The database of spectroscopic constants of diatomic molecules (DSCDM): A dynamic and user-friendly interface for molecular physics and spectroscopy

Yueqian Wang [a],[1], Daniel Julian [a],[1], Mahmoud A.E. Ibrahim [a], Connor Chin [a], Saketh Bhattiprolu [a], Ethan Franco [a], Jesús Pérez-Ríos [a],[b],[*]

[a] *Department of Physics and Astronomy, Stony Brook University, Stony Brook, 11794, NY, USA*
[b] *IACS, Stony Brook University, Stony Brook, 11794, NY, USA*

ARTICLE INFO

ABSTRACT

We present the database of spectroscopic constants of diatomic molecules (DSCDM), a website dedicated to the spectroscopic constants of ground and excited states of diatomics (https://dscdm.physics.stonybrook.edu). The database can be improved based on community feedback: the community as contributors may upload new data. Additionally, it counts on an Application Programming Interface (API) for molecular spectroscopy data retrieval to make the search process more efficient. On the other hand, the website presents a machine learning predictor of spectroscopic constants based on atomic properties and a new plotting tool to study the statistical distribution of spectroscopic constants.

1. Introduction

Most quantum information sciences platforms based on ultracold molecules, and the most interesting scenarios in cold and ultracold chemistry, require molecules suitable for laser cooling, i.e., molecules with diagonal Franck–Condon factors. Hence, it is key to have spectroscopy data for the ground and excited states of diatomic molecules to find the most prominent candidates. On the other hand, we are living in the Big Data era, and with it comes the time for cloud computing. As a result, the current scientific paradigm is shifting towards data-driven science. Indeed, within the umbrella of artificial intelligence, data-driven science has irrupted in chemical physics, particularly in spectroscopy [1–6]. Hence, the need for reliable data has proliferated. A few databases relevant to molecular spectroscopy of diatomics have been developed, as it has been recently reviewed in Ref. [7], such as the Cologne Database for Molecular Spectroscopy [8], the NIST Diatomic Microwave Database [9,10] or DiRef. [11].

In the case of spectroscopic constants of diatomic molecules, the community counts on:

- The NIST chembook website [12], containing all spectroscopic data from Huber–Herzberg [13].

- The diatomic molecular spectroscopy database, a more recent website, contains 173 heteronuclear molecules from Huber–Herzberg [13], including ground state and at least one electronic excited state spectroscopic constants [14].

Similarly, DiRef counts with a large selection of references where users can find spectroscopic information on diatomic molecules. The NIST website is very useful, although outdated, and does not support direct data download, which makes it a robust but not user-friendly website. On the contrary, the diatomic molecular spectroscopy database incorporates most capabilities offered by modern websites, like an API, and it is possible to download the whole database and the particular data required. Furthermore, it provides Franck–Condon calculations and plotting options. However, it only contains data on 173 diatomic polar molecules.

Here, we present the database of spectroscopic constants of diatomic molecules (DSCDM), including spectroscopic information for the ground state for 343 diatomic molecules–polar and homonuclear, and 236 of them count with information of the ground electronic state and at least one excited electronic state. As a result, we deliver the most extensive database for spectroscopic constants of diatomic molecules up to date. The database is linked to a user-friendly website showing

---

* Corresponding author at: Department of Physics and Astronomy, Stony Brook University, Stony Brook, 11794, NY, USA.
E-mail address: jesus.perezrios@stonybrook.edu (J. Pérez-Ríos).
[1] Contributed equally to the work.

extra features: a general plotting tool for spectroscopic constants and a machine learning predictor of spectroscopic constants based on atomic properties, fueled by the study presented on Ref. [1]. Specifically, we deliver an online tool dedicated to predicting spectroscopic constants of diatomic molecules that can be used to identify potential good candidates for different applications of diatomics, such as laser cooling.

2. Data bases for diatomic molecules

2.1. Data gathering and cleaning

In this project, the spectroscopic constants of diatomic molecules have been gathered from various published books, papers, and online accessible databases (e.g., The Diatomic Molecular Spectroscopy Database [14] and NIST [12]). As an extension to the diatomic molecular spectroscopy database, we include the ground state spectroscopic constants of 32 homonuclear and 54 extra heteronuclear diatomic molecules from Refs. [12,13,15–84]. In particular, 340 excited states of 53 molecules (As_2, Au_2, B_2, Bi_2, Br_2, C_2, Cl_2, Cs_2, Cu_2, F_2, H_2, D_2, T_2, He_2, I_2, K_2, Kr_2, Li_2, Mg_2, N_2, Na_2, Ne_2, O_2, P_2, Pb_2, Po_2, Rb_2, S_2, Sb_2, Se_2, Si_2, Te_2, Xe_2 [13], AuF [15–20], CaD [13,21], CoF [23–28,85–87], CrC [34], CrCl [13,88], CrF [32,88], InBr, InCl, InF [33], LiBe [36], LiCa [38–41,89,90], LiCs [42–44,91], LiK [92], LiMg [48], LiNa [93], MgCa [50,51], MgH [52], NaK [94], NaRb [56–58], SrI [13,66–71]) were gathered, including molecules containing deuterium (D) and tritium (T). Most references are experimental studies, and some are theoretical studies. Theoretical studies usually compare their results with previously published experimental results, helping to find several experimental studies regarding molecules of theoretical and experimental interest. Both experimental and theoretical studies that led us to them are cited in this work.

Experimental data from cited books, papers, or online accessible databases could contain errors. Therefore, a careful reading of the text describing the raw data is required. For instance, footnotes from the book of Huber and Herzberg helped to curate the data and initiate a search to find more recent experimental results. In this search, we encountered cases in which different experimental values were reported for the spectroscopic constants of the same molecule. For example, in the case of AuF, we found two values for ω_e: Saenger et al. reported $\omega_e = 560$ cm^{-1} in 1992 [20], while Andreev et al. reported $\omega_e = 448$ cm^{-1} in 2000 [16]. Similarly, for 30 years, there was a discrepancy in the value of ω_e of ZnBr. Gosavi et al. reported $\omega_e \approx 319$ cm^{-1} in 1971 [95]. Next, Givan et al. reported $\omega_e \approx 198$ cm^{-1} in 1982 [96]. Finally, Burton et al. experimentally reported $\omega_e = 284$ cm^{-1} in 2019 [75]. After finding these situations, we designed a strategy based on looking into theoretical studies to gain insight into the most probable value of a given spectroscopic constant. As a result, the data is clean to the best of the author's knowledge.

During the development of this work, we have realized that, historically, uncertainties about the dissociation energy experimental values had restrained the development of data-based studies concerning D_0 and have led several authors to focus their efforts on the ω_e and R_e due to the availability of [97–99]. The database of spectroscopic constants of diatomic molecules (DSCDM) is primarily collected from Huber and Herzberg's constants of diatomic molecules, first published in 1979 [13]. However, unlike experimental values of R_e and ω_e, since 1980, a significant number of D_0 values have been updated [100]. To name a few, MgD, MgBr, MgO, CaCl. CaO, SrI, SrO, TiS, NbO, AgF, AgBr, and BrF all have their experimental values updated with at least ±2.3 kcal/mol difference from their values in Huber and Herzberg [13,100]. Moreover, the uncertainties in D_0 experimental values are not within chemical accuracy for some molecules. For instance, MgH, CaCl, CaO, CaS, SrH, BaO, BaS, ScF, Tif, NbO, and BrF have uncertainties ranging from ±1 kcal/mol up to ±8 kcal/mol [100]. The authors are working in a parallel effort to update the dissociation energies of diatomic molecules, in which the active thermochemical database [101] could be very relevant, and it will be published elsewhere.
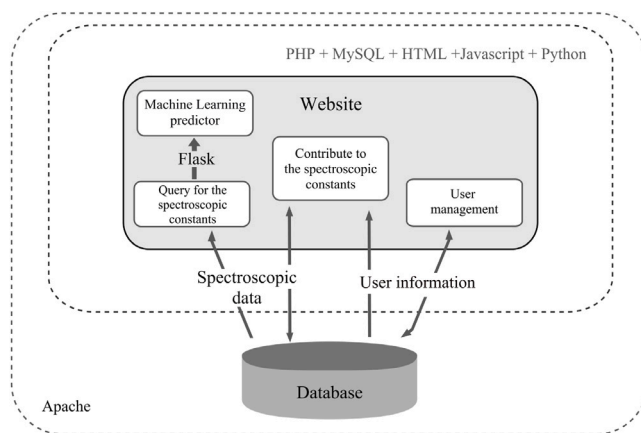

Fig. 1. The infrastructure of the website.

