The Bernese Atmospheric Multiple Catalog Access Tool (BEAMCAT): a tool for users of popular spectral line catalogs

Dietrich G. Feist [1]

Institute of Applied Physics, University of Bern, Switzerland
Email: dietrich.feist@mw.iap.unibe.ch

Abstract

Users of spectroscopic data bases in the microwave region quickly realize that each existing spectral line catalog provide only part of the information that they would like to have. As a workaround for this problem, several merged spectral line data bases have been created by different groups. However, these merged data bases are usually very specific for a certain application and are difficult to maintain. The BEAMCAT data base takes a totally new approach that makes it possible to generate merged spectral line catalogs from any number of source catalogs in multiple user-defined formats. The current version of BEAMCAT contains the complete JPL and HITRAN catalog. Other catalogs like GEISA will soon be included, too.

As a first application of the BEAMCAT data base, the author conducted a thorough intercomparison of spectral parameters for all the transitions that the JPL catalog and HITRAN have in common. The intercomparison shows that the spectral parameters in the catalogs are by no means identical. While the difference in center frequency is usually small, the differences in line intensity reach from almost exact match to discrepancies of several orders of magnitude. While it cannot be ruled out that some of the lines were matched incorrectly, this intercomparison might be helpful to identify problems with the original catalogs.

Key words: Spectral line catalogs, JPL catalog, HITRAN

---

1 Introduction

Radiative transfer models for the atmosphere need information about the spectroscopic properties of atmospheric constituents. The molecules of many atmospheric

---

[1] Supported by the Swiss National Science Foundation under grant 2000-063793.00

trace gases show rotational transitions in the millimeter and submillimeter range as well as vibrational transitions at infrared wavelengths. Information about these transitions is collected in large spectroscopic databases like the JPL catalog [1], HITRAN [2], and GEISA [3]. Among these databases, the JPL catalog is the only one that is specialized on the millimeter and submillimeter domain. The others contain mostly transitions in the IR region.

The problem with the JPL catalog is that it does not contain all the information that is necessary to model atmospheric spectra. Important information, like line-broadening parameters, is missing completely. This data has to be supplied from other sources. There have been other attempts to create merged databases [4–6] but none of them satisfies all the needs of the microwave community. Several databasess, like SAO [4] or the MASTER-SOPRANO database [5], were created for certain instruments and cover only a limited frequency range. Besides, many databases use the same format as HITRAN. This format is sufficient for IR applications but provides a frequency resolution which is undesirably low in the millimeter wave range.

At the Institute of Applied Physics of the University of Bern, Switzerland, another merged spectral line catalog called BEAMCAT was created as a supplement for the author's spectroscopic simulation code BEAM (Bernese Atmospheric Model). Both BEAM and BEAMCAT were part of the author's PhD work [7]. The first version of BEAMCAT was similiar to previous merged spectral line data bases. After laying dormant for several years, a new version of BEAMCAT was produced in 2002. The new version follows a very different approach than other merged catalogs. It aims at providing a universal tool for the users of all spectral line data bases.

2 Meta database vs. merged spectral line catalogs

Previous attempts to combine spectral line data from several sources [4–6], including our own JPLHITRAN catalog, tried to produce merged databases. While this appears to be a promising approach at first sight, it is far from an ideal solution for a number of reasons. The merged data bases are usually created in a complicated one-time process. Since the number of spectral lines that have to be merged is huge, the merging process has to be completely automatic. Three severe problems arise:

Errors: The merging process will certainly introduce some errors and mismatches. These errors could be corrected manually but the corrections would be lost when the merging process has to be repeated.

Updates: Updates become a difficult task as the merging process would have to be repeated with new data.

Lost information: The existing merged catalogs throw away information during

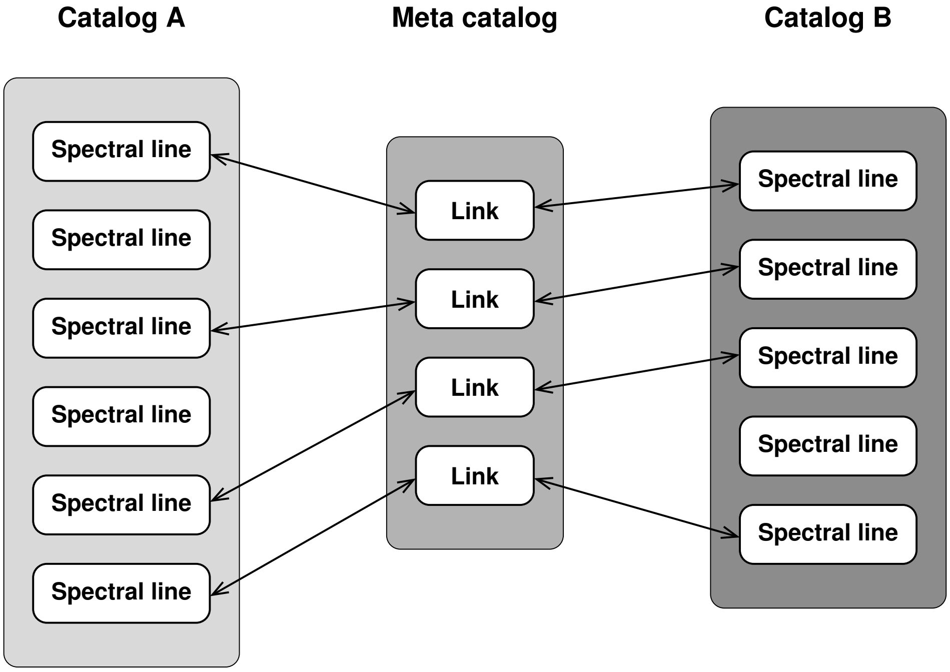


Fig. 1. Concept of the meta catalog. Unlike a merged spectral line catalog, the meta catalog provides links between identical quantum-mechanical transitions listed in different catalogs (A and B in this case). The meta catalog itself does not contain any spectral parameters. It rather provides an index to where the parameters for a certain spectral line can be found in each of the catalogs A and B. The concept works with any number of catalogs.

the merging process by prefering parameters from one catalog over the corresponding parameters from the other catalog. This lost information can not be recovered from the merged catalog any more.

An alternative to a merged catalog is a meta catalog. A meta catalog does not contain any spectral parameters itself. Instead, it contains links between corresponding spectral lines in different catalogs. The concept is illustrated in Figure 2. By following such a link, one can quickly look up spectral line data for the same transition in several catalogs.

With the help of a meta catalog merged catalogs can be produced on the fly again and again. Errors in the meta catalog – for example mismatched or missed spectral lines – can be corrected or added any time. Updates are no problem since only the spectral parameters are updated, the meta information stays the same. No information is lost since the source catalogs contain the complete information. The concept is not limited to two source catalogs in any way. The meta catalog can store links between any number of spectral line catalogs.

3 Database design

3.1 File-based approach vs. SQL data base server approach

All the spectral line catalogs mentioned above store their data in large fixed-width column text files. These files are usually broken into separate files by molecule or isotope type that are sorted by line frequency. This approach has its roots in earlier days when computer memory was very limited and data base systems were unavailable. It is certainly not state-of-the-art any more. In fact, it has serious drawbacks:

Sorting and searching: While one can quickly locate a certain spectral transition if the molecule/isotope and frequency are known, it is a rather difficult to locate it by some other parameter, for example by its quantum numbers. It is also very difficult to sort the data base in any other way. For example, the seemingly simple task of producing a frequency-sorted list of all transitions in a specified frequency region becomes a huge effort.

Restrictive format: The traditional catalogs use very stringent formats with fixed-width data fields in fixed-length lines. In addition to that, the same format is used for all transitions. This restrictive format becomes a problem when the physics does not fit into the tight envelope of the data format. For example, the JPL catalog has to use an awkward coding scheme to fit positive three-digit or negative two-digit quantum numbers into the corresponding two-character-wide data fields. For HITRAN, the data format limits the frequency resolution of spectral lines to roughly 30 kHz, even though many transitions in the micrwave region are known with much higher accuracy.

Additional parameters: The fixed line format makes it next to impossible to include additional line parameters without breaking the format. Since a new format would render most of the existing software for the catalogs obsolete, the catalog maintainers tend to avoid this step for as long as possible, even if there is genuine need for these changes.

BEAMCAT avoids all the above mentioned problems by storing the spectral line information along with other information in a modern SQL-based data base management system (DBMS). The BEAMCAT data base server runs on MySQL [8], a widely used DBMS provided by the free software community. MySQL is well known for its performance and can easily handle data bases that are much bigger than all the existing spectral line catalogs combined. The client software that is needed to access the data base server as well as application programming interfaces (API) for many popular scripting and programming languages are also available as free software. They exist for most platforms and operating systems. An SQL data base server has many advantages compared to the file-based approach:

- Sorting and searching is possible for any parameter or even combinations of parameters. By setting up index columns for parameters that are often used for sorting and searching, the search process can be sped up by several orders of magnitude.

- There is no restriction by data format for an SQL data base server. The output format is determined by the user on each request. It is even possible to request arithmetic expressions consisting of several parameters in a request. For example, one can retrieve a parameter and change its units on-the-fly with an appropriate request.
- Columns in the data base tables can be added, removed, and changed easily. As long as the names of existing columns do not change, it is not even necessary to change software that accesses the SQL server when new parameters are added.

3.2 BEAMCAT structure

An SQL data base consists of tables that are made up of rows and columns. At the time of this writing, BEAMCAT contains three tables that are important to the user: one table for HITRAN, one table for the JPL catalog and one table for the meta catalog. Other tables exist for internal use only and are not relevant for the general user.

3.2.1 The HITRAN table

This SQL table contains the complete HITRAN 2000 spectral line catalog. Its structure is described in detail in Table 1. It has one column for each spectral parameter in HITRAN and one row for each spectral line. The columns carry the same name as the corresponding HITRAN parameters as far as possible. The catalog has two additional columns: HITRAN_ID and TIMESTAMP. HITRAN_ID is a unique identifier for each spectral line. TIMESTAMP provides the date and time when a spectral line was last updated. The initial value of TIMESTAMP was set to the date and time of the file on the HITRAN FTP site that contained the corresponding spectral line.

3.2.2 The JPL table

With a few execptions, this SQL table contains almost the complete JPL spectral line catalog. The transitions for the JPL tag numbers 25002, 27004, 38002, 39001, 39002, 39003, and 53001 are not yet included because the corresponding JPL files cpntained format errors that could not be resolved. However, these JPL tags are not critical since they have no counterpart in HITRAN. The structure of the JPL table is described in detail in Table 2. It has one column for each spectral parameter in JPL and one row for each spectral line. The columns carry the same name as the corresponding JPL parameters as far as possible. The catalog has two additional

Table 1
Structure of the HITRAN table in BEAMCAT: the first two columns provide the SQL column name and data type, the third column lists the corresponding field in HITRAN.

SQL column	Data type	HITRAN	Description
HITRAN_ID	mediumint(7)	—	Unique index into the HITRAN table
Mol	tinyint(2)	Mol	Molecule number
Iso	tinyint(1)	Iso	Isotope number
nu	double(12,6)	$\nu_{\eta\eta'}$	Frequency in cm^{-1}
S	double	$S_{\eta\eta'}$	Intensity in $\mathrm{cm}^{-1}/(\mathrm{molecule}\cdot\mathrm{cm}^{-2})$ at 296 K
R	double	$\mathcal{R}_{\eta\eta'}$	Weighted transition moment-sqared in Debye^2
gamma_air	float(6,4)	γ_{air}	Air-broadened halfwidth (HWHM) in $\mathrm{cm}^{-1}/\mathrm{atm}$ at 296 K
gamma_self	float(6,4)	γ_{self}	Self-broadened halfwidth (HWHM) in $\mathrm{cm}^{-1}/\mathrm{atm}$ at 296 K
E_low	double(10,4)	E''	Lower state energy in cm^{-1}
n	float(4,2)	n	Coefficient of temperature dependence of air-broadened halfwidth
delta	float(9,6)	δ	Air broadened pressure shift of line transition in cm^{-1}/atm at 296 K
iv_up	smallint(3)	iv'	Upper state global quanta index
iv_low	smallint(3)	iv''	Lower state global quanta index
q_up	char(9)	q'	Upper state local quanta
q_low	char(9)	q''	Lower state local quanta
ierr_freq	tinyint(1)	ierr[1]	Accuracy index for frequency
ierr_int	tinyint(1)	ierr[2]	Accuracy index for intensity
ierr_hw	tinyint(1)	ierr[3]	Accuracy index for air-broadened halfwidth
iref_freq	tinyint(2)	iref[1]	Reference index for frequency
iref_int	tinyint(2)	iref[2]	Reference index for intensity
iref_hw	tinyint(2)	iref[3]	Reference index for halfwidth
TIMESTAMP	timestamp(8)	—	Time and date of last update

Table 2
Structure of the JPL table in BEAMCAT: the first two columns provide the SQL column name and data type, the third column lists the corresponding field in the JPL catalog.

SQL column	Data type	JPL	Description
JPL_ID	mediumint(7)	—	Unique index into the JPL table
FREQ	double(13,4)	FREQ	Center frequency in MHz
ERR	double(8,4)	ERR	Frequency error in MHz
LGINT	double(8,4)	LGINT	$\log_{10}$ of integrated intensity in nm^{2}MHz at 300 K
DR	tinyint(2)	DR	Degrees of freedom in the rotational partition function
ELO	double(10,4)	ELO	Lower state energy in cm^{-1} relative to the ground state
GUP	smallint(3)	GUP	Upper state degeneracy
TAGSIGN	tinyint(4)	sign(TAG)	Sign of JPL TAG
TAG	mediumint(7)	abs(TAG)	Absolute value of JPL TAG
Q	tinyint(2)	Q	Type of molecule
H	tinyint(1)	H	Coding of half-integer quantum numbers
NQN	tinyint(1)	NQN	Number of quantum numbers for each state
QN_up1	smallint(3)	QN$'$[1]	Upper state quantum number 1
QN_up2	smallint(3)	QN$'$[2]	Upper state quantum number 2
QN_up3	smallint(3)	QN$'$[3]	Upper state quantum number 3
QN_up4	smallint(3)	QN$'$[4]	Upper state quantum number 4
QN_up5	smallint(3)	QN$'$[5]	Upper state quantum number 5
QN_up6	smallint(3)	QN$'$[6]	Upper state quantum number 6
QN_low1	smallint(3)	QN$''$[1]	Lower state quantum number 1
QN_low2	smallint(3)	QN$''$[2]	Lower state quantum number 2
QN_low3	smallint(3)	QN$''$[3]	Lower state quantum number 3
QN_low4	smallint(3)	QN$''$[4]	Lower state quantum number 4
QN_low5	smallint(3)	QN$''$[5]	Lower state quantum number 5
QN_low6	smallint(3)	QN$''$[6]	Lower state quantum number 6
TIMESTAMP	timestamp(8)	—	Time and date of last update

Table 3
Structure of the meta_lines table in BEAMCAT: the table stores the values of HITRAN_ID and JPL_ID for corresponding spectral lines. The columns meta_ID and TIMESTAMP have similiar function as in Tables 1 and 2. The column rule is only used for debugging.

SQL column	Data type	Description
meta_ID	mediumint(9)	Unique index into the meta_lines table
rule	smallint(6)	Number of rule that was used to identify this match
HITRAN_ID	mediumint(9)	Value of the HITRAN_ID field in the HITRAN table
JPL_ID	mediumint(9)	Value of the JPL_ID field in the JPL table
TIMESTAMP	timestamp(14)	Date and time of last update

columns: JPL_ID and TIMESTAMP. JPL_ID is a unique identifier for each spectral line. TIMESTAMP provides the date and time when a spectral line was last updated. The initial value of TIMESTAMP was set to the date and time of the file on the JPL FTP site that contained the corresponding spectral line.

Note that there is a subtle difference between the TAG definition in the original JPL catalog and the JPL table in BEAMCAT. In the JPL catalog tag numbers are listed as negative numbers to indicate that a spectral line has been measured in the laboratory [1]. However, this definition makes sorting or searching by tag numbers very inefficient. Therefore, the BEAMCAT table column TAG contains the absolute value of the JPL tag and the TAGSIGN column contains the sign of TAG so that $\mathrm{TAG(JPL)} = \mathrm{TAGSIGN} \times \mathrm{TAG}$.

3.2.3 The meta_lines table

The meta_lines table contains the information that is necessary to look up corresponding spectral lines in several catalogs. If a spectral line is listed in both the JPL catalog and HITRAN its HITRAN_ID and JPL_ID values are stored in the same row of the meta_lines table. This simple scheme allows the MySQL server to produce merged tables with spectral line data from different catalogs on the fly. The scheme can easily be extended to more than two spectral line catalogs.

