A room temperature CO₂ line list with ab initio computed intensities

Emil Zaka, Jonathan Tennyson, Oleg L. Polyansky, Lorenzo Lodi, Nikolay F. Zobov, Sergey A. Tashkun, Valery I. Perevalov

a Department of Physics and Astronomy, University College London, London WC1E 6BT, UK
b Institute of Applied Physics, Russian Academy of Sciences, Ulyanov Street 46, Nizhny Novgorod 603950, Russia
c V.E. Zuev Institute of Atmospheric Optics, SB RAS, 1, Academician Zuev Square, Tomsk 634021, Russia

ABSTRACT

Atmospheric carbon dioxide concentrations are being closely monitored by remote sensing experiments which rely on knowing line intensities with an uncertainty of 0.5% or better. We report a theoretical study providing rotation–vibration line intensities substantially within the required accuracy based on the use of a highly accurate ab initio dipole moment surface (DMS). The theoretical model developed is used to compute CO₂ intensities with uncertainty estimates informed by cross comparing line lists calculated using pairs of potential energy surfaces (PES) and DMS’s of similar high quality. This yields lines sensitivities which are utilized in reliability analysis of our results. The final outcome is compared to recent accurate measurements as well as the HITRAN2012 database. Transition frequencies are obtained from effective Hamiltonian calculations to produce a comprehensive line list covering all ¹²C¹⁶O₂ transitions below 8000 cm⁻¹ and stronger than 10⁻²⁰ cm/molecule at T = 296 K.

1. Introduction

The quantity of carbon dioxide (CO₂) in Earth’s atmosphere is thought to have a key role in climate change and is therefore being closely monitored. Several agencies are flying experiments or whole missions, for example GOSAT [1], OCO-2 [2] and ASCENDS [3], to explicitly monitor the atmospheric CO₂ content. Similarly, international ground-based networks such as TCCON [4] and NDACC [5] are also dedicated to monitoring atmospheric CO₂ content. A major aim of this activity is to establish CO₂ concentrations at the parts per million (ppm) level or, preferably, better. These projects aim to look at overall CO₂ concentration and its variation; it is of particular interest to pinpoint where CO₂ is being produced (sources) and where it is going (sinks). This activity is clearly vital to monitoring and hopefully controlling CO₂ and hence climate change [6].

All CO₂ remote sensing activities, both from the ground and space, rely on monitoring CO₂ vibration–rotation spectra and therefore are heavily dependent on laboratory spectroscopy for reliable parameters; it is only through these parameters that atmospheric spectroscopic measurements can be interpreted. These spectroscopic parameters are of three types: line centers, line profiles and line intensities. Line centers or positions are established to high accuracy in many laboratory high resolution spectroscopy studies and in general do not require significant improvement for studies of Earth’s atmosphere. Line profiles are more difficult but significant progress on these has been made in recent years with, for example, the inclusion...
of line mixing in both the HITRAN database [7] and many retrieval models, and the move beyond Voigt profiles [8]. Here we focus on line intensities for the main isotopolog of carbon dioxide, $^{12}\text{C}^{16}\text{O}_2$.

In the laboratory it is much harder to determine accurately line intensities than line frequencies. Typical accuracies for experimental line intensity data used in atmospheric models and retrievals are only 3–10% [9–12] and, until recently, the best published measurements, e.g. Boudjaadar et al. [13], only provide accuracies in the 1–3% range, still very significantly worse than the precision of 0.3–1% required by the modern remote sensing experiments [14–16].

Recently there have been a number of laboratory measurements aimed at measuring absolute CO2 line intensities with the high accuracy needed for remote sensing [17–19,12,20–22]. With the exception of recent work by Devi et al. [21], these studies have all focused on obtaining the highest possible accuracy for a few lines or even a single line. These investigations will be discussed further below. While they clearly do not provide the volume of data needed for remote sensing studies, they do provide benchmarks that can be used to assess calculated intensities such as those provided here. Approximately 20 000 transitions of $^{12}\text{C}^{16}\text{O}_2$ have been measured experimentally; the experiments up to 2008 were reviewed by Perevalov et al. [10] and more recently by Tashkun et al. [23].

There have been a number of attempts to use theory to provide intensities for CO2. Wattson et al. [24,25] produced line lists using variational nuclear motion calculations. More recently, Huang et al. have performed a series of quantum mechanical studies giving line positions and intensities for CO2 [26–28]. In particular, Huang et al. provide the most accurate currently available potential energy surface (PES) for the CO2 system. A widely-used alternative theoretical approach is based on effective operators for the Hamiltonian and the spectroscopic dipole moment [29]. Currently, the effective Hamiltonian approach achieves one order of magnitude better accuracy for $^{12}\text{C}^{16}\text{O}_2$ frequencies than the best available PES [26]. Within this framework, the calculation of intensities requires eigenfunctions of an effective Hamiltonian whose parameters were fitted to observed positions of rotation–vibration lines as well as dipole moment operators were tuned to observed transition intensities. This approach has been used to create dedicated versions of the carbon dioxide spectroscopic databank (CDSD) for room-temperature [23] and high-temperature [30,31] applications.

