



EPN-2024-RI

EUROPLANET2024 Research Infrastructure

H2020-INFRAIA-2018-2020

Grant agreement no: 871149

Milestone MS83 Final SSHADE Band List Database

Due date of deliverable: 31/05/2023

Actual submission date: 31/05/2023

Start date of project: 01 February 2020

Duration: 48 months

Responsible WP Leader: Observatory of Paris, Stephane Erard

| Project funded by the European Union's Horizon 2020 research and innovation programme | | |
|---|---|---|
| Dissemination level | | |
| PU | Public | X |
| PP | Restricted to other programme participants (including the Commission Service) | |
| RE | Restricted to a group specified by the consortium (including the Commission Services) | |
| CO | Confidential, only for members of the consortium (excluding the Commission Services) | |

| | |
|-------------------------|--|
| Project Number | 871149 |
| Project Title | EPN2024 - RI |
| Project Duration | 54 months: 01 February 2020 – 31 July 2024 |

| | |
|--------------------------------------|--|
| Milestone Number | MS83 |
| Contractual Delivery date | 31.05.2023 |
| Actual delivery date | 31.05.2023 |
| Title of Deliverable | Final SSHADE Band List Database |
| Contributing Work package (s) | WP 6 |
| Dissemination level | Public |
| Author (s) | Bernard Schmitt, Damien Albert, Manon Furrer, Philippe Bollard |

Abstract: “SSHADE Band List” (BL) is a common database of the SSHADE infrastructure in experimental solid spectroscopy that complement and supplement the set of spectral databases of the SSHADE consortium, a VESPA sub-network of about 25 European contributor groups (11 countries) + 5 non-European (IN, TW, CN, BR, CA). This new band list database started to put online a large set of ‘list of bands’ (absorption, Raman scattering) of ices, minerals, and other simple solids for the planetary sciences and astronomical/astrophysical user communities to help interpret VUV to far-IR spectroscopic observations of planets and small body surfaces, aerosols and grains. In particular, it will allow to easily search the ice or mineral that produce an unknown absorption band in an astronomical spectrum just by providing its position and optionally other known band/physical/chemical constraints. It will then provide the atom/molecule/chemical function or bond and the phase of the solid constituent at the origin of the absorption and list all other bands that should be also present.

The SSHADE-BL interface (<https://www.sshade.eu/>) with its supporting relational Solid Spectroscopy Data Model (SSDM-BL) has been developed in order to store, search, display and export a large number of BL data sets. It will also allow to develop an EPN/TAP access for use in a VO context, in particular with the VESPA VO portal (based on this final version).

The SSDM-BL data model partly inherited from SSHADE-Spectra was greatly expanded to describe the various parameters of a band list and its bands (position, width, intensities, errors, mode attribution, environment, references, ...). The development of the infrastructure of SSHADE-BL aimed to design and build the tools dedicated to data providers (data preparation, import and management, data search, and visualization) and to users (data search, visualization and export). An automatic DOI generator has been also set up to provide a data reference and a direct link to the cited data. At the time of this delivery the SSHADE-BL database has been already filled with 42 band lists from 24 different ice phases and 7 minerals, with a total of more than 1100 bands. The capacity of the SSHADE-BL to provide a world unique set of spectroscopy band lists of solids is thus confirmed.

Table of Contents

| | | |
|-----|---|----|
| 1 | Introduction and aims..... | 3 |
| 1.1 | The SSHADE project | 3 |
| 1.2 | The SSHADE Band List..... | 5 |
| 3 | The band List Database..... | 7 |
| 3.1 | Technical choices | 7 |
| 3.2 | Server infrastructure | 7 |
| 4 | Development of tools for data providers..... | 8 |
| 4.1 | Data ingestion tools | 8 |
| 4.2 | Band list and band searches and visualization..... | 9 |
| 4.3 | Band List review Excel table..... | 10 |
| 4.4 | Band List XML generator..... | 10 |
| 5 | Development of the interface for Users | 11 |
| 5.1 | Data search for users..... | 11 |
| 5.2 | Data visualization | 14 |
| 5.3 | Data export | 17 |
| 6 | DOI generator for bandlists..... | 19 |
| 7 | Band List Database filling..... | 19 |
| 8 | Milestone..... | 22 |
| 9 | Future developments of the SSHADE-BL database | 22 |
| 9.1 | Future SSDM-BL developments | 22 |
| 9.2 | Future SSHADE-BL interface and tool developments..... | 23 |
| 9.3 | Future SSHADE-BL content | 23 |

1 Introduction and aims

1.1 The SSHADE project

SSHADE (lead: IPAG, Consortium of contributor partners: <https://wiki.sshade.eu/doku.php?id=sshade:databases>) is a multi-databases infrastructure in experimental solid spectroscopy for a VESPA sub-network of European and non-European contributor groups, currently from 30 different laboratories in 16 different countries. The spectral databases of the SSHADE consortium are now complemented and supplemented by a common “Band List Database” that put online a world-unique set of ‘list of bands’ (absorption, Raman scattering) of ices, minerals, and other simple organic or inorganic solids (or liquids) for the planetary sciences and astronomical/astrophysical user communities to help interpret VUV to far-IR spectroscopic observations of planets and small body surfaces, aerosols and grains.

Our **definition** of a band list of a solid (or liquid):

It is a list of band parameters and vibration modes of:

- a ‘primary’ molecule in a simple molecular constituent (2-3 species maxi)
- or a mineral, or a ionic/covalent solid

with a defined phase (amorphous or crystalline) and composition (fixed or over a small range)

A band list includes:

- the bands of all the isotopes (sub-band lists)
- the band parameters for different environment conditions (T, P, ...)

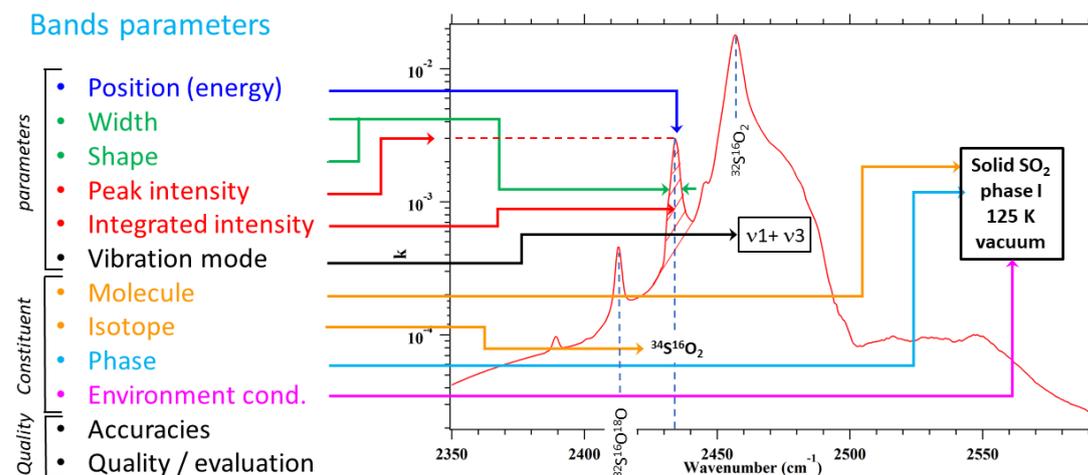


Figure 1. The parameters of the bands, the constituent, and the environment

The **SSHAD** Band lists:

- provide the various parameters of a band list (solid composition, primary specie, temperature, pressure, excitation for Raman, sample and polarization orientation, ...) and its bands (position, width, shape, overlap mode, intensities, errors, mode attribution, ...)
- provide a complete set of data references and links to spectra for each band
- can be of different types: currently ‘absorption band list’ or ‘Raman scattering band list’
- concern several types of simple solids, mixtures and compounds: ices and molecular solids, minerals, inorganic solids, simple organic solids.
- is fed from a critical review of all, or most of, the available literature on experimental spectroscopy of solids, complemented by the detailed analysis of SSHAD original data provided by the consortium partners
- cover a wide range of wavelengths: currently from VUV to far-IR

The resulting service started to put online a large set of critically reviewed band lists of ices, minerals, and other organic and inorganic solids for the astronomical/astrophysical user community (and well beyond: geology, glaciology, material sciences, optical design, chemistry, ...) as reference data to identify absorption or emission bands from solids observed in various astrophysical environments or in laboratory simulations.

