Ore Geology Reviews 175 (2024) 106347

Contents lists available at ScienceDirect

Ore Geology Reviews

journal homepage: www.elsevier.com/locate/oregeorev

Review

Optimizing Exploration: Synergistic approaches to minimize false positives in pegmatite prospecting – A comprehensive guide for remote sensing and mineral exploration

Douglas Santos [a,b], Antônio Azzalini [a], Ariane Mendes [a], Joana Cardoso-Fernandes [a,b], Alexandre Lima [a,b], Axel Müller [c,d], Ana C. Teodoro [a,b,*]

[a] Department of Geosciences, Environment and Spatial Planning, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal
[b] Institute of Earth Sciences, FCUP Pole, 4169-007 Porto, Portugal
[c] Natural History Museum, University of Oslo, 0318 Oslo, Norway
[d] Natural History Museum, London SW7 5BD, UK

ARTICLE INFO

Keywords:
Remote sensing
Machine learning
Fuzzy logic
LightGBM
MTMF
Mineral prospectivity mapping
Critical raw materials

ABSTRACT

Pegmatites represent a critical source of raw materials for the European Union's future, emphasizing the need to explore new deposits. Remote sensing plays a pivotal role in early-stage pegmatite exploration, generating prospectivity maps to facilitate field validation. The higher the false positive rate, the higher the time and costs required for field campaigns. This study aims to devise an approach that minimizes false positives, optimizing the work of the exploration field team. The proposed method seeks to create mineral prospectivity maps that accurately pinpoint high potential areas for pegmatite occurrences with minimal false positives occurrence through the application of fuzzy logic and leveraging the best mapping methods identified in prior research, in the Tysfjord region, northern Norway. Data from two multispectral satellites, namely WorldView-3 (WV3) and Sentinel-2, as well as light detection and ranging (LiDAR) data point cloud, were used, the last allowing the creation of a high-resolution hillshade map. Two classification methods were applied to the satellite images to leverage their individual advantages while minimizing possible weaknesses: the Mixture Tunned Matched Filtering (MTMF) using the Spectral Hourglass Wizard (SHW) Workflow and boosting through the LightGBM (LGBM) algorithm. These classification methods were employed to identify potential points for pegmatite exploration as these rocks have gained economic importance for being sources of raw materials such as high-purity quartz, ceramic feldspars, and Rare Earth Elements (REE). The high-resolution hillshade map was used to extract geological structures in the study area. The results of the fuzzy logic approach indicate potential locations of interest for pegmatite prospecting, providing a more comprehensive analysis of the remote sensing methods in the Tysfjord area. The resulting map seamlessly integrates into reports, streamlining field validation and supports informed decision-making. The methodology proposed in this study can be adaptable to other targets (minerals and rocks) and can be used as a guide for exploration worldwide.

1. Introduction

Pegmatites play a crucial role as economically significant sources of various rare elements and for sustainable development (Chaves et al., 2021; Gourcerol et al., 2019). They are characterized as coarse-grained magmatic rocks, featuring crystals that exceed 2 cm in size (London, 2016). Additionally, these rocks host minerals of industrial importance, including potassium and sodium feldspar, quartz, and mica. According to Cerny (1991), rare-element pegmatites can be classified into two families: Niobium, Yttrium, Fluorine (NYF), and Lithium, Cesium, and Tantalum (LCT). Therefore, identifying novel pegmatite deposits stands as a pivotal step in the exploration process, where remote sensing techniques emerge as a pivotal tool, particularly in the preliminary phases of mineral exploration (Chirico et al., 2023; Sabins, 1999; Zhao et al., 2021). Remote sensing techniques provide an efficient means for detecting pegmatite outcrops over extensive terrains, offering a non-intrusive alternative to conventional methodologies (Cardoso-Fernandes et al., 2019; Ding et al., 2023).

