17}O$ transitions are marked as taken from the W2020 line list [20]. But the positions of all these twelve transitions do not coincide with the original W2020 values [20]. The position differences reach 0.03 cm^{-1}. All the 442 transitions from Lodi & Tennyson [40] are provided with calculated positions and incomplete vibration-rotation assignment.

All HITRAN line positions of $HD^{16}O$ are due to Kyuberis et al. [41]. 1273 of them are empirical values obtained from updated IUPAC energy levels [31] and 53 positions are *ab initio* values.

All positions and intensities of $HD^{17}O$, $HD^{18}O$ and $D_2{}^{16}O$ are taking from Kyuberis et al. [41]. According to HITRAN reference code, the line positions of $HD^{17}O$ and $HD^{18}O$ are empirical values obtained from updated IUPAC energy levels [31] and *ab initio* values for $D_2{}^{16}O$.

As mentioned above, an overall satisfactory agreement is observed between our line positions and HITRAN2020 values. The average ($\nu_{TW} - \nu_{HITRAN2020}$) deviation for 5238 line positions is -2.05×10^{-3} cm^{-1}. Excluding lines marked as "noisy", "blended", "overlapped" in our experimental list and

excluding differences larger than 0.05 cm^{-1}, it leads to an average deviation of -1.06×10^{-3} cm^{-1} for 5141 line positions

The graphical overviews of the line position differences are presented on **Fig. 6** (left panels) for $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$. The most noticeable global observation is the clear increase of outliers for $H_2{}^{17}O$ above 8340 cm^{-1} corresponding to the change of the main source for HITRAN position from Ref. [38] to Ref. [40], below and above 8340 cm^{-1}, respectively. In the following, we discuss some of the most significant positions differences ($\left| \nu_{TW} - \nu_{HITRAN2020} \right| \geq 0.015$ cm^{-1}) between the two line lists.

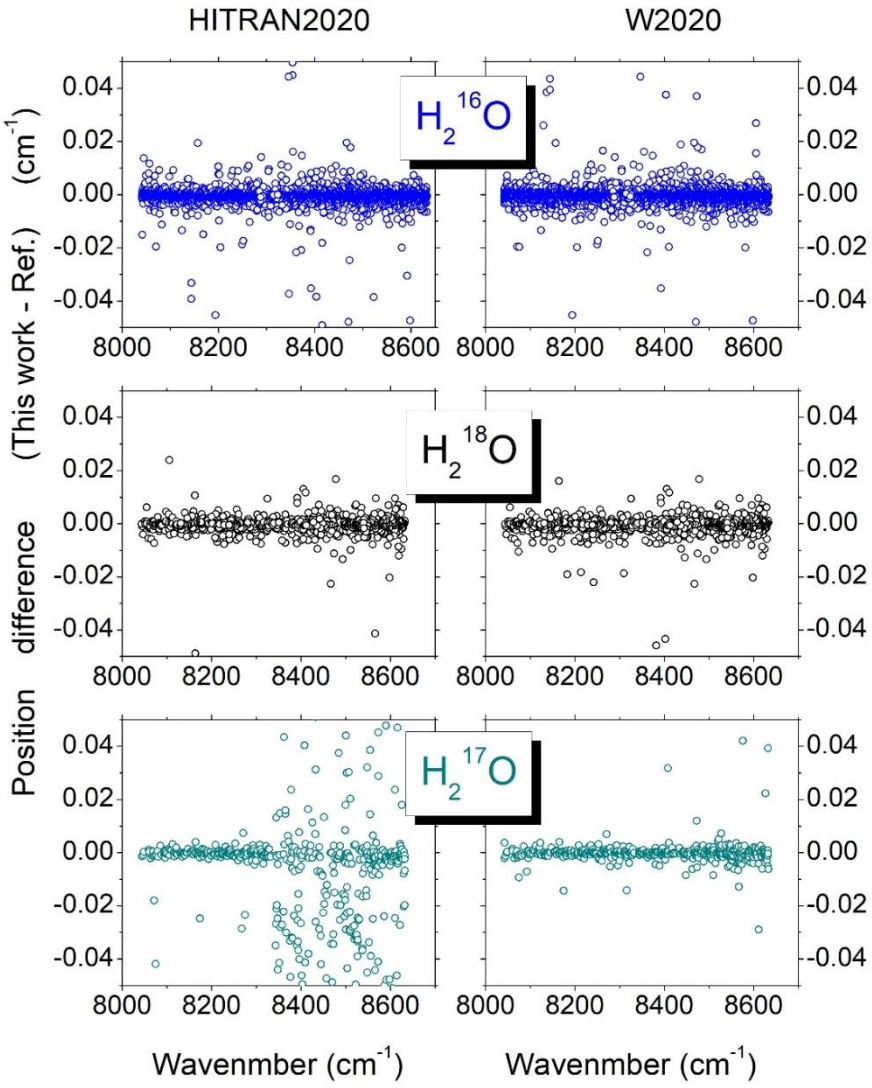

Fig.6.
Comparison of the $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$ line positions retrieved in this work (TW) between 8041.45 and 8633.41 cm^{-1} and the HITRAN2020 [9] and W2020 values [20] (left and right panels, respectively).

As mentioned above, the largest deviation (0.278 cm^{-1}) is observed for the $H_2{}^{16}O$ $2\nu_3$ $17_{5\,12} - 16_{4\,13}$ (8065.53458 cm^{-1}, S= 2.491×10^{-29} cm/molecule instead of 8065.25653 cm^{-1}, S= 2.913×10^{-29} cm/molecule in HITRAN). It is due to an incorrect assignment of the upper level (030 $17_{13\,4}$) and an incorrect value of the line position. The HITRAN reference to the position of this line is a "private communication (2008)". The W2020 line list gives a very close line position (8065.25729 cm^{-1}) and

the same incorrect assignment ($3\nu_2$ $17_{13\,4} - 16_{4\,13}$). The W2020 energy value corresponding to our vibration-rotation assignment (002 $17_{5\,12}$ instead of 030 $17_{13\,4}$) is 11504.775(10) cm^{-1}, about 0.066 cm^{-1} below our value (11504.841(3) cm^{-1}) but the $H_2{}^{16}O$ $2\nu_3$ $17_{5\,12} - 16_{4\,13}$ transition is absent in the W2020 list.

The fifteen HITRAN2020 positions showing position deviations larger than 0.1 cm^{-1} compared to our measurements correspond to $H_2{}^{16}O$ lines of high bending bands ($3\nu_2+\nu_3$, $4\nu_2+\nu_3$, $5\nu_2$, $6\nu_2$, and $7\nu_2$) and three transitions of the $2\nu_3$, $\nu_2+2\nu_3$, and $2\nu_1+\nu_2$ bands. Eleven of these deviations are due to incorrect values of the upper energies in W2020 list.

The line position of the $H_2{}^{16}O$ $2\nu_3$ $14_{8\,7} - 13_{5\,8}$ transition is given at 8134.50567 cm^{-1} – (ν_{TW} - $\nu_{HITRAN2020}$) = -0.1275 cm^{-1} – with position reference to Mikhailenko et al. [38]. This is not correct: in our line list covering the 5850 – 8340 cm^{-1} region [38], this line position (calculated from empirical energies) was given at 8134.3788 cm^{-1}. This value is close to the CRDS line position measured in Ref. [6] (8134.3807 cm^{-1}) which is itself in excellent agreement with the 8134.37816 cm^{-1} position determined in the present study.

The $H_2{}^{16}O$ $2\nu_3$ $17_{4\,14} - 16_{1\,15}$ transition with measured line position at 8173.7364 cm^{-1} – (ν_{TW} - $\nu_{HITRAN2020}$) = -0.14 cm^{-1} – is given without complete VR assignment and with variational position in HITRAN2020 (8173.87670 cm^{-1}, Ref. [37]) and in W2020 (8173.44656 cm^{-1}, Ref. [42]). This transition which was assigned by Campargue et al. [6] to the line at 8173.7328 cm^{-1} but was excluded from the W2020 transition database probably due to its superposition with an ammonia line.

The $H_2{}^{16}O$ $2\nu_3$ $16_{7\,10} - 15_{4\,11}$ transition – (ν_{TW} - $\nu_{HITRAN2020}$) = -0.068 cm^{-1} – is measured in the present work at 8077.75668 cm^{-1} instead of 8077.82494 cm^{-1} in the HITRAN/W2020 lists. This last value relies on the W2020 upper term value [20] derived from four emission positions reported by Zobov et al. [14,43] and Rutkowski et al. [44], three of them having multiple VR assignments.

The measured position of the $6\nu_2$ $8_{2\,7} - 9_{5\,4}$ transition at 8569.95452 cm^{-1} deviates from the HITRAN/W2020 value by (ν_{TW} - $\nu_{HITRAN2020}$) = -0.056 cm^{-1}. HITRAN/W2020 line position was calculated from W2020 upper energy level derived in Ref. [23] from four emission line positions given by Coheur et al. [45] and Zobov et al. [14,43], two of them having multiple VR assignments.

