

# New Studies of the Visible and Near-Infrared Absorption by Water Vapour and Some Problems with the HITRAN Database

Djedjiga Belmiloud<sup>1</sup>, Roland Schermaul<sup>2</sup>, Kevin M. Smith<sup>3</sup>, Nikolai F. Zobov<sup>1</sup>, James W. Brault<sup>4</sup>, Richard C. M. Learner<sup>2</sup>, David A. Newnham<sup>3</sup>, and Jonathan Tennyson<sup>1</sup>

**Abstract.** New laboratory measurements and theoretical calculations of integrated line intensities for water vapour bands in the near-infrared and visible (8500-15800 cm<sup>-1</sup>) are summarised. Band intensities derived from the new measured data show a systematic 6 to 26 % increase compared to calculations using the HITRAN-96 database. The recent corrections to the HITRAN database [Giver *et al.*, *J. Quant. Spectrosc. Radiat. Transfer*, 66, 101-105, 2000] do not remove these discrepancies and the differences change to 6 to 38 %. The new data is expected to substantially increase the calculated absorption of solar energy due to water vapour in climate models based on the HITRAN database.

## 1. Introduction

Understanding the amount and wavelength dependence of absorption of solar radiation by water vapour is essential for climate studies. Model calculations of the Earth's climate rely on databases, of which HITRAN-96 [Rothman *et al.*, 1998] is the most used, to provide information about the spectral properties of molecules in the atmosphere. Recently a problem has been identified in model calculations of atmospheric absorption in both clear and cloudy skies [Arking, 1999]. Many climate models substantially underestimate the globally averaged short-wavelength absorption compared to atmospheric observations, by as much as 30 % of the total atmospheric absorption in the case of clear skies. Much research has been directed at identifying atmospheric absorbers additional to those already in the molecular databases to make up the discrepancy.

The authors are all at present working on the spectroscopic properties of water vapour in the visible and near-infrared (8500-15800 cm<sup>-1</sup>), and each of the groups concerned has become aware of important differences between their results and the data in the most recent version of HITRAN, which includes the revisions set out by Giver *et al.*

[2000] (referred to as HITRAN-COR in the following). We summarise in this paper a collective view of the principal differences between the new work and the database covering the four water polyads in this spectral region, which are designated as  $2\nu + \delta$ ,  $3\nu$ ,  $3\nu + \delta$  and  $4\nu$  (located around 8800 cm<sup>-1</sup>, 10600 cm<sup>-1</sup>, 12200 cm<sup>-1</sup> and 13800 cm<sup>-1</sup>, respectively), where  $\nu$  denotes a quantum of the stretching modes and  $\delta$  of the bending mode. Four different approaches to the problem have been employed: two use different methods to analyse extensive new laboratory data, one relies on sophisticated *ab initio* theory and the fourth exploits the fact that the laboratory data used in the 1980s to generate the current database are in the Kitt Peak FTS archive of the National Solar Observatory (NSO, Tucson, Arizona, U.S.A.) and so are available for examination and review. This paper concentrates on the overall picture; more detailed accounts with line-by-line data will be published separately.

## 2. Laboratory measurements

All the observations, both the 1980s NSO data and measurements reported here, use high resolution Fourier transform spectrometry and multiple-reflection White cells to study the long path absorption spectra of water vapour. Four problems dominate the accuracy of the final data and all affect the systematic errors in the analysis. These are dynamic range, the measurement of the partial pressure of water vapour in the absorption cell, the determination of the background intensity and the modelling of the line profile.

The problem of dynamic range derives from the intensity distribution of the lines in the water spectrum. The relation between transmittance and absorption coefficient is not linear and bad data results both for lines that are too weak relative to the noise level and for lines strong enough to be opaque at the line centre. One observation at a given water vapour pressure - path length product, using a baseline *rms*-signal-to-noise ratio approaching 1000:1, will generate precise data (i.e., data with random errors of less than 1 %) over a range of only about a factor of four in absorption (lines of approximately 18-72 % peak absorption) and a factor of 6.5 in absorbance. In order to measure precisely line intensities outside this range, i.e. weaker and stronger lines, measurements at different experimental conditions have to be made. The modelling of the absorption in the atmosphere in the range 8500-15800 cm<sup>-1</sup> includes line intensities that cover a range of five orders of magnitude. We have recorded an extensive series of laboratory spectra covering a dynamic range of a factor 650 from the weakest (usable) line at the longest path to the strongest (usable) line at the shortest path. This covers all the strong and medium strong water

<sup>1</sup>Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, U.K.

