A spectroscopic database for water vapor between 5850 and 8340 cm⁻¹

S.N. Mikhailenko^{a,b}, S. Kassi^{d,e}, D. Mondelain^{d,e}, R. R. Gamache^c, and A. Campargue^{d,e*}

^a *Laboratory of Theoretical Spectroscopy, V.E. Zuev Institute of Atmospheric Optics, SB, Russian Academy of Science, 1, Akademician Zuev square, 634021 Tomsk, Russia*

^b *Mathematical Physics Department, Tomsk Polytechnic University 30, Lenin av., 634050, Tomsk, Russia*

^c *Office of Academic Affairs, Student Affairs, International Relations, University of Massachusetts, One Beacon Street, Boston, MA 02108, USA*

^d *Univ. Grenoble Alpes, LIPhy, F-38000 Grenoble, France*

^e *CNRS, LIPhy, F-38000 Grenoble, France*

Key words: water vapor; H₂O; absorption spectroscopy; Cavity Ring Down Spectroscopy; HITRAN; GEISA; spectroscopic database

Number of Pages: 36

Number of Figures: 12

Number of Tables: 8

Corresponding author.

E-mail address: Alain.Campargue@ujf-grenoble.fr

Tel.: 33 4 76 51 43 19 Fax. 33 4 76 63 54 95

Abstract

An accurate and complete empirical spectroscopic database is constructed for natural water in the 5850–8340 cm^{-1} near infrared region (1.71 – 1.20 μm). The global list includes more than 46500 transitions of the six most abundant isotopologues (H_2^{16}O , H_2^{18}O , H_2^{17}O , HD^{16}O , HD^{17}O , HD^{18}O) in natural isotopic abundance with intensity greater than 1×10^{-29} $\text{cm}/\text{molecule}$. *All* line positions were obtained by difference of empirically determined energy values. These empirical determinations are mostly based on extensive investigations of water spectra by high sensitivity CRDS complemented with literature data. In particular, the sets of energy levels of the minor isotopologues have been significantly enlarged from very recent analysis of CRDS spectra of D, ^{18}O and ^{17}O enriched water samples. The empirical line positions were associated to experimental intensity values from the literature obtained by FTS for the strong lines and CRDS for weaker lines. The list is made mostly complete by including a large number of weak transitions not yet detected, with empirical positions calculated from empirical levels associated with variational intensities. Air- and self-broadening coefficients have been added to each transition using the most recent version of the algorithm used for the HITRAN and GEISA databases. Examples of comparison of CRDS spectra to simulations based on the HITRAN2012 list illustrate the advantages of our database.

The relative merit of experimental and variational line intensities is discussed. A separate line list combining empirical line positions and variational intensities is provided. Future improvements from combination of experimental and variational intensities are suggested.

1. Introduction

In 2014, we constructed an empirical line list for water vapor in natural isotopic abundance between 5850 and 7920 cm^{-1} [1]. The list contained 40290 transitions of the six most abundant water isotopologues (H_2^{16}O , H_2^{18}O , H_2^{17}O , HD^{16}O , HD^{18}O and HD^{17}O) with an intensity cut off of 1×10^{-29} $\text{cm}/\text{molecule}$, identical to the HITRAN intensity cut off. Each transition has complete spectroscopic assignment ($V_1 V_2 V_3 J K_a K_c$). Our list (namely “2014 list”, hereafter) consists in gathering spectroscopic parameters from mainly three sources:

- (i) FTS data by Toth [2] for the strongest lines,
- (ii) CRDS data of Refs. [3, 4, 5] for the weak and medium lines,
- (iii) (not yet measured) transitions with frequency calculated using known empirical energy levels and intensities obtained from variational calculations by Schwenke and Partridge (SP) for the unobserved lines [6-8].

Our 2014 list was adopted for the current version of the GEISA database [9] improving significantly the previous version [10]. The current version of the HITRAN list [11] is a composite of experimental and variational line parameters depending on the spectral range and on the isotopologue (see Ref. [1]). A laborious comparison of the CRDS spectra to simulation using the 2014 list and the corresponding HITRAN2012 list allowed identifying a number of deficiencies and errors in the HITRAN list concerning mostly lines of small or intermediate intensities. In Ref. [1], the various sources of data adopted in HITRAN for positions and intensities were examined in details and we could identify the following problems:

- (i) missing HD^{16}O lines; this isotopologue in natural abundance has a strong contribution in particular in the transparency windows near 6200 and 8000 cm^{-1} ,
- (ii) absence of HD^{18}O and HD^{17}O transitions; these isotopologues have lines with intensity above the HITRAN intensity cut off,
- (iii) duplicated or non traceable lines,
- (iv) inaccurate line positions obtained from variational calculations or derived from inaccurate empirical energy levels obtained by an IUPAC-TG¹.
- (v) inaccurate line intensities from experiment or variational calculations in particular for bands involving a highly bending excitation.

¹ A few years ago, under the project “A database of water transitions from experiment and theory”, an IUPAC task group (IUPAC-TG) reported an exhaustive review and evaluation of rovibrational line positions, energy levels and assignments for all the main isotopologues of water. By using the procedure and code MARVEL (Measured Active Rotational-Vibrational Energy Levels) [12, 13], all the high quality absorption or emission line positions available in the literature were used to determine and validate the energy levels together with their self consistent uncertainties. The results for the different isotopologues can be found in Refs. [14-17].

The present work is devoted to an extension and an update of the 2014 list mainly focused on the following aspects:

(i) On the basis of a recent CRDS study of the water vapor spectrum, the upper wavenumber limit of the list is increased up to 8340 cm^{-1} , thus covering entirely the important $1.25\text{ }\mu\text{m}$ atmospheric window [18];

(ii) As said above, the 2014 list includes only lines of which the positions have experimental origin, not only transitions which were directly observed by experiment (FTS and CRDS) but also not yet observed transitions between energy levels, experimentally determined as lower or upper states of different transitions. Despite its completeness, about 3100 weak transitions (over a total of 40290) with intensity larger than $1\times 10^{-29}\text{ cm/molecule}$ could not be included in the 2014 list because their upper energy levels lacked experimental determination [1]. These missing transitions are mostly due to H_2^{17}O (1400 transitions) and HD^{16}O (1090 transitions). In the present work, the amount of missing transitions is significantly reduced using recent analysis of CRDS spectra of D [19], ^{18}O [20] and ^{17}O [21] enriched water samples which provided a large amount of new energy levels for HD^{16}O , H_2^{18}O and H_2^{17}O ;

(iii) While in the 2014 list, the CRDS line position values were kept as measured, in the present list, all frequencies were obtained by differences of empirically determined energy values.

(iv) In the present list, air- and self broadened half-widths have been added to each transition using the algorithm which has been used for the last edition of both the HITRAN [11] and GEISA [9] databases. The resulting database allows for a direct comparison of experimental spectra to simulations based on our empirical list.

The rest of the paper is organized as follows. Some aspects of the CRDS recordings of water vapor spectra used for the line list are recalled in Section 2. The line list construction is detailed in Section 3 with sub-section devoted to line positions, line intensities and line profile parameters. In the Discussion (Section 4), we consider in detail the advantages of experimental and variational intensities. Finally, in the Conclusion (Section 5), we illustrate the advantages of the new line list by comparison to simulations based on the HITRAN line list and with CRDS spectra and we suggest some possible further improvements.

2. Some aspects of the CRDS spectra of water vapor

As said above, a great amount of our empirical database relies on CRDS recordings performed since 2004 by high-sensitivity Continuous Wave Cavity Ring Down Spectroscopy (CW-CRDS) [3-5, 18-23]. The light sources used for these recordings consists of (i) a collection of about 90 Distributed Feedback (DFB) laser diodes, each of them with 30 cm^{-1} tuning range, allowing us to cover the $5850\text{-}7920\text{ cm}^{-1}$ range and (ii) an External Cavity Laser Diode (ECDL) giving access to the $7920\text{-}8340\text{ cm}^{-1}$ region. While this set up allowed

achieving a record sensitivity of $\alpha_{min} \sim 5 \times 10^{-13} \text{ cm}^{-1}$ by averaging over a small spectral interval [24], the typical noise equivalent absorption of the water spectra was between $\alpha_{min} \sim 2 \times 10^{-11}$ and $2 \times 10^{-10} \text{ cm}^{-1}$ depending on the reflectivity of the CRDS super mirrors. The spectra of “natural” water were recorded over the entire 5850-8340 cm^{-1} range except for a few narrow spectral gaps. In addition, the present database takes into account the recent analysis of CRDS spectra of D [19], ^{18}O [20] en ^{17}O [21] enriched water samples recorded from 5850 to about 6750 cm^{-1} .

In order to illustrate the sensitivity and high dynamics on the intensity scale allowed by the CRDS technique, we present in **Fig. 1** four successive zooms of a sample of the spectrum recorded at 1.8 Torr with an ECDL near 8082 cm^{-1} [18]. Absorption coefficients are measureable from the noise level on the order of $\alpha_{min} \sim 1.5 \times 10^{-11} \text{ cm}^{-1}$ up to $3 \times 10^{-6} \text{ cm}^{-1}$. By using pressure ranging between 0.1 and 10 Torr, lines with intensities ranging over six orders of magnitude – from to 5×10^{-30} up to $1 \times 10^{-24} \text{ cm/molecule}$ – could be measured.

On the upper panel of **Fig. 1**, the few lines of the region that are included in the HITRAN list with experimental line intensities have been highlighted. They correspond roughly to lines with intensity larger than $3 \times 10^{-26} \text{ cm/molecule}$, measured by FTS by Toth [2] and Régalia et al. [25], below and above 8000 cm^{-1} , respectively. In fact, the routine sensitivity of our CRDS set up allowed us to fill the four orders of magnitude gap of the line intensities between the FTS detectivity threshold and the $1 \times 10^{-29} \text{ cm/molecule}$ intensity cut off of the HITRAN list. For instance, the weakest line on the lower panel of **Fig. 1** (near 8082.27 cm^{-1}) observed with a good signal to noise ratio at 1.8 Torr has an intensity of about $3 \times 10^{-28} \text{ cm/molecule}$. In fact, in the 1.5 and 1.25 μm transparency windows, the smallest experimental intensities measured by CRDS are on the order of $5 \times 10^{-30} \text{ cm/molecule}$ [5, 18] i.e. below the HITRAN intensity cut off. Another illustration of the achieved sensitivity was the first laboratory detection of lines due to HD^{18}O and HD^{17}O in natural isotopic abundance (6.23003×10^{-7} and 1.15853×10^{-7} , respectively) [5]. **Fig. 2** presents the spectrum in a spectral interval of the 1.5 μm window where lines of five water isotopologues are identified, in particular the $4_{04} - 5_{15}$ transition of the $\nu_1 + \nu_3$ band of HD^{17}O with a line intensity of about $3.6 \times 10^{-29} \text{ cm/molecule}$.