The (ν_{TW} - $\nu_{HITRAN2020}$) differences are -0.052 $cm^{-1}$and -0.047 cm^{-1} for the $3\nu_2+\nu_3$ $12_{5\,8} - 12_{5\,7}$ and $3\nu_2+\nu_3$ $12_{5\,8} - 13_{3\,11}$ transition, respectively. These two transitions at 8545.77779 and 8598.40303 cm^{-1}, respectively, reach the same 031 $12_{5\,8}$ upper level. Our experimental uncertainties on the line positions of these weak lines are about 0.003 cm^{-1}. The W2020 energy of the level (10846.51486 cm^{-1}) was determined from nine transitions assigned in emission spectra [14,43,44,46] and one CRDS transition [6]. Eight of the emission lines have multiple VR assignments. Interestingly, three transitions observed by absorption by Régalia et al. [8] were weighted with a very large uncertainty between 0.075 and 0.1 cm^{-1} by the xMARVEL procedure which is equivalent to discard these data. The erroneous assignment of the $3\nu_2+\nu_3$ $12_{5\,8} - 13_{5\,9}$ transition to the line at 8259.9848 cm^{-1} reported

in Ref. [6] relied on the inaccurate IUPAC upper energy of the 031 $12_{5\,8}$ (10846.513718 cm^{-1}) [36]. This IUPAC energy value relied itself on emission line positions of Refs. [14,43]. Taking into account the W2020 uncertainty of different positions, only two of them (8259.9848 cm^{-1} [6] and 6887.261822 cm^{-1} of the $3\nu_2+\nu_3-\nu_2$ $12_{5\,8} - 12_{5\,7}$ [36]) were really used for the determination of the W2020 term value. In summary, only emission data was used for the both IUPAC [36] and W2020 [20] energy values. This is a typical example of biases of the xMARVEL procedure used to derive the W2020 energy levels [20]: the W2020 energy value of the considered 031 $12_{5\,8}$ upper level relies exclusively on emission data while three reliable transitions observed by absorption by Régalia et al. [8] were not taken into account for the energy determination. Our line positions of the $3\nu_2+\nu_3$ $12_{5\,8} - 12_{5\,7}$ and $3\nu_2+\nu_3$ $12_{5\,8} - 13_{3\,11}$ transitions are very consistent with the $3\nu_2+\nu_3$ $12_{5\,8} - 11_{5\,7}$ FTS position reported at 8860.6800 cm^{-1} [8] and in satisfactory agreement with two additional transitions reported by Régalia et al. [8]: our term value is 10846.4652 cm^{-1} to be compared to the FTS values of 10846.4649 and 10846.4785 cm^{-1}, respectively.

For the $H_2{}^{17}O$ isotopologue, the (ν_{TW} - $\nu_{HITRAN2020}$) differences range from -0.0764 to +0.0897 cm^{-1}. The absolute values of the differences exceed 0.005 cm^{-1} for 181 transitions. The main part of these large differences is due to variational positions from Lodi & Tennyson [40], according to HITRAN reference code. Sixteen of the line positions with large (ν_{TW} - $\nu_{HITRAN2020}$) differences rely on an (unpublished) update of the IUPAC database [47]. In addition, four transitions ($3\nu_2+\nu_3$ $5_{2\,4} - 5_{2\,3}$, $\nu_1+3\nu_2$ $3_{2\,2} - 3_{1\,3}$, $\nu_1+3\nu_2$ $4_{3\,1} - 4_{2\,2}$, and $\nu_1+3\nu_2$ $4_{2\,3} - 3_{1\,2}$) are referenced to W2020 line list while their HITRAN's positions differ from their W2020 counterpart by a value up to 0.017 cm^{-1}.

The same situation of incorrect sourcing is found for eleven $H_2{}^{18}O$ transitions. All of them (according to HITRAN reference code) are coming from our 5850 – 8340 cm^{-1} line list [38]. But in fact, none of these line positions coincide with the values of Ref. [38]. The deviations reach a value of 0.05 443 cm^{-1} and the position reported in Ref. [38] are confirmed in the present study.

Table 3.
Term values of the $HD^{16}O$ isotopologue corrected compared to published values [31,48].

$V_1V_2V_3$	J	K_a	K_c	Energy	dE	NT	E [31]	unc	ΔE	Ratio	E [48]	$d2$
012	9	8	2	10224.16027	337		10224.06221	500	9806	19.61		
012	9	8	1	10224.16030	337		10224.06238	500	9792	19.58		
012	11	6	6	10162.17475	208		10162.16664	500	811	1.62		
012	11	6	5	10162.25323	201		10162.24747	500	576	1.15		
012	14	3	11	10453.35835	171		10453.33780	500	2055	4.11		
012	15	0	15	10191.15764	269		10191.11790	514	3974	7.73		
022	6	3	3	10438.87664	427		10438.88316	570	-652	1.14		
111	8	7	2	9096.20106	434		9096.21122	500	-1016	2.03		
111	8	7	1	9096.20117	434		9096.21111	500	-994	1.99		
220	7	1	6	8566.01312	208	2	8484.16639	500	-15327	30.65	8566.0068	632
220	7	3	4	8697.53845	240						8697.5323	615
220	9	0	9	8702.17405	239		8702.17946	500	-541	1.08	8702.1731	95
300	11	7	4	9599.24330	354		9599.23597	500	733	1.47		
300	13	3	10	9447.67038	116		9447.66474	500	564	1.13		

Notes

$V_1V_2V_3$ – vibration quantum numbers; $J\ K_a\ K_c$ – rotation quantum numbers; Energy/cm^{-1} – empirical term values; *dE* – estimated uncertainty on the term value in 10^{-5} cm^{-1} units; *NT* – number of line positions used for the energy determination if it is larger than 1; *E* [31] and *E* [48] – empirical term values published in corresponding references in cm^{-1}; unc – IUPAC term value uncertainties in 10^{-5} cm^{-1} units; ΔE – term value differences between empirical values of this work and those of Ref. [31] in 10^{-5} cm^{-1} units; Ratio = ΔE / unc; *d2* – term value differences between empirical values of this work and those of Ref. [48] in 10^{-5} cm^{-1} units.

In **Table 3**, we list 14 corrected term values of the $HD^{16}O$ molecule corresponding to $|\nu_{TW} - \nu_{HITRAN2020}|$ deviations larger than 0.005 cm^{-1}. According to HITRAN reference code, HITRAN positions rely on an update of the IUPAC-TG levels [31] reported by Kyuberis et al. [41]. The table compares our energy values to the IUPAC-TG values [31] and to the values given by Liu et al. [48].

4.5. Comparison to the W2020 line lists

Following the approach developed a decade ago by a task group (TG) of the International Union of Pure and Applied Chemistry (IUPAC) [31,36,47], the xMARVEL procedure and code were applied to an exhaustive catalog of absorption and emission measured line positions collected in the literature, in order to derive accurate sets of empirical energy levels of $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$. For the main isotopologue, the collected W2020-$H_2{}^{16}O$ transition dataset gathers 286,987 non-redundant rovibrational transitions, and 19,225 empirical energy levels were determined [20]. Most of the line positions of the W2020 lists (tagged with "M") were obtained by difference of empirical energy levels and released with their self-consistent uncertainties [20]. In absence of empirically determined energy levels, less accurate calculated values were used (tag "C").

All 4687 transitions of the $H_2{}^xO$ (x=16,17,18) isotopologues assigned in our spectrum are present in the W2020 line list [20]. Note that the complete VR assignment of 71 transitions missing in the W2020 list is included in the list provided as Supplementary Material. These transitions correspond to the 60 $H_2{}^xO$ newly determined upper energy levels listed in **Table 2**. The $H_2{}^{16}O$ 002 $17_{\ 4\ 14}$ and $H_2{}^{18}O$ 031 $11_{\ 1\ 10}$ levels included in the table are "new" compared to the W2020 energy lists but they were previously reported in Ref. [6] from the $H_2{}^{16}O$ 002 $17_{\ 4\ 14}$ – 000 $16_{\ 1\ 15}$ and $H_2{}^{18}O$ 031 $11_{\ 1\ 10}$ – 000 $12_{\ 1\ 11}$ transitions at 8173.7328 and 8130.8295 cm^{-1}, respectively. These positions have not been used for W2020 input file probably due to the overlapping with ammonia lines mentioned in Ref. [6]. In the present work, the positions of these very weak lines ($S \sim 4\times10^{-29}$ cm/molecule) is found at 8173.73640 and 8130.82895 cm^{-1}, respectively, in agreement with the values of Ref. [6].

As mentioned above, the general agreement with the W2020 line position is very good (see **Fig. 4**) with average deviations between -1.7×10^{-4} cm^{-1} and -3.4×10^{-4} cm^{-1}, depending on the set of outliers excluded from the comparison.

