ELSEVIER

Journal of Quantitative Spectroscopy & Radiative Transfer 107 (2007) 349–359

Journal of Quantitative Spectroscopy & Radiative Transfer

www.elsevier.com/locate/jqsrt

Line parameter validation using ground-based solar occultation measurements: Water vapor—A case study ☆

B. Veihelmann[a,b,*], A.N. Maurellis[b], K.M. Smith[c], R.N. Tolchenov[d], J. Tennyson[d], W.J. van der Zande[a]

[a]*Institute for Molecules and Materials, Radboud University Nijmegen, The Netherlands*
[b]*Earth Oriented Science Department, National Institute for Space Research (SRON), Utrecht, The Netherlands*
[c]*Space Science and Technology Department, Rutherford Appleton Laboratory, Didcot, Oxfordshire, UK*
[d]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*

Received 16 May 2006; received in revised form 1 February 2007; accepted 2 February 2007

Abstract

Water vapor spectroscopy data for the 720 nm absorption band (4*v* polyad) are validated in the context of atmospheric radiative transfer calculations. We validate line parameters from the HITRAN-2000 database and from the ULB-UFR-BIRA database which have been used for the 2004 release of HITRAN. For this purpose, ground-based high resolution observations of the direct solar radiation are compared with simulated spectra. Line parameters are selected for validation based on a full error characterization including instrument noise and uncertainties in the atmospheric profile data. For a subset of the line parameters validated in the present study independent high-accuracy data are available. The comparison with these independent reference data confirms the capability of our approach to improve water vapor spectroscopic data.

PACS: 95.75.Fg; 33.20.Kf; 92.60.Jq; 42.68.Wt; 94.10.Fg; 94.10.Gb; 92.60.Vb

Keywords: Spectroscopic database; Validation; Water vapor; Atmospheric absorption; Line parameter; Ground based measurement; Solar occultation

1. Introduction

Water vapor is an atmospheric trace gas with a significant impact on the global radiation balance. The considerable radiative forcing properties of atmospheric water vapor are well documented (cf. [1]). Atmospheric water vapor is routinely monitored using spectral reflectance measurements from satellite instruments like GOME-2 (Global Ozone Monitoring Experiment) on the MetOp satellite or SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) on ENVISAT. The water

☆ Supplementary data for this article are available on ScienceDirect.

*Corresponding author. The Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands. Tel.: +31 30 2206699; fax: +31 30 2210407.

E-mail address: veihelma@knmi.nl (B. Veihelmann).

0022-4073/$ - see front matter © 2007 Published by Elsevier Ltd.
doi:10.1016/j.jqsrt.2007.02.005

vapor abundance is retrieved from water vapor absorption signatures in the visible [2,3] and in the infrared [4–6]. It has been shown that the water vapor column retrievals are strongly dependent on the accuracy of the spectroscopic parameters [7]. We focus in this study on the validation of spectroscopic line parameters of water vapor taken from spectroscopic databases such as HITRAN-2000 [8] and a database released by Coheur et al. [9] referred to as ULB-UFR-BIRA in the following, which has been used to update the HITRAN-2004 database [10].

The construction of spectroscopic databases containing line parameters is difficult. Water vapor line intensities cover many orders of magnitude [11]. Long-path absorption cell experiments are required for weak lines. Such experiments can be complicated by condensation inside the absorption cell and by difficulties in determining the absorption path length [9]. Short-path cells are required for strong lines. *A priori* data taken from existing databases and from theory are used for the retrieval of line parameters from measured absorption spectra [12]. Theoretical data are required as well for the line assignment [13] which is necessary for the determination of the temperature-dependence of line intensities (e.g. [14]). Such spectroscopic data in the visible and the near infrared are usually less accurate than at longer wavelengths [15]. The characterization of water lines in the visible and the near infrared is experimentally more challenging because the line intensities in this part of the spectrum are on average weaker than in the thermal and far infrared. The spectroscopic data relevant to atmospheric radiative transfer simulations thus come from a variety of sources including experiments made under conditions that may differ significantly from atmospheric conditions. It is therefore essential to validate these spectroscopic data in the context of atmospheric radiative transfer calculations on real measurements.

It has been shown in a number of past studies that spectroscopic data can be verified in general terms using ground-based observations of direct solar radiation under cloud-free conditions. For water vapor this has been done by Smith et al. [16], Veihelmann et al. [17], Sierk et al. [18], Coheur et al. [19] and Albert et al. [20] for a number of bands between 10,000 and 20,000 cm^{-1}. In these studies, ground-based Fourier transform and grating spectrometer measurements are compared with radiative transfer simulations based on the spectroscopic database to be evaluated. The approach thus far has generally been to identify overall scaling problems of band-integrated intensities or to make broad statements about the overall quality of the database under study. What makes our approach here more general is that a full characterization of instrument and trace gas profile uncertainties are used to make specific statements about individual line parameters.

As a case study we evaluate spectroscopic data for the 720 nm absorption band (4ν polyad). The data are taken from two sources: HITRAN-2000 and ULB-UFR-BIRA (Section 5). These are relevant both for water vapor retrievals from satellite observations and for radiative transfer models (cf. [21,22]). The reference spectrum is derived from ground-based Fourier Transform Radiometer (FTR) measurements of direct sunlight. This measurement has been made using an FTR with a resolution high enough to resolve individual lines (0.05 cm^{-1}) [23]. The measurement and the atmospheric data used for the simulation are discussed in Section 2.

Our approach is first to identify the limiting factors of the validation of line intensity parameters S and the air-broadening parameters γ_{air} (Section 3). Then a general expression is derived that allows to estimate whether a line parameter with a given accuracy can be validated depending on the uncertainty of the atmospheric water vapor abundance and the signal-to-noise ratio of the FTR measurement. Line parameters are identified in ULB-UFR-BIRA and HITRAN-2000 that can be validated with the present FTR solar occultation measurement.