Recently a number of studies have shown that it is possible to compute line intensities using dipoles from ab initio electronic structure calculations with an accuracy comparable to, or even better than, available measurements [20,32–35]. The intensity of a line depends on the transition line strength which is obtained quantum-mechanically from the integral

$$S_f = \left| \sum_\alpha \langle \mu_\alpha | f \rangle \right|^2$$  \hspace{1cm} (1)

where here $|i\rangle$ and $|f\rangle$ are the initial and final state rovibrational wavefunctions of the molecule and $\mu_\alpha$ is component of the dipole moment surface (DMS) respectively. The requirements for accurate linestrengths are therefore high quality nuclear motion wavefunctions and DMSs. Lodi and Tennyson [33] developed a procedure which provides estimated uncertainty on a transition-by-transition basis based on the evaluation of multiple line lists. They initially applied this procedure to water vapor spectra. Their data were used to replace all $^{2}\text{H}_2^{17}\text{O}$ and $^{2}\text{H}_2^{18}\text{O}$ intensities for water in the 2012 release of HITRAN [36]. These data have since been critically assessed and verified empirically for the 6450–9400 cm$^{-1}$ region [37]. The present study combines the high accuracy ab initio DMS presented by Polyansky et al. [20] and the methodology of Lodi and Tennyson, which required some extension for the CO2 problem. This is discussed in the following section.

The current release of HITRAN [36] takes its CO2 line intensities substantially from two sources: the Fourier transform measurements of Toth et al. [38] and an unpublished version of CDSD [39]. The CDSD, whose intensities are accurate to about 2–20% depending on the vibrational band, has recently been updated and released as CDSD-296 [23]. The uncertainty estimate is up to 20% for many transitions and is probably rather conservative. Recently some of us computed a new, high accuracy DMS for CO2 which we compared with new high-accuracy experiments [20] and the data in HITRAN. The comparisons suggested that the new DMS is indeed excellent. In this work we construct a new line list for $^{12}\text{C}^{16}\text{O}_2$ which we suggest will significantly improve the precision of the intensity parameters. Due to considerations associated with the DMS, this line list is restricted to transition wavenumbers below 8000 cm$^{-1}$. However, in this range the list should be comprehensive and includes transitions which have yet to be quantified experimentally. The next section presents the methodology used to construct the line list. Section 3 presents the final line list and compares our results with those from other sources. The final section gives our conclusions and plans for future work.

2. Methodology

The Lodi-Tennyson method [33] for validating linelists on a purely theoretical basis relies on the use of accurate, $ab$ initio transition intensity calculations require an accurate procedures for obtaining nuclear motion wavefunctions together with the use of at least two DMSs and two PESs. These aspects are described below.

2.1. $ab$ initio surfaces

The first stage in the molecular linelist evaluation process involves computing energy levels and rotational–vibrational wavefunctions. Our approach utilizes an exact nuclear kinetic energy operator following the framework proposed by Tennyson and Sutcliffe [40–43] and implemented in DVR3D suite [44]; the quality of the electronic PES provided is of primary importance. Energy levels and rotational–vibrational wavefunctions obtained in this way are further used in intensity calculations, requiring additionally a DMS function as an input. The accuracy of the resulting line positions depends strongly on the quality of
the PES, while line intensities are dependent on both the PES and the DMS. Therefore, in order to generate high accuracy line intensities, it is necessary to provide those two essential functions with the highest possible accuracy. The present state-of-the-art *ab initio* PESs are capable of reproducing experimental energy levels to 1 cm\(^{-1}\) accuracy, which still remains insufficient for high resolution spectroscopy purposes. Hence empirical fitting of *ab initio* surfaces has become a standard procedure. This semi-empirical approach is much less successful in the case of DMSs, partly due to technical difficulties in obtaining accurate experimental data, suggesting the use of *ab initio* DMSs is a better choice [45]. It is natural to ask how different PESs and DMSs affect energy levels and line intensities. Answering this, in turn, can shed some light on the reliability of line intensities provided by our theoretical scheme. Accordingly, the present study involves 6 independent runs of nuclear motion calculations using the inputs presented below.