Fig. 1.

CW-CRDS spectrum of water vapor near 8082 cm⁻¹ recorded with an ECDL [18]. The sample pressure was about 1.8 Torr. Four successive enlargements illustrate the sensitivity and high dynamics of the recordings allowing for the measurement of absorption coefficient spanning more than four orders of magnitude. The insert on the lower panel illustrates the achieved noise equivalent absorption, $\alpha_{min} \sim 1.5 \times 10^{-11} \text{ cm}^{-1}$. The four transitions marked with blue circles on the upper panel are the only lines with experimental intensities provided in the HITRAN database [11] in the considered interval.

Fig. 2.

CW-CRDS spectrum of natural water vapor near 6324 cm^{-1} [5]. The sample pressure was 10.0 Torr. In the displayed spectral interval, transitions of five water isotopologues are identified. The line of HD^{17}O in natural isotopic abundance (1.15853×10^{-7}) has an intensity of $3.6 \times 10^{-29}\text{ cm/molecule}$.

The line centers and intensities were determined using an interactive least squares multi-line fitting program assuming a Voigt profile, the HWHM of the Gaussian component being generally fixed to the theoretical value of the Doppler width. The calibration of the frequency axis of the CRDS spectra relied on the wavelength values provided by a wavemeter. It was refined using reference line positions of H_2O from the HITRAN database [11]. We estimate the average uncertainty on the line centers of isolated lines to be $\pm 1 \times 10^{-3}\text{ cm}^{-1}$. It is more difficult to give an average uncertainty value for line intensities. Comparisons with literature values and between various determinations from different CRDS spectra indicate that a value of 2-5% is a reasonable estimate for well-isolated H_2^{16}O lines of intermediate intensity (see Section 4). Larger error bars should be attached to blended lines and the relative abundance of the minor isotopologues adds uncertainty to the line intensities of these species. In Section 4, we give additional information about the quality of the intensity values included in the database.

3. Line list construction

The construction of our line list presented on the upper panel of **Fig. 3**, uses as starting point the variational line lists [8] computed by S.A. Tashkun based on the variational calculation results of Schwenke and Partridge [6, 7] for each of the six most abundant isotopologues (H_2^{16}O , H_2^{18}O , H_2^{17}O , HD^{16}O , HD^{18}O , and HD^{17}O). These variational line lists are probably not the best to date variational lists (in particular for line positions) but they have the advantage of including reasonable rovibrational assignment and to be publicly available at <http://spectra.iao.ru/> for all the six required isotopologues. The general procedure consisted of replacing as many variational line positions and line intensities as possible by their corresponding empirical values.

3.1 Line positions

In a first step, we replaced the variational line positions by their corresponding values calculated using empirical values of the lower and upper energy levels. As shown below, in the considered $5850 - 8340 \text{ cm}^{-1}$ region, this is possible for the quasi totality of the transitions with intensity larger than the $1 \times 10^{-29} \text{ cm/molecule}$ cut off. In the present list, this method was applied to both the not yet detected transitions and to the positions of lines measured by FTS or CRDS. This was not the case for the 2014 list where CRDS line positions were kept as measured [1]. The adopted procedure together with the correction of a number of energy levels is expected to improve the accuracy of the transition frequencies in particular for blended lines. Improved sets of energy levels were constructed for the various isotopologues. The main sources of the lower and upper energies are summarized in **Table 1** and presented below.

Table 1.

Main sources used for determining the lower and upper energy levels adopted for the construction of the line list of water between 7920 and 8350 cm^{-1} and approximate number of FTS and CRDS transition frequencies.

Isotopologue	Lower levels	Upper levels	
		Origin	N^b
H_2^{16}O	IUPAC-TG [16]	CRDS, IUPAC-TG [16], Sironneau & Hodges [26]	15300
H_2^{18}O	Mellau & Mikhailenko [27]	CRDS, IUPAC-TG [14]	6450
H_2^{17}O	Mellau & Mikhailenko [27]	CRDS, IUPAC-TG [14]	4520
HD^{16}O	IUPAC-TG [15]	CRDS, IUPAC-TG [15]	11380
HD^{18}O	Calculated values [28] ^a	CRDS + FTS [30]	3920
HD^{17}O	Calculated values [29] ^a	CRDS + FTS [30]	7290

Note

^a Lower state energy calculated using the effective Hamiltonian approach.

^b Approximate number of FTS and CRDS transition frequencies used to derive the upper energy levels

Fig. 3. Overview comparison between our list and the HITRAN2012 line list for water vapor between 5850 and 8340 cm^{-1}

Lower energy levels

For the lower state energy levels of H_2^{16}O and HD^{16}O , we used the values recommended by the IUPAC-TG [15, 16]. In the case of H_2^{18}O and H_2^{17}O , some levels of the (000) and (010) states are missing or inaccurate in the IUPAC-TG dataset [14], in particular those corresponding to high J and K_a values of the vibrational ground state. Thus, for H_2^{18}O and H_2^{17}O , we preferred the lower state energy levels of Ref. [27] recently derived from hot emission spectra of ^{18}O and ^{17}O enriched water.

The most extended energy sets of the ground and first excited states of HD^{18}O and HD^{17}O were reported by Liu et al. [28] and Mellau et al. [29], respectively. The lower state energy values obtained using the effective Hamiltonian model developed in Ref. [28] and Ref.

[29], were adopted for HD¹⁸O and HD¹⁷O, respectively. Their accuracy is estimated to be better than 5×10⁻⁴ and 10⁻³ cm⁻¹, respectively.

Upper energy levels

The Ritz combination principle was used to determine the upper energy levels from measured line positions. **Table 1** includes the approximate number of line positions used as input data for the energy level derivation of the six isotopologues. All our extensive CRDS studies of water absorption between 5850 and 8340 cm⁻¹ using normal [3, 4, 5, 18, 22, 23] or isotopically enriched [19, 20, 21] samples were used in order to get improved sets of upper energy levels for H₂¹⁶O, H₂¹⁸O, H₂¹⁷O, HD¹⁶O, HD¹⁸O and HD¹⁷O. The FTS study of deuterated water vapor enriched in ¹⁸O [30] was also considered for HD¹⁸O and HD¹⁷O. Overall, more than 42500 CRDS line positions of the four most abundant isotopologues and about 6300 FTS line positions of HD¹⁸O and HD¹⁷O [30] were compiled to build the most complete and precise energy levels sets for the region of the interest. Note that the CRDS entries also include 71 accurate H₂¹⁶O line positions reported by Sironneau and Hodges with a 3 MHz (10⁻⁴ cm⁻¹) accuracy in the 7710-7920 cm⁻¹ interval [26] (see **Table 1**).

Table 2 gives, for each isotopologue, the number of upper energy levels used to construct the line list in the 5850 – 8340 cm⁻¹ region. This Table includes in the third column, the number of additional energy levels compared to those of the IUPAC-TG datasets [14-16]. About eighty new levels were derived for H₂¹⁶O and H₂¹⁸O. Our recent CRDS investigations with isotopically enriched samples [19, 20, 21] allowed for a significant extension of the sets of energy levels of H₂¹⁷O, HD¹⁶O, HD¹⁸O and HD¹⁷O. For instance, compared to the IUPAC-TG sets, the number of levels required to construct the list in our region is duplicated for H₂¹⁷O and increased from 1261 to 1685 for HD¹⁶O. None of the 191 required upper levels of HD¹⁷O is provided by the IUPAC-TG [15].

Table 2.

Upper energy levels used for the construction of the line list of water between 7920 and 8350 cm⁻¹ and comparison to the IUPAC-TG energy levels [14-16].

Isotopologue	Total number	Upper levels absent in the IUPAC-TG sets	Number of levels according to the difference $\delta = E_{\text{This work}} - E_{\text{IUPAC}}$ in 10 ⁻³ cm ⁻¹ unit				
			$\delta = 0$ ^a	$0 < \delta < 2$	$2 < \delta < 10$	$10 < \delta < 100$	$100 < \delta$
H ₂ ¹⁶ O	3828	83	1108	2001	521	79	36
H ₂ ¹⁸ O	1645	74	386	1058	108	16	2
H ₂ ¹⁷ O	1278	637	108	412	99	21	1
HD ¹⁶ O	1685	424	240	809	198	13	1
HD ¹⁸ O	408	20	81	255	42	10	0
HD ¹⁷ O	191	191					

Note

^a $\delta = 0$ indicates that the upper energy level value provided by the IUPAC-TG was adopted for the line list construction.

Our energy values were systematically compared to the corresponding IUPAC-TG value, when available. The overall agreement is very good; for instance for the main isotopologue, excluding outliers with deviations greater than 0.1 cm⁻¹, the average and standard deviations are 6.3×10⁻⁴ cm⁻¹ and 5.4×10⁻⁴ cm⁻¹, respectively. It is worth mentioning that our sets of transition frequencies rely mostly on line positions measured in the range of

our database (5850 – 8340 cm^{-1}) and are thus more limited than the one collected by the IUPAC-TG. The IUPAC-TG considered all the transition frequencies reported in the literature from the microwave to the visible range, including many hot band transitions from emission spectra in the medium infrared and part of our CRDS results. In case of significant differences *i.e.* $\delta = E_{\text{This work}} - E_{\text{IUPAC}} > 0.5 \times 10^{-3} \text{ cm}^{-1}$, the energy level determination was examined in details using the CRDS spectra at disposal and in a number of cases the IUPAC-TG value was preferred (for instance, 29% and 23% of the H_2^{16}O and H_2^{18}O levels, respectively). We have included in **Table 2** the number of levels for various intervals of the deviation values, δ . For thirty six H_2^{16}O energy levels, the deviation exceeds 0.1 cm^{-1} . In general, these large deviations are due to IUPAC-TG energy levels relying on a single line reported with erroneous assignment or inappropriate weighting of transition frequencies obtained from different sources, in particular emission data. In Refs. [1, 5, 18], we have discussed in details a number of these situations (see in particular one of the files provided as Supplementary material attached to Ref. [5]). We give here a single example of an improper weighting of the experimental input data used by the MARVEL procedure leading to an inaccurate energy value. One of the largest deviations (-0.646 cm^{-1}) between our and IUPAC-TG energy levels concerns the (021) 16_{215} level of H_2^{16}O . A detailed analysis indicates that the IUPAC-TG value (9814.4500 cm^{-1} [16]) relies on three transition frequencies measured by emission [31, 32]. For an unknown reason, our previous observation from the $2\nu_2 + \nu_3$ $16_{215} - 17_{216}$ absorption line at 6522.6544 cm^{-1} leading to a (correct) value of 9813.807 cm^{-1} [22] was discarded by the MARVEL procedure [16]. The uncertainties on the emission data were increased up to 0.03 cm^{-1} to achieve consistency between them. Our present energy value (9813.804 cm^{-1}) relies on two recent CRDS determinations [5]: 9813.8046 cm^{-1} from the $2\nu_2 + \nu_3$ $16_{215} - 17_{216}$ line at 6522.6558 cm^{-1} and 9813.8033 cm^{-1} from the $2\nu_2 + \nu_3$ $16_{215} - 16_{214}$ line at 6602.7485 cm^{-1} [5]. Interestingly, these two transitions are included in the HITRAN database with correct line positions calculated by Gordon (HITRAN upper state energy of 9813.8070 cm^{-1}) and variational intensities.

