scientific data

OPEN

DATA DESCRIPTOR

The W2024 database of the water isotopologue $H_2{}^{16}O$

Tibor Furtenbacher[1], Roland Tóbiás[1,2] ✉, Jonathan Tennyson[3], Robert R. Gamache[4] & Attila G. Császár[1,2,5] ✉

The rovibrational spectrum of the water molecule is the crown jewel of high-resolution molecular spectroscopy. While its significance in numerous scientific and engineering applications and the challenges behind its interpretation have been well known, the extensive experimental analysis performed for this molecule, from the microwave to the ultraviolet, is admirable. To determine empirical energy levels for $H_2{}^{16}O$, this study utilizes an improved version of the MARVEL (Measured Active Rotational-Vibrational Energy Levels) scheme, which now takes into account multiplet constraints and first-principles energy-level splittings. This analysis delivers 19027 empirical energy values, with individual uncertainties and confidence intervals, utilizing 309 290 transition wavenumbers collected from 189 (mostly experimental) data sources. Relying on these empirical, as well as some computed, energies and first-principles intensities, an extensive composite line list, named CW2024, has been assembled. The CW2024 dataset is compared to lines in the canonical HITRAN 2020 spectroscopic database, providing guidance for future experimental investigations.

Background & Summary

Over the last half of a century, collecting accurate, line-by-line spectroscopic data for isotopologues of the water molecule has been a major research activity in a large number of spectroscopic laboratories (see, for instance, Refs. [1–10] and references cited therein). An important contribution toward the detailed understanding of high-resolution spectra recorded for several water isotopologues, beyond selecting particular data for a particular database, started two decades ago, when a Task Group (TG) was set up by the International Union of Pure and Applied Chemistry (IUPAC) on "A Database of Water Transitions from Experiment and Theory" (Project No. 2004-035-1-100). This TG, formed by experimental and computational spectroscopists, published validated sets of measured rovibrational transitions and empirical energy levels on nine water isotopologues, $H_2{}^xO$[1–3], HD^xO[2], and $D_2{}^xO$[4] ($x = 16, 17, 18$).

A significant update of the IUPAC TG water data[5] was published by four of the five authors of this paper in 2020, in the form of the W2020 database[7,8], for the three $H_2{}^xO$ species. During the development of the W2020 datasets[7,8], the spectroscopic data of the three isotopologues were considered jointly[8], allowing improvements to be made for the individual datasets. The W2020-$H_2{}^xO$ line lists were successfully employed in the latest edition of HITRAN[10], the canonical source of line-by-line spectroscopic information for species of atmoshperic interest, representing about 85% of the $\approx$ 233 000 lines with complete assignment in the HITRAN-$H_2{}^xO$ catalogs.

Accurate, high-resolution spectroscopic information on various water isotopologues is required by numerous complex applications, including climate-change and atmospheric research, astronomy, combustion chemistry, metrology, planetary science, and remote sensing[6,11–13], all with vastly different environments. The experimental studies of water spectra have been aided by the development of high-resolution and ultrahigh-precision techniques, such as cavity ring-down spectroscopy (CRDS)[14] and noise-immune cavity-enhanced optical heterodyne molecular spectroscopy (NICE-OHMS)[15–22]. Theoretical interpretation of complex (ultra)high-resolution spectra also had to be improved. With the arrival of the fourth age of quantum chemistry[23] came the ability to compute nearly complete line lists for molecules, like those constructed under the aegis of the ExoMol project[24–27]. Novel algorithms have also been devised which can cope with experimental data of vastly different accuracy[23,28,29].

[1]HUN-REN–ELTE Complex Chemical Systems Research Group, P.O. Box 32, H-1518, Budapest 112, Hungary. [2]Institute of Chemistry, ELTE Eötvös Loránd University, H-1117 Budapest, Pázmány Péter sétány 1/A, Hungary. [3]Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, United Kingdom. [4]Department of Environmental, Earth, and Atmospheric Sciences, University of Massachusetts Lowell, 365 Riverside Street, Lowell, MA, 01854, USA. [5]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA. ✉e-mail: tobiroli.new@gmail.com; attila.csaszar@ttk.elte.hu

A notable theoretical advancement in high-resolution spectroscopy was the introduction of the concept of spectroscopic networks (SN)[30–34]. SNs form the basis of the MARVEL (Measured Active Rotational-Vibrational Energy Levels) procedure[30–32,34–37], a global spectrum analysis tool[7,33,34,38–40]. MARVEL inverts the information contained in experimental line positions and delivers empirical energy levels[41] with individual uncertainties. MARVEL has been used to study the spectra of several diatomic[42–47], triatomic[48–52], tetratomic[53–55], and larger[56,57] species.

A number of developments since the publication of the extensive W2020-$H_2{}^{16}O$ dataset have made its reexamination desirable. Most importantly, results from carefully designed precision-spectroscopy experiments have become available for four water isotopologues[29,58–63]. These studies, in particular, yielded empirical energies, accurate to a few kHz, for a large number of lower states in the experimental SN of $H_2{}^{16}O$. Further new experimental studies have also appeared[64–69]. In particular, experimentalists published a number of measured lines, with a typical uncertainty of 10^{-3}–10^{-5} cm^{-1}[64,66,68–70], challenging certain database entries of the W2020-$H_2{}^xO$ lists. Avoiding the criticism directed towards a subset of empirical W2020-$H_2{}^xO$ energy levels[66], some already refuted in ref. [65], requires further improvements on how the experimental information is handled during a MARVEL-type analysis.

The research behind in this paper focused on the (a) refinement of the W2020-$H_2{}^{16}O$ database *via* an improved MARVEL methodology, leading to the W2024 dataset, and (b) construction of a large composite line list, called CW2024, for the $H_2{}^{16}O$ molecule. The (C)W2024 datasets are compared to HITRAN 2020[10], to expedite the inclusion of the (C)W2024 data in spectroscopic information systems.

Methods

Improved MARVEL methodology. During our studies devoted to MARVEL-based analyses of high-resolution rovibronic spectra of small, usually atmospherically and astronomically relevant molecules[42–57], novel aspects and analysis tools have constantly been introduced. These became essential features in later versions of the MARVEL approach and they are described in a number of publications[7,33,34,38–40]. Nevertheless, the ever-expanding spectroscopic information available for $H_2{}^{16}O$ made it necessary to further improve our MARVEL-based analysis technique, as outlined below.

In what follows, the $H_2{}^{16}O$ energy levels are labelled as $(v_1\ v_2\ v_3)J_{K_a,K_c}$, whereby v_1, v_2, and v_3 are the normal-mode quantum numbers of the symmetric stretch, bend, and antisymmetric stretch motions, respectively, J is the overall rotational quantum number, while K_a and K_c symbolize the conventional prolate- and oblate-top rotational quantum numbers, respectively. As usual, $(v_1'\ v_2'\ v_3')J'_{K_a',K_c'} \leftarrow (v_1''\ v_2''\ v_3'')J''_{K_a'',K_c''}$ denotes a rovibrational transition, where $'$ and $''$ signify the upper and lower states, respectively[71].

Multiplet constraints. Under favorable circumstances, a spectral line representing a transition between two states is well separated from all neighboring lines, yielding a unique position for it. If two or more transitions are closer to each other than can be resolved by a particular experiment, the lines form an unresolved multiplet. In such cases, (a) the spectral line shape might become distorted, (b) the observed intensity corresponds to the sum of intensities of the individual transitions, and (c) the measured position will be an intensity-weighted average of the unknown individual positions. How the treatment of these multiplets was introduced to MARVEL is described next.

Let the wavenumber of the ith line in the dataset, σ_i, be represented with the following expression:

$$\sigma_i \approx S_i \equiv \sum_{j=1}^{N_T} u_{ij} s_j, \tag{1}$$

whereby N_T is the number of transitions within the database, s_i means the exact (unknown) position of the ith transition, and the u_{ij} entries are the relative weights satisfying $u_{11} + u_{12} + \ldots + u_{1N_T} = 1$. When the (i, j) line pair is part of the same multiplet, u_{ij} will be the relative intensity of the jth line in this multiplet; otherwise, $u_{ij} = 0$. For example, if (1, 2) means an unresolved (*ortho*, *para*) doublet of $H_2{}^{16}O$, then $\sigma_1 = \sigma_2 \approx 0.75s_1 + 0.25s_2$. Note that equation (1) holds for an isolated line, as well, then $u_{ii} = 1$ and $u_{ij} = 0$ for all $j \neq i$.

According to quantum mechanics, the s_j wavenumber of a transition is subject to the Ritz principle[72],

$$s_j = E_{\mathrm{up}(j)} - E_{\mathrm{low}(j)}, \tag{2}$$

where up(j) and low(j) symbolize the indices of the upper and lower states of the jth line, respectively, and E_k is the (unknown) energy value of the kth quantum state within the transition dataset. Combining equation (1) and equation (2), the following least-squares objective function can be prescribed for MARVEL:

$$\Omega(\mathbf{E}) = \sum_{i=1}^{N_T} w_i \left[\sigma_i - S_i(\mathbf{E})\right]^2, \tag{3}$$

whereby w_i is the statistical (MARVEL) weight of the ith transition and $\mathbf{E}$ is the vector of unknown energy values (variables) in the $S_i(\mathbf{E}) \equiv S_i$ sums. If an $\overline{\mathbf{E}}$ vector minimizes the (quadratic) objective function $\Omega(\mathbf{E})$, its entries are called empirical (MARVEL) energies [for details on how to calculate these MARVEL energies, see Supplementary Information (A)].

A drawback of applying multiplet constraints is that they reduce the number of statistical degrees of freedom, n_{DOF}, in the database. In effective Hamiltonian (EH) fits, where such constraints are often employed, *e.g.*, within the SPFIT code[73], this is not a problem, as EH models contain much fewer fitting parameters than MARVEL;

thus, they can tolerate a decreased n_{DOF} value. Accordingly, to make the MARVEL equations solvable, the input dataset must be complemented with accurate estimates for the relative positions of the individual lines within unresolved multiplets. A feasible way on how to find such estimates is proposed in the next subsection.

Use of computed energy-level splittings and relative positions. As evidenced multiple times, also for water isotopologues[29,58,60–62], energy differences of rovibrational state pairs pertaining to the same vibrational band can be accurately derived from first-principles solution of the nuclear Schrödinger equation. This favorable state of affairs is due to the utilization of exact kinetic energy operators and the fact that discrepancies arising from deficiencies, such as local inaccuracies in the model potential energy surface (PES) employed, are largely systematic, leading to considerable error cancellation when energy differences between highly similar state pairs are formed. The same holds for the relative position of two lines sharing their upper and lower vibrational parents, as it can be obtained from the (signed) splittings of their upper and lower states, d'_{ij} and d''_{ij}, respectively:

$$\rho_{ij} = s_i - s_j$$

Thus, computed energy-level splittings, which can be added as wavenumbers of "virtual" lines to the input file, are able to eliminate the underdeterminacy induced by multiplet constraints. Note that resonance interactions among closely-spaced levels in the same *J*/symmetry block may decrease the accuracy of these computed splittings at high *J*s, which must be accounted for in the final uncertainty budget.

Within the W2024 dataset, the virtual transitions defined above are placed into a segment called "24virt" and correspond only to energy splittings of *ortho–para* state pairs, whose assignments differ solely in their K_a or K_c quantum numbers (the error cancellation seems to work exceptionally well for these state pairs). The splitting values included in the 24virt segment are taken from the first-principles POKAZATEL[74] energy list, for which

$$\mathcal{U}(d^{\mathrm{POK}}) \approx \max\left[|d^{\mathrm{POK}} - d^{\mathrm{BT2}}|,\ \min\left(0.1\,|d^{\mathrm{POK}}|,\ 0.025\ \mathrm{cm}^{-1}\right),\ 10^{-6}\ \mathrm{cm}^{-1}\right] \quad (5)$$

is employed as an (initial) uncertainty approximation, whereby d^{POK} and d^{BT2} are the POKAZATEL[74] and BT2[75] estimates for the same splitting, respectively. A description of how uncertainties of relative positions are taken into account in the uncertainties of energies and predicted wavenumbers is offered in Supplementary Information (B).

Confidence intervals. To appreciate the importance of confidence intervals, a new concept introduced here to MARVEL, one needs to understand the limitations of network-based procedures for the recognition of outliers. During the analysis of SNs, outliers are lines with faulty wavenumbers, uncertainties, or assignments. As shown before[40], outlier-detection tools designed for SNs are built upon the notion of network cycles (that is, sequences of connected lines and states, where each state has exactly two neighboring states) and network (in)consistency. It must also be stressed that there are a few misconceptions surrounding outlier detection in high-resolution spectroscopy[40]. One of them is related to latent outliers, which cannot be detected *via* network-theoretical means, as they do not violate the consistency of the SN.

Owing to potential error compensation in cycles, see misconception M5 in Ref. [40], in principle any transition might be a latent outlier. In practice, a latent outlier is typically (a) a bridge (i,e., a line without cycles) or (b) a transition whose uncertainty is smaller than the threshold (that is, the sum of uncertainties) in all of its cycles. For instance, if a line has an uncertainty of 10^{-4} cm^{-1}, but it participates only in cycles with thresholds being 10^{-3} cm^{-1}, the accuracy of this transition can be validated by MARVEL only to 10^{-3} cm^{-1}.

