https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Spectral Library of European Pegmatites, Pegmatite Minerals and Pegmatite Host-Rocks – The GREENPEG Project Database

Joana Cardoso-Fernandes[1,2], Douglas Santos[1], Cátia Rodrigues de Almeida[1,2], Alexandre Lima[1,2], Ana C. Teodoro[1,2] and the GREENPEG project team[+]

[1]Department of Geosciences, Environment and Spatial Planning, Faculty of Sciences, University of Porto, Rua Campo Alegre, 4169-007 Porto, Portugal
[2]ICT (Institute of Earth Sciences) – Porto Pole (Portugal), Rua Campo Alegre, 4169-007 Porto, Portugal
[+] A full list of authors appears at the end of the paper.

Correspondence to: Joana Cardoso-Fernandes (joana.fernandes@fc.up.pt)

Abstract. The GREENPEG spectral database contains the spectral signature, obtained through reflectance spectroscopy studies, of European pegmatites and minerals, as well as their host rocks. Samples include Nb-Y-F (NYF) and Li-Cs-Ta (LCT)-type pegmatites and host rocks from pegmatite locations in Austria, Ireland, Norway, Portugal, and Spain. The database contains the reflectance spectra (raw and with continuum removed), sample photographs, and main absorption features automatically extracted by a self-proposed Python routine. Whenever possible, spectral mineralogy was interpreted based on the continuum-removed spectra. A detailed description of the database, its content and structure, the measuring instrument, and interoperability with Geographic Information Systems (GIS) is available in this database report. Moreover, examples of how the data can be used and interpreted are also provided. The advantages and added value of the presented dataset reside on its European scale with representative samples from pegmatites with distinct genesis, mineralogy, structure, and host rocks that can be used as a reference for pegmatite exploration at a global scale through satellite image processing, for example. The reported spectral mineral assemblages can also be of interest when considering resource estimation or ore processing. Thus, it is expected that this open dataset, available on the Zenodo platform https://doi.org/10.5281/zenodo.6518319 (Cardoso-Fernandes et al., 2022), will be a reference for distinct types of users ranging from academia to industry.

1. Introduction

This spectral database was built in the frame of the GREENPEG project (https://www.greenpeg.eu/), funded by the European Commission Horizon 2020, which aims to develop multi-method exploration toolsets for the identification of European, buried, small-scale (0.01-5 million m^3) pegmatite ore deposits of the Nb-Y-F (NYF) and Li-Cs-Ta (LCT) chemical types (Müller et al., 2022a). Moreover, the GREENPEG project also aims to enhance European databases of petrophysical and reflectance properties, for example adding new data on the properties of pegmatites and their green raw materials, including their spectral signature obtained through reflectance spectroscopy studies. This database is publically available on the Zenodo platform https://doi.org/10.5281/zenodo.6518319 (Cardoso-Fernandes et al., 2022).

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

The main objectives of this work are to present (i) the spectral library composed under the GREENPEG project, (ii) the formats in which the database is made available and the interoperability with Geographic Information Systems (GIS) software, (iii) how the spectral library is organized, and (iv) how the data can be used and interpreted.

Previous databases have been published concerning pegmatite exploration. Cardoso-Fernandes et al. (2021) provided a dedicated spectral library only for the Fregeneda–Almendra Aplite–Pegmatite Field in Central Iberia. The samples corresponded to LCT pegmatites considered to be the result of fractional crystallization of peraluminous granites derived from partial melting of highly peraluminous, calcium-poor, and phosphorus-rich metasedimentary rocks of Neoproterozoic age, during the Variscan orogeny (Roda-Robles et al., 2018). Despite its relevance to the state of the art, the previous database lacks spectral reflectance data for other types of pegmatites, with distinct mineralogy, textures, and genesis. Fabre et al. (2021) presented a broader database of laser-induced breakdown spectroscopy (LIBS) spectra from multiple locations around the world, such as the Fregenda–Almendra and Gonçalo pegmatite fields in the Iberian Peninsula, and also from Canada and Brazil. Although this database comprised samples from distinct pegmatites, the main focus was on Li-minerals, and therefore samples from LCT pegmatites.

Taking this into account, the GREENPEG spectral reflectance database presents several advantages and high added value when compared with the already available datasets: (i) it is the first database of this kind built at a European scale; (ii) it includes samples from distinct pegmatites with different mineralogy, structure, host-rocks, and genesis (anatectic and granite-related); and (iii) it is the first open database providing data on pegmatites of the NYF chemical type. Samples include LCT- and NYF-type pegmatites and host rocks from pegmatite locations in Austria, Ireland, Norway, Portugal, and Spain.

This extensive database provided in multi-formats represents, therefore, a high-quality dataset to be used by multiple users of different backgrounds, from the academia to the mining industry. For better comprehension of the dataset, representative spectra were selected as an example to demonstrate how the spectral mineralogy included in the database was interpreted. To complement the metadata already provided within the spectral library (Cardoso-Fernandes et al., 2022), this work presents in detail its structure, content, the process of sample preparation and data acquisition, as well as useful tips on usability and accessibility of the data.

1.1. Case studies

The GREENPEG project develops research at various scales in three European demonstration sites (Fig. 1): (i) Leinster (Ireland); (ii) Wolfsberg (Austria); and (iii) Tysfjord (Norway). Complementary prospective areas in Spain and Portugal were selected for testing the developed methodologies within the GREENPEG project, including the usefulness of reference spectral data (Müller et al., 2022a).

The pegmatite field in the Tysfjord area (Norway) has an extension of about 20 km^2 with 22 known NYF dikes emplaced in granite-type rocks (Müller et al., 2022a). Two groups of pegmatites were mapped, both emplaced in the Tysfjord granite gneiss (Müller et al., 2022b). The first group consists of older Paleoproterozoic and metamorphosed pegmatites with lens- to cigar-shapes reaching up to 400 m in size, formed from residual melts of the hosting granitic gneiss; while the second group is

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Open Access Earth System Science Data Discussions

composed of younger, smaller undeformed pegmatites (400-379 Ma) formed by anatexis due to partial melting of the granitic gneiss during late Caledonian events (Müller et al., 2022b). The pegmatite samples for the spectral library were collected in 14 distinct pegmatite locations (either *in situ* or from historical drill cores), but with an emphasis on the Håkonhals and Jennyhaugen pegmatites, both exposed in large open-pit areas allowing for detailed remote sensing studies to be carried out in the scope of the GREENPEG project (Santos et al., 2022; Teodoro et al., 2021).