* Corresponding author.
E-mail addresses: douglas.santos@fc.up.pt (D. Santos), amteodor@fc.up.pt (A.C. Teodoro).

https://doi.org/10.1016/j.oregeorev.2024.106347
Received 29 March 2024; Received in revised form 8 November 2024; Accepted 13 November 2024
Available online 17 November 2024

Recent studies in the realm of remote sensing applications for pegmatite exploration have yielded noteworthy contributions, including the application of machine learning algorithms (Abdelkader et al., 2022; Jiang et al., 2023), supervised classification methods (Gemusse et al., 2023; Mashkoor et al., 2022), spectral data collection (Bai et al., 2024; Cardoso-Fernandes et al., 2023a), and the synergistic use of techniques for lithological units and lineament structures mapping (Ali et al., 2023; Dao et al., 2022; Forson et al., 2021; Yamusa et al., 2018). However, these methods are predominantly tailored to Lithium-Cesium-Tantalum pegmatite studies. Regarding the prospecting of NYF pegmatites, key work has been conducted by Santos et al. (2022b), who adapted different image processing methods, such as Band Ratio (BR) and tree-based machine learning algorithms, to map NYF-type pegmatites. In Tysfjord, northern Norway, obstacles to the use of remote sensing data for pegmatite exploration have been identified, such as vegetation cover, subpixel occurrence, spectrally similar rocks (granite), and false positives related to urban areas. Previous studies (Santos et al., 2022b; Santos et al., 2023b) have revealed additional factors in Tysfjord, that can exacerbate false positives, including snow and coastal areas. Despite efforts to refine powerful processing methods like boosting through the LightGBM – (Santos et al., 2022b) and Spectral Unmixing (Santos et al., 2023b), these obstacles persist in Tysfjord. Furthermore, the relatively small size of European pegmatite bodies complicates their identification using open satellite data, where pixel size often exceeds the spatial extent of outcropped pegmatites (Santos et al., 2023b).

While spectral analyses can assist in distinguishing pegmatites from granites, and spectral unmixing has shown effectiveness in mitigating subpixel occurrences, addressing false positives in Tysfjord remains an ongoing problem to solve for both spectral unmixing and machine learning algorithms. The issues related to false positives can be divided into three sets: (i) Distinguishing false positives associated with materials sharing a chemical composition similar to pegmatites, like granites, proves to be challenging in satellite imagery. Addressing false positives with granite often necessitates on-site fieldwork, incurring both time and financial investments, (II) Materials exhibiting high albedo, such as snow, slopes, and coastal lines, contribute to false positives. Processing analyses employing a snow mask effectively eliminate false positives related to snow (Adiri et al. 2020; Santos et al. 2022). However, mitigating false positives in coastal areas and slopes proves to be more complex, requiring a meticulous manual analysis. This detailed process demands a significant time commitment from the analyst; (iii) False positives arising from urban features, including houses, roads, and streets, present another challenge. Although conventional urban masks can be employed to mask such false positives, they also risk masking pegmatite occurrences. Past studies show that while Mixture Tuned Matched Filtering (MTMF) demonstrates effectiveness in minimizing false positives with granite, signal confusion persists in both MTMF and boosting through the LightGBM (LGBM) classifications regarding urban areas, slopes, and coastal lines (Santos et al., 2022b; Santos et al., 2023b).

The main objective of this study is to develop a prospecting map for NYF pegmatites in the Tysfjord area, aiming to minimize (or eliminate) false positives to ensure the reliability of the results. The fuzzy logic approach has been employed to integrate various types of data with remote sensing for generating prospecting maps, including geochemical and geophysical data, along with basic remote sensing processing methods (Abdelkareem and Al-Arifi, 2021; Boadi et al., 2022; Forson et al., 2021; Ghoneim, 2018). However, limited research has been conducted on NYF pegmatite prospecting, often relying on data with prior knowledge of the study area, which excludes greenfield sites. Thus, the question arises: what steps should be taken to create prospecting maps with minimal false positives that can effectively be applied to NYF pegmatite exploration in greenfield sites? This study introduces a novel fuzzy logic-based approach to mineral prospectivity mapping, intending to provide a clear and objective indication of potential pegmatite locations while minimizing false positives. All input layers used in this study were tested in previous works (Santos et al., 2023b; Santos et al., 2023a, 2022b) and are products of remote sensing that can be obtained without extensive prior knowledge of the study area. These products include Worldview-3 (WV-3) and Sentinel-2 data, and LiDAR point cloud data, which was used to generate the Digital Terrain Model (DTM) and Line Density Map (LDM).