<sup>2</sup>Laser Optics & Spectroscopy, Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BW, U.K.

<sup>3</sup>Atmospheric Science Division, Space Science and Technology Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, U.K.

<sup>4</sup>Laboratory for Atmospheric and Space Physics, University of Colorado, 325 Broadway, Boulder, CO 80303-0590, U.S.A.

lines, while the shortfall of another factor of 150 in dynamic range can be provided by theoretical calculations.

The spectra of pure water vapour and water vapour in air were recorded at path lengths of 5, 32, 128 and 512 metres, at two temperatures (296 and 252 K), and high spectral resolution ( $0.03 \text{ cm}^{-1}$ ) at the Rutherford Appleton Laboratory (RAL) [Ballard *et al.*, 1994; Newnham and Ballard, 1998]. Signal-to-noise ratios in excess of 725:1 were achieved at the peak of the background signal around  $10500 \text{ cm}^{-1}$ .  $\text{H}_2\text{O}$  partial pressures were measured using humidity sensors with a stated accuracy of 1%. Several hours were allowed for adequate mixing of the water and air, and the attainment of equilibrium confirmed by repeated absorbance measurement. "Zero" grade synthetic air (Air Products Ltd, 21% oxygen, 79% nitrogen) was used to make mixtures of 10 hPa water vapour with air at a total pressure of 1000 hPa.

The third problem that arises is to define precisely the background continuum. The accuracy of the determination of the baseline is critical, especially in regions of weak absorption, and the method of using the ratio of sample- and empty-cell measurements does not achieve the necessary precision, though it is useful in helping to characterise the light source and filters which determine the passband. Both the observation and pump-out times are long and the systematic effects of lamp ageing are much greater than the random noise. These effects leave discrepancies of up to 2% between the measured and real transmittance baseline. A much improved baseline level must therefore be derived from points between the spectrum lines where there is no obvious absorption. In strong bands, where the line wings overlap, these regions may be some hundreds of  $\text{cm}^{-1}$  apart. We have quantified the problem by having two laboratories generate independent estimates of the background intensities using the same experimental observations. Where there are good window regions we find we can set a continuum level to about one-fifth of the *rms*-noise.

The problem that arises during the analysis of many thousands of lines, is to strike a balance between fast but less accurate algorithms and more sophisticated models, which are much slower, but provide stable and reliable descriptions of the spectral lines (including blended lines). We have used two approaches. The first assumes no knowledge of the line profiles and merely integrates the spectral absorbance over a range of wavenumbers. This method is, necessarily, sensitive to the *rms*-noise level and the accurate definition of the baseline. The second used least-squares spectral line fits, which generate stable analyses of the lines, including blends. This method assumes a Voigt line profile, i.e. a convolution of Lorentzian (pressure-broadened) and Gaussian (Doppler-broadened) contributions. Both line shape components were left free during the fitting process in order to take account of collisional narrowing effects [Ponsardin and Browell, 1997]. The line intensity is evaluated as the integral over the fitted Voigt profile. Comparison of the two data reduction methods for  $100 \text{ cm}^{-1}$  intervals throughout the  $3\nu$  polyad agree to within  $\pm 6\%$  of their means, which is within the range expected due to random errors.

### 3. Theoretical calculation

*Ab initio* theory based on a variational treatment of the nuclear motion problem is now able to treat water spectra with considerable precision [Partridge and Schwenke, 1997;

**Table 1.** Comparison of Four Different Sets of Intensity Data with the HITRAN-96 Database for the Four Water Polyads in the Spectral Region  $8500\text{--}15800 \text{ cm}^{-1}$ . Each Entry gives the Ratio of the Total Intensity Derived for a Polyad to HITRAN-96 for  $T = 296 \text{ K}$  Using the Following Methods: 1: Integration of Spectral Absorbances, 2: Spectral Linefits to New Spectra, 3: Empirically Adjusted *Ab Initio* Theory, 4: Data from HITRAN-COR [Giver *et al.*, 2000]. The Final Column gives our Recommended (Interim) Correction Factors for the Band Intensities of Giver *et al.* – HITRAN-COR.

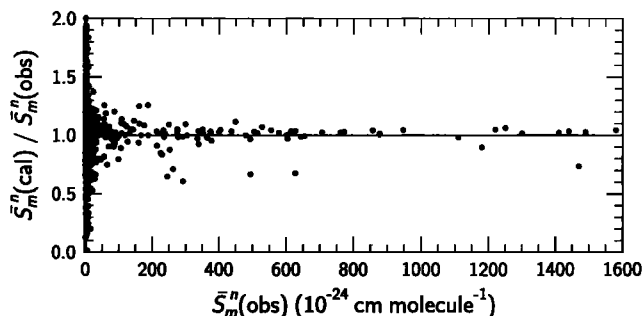
Polyad	1	2	3	4	Final
$2\nu + \delta$	–	1.26	1.31	0.92	1.38
$3\nu$	1.19	1.21	1.04	1.14	1.06
$3\nu + \delta$	1.26	1.25	1.25	1.09	1.15
$4\nu$	–	1.06	1.04	0.96	1.10

Note: Only lines that are common to both HITRAN and the new measurements were included in the comparisons (except in method 1). The ratios for the individual lines within a polyad vary, and only the combined effect is given. These ratios are temperature dependent. The recommended correction factors treat the total strength of the polyads and are not the mean of the line ratios as plotted in Figure 2.

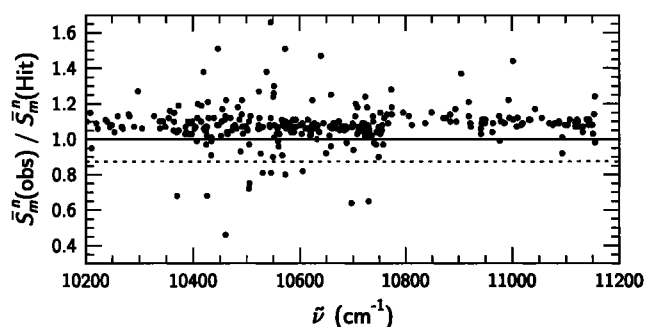
Polyansky *et al.*, 1997]. Although theoretical calculations are intrinsically less accurate, they have the advantage that they can produce complete, unbiased results for a large range of frequencies and intensities.

To estimate the errors in a variational treatment of the problem, three aspects must be considered: the potential energy surface used to obtain frequencies and wave functions, the dipole surface used to obtain line strengths and the computational procedure used for the variational calculations. In practice only the first two are a source of significant error in the present study.

Completely *ab initio* theory cannot yet give a Born-Oppenheimer potential energy surface for water of spectroscopic accuracy [Partridge and Schwenke, 1997], so in the present study we have compared results obtained using the PS [Partridge and Schwenke, 1997] linelist which uses an *ab initio* potential which has been adjusted using spectroscopic data (for energy levels, not intensities) to give frequencies close to the observed ones. Similar results are obtained with fully *ab initio* procedures.



**Figure 1.** Line-by-line ratio of experimental integrated line intensities ( $\bar{S}_m^n(\text{obs})$ ), derived from line fitting of the new spectra, and theoretical values ( $\bar{S}_m^n(\text{cal})$ ), as a function of intensity for lines within the  $2\nu + \delta$  polyad.



**Figure 2.** The ratio of integrated line intensities from line fitting of the new spectra ( $\bar{S}_m^n(\text{obs})$ ) to the corrected HITRAN data of Giver *et al.* [2000] ( $\bar{S}_m^n(\text{Hit})$ ) for the water  $3\nu$  polyad. Only the ratios for lines with intensities of more than  $10^{-23}$  cm molecule $^{-1}$  are shown. The level of the unit ratio is indicated by the continuous line and the dotted line shows the corresponding ratio for the uncorrected HITRAN-96 data.

Previous work [Lynas-Gray *et al.*, 1995] has shown that dipole surfaces determined using measured line intensities give worse results than purely *ab initio* dipole surfaces. For this reason both linelists used here employed the most accurate available *ab initio* dipole surface, which is that due to Partridge and Schwenke [1997]. Very recent work [Schwenke and Partridge, 2000] has confirmed the accuracy of the *ab initio* dipole calculations but highlighted problems with fitting to *ab initio* data. These problems particularly affect the excited bending states in the surface used here.

#### 4. Method intercomparison and comparison to HITRAN

Table 1 compares the results of the different methods employed with the HITRAN-96 database. The total intensity for each water polyad in the spectral region 8500–15800  $\text{cm}^{-1}$  was determined using the following methods: 1: integration of spectral absorbances using a fast algorithm, 2: spectral linefits to the same measurements, 3: variational calculations and 4: using HITRAN-COR [Giver *et al.*, 2000]. Figure 1 plots the ratio of the experimental (method 2) and theoretical line intensities for the component bands of the  $2\nu + \delta$  polyad. This comparison is chosen as it provides a good test of the dipole surface. The scatter of the data is typical of such comparisons: showing very good agreement for strong lines and increasing uncertainty for the weak ones.

Figure 2 plots the ratio of our measured room temperature line intensities (method 2) for the  $3\nu$  polyad against those in HITRAN-COR [Giver *et al.*, 2000]. The dominant effect is a systematic difference. Although the correction to Hitran-96 increased the intensities systematically by about 14% for this polyad, a further increase of about 10% is required to bring HITRAN-COR into reasonable agreement with the measurements. The corrections applied to the other polyads are less systematic and are not discussed in detail here. In summary, the new measurements suggest an overall increase in HITRAN-96 values (Table 1, columns 1–3) compared with the measurements relative to HITRAN-COR (column 4), which suggests both an increase ( $3\nu$  and  $3\nu + \delta$  polyads) and a decrease ( $2\nu + \delta$  and  $4\nu$  polyads) in intensities.

Attempts to explain the difference between the database and the new data necessarily lead back to the experimental observations upon which the database is built. These are an extensive series of measurements made using the NSO FTS at Kitt Peak in the 1980s [Mandin *et al.*, 1986, 1988; Chevillard *et al.*, 1989]. Returning to the problems set out at the beginning of Section 2, the 1980s data reductions used simple algorithms and were most prone to error when treating saturated lines and strong blends. The current work, with the computing speed needed to fit line profiles to over ten thousand lines and extensive checks of run-to-run consistency, should be considerably more robust. Most significant, judging by the uniformity of line intensity ratios for lines of moderate strength shown in Figure 2, are systematic errors arising either from the approximations of the 1980s line fitting procedure or from the determination of the concentration of water vapour in the absorption cell. We believe that our new data, taken from a wider range of experimental conditions and buttressed by line-by-line profile fitting and by sophisticated theoretical calculation, are an improvement on the current HITRAN listing. The crucial test that supports this view is that both the University College London (UCL) and Imperial College (IC) groups have taken the line-by-line data generated at IC and used it to construct synthetic spectra. The reconstruction agrees with the original spectrum to better than  $\pm 2\%$ . Similar attempts to reconstruct the Kitt Peak spectra, using the line-by-line data of Mandin *et al.* [1986, 1988] and Chevillard *et al.* [1989], do not reproduce their observed spectra.

#### 5. Conclusions

Table 1 (Final column) also sets out values for the comparison of our “best” total intensities of the water polyads with those given in HITRAN-COR. It should be stressed that the line intensities of our observations and the HITRAN database differ from line to line and that the given values are only valid for room temperature. The measurements at 252 K yielded slightly different ratios, but are omitted here for clarity. For a major fraction of the lines, the principal part of the change takes the form of a re-scaling of the data on a polyad-by-polyad basis, we recommend the use of the factors set out in Table 1 as an interim solution. Other databases, such as GEISA [Jacquinet-Husson *et al.*, 1999], are based on the same laboratory data and will therefore require the same correction. Detailed line-by-line data including both experiment and theory will be published [Schermaul *et al.*, 2000].

There is another lesson to be learned. Making sure the database is valid is necessary foundation for all modelling of atmospheric radiation transfer, especially so when theory and observation fail to agree.

**Acknowledgments.** We thank the UK Natural Environment Research Council (Grants GR3/11097, GR3/11674) and the European Space Agency (ESTEC Contract No. 13312/99/NL/SF) for support. Technical assistance by R. G. Williams and helpful discussions with J. Ballard are acknowledged. We have benefited greatly from the work of G. S. Jones-Gilby, who carried out a large part of the basic line fitting computation. We wish to thank O. L. Polyansky for help and advice, D. W. Schwenke for supplying results in advance of publication and M. Wickett of Serco Europe Ltd. for undertaking the project management.

## References

- Arking, A., Bringing climate models into agreement with observations of atmospheric absorption, *J. Climate*, *12*, 1589-1600, 1999.
- Ballard, J., K. Strong, J. J. Remedios, M. Page, and W. B. Johnston, A coolable long path absorption cell for laboratory spectroscopic studies of gases, *J. Quant. Spectrosc. Radiat. Transfer*, *52*, 677-691, 1994.
- Chevillard, J.-P., J.-Y. Mandin, J.-M. Flaud, and C. Camy-Peyret,  $H_2^{16}O$ : 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.
- Giver, L. P., C. Chackerian Jr., and P. Varanasi, Visible and near-infrared  $H_2^{16}O$  line intensity corrections for Hitran-96, *J. Quant. Spectrosc. Radiat. Transfer*, *66*, 101-105, 2000.
- Jacquinet-Husson, N., et al., The 1997 Spectroscopic GEISA databank, *J. Quant. Spectrosc. Radiat. Transfer*, *62*, 205-254, 1999.
- Lynas-Gray, A. E., S. Miller, and J. Tennyson, Infra red transition intensities for water: a comparison of *ab initio* and fitted dipole moment surfaces, *J. Mol. Spectrosc.*, *169*, 458-467, 1995.
- Mandin, J.-Y., J.-P. Chevillard, C. Camy-Peyret, and J.-M. Flaud, The high-resolution spectrum of water-vapor between 13200 and 16500  $cm^{-1}$ , *J. Mol. Spectrosc.*, *116*, 167-190, 1986.
- Mandin, J.-Y., J.-P. Chevillard, J.-M. Flaud, and C. Camy-Peyret,  $H_2O$ -16 - line positions and intensities between 8000 and 9500  $cm^{-1}$ . The 2nd hexad of interacting vibrational-states [(050), (130), (031), (210), (111), (012)], *Can. J. Phys.*, *66*, 997-1011, 1988.
- Newnham, D. A., and J. Ballard, Visible absorption cross-sections and integrated absorption intensities of molecular oxygen ( $O_2$  and  $O_4$ ), *J. Geophys. Res.*, *103*, 28,801-28,816, 1998.
- Partridge H., and D. W. Schwenke, The determination of an accurate isotope dependent potential energy surface for water from extensive *ab initio* calculations and experimental data, *J. Chem. Phys.*, *106*, 4618-4639, 1997.
- Polyansky, O. L., N. F. Zobov, S. Viti, J. Tennyson, P. F. Bernath, and L. Wallace, Water in the Sun: line assignments based on variational calculations, *Science*, *277*, 346-349, 1997.
- Ponsardin, P. L., and E. V. Browell, Measurement of  $H_2^{16}O$  line strengths and air-induced broadenings and shifts in the 815-nm spectral region, *J. Mol. Spectrosc.*, *185*, 58-70, 1997.
- Rothman L. S., C. P. Rinsland, A. Goldman, S. T. Massie, D. P. Edwards, J.-M. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.-Y. Mandin, J. Schroeder, A. McCann, R. R. Gamache, R. B. Wattson, K. Yoshino, K. V. Chance, K. W. Jucks, L. R. Brown, V. Nemtchinov, and P. Varanasi, The HITRAN molecular spectroscopic database and HAWKS (HITRAN Atmospheric Workstation): 1996 edition, *J. Quant. Spectrosc. Radiat. Transfer*, *60*, 665-710, 1998.
- Schermaul, R., R. C. M. Learner, D. A. Newnham, J. Ballard, N. F. Zobov, D. Belmiloud, and J. Tennyson, The water vapour spectrum in the region 8600-15000  $cm^{-1}$ : experimental and theoretical studies for a new spectral line database I: Laboratory measurements, *J. Mol. Spectrosc.*, in preparation, 2000.
- Schwenke D. W., and Partridge H., On using *ab initio* dipole moment functions to predict intensities for polyatomic molecules, *J. Chem. Phys.*, in press, 2000.
- R. Schermaul and R. C. M. Learner, Laser Optics & Spectroscopy, Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BW, U.K. (e-mail: r.schermaul@ic.ac.uk; r.learner@ic.ac.uk)
- K. M. Smith and D. A. Newnham, Atmospheric Science Division, Space Science and Technology Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, U.K. (e-mail: K.M.Smith@rl.ac.uk; david.newnham@rl.ac.uk)
- D. Belmiloud, N. F. Zobov, and J. Tennyson, Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, U.K. (e-mail: djedjiga@theory.phys.ucl.ac.uk; kolya@theory.phys.ucl.ac.uk; j.tennyson@ucl.ac.uk)
- J. W. Brault, Laboratory for Atmospheric and Space Physics, University of Colorado, 325 Broadway Boulder, CO 80303-0590, U.S.A. (e-mail: brault@noao.edu)

(Received June 16, 2000; revised July 26, 2000; accepted July 31, 2000.)