The overviews of the ($\nu_{TW} - \nu_{W2020}$) position differences are included in **Fig. 6** for $H_2{}^{16}O$, $H_2{}^{18}O$ and $H_2{}^{17}O$ (right panels). Let us mention that 27 transitions have their W2020 positions showing a large deviation compared to the present measurements (between 0.0052 and 0.0260 cm^{-1}) while HITRAN value agree with experiment within 0.004 cm^{-1}. Three of such examples are illustrated on the upper panels of **Fig. 7**. These lines are tagged as "bad position in W2020" in the experimental list

provided as supplementary material. All but two of these (good) HITRAN positions originate from our 5850 – 8340 cm^{-1} list [38]. The most inaccurate W2020 position from this group of transitions, (ν_{TW}-ν_{W2020}) = 0.026 cm^{-1}, is that of the $H_2{}^{16}O$ $5\nu_2$ $14_{4\,11} - 15_{1\,14}$ transition (see Fig. 7 (b)). The W2020 position is at 8130.01045 cm^{-1} while the present measurement gives a value at 8130.03640 cm^{-1}, in very good agreement with our previous determination at 8130.0371 cm^{-1} [6] and with HITRAN value (8130.0391 cm^{-1}). Interestingly, the W2020 energy derivation of the 050 $14_{4\,11}$ upper level relies on the $5\nu_2$ $14_{4\,11} - 15_{1\,14}$ position of Ref. [6] and on the $5\nu_2$-$4\nu_2$ $14_{4\,11} - 15_{5\,10}$ emission transition reported with an inaccurate value at 928.52266 cm^{-1} [45]. For an unknown reason, the xMARVEL procedure attached the same uncertainty of 0.02665 cm^{-1} to the absorption and emission line positions [20]. This is a typical example of a good absorption data "spoiled" by inaccurate emission data leading to an inaccurate W2020 energy level.

Overall, 191 of our positions deviate from the W2020 positions by more than 0.005 cm^{-1}. Our recommended values of the corresponding upper energy levels are listed in **Table 4** together with W2020 values. This table includes the ratio $R = \frac{|E_{TW} - E_{W2020}|}{Unc_{W2020}}$ which compares the absolute deviation of the W2020 energy level from our value to the claimed W2020 uncertainty (Unc_{W2020}). Among the 127 corrected levels, 48 have a term value which deviates from our observations by more than ten times their W2020 uncertainty (*R* larger than 10, up to 375). Six examples of inaccurate W2020 line positions with deviations largely exceeding the W2020 error bars are displayed in **Fig. 7** (the very small W2020 error bars plotted on the figure are hardly visible at the scale of the graphs). The assignment of the problematic lines is given in the caption of the figure.

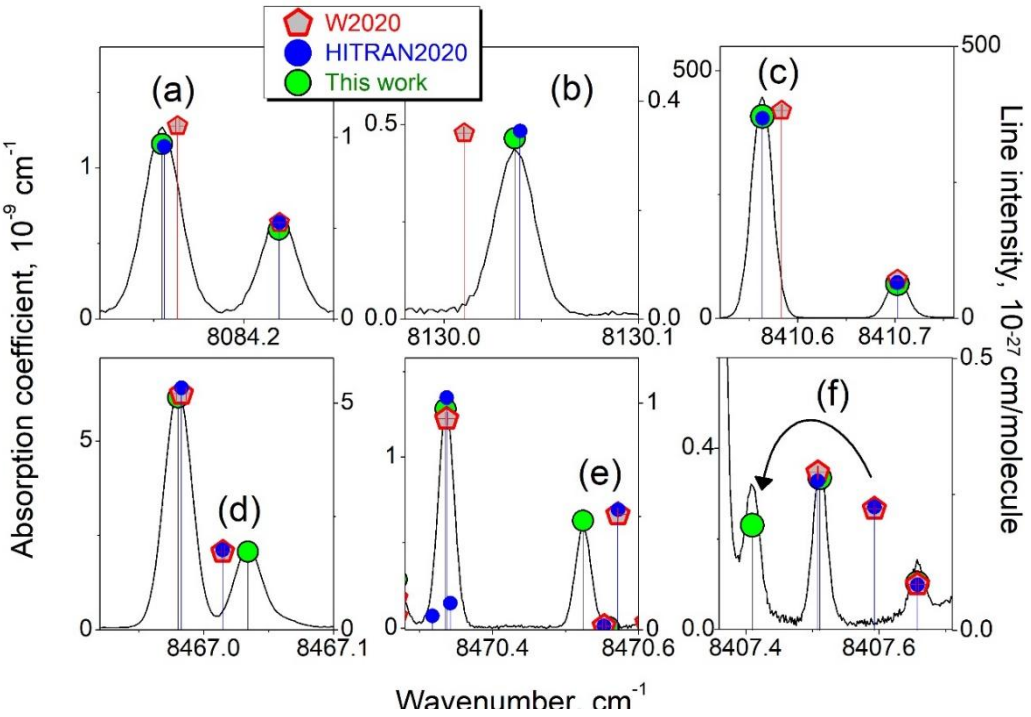

Fig. 7.
Examples of comparison of the CRDS spectrum of water vapor and corresponding line list (green circles) to the W2020 line list of $H_2{}^{16}O$ [20] (grey pentagons) and the HITRAN2020 list of natural water [19] (blue circles).

The W2020 line positions are based on empirically determined energy levels. The (very small) W2020 error bars displayed on the different panels are considerably smaller than the observed deviations.
The right-hand intensity scale is adjusted to correspond approximately to the peak heights.
In the three examples displayed on the upper panels [(a) $4\nu_2+\nu_3-\nu_2\ 2_{0\,2} - 3_{2\,1}$, (b) $5\nu_2\ 14_{4\,11} - 15_{1\,14}$, (c) $3\nu_2+\nu_3\ 6_{5\,2} - 7_{5\,3}$], the W2020 positions deviate from the observation while the HITRAN positions show a good agreement. On the three lower panels [(d) $\nu_1+\nu_2+\nu_3\ 14_{1\,14} - 14_{1\,13}$, (e) $5\nu_2\ 11_{6\,5} - 11_{5\,6}$, (f) $6\nu_2\ 7_{1\,7} - 8_{4\,4}$], the W2020 and HITRAN2020 positions coincide and deviate both from the recorded spectrum.

The largest discrepancy (about 0.38 cm^{-1}) concerns the $15_{0\,15} - 16_{1\,16}$ and $15_{1\,15} - 16_{0\,16}$. doublet of the $2\nu_1+\nu_2$ band of $H_2{}^{16}O$ that we assign to the line at 8354.08116 cm^{-1} in accordance with the *J*-dependence of the $(\nu^{OBS} - \nu^{SP})$ deviations of the experimental (ν^{OBS}) and calculated (ν^{SP}) positions for the $2\nu_1+\nu_2\ J_{0\,J} - J+1_{1\,J+1}$ series of transitions. The 210 $15_{0\,15}$ and $15_{1\,15}$ term values of W2020 were derived from the $2\nu_1+\nu_2\ 15_{0\,15} - 14_{1\,14}$ absorption line reported at 8941.888 cm^{-1} given by Tolchenov & Tennyson [24] which is believed to be erroneously assigned. In the HITRAN2020 line list, this doublet is assigned the $\nu_1+\nu_2\ 15_{15\,1} - 16_{0\,16}$ transition and to a $15_{15\,0} - 16_{1\,16}$ transition of an unidentified band with W2020 position (8354.03627 cm^{-1}) and variational position (8354.03153 cm^{-1} [37]), respectively. Note that this last assignment is incorrect as the $15_{15\,0} - 16_{1\,16}$ rotational assignment is forbidden for both *A* and *B* type bands.

We have examined in details the origin of the inaccuracy of the W2020 energy value for a sample of $H_2{}^{16}O$ levels of **Table 4**. Obviously, these situations result from the existence of an inaccurate line position in the W2020 transition dataset which impacts the resulting W2020 energy level value but the amplitude of the impact depends on the W2020 weighting of the experimental line positions. (Note that among the 127 corrected levels of $H_2{}^{16}O$, only 36 involve CRDS line positions, a small fraction of them being inaccurate or wrong as discussed in section 4.3). The point is that in a number of cases where conflicting experimental values exist, the xMARVEL procedure fails in discriminating the good data from the less accurate data. The difficulty of the process is due to the lack of reliable line-by-line error bars in most of the experimental sources. In some problematic situations, for an unknown reason, the MARVEL procedure attached unrealistically small uncertainty value to some inaccurate experimental values which have then a considerable weight on the resulting energy value (see the example of the 031 $12_{5\,8}$ discussed in section 4.4). Another typical situation is when a uniform error bar is attached to all the position values of a given source, while obviously the position of weak blended lines is much less accurate than that of "an isolated line of intermediate intensity". Again, this is due the fact that the error bars needed by the xMARVEL procedure are not available and can hardly be guessed. In the number of problematic cases examined in details, we found examples where our present measurements coincide with some previous results which were *de facto* excluded from the energy determination due to the W2020 decision to attach them excessive error bars.

Table 4

Term values of $H_2{}^xO$ (x= 16,17,18) water isotopologues corrected by more than 0.005 cm^{-1} compared to the W2020 empirical energy levels [20].