4 Semi-automatic identification of HITRAN and JPL catalog lines

4.1 Approaches and problems with automatic identification

The first step in identifying spectral lines in several catalogs is to map the species, that is molecules and isotopes, from one catalog to the other. The first BEAMCAT report [9] contains information on how this was done. This information is contained in more than 370 rules in the merge_rules table in the BEAMCAT data base. The meta_lines table contains an identifier for the rule that was used to match each transition. This way, matching errors that are caused by a mistake in the species mapping can always be corrected later.

The second – and much bigger – problem is the identification of common transitions for a single species in both catalogs. There are basically two approaches towards that goal: the naive way of comparing line center frequencies and the more formal way of comparing quantum numbers. Unfortunately, both ways fail in practice. Comparing by frequency alone fails for several reasons:

- The frequencies in the catalogs come from different sources
- The HITRAN data format limits the frequency resolution to about 30 kHz compared to 100 Hz in the JPL catalog
- Lines with fine structure may be listed individually in one catalog and as a single averaged line in the other

Comparing by quantum numbers alone works well for some species but fails for others because of the following reasons:

- Quantum numbers are incomplete or missing
- The quantum numbers in the catalog are not sufficient to describe a transition unambiguously. In these cases, the same set of quantum numbers appears more than once for the same species but for different transitions.
- The quantum numbers are not comparable because different sets of quantum numbers are listed in the two catalogs for the same species

4.2 Combined search strategy

Most of the problems mentioned above can be overcome by a combined strategy that relies on frequency as well as quantum numbers. The scheme is illustrated in Figure 2. A line from the JPL catalog is first compared to the HITRAN line with the closest frequency match. Then the quantum numbers are compared. If the quantum numbers match the line will have been identified. If the quantum numbers do not match, neighboring HITRAN lines with lower and higher frequencies will be tested

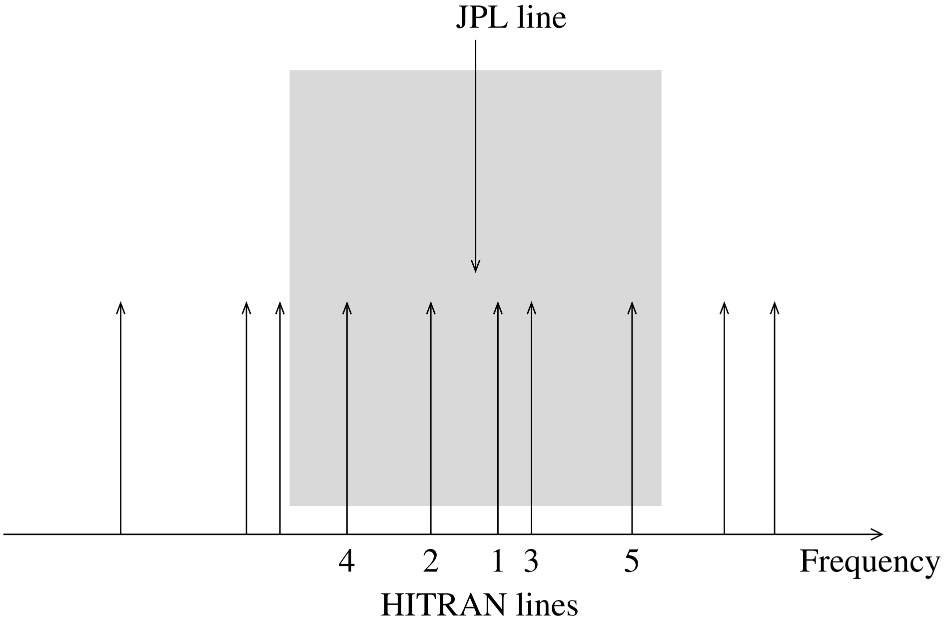


Fig. 2. Illustration of the search strategy that is used to identify a JPL line in HITRAN. Only lines within a given frequency range (shaded region) are tested. The line is first compared to the HITRAN line with the closest frequency match (No. 1). If the quantum numbers do not match, the line will be tested against the other HITRAN lines in ascending order.

alternately until either a match is found or all the lines in a given frequency range have been tested.

4.3 Matching by quantum numbers

The actual comparison of quantum numbers is rather difficult. The quantum number format depends on the type of molecule and its symmetries. Unfortunately, the two catalogs do not always place the same molecule type in the same category and they do not necessarily list the same quantum numbers. In addition to this, the quantum numbers are often incomplete. The question, whether two quantum states found in the two catalogs are the same or not, can in be decided in two ways.

1. **Positive match:** a match of quantum states is assumed if *all* the quantum numbers *agree*
2. **Negative mismatch:** a match of quantum states is assumed if *none* of the quantum numbers *disagree*

There is a subtle but important difference between the two interpretations. The second case is more robust against incomplete or missing quantum numbers. If, in the worst case, quantum numbers are missing completely but the frequency of the two entries agrees, the entry will be matched. Such cases occur frequently. There are whole catalog files with unusable or missing quantum numbers in one catalog but exactly matching frequencies and the same number of listed transitions. In such a case matching by frequency is a practical approach. After all, it is not important how a transition was identified in both catalogs — as long as it is indeed the same transition.

The *negative mismatch* scheme was adopted for the automatic identification of lines in BEAMCAT because it is more robust against incomplete or ambiguous quantum numbers. However, only the initial matching process for BEAMCAT was done this way. The BEAMCAT data base was designed to make it very easy to fix matching errors later. Mismatched transitions can be removed and overlooked matches can be added at any time without ever repeating the automatic matching process.

4.4 Handling different quantum number format cases

The exact quantum number format for a species is defined in the reference publications and documentation for the JPL catalog [1] and HITRAN [2]. In both catalogs, the molecules are divided into groups that reflect their symmetries and degrees of freedom. In many cases, the mapping between groups in the JPL catalog and HITRAN is obvious, for example all asymmetric rotors in the JPL catalog belong to HITRAN group 1. There are, however, some exceptions where the catalogs' classifications disagree, for example there are some symmetric rotors in the JPL catalog that are classified as group 2 in HITRAN. The JPL catalog contains some ClO lines with an undocumented quantum number format.

To compare quantum states, several empirical cases were defined from experience with the two catalogs. Each case defines a distinct combination of JPL/HITRAN quantum number groups. They cover all the species that exist in both catalogs. Table 4 shows which quantum numbers should be compared against each other for molecules of a given geometry.

5 Intercomparison of JPL and HITRAN spectral line data

BEAMCAT provides a simple way to intercompare the spectral parameters of different spectral line catalogs. Since the JPL catalog and HITRAN were the first catalogs in BEAMCAT, this work features a first intercomparison of the data in these two catalogs. The intercomparison focused on the three main parameters frequency, lower state energy and intensity, which are listed in both catalogs.

The JPL catalog data in BEAMCAT was downloaded from the JPL catalog's FTP site. The most recent files on this FTP site were updated on April 22, 2002. The JPL table in BEAMCAT contains 1777426 spectral lines. The HITRAN table in

BEAMCAT was downloaded from the HITRAN FTP site and contains 1080174 spectral lines. The most recent ones were updated on December 22, 2000.

Of the more than one million spectral lines in HITRAN, only about 420000 belong to species that exist in both catalogs and are within the frequency cutoff of the JPL catalog. Roughly 148000 of these lines were identified in both catalogs by the semi-automatic matching algorithm. About 170 of these lines were obvious mismatches with large frequency differences of more than 10%.

For the intercomparison of center frequency it was only necessary to change the units of the HITRAN values from cm^{-1} to MHz. For the lower state energy both catalogs list the same units, so no changes were necessary. Intensity comparisons were more difficult since both catalogs use different reference temperatures: JPL lists the values at T =300 K while HITRAN lists them at T_0 =296 K. The following correction was applied to the HITRAN intensity $S(T_0)$ to get the intensity at the JPL reference temperature T:

$$S(T) = S(T_0)[Q_{rs}(T_0)/Q_{rs}(T)]\frac{e^{-E_l/kT} - e^{-E_u/kT}}{e^{-E_l/kT_0} - e^{-E_u/kT_0}}. \tag{1}$$

In this equation, E_u and E_l are the upper and lower state energy of the transition and $Q_{rs}(T_0)$ and $Q_{rs}(T)$ are the partition function values at the HITRAN and JPL reference temperatures, respectively. The values for Q_{rs} were taken from a table on the HITRAN web page. In addition to that, HITRAN intensities are scaled with isotopic corrections that had to be removed first.

For frequency and lower state energy the following plots show the relative difference in per cent of the values listed in JPL and HITRAN. This relative difference is plotted over frequency. The values in the JPL catalog were used as the reference, so for a parameter P the relative difference $\Delta P = \left|\frac{P_{\mathrm{JPL}} - P_{\mathrm{HITRAN}}}{P_{\mathrm{JPL}}}\right|$.

For intensity, the absolute values are plotted against each other on logarithmic scales. To the right of each plot, a bar graph shows the statistical distribution of the relative differences for each parameter. Unless noted otherwise, the distribution plots are cut off at a relative difference of more than 100%.

5.1 H_2O

The comparison for H_2O is shown in Figure 3. Frequency and lower state energy match very well between both catalogs. The differences appear to be random without any obvious structure. There are two interesting features in the plots: first of all only transitions from the main isotope appear above 10 THz. Secondly, above 15 THz, the relative frequency difference drops sharply by more than six orders of magnitude. The agreement in frequency is so good that it is likely that the parameters were taken from the same source and that the small residue scattering is the result of the unit conversion. The line intensity agreement is good even though differences between intensity values may very well reach a factor of 2-3. There are four lines of the main isotope where the intensity difference is larger than one order of magnitude.

5.2 O_3

The comparison for O_3 is shown in Figure 4. The agreement in frequency is very good except for a handful of lines with differences above one percent. For the lower state energy, the agreement is typically better than 0.01% but there is a group of several hundred lines with differences between 1–2%. However, there is a lot of structure in the intensity plots. While the bulk of the roughly 100000 identified lines show intensity differences of less than $\pm 20\%$, there are several hundred lines, mostly from the isotope 4 and 5 that disagree by several orders of magnitude. It is possible that these are the same lines that show the unusual disagreement in lower state energy.

5.3 N_2O

The comparison for N_2O is shown in Figure 5. In general, the agreement is very good for frequency as well as lower state energy. The intensity values are well-correlated with very little spread. However, there seems to be a general offset of about 5% in intensity for practically all the lines.

5.4 CO

The comparison for CO is shown in Figure 6. Frequency as well as lower state energy agree very well. The same is true for intensity.

5.5 CH_4

The comparison for CH_4 is shown in Figure 7. Only lines of isotope 3, the only one with rotational transitions, were identified in both catalogs. The agreement in frequency and lower state energy is very good. The intensities correlate very well but there is a 4.5%-offset between the JPL catalog and HITRAN.

5.6 O_2

The comparison for O_2 is shown in Figure 8. There is a large spread in relative frequency difference which might imply that some of the lines were actually mismatched. Especially the isotopes 1 and 2 show these large differences, while the agreement in frequency for isotope 3 is very good. The same picture appears in the intensity distribution. While the agreement for isotope 3 is good, the other two isotopes show huge offsets.

5.7 NO

The comparison for NO is shown in Figure 9. All the plots basically show a threefold structure: there is one branch where agreement is poor, one where it is good and one where it is excellent. There is no obvious explanation for this.

5.8 SO_2

The comparison for SO_2 is shown in Figure 10. There is a lot of structure in the frequency and lower state energy plots but the agreement is very good. The intensity agreement is good but it is interesting to note that the HITRAN intensities are always higher.

5.9 NO_2

The comparison for NO_2 is shown in Figure 11. There is only one spectral line for this molecule and the values in both catalogs agree very well.

5.10 NH_3

The comparison for NH_3 is shown in Figure 12. Even though there is some spread and structure in the frequency comparison plot, the agreement for all parameters is very good. The intensity differences are only between -10% to +20%.

5.11 HNO_3

The comparison for HNO_3 is shown in Figure 13. There is good agreement in frequency as well as lower state energy. However, the intensity plot shows relative offsets from a factor of 2–10 for about 35-40% of the matched spectral lines.

5.12 OH

The comparison for OH is shown in Figure 14. This seems to be a similiar case as NO in Figure 9. Three different branches are visible in all plots that show poor, satisfactory and excellent agreement.

5.13 HF

The comparison for HF is shown in Figure 15. The agreement for all of the 8 spectral lines that were matched is very good.

5.14 HCl

The comparison for HCl is shown in Figure 16. While the agreement on all parameters is very good for most lines, there are five lines of each isotope that show an offset of almost three orders of magnitude. Interestingly, this offset appears at exactly the same JPL intensities for both isotopes and the HITRAN values for each isotope are also exactly the same.

5.15 HBr

The comparison for HBr is shown in Figure 17. This is a similiar case as Figure 16 for HCl. There is the same correlation of JPL and HITRAN intensities with large offsets.

5.16 ClO

The comparison for ClO is shown in Figure 18. The general agreement is very good on all parameters for the largest part of the roughly 5100 spectral lines. However, about 200 of these lines show large disagreement in intensity with no obvious structure.

5.17 OCS

The comparison for OCS is shown in Figure 19. The agreement on frequency and lower state energy is very good. The HITRAN intensity values are generally higher by an offset of up to 25%. This offset seems to decrease with intensity.

5.18 H_2CO

The comparison for H_2CO is shown in Figure 20. There is good agreement on all parameters. As with OCS in Figure 19, the HITRAN intensity values appear to be higher than the ones listed in the JPL catalog by up to 30%.

5.19 HOCl

The comparison for HOCl is shown in Figure 21. There is a lot of structure in all plots but frequency and lower state energy agree well for all spectral lines. However, while the majority of intenstity values agree well, there is a large number of lines, where the HITRAN values are far higher. This affects both isotopes.

5.20 HCN

The comparison for HCN is shown in Figure 22. The agreement is very good except for three lines of isotope 2 and one line of isotope 1 where the intensity values of HITRAN disagree by up to one order of magnitude.

5.21 H_2O_2

The comparison for H_2O_2 is shown in Figure 23. The agreement on frequency is good with a handful of lines above one percent relative difference. Even though a number of lines agree very well in intensity the majority shows a general offset of about 35% between JPL and HITRAN.

5.22 H_2S

The comparison for H_2S is shown in Figure 24. There is a lot of spread in frequency and lower state energy differences but no obvious structure is visible. The agreement on intensity is very good without any outliers.

5.23 HO_2

The comparison for HO_2 is shown in Figure 25. There is a lot of structure in all the parameter plots. Some lines show large frequency offsets which might imply that they have been mismatched. However, the number of these lines or lines with poorly matching lower state energy is very small. However, the intensity differences are only small for about half of the lines. For the other half, intensity differences can reach several orders of magnitude.

5.24 HOBr

The comparison for HOBr is shown in Figure 26. The agreement on frequency and lower state energy is very good. However, for the largest number of lines the disagreement on intensity is large. For the largest number of lines, the HITRAN values are consistently larger by a factor of roughly 4. For the rest of the lines, this factor reaches several orders of magnitude.

Table 4
Complete list of possible combinations of HITRAN and JPL catalog quantum numbers that have to be compared. The definitions for the HITRAN groups and JPL classifications can be found in [2] and [1], respectively.