The preceding example is a typical case where a single absorption datum should be preferred to various emission data. In fact, in addition to the energy levels corrected from their IUPAC-TG values on the basis of the CRDS spectra (see **Table 2**), twenty-two IUPAC-TG levels of H_2^{16}O relying almost exclusively from emission data, were also excluded (see **Table 3**). These levels were excluded considering their deviations from both their variational values and effective Hamiltonian calculations, which were compared to those of the nearby rotational levels of the same vibrational band. These excluded levels are mostly high rotational levels of pure bending vibrational states, mostly (040), or of the $\nu_1 + \nu_3$ band. Note all but three energies of **Table 3** are graded as *C*. As mentioned in Ref. [16], “This means that

there are energy levels with the worst grade, *C*, which may be accurately known. Nevertheless, it is expected that since many of the grade *C* energy levels come from a single source, their uncertainty may not be dependable”.

Table 3.

IUPAC-TG energy levels of H_2^{16}O [16] which were excluded from the line position calculation.

Vib	<i>J</i>	<i>K_a</i>	<i>K_c</i>	<i>E</i> / cm^{-1}	<i>dE</i>	<i>Nt</i>	Grade	Transition Refs.	
040	11	10	2	10193.838125	5000	2	C	[31]	[32]
040	11	10	1	10193.838125	5000	3	C	[31]	[32]
040	11	11	1	10533.022713	2058	3	C	[31]	[33]
040	11	11	0	10533.022713	4010	3	C	[31]	[33]
040	12	10	3	10489.766330	5000	2	C	[31]	[33]
040	12	10	2	10489.766330	5000	2	C	[31]	[33]
040	12	11	2	10833.515904	2403	5	C	[31]	[33]
040	12	11	1	10833.519495	5000	3	C	[31]	[33]
040	14	9	6	10827.162638	1371	5	C	[31]	[32]
040	14	9	5	10827.162441	50000	1	C	[32]	
050	13	7	6	11404.902413	8653	2	C	[33]	[34]
060	0	0	0	8869.950054	5000	1	C	[34]	
060	6	0	6	9339.638670	5000	1	C	[33]	
101	15	9	7	11078.329475	2063	4/9	A+	[32]	[34]
101	15	10	6	11328.487600	6463	3	C	[32]	[34]
101	15	11	5	11556.066789	2737	4	B-	[32]	[34]
101	15	11	4	11556.070231	3261	4	C	[32]	[34]
101	19	3	16	11694.160362	2260	5	B-	[32]	[34]
101	20	3	18	11819.333261	5677	1	C	[32]	
111	12	8	5	11560.519711	8000	1	C	[36]	
111	13	8	6	11868.377421	775	8	C	[34]	[36]
210	15	0	15	11015.401680	952	3	C	[34]	[37]

Notes: Vib *J Ka Kc* – vibrational and rotational quantum numbers; *E* / cm^{-1} – empirical energy value; *dE* – energy uncertainty in 10^{-6} cm^{-1} unit; *Nt* – number of transitions used for the energy level determination; Grade – energy level grade (see Ref. [16]); Transition Refs – reference of the data source used by the IUPAC-TG [16].

Transition frequencies

Using the derived sets of empirical energy levels, the frequencies of a total of 46 813 transitions with variational intensity larger than $1 \times 10^{-29} \text{ cm/molecule}$ were calculated for H_2^{16}O , H_2^{18}O , H_2^{17}O , HD^{16}O , HD^{17}O and HD^{18}O . According to the variational calculations, 144, 187, 127 and 191 empirical energy levels are missing thus we cannot compute empirical line positions of a total of 1198 very weak transitions of H_2^{16}O , H_2^{18}O , H_2^{17}O and HD^{16}O . In spite of a larger spectral extension, this number of missing lines is significantly reduced (1198 compared to 3161) compared to the 2014 list (see **Fig. 4**). As indicated in **Table 4**, this improvement concerns mostly the HD^{16}O and H_2^{17}O species, which benefited from recent CRDS investigations in Ref. [19] and Ref. [21], respectively.

Table 4

Comparison of the number of transitions calculated from empirical energy levels (*N1*) or reaching an upper level which is not yet experimentally determined (*N2*), in the 2014 list [1] and in the present database.

Isotopologue	2014 list [1] (5850 – 7920 cm ⁻¹)		This work (5850 – 8340 cm ⁻¹)	
	<i>N1</i>	<i>N2</i>	<i>N1</i>	<i>N2</i>
H ₂ ¹⁶ O	18757	346	21319	293
H ₂ ¹⁸ O	6641	328	7248	322
H ₂ ¹⁷ O	3659	1373	5112	248
HD ¹⁶ O	9261	1114	11133	333
HD ¹⁸ O	1372	0	1397	0
HD ¹⁷ O	600	0	604	0
Total	40290	3161	46813	1196

Notes

N1 - Number of variational transitions calculated from empirical energy levels

N2 - Number of variational transitions with SP intensity above 1×10⁻²⁹ cm/molecule reaching an upper level which is not yet experimentally determined (and then absent from our list). [These 1196 transitions are included and specifically indicated in our variational list (Sup Mat 2)].

3.2 Line intensities

The different sources used for the line intensities are listed in **Table 5** and illustrated in the overview displayed in **Fig. 5**. The line position values calculated from empirical energy levels were combined with FTS or CRDS experimental intensities when available and otherwise the intensities were from variational calculations. For the main isotopologue, FTS intensity values were taken from Toth [2] for the lines with intensities larger than 1×10⁻²³ cm/molecule. It is worth underlying that while Toth's intensity values of the strong lines are known to be underestimated by about 15 % above 8000 cm⁻¹ [25], this is not the case in our region below 8000 cm⁻¹ (the average intensity ratio with recent FTS values by Régalia et al below 8000 cm⁻¹ is 0.99 [38]). In the 1×10⁻²³ – 3×10⁻²⁶ cm/molecule interval, the intensity values reported by Régalia et al. [38] which improve those of Oudot et al. [25], were preferred to those of Toth (included in the HITRAN2012 list). The increased pressure×path length values used by Régalia et al. [38] have allowed for a better determination of the intensities of the medium and weak lines (see below). We also adopted Toth's FTS intensity values [2] for the strongest lines of the H₂¹⁸O, H₂¹⁷O and HD¹⁶O minor isotopologues measured with highly isotopically enriched samples.

Fig. 4

Comparison of the amount of lines “missing” (black dots) in the 2014 list [1] and in the present list for the four major water isotopologues. The “missing” lines are predicted by variational calculations [6, 8] and correspond to transitions reaching an upper level which was not yet experimentally determined.

The intensities of the weaker lines were taken from six of our CRDS studies of natural water spectra [3, 4, 5, 18, 22, 23] and from a recent report by Sironneau and Hodges [26] who measured 71 intensity values by CRDS with accuracy on the order of 0.2% (see below). In all our CRDS analysis, we systematically compared our intensity values to the variational values which are known to be correct *on average* within a few % in our spectral range. From this comparison, a few cases of systematic deviations were observed, which allowed us to correct some of our CRDS intensities. The deviations concern line intensities of minor isotopologues whose sample relative abundances deviate from their natural values or to an inaccurate calibration of the pressure gauge.

The following corrections were performed:

- (a) In Ref. [3], the average intensity ratio of the CRDS and variational intensities of H_2^{16}O was found to be 0.95 in the 7408 – 7920 cm^{-1} region, indicating a systematic error due to the pressure gauge calibration. This systematic bias was confirmed by Sironneau and Hodges, who suggested multiplying our absolute values by a factor 1.044 [26]. This multiplication factor was applied to all the intensities of the 7408-7920 cm^{-1} region,

- (b) As a result of a D depletion in the recordings in the 5850 – 6607 cm⁻¹ [5] and 7911 – 8337 cm⁻¹ [18] region, HD¹⁶O intensities were divided by the factor 0.85,
- (c) As a result of a ¹⁸O enrichment in water sample used for the recordings in the 6886 – 7406 cm⁻¹ [4], H₂¹⁸O intensities were multiplied by a factor 0.80.

Table 5

Origin of the intensity values, corresponding experimental technique and number of transitions included in the water database between 5850 and 8340 cm⁻¹.

Line intensity source	Spectral range (cm ⁻¹)	Source ^a	Number of transitions	Method
Toth [2]	5852 – 7690	Toth	4382	FTS
Régalia et al. [38]	6460 – 8340	14ReOuMi	2121	FTS
Macko et al. [22]	6606 – 6747	04MaRoMi	786	CRDS
Mikhailenko et al. [23]	6745 – 6885	07MiWaKa	372	CRDS
Lisak et al. [39]	7170 – 7183	09LiHaHo	15	CRDS
Mikhailenko et al. [3]	7409 – 7919	11MiKaWa	2923	CRDS
Leshchishina et al. [4]	6886 – 7406	12LeMiMo	1926	CRDS
Leshchishina et al. [5]	5852 – 6606	13LeMiMo	4977	CRDS
Campargue et al. [18]	7911 – 8337	15CaMiLo	2808	CRDS
Sironneau and Hodges [26]	7714 – 7920	15SiHo	71	CRDS
Schwenke and Partridge [8]	5850 – 8340		26176	Variational
Total			46557	

Notes

^a Tag indicating the intensity source used in the spectroscopic database provided as Sup. Mat. 1.