$V_1V_2V_3$	J	K_a	K_c	Energy (cm^{-1})	dE (10^{-5} cm^{-1})	NT	E_{W2020} [20] (cm^{-1})	Unc_{W2020} (10^{-5} cm^{-1})	ΔE (10^{-5} cm^{-1})	Ratio
$H_2{}^{16}O$										
002	16	7	10	11322.35736	297		11322.42562	1710	-6826	3.99
002	17	5	12	11504.84120	268		11504.77535	1080	6585	6.10
002	18	4	15	11503.51499	107		11503.38094	389	13405	34.50
012	8	8	1	10786.81160	12		10786.81774	13	-614	48.58
012	8	8	0	10786.81160	12		10786.81774	13	-614	48.46
012	10	1	10	10074.46086	81		10074.46764	51	-678	13.42
012	10	7	4	11043.78788	128		11043.78215	50	573	11.40
012	13	2	12	10988.99744	746		10989.01615	105	-1871	17.87
012	13	4	9	11518.89796	163		11518.90828	793	-1032	1.30
012	13	5	9	11566.01465	47		11565.78065	1333	23400	17.55
012	13	8	5	12094.78477	648		12094.73357	632	5120	8.11
012	14	4	11	11701.36744	228		11701.36188	609	556	0.91
012	15	1	14	11558.06940	501		11558.05387	1460	1553	1.06
012	15	2	14	11557.73092	103		11557.70404	1745	2688	1.54
022	7	3	4	11386.13814	338		11386.14398	228	-584	2.56
031	6	5	2	9470.20970	19	3	9470.22957	60	-1987	32.99
031	9	7	2	10519.22737	178	2	10519.22222	36	515	14.26
031	10	6	5	10517.31559	11		10517.30864	50	695	13.83
031	11	4	8	10356.52419	18	4	10356.53216	51	-797	15.67
031	12	4	9	10638.92985	66	4	10638.93901	51	-916	18.13
031	12	5	8	10846.46518	251	2	10846.51486	134	-4968	37.05
031	13	6	7	11383.69381	42		11383.80423	601	-11042	18.38
031	14	2	13	10738.00571	52		10737.99657	922	914	0.99
031	14	4	10	11373.27104	107		11373.27805	536	-701	1.31
031	14	6	8	11719.39308	364		11719.37540	601	1768	2.94
031	15	2	13	11335.35830	292		11335.37176	2461	-1346	0.55
031	15	6	10	12063.93046	570		12063.92157	88	889	10.07
031	16	1	15	11356.00250	126	2	11356.00882	414	-632	1.53
031	17	0	17	11261.86083	41		11261.86776	347	-693	2.00
041	6	3	3	10633.97536	11	5	10633.98201	51	-665	12.97
041	9	4	5	11413.20972	179		11413.19366	276	1606	5.81
041	10	6	5	12137.16876	494	2	12137.17758	879	-882	1.00
041	11	2	9	11640.91011	143		11640.92096	1388	-1085	0.78
041	11	3	8	11775.41233	219		11775.28222	1144	13011	11.37
041	12	4	8	12226.36835	464		12226.37946	995	-1111	1.12
041	14	0	14	11850.96054	368		11850.95208	341	846	2.48
050	5	5	1	8906.91550	32	5	8906.92237	100	-687	6.85
050	8	4	5	9117.28554	9	4	9117.27813	25	741	29.35
050	9	6	4	9961.82189	14	2	9961.83034	51	-845	16.50
050	10	5	5	9869.74092	45		9869.74602	407	-510	1.25
050	11	4	8	9822.62023	28	4	9822.61334	51	689	13.38
050	11	5	6	10135.02266	35	3	10135.01166	51	1100	21.61
050	11	6	5	10469.51933	18	2	10469.56722	226	-4789	21.20
050	13	3	10	10230.53718	152		10230.72899	323	-19181	59.38
050	14	4	11	10761.30524	25		10761.27930	5331	2594	0.49
060	7	1	7	9539.18577	69		9539.36866	179	-18289	102.12
060	7	1	6	9714.49067	86		9714.75867	183	-26800	146.77
060	7	5	2	10837.70631	563		10837.71662	696	-1031	1.48
060	8	2	7	10047.25190	164		10047.30825	1068	-5635	5.28
060	9	0	9	9837.80195	23	2	9837.81511	1800	-1316	0.73
060	9	5	4	11250.37347	547		11250.39088	2001	-1741	0.87
060	9	6	3	11613.17589	174		11613.08115	408	9474	23.22

060	10	0	10	10039.61145	140	2	10039.60106	1002	1039	1.04
060	10	4	7	11143.39648	630		11143.41604	572	-1956	3.42
060	13	0	13	10756.15497	56		10756.17475	3654	-1978	0.54
070	3	0	3	10223.87421	17		10223.86799	300	622	2.07
070	3	1	3	10435.52253	318	2	10435.52787	184	-534	2.90
070	5	0	5	10428.00278	27	2	10428.15128	506	-14850	29.34
070	7	0	7	10718.88596	148		10718.89141	530	-545	1.03
070	10	1	10	11425.02327	424		11425.03050	300	-723	2.41
070	11	0	11	11573.11404	553		11573.24735	1323	-13331	10.08
111	11	9	2	11526.52882	80		11526.54096	1783	-1214	0.68
111	12	8	5	11560.51311	29		11560.51929	276	-618	2.24
111	12	9	4	11813.62903	70		11813.63488	1172	-585	0.50
111	13	8	6	11868.38386	37		11868.37523	1369	863	0.63
111	13	9	5	12122.69489	166		12122.68722	1091	767	0.70
111	14	1	14	10794.91759	39	2	10794.89846	128	1913	14.98
111	14	6	9	11861.29848	72		11861.22782	1142	7066	6.19
111	14	8	7	12198.09808	223		12198.08346	1013	1462	1.44
111	15	3	12	11829.52983	37		11829.56510	1600	-3527	2.20
111	15	4	11	12006.20806	81		12006.21315	604	-509	0.84
111	16	3	13	12177.70070	74		12177.70741	1013	-671	0.66
111	16	4	13	12173.85277	165		12173.87453	1877	-2176	1.16
111	17	2	16	11995.91261	119		11995.91860	729	-599	0.82
111	17	2	15	12286.98353	291		12286.93924	603	4429	7.35
121	14	1	14	12305.58544	271		12305.59729	1436	-1185	0.82
130	8	4	4	9557.47160	10	5	9557.47698	60	-538	8.92
130	9	7	2	10481.20100	25	2	10481.19377	54	723	13.42
130	10	6	5	10458.18408	65	3	10458.18948	52	-540	10.31
130	11	6	5	10721.85103	129		10721.85762	467	-659	1.41
130	14	5	10	11401.98039	161		11401.96108	103	1931	18.71
140	1	1	0	9791.61828	7		9791.60323	50	1505	29.95
140	10	3	7	11389.66218	331		11389.65099	300	1118	3.73
140	12	3	10	11841.26194	320		11841.27165	555	-971	1.75
210	7	5	3	9840.83208	14	2	9840.83768	60	-560	9.30
210	10	9	2	11277.81268	291		11277.80304	1003	964	0.96
210	10	9	1	11277.81268	291		11277.80307	1006	961	0.96
210	11	1	11	10031.64566	5	2	10031.65114	51	-548	10.83
210	12	4	9	10848.17535	13	3	10848.18042	52	-507	9.66
210	13	3	10	11140.37874	66		11140.53272	1701	-15398	9.05
210	14	1	14	10743.70487	94		10743.69587	611	900	1.47
210	15	0	15	11015.02598	56		11015.40361	101	-37763	375.38
210	15	1	15	11015.02716	100		11015.41085	1006	-38369	38.14
210	15	4	11	11944.06956	381		11944.06072	1003	884	0.88
$H_2{}^{17}O$										
111	7	6	1	10016.18577	67		10016.21475	10	-2898	287.54
111	11	0	11	10064.14958	462		10064.15817	3351	-859	0.26
111	13	2	12	10771.07055	213		10771.07912	504	-857	1.70
210	5	1	5	9064.74820	48	2	9064.75439	387	-619	1.60
$H_2{}^{18}O$										
012	7	1	7	9532.06362	129		9532.06939	163	-577	3.55
031	7	4	4	9403.31670	57	2	9403.32273	50	-603	12.05
031	8	5	4	9791.60152	170		9791.59215	1296	937	0.72
031	8	6	3	10010.64952	48		10010.63985	502	967	1.93
031	9	4	5	9824.45326	141	2	9824.46046	51	-720	14.01
031	9	6	3	10228.58454	125		10228.59167	500	-713	1.43
031	10	4	7	10054.74314	185		10054.73716	501	598	1.19
031	10	7	3	10710.96180	229		10710.94880	502	1300	2.59
031	12	2	11	10147.65947	235		10147.65335	56	612	10.89
050	7	5	2	9182.10769	69	4	9182.11391	10	-622	60.06
111	6	5	2	9670.59957	80		9670.61163	11	-1206	110.04
111	7	6	1	9997.65723	23		9997.67753	638	-2030	3.18

111	7	7	0	10187.36809	25		10187.36267	176	542	3.09
111	8	8	0	10588.71087	70		10588.71610	595	-523	0.88
111	9	4	6	10103.21716	26	2	10103.22282	10	-566	54.65
111	9	4	5	10123.49012	166		10123.49996	101	-984	9.71
111	13	1	13	10505.27762	61		10505.26993	1426	769	0.54
111	13	3	11	10972.77924	183		10972.78524	501	-600	1.20
111	14	1	14	10759.33413	325		10759.32248	505	1165	2.31
130	0	0	0	8249.03287	84		8249.03837	50	-550	10.98
130	5	5	1	9219.57938	199		9219.60199	51	-2261	44.52
130	9	2	7	9499.30103	58	3	9499.30665	838	-562	0.67
130	9	3	6	9608.77167	101		9608.70833	100	6334	63.34
210	4	1	3	9010.25415	22		9010.24725	10	690	67.34
210	5	4	2	9358.20451	161		9358.18782	3749	1669	0.45
210	5	5	0	9498.99357	81		9498.98782	472	575	1.22
210	7	1	7	9301.25496	21	2	9301.26294	50	-798	15.81
210	8	1	7	9599.67536	88		9599.66811	144	725	5.03
210	8	2	6	9704.95432	478		9704.96775	500	-1343	2.69

Notes

$V_1V_2V_3$ – vibration quantum numbers; $J\ K_a\ K_c$ – rotation quantum numbers; Energy/cm^{-1} – empirical term values; *dE* – term value uncertainties in 10^{-5} cm^{-1} units; *NT* – number of line positions used for the energy determination if it is larger than 1; E_{W2020} – W2020 empirical term values [20] in cm^{-1}; Unc – W2020 term value uncertainties in 10^{-5} cm^{-1} units; ΔE – term value differences between empirical values of this work and those of Ref. [20] in 10^{-5} cm^{-1} units; Ratio = ΔE/Unc.