Case A: HITRAN group 1 vs. JPL asymmetric rotor with vibration

HITRAN		JPL
iv', iv''	$\longrightarrow$	v', v'' [a]
J', J''	$\longrightarrow$	N', N''
K'_a, K''_a	$\longrightarrow$	K'_{-1}, K''_{-1}
K'_c, K''_c	$\longrightarrow$	K'_{+1}, K''_{+1}

[a] from table of corresponding iv and v

Case B: HITRAN group 1 vs. JPL asymmetric rotor

HITRAN		JPL
J', J''	$\longrightarrow$	N', N''
K'_a, K''_a	$\longrightarrow$	K'_{-1}, K''_{-1}
K'_c, K''_c	$\longrightarrow$	K'_{+1}, K''_{+1}
F', F''	$\longrightarrow$	J', J''

Case C: HITRAN group 2 vs. JPL linear Σ

HITRAN		JPL
J''	$\longrightarrow$	N''
$F''+1$	$\longrightarrow$	J''

Case D: HITRAN group 4 vs. JPL symmetric rotor

HITRAN		JPL
J', J''	$\longrightarrow$	N', N''
K', K''	$\longrightarrow$	K', K''

Case E: HITRAN group 5 vs. JPL linear Σ

HITRAN		JPL
N'', J''	$\longrightarrow$	N', J' [a]
N'', J''	$\longrightarrow$	N'', J'' [b]
$F''+\frac{1}{2}$	$\longrightarrow$	F''_1

[a] if N' and J' are defined
[b] if N' and J' are not defined

Case F: HITRAN group 6 vs. JPL linear case a (2S+1 even)

HITRAN		JPL
$J''+\frac{1}{2}$	$\longrightarrow$	J''
$F''+1$	$\longrightarrow$	F''_1
Sym''	$\longrightarrow$	Λ'' [a]

[a] '+' corresponds to 1, and '-' to -1

Case G: HITRAN group 6 vs. JPL undocumented case (ClO species)

HITRAN		JPL
$J''+\frac{1}{2}$	$\longrightarrow$	J''
F''	$\longrightarrow$	F''_1
Sym''	$\longrightarrow$	Λ [a]

[a] '+' corresponds to 1, and '-' to -1

Case H: HITRAN group 6 vs. JPL symmetric rotor

HITRAN		JPL
$J''+\frac{1}{2}$	$\longrightarrow$	J''
$F''+\frac{1}{2}$	$\longrightarrow$	F''_1
Sym''	$\longrightarrow$	K'' [a]

[a] '+' corresponds to 1, and '-' to -1

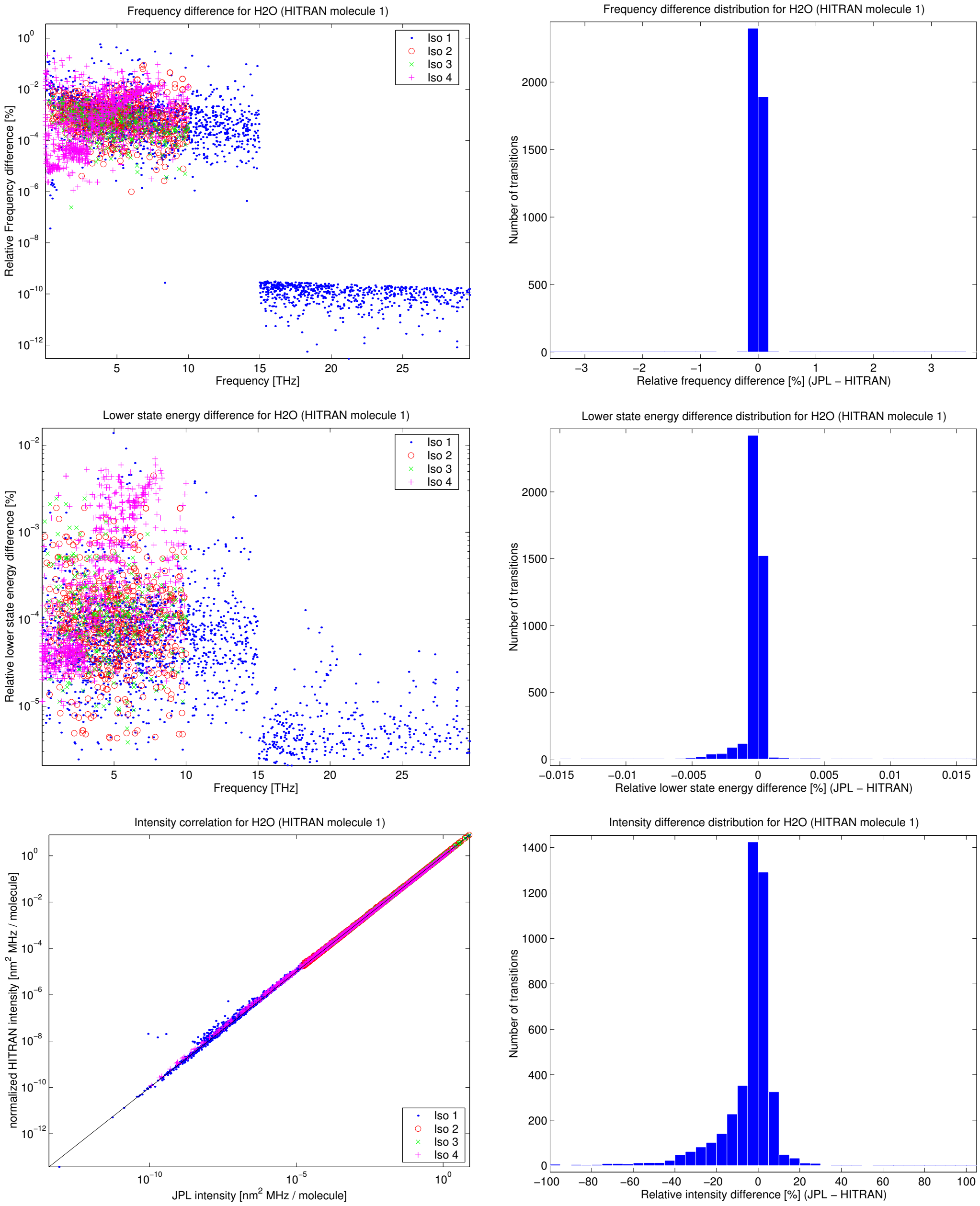


Fig. 3. Spectral line data intercomparison for H_2O. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

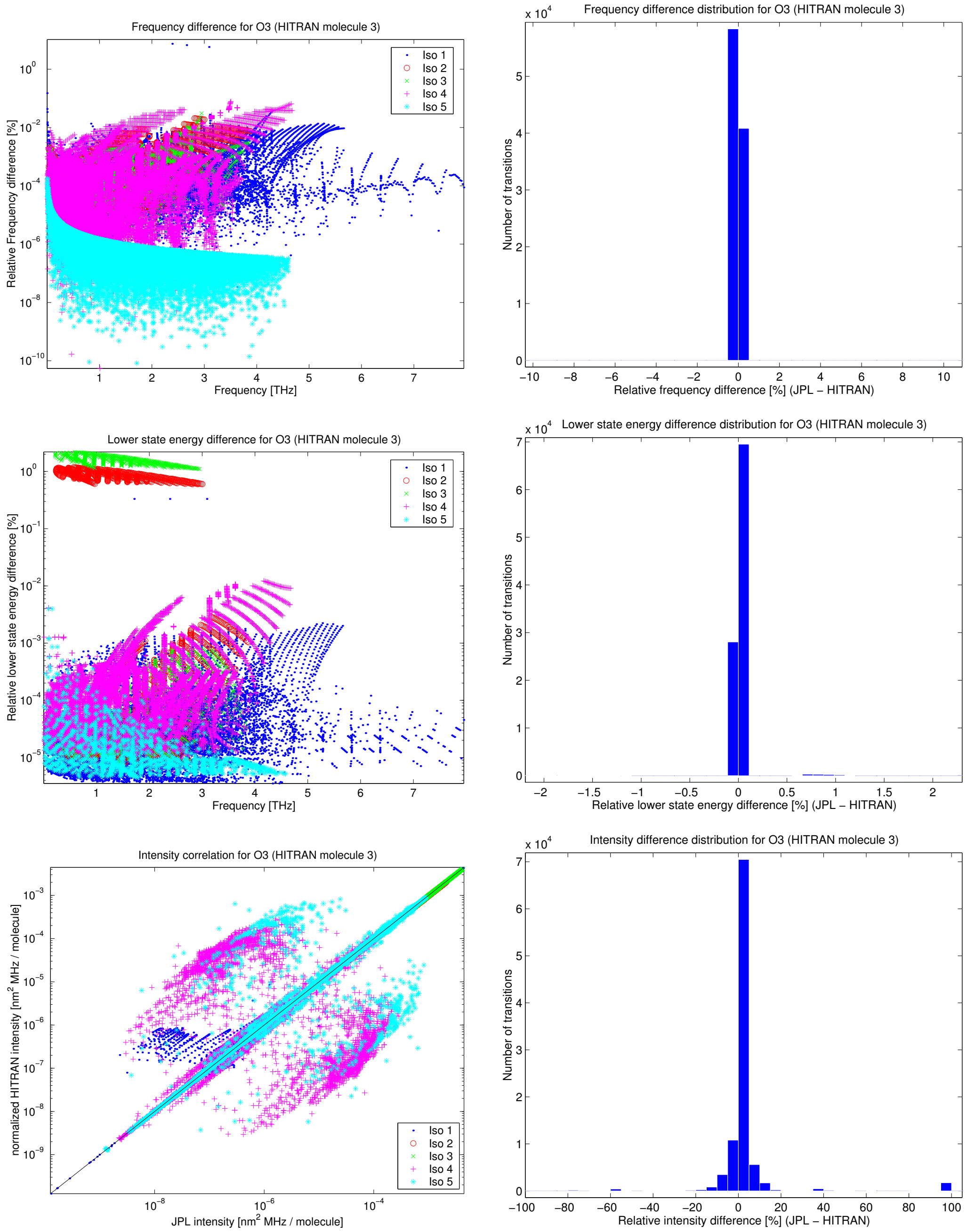


Fig. 4. Spectral line data intercomparison for O_3. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

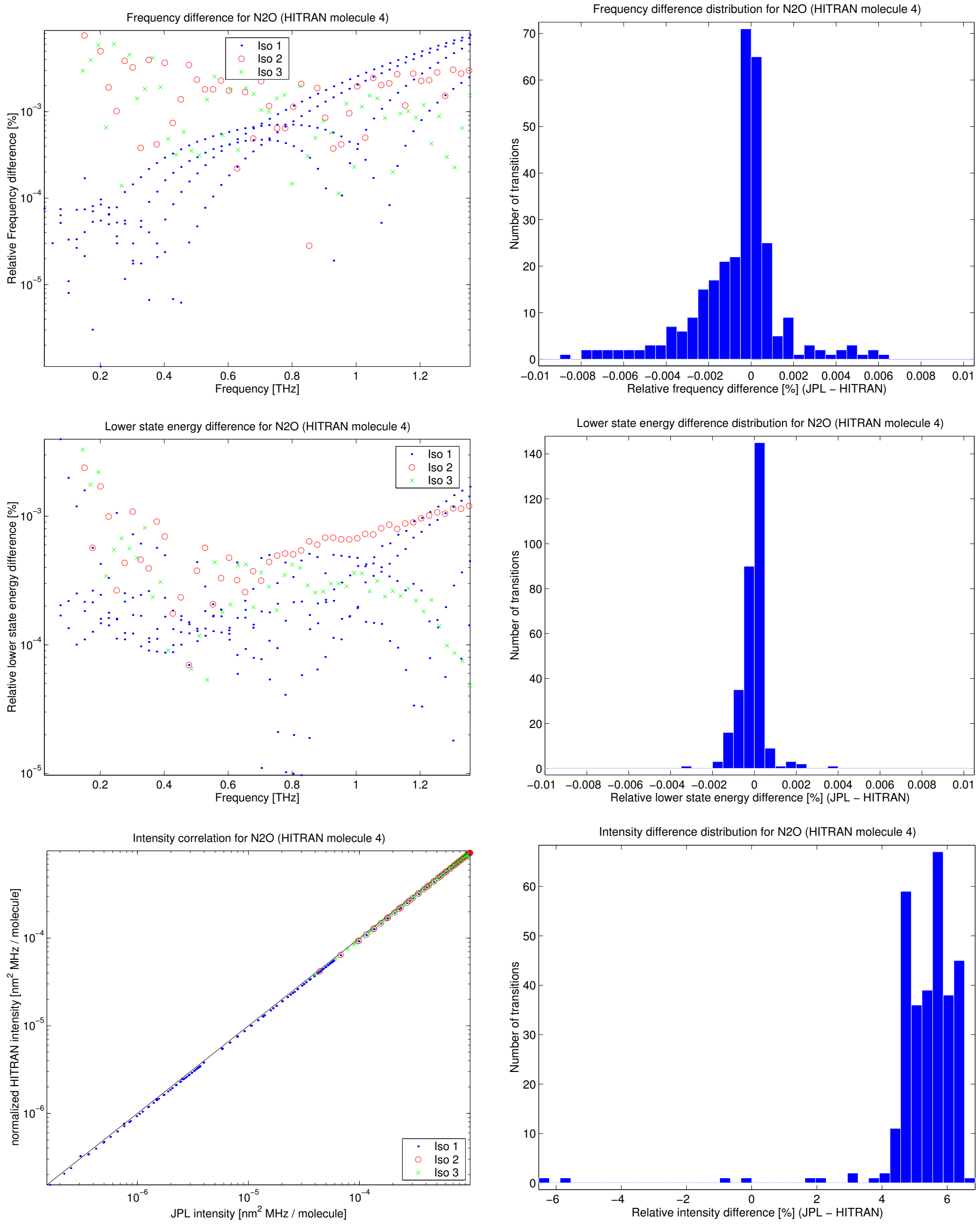


Fig. 5. Spectral line data intercomparison for N_2O. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

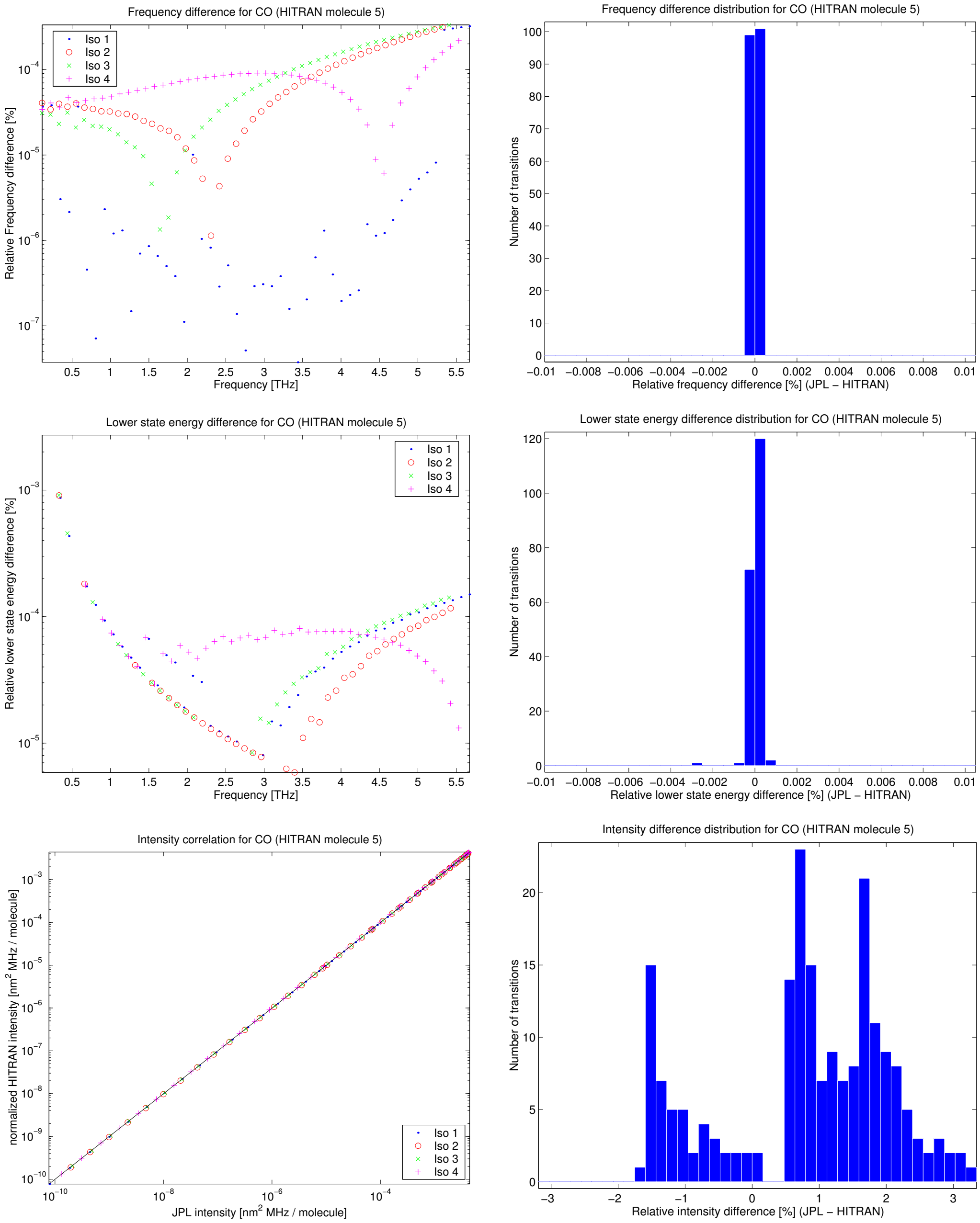


Fig. 6. Spectral line data intercomparison for CO. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