3. Database structure

The structure of the database is centered around the molecule table, which uses a non-null integer "mol_id" as the primary key for identifying individual molecules. The name and electronic state of each molecule are stored as variable character strings in the "molecule_name" and "electronic_state" columns, respectively. Each molecule's reduced mass is saved as a floating point number under the "reduced_mass" column. The complete data structure is shown in Table 1.

The spectroscopic constants (T_e, ω_e, $\omega_e x_e$, B_e, α_e, D_e, R_e, D_0, IP) can either be floating point values or NULL. Information about the data source is also stored in the database, including the original reference as a text string in the reference column and the reference date stored as a variable character string in "date_of_reference". Hence, the user can check the data accuracy and experimental techniques employed.

4. Website structure

The web platform is powered by Red Hat Enterprise Linux 9.0, Apache HTTP Server, MySQL, PHP, and Flask framework for back-end operations. The front-end development employs Hypertext Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript, facilitating the creation of responsive and user-friendly websites with modern interfaces. The infrastructure of the website is shown in Fig. 1. The platform is designed to adapt to the user's device screen size, optimizing the layout accordingly. On larger screens (max-width is larger than 868px), all features, including tooltips and detailed diagrams designed for larger screens (Fig. 4) are displayed to provide comprehensive information. Conversely, larger icons and simpler layouts are employed on smaller screens like mobile devices to enhance user experience.

HTML, PHT, and JavaScript power the dynamic elements of the website. In addition, critical features like user authentication and file uploads are implemented using PHP to ensure robust database security.
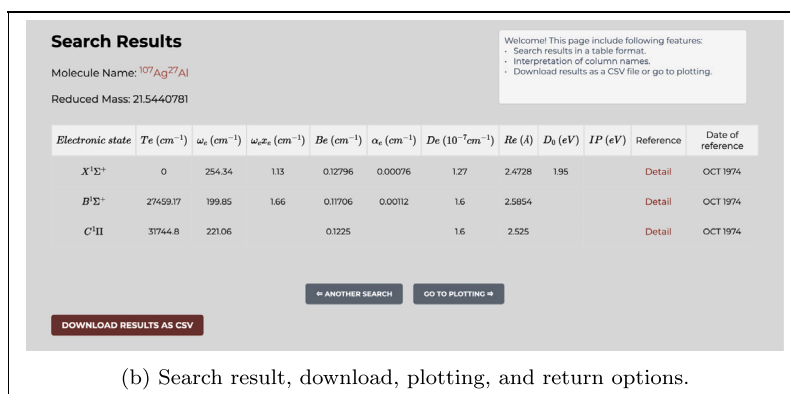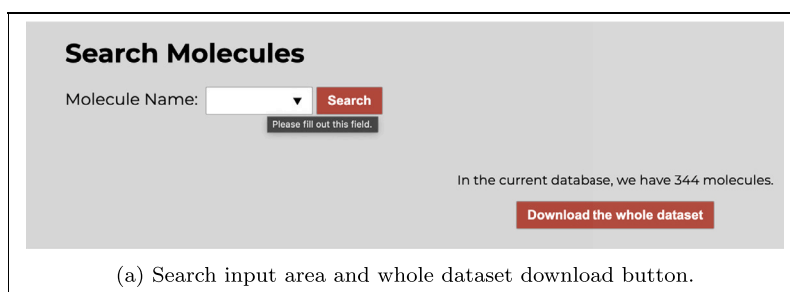
4.1. Search

The website's search function includes two pages: a homepage for entering the search query and a result page to display the retrieved spectroscopic constants. In Fig. 2, a zoomed-in screenshot of the search home page and result page shows the main features. Users can enter a molecule name manually or select from the drop-down list. Before processing the user's input, the code checks whether the entered molecule name meets a specific pattern since the chemical formula of molecules is case-sensitive. The corresponding result will only be retrieved if the

Table 1
Main database structure.

Parameter	Units	Identifier	Description
Molecule name	–	molecule_name	Name of the molecule
Electronic state	–	electronic_state	Electronic state
Molecule ID	–	mol_id	Unique integer identifying the molecule
Reduced mass	amu	reduced_mass	Reduced mass
T_e	cm^{-1}	Te	Minimum electronic energy
ω_e	cm^{-1}	omega_e	Vibrational harmonic frequency
$\omega_e x_e$	cm^{-1}	omega_ex_e	First anharmonic correction
B_e	cm^{-1}	Be	Equilibrium rotational constant
α_e	cm^{-1}	alpha_e	First correction of the rotational constant
D_e	cm^{-1}	De	Centrifugal distortion constant
R_e	Å	Re	Equilibrium internuclear distance
D_0	eV	D0	Binding energy
IP	eV	IP	Ionization potential
Reference	–	reference	Reference of the data
Date of reference	–	date_of_reference	Date of the reference

(a) Search input area and whole dataset download button.

(b) Search result, download, plotting, and return options.

**Fig. 2.** Screenshots of the database search page.

input molecule name consists of all characters in the correct order, including uppercase and lowercase letters, as shown in Fig. 3.

To enhance security, the website employs additional precautions:

- Utilizing prepared statements for SQL query execution helps prevent SQL injection attacks.
- Implementing regular expressions to validate user input aids in thwarting potential threats like Cross-Site Scripting (XSS).

These measures contribute to a robust, efficient, and user-friendly searchable database interface.

The search page presents a statistical analysis of the whole database. The results are presented as a pie chart or a bar plot, as shown in Fig. 4. Panel (a) displays the number of molecules with $^1\Sigma^+$ ground state as an example. Indeed, the list of molecules is updated once the user hovers over a different slice of the pie chart. A bar plot is shown in panel (b) of the same Figure, where molecules are classified based on the value of Λ of the ground electronic state. This information helps identify what kind of molecules should be further included in the database to have a better sample of the diatomics. For example, having more Φ ground
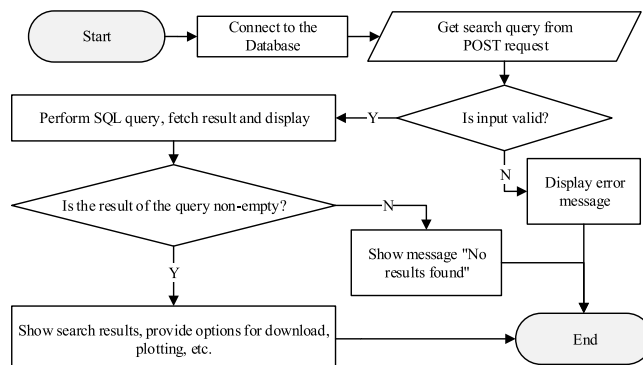
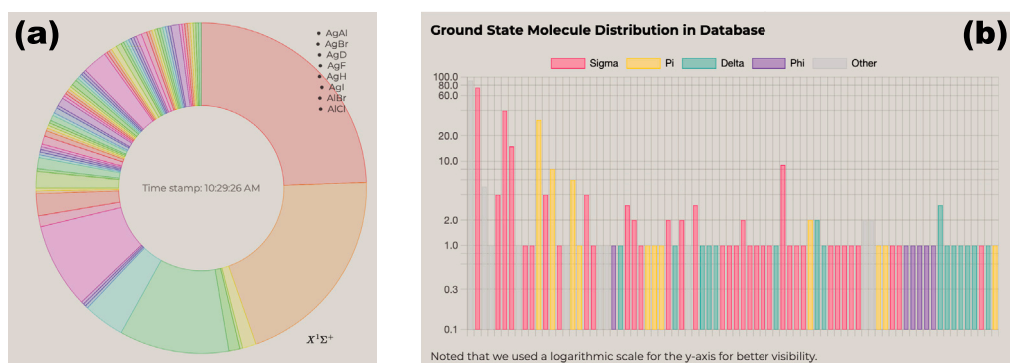**Fig. 3.** Flowchart of the search_php.