5. Line intensities - Comparison with literature

5.1. Overview comparison

The S_{Ref} / S_{TW} line intensity ratios comparing our measured values to the FTS values of Ref. [8], to the CRDS values of Ref. [6] and to the HITRAN2020 [9] and W2020 [20] intensities are displayed on **Fig. 8**. This plot is limited to the main isotopologue, $H_2{}^{16}O$. Overall, the agreement appears to be reasonable, although significantly better between the experimental datasets. Let us recall that all the W2020 intensities are calculated values from the POKAZATEL list [42]. In our region of interest, most of the HITRAN line intensities are calculated values from Conway et al. [49] (6036 of 7141) and experimental values of Campargue et al. below 8340 cm^{-1} [6] (927 entries). 121 line intensities are from Refs. [8,19,50]. The W2020 and Conway et al. intensities were computed from different variants of the semi-empirical potential energy and *ab initio* dipole moment surfaces of the water molecules. Differences between the calculated intensities of the HITRAN2020 and W2020 lists reflect the sensitivity of the calculations to small changes in the used surfaces in the considered region. According to **Fig. 8,** the W2020 and HITRAN2020 intensity values are validated by experiment within 10-15% for most of the lines although a number of outliers are observed (see below). We also note that, contrary to HITRAN values, the POKAZATEL intensities of the W2020 list present a systematic overestimation by about 10% in the considered region.

In order to examine the situation in more details, the intensity ratios of the $\nu_1+3\nu_2$, $3\nu_2+\nu_3$ and $2\nu_1+\nu_2$ bands of the second hexade have been separated in different panels in **Fig. 9**. The differences between the general appearance of the upper and lower panels (corresponding to W2020 and HITRAN2020, respectively) reflect the differences between the POKAZATEL intensities and those of

Conway et al. [49]. The 10% systematic overestimation of the W2020 intensities is clearly apparent for the $\nu_1+3\nu_2$ and $3\nu_2+\nu_3$ bands while HITRAN average values mostly coincide with experiment, although with a larger dispersion than W2020 dataset in the case of the first band. Interestingly, a larger dispersion of the intensity ratios is observed for the $2\nu_1+\nu_2$ band both for HITRAN2020 and W2020. This situation is unusual as, in general, calculated intensities are known to be less accurate in the case of bands involving a high vibrational excitation of the bending mode V_2 (see discussion and Figs. 7 and 8 in Ref. [19]).

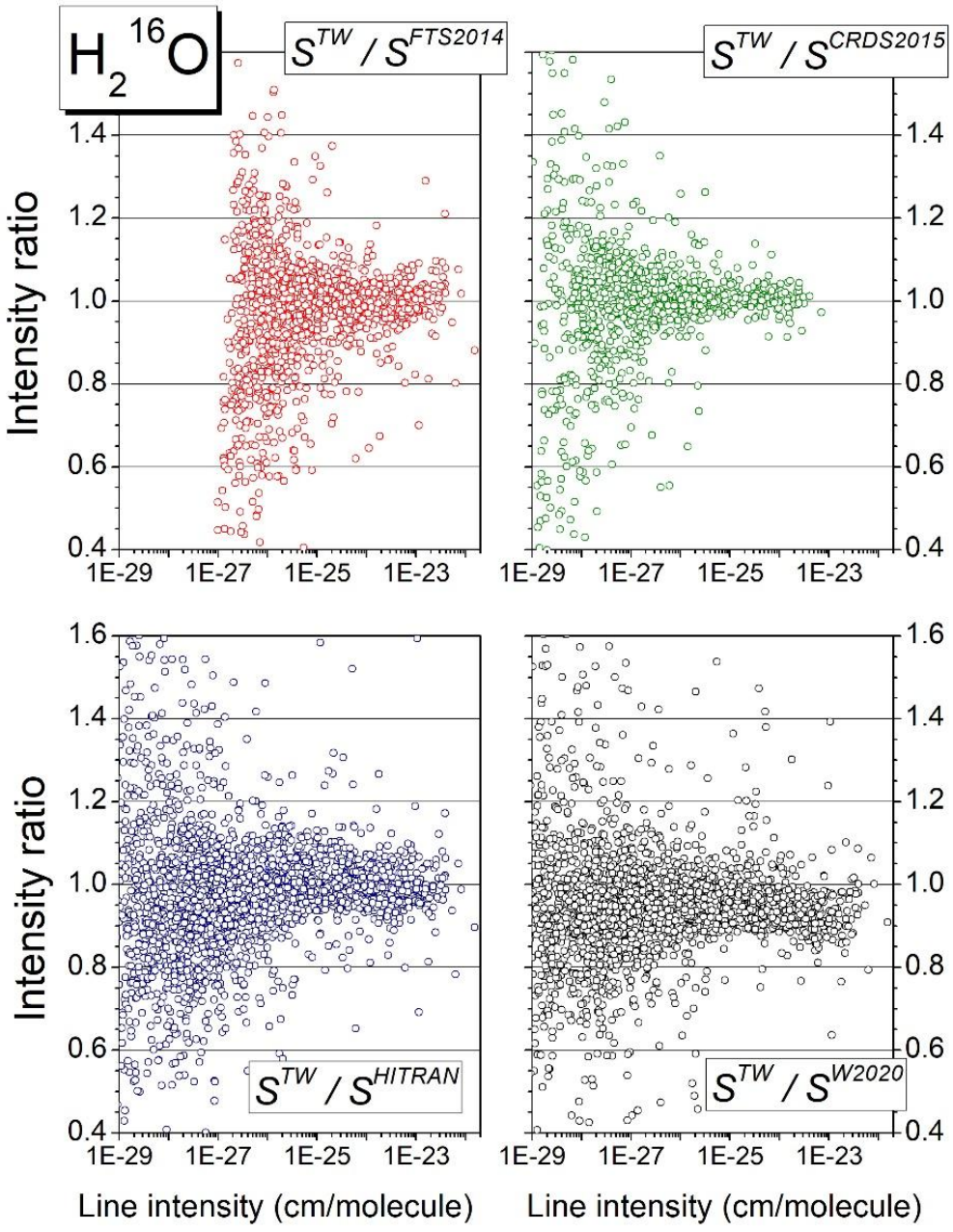

Fig. 8.
Ratios of the CRDS intensity values to the FTS values of Ref. [8] (FTS2014, red circles), to the CRDS values of Ref. [6] (CRDS2015, green circles), and to the HITRAN2020 and W2020 intensities (violet and black circles, respectively). The plots are limited to the transitions of the main isotopologue, $H_2{}^{16}O$, measured in the present work in the 8041 – 8633 cm^{-1} region.

Six spectral intervals showing disagreement between the experimental spectrum and the HITRAN2020 and W2020 line intensities are displayed in **Fig. 10** (the rovibrational assignment of the problematic lines is given in the caption of the figure). In the first panels (a)-(d), HITRAN intensity values are clearly overestimated while W2020 intensities agree with experiment except in case (c) where the W2020 intensity of the $2\nu_1+\nu_2$ $9_{1\,9} - 10_{0\,10}$ transition is strongly underestimated. This situation contrasts with that shown on panel (e) where the $10_{1\,10} - 11_{0\,11}$ of the same $2\nu_1+\nu_2$ band shows a W2020 line intensity largely overestimated. (Note that in (c), the HITRAN and W2020 intensities differ by a factor of 6).

The above examples illustrate the fact that it is difficult to propose empirical corrections of the calculated intensity values and that successive calculation might lead to large variation of the intensities in the case of "unstable" transitions very sensitive to small changes of the potential energy surface used for the calculations.