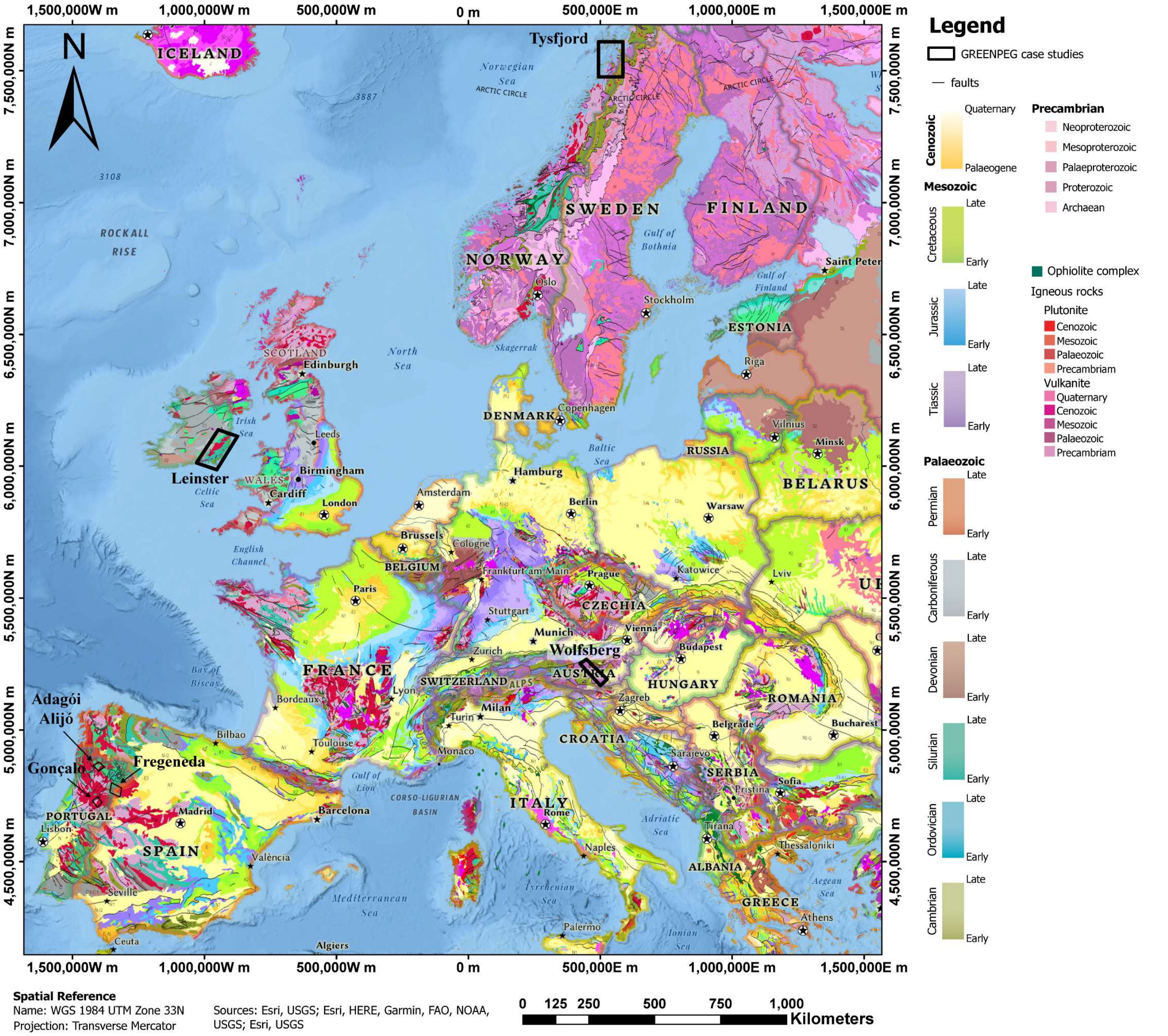


Figure 1: Location of the distinct sampled pegmatite fields in Norway (Tysfjord), Austria (Wolfsberg), Ireland (South Leinster), Portugal (Adagói, Alijó, Gonçalo), and Spain (Fregeneda) over the International Geological Map of Europe and Adjacent Areas (IGME 5000) at the scale 1:5,000,000 (adapted from Asch (2005) available at the European Geological Data Infrastructure (EGDI) platform - https://www.europe-geology.eu/). Basemap provided by ESRI, NOAA, USGSS, FAO, Copyright © 1996-2022 Garmin Ltd., © 2022 HERE.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Open Access Earth System Science Data Discussions

The pegmatite field in the Leinster area in Ireland shows around 70 km^2 of extension and with 18 known LCT pegmatites emplaced in either Paleozoic granitic or metasedimentary rocks (Barros and Menuge, 2016; Müller et al., 2022a). The pegmatites are mainly hosted along a 3 km wide NE–SW regional structure along the eastern margin of the S-type Leinster Batholith (emplaced around 400 Ma) known as the East Carlow Deformation Zone (Barros et al., 2020; Barros and Menuge, 2016). The unzoned to weakly zoned pegmatite dikes can range from a few meters to up ~20 m in thickness and are considered to be of anatectic origin (Barros et al., 2020; Barros and Menuge, 2016). Due to a lack of pegmatite outcrops, some samples were collected from boulders dispersed in agricultural fields, but most of the samples come from a diamond drilling campaign conducted in the Moylisha area.

The Wolfsberg pegmatite field in Austria corresponds to an area of 25 km^2, where 14 known LCT dikes are emplaced in either amphibolite or mica-schist rocks (Gourcerol et al., 2019; Müller et al., 2022a). The LCT pegmatites are spatially associated with simple pegmatites and leucogranites formed during an extensional event in Permian times (Knoll et al., 2018). The pegmatites are mostly unzoned bodies (Göd, 1989). Recent studies seem to indicate that the LCT pegmatites resulted from anatectic melts produced from the simple pegmatites and leucogranites, followed by fractionated crystallization (Knoll et al., 2018). Both pegmatites and leucogranites were overprinted during the Alpine orogeny with different intensities, affecting both structure and chemistry (Göd, 1989; Knoll et al., 2018). All samples analysed in this study were collected in recent drilling campaigns.

The samples from complementary test sites were collected in three distinct LCT pegmatite fields in Iberia, namely the Fregeneda-Almendra, Barroso-Alvão, and Gonçalo fields (Roda-Robles et al., 2018). All samples from Spain come from the Fregeneda side of the Fregeneda-Almendra pegmatite field (Roda-Robles et al., 1999; Vieira, 2010). The aforementioned pegmatite field spreads across a ~30 km-long and ~7 km-wide W–E belt of Neoproterozoic to Cambrian metasediments, with 11 distinct types of dykes identified, but with only five types of metasediment-hosted pegmatites worth mentioning, namely (Errandonea-Martin et al., 2022; Roda-Robles et al., 1999; Vieira et al., 2011): (i) concordant barren pegmatites; (ii) petalite-bearing pegmatites; (iii) spodumene-bearing pegmatites; (iv) lithium (Li)-mica pegmatites; and (v) spodumene+Li-mica pegmatites.

The samples collected in Portugal are distributed among the remaining pegmatite fields. The Barroso-Alvão pegmatite field comprises both the Adagói and Alijó pegmatite sites. Martins (2009) defined distinct types of dykes in the region: (i) intragranitic pegmatites; (ii) quartz-andalusite dykes, (iii) barren pegmatites, (iv) spodumene pegmatites, (v) petalite pegmatites, and (vi) lepidolite pegmatites, all intruding metasedimentary rocks of Silurian-age. While in Alijó only spodumene crystals were identified, in Adagói both spodumene, petalite and eucryptite are found (Lima, 2000). Finally, the Gonçalo pegmatite field is characterised by granite-hosted sub-horizontal dykes with complex rhythmic layering with lepidolite-rich, albite-rich and quartz-rich layers, alternating with different textures (Neiva and Ramos, 2010; Roda-Robles et al., 2018). The aplite-pegmatite sills belong to the lepidolite and amblygonite subtypes, from the complex type of the LCT family of the rare-element class of pegmatites (Černý and Ercit, 2005).

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

In general, the LCT pegmatites from Iberia display different dips, from sub-vertical to sub-horizontal, reduced thicknesses (from <50 cm up to ≈30 m), and lengths usually lesser than 1 km (Roda-Robles et al., 2022). Moreover, these dykes do not present internal zoning, but show variable grain size, from aplitic to pegmatitic textures, with the biggest crystals being usually less than 12 cm long (Roda-Robles et al., 2022; Roda-Robles et al., 2018). Textures indicative of unidirectional solidification are commonly observed, mainly comb crystals of alkali feldspars and/or Li aluminosilicates, and/or a layering parallel to the contacts with the host rocks, while graphic textures or quartz cores are not commonly observed (Roda-Robles et al., 2022; Roda-Robles et al., 2018). Roda-Robles et al. (2018) proposed a petrogenetic link between peraluminous granites and the LCT aplite-pegmatite bodies of the Central Iberian Zone.

2. Spectral library formats, structure and use

The spectral library is mainly focused on the spectral signature of the GREENPEG project demonstration sites: Norway (Tysfjord), Ireland (Leinster), and Austria (Wolfsberg). The spectral measurements were conducted on surface-collected and drill core samples provided by the GREENPEG project partners. Field campaigns allowed obtaining the spectral signature from representative samples from the test site in Portugal (Adagói). Additional samples from Portugal (Alijó, Gonçalo) and Spain (Fregeneda) previously collected were analysed and included in the spectral database. For further details on sample collection, please refer to Haase et al. (2022).

The database was originally created in a Microsoft Access database format (see chapter 2.2), but then converted to geodatabase/geopackage formats for interoperability with GIS software (see chapter 4), ArcGIS (Esri, Redlands, CA, USA) and QGIS, respectively. Therefore, users should: 1) Select the folder(s) of interest, 2) download the zip file, 3) extract the files, and 4) if necessary, follow the available tutorials. The structure of the database is presented in Table 1. Each database level is separated according to the country of collection site: (i) Norway (Tysfjord), (ii) Austria (Wolfsberg), (iii) Ireland (Leinster), (iv) Portugal (Adagói, Alijó, Gonçalo), and (v) Spain (Fregeneda). The detailed content of the database is summarized in Appendix A (Table A1).