### 2.1. Ames PES

As a primary choice we decided to use the semi-empirical Ames-1 PES from Huang et al. [26], which is probably most accurate available. The fit of this PES started from a series of CCSD(T) *ab initio* calculations with scaled averaged coupled-pair function (ACPF) corrections, which also accounts for relativistic effects. No non-Born–Oppenheimer effects were included, resulting in an isotope-independent PES. In addition to this a two-step refinement was performed: first using a subset of HITRAN2008 \(J = 0\)–4 energy levels, second with the use of fully experimental levels for chosen \(J\)’s up to 85. The resultant PES was then rigorously tested against HITRAN2008 and HITRAN2012 as well as against more recent experiments [26,28]. The best fit root-mean-square-deviation (RMSD) with respect to purely experimental energy levels for the final Ames-1 PES for the CO\(_2\) main isotopolog was equal to 0.0156 cm\(^{-1}\) in \(J = 0\)–117 range. Comparison with HITRAN2012 database frequencies gave an average overall shift of −0.0456 cm\(^{-1}\) and a spread (rms) of 0.0712 cm\(^{-1}\), which is consistent with our own calculation based on this PES. The relatively large discrepancy between Ames-296 and HITRAN2012 was the reason to exclude most of HITRAN energy levels from the fitting procedure. It also points to inconsistencies in the current release of the database.

### 2.1.2. *Ab initio* PES

To aid the line sensitivity analysis, we independently constructed a fully *ab initio* CO\(_2\) PES using the energy points used by Polyansky et al. [20] to compute their DMS. MOLPRO [46] multi-reference configuration interaction theory (MRCI) calculations with the aug-cc-pVQZ basis were augmented by relativistic corrections at the one-electron mass–velocity Darwin (MVD1) level. For more details see the supplementary materials in Ref. [20].

### 2.1.4. Ames DMS

The dipole moment surface ‘DMS-N2’ was based on 2531 CCS3(T)/aug-cc-pVQZ dipole vectors [27]. The linear least-squares fits were performed with 30 000 cm\(^{-1}\) energy cutoff and polynomial expansion up to 16-th order with 969 coefficients, which gave a RMSD of 3.2 \(\times 10^{-6}\) a.u. and 8.0 \(\times 10^{-6}\) a.u. for respective dipole vector components. Comparison with recent experiments [27] and CDSD data leads to the general conclusion that the Ames DMS, while reliable, still does not meet requirements for remote sensing accuracy.

### 2.1.5. UCL DMS

Our dipole moment surface was calculated using the finite field method. Both positive and negative electric field vector directions were considered for the \(x\) (perpendicular to molecular long axis) and \(y\) (along molecular long axis) components of the dipole moment, requiring 4 independent runs for each *ab initio* point. Finally the dipole moment was computed as first derivative of electronic energy with respect to a weak uniform external electric field \((3 \times 10^{-4}\) a.u.\); a two-point numerical finite difference approximation was used. Previous research suggests that in general derivative method yields more reliable dipole moments than those obtained from simple expectation value evaluation [47]. Randomly distributed *ab initio* points were then fitted with a polynomial in symmetry adapted bond-lengths and bond angle coordinates. This resulted in an expansion up to fifth order. Points above 15 000 cm\(^{-1}\) were rejected from the fit, leaving 1963 points for the \(x\) component fitted with 17 constants giving a RMSD of 2.25 \(\times 10^{-5}\) a.u.; and 1433 points for the \(y\) component fitted with 19 constants giving RMSD of 1.85 \(\times 10^{-5}\) a.u.

### 2.2. Nuclear motion calculations

Nuclear-motion calculations were performed using the DVR3D suite [44]. Symmetrized Radau coordinates in bisector embedding were applied to represent nuclear degrees of freedom. Rovibrational wavefunctions and energy levels were computed utilizing exact kinetic energy
operator (in the Born–Oppenheimer approximation) with nuclear masses for carbon (11.996709 Da) and oxygen (15.999492 Da).

As a first preliminary step in our procedure basis set parameters were optimized with respect to energy levels convergence using the Ames-1 PES. The final set of parameters for Morse-like basis functions [44,48], describing stretching and bending motions, was considered as \( r_e = 2.95 \) a₀, \( D_e = 0.30 \) \( E_h \), and \( \alpha = 0.0085 \) \( E_h \). These values were chosen in a careful scan of parameter space with convergence speed as a criterion. The contracted DVR basis set associated with Gauss–Legendre quadrature points consisted of 30 radial and 120 angular functions, respectively. The appropriate choice of basis set parameters allowed us to reduce the size of the basis needed to converge energy levels, hence speeding up calculations. The same set of parameters was used for rovibrational energy evaluation with \textit{ab initio} and fitted PESs.