Fig. 4. Statistics section of the search page. These two graphs employ the electronic state properties of the diatomic molecules to classify the database.

state molecules may help to understand the spectroscopy of diatomics since we only count on a handful of them.

4.2. API

Our server provides an Application Programming Interface (API) for molecular spectroscopy data retrieval. It offers an efficient, direct, and secure method for accessing the spectroscopic constants of various molecules. The page is well structured with sufficient guidance that introduces user query commands. Users can obtain JSON output with their desired schema using three commends and setting two parameters. Furthermore, users can click the copy button of any commands and paste them into the query box.

General scheme

- **accessed**: An ISO 8601 timestamp indicating when the data was accessed.
- **n_records**: A numerical representation indicating the number of records retrieved.
- **data**: An array containing the actual data records, with each record being a JSON object. The structure of the objects within this array depends on the query made.

List molecules endpoint

- **id_molecule**: The unique identifier of a molecule.
- **chemical_formula**: The chemical formula of the molecule.

Here is the output format:

```
{
    "accessed": "ISO-8601-DATE",
    "n_records": INTEGER,
    "data": [
        {
            "id_molecule": INTEGER,
            "chemical_formula": "MOLECULE_NAME"
        },
        ...
    ]
}
```

Retrieve all spectroscopy constants endpoint

- **chemical_formula**: The molecule's chemical formula.
- **state**: The electronic state of the molecule.
- Multiple keys representing different spectroscopy constants, such as Te, ω_e, $\omega_e x_e$, etc.

Here is the output format:

```
{
    "accessed": "ISO-8601-DATE",
    "n_records": INTEGER,
    "data": [
        {
            "chemical_formula": "MOLECULE_NAME",
            "state": "STATE_NAME",
            "Te": FLOATING_POINT_OR_NULL,
            "omega_e": FLOATING_POINT_OR_NULL,
            ...
        },
        ...
    ]
}
```

Retrieve specific spectroscopy constant endpoint

- **id_molecule**: Unique identifier.
- **chemical_formula**: Chemical formula of the molecule.
- **reference**: Reference source for the spectroscopy constant.
- **reference_date**: Date associated with the reference.
- **state**: Electronic state.
- **mass**: Reduced mass of the molecule.
- **name_of_spectroscopy_constant**: Name of the specific spectroscopy constant requested.
- **value_of_spectroscopy_constant**: Value of the specific spectroscopy constant.

Here is the output format:

```
{
    "accessed": "ISO-8601-DATE",
    "n_records": INTEGER,
    "data": [
        {
            "id_molecule": INTEGER,
            "chemical_formula": "MOLECULE_NAME",
            "reference": "REFERENCE",
            "reference_date": "DATE",
            "state": "STATE_NAME",
            "mass": FLOATING_POINT_OR_NULL,
            "name_of_spectroscopy_constant": "
                ↪ CONSTANT_NAME",
            "value_of_spectroscopy_constant":
                ↪ FLOATING_POINT_OR_NULL
        },
        ...
    ]
}
```
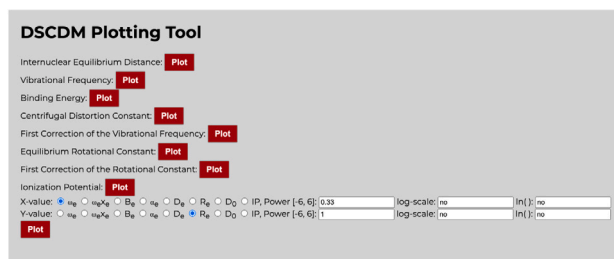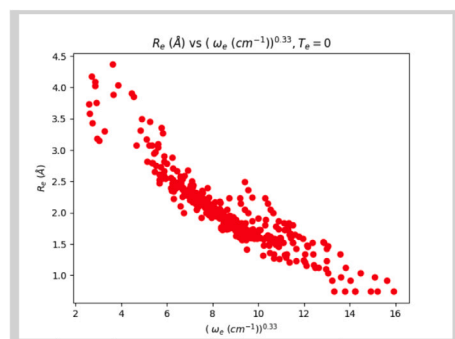
Fig. 5. User interface for our plotting tool.

Fig. 6. Output of the plotting tool. The input to generate this specific plot is shown in Fig. 5.

Error handling

- **error**: A boolean value set to true.
- **message**: An error message detailing the reason for the failure.

Here is the output format:

```
{
    "error": true,
    "message": "ERROR_MESSAGE"
}
```

The program handles different queries with three API endpoints. These endpoints have been designed to accommodate user-specific needs, from fetching a comprehensive list of available molecules to detailed constant-specific information. Here are the five possible combinations of queries and the results:

1. List All Molecules:
   Users specify the query type as `list_molecules`:

   ```
   { "query": "list_molecules" }
   ```

   The output provides a list containing all molecules available in the database.

2. Get Spectroscopy Constants for All Molecules:
   By sending the query `get_all_spectroscopy_constants` without a specific `molecule_name`:

   ```
   { "query": "
     ↪ get_all_spectroscopy_constants" }
   ```

   The output provides all spectroscopy constants for all listed molecules.

3. Get Spectroscopy Constants for a Specific Molecule:
   The query type `get_all_spectroscopy_constants`, along with a specified `molecule_name`:

   ```
   { "query": "
     ↪ get_all_spectroscopy_constants",
     ↪ "molecule_name": "
     ↪ molecule_name_here" }
   ```

   The output provides all spectroscopy constants related to the specified molecule.

4. Get a Particular Constant's Value for All Molecules:
   Users can specify the query
   `get_specific_spectroscopy_constant` with a defined constant but without the `molecule_name`:

   ```
   { "query": "
     ↪ get_specific_spectroscopy_constant
     ↪ ", "constant": "constant_name_here"
     ↪ }
   ```

   The output provides only the specified constant data of all molecules.

5. Get a Particular Constant's Value for a Specific Molecule:
   The user specify both the `constant` and `molecule_name`:

   ```
   { "query": "
     ↪ get_specific_spectroscopy_constant
     ↪ ", "constant": "constant_name_here
     ↪ ", "molecule_name": "
     ↪ molecule_name_here" }
   ```

   The response provides the value of the mentioned constant for the given molecule.

4.3. Plotting tool

The website provides a plotting tool to quickly visualize the spectroscopic constants of the ground electronic state of diatomic molecules ($T_e = 0$ cm^{-1}) available in our database. The plots provided by our tool are generated using the Python package Matplotlib [102]. There are two options for data visualization using this tool:

- By clicking the "Plot" buttons next to the menu attached to each of the spectroscopic constants, as displayed in Fig. 5, users can request a histogram plot, along with the mean and standard deviation, of a particular spectroscopic constant (R_e, ω_e, $_0$, D_e, $\omega_e x_e$, B_e, α_e, and IP). These histograms show the distribution of the spectroscopic constants, which may be helpful for those not acquainted with these constants and their typical values.
- Users can request scatter plots between particular functions of two spectroscopic constants–x^α, with $\alpha \in [-6, 6]$ and $\ln(x)$, displaying their relationship in log–log, log–linear, or linear–linear scale. This tool may be helpful to those who wish to investigate the possible relationships between the spectroscopic constants. To use this option, the user needs to click one of the possible X-values and select the function to represent. Then, a possible Y-value and a function must be chosen, and finally, push the button plot underneath, as displayed in Fig. 5. An example output using $\omega_e^{1/3}$ as the X-value and R_e for the Y-value is shown in Fig. 6. The results are displayed in linear scale as default. However, it is possible to change this by typing yes on the pertinent box representing the X or Y axis scale (see Fig. 5), as the user requires.