In particular, it starts to allow to easily search the ice or mineral that produce an unknown absorption band in an astronomical spectrum just by providing its position and optionally other known or suspected band/physical/chemical constraints. It thus provides the atom/molecule/chemical function or bond and the phase of the solid constituent at the origin of the absorption and list all other bands that should be also present. The only limitation is the current number of band lists and bands available, the search interface being already very evolved.

When the EPN/TAP access for use in a VO context will be developed (starting soon version), it will be a powerful tool for query systems (VESPA portal and alternative accesses), and in particular for workflow platforms such as OPUS that allow for mass processing.

It will help interpret spectroscopic observations of planets and small bodies surfaces, aerosols and grains, in particular from spacecraft missions with spectrometers and hyperspectral imagers on board, in terms of composition and environment of the constituting ices or minerals (composition, isotopes, mixing, phase, T, P, ...). It will also help select the best spectra to compare with observations, or to use in radiative transfer models (direct link between SSHADE band list database and spectra in the SSHADE spectral databases, as well as in some other external relevant databases).

1.2 The SSHADE Band List

The developed SSHADE-BL interface includes tools for band list and bands search, visualization and export for the users. It also includes tools for data preparation, import, search and visualization for the data providers in order to allow them to efficiently prepare, import and validate data. SSHADE-BL was developed to get the capacity to deal with a large number of datasets.

To do so, an extended data model, adapted to the description of band list and its band parameters, had been first developed building on some parts of the existing SSDM-Spectra data model. Then on this data model a database, complemented by a series of tools, was designed and implemented using the state of the art of available free software and technologies. Hardware at the "OSUG-Data Center" (OSUG-DC) of the "Observatoire des Sciences de l'Univers de Grenoble, CNRS-INSU/UGA" has been used in order to provide a flexible development platform and an efficient access for the data provider and managers, and the users. Finally, a wide range of band list datasets (over 40 band lists and 1100 bands, as for 30th May 2023) have been already prepared and imported in the SSHADE-BL database. This allows us to demonstrate the ability of SSHADE-BL to provide access to a world unique set of band lists of solids.

This phase of intensive literature review, data analysis and validation (to ensure consistency and data quality, but also to improve the completeness of the published data) and integration of band list data in SSHADE-BL was carried out by the SSHADE manager.

2 SSDM-BL data model

| | |
|---------|--|
| Authors | Development: Bernard Schmitt (<i>scientific manager</i>), Damien Albert, Philippe Bollard, Olivier Poch, Lydie Bonal |
|---------|--|

Abstract: This part presents the development of the SSDM-BL data model needed to support the SSHADE-BL database.

The relational Solid Spectroscopy Data Model, SSDM-BL (currently version 0.9.2a) partly inherited from SSHADE-Spectra SSDM data model was greatly expanded to describe the various parameters of a band list and its bands (position, width, intensities, errors, mode attribution, ...). This new version of SSDM-BL is now fully implemented in the SSHADE-BL database structure.

Some points specifically developed:

- The ‘band list constituent’ data model includes:
 - A description of the solid, mineral and liquid fundamental phases and its main physical properties, + links
- The ‘band list’ data model includes:
 - Type of band list and general description,
 - Spectral range covered and reference wavelengths,
 - review process description, quality flag...
 - version, history, ...
 - links to relevant SSHADE spectra or to data in external databases *
 - review file
 - preview settings and band list organization (*)
- The ‘band’ data model, includes:
 - Mode assignment information of each band (Label, type of transition, bond involved, crystal site...)
 - Mode multiplicity, degeneracy, resonances and splitting
 - Environment parameters: T, P
 - Excitation wavelength (for Raman and fluorescence) *
 - Sample orientation, state of polarization and orientation *
 - Analysis methods
 - Band characteristics:
 - ✓ position (peak and centre values +uncertainty),
 - ✓ width (fwhm +uncertainty, shape and asymmetry factor)
 - ✓ peak intensity (in absorption coefficient, specific absorption coefficient and relative +uncertainty, band strength)
 - ✓ integrated intensity (in absorption coefficient, specific absorption coefficient and relative +uncertainty, band strength)
 - ✓ each with method, evaluation level and comment
 - Main publications reviewed and selected for each band

* : fully new features in the final version. Many other improvements have been also done on most of the other parts of the datamodel

3 The band List Database

| | |
|---------|--|
| Authors | Development: Manon Furrer (<i>developer from September 2020 to August 2021</i>), Philippe Bollard (<i>project manager, up to November 2020</i>), Damien Albert (<i>project manager and developer, from December 2020</i>), Bernard Schmitt (<i>scientific manager</i>) |
|---------|--|

Abstract: This part presents the technical choices of software and hardware implementations and the major steps of the database and interface development

During the first 6 months the major actions of the development of the SSHADE band list database were the design of the database architecture from the SSDM-BL data model and to build the tools dedicated to data providers (import, search, and visualization of data). The next year was mostly dedicated to develop the basic tools for users (search, visualization and export of data). The last 1.5 years were dedicated to improve most of the user tools, i.e. expand the capability of the search, visualization and export tools, and upgrade the datamodel considering the data manager and user experiences. We also developed a DOI generator that automatically do the request to DataCite for each band list that is pushed online.

An easy-to-fill excel file for collecting and selecting the various data during the bibliographic review process and a tool to convert the selected data into an import xml file have been also developed with the help of participative software developers of the civil society.

3.1 Technical choices

The technical base of the SSHADE-BL database is the same as for the SSHADE infrastructure, in order to integrate it in SSHADE and ease the future VO integration with other software parts of the Europlanet project, particularly VESPA. Only state of the art and stable free software were used:

PostgreSQL, programming language Python 3, framework Pyramid, the ORM (Object-relational mapping) SQLAlchemy. They provide reliability and access to several scientific libraries that can be used for calculation and visualization. Elasticsearch is also used as search engine in order to provide a ‘Google-like’ search tool based on an efficient data indexation.

3.2 Server infrastructure

The OSUG-DC platform provides two virtual servers for the SSHADE infrastructure and its databases.

A first VM hosts the "production" web-services:

- SSHADE interface at <https://www.sshade.eu/>
- SSHADE Wiki at <https://wiki.sshade.eu/> (public, but with a restricted area for partners)

A second VM is dedicated to development and hosts:

- A forge (for project management, tests and continuous integration) at <https://forge.sshade.eu/> (access restricted to the development team).
- A development version of SSHADE at <https://dev.sshade.eu/> (access restricted to the development team)

The OSUG-DC hosting infrastructure is based on three physical servers distributed over two buildings and connected each other by dedicated optical fibers. These servers are operated together through a redundant VMWare cluster providing service continuity and regular backups for all hosted virtual servers (VM). The VMWare cluster uses the “SUMMER” infrastructure provided by University Grenoble Alpes (UGA) for distributed storage and backups.

4 Development of tools for data providers

| | |
|---------|---|
| Authors | Development: Manon Furrer, Damien Albert, Philippe Bollard, Bernard Schmitt + free contributors |
|---------|---|

Abstract: This part presents the data ingestion tools, the data search and visualization tools for data providers and managers.