At room temperature the highest initial energy level that can be populated enough to give a transition above 300 cm\(^{-1}\) is \( 42 \) \( J \rightarrow 51 \), \( 42 \rightarrow 50 \) transitions which were not fully characterized by the underlying PES. Such resonance interactions can lead to intensity variations where the wavefunctions do play an important role is in capturing the interaction between different vibrational states. Such resonance interactions can lead to intensity stealing and, particularly for so-called dark states, huge changes in transition intensities. The Lodi–Tennyson methodology was designed to capture accidental resonances which were not fully characterized by the underlying PES. Under these circumstances calculations with different procedures should give markedly different results. Lodi and Tennyson monitored the effects of changes to the PES and fit of the DMS. The procedure does not yield an uncertainty as such, it simply establishes which PES is relatively inaccurate and which are not, in which case the predictions were deemed as unreliable and alternative sources of intensity information was recommended.

In other words, trustworthy lines should be stable under minor PES/DMS modifications. One problem with this strategy is that if the alternate PES (or DMS) differs too much from the best PES then large intensity variations can be found which do not reflect problems with the best calculation. This issue already arose in a study on HDO [49] where the \textit{ab initio} and fitted surfaces showed significant differences. For \( \text{CO}_2 \) our \textit{ab initio} PES is relatively inaccurate and hence far from the high quality Ames-1 fitted PES; it was for this reason we constructed a third PES by performing our own, light-touch fit.

Here we therefore follow the Lodi–Tennyson strategy [33] but constructed and evaluated six linest of
the three different PESs and two different DMSs introduced above. For this purpose three sets of nuclear-motion wavefunctions were produced: the first based on the Ames-1 semi-empirical PES, second based on the UCL \textit{ab initio} PES and the third on our new fitted PES. Those three sets of wavefunctions were combined with the two \textit{ab initio} DMSs described above, to give line intensities. Having six line lists, the next step was to match line-by-line pairs of respective line lists: (Ames PES & Ames DMS, Ames PES & UCL-DMS)=(AA,AU), (UCL-\textit{ab initio} & Ames DMS, UCL-\textit{ab initio} & UCL-DMS)=(UA,UU), (fitted PES & Ames DMS, fitted PES & UCL-DMS)=(FA,FU). This stage was straightforward, yielding almost 100\% match as the line lists being compared differ only in DMS, which does not affect energy levels. The second stage involved matching the Ames-PES based with UCL-PESs based line lists, i.e. (AA, AU) vs. (UA,UU) and (AA,AU) vs. (FA,FU). In both cases line-by-line matching was preceded by matching of energy levels. In the case of Ames vs. UCL we managed to match 90\% of lines stronger than \(10^{-30}\) cm/molecule, while Ames vs. fitted resulted in high 99\% correspondence. This confirms that reducing the 6.2 cm\(^{-1}\) RMSD to 1.4 cm\(^{-1}\) makes a significant difference. Note that since the (AU) line list provides our best estimates of the intensities, there is no benefit in performing a (UA,UU,FA,FU) scatter factor analysis.

For each ‘matched’ line, the ratio of strongest to weakest transition intensity was calculated, yielding a scatter factor \(\rho\). Fig. 1 shows scatter factors statistics for the two sets of interest. We can clearly see that (AA,AU,UA, UU) set has more uniform and compact distribution of \(\rho\). However statistics for the \textit{ab initio} UCL PES are based on an incomplete match, leaving around 10\% of unmatched lines with an unknown scatter factor. On the other hand, cumulative distribution function for (AA,AU,FA,FU) set reaches plateau at higher percentage of all lines, indicating the advantage of fitted PES over UCL-\textit{ab initio}.

This leaves the problem of the choice of a critical value for the scatter factor. Lodi and Tennyson chose the arbitrary value of \(\rho = 2\). Here we used the scatter factor statistics to help inform our choice for this number. Fig. 1 suggests that \(\rho = 2.5\) is a reasonable value for this descriptor. Our more detailed analysis of individual bands, given below, suggests that this is indeed an appropriate value.

Detailed band-by-band comparisons revealed another feature of (AA,AU,UA, UU) set: for a number of bands for which the AU intensities gave excellent agreement with the measurements for all transitions, but an arbitrary proportion of the transitions was identified as being unstable. These false negatives are unhelpful and lead to the risk of good results being rejected. For the (AA,AU,FA, FU) set we found that provided the scatter factor was taken to be high enough, \(\rho > 2.5\), this problem was not encountered. Hence we decided to use fitted PES as a working set for further analysis.

For \(J \geq 25\) it is quite common to have almost degenerate transitions, that is transitions from exactly the same lower energy level to upper energy levels with same \(J\) and \(e/f\) symmetries, and as close as \(0.1\) cm\(^{-1}\). Therefore sometimes even the energetically best match is not correct which leads to very inflated scatter factors. In this case, manual matching based on intensity considerations, eliminates this problem for stronger bands \((1 > 10^{-26}\) cm/molecule\) and leaves only true \(J\)-localized instabilities. Due to this issue with almost degenerate transitions, we should note that the numerical values of \(\rho\) for unstable lines may in some cases be caused by mis-assignments which leaked through our tests. In particular, such a situation can occur when almost degenerate transitions have similar line intensities.