4.4. Machine learning tool

The website provides predictors for the internuclear equilibrium separation R_e, harmonic frequency ω_e, binding energy D_0, and the first anharmonic correction $\omega_e x_e$ to the harmonic frequency, as shown in Fig. 7. To retrieve the R_e prediction, the user only needs to type the chemical formula. The same applies to predicting ω_e, although it is necessary to provide the value of R_e. On the contrary, the user must

Fig. 7. Spectroscopic constants predictor page.

provide the chemical formula, R_e, and ω_e for the prediction of D_0 and the first anharmonic correction $\omega_e x_e$.

The featurization of our machine learning models for R_e, ω_e, and D_0 follows Ref. [14]. However, in the present case, apart from Gaussian process regression (GPR) models, we employ neural-networks (NN) and develop a predictor for $\omega_e x_e$.

4.5. Contributions

Users are allowed to upload new molecule data into our database after registration. Users who intend to contribute new data will be redirected to the upload page, as Fig. 8 displays. By completing the form and typing in any additional message, all information will be sent to the administrator for review. Once new data are approved, we will update the main database and machine learning models.

5. Machine learning

GPR is a non-parametric fitting model approach that assumes no prior functional form of the target and learns the mean function and covariance [103], so every predicted result has attached a confidence interval. Specifically, GPR aims to predict target value(s) by learning the relationship between inputs (features) and outputs (targets) using a training data set–GPR learns the probability distribution of the target value(s) conditioned on the features and the training data.

For three of our predictors, R_e, ω_e, and $\omega_e x_e$, we also provide predictions made by neural networks (NN). Neural networks are a popular machine learning model type with vast and varied applications, such as in data-driven regression tasks [104]. NN models have parameters arranged in a series of layers, each with a variable number of neurons, as depicted in Fig. 9. This Figure displays the neural network architecture for the R_e predictor: 21 layers with a variable number of neurons. Each neuron contains a set of parameters to perform a linear transformation on its inputs; the result of this transformation is sent to a non-linear activation function to produce a neuron's output. All of the activations produced by a layer's neurons are sent to the following layer, where the same process is repeated until the final layer is reached and delivers the final output, which, in our case, is the result of the regression.

The featurizations between the two model types (NN and GPR) for a given spectroscopic constant are similar. In the case of NN, some extra adjustments are necessary to ensure permutational invariance of the chemical formula since the location of features in the input vector matters. On the contrary, GPR easily learns to treat permutations,

Table 2
Featurizations for the different models. The rows are the model types, Gaussian process regression (GPR) or neural network (NN), and the columns represent a spectroscopic constant supported by our predictor. g_i and p_i stand for the group and period of the ith atom, respectively, and μ is the reduced mass. "…" denote the features used in learning R_e when employed on the prediction of other spectroscopic constants. Note that there is no NN for predicting D_0 values, and the features are absent.

	R_e	ω_e	$\omega_e x_e$	D_0
GPR	(g_1, g_2, p_1, p_2)	$(\ldots, \mu, R_e)$	$(\ldots, e^{-R_e}, \ln(\omega_e))$	$(\ldots, R_e, \frac{1}{R_e^{12}} - \frac{1}{R_e^6}, \ln(\omega_e))$
NN	$(\frac{g_1+g_2}{2}, \sqrt{g_1 g_2}, \frac{p_1+p_2}{2}, \sqrt{p_1 p_2}, \min(g_1, g_2), \max(g_1, g_2), \min(p_1, p_2), \max(p_1, p_2))$	$(\ldots, R_e, \mu)$	$(\ldots, e^{-R_e}, \mu, \ln(\omega_e))$	

such as HCl versus ClH, equally because it uses the distances between inference points and training points to make predictions.

All machine learning models are trained on the maximum amount of data available in the database for a given task except for GPR predictions for heteronuclear molecules' R_e values, for which a model trained on a subset of heteronuclear molecules performs better. Using all of the relevant data in our database limits the amount of test data available but also provides the models with more information, which generally leads to better predictions. In the case of GPR, our predictors return the mean of the predicted Gaussian distribution with plus or minus one standard deviation.

Featurization and user workflow

The complete set of features used across all seven models in our spectroscopic constant predictor tool is provided in Table 2. The R_e predictor is the simplest in terms of featurization: it only uses atomic properties. It should be noted that since this predictor, as well as all others, use the groups and periods of the constituent atoms as features, the predictors are restricted to making predictions for molecules that do not contain lanthanides or actinides, for the sake of simplicity, although it can be included in more complex models [1,2,14].

We present two models for the R_e predictor. One of these predictions uses GPR, and the other a feed-forward NN, whose structure is displayed in Fig. 9. Due to the simplicity of this model's featurization, it is the first model presented to the user (the top-most one) as shown in Fig. 7. All this model requires is for the user to submit the chemical formula of the diatomic molecule. This model is presented first such that users can use its prediction in later predictors. For example, if one wishes to predict ω_e for a molecule that does not have a measured R_e, they would first need to use the R_e predictor to get a value of R_e for the ω_e predictor.

The second predictor makes inferences for the harmonic frequency ω_e of diatomic molecules based on atomic properties and the molecule's R_e value. All the user needs to enter is the chemical formula of the molecule and its R_e in Angstroms, and two predictions are provided: one made by a GPR model and the other by a feed-forward NN.

The third predictor makes predictions for the binding energy D_0 of a given diatomic molecule. Unlike the previous predictors, this model provides just one prediction made by a GPR model. This model requires the user to input the molecule's chemical formula, its R_e value in Å, and its ω_e value in cm^{-1}.

The fourth and final model of the database's machine-learned spectroscopic constant predictor makes two predictions for $\omega_e x_e$: one from a GPR model and a second from a neural network. To make a prediction for $\omega_e x_e$ the user needs only to input the molecule's chemical formula, its R_e in Å, and its ω_e in cm^{-1}.

Fig. 8. The upload form and message box.
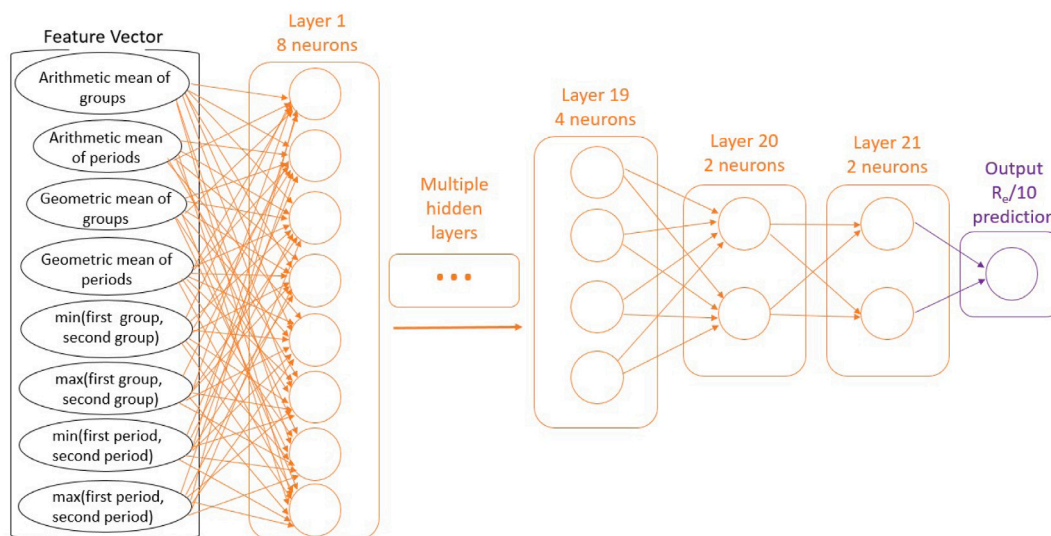
Fig. 9. Neural network architecture. Scheme of the feed-forward neural network model architecture used to make predictions for R_e values.

Model performance

Figs. 10 and 11 show the website predictor's performance on a small number of diatomic molecules not present in the database. The associated root mean square deviation (RMSD) errors are provided in Table 3, showing the model's performance on molecules for which it has no previous data. The small number of test points provided here can be explained by the lack of experimental data or the difficulty of finding molecules not already in the database for which sufficient data is available. The "actual" data for evaluating the machine learning predictions was gathered from the NIST chemistry webbook, Ref. [12], for all spectroscopic constants except D_0. The "actual" D_0 data was gathered from Ref. [105], except for Al_2, taken from Ref. [106].