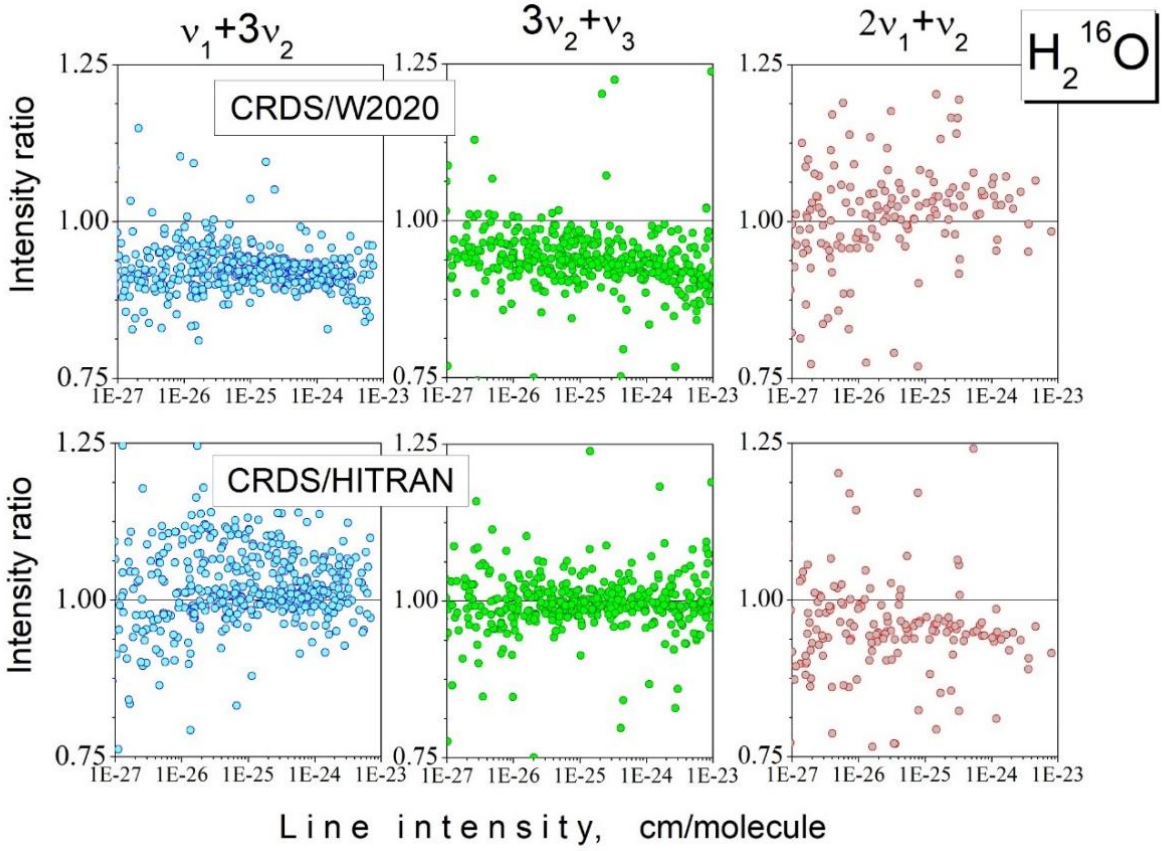

Fig. 9.
Ratios of the CRDS intensity values to the W2020 [20] and HITRAN2020 [9] intensities for $H_2{}^{16}O$ transitions of the $\nu_1+3\nu_2$, $3\nu_2+\nu_3$ and $2\nu_1+\nu_2$ bands measured in the 8041 - 8633 cm^{-1} interval.

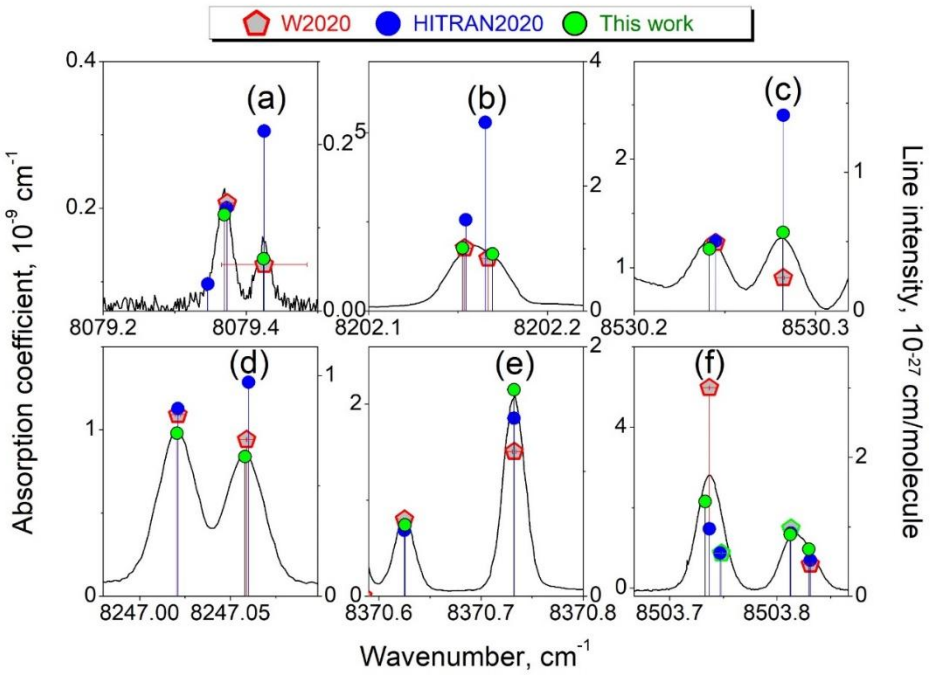

Fig. 10.
Examples of comparison of the CRDS spectrum of water vapor and corresponding line list (green circles) to the W2020 line list [20] (grey pentagons) and to the HITRAN2020 list of natural water [9] (blue circles). The right-hand intensity scale is adjusted to correspond approximately to the peak heights.
The rovibrational of the problematic lines is the following: (a) $\nu_1+\nu_3\ 13_{6\,8} - 13_{0\,13}$, (b) $\nu_1+3\nu_2\ 12_{3\,10} - 12_{4\,9}$, (c) $2\nu_1+\nu_2\ 9_{1\,9} - 10_{0\,10}$], (d) $\nu_1+4\nu_2\ 5_{1\,4} - 4_{2\,3}$, (e) $3\nu_2+\nu_3\ 6_{0\,6} - 5_{2\,3}$, and (f) $2\nu_1+\nu_2\ 10_{1\,10} - 11_{0\,11}$].

6. Conclusion

The room temperature absorption spectrum of water vapor has been recorded with unprecedented sensitivity in the high energy range of the 1.25 µm atmospheric transparency window (8040 – 8620 cm^{-1}). The use of a comb-referenced cavity ring-down spectrometer allowed for an absolute frequency calibration of the spectra. A list of about 5200 lines with intensity as weak as a few 10^{-30} cm/molecule was constructed and rovibrationally assigned to more than 5400 transitions of the first six water isotopologues ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, $HD^{16}O$, $HD^{18}O$ and $HD^{17}O$). About one third of the assigned transitions are newly measured and the first experimental determination of 81 rovibrational levels of ($H_2{}^{16}O$, $H_2{}^{18}O$, $H_2{}^{17}O$, and $HD^{16}O$) is reported.

This large dataset with high position accuracy (on the order of 10^{-4} cm^{-1} for isolated lines of intermediate intensity), provides stringent validation tests for previous experimental investigations and spectroscopic databases in the region. The comparison shows an overall satisfactory agreement but a systematic shift on the order of -8×10^{-4} cm^{-1} is evidenced compared to the most relevant FTS [8] and CRDS [6] studies in the region. As concerns the current version of the HITRAN database, the overall agreement is satisfactory but several issues are pointed, some of them having been already mentioned in other spectral regions from the far infrared [51] to the near infrared [32,33,52].

(i) Part of the HITRAN line positions lacks traceability. For instance, a large part of HITRAN positions is given with W2020 source but their positions may differ by a few 10^{-3} cm^{-1} from the

published W2020 position values [20]. Interestingly, we found examples where HITRAN positions with W2020 source agree much better with experiment than the original W2020 position values **(Fig. 7)**. In addition, for an unknown reason, old position sources [31,36,37,47] have been kept for several transitions while they have been superseded by recent works, in particular the W2020 lists [20]. Finally, some examples discussed above concern inaccurate positions given in HITRAN with our Ref. [38] as source while this reference provides a correct position value,

(ii) A few weak lines are missing in the HITRAN database. The strongest one belonging to the main isotopologue is assigned to $2\nu_1+\nu_2$ $12_{1\ 11} - 12_{2\ 10}$ and has an intensity larger than 2×10^{-27} cm/molecule (see **Fig. 5**),

(iii) A series of important deviations is noted for the $H_2{}^{17}O$ line positions above 8340 cm^{-1} where Ref. [40] is used as main HITRAN source. The comparison to the measurements (**Fig. 6**) indicates that the W2020 positions should have been preferred.

(iv) As concerns line intensities, most HITRAN values are calculated values by Conway et al. [49]. Overall, the comparison to our measurements indicates that these more recent calculations improve the POKAZATEL intensities of the W2020 list. We noticed that the observed deviations from our intensity measurements do not show systematic tendency which prevents empirical band-by-band corrections. As a large part of the observed deviations exceeds our experimental uncertainty, we believe that experimental intensity values should be preferred for most of the lines in the region.