4.1 Data ingestion tools

The first part of the development focused on data ingestion. The GUI for the data ingestion was adapted to band lists and bands and fully tested. The import tool includes a validator that checks the correctness of the xml structure, the validity of the terms used in keywords with enumerations, the existence in the database of the linked data, the presence of "mandatory" data or optional blocks when specific conditions are satisfied... On error, the validator stops the ingestion and provides guidance to solve the problem in the xml file without recording any wrong data into the database.

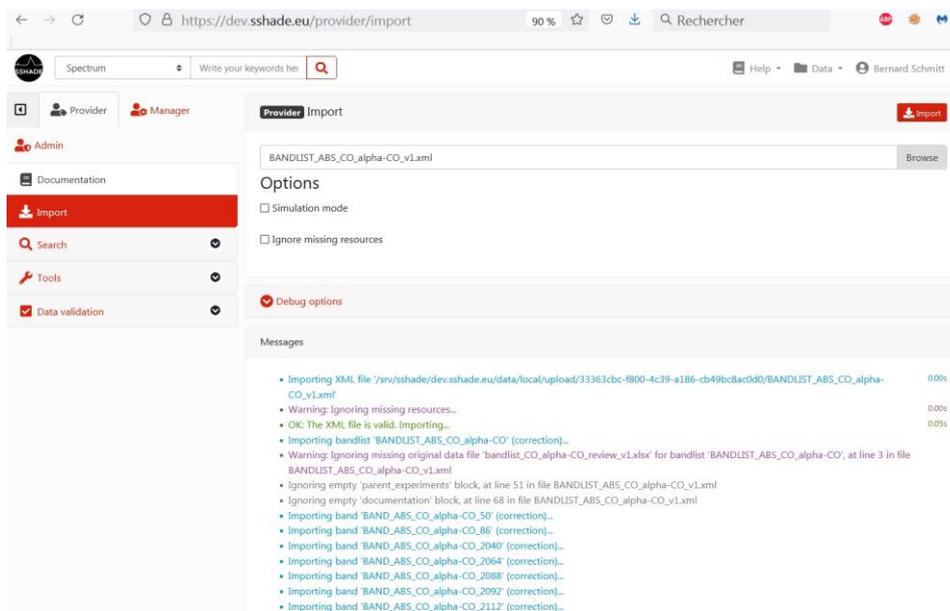


Figure 2. Import interface for providers, displaying import errors or the list of data import upon successful validation of the content of the xml import file

An history is stored for every successful import and provides a way to download the imported files for possible later correction. Data unit conversions are operated on most numerical data during import in order to homogenise the content of the database.

The import tool manages the right of data import and correction by accredited data providers.

A set of 2 specific import XML template files are provided to the SSHADE-BL managers to prepare and validate the different types of data (Constituent, Band list + bands). They provide all necessary data information and import rules (complemented by the SSDM-BL reference document)

4.2 Band list and band searches and visualization

The second step of the development focused on the GUIs for band list data providers: data search and visualization. A data provider can search all kinds of useful data stored in the SSHADE common database: fundamental species (atoms, ions, chemical bonds, chemical functions, and molecules) and phases (minerals, solid phases, liquid phases), spectral data (constituents, experiments, spectra), publications, other band lists and bands, that he will need to link with his data. For this, the data provider has access to several forms using a ‘Google-style’ search toolbar that can be complemented with a large and powerful set of specialized filters. We specifically extended this search tool for ‘band lists’, ‘bands’ and their ‘constituents’, using a dedicated search engine based on ‘Elasticsearch’ tool.

Figure 3. Search page of ‘Bands’ for data provider showing some of the different search options

All metadata can be visualized with all their relational links to other data in SSHADE (e.g. publications, linked spectra...) and with external sources of information (e.g. Wikipedia, WebMineral web pages, ...).

4.3 Band List review Excel table

To help the data managers to perform the complex process of band list data compilation and critical review a formatted Excel table template (~140 keywords) has been developed to compile the whole set of band list and band parameters values of all the data available in the literature. This allows an easier selection or synthesis of the best ones.

When a data is selected or synthesized on a dedicated line, a flag is added to the line or to the group of parameters in order to allow the tool described below to convert only the selected data in an XML import file. The most relevant publications for each band are also selected to be included in the database.

Figure 4. Part of the data information compiled for a few bands of the alpha-CO band list

4.4 Band List XML generator

A second tool to help the data managers to prepare band list data is a specific stand-alone software, developed with the help of participative software developers of the civil society. It allows to parse the content of a band list review Excel table, check if all mandatory data are present and if some condition are met and then convert the selected data flagged in the Excel table to a band list import XML file fully ready to be ingested in SSHADE-BL by the import tool. It has been recently upgraded to the SSDM-BL version 0.9.2.

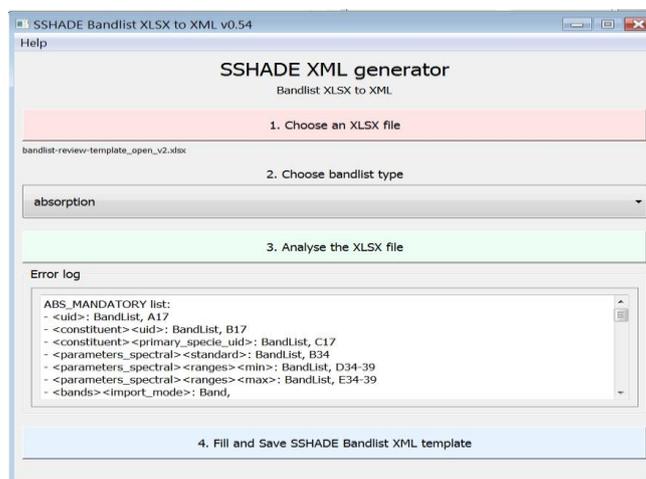


Figure 5. Interface of the converter of the selected data in the Excel table to an import XML file

5 Development of the interface for Users

| | |
|---------|---|
| Authors | Development: Damien Albert, Manon Furrer, Philippe Bollard, Bernard Schmitt |
|---------|---|

| |
|--|
| Abstract: This part presents the data search, visualization and export interface for users. |
|--|

The third step of the development focused on the GUIs for users: data search, visualization and export.

5.1 Data search for users

It uses the same ‘Elasticsearch’ engine as the data search for providers but differs mainly on the properties selected to be queried and displayed. From the front page of SSHADE (<https://www.sshade.eu>) an user can currently search either ‘band list’ or ‘band’ data (in addition to spectral data and publications) through two distinct forms using a simple ‘Google-style’ search tool that can be complemented with a number of specialized filters to refine the search.

Figure 6. User search page for ‘Band list’ showing the different filters (with additional search ‘options’ folded)

For the band lists the user can currently search according to a series of topics:

- by band list (type, spectral range (unit options), transition category)
- by band position (up to 3 bands) with constraint on their strength
- by molecule (for molecular solids and liquids: name, formula, code, type of bond or chemical function)
- by constituent (name, type, code, constituting atom or molecule),
- by environment (temperature and pressure ranges (unit options),

For the band data the user can filter the search according to the following topics:

- by band (position wavelength + uncertainty (unit options), width and strength ranges, type of transition and specific mode, constraint from a second band),
- by molecule (name, formula, code, type of bond),
- by constituent (name, type, code, constituting atom or molecule),
- by environment (temperature and pressure ranges (unit options),

The screenshot shows a web interface for searching bands. At the top, there is a search bar with the text "Write your keywords here" and a search icon. Below this is a "Bands search" section with a "Reset" button and a search bar containing "in user S1b" and a "Search" button. The main section is titled "Filters" and contains several filter categories:

- By Band:**
 - Type: "in" dropdown, value "Raman scattering"
 - Band position: "-" dropdown, value "1100", followed by a range operator "=", a value "50", and a unit dropdown "cm-1".
 - Options: expandable section.
- By Molecule:**
 - Name: "contains all words" dropdown, empty input field.
 - Formula: "contains all" dropdown, value "C O".
 - InChI + key: "is" dropdown, empty input field.
 - Options: expandable section.
- By Constituent:**
 - Name: "contains all words" dropdown, empty input field.
 - Liquid compound type: "in" dropdown, value "Nothing selected".
 - Solid (synthetic) compound type: "in" dropdown, value "Nothing selected".
 - Mineral compound type: "in" dropdown, value "carbonate".
 - CAS number: "is" dropdown, empty input field.
 - Options: expandable section.
- By Atoms and Radicals:**
 - Element: "is" dropdown, value "Nothing selected".
 - Radical / Molecule: "contains all words" dropdown, empty input field.