The molecules used for testing the performance of the R_e models are the following: HCl, HBr, HI, AsH, HSn, KO, OF, Ca_2, and Al_2. For evaluating the ω_e models' performance, all of the molecules mentioned above except HSn are used. For $\omega_e x_e$ models, the following molecules were used: HCl, HBr, HI, Al_2, OF, and Ca_2, and the results are shown in Figs. 10 and 11, and the associated errors of the models are presented in Table 3. The R_e and ω_e models' results indicate relatively accurate predictions (GPR predictions are more accurate than NN), so users should expect that generated predictions are close to actual measurements for molecules similar to those in our database. On the contrary, the performance of the $\omega_e x_e$ predictor varies. In this case, the NN model for $\omega_e x_e$ performs much better than GPR. Admittedly, the performance of the GPR model for $\omega_e x_e$ is lacking for select test molecules.

The molecules used for the D_0 model evaluation are HCl, HBr, HI, Al_2, Ca_2, KO, and HSn. The results of the model evaluation can be found in Fig. 11, and the associated errors of the model are presented in

Table 3
Root mean square deviations (RMSD) of the different machine learning models' predictions in comparison to available data. All of these evaluation points are done on molecules absent from the database and the training data set. When applicable, both the error results of neural networks (NN) and Gaussian process regression (GPR) are shown.

Model	R_e (Å)	ω_e (cm^{-1})	$\omega_e x_e$ (cm^{-1})	D_0 (eV)
GPR	0.0788	155	9.65	0.510
NN	0.169	79.1	2.62	

Table 3. The D_0 model makes accurate predictions for most molecules. It should be noted that for the prediction of D_0 of HSn, a predicted value of ω_e from our ω_e predictor's NN model had to be used due to the absence of experimental data. However, the D_0 model fails on certain molecules, which we can consider outliers in performance.

These outliers in performance could represent instances in the test data that are too different from training data to expect decent results, specific regions in the feature space where there are faults in the predicting model, errors in the data, or potentially combinations of these. Explanations for faulty model performance on some molecules could include the lack of sufficient and representative training data, as is typical for poor ML models, and insufficient hyperparameter optimization. Additionally, improvement may be found by developing multiple models that are tailored to specific kinds of molecules, such as a homonuclear D_0 model separate from a heteronuclear D_0 one, because the relationship between features and targets may be more apparent to the models when the data is meaningfully separated. However, that separation may be more complex than heteronuclear versus homonuclear molecules [2].
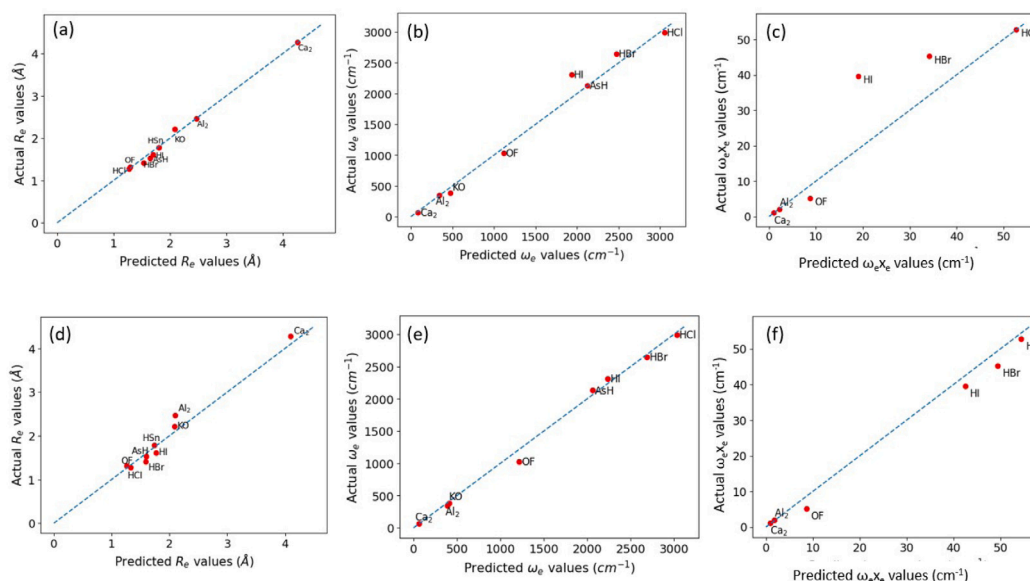
Fig. 10. Machine learning predictor performance. Actual versus predicted data for different spectroscopic constants. The blue dashed line represents the $y = x$ graph, and the data (red dots) would fit on the dashed line for a perfect predictor. Panels (a), (b), and (c) represent prediction data for the GPR models, whereas panels (d), (e), and (f) represent prediction data for the same constants as (a), (b), and (c), respectively, except with results for the NN models.
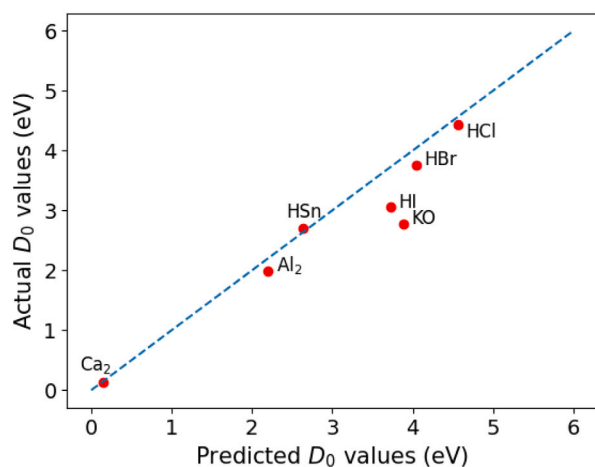

Fig. 11. D_0 from our predictor on a set of molecules not included in the database nor in the model's training data set. The blue, dashed line represents the graph of $y = x$ and a perfect model would place every data point on said line.

6. Conclusions

We have presented "The database of spectroscopic constants of diatomic molecules (DSCDM)", including new capabilities and a more extensive database than others to maximize the user's experience. The database, containing 344 diatomic molecules with spectroscopic constants and 236 with at least one excited electronic state, is the back-end of a modern website. The website shows a general plotting tool, allowing the user to plot all the data in the database (for ground electronic state properties) in different ways. On the other hand, the website shows a spectroscopic constant predictor using GPR and NN machine learning models based on the current database. Despite some prediction flaws on specific molecules, our machine learning predictors offer a user-friendly way to supplement the database by allowing users to generate accurate predictions for select spectroscopic constants when the corresponding experimental data is absent from our database.

Any data from the database can be searched via an API and through the search engine incorporated into the website. Similarly, all the spectroscopic information can be downloaded in the user-friendly .csv format. In addition, users can contribute to the database after a screening

by the website manager. In this way, we aim to deliver a community-driven database.

Finally, it is worth emphasizing that this database contains only 4.9% of the possible diatomic molecules. As a result, the need for more spectroscopic data is evident for a better understanding of diatomic molecules, the basic building block of larger and more relevant chemical complexes.

CRediT authorship contribution statement

Yueqian Wang: Designed the website and contributed to the writing of the manuscript. **Daniel Julian:** Designed the whole ML and plotting sections of the website and contributed to the writing of the manuscript. **Mahmoud A.E. Ibrahim:** Gathered the data, Curated the data, Wrote a part of the manuscript. **Connor Chin:** Designed the login page of the website and the security of the website. **Saketh Bhattiprolu:** Designed the login page of the website and the security of the website. **Ethan Franco:** Contributed to the ML part of the project. **Jesús Pérez-Ríos:** Envisioned the idea, Supervised the project, Wrote the manuscript.

Declaration of competing interest

The authors declare no conflict of interest

Data availability

Data will be made available on request.