The comparison of the measurements to the W2020 line positions has confirmed that 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 350) (see **Fig. 7**). The complex procedure elaborated to determine the empirical W2020 energy levels uses as basis a transition database including all the experimental sources available in the literature. As a result, the accuracy of the W2020 energy levels should ideally supersede the accuracy of all the sources used as input data. This is not the case [32,33,51,52]. For instance, in Ref. [52], we gave series of examples where the W2020 line positions differ from our previous CRDS measurements [6] while these measurements, confirmed by the present spectra, were included in the W2020 transition database. A serious issue identified in the derivation of the W2020 energy levels is related to the large datasets from emission spectra incorporated in the transition databases used to determine the W2020 energy levels. These emission line positions are generally reported with a poor accuracy compared to absorption data but they have a strong impact on the resulting energy values. In addition, many of these emission lines were assigned to several transitions. In the present study, some errors of the W2020 line positions were identified as due to the determination of the upper energy level from emission lines while (more accurate) absorption data were excluded. We are convinced that the overall quality of most W2020 energy levels and line positions would benefit from the exclusion of emission data.

The W2020 energy levels may also benefit from a stricter selection of the experimental sources. If we consider for instance a same spectral region studied successively by a same group, it would be reasonable to exclude the first measurements from the transition database. This is for example the case of the 8110 – 8340 cm^{-1} region in common between the present study and Ref. [6] (CRDS2015). Besides the more accurate frequency calibration of the present spectra, we corrected a few position values and assignments (see Section 3). For future derivation of the empirical energy levels, the best choice would be to keep only the transition frequencies of the present work and exclude those of Ref. [6]. The inclusion of the CRDS2015 line positions can only have a negative impact on the accuracy of the resulting energy levels.

As a final conclusion, let us underline that water calculated line lists have unique advantages in terms of spectral coverage and completeness. Our recent studies [32,33,51,52] and the present work have demonstrated that validation tests of the resulting calculated line lists against high quality experimental data are highly suitable to point deficiencies and bring hints for further improvements of calculated line lists.

Acknowledgements

The support of CNRS (France) in the frame of the International Research Project SAMIA is acknowledged. SNM activity was supported in the frame of the Russian Science Foundation, grant no. 18-11-00024-П.

References

1. Macko P, Romanini D, Mikhailenko SN, Naumenko OV, Kassi S, Jenouvrier A, Tyuterev VlG, Campargue A. High sensitivity CW-cavity ring down spectroscopy of water in the region of the 1.5 μm atmospheric window. J Mol Spectrosc 2004;227:90-108. https://doi.org/10.1016/j.jms.2004.05.020
2. Mikhailenko SN, Wang Le, Kassi S, Campargue A. Weak water absorption lines around 1.455 and 1.66 μm by CW-CRDS. J Mol Spectrosc 2007;244:170-8. https://doi.org/10.1016/j.jms.2007.05.013
3. Mikhailenko S, Kassi S, Wang Le, Campargue A. The absorption spectrum of water in the 1.25 μm transparency window (7408 – 7920 cm^{-1}). J Mol Spectrosc 2011;269:92-103. https://doi.org/10.1016/j.jms.2011.05.005
4. Leshchishina O, Mikhailenko S, Mondelain D, Kassi S, Campargue A. CRDS of water vapor at 0.1 Torr between 6886 and 7406 cm^{-1}. J Quant Spectrosc Radiat Transf 2012;113:2155-66. https://doi.org/10.1016/j.jqsrt.2012.06.026
5. Leshchishina O, Mikhailenko S, Mondelain D, Kassi S, Campargue A. An improved line list for water vapor in the 1.5 μm transparency window by highly sensitive CRDS between 5852 and 6607 cm^{-1}. J Quant Spectrosc Radiat Transf 2013;130:69-80. https://doi.org/10.1016/j.jqsrt.2013.04.010
6. Campargue A, Mikhailenko SN, Lohan BG, Karlovets EV, Mondelain D, Kassi S. The absorption spectrum of water vapor in the 1.25 μm atmospheric window (7911 – 8337 cm^{-1}). J Quant Spectrosc Radiat Transf 2015;157:135-52. https://doi.org/10.1016/j.jqsrt.2015.02.011
7. Mikhailenko SN, Karlovets EV, Vasilchenko S, Mondelain D, Kassi S, Campargue A. New transitions and energy levels of water vapor by high sensitivity CRDS near 1.73 and 1.54 μm. J Quant Spectrosc Radiat Transf 2019;236:106574. https://doi.org/10.1016/j.jqsrt.2019.106574
8. Régalia L, Oudot C, Mikhailenko S, Wang L, Thomas X, Jenouvrier A, Von der Heyden P. Water vapor line parameters from 6450 to 9400 cm^{-1}. J Quant Spectrosc Radiat Transf 2014;136:119-36. https://doi.org/10.1016/j.jqsrt.2013.11.019
9. Gordon IE, Rothman LS, Hargreaves RJ, Hashemi R, Karlovets EV, Skinner FM, Conway EK, Hill C, Kochanov RV, Tan Y, Wcisło P, Finenko AA, Nelson K, Bernath PF, Birk M, Boudon V, Campargue A, Chance KV, Coustenis A, Drouin BJ, Flaud J-M, Gamache RR, Hodges JT, Jacquemart D, Mlaver EJ, Nikitin AV, Perevalov VI, Rotger M, Tennyson J, Toon GC, Tran H, TyuterevVlG, Adkins EM, Baker A, Barbe A, Canè E, Császár AG, Egorov O, Fleisher AJ, Fleurbaey H, Foltynowicz A, Furtenbacher T, Harrison JJ, Hartmann J-M, Horneman V-M, Huang X, Karman T, Karns J, Kassi S, Kleiner I, Kofman V, Kwabia-Tchana F, Lee TJ, Long DA, Lukashevskaya AA, Lyulin OM, Makhnev VYu, Matt W, Massie ST, Melosso M, Mikhailenko SN, Mondelain D, Müller HSP, Naumenko OV, Perrin A, Polyansky OL, Raddaoui E, Raston PL, Reed ZD, Rey M, Richard C, Tóbiás R, Sadiek I, Schwenke DW, Starikova E, Sung K, Tamassia F, Tashkun SA, Vander Auwera J, Vasilenko IA, Vigasin AA, Villanueva GL, Vispoel B, Wagner G, Yachmenev A, Yurchenko SN. The HITRAN2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf 2022;277:107949. https://doi.org/10.1016/j.jqsrt.2021.107949
10. Delahaye T, Armante R, Scott NA, Jacquinet-Husson N, Chédin A, Crépeau L, Crevoisier C, Douet V, Perrin A, Barbe A, Boudon V, Campargue A, Coudert LH, Ebert V, Flaud J-M, Gamache RR, Jacquemart D, Jolly A, KwabiaTchana F, Kyuberis A, Li G, Lyulin OM, Manceron L, Mikhailenko S, Moazzen-Ahmadi N, Müller HSP, Naumenko OV, Nikitin A, Perevalov VI, Richard C, Starikova E, Tashkun SA, Tyuterev VlG, Vander Auwera J, Vispoel B, Yachmenev A, Yurchenko S. The 2020 edition of the GEISA spectroscopic database. J Mol Spectrosc 2021;380:111510. https://doi.org/10.1016/j.jms.2021.111510
11. Mondelain D, Sala T, Kassi S, Romanini D, Marangoni M, Campargue A. Broadband and highly sensitive comb-assisted cavity ring down spectroscopy of CO near 1.57 μm with sub-MHz frequency accuracy. J Quant Spectrosc Radiat Transf 2015;154:35-43. http://dx.doi.org/10.1016/j.jqsrt.2014.11.021