There are two main cases when \textit{ab initio} based intensities may lose their reliability: energy levels crossing and intensity borrowing by a weak band from a very strong band via resonance interactions.

The latter is just the case for \(1110i-00001\) \((i = 1, 2, 3)\) bands. They borrow intensities from very strong asymmetric stretching fundamental via second order Coriolis interaction. This appears as a sharp peak around \(2000\) cm\(^{-1}\) (upper energy level) as depicted in Fig. 2. In this case, reproducing the line intensities with high accuracy requires very precise wavefunctions. We describe these lines as being associated with a \(J\)-localized instability.

![Fig. 1. Scatter factor, \(\rho\), statistics for two sets of PES-DMS combination. Inset: cumulative distribution function. See text for further details.](image1)

![Fig. 2. Scatter factor as a function of lower energy level for the 11102-00001 band. The line denotes critical value of the scatter factor (\(\rho = 2.5\)).](image2)
Together with bands for which scatter factor has peaks concentrated around certain energetic region we also encountered entire vibrational bands with $\rho > 2.5$, which we shall name as ‘sensitive’.

The 40011-00001 and 40012-00001 bands are good examples of combination of these two effects. Firstly both bands have their upper energy levels around 8000 cm$^{-1}$ and the whole bands are uniformly unstable. Moreover we observe peaking of the scatter factor around $J = 76$ ($E_i = 2278$ cm$^{-1}$), which we attribute to a strongly mixed lower energy level involved in transition. The 40012-00001 band is much stronger, therefore the $J = 76$ transitions are still above our intensity threshold, and we find a maximum in the scatter factor.

The Lodi–Tennyson approach was based on the idea of occasional, accidental resonances. However it is well-known that CO$_2$ has a series of systematic, Fermi-resonances. We found that some of the bands gave large $\rho$ values for all transitions. There are two possible causes for this: incomplete representation of the resonance interaction in the PESs used or that the compared PESs differ significantly for this band. Comparisons also suggested that some of the predicted intensities for these bands may not be as reliable as one would expect for the accurate UCL DMS. We therefore adjusted our strategy and introduced an intermediate category of lines between stable and unstable for which the uncertainty of our intensity predictions was increased.

2.4. Line positions

Lodi and Tennyson’s water line list was based on the use of experimental energy levels [50–52] based on the MARVEL procedure [53,54]. For CO$_2$ an effective Hamiltonian model was able to reproduce all published observed line positions with accuracy compatible with measurement uncertainties [23]. Specifically, the fitted model of $H_{eff}$ was able to reproduce 44 917 observed line positions of $^{12}$C$^{16}$O$_2$ having measurement uncertainties in the $3.0 \times 10^{-9}$ [55] to 0.02 cm$^{-1}$ range with a dimensionless standard deviation 2.0. This means that, on average, the obs–calc residuals exceed the measured uncertainties by only a factor of two. This makes these calculated line positions appropriate for our new $^{12}$C$^{16}$O$_2$ line list.

3. Results

Our final line list given in the supplementary information, includes the $\rho$ parameter, determined from (AA, AU, FA, FU), as one of the fields; $\rho$ is set to $–1.0$ whenever it could not be extracted. For the most intense bands this automatic procedure was followed by manual matching and double-check, see Table 1.

3.1. Scatter factors

In order to appreciate the landscape of scatter factor distributions, it is instructive to introduce scatter factor maps as a function of lower and upper energy level. Fig. 3 shows a map where color codes represent values of the scatter factor for a given transition. The fundamental bands are easily identified as straight lines originating at 0 cm$^{-1}$ lower energy. The lowest hot bands originate at around 668 cm$^{-1}$, complicating the whole picture. A general conclusion from Fig. 3 is that the higher energy of a level involved in a transition, the higher tendency for the transition to be unstable. The color coding in the figure divides scatter factor space into 3 regions of increasing instability, marked blue, orange and red, respectively. The blue region is considered to be stable and corresponding intensities are reliable. The orange region is intermediate between stable and unstable, hence transitions marked orange need careful consideration. The red region contains highly unstable lines whose computed line intensities should not be trusted. There are a few super-unstable transitions ($\rho > 10$) which are not shown on the plots; these lines are usually associated with a strong resonance interaction with some other energetically-close level. Analysis of scatter factors for individual bands can yield insight. By zooming in an energetic region of interest it is straightforward to pick up entirely unstable bands or single transitions which happen to fall into resonance. Altogether we find 5400 transitions we classify as unstable.

For example, as can be seen from Fig. 4 (which considers only lines stronger than $10^{-25}$ cm/molecule) while majority of bands have completely uniform scatter factors below the critical value of 2.5, there are entire bands (marked orange) systematically shifted by change of the underlying PES by a factor of more than 2.5. Those bands involving vibrational states which appear to be influenced by strong resonance interactions are called ‘sensitive’ bands below. A completely different behavior may be observed for example for the 11101, 11102 and 11103 series of bands (indicated with arrows). Here a fairly uniform scatter factor is disturbed by $J$-localized peak. Fig. 2 illustrates such behavior, which is explained by inter-band intensity borrowing via rotational–vibrational (Coriolis) interaction terms in molecular Hamiltonian. A summary of stability analysis for selected bands is given in Table 1. 108 out of 116 bands stronger than $10^{-25}$ cm/molecule are stable. Bands involving bending excitations are also very stable. For some bands, such as 32203-03301 and 42201-03301 $J$-localized instabilities appear only weakly, generating peaks which do not exceed the critical value.

3.2. Comparison with high-accuracy measurements

Polyansky et al. [20] showed that transition intensities based on the (A,U) model gave excellent agreement with new, high accuracy measurements of the 30013-00001 band in the 6200–6258 cm$^{-1}$ reported in the same paper. Polyansky et al. also compared their predictions with the high accuracy measurements of Casa et al. [17,18] and Wuebbeler et al. [19] for the 20012-00001 band. While their results were in excellent agreement with the single line intensity measured by Wuebbeler et al., they suggested that the results of Casa et al. were significantly less accurate than claimed. This assertion has since been confirmed by new high-accuracy measurements performed by Brunzendorf et al. [22] which show almost no systematic
shift and average deviation of only 0.35% with respect to our (and Polyansky et al.’s) predictions.

Two more lines in \( \nu_1 + \nu_3 \) band (P34,P36) were measured by Pogany et al. [56] with reported 1.1 and 1.3% uncertainty. The corresponding UCL intensities deviate by 2.0% and 2.5% respectively. Nevertheless these are on average 1% closer to experimental values than the intensities obtained from either Ames-296 or CDSD-296.

Very recently Devi et al. [21] performed a detailed study at 1.6 \( \mu \)m. The strongest band in this region is 30013-00001.

A comparison between their measured line intensities and our predictions is given in Fig. 5.

From Fig. 5 it is evident that a majority of our line intensities lie within 1% of the new measurements. The bow-like behavior seen particularly at high J’s here, and in other comparisons discussed below, is unlikely to be due to our calculations. Instead we expect it is an artifact associated with Herman–Wallis factors used to parameterize the experimental data, which tend to overestimate line intensities for high-J transitions. If this is so,
then it is likely that our results match Devi et al.’s at the sub-percent level.

Low J HITRAN2012 line intensities for the 30013-00001 band originate from the JPL OCO linelist of Toth et al. [38] and lie on average 0.5% above the value of Devi et al. These lines are marked with a 7 as the HITRAN intensity uncertainty code which means that these line intensities are accurate within 2%. The high-J line intensities (J > 45) are all calculated values based on a fitted effective dipole moment model [39]. They have 3 as the HITRAN intensity uncertainty code but may have errors in the intensity greater than 20%. The clearly visible jump in HITRAN points in Fig. 5 is located at the meeting point of the two data sources.

3.3. Comparison with other line lists

3.3.1. Ames-296

Huang et al. [28] published infrared line lists for 12 stable and 1 radioactive isotopologs of CO2. These linelists were calculated with Ames-1 PES [26] and DMS-N2 [27], or (A,A) in our notation above. We generated from their data a 12C16O2 line list for its natural abundance, T = 296 K and with an intensity cutoff of 10−30 cm/molecule, which we refer to as Ames-296. Ames-296 contains 162 558 lines in the 0–8000 cm−1 range. To facilitate comparison with other line lists we performed a spectroscopic assignment of this line list. As a first step, for the sake of consistency, it was necessary to compare energy levels from original Ames-296 linelist with our DVR3D recalculation. Energy levels up to 6000 cm−1 gave a RMSD of 0.05 cm−1 and 0.06 cm−1 up to 10 000 cm−1. This is slightly more than we would have expected on the basis of previous comparisons [57] and appears to be due a slightly non-optimal choice integration grids in Huang et al.’s calculations (Huang and Lee, 2015, private communication).

3.3.2. CDSD-296

The effective operator approach enables one to reproduce all published observed positions and intensities with accuracies compatible with measurement uncertainties. Based on fitted H eff and D eff models Tashkun et al. [23] created a high resolution spectroscopic databank CDSD-296 aimed at atmospheric applications. The databank contains the calculated line parameters (positions, intensities, air-and self-broadened half-widths, coefficients of temperature dependence of air-broadened half-widths and air pressure-induced lineshifts) of the twelve stable isotopic species of CO2. The reference temperature is 296 K and the intensity cutoff is 10−30 cm/molecule.