References

[1] X. Liu, G. Meijer, J. Pérez-Ríos, On the relationship between spectroscopic constants of diatomic molecules: a machine learning approach, RSC Adv. 11 (2021) 14552–14561, http://dx.doi.org/10.1039/D1RA02061G.

[2] M.A.E. Ibrahim, X. LiU, J. Pérez-Ríos, Spectroscopic constants from atomic properties: a machine learning approach, 2023, arXiv:2308.08933.

[3] X. Liu, G. Meijer, J. Pérez-Ríos, A data-driven approach to determine dipole moments of diatomic molecules, Phys. Chem. Chem. Phys. 22 (2020) 24191–24200, http://dx.doi.org/10.1039/D0CP03810E.

[4] J. Fu, Z. Wan, Z. Yang, L. Liu, Q. Fan, F. Xie, Y. Zhang, J. Ma, Combining ab initio and machine learning method to improve the prediction of diatomic vibrational energies, Int. J. Quantum Chem. 122 (18) (2022) e26953, http://dx.doi.org/10.1002/qua.26953, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.26953. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.26953.

[5] J. Fu, S. Long, J. Jian, Z. Fan, Q. Fan, F. Xie, Y. Zhang, J. Ma, A joint data and model driven method for study diatomic vibrational spectra including dissociation behavior, Spectrochim. Acta A 239 (2020) 118363.

[6] I.C. Stevenson, J. Pérez-Ríos, Genetic based fitting techniques for high precision potential energy curves of diatomic molecules, J. Phys. B: At. Mol. Opt. Phys. 52 (10) (2019) 105002, http://dx.doi.org/10.1088/1361-6455/ab0c4b.

[7] L.K. McKemmish, Molecular diatomic spectroscopy data, WIREs Comput. Mol. Sci. 11 (5) (2021) e1520, http://dx.doi.org/10.1002/wcms.1520, arXiv:https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1520. URL https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1520.

[8] C.P. Endres, S. Schlemmer, P. Schilke, J. Stutzki, H.S. Müller, The cologne database for molecular spectroscopy, cdms, in the virtual atomic and molecular data centre, vamdc, J. Mol. Spectrosc. 327 (2016) 95–104, http://dx.doi.org/10.1016/j.jms.2016.03.005, new Visions of Spectroscopic Databases, Volume II. URL https://www.sciencedirect.com/science/article/pii/S0022285216300340.

[9] F.J. Lovas, E. Tiemann, Microwave spectral tables I. Diatomic molecules, J. Phys. Chem. Ref. Data 3 (3) (2009) 609–770, http://dx.doi.org/10.1063/1.3253146, arXiv:https://pubs.aip.org/aip/jpr/article-pdf/3/3/609/11480100/609_1_online.pdf.

[10] F.J. Lovas, E. Tiemann, J.S. Coursey, S.A. Kotochigova, J. Chang, K. Olsen, R.A. Dragoset, Molecular microwave spectral databases, 2005, [online].

[11] P. Bernath, S. McLeod, Diref, a database of references associated with the spectra of diatomic molecules, J. Mol. Spectrosc. 207 (2) (2001) 287, http://dx.doi.org/10.1006/jmsp.2001.8345, URL https://www.sciencedirect.com/science/article/pii/S0022285201983456.

[12] K. Huber, G. Herzberg, Constants of diatomic molecules (data prepared by jean w. gallagher and russell d. johnson, iii), in: NIST Chemistry WebBook, NIST Standard Reference Database Number 69, 2021.

[13] K.-P. Huber, Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules, Springer Science & Business Media, 2013.

[14] X. Liu, S. Truppe, G. Meijer, J. Pérez-Ríos, The diatomic molecular spectroscopy database, J. Cheminform. 12 (1) (2020) 31, http://dx.doi.org/10.1186/s13321-020-00433-8.

[15] C.J. Evans, M.C. Gerry, Confirmation of the existence of gold (i) fluoride, auf: Microwave spectrum and structure, J. Am. Chem. Soc. 122 (7) (2000) 1560–1561.

[16] S. Andreev, J.J. BelBruno, Detection of auf by emission spectroscopy in a hollow cathode discharge, Chem. Phys. Lett. 329 (5–6) (2000) 490–494.

[17] D. Schröder, J. Hrušák, I.C. Tornieporth-Oetting, T.M. Klapötke, H. Schwarz, Definitiver beweis für die existenz von auf, Angew. Chem. 106 (2) (1994) 223–225.

[18] C. van Wüllen, Molecular density functional calculations in the regular relativistic approximation: Method, application to coinage metal diatomics, hydrides, fluorides and chlorides, and comparison with first-order relativistic calculations, J. Chem. Phys. 109 (2) (1998) 392–399.

[19] D. Figgen, G. Rauhut, M. Dolg, H. Stoll, Energy-consistent pseudopotentials for group 11 and 12 atoms: adjustment to multi-configuration dirac–hartree–fock data, Chem. Phys. 311 (1–2) (2005) 227–244.

[20] K. Saenger, C. Sun, Yellow emission bands produced during gold etching in o 2-cf 4 rf glow-discharge plasmas: Evidence for gas-phase auf, Phys. Rev. A 46 (1) (1992) 670.

[21] E. GharibNezhad, A. Shayesteh, P.F. Bernath, Fourier transform emission spectra of the a2π→ x2σ+ and b2σ+→ x2σ+ transitions of cad, J. Mol. Spectrosc. 281 (2012) 47–50.

[22] T.C. Steimle, T. Ma, A.G. Adam, W.D. Hamilton, A.J. Merer, High resolution laser induced fluorescence spectroscopy of the [18.8] φ i 3- x φ i 3 (0, 0) band of cobalt monofluoride, J. Chem. Phys. 125 (6) (2006) 064302.

[23] Q. Nadhem, S. Behere, S. Behere, Franck-condon factors and r-centroids for several systems of cof molecule, Int. Lett. Chem. Astron. 58 (2015) 91.

[24] R. Ram, P. Bernath, S. Davis, The low-lying electronic states of cof, J. Chem. Phys. 104 (18) (1996) 6949–6955.

[25] A.I. Boldyrev, J. Simons, Periodic Tables of Diatomic Molecules, John Wiley & Sons, 1997.

[26] K.P. Jensen, B.O. Roos, U. Ryde, Performance of density functionals for first row transition metal systems, J. Chem. Phys. 126 (1) (2007) 014103.

[27] H. Wang, X. Zhuang, T.C. Steimle, The permanent electric dipole moments of cobalt monofluoride, cof, and monohydride, coh, J. Chem. Phys. 131 (11) (2009) 114315.

[28] M.M.F. de Moraes, Y.A. Aoto, Multi-reference and multi-occupancy character of the cobalt monofluoride, J. Mol. Spectrosc. 385 (2022) 111611.

[29] M. Barnes, D. Clouthier, P. Hajigeorgiou, G. Huang, C. Kingston, A. Merer, G. Metha, J. Peers, S. Rixon, The electronic spectrum of gaseous coo in the visible region, J. Mol. Spectrosc. 186 (2) (1997) 374–402.

[30] F. Liu, F.-X. Li, P. Armentrout, Guided ion-beam studies of the reactions of co n+(n=2–20) with o 2: Cobalt cluster-oxide and-dioxide bond energies, J. Chem. Phys. 123 (6) (2005) 064304.

[31] S. McLamarrah, P. Sheridan, L. Ziurys, The pure rotational spectrum of coo (x4δi): Identifying the high-spin components, Chem. Phys. Lett. 414 (4–6) (2005) 301–306.

[32] O. Launila, Spectroscopy of crf: Rotational analysis of the a6σ+-x6σ+ band system in the 1-µm region, J. Mol. Spectrosc. 169 (2) (1995) 373–395.

[33] S. Mishra, R.K. Yadav, V. Singh, S. Rai, Spectroscopic studies of diatomic indium halides, J. Phys. Chem. Ref. Data 33 (2) (2004) 453–470.

[34] D.J. Brugh, M.D. Morse, A. Kalemos, A. Mavridis, Electronic spectroscopy and electronic structure of diatomic crc, J. Chem. Phys. 133 (3) (2010) 034303.