Through a synergistic integration of distinguishing input layers, this study utilizes Sentinel-2 and WV-3 satellite data, along with spectral analyses on laboratory-collected spectral data and satellite imagery. The input layers consist of MTMF, LGBM, and LDM. Each method has its strengths and weaknesses, and by combining them, it is possible to leverage their respective advantages to overcome limitations and enhance the final results. MTMF may excel at certain aspects of classification (subpixel classification, abundance mapping), while LGBM could provide additional insights or address different aspects of the data that MTMF might miss. This synergistic approach helps to ensure a more comprehensive and accurate analysis of the data, ultimately leading to a more robust exploration strategy. The resulting prospectivity maps, corresponding to each satellite data type, offer significant potential contributions to the exploration of NYF pegmatites in Tysfjord, providing valuable insights for further research and practical applications.

2. Geological setting

The study area is situated in Nordland, northern Norway, specifically within the Tysfjord pegmatite field, identified as an NYF-type pegmatite field (Fig. 1). According to Zhou et al. (2023), more than 30 NYF pegmatites are hosted by the Tysfjord granite gneiss. These pegmatites predominantly consist of potassium feldspar, plagioclase, biotite, and quartz. Notable accessory minerals encompass allanite-(Ce), columbite-(Fe), gadolinite-(Y), fergusonite-(Y), thalénite-(Y), yttrium-rich fluorite, zircon, beryl, and various sulfides. Within this pegmatite field, we highlight two main bodies: Jennyhaugen and Håkonhals (Fig. 1).

In terms of historical pegmatite mining, activities in the Tysfjord area date back to 1906, persisting until around 1970. The primary focus during this period was on extracting ceramic feldspar. However, a notable shift occurred in 1996, marking a transition towards the extraction of high-purity quartz used for photovoltaics, semiconductors, and the manufacturing of quartz glass.

The Tysfjord pegmatites and granite gneisses are exposed in the tectonic Tysfjord basement window, belonging to the Svecofennian basement of the ancient Baltic continent (Müller et al. 2022). The window is surrounded by the complexes of Caledonian nappes. The Svecofennian basement underwent high-grade metamorphism up to amphibolite facies during the Caledonian collisional orogeny resulting in the recrystallization and deformation of the granites and pegmatites (Zhou et al., 2022).

In the genetic context, the pegmatites are intrinsically linked to the Tysfjord granite gneiss, which belongs geotectonically to the Trans-Scandinavian Igneous Belt – TIB (Fig. 1b). The magmatism of the TIB underwent two distinct stages (Andersson et al., 2004). The first stage occurred between 1.81 and 1.77 Ga ago and the second stage took place between 1.71 and 1.67 Ga ago (Romer et al. 1992). Based on their deformation, age, and size, two types and generations of pegmatites were identified in the area (Müller et al. 2022). The first generation, denominated as “meta-pegmatites” (Husdal, 2008; Müller et al., 2022), comprises large and lenticular bodies, dating back to the Paleoproterozoic (1772–1755 Ma). This generation represents residual melts of the first stage of TIB granites and underwent deformation during the Caledonian collisional orogeny, evidenced by pegmatite shearing and mineral recrystallization as highlighted by Zhou et al. (2022). While it is possible to visualize the contact between the pegmatite bodies and the granite gneiss, the primary magmatic zoning of the *meta*-pegmatites, including the border, wall, intermediate, and core zones, is locally preserved (Müller et al. 2022). This occurs due to intense deformation,