12. Modelain D, Mikhailenko SN, Karlovets EV, Béguer S, Kassi S, Campargue A. Comb-assisted cavity ring down spectroscopy of ^{17}O enriched water between 7443 and 7921 cm^{-1}. J Quant Spectrosc Radiat Transf 2017;203:206-12. https://doi.org/10.1016/j.jqsrt.2017.03.029
13. Mikhailenko SN, Mondelain D, Karlovets EV, Kassi S, Campargue A. Comb-assisted cavity ring down spectroscopy of ^{17}O enriched water between 6667 and 7443 cm^{-1}. J Quant Spectrosc Radiat Transf 2018;206:163-71. https://doi.org/10.1016/j.jqsrt.2017.10.023
14. Zobov NF, Shirin SV, Ovsyannikov RI, Polyansky OL, Barber RJ, Tennyson J, Bernath PF, Carleer M, Colin R, Coheur P-F. Spectrum of hot water in the 4750 – 13 000 cm^{-1} wavenumber range (0.769 – 2.1 μm). Mon Not R Astron Soc 2008;387:1093-8. https://doi.org/10.1111/j.1365-2966.2008.13234.x
15. Mikhailenko SN, Naumenko OV, Nikitin AV, Vasilenko IA, Liu A-W, Song K-F, Ni H-Y, Hu S-M. Absorption spectrum of deuterated water vapor enriched by ^{18}O between 6000 and 9200 cm^{-1}. J Quant Spectrosc Radiat Transf 2012;113:653-69. https://doi.org/10.1016/j.jqsrt.2012.02.009
16. https://spectra.iao.ru/molecules/simlaunch?mol=1
17. Partridge H, Schwenke DW. The determination of an accurate isotope dependent potential energy surface for water from extensive *ab initio* calculations and experimental data. J Chem Phys 1997;106:4618-39. https://doi.org/10.1063/1.473987
18. Schwenke DW, Partridge H. Convergence testing of the analytic representation of an *ab initio* dipole moment function for water: Improved fitting yields improved intensities. J Chem Phys 2000;113:6592-7. https://doi.org/10.1063/1.1311392
19. Mikhailenko SN. Kassi S, Mondelain D, Campargue A. Water vapor absorption between 5690 and 8340 cm^{-1}: Accurate empirical line centers and validation tests of calculated line intensities. J Quant Spectrosc Radiat Transf 2020;245:106840. https://doi.org/10.1016/j.jqsrt.2020.106840
20. Furtenbacher T, Tóbiás R, Tennyson J, Polyansky OL, Kyuberis AA, Ovsyannikov RI, Zobov NF, Császár AG. The W2020 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
21. Flaud J-M, Camy-Peyret C, Narahari Rao K, Chen D-W, Hoh Y-S, Maillard J-P. Spectrum of water vapor between 8050 and 9370 cm^{-1}. J Mol Spectrosc 1979;75:339-62. https://doi.org/10.1016/0022-2852(79)90081-X
22. Mandin J-Y, Chevillard J-P, Flaud J-M, Camy-Peyret C. $H_2{}^{16}O$: line positions and intensities between 8000 and 9500 cm^{-1}: the second hexad of interacting vibrational states {(050), (130), (031), (210), (111), (012)}. Can J Phys 1988;66:997-1011. https://doi.org/10.1139/p88-162
23. Bykov A, Naumenko O, Sinitsa L, Voronin B, Flaud J-M, Camy-Peyret C, Lanquetin R. High-order resonances in the water molecule. J Mol Spectrosc 2001;205:1-8. https://doi.org/10.1006/jmsp.2000.8231
24. Tolchenov RN, Tennyson J. Water line parameters for weak lines in the range 7400 – 9600 cm^{-1}. J Mol Spectrosc 2005;231:23-7. https://doi.org/10.1016/j.jms.2004.12.001
25. Oudot C, Wang Le, Thomas X, Von der Heyden P, Daumont L, Régalia L. Intensity measurements of $H_2{}^{16}O$ lines in the spectral region 8000 – 9350 cm^{-1}. J Mol Spectrosc 2010;262:22-9. https://doi.org/10.1016/j.jms.2010.04.011
26. Liu A-W, Hu S-M, Camy-Peyret C, Mandin J-Y, Naumenko O, Voronin B. Fourier transform absorption spectra of $H_2{}^{17}O$ and $H_2{}^{18}O$ in the 8000 – 9400 cm^{-1} spectral region. J Mol Spectrosc 2006;237:53-62. https://doi.org/10.1016/j.jms.2006.02.008
27. Hu S-M, He S-G, Zheng J-J, Wang X-H, Ding Y, Zhu Q-S. High-resolution analysis of the $\nu_2+2\nu_3$ band of HDO. Chinese Physics 2001;10:1021-7. https://doi.org/10.1088/1009-1963/10/11/306
28. Naumenko OV, Voronina S, Hu S-M. High resolution Fourier transform spectrum of HDO in the 7500 – 8200 cm^{-1} region: revisited. J Mol Spectrosc 2004;227:151-7. https://doi.org/10.1016/j.jms.2004.06.002
29. Mikhailenko SN, Tashkun SA, Putilova TA, Starikova EN, Daumont L, Jenouvrier A, Fally S, Carleer M, Hermans C, Vandaele AC. Critical evaluation of rotation-vibration transitions and an experimental dataset of energy levels of $HD^{18}O$.J Quant Spectrosc Radiat Transf 2009;110:597-608. https://doi.org/10.1016/j.jqsrt.2009.01.012

30. Mikhailenko SN, Tashkun SA, Daumont L, Jenouvrier A, Carleer M, Fally S, Vandaele AC. Line positions and energy levels of the 18-O substitutions from the HDO/D_2O spectra between 5600 and 8800 cm^{-1}. J Quant Spectrosc Radiat Transf 2010;111:2185-96. https://doi.org/10.1016/j.jqsrt.2010.01.028
31. Tennyson J, Bernath PF, Brown LR, Campargue A, Császár AG, Daumont L, Gamache RR, Hodges JT, Naumenko OV, Polyansky OL, Rothman LS, Toth RA, Vandaele AC, Zobov NF, Fally S, Fazliev AZ, Furtenbacher T, Gordon IE, Hu S-M, Mikhailenko SN, Voronin BA. 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
32. 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
33. 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
34. Bordet B, Kassi S, Campargue A. Line parameters of the 4-0 band of carbon monoxide by high sensitivity cavity ring down spectroscopy near 1.2 μm. J Quant Spectrosc Radiat Transf 2021;260:107453. https://doi.org/10.1016/j.jqsrt.2020.107453
35. Sironneau VT, Hodges JT. Line shapes, positions and intensities of water transitions near 1.28 μm. J Quant Spectrosc Radiat Transf 2015;152:1-15. https://doi.org/10.1016/j.jqsrt.2014.10.020
36. Tennyson J, Bernath PF, Brown LR, Campargue A, Császár AG, Daumont L, Gamache RR, Hodges JT, Naumenko OV, Polyansky OL, Rothman LS, Vandaele AC, Zobov NF, Al Derzi AR, Fábri C, Fazliev AZ, Furtenbacher T, Gordon IE, Lodi L, Mizus II. 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
37. Bubukina II, Zobov NF, Polyansky OL, Shirin SV, Yurchenko SN. Optimized semiempirical potential energy surface for $H_2{}^{16}O$ up to 26000 cm^{-1}. Opt Spectrosc 2011;110:160-6. https://doi.org/10.1134/S0030400X11020032
38. Mikhailenko SN, Kassi S, Mondelain D, Gamache RR, Campargue A. A spectroscopic database for water vapor between 5850 and 8340 cm^{-1}. J Quant Spectrosc Radiat Transf 2016;179:198-216. https://doi.org/10.1016/j.jqsrt.2016.03.035
39. Campargue A, Kassi S, Yachmenev A, Kyuberis AA, Küpper J, Yurchenko SN. Observation of electric-quadrupole infrared transitions in water vapor. Phys Rev Res 2020;2:023091. https://doi.org/10.1103/PhysRevResearch.2.023091
40. 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
41. Kyuberis AA, Zobov NF, Naumenko OV, Voronin BA, Polyansky OL, Lodi L, Liu A, Hu S-M, Tennyson J. 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
42. Polyansky OL, Kyuberis AA, Zobov NF, Tennyson J, Yurchenko SN, Lodi L. ExoMol molecular line lists XXX: a complete high-accuracy line list for water. Mon Not R Astron Soc 2018;480:2597-608. https://doi.org/10.1093/mnras/sty1877
43. Zobov NF, Shirin SV, Polyansky OL, Barber RJ, Tennyson J, Coheur P-F, Bernath PF, Carleer M, Colin R. 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
44. Rutkowski L, Foltynowicz A, Schmidt FM, Johansson AC, Khodabakhsh A, Kyuberis AA, Zobov NF, Polyansky OL, Yurchenko SN, Tennyson J. 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
45. Coheur P-F, Bernath PF, Carleer M, Colin R, Polyansky OL, Zobov NF, Shirin SV, Barber RJ, Tennyson J. A 3000 K laboratory emission spectrum of water. J Chem Phys 2005;122:074307. https://doi.org/10.1063/1.1847571

46. Czinki E, Furtenbacher T, Csaszar AG, Eckhardt AK, MellauGCh. 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
47. Tennyson J, Bernath PF, Brown LR, Campargue A, Carleer MR, Császár AG, Gamache RR, Hodges JT, Jenouvrier A, Naumenko OV, Polyansky OL, Rothman LS, Toth RA, Vandaele AC, Zobov NF, Daumont L, Fazliev AZ, Furtenbacher T, Gordon IE, Mikhailenko SN, Shirin SV. 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
48. Liu A-W, Naumenko OV, Kassi S, Campargue A. CW-Cavity Ring Down Spectroscopy of deuterated water in the 1.58 μm atmospheric transparency window. J Quant Spectrosc Radiat Transf 2014;138:97-106. https://doi.org/10.1016/j.jqsrt.2014.02.002
49. Conway EK, Gordon IE, Kyuberis AA, Polyansky OL, Tennyson J, Zobov NF. Calculated line lists for $H_2{}^{16}O$ and $H_2{}^{18}O$ with extensive comparisons to theoretical and experimental sources including the HITRAN2016 database. J Quant Spectrosc Radiat Transf 2020;241:106711. https://doi.org/10.1016/j.jqsrt.2019.106711
50. Barber RJ, Tennyson J, Harris GJ, Tolchenov RN. A high-accuracy computed water line list. Mon Not R Astron Soc 2006;368:1087-94. https://doi.org/10.1111/j.1365-2966.2006.10184.x
51. Toureille M, Koroleva AO, Mikhailenko SN, Pirali O, Campargue A. Water vapor absorption spectroscopy in the far-infrared (50-720 cm^{-1}). Part 1: Natural water. J Quant Spectrosc Radiat Transf 2022;291:108326. https://doi.org/10.1016/j.jqsrt.2022.108326
52. 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