Based on all these considerations, it is worth defining a measure of "validity", what is called here a confidence interval (CI), characterizing each transition of the dataset. A CI value provides a lower limit, below which no error can be recognized by MARVEL in a line position or its uncertainty. The emphasis is on the lower-limit property of CI, because there is no upper limit for the magnitude of hidden errors (again, due to possible error cancellation). Intuitively, the CI of a line can be defined as the accuracy of its most accurate, non-trivial cycle (a trivial cycle has only two transitions with the same assignment). This specification leaves CI undefined for a line which is not part of any non-trivial cycles. Actually, for such transitions it is not meaningful to speak of a MARVEL-based validation. For a formal definition of the CI parameter and its extension to energy levels, see Supplementary Information (C).

Data sources and their treatment. There are only a limited number of data sources[60,62–64,66–69,76–81] which are available today but were not handled during the construction of the W2020-$H_2{}^{16}O$ dataset. Apart from five publications[76–80], these sources were published after 2020. In the W2020 input file, 93GuRa[82] was mistakenly referred to as '86GuRa'; this tag should refer to one of the new sources, Ref. [78]. Seven W2020 sources, 67HaDo[83], 73PuRa[84], 09GrBoRiMa[85], 12Boyarkin[86], 20virt[7], 20extra[7], and 20compl[7], have been fully removed from the present analysis. Of these sources, 09GrBoRiMa[85] and 12Boyarkin[86], which utilize multiphoton techniques to probe highly-lying states of $H_2{}^{16}O$, may well be included in a future update of W2024, when more accurate first-principles energies will be available above 30 000 cm^{-1}, allowing a reliable validation of the 09GrBoRiMa[85] and 12Boyarkin[86] lines.

Table 1 contains segments constructed from the new sources considered during this study. While certain sources not divided up in W2020 into segments were divided into multiple segments in W2024, for the sake of simplicity these segments are not specified in Table 1. Furthermore, the 24virt segment, which substitutes its predecessors in W2020, 20virt, 20virt_S2, 20virt_S3, and 20virt_S4, is not given in Table 1 either. The 14 transitions which had to be deleted from the new sources are listed in Table 2.

During the construction of the W2024 database, it became necessary to add short comments to lines which, in certain aspects, must be distinguished from other transitions of the input dataset. Accordingly, the standard format of the line tags[1], which consists of the segment name and a serial number, has been extended in this

Segment tag	Range	*A*/*D*	ESU	MSU	LSU
22KaLaChCa[63]	7168.4–7168.4	1/0/0	3.34e-09	3.34e-09	3.34e-09
24ToDiCoUb[62]	6755.3–8670.3	135/0/0	1.90e-07	1.90e-07	1.28e-06
22DiToScCo[60]	7008.4–7346.8	71/0/0	2.54e-07	2.54e-07	1.33e-06
24MiKaKoCa[256]	51.434–705.36	679/0/0	3.00e-05	4.00e-05	1.08e-03
22ToKoMiPi[81]	51.435–718.43	1310/0/0	5.00e-05	5.00e-05	3.00e-03
24KaMiKoCa[69]	51.434–721.42	1130/0/0	5.00e-05	5.00e-05	1.50e-03
86GuRa[78]	25.085–349.76	265/0/0	1.20e-04	1.50e-04	7.70e-04
86GuRa_S2[78]	501.57–713.79	85/0/1	2.30e-04	2.80e-04	8.20e-04
46ToMe[76]	0.741 75–0.741 75	1/0/0	1.67e-04	1.67e-04	1.67e-04
23KoMiKaCa[68]	8041.5–8633.4	3246/2/1	3.60e-04	4.30e-04	2.31e-01
23KoMiKaCa_S2[68]	8058.5–8065.5	1/0/0	1.00e-02	1.00e-02	1.00e-02
82EsHuSaVa[77]	1601.2–2001.0	703/2/11	4.00e-04	9.65e-04	2.12e-02
20MiKaMoCa[80]	5789.5–7841.7	11/0/0	1.00e-03	1.00e-03	1.00e-03
21VaMiCa[64]	12 969–13 172	542/0/9	1.00e-03	1.00e-03	3.48e-02
21VaMiCa_S2[64]	13 015–13 100	6/0/6	1.00e-03	1.00e-03	1.00e-03
22SoPeSoDe[65]	9344.3–9787.6	783/0/0	1.00e-03	1.00e-03	2.58e-02
97Lanqueti[79]	2373.3–3722.3	114/5/0	2.00e-03	2.00e-03	1.00e-02
22VaMiCa[66]	13 171–13 418	1265/5/15	3.00e-03	3.00e-03	3.00e-02
22VaMiCa_S2[66]	13 191–13 193	3/0/3	3.00e-03	3.00e-03	3.00e-03
22YaCoLiGo[67]	24 064–24 124	21/0/1	3.00e-03	3.00e-03	9.20e-03

Table 1. List of segments involving data sources new to W2024 compared to W2020-$H_2{}^{16}O$. [a] Tags designate data-source segments. The range, in cm^{-1}, corresponds to line positions of the segments. *A*, *D*, and *R* represent the number of assigned, deleted, and reassigned transitions within a segment, respectively. ESU, MSU, and LSU, in cm^{-1}, stand for the estimated, the median, and the largest segment uncertainties, respectively (the last two quantities characterize the final W2024 database). For the "main" segments, the '_S1' identifier is neglected from the segment tags. The introduction of the '_S2' segments is justified by two reasons: (*i*) the underlying transitions have been recorded with a different technique than those of the main segment (86GuRa_S2), and (*ii*) they contain lines which were left unassigned in the original publications (23KoMiKaCa_S2, 21VaMiCa_S2, and 22VaMiCa_S2). Note also that 23KoMiKaCa_S2 carries a line which is given only in the main text of the source 23KoMiKaCa[68], but not listed in the Supplementary Material of that paper. The *R* values in the third column include the originally unassigned transitions, as well. Some LSUs are large due to (*i*) two large original uncertainties reported in 22SoPeSoDe[65] and 23KoMiKaCa[68], as well as (*ii*) the disturbing presence of nearby lines in 21VaMiCa[64] and 22VaMiCa[64]. For the source 82EsHuSaVa[77], an optimized recalibration factor of 0.999 999 598 was determined *via* the MARVEL protocol (for details, see ref. [34]).

study with so-called markers. The principal markers used in the W2024 input file are listed in Supplementary Information (D).

As to the MARVEL treatment of the input transitions dataset, it is worth emphasizing a few important aspects. First, an *ortho–para* doublet of a segment, observed under Doppler-limited conditions, was deemed to be unresolved if the separation of the reported experimental positions were smaller than one third of the associated Doppler half width at the actual measurement temperature. Second, when the *ortho*/*para* complement of a *para*/*ortho* line was not published, then it was added to the W2024 input [see also Supplementary Information (D)]. Third, lines within unresolved multiplets other than *ortho–para* doublets were not subject to multiplet constraints, as their relative positions are usually not known accurately from first-principles computations; their confidence intervals have been increased to reflect their potential inaccuracy. Fourth, a set of empirical positions and energy levels, taken from the literature[64–66,68,70,80], was used during the refinement of the wavenumber uncertainties, to reach better agreement, wherever possible, with these auxiliary data. Upon termination of the refinement process, MARVEL was re-executed by eliminating all but 11 lines of this auxiliary dataset from the final W2024 input. The 11 empirical transitions preserved come from the source 20MiKaMoCa[80], see Table 1, which seem to rely partially on accurate unpublished experimental lines.

Data Records

The W2024 database is available in an OSF (Open Science Framework) repository[87], which contains validated transitions, empirical energy levels, and an extensive line list for the $H_2{}^{16}O$ isotopologue. In the rest of this section, a brief summary is provided about the nine files located in the file "W2024.zip" within the W2024 repository.

As customary, the W2024 repository is accompanied with a "README.txt" file, including a concise description of the content of the other files. In "README.txt", the file names are arranged in the order of their importance.

The entire collection of the 212 segments created from the 189 sources[29,41,60,62–69,74,76–82,86,88–256] are presented in "W2024_segment_table.pdf", where a couple of important statistical parameters are given for each segment, in a form similar to that of Table 1 (the difference is only that *R* is missing from "W2024_segment_table.pdf").

Position/cm^{-1}	Original assignment	Line tag	Comment
1648.265 237(400)	$(020)3_{2,2} \leftarrow (010)3_{1,3}$	82EsHuSaVa.61[77]	—
1963.728 110(400)	$(100)7_{2,5} \leftarrow (010)7_{3,4}$	82EsHuSaVa.626[77]	—
2629.948 4(20)	$(001)30_{0,30} \leftarrow (010)29_{0,29}$	97Lanqueti.19[79]	—
3513.305 6(20)	$(011)15_{15,1} \leftarrow (010)15_{15,0}$	97Lanqueti.53[79]	—
3490.139 6(20)	$(011)16_{16,0} \leftarrow (010)16_{16,1}$	97Lanqueti.54[79]	—
3467.102 8(20)	$(011)17_{17,1} \leftarrow (010)17_{17,0}$	97Lanqueti.55[79]	—
2649.340(10)	$(100)20_{20,1} \leftarrow (000)21_{21,0}$	97Lanqueti.56[79]	—
8203.349 62(260)	$(051)5_{1,5} \leftarrow (020)4_{1,4}$	23KoMiKaCa.783[68]	HD^{16}O line
8529.592 96(29)	$(050)11_{4,8} \leftarrow (000)10_{1,9}$	23KoMiKaCa.2547[68]	bad multiplet
13 183.545 9(30)	$(122)11_{2,9} \leftarrow (000)12_{7,6}$	22VaMiCa.45[66]	HD^{16}O line
13 214.22 73(30)	$(240)8_{7,1} \leftarrow (000)9_{8,2}$	22VaMiCa.148[66]	bad multiplet
13 280.38 49(30)	$(061)6_{3,4} \leftarrow (000)5_{1,5}$	22VaMiCa.393[66]	bad multiplet
13 306.91 39(30)	$(042)12_{3,10} \leftarrow (000)13_{2,11}$	22VaMiCa.533[66]	bad multiplet
13 324.30 68(30)	$(061)5_{5,0} \leftarrow (000)5_{5,1}$	22VaMiCa.636[66]	bad multiplet

Table 2. Complete list of experimental lines deleted from the new sources given in Table 1. [a]The values in the first column are observed line positions, with the uncertainties of the last few digits in parentheses. The second column provides the original assignments of the deleted lines which could not be corroborated during the validation of the W2024-$H_2{}^{16}O$ input dataset. In the third column, the line tags used within the W2024 input file are itemized. Where applicable, the last column contains comments on why these lines proved to be incorrect. 'Bad multiplet' means that the given line cannot be part of an unresolved multiplet, as incorrectly indicated in its data source.

The file "W2024_segments.txt" is the segment input file for the MARVEL code, where the unit of the line positions and their uncertainties are specified for each segment. The file "W2024_transitions.txt" contains the 309 290 input transitions collected for the MARVEL procedure. In this file, each input transition is associated with (a) a line position, (b) an initial and an adjusted line-position uncertainty, (c) a $(v_1' \; v_2' \; v_3')J'_{K_a',K_c'} \leftarrow (v_1'' \; v_2'' \; v_3'')J''_{K_a'',K_c''}$ rovibrational assignment, and (d) a line tag representing a unique identifier.

The empirical energy values, obtained for 19 027 rovibrational states in the 0–26 268 cm^{-1} range, are placed in the file "W2024_energy_levels.txt". Each state of this data file is supplied with (a) a $(v_1 \; v_2 \; v_3)J_{K_a,K_c}$ label, (b) an empirical (MARVEL) energy, (c) an energy uncertainty followed by a (relative) confidence interval in parentheses, (d) the number of transitions incident to this state, and (e) the index of the respective POKAZATEL[74] state.

The file "W2024-24MiVaCa_comparison.xls" lists 57 states, for which the W2024 and the 24MiVaCa[70] energies deviate by more than 0.005 cm^{-1} or their assignments are different. For each line a short comment is given indicating a potential reason for the discrepancy.

Using empirical (W2024) and first-principles (POKAZATEL[74]) energies, a composite line list, named CW2024, was constructed, forming part of the file "CW2024_line_list.txt". This line list consists of more than 490 000 dipole-allowed transitions in the 0–41 200 cm^{-1} range, with room-temperature intensities down to 10^{-31} cm molecule^{-1}. For almost half of the CW2024 entries, that is for about 231 000 lines in the 0.07–25681.5 cm^{-1} region, empirical positions are reported; all of them are augmented with individual wavenumber uncertainties and (relative) confidence intervals. For all of the CW2024 lines, the intensities are taken from the POKAZATEL line list, complemented with their BT2[75] counterparts, whenever applicable. For the empirical transitions of this list, essential walks are also provided in "CW2024_walk_file.txt". These walks help to understand how the empirical positions and their uncertainties can be approximately extracted from a handful of W2024 input lines [for details on the use of walks, see Supplementary Information (B)]. A line-by-line comparison between HITRAN 2020 and the (C)W2024 dataset is presented in the file "HITRAN_comparison.txt", which will be discussed in the "Technical Validation" part of this paper.

Finally, the "MARVEL.zip" file contains a developer version of the MARVEL code, written in the C++ language. This version of the MARVEL code, distributed with the necessary input files, was used to generate the numerical data in the TXT files of the W2024 repository (except the input data listed in "W2024_transitions.txt"). The novel MARVEL features, implemented in this code version and described in the "Methods" section, will form part of the http://kkrk.chem.elte.hu/marvelonline/MARVELOnline web application in the future.

Technical Validation

Validation of the W2024 energy levels. The principal validation of the W2024 energy levels was performed *via* MARVEL, by checking the consistency of the input transitions in relation to their assignments, wavenumbers, and uncertainties. This process resulted in a self-consistent energy-level dataset with individual uncertainties and confidence intervals.

The W2024 energy levels have been matched with their first-principles (BT2[75], POKAZATEL[74], and VoTe[257]) counterparts, making use of the $|E^{\mathrm{W2024}} - E^{\mathrm{comp}}| \leq 10^{-4}E^{\mathrm{comp}}$ criterion, where E^{W2024} and E^{comp} denote empirical and computed energy values, respectively. Despite previous efforts[74,75,257–259], no unambiguous labelling scheme exists for water isotopologues and, indeed, it is unlikely that such a scheme could be developed[259,260]. Owing to a number of notable differences in rovibrational assignments across the three datasets, only the J/symmetry labels were used during the formation of the (E^{W2024}, E^{comp}) pairs.

Source tag	New states	Most accurate new state	
	All [*ortho*/*para*]	Assignment	Energy / cm^{-1}
20MiKaMoCa[80]	3 [2/1]	$(0\,6\,0)1_{1,0}$	9 004.617 0(10)
21VaMiCa[64]	119 [68/51]	$(0\,9\,0)2_{1,2}$	13 093.354 7(10)
22VaMiCa[66]	147 [67/80]	$(0\,4\,2)2_{1,1}$	13 566.440 4(30)
22YaCoLiGo[67]	18 [14/4]	$(2\,2\,4)3_{3,0}$	24 200.455 7(22)
22SoPeSoDe[65]	27 [11/16]	$(0\,6\,0)7_{3,5}$	10 158.506 0(10)
23KoMiKaCa[68]	27 [9/18]	$(1\,3\,0)10_{6,4}$	10 458.465 32(12)
24virt (Ref. [74]; this work)	211 [15/196]	$(1\,4\,0)9_{7,3}$	12 148.651 8(10)

Table 3. Contribution of the new sources to the set of new W2024 energy levels. [a] If a new state occurs in more than one of the new sources, it is attributed to the oldest source. When a source contains multiple matches for the most accurate new state, the lowest-energy match is given in the last two columns.

Assignment	Rovibrational energy/cm^{-1}			Source of the best incident line		
	Theory[74]	W2024	W2020[8]	W2024	W2020[8] (comment on best line)	
$(0\,6\,0)7_{1,7}$	9 539.20	9 539.185 77(69)	9 539.368 7(18)[−0.18]	23KoMiKaCa[68]	14ReOuMiWa[225]	(deleted)
$(0\,6\,0)7_{1,6}$	9 714.54	9 714.490 67(86)	9 714.758 8(18)[−0.27]	23KoMiKaCa[68]	14ReOuMiWa[225]	(reassigned)
$(0\,6\,0)11_{2,9}$	10 826.06	10 825.988 1(10)	10 826.152 2(33)[−0.16]	22SoPeSoDe[65]	05CoBeCaCo[180]	(deleted, bmult)
$(0\,9\,0)5_{2,3}$	13 894.80	13 894.760 9(30)	13 895.032 4(40)[−0.27]	22VaMiCa[66]	08CaMiLi[200]	(reassigned)
$(0\,4\,2)7_{0,7}$	14 025.94	14 025.951 1(30)	14 026.309 3(18)[−0.36]	22VaMiCa[66]	08ToTe[203]	(reassigned)
$(0\,4\,2)6_{3,3}$	14 228.79	14 228.816 3(30)	14 229.149 9(40)[−0.33]	22VaMiCa[66]	08CaMiLi[200]	(reassigned)
$(4\,0\,0)10_{6,4}$	15 539.41	15 539.547 6(30)	15 539.209 1(80)[0.34]	22VaMiCa[66]	11BeMiCa[208]	(deleted)
$(2\,2\,1)13_{2,12}$	15 615.18	15 615.198 4(30)	15 615.497 2(40)[−0.30]	22VaMiCa[66]	08CaMiLi[200]	(deleted, typo)
$(2\,0\,2)11_{2,9}$	15 788.77	15 788.944 0(30)	15 789.218 7(80)[−0.27]	22VaMiCa[66]	11BeMiCa[208]	(deleted)
$(3\,0\,1)11_{7,5}$	15 905.13	15 905.325 7(30)	15 904.862(18)[0.46]	22VaMiCa[66]	08ZoShOvPo[204]	(deleted, bmult)

Table 4. Ten selected W2024 energy levels deviating by more than 0.1 cm^{-1} from their W2020 counterparts. [a]The rovibrational energies, tabulated in columns 2–4, are extracted from the POKAZATEL[74], W2024, and W2020[8] energy lists. In all cases, the POKAZATEL energies agree better with the W2024 estimates than with the W2020 ones. The values in parentheses represent last-two-digit uncertainties, whereas those in brackets are energy shifts characterizing the W2020 energies. In the last two columns, each state is associated with the source of its most accurate line in the W2024 and W2020 databases. As the comments in the last column suggest, the lines yielding the incorrect W2020 energies were either deleted or reassigned in W2024. In the last column, 'typo' indicates that the problematic 08CaMiLi[200] entry probably contains a typographical error, while 'bmult' denotes that the lines leading to the incorrect W2020 energies form part of 'bad' (spurious) multiplets.

Consistency of the W2024 energy levels was also checked *via* the pair identity and smooth variation rules of Ma *et al.*[261]. For each vibrational state and J, a plot was made of the energy versus the K_a quantum number. These plots were studied to insure their correct pairing structures and smooth variations. Everything checked out correctly, giving further confidence in the correctness of the W2024 assignments and empirical rovibrational energies.

Compared to W2020-$H_2{}^{16}O$, the W2024-$H_2{}^{16}O$ dataset deals with only a small number of new data sources. Nevertheless, it contains more than 500 new empirical rovibrational energy levels. How each new source contributes to the set of new energy levels is given in Table 3. Not too surprisingly, the largest contributor is the source 24virt, yielding empirical energies for the *ortho*/*para* complements of over 200 *para*/*ortho* state pairs. Note also that a few additional energy levels, not reflected in the numbers given in Table 3, were also obtained from the set of more than 1000 transitions reassigned during this study. Consideration of new sources has particular relevance when they provide new energy levels or help to determine improved empirical energy values and/or uncertainties for states already available. For the latter case, the highly accurate sources listed in the first few rows of Table 1 proved to be particularly useful.

A detailed comparison of the W2024 energy levels with their W2020 counterparts reveals occasional significant shifts, displayed in Table 4, in previously known energy values. As shown there, not only the less accurate emission sources, like 08ZoShOvPo[204] and 05CoBeCaCo[180], but some of the more dependable absorption sources, namely 08ToTe[203], 08CaMiLi[200], 11BeMiCa[208], and 14ReOuMiWa[225], produced a few unreliable energies, as well. All in all, there are only about 500 cases where the W2024 – W2020 deviations fall outside of the W2020/W2024 uncertainties.

To provide a comprehensive picture about the collection of W2024 energy levels, their distributions are plotted against the rovibrational energies and their uncertainties in Fig. 1. For reference purposes, the energy distribution of the first-principles POKAZATEL[74] states used, forming a complete set in the 0–26 268 cm^{-1} range investigated, is also given. As obvious from Fig. 1, (a) all states are known in W2024 up to 8995 cm^{-1}, (b) the number of missing empirical states increases rapidly as the energy increases, (c) for a significant number of

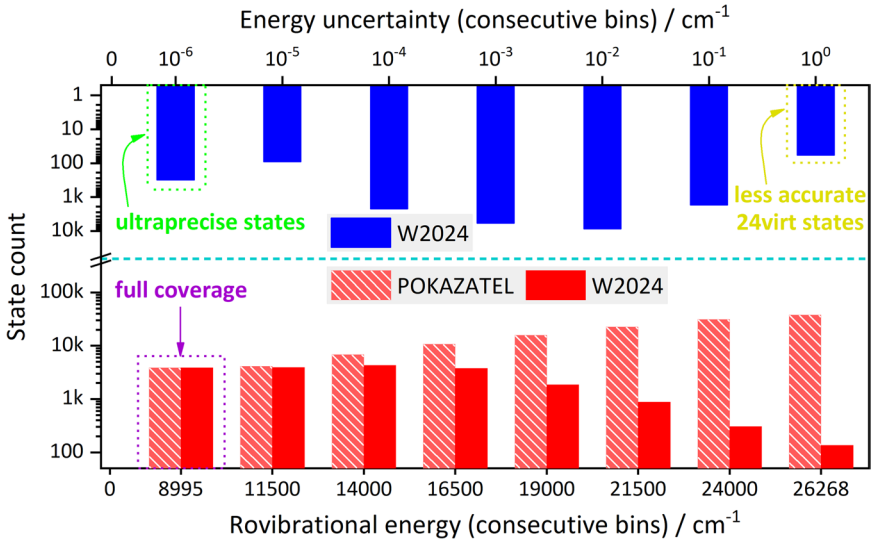

Fig. 1 Distribution of the W2024 and POKAZATEL states along the rovibrational energies and their uncertainties. The lower panel gives the distribution of the energy values for the W2024 and the POKAZATEL datasets. The upper panel provides the distribution of W2024 states by uncertainties. The range represented by a bin is given by the actual and the previous axis ticks (*e.g.*, the blue bin at 10^{-4} cm^{-1} contains empirical states with an uncertainty of 10^{-5}–10^{-4} cm^{-1}). The state counts, that is the bin sizes, are plotted on a unified bi-directed vertical axis for both distributions. Note the logarithmic scale on both parts of the vertical axis.

Indicator	Data type	Indicator value			
		A	B	C	D
Q_1	line match	**unique**	**ambiguous**	**expt. only**	**none**
Q_2	position	**approved**	**acceptable**	**change/conflict**	**unverified**
Q_3	uncertainty code (position)	**approved**	**undefined**	**change/conflict**	**unverified**
Q_4	intensity	**approved**	**acceptable**	**change/conflict**	**unverified**
Q_5	assignment	**approved**	**undefined**	**change/conflict**	**unverified**
Q_6	lower-state energy	**approved**	**acceptable**	**change/conflict**	**unverified**
Comments to Q_1[b]:		**BADSYM**	**QUADRUP**	**SMALLINT**	**DUPSTATE**

Table 5. Quality indicators and comment categories used to characterize the HITRAN-$H_2{}^{16}O$ lines. [a]The uncertainty codes of the positions are specified on the HITRANonline page. The term "expt. only" refers to a unique match with a measured W2024 line not present in the CW2024 list. [b]Meaning of comments made for certain ill-matched HITRAN lines: (*i*) BADSYM = symmetry violation, (*ii*) QUADRUP = quadrupole transition, (*iii*) SMALLINT = a line with too small ($<10^{-31}$cm molecule^{-1}) intensity, and (*iv*) DUPSTATE = a transition whose upper state is deleted from W2024.

states the energies are known with an accuracy better than 10^{-6} cm^{-1}, and (d) a few W2024 states, deduced from some less accurate 24virt lines, have relatively large, 0.1 – 0.7 cm^{-1}, uncertainties (such states could be targeted by future measurements).

Comparison with HITRAN 2020. Comparing the (C)W2024 and the HITRAN 2020-$H_2{}^{16}O$ line catalogs is particularly important, as it allows additional validation of the W2024 database; furthermore, it might reveal HITRAN entries which require further verification/modification. Results of this comparison are discussed next, without reliance on BT2 intensities.

To facilitate the comparison of the HITRAN 2020-$H_2{}^{16}O$ line list with the (C)W2024 dataset, an attempt was made to set up a simple quality-assessment scheme for six HITRAN data types which are also present in CW2024. A six-character quality sequence, Q_1-Q_2-Q_3-Q_4-Q_5-Q_6, has been introduced, where Q_p symbolizes the pth quality indicator (QI). Intuitive definitions for the four possible values, A–D, of the six QIs are included in Table 5. Briefly, (*i*) "A" is the best category, (*ii*) "B" means acceptable, given the present knowledge, (*iii*) "C" indicates a conflict between (C)W2024 and HITRAN, which is most probably due to the incorrectness of the HITRAN entry, and (*iv*) "D" means that no verification was possible for a data type. Table 5 also lists four "comments" attached by us to a few peculiar HITRAN 2020 lines. For further details on the QI values, see Supplementary Information (E).

HITRAN position/cm^{-1}	Unc. code	Intensity/ cm molecule^{-1}	Lower-state energy / cm^{-1}	Assignment	Quality sequence
2 388.014 241	3	2.560e-27	3 266.76 43	$(0\,1\,0)13_{11,3} \leftarrow (0\,0\,0)12_{10,2}$	A-B-C-A-A-B
9 055.063 450	4	1.312e-29	4 021.217 9	$(1\,1\,1)20_{2,19} \leftarrow (0\,0\,0)19_{2,18}$	A-C-C-A-A-C
12 058.339 388	7	1.902e-29	816.694 2	$(1\,6\,0)6_{0,6} \leftarrow (0\,0\,0)7_{3,5}$	D-D-D-D-D-D(DUPSTATE)
13 990.042 460	3	3.138e-27	756.724 8	$(3\,2\,0)7_{5,2} \leftarrow (0\,0\,0)6_{4,3}$	A-C-C-C-A-A
15 344.739 600	7	3.236e-28	212.156 4	$(4\,1\,0)4_{-4,-4} \leftarrow (0\,0\,0)3_{2,1}$	A-C-C-C-C-A
16 823.549 400	7	6.640e-25	285.219 3	$(3\,2\,1)3_{1,2} \leftarrow (0\,0\,0)3_{3,1}$	A-C-C-A-C-A

Table 6. Typical examples for problematic transitions in the HITRAN 2020-$H_2{}^{16}O$ list. [a]The first two columns give HITRAN positions and their associated uncertainty codes. The intensity values, the lower-state energies, and the assignments are taken from HITRAN. The negative subscripts in the penultimate row mean that those K_a and K_c values are undefined in HITRAN. For these six lines, the quality sequences, reflecting only the level of agreement between the CW2024 line list and HITRAN 2020, are also exhibited in the last column. Although all these lines are claimed to be extracted from W2020[8], the underlined digits of this table do not correspond to W2020 information.

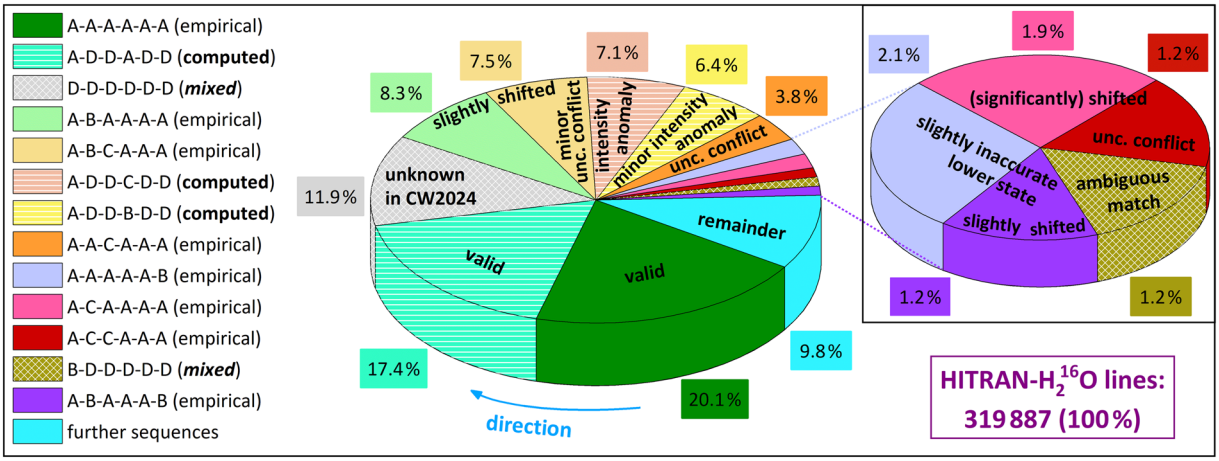

Fig. 2 Distribution of the most frequent quality sequences, covering 90 % of the HITRAN 2020-$H_2{}^{16}O$ lines. The color codes applied for the quality sequences are shown on the left-hand side of this figure, where slices formed by empirical, computed, or mixed (empirical plus computed) HITRAN 2020 transitions are clearly distinguished. The blue arrow indicates the direction whereby the slices follow the ordering utilized in the color legend. To highlight its most important characteristics, each slice is supplied with short stamps, displayed on the pie charts (notice that there are stamps shared between two slices). The relative fractions of the individual slices, with respect to the complete HITRAN 2020-$H_2{}^{16}O$ database, are included in colored boxes. Those lines which could not be matched with the (C)W2024 line list appear in the gray slice. The remaining transitions, whose quality sequences differ from those exhibited in the color legend, are collected in the cyan slice. For improved transparency, slices with less than 10 000 lines are enlarged in an inset.

During this comparison, more than 2500 "W2020" transitions have been identified in HITRAN 2020 for which either the position or the rovibrational assignment differ significantly from that contained in the W2020 database[8]. Table 6 gives six characteristic examples for such questionable HITRAN 2020 transitions. For these lines, certain parameter values were seemingly incorrectly transcribed from W2020.

Figure 2 depicts the distribution of the 13 most common quality sequences, corresponding to 90% of the HITRAN 2020-$H_2{}^{16}O$ lines. The good news is that the leading sequence is "6A", see the dark green slice in Fig. 2, where all the six data types of Table 5 are corroborated by (C)W2024. Nevertheless, there is a considerable number of transitions which require particular attention, and may lead, after additional validation, to corrections of certain HITRAN 2020 entries. Lines falling into the "gray zone", with a sequence "6D", for which none of the six HITRAN parameters could be affirmed by (C)W2024, must be investigated carefully. Note in this respect that POKAZATEL intensities are highly accurate in the infrared, but get increasingly inaccurate as one moves toward visible wavelengths[262,263]. For the full list of quality sequences attached to the HITRAN 2020-$H_2{}^{16}O$ lines, see ref. [87].

Usage Notes

The empirical energy levels derived during this study are associated with individual uncertainties and confidence intervals, giving numerical characterization of the trust we have in the W2024 energies, as well as in the predicted transition wavenumbers. These important statistical parameters must be taken into account in

applications using the (C)W2024 datasets. The set of new and corrected empirical energies of this study could, for example, prove useful for adjusting existing potential energy surfaces of the $H_2{}^{16}O$ molecule, reducing the discrepancies between the results of variational nuclear-motion computations and experiment.

The CW2024 database could be helpful for experimental spectroscopists, who wish to (re)analyze their new and old spectra, especially when looking for new energy levels absent from the W2024 energy list. This CW2024 catalog would also provide support for the validation and occasional correction of $H_2{}^{16}O$ transitions present in line-by-line spectroscopic databases.

Updated versions of the database files will be made available at the website https://respecth.elte.hu/. Version history will be provided in a file called "NOTES.txt" under the W2024 repository[87].

Code availability

The developer version of the MARVEL code, used during the compilation and validation of the W2024 database, is freely available[87].

Received: 13 June 2024; Accepted: 28 August 2024;
Published online: 28 September 2024

References

1. Tennyson, J. *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. Rad. Transf.* **110**, 573–596 (2009).
2. Tennyson, J. *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. Rad. Transf.* **111**, 2160–2184 (2010).
3. Tennyson, J. *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. Rad. Transf.* **117**, 29–80 (2013).
4. Tennyson, J. *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. Rad. Transf.* **142**, 93–108 (2014).
5. Tennyson, J. *et al.* A database of water transitions from experiment and theory (IUPAC technical report). *Pure Appl. Chem.* **86**, 71–83 (2014).
6. Gordon, I. E. *et al.* The HITRAN 2016 molecular spectroscopic database. *J. Quant. Spectrosc. Rad. Transf.* **203**, 3–69 (2017).
7. Furtenbacher, T., Tóbiás, R., Tennyson, J., Polyansky, O. L. & Császár, A. G. W2020: A database of validated rovibrational experimental transitions and empirical energy levels of $H_2{}^{16}O$. *J. Phys. Chem. Ref. Data* **49**, 033101 (2020).
8. Furtenbacher, T. *et al.* The W2020 Database of Validated Rovibrational Experimental Transitions and Empirical Energy Levels of Water Isotopologues. Part II. $H_2{}^{17}O$ and $H_2{}^{18}O$ with an Update to $H_2{}^{16}O$. *J. Phys. Chem. Ref. Data* **49**, 043103 (2020).
9. Delahaye, T. *et al.* The 2020 edition of the GEISA spectroscopic database. *J. Mol. Spectrosc.* **380**, 111510 (2021).
10. Gordon, I. E. *et al.* The HITRAN2020 molecular spectroscopic database. *J. Quant. Spectrosc. Rad. Transf.* **276**, 107949 (2022).
11. Wayne, R. P. Chemistry of Atmospheres (Oxford University Press, New York, 2000).
12. Bernath, P. F. The spectroscopy of water vapour: experiment, theory and applications. *Phys. Chem. Chem. Phys.* **4**, 1501–1509 (2002).
13. Boone, C. D., Walker, K. A. & Bernath, P. F. Speed-dependent Voigt profile for water vapor in infrared remote sensing applications. *J. Quant. Spectrosc. Rad. Transf.* **105**, 525–532 (2007).
14. Berden, G., Peeters, R. & Meijer, G. Cavity ring-down spectroscopy: Experimental schemes and applications. *Int. Rev. Phys. Chem.* **19**, 565–607 (2000).
15. Hall, J. L. Nobel lecture: defining and measuring optical frequencies. *Rev. Mod. Phys.* **78**, 1279–1295 (2006).
16. Hänsch, T. W. Nobel lecture: passion for precision. *Rev. Mod. Phys.* **78**, 1297–1309 (2006).
17. Quack, M. & Merkt, F. (Eds.), Handbook of High-Resolution Spectroscopy (Wiley, Chichester, 2011).
18. Twagirayezu, S., Cich, M. J., Sears, T. J., McRaven, C. P. & Hall, G. E. Frequency-comb referenced spectroscopy of ν_4- and ν_5-excited hot bands in the 1.5 μm spectrum of C_2H_2. *J. Mol. Spectrosc.* **316**, 64–71 (2015).
19. Santamaria, L. *et al.* Comb-assisted cavity ring-down spectroscopy of a buffer-gas-cooled molecular beam. *Phys. Chem. Chem. Phys.* **18**, 16715–16720 (2016).
20. Gatti, D. *et al.* Comb-locked Lamb-dip spectrometer. *Sci. Rep.* **6**, 27183 (2016).
21. Wang, J. *et al.* Comb-locked cavity ring-down saturation spectroscopy. *Rev. Sci. Instrum.* **88**, 043108 (2017).
22. Reed, Z. D., Long, D. A., Fleurbaey, H. & Hodges, J. T. SI-traceable molecular transition frequency measurements at the 10^{-12} relative uncertainty level. *Optica* **7**, 1209–1220 (2020).
23. Császár, A. G. *et al.* The fourth age of quantum chemistry: molecules in motion. *Phys. Chem. Chem. Phys.* **14**, 1085–1106 (2012).
24. Tennyson, J. *et al.* The ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Mol. Spectrosc.* **327**, 73–94 (2016).
25. Tennyson, J. & Yurchenko, S. N. The ExoMol project: Software for computing large molecular line lists. *Int. J. Quant. Chem.* **117**, 92–103 (2017).
26. Tennyson, J. *et al.* The 2020 release of the ExoMol database: Molecular line lists for exoplanet and other hot atmospheres. *J. Quant. Spectrosc. Rad. Transf.* **255**, 107228 (2020).
27. Tennyson, J. *et al.* The 2024 release of the ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Quant. Spectrosc. Rad. Transf.* **326**, 109083 (2024).
28. Pachucki, K. & Komasa, J. Rovibrational levels of HD. *Phys. Chem. Chem. Phys.* **12**, 9188–9196 (2010).
29. Tóbiás, R. *et al.* Spectroscopic-network-assisted precision spectroscopy and its application to water. *Nat. Commun.* **11**, 1708 (2020).
30. Császár, A. G., Czakó, G., Furtenbacher, T. & Mátyus, E. An active database approach to complete spectra of small molecules. *Annu. Rep. Comput. Chem.* **3**, 155–176 (2007).
31. Furtenbacher, T., Császár, A. G. & Tennyson, J. MARVEL: measured active rotational-vibrational energy levels. *J. Mol. Spectrosc.* **245**, 115–125 (2007).
32. Császár, A. G. & Furtenbacher, T. Spectroscopic networks. *J. Mol. Spectrosc.* **266**, 99–103 (2011).
33. Császár, A. G., Furtenbacher, T. & Árendás, P. Small molecules – big data. *J. Phys. Chem. A* **120**, 8949–8969 (2016).
34. Tóbiás, R., Furtenbacher, T., Tennyson, J. & Császár, A. G. Accurate empirical rovibrational energies and transitions of $H_2{}^{16}O$. *Phys. Chem. Chem. Phys.* **21**, 3473–3495 (2019).
35. Furtenbacher, T. & Császár, A. G. On employing $H_2{}^{16}O$, $H_2{}^{17}O$, $H_2{}^{18}O$, and $D_2{}^{16}O$ lines as frequency standards in the 15–170 cm^{-1} window. *J. Quant. Spectrosc. Rad. Transf.* **109**, 1234–1251 (2008).
36. Furtenbacher, T. & Császár, A. G. MARVEL: measured active rotational-vibrational energy levels. II. Algorithmic improvements. *J. Quant. Spectrosc. Rad. Transf.* **113**, 929–935 (2012).

37. Furtenbacher, T. & Császár, A. G. The role of intensities in determining characteristics of spectroscopic networks. *J. Mol. Spectrosc.* **1009**, 123–129 (2012).
38. Furtenbacher, T., Árendás, P., Mellau, G. & Császár, A. G. Simple molecules as complex systems. *Sci. Rep.* **4**, 4654 (2014).
39. Tóbiás, R., Furtenbacher, T. & Császár, A. G. Cycle bases to the rescue. *J. Quant. Spectrosc. Rad. Transf.* **203**, 557–564 (2017).
40. Tóbiás, R., Bérczi, K., Szabó, C. & Császár, A. G. autoECART: automatic energy conservation analysis of rovibronic transitions. *J. Quant. Spectrosc. Rad. Transf.* **272**, 107756 (2021).
41. Flaud, J.-M., Camy-Peyret, C. & Maillard, J. P. Higher ro-vibrational levels of H_2O deduced from high resolution oxygen-hydrogen flame spectra between 2800–6200 cm^{-1}. *Mol. Phys.* **32**, 499–521 (1976).
42. Furtenbacher, T. *et al*. Experimental energy levels and partition function of the $^{12}C_2$ molecule. *Astrophys. J. Suppl. S.* **224**, 44 (2016).
43. McKemmish, L. K. *et al*. MARVEL analysis of the measured high-resolution rovibronic spectra of $^{48}Ti^{16}O$. *Astrophys. J. Suppl. S.* **228**, 15 (2017).
44. McKemmish, L. K. *et al*. MARVEL analysis of the measured high-resolution rovibronic spectra of $^{90}Zr^{16}O$. *Astrophys. J.* **867**, 33 (2018).
45. Furtenbacher, T. *et al*. MARVEL analysis of the measured high-resolution rovibronic spectra and definitive ideal-gas thermochemistry of the $^{16}O_2$ molecule. *J. Phys. Chem. Ref. Data* **48**, 023101 (2019).
46. Darby-Lewis, D. *et al*. MARVEL analysis of the measured high-resolution spectra of ^{14}NH. *J. Mol. Spectrosc.* **362**, 69–76 (2019).
47. McKemmish, L. K. *et al*. An update to the MARVEL data set and ExoMol line list for $^{12}C_2$. *Mon. Not. R. Astron. Soc.* **497**, 1081–1097 (2020).
48. Furtenbacher, T., Szidarovszky, T., Fábri, C. & Császár, A. G. MARVEL analysis of the rotational-vibrational states of the molecular Ions H_2D^+ and D_2H^+. *Phys. Chem. Chem. Phys.* **15**, 10181–10193 (2013).
49. Furtenbacher, T., Szidarovszky, T., Mátyus, E., Fábri, C. & Császár, A. G. Analysis of the rotational-vibrational states of the molecular ion H_3^+. *J. Chem. Theor. Comput.* **9**, 5471–5478 (2013).
50. Chubb, K. L. *et al*. MARVEL analysis of the measured high-resolution rovibrational spectra of $H_2{}^{32}S$. *J. Quant. Spectrosc. Rad. Transf.* **218**, 178–186 (2018).
51. Tóbiás, R. *et al*. Critical evaluation of measured rotational-vibrational transitions of four sulphur isotopologues of $S^{16}O_2$. *J. Quant. Spectrosc. Rad. Transf.* **208**, 152–163 (2018).
52. Tennyson, J., Furtenbacher, T., Yurchenko, S. N. & Császár, A. G. Empirical rovibrational energy levels for nitrous oxide. *J. Quant. Spectrosc. Rad. Transf.* **316**, 108902 (2024).
53. Al Derzi, A. R., Furtenbacher, T., Yurchenko, S. N., Tennyson, J. & Császár, A. G. MARVEL analysis of the measured high-resolution spectra of $^{14}NH_3$. *J. Quant. Spectrosc. Rad. Transf.* **161**, 117–130 (2015).
54. Chubb, K. L. *et al*. MARVEL analysis of the measured high-resolution spectra of C_2H_2. *J. Quant. Spectrosc. Rad. Transf.* **204**, 42–55 (2018).
55. Furtenbacher, T. *et al*. Empirical rovibrational energy levels of ammonia up to 7500 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **251**, 107027 (2020).
56. Fábri, C. *et al*. Variational quantum mechanical and active database approaches to the rotational-vibrational spectroscopy of ketene, H_2CCO. *J. Chem. Phys.* **135**, 094307 (2011).
57. Kefala, K., Boudon, V., Yurchenko, S. N. & Tennyson, J. Empirical rovibrational energy levels for methane. *J. Quant. Spectrosc. Rad. Transf.* **316**, 108897 (2024).
58. Diouf, M. L. *et al*. Network-based design of near-infrared Lamb-dip experiments and the determination of pure rotational energies of $H_2{}^{18}O$ at kHz accuracy. *J. Phys. Chem. Ref. Data* **50**, 023106 (2021).
59. Melosso, M. *et al*. Hyperfine-resolved near-infrared spectra of $H_2{}^{17}O$. *J. Phys. Chem. A* **125**, 7884–7890 (2021).
60. Diouf, M. L. *et al*. Ultraprecise relative energies in the (2 0 0) vibrational band of $H_2{}^{16}O$. *Mol. Phys.* **120**, e2050430 (2022).
61. Diouf, M. L. *et al*. Parity-pair-mixing effects in nonlinear spectroscopy of HDO. *Opt. Express* **30**, 46040–46059 (2022).
62. Tóbiás, R., Diouf, M. L., Cozijn, F. M., Ubachs, W. & Császár, A. G. All paths lead to hubs in the spectroscopic networks of water isotopologues $H_2{}^{16}O$ and $H_2{}^{18}O$. *Commun. Chem.* **7**, 34 (2024).
63. Kassi, S., Lauzin, C., Chaillot, J. & Campargue, A. The (2-0) R(0) and R(1) transition frequencies of HD determined to a 10^{-10} relative accuracy by Doppler spectroscopy at 80 K. *Phys. Chem. Chem. Phys.* **24**, 23164–23172 (2022).
64. Vasilchenko, S., Mikhailenko, S. N. & Campargue, A. Water vapor absorption in the region of the oxygen A-band near 760 nm. *J. Quant. Spectrosc. Rad. Transf.* **275**, 107847 (2021).
65. Solodov, A. M., Petrova, T. M., Solodov, A. A., Deichuli, V. M. & Naumenko, O. V. FT spectroscopy of water vapor in the 0.9 μm transparency window. *J. Quant. Spectrosc. Rad. Transf.* **293**, 108389 (2022).
66. Vasilchenko, S., Mikhailenko, S. N. & Campargue, A. Cavity ring down spectroscopy of water vapour near 750 nm: a test of the HITRAN2020 and W2020 line lists. *Mol. Phys.* **120**, e2051762 (2022).
67. Yang, Q.-Y. *et al*. Cavity ring-down spectroscopy of water vapor in the deep-blue region. *Atmos. Meas. Tech.* **15**, 4463–4472 (2022).
68. Koroleva, A., Mikhailenko, S. N., Kassi, S. & Campargue, A. Frequency comb-referenced cavity ring-down spectroscopy of natural water between 8041 and 8633 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **298**, 108489 (2023).
69. Karlovets, E. V., Mikhailenko, S. N., Koroleva, A. O. & Campargue, A. Water vapor absorption spectroscopy and validation tests of databases in the far-infrared (50–720 cm^{-1}). Part 2: $H_2{}^{17}O$ and HD ^{17}O. *J. Quant. Spectrosc. Rad. Transf.* **314**, 108829 (2024).
70. Mikhailenko, S., Vasilchenko, S. & Campargue, A. A recommended line list for water vapor in the 12969–13418 cm^{-1} interval. *J. Quant. Spectrosc. Rad. Transf.* **326**, 109099 (2024).
71. Brett, C. M. A. *et al*. Quantities, Units and Symbols in Physical Chemistry: 4th Edition, Abridged Version (Royal Society of Chemistry, 2023).
72. Ritz, W. On a new law of series spectra. *Astrophys. J.* **28**, 237–243 (1908).
73. Pickett, H. M. The fitting and prediction of vibration-rotation spectra with spin interactions. *J. Mol. Spectrosc.* **148**, 371–377 (1991).
74. Polyansky, O. L. *et al*. ExoMol molecular line lists XXX: A complete high-accuracy line list for water. *Mon. Not. R. Astron. Soc.* **480**, 2597–2608 (2018).
75. Barber, R. J., Tennyson, J., Harris, G. J. & Tolchenov, R. N. A high-accuracy computed water line list. *Mon. Not. R. Astron. Soc.* **368**, 1087–1094 (2006).
76. Townes, C. H. & Merritt, F. R. Water spectrum near one-centimeter wave-length. *Phys. Rev.* **70**, 558–559 (1946).
77. Esplin, M. P., Huppi, R. J., Sakai, H., Vanasse, G. A. & Rothman, L. S. Absorption measurements of CO_2 and H_2O at high resolution and elevated temperatures (1982). Tech. Rep. AFGL-TR-82-0057 (Utah State University).
78. Guelachvili, G. & Rao, K. N.Handbook of Infrared Standards (Academic Press, Orlando, 1986).
79. Lanquetin, R. Spectrométrie de Fourier de flammes à basse pression et analyse des spectres d'émission de la vapeur d'eau à haute température (1997). Universite de Paris Nord, Paris.
80. Mikhailenko, S. N., Kassi, S., Mondelain, D. & Campargue, A. Water vapor absorption between 5690 and 8340 cm^{-1}: accurate empirical line centers and validation tests of calculated line intensities. *J. Quant. Spectrosc. Rad. Transf.* **245**, 106840 (2020).
81. Toureille, M., Koroleva, A. O., Mikhailenko, S. N., Pirali, O. & Campargue, A. Water vapor absorption spectroscopy and validation tests of databases in the far-infrared (50–720 cm^{-1}). Part 1: Natural water. *J. Quant. Spectrosc. Rad. Transf.* **291**, 108326 (2022).
82. Guelachvili, G. & Rao, K. N.Handbook of Infrared Standards II (Academic Press, Orlando, 1993).
83. Hall, R. T. & Dowling, J. M. Pure rotational spectrum of water vapor. *J. Chem. Phys.* **47**, 2454–2461 (1967).

84. Pugh, L. A. & Narahari Rao, K. Spectrum of water vapor in the 1.9 and 2.7 μm regions. *J. Mol. Spectrosc.* **47**, 403–408 (1973).
85. Grechko, M. *et al*. State-selective spectroscopy of water up to its first dissociation limit. *J. Chem. Phys.* **131**, 221105 (2009).
86. Boyarkin, O. V. Private communication (2012).
87. Furtenbacher, T., Tóbiás, R., Tennyson, J., Gamache, R. R. & Császár, A. G. Repository for the W2024 database of the water isotopologue $H_2^{16}O$. https://doi.org/10.17605/OSF.IO/4E237.
88. Golden, S., Wentink, T., Hillger, R. & Strandberg, M. W. P. Stark spectrum of H_2O. *Phys. Rev.* **73**, 92–93 (1948).
89. Jen, C. K. Rotational magnetic moments in polyatomic molecules. *Phys. Rev.* **81**, 197–203 (1951).
90. King, W. C. & Gordy, W. One-to-two millimeter wave spectroscopy. IV. Experimental methods and results for OCS, CH_3F, and H_2O. *Phys. Rev.* **93**, 407–412 (1954).
91. Fraley, P. E. & Rao, K. N. High resolution infrared spectra of water vapor. *J. Mol. Spectrosc.* **29**, 348–364 (1969).
92. Kukolich, S. G. Measurement of the molecular *g* values in H_2O and D_2O and hyperfine structure in H_2O. *J. Chem. Phys.* **50**, 3751–3755 (1969).
93. Evenson, K. M., Wells, J. S., Matarrese, L. M. & Elwell, L. B. Absolute frequency measurements of the 28- and 78-μm cw water vapor laser lines. *Appl. Phys. Lett.* **16**, 159–162 (1970).
94. Stephenson, D. A. & Strauch, R. G. Water vapor spectrum near 600 GHz. *J. Mol. Spectrosc.* **35**, 494–495 (1970).
95. Huiszoon, C. A high resolution spectrometer for the shorter millimeter wavelength region. *Rev. Sci. Instrum.* **42**, 477–481 (1971).
96. Steenbeckeliers, G. & Bellet, J. Spectre micro-onde de molecules $H_2^{16}O$, $H_2^{17}O$ et $H_2^{18}O$. *C. R. Acad. Sci. B Phys.* **273**, 471–474 (1971).
97. Flaud, J.-M., Camy-Peyret, C. & Valentin, A. Spectre infrarouge a haute résolution des bandes $\nu_1 + \nu_2$ et $\nu_2 + \nu_3$ de $H_2^{16}O$. *J. Phys. (Paris)* **33**, 741–747 (1972).
98. de Lucia, F. C., Helminger, P., Cook, R. L. & Gordy, W. Submillimeter microwave spectrum of $H_2^{16}O$. *Phys. Rev. A* **5**, 487–490 (1972).
99. Blaney, T. G., Bradley, C. C., Edwards, G. J. & Knight, D. J. E. Absolute frequency measurement of a Lamb-dip stabilised water vapour laser oscillating at 10.7 THz (28 μm). *Phys. Lett. A* **43**, 471–472 (1973).
100. Camy-Peyret, C., Flaud, J.-M., Guelachvili, G. & Amiot, C. High resolution Fourier transform spectrum of water between 2930 and 4255 cm^{-1}. *Mol. Phys.* **26**, 825–855 (1973).
101. Toth, R. A. & Margolis, J. S. Line positions of H_2O in the 1.33 to 1.45 micron region. *J. Mol. Spectrosc.* **55**, 229–251 (1975).
102. Fleming, J. W. & Gibson, M. J. Far-infrared absorption spectra of water vapor $H_2^{16}O$ and isotopic modifications. *J. Mol. Spectrosc.* **62**, 326–337 (1976).
103. Flaud, J.-M., Camy-Peyret, C., Maillard, J.-P. & Guelachvili, G. The H_2O spectrum between 4200 and 5000 cm^{-1}. *J. Mol. Spectrosc.* **65**, 219–228 (1977).
104. Kauppinen, J., Kärkkäinen, T. & Kyrö, E. High-resolution spectrum of water vapour between 30 and 720 cm^{-1}. *J. Mol. Spectrosc.* **71**, 15–45 (1978).
105. Flaud, J.-M. *et al*. Spectrum of water vapor between 8050 and 9370 cm^{-1}. *J. Mol. Spectrosc.* **75**, 339–362 (1979).
106. Herman, M., Johns, J. W. C. & McKellar, A. R. W. High resolution laser Stark and infrared-radiofrequency double resonance spectroscopy of $H_2^{16}O$ at 6 μm. *Can. J. Phys.* **57**, 397–401 (1979).
107. Wilkerson, T. D., Schwemmer, G., Gentry, B. & Giver, L. P. Intensities and N_2 collision-broadening coefficients measured for selected H_2O absorption lines between 715 and 732 nm. *J. Quant. Spectrosc. Rad. Transf.* **22**, 315–331 (1979).
108. Camy-Peyret, C., Flaud, J. M. & Maillard, J.-P. The $4\nu_2$ band of $H_2^{16}O$. *J. Phys. Lett (Paris)* **41**, 23–26 (1980).
109. Kauppinen, J. & Kyrö, E. High resolution pure rotational spectrum of water vapor enriched by $H_2^{17}O$ and $H_2^{18}O$. *J. Mol. Spectrosc.* **84**, 405–423 (1980).
110. Kuze, H. Microwave spectrum of water in the ν_2 excited vibrational state. *Astrophys. J.* **239**, 1131–1133 (1980).
111. Partridge, R. H. Far-infrared absorption spectra of $H_2^{16}O$, $H_2^{17}O$, and $H_2^{18}O$. *J. Mol. Spectrosc.* **87**, 429–437 (1981).
112. Kauppinen, J., Jomena, K. & Horneman, V.-M. New wavenumber calibration tables for H_2O, CO_2 and OCS lines between 400 cm^{-1} and 900 cm^{-1}. *Appl. Optics* **21**, 3332–3336 (1982).
113. Burenin, A. V., Fevralskikh, T. M., Karyakin, E. N., Polyansky, O. L. & Shapin, S. M. Effective Pade Hamiltonian operator and its application for treatment of $H_2^{16}O$ rotational spectrum in the ground state. *J. Mol. Spectrosc.* **100**, 182–192 (1983).
114. Guelachvili, G. Experimental Doppler-limited spectra of the ν_2-bands of $H_2^{16}O$, $H_2^{17}O$, $H_2^{18}O$, and HDO by Fourier-transform spectroscopy-secondary wavenumber standards between 1066 and 2296 cm^{-1}. *J. Opt. Soc. Am.* **73**, 137–150 (1983).
115. Helminger, P., Messer, J. K. & de Lucia, F. C. Continuously tunable coherent spectroscopy for the 0.1-1.0 THz region. *Appl. Phys. Lett.* **42**, 309–10 (1983).
116. Messer, J. K., de Lucia, F. C. & Helminger, P. The pure rotational spectrum of water vapor-a millimeter, submillimeter, and far infrared analysis. *Int. J. Infrared Milli.* **4**, 505–539 (1983).
117. Pine, A. S., Coulombe, M. J., Camy-Peyret, C. & Flaud, J.-M. Atlas of the high-temperature water vapor spectrum in the 3000 to 4000 cm^{-1} region. *J. Phys. Chem. Ref. Data* **12**, 413–465 (1983).
118. Toth, R. A. & Brault, J. W. Line positions and strengths in the (001), (110), and (030) bands of HDO. *Appl. Optics* **22**, 908–926 (1983).
119. Brown, L. R. & Toth, R. A. Comparison of the frequencies of NH_3, CO_2, H_2O, N_2O, CO, and CH_4 as infrared calibration standards. *J. Opt. Soc. Am. B* **2**, 842–856 (1985).
120. Camy-Peyret, C. *et al*. The high-resolution spectrum of water vapor between 16500 and 25250 cm^{-1}. *J. Mol. Spectrosc.* **113**, 208–228 (1985).
121. Johns, J. W. C. High-resolution far-infrared (20–350 cm^{-1}) spectra of several isotopic species of H_2O. *J. Opt. Soc. Am. B* **2**, 1340–1354 (1985).
122. Mandin, J.-Y., Chevillard, J.-P., Camy-Peyret, C. & Flaud, J.-M. Line intensities in the $\nu_1 + 2\nu_2$, $2\nu_2 + \nu_3$, $2\nu_1$, $\nu_1 + \nu_3$, $2\nu_3$, and $\nu_1 + \nu_2 + \nu_3 - \nu_2$ bands of $H_2^{16}O$, between 6300 and 7900 cm^{-1}. *J. Mol. Spectrosc.* **118**, 96–102 (1986).
123. Mandin, J.-Y., Chevillard, J.-P., Camy-Peyret, C. & Flaud, J.-M. The high-resolution spectrum of water-vapor between 13 200 and 16 500 cm^{-1}. *J. Mol. Spectrosc.* **116**, 167–190 (1986).
124. Baskakov, O. I., Alekseev, V. A., Alekseev, E. A. & Polevoi, B. I. New submillimeter rotational lines of water and its isotopes. *Opt. Spectrosc.* **63**, 1016–1018 (1987).
125. Belov, S. P., Kozin, I. N., Polyansky, O. L., Tretyakov, M. Y. & Zobov, N. F. Rotational spectrum of the $H_2^{16}O$ molecule in the (0 1 0) excited vibrational state. *J. Mol. Spectrosc.* **126**, 113–117 (1987).
126. Mandin, J. Y., Chevillard, J. P., Flaud, J.-M. & Camy-Peyret, C. $H_2^{16}O$: line positions and intensities between 8000 and 9500 cm^{-1}: the second hexad of interacting vibrational states: ((050), (130), (031), (210), (111), (012)). *Can. J. Phys.* **66**, 997–1011 (1988).
127. Bauer, A., Godon, M., Kheddar, M. & Hartmann, J. Temperature and perturber dependences of water vapor line-broadening. Experiments at 183 GHz; calculations below 1000 GHz. *J. Quant. Spectrosc. Rad. Transf.* **41**, 49–54 (1989).
128. Chevillard, J.-P., Mandin, J.-Y., Flaud, J.-M. & Camy-Peyret, C. $H_2^{16}O$ Line positions and intensities between 9500 and 11500 cm^{-1}. The interacting vibrational states (041), (220), (121), (022), (300), (201), (102), and (003). *Can. J. Phys.* **67**, 1065–1084 (1989).
129. Amano, T. & Scappini, F. Millimeter-wave spectrum of rotationally excited H_2O. *Chem. Phys. Lett.* **182**, 93–95 (1991).
130. Pearson, J. C., Anderson, T., Herbst, E., de Lucia, F. C. & Helminger, P. Millimeter- and submillimeter-wave spectrum of highly excited states of water. *Astrophys. J.* **379**, L41–L43 (1991).

131. Toth, R. A. ν_2 band of $H_2^{16}O$: line strengths and transition frequencies. *J. Opt. Soc. Am. B* **8**, 2236–2255 (1991).
132. Dana, V., Mandin, J. Y., Camy-Peyret, C., Flaud, J.-M. & Rothman, L. S. Rotational and vibrational dependences of collisional linewidths in the $n\nu_2 - (n-1)\nu_2$ hot bands of H_2O from Fourier-transform flame spectra. *Appl. Optics* **31**, 1179–1184 (1992).
133. Dana, V. *et al*. Measurements of collisional linewidths in the ν_2 band of H_2O from Fourier-transform flame spectra. *Appl. Optics* **31**, 1928–1936 (1992).
134. Mandin, J. Y., Dana, V., Camy-Peyret, C. & Flaud, J.-M. Collisional widths of pure rotational transitions of H_2O from Fourier-transform flame spectra. *J. Mol. Spectrosc.* **152**, 179–184 (1992).
135. Toth, R. A. $2\nu_2 - \nu_2$ and $2\nu_2$ bands of $H_2^{16}O$, $H_2^{17}O$, and $H_2^{18}O$: line positions and strengths. *J. Opt. Soc. Am. B* **10**, 1526–1544 (1993).
136. Toth, R. A. $\nu_1 - \nu_2$, $\nu_3 - \nu_2$, ν_1 and ν_3 bands of $H_2^{16}O$: line positions and strengths. *J. Opt. Soc. Am. B* **10**, 2006–2029 (1993).
137. Toth, R. A. Measurements of $H_2^{16}O$: line positions and strengths: 11610 to 12861 cm^{-1}. *J. Mol. Spectrosc.* **166**, 176–183 (1994).
138. Toth, R. A. Extensive measurements of $H_2^{16}O$: frequencies and strengths: 5750 to 7965 cm^{-1}. *Appl. Optics* **33**, 4851–4867 (1994).
139. Markov, V. N. & Krupnov, A. F. Measurements of the pressure shift of the (110)-(101) water line at 556 GHz produced by mixtures of gases. *J. Mol. Spectrosc.* **172**, 211–214 (1995).
140. Matsushima, F., Odashima, H., Iwasaki, T., Tsunekawa, S. & Takagi, K. Frequency measurement of pure rotational transitions of H_2O from 0.5 to 5 THz. *J. Mol. Spectrosc.* **352**, 371–378 (1995).
141. Paso, R. & Horneman, V.-M. High-resolution rotational absorption spectra of $H_2^{16}O$, $HD^{16}O$, and $D_2^{16}O$ between 110 and 500 cm^{-1}. *J. Opt. Soc. Am. B* **12**, 1813–1838 (1995).
142. Pearson, J. C. Ph.D. thesis, Duke University (1995).
143. Belov, S. P. Private communication, data from 00ChPePiMa (1996).
144. Brown, L. R. & Margolis, J. S. Empirical line parameters of NH_3 from 4791 to 5294 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **56**, 283–294 (1996).
145. Brown, L. R. & Plymate, C. H_2-broadened $H_2^{16}O$ in four infrared bands between 55 and 4045 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **56**, 263–282 (1996).
146. Polyansky, O. L., Busler, J. R., Guo, B. J., Zhang, K. Q. & Bernath, P. F. The emission spectrum of hot water in the region between 370 and 930 cm^{-1}. *J. Mol. Spectrosc.* **176**, 305–315 (1996).
147. Flaud, J.-M. *et al*. The high-resolution spectrum of water vapor between 11 600 and 12 750 cm^{-1}. *J. Mol. Spectrosc.* **183**, 300–309 (1997).
148. Mikhailenko, S. N. *et al*. The $2\nu_2$ band of water: analysis of new FTS measurements and high-K_a transitions and energy Levels. *J. Mol. Spectrosc.* **184**, 330–349 (1997).
149. de Natale, P. *et al*. Accurate frequency measurement for H_2O and $^{16}O_3$ in the 119 cm^{-1} OH atmospheric window. *Appl. Optics* **36**, 8526–8532 (1997).
150. Polyansky, O. L., Tennyson, J. & Bernath, P. F. The spectrum of hot water: rotational transitions and difference bands in the (020), (100), and (001). *J. Mol. Spectrosc.* **186**, 213–221 (1997).
151. Polyansky, O. L., Zobov, N. F., Tennyson, J., Lotoski, J. A. & Bernath, P. F. Hot bands of water in the ν_2 manifold up to $5\nu_2 - 4\nu_2$. *J. Mol. Spectrosc.* **184**, 35–50 (1997).
152. Polyansky, O. L. *et al*. High temperature rotational transitions of water in sunspot and laboratory spectra. *J. Mol. Spectrosc.* **186**, 422–447 (1997).
153. Bauer, A., Godon, M., Carlier, J. & Gamache, R. Continuum in the windows of the water vapor spectrum. Absorption of H_2O-Ar at 239 GHz and linewidth calculations. *J. Quant. Spectrosc. Rad. Transf.* **59**, 273–285 (1998).
154. Chance, K. V., Park, K. & Evenson, K. M. Pressure broadening of far infrared rotational transitions: 88.65 cm^{-1} H_2O and 114.47 cm^{-1} O_3. *J. Quant. Spectrosc. Rad. Transf.* **59**, 687–688 (1998).
155. Esplin, M. P., Wattson, R. B., Hoke, L. B. & Rothman, L. S. High-temperature spectrum of H_2O in the 720-1400 cm^{-1} region. *J. Quant. Spectrosc. Rad. Transf.* **60**, 711–739 (1998).
156. Polyansky, O. L., Zobov, N. F., Viti, S. & Tennyson, J. Water vapour line assignments in the near infrared. *J. Mol. Spectrosc.* **189**, 291–300 (1998).
157. Toth, R. A. Water vapor measurements between 590 and 2582 cm^{-1}: line positions and strengths. *J. Mol. Spectrosc.* **190**, 379–396 (1998).
158. Carleer, M. *et al*. The near infrared, optical and near ultraviolet overtone spectrum of water. *J. Chem. Phys.* **111**, 2444–2450 (1999).
159. Toth, R. A. Analysis of line positions and strengths of $H_2^{16}O$ ground and hot bands connecting to interacting upper states: (020), (100), and (001). *J. Mol. Spectrosc.* **194**, 28–42 (1999).
160. Valentin, A. *et al*. The water-vapor ν_2 band lineshift coefficients induced by nitrogen pressure. *J. Mol. Spectrosc.* **198**, 218–229 (1999).
161. Zobov, N. F. *et al*. Hot bands of water up to $6\nu_2 - 5\nu_2$ in the 933–2500 cm^{-1} region. *J. Mol. Spectrosc.* **193**, 118–136 (1999).
162. Brown, L. R. & Plymate, C. Experimental Line Parameters of the Oxygen *A* Band at 760 nm. *J. Mol. Spectrosc.* **199**, 166–179 (2000).
163. Bykov, A. D., Naumenko, O. & Vorobyova, L. Visible water vapor spectrum from 17358 to 21400 cm^{-1} reanalysis. In *Seventh International Symposium on Atmospheric and Ocean Optics*, vol. 4341, 31–37 (2000).
164. Chen, P., Pearson, J. C., Pickett, H. M., Matsuura, S. & Blake, G. A. Submillimeter-wave measurements and analysis of the ground and $v_2 = 1$ states of water. *Astrophys. J. Suppl. S.* **128**, 371–385 (2000).
165. Zobov, N. F. *et al*. The near ultraviolet rotation-vibration spectrum of water. *J. Chem. Phys.* **113**, 1546–1552 (2000).
166. Zobov, N. F. *et al*. Using laboratory spectroscopy to identify lines in the K- and L-band spectrum of water in a sunspot. *Astrophys. J.* **530**, 994–998 (2000).
167. Bykov, A. *et al*. High-order resonances in the water molecule. *J. Mol. Spectrosc.* **205**, 1–8 (2001).
168. Naus, H. *et al*. Cavity-ring-down spectroscopy on water vapor in the range 555–604 nm. *J. Mol. Spectrosc.* **205**, 117–121 (2001).
169. Brown, L. R., Toth, R. A. & Dulick, M. Empirical line parameters of $H_2^{16}O$ near 0.94 μm: positions, intensities, and air-broadening coefficients. *J. Mol. Spectrosc.* **212**, 57–82 (2002).
170. Mikhailenko, S. N. *et al*. Water spectra in the region 4200-6250 cm^{-1}: extended analysis of $\nu_1 + \nu_2$, $\nu_2 + \nu_3$, and $3\nu_2$ bands and confirmation of highly excited states from flame spectra and from atmospheric long-path observations. *J. Mol. Spectrosc.* **213**, 91–121 (2002).
171. Schermaul, R. *et al*. Weak line water vapor spectrum in the regions 13 200–15 000 cm^{-1}. *J. Mol. Spectrosc.* **211**, 169–178 (2002).
172. Tereszchuk, K. *et al*. Laboratory spectroscopy of hot water near 2 microns and sunspot spectroscopy in the H-band region. *Astrophys. J.* **577**, 496–500 (2002).
173. Tolchenov, R. N., Tennyson, J., Brault, J. W., Canas, A. A. D. & Schermaul, R. Weak line water vapor spectrum in the 11 787-13 554 cm^{-1} region. *J. Mol. Spectrosc.* **215**, 269–274 (2002).
174. Naumenko, O. & Campargue, A. Rovibrational analysis of the absorption spectrum of H_2O around 1.02 μm by ICLAS-VECSEL. *J. Mol. Spectrosc.* **221**, 221–226 (2003).
175. Tolchenov, R. N. *et al*. Water line parameters for weak lines in the range 9000–12 700 cm^{-1}. *J. Mol. Spectrosc.* **221**, 99–105 (2003).
176. Zou, Q. & Varanasi, P. Laboratory measurement of the spectroscopic line parameters of water vapor in the 610–2100 and 3000–4050 cm^{-1} regions at lower-tropospheric temperatures. *J. Quant. Spectrosc. Rad. Transf.* **82**, 45–98 (2003).

177. Coudert, L. H., Pirali, O., Vervloet, M., Lanquetin, R. & Camy-Peyret, C. The eight first vibrational states of the water molecule: measurements and analysis. *J. Mol. Spectrosc.* **228**, 471–498 (2004).
178. Macko, P. *et al*. High sensitivity CW-cavity ring down spectroscopy of water in the region of the 1.5 μm atmospheric window. *J. Mol. Spectrosc.* **227**, 90–108 (2004).
179. Shirin, S. V. *et al*. Analysis of hot D_2O emission using spectroscopically determined potentials. *J. Chem. Phys.* **120**, 206–210 (2004).
180. Coheur, P.-F. *et al*. A 3000 K laboratory emission spectrum of water. *J. Chem. Phys.* **122**, 074307 (2005).
181. Dupré, P., Gherman, T., Zobov, N. F., Tolchenov, R. N. & Tennyson, J. Continuous-wave cavity ringdown spectroscopy of the 8ν polyad of water in the 25195–25340 cm^{-1} range. *J. Chem. Phys.* **123**, 154307 (2005).
182. Golubiatnikov, G. Y. Shifting and broadening parameters of the water vapor 183-GHz line (313 − 220) by H_2O, O_2, N_2, CO_2, H_2, He, Ne, Ar, and Kr at room temperature. *J. Mol. Spectrosc.* **230**, 196–198 (2005).
183. Horneman, V.-M., Anttila, R., Alanko, S. & Pietilä, J. Transferring calibration from CO_2 laser lines to far infrared water lines with the aid of the ν_2 band of OCS and the ν_2, $\nu_1 - \nu_2$, and $\nu_1 + \nu_2$ bands of $^{13}CS_2$: Molecular constants of $^{13}CS_2$. *J. Mol. Spectrosc.* **234**, 238–254 (2005).
184. Kassi, S., Macko, P., Naumenko, O. & Campargue, A. The absorption spectrum of water near 750 nm by CW-CRDS: contribution to the search of water dimer absorption. *Phys. Chem. Chem. Phys.* **7**, 2460–2467 (2005).
185. Ptashnik, I. V., Smith, K. M. & Shine, K. P. Self-broadened line parameters for water vapour in the spectral region 5000–5600 cm^{-1}. *J. Mol. Spectrosc.* **232**, 186–201 (2005).
186. Tolchenov, R. N. *et al*. Water vapor line assignments in the 9250–26000 cm^{-1} frequency range. *J. Mol. Spectrosc.* **233**, 68–76 (2005).
187. Tolchenov, R. N. & Tennyson, J. Water line parameters for weak lines in the range 7400-9600 cm^{-1}. *J. Mol. Spectrosc.* **231**, 23–27 (2005).
188. Toth, R. A. 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. Rad. Transf.* **94**, 51–107 (2005).
189. Golubiatnikov, G. Y., Markov, V. N., Guarnieri, A. & Knochel, R. Hyperfine structure of $H_2{}^{16}O$ and $H_2{}^{18}O$ measured by Lamb-dip technique in the 180–560 GHz frequency range. *J. Mol. Spectrosc.* **240**, 251–254 (2006).
190. Joly, L., Parvitte, B., Zéninari, V., Courtois, D. & Durry, G. A spectroscopic study of water vapor isotopologues $H_2{}^{16}O$, $H_2{}^{18}O$ and HDO using a continuous wave DFB quantum cascade laser in the 6.7 μm region for atmospheric applications. *J. Quant. Spectrosc. Rad. Transf.* **102**, 129–138 (2006).
191. Mazzotti, F., Naumenko, O. V., Kassi, S., Bykov, A. D. & Campargue, A. ICLAS of weak transitions of water between 11 300 and 12 850 cm^{-1}: comparison with FTS databases. *J. Mol. Spectrosc.* **239**, 174–181 (2006).
192. Matsushima, F., Tomatsu, N., Nagai, T., Moriwaki, Y. & Takagi, K. Frequency measurement of pure rotational transitions in the $v_2 = 1$ state of H_2O. *J. Mol. Spectrosc.* **235**, 190–195 (2006).
193. Petrova, T., Poplavskii, Y., Serdukov, V. & Sinitsa, L. Intracavity laser spectroscopy of high-temperature water vapour in the range 9390-9450 cm^{-1}. *Mol. Phys.* **104**, 2692–2700 (2006).
194. Zobov, N. F. *et al*. Infrared emission spectrum of hot D_2O. *J. Mol. Spectrosc.* **240**, 112–119 (2006).
195. Zobov, N. F. *et al*. Spectrum of hot water in the 2000–4750 cm^{-1} frequency range. *J. Mol. Spectrosc.* **237**, 115–122 (2006).
196. Jenouvrier, A. *et al*. Fourier transform measurements of water vapor line parameters in the 4200-6600 cm^{-1} region. *J. Quant. Spectrosc. Rad. Transf.* **105**, 326–355 (2007).
197. Koshelev, M. A. *et al*. Broadening and shifting of the 321-, 325- and 380 GHz lines of water vapor by pressure of atmospheric gases. *J. Mol. Spectrosc.* **241**, 101–108 (2007).
198. Mazzotti, F., Tolchenov, R. N. & Campargue, A. High sensitivity ICLAS of $H_2{}^{18}O$ in the region of the second decade (11 520–12 810 cm^{-1}). *J. Mol. Spectrosc.* **243**, 78–89 (2007).
199. Mikhailenko, S. N., Le, W., Kassi, S. & Campargue, A. Weak water absorption lines around 1.455 and 1.66 μm by CW-CRDS. *J. Mol. Spectrosc.* **244**, 170–178 (2007).
200. Campargue, A., Mikhailenko, S. & Liu, A. W. ICLAS of water in the 770 nm transparency window (12 746–13 558 cm^{-1}). Comparison with current experimental and calculated databases. *J. Quant. Spectrosc. Rad. Transf.* **109**, 2832–2845 (2008).
201. Lisak, D. & Hodges, J. T. Low-uncertainty H_2O line intensities for the 930 nm region. *J. Mol. Spectrosc.* **249**, 6–13 (2008).
202. Mikhailenko, S. N. *et al*. Water vapor absorption line intensities in the 1900-6600 cm^{-1} region. *J. Quant. Spectrosc. Rad. Transf.* **109**, 2687–2696 (2008).
203. 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. Rad. Transf.* **109**, 559–568 (2008).
204. Zobov, N. F. *et al*. Spectrum of hot water in the 4750–13 000 cm^{-1} frequency range (0.769–2.1 μm). *Mon. Not. R. Astron. Soc.* **387**, 1093–1098 (2008).
205. Cazzoli, G., Puzzarini, C., Buffa, G. & Tarrini, O. Pressure-broadening of water lines in the THz frequency region: improvements and confirmations for spectroscopic databases. Part II. *J. Quant. Spectrosc. Rad. Transf.* **110**, 609–618 (2009).
206. Cazzoli, G., Puzzarini, C., Harding, M. E. & Gauss, J. The hyperfine structure in the rotational spectrum of water: Lamb-dip technique and quantum-chemical calculations. *Chem. Phys. Lett.* **473**, 21–25 (2009).
207. Liu, A., Naumenko, O., Kassi, S. & Campargue, A. High sensitivity CW-CRDS of ^{18}O enriched water near 1.6 μm. *J. Quant. Spectrosc. Rad. Transf.* **110**, 1781–1800 (2009).
208. Béguier, S., Mikhailenko, S. & Campargue, A. The absorption spectrum of water between 13 540 and 14 070 cm^{-1}: ICLAS detection of weak lines and a complete line list. *J. Mol. Spectrosc.* **265**, 106–109 (2011).
209. Drouin, B. J., Yu, S., Pearson, J. C. & Gupta, H. Terahertz spectroscopy for space applications: 2.5–2.7 THz spectra of HD, H_2O and NH_3. *J. Mol. Struct.* **1066**, 2–12 (2011).
210. Koshelev, M. A. Collisional broadening and shifting of the 2_{11}–2_{02} transition of $H_2{}^{16}O$, $H_2{}^{17}O$, and $H_2{}^{18}O$ by atmosphere gases. *J. Quant. Spectrosc. Rad. Transf.* **112**, 550–552 (2011).
211. Mikhailenko, S., Kassi, S., Wang, L. & Campargue, A. The absorption spectrum of water in the 1.25 μm transparency window (7408–7920 cm^{-1}). *J. Mol. Spectrosc.* **269**, 92–103 (2011).
212. Down, M. J., Tennyson, J., Orphal, J., Chelin, P. & Ruth, A. A. Analysis of an ^{18}O and D enhanced water spectrum and new assignments for $HD^{18}O$ and $D_2{}^{18}O$ in the near-infrared region (6000–7000 cm^{-1}) using newly calculated variational line lists. *J. Mol. Spectrosc.* **282**, 1–8 (2012).
213. 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. Rad. Transf.* **113**, 2155–2166 (2012).
214. Mikhailenko, S. N. *et al*. Absorption spectrum of deuterated water vapor enriched by ^{18}O between 6000 and 9200 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **113**, 653–669 (2012).
215. Vasilchenko, S. S., Mikhailenko, S. N., Serdyukov, V. I. & Sinitsa, L. N. The absorption spectrum of $H_2{}^{18}O$ in the range 13400–14460 cm^{-1}. *Opt. Spectrosc.* **113**, 499–504 (2012).
216. Yu, S. *et al*. Measurement and analysis of new terahertz and far-infrared spectra of high temperature water. *J. Mol. Spectrosc.* **279**, 16–25 (2012).
217. Leshchishina, O., Mikhailenko, S. N., 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. Rad. Transf.* **130**, 69–80 (2013).

218. Lu, Y., Li, X.-F., Wang, J., Liu, A.-W. & Hu, S.-M. H_2O line positions in the 784–795 nm region with 10^{-9} accuracy. *J. Quant. Spectrosc. Rad. Transf.* **118**, 96–101 (2013).
219. Mikhailenko, S. N., Serdyukov, V. I., Sinitsa, L. N. & Vasilchenko, S. S. LED-based Fourier-transform spectroscopy of $H_2^{18}O$ in the range 15 000–15 700 cm^{-1}. *Opt. Spectrosc.* **115**, 912–921 (2013).
220. Tretyakov, M. Y., Koshelev, M. A., Vilkov, I. N., Parshin, V. V. & Serov, E. A. Resonator spectroscopy of the atmosphere in the 350–500 GHz range. *J. Quant. Spectrosc. Rad. Transf.* **114**, 109–121 (2013).
221. Yu, S., Pearson, J. C. & Drouin, B. J. Terahertz spectroscopy of water in its second triad. *J. Mol. Spectrosc.* **288**, 7–10 (2013).
222. Coudert, L. H., Martin-Drumel, M.-A. & Pirali, O. Analysis of the high-resolution water spectrum up to the second triad and to $J = 30$. *J. Mol. Spectrosc.* **303**, 36–41 (2014).
223. Liu, A.-W., Naumenko, O. V., Kassi, S. & Campargue, A. CW-cavity ring down spectroscopy of deuterated water in the 1.58 μm atmospheric transparency window. *J. Quant. Spectrosc. Rad. Transf.* **138**, 97–106 (2014).
224. Osipov, K. Y., Kapitanov, V. A., Protasevich, A. E., Pereslavtseva, A. A. & Ponurovsky, Y. Y. Diode laser spectroscopy of $H_2^{16}O$ spectra broadened by N_2 and He in 1.39 mm region. *J. Quant. Spectrosc. Rad. Transf.* **142**, 1–8 (2014).
225. Regalia, L. *et al.* Water vapor line parameters from 6450 to 9400 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **136**, 119–136 (2014).
226. Sinitsa, L. N. *et al.* Fourier-transform absorption spectrum of H_2O in the region of 15500–16000 cm^{-1}. In Romanovskii, O. A. (ed.) *20th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*, 92920J (Society of Photographic Instrumentation Engineers, Bellingham, 2014).
227. Campargue, A. *et al.* The absorption spectrum of water vapor in the 1.25 μm atmospheric window (7911-8337 cm^{-1}). *J. Quant. Spectrosc. Rad. Transf.* **157**, 135–152 (2015).
228. Hu, S.-M. *et al.* Water line positions in the 782–840 nm region. *J. Quant. Spectrosc. Rad. Transf.* **164**, 37–44 (2015).
229. Mikhailenko, S. N., Serdyukov, V. I. & Sinitsa, L. N. LED-based Fourier transform spectroscopy of $H_2^{18}O$ in the 15 000–16 000 cm^{-1} range. *J. Quant. Spectrosc. Rad. Transf.* **156**, 36–46 (2015).
230. Sironneau, V. T. & Hodges, J. T. Line shapes, positions and intensities of water transitions near 1.28 μm. *J. Quant. Spectrosc. Rad. Transf.* **152**, 1–15 (2015).
231. Mikhailenko, S. N. *et al.* 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. Rad. Transf.* **177**, 108–116 (2016).
232. Birk, M. *et al.* Accurate line intensities for water transitions in the infrared: comparison of theory and experiment. *J. Quant. Spectrosc. Rad. Transf.* **203**, 88–102 (2017).
233. Campargue, A. *et al.* The absorption spectrum of water vapor in the 2.2 μ transparency window: high sensitivity measurements and spectroscopic database. *J. Quant. Spectrosc. Rad. Transf.* **189**, 407–416 (2017).
234. Loos, J., Birk, M. & Wagner, G. Measurement of air-broadening line shape parameters and temperature dependence parameters of H_2O lines in the spectral ranges 1850–2280 cm^{-1} and 2390–4000 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **203**, 103–118 (2017).
235. Mondelain, D. *et al.* Comb-assisted cavity ring down spectroscopy of ^{17}O enriched water between 7443 and 7921 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **203**, 206–212 (2017).
236. Chen, J. *et al.* Absolute frequencies of water lines near 790 nm with 10^{-11} accuracy. *J. Quant. Spectrosc. Rad. Transf.* **205**, 91–95 (2018).
237. Czinki, E., Furtenbacher, T., Császár, A. G., Eckhardt, A. K. & Mellau, G. C. The 1943 K emission spectrum of $H_2^{16}O$ between 6600 and 7050 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **206**, 46–54 (2018).
238. Devi, V. M. *et al.* Line shape parameters of air-broadened water vapor transitions in the ν_1 and ν_3 spectral region. *J. Mol. Spectrosc.* **348**, 13–36 (2018).
239. Kassi, S., Stoltmann, T., Casado, M., Daeron, M. & Campargue, A. Lamb dip CRDS of highly saturated transitions of water near 1.4 μm. *J. Chem. Phys.* **148**, 054201 (2018).
240. Mikhailenko, S. N., Mondelain, D., Karlovets, E. V., Kassi, S. & Campargue, A. Comb-assisted cavity ring down spectroscopy of ^{17}O enriched water between 6667 and 7443 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **206**, 163–171 (2018).
241. Mikhailenko, S. N., Serdyukov, V. I. & Sinitsa, L. N. Study of $H_2^{16}O$ and $H_2^{18}O$ absorption in the 16 460–17 200 cm^{-1} range using LED-based Fourier transform spectroscopy. *J. Quant. Spectrosc. Rad. Transf.* **217**, 170–177 (2018).
242. Rutkowski, L. *et al.* An experimental water line list at 1950 K in the 6250–6670 cm^{-1} region. *J. Quant. Spectrosc. Rad. Transf.* **205**, 213–219 (2018).
243. Schroeder, P. J. *et al.* Speed-dependent Voigt lineshape parameter database from dual frequency comb measurements up to 1305 K. Part I: pure H_2O absorption, 6801–7188 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **210**, 240–250 (2018).
244. Serdyukov, V. *et al.* Study of H_2O line broadening and shifting by N_2 pressure in the 16 600–17 060 cm^{-1} region using FT-spectrometer with LED source. *J. Quant. Spectrosc. Rad. Transf.* **219**, 213–223 (2018).
245. Sinitsa, L. N., Serdyukov, V. I., Polovtseva, E. R., Bykov, A. D. & Shcherbakov, A. P. Study of the water vapor absorption spectrum in the visible spectral region from 19 480 to 20 500 cm^{-1}. *Atmospheric Ocean. Opt.* **31**, 329–334 (2018).
246. Tan, Y. *et al.* CRDS absorption spectrum of ^{17}O enriched water vapor in the 12 277–12 894 cm^{-1} range. *J. Quant. Spectrosc. Rad. Transf.* **221**, 233–242 (2018).
247. Liu, A.-W. *et al.* Cavity ring-down spectroscopy of ^{17}O-enriched water vapor between 12 055 and 12 260 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **239**, 106651 (2019).
248. Mikhailenko, S. N. *et al.* New transitions and energy levels of water vapor by high sensitivity CRDS near 1.73 and 1.54 μm. *J. Quant. Spectrosc. Rad. Transf.* **236**, 106574 (2019).
249. Mikhailenko, S. N., Mondelain, D., Karlovets, E. V., Kassi, S. & Campargue, A. Cavity ring down spectroscopy of ^{17}O enriched water vapor near 1.73 μm. *J. Quant. Spectrosc. Rad. Transf.* **222**, 229–235 (2019).
250. Régalia, L., Thomas, X., Rennesson, T. & Mikhailenko, S. Line parameters of water vapor enriched by ^{18}O from 6525 to 8011 cm^{-1}. *J. Quant. Spectrosc. Rad. Transf.* **235**, 257–271 (2019).
251. Serdyukov, V. I., Sinitsa, L. N., Lavrentieva, N. N. & Dudaryonok, A. S. Measurements of N_2-broadening and -shifting parameters of the water vapor spectral lines in the 19 560–19 920 cm^{-1} region using FT-spectrometer with LED source. *J. Quant. Spectrosc. Rad. Transf.* **234**, 47–54 (2019).
252. Vasilchenko, S., Tran, H., Mondelain, D., Kassi, S. & Campargue, A. Accurate absorption spectroscopy of water vapor near 1.64 μm in support of the MEthane Remote LIdar missioN (MERLIN). *J. Quant. Spectrosc. Rad. Transf.* **235**, 332–342 (2019).
253. Campargue, A. *et al.* Observation of electric-quadrupole infrared transitions in water vapor. *Phys. Rev. Res.* **2**, 023091 (2020).
254. Campargue, A., Solodov, A. M., Solodov, A. A., Yachmenev, A. & Yurchenko, S. N. Detection of electric-quadrupole transitions in water vapour near 5.4 and 2.5 μm. *Phys. Chem. Chem. Phys.* **22**, 12476–12481 (2020).
255. Mikhailenko, S. N. *et al.* The far-infrared spectrum of ^{18}O enriched water vapour (40–700 cm^{-1}). *J. Quant. Spectrosc. Rad. Transf.* **253**, 107105 (2020).
256. Mikhailenko, S. N., Karlovets, E., Koroleva, A. & Campargue, A. The far infrared absorption spectrum of $D_2^{16}O$, $D_2^{17}O$, and $D_2^{18}O$: Experimental line positions, empirical energy levels and recommended line lists. *J. Phys. Chem. Ref. Data* **53** (2024).
257. Voronin, B. A., Tennyson, J., Lodi, L. & Kozodoev, A. V. The VoTe Room Temperature $H_2^{16}O$ Line List up to 25 000 cm^{-1}. *Opt. Spectrosc.* **127**, 967–973 (2019).
258. Mátyus, E. *et al.* Assigning quantum labels to variationally computed rotational-vibrational eigenstates of polyatomic molecules. *J. Chem. Phys.* **133**, 034113 (2010).

259. Szidarovszky, T., Fábri, C. & Császár, A. G. The role of axis embedding on rigid rotor decomposition analysis of variational rovibrational wave functions. *J. Chem. Phys.* **136**, 174112 (2012).
260. Child, M. S., Weston, T. & Tennyson, J. Quantum monodromy in the spectrum of H_2O and other systems: new insight into the level structures of quasi-linear molecules. *Mol. Phys.* **96**, 371–379 (1999).
261. Ma, Q., Tipping, R. H. & Lavrentieva, N. N. Pair identity and smooth variation rules applicable for the spectroscopic parameters of H_2O transitions involving high-*J* states. *Mol. Phys.* **109**, 1925–1941 (2011).
262. Conway, E. K. *et al*. Accurate line lists for $H_2{}^{16}O$ and $H_2{}^{18}O$ with extensive comparisons to theoretical and experimental sources including the HITRAN 2016 database. *J. Quant. Spectrosc. Rad. Transf.* **241**, 106711 (2020).
263. Conway, E. K. *et al*. A semi-empirical potential energy surface and line list for $H_2{}^{16}O$ extending into the near-ultraviolet. *Atmos. Chem. Phys.* **20**, 10015–10027 (2020).

Acknowledgements

The work performed in Budapest received support from NKFIH (grant no. K138233 to AGC and PD145972 to RT). The research done in the United Kingdom received support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme through Advance Grant number 883830. The joint work between the Budapest and London groups received support from the COST Action COSY (Confined Molecular Systems, CA21101). AGC is grateful to the Fulbright Foundation for funding his stay at the California Institute of Technology, Jet Propulsion Laboratory and to his host, Dr. Brian J. Drouin, for useful discussions; BJD is funded under contract with the National Aeronautics and Space Administration (NASA).

Author contributions

T.F., A.G.C., and R.T. designed the study, with important contributions from J.T. and R.R.G. T.F., R.T., and R.R.G. checked different aspects of the datasets (transitions and energy levels). R.T. draw the figures and prepared the tables. T.F., R.T., and A.G.C. wrote the original version of the manuscript. All of the authors have contributed to writing the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-024-03847-3.

Correspondence and requests for materials should be addressed to R.T. or A.G.C.