ELSEVIER

Line lists for $H_2{}^{18}O$ and $H_2{}^{17}O$ based on empirical line positions and *ab initio* intensities

Lorenzo Lodi, Jonathan Tennyson *

Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom

ARTICLE INFO

ABSTRACT

New line lists for isotopically substituted water are presented. Most line positions were calculated from experimentally determined energy levels, while all line intensities were computed using an *ab initio* dipole moment surface. Transitions for which experimental energy levels are unavailable use calculated line positions. These line lists cover the range 0.05–$20\,000\ \mathrm{cm}^{-1}$ and are significantly more complete and potentially more accurate than the line lists available via standard databases. All lines with intensities (scaled by isotopologue abundance) greater than 10^{-29} cm/molecule at 296 K are included, augmented by weaker lines originating from pure rotational transitions. The final line lists contain 39 918 lines for $H_2{}^{18}O$ and 27 546 for $H_2{}^{17}O$ and are presented in standard HITRAN format. The number of experimentally determined $H_2{}^{18}O$ and $H_2{}^{17}O$ line positions is, respectively, 32 970 (83% of the total) and 17 073 (62%) and in both cases the average estimated uncertainty is $2 \times 10^{-4}\ \mathrm{cm}^{-1}$. The number of *ab initio* line intensities with an estimated uncertainty of 1% is 16 621 (42%) for $H_2{}^{18}O$ and 13 159 (48%) for $H_2{}^{17}O$.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Accurate characterization of the spectrum of water vapor from the microwave to the near ultraviolet is of great importance for modeling, observation and understanding of many fields in chemistry, physics and engineering [1]. Key topics include atmospheric modeling, with emphasis on the definitive understanding of global warming, as water vapor is responsible for about 70% of the known atmospheric absorption of sunlight and the majority of the greenhouse effect; communication-related fields having to deal with the Earth's atmosphere, such as satellites and telecommunication; astrophysics, such as the atmospheres of most cool stars and brown dwarfs where hot water is a major constituent; and combustion research, with applications, e.g., to rocket exhausts and

turbine engines, as hot steam is a major product of most combustion processes. These studies all rely on access to reliable line lists of water transitions.

This paper presents new line lists for the water isotopologues $H_2{}^{18}O$ and $H_2{}^{17}O$. These line lists rely on the use of two recent, independent developments: the availability of precise, experimentally determined energy levels for these systems [2,3] produced by a IUPAC task group on water spectroscopy and availability of a high accuracy dipole moment surface (DMS) for the water molecule based on high level *ab initio* calculations [4]. The IUPAC energy levels were generated using a MARVEL analysis [5] of the entire available experimental dataset of line positions and therefore have the virtue of being both obtained to experimental accuracy and of being self-consistent. Combining these developments not only allows one to give reliable line positions and intensities for many previously observed transitions, it also permits one to make secure predictions for these parameters for a large number of yet to be observed transitions.

---

* Corresponding author.
  E-mail address: j.tennyson@ucl.ac.uk (J. Tennyson).

We have chosen to employ this methodology initially for the water isotopologues $H_2^{18}O$ and $H_2^{17}O$ since the available, measured transition data for them are sparse because of the greater experimental effort needed with respect to $H_2^{16}O$ due to the use of enriched samples or longer optical paths. Furthermore working with trace species makes the measured intensities generally less reliable because of the additional complexity of the experimental setup. As this is not an issue for *ab initio* calculation, there is clearly a larger potential benefit in using a validated *ab initio* procedure for calculating intensities for these species. Our aim is to replace the transition data in standard databases such as HITRAN [6] with our new line lists.

2. Procedure for line intensity determination

There are two requirements for computing accurate line intensities from first principles: a high accuracy dipole moment surface (DMS) and reliable nuclear-motion wavefunctions. We have expended considerable effort developing procedures to calculate, *ab initio*, reliable DMSs for water [4,7]. In particular, we recently published a DMS, referred to as LTP2011, which we believe to be accurate to 1% in regions of interest to this paper. This claimed accuracy was based on theoretical grounds and supported by comparison with direct measurements of the water dipole [8]. The accuracy of predicted line intensities relies on both the DMS and the wavefunctions used, as discussed below. Comparisons with precise measurements for $H_2^{16}O$ [9] given in our previous paper suggest that when the LTP2011 DMS, when combined with good wavefunctions, is capable of giving line intensities accurate to about 1% for most transitions. Support for this view comes from recent work on measured Stark coefficients [10] for excited states of water. The original study of Stark coefficients for excited stretching states of water [11] showed that the best then available DMSs could only reproduce the measured dipoles to within about 5%. The recent study [10] shows that the LTP2011 DMS gives a substantial improvement, reproducing the measured coefficients to 1% accuracy in all cases of interest to this paper. Above 25 000 cm^{-1}, the results are less good but this can be directly attributed to the lack of completely reliable wavefunctions at these higher energies [10].

The strategy employed to produce line intensities and associated error bars is as follows:

1. Compute two sets of nuclear-motion wavefunctions and energy levels using two different, high quality PESs [12,13].
2. Using two different, high-quality DMSs [4] and the two sets of wavefunctions compute four sets of line intensities. For each transition, evaluate the reliability of the calculated intensities by computing the ratio of strongest to the weakest computed intensities (hereafter designated the scatter ϱ). The transition is marked as 'sensitive' if the ratio is greater than 2.0.
3. Improve the accuracy of line positions by replacing theoretically predicted energy levels with experimentally derived values from Ref. [3].

The quality of the nuclear-motion wavefunctions used to compute the line-strength for any given transition is almost entirely determined by the underlying potential energy surface (PES) used to generate these wavefunctions. In particular, intensity stealing by weak lines, which usually occurs via resonant interactions with states involved in strong transitions, is very sensitive to the choice of the underlying PES and the corresponding nuclear-motion wavefunctions.

As described in point 2 above, we adopted a pragmatic strategy to address the issue of line intensity sensitivity; namely, we computed line lists with several, similar-quality combinations of PESs and DMSs and registered the scatter of the computed line intensities. Lines which show a large scatter in computed intensity are very sensitive and the predicted intensities cannot be trusted. Conversely, lines which show stability upon changes of the DMSs and PESs are probably stable and there is good reason to believe they should be accurate.

Our procedure will naturally give rise to four disjoint line lists:

1. Stable lines, accurate line position available.
2. Stable lines, accurate line position not available.
3. Unstable lines, accurate line positions available.
4. Unstable lines, accurate line positions not available.

Clearly it is the first set we are most interested in. However, for the sake of completeness, our final linelists includes lines originating from all sets. As described in detail is Sections 5 and 6 all transitions have been appropriately labelled with uncertainty codes to reflect the accuracy of the data. This makes it straighforward to filter out lines from unwanted sets, if desired.

Before discussing our procedure in detail it is worthwhile to be specific about the intensity information being considered. Our variational calculations yield temperature-independent transition dipole moments (sometimes called linestrengths), temperature-independent Einstein A coefficients and temperature-dependent integrated effective cross sections ('line intensities'). All these parameters are simply inter-related by analytical formulae (see, e.g., Formulae (55)–(61) of Ref. [14] or Ref. [15]). In keeping with HITRAN conventions we report line intensities at 296 K scaled by natural abundance [15]: 0.00199983 for $H_2^{18}O$ and 0.000371884 for $H_2^{17}O$.

3. Technical details of the calculations

Comprehensive theoretical line lists for $H_2^{16}O$, $H_2^{17}O$ and $H_2^{18}O$ were recently computed by Shirin et al. [12]; these authors construct new semi-empirical potential energy surfaces (PESs) for each water isotopologue, obtained by fitting experimentally derived energy levels up to 18 000 cm^{-1}. We will collectively call these PESs SHIRIN2008 in the following. Transition intensities calculated by Shirin et al. used the CVR DMS of Lodi et al. [7].

The wavefunction data of Shirin et al. are no longer available, and new sets of nuclear-motion wavefunctions for $H_2^{17}O$ and $H_2^{18}O$ were produced using the DVR3D program suite [14] and the semi-empirical SHIRIN2008

PESs. For our analysis of the sensitivity of line intensities we also computed nuclear-motion wavefunctions both for $H_2^{17}O$ and $H_2^{18}O$ using the *ab initio* PESs cvrqd [13]; note that the cvrqd PESs include a mass-dependent adiabatic correction and therefore are slightly different for different water isotopologues. Shirin et al. [12] introduced some small terms in the Hamiltonian for $J \geq 1$ to allow for non-adiabatic effects in highly excited rotational levels. The effect of these extra terms on transition intensities is unknown (but is likely to be very small) and we did not include these correction terms in the present calculations.

3.1. Calculation of the nuclear-motion wavefunctions

The nuclear-motion calculations used symmetrized Radau coordinates and the bisector embedding [16]. The following atomic masses were used: $m(H) = 1.00727646677$ u, $m(^{18}O) = 17.99477385$ u, $m(^{17}O) = 16.994745$ u. Other parameters for dvr3d [14] were as follows. The radial basis used Morse-like oscillator functions with $r_e = 2.55$ a$_0$, $D_e = 0.25$ E$_h$ and $\omega = 0.007$E$_h$. The DVR grid consists in 28 points in the Radau radial coordinate and 44 points in the Radau angular coordinate; the wavefunctions were forced to vanish at the last DVR point.

In the first step of the calculation (program dvr3drjz) the dimension of the Hamiltonian was truncated at 2500 (parameter MAX3D). In the following step of the calculation (program rotlev3b) the parameter controlling truncation of the ro-vibrational Hamiltonian, IBASS, was set to $530(J+1)$. These input parameters should be sufficient to converge computed rotational–vibrational energies to about 0.05 cm^{-1}. Wavefunctions were computed for $J = 0$ to $J = 19$ for $H_2^{18}O$ and for $J = 0$ to $J = 18$ for $H_2^{17}O$. This procedure generated 26 600 energy levels with energy under 19 000 cm^{-1} and $J \leq 18$ for $H_2^{18}O$ and 26 485 such levels for $H_2^{17}O$.

The energy levels both for $H_2^{18}O$ and $H_2^{17}O$ produced with these settings and the semi-empirical shirin2008 PESs agree closely with those reported by Shirin et al. [12]; for $J = 0$ to 6 the average of the absolute differences between the two sets of energy levels is of 0.06 cm^{-1} and 0.01 cm^{-1} for, respectively, for $H_2^{18}O$ and $H_2^{17}O$. These values are comparable with the estimated error due to incomplete convergence and should be considered very small. At higher Js differences grow larger and reach 0.25 cm^{-1} for some $J = 18$ energy levels for both isotopologues. These small differences are attributable to the non-adiabatic correction terms included by Shirin et al. and, to a lesser extent, to the slightly different settings used for the grid size and other parameters.

The absolute differences between energy levels computed with the *ab initio* cvrqd PESs and those computed with shirin2008 grow approximately linearly with the energy of the level; for example, levels with $E \approx 5000$ cm^{-1} show typical differences of about 0.3 cm^{-1} and levels with $E \approx 10\,000$ cm^{-1} show typical differences of about 0.6 cm^{-1}. About 10% of the levels with energies in the range 10 000–20 000 cm^{-1} show relatively large differences (2 to 4.5 cm^{-1}) between the two PESs. Note that differences of a few cm^{-1} for such high-energy levels are large by spectroscopic standards but are exceptionally small for a fully *ab initio* procedure [13].
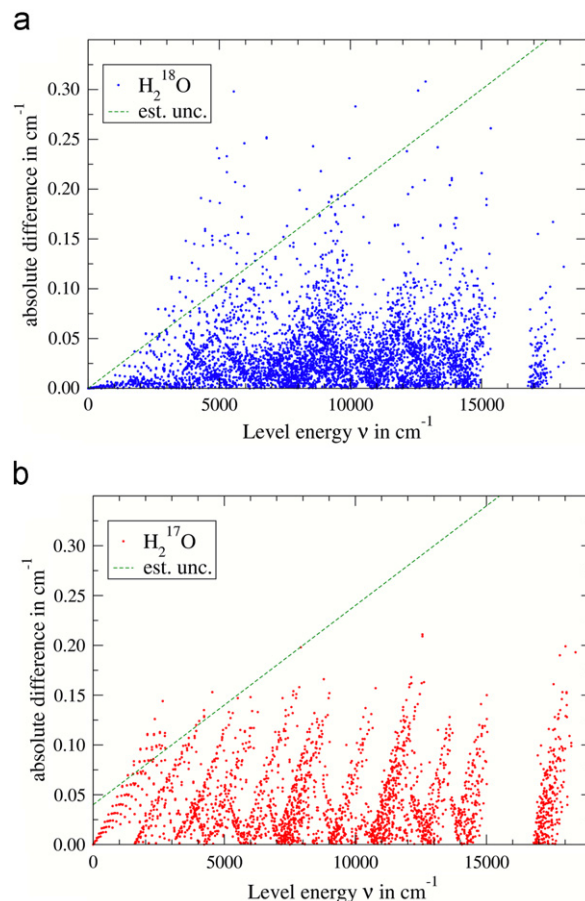
More relevant are the absolute differences of the energy levels computed with the shirin2008 PESs [12] and the experimentally determined IUPAC ones. These are reported as Fig. 1.

The energy levels, particularly those for $H_2^{17}O$, show a structure in J, but as the focus of this work is in line intensities and not line positions it was not investigated further. From the analysis of these differences we estimated the error of our theoretical energy levels computed with the shirin2008 PES to be approximately given by $2 \times 10^{-5}\, \bar{v}$ for $H_2^{18}O$ and $2 \times 10^{-5}\, \bar{v} + 0.04$ cm^{-1} for $H_2^{17}O$. These values were chosen so that over 98% of the IUPAC energy levels were consistent with the chosen error bars and should be indicative of the actual accuracy of the shirin2008 computed energy levels. A few sensitive levels are expected to differ more than indicated by the given error bars.

3.2. Calculation of the partition function

Line intensities calculation requires values for the total internal partition function [15,17]

$$Q(T) = d_i \sum_k g_k (2J_k + 1) e^{-E_k/(k_B T)} \tag{1}$$

Fig. 1. Plot of the absolute differences of theoretical energy levels computed in this work using the shirin2008 PES [12] and experimentally derived energy levels from IUPAC [2]. The dashed lines represent our estimated error for the theoretical energy levels; the lines are $y = 2 \times 10^{-5} v$ cm^{-1} for $H_2^{18}O$ and $y = 2 \times 10^{-5} v + 0.04$ cm^{-1} for $H_2^{17}O$.

where $d_i = 1$ and $d_i = 6$, respectively, for $H_2{}^{18}O$ and $H_2{}^{17}O$ and $g_k = 1$ for para states and $g_k = 3$ for ortho states. Our calculations for $H_2{}^{18}O$ and $H_2{}^{17}O$ give for $Q(276\ K)$ values of, respectively, 176.04 and 1052.1 for both the SHIRIN2008 and CVRQD calculations. The IUPAC experimental energy levels [2], 5131 for $H_2{}^{18}O$ and 2723 for $H_2{}^{17}O$, lead to almost the same values, $Q(276\ K) = 176.05$ and 1052.2; the corresponding values used in HITRAN [17] are slightly lower, namely 175.11 and 1047.9. We believe our values to be more accurate, because our values originate from actual, accurate energy levels, while the values used in HITRAN derive from analytical models of the energy levels [17]. The close agreement between the values for Q derived from theoretical energy levels and the IUPAC ones also supports the previous statement. Furthermore, applying to Eq. (1) the standard error-propagation formula

$$\text{error}^2 = \sum_k \left(\frac{\partial Q}{\partial E_k}\right)^2 (\Delta E_k)^2 \tag{2}$$

and using for ΔE_k the estimated uncertainties of the SHIRIN2008 theoretical energy levels given in the previous section results in an estimated uncertainty in Q of less than 0.01 for both isotopologues.

3.3. Analysis of the theoretical intensities

As described previously, the quality of computed line intensities was judged using the scatter ϱ (intensity ratio of the strongest to weakest computed intensity for each transition). For this purpose two sets of nuclear-motion wavefunctions were produced for each isotopologue using the SHIRIN2008 and the CVRQD PESs. Intensities for each isotopologue were then computed using the nuclear-motion wavefunctions produced from the SHIRIN2008 and the CVRQD PESs and the LTP2011 and LTP2011S DMS [4], giving a total of four possible intensities (SHIRIN2008/LTP2011, SHIRIN2008/LTP2011S, CVRQD/LTP2011 and CVRQD/LTP2011S) for each transition.

3.4. Line intensity threshold

Many of the transitions we compute are very weak so it is necessary to select those strong enough to be significant. For this we used the following threshold function S (which differs from that given by Eq. (1) of the HITRAN2004 paper [18])

$$S_{\text{cut}} = S_{\text{crit}} \tanh^2\left(\frac{hc\tilde{v}}{2k_BT}\right) \tag{3}$$

where $S_{\text{crit}} = 10^{-29}$ cm molecule^{-1}; computed line intensities were multiplied by the appropriate isotopic abundance factor before applying the formula. Even though this function is lower than the one quoted for HITRAN, there are still 113 lines in HITRAN weaker than this (for all water isotopologues); of these, four are $H_2{}^{18}O$ lines and 16 $H_2{}^{17}O$ lines. Eleven of these $H_2{}^{17}O$ lines lying in the range 11 804–14 172 cm^{-1} belong to a larger set of 18 very weak lines from Ref. [19] all of which are unlabelled lines missing line shape data and energy of the lower state and with computed intensities. These lines were consequently dropped. The remaining four $H_2{}^{18}O$ and five

$H_2{}^{17}O$ lines were carried over from HITRAN2008 in our final line list.

3.5. Overview of the final line position and line intensity data sets

For $H_2{}^{18}O$ our final theoretical line list contains 39 918 lines with $J=0$ to 19 and with intensities (scaled by the isotopic abundance factor, 0.00199983) down to S_{cut} given by Eq. (3). Of these, 32 970 have been matched to the experimentally derived energy levels [2] while 6948, mostly very weak, remained unmatched. A total of 1268 lines (3.2%) have an intensity scatter larger than 2.0 and 403 lines (1.0%) greater than 10.0.

For $H_2{}^{17}O$ the situation is very similar but, because fewer experimentally derived energy levels [2] are available, more lines remain unmatched. The final theoretical linelist for $H_2{}^{17}O$ contains 27 546 lines with $J=0$ to 18 and with intensity higher than S_{cut} (the isotopic abundance factor used is 0.000371884). Of these, 17 073 have been matched to the experimentally derived energy levels [2] while 10 473 remained unmatched. A total of 689 lines (2.5%) have a scatter larger than 2.0 and 238 lines (0.9%) greater than 10.0.

These data are displayed graphically in Fig. 2, where transitions which were fully matched are plotted separately. From the plots it is clear that virtually all strong lines were fully matched.

The uncertainty on line positions was straightforwardly derived from the stated uncertainties on the IUPAC energy levels [2] as the square root of the sum of the squares of the uncertainties for the upper and lower level. The corresponding HITRAN uncertainty flag was then opportunely set (see Table 5 of Ref. [18]). For unmatched transitions line positions computed with the SHIRIN2008 PES are provided; the uncertainty flag was conservatively set to zero, corresponding to an uncertainty ≥ 1 cm^{-1}; only computed line positions have the line position uncertainty flag set to zero, making it easy to filter them out if desired. As discussed in the previous section more realistic uncertainties for theoretical lines are given by $2 \times 10^{-5}\ \bar{v}$ for $H_2{}^{18}O$ and by $2 \times 10^{-5}\ \bar{v} + 0.04$ cm^{-1} for $H_2{}^{17}O$.

The scatter factor ϱ was used to set the uncertainty flag for intensities. Lines with large scatters ($\varrho > 2$) were also included for completeness but the corresponding uncertainty flag was set 2, corresponding to 'estimate' (and not 'greater than 20%'). Use of such line intensities should be avoided if possible as they may be very inaccurate.

4. Labelling of the energy levels

In the theoretical calculations using the DVR3D suite levels are labelled only by exact quantum number, namely by the total angular momentum J, the rotationless parity e/f [20], the para/ortho nuclear exchange symmetry and a further number labelling the levels in ascending energetic order within a symmetry block. Taking by convention $p = +1/-1$ for e/f and $q = +1/-1$ for para/ortho these exact quantum numbers are connected to the
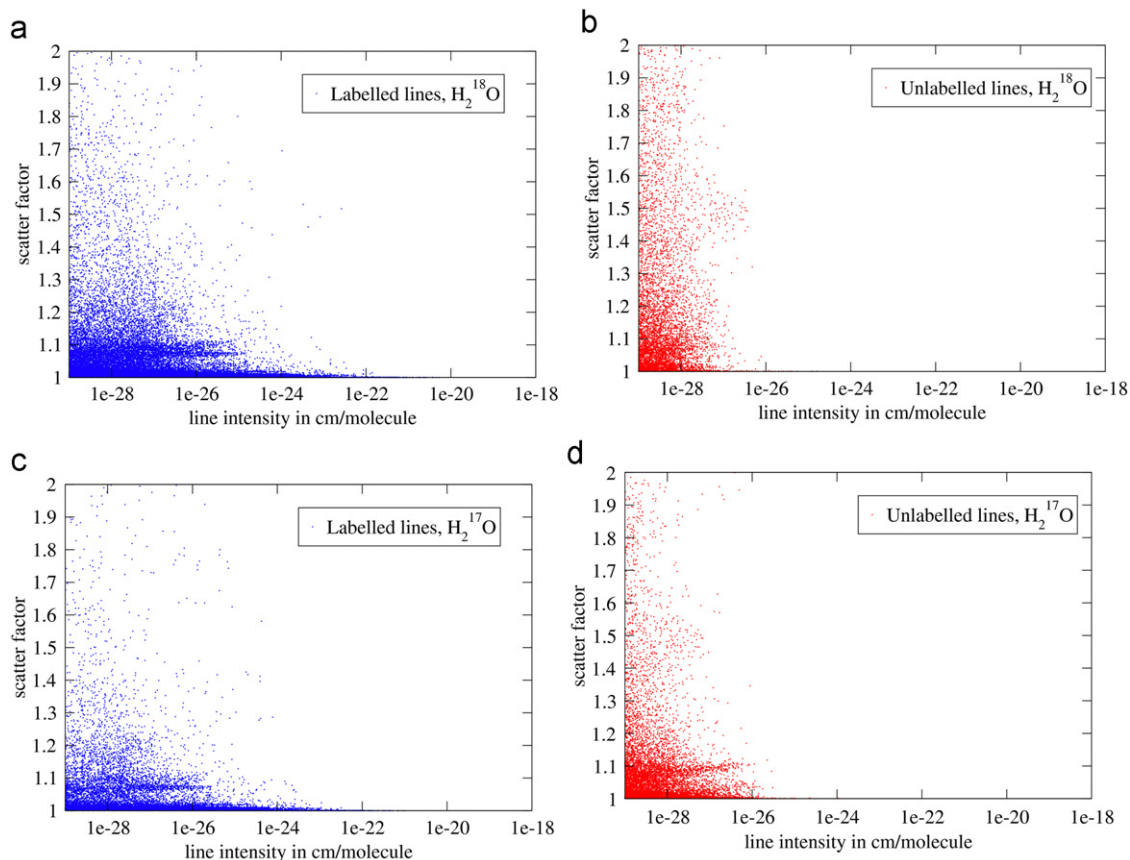
Fig. 2. Plot of the intensity scatter ϱ (intensity ratio of the strongest to weakest computed intensity for each transition) for computed $H_2{}^{18}O$ and $H_2{}^{17}O$ lines (see text).

standard approximate ones for water by $p = (-1)^{J+K_c}$ and $q = (-1)^{v_3+K_a+K_c}$.

We systematically matched the $H_2{}^{18}O$ and $H_2{}^{17}O$ energy levels obtained using the SHIRIN2008 PES with the experimentally derived ones presented in [2]. A summary is reported as Table 1. The energy levels for $H_2{}^{18}O$ and $H_2{}^{17}O$ presented by Shirin et al. [12] are fully labelled with the approximate quantum numbers J, K_a, K_c and v_1, v_2, v_3. However, in many cases the labelling differed from the one recommended by the IUPAC Task Group [2] and adopted by us (see Table 1).

Computed energy levels have a typically accuracy of better than $0.2\ cm^{-1}$ and it was generally straightforward to match our theoretical energy levels with the IUPAC ones using the exact quantum numbers J, p, q and imposing a threshold of about $1\ cm^{-1}$ on the energy itself. However, for both isotopologues there were about a dozen cases where more than one computed level had the correct symmetries and lay very close (closer than $0.1\ cm^{-1}$) to the experimental one. In these cases if possible the matching was decided using of the approximate quantum numbers reported by Shirin et al. [12]. In a few cases the approximate quantum numbers of Shirin *et al.* differed for all the candidate levels from the IUPAC study [2], and the level was matched to the one energetically closest. It is therefore possible that in a few cases the wrong quantum number labels were chosen.

Table 1
Summary of the matching of experimentally derived and theoretical energy levels.

	$H_2{}^{18}O$	$H_2{}^{17}O$
IUPAC levels [2]	5132	2723
Matched to calculated levels	5131	2723
Quantum number mismatches[a]	686	131
Mean absolute error in cm^{-1}[b]	0.04	0.04
Levels with $|\Delta E| > 0.1\ cm^{-1}$[c]	262	189

[a] Levels which have different approximate quantum number assignments in Refs. [2] and [12].
[b] Theoretical energy levels used the SHIRIN2008 PES.
[c] Matched levels with absolute energy differences greater that $0.1\ cm^{-1}$.

5. Analysis of HITRAN2008 data for $H_2{}^{18}O$

HITRAN2008 contains data for 9752 $H_2{}^{18}O$ transitions which are summarized in Table 2; a full list of experimental studies on this isotopologue has been given by the IUPAC Task Group [2]. Is should be noted that transitions intensities coming from the major contributor, Toth's SISAM website [21], are obtained by a semi-empirical procedure using effective Hamiltonians based on extensive experimental studies of water spectra including work on $H_2{}^{18}O$ and $H_2{}^{17}O$ [22–26].

Table 2
HITRAN2008 references for $H_2{}^{18}O$ line positions and intensities.

Reference[a]	Range (cm^{-1})	Number of lines
HITRAN86 [27]	0.05–9290	627
Jucks [28,29]	11–500	778
SISAM [21]	501–7732	5792
SISAM [21][b]	4941–7512	41
Barber et al. [30]	7055	1
Partridge and Schwenke [31][c]	9658–11 184	190
Chevillard et al. [32,31][d]	9677–11 246	1171
Tolchenov and Tennyson [33]	10 096–12 313	138
Brown et al. [34]	10 169	1
Tanaka et al. [19]	12 405–14 518	1013
Total	0.05–14 518	9752

   [a] When two references are given, the first refers to line positions and the second to line intensities.
   [b] These lines are doubled with the weaker unassigned line hidden beneath the stronger component.
   [c] Intensities from [31] were scaled by a band-dependent factor determined by comparison with selected experimental intensities from [34].
   [d] Intensities from [31] were scaled by a band-dependent factor determined by comparison with selected experimental intensities from [32].

Clearly HITRAN2008 is far from giving a complete coverage for $H_2{}^{18}O$, and there are several energy windows where data are completely missing, see Fig. 3.

HITRAN2008 lines were compared with our dataset by matching the exact quantum numbers *J,p,q*. This approach is a compromise. There are several issues which lead to differing approximate quantum number assignments and therefore it was considered undesirable to match HITRAN transitions only on the basis of the assigned quantum numbers. On the other hand trying to match HITRAN transitions to theoretical ones only on the basis of line positions and line intensities generally leads to far too many possible matches. Tests showed that extracting the exact quantum numbers from the approximate ones listed in HITRAN2008 and imposing a match on these works well. For each given HITRAN2008 line, possible candidate matches in the theoretical dataset were then selected by energy and intensity criteria; indicating with HT HITRAN2008 data and with LT our calculations we used the following criteria to establish a match

$$\left| \tilde{\nu}_{LT} - \tilde{\nu}_{HT} \right| \le \sigma_{HT} + 0.2 \ \mathrm{cm}^{-1}$$

$$r/(\varrho \varrho_{HT}) \le 2 \tag{4}$$

where: σ_{HT} is the declared uncertainty in HITRAN2008 line positions (a value $\sigma_{HT} = 1 \ \mathrm{cm}^{-1}$ was used when the position uncertainty was marked as unavailable); ϱ_{HT} is the declared HITRAN2008 uncertainty in line intensity, expressed as 1.2 for 20% etc.; ϱ is the theory scatter factor introduced in Section 3. *r* is the intensity ratio between max(int$_{HT}$,int$_{LT}$) and min(int$_{HT}$,int$_{LT}$). When the HITRAN intensity uncertainty was marked as 'unavailable', 'default' or 'average' the value $\varrho_{HT} = 1.5$ was used.
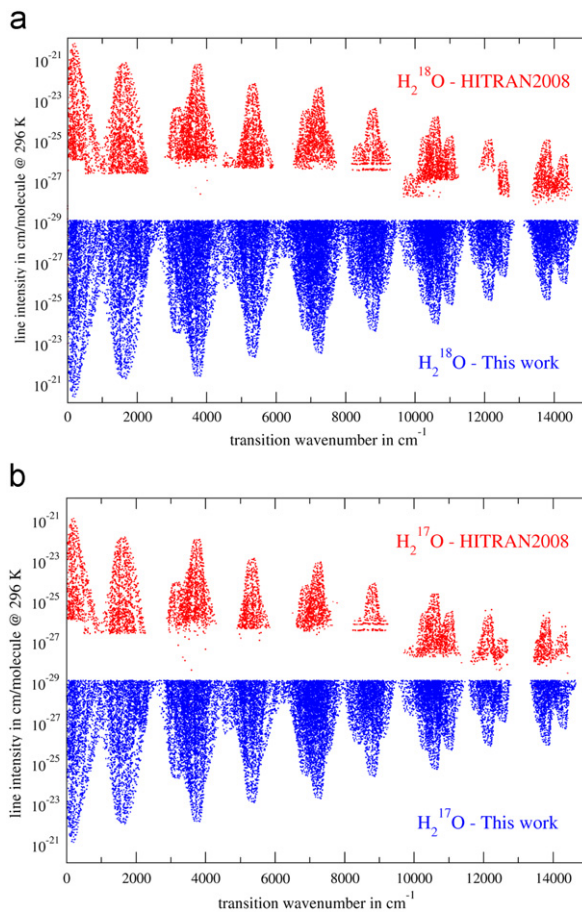
Fig. 3. Overview of the $H_2{}^{18}O$ and $H_2{}^{17}O$ spectra for HITRAN2008 and for the present calculation using the S$_{HIRIN}$2008 PES and the LTP2011 DMS (see text).

This automatic procedure matched 9233 HITRAN lines; these data were then inspected manually to resolve multiple-matches. A further set of 475 lines were matched by a more careful one-by-one analysis. Finally we matched 9708 HITRAN2008 lines, while a small set of 44 lines could not be matched at all to our theoretical lines. For the subset of matched lines accurate line positions derived from the IUPAC data [3] are available for all but 2 lines; in turn, for this set accurate intensities (scatter less than 2) are available for 9602 lines. For this subset HITRAN2008 intensity and line position data are compatible with the theoretical one ('compatible' is taken to indicate here a discrepancy of less than three times the declared HITRAN uncertainty) for 8959 lines. The complementary set of $9602 - 8959 = 643$ lines show large discrepancies with the theoretical data either for line positions or intensities. A plot of absolute differences of transition wavenumbers and of intensity ratios are given in Fig. 4.

The swarm of data at about $8600 \ \mathrm{cm}^{-1}$, showing differences up to several wavenumbers, come unattributed from HITRAN86 [27]; line positions have a reported uncertainty of worse than $1 \ \mathrm{cm}^{-1}$. Part of the HITRAN86 sources can be inferred from the 1973 AFGL report [35] along with the successive AFGL compilations [36,37].
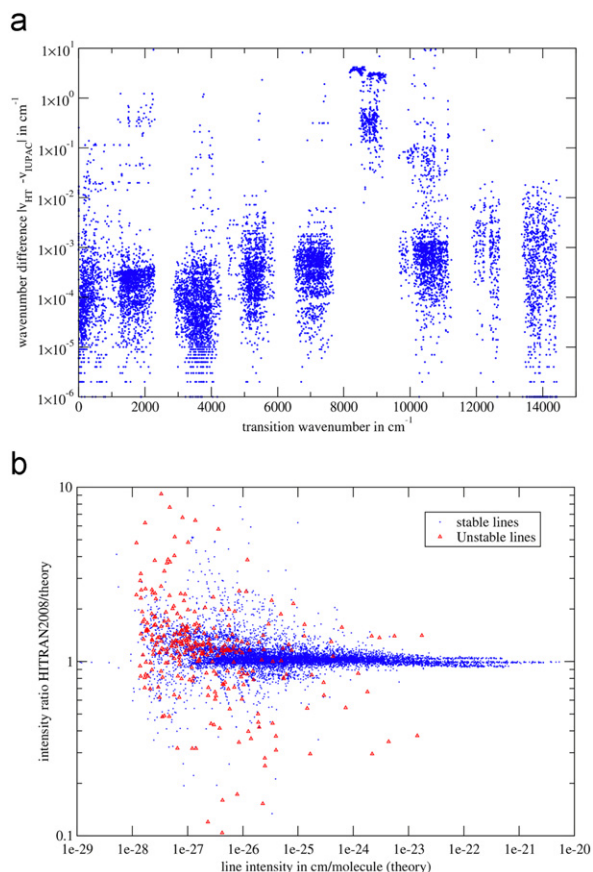
Fig. 4. Plots relative to $H_2^{18}O$ HITRAN lines of the wavenumber differences with IUPAC [3] and of the intensity ratios with calculated ones using the LTP2011 DMS and wavefunctions based on the SHIRIN2008 PES.

6. Analysis of HITRAN2008 data for $H_2^{17}O$

An analogous strategy was followed for $H_2^{17}O$. HITRAN2008 reports 6992 transitions for the $H_2^{17}O$ isotopologue whose provenance is summarised in Table 3. As mentioned in Section 3 18 very weak lines, all from Ref. [19], are unassigned and report very incomplete data and were excluded from the analysis. Of the remaining transitions, 107 could not be matched to theoretical lines and are probably incorrect; 100 of these lines come from HITRAN86 [27] and 6 from the SISAM web site [21]. The remaining 6867 could be matched to our linelist; however, 135 of these involved energy levels not available in the list of 2 723 experimentally derived IUPAC energy levels [3] and therefore for these lines comparison with accurate line positions was not available (108 of these lines have computed line positions and intensities, mostly from SISAM [21]).

Plot relative to the remaining 6732 lines of absolute differences of transition wavenumber and of intensity ratios with theoretical intensities are presented as Fig. 5.

Note that the group of lines situated at 8560–9128 cm^{-1} showing large deviations comes entirely from HITRAN86 and carry an uncertainty of 1 cm^{-1}.

Table 3
HITRAN2008 references for $H_2^{17}O$ line positions and intensities.

Reference[a]	Range (cm^{-1})	Number of lines
Jucks [28,29]	11–499	501
Jucks [28,27]	22–448	117
SISAM [21]	501–7782	4101
SISAM [21][b]	6794–7443	19
HITRAN86 [27]	8241–9241	319
Camy-Peyret et al. [38]	9712–11 335	1062
Tolchenov and Tennyson [33]	10 371–13 727	36
Tanaka et al. [19]	11 365–14 472	835
Tolchenov and Tennyson [33,19]	12 109–12 200	2
Total		6992

[a] When two references are given, the first refers to line positions and the second to line intensities.
[b] These lines are doubled with the weaker unassigned line hidden beneath the stronger component.
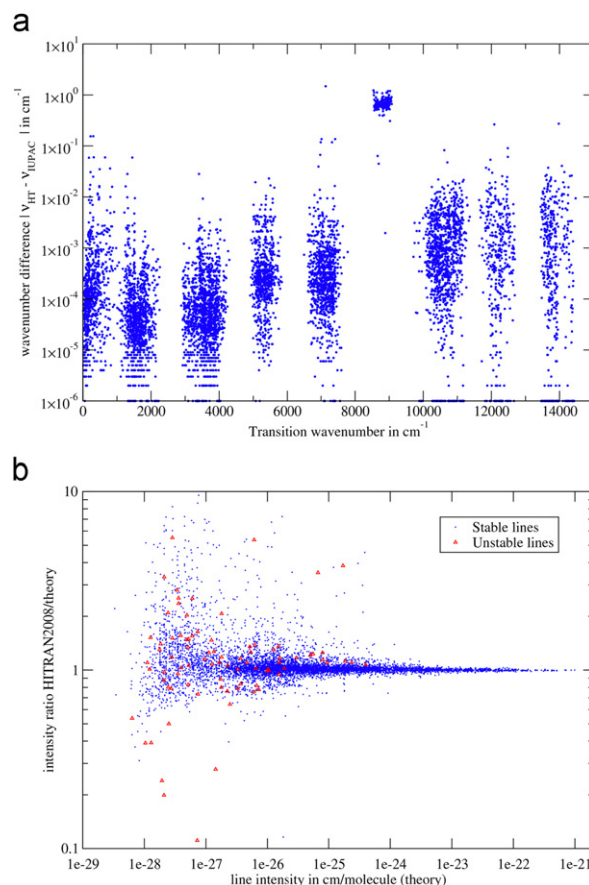
Fig. 5. Plots relative to $H_2^{17}O$ HITRAN lines of the wavenumber differences with IUPAC [3] and of the intensity ratios with calculated ones using the LTP2011 DMS and wavefunctions based on the SHIRIN2008 PES.

The agreement of HITRAN line positions for the 6732 lines for which IUPAC energy levels are available is as follows. The number of lines for which the HITRAN and

IUPAC line position differ by more than three standard deviations (calculated using the HITRAN declared uncertainties) is 184; almost all of these are lines for which the difference is small in absolute terms but which are flagged both by HITRAN and by the IUPAC levels as very accurate (uncertainties $\leq 10^{-4}$ cm^{-1}). This probably indicates that in the case of accurate lines systematic errors have been somewhat underestimated.

The agreement of the theoretical intensities with the HITRAN data can be considered fair, with 81% of the lines agreeing within 20%. There are also 30 transitions not visible in Fig. 5 which differ by more than a factor 10 from the theoretical data; only four of these lines have been marked as sensitive (scatter greater than 2), indicating that the HITRAN intensity data for the remaining 26 lines is probably very inaccurate. Overall, there are 90 lines for which the theoretical scatter is greater than a factor of 2 (unstable lines). Most (but not all) of the lines visible in Fig. 5 and displaying large discrepancies come from Toth's website [21]; each of these and a few other outliers were inspected one-by-one to see if the discrepancy could be put down to labelling issues but with no success. Note that all of the transitions intensities for which theory and HITRAN differ significantly are weak or medium-weak lines with intensities no more than 2×10^{-24} cm/molecule.

7. Description of the final line lists

Our final recommended line lists for $H_2^{18}O$ and $H_2^{17}O$ are given in the supplementary data. These linelists are given in HITRAN format and with intensities scaled by natural abundance. This form requires four line shape parameters (air-broadened and self-broadened half widths, temperature-dependence exponents for air-broadened half widths and air-pressure-induced line shifts) which are essentially unknown for isotopically substituted water. We have therefore assumed that one could use the parameters for the corresponding $H_2^{16}O$ transition having the same approximate quantum number assignments. Following this assumption we used for our line lists the $H_2^{16}O$ line shape parameters given in the recent release of HITEMP [39]. These parameters are based on an improved version of the algorithm of Gordon et al. [40]. Line shape parameters for unlabelled lines were also taken from HITEMP2010, using the default values (dependent on J_{upper} and J_{lower}) employed in HITEMP for unassigned lines. Tables giving the default line shape parameters used for unlabelled lines can be found in the readme file accompanying the supplementary data.

For unlabelled lines the approximate quantum numbers were set to conventional negative values. The vibrational quantum numbers $v_1 v_2 v_3$ were always set to -2. The angular momentum quantum number J was set to its correct value, while K_a and K_c were conventionally set to either -2 or -1 in such a way that application of the formulae mentioned in Section 4 permit to derive the exact quantum numbers p and q.

All transitions computed by us respect the threshold condition on intensities given by formula (3); however, as mentioned at the end of Section 3, four HITRAN2008 lines violate this threshold and were simply carried over from HITRAN.

We have used the HITRAN accuracy flags to distinguish between those transitions we have confidence in (see Section 3.5): line positions from the IUPAC data and stable intensities, and those which we were forced to use calculated line positions and/or unstable intensities. Where possible, available measured intensities have been used instead of unstable ones.

Our final line lists contain 39 918 lines for $H_2^{18}O$ and 27 546 for $H_2^{17}O$.

8. Conclusion

We have applied a mixed empirical—*ab initio* procedure to produce reliable line lists for $H_2^{18}O$ and for $H_2^{17}O$. We suggest these line lists are both more accurate and more complete than those currently available in standard databases [6,41].

We note that a comprehensive list of experimentally determined energy levels are also available for HDO [3], so a similar procedure could be applied to deuterated water. A similar compilation of energy levels for the major isotopologue $H_2^{16}O$ is almost complete. This study is complicated by the large amount of transition data available, which means that fewer important new lines will be predicted from the empirical energy levels. However, given the great importance of $H_2^{16}O$, synthesizing a line list using the techniques developed here would undoubtedly be of use.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jqsrt.2012.02.023.

References

[1] Bernath PF. The spectroscopy of water vapour: experiment, theory and applications. Phys Chem Chem Phys 2002;4:1501–9.
[2] 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 Trans 2009;110:573–96.

[3] 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 Trans 2010;111:2160–84.

[4] 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.

[5] Furtenbacher T, Császár AG, Tennyson J. MARVEL: measured active rotational–vibrational energy levels. J Mol Spectrosc 2007;245:115–25.

[6] Rothman LS, Gordon IE, Barbe A, Benner DC, Bernath PF, Birk M, et al. The HITRAN 2008 molecular spectroscopic database. J Quant Spectrosc Radiat Trans 2009;110:533–72.

[7] Lodi L, Tolchenov RN, Tennyson J, Lynas-Gray AE, Shirin SV, et al. A new ab initio ground-state dipole moment surface for the water molecule. J Chem Phys 2008;128:044304.

[8] Clough SA, Beers Y, Klein GP, Rothman LS. Dipole-moment of water from Stark measurements of H_2O, HDO, and D_2O. J Chem Phys 1973;59:2254–9.

[9] Lisak D, Harvey DK, Hodges JT. Spectroscopic line parameters of water vapor for rotation–vibration transitions near 7180 cm^{-1}. Phys Rev A 2009;79:052507.

[10] Boyarkin OV, Grechko M, Aseev O, Rizzo TR, Zobov NF, et al. Observation and calculation of Stark coefficients for states of the water molecule up to dissociation. J Chem Phys., to be submitted.

[11] Callegari A, Theule P, Tolchenov RN, Zobov NF, Polyansky OL, Tennyson J, et al. Dipole moments of highly vibrationally excited water. Science 2002;297:993–5.

[12] 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 (H2O)-O-16, (H2O)-O-17, and (H2O)-O-18. J Chem Phys 2008;128:224306.

[13] Barletta P, Shirin SV, Zobov NF, Polyansky OL, Tennyson J, Valeev EF, et al. CVRQD ab initio ground-state adiabatic potential energy surfaces for the water molecule. J Chem Phys 2006;125:204307.

[14] Tennyson J, Kostin MA, Barletta P, Harris GJ, Polyansky OL, Ramanlal J, et al. DVR3D: a program suite for the calculation of rotation–vibration spectra of triatomic molecules. Comput Phys Commun 2004;163:85–116.

[15] Šimečková M, Jacquemart D, Rothman LS, Gamache RR, Goldman A. Einstein A-coefficients and statistical weights for molecular absorption transitions in the HITRAN database. J Quant Spectrosc Radiat Trans 2006;98:130–55.

[16] Tennyson J, Sutcliffe BT. Discretisation to avoid singularities in vibration–rotation hamiltonians: a bisector embedding for AB$_2$ triatomics. Int J Quantum Chem 1992;42:941–52.

[17] Fisher J, Gamache RR, Goldman A, Rothman LS, Perrin A. Total internal partition sums for molecular species in the 2000 edition of the HITRAN database. J Quant Spectrosc Radiat Trans 2003;82:401–12.

[18] Rothman LS, Jacquemart D, Barbe A, Benner DC, Birk M, Brown LR, et al. The HITRAN 2004 molecular spectroscopic database. J Quant Spectrosc Radiat Trans 2005;96:139–204.

[19] Tanaka M, Naumenko O, Brault J, Tennyson J. Fourier transform absorption spectra of $H_2{}^{18}$O and $H_2{}^{17}$O in the $3\nu+\delta$ and 4ν polyad region. J Mol Spectrosc 2005;234:1–9.

[20] Brown JM, Hougen JT, Huber K-P, Johns JWC, Kopp I, Lefebvre-Brion H, et al. The labeling of parity doublet levels in linear molecules. J Mol Spectrosc 1975;55:500–3.

[21] Toth RA. Linelist of water vapor parameters from 500 to 8000 cm^{-1}. Available from ⟨http://mark4sun.jpl.nasa.gov/data/spec/H2O⟩.

[22] Toth RA. Transition frequencies and absolute strengths of $H_2{}^{17}$O and $H_2{}^{18}$O in the 6.2-μm region. J Opt Soc Am B 1992;9:462–82.

[23] Toth RA. The ν_1 and ν_3 bands of $H_2{}^{17}$O and $H_2{}^{18}$O: line positions and strengths. J Mol Spectrosc 1994;166:184–203.

[24] Toth RA. Transition frequencies and strengths of $H_2{}^{17}$O and $H_2{}^{18}$O: 6600 to 7640 cm^{-1}. Appl Opt 1994;33:4868–79.

[25] Toth RA. Water vapor measurements between 590 and 2582 cm^{-1}: line positions and strengths. J Mol Spectrosc 1998;190:379–96.

[26] Toth RA. Measurements of positions, strengths and self-broadened widths of H_2O from 2900 to 8000 cm^{-1}: line strength analysis of the 2nd triad bands. J Quant Spectrosc Radiat Trans 2005;94:51–107.

[27] Rothman LS, Gamache RR, Goldman A, Brown LR, Toth RA, Pickett HM, et al. The HITRAN database: 1986 edition. Appl Opt 1987;26:4058–97.

[28] Jucks KW. Calculations. Private communication, 2000.

[29] Pearson JC. Private communication, 2000.

[30] Barber RJ, Tennyson J, Harris GJ, Tolchenov RN. A high-accuracy computed water line list. Mon Not Roy Astron Soc 2006;368:1087–94.

[31] 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.

[32] Chevillard J, Mandin J-Y, Flaud J-M, Camy-Peyret C. $H_2{}^{18}$O: line positions and intensities between 9500 and 11 500 cm^{-1}. The (041), (220), (121), (201), (102), and (003) interacting states. Can J Phys 1987;65:777–89.

[33] Tolchenov R, Tennyson J. Water line parameters from refitted spectra constrained by empirical upper state levels: study of the 9500–14 500 cm^{-1} region. J Quant Spectrosc Radiat Trans 2008;109:559–68.

[34] Brown LR, Toth RA, Dulick M. Empirical line parameters of $H_2{}^{16}$O near 0.94 μm: positions, intensities and air-broadening coefficients. J Mol Spectrosc 2002;212:57–82.

[35] McClatchey RA, Benedict WS, Clough SA, Burch DE, Calfee RF, Fox K, et al. AFCRL atmospheric absorption line parameters compilation AFCRL-TR-73-0096, 1973.

[36] Rothman LS. AFGL atmospheric absorption line parameters compilation: 1980 version. Appl Opt 1981;20:791–5.

[37] Rothman LS, Gamache RR, Barbe A, Goldman A, Gillis JR, Brown LR, et al. AFGL atmospheric absorption line parameters compilation: 1982 version. Appl Opt 1983;22:2247–56.

[38] Camy-Peyret C, Flaud J-M, Mandin J-Y, Bykov A, Naumenko O, Sinitsa L, et al. Fourier-transform absorption spectrum of the $H_2{}^{17}$O molecule in the 9711–11 335 cm^{-1} spectral region: the first decade of resonating states. J Quant Spectrosc Radiat Trans 1999;61:795–812.

[39] 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 Trans 2010;111:2139–50.

[40] Gordon IE, Rothman LS, Gamache RR, Jacquemart D, Boone C, Bernath PF, et al. Current updates of the water–vapor line list in HITRAN: a new "diet" for air-broadened half-widths. J Quant Spectrosc Radiat Trans 2007;108:389–402.

[41] Jacquinet-Husson N, Crepeau L, Armante R, Boutammine C, Chédin A, Scott NA, et al. The 2009 edition of the GEISA spectroscopic database. J Quant Spectrosc Radiat Trans 2011;112:2395–445.