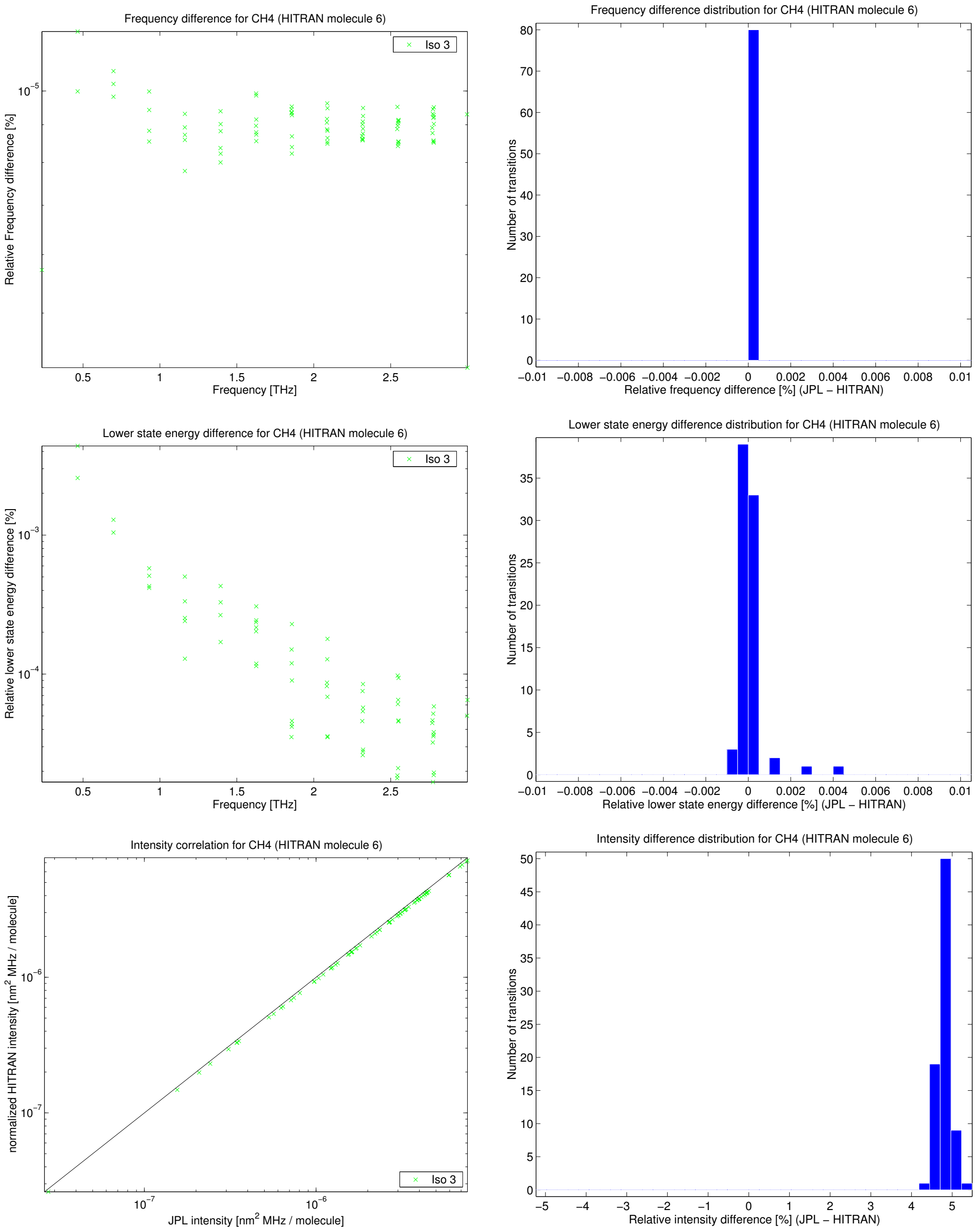
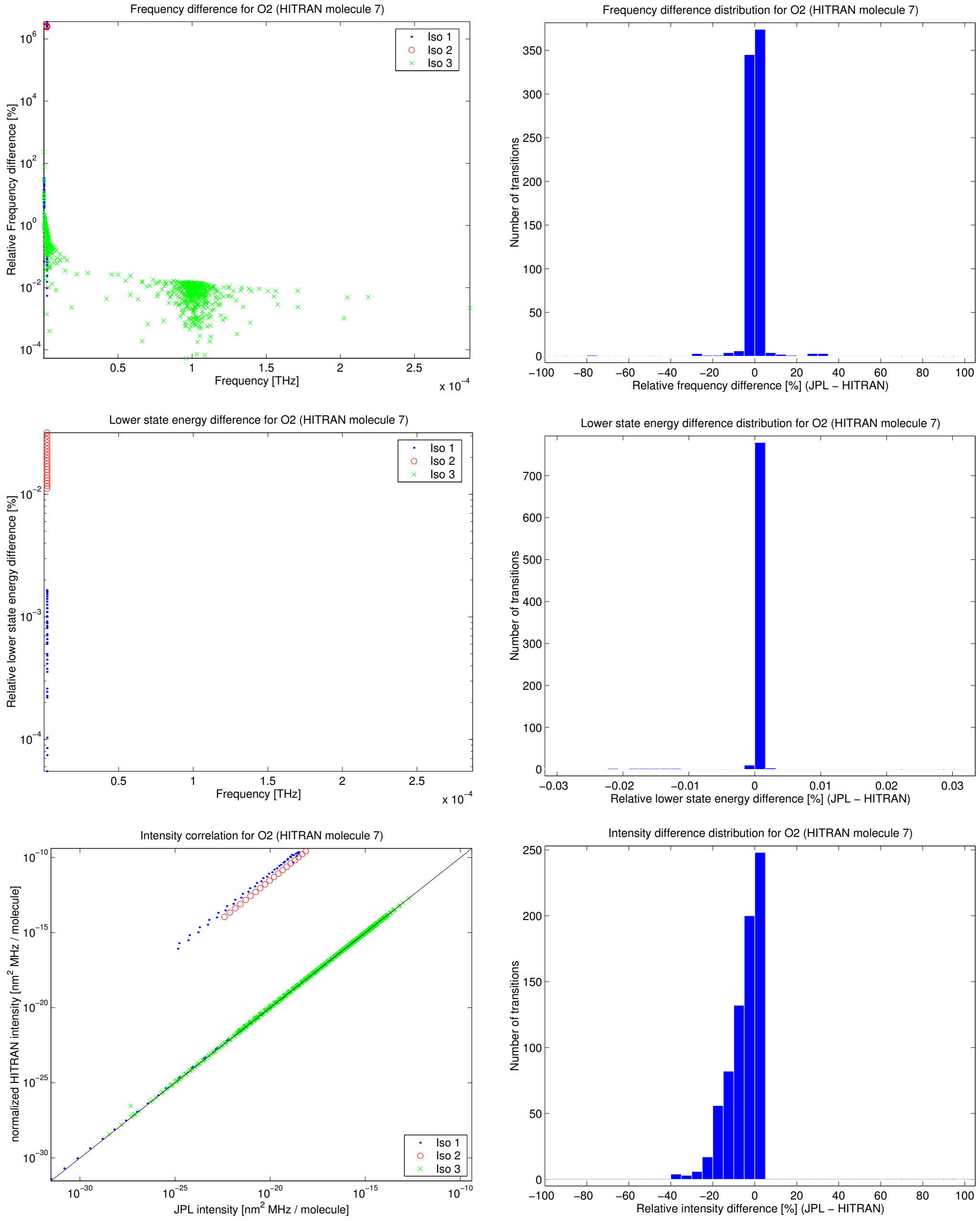


Fig. 7. Spectral line data intercomparison for CH_4. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.


Frequency difference for O2 (HITRAN molecule 7)

Frequency difference distribution for O2 (HITRAN molecule 7)

Lower state energy difference for O2 (HITRAN molecule 7)

Lower state energy difference distribution for O2 (HITRAN molecule 7)

Intensity correlation for O2 (HITRAN molecule 7)

Intensity difference distribution for O2 (HITRAN molecule 7)

Fig. 8. Spectral line data intercomparison for O_2. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

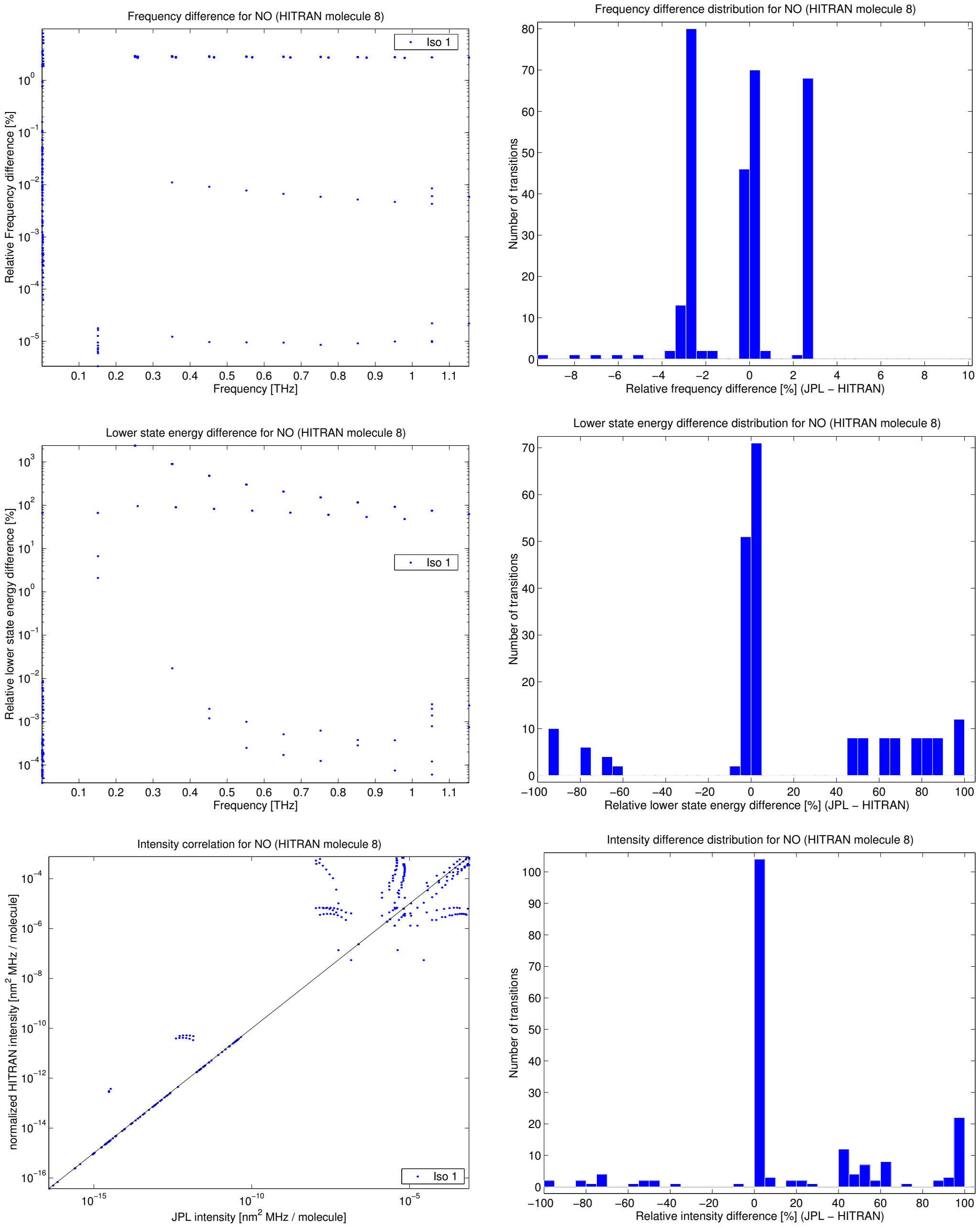


Fig. 9. Spectral line data intercomparison for NO. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

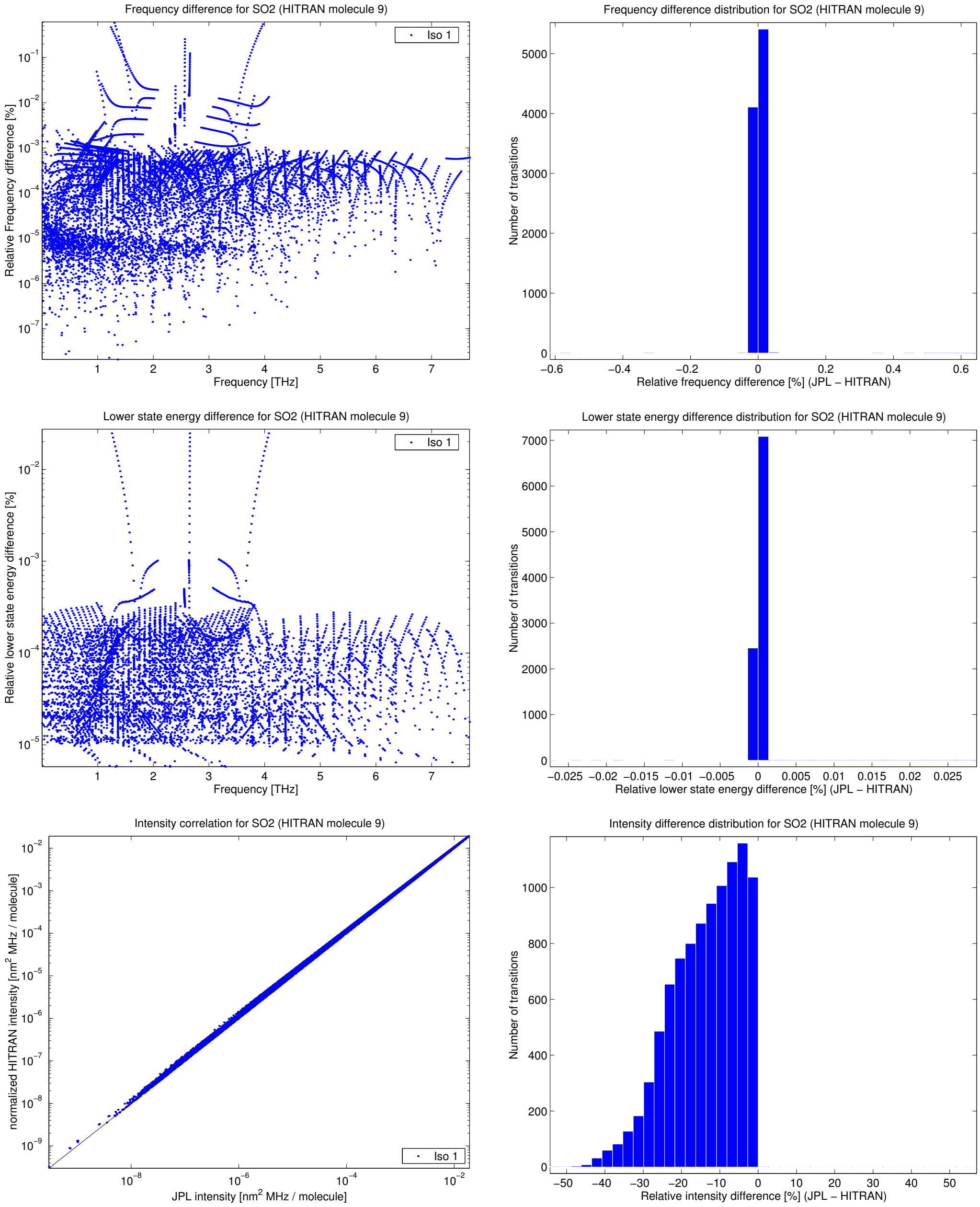


Fig. 10. Spectral line data intercomparison for SO_2. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