In absence of FTS and CRDS intensity values, the variational intensity values [6-8] were added to the empirical line positions of the not yet measured absorption lines with variational intensity greater than 1×10⁻²⁹ cm/molecule. It leads to a large set of weak transitions displayed on the lower panel of **Fig. 5**. They represent more than half the total number of transitions of the database but less than 0.025% of the sum of the intensities in the region.

Fig. 5.

Sources of the intensities added to the empirical line position values in the water spectroscopic database in the 5850-8340 cm^{-1} range.

Upper panel: FTS data by Toth [2] (black circles) and Régalia et al. [38] (red circles) above 1×10^{-23} cm/molecule and in the $1 \times 10^{-23} - 3 \times 10^{-26}$ cm/molecule range, respectively. Cyan circles correspond to Toth's intensity values of the H_2^{18}O , H_2^{17}O and HD^{16}O minor isotopologues [2].

Middle panel: CRDS data obtained in Grenoble (blue and red circles) [3, 4, 5, 18, 22, 23], by Lisak et al. [39] near 7180 cm^{-1} (dark green dots) and by Sironneau and Hodges [26] near 7800 cm^{-1} (green open circles).

Lower panel: Variational intensities by Schwenke and Partridge [6-8].

Our list of 46557 transitions is provided as Supplementary Material. It includes 6539 FTS, 13878 CRDS and 26140 variational intensities; all associated to line positions calculated from experimental energy values. A sample of the list is presented in **Table 6**.

Table 6

Sample of the spectroscopic database of 46557 water transitions between 5850 and 8340 cm^{-1} as given as Supplementary Material 1. All line positions are empirical values. Intensities are mostly experimental values.

Nu	S	Broadening						E_lower	n	shift				Molec	Upper state				Lower state				Tag	
5855.0698	5.53E-29	0.0673	4	28	0.298	7	57	883.6522	0.57	1	22	0.000183	3	8	H2O17	110	9	5	4	000	8	2	7	13LeMiMo
5855.0987	2.36E-28	0.0920	4	25	0.412	1	71	1908.0163	0.69	1	22	-0.004932	3	7	H2O	050	3	1	2	010	4	2	3	
5855.5417	5.05E-28	0.0746	4	28	0.353	5	3	1128.9384	0.57	1	22	-0.014826	3	8	H2O17	011	9	6	3	000	8	4	4	
5855.6549	6.47E-27	0.0727	4	25	0.265	4	57	2572.1393	0.61	1	22	-0.010628	3	7	H2O	021	8	6	2	010	7	4	3	
5855.6878	5.89E-27	0.0766	7	82	0.375	5	66	1355.1990	0.55	1	22	-0.009200	1	34	H2O18	011	10	6	4	000	9	4	5	13LeMiMo
5855.7635	1.49E-26	0.0261	4	29	0.160	6	57	2254.2833	0.51	1	22	-0.019364	3	9	H2O	110	11	2	2	000	10	8	3	13LeMiMo
-5855.7635		0.0232	4	29	0.298	1	71	2254.2840	0.51	1	22	-0.017001	3	9	H2O	110	11	9	3	000	10	8	2	
5856.1804	3.56E-24	0.0644	5	24	0.307	4	3	927.7439	0.61	1	22	-0.012100	1	6	H2O	011	8	6	3	000	7	4	4	Toth
5856.5418	1.76E-28	0.0915	4	25	0.391	1	71	503.9681	0.67	1	22	-0.009108	3	7	H2O	040	4	0	4	000	5	3	3	13LeMiMo
5856.6289	1.12E-29	0.0778	4	28	0.315	1	71	2998.7660	0.55	1	22	-0.007348	3	8	H2O	040	10	7	4	010	9	4	5	
5857.0873	2.34E-28	0.0620	4	25	0.333	1	71	2764.6973	0.57	1	22	-0.011895	3	7	H2O	200	9	5	4	010	8	4	5	
5857.3142	1.93E-29	0.0497	4	29	0.221	1	71	2631.2671	0.37	1	22	-0.019272	3	9	H2O	110	15	4	11	000	15	1	14	
5857.4110	1.43E-28	0.0872	4	25	0.371	1	71	2271.7120	0.63	1	22	-0.002608	3	7	H2O	050	5	2	3	010	6	3	4	
5857.7870	2.74E-29	0.0569	4	31	0.394	6	57	2102.5542	0.44	1	22	-0.003559	3	11	H2O17	011	13	5	8	000	12	3	9	
5858.2639	2.21E-27	0.0538	4	28	0.329	3	57	2569.5079	0.61	1	22	-0.012364	3	8	H2O	021	8	6	3	010	7	4	4	13LeMiMo
5858.2837	5.39E-29	0.0380	4	28	0.281	1	71	2915.8738	0.47	1	22	-0.012084	3	8	H2O	120	12	3	10	010	11	0	11	
5858.5066	3.56E-29	0.0740	4	28	0.492	3	57	1279.7974	0.55	1	22	-0.005844	3	8	H2O18	110	10	6	5	000	9	3	6	
5858.5843	1.92E-24	0.0560	5	24	0.329	4	3	1216.2311	0.55	1	22	-0.000639	1	6	H2O	011	10	5	6	000	9	3	7	Toth
5858.7432	1.51E-26	0.0090	4	30	0.218	4	57	1327.1096	0.47	1	22	-0.013284	3	10	H2O	011	11	4	8	000	11	0	11	13LeMiMo

Notes:

The different columns are:

- Nu: Empirical line position, cm^{-1} ; In the case of doublets, the line position is duplicated with a minus sign for the second component.
- S: Experimental value of the line intensity (in $\text{cm}/\text{molecule}$ at 296K) corresponding to the reference given in the last column, otherwise variational intensity value from (SP, <http://spectra.iao.ru>); In the case of doublets, no experimental intensity value is given for the second component; the experimental intensity given for the first component corresponds to the sum of the two intensities;
- Broadening: Air- and self-broadening parameters, $\text{cm}^{-1}\text{atm}^{-1}$ (see Section 3.3), ;
- E_lower: Empirical lower energy level, cm^{-1} ;
- n: Temperature dependence parameter of air broadening parameter (see Section 3.3);
- Air shift: Pressure air shift parameter, $\text{cm}^{-1}/\text{atm}$;
- Mol: Isotopologue;
- Upper state: Vibration and rotation numbers of upper state;
- Lower state: Vibration and rotation numbers of lower state;
- Tag: Reference of the experimental line intensity (see Table 5).
- The line shape parameters (broadening and shift) include an uncertainty code and reference following the HITRAN tables (see cfa.harvard.edu/hitran/HITRAN FACTS - uncertainty indices for the uncertainty code and Documentation - Sources of data in HITRAN for the references).

3.3 Line profile parameters

The air-broadened half-width, γ_{air} , its temperature dependence, n_{air} , the air-induced line shift, δ_{air} , and the self-broadened half-width, γ_{self} were added to each transition. These data are from a number of sources as explained below. For the three most abundant isotopologues of water, H_2^{16}O , H_2^{18}O , and H_2^{17}O , γ_{air} , δ_{air} , and γ_{self} were added using a rather sophisticated scheme explained in Ref. [40] that determines and eliminates the experimental outliers and then either uses their averages from the measurement databases [41], if they exist, or values from single measurements, or theoretical complex Robert-Bonamy values (see for instance Refs.[42, 43]), or, if no experimental or theoretical value exists, semi-empirical values from Ref. [44]. When none of the above data exist the half-width data are obtained from J -dependent polynomials [45]. The temperature dependence of the half-width is taken from measured values, if no data exist the data are obtained from a polynomial in J'' that was

developed [46] using the data of Birk and Wagner [47] smoothened and extrapolated to $J'' = 50$. Note, when no data are available for H_2^{18}O or H_2^{17}O , the corresponding data for H_2^{16}O (if they exist) are used.

For the principal deuterated isotopologue, HD^{16}O , an algorithm was developed using the measurement database [41]. When measured HD^{16}O line shape data are available they are added to the transitions or the theoretical values of Gamache and Fischer [48] are used. There are far fewer measurements for HD^{16}O than for the principal isotopologue. To increase the database the following actions were taken. First, the roughly 1500 measured HDO-air half-widths were compared with data from the semi-empirical γ_{air} algorithm data scaled from H_2^{16}O to HD^{16}O and with the smoothened HDO-air half-widths of Toth [49]. Overall the data of Toth gave better agreement: an average percent difference of ~ 20 . These data were used, neglecting vibrational dependence, to help fill in data for transitions for which there were no measurements. Next, the measurement database [41] was taken and the ratio $\gamma_{\text{air}}(\text{H}_2^{16}\text{O})/\gamma_{\text{air}}(\text{HD}^{16}\text{O})$ determined for 921 transitions. In **Fig. 6**, the ratio of the half-widths is plotted *versus* the $\gamma_{\text{air}}(\text{H}_2^{16}\text{O})$ value (blue + symbols), the red horizontal dashed line is the average of the ratios, 0.8502, and the black solid line is a straight line fit of $\text{Ratio} = a + b\gamma_{\text{air}}(\text{H}_2^{16}\text{O})$. The least-squares fit of the data gives $a = 0.3385$ and $b = 7.905 \text{ cm atm}$. The straight-line equation allows the calculation of the scaled HDO value from the H_2O value. Scaling generated more than 1500 additional HDO-air half-widths that were added to the database. For scaling of the line shifts of HDO from H_2O it was found the constant average ratio of 0.7772 determined from 392 transitions worked well. Scaling H_2O line shifts produced over 1000 additional HDO line shifts. Note, there are not enough data to repeat this procedure for n_{air} , or γ_{self} .

Fig. 6.

Ratio $\gamma_{\text{air}}(\text{H}_2\text{O})/\gamma_{\text{air}}(\text{HDO})$ versus $\gamma_{\text{air}}(\text{H}_2\text{O})$ in $\text{cm}^{-1}\text{atm}^{-1}$; dashed red line is the average ratio, solid black line is a straight line fit to the data (see text).

Finally if there are no air-broadened HDO half-width data from the above procedure the data are taken from a polynomial in J'' fit to the J'' averaged data from the database [41] extrapolated to $J'' = 50$.

The temperature dependence of the air-broadened HDO half-width is taken from the measurement database, if available; otherwise the polynomial value, $n_{\text{air}}(\text{H}_2^{16}\text{O})$, is used.