For these lines, errors in the line parameter are detected from the residuals, i.e. the difference between simulated and measured spectra. We specify the error type for cases when the shape of the residuals allows the distinction between errors in the line intensity, the line width and the line position (Section 5). The identification of line intensity parameter errors is validated using as reference an independent measurement dedicated to the characterization of weak lines with an exceptionally high signal-to-noise ratio [24]. The line intensity data of the ULB-UFR-BIRA database are identical with the line intensity data in HITRAN-2004 for the frequency range considered in this study. The results of the study presented concerning line intensities are therefore directly relevant to HITRAN-2004.

2. Experiment and model

In this study we evaluate spectroscopic data from HITRAN-2000 [8] and the ULB-UFR-BIRA databases [9], much of which form the input for the 2004 release of HITRAN. The water vapor spectroscopy data given

in HITRAN-2000 for the wavelength region studied here originally come from laboratory measurements by Mandin et al. [25]. The 2001 update of HITRAN does not affect the water vapor spectroscopy in this wavelength region and is not considered separately. The ULB-UFR-BIRA database contains line positions, integrated intensities and broadening parameters for more than 9000 lines in the region from 13,000 to 26,000 cm^{-1}. In the spectral region studied here, ranging from 13,500 to 14,300 cm^{-1}, the database contains about 850 weak lines that are absent in HITRAN-2000.

The field measurement has been made on 21 August 1999 between 13:25 and 13:57 (GMT) using a Bomem DA3.002 Fourier transform spectrometer that was located at 51.6°N, 1.3°W at an altitude of 135 m above sea-level. The measurement consisted of 20 co-added scans. The instrumental setup has been used in a series of measurement campaigns that have been conducted since 1997 [23]. The effects of any misalignment present in the optical system are negligible. This has been shown in a dedicated absorption cell measurement using the well characterized iodine absorption line at 18,628.93 cm^{-1}. Therefore we approximate the instrument line shape by a sinc function calculated for the appropriate maximum optical path difference. Instrument effects are accounted for by convolving the simulated spectra at the bottom of the atmosphere with this instrumental line shape function. We convolve both the measured as well as the simulated spectra with a Norton Beer strong apodization function [26] in order to facilitate the interpretation of the spectra and the residuals.

The instrument is equipped with a telescope with a circular field of view pointing at the sun. The radius of the disc viewed by the instrument is about a fifth of the radius of the total solar disc. Multiple scattering can be neglected in radiative transfer simulations in this observation geometry [27]. The instrument was operated at a maximum unapodized spectral resolution of 0.05 cm^{-1}, allowing individual atmospheric absorption lines to be resolved in the recorded spectra. The atmospheric spectra used in this study have not been published elsewhere.

The solar spectrum used for the simulation of the measured spectra is derived from high resolution measurements made using the McMath–Pierce Telescope at the National Solar Observatory (NSO) on the Kitt Peak, Arizona, USA on 2100 m above sea level. This spectrum has been corrected for atmospheric absorption by water vapor above the telescope [28] using a Langley-plot technique. In spectral regions with strong absorption by water vapor the extrapolation to zero air-mass is unstable and the solar spectrum is not given. This is the case close to lines with intensities larger than 10^{-24} cm/mol. In order to exclude errors due to Fraunhofer lines, we also discard spectral regions where the corrected solar spectrum has an intensity of less than 0.95 times the background intensity. Weak solar lines are not masked out.

The atmosphere is modeled at 25 homogeneous layers with a thickness of 1 km accounting for molecular absorption by water vapor and O_2 as well as by Rayleigh scattering. Absorption by water vapor is simulated based on the spectroscopic data given in either of HITRAN-2000 and ULB-UFR-BIRA. The profiles of the molecular abundance as well as pressure and temperature are interpolated over the slant path defined by the line-of-sight of the measurement. The interpolation is based on data from the European Center for Medium-range Weather Forecast (ECMWF) and from a radiosonde measurement at the weather station Larkhill which is 48 km south of the measurement site. The uncertainty of the water vapor abundance thus obtained is between 10% and 30% according to the source of these data. A comparison of simulated and measured spectra strongly suggests that the true water vapor column density is significantly lower. Therefore, we retrieve the water vapor column from the measured spectra using a non-linear, least squares fitting routine which minimizes the norm of the residuals independently for eight adjacent 100 cm^{-1}-broad spectral regions of one measurement. The columns obtained from these independent retrievals agree well with each other (with a standard deviation of about 4% of the mean values) but differ from the initial data, by approximately 30%. An improved water vapor profile is hence derived by scaling the original ECMWF data combined with radiosonde measurements by the optimized column value. The fitting of the water vapor column compensates for any scaling of the global level of line intensities within the band. As a consequence, line intensity errors derived from the residuals have to be regarded as relative to the overall level of line intensities.

The molecular absorption cross section is calculated using a Voigt line shape, notwithstanding the possibility of collisional narrowing effects [29]. According to a continuum model of Ma et al. [30], the Voigt line shape representation is sufficiently accurate for our purposes.

The unknown instrument gain as well as broad-band absorption contributions that are not modeled are accounted for by a multiplicative baseline that is fitted to the measurement.

3. Sensitivity analysis

In this section, the importance of uncertainties of the atmospheric profile data for line parameter validation is discussed. The sensitivity of the spectrum to the atmospheric pressure, temperature and water vapor abundance is compared with the sensitivity of the spectrum to line intensity and air-broadening parameters S and γ_{air}. The sensitivities are determined using error propagation calculations that are based on the Jacobian matrix $\mathbf{K}$ of the linearized forward model

$$(\mathbf{I}-\mathbf{I}_0)=\mathrm{K}(\mathbf{x}-\mathbf{x}_0) \quad \text{with } K_{ij}=\frac{\partial I_i}{\partial x_j}. \tag{1}$$

The vector $\mathbf{I}$ denotes the discrete intensity spectrum, while the vector $\mathbf{x}$ represents either a discrete representation of one of the atmospheric profiles, or one of the line parameters for the series of lines considered. $\mathbf{I}_0$ and $\mathbf{x}_0$ denote the reference state, defined by the conditions of the ground-based measurement. The matrix $\mathbf{K}$ relates the error covariance matrix $\mathbf{C}_x$ of the vector $\mathbf{x}$ with the error covariance matrix $\mathbf{C_I}$ of the simulated discrete intensity spectrum $\mathbf{I}$, viz.

$$\mathbf{C_I}=\mathbf{K}\mathbf{C}_x\mathbf{K}^{\mathbf{T}}. \tag{2}$$

The square root of the diagonal elements of $\mathbf{C_I}$ are the standard deviations of the measured intensities $\sigma_{I,x}$.

Fig. 1 shows the sensitivity analysis of the simulated spectrum for a $10\,\mathrm{cm}^{-1}$ wide window that contains two very strong absorption lines ($S=3\times10^{-23}\,\mathrm{cm/mol}$) that are saturated at the line center as well as weaker lines that show no saturation. The simulated absorption spectrum is depicted in the upper left graph (a). The sensitivities of the spectrum to the line intensities $\sigma_{I,S}$ (solid) and the broadening parameters $\sigma_{I,\gamma}$ (dashed) are shown as a function of the wavenumber (b) assuming an uncertainty of 10% for both parameters. This is a realistic estimate for strong lines. The relative uncertainty of the line parameters of weak lines can be much larger. The sensitivity $\sigma_{I,S}$ peaks at the line centers of unsaturated lines. For saturated lines $\sigma_{I,S}$ reaches a maximum at the rising edges of the line and decreases to a minimum at the center. According to Beer–Lambert's Law, the sensitivity of the simulated spectrum with respect to the line intensity is largest where the optical thickness equals unity. The error signature of line broadening parameters $\sigma_{I,\gamma}$ usually has distinct contributions at the line center and in the line wings. This is hardly visible due to the blending of lines. The amplitude of $\sigma_{I,\gamma}$ is about 20% less than the amplitude of $\sigma_{I,S}$ (at line centers of unsaturated lines).

The sensitivities $\sigma_{I,p}$ and $\sigma_{I,T}$ are almost an order of magnitude smaller than the sensitivities $\sigma_{I,S}$ and $\sigma_{I,\gamma}$ assuming statistically independent errors of 2% for both the pressure and the temperature values in each height level. This error magnitude is typical for commonly available pressure and temperature data.

The impact of a water vapor profile error $\sigma_{I,\mathrm{w_p}}$ is depicted as a dashed line in the lower left graph (c) assuming statistically independent errors of 10% in the sub-columns. This error magnitude is representative for the uncertainty radiosonde measurements and also for the water vapor data used in this study. The impact of an error of 10% in the total integrated column density $\sigma_{I,\mathrm{w_c}}$ is shown as a solid line. Column errors of this magnitude are common to ECMWF data.

The sensitivities $\sigma_{I,\mathrm{w_c}}$ and $\sigma_{I,S}$ have almost the same shape. We note that the amplitude of the sensitivities is generally similar when the same relative error is assumed since both the column and the line intensity appear as a product in Beer–Lambert's Law. Hence, only line intensity parameters known with an uncertainty larger than the uncertainty of the water vapor column density can be validated. In the present study the validation is limited by uncertainties in the water vapor data rather than by uncertainties in the pressure and temperature data. This is also true for studies using other commonly available atmospheric data.

4. Accuracy limits and line intensity

The validation of the line parameters using a ground-based FTR spectrum is also limited by instrument noise. In this section a general expression is derived that allows to estimate whether a line parameter with a given accuracy can be validated or not depending on the uncertainty of the water vapor data and the signal-to-noise ratio of the instrument. The sensitivities of the line parameters $\tilde{\sigma}_S$ and $\tilde{\sigma}_{\gamma\mathrm{air}}$ to instrument noise and to uncertainties in the atmospheric water vapor data are a measure for how much information the ground-based

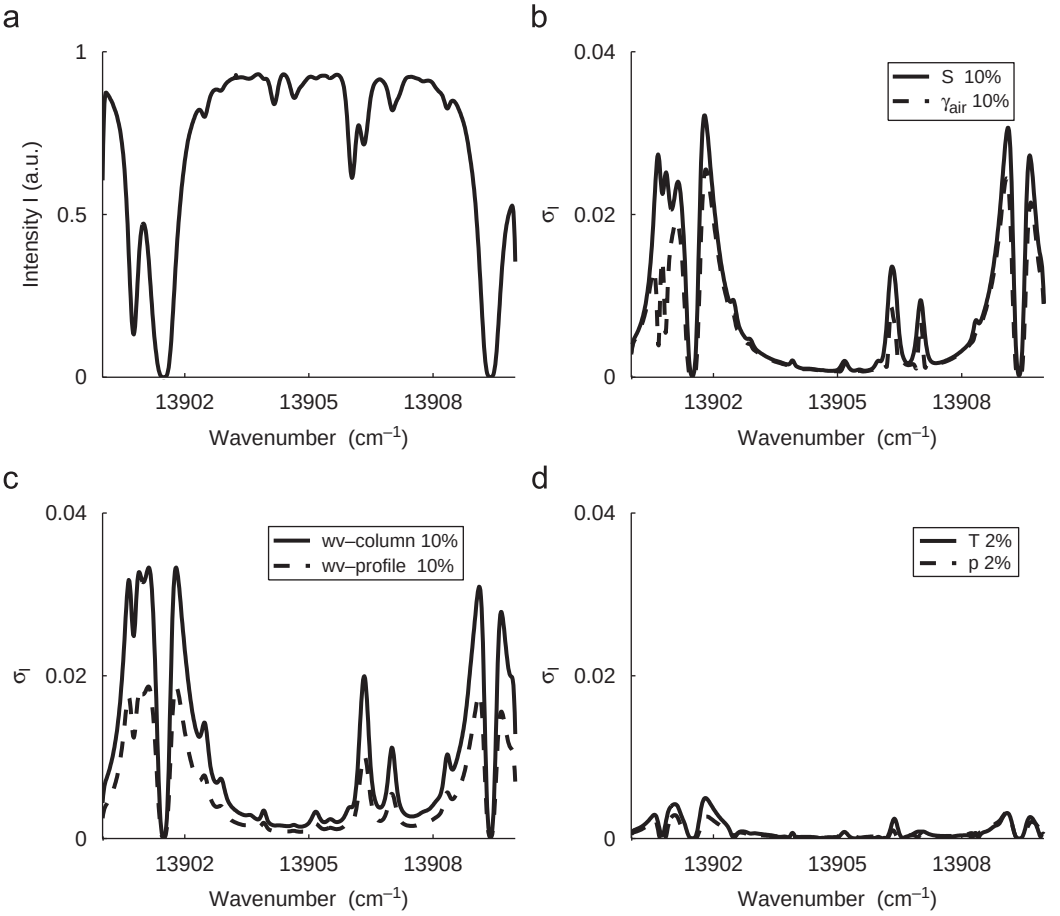

Fig. 1. The simulated intensity spectrum I (a) and its sensitivity σ_I with respect to various parameters (b, c, d). The amplitudes of the sensitivities with respect to line parameter errors (b) and with respect to errors in the atmospheric water vapor data (c) are comparable. We consider errors of 10% in the line intensity (b, solid) and air-broadening (b, dashed) parameters, the water vapor column (c, solid) and the water vapor profile (c, dashed). The sensitivities with respect to errors of 2% in the atmospheric pressure (d, dashed) and temperature data (d, solid) have a lower amplitude.

measurement can add to the line parameters S and γ_{air}. Line parameters given with an accuracy lower than the related sensitivity can be validated. We consider the ground-based measurement to be a significant test also for line parameters for which the difference of values given in different databases is larger than the related sensitivity. The sensitivities of the line parameters $\tilde{\sigma}_S$ and $\tilde{\sigma}_{\gamma\mathrm{air}}$ are determined as follows: The error covariance matrix of the line parameters $\mathbf{C}_{\mathrm{p}}$ is derived from the covariance matrix of the measurement $\mathbf{C_I}$ using (e.g. [31])

$$\mathbf{C}_{\mathrm{p}} = (\mathbf{K}_{\mathrm{p}}^{\mathbf{T}}\mathbf{C}_{\mathbf{I}}^{-1}\mathbf{K}_{\mathrm{p}})^{-1}. \tag{3}$$

The Jacobian matrix $\mathbf{K}_{\mathrm{p}}$ is taken from the linearized forward model. The error covariance matrix $\mathbf{C_I}$ represents alternatively the instrument noise or the uncertainty of the simulated spectrum due to errors in the water vapor abundance (see Section 3). The errors in line intensity and air-broadening parameters are treated together in one combined error covariance matrix $\mathbf{C}_{\mathrm{p}}$ which contains both variance as well as covariance terms. The cross-correlation of errors in S and γ_{air} for each line are thus taken into account. If n lines are considered, the variances of the line parameters S and γ_{air} are the diagonal elements,

$$\tilde{\sigma}_S^2(i) = \mathbf{C_p}(i,i) \quad \text{and} \quad \tilde{\sigma}_{\gamma\mathrm{air}}^2(i) = \mathbf{C_p}(n+i,n+i), \tag{4}$$

of the matrix $\mathbf{C_p}$ with the dimensions $(2n \times 2n)$.

In Fig. 2 the sensitivities of the line parameters S (upper graph) and γ_{air} (lower graph) are depicted as relative errors $\tilde{\sigma}_S/S$ and $\tilde{\sigma}_{\gamma\mathrm{air}}/\gamma_{\mathrm{air}}$ as a function of the line intensity S. The instrument noise is assumed to be constant for the spectral region considered with an amplitude of 0.001 times the baseline intensity of the spectrum. A water vapor profile error of 10% is assumed, which corresponds to a column error of about 5.6%. The relative errors of the line parameters due to instrument noise (dots) show a decreasing trend with increasing line intensity. This behavior is sketched with the slant line determined from a linear regression made on double-logarithmic scales. The relative errors due to the water vapor profile errors (gray crosses clustered around horizontal line) are largely independent of the line intensities. The large spread of the errors is due to the blending of lines, which can cause large and strongly correlated errors of line parameters.

For lines with an intensity lower than a threshold intensity $S_{\mathrm{thr}} \approx 2 \times 10^{-25}\,\mathrm{cm/mol}$ the validation of line intensity parameters is mainly limited by noise. For stronger lines it is mainly limited by uncertainties in the water vapor profile (horizontal line). The validation of air-broadening parameters is mainly limited by instrumental noise for lines with an intensity up to $S_{\mathrm{thr}} \approx 2 \times 10^{-24}\,\mathrm{cm/mol}$ and by uncertainties in the water vapor data for stronger lines (lower plot in Fig. 2).

Based on the fitted lines (in Fig. 2) we obtain a formula relating $\tilde{\sigma}_S/S$ to the line intensity

$$\frac{\tilde{\sigma}_S}{S} = \begin{cases} 10^{-26.2} \cdot S^{-0.94}: & S < 2 \times 10^{-25}\,\mathrm{cm/mol}, \\ 0.053: & S \geqslant 2 \times 10^{-25}\,\mathrm{cm/mol}. \end{cases} \tag{5}$$

This formula is derived from a case with a signal-to-noise ratio of $\xi = 1000$ and a water vapor column $w_{\mathrm{c}} = 5.8 \times 10^{22}\,\mathrm{molec/cm^2}$ with a relative uncertainty $\sigma_{w_{\mathrm{c}}}/w_{\mathrm{c}} = 0.056$. The relative error $\tilde{\sigma}_S/S$ due to noise is inversely proportional to the signal-to-noise ratio ξ and inversely proportional to the integrated water vapor column w_{c} given in molec/cm^2. The relative error $\tilde{\sigma}_S/S$ due to uncertainties in the water vapor information scales with the relative error of the total water vapor column $\sigma_{w_{\mathrm{c}}}/w_{\mathrm{c}}$. Using these relations, we generalize Eq. (5) to the approximation

$$\tilde{\sigma}_S \approx \begin{cases} 0.36 \cdot \dfrac{1}{\xi w_{\mathrm{c}}}: & S < 2 \times 10^{-25}\,\mathrm{cm/mol}, \\ 0.95 \cdot S \dfrac{\sigma_{w_{\mathrm{c}}}}{w_{\mathrm{c}}}: & S \geqslant 2 \times 10^{-25}\,\mathrm{cm/mol}. \end{cases} \tag{6}$$

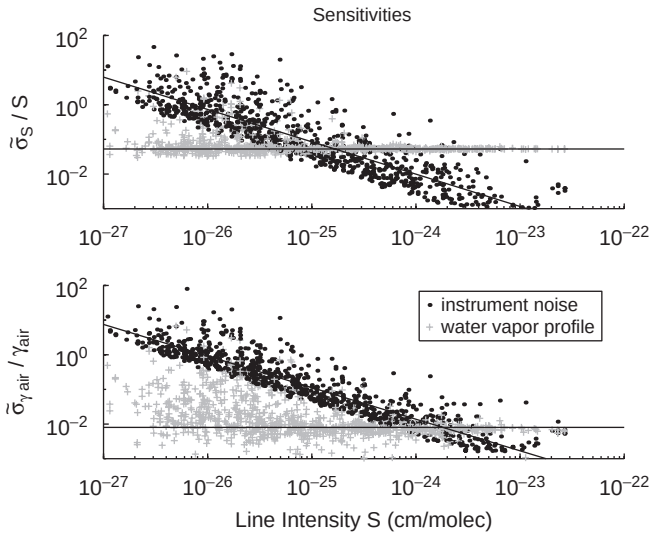

Fig. 2. The sensitivities of line intensities S (upper graph) and air-broadening parameters γ_{air} (lower graph) to instrument noise and to the water vapor profile are shown as relative errors $\tilde{\sigma}_S/S$ and $\tilde{\sigma}_{\gamma\mathrm{air}}/\gamma_{\mathrm{air}}$. Errors due to instrument noise with a signal-to-noise ratio of 1000 are shown as dots (clustered around the slanted line), errors due to an uncertainty of the water vapor profile of 10% are shown as gray crosses (clustered around the horizontal line).

For the air-broadening parameter we obtain a similar empirical relation

$$\tilde{\sigma}_{\gamma_{\mathrm{air}}} \approx \begin{cases} 0.36 \cdot \dfrac{1}{\xi w_{\mathrm{c}}}: & S < 2 \times 10^{-24}\,\mathrm{cm/mol}, \\ 0.14 \cdot S\dfrac{\sigma_{w_{\mathrm{c}}}}{w_{\mathrm{c}}}: & S \geqslant 2 \times 10^{-24}\,\mathrm{cm/mol}. \end{cases} \tag{7}$$

With these relations a first-order estimate for the sensitivities $\tilde{\sigma}_S$ and $\tilde{\sigma}_{\gamma_{\mathrm{air}}}$ as functions of the line intensity can be inferred for experimental conditions that deviate from the conditions of the measurement presented here.

The accuracies of line parameters given in the ULB-UFR-BIRA database for the spectral region from $13,500$ to $14,300\,\mathrm{cm}^{-1}$ are depicted as relative errors σ_S/S (upper graph) and $\sigma_{\gamma\mathrm{air}}/\gamma_{\mathrm{air}}$ (lower graph) in Fig. 3. The comparison with the trends of $\tilde{\sigma}_S/S$ and $\tilde{\sigma}_{\gamma\mathrm{air}}/\gamma_{\mathrm{air}}$ (solid lines) indicates that the ground-based measurement adds information about the line intensity parameters of ~10% of the lines (black dots, upper graph) and about air-broadening parameters of ~50% of the lines (black dots, lower graph) is possible. The accuracies of line parameters that cannot be validated are denoted by gray dots.

In Fig. 4 the differences between line parameters given in HITRAN-2000 and the ULB-UFR-BIRA database are shown for the spectral region from $13,500$ to $14,300\,\mathrm{cm}^{-1}$. The relative differences in the line intensity S (upper plot) and the air-broadening parameters γ_{air} (lower plot) are smaller than the relative errors $\tilde{\sigma}_S/S$ and $\tilde{\sigma}_{\gamma\mathrm{air}}/\gamma_{\mathrm{air}}$ for about 50% of the lines (gray dots). For about 50% of the lines the measurement is thus a significant test (black dots).

5. Interpretation of the residuals

The spectroscopic data for the 720 nm band given in HITRAN-2000 and the ULB-UFR-BIRA database are evaluated based on details of the residuals, i.e. the difference between the simulated and measured spectra (Figs. 5 and 6). For the validation of line parameters, lines are selected according to the limitations derived in the previous section. With this we exclude line parameters that are sensitive to errors in the atmospheric water vapor profile derived in Section 2. The intensity of the solar spectrum used in this study is not available close to water vapor absorption lines with intensities larger than $S = 10^{-24}\,\mathrm{cm/mol}$ (see Section 2). This poses an upper limit to the intensity of lines that we validate. Furthermore, we mask out strong Fraunhofer lines (black box). For ULB-UFR-BIRA, we only consider lines where all line parameters are given. With this selection we

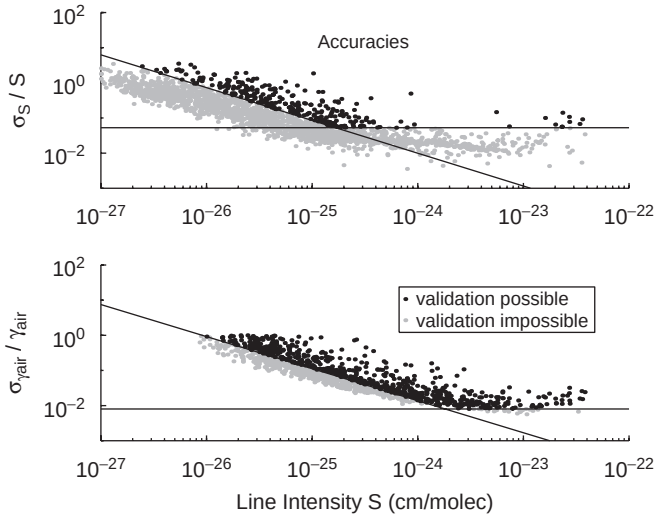

Fig. 3. Accuracy of line parameters given in the ULB-UFR-BIRA database for the spectral region from $13,500$ to $14,300\,\mathrm{cm}^{-1}$. The accuracies are shown as relative errors, σ_S/S (upper graph) and $\sigma_{\gamma\mathrm{air}}/\gamma_{\mathrm{air}}$ (lower graph). The line parameters that can be validated using the ground-based measurement are shown as black dots. Line parameters shown in gray cannot be validated with the current measurement uncertainties. The solid lines are the same as the ones shown in Fig. 2.

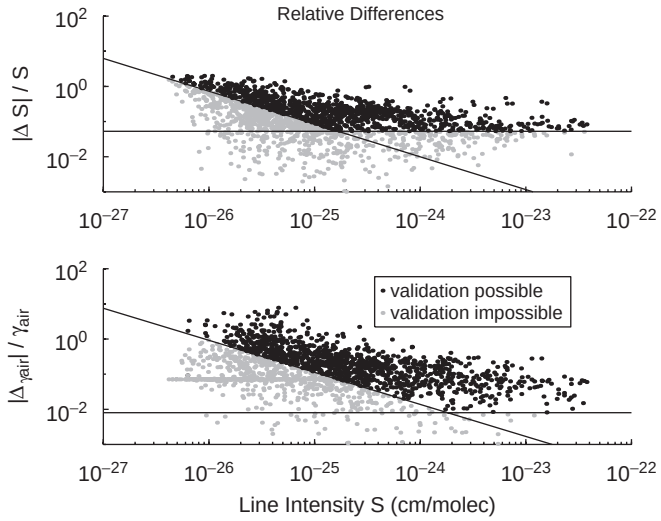

Fig. 4. The amplitude of relative differences of line parameters $|S_{\mathrm{hit01}} - S_{\mathrm{ulb04}}|/S_{\mathrm{ulb04}}$ (upper graph) and $|\gamma_{\mathrm{air,hit01}} - \gamma_{\mathrm{air,ulb04}}|/\gamma_{\mathrm{air,ulb04}}$ (lower graph) given in HITRAN-2000 and ULB-UFR-BIRA. Line parameter differences that exceed the uncertainty of the ground-based measurement are shown as black dots. The other line parameter differences are shown gray. The solid lines are the same as in Fig. 2.

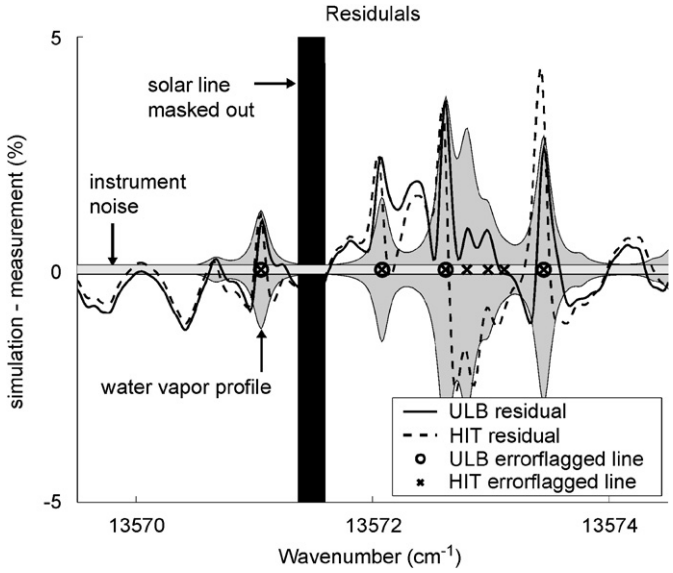

Fig. 5. Residuals, i.e. differences between measured spectra and simulations based on ULB-UFR-BIRA (solid) and HITRAN-2000 (dashed). Error-flagged lines are marked with a circle (ULB) and a cross (HITRAN-2000). The intensity error due to instrument noise (signal-to-noise ratio 1000, light gray) and due to an uncertainty of a 20% water vapor column error (dark gray) is depicted as 1-σ-environment. Spectral regions that may be affected by errors related to the solar spectrum are masked out (black).

obtain 786 lines between 13, 500 and 14, 300 cm^{-1} for which we evaluate the residuals. For HITRAN-2000 we include lines lacking a pressure shift parameter, which is missing for all lines in the spectral region studied. The result is that 830 lines are selected in the same spectral range.

Line parameter errors are identified from the residual close to the center of selected lines. The residual is regarded as significant where it exceeds the uncertainty of the simulated (light gray area) assuming a 20% water vapor column error and the uncertainty of the measured spectrum (dark gray area) with a signal-to-noise ratio of 1000.

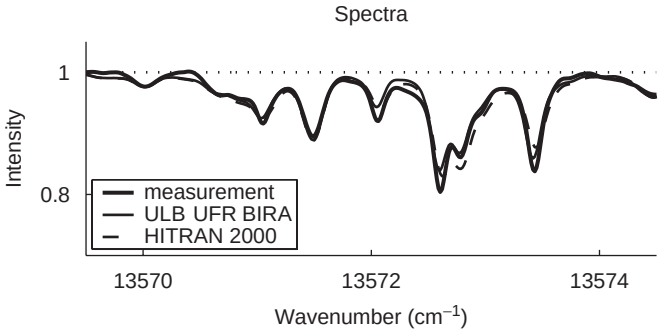

Fig. 6. Measured spectrum (thick solid) and simulations based on ULB-UFR-BIRA (solid) and HITRAN-2000 (dashed) corresponding to the residuals shown in Fig. 5. The background is normalized to unity.

The residuals of both ULB-UFR-BIRA (dashed) and HITRAN 2000 (solid) show signatures of typical line parameter errors (e.g. at $13,571.1$ and $13,572.1\,\mathrm{cm}^{-1}$) as well as signatures that we cannot assign to any line parameter error (e.g. at $13,570.5\,\mathrm{cm}^{-1}$). These signatures may be due to instrument effects. The solar spectrum used is another possible error source. Unknown weak lines may also contribute to the residuals. In the table that is given as electronic supplement to this article we list 129 corresponding pairs of lines of the two databases where at least one line is flagged as erroneous. 82 lines of ULB-UFR-BIRA and 100 lines of HITRAN-2000 are flagged. The wavenumber, the line intensity S and the air-broadening parameter γ_{air} of lines are listed together with an error key indicating the type of error. We distinguish line width errors (w), line intensity errors (i) and line position errors (p). Combined errors of the three error types are marked with a 'c', unclassified signatures in the residuals with a significant amplitude are marked with an 'x' and lines with no error signature are marked with a '+'. We observe that many line intensity errors in HITRAN-2000 have been improved in ULB-UFR-BIRA. Nevertheless, we recognize characteristic error signatures in both databases. This number of error flagged lines is too low to reflect the relative quality of the databases with statistical significance. According to the root mean square values of the residuals the ULB-UFR-BIRA (0.0105) database is slightly better than HITRAN (0.0122) in the spectral region from $13,500$ to $14,300\,\mathrm{cm}^{-1}$.

The results of the classification of line intensity parameters are compared with independent high-accuracy data. Schermaul et al. [24] have measured the line intensity parameters of pure water vapor in the wavelength region $13,200$ to $15,000\,\mathrm{cm}^{-1}$ using a long-path absorption cell in combination with an FTR. This measurement is dedicated to the characterization of weak lines with an intensity $S \leqslant 6 \times 10^{-24}\,\mathrm{cm/mol}$ and has an exceptionally high signal-to-noise ratio (3600 as compared to usually 1200). The Schermaul et al. data do not contain the line parameters of the strong lines. That is why we do not perform radiative transfer simulations based on these data in the same way as we do for HITRAN-2000 and ULB-UFR-BIRA database. The data given by Schermaul et al. [24] are independent from the water–air measurements used in the ESA-WVR database [32]. The line positions and intensity parameters are included in the table (see electronic supplement). The error identification and classification for lines given in HITRAN-2000 and in the ULB-UFR-BIRA database is confirmed in 74% of the cases when a corresponding line occurs in the data given by Schermaul et al. [24].

We have decided not to derive an improved set of line parameters since the lines in the spectrum evaluated here are mostly blended and since the signatures in the residual often indicate a combination of line intensity, line width and line position errors. The list of error-flagged line parameters provides an input for future evaluations of long-path absorption cell experiments or theoretical studies.

6. Conclusions

We have validated spectroscopic data for the water vapor band close to 720 nm given in HITRAN-2000 and the ULB-UFR-BIRA database in the context of atmospheric radiative transfer simulations using spectra from a ground-based Fourier transform spectrometer. Atmospheric profile data from ECMWF and a nearby radiosonde measurement have been used for radiative transfer simulations.

Line parameters have been selected for validation based on a full characterization of the inherent uncertainties in the measurement situation on a line specific basis. This includes the atmospheric profile information used for the simulation and the contamination of the measurement by instrument noise. According to this error analysis, the validation of line intensity parameters is limited by instrument noise for lines with intensities $S<2\times10^{-25}$ cm/mol. For lines with intensities $S>2\times10^{-25}$ cm/mol the validation of line intensity is limited by uncertainties in the atmospheric water vapor profile. The validation of air-broadening parameters is limited by instrument noise for lines with intensities lower than $S_{\mathrm{thr}}=2\times10^{-24}$ cm/mol and by uncertainties in the water vapor data for stronger lines. The temperature and pressure data do not pose a limit to the line parameter validation using commonly available atmospheric data. Generalizing the results of this analysis we have derived relations to estimate whether line parameters with a given accuracy can be validated also for experimental conditions that deviate from the conditions of the measurement presented here.

We have found that the ground-based measurement adds information to about 10% of the line intensity parameters of databases such as HITRAN-2000 and the ULB-UFR-BIRA. The lines that can be validated have mostly line intensities lower than 5×10^{-25} cm/mol. The air-broadening parameters of as much as 50% of the lines can be validated. This includes both strong as well as weak lines. For about half of the line parameters it is possible to decide which database of HITRAN-2000 and ULB-UFR-BIRA provides a value that is more consistent with the atmospheric measurement. The line parameter errors that have been identified in these database are listed in the electronic supplement.

The results are confirmed based on independent data from an experiment that is dedicated to the characterization of weak lines. The results of the validation study presented here therefore provide a useful input for future evaluations of laboratory or theoretical data for the improvement of the water vapor spectroscopy. Recently the ULB-UFR-BIRA spectra have been re-analyzed over the entire range from 9250 to $26,000\,\mathrm{cm}^{-1}$ [33]. This re-analysis, which both re-assigned and dealt with many problems associated with line widths, also aimed to address the issues raised in the presented study and has been used as input for HITRAN-2004. The results of the validation study presented here has thus already facilitated a line list compilation used as input for HITRAN-2004.

Acknowledgments

The Natural Environment Research Council (NERC, UK) supported the field measurements through Grant Ref. GR3/11097, and provided access to the British Atmospheric Data Centre (BADC). This work is part of the Research Program of the Stichting voor Fundamenteel Onderzoek der Materie (FOM) and has been partially funded by the Stichting voor Nederlands Wetenschappelijk Onderzoek (NWO) through projects EO-023 and EO-046 of the Gebruikers Ondersteuning (GO) program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jqsrt.2007.02.005.

References

[1] Kiehl JT, Trenberth KE. Earth's annual mean global energy budget. Bull Amer Meteorol Soc 1997;78:197–208.
[2] Lang R, Maurellis AN, van der Zande WJ, Aben I, Landgraf J, Ubachs W. Forward modelling and retrieval of water vapor from GOME: treatment of narrow band absorption spectra. J Geo Res 2002;107.
[3] Maurellis AN, Lang R, van der Zande WJ, Aben I, Ubachs W. Precipitable water column retrieval from GOME data. Geophys Res Lett 2000;27:903–6.
[4] Noël S, Buchwitz M, Bovensmann H, Hoogen R, Burrows JP. Atmospheric water vapor amounts retrieved from GOME satellite data. Geophys Res Lett 1999;26:1841–4.
[5] Casadio S, Zehner C, Piscane G, Putz E. Empirical retrieval of atmospheric air mass factor (ERA) for the measurement of water vapor vertical content using GOME data. Geophys Res Lett 2000;27:1483–6.

[6] Wagner T, Heland J, Zöger M, Platt U. Atmos Chem Phys 2003;3:651–63.
[7] Maurellis AN, Lang R, Williams JE, van der Zande WJ, Smith K, Newnham DA, et al. The impact of new water vapor spectroscopy on satellite retrievals. In: Camy-Peyret C, Vigasin A, editors. Proceedings of the NATO advanced research workshop "Weakly interacting molecular pairs: unconventional absorbers of radiation in the atmosphere"; 2003. p. 259–72.
[8] Rothman LS, et al. The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001. JQSRT 2003; 82:5–44.
[9] Coheur P-F, Fally S, Carleer M, Clerbaux C, Colin R, Jenouvrier A, et al. New water vapor line parameters in the 26 000–13 000 cm^{-1} region. JQSRT 2002;74:493–510.
[10] Rothman LS, et al. The HITRAN 2004 molecular spectroscopic database. JQSRT 2005;96:139–204.
[11] Tolchenov RN, Tennyson J, Shirin SV, Zobov NF, Polyanski OL, Maurellis AN. Water line parameters for weak lines in the range 9000–12 700 cm^{-1}. J Mol Spectrosc 2003;221:99–105.
[12] Schermaul R, Learner RCM, Newnham DA, Ballard J, Zobov NF, Belmiloud D, et al. The water vapor spectrum in the region 8600–15, 000 cm^{-1}: experimental and theoretical studies for a new spectral line database. II. Linelist construction. J Mol Spectrosc 2001;208:43–50.
[13] Polyansky OL, Zobov NF, Viti S, Tennyson J. Water vapor line assignments in the near infrared. J Mol Spectrosc 1998;189:291–300.
[14] Lenoble J. Atmospheric radiative transfer. Hampton: Deepak; 1993.
[15] Polyansky OL, Császár AG, Shirin SV, Zobov NF, Barletta P, Tennyson J, et al. High-accuracy ab initio rotation–vibration transitions for water. Science 2003;299:539–42.
[16] Smith KM, Newnham DA. High-resolution atmospheric absorption by water vapor in the 830–985 nm region: evaluation of spectral databases. Geophys Res Lett 2001;28:3115–8.
[17] Veihelmann B, Lang R, Smith KM, Newnham DA, van der Zande WJ. Evaluation of spectroscopic databases of water vapor between 585 and 600 nm. Geophys Res Lett 2002;29.
[18] Sierk B, Solomon S, Daniel JS, Portmann RW, Gutman SI, Langford AO, et al. Field test of spectral line intensity parameters for tropospheric water vapor. J Geo Res 2003;108(D12):4351.
[19] Coheur P-F, Clerbaux C, Carleer M, Fally S, Hurtmans D, Colin R, et al. Retrieval of atmospheric water vapor columns from FT visible solar absorption spectra and evaluation of spectroscopic databases. JQSRT 2003;82:133–50.
[20] Albert P, Smith KM, Bennartz R, Newnham DA, Fischer J. Satellite- and ground-based observations of atmospheric water vapor absorption in the 940 nm region. JQSRT 2004;84:181–93.
[21] Bennartz R, Lohmann U. Impact of improved near infrared water vapor line data on absorption of solar radiation. Geophys Res Lett 2001;28:4591–4.
[22] Smith KM, Ptashnik I, Newnham DA, Shine KP. Absorption by water vapor in the 1 to 2 μm region. JQSRT 2004;83:735–49.
[23] Smith KM, Newnham DA, Williams RG. Collision induced absorption of solar radiation in the atmosphere by molecular oxygen at 1.2 μm: field observations and model calculations. J Geo Res 2001;106:7541–52.
[24] Schermaul R, Learner RCM, Canas AAD, Brault JW, Polyansky OL, Belmiloud D, et al. Weak line water vapor spectra in the region 13 200–15 000 cm^{-1}. J Mol Spectrosc 2002;211:169–78.
[25] Mandin JY, Chevillard JP, Camy-Peyret C, Flaud J, Brault JW. The high-resolution spectrum of water vapor between 13,200 and 16, 500 cm^{-1}. J Mol Spectrosc 1986;116:167–90.
[26] Norton RH, Beer R. New apodizing functions for Fourier spectrometry. J Opt Soc Am 1976;66:259–64.
[27] Box MA, Deepak A. Atmospheric corrections to solar radiometry. Appl Opt 1979;12:1941–9.
[28] Wallace L, Hinkle K, Livingston W. An atlas of the spectrum of the solar photosphere from 13,500 to 28, 000 cm^{-1} (3570 to 7405 Å). Technical Report, Kitt Peak National Observatory/National Solar Observatory, National Optical Astronomy Observatories; 1996.
[29] Lepère M, Henry A, Valentin A, Camy-Peyret C. Diode-laser spectroscopy: line profiles of H_2O in the region of 1.38 μm. J Mol Spectrosc 2001;208:25–31.
[30] Ma Q, Tipping RH. The averaged density matrix in the coordinate representation: application to the calculation of the far-wing line shape for H_2O. J Chem Phys 1999;111:5909–21.
[31] Rodgers CD. Inverse methods for atmospheric sounding: theory and practice. London: World Scientific; 2000.
[32] Schermaul R, Learner RCM, Newnham DA, Williams RG, Ballard J, Zobov NF, et al. The water vapour spectrum in the region 8600–15 000 cm^{-1}: experimental and theoretical studies for a new spectral line database. I: Laboratory measurements. J Mol Spectrosc 2001;208:32–42.
[33] Tolchenov RN, Naumenko O, Zobov NF, Shirin SV, Polyanski OL, Tennyson J, et al. Water vapor line assignments in the 9250–2600 cm^{-1} frequency range. J Mol Spectrosc 2005;233(1):68–76.