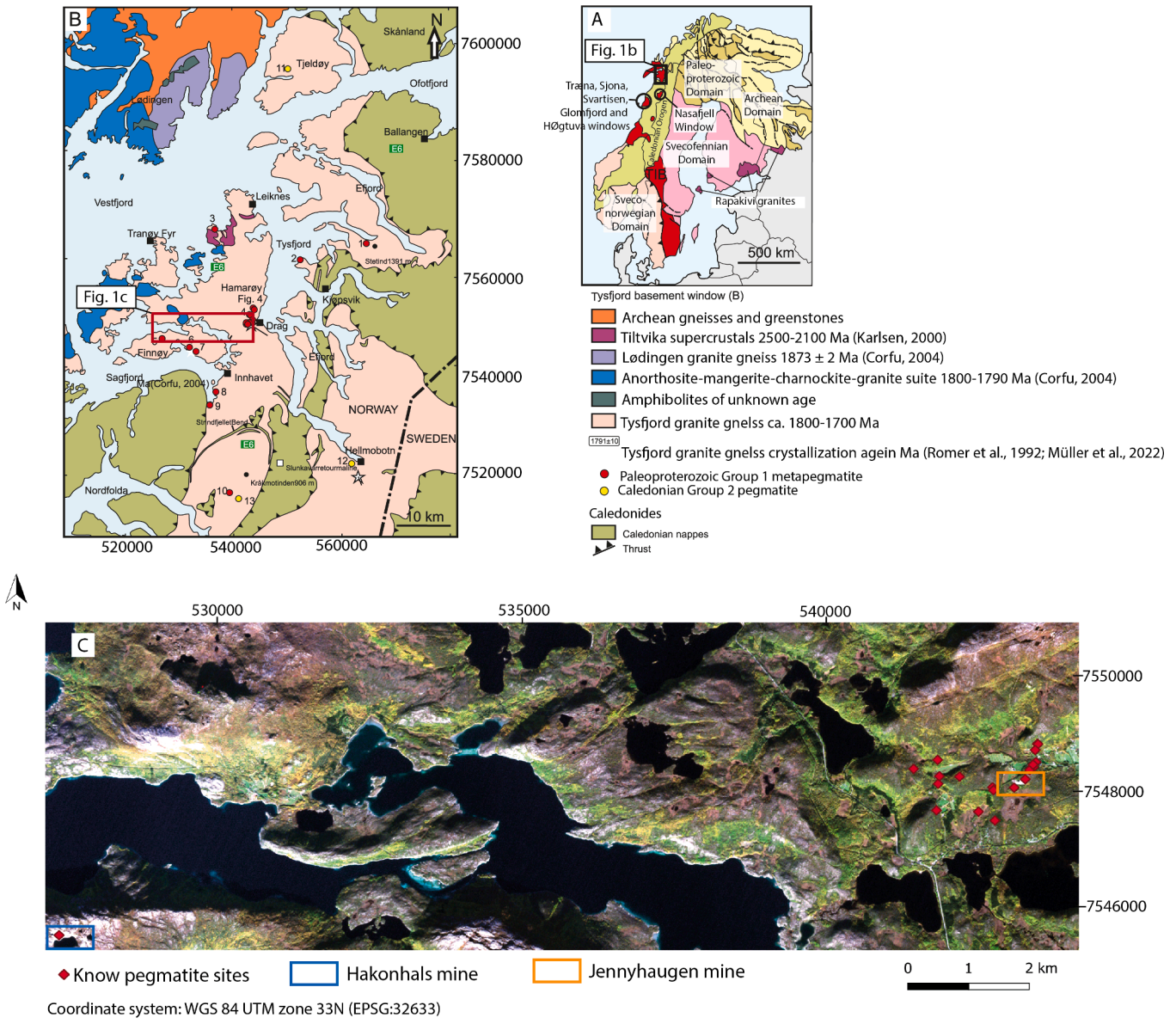

Fig. 1. Location of the study area. A) Detailed geological map and location of the study area in Norway. B) Geologic map of Tysfjord. C). Sentinel-2 true color composite of the study area. The Håkonhals pegmatite is marked with a blue rectangle, Jennyhaugen pegmatite with an orange rectangle, and other pegmatites with red rhombus. TIB: Trans-Scandinavian Igneous Belt.
Adapted from Müller et al (2022).