Fig. 6 compares Ames and UCL line intensities with the semi-empirical CDSD-296 results. For the sake of clarity only strong bands with intensities greater than 10−23 cm/molecule are plotted.

For the strongest bands UCL linelist agrees much more closely with CDSD-296 than Ames-296 does. The only real exception to this are the 00031-00001 and 01 131-01 101 bands. For this family of bands, whose intensity derives from the same dipole moment derivative, the deviations from Ames-296 are three times less than UCL ones. We identified 3170 transitions belonging to this family.
3.4. HITRAN2012

HITRAN2012 [36] contains 160,292 $^{12}$C$^{16}$O$_2$ lines in the 0–8000 cm$^{-1}$ region. A matching procedure for our Ames-1 PES energy levels to those originally extracted from HITRAN2012 database was conducted by imposing rigorous restrictions on rotational quantum numbers and rotationless parities as well as a 0.3 cm$^{-1}$ tolerance for energy difference. This scheme matched all 16,777 unique energy levels present in HITRAN2012 covering $J$ values from 0 to 129 with RMSD of 0.07 cm$^{-1}$. The largest deviation found between two levels was roughly 0.2 cm$^{-1}$.

The next step was to match transition lines between HITRAN2012 and UCL linelists. The procedure relied on a simple algorithm, where corresponding lines were matched using already matched energy levels list. As a result all 160,292 lines up to 8000 cm$^{-1}$ were matched with a RMSD of 0.08 cm$^{-1}$ in line positions.

There are two main sources of HITRAN2012 data for CO$_2$ main isotopolog: a small set of 605 lines in 4800–6989 cm$^{-1}$ range originating from experiment (JPL OCO line list) by Toth et al. [38] and the majority of transitions from a previous version of CDSD. In general, data from latest version of CDSD-296 are very close to line positions and intensities given in HITRAN2012.

The estimated uncertainties for all CDSD intensities is given as 20% or worse in HITRAN (uncertainty code 3). On the other hand, Toth et al.’s intensities are supposed be accurate to better than 2% (uncertainty code 7) or 5% (code 6). This reveals two issues with the current version of HITRAN:

(a) The stated uncertainty estimate of all current entries are insufficiently accurate for remote sensing applications. Our previous study [20] already showed that for a number of important bands the actual accuracy of the intensities in HITRAN is much higher than suggested by their estimated uncertainties.

(b) Line intensity accuracies are not uniform throughout the spectral region. Our experience from studies on several molecules is that the ratio of observed to variational line intensities should be roughly constant for a given band unless there is an isolated resonance (see below). For CO$_2$, comparing HITRAN intensities with our predictions we would expect the same, but detailed analysis (cf. Fig. 11), that such jumps in accuracy cause artificial patterns in line intensities within a single vibrational band.

All HITRAN2012 entries taken from a pre-release version of CDSD have been tagged with uncertainty code 3 (20% or worse). However, this number does not reflect actual uncertainties of the intensities. Most of the HITRAN intensities have the uncertainties much better than 20%. More detailed information about the actual uncertainties can be found in the official release of CDSD [58]. The reader should use this work in order to get a realistic information about the uncertainties of the line parameters.

Fig. 7 gives a general overview of the two linelists. Overall the agreement is excellent with more than 98% of entries common between both lists and very similar intensities. However, there is some incomplete coverage by HITRAN2012 with several artificial windows, especially for low intensity transitions. There are also a few missing medium-intensity transitions around 400 cm$^{-1}$ and 1600 cm$^{-1}$.

Intensities of all assigned UCL lines relative to HITRAN2012 are depicted in Fig. 8. As expected discrepancies between the two linelists grow as lines get weaker, which results in a funnel-like shape in the plot which characteristic of such comparisons (e.g. [59]). The dependence of the UCL lines on the scatter factors are also shown; as could be anticipated stable lines predominate at higher intensities.

It is instructive to divide HITRAN2012 data into subsets of a given intensity accuracy code. Each of those sets can be then compared to our results separately to provide an estimate for compatibility of two linelists at different levels of accuracy. To achieve that we plotted HITRAN intensities with the accuracy code found for CO$_2$ which is 7 (2% or better uncertainty) against the UCL ones. This set of lines encompass the important 20011, 20012, 20013, 30011, 30012, 30013 and 30014 bands as well as the
asymmetric stretching second overtone 00031. All bands except intermediate 30011 band are stable. Comparisons with high accuracy measurements above have already shown that our results for the 30013 band are accurate to about 1% or better Fig. 9.

Again one can see characteristic bow-like structures corresponding to particular rotational transitions within a vibrational band, with the peak of an arc refers to most intense, low J transition. We suggest that these structures are artifacts which originate from the semi-empirical treatment of the intensities.

A similar situation occurs for bands with HITRAN uncertainty code 6, see Fig. 10; here very good agreement is spoiled by 01131-01101 band.

Fig. 9. Comparison of HITRAN2012 most accurate intensities and UCL line intensities. The dashed line indicates the stated HITRAN uncertainty, i.e. 2%. Arrows label vibrational bands, which all start from the ground 00001 state.

Fig. 10. Comparison of HITRAN2012 medium-accuracy intensities and UCL line intensities. The dashed line indicates the stated HITRAN uncertainty, i.e. 5%. Arrows label vibrational bands.

Fig. 11 gives an intensity comparison for the 20012 band. HITRAN 2012 used two separate data sources for this band. This is clearly visible which means, despite the overall good agreement with present results, there is an abrupt change in intensity trends at $J = 64$. This is the point where the experimental data finished and the database had to rely on results from a the CDSD effective Hamiltonian calculations.

3.5. A HITRAN-style line list

The final UCL-296 line list, given as supplementary data, contains 162 468 line positions, intensities scaled by natural abundance (0.98420), quantum numbers and scatter factors taken from our computation.

Our final recommended line list for $^{12}$C$^{16}$O$_2$ is also given in supplementary data. This list contains 162 260 lines in HITRAN format with intensities scaled by natural abundance and uniformly cut off at $10^{-30}$ cm/molecule. Vibrational quantum numbers were taken from CDSD and
cross-checked with HITRAN2012 assignments. Line positions were transferred from CDSD-296 with appropriate uncertainties.

The majority of line intensities (151 602) were taken from our “AD” calculations; we assign HITRAN uncertainty code 8 for stable bands with at least one transition stronger than $10^{-25}$ cm/molecule and 7 for stable bands weaker than $10^{-25}$ cm/molecule together with 8647 lines from intermediate bands.

Whenever our line intensity turned out to be unreliable (i.e. was either unstable and no additional tests confirmed its high accuracy or belonged to $3σ$ family of bands) it was replaced by CDSD-296 value. This was the case for 10 080 (6%) lines.

4. Conclusion

We present a new mixed ab initio-empirical linelist providing reliable intensities for $^{12}$C$^{16}$O$_2$ up to 8000 cm$^{-1}$. We believe this line list is more complete and the intensities more accurate than in HITRAN2012 [36]. A detailed analysis shows that our line intensities generally are accurate at the sub-percent level when compared to recent, high-accuracy measurements, consequently validating our approach; furthermore we find that intensity uncertainties stated in HITRAN2012 are probably too conservative. We believe these improved intensities should assist in improving CO$_2$ monitoring in remote atmospheric sensing studies, and in other applications. Furthermore this new line lists fills in the small gaps in the HITRAN2012 list. Of course for use in atmospheric conditions this line list needs to be supplemented by both line profile parameters and consideration of line-mixing [60].

One issue that we should raise concerns perpendicular transitions (those with $Δε = ±1$ and $±2$). The majority of the perpendicular bands borrow intensity from the considerably stronger parallel ($Δε = 0$) bands via Coriolis resonance or anharmonic plus $ε$-type interactions. To describe this process it is necessary to have very precise wavefunctions. So far, the very high accuracy of the line intensity calculations presented here is confirmed experimentally only for parallel bands. All weaker bands have been given a lower accuracy rating in our line list; nonetheless it would be very helpful to have some high accuracy experimental measurements of perpendicular bands to help to independently validate our results.

Future work will focus on two aspects of the problem. First, it is apparent that our ab initio dipole moment surface is less accurate for transitions involving changes of 3 or more quanta in $ν_3$. This problem will be the subject of future theoretical investigation which will also aim to extend our model to frequencies higher than 8000 cm$^{-1}$. Second, a major advantage of our methodology is that theoretical calculations can be used to give intensities for all isotopologs of CO$_2$ with essentially the same accuracy as the $^{16}$O$^{12}$C$^{16}$O results presented here. This should be particularly used in providing accurate intensities for trace species such as $^{18}$O$^{12}$C$^{16}$O, which are important for monitoring purposes [61]. Line lists for isotopically substituted CO$_2$ will be published in the near future.

Acknowledgments

This work is supported by the UK Natural Environment Research Council (NERC) through grant NE/J010316, the ERC under the Advanced Investigator Project 267219 and the Russian Fund for Fundamental Science. We would like to thank Jens Brunzendorf and V. Malathy Devi for sharing their results with us prior to publication.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.jms.2015.12.022.

References


Acknowledgments

This work is supported by the UK Natural Environment Research Council (NERC) through grant NE/J010316, the ERC under the Advanced Investigator Project 267219 and the Russian Fund for Fundamental Science. We would like to thank Jens Brunzendorf and V. Malathy Devi for sharing their results with us prior to publication.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.jms.2015.12.022.

References

[1] Butz A, Gue...