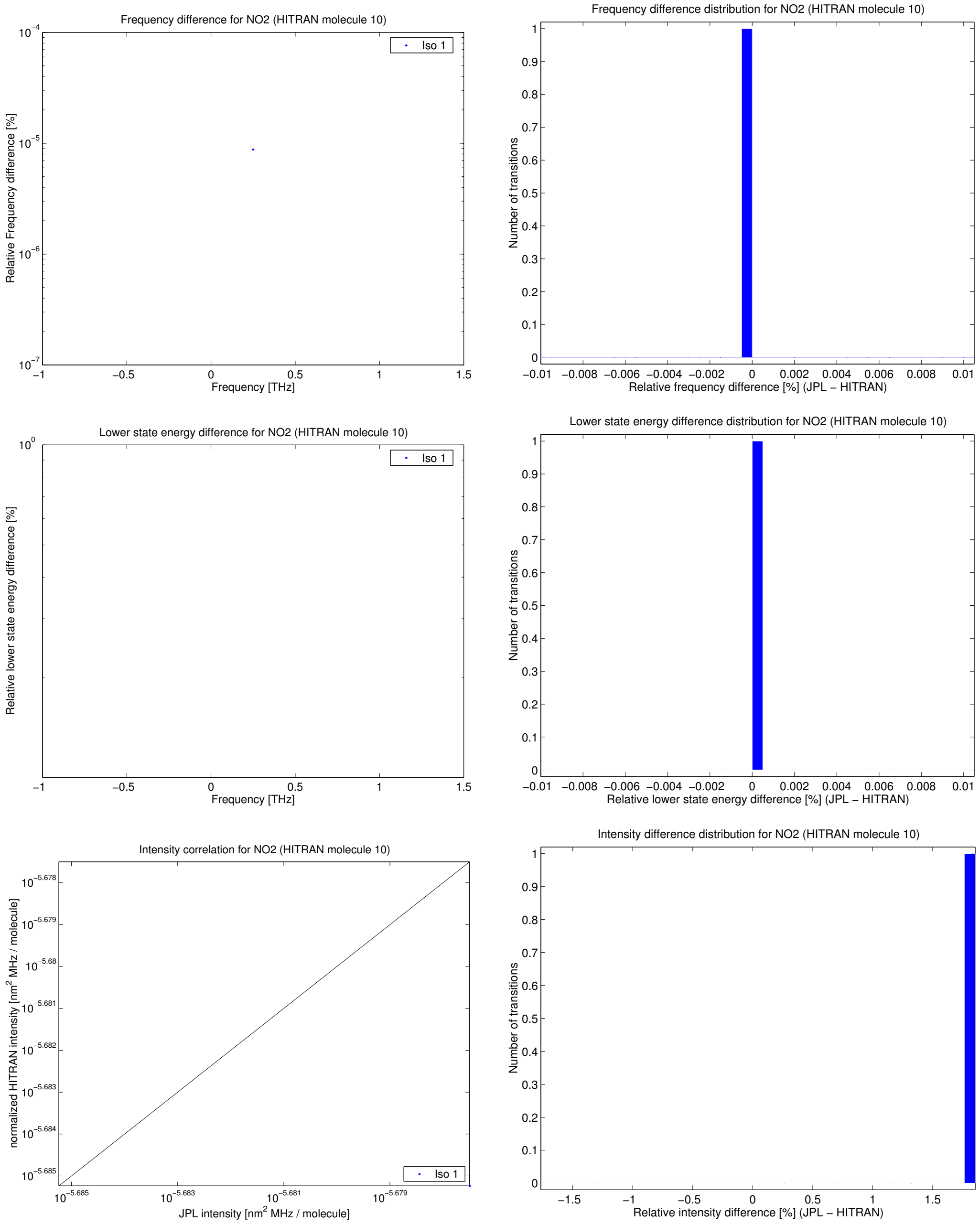


Fig. 11. Spectral line data intercomparison for NO_2. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

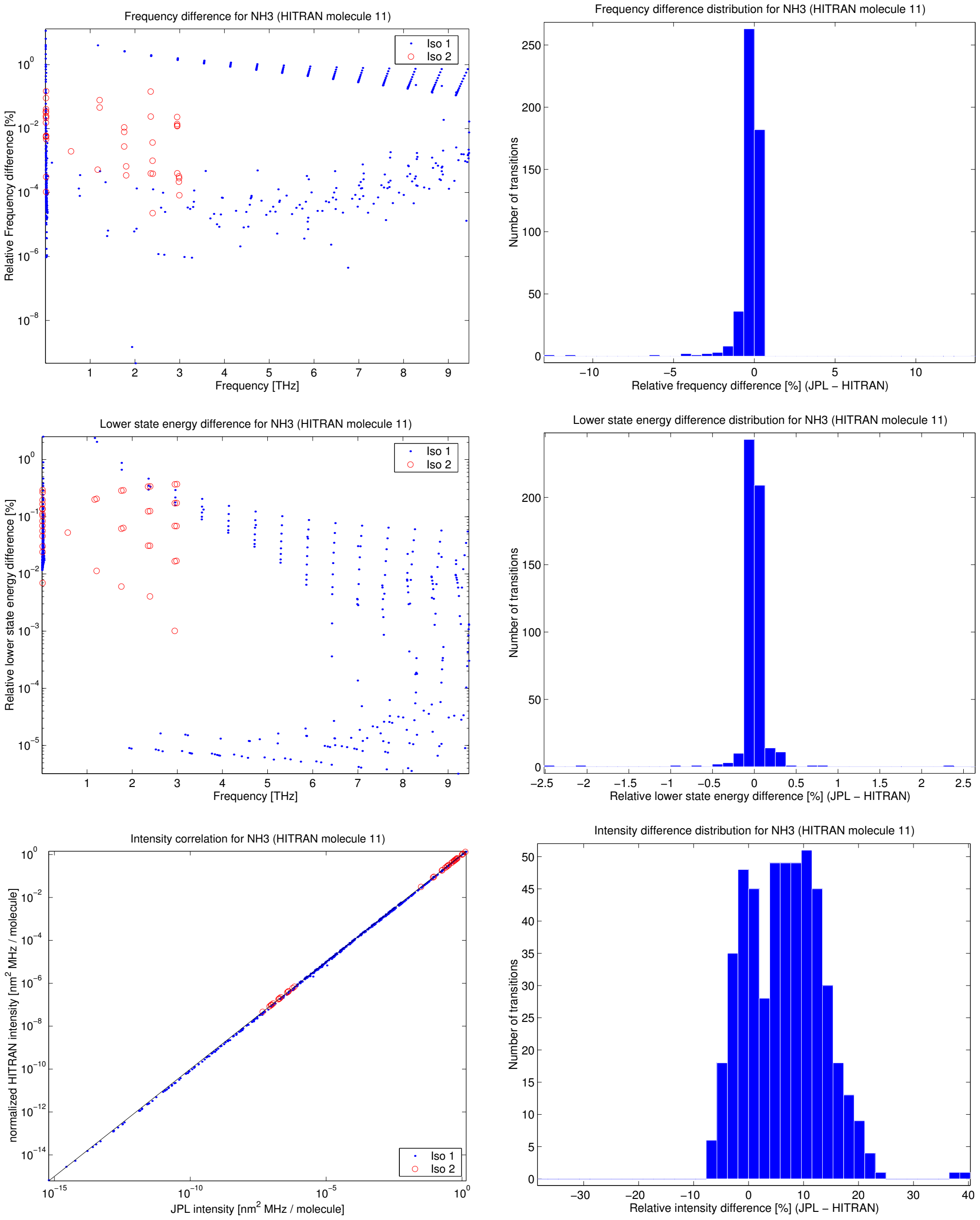


Fig. 12. Spectral line data intercomparison for NH_3. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

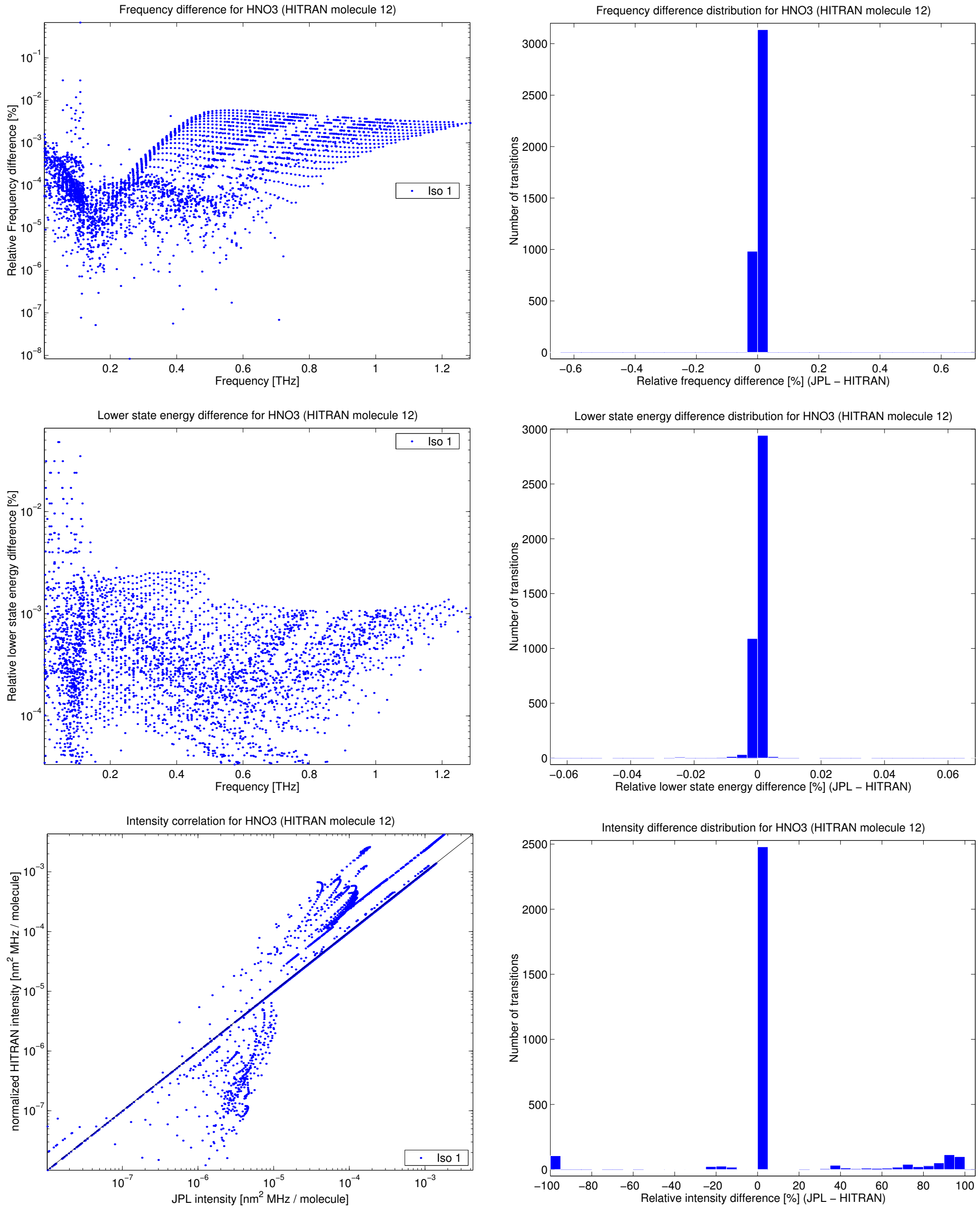


Fig. 13. Spectral line data intercomparison for HNO_3. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

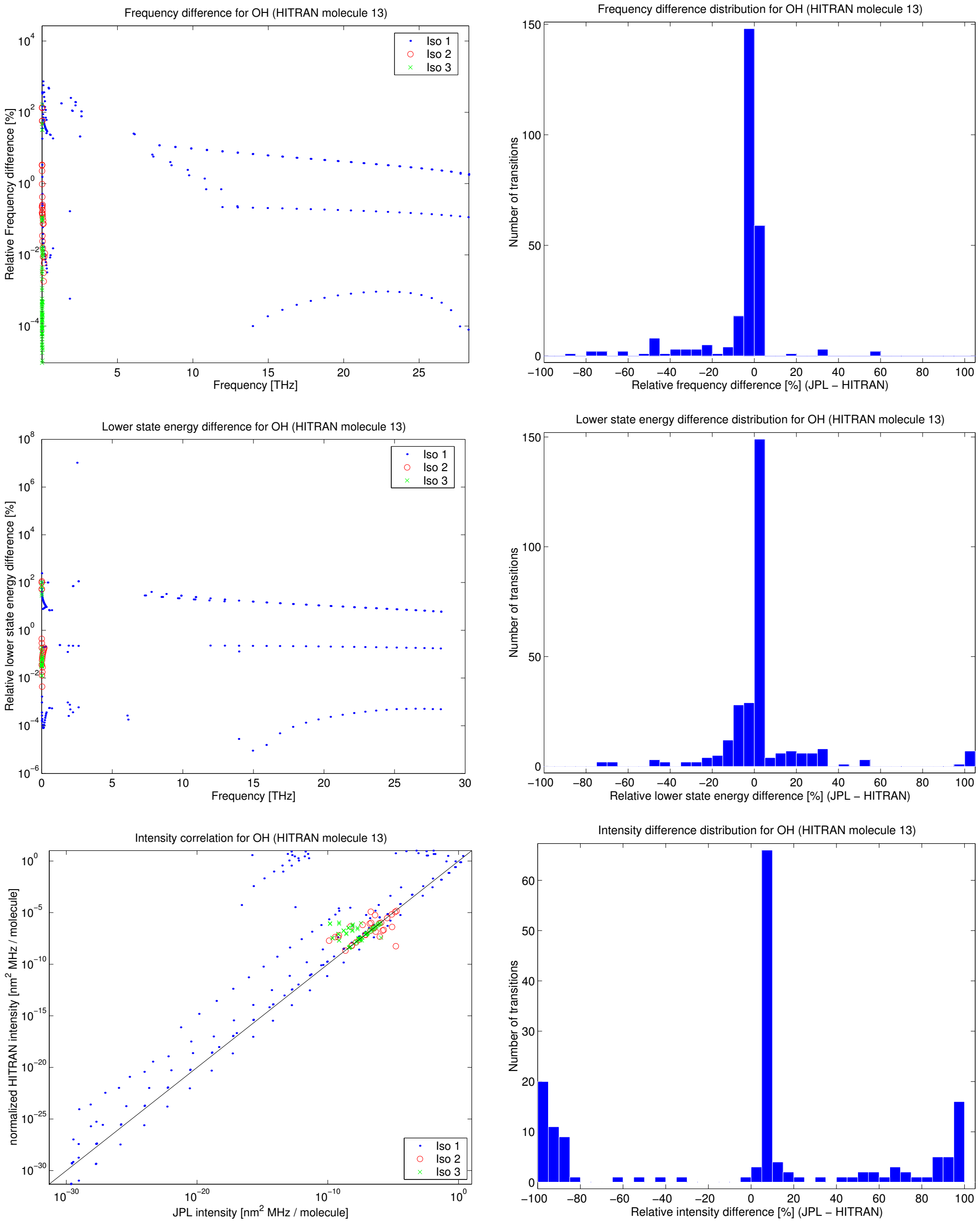


Fig. 14. Spectral line data intercomparison for OH. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

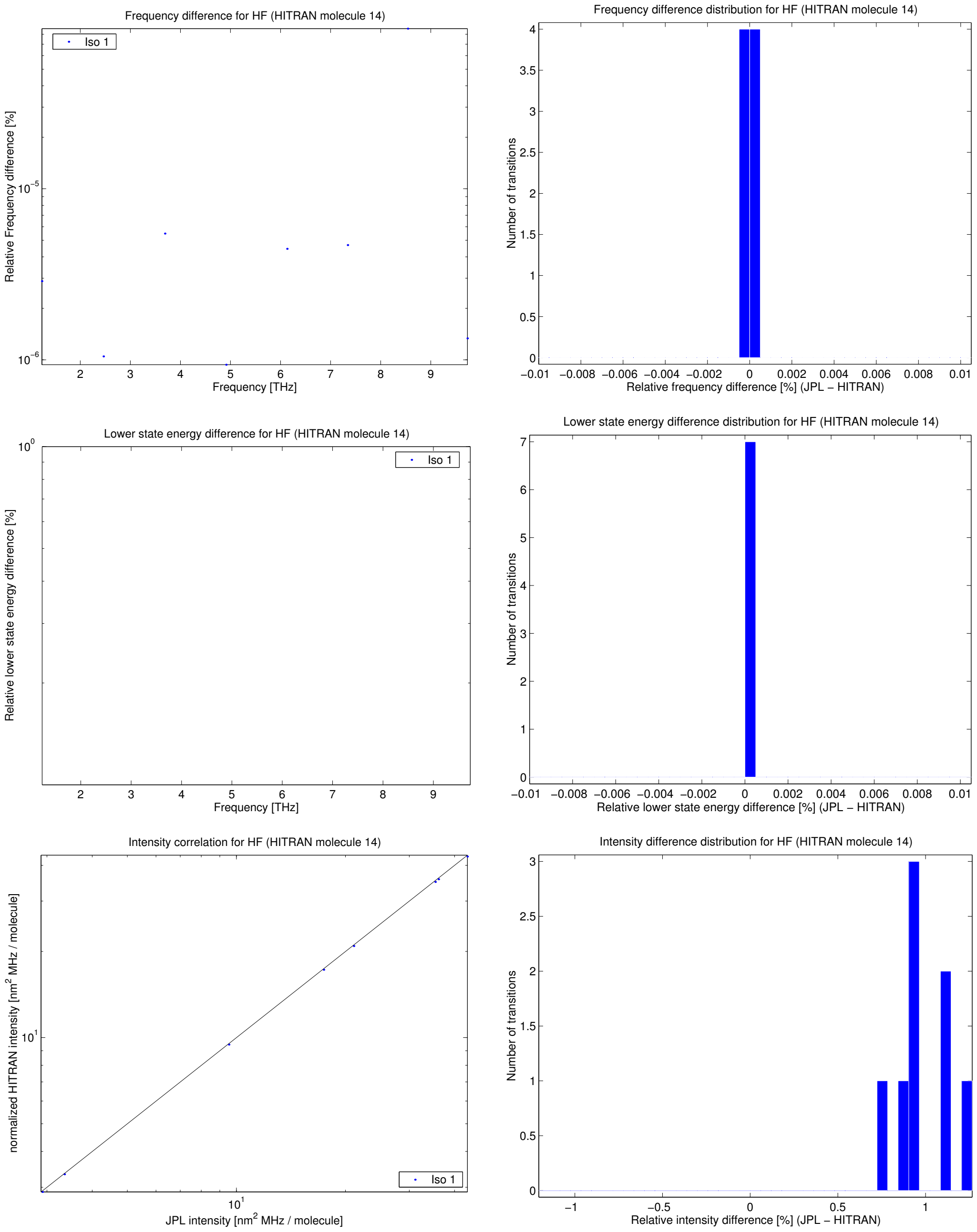


Fig. 15. Spectral line data intercomparison for HF. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