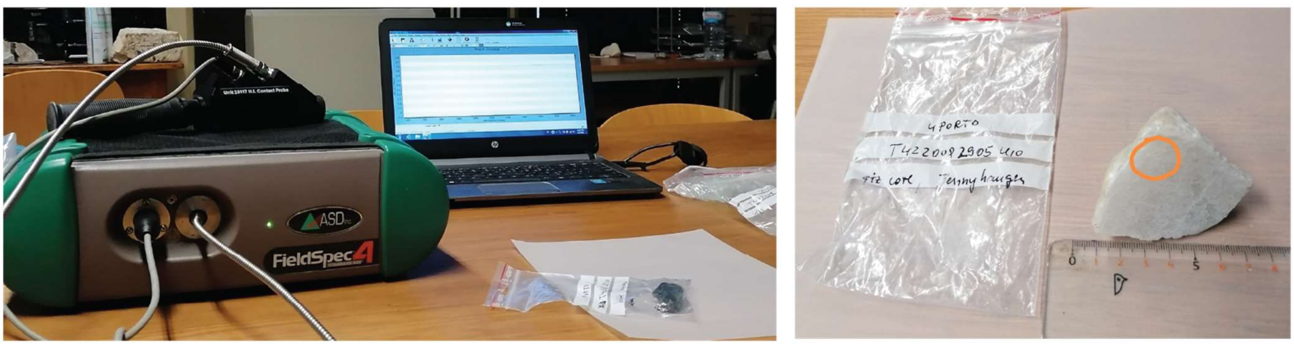
Table 1: Structure of the data available in the Zenodo database and respective content.

Database Level	Folder Name	Content
0	Database_files	Individual spectra and image files, stored by each demonstration site.
1	Microsoft_Access_database	Complete database with sample description and attachments.
2	Geodatabase	Geodatabase files to be displayed in ArcGIS.
3	Geopackage	Geopackage files to be displayed in QGIS, folders containing the attachment files to be linked to each geopackage file, and related tutorial.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

2.1. Sample preparation and data acquisition

All samples were dried at ≈ 50ºC in a muffle furnace to remove any moisture from the sample that can influence the spectral behavior. The measurements were conducted in a dark room using the FieldSpec 4 (ASD Inc., Boulder, CO, USA) standard resolution spectroradiometer covering the spectral range between 350 and 2500 nm with the following resolution: 3.0 nm @ 700 nm, 10.0 nm @ 1400 nm, and 10.0 nm @ 2100 nm (Fig. 2-a). A contact probe with an internal light source provided by a halogen bulb and a spot size of 10 mm was used for the laboratory measurements. Reflectance calibration was achieved with a Spectralon (Labsphere) plate with a maximum reflectance higher than 95% for the 250 to 2500 nm range and higher than 99% for the 400 to 1500 nm range. To increase the signal-to-noise ratio, each measurement comprises an average of 40 scans with four additional measurements acquired in each analysed spot that were later averaged into a final spectrum (Cardoso-Fernandes et al., 2021). For each spot analysed, a description was made regarding the sample color and type of surface (weathered, freshly broken). The location of each measurement was annotated and photographed (Fig. 2-b). For spectral post-processing, a self-proposed Python routine, using the *pysptools* library (Therien, 2013), was applied to eliminate the spectra continuum (normalisation) and to extract automatically the main absorption features and associated statistics (Cardoso-Fernandes et al., 2021).



(a) (b)

Figure 2: (a) FieldSpec 4 (ASD, Inc.) standard resolution spectroradiometer. (b) Example of an annotated photograph of the measurement spot in a sample from Tysfjord. A similar photograph accompanies each analysed spot in the spectral library.

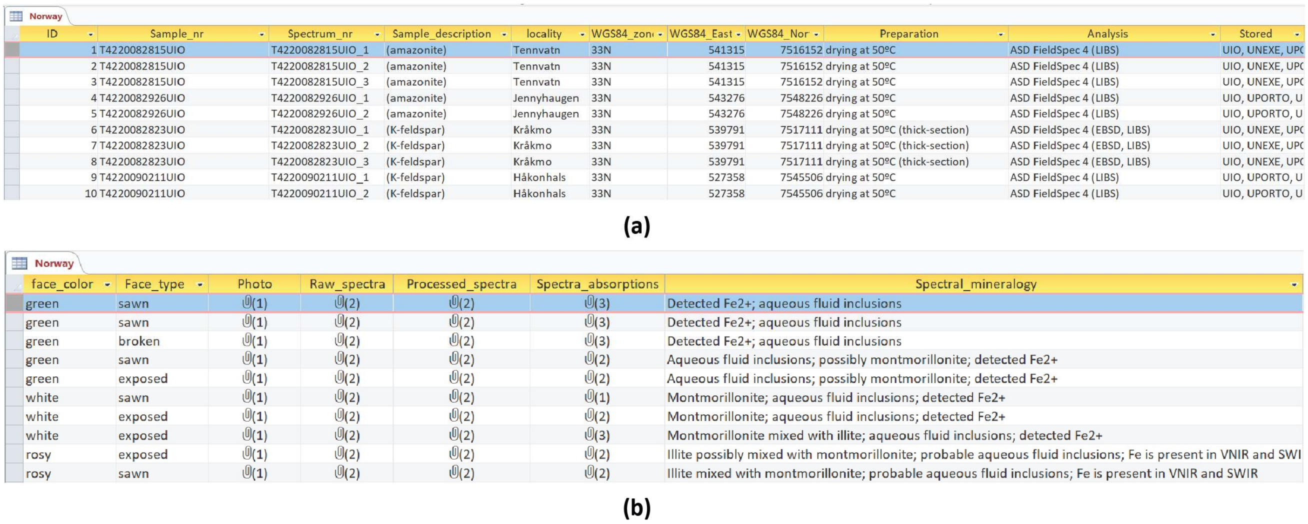
2.2. Microsoft Access database

The database contains a table for each demonstration site and additional tables for the Portuguese and Spanish test sites. Each table follows the recommended nomenclature and content for hard rock samples defined in the GREENPEG Project Management Plan (Greenpeg D1.1, 2020), namely: (i) sample number (nr); (ii) sample description; (iii) locality; (iv) WGS84 Zone; (v) WGS84 Easting coordinate; (vi) WGS84 Northing coordinate; (vii) preparation; (viii) analysis; and (ix) place where the samples are stored (Fig. 3-a). The columns also provide information for each measurement: (i) a photograph of the sample; (ii) the observed sample color; (iii) the type of sample surface; (iv) the raw spectrum (as image and universal text file); (v) the normalised spectrum (either in an image and universal text file); (vi) the automatically identified absorption features; and (vii)

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Open Access Earth System Science Data Discussions

the interpreted spectral mineralogy (Fig. 3-b). Several database columns contain attachments (paper-clip icons) that can be previewed or saved to a local desktop. The number next to these icons indicates the number of available attachments (Fig. 3-b). Additionally, Microsoft Access automatically adds an ID column, where a primary key (number) is attributed to each entry in the database. The content of the database and the type of attachments is summarized in Table 2.


Norway

ID	Sample_nr	Spectrum_nr	Sample_description	locality	WGS84_zon	WGS84_East	WGS84_Nor	Preparation	Analysis	Stored
1	T4220082815UIO	T4220082815UIO_1	(amazonite)	Tennvatn	33N	541315	7516152	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UNEXE, UPC
2	T4220082815UIO	T4220082815UIO_2	(amazonite)	Tennvatn	33N	541315	7516152	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UNEXE, UPC
3	T4220082815UIO	T4220082815UIO_3	(amazonite)	Tennvatn	33N	541315	7516152	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UNEXE, UPC
4	T4220082926UIO	T4220082926UIO_1	(amazonite)	Jennyhaugen	33N	543276	7548226	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UPORTO, U
5	T4220082926UIO	T4220082926UIO_2	(amazonite)	Jennyhaugen	33N	543276	7548226	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UPORTO, U
6	T4220082823UIO	T4220082823UIO_1	(K-feldspar)	Kråkmo	33N	539791	7517111	drying at 50ºC (thick-section)	ASD FieldSpec 4 (EBSD, LIBS)	UIO, UNEXE, UPC
7	T4220082823UIO	T4220082823UIO_2	(K-feldspar)	Kråkmo	33N	539791	7517111	drying at 50ºC (thick-section)	ASD FieldSpec 4 (EBSD, LIBS)	UIO, UNEXE, UPC
8	T4220082823UIO	T4220082823UIO_3	(K-feldspar)	Kråkmo	33N	539791	7517111	drying at 50ºC (thick-section)	ASD FieldSpec 4 (EBSD, LIBS)	UIO, UNEXE, UPC
9	T4220090211UIO	T4220090211UIO_1	(K-feldspar)	Håkonhals	33N	527358	7545506	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UPORTO, U
10	T4220090211UIO	T4220090211UIO_2	(K-feldspar)	Håkonhals	33N	527358	7545506	drying at 50ºC	ASD FieldSpec 4 (LIBS)	UIO, UPORTO, U

(a)