The line shape parameters (γ_{air} , n_{air} , δ_{air} , and γ_{self}) for each line include an uncertainty code and reference following the HITRAN tables (see cfa.harvard.edu/hitran/HITRAN FACTS – uncertainty indices for the uncertainty code and Documentation – Sources of data in HITRAN for the references.)

4. Discussion: experimental *versus* variational line intensities

In spite of considerable progress, variational calculations are not able to reach the typical $1 \times 10^{-3} \text{ cm}^{-1}$ accuracy on transition frequency provided routinely by FTS and CRDS in the region. We are confident that line positions included in the present database represent a significant improvement compared to all other lists for water vapor available in the region but the situation for line intensities is different. The question of the most suitable line intensities to be included in the atmospheric spectroscopic databases such as HITRAN or GEISA is open. Up to now, as a rule, variational intensity values were adopted for the water HITRAN list when experimental values were missing. During the last years, state of the art experimental set ups have been used to determine accurate intensity values of water lines and

variational intensity predictions have been demonstrated to be very good [6, 7, 50-54]. Nevertheless, both in cases of agreements or disagreements, it is always hazardous to generalize to large databases conclusions based on a very limited number of measurements. In the following, we discuss the uncertainty on the CRDS and FTS intensity values of water lines in our spectral region and provide a comparison to variational values available in the literature. These elements of discussion may help for future choice of the best intensity values to be adopted for the water line list in the near infrared.

4.1 Experimental line intensities

The experimental line intensities included in our database were derived from FTS or CRDS spectra using an interactive least squares multi-line fit of the absorption feature assuming a Voigt type for the line shape. The good quality of the reproduction of the CRDS spectra was illustrated in several of our papers (see Fig. 2 of Ref. [5], for instance) but this does not prevent systematic errors on the retrieved intensities related to the pressure value or to deviations from the assumed HITRAN abundance value of the minor isotopologues in the sample. When detected from comparison with average variational values, the CRDS intensity values were corrected accordingly (see Section 3.2).

As a result of the high sensitivity of the CRDS recordings, the spectral congestion is high, about 10 lines/cm⁻¹ and line overlapping is mostly systematic (in particular, at 10 Torr). Weak lines are usually observed on the wings of strong lines, which may have intensity larger by 4 or 5 orders of magnitude (see **Figs. 1 and 2**). In these circumstances, the line parameter retrieval from the multi-line fit of the spectrum was particularly difficult and laborious. Overall, considering the recordings with natural and isotopically enriched samples, line parameters of more than 20000 absorption lines had to be retrieved. It is impossible to provide a comprehensive and transparent uncertainty analysis of the intensities for such a large number of transitions. In the most favorable case of a well-isolated H₂¹⁶O line of intermediate intensity, a 2-5% relative accuracy is a reasonable estimate for the error bar on our CRDS intensities. Thus larger error bars have to be attached to the various components of a blended absorption feature, the relative error on the total intensity of the blend being significantly smaller than that on the individual components.

A few laser studies have been recently dedicated to optimizing the accuracy of the intensity of water transitions in our region. These accurate intensity values are particularly valuable for testing the quality of both extensive FTS and CRDS measurements and variational intensity calculations (see below). In order to lower relative uncertainties below 1%, special experimental efforts are required including very high signal to noise ratio of the recordings with high accuracy of the frequency axes and high linearity of the intensity axes. The preparation and control of the water sample make suitable the implementation of a primary standard humidity generator, mK-level temperature control, and traceable pressure measurements. The simultaneous treatment of various spectra recorded at different pressures

is suitable and the choice of the line shape function is critical. The Voigt type function is known to be insufficient to account for the correct line shape and several mechanisms including collisional narrowing and speed-dependent effects have to be modeled (see for instance Refs. [55-57]). Such demanding efforts cannot be applied to thousands of lines as those measured in our extensive CRDS measurements. The most accurate intensities of water lines in our region were reported in Refs. [26, 39] for relatively strong, isolated lines but the number is limited to a few tens at most. For instance, using frequency stabilized CRDS (FS-CRDS), Lisak et al. [39] reported line parameters of 15 strong lines used in optical hygrometers in the $7170 - 7183 \text{ cm}^{-1}$ range. The intensities reported with a relative combined uncertainty better than 0.4% are incorporated in our database (see **Fig. 5**). An average value of 0.991(22) was obtained for the FS-CRDS/HITRAN intensity ratio [39]. This good agreement indicates that the 2-5% error bar attached in the HITRAN database to the Toth's intensity values of the considered strong lines is conservative. Of particular relevance is the work of Sironneau and Hodges (SH) who reported very accurate line shape analysis and spectroscopic parameters for about 70 lines by FS-CRDS in the $7710 - 7920 \text{ cm}^{-1}$ region [26]. The relative uncertainties on the line intensities derived using the partially correlated, quadratic-speed-dependent Nelkin-Ghatak profile was 0.2%, which, to the best of our knowledge, is the smallest intensity uncertainty reported so far for near infrared water line intensities. This study is particularly valuable because the reported line intensities vary from 6.8×10^{-27} to $1.2 \times 10^{-24} \text{ cm/molecule}$, which makes possible a comparison over a significant intensity range. **Fig. 7** shows the ratio of intensities from the literature to SH values for the 71 lines reported in Ref. [26], *versus* the line intensity. The comparison with FTS intensities by Toth included in HITRAN and the recent FTS measurements performed by Régalia et al. [38] show good average values of 1.034 and 0.990, respectively with a larger dispersion for Toth's values in particular for the weakest lines (standard deviations of 8.5 and 6.8%, respectively). This is one of the reasons why we adopted Régalia's FTS values instead of Toth's values in the $1 \times 10^{-23} - 3 \times 10^{-26} \text{ cm/molecule}$ intensity interval in our database. Sironneau and Hodges performed a similar comparison and noted "the smallest standard deviation of the fractional intensity differences (3%) occurs when comparing to the CRDS measurements by Mikhailenko et al." [3] and that "the most comprehensive line list in the region could be realized by multiplying the reported intensities by a factor 1.044" – what was done as explained in Section 3.2 and corresponds to the CRDS data displayed in **Fig. 7**. Even if only partially representative (for instance, the range of SH line intensities corresponds to the strongest lines that we measured by CRDS and SH lines are mostly isolated lines, all belonging to the main isotopologue), this comparison indicates that the uncertainty on the FTS values by Régalia et al. [38] and our CRDS values is generally better than 5% for non blended lines of H_2^{16}O .

Fig. 7.

Ratio of the experimental and variational intensities to the FS-CRDS values reported by Sironneau and Hodges [26] with 0.2% relative uncertainty, for 71 H_2^{16}O lines in the 7710 – 7920 cm^{-1} region.

Upper panels, from left to right: FTS by Toth [2] included in the HITRAN database, FTS by Régalia et al. [38], CRDS by Mikhailenko et al. [3] with a correction factor of 1.044 suggested by Sironneau and Hodges [26],

Lower panels, from left to right: Variational calculations by Schwenke and Partridge (SP) [7, 8], Barber, Tolchenov and Tennyson (BT2) [52], and Polyansky et al. (UCL2016) [58]. The red open circle marks the $4_{23} - 4_{14}$ transition of the $5\nu_2$ band of H_2^{16}O (see Text).

(The three arrows indicate data out of frame).

4.2 Variational line intensities

We have included in **Fig. 7** the ratio between variational intensities and SH values. We consider 3 sources: variational line list computed by S.A. Tashkun [8] using the results of the variational calculations by Schwenke and Partridge [6, 7] (SP), calculations by Barber et al. [52] (BT2) and very recent line list from University College of London provided by Polyansky et al. [58] (UCL2016). The overall agreement is very good. We note that SP and BT2 intensity values are in general very close because they use the same underlying dipole moment surface and that the dispersion of the ratios is more limited than for the experimental values. The improvement of the UCL2016 list compared to SP and BT2 is not clear. For the largest intensities, there is a clear propensity of the UCL2016 ratios to be distributed around the 1.0 and 0.96 values. This observation was previously seen for BT2 values by Sironneau and Hodges [26] who noted that the two mostly constant 1.0 and 0.96 values correspond to the $\nu_1 + \nu_3$ and $2\nu_3$ vibrational bands, respectively. Surprisingly, our SH/BT2 intensity ratios do

not show this clear vibrational dependence. It may indicate that the BT2 intensity values used by Sironneau and Hodges differ from our values which were obtained from the UCL website indicated in Ref. [52].

A similar variational dependence of the variational intensity ratios was also seen and discussed for the case of the $\nu_1+3\nu_2$ band of H_2^{16}O . Systematic overestimations by 8% and 6% were found for SP and BT2 intensity values, respectively [18].

Another interesting observation concerns the line at 7770.3732 cm^{-1} (intensity = $1.19\times 10^{-26}\text{ cm/molecule}$ [26]) whose variational intensity shows the largest deviation compared with the SH intensity value (red open circle in **Fig. 7**). This line is assigned to the $4_{23} - 4_{14}$ transition of the $5\nu_2$ band of H_2^{16}O and corresponds to an intensity ratio of 0.44, 0.78 and 0.89 for SP, BT2 and UCL2016, respectively. These values indicate that the experimental intensity of this transition – which is confirmed by our CRDS value (ratio=1.0098) – is poorly predicted by theory and that UCL2016 improves significantly the prediction. This improvement probably reflects the better potential energy surface used for the most recent UCL calculations [59]. Indeed, the transitions belonging to the purely bending vibrational bands such as the $4\nu_2$ and $5\nu_2$ bands located in our region are known to have intensity very sensitive to local resonances. For instance, we have seen several examples of transitions of the $4\nu_2$ band near 6350 cm^{-1} with variational intensities overestimated by a factor of 2 compared with the measurements [5].

The quality of the variational intensities depends on the accuracy of the nuclear motion wavefunctions computed from the potential energy surface (PES) and of the dipole moment surface (DMS). As a rule, systematic deviations on the order of a few percent associated with some vibrational bands are due to inaccuracies of the DMS while larger and irregular deviations are mostly related to inaccuracy of the wavefunctions that do not account satisfactorily for some resonance interaction. The resulting uncertainty on the computed intensities may be very large in the case of an intensity transfer to a dark state. In those situations, the calculated intensity value is very sensitive to small changes in the PES and DMS. The above example of the $4_{23} - 4_{14}$ $5\nu_2$ transition and the $4\nu_2$ transitions discussed in Ref. [5] are examples of these “unstable” transitions (see Ref. [51], for more details about unstable transitions in H_2^{17}O and H_2^{18}O). This is reflected by the important difference between SP, BT2 and UCL2016 intensities, the best value being still 11% below the SH experimental value. Note that the HITRAN database uses the BT2 intensity with an error code 4 (between 10% and 20%).