often resulting in the absence of a linear geological structure but rather in ductile reactions, meaning irregular, rounded, and undulating structures. The pegmatites used as ground truth, Jennyhaugen and Håkonhals (Fig. 1c), fit into this classification. The second generation of pegmatites has an emplacement age that ranges from 400 to 380 Ma (Müller et al. 2022). Originating from anatectic melts and emplaced during the Caledonian post-collisional extension, this generation remains undeformed (Müller et al. 2022). They form sheet to dyke-like bodies with sharp contacts that cut across the foliation of the host granite gneiss. These details underscore the temporal complexity and diverse geological conditions that influenced the formation of these distinct generations of pegmatites in Tysfjord.

According to Müller et al. (2022), the Proterozoic Jennyhaugen pegmatite corresponds to a lens-shaped body, approximately 40 m wide and at least 200 m long. The body is parallel to the foliation of the host rock, Tysfjord granite gneiss, and has a dip of 70°. Its mineral assemblage includes quartz, biotite, feldspars and the color variety amazonite, as well as accessory minerals like allanite-(Ce), garnet, fluorite, monazite-(Ce), and zircon, among others.

As for the Proterozoic Håkonhals pegmatite, it has a dip of 36° and boasts dimensions of approximately 200 x 400 x 25 m, earning recognition as the world's largest intraplutonic NYF pegmatite (Zhou et al., 2023). The pegmatite displays significant ductile deformation, and although the contacts with the Tysfjord granite are distinct, extreme deformation led to pronounced mylonitization, with the primary magmatic zoning preserved locally (Müller et al., 2022). The predominant mineral composition in the Håkonhals pegmatite includes quartz, potassium feldspar, plagioclase, and biotite. Furthermore, different varieties of yttrofluorite are identified, alongside various accessory minerals such as beryl, thorite, gadolinite, bastnäsite-(Ce), allanite-(Ce), monazite-(Y), xenotime-(Yb) and kainosite-(Y).

The study area is home to more than 30 identified pegmatite sites, including prominent locations like Håkonhals and Jennyhaugen. Examining Fig. 1c reveals that nearly all these known pegmatite sites are concentrated in the eastern sector of the study area, right where the Jennyhaugen pegmatite is situated. Among the 30 pegmatite sites, eight

of them showcase visible outcrops detectable through orbital satellite imagery, such as WV3, in the study area.

3. Methodology

3.1. Data

Data from two multispectral sensors were used in this study, namely WV-3 from Maxar Technologies and Sentinel-2 from the European Space Agency. The LiDAR point cloud data, which was collected and classified by Terratec AS was used to generate the DTM.

A WV-3 image was acquired on May 23, 2022. The WV-3 satellite was launched in 2014 on an Atlas-V 401 vehicle designed by Lockheed Martin Commercial Launch Services (LMCLS). The platform carries the high-resolution WV-3 Imager Multispectral sensor, which has 8 bands between 400 nm and 1040 nm in the VNIR, with a spatial resolution of 1.24 m at nadir, and 8 bands between 1195 nm and 2365 nm in the SWIR, with a spatial resolution of 3.7 m at nadir.

A Sentinel-2 image collected on 28 September 2019 was acquired from the GREENPEG project database. The Copernicus Sentinel-2 mission comprises a constellation of two identical satellites, the Sentinel-2A and Sentinel-2B, launched in 2015 and 2017, respectively, both on a Vega Rocket. The multispectral sensor operates in push broom mode and has a swath width of 290 km. The instrument comprises a total of 13 spectral bands. There are three bands situated in the Visible region range, spanning from 468 nm to 680 nm, each featuring a spatial resolution of 10 m. In the Red Edge region, three bands extend from 698 nm to 793 nm, providing a Ground Sample Distance (GSD) of 20 m. Additionally, the Near Infrared (NIR) region is covered by two bands ranging from 785 nm to 875 nm, offering spatial resolutions of 10 m and 20 m. Lastly, the SWIR region is represented by two bands spanning from 1565 nm to 2280 nm, with a GSD of 20 m. Bands 1 (coastal aerosol), 9 (water vapor) and 10 (cirrus) have no surface information.