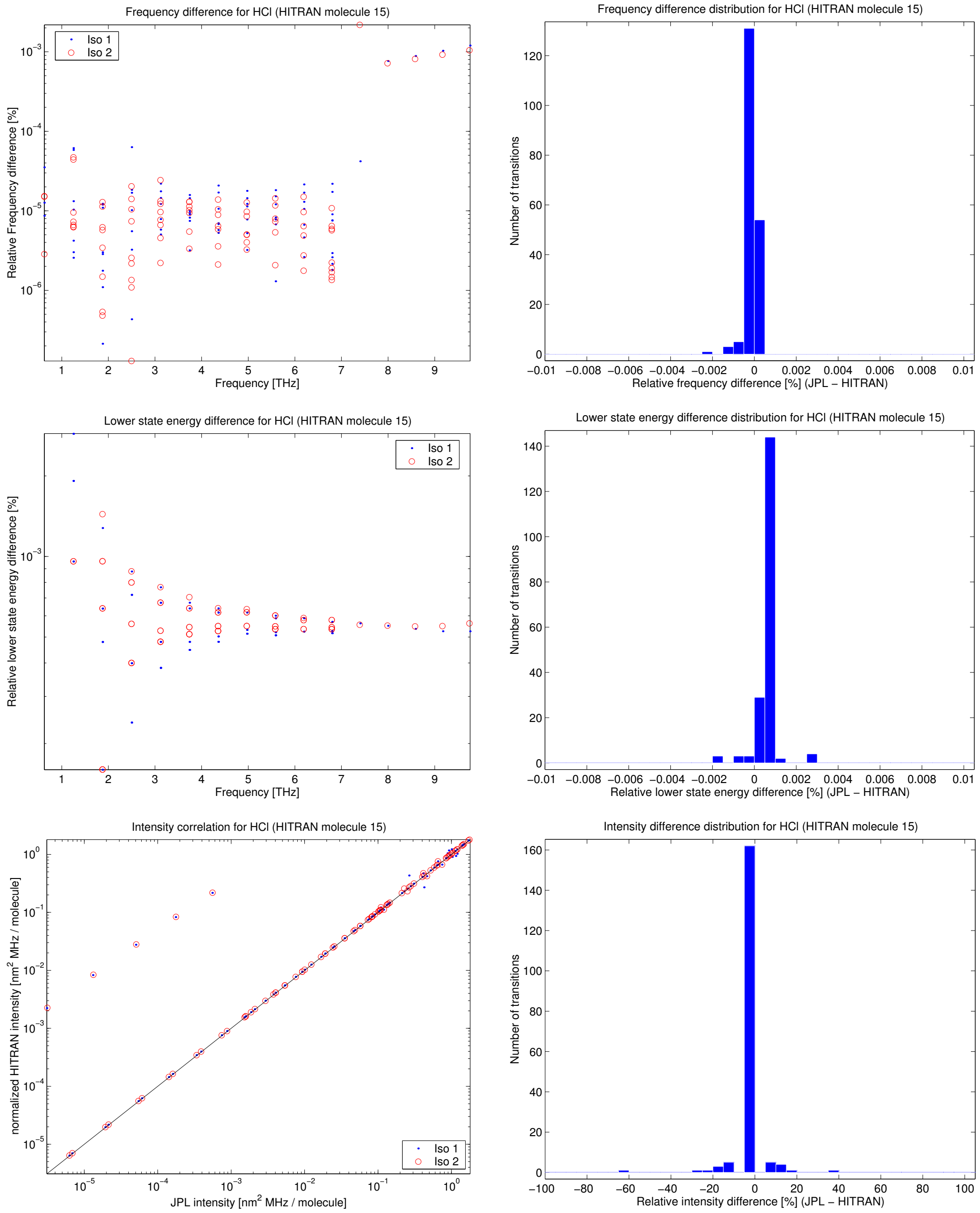


Fig. 16. Spectral line data intercomparison for HCl. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

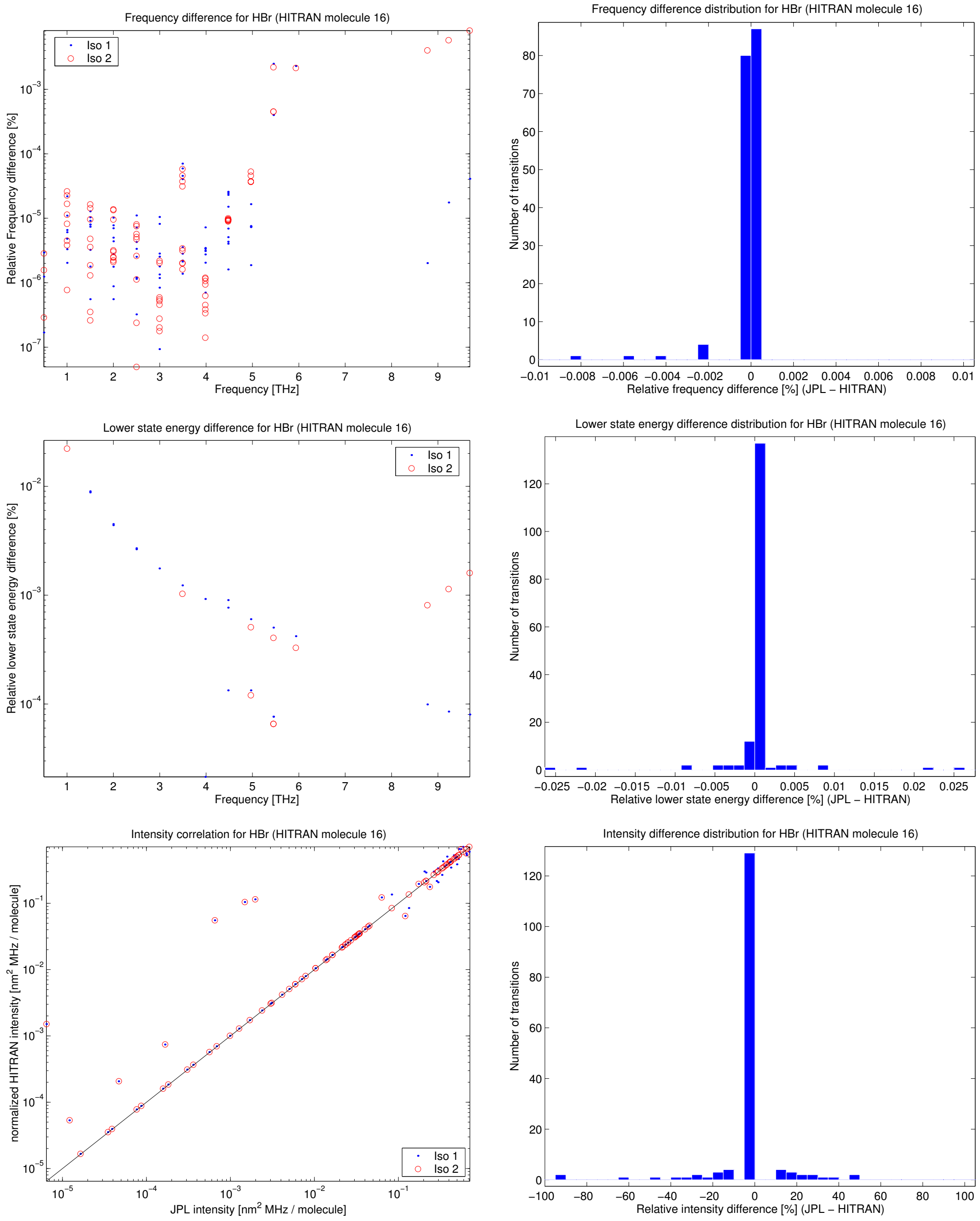


Fig. 17. Spectral line data intercomparison for HBr. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

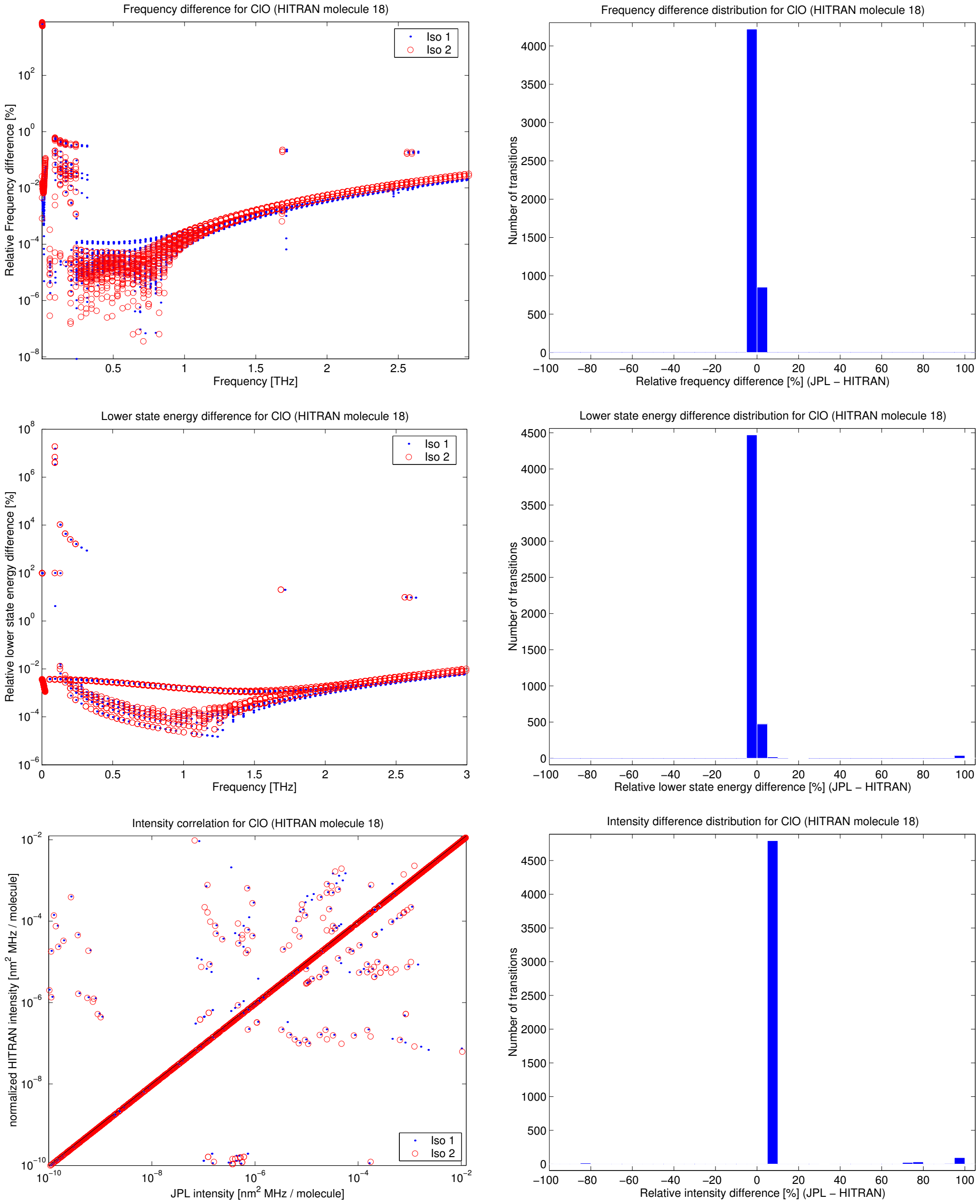


Fig. 18. Spectral line data intercomparison for ClO. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

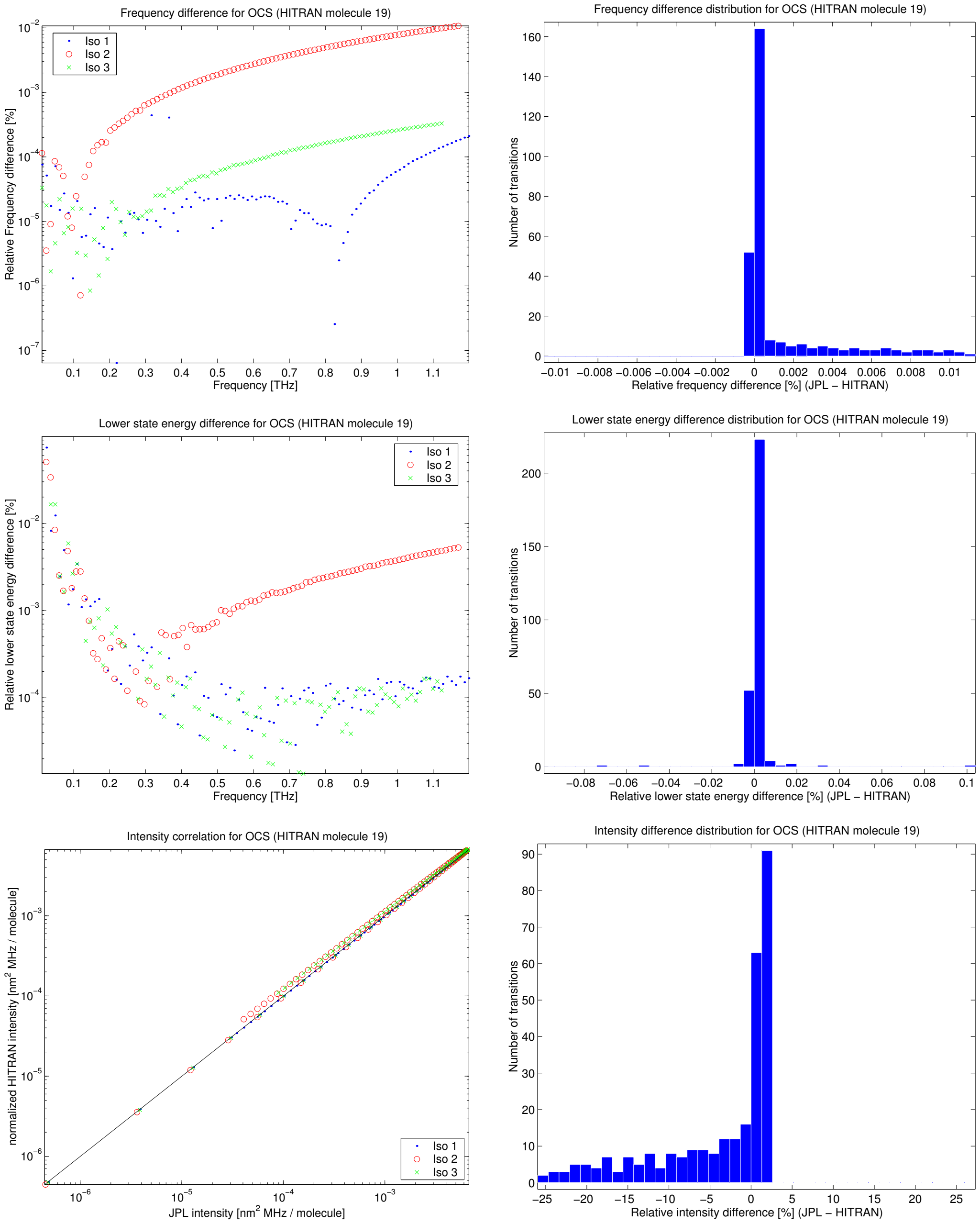


Fig. 19. Spectral line data intercomparison for OCS. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

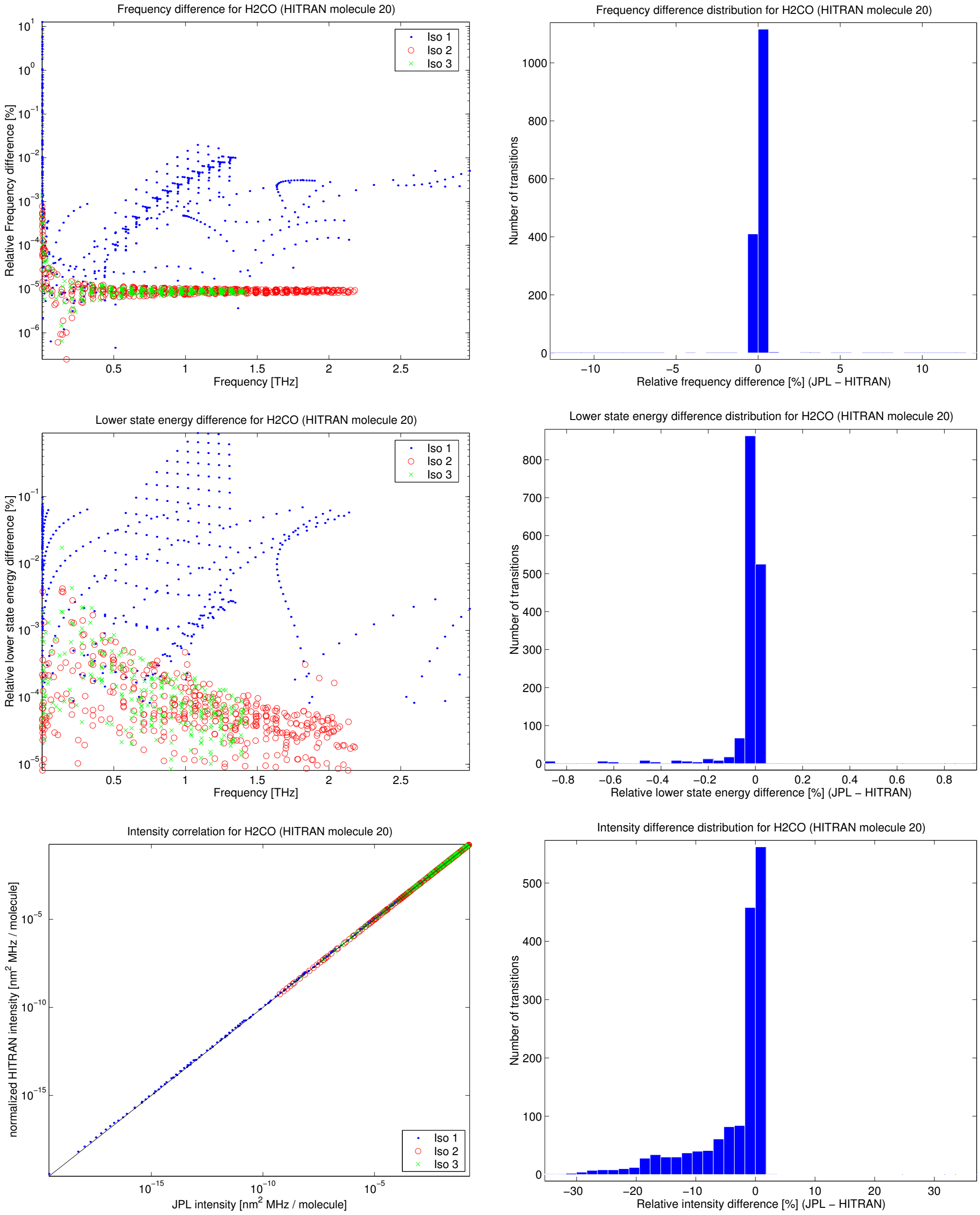


Fig. 20. Spectral line data intercomparison for H_2CO. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

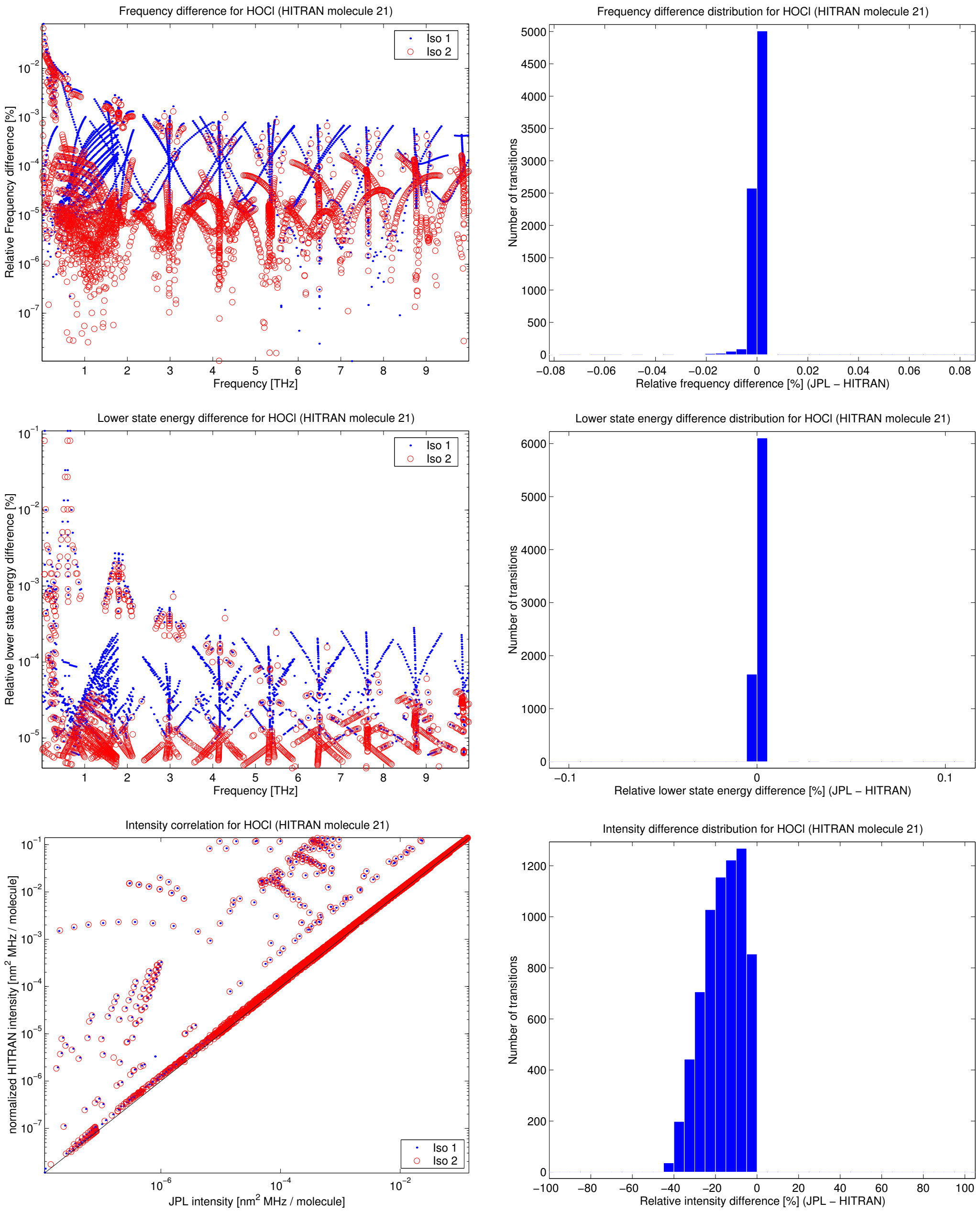


Fig. 21. Spectral line data intercomparison for HOCl. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

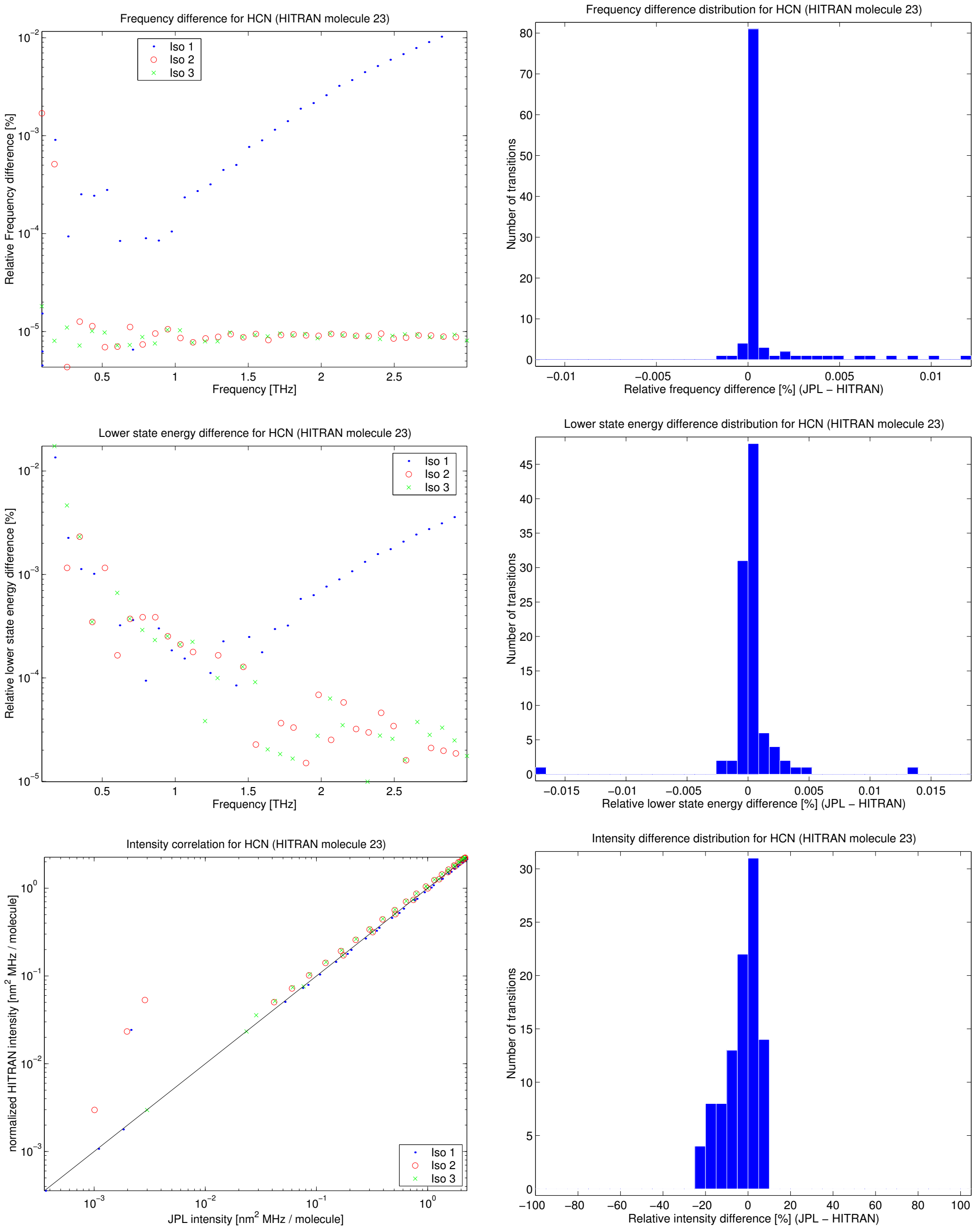


Fig. 22. Spectral line data intercomparison for HCN. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

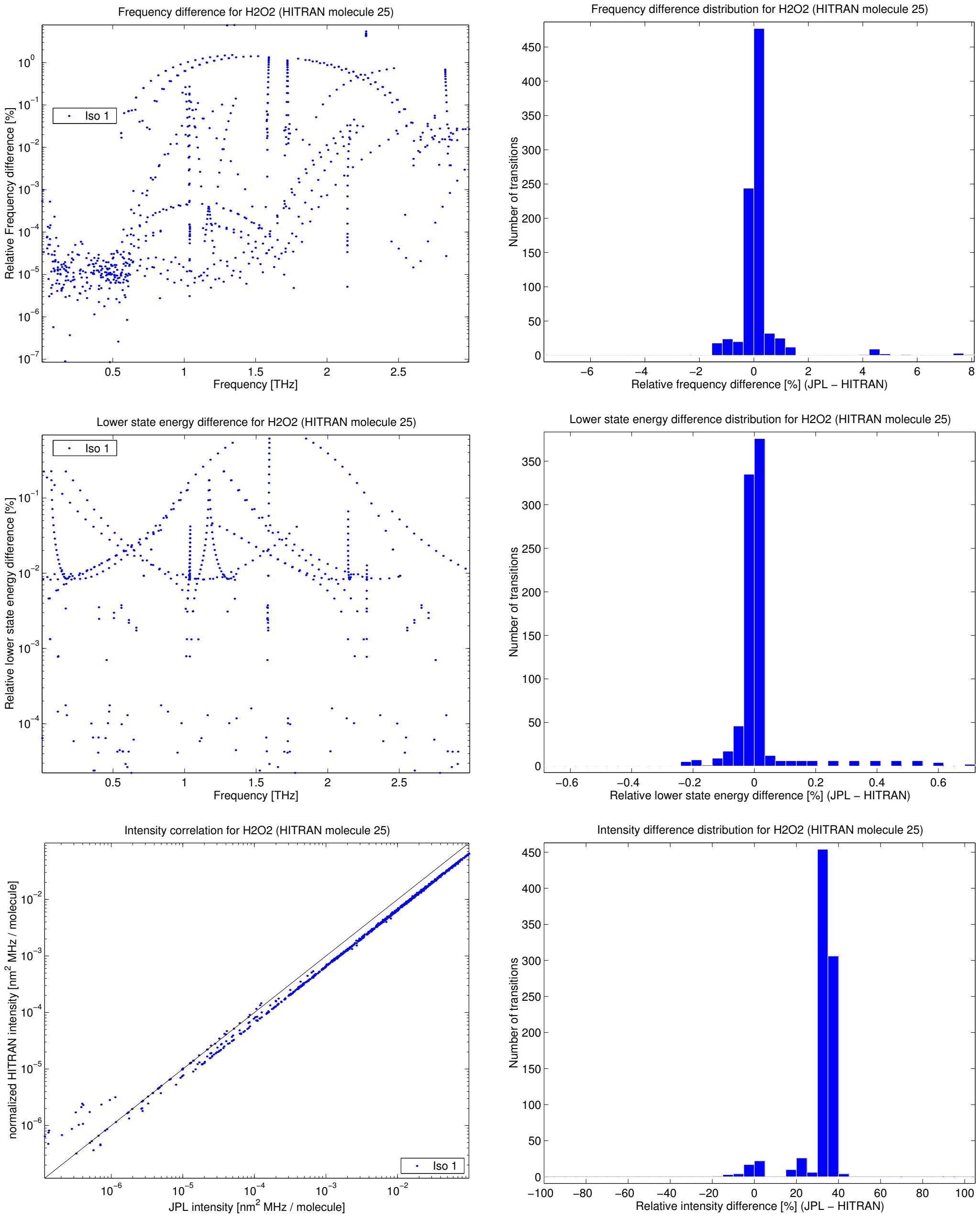


Fig. 23. Spectral line data intercomparison for H_2O_2. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

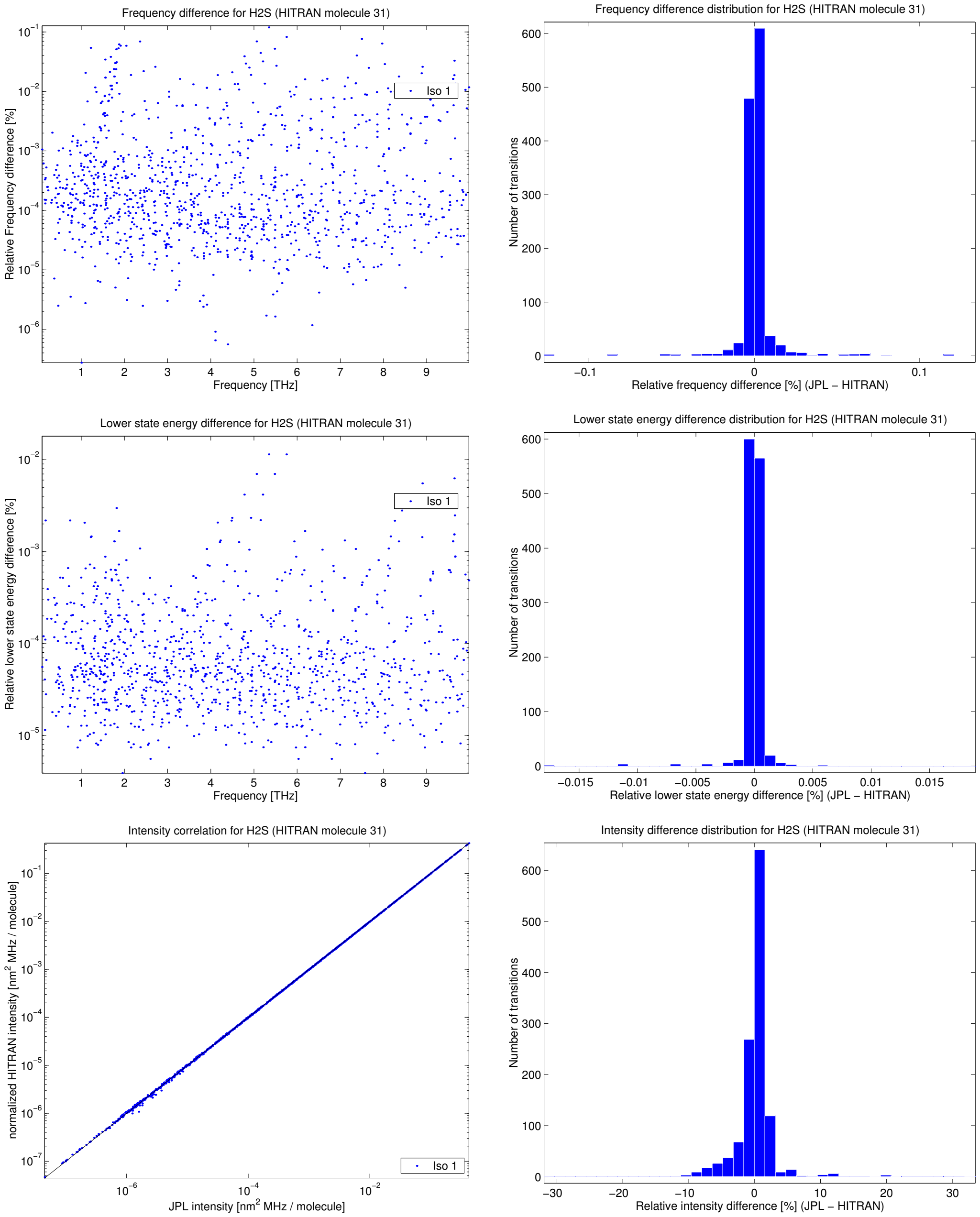


Fig. 24. Spectral line data intercomparison for H_2S. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

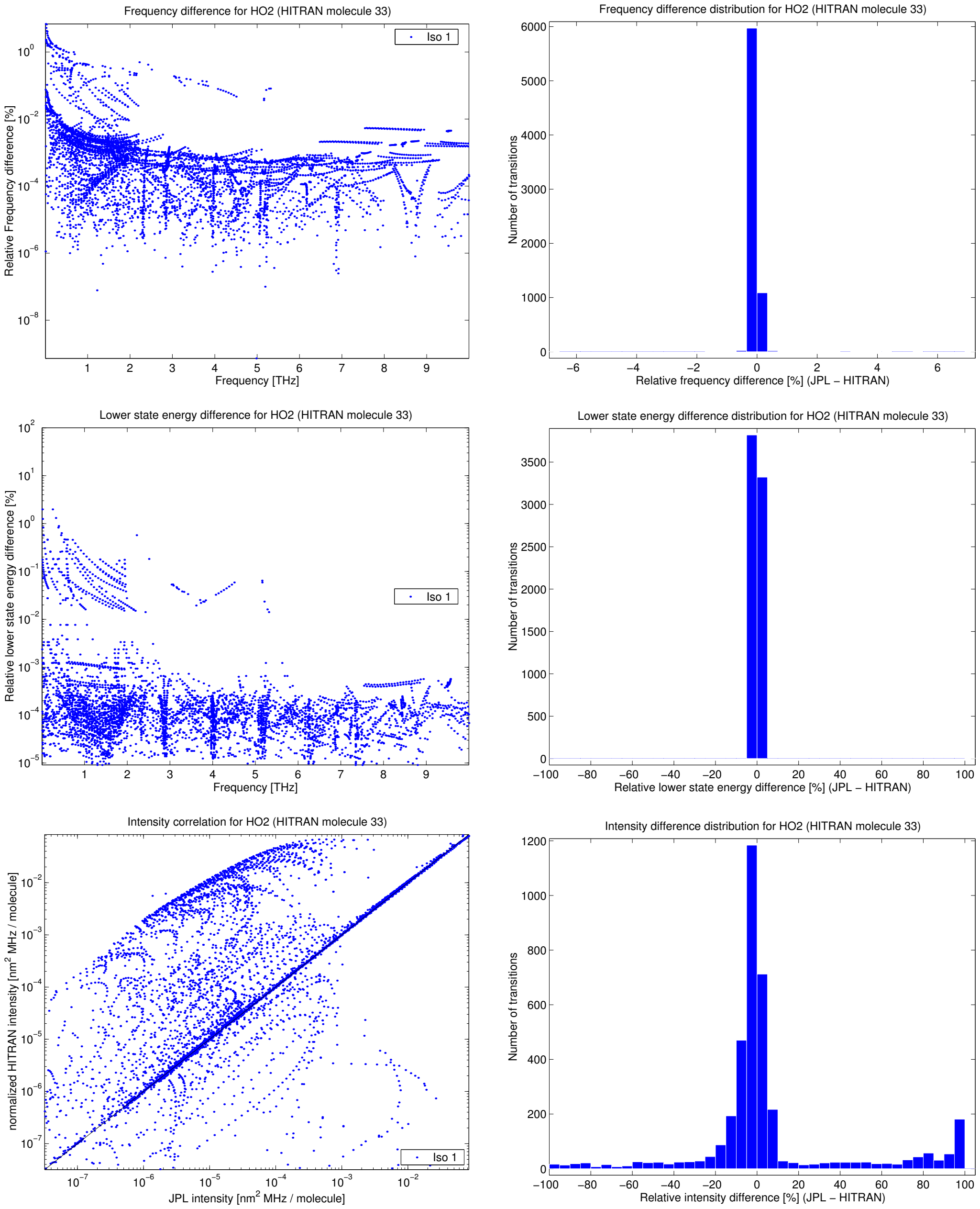


Fig. 25. Spectral line data intercomparison for HO_2. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

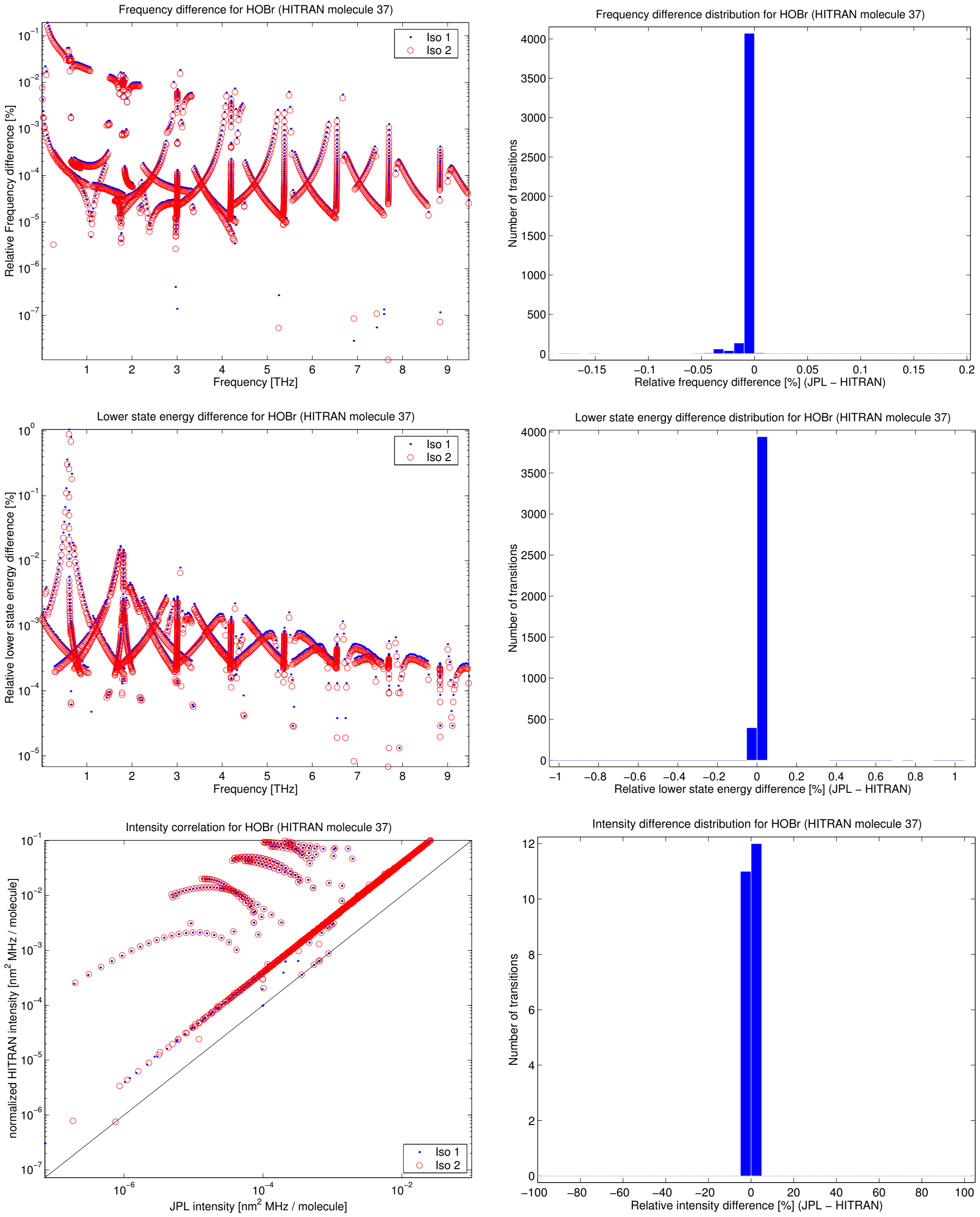


Fig. 26. Spectral line data intercomparison for HOBr. Top row: frequency difference. Center row: lower state energy difference. Bottom row: intensity difference.

6 Conclusions

The BEAMCAT meta data base provides a new way for users of spectral line catalogs to access the best available data from different catalogs. Instead of listing spectral line information itself, BEAMCAT provides links between the same spectral lines in different catalogs. With this link information, the BEAMCAT data base server can extract and combine information from several spectral line catalogs and output them in almost any format. This concept overcomes the limitations of traditional file-based spectral line catalogs. It allows the user to select spectral line information from different sources according to his/her quality criteria and in a format of his/her choice.

The first catalogs that were integrated into BEAMCAT were HITRAN and the JPL catalog. A large number of common spectral lines were identified in the two catalogs. A first intercomparison of key parameters of these spectral lines shows both remarkable agreement as well as hard-to-explain discrepancies. In general, the center frequencies agree very well, typically on the order of much less than $\pm 1\%$. The same is true for the lower state energy of these transitions. However, the agreement of spectral line intensities looks very different. A good agreement for intensities is in the order of $\pm 10\%$ while differences of several orders of magnitude exist, too. It is not clear to the author why the discrepancy becomes so large for so many spectral lines. It is likely that some spectral lines were not matched correctly between the two catalogs during the building process of BEAMCAT. However, the extremely good agreement in frequency suggests that less than two hundred spectral lines out of more than 148000 are obvious mismatches. Further investigation of this issue is necessary.

The JPL and HITRAN will not be the only spectral line catalogs in BEAMCAT for long. Among the large catalogs, GEISA is a very good candidate to be included in the next release. However, a catalog like BEAMCAT will become even more useful when it also contains a number of small catalogs with very high quality information or even single measurements from the literature. This would provide the optimum combination of highest quality for selected lines with the completeness of the large spectral line catalogs. The scientific community is welcome to contribute to future releases of the BEAMCAT data base.

References

[1] H. M. Pickett, R. L. Poynter, E. A. Cohen, M. L. Delitsky, J. C. Pearson, H. S. P. Müller, Submillimeter, millimeter, and microwave spectral line catalog, J. Quant. Spectrosc. Radiat. Transfer 60 (5) (1998) 883–890.

[2] L. S. Rothman, 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, B. B. Wattson, K. Yoshino, K. V. Chance, K. W. Jucks, L. R. Brown, V. Nemtchinov, P. Varanasi, The HITRAN molecular spectroscopic database and HAWKS (HITRAN Atmospheric Workstation): 1996 edition, J. Quant. Spectrosc. Radiat. Transfer 60 (5) (1998) 665–710.

[3] N. Jacquinet-Husson, E.Arié, J. Ballard, A. Barbe, G. Bjoraker, B. Bonnet, L. R. Brown, C. Camy-Peyret, J. P. Champion, A. Chédin, A. Chursin, C. Clerbaux, G. Duxbury, J.-M. Flaud, N. Fourrié, A. Fayt, G. Graner, R. Gamache, A. Goldman, V. Golovko, G. Guelachvili, J. M. Hartmann, J. C. Hilico, J. Hillman, G. Lefèvre, E. Lellouch, S. N. Mikhaïlenko, O. V. Naumenko, V. Nemtchinov, D. A. Newnham, A. Nikitin, J. Orphal, A. Perrin, D. C. Reuter, C. P. Rinsland, L. Rosenmann, L. S. Rothman, N. A. Scott, J. Selby, L. N. Sinitsa, J. M. Sirota, A. M. Smith, K. M. Smith, V. G. Tyuterev, R. H. Tipping, S. Urban, P. Varanasi, M. Weber, The 1997 spectroscopic GEISA databank, J. Quant. Spectrosc. Radiat. Transfer 62 (2) (1999) 205–254.

[4] K. Chance, K. W. Jucks, D. G. Johnson, W. A. Traub, The Smithsonian Astrophysical Observatory database SAO92, J. Quant. Spectrosc. Radiat. Transfer 52 (3/4) (1994) 447–457.

[5] A. Bauer, M. Birk, G. Wagner, J. M. Colmont, D. Priem, G. Wlodarczak, S. Bühler, A. von Engeln, K. Künzi, A. Perrin, Study on a spectroscopic database for millimeter and submillimeter wavelengths, ESA Contract 11581/95/NL/CN, European Space Agency (1998).

[6] J. P. Eriksson, F. Merino, On simulating passive observations of the middle atmosphere in the range of 1–1000 GHz, Tech. Rep. No. 179, Department of Radio and Space Science, Chalmers University of Technology, Göteborg, Sweden (Aug. 1997).

[7] D. G. Feist, Analysis and simulation of atmospheric microwave spectra measured with a space borne limb-sounding instrument, Ph.D. thesis, University of Berne, Berne, Switzerland (Apr. 1999).

[8] M. Widenius, D. Axmark, MySQL reference manual, O'Reilly & Associates, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, USA, 2002.

[9] D. G. Feist, BEAMCAT 1.0: A spectral line database for millimeter and submillimeter wave propagation in the Earth's atmosphere, Research report 99-1, Institute of Applied Physics, Dept. of Microwave Physics, Sidlerstr. 5, 3012 Bern, Switzerland (Mar. 1999).

A Access to the BEAMCAT database

The main part of this article focuses on general aspects of the BEAMCAT data base and on the intercomparison between the JPL and HITRAN catalog data. In addition to that, this appendix provides more practical information on how the BEAMCAT data base can be used by interested scientists.

BEAMCAT is provided as a service for the scientific community. It can be accessed by anyone at no charge. For security reasons, some services require users to register first.

A.1 BEAMCAT web site

The typical entry point for BEAMCAT users is the BEAMCAT web site. Its address is listed in section A.3. The web site provides the following services to BEAMCAT users:

- Information and documentation
- Simple access to the catalog data
- Registration service for frequent users

Users can access the BEAMCAT data base and all the spectral line catalogs within it through a web interface. Through the web interface, spectral line data from several catalogs can be merged and downloaded according to the user's selection criteria. The data can be downloaded in several data formats through a standard web browser without any knowledge of the SQL language. Registration is not necessary to access the web interface.

Frequent users of BEAMCAT are asked to register. Registered users will be notified about new developments around the BEAMCAT data base. They will also get a username and password to directly access the BEAMCAT data base server through a MySQL client application.

For the future, an advanced web interface for registered users is planned. This advanced web interface will provide more possibilities to select and combine data from different catalogs.

A.2 Direct access to the MySQL data base server

Advanced users are able to access the BEAMCAT MySQL server directly. While this requires some knowledge of SQL, it allows the user to access the data in any way they need it. For example, through direct SQL queries, users can change physical units and define their own data format. It is also possible to access the data base with self-written programs and scripts without human intervention.

Table A.1
Internet addresses for the BEAMCAT data base

Service	Address or Universal Resource Locator (URL)
Web site	`http://www.iapmw.unibe.ch/research/projects/BEAMCAT/`
Support	`mailto:beamcat@mw.iap.unibe.ch`
MySQL server	`beamcat.iapmw.unibe.ch`

A.3 Overview of BEAMCAT internet addresses

Table A.1 summarizes the most important internet addresses for the BEAMCAT data base. Care will be taken that these addresses will stay valid for as long as possible.