At the bottom, there is a "Results" section showing "6 bands".

Figure 7. User search page for 'Bands' showing the different filters (with search 'options' folded)

Band lists search

carbonate raman

Filters

Results 6 band lists

| Title | Bands number | Molecule/Atom | Constituent |
|---|--------------|---------------|---------------|
| Raman bandlist of natural Calcite | 7 | | Calcite |
| Raman bandlist of natural Rhodochrosite | 7 | | Rhodochrosite |
| Raman bandlist of natural Magnesite | 5 | | Magnesite |
| Raman bandlist of natural Aragonite | 23 | | Aragonite |
| Raman bandlist of natural Dolomite | 10 | | Dolomite |
| Raman bandlist of natural Siderite | 5 | | Siderite |

Figure 8. Example of results for a “Band list” search of ‘Raman spectra of carbonates’. The band lists display their title and the number of bands they contain

The results of the band list search are displayed as a list of title of band lists with the number of bands they contain, the molecule and constituent involved. The list can be ordered according to these parameters. The user can then have a quick look at each band list with a preview popup and then go to its detail page, but s-he can also decide to immediately download a band list or send it to the shopping basket.

The results of the band search are displayed as a list of bands with their main parameters (position, width, strength, assignment, primary species...) that can be also used to order the list. The user can then display the band information or go up to its band list. S-he can also decide to immediately download a band or send it to the shopping basket.

Bands search

Write your keywords here or leave it empty to get all the data...

Filters

Results 4 bands

| | Position (cm-1) | Width (cm-1) | Strength | Assignment | Type | Species | Molecule/Atom | Constituent | Temperature (K) | Pressure (bar) |
|---|-----------------|--------------|----------|--------------------------------|------------|---|-----------------|-----------------------------------|-----------------|----------------|
| 👁 | 2412.5 ± 0.3 | - | w | ν ₁ +ν ₂ | absorption | ³² S ¹⁶ O ¹⁶ O | SO ₂ | natural SO ₂ - phase I | 90 | - |
| | 2413 ± 0.3 | 2.4 ± 0.2 | w | | | | | | | |
| 👁 | 2433.5 ± 0.3 | 4 ± 0.2 | m | ν ₁ +ν ₂ | absorption | ³⁴ S ¹⁶ O ₂ | SO ₂ | natural SO ₂ - phase I | 90 | - |
| | 2434.5 ± 0.3 | 3.2 ± 0.2 | m | | | | | | | |
| 👁 | 2445 ± 0.3 | - | w | ν ₁ +ν ₂ | absorption | ³² S ¹⁶ O ₂ | SO ₂ | natural SO ₂ - phase I | 90 | - |
| | 2445.3 ± 0.5 | - | w | | | | | | | |
| 👁 | 2456 ± 0.5 | 4 ± 0.2 | s | ν ₁ +ν ₂ | absorption | ³⁴ S ¹⁶ O ₂ | SO ₂ | natural SO ₂ - phase I | 90 | - |
| | 2457 ± 0.3 | 6.3 ± 0.3 | s | | | | | | | |

Figure 9. Example of results for a “Band” search of ‘SO₂ between 2400 and 2460 cm⁻¹’. The main parameters of the selected bands are displayed

5.2 Data visualization

The user can select and visualise a band list, and will then get a page with the collapsible structure of the band list/bands on the left part of the page, and the information on the constituent in another tab. While on the right part of the page one or two previews (in absorption coefficient and/or relative intensity, whichever available) of the band list is displayed together with the main information on the band list and its constituent. These previews simulate the spectrum of the selected molecule by using the available list of position, width and intensity of all its bands.

The user can then decide either to visualize the band list interactively together with all its associated information, or to look at the detailed information of any band, or of the band list constituent.

The detailed page of each level of the band list, band and constituent structure contains all the relevant metadata with different types of links either to another level of the structure, to other information stored in SSHADE (such as publications) or to external pages (such as Wikipedia, ...). These links either create a popup window with the information (internal links) or open the information in another tab of the browser (external links).

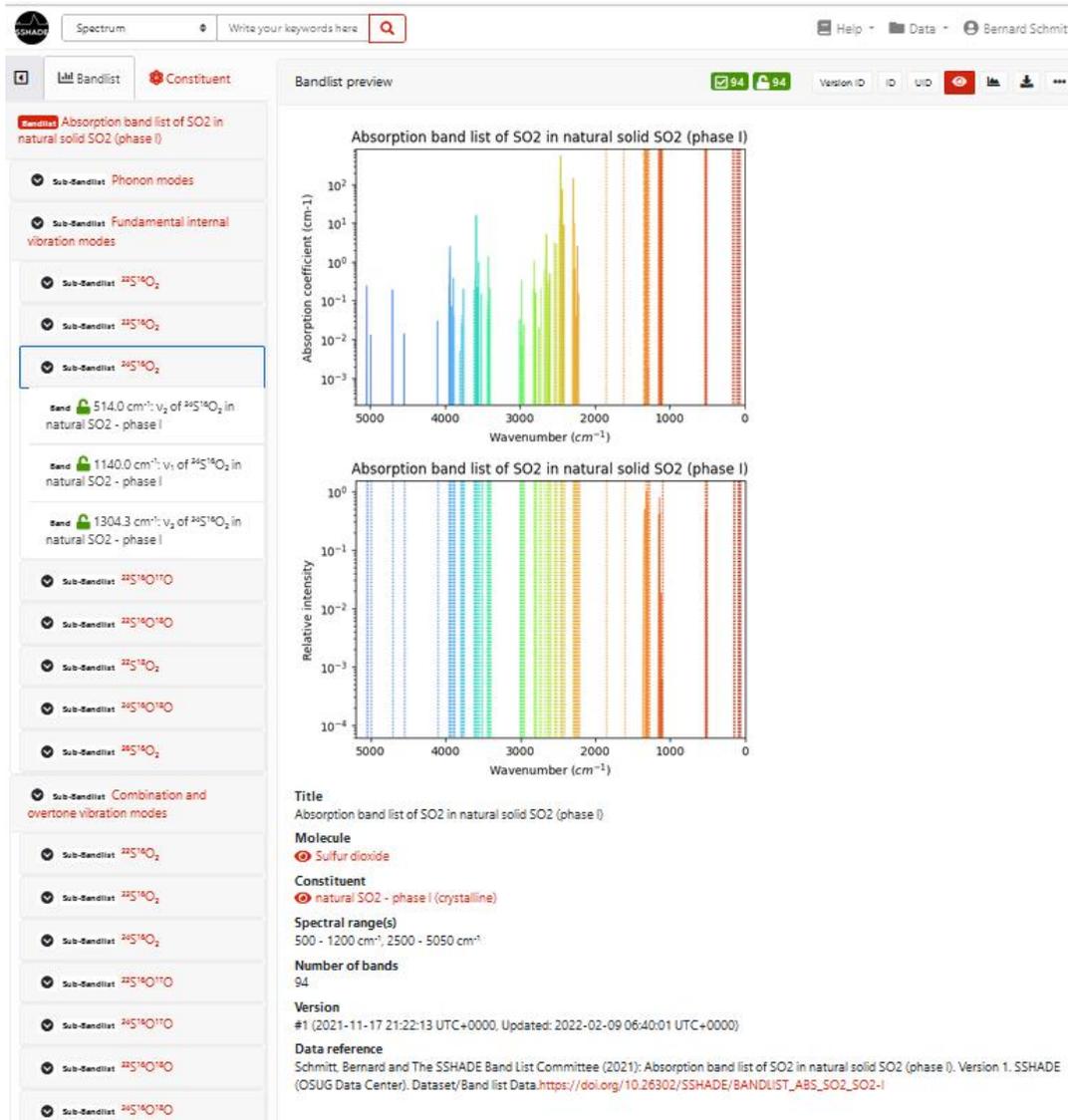


Figure 10. Collapsible structures of the band list/(sub-band list)/bands and of the constituent (left) and absorption coefficient and relative intensity band list previews with the main information below