[35] M.A. Garcia, C. Vietz, F. Ruipérez, M.D. Morse, I. Infante, Electronic spectroscopy and electronic structure of diatomic irsi, J. Chem. Phys. 138 (15) (2013) 154306.

[36] R. Schlachta, I. Fischer, P. Rosmus, V. Bondybey, The simplest heteronuclear metal cluster: Libe, Chem. Phys. Lett. 170 (5–6) (1990) 485–491.

[37] T.D. Persinger, J. Han, M.C. Heaven, Electronic spectroscopy and photoionization of libe, J. Phys. Chem. A 125 (37) (2021) 8274–8281.

[38] A. Stein, M. Ivanova, A. Pashov, H. Knöckel, E. Tiemann, Spectroscopic study of the 22σ+ and the 42σ+ excited states of lica, J. Chem. Phys. 138 (11) (2013) 114306.

[39] C. Wu, H. Ihle, K. Gingerich, A mass spectrometric study of the alkaline earth diatomic molecules. dissociation energies of mg2, ca2 and cali, Int. J. Mass Spectrom. Ion Phys. 47 (1983) 235–238.

[40] J.V. Pototschnig, R. Meyer, A.W. Hauser, W.E. Ernst, Vibronic transitions in the alkali-metal (li, na, k, rb)–alkaline-earth-metal (ca, sr) series: A systematic analysis of de-excitation mechanisms based on the graphical mapping of frank-condon integrals, Phys. Rev. A 95 (2) (2017) 022501.

[41] G. Krois, J.V. Pototschnig, F. Lackner, W.E. Ernst, Spectroscopy of cold lica molecules formed on helium nanodroplets, J. Phys. Chem. A 117 (50) (2013) 13719–13731.

[42] A. Grochola, J. Szczepkowski, W. Jastrzebski, P. Kowalczyk, Study of the a1σ+ and b3π0 states in lics by a polarization labelling spectroscopy technique, J. Quant. Spectrosc. Radiat. Transfer 145 (2014) 147–152.

[43] N. Mabrouk, H. Berriche, H.B. Ouada, F.X. Gadéa, Theoretical study of the lics molecule: adiabatic and diabatic potential energy and dipole moment, J. Phys. Chem. A 114 (24) (2010) 6657–6668.

[44] P. Staanum, A. Pashov, H. Knöckel, E. Tiemann, X σ+ 1 and a σ+ 3 states of lics studied by fourier-transform spectroscopy, Phys. Rev. A 75 (4) (2007) 042513.

[45] W. Müller, W. Meyer, Ground-state properties of alkali dimers and their cations (including the elements li, na, and k) from abinitio calculations with effective core polarization potentials, J. Chem. Phys. 80 (7) (1984) 3311–3320.

[46] E.J. Breford, F. Engelke, G. Ennen, K.H. Meiwes, Crossed laser and molecular beam studies of mixed alkali dimer: preparation, perturbation and predissociation, Faraday Discuss. Chem. Soc. 71 (1981) 233–252.

[47] K.F. Zmbov, C. Wu, H. Ihle, A mass spectrometric study of heteronuclear diatomic alkali metal molecules. dissociation energies and ionization potentials of nali, kli, and nak, J. Chem. Phys. 67 (10) (1977) 4603–4607.

[48] T.D. Persinger, J. Han, M.C. Heaven, Electronic spectroscopy and photoionization of limg, J. Phys. Chem. A 125 (17) (2021) 3653–3663.

[49] F. Engelke, G. Ennen, K. Meiwes, Laser induced fluorescence spectroscopy of nali in beam and bulk, Chem. Phys. 66 (3) (1982) 391–402.

[50] H. Atmanspacher, H. Scheingraber, C. Vidal, Laser induced fluorescence of the mgca molecule, J. Chem. Phys. 82 (8) (1985) 3491–3501.

[51] G.C. Rizkallah, A.A. Assaf, S.N. Tohme, Molecular structure and properties of mgca molecule, Chem. Phys. 550 (2021) 111316.

[52] L.B. Knight Jr., W. Weltner Jr., Hyperfine interaction and chemical bonding in mgh, cah, srh, and bah molecules, J. Chem. Phys. 54 (9) (1971) 3875–3884.

[53] R.S. DaBell, R.G. Meyer, M.D. Morse, Electronic structure of the 4d transition metal carbides: Dispersed fluorescence spectroscopy of moc, ruc, and pdc, J. Chem. Phys. 114 (7) (2001) 2938–2954.

[54] D.J. Brugh, T.J. Ronningen, M.D. Morse, First spectroscopic investigation of the 4d transition metal monocarbide moc, J. Chem. Phys. 109 (18) (1998) 7851–7862.

[55] S. Leutwyler, M. Hofmann, H.-P. Harri, E. Schumacher, The adiabatic ionization potentials of the alkali dimers na2, nak and k2, Chem. Phys. Lett. 77 (2) (1981) 257–260.

[56] M. Chaieb, H. Habli, L. Mejrissi, B. Oujia, F.X. Gadéa, Ab initio spectroscopic study for the narb molecule in ground and excited states, Int. J. Quantum Chem. 114 (11) (2014) 731–747.

[57] O. Docenko, M. Tamanis, R. Ferber, A. Pashov, H. Knöckel, E. Tiemann, Potential of the ground state of narb, Phys. Rev. A 69 (4) (2004) 042503.

[58] N. Takahashi, H. Katô, Laser-induced fluorescence of the narb molecule, J. Chem. Phys. 75 (9) (1981) 4350–4356.

[59] B. Simard, P.I. Presunka, H.P. Loock, A. Bérces, O. Launila, Laser spectroscopy and density functional calculations on niobium monocarbide, J. Chem. Phys. 107 (2) (1997) 307–318.

[60] J. Ogilvie, F.Y. Wang, Potential-energy functions of diatomic molecules of the noble gases i. like nuclear species, J. Mol. Struct. 273 (1992) 277–290.

[61] D.J. Brugh, M.D. Morse, Resonant two-photon ionization spectroscopy of nic, J. Chem. Phys. 117 (23) (2002) 10703–10714.

[62] R. Ram, P. Bernath, Fourier transform infrared emission spectroscopy of a new a3πi-x3σ- system of nio, J. Mol. Spectrosc. 155 (2) (1992) 315–325.

[63] R. Ram, S. Yu, I. Gordon, P. Bernath, Fourier transform infrared emission spectroscopy of new systems of nis, J. Mol. Spectrosc. 258 (1–2) (2009) 20–25.

[64] C.J. Evans, L.-M.E. Needham, N.R. Walker, H. Köckert, D.P. Zaleski, S.L. Stephens, The pure rotational spectra of the open-shell diatomic molecules pbi and sni, J. Chem. Phys. 143 (24) (2015) 244309.

[65] J.D. Langenberg, L. Shao, M.D. Morse, Resonant two-photon ionization spectroscopy of jet-cooled pdc, J. Chem. Phys. 111 (9) (1999) 4077–4086.

[66] J. Schroeder, C. Nitsch, W. Ernst, Polarization spectroscopy of sri in a heat pipe: The b2ε+-x2ε+(0, 0) system, ChemInform 20 (10) (1989) no.

[67] W. Ernst, J. Schröder, B. Zeller, Rotational analysis and deperturbation of the sri a2π-x2σ+(0, 0) system, J. Mol. Spectrosc. 135 (1) (1989) 161–168.

[68] S. Antrobus, D. Husain, J. Lei, F. Castaño, M.S. Rayo, Kinetic study of the molecular chemiluminescence sri (a2ii1/2, 3/2, b2σ+$\rightarrow$ x2σ+) and the atomic resonance fluorescence sr (53p1$\rightarrow$ 51so) following the pulsed dye laser generation of sr (53pj) in the presence of ch3i, Z. Phys. Chem. 190 (2) (1995) 267–287.

[69] A. Bernard, C. Effantin, J. d'Incan, A. Topouzkhanian, G. Wannous, Laser-excited fluorescence spectra of strontium monoiodide, J. Mol. Spectrosc. 195 (1) (1999) 11–21.