Norway

face_color	Face_type	Photo	Raw_spectra	Processed_spectra	Spectra_absorptions	Spectral_mineralogy
green	sawn	(1)	(2)	(2)	(3)	Detected Fe2+; aqueous fluid inclusions
green	sawn	(1)	(2)	(2)	(3)	Detected Fe2+; aqueous fluid inclusions
green	broken	(1)	(2)	(2)	(3)	Detected Fe2+; aqueous fluid inclusions
green	sawn	(1)	(2)	(2)	(2)	Aqueous fluid inclusions; possibly montmorillonite; detected Fe2+
green	exposed	(1)	(2)	(2)	(2)	Aqueous fluid inclusions; possibly montmorillonite; detected Fe2+
white	sawn	(1)	(2)	(2)	(1)	Montmorillonite; aqueous fluid inclusions; detected Fe2+
white	exposed	(1)	(2)	(2)	(2)	Montmorillonite; aqueous fluid inclusions; detected Fe2+
white	exposed	(1)	(2)	(2)	(3)	Montmorillonite mixed with illite; aqueous fluid inclusions; detected Fe2+
rosy	exposed	(1)	(2)	(2)	(2)	Illite possibly mixed with montmorillonite; probable aqueous fluid inclusions; Fe is present in VNIR and SWI
rosy	sawn	(1)	(2)	(2)	(2)	Illite mixed with montmorillonite; probable aqueous fluid inclusions; Fe is present in VNIR and SWIR

(b)

Figure 3: Database table preview for the Tysfjord demonstration site: (a) recommended content of hard rock sample list and (b) content specifically related to the spectral database (© Microsoft 2022).

Table 2: Content of the database in brief.

Field	Description	Attachments files
Sample number (nr)	Sample identification following GREENPEG's nomenclature	–
Spectrum number	Spectrum number (sample number + number of the analysed spot within the sample)	–
Sample description	Description of the sample (provided by the partners or taken in the field by the authors)	–
Locality	Place where the samples were collected	–
WGS84 Zone	UTM zone	–
WGS84 Easting	X-coordinate in UTM (Easting)	–
WGS84 Northing	Y-coordinate in UTM (Northing)	–
Preparation	Sample preparation (for the spectral library and other parallel studies, the latter between brackets)	–
Analysis	Analytical methods employed (complementary studies are between brackets)	–

Stored	Where the sample and respective duplicates are stored (names represent the project partners)	–
Face color	The sample color in the measured spot	–
Face type	The sample face type in the measured spot	–
Photo	The sample photograph (measured spots are highlighted)	.png/.jpg
Raw spectra	The raw spectra (either in an image and universal text file)	.txt/.pdf
Processed spectra	The continuum removed spectra (either in an image and universal text file)	.txt/.pdf
Spectra absorptions	The automatically identified absorption features	.png/.csv
Spectral mineralogy	Interpreted spectral mineralogy	–

2.2.1. Data interpretation and use

The identified spectral mineralogy of the investigated samples was achieved through the comparison of the measured absorption spectrum with reference material spectra of the United State Geological Survey (USGS) spectral library (Clark, 1999; Clark et al., 2003; Kokaly et al., 2017) and other public libraries (Clark et al., 1990; Hunt, 1977; Pontual et al., 2008). Our results show that the spectral mineralogy identified does not necessarily match the minerals identified by observation of hand specimens and optical microscopy. This is because some silicates do not present necessarily diagnostic absorption features (Spatz, 1997) or because the spectra are dominated by alteration minerals that are spectrally very active due to the presence of water/hydroxyl group and superimpose unaltered mineral domains (Hunt and Ashley, 1979).

The representative reflectance spectra stored in the libraries can be utilised for satellite image processing, namely in the image classification tasks. To do so, the acquired spectra can be resampled to match the satellite sensors' spectral resolution and used as a target for algorithm training instead of the image pixels. This approach is, for example, an alternative satellite image processing method for the Leinster and Wolfsberg demonstration sites, where there are insufficient outcrops of pegmatite in the target area to serve as training areas in the satellite image processing, thus preventing the application of machine learning algorithms (Greenpeg D2.3, 2021). Moreover, by comparing the location of the reflectance anomalies of the reference spectra with the location of the satellite sensor bands, it is possible to select the most indicative bands for image processing and improve the previously employed approaches (Santos et al., 2022). Thus, the spectral library will be extremely helpful for the processing of Worldview-3 images, in particular, due to the increased spatial and spectral resolution.

3. Representative spectra of the demonstration sites

This section selects representative spectra of each demonstration site to demonstrate how the data available in the database can be interpreted with some examples of the spectra figure files.

Spectra of pegmatite samples from Tysfjord (Norway) are characterised by biotite/chlorite features mixed with crystallographic water, microfluid inclusions and alteration minerals, such as montmorillonite and illite (Fig. 4-a, -b). Granitic gneiss samples

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Open Access Earth System Science Data Discussions

are characterised by strong Fe^{2+} ramp-like absorptions caused by biotite/chlorite mixed with white mica, illite and/or montmorillonite (Fig. 4-c, -d).

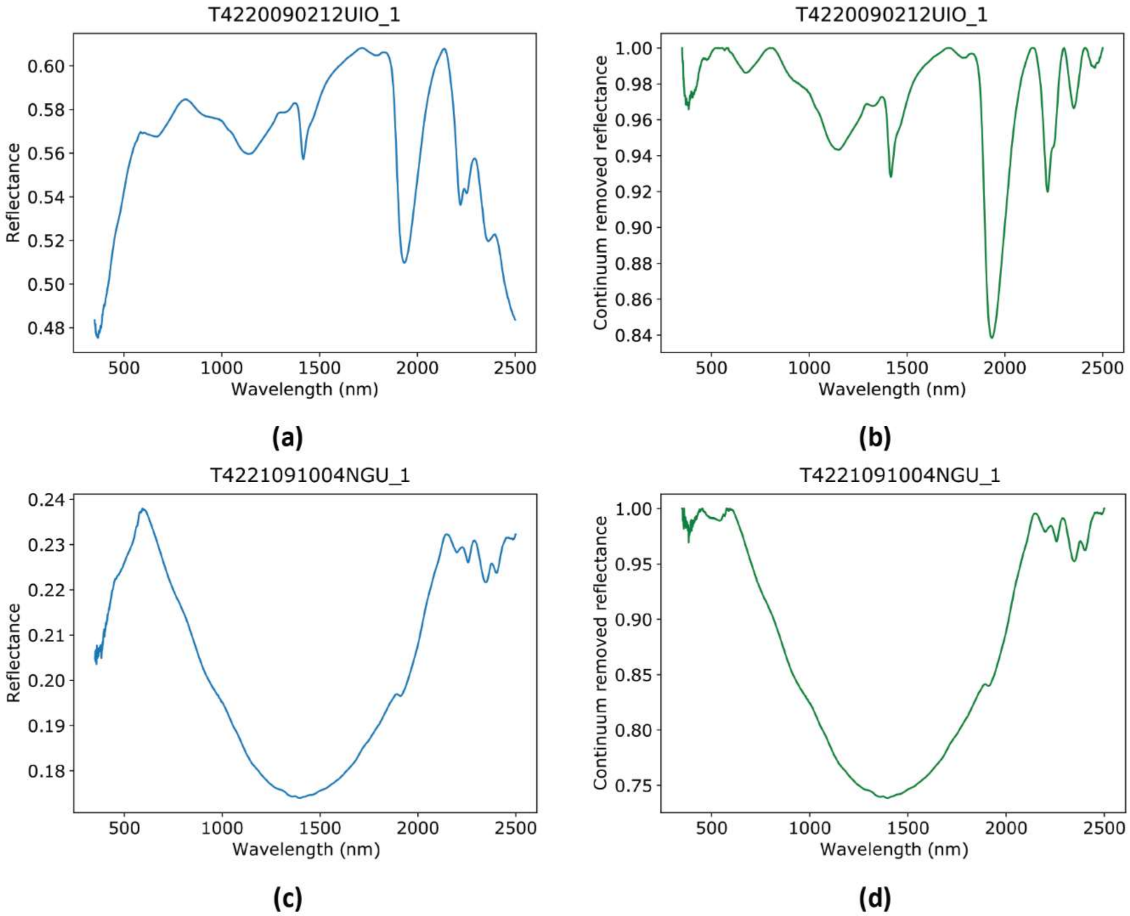


Figure 4: Representative spectra of samples from the Tysfjord demonstration site. Raw (a) and normalised (b) spectra of pegmatitic plagioclase showing features of illite mixed with montmorillonite and biotite and crystallographic water and/or water of microfluid inclusions. Raw (c) and normalised (d) spectra of granitic gneiss displaying a strong ramp-like Fe^{2+} absorption and chlorite/biotite mixed with white mica.