The interactive display has numerous options and also manages the missing information on width and / or intensity on some bands.

It can display either:

- Individual bands
- Sum of all bands
- Sum of bands by isotopes
- Sum of bands of an individual isotope

in two intensity units (when available):

- Absorption coefficient
- Relative intensity

with different band representations:

- Peaks: simple line (position + intensity)
- Triangles (position + width + intensity)
- Gaussian (position + width + asymmetry + intensity)

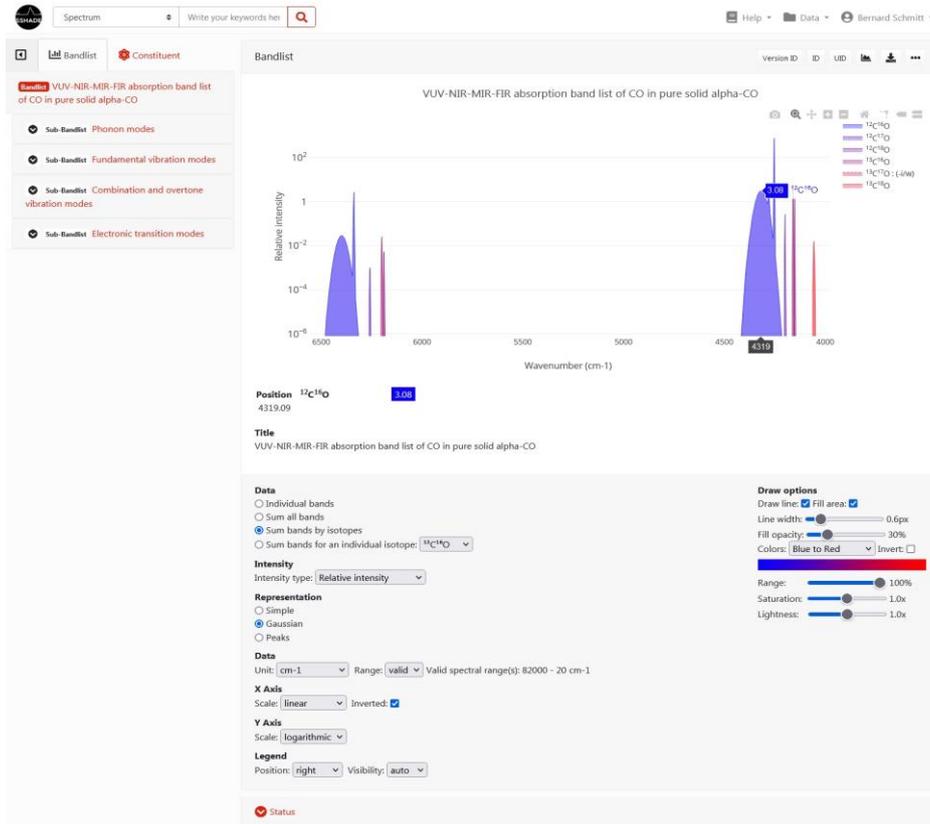


Figure 11. Example of dynamic display of part of the alpha-CO absorption band list, with the different display options below (left: band list/ band structure).

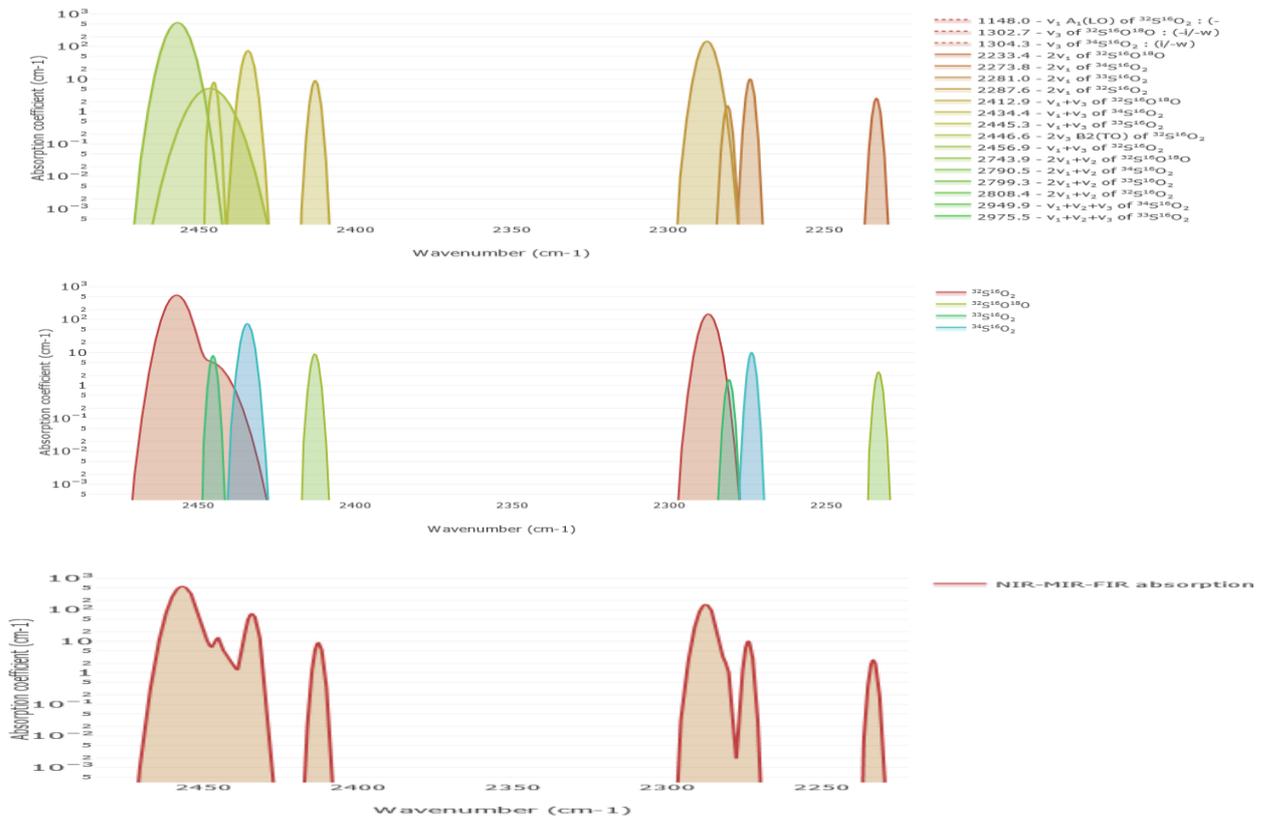


Figure 12. Example of different types of displays of the same part of the SO₂-phase I absorption band list (top: individual bands, middle: sum by isotope, bottom: sum of all bands)