In summary, the above comparison with state of the art intensity measurements indicates that at present theory is not yet able to provide intensities with a uniform accuracy better than a few percent in our region. Nevertheless, since the pioneering extensive

calculations by Partridge and Schwenke [6], the average accuracy of variational intensity in our spectral region is known to be very good. This is why we provide as a second Supplementary Material, three line lists for natural water in the 5850 – 8340 cm^{-1} region, with our empirical transition frequencies combined with variational intensities (1×10^{-29} $\text{cm}/\text{molecule}$ intensity cut off):

(i) intensities by Schwenke and Partridge for the six isotopologues as calculated by Tashkun [8];

(ii) intensities by Barber et al. [52] (BT2) for H_2^{16}O , by Lodi and Tennyson (LT) for H_2^{18}O and H_2^{17}O [51] and by Voronin et al. [53] (VTT) for HD^{16}O ;

(iii) very recent intensity values obtained at University College London for H_2^{16}O , H_2^{18}O and H_2^{17}O [58];

The construction of these lists uses the rovibrational quantum numbers of the lower and upper levels to match the variational intensities to the empirical frequencies. This step was not straightforward in the case of the BT2, VTT, LT, and UCL2016 variational lists because of missing or distinct vibrational assignments and incomplete rotational assignments.

Note that for unknown reasons, about 500 transitions of H_2^{16}O , H_2^{18}O and H_2^{17}O are missing in the UCL2016 list, about 50 of them with intensities larger than 1×10^{-25} $\text{cm}/\text{molecule}$. The two other lists are complete in the sense that, for each isotopologue, all the lines with variational intensities above the chosen intensity cut off are included, in particular the weak lines for which an empirical value of the upper state energy is not yet experimentally determined. These lines are provided with SP variational line center and are indicated in Sup Mat 2. They are absent in Sup Mat 1 (see **Fig. 4**). Line broadening parameters as calculated in Section 3.3 are incorporated. A sample of the lists provided as Supplementary Material 2 is presented in **Table 7**.

Fig. 8.

Comparison of experimental line positions H_2^{16}O as included in the database and their variational values between 5850 and 8340 cm^{-1} .

Upper panels: differences of the line positions *versus* the positions SP [8], BT2 [52] and UCL2016 [58].

Lower panels: corresponding histograms

Table 7.

Sample of the spectroscopic database of 48009 water transitions between 5850 and 8340 cm^{-1} as given as Supplementary Material 2. Except for 1196 variational positions [7,8], all line positions are empirical values. Intensities are variational values from various sources.

Nu	S	Broadening	E_lower	n	Air shift	UCL-2016	BT2,L&T,VTT	Molec	Upper state	Lower state	Comment
6482.0908	2.705E-26	0.0930 8 77 0.411 5 57	29.8086	0.73 1	22-0.007000 2 38		2.72E-26	HDO	021 2 1 2	000 1 1 1	
6482.1197	1.966E-29	0.0343 4 31 0.265 1 71	2813.5273	0.44 1	22-0.023920 3 11	1.90E-29	1.97E-29	H2O	101 12 4 9	000 12 8 4	
6482.1467	1.654E-29	0.0727 4 28 0.352 1 71	2572.1393	0.61 1	22-0.006996 3 7	1.54E-29	1.60E-29	H2O	050 6 5 2	010 7 4 3	
6482.2818	4.216E-29	0.0856 4 36 0.391 1 71	1630.4932	0.67 1	22 0.000000 0 0		4.29E-29	HDO	111 6 1 6	010 5 1 5	
6482.2840	3.141E-26	0.1070 6 77 0.457 5 57	308.6157	0.63 1	22-0.008000 1 38		2.55E-26	HDO	101 6 2 5	000 6 1 6	
6482.3136	1.040E-29	0.0636 4 29 0.281 1 71	4038.4028	0.47 1	22-0.011120 3 9	1.04E-29	1.03E-29	H2O	111 10 5 5	010 11 7 4	
6482.4037	3.731E-27	0.0814 1 33 0.412 1 71	295.4873	0.69 1	22-0.014671 1 1		3.32E-27	HDO	101 5 3 3	000 4 3 2	
6482.4182	3.321E-26	0.0880 1 33 0.480 1 71	15.5082	0.73 1	22 0.000064 1 1		3.36E-26	HDO	021 2 0 2	000 1 0 1	
6482.4551	1.361E-29	0.0261 4 25 0.298 1 71	4260.3522	0.51 1	22-0.016152 3 7	1.38E-29	1.36E-29	H2O	041 9 0 9	020 10 0 10	
6482.4662	2.541E-28	0.0710 4 25 0.371 1 71	2271.7120	0.63 1	22-0.003644 3 7	2.43E-28	2.58E-28	H2O	050 6 4 3	010 6 3 4	
6482.6734	9.824E-29	0.0940 4 36 0.371 1 71	469.6636	0.63 1	22 0.000000 0 0		9.29E-29	HDO	021 7 2 6	000 6 3 3	
6482.6820	5.415E-29	0.0869 4 28 0.333 1 71	711.7953	0.57 1	22 0.000000 0 0		5.47E-29	HDO	021 9 1 8	000 8 3 5	
6482.6881	3.420E-29	0.0637 4 28 0.352 1 71	921.8957	0.61 1	22-0.008116 3 8	3.32E-29	3.40E-29	H2O18	040 6 5 1	000 7 4 4	
6482.7007	6.527E-28	0.0812 4 29 0.498 4 57	1538.1488	0.51 1	22 0.001876 3 8	6.00E-28	6.61E-28	H2O	040 10 4 6	000 10 3 7	
6482.7777	1.672E-29	0.0979 4 25 0.470 6 3	46.0284	0.72 1	22-0.000476 3 7			HD17O	021 3 0 3	000 2 0 2	
6482.8447	7.836E-27	0.0388 4 29 0.347 3 57	2813.5114	0.44 1	22-0.021260 3 9	7.86E-27	7.83E-27	H2O	101 11 6 6	000 12 8 5	
6482.8551	6.114E-27	0.0348 4 29 0.250 1 71	2426.1953	0.41 1	22-0.026080 3 9	6.03E-27	6.08E-27	H2O	200 12 1 11	000 13 4 10	
6482.8709	4.173E-27	0.0780 6 77 0.431 5 57	156.3823	0.69 1	22-0.012000 0 38		4.16E-27	HDO	021 4 1 3	000 4 1 4	
6482.9306	1.450E-28	0.0236 4 29 0.235 1 71	3922.3294	0.40 1	22-0.024760 3 9	1.43E-28	1.44E-28	H2O	200 13 9 4	000 14 10 5	
6482.9307	4.835E-29	0.0236 4 29 0.235 1 71	3922.3293	0.40 1	22-0.024760 3 9	4.76E-29	4.83E-29	H2O	200 13 9 5	000 14 10 4	
6482.9806	4.595E-29	0.0850 6 24 0.514 6 3	107.6303	0.72 1	22-0.006060 1 6			HD18O	021 3 2 2	000 2 2 1	
6483.0315	9.973E-28	0.0382 4 28 0.298 1 71	2904.4280	0.51 1	22-0.020336 3 8	1.02E-27	1.01E-27	H2O	130 9 1 8	010 10 2 9	
6483.0567	3.731E-27	0.0740 5 77 0.437 4 57	295.6775	0.69 1	22-0.014000 0 38		3.32E-27	HDO	101 5 3 2	000 4 3 1	
6483.0697	2.122E-29	0.0224 4 36 0.208 1 71	1839.7590	0.35 1	22 0.000000 0 0		2.06E-29	HDO	210 15 0 15	000 16 1 16	var.pos.
6483.1483	3.539E-29	0.0699 4 29 0.265 1 71	2105.8666	0.44 1	22 0.002704 3 9	3.21E-29	3.60E-29	H2O	040 12 4 8	000 12 3 9	
6483.1641	2.778E-27	0.1010 6 77 0.411 6 57	576.9047	0.61 1	22-0.005000 0 38		2.79E-27	HDO	021 7 3 4	000 7 3 5	
6483.1766	1.120E-28	0.0641 4 29 0.281 1 71	2144.0458	0.47 1	22-0.027160 3 9	1.21E-28	1.12E-28	H2O	120 11 3 8	000 11 6 5	
6483.1836	4.074E-29	0.0313 4 25 0.298 1 71	4260.4672	0.51 1	22-0.016580 3 7	4.12E-29	4.08E-29	H2O	041 9 1 9	020 10 1 10	
6483.1847	2.093E-29	0.0224 4 36 0.208 1 71	1839.7560	0.35 1	22 0.000000 0 0		7.15E-30	HDO	210 15 1 15	000 16 0 16	
6483.2876	1.025E-29	0.0893 4 25 0.412 1 71	3482.0646	0.69 1	22-0.010952 3 7	1.10E-29	1.05E-29	H2O	140 4 1 4	020 4 2 3	
6483.3562	2.925E-26	0.0899 4 36 0.351 4 57	157.0647	0.71 1	22 0.000000 0 0		2.81E-26	HDO	101 3 3 0	000 3 2 1	
6483.4938	1.002E-28	0.1020 6 77 0.399 6 57	90.7654	0.71 1	22-0.005000 0 38			HD18O	021 4 0 4	000 3 0 3	
6483.5420	1.803E-28	0.0839 1 33 0.371 1 71	384.8756	0.63 1	22-0.002186 0 1		8.20E-29	HDO	021 6 2 4	000 6 2 5	

Notes:

The different columns are:

- Nu: Empirical line position, cm^{-1} ; if no comment in the last column otherwise variational line position (SP, <http://spectra.iao.ru>);
- S: Variational (SP, <http://spectra.iao.ru>) line intensity, cm/molecule at 296K;
- Broadening: Air- and self-broadening parameters, $\text{cm}^{-1}\text{atm}^{-1}$ (see Section 3.2);
- E_lower: Empirical lower energy level, cm^{-1} ;
- n: Temperature dependence parameter of air broadening parameter (see Section 3.3);
- Air shift: Pressure air shift parameter, $\text{cm}^{-1}/\text{atm}$;
- UCL2016: line intensity, cm/molecule at 296K calculated at UCL for London for H_2^{16}O , H_2^{18}O and H_2^{17}O [58];
- BT2,L&T,VTT: line intensity, cm/molecule at 296K calculated by Barber et al. [52] (BT2) for H_2^{16}O , by Lodi and Tennyson (LT) for H_2^{18}O and H_2^{17}O [51] and by Voronin et al. [53] (VTT) for HD^{16}O ;
- Molec: Isotopologue;
- Upper state: Vibration and rotation numbers of upper state;
- Lower state: Vibration and rotation numbers of lower state;
- Comment: "variational position" indicates the variational line position from SP, <http://spectra.iao.ru> which are given for the 1196 transitions reaching an upper energy level which was not yet experimentally determined.