[70] M.W. Chase, N.I.S.O. (US), NIST-JANAF Thermochemical Tables, Vol. 9, American Chemical Society, Washington, DC, 1998.

[71] V. Belyaev, I. Gotkis, N. Lebedeva, K. Krasnov, Ionization potentials of mx molecules (m=ca, sr, ba; x=f, cl, br, i, oh, o), Russ. J. Phys. Chem. 64 (1990) 773.

[72] T. Imajo, Y. Kobayashi, Y. Nakashima, K. Tanaka, T. Tanaka, Fourier transform emission spectroscopy of the tif radical in the 407 nm region, J. Mol. Spectrosc. 230 (2) (2005) 139–148.

[73] R. Ram, P. Bernath, Fourier transform emission spectroscopy of the f4δ–x4ϕ system of tif, J. Mol. Spectrosc. 231 (2) (2005) 165–170.

[74] R. Ram, J. Peers, Y. Teng, A. Adam, A. Muntianu, P. Bernath, S. Davis, Laser and fourier transform emission spectroscopy of theg4ϕ–x4ϕ system of tif, J. Mol. Spectrosc. 184 (1) (1997) 186–201.

[75] M. Burton, L. Ziurys, The pure rotational spectrum of the znbr radical (x2σ+): Trends in the zinc halide series, J. Chem. Phys. 150 (3) (2019) 034303.

[76] I.O. Antonov, M.C. Heaven, Spectroscopic and theoretical investigations of uf and uf+, J. Phys. Chem. A 117 (39) (2013) 9684–9694.

[77] R. Ram, P. Bernath, S. Davis, Infrared emission spectroscopy of the [10.5] 5 δ–x 5 δ system of vf, J. Chem. Phys. 116 (16) (2002) 7035–7039.

[78] L.A. Kaledin, J.E. McCord, M.C. Heaven, Laser spectroscopy of uo: Characterization and assignment of states in the 0-to 3-ev range, with a comparison to the electronic structure of tho, J. Mol. Spectrosc. 164 (1) (1994) 27–65.

[79] B. Simard, P.A. Hackett, W.J. Balfour, Jet-cooled optical spectroscopy of yttrium monocarbide: Evidence for a 4πi ground state, Chem. Phys. Lett. 230 (1–2) (1994) 103–109.

[80] I. Shim, M. Pelino, K.A. Gingerich, Electronic states and nature of bonding in the molecule yc by all electron ab initio multiconfiguration self-consistent-field calculations and mass spectrometric equilibrium experiments, J. Chem. Phys. 97 (12) (1992) 9240–9248.

[81] M. Sievers, Y.-M. Chen, P. Armentrout, Metal oxide and carbide thermochemistry of y+, zr+, nb+, and mo+, J. Chem. Phys. 105 (15) (1996) 6322–6333.

[82] M. Bobetic, J. Barker, Vibrational levels of heteronuclear rare gas van der waals molecules, J. Chem. Phys. 64 (6) (1976) 2367–2369.

[83] K. Tang, J. Toennies, The van der waals potentials between all the rare gas atoms from he to rn, J. Chem. Phys. 118 (11) (2003) 4976–4983.

[84] L. Piticco, F. Merkt, A.A. Cholewinski, F.R. McCourt, R.J. Le Roy, Rovibrational structure and potential energy function of the x0+ ground electronic state of arxe, J. Mol. Spectrosc. 264 (2) (2010) 83–93.

[85] R. Ram, P. Bernath, S. Davis, Fourier transform emission spectroscopy of the [10.3] 3ϕi-x3ϕi system of cof, J. Mol. Spectrosc. 173 (1) (1995) 158–176.

[86] X. Zhang, J. Guo, T. Wang, L. Pei, Y. Chen, C. Chen, Visible laser spectroscopy of cobalt monofluoride, J. Mol. Spectrosc. 220 (2) (2003) 209–213.

[87] A. Adam, W. Hamilton, A hund's case (a) analysis of the [18.8] 3ϕi–x3ϕi electronic transition of cof, J. Mol. Spectrosc. 206 (2) (2001) 139–142.

[88] M. Bencheikh, R. Koivisto, O. Launila, J. Flament, The low-lying electronic states of crf and crcl: Analysis of the a 6 σ+$\rightarrow$ x 6 σ+ system of crcl, J. Chem. Phys. 106 (15) (1997) 6231–6239.

[89] J. Gerschmann, High-resolution spectroscopy of the alkali-earth-alkaline molecules kca and lica, 2021.

[90] L. Russon, G. Rothschopf, M. Morse, A. Boldyrev, J. Simons, Two-photon ionization spectroscopy and all-electron ab initio study of lica, J. Chem. Phys. 109 (16) (1998) 6655–6665.

[91] A. Stein, A. Pashov, P. Staanum, H. Knöckel, E. Tiemann, The b 1 π and d 1 π states of lics studied by fourier-transform spectroscopy, Eur. Phys. J. D 48 (2008) 177–185.

[92] A. Grochola, J. Szczepkowski, W. Jastrzebski, P. Kowalczyk, The a1σ+ electronic state of kli molecule, Chem. Phys. Lett. 535 (2012) 17–20.

[93] P. Neogrády, P.G. Szalay, W.P. Kraemer, M. Urban, Coupled-cluster study of spectroscopic constants of the alkali metal diatomics: Ground and the singlet excited states of na 2, nali, nak, and narb, Collect. Czechoslov. Chem. Commun. 70 (7) (2005) 951–978.

[94] A. Ross, P. Crozet, I. Russier-Antoine, A. Grochola, P. Kowalczyk, W. Jastrzebski, P. Kortyka, On the c1σ+ state of nak, J. Mol. Spectrosc. 226 (1) (2004) 95–102.

[95] R. Gosavi, G. Greig, P. Young, O. Strausz, Reactions of metal atoms. iv. the uv spectra of cdbr, cdi, znbr, and zni, J. Chem. Phys. 54 (3) (1971) 983–991.

[96] A. Givan, A. Loewenschuss, A matrix-isolation raman study of the mx2+ zn$\rightarrow$ mx+ znx reaction: the vibrational spectra of zncl, znbr, hgcl and hgbr, J. Mol. Struct. 78 (3–4) (1982) 299–301.

[97] R.M. Badger, A relation between internuclear distances and bond force constants, J. Chem. Phys. 2 (3) (1934) 128–131.

[98] G.B. Sutherland, The determination of internuclear distances and of dissociation energies from force constants, J. Chem. Phys. 8 (2) (1940) 161–164.

[99] K.S. Jhung, I.H. Kim, K.-H. Oh, K.B. Hahn, K.H.C. Jhung, Universal nature of diatomic potentials, Phys. Rev. A 42 (11) (1990) 6497.

[100] Y.-R. Luo, Comprehensive Handbook of Chemical Bond Energies, CRC Press, 2007.

[101] B. Ruscic, R.E. Pinzon, M.L. Morton, G. von Laszevski, S.J. Bittner, S.G. Nijsure, K.A. Amin, M. Minkoff, A.F. Wagner, Introduction to active thermochemical tables: Several keyenthalpies of formation revisited, J. Phys. Chem. A 108 (45) (2004) 9979–9997, http://dx.doi.org/10.1021/jp047912y.

[102] J.D. Hunter, Matplotlib: A 2d graphics environment, Comput. Sci. Eng. 9 (3) (2007) 90–95, http://dx.doi.org/10.1109/MCSE.2007.55.

[103] C.K. Williams, C.E. Rasmussen, Gaussian Processes for Machine Learning, Vol. 2, MIT Press, Cambridge, MA, 2006.

[104] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: A survey, Heliyon 4 (11) (2018) e00938, http://dx.doi.org/10.1016/j.heliyon.2018.e00938, URL https://www.sciencedirect.com/science/article/pii/S2405844018332067.

[105] Y.-R. Luo, Comprehensive Handbook of Chemical Bond Energies, CRC Press, 2007.

[106] B. Rosen, Spectroscopic Data Relative to Diatomic Molecules, Pergamon Press, 1970.