The spectra of pegmatite samples from Leinster (Ireland) are dominated by illite and/or montmorillonite features (Fig. 5-a, -b), sometimes mixed with white mica or orthoclase. Granite samples are characterised by Fe^{2+} ramp-like absorptions caused by biotite/chlorite mixed with white mica, illite and/or montmorillonite (Fig. 5-c, -d). The schist samples have spectral features of white mica or hydrated white mica, probably sericite, mixed with chlorite/biotite (Fig. 5-e, -f). Altered samples are dominated by alteration minerals, such as illite or montmorillonite.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

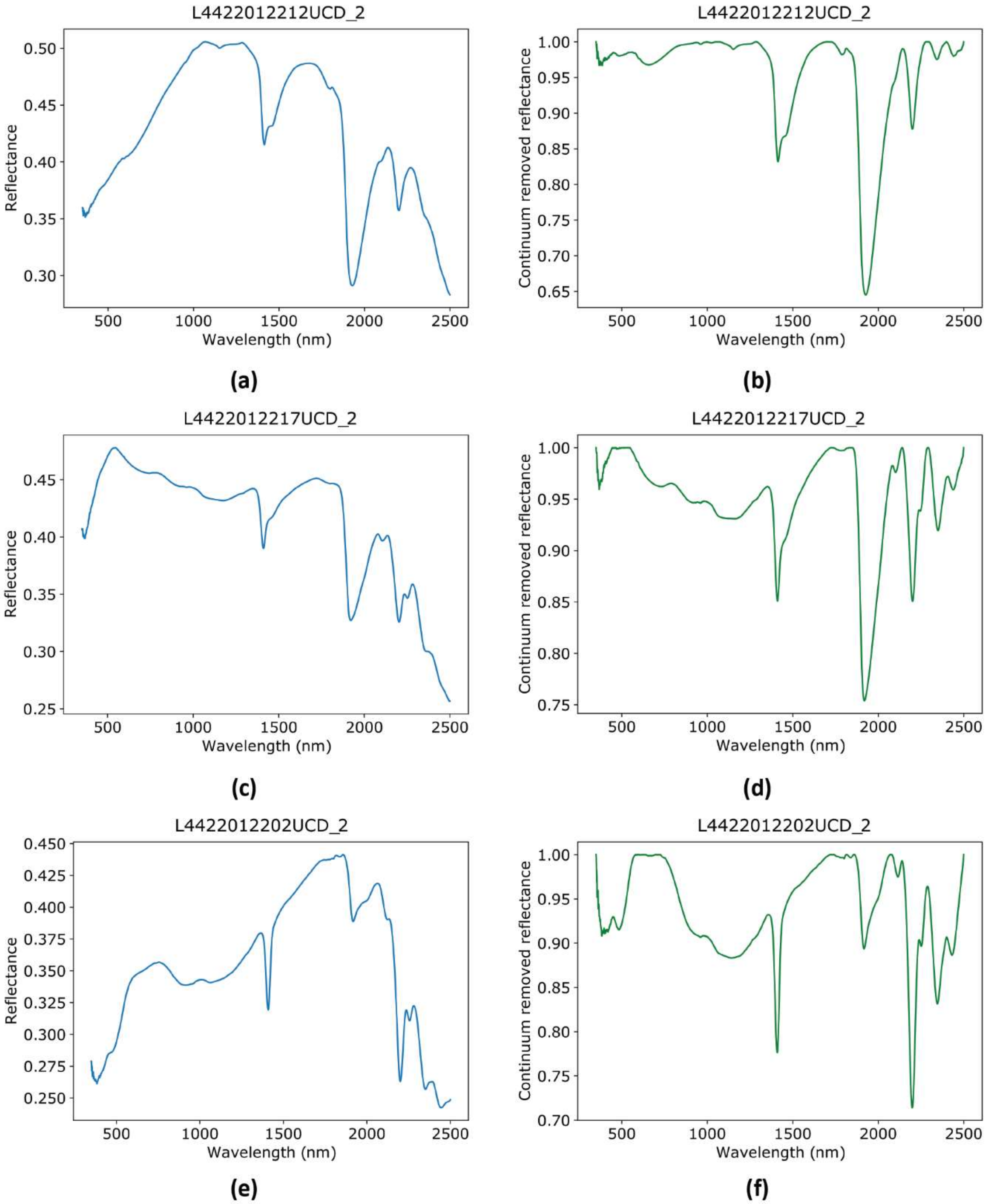


Figure 5: Representative spectra of samples from the Leinster demonstration site. Raw (a) and normalised (b) spectra of an albitised pegmatite sample showing spectral features of montmorillonite mixed with white mica/illite. Raw (c) and normalised (d) spectra of a granite sample showing a ramp-like Fe^{2+} absorption and features of biotite mixed with chlorite, illite and montmorillonite. Raw (e) and normalised (f) spectra of a schist sample showing spectral features of hydrated white mica, probably sericite mixed with chlorite/biotite.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

The spectra of pegmatite samples from Wolfsberg (Austria) are dominated by white mica or hydrated white mica (probably sericite) sometimes mixed with illite or montmorillonite (Fig. 6-a, -b). The pegmatite host rock samples display features diagnostic for white mica, chlorite, biotite, phlogopite and/or carbonates, such as siderite and magnesite (Fig. 6-c, -d).

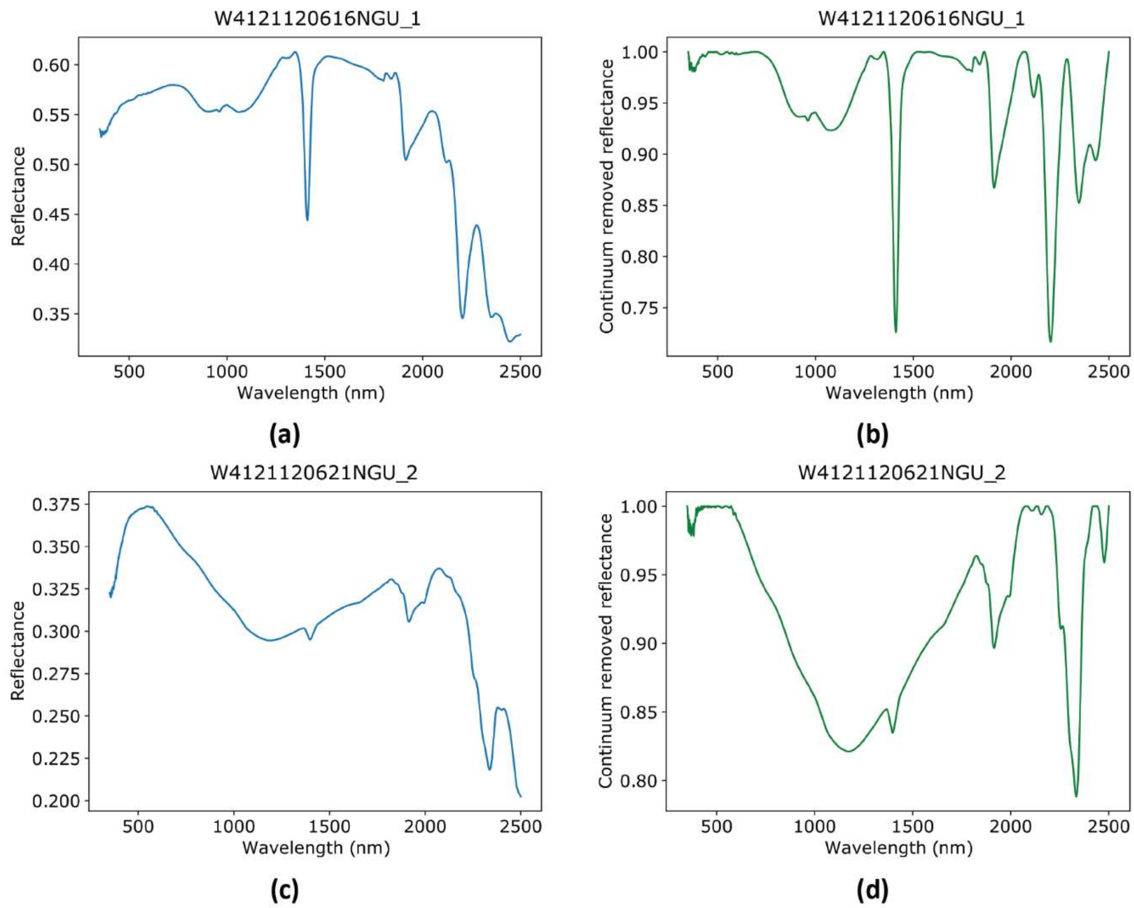


Figure 6: Representative spectra of samples from the Wolfsberg demonstration site. Raw (a) and normalised (b) spectra of a pegmatite sample showing spectral features of hydrated white mica, probably sericite. Raw (c) and normalised (d) spectra of the pegmatite-host rock (probable amphibolite) sample showing ramp-like Fe^{2+} absorption spectral features diagnostic for carbonates (siderite) possibly mixed with chlorite.