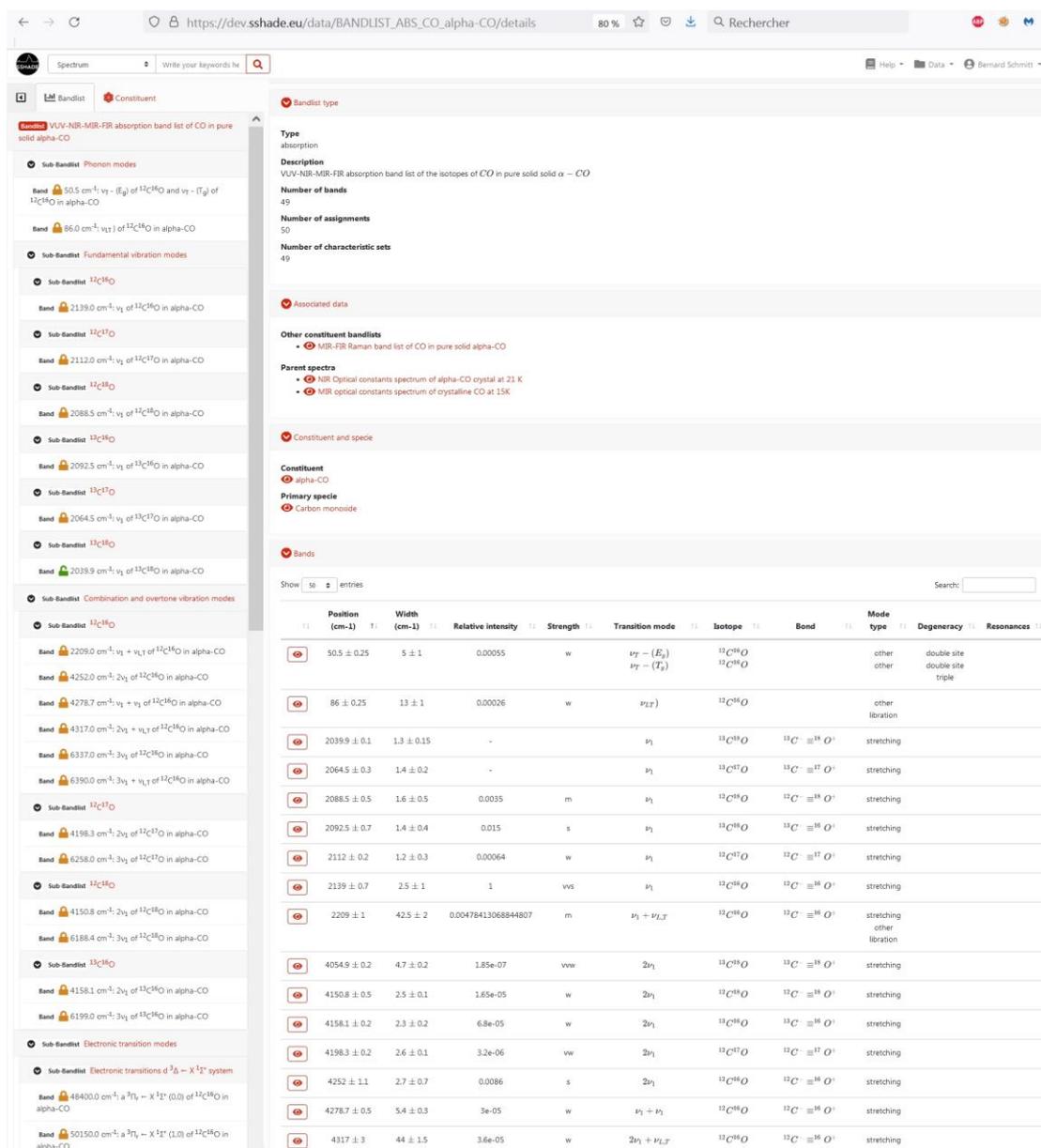


Figure 13. Example of part of the detailed information of the alpha-CO absorption band list. In particular a synthetic table of all the bands of the band list, with their main parameters, is provided.

5.3 Data export

The users can download the information on a whole band list or of a single band from the export page. The users may also add a band list or a band in the ‘shopping basket’ for future export.

Currently the export tool allows downloading:

- ✓ the band list data as a simple 25-columns ascii files (with or without short header, CSV or TSV data) containing most of the band parameters;
- ✓ the full metadata of the band list, of all its bands and of the constituent (html format)

The export setting page also allows converting the spectral units, limit the spectral range and set the number format. These settings can be stored in the user dashboard.

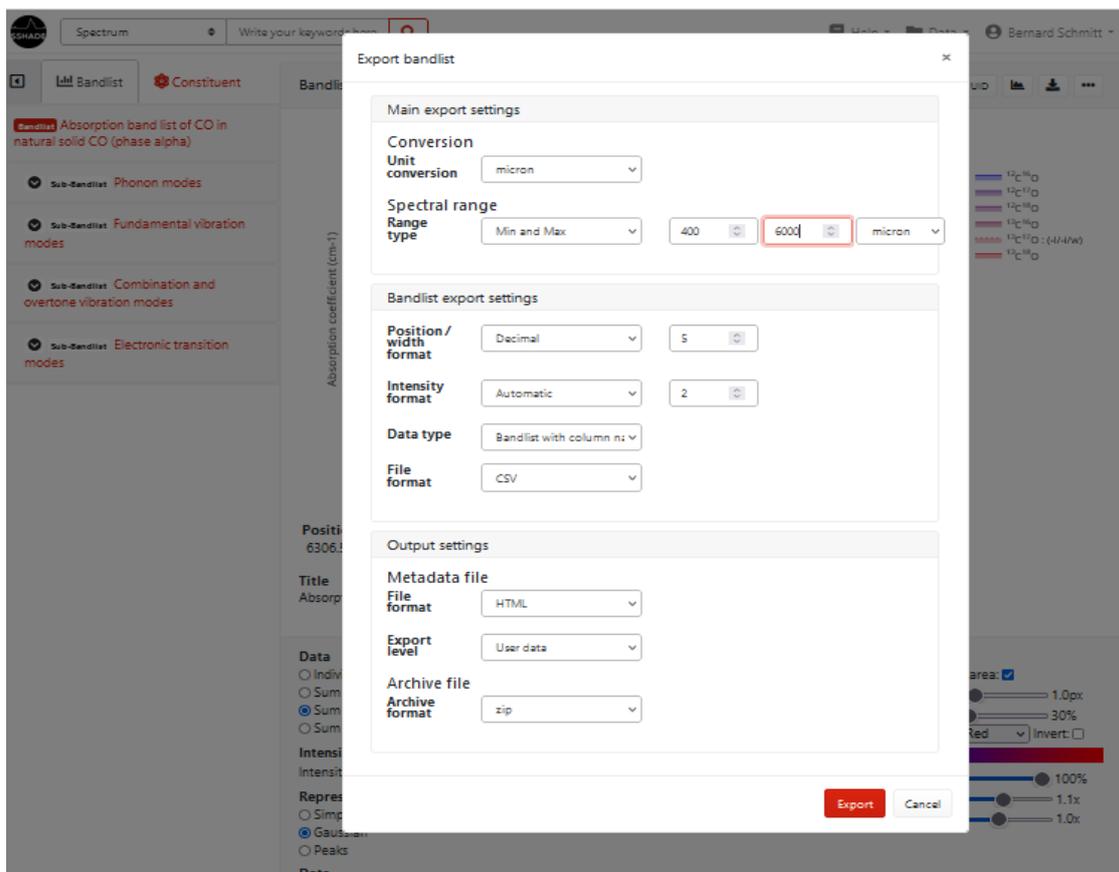


Figure 14. The export customization interface

All data and metadata are delivered in a zip file that also contains:

- ✓ one or two previews of the data,
- ✓ a detailed description of the data files
- ✓ a ‘citation file’ providing the reference of the main paper(s) used for the review of the band list and the ‘data reference’ of the band list (that will be complemented with its DOI in the final version), as well as guidelines on how to use and cite the data.

| Export | UID | Title | Export date | Size | Steps | Progression | ETA |
|--------|-----|---------------------------------|--|------------|----------|-------------|-----|
| | | BANDLIST_ABS_CO_alpha-CO | Absorption band list of CO in natural solid CO (phase alpha) | 2023-05-24 | 1.7 MB | done | 0s |
| | | BANDLIST_ABS_CH3CN_alpha-CH3CN | Absorption band list of CH3CN in natural solid CH3CN (phase alpha) | 2023-04-27 | | done | 0s |
| | | BAND_ABS_CH3CN_alpha-CH3CN_1036 | 1036.0 cm ⁻¹ v ₁ of CH ₃ CN in natural CH ₃ CN - phase alpha | 2023-04-27 | 59.8 kB | done | 0s |
| | | BANDLIST_ABS_N2_alpha-N2 | Absorption band list of N ₂ in natural solid N ₂ (phase alpha) | 2023-04-19 | 595.2 kB | done | 0s |
| | | BANDLIST_ABS_HCN_HCN-II | Absorption band list of HCN in natural solid HCN (phase II) | 2023-04-19 | 376.8 kB | done | 0s |
| | | BAND_RAMAN_N2_alpha-N2_32 | 32.5 cm ⁻¹ v ₁ (E _g) of ¹⁴ N ₂ in natural N ₂ - phase alpha | 2023-04-19 | 59.3 kB | done | 0s |

Figure 15. The export dashboard interface of the users

6 DOI generator for bandlists

| | |
|---------|--------------------------------|
| Authors | Damien albert, Bernard Schmitt |
|---------|--------------------------------|

Abstract: This part presents the DOI of SSHADE-BL

A mapping of the DataCite keywords with the Band list ones allowed us to design a tool that collect all the necessary information in the band list database to automatically create the XML file that need to be submitted to DataCite to get an active DOI as well as a data reference. This creation occurs when a new band list is put online. The band list data references can be cited in publications and its DOI allows the reader to go directly, with a single click, to the relevant band list data, through a landing page providing general info on the band list.

SSHADE is already registered as a DOI provider (doi: 10.26302/)

A typical band list data reference is given below:

Schmitt, Bernard and The SSHADE Band List Committee (2021): Absorption band list of CH₄ in natural solid CH₄ (phase I). Version 2. SSHADE (OSUG Data Center). Dataset/Band list data. https://doi.org/10.26302/SSHADE/BANDLIST_ABS_CH4_CH4-I

7 Band List Database filling

| | |
|---------|---|
| Authors | Bernard Schmitt, with Lucia Mandon, Apolline Leclef |
|---------|---|

Abstract: This part presents the process of content selection and the current state of data ingestion in SSHADE-BL

The SSHADE-BL database has been already filled with more than 40 absorption and Raman band lists from 24 different ice phases and 7 minerals, with a total of more than 1100 bands.

The preparation of each data set first need a careful scientific process including:

- a bibliographic survey of all publications available for a specific band list (Absorption or Raman data on either one molecule in a well-defined and simple constituent phase, or on a single mineral)
- a detailed analysis of all papers, in particular the tables with list of band parameters and assignment (with sometimes a complementary extraction of band parameters on scaled figures)
- the extraction of all the relevant information and compilation in the review Excel file template.
- in some case the detailed analysis (retrieval of all band parameters) of published or unpublished spectra available either in one of the SSHADE-Spectra databases, or in an external reliable database (such as ‘The Cosmic Ice Laboratory’, NASA/GSFC) in order to verify or complement the data missing in publications.
- In some cases a theoretical spectroscopic analysis of the vibration or phonon modes attributions in order to confirm the assignments to specific modes or

isotopes of the molecule, or to determine those of the bands with still unknown attribution (for example: still to be done for most of the combination bands of CO₂ ice !)

- a critical review and accuracy/quality assessment, band-by-band (and also globally), of all the parameter values and associated information.

Due to the large amount of time needed for the critical review of a single band list (for example the band-list of alpha-CO is based on the compilation of more than 35 papers scattered in the literature and spanning the 1961-2022 years), plus additional data analysis, this whole process has been completed only for part of the band lists imported in the database, but for all other band lists at least 3 of the most comprehensive data sources (sometimes containing review of older data) have been compiled and compared.

Currently (as 30/05/2023) the 42 following data sets have been put online:

- Absorption band lists:
 - For 24 different ices: N₂ (phases alpha and beta), CO (phase alpha), CO₂ (phase I), CH₄ (phase I), C₂H₂ (amorphous, metastable and phase I), C₄H₂ (amorphous, and phase I), SO (in Ar matrix), SO₂ (phase I), S₂O (in SO₂ and Ar matrices), H₂O (amorphous Ia and phase Ih), NH₃ (phase I), CH₃OH (amorphous and phase alpha) HCN (phase II), HC₃N (phase I), CH₃CN (phases alpha and beta), C₂N₂ (phase I) ...
 - For 2 minerals: Calcite, Hematite

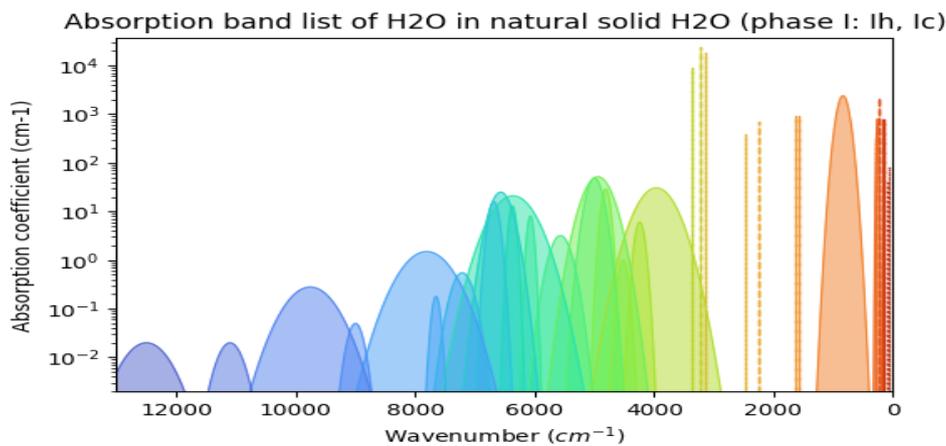


Figure 16. Absorption band list of H₂O (phase Ih)

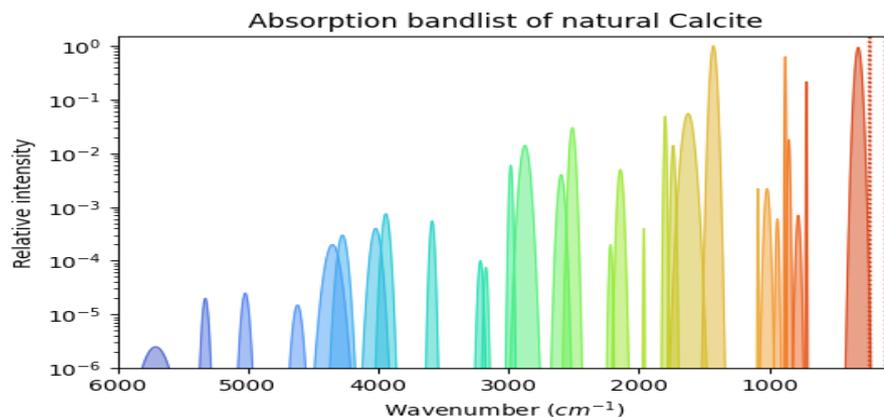


Figure 17. Absorption band list of Calcite

- Raman band lists:
 - For 10 different ices: N₂ (phases alpha and beta), CO (phase alpha), CH₄ (phase I), SO₂ (phase I), S₂O (in SO₂ and Ar matrices), HCN (phase II), CH₃CN (phases alpha), C₂N₂ (phase I),
 - For 6 carbonate minerals: Calcite, Aragonite, Dolomite, Magnesite, Siderite, Rhodochrosite