The line shape parameters (broadening and shift) include an uncertainty code and reference following the HITRAN tables (see cfa.harvard.edu/hitran/HITRAN FACTS - uncertainty indices for the uncertainty code and Documentation - Sources of data in HITRAN for the references).

Fig. 9.

Comparison of the intensities of the H_2^{16}O lines between 5850 and 8340 cm^{-1} .

Upper panels: ratios of the variational intensities to the experimental intensities of the database (FTS and CRDS, blue and red circles, respectively): SP [8], BT2 [52] and UCL2016 [58].

Lower panels: ratios of the different variational intensities.

The combination of empirical line positions and variational intensities is a powerful way to improve and complete spectroscopic databases beyond the observations. It is nevertheless important to note that shifting a variational line position while keeping unchanged the variational intensity may induce some errors. Indeed, as intensity transfers depend on the separation of the energy levels, larger errors on the variational intensities are expected for variational energy levels deviating importantly from their experimental values. These types of errors (which are difficult to quantify) are probably more pronounced in SP calculations, the water PES having been significantly improved by the recent UCL PES [59]. Surprisingly, in the considered spectral region, this improvement is not reflected by the overview of the deviations between the variational and experimental positions (**Fig. 8**)

Fig. 9 shows an overview of the ratios of the H_2^{16}O experimental intensities to the SP, BT2 and UCL2016 variational values for the whole list in the 5850 – 8340 cm^{-1} . At the scale of the plots, no clear tendency is observed, the largest dispersion of the ratios for low intensities being dominated by the larger uncertainties on the experimental values. An even larger dispersion is expected for the minor isotopologues whose intensities are less accurate because they were partly retrieved from weak absorption lines in natural water spectra. As mentioned above, for the minor isotopologues, an additional error source is due to possible isotopic fractionation leading to significant deviations between the sample isotopic abundance

and the “natural” abundance value adopted in the HITRAN database. Obviously, theory has the advantage of providing intensity predictions of similar quality for all the isotopologues.

5. Concluding remarks

As a result of long term experimental efforts initiated in 2004, we have constructed an improved spectroscopic database for natural water at 296 K covering a significant part of the near infrared region (5850 – 8340 cm^{-1} or 1.71 – 1.20 μm). A large part of the spectroscopic parameters and assignments included in the database relies on the analysis of high sensitivity CRDS spectra recorded using either DFB diode lasers or an ECDL. Our list (Sup. Mat. 1) includes 46557 lines of the six most abundant water isotopologues (H_2^{16}O , H_2^{18}O , H_2^{17}O , HD^{16}O , HD^{17}O and HD^{18}O) with intensity greater than 1×10^{-29} $\text{cm}^2/\text{molecule}$. All the experimental positions included in the list have experimental accuracy (better than 2×10^{-3} cm^{-1}). Their accuracy and completeness are believed to represent an important improvement and update compared to literature results, in particular to the present version of the water list in the HITRAN database (see below) and to the energy levels reported by an IUPAC-TG, a few years ago. Line intensities are experimental values obtained by FTS or CRDS when available. Otherwise, we adopted the variational values by Schwenke and Partridge for a large set of more than 24400 weak and very weak lines that were not measured so far but represent in total no more than 0.025% of the total absorption in the region.

We have used the recent set of 71 line intensities reported by Sironneau and Hodges with a 0.2% relative accuracy [26] to illustrate the advantages and drawbacks of experimental and variational intensities. While, in the case of the water molecule, theory does not compete with experiment for line positions, variational calculations provide line intensities with few percent accuracy for the majority of the transitions of all the isotopologues. The well-known importance of resonance interaction may nevertheless induce larger errors for a small fraction of the lines, the so-called “unstable” lines which are very sensitive to small changes of the PES used to compute the variational line list. On the experimental side, several uncertainty sources including pressure, temperature, isotopic abundance, presence of impurity lines and data treatment (line overlapping, line profile issue...) do not allow one to achieve a few % accuracy on intensity values for a set of thousands of absorption lines. In those circumstances, a future improvement of our line list may consist in using the variational intensity values at least for the weak lines, in particular those of the minor isotopologues, which are frequently overlapped by much stronger H_2^{16}O lines. For the stronger lines, a detailed band-by-band analysis should help to identify systematic over- or underestimation of the variational intensities, which could be corrected using experimental results. A procedure proposed in Ref. [51] may allow identifying “unstable” lines having less accurate variational intensity. The corresponding experimental intensity should be preferred for these unstable lines. In summary, even if the line list constructed in this work (Sup. Mat 1) represents significant progress, there is probably room for further improvement of the accuracy of the intensities by

combining experimental and variational intensities. On the experiment side, a significant improvement on the line centers accuracy can now be routinely achieved over wide spectral regions by combining a self-referenced frequency comb with a CRDS spectrometer. For instance, a sub-MHz accuracy ($\sim 10^{-5} \text{ cm}^{-1}$) was recently demonstrated for CO transitions of the $3 - 0$ band over the $6172.5 - 6418.0 \text{ cm}^{-1}$ interval, without any degradation of the data acquisition rate [60].

Finally, we would like to present some examples of comparison to the water list in the 2012 HITRAN database [11]. Our list includes a total of 46557 transitions while the HITRAN list [11] includes 34541 transitions, the main difference being due to the absence of many HD^{16}O lines in the HITRAN list (see **Fig. 3** and **Table 8**). It is worth noting that the total sum of the line intensities is identical in the two databases (6.7480×10^{-19} and $6.7497 \times 10^{-19} \text{ cm/molecule at 296 K}$), as both databases use Toth's intensity data for the stronger lines. Our line positions relying on experimental energy levels has a more homogeneous quality those of the HITRAN list, which mixes experimental and variational line positions. In addition, about 3300 transitions of the HITRAN2012 line list from BT2 or LT variational calculations are provided without vibrational and with incomplete rotational labeling. They are listed with an artificial $(-2-2-2; J K_a K_c)$ assignment of the upper state where K_a and K_c are in fact indexes of symmetry (1 – 4) accompanied with a negative sign [61].

Table 8

Number of transitions in the 5850 – 8340 cm^{-1} range for the six major isotopologues of water included in our list (Sup Mat 1) and in the HITRAN list [11].

	This work	HITRAN [11]
H_2^{16}O	21243	20279
H_2^{18}O	7213	7579
H_2^{17}O	5095	5353
HD^{16}O	11018	1330
HD^{18}O	1384	0
HD^{17}O	604	0
Total	46557	34541

As mentioned in the Introduction, our line list incorporates a large amount of information retrieved from high sensitivity CRDS spectra. The detectivity threshold of the recordings allowed measuring lines with intensity at the level of the intensity cut off i.e. three orders of magnitude lower than the smallest experimental intensity included in the HITRAN database in the region. Thus, the CRDS spectra at disposal allow for a critical test of the large set of lines with intensities in the $10^{-29} - 10^{-26}$ cm/molecule range included (or missing) in the HITRAN database (see **Fig. 1**). The inclusion of the line profile parameters in our list (Section 3.3) allows for performing simulations for systematic comparison to the CRDS spectra in the experimental conditions of the recordings. Considering the variety of sources used for the HITRAN list of water in the region, we will not repeat here the more systematic and detailed discussion of the weaknesses of the HITRAN list, which was presented in Refs. [1, 5, 18]. We just illustrate the main issues in **Figs. 10-12** where the CRDS spectrum is compared with simulations based on HITRAN and on our lists: (i) absence of many HD^{16}O transitions which contribute greatly to the absorption in the 1.6 and 1.25 μm windows (**Figs. 10-12**), (ii) inaccurate line positions with variational, IUPAC-TG or untraceable origin (**Fig. 11**) and (iii) inaccurate line intensities from experiment, variational or untraceable origin (**Fig. 12**).

As a final conclusion, we would like to temper our criticism of the present HITRAN water list by mentioning that, except for the missing HDO lines that are relatively strong in the transparency windows, most of the other issues concern line of intermediate or small intensities with small or negligible impact for most of atmospheric applications. In addition, most of the inaccuracies illustrated in **Figs. 10-12** are consistent with the HITRAN error bars attached to the problematic positions or intensities. Our CRDS sensitivity and detailed analysis have allowed us to detect and hopefully correct most of these problems. We hope that our work will help to further improve the HITRAN database in a spectral range of particular importance for many applications.

Figure 10

Examples of spectral interval where HDO lines are missing in the HITRAN database

Upper panels: Comparison between the CRDS spectrum of water at 10.0 Torr and simulation using our spectroscopic database

Lower panels: Corresponding simulation using the HITRAN2012 spectroscopic database

On the right-hand panel, the problematic line with inaccurate line position near 7872.5 cm^{-1} corresponds to the $11_{29} - 12_{112}$ transition of the $5\nu_2$ band of H_2^{16}O . Its HITRAN position was calculated from an inaccurate upper energy level.

Figure 11

Examples of inaccurate line parameters in the HITRAN database

Left hand panels: The problematic line corresponds to the $12_{310} - 11_{011}$ transition of the $2v_3$ band of H_2^{16}O . The origin of its HITRAN position and intensity are unclear (HITRAN1986 is given as source for both).

Right hand panels: The $9_{81} - 8_{72}$ transition of the $2v_3$ band of H_2^{16}O is observed shifted by about 0.43 cm^{-1} from its HITRAN position due to BT2 variational calculations [52].

Figure 12

Examples of inaccurate line intensity in the HITRAN database

Left hand panels: The $4_{23} - 3_{12}$ transition of the $4v_2$ band of H_2^{16}O is observed with an intensity three times smaller than BT2 variational value adopted in the HITRAN database (2.06×10^{-27} [5] and 6.37×10^{-27} cm/molecule [52], respectively).

Right hand panels: The problematic line near 8002.72 cm^{-1} corresponds to the $11_{66} - 10_{29}$ transition of the $\nu_1 + \nu_3$ band of H_2^{16}O . Its HITRAN position relies on the IUPAC-TG energy levels [16]. The origin of its HITRAN intensity is unclear (HITRAN1986 is given as source).

Acknowledgements

This work is jointly supported by CNRS (France) and RFBR (Russia) in the frame of the International Associated Laboratory SAMIA. This project takes place in the frame of the LabexOSUG@2020 (ANR10 LABX56). RRG is pleased to acknowledge support of this research by the National Science Foundation through Grant No. AGS-1156862. We thank OL Polyansky (University College London) for kindly providing us the most recent UCL variational line lists of H_2^{16}O , H_2^{18}O and H_2^{17}O before publication. S. Béguier (LIPhy Grenoble) is thanked for the development of the software used for the spectra simulations.

References

1. Mikhailenko SN, Mondelain D, Kassi S, Campargue A. An accurate and complete empirical line list for water vapor between 5850 and 7920 cm^{-1} . J Quant Spectrosc Radiat Transfer 2014;140:48-57.
2. Toth RA. Line list of water vapor parameters from 500 to 8000 cm^{-1} . <http://mark4sun.jpl.nasa.gov/h2ohtml>
3. Mikhailenko S, Kassi S, Le Wang, 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.
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 Transfer 2012;113:2155-66.
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 Transfer 2013;130:69-80.
6. 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.
7. 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.
8. <http://spectra.iao.ru/1536x765/en/mol/survey/1/>
9. Jacquinet-Husson N, Armante R, Crépeau N, Chédin A, Scott NA, Boutammine C, et al. The 2015 edition of the GEISA spectroscopic database. J Mol Spectrosc 2016 (submitted).
10. Jacquinet-Husson N, Crepeau N, Armante R, Boutammine C, Chédin A, Scott NA et al. The 2009 edition of the GEISA spectroscopic database. J Quant Spectrosc Radiat Transfer 2011;112:2395-445.
11. Rothman LS, Gordon IE, Babikov Y, Barbe A, Benner DC, Bernath P.F, et al. The HITRAN2012 molecular spectroscopic database. J Quant Spectrosc Radiat Transfer 2013;130:4-50.
12. Furtenbacher T, Császár AG, Tennyson J. MARVEL: measured active rotational-vibrational energy levels. J Mol Spectrosc 2007;245:115-25.
13. Furtenbacher T, Császár AG. MARVEL: measured active rotational-vibrational energy levels. II. Algorithmic improvements. J Quant Spectrosc Radiat Transfer 2012;113:929-35.
14. 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 Transfer 2009;110:573-96.
15. 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 Transfer 2010;111:2160-84.
16. 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 Transfer 2013;117:29-58.
17. 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 Transfer 2014;142:93-108.
18. 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 Transfer 2015;157:135-52.
19. 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 Transfer 2014;138:97-106.
20. Liu A, Naumenko O, Kassi S, Campargue A. High sensitivity CW-CDRS of ^{18}O enriched water near 1.6 μm . J Quant Spectrosc Radiat Transfer 2009;110:1781-800.
21. Mikhailenko SN, Leshchishina O, Karlovets EV, Mondelain D, Kassi S, Campargue A. CRDS of ^{17}O enriched water between 5850 and 6671 cm^{-1} : more than 1000 energy levels of H_2^{17}O and HD^{17}O newly determined. J Quant Spectrosc Radiat Transfer <http://dx.doi.org/10.1016/j.jqsrt.2015.11.016>
22. Macko P, Romanini D, Mikhailenko SN, Naumenko OV, Kassi S, Jenouvrier A et al. 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.
23. Mikhailenko SN, Le Wang, 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.
 24. Kassi S, Campargue A. Cavity Ring Down Spectroscopy with $5 \times 10^{-13} \text{ cm}^{-1}$ sensitivity *J Chem Phys* 2012;137:234201.
 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.
 26. Sironneau VT, Hodges JT. Line shapes, positions and intensities of water transitions near 1.28 μm . *J Quant Spectrosc Radiat Transfer* 2015;152:1-15.
 27. Mellau GCh, Mikhailenko SN. Rotational levels of five lowest states of H_2^{18}O and H_2^{17}O from hot emission spectra between 320 and 2500 cm^{-1} , in preparation.
 28. Liu A-W, Song K-F, Ni H-Y, Hu S-M, Naumenko OV, Vasilenko IA, Mikhailenko SN. (000) and (010) energy levels of the HD^{18}O and D_2^{18}O molecules from analysis of their ν_2 bands. *J Mol Spectrosc* 2011;265:26-38.
 29. Mellau GCh, Mikhailenko SN, Tyuterev VIG. Hot water emission spectra: Rotational energy levels of the (000) and (010) states of HD^{17}O . *J Mol Spectrosc* 2015;308-309:6-19.
 30. Mikhailenko SN, Naumenko OV, Nikitin AV, Vasilenko IA, Liu A-W, Song K-F et al. Absorption spectrum of deuterated water vapor enriched by ^{18}O between 6000 and 9200 cm^{-1} . *J Quant Spectrosc Radiat Transfer* 2012;113:653-69.
 31. 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.
 32. 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.
 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.
 34. 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.
 35. Tereszchuk K, Bernath PF, Zobov NF, Shirin SV, Polyansky OL, et al. Laboratory spectroscopy of hot water near 2 microns and sunspot spectroscopy in the *H*-band region. *Astrophys J* 2002;577:496-500.
 36. Petrova T, Poplavskii Yu, Serdyukov V, Sinitsa L. Intracavity laser spectroscopy of high-temperature water vapour in the range 9390-9450 cm^{-1} . *Mol Phys* 2006;104:2691-2700.
 37. Tolchenov RN, Tennyson J. Water line parameters for weak lines in the range 7400–9600 cm^{-1} . *J Mol Spectrosc* 2005;231:23-7.
 38. 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 Transfer* 2014;136:119-36.
 39. Lisak D, Havey DK, Hodges JT. Spectroscopic line parameters of water vapor for rotation-vibration transitions near 7180 cm^{-1} . *Phys Rev A* 2009;79:052507.
 40. Gordon IE, Rothman LS, Gamache RR, Jacquemart D, Boone C, Bernath PF, et al. Current updates of water vapor linelist in HITRAN2004: A new “Diet” for air-broadened half-widths. *J Quant Spectrosc Radiat Transfer* 2007;108:389-402.
 41. Gamache RR, Hartmann J-M. An intercomparison of measured pressure-broadening and pressure-shifting parameters of water vapor. *Can J Chem* 2004;82:1013-27.
 42. Gamache RR, Laraia AL. N_2 -, O_2 -, and air-broadened half-widths, their temperature dependence, and line shifts for the rotation band of H_2^{16}O . *J Mol Spectrosc* 2009;257:116-27.
 43. Gamache RR. Line Shape parameters for water vapor in the 3.2 to 17.76 μm region for atmospheric applications. *J Mol Spectrosc* 2005;229:9-18.
 44. Jacquemart D, Gamache R, Rothman LS. Semi-empirical calculation of air-broadened half-widths and air pressure-induced frequency shifts of water-vapor absorption lines. *J Quant Spectrosc Radiat Transfer* 2005;96:205-39.
 45. Rothman LS, Gordon IE, Barber RJ, Dothe H, Gamache RR, Goldman A, et al. HITRAN, the High-Temperature Molecular Spectroscopic Database. *J Quant Spectrosc Radiat Transfer* 2010;111:2139-50.
 46. Gamache RR. Unpublished results. 2014, University of Massachusetts Lowell. [The reference is not cited in the text.]
 47. Birk M, Wagner G. Temperature-dependent air broadening of water in the 1250-1750 cm^{-1} range. *J Quant Spectrosc Radiat Transfer* 2012;113:889-928.

48. Gamache RR, Fischer J. Half-widths of H_2^{16}O , H_2^{18}O , H_2^{17}O , HD^{16}O , and D_2^{16}O : I Comparison between isotopomers. *J Quant Spectrosc Radiat Transfer* 2003;78:289-304.
49. Toth RA. Smoothed HDO half-width data - Line lists of water vapor parameters for ν_2 . see <http://mark4sun.jpl.nasa.gov.data/spec/HDO>. Files HDOWID.AIR and HDOSHFT.AIR. 2004 Available from: <http://mark4sun.jpl.nasa.gov.data/spec/HDO>.
50. Lodi L, Tennyson J, Polyansky OL. A global, high accuracy *ab initio* dipole moment surface for the electronic ground state of the water molecule. *J Chem Phys* 2011;135:034113.
51. 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 Transfer* 2012;113:850-8.
52. Barber RJ, Tennyson J, Harris GJ, Tolchenov RN. A high-accuracy computed line list. *Mon Not R Astron Soc* 2006;368:1087-94.
53. Voronin BA, Tolchenov RN, Tennyson J, Lugovskoy AA, Yurchenko SN. A high accuracy computed line list for the HDO molecule. *Mon Not R Astr Soc* 2009;402:492-6.
54. Shirin SV, Zobov NF, Ovsyannikov RI, Polyansky OL, Tennyson J. Water line lists close to experimental accuracy using a spectroscopically determined potential energy surface for H_2^{16}O , H_2^{17}O , and H_2^{18}O . *J Chem Phys* 2008;128:224306.
55. Ngo NH, Tran H, Gamache RR. A pure H_2O isolated line-shape model based on classical molecular dynamics simulations of velocity changes and semi-classical calculations of speed-dependent collisional parameters. *J Chem Phys* 2012;136:154310.
56. Tennyson J, Bernath PF, Campargue A, Csaszar AG, Daumont L, Gamache RR, et al. Recommended isolated-line profile for representing high-resolution spectroscopic transitions (IUPAC Technical Report). *Pure Appl Chem* 2014;86:1931-43.
57. Lisak D, Cygan A, Bermejo D, Domenech JL, Hodges JT, Tran H. Application of the Hartmann-Tran profile to analysis of H_2O spectra. *J Quant Spectrosc Radiat Transfer* 2015;164:221-230.
58. Polyansky OL, Kyuberis AA, Zobov NF, Tennyson J, Lodi L. Monthly Notices, to be submitted.
59. Polyansky OL, Ovsyannikov RI, Kyuberis AA, Lodi L, Tennyson J, Zobov NF. Calculation of rotation-vibration energy levels of the water molecule with near-experimental accuracy based on an *ab initio* potential energy surface. *J Phys Chem A* 2013;117:9633-43.
60. 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 Transfer* 2015; 154:35-43.
61. Rothman LS, Gordon IE, Barber RJ, Dothe H, Gamache RR, Goldman A, et al. HITEMP, the high-temperature molecular spectroscopic database. *J Quant Spectrosc Radiat Transfer* 2010;111:2139-50.