4. Interoperability with GIS

Interoperability of the spectral database with a GIS environment is crucial for the successful exploitation of the spectral library and established dataset. Therefore, the spectral database is provided in two additional formats: a geodatabase file format (.gdb) and a geopackage file format (.gpk). The geodatabase file can be opened in ArcMap or ArcGIS Pro (Esri, Redlands, CA, USA) and the attachments are automatically previewed or opened since they are stored within the geodatabase (Fig. 7). The information from the spectral library, including attachments, can be easily previewed using (i) the 'Explore' tool, which opens

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

a pop-up window (left panel of Fig. 7), or (ii) by accessing the 'Attributes' tab (right panel of Fig. 7). In this panel, the user can double-click each file to open them in the system default program according to the file type.

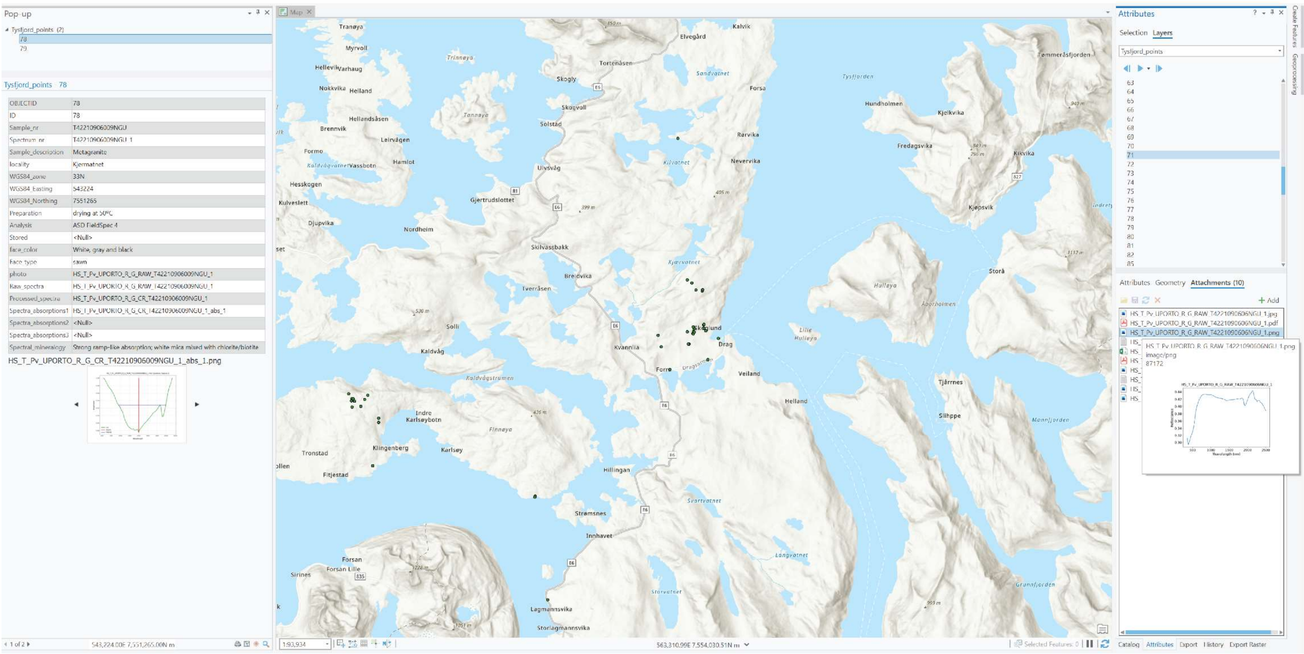


Figure 7: Illustration of how to use the geodatabase in ArcGIS Pro on the example of the Tysfjord demonstration site: the 'Explore' tool is shown on the left panel and the 'Attributes' tab in the right panel. Basemap provided by ESRI.

For QGIS users, the processing is manual and for that, the files are stored externally. For this approach, the geopackage file is provided together with a folder with related attachments. The QGIS user needs to re-link the attachments in the folder with the paths stored in the columns of the layer's attribute table of the geopackage file. A tutorial for this latter step is also provided within the respective database level together with the geopackage file. After linking the files, the attachments can be previewed in QGIS (Fig. 8). The information from the spectral library, including attachments, can be easily previewed using the 'Identify' tool that opens a pop-up window (right panels of Fig. 8). It should be noticed that unlike with the geodatabase, the geopackage file only allows previewing image files (.png, .jpeg, etc.).

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Open Access Earth System Science Data Discussions

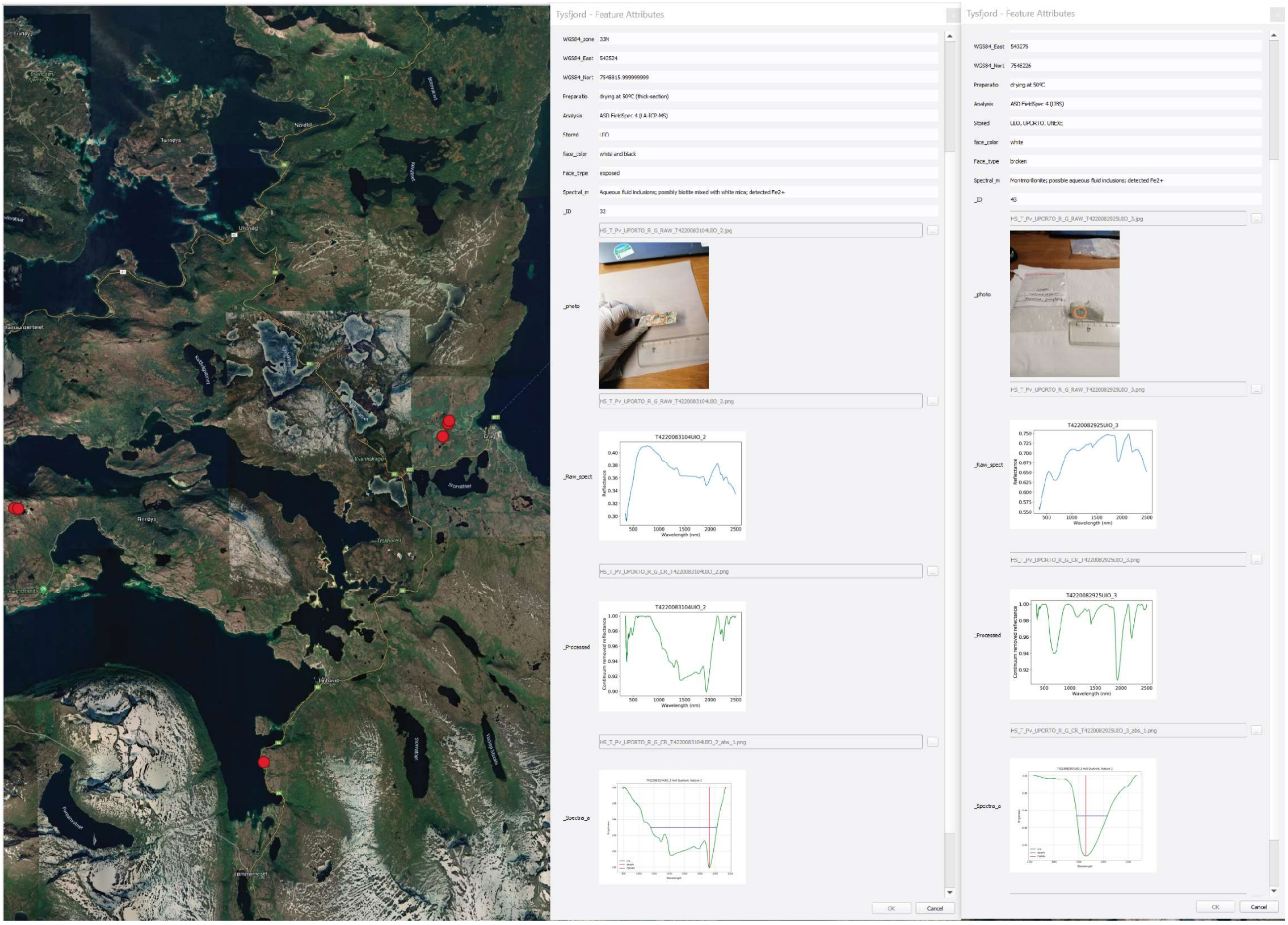


Figure 8: Illustration of how to use the geopackage file in QGIS on the example of the Tysfjord demonstration site: the 'Identify' tool opens a pop-up window with the attachments (right panels). Basemap provided by © Google Earth 2022.

5. Data availability

Since the GREENPEG project is committed to an Open Research Data Pilot (ORDP), the spectral library is openly available on the Zenodo platform (https://www.zenodo.org/communities/greenpeg-project) with the DOI identifier https://doi.org/10.5281/zenodo.6518319 (Cardoso-Fernandes et al., 2022), following the GREENEPG ORDP (Greenpeg D8.1, 2020). All formats (Microsoft Access, geodatabase, geopackage) and individual spectral library files are stored as previously described for easy accessibility and use. File naming and metadata records follow the rules established in the Project Management Plan and Data Management Plan (Greenpeg D1.1, 2020; Greenpeg D8.1, 2020). However, the spectral library version available in the Zenodo platform has confidential and sensitive information redacted, as requested by GREENPEG's industry partners.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

6. Code availability

No original Python code was produced. The Python routine is available as supplementary material to the work of Cardoso-Fernandes et al. (2021).

7. Conclusions

The GREENPEG European spectral database presented in this works aims to add new data on the properties of pegmatites and their green raw materials as well as different host rocks, allowing the evaluation of the potential for discriminating the two. The advantages and added value of the presented dataset reside on its European scale, therefore presenting reference spectra for pegmatites of both NYF and LCT chemical types, and also representative samples from pegmatites with distinct genesis, mineralogy, structure, and host rocks.

This spectral database is also relevant because the results show that the spectral mineralogy identified does not necessarily match the minerals identified by observation of hand specimens and optical microscopy. Such information is crucial for users trying to detect other pegmatites worldwide. The reported spectral mineral assemblages can also be of interest when considering resource estimation or ore processing due to the large information provided for the distinct pegmatites sampled.

Since this dataset is aimed to be used as a reference for pegmatite exploration at a global scale, and distinct types of users are expected to benefit from its content, data usability and accessibility were a priority, thus ensuring a multi-format database and clear interoperability with a GIS environment to fully exploit the data provided.

The reflectance spectra stored in the database can be therefore utilised for satellite image processing such as image classification in the early stages of pegmatite exploration. Currently, the spectral library is being used for the processing of Worldview-3 images over the Tysford (Norway) case study.

8. Appendices

Table A1: Detailed content of the spectral library.

Database	Spectra file name	No. spectra/ files	File type
Tysfjord (Norway)	HS_T_Pv_UPORTO_R_G_RAW_T4220082815UIO_1 and HS_T_Pv_UPORTO_R_G_CR_T4220082815UIO_1 … HS_T_Pv_UPORTO_R_G_RAW_T4221090611NGU_2 and HS_T_Pv_UPORTO_R_G_CR_T4221090611NGU_2	131/1309	.png .jpg .txt .pdf .csv
Leister (Ireland)	HS_L_Pv_UPORTO_R_G_RAW_L4422012201UCD_1 and HS_L_Pv_UPORTO_R_G_CR_L4422012201UCD_1	69/860	

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

	… HS_L_Pv_UPORTO_R_G_RAW_L2321102001UPORTO_2 and HS_L_Pv_UPORTO_R_G_CR_L2321102001UPORTO_2	
Wolfsberg (Austria)	HS_W_Pv_UPORTO_R_G_RAW_W4121120601NGU_1 and HS_W_Pv_UPORTO_R_G_CR_W4121120601NGU_1 … HS_W_Pv_UPORTO_R_G_RAW_W4121120636NGU_2 and HS_W_Pv_UPORTO_R_G_CR_W4121120636NGU_2	76/877
Portugal (Adagói, Alijó, Gonçalo)	HS_B_Pv_UPORTO_R_G_RAW_B4220120902UPV_1 and HS_B_Pv_UPORTO_R_G_CR_B4220120902UPV_1 … HS_P_Pv_UPORTO_R_G_RAW_P23220207008UPORTO_3 and HS_P_Pv_UPORTO_R_G_CR_P23220207008UPORTO_3	34/424
Spain (Fregeneda)	HS_F_Pv_UPORTO_R_G_RAW_F4220092536UPV_1 and HS_F_Pv_UPORTO_R_G_CR_F4220092536UPV_1 … HS_F_Pv_UPORTO_R_G_RAW_F4220121104UPV_2 and HS_F_Pv_UPORTO_R_G_CR_F4220121104UPV_2	18/212

Team list

University of Porto: Joana Cardoso-Fernandes, Ana C. Teodoro, Douglas Santos, Cátia Rodrigues de Almeida, Alexandre Lima. University of Oslo: Axel Müller, William Keyser. Geological Survey of Norway: Kerstin Saalmann, Claudia Haase and Marco Brönner. University College Dublin: Julian Menuge. Geo Unterweissacher GmbH: Ralf Steiner. Universidad del País Vasco-UPV/EHU: Encarnación Roda Robles, Jon Errandonea-Martin and Idoia Garate-Olave.

Author contribution

Conceptualization: JC-F, ACT, AL; Data curation: JC-F, DS, CRdA; Formal analysis: JC-F; Funding acquisition: AL; Investigation: JC-F, DS, CRdA, AL; Methodology: JC-F, ACT; Project administration: AL; Resources: JC-F, ACT, AL; Software: JC-F; Supervision: JC-F, ACT, AL; Validation: JC-F, ACT, DS; Visualization: JC-F; Writing – original draft preparation: JC-F; Writing – review & editing: ACT, DS, CRdA, AL.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This study is funded by European Commission's Horizon 2020 innovation programme under grant agreement no. 869274, project GREENPEG New Exploration Tools for European Pegmatite Green-Tech Resources. The GREENPEG project is coordinated by Axel Müller. The Portuguese partners also acknowledge the support provided by Portuguese National Funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P. (Portugal) projects UIDB/04683/2020 and UIDP/04683/2020 (Institute of Earth Sciences). Cátia Rodrigues de Almeida benefits from a Ph.D. grant provided by FCT with reference PRT/BD/153518/2021. The authors thank European Lithium ECM in Austria and Blackstairs Lithium Limited in Ireland for providing us with samples from their drill cores. Thanks to Axel Müller, Julian Menuge, Kerstin Saalmann, Ralf Steiner, William Keyser, Encarnación Roda Robles, Claudia Haase, and Marco Brönner for collecting and/or sharing the samples.

References

Asch, K.: IGME 5000: 1 : 5 Million International Geological Map of Europe and Adjacent Areas, Federal Institute for Geosciences and Natural Resources (BGR) Hannover, Germany, 2005.

Barros, R., Kaeter, D., Menuge, J. F., and Škoda, R.: Controls on chemical evolution and rare element enrichment in crystallising albite-spodumene pegmatite and wallrocks: Constraints from mineral chemistry, Lithos, 352-353, 105289, https://doi.org/10.1016/j.lithos.2019.105289, 2020.

Barros, R. and Menuge, J. F.: The Origin of Spodumene Pegmatites Associated with the Leinster Granite in Southeast Ireland, The Canadian Mineralogist, 54, 847-862, 10.3749/canmin.1600027, 2016.

Cardoso-Fernandes, J., Teodoro, A. C., Santos, D., Rodrigues de Almeida, C., and Lima, A.: Spectral Library of European Pegmatites, Pegmatite Minerals and Pegmatite Host-Rocks – The Greenpeg Database (Version 1) [dataset], https://doi.org/10.5281/zenodo.6518319, 2022.

Cardoso-Fernandes, J., Silva, J., Dias, F., Lima, A., Teodoro, A. C., Barrès, O., Cauzid, J., Perrotta, M., Roda-Robles, E., and Ribeiro, M. A.: Tools for Remote Exploration: A Lithium (Li) Dedicated Spectral Library of the Fregeneda–Almendra Aplite–Pegmatite Field, Data, 6, 33, 10.3390/data6030033, 2021.

Černý, P. and Ercit, T. S.: Classification of granitic pegmatites revisited, Can. Mineral., 43, 2005–2026, 2005.

Clark, R. N., King, T. V. V., Klejwa, M., Swayze, G. A., and Vergo, N.: High spectral resolution reflectance spectroscopy of minerals, Journal of Geophysical Research: Solid Earth, 95, 12653-12680, 10.1029/JB095iB08p12653, 1990.

Clark, R. N., Swayze, G. A., Wise, R., Livo, K. E., Hoefen, T. M., Kokaly, R. F., and Sutley, S. J.: USGS digital spectral library splib05a: https://pubs.usgs.gov/of/2003/ofr-03-395/, last access: 8 January 2021.

Clark, R. N.: Spectroscopy of rocks and minerals and principles of spectroscopy: Chapter 1, in: Remote Sensing for the Earth Sciences: Manual of Remote Sensing,, 3rd ed., edited by: Ryerson, R. A., 3, John Wiley & Sons, Inc., 1999.

Errandonea-Martin, J., Garate-Olave, I., Roda-Robles, E., Cardoso-Fernandes, J., Lima, A., Ribeiro, M. d. A., and Teodoro, A. C.: Metasomatic effect of Li-bearing aplite-pegmatites on psammitic and pelitic metasediments: Geochemical constraints on critical raw material exploration at the Fregeneda–Almendra Pegmatite Field (Spain and Portugal), Ore Geology Reviews, 150, 105155, https://doi.org/10.1016/j.oregeorev.2022.105155, 2022.

Fabre, C., Ourti, N. E., Mercadier, J., Cardoso-Fernandes, J., Dias, F., Perrotta, M., Koerting, F., Lima, A., Kaestner, F., Koellner, N., Linnen, R., Benn, D., Martins, T., and Cauzid, J.: Analyses of Li-Rich Minerals Using Handheld LIBS Tool, Data, 6, 68, 10.3390/data6060068, 2021.

Göd, R.: The spodumene deposit at "Weinebene", Koralpe, Austria, Mineralium Deposita, 24, 270-278, 10.1007/BF00206391, 1989.

Gourcerol, B., Gloaguen, E., Melleton, J., Tuduri, J., and Galiegue, X.: Re-assessing the European lithium resource potential – A review of hard-rock resources and metallogeny, Ore Geology Reviews, 109, 494-519, https://doi.org/10.1016/j.oregeorev.2019.04.015, 2019.

GREENPEG D1.1: Deliverable 1.1; Project Management Plan – July 2020, GREENPEG Horizon2020 GA 869274, 47, 2020.

GREENPEG D2.3: Deliverable 2.3; New algorithms for satellite image processing, GREENPEG Horizon2020 GA 869274, 29, 2021.

GREENPEG D8.1: Deliverable 8.1; Communication Plan, Data Management Plan and Project Identity kit, GREENPEG Horizon2020 GA 869274, 39, 2020.

Haase, C., Brönner, M., Pohl, C. M., Osinska, M., and Gellein, J.: EuroPeg_PetroDB: A petrophysical database of European pegmatite ores and wall rocks [dataset], 10.5281/ZENODO.6497985, 2022.

Hunt, G. R. and Ashley, R. P.: Spectra of altered rocks in the visible and near infrared, Economic Geology, 74, 1613-1629, 10.2113/gsecongeo.74.7.1613, 1979.

Hunt, G. R.: Spectral signatures of particulate minerals in the visible and near infrared, Geophysics, 42, 501-513, 10.1190/1.1440721, 1977.

Knoll, T., Schuster, R., Huet, B., Mali, H., Onuk, P., Horschinegg, M., Ertl, A., and Giester, G.: Spodumene Pegmatites and Related Leucogranites from the AustroAlpine Unit (Eastern Alps, Central Europe): Field Relations, Petrography, Geochemistry, and Geochronology, The Canadian Mineralogist, 56, 489-528, 10.3749/canmin.1700092, 2018.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Kokaly, R. F., Clark, R. N., Swayze, G. A., Livo, K. E., Hoefen, T. M., Pearson, N. C., Wise, R. A., Benzel, W. M., Lowers, H. A., Driscoll, R. L., and Klein, A. J.: USGS Spectral Library Version 7, Reston, VA, Report 1035, 68, 10.3133/ds1035, 2017.

Lima, A.: Estrutura, mineralogia e génese dos filões aplitopegmatíticos com espodumena da região do Barroso-Alvão, PhD thesis, Universidade do Porto, Porto, Portugal, 270 pp., 2000.

Martins, T.: Multidisciplinary study of pegmatites and associated Li and Sn-Nb-Ta mineralisation from the Barroso-Alvão region, PhD thesis, Faculdade de Ciências da Universidade do Porto, Porto, Portugal, 2009.

Müller, A., Reimer, W., Wall, F., Williamson, B., Menuge, J., Brönner, M., Haase, C., Brauch, K., Pohl, C., Lima, A., Teodoro, A., Cardoso-Fernandes, J., Roda-Robles, E., Harrop, J., Smith, K., Wanke, D., Unterweissacher, T., Hopfner, M., Schröder, M., Clifford, B., Moutela, P., Lloret, C., Ranza, L., and Rausa, A.: GREENPEG - Exploration for pegmatite minerals to feed the energy transition: First steps towards the Green Stone Age, Geological Society, London, Special Publications, 526, SP526-2021-2189, 10.1144/SP526-2021-189, 2022a.

Müller, A., Romer, R. L., Augland, L. E., Zhou, H., Rosing-Schow, N., Spratt, J., and Husdal, T.: Two-stage regional rare-element pegmatite formation at Tysfjord, Norway: implications for the timing of late Svecofennian and late Caledonian high-temperature events, International Journal of Earth Sciences, 111, 987-1007, 10.1007/s00531-022-02166-5, 2022b.

Neiva, A. M. R. and Ramos, J. M. F.: Geochemistry of granitic aplite-pegmatite sills and petrogenetic links with granites, Guarda-Belmonte area, central Portugal, European Journal of Mineralogy, 22, 837-854, 10.1127/0935-1221/2010/0022-2072, 2010.

Pontual, S., Merry, N. J., and Gamson, P.: Spectral interpretation field manual. GMEX Spectral analysis guides for mineral exploration, 3rd, AusSpec International Ltd., Victoria2008.

Roda-Robles, E., Gil-Crespo, P. P., Pesquera, A., Lima, A., Garate-Olave, I., Merino-Martínez, E., Cardoso-Fernandes, J., and Errandonea-Martin, J.: Compositional Variations in Apatite and Petrogenetic Significance: Examples from Peraluminous Granites and Related Pegmatites and Hydrothermal Veins from the Central Iberian Zone (Spain and Portugal), Minerals, 12, 1401, 10.3390/min12111401, 2022.

Roda-Robles, E., Pesquera, A., Velasco, F., and Fontan, F.: The granitic pegmatites of the Fregeneda area (Salamanca, Spain): characteristics and petrogenesis, Mineral. Mag., 63, 535–558, 1999.

Roda-Robles, E., Villaseca, C., Pesquera, A., Gil-Crespo, P. P., Vieira, R., Lima, A., and Garate-Olave, I.: Petrogenetic relationships between Variscan granitoids and Li-(F-P)-rich aplite-pegmatites in the Central Iberian Zone: Geological and geochemical constraints and implications for other regions from the European Variscides, Ore Geology Reviews, 95, 408-430, https://doi.org/10.1016/j.oregeorev.2018.02.027, 2018.

https://doi.org/10.5194/essd-2022-386
Preprint. Discussion started: 29 November 2022

Santos, D., Cardoso-Fernandes, J., Lima, A., Müller, A., Brönner, M., and Teodoro, A. C.: Spectral Analysis to Improve Inputs to Random Forest and other Boosted Ensemble Tree-Based Algorithms for Detecting NYF Pegmatites in Tysfjord, Norway, Remote Sensing, 14, 3532, 2022.

Spatz, D. M.: Remote sensing characteristics of the sediment- and volcanic-hosted precious metal systems: Imagery selection for exploration and development, International Journal of Remote Sensing, 18, 1413-1438, 10.1080/014311697218205, 1997.

Teodoro, A. C., Santos, D., Cardoso-Fernandes, J., Lima, A., and Brönner, M.: Identification of pegmatite bodies, at a province scale, using machine learning algorithms: preliminary results, Proc. SPIE 11863, Earth Resources and Environmental Remote Sensing/GIS Applications XII, SPIE Remote Sensing, 12 September 2021, https://doi.org/10.1117/12.2599600, 2021.

Therien, C.: Welcome to the PySptools Documentation: https://pysptools.sourceforge.io/, last access: 22 July 2020.

Vieira, R.: Aplitopegmatitos com elementos raros da região entre Almendra (V.N. de Foz Côa) e Barca d'Alva (Figueira de Castelo Rodrigo). Campo aplitopegmatítico da Fregeneda- Almendra, PhD thesis, Faculdade de Ciências da Universidade do Porto, Porto, Portugal, XXVI, 273 pp., 2010.

Vieira, R., Roda-Robles, E. n., Pesquera, A., and Lima, A.: Chemical variation and significance of micas from the Fregeneda-Almendra pegmatitic field (Central-Iberian Zone, Spain and Portugal), American Mineralogist, 96, 637-645, 10.2138/am.2011.3584, 2011.