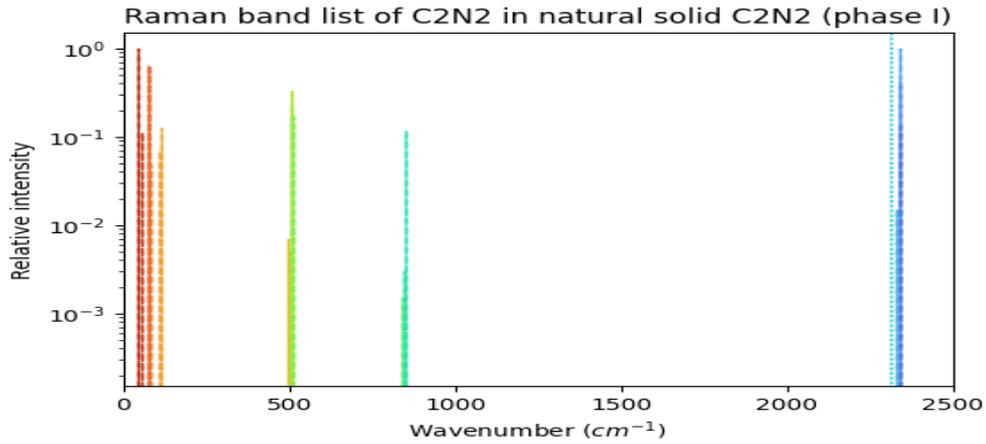


Figure 18. Raman band list of C₂N₂ (phase I)

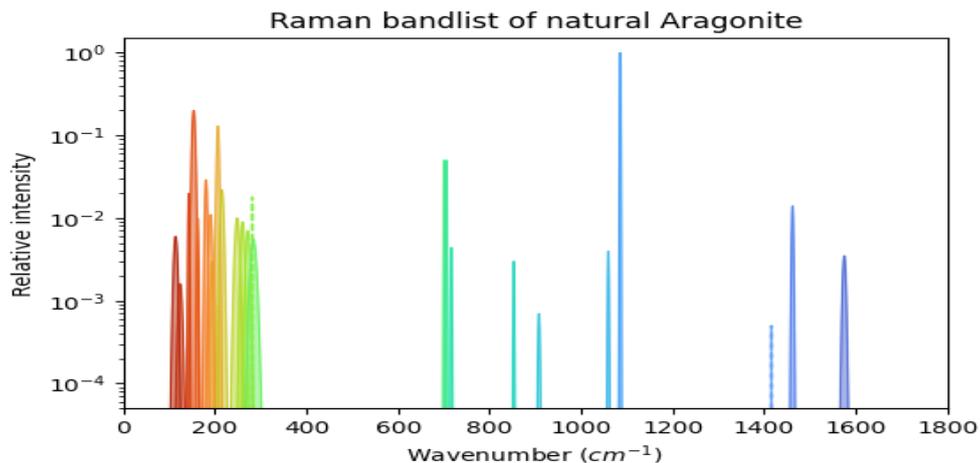


Figure 19. Raman band list of Aragonite

- Linked publications (~150). Our searchable database on solid spectroscopy publications has now over 520 references.

A dozen additional absorption band lists (containing about 300 bands) are still in the review process or at the ingestion step (C₂H₂, C₂H₄, C₆H₆, H₂O₂, H₂S, CO₂ in various matrices, S₂, S₃, S₈, Dolomite, Magnesite, Siderite, ...). They are expected to be all online by the end of summer 2023.

The current effort is centred on the preparation of data useful for the interpretation of current space missions to Mars and asteroids, for the preparation of the JUICE mission and for the interpretation of JWST solar system observations.

8 Milestone

The SSHADE-Bandlist database, online for the user since end of September 2021, contain already more than 1100 band data

- The SSHADE-BL database is at <https://www.sshade.eu/>
- Source code of the SSHADE-BL interface is available at: https://dev.sshade.eu/sources/ssshade-bl_latest_develop.zip
- The SSHADE wiki, with provider area, is at <https://wiki.sshade.eu/>
- The Excel band list review template and both band list/bands and constituent XML templates for data ingestion in SSHADE are at <https://wiki.sshade.eu/doku.php?id=ssshade:provider> (restricted access).
- The XML Band list generator is at <https://wiki.sshade.eu/doku.php?id=ssshade:provider> (restricted access).

The ‘provider area’ of the SSHADE infrastructure and of the SSHADE wiki have restricted access. A temporary login/password with ‘partial data provider rights’ has been set. It allows accessing the SSHADE provider menu with all the specialized search and tools for providers, but does not allow to import/correct data. In the wiki, it also provides access to the ‘Provider documentation’ section. It is strictly restricted to the European Commission service and the Europlanet Management Committee.

- login/password: EPN@2024.eu/SSHADE-BL

Contacts: Bernard Schmitt (bernard.schmitt@univ-grenoble-alpes.fr)

9 Future developments of the SSHADE-BL database

After this delivery of the ‘final version’ of the SSHADE Band List Database (MS84 - 30/05/2023) we will mostly develop an EPN/TAP access for use in the VO context, in particular with the VESPA VO portal. We will also very significantly increase the amount of critically reviewed and ingested band list datasets.

We will also continue to improve over the long term the current interface and its underlying data model, depending on the data manager and user requests.

The following main SSDM-BL and software developments will be continued within the VA VESPA (WP6) activity for the last 14 months of the Europlanet-2024 RI program.

9.1 Future SSDM-BL developments

Possible improvements/extensions of the SSDM-BL datamodel could be:

- Addition of the possibility to describe band list of molecular species adsorbed on the surface or inside the porosity of solids

9.2 Future SSHADE-BL interface and tool developments

- Development of EPN-TAP access to the Band List Database for use in a VO context. Development of a specific VO interface with VESPA.
- Export options: addition of the ‘VO-Table’ export format

This interface and extension to VESPA will be delivered in D6.15: “VESPA VA activity report, 4”.

Other possible improvements/extensions of the interface could be:

For data import:

- Implementation of the ‘adsorbed species’ extensions for constituents to describe the band list of a molecule adsorbed on a solid

For the users:

- Extension of the export options (data file formats: FITS.... choice of data columns, intensity unit conversion, gaussian simulations, selection of specific isotope, ...)

9.3 Future SSHADE-BL content

The filling of the Band List database and its outreach activities (including SSHADE-BL documentation, tutorials and user-cases, SSHADE-BL talks and demonstrations at conferences, ...) will continue and will be reported in D6.15: “VESPA VA activity report, 4”. A 6-months Post-doc will be hired in January 2024 for these tasks.

It is planned that by the end of the Europlanet-2024 program in July 2024 the SSHADE band list database will contain over 75 absorption and Raman band lists of molecular solids and minerals in various phases, including more than 1600 bands. Hopefully a position of an ‘associate astronomer’ with a ‘service task’ on SSHADE will be obtained at the end of 2024 from CNAP/INSU/CNRS. S-he will greatly help to increase the content of the database much faster.

We also started to ask a few SSHADE partners to participate to a “SSHADE band list review committee” to contribute to the compilation and review process of specific sets of band list data.