Two sets of satellite data were chosen because each offers unique advantages that enhance our analysis. Despite WV-3 providing superior spectral and spatial resolution, which is crucial for capturing smaller geological features and improving pegmatite detection accuracy, WV-3 data is not open source. Testing our method with Sentinel-2 allows us to demonstrate its applicability using open-source data. Sentinel-2, with its broader accessibility, enables the method to be more widely applicable and feasible for various users, extending the study's relevance beyond proprietary data constraints.

The airborne laser scan LiDAR data, collected by Terratec AS, consisted of a classified point cloud with a minimum point density of 2 points/m^2. The LiDAR acquisition mission happened in August 2016 and used a Leica ALS80 sensor to capture the data mounted on a Cessna C208 aircraft. The mission covered an area of 598,4 km^2.

The reflectance spectral data used in this study were obtained through the GREENPEG Spectral Library, which is made available in open access (Cardoso-Fernandes et al., 2022).

3.2. Software and equipment

The study utilized PCI Geomatica (version 2018) to assess class separability in the classification process for the LGBM algorithm and to extract lineaments. ENVI (version 5.6) software supported satellite image pre-processing, including layer stacking, pixel size resampling, and atmospheric correction, as well as image processing via the HSW workflow. ENVI was also employed in post-processing steps (e.g., color slicing), extracting pixel spectra, and resampling lab-acquired hyperspectral spectra to match the resolution of multispectral satellite images. The Python programming language was used to develop and implement the LGBM algorithm. Lastly, ArcGIS Pro (version 3.2) enabled pixel data extraction for Python processing and contributed to fuzzification processes, including fuzzy membership and overlay analyses.

The spectral data retrieved from the GREENPEG spectral library were collected with the support of Analytical Spectral Devices (ASD). The ASD FieldSpec 4 spectroradiometer is a transportable battery-powered spectrometer with a spectral range of 350–2500 nm, a spectral resolution of 3 nm at 700 nm (VNIR), 10 nm at 1400 nm (SWIR 1), and 10 nm at 2100 nm (SWIR 2), with a scanning time of 100 ms (Cardoso-Fernandes et al., 2022). The analyses of laboratory spectra were made through OriginPro (version 2018) software.

3.3. Satellite data pre-processing

The layer stacking tool in ENVI was used on the images from both satellites (Sentinel-2 and WV3) to merge the VNIR and SWIR bands into a single layer and resample the pixel size. As the bands have different spatial resolutions, all the Sentinel-2 bands were resampled to 10 m and all the WV-3 bands were resampled to 2 m using the function nearest neighbor. After the stacking, the images from both satellites the images were converted into surface reflectance using Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) atmospheric correction.

3.4. Analysis of laboratory reflectance spectra

Spectral analysis such as reflectance spectroscopy is an important tool in remote sensing and mineral exploration, allowing us to connect spectral data obtained in the laboratory with orbital satellite sensor data (Santos et al., 2022; Cardoso-Fernandes et al. 2023a). Many geological studies were conducted using reflectance spectral analyses (Badr, 2021; Santos et al., 2022b; Sekandari et al., 2020; Roger, 1999). Recently, Cardoso-Fernandes et al. (2023b) conducted research based on acquiring, processing, and organizing spectral data of pegmatite minerals in a spectral library for several study areas including Tysfjord (Cardoso-Fernandes et al., 2023b). This research exploits information from this database and follows the method applied by Santos et al. (2022), who use spectral analyses to select the most important bands to be used as input in the image processing methods.

First, spectral data of the principal minerals that compose the ground truth sites (Håkonhals and Jennyhaugen mines) were analyzed. The minerals chosen were biotite, plagioclase, K-feldspar, and quartz. These spectra were collected using the ASD Fieldspec 4 spectrometer and are available in GREENPEG's spectral library. Here we should point out that the samples can be divided into two types. The large monomineralic samples and the samples in which the grain size is smaller, and there is a variety of minerals within the field of view (FOV) of the spectrometer. The GREENPEG samples whose spectra are available in the spectral library are mostly of the smaller grain size type, i.e., with a variety of minerals representative of the pegmatite or granite body. This means that by passing the pegmatite samples through the spectroradiometer, the information will be collected on the minerals present in the equipment's FOV. This generates a mixed spectrum, which is analyzed later (Annex 1). After, the absorption features and reflectance peaks were identified and evaluated. The minerals in question are representative of the pegmatites in Tysfjord. Using minerals as pathfinders is a very common approach in remote sensing applied to mineral exploration (Rajan Girija and Mayappan, 2019; RAJESH, 2004). Those spectra were analyzed with the continuum removed (already provided in the database). The continuum removal process, in the broad band technique, ensures that the absorptions are more prominent, providing a more efficient analysis and making the data comparable with other study cases (Cardoso-Fernandes et al., 2022).

On the other hand, to compare those spectra with the multispectral sensors data, both absorptions and reflectance peaks need to be analyzed. For this, raw spectra were compared with the orbital satellite data, ensuring that reflectance peaks and absorptions were well correlated with the spectral range of multispectral sensors. The spectra chosen for the analyses were from amazonite, plagioclase, K-feldspar, and quartz. As granite and pegmatite are intrinsically correlated in Tysfjord,

a spectral analysis was also conducted to assess the potential to differentiate these two rock types using the selected satellite sensors.

A total of 48 laboratory reference reflectance spectra were analyzed for this research, but only the most representative were selected for illustration purposes. The most representative spectra were selected by observing the spectral behavior that is most frequent among the minerals studied. Out of the 48 samples, 5 were amazonite, 5 were biotite and granite, 3 were granite, 10 were K-feldspar, 9 were plagioclase, 8 were quartz, and 8 were pegmatite. Following the methodology applied by Santos et al. (2022b), an average of each mineral studied was taken to compare the laboratory spectra with the multispectral data. Using the information available in the "sample_description" field of the GREENPEG spectral library, it was possible to select the mineral spectra used in this study. The spectral analysis was then carried out on the basis of the data available in the spectral library, as well as a comparison with available literature.

3.5. Analysis of satellite image spectra

Pegmatite spectra were also collected directly from the WV-3 images using the Z profile (spectrum) tool in ENVI software. These spectra were chosen where the pegmatite is most exposed to ensure the maximum purity possible (Santos et al., 2022a). An average of 233 spectra was made for Jennyhaugen, and for *Håkonhals* 917 spectra were averaged. This analysis is the most direct way to identify the key bands to be used as input in the image processing methods. Santos et al. (2022b) conducted a robust spectral analysis for Sentinel-2 satellite data, so this research will present new information about the WV-3 satellite.

The spectra collected with the spectroradiometer were only used for analysis or comparison with the multispectral sensors and to identify the most important bands to be used in LGBM classification. To map the pegmatites in the study area using the MTMF, endmembers obtained through the Spectral Hourglass Wizard (SHW) workflow were utilized.

3.6. Processing spectral Hourglass Wizard (SHW) workflow

In this research, spectral unmixing was applied through the SHW workflow that directs the user through diverse methods designed to identify, select, and extract endmembers from an image (L3Harris n.d.; Santos et al 2022a; Santos et al., 2023b). These extracted endmembers are subsequently employed in the MTMF process for effective mapping of pegmatite abundance. Data processing consists of three steps (Fig. 2).

1st step: Data reduction – After applying the Minimum Noise Fraction (MNF), the SHW workflow allows us to either accept the spectral reduction performed by the MNF or carry out this reduction using the Data Dimensionality Panel tool. The Data Dimensionality Panel tool works by separating the noise from the data of interest through the adjustment of a threshold value, graphically represented as a red line (Santos et al. 2023; Wolfe and Black 2018). Everything above the red line is considered data of interest (>0.85 for WV-3 and >0.90 for Sentinel-2), i.e., the bands we want to use, while everything below the red line is noise (Annex 2). Seven MNF bands were extracted from the noise for WV-3 imagery (Annex 2a), while three MNF bands were selected for Sentinel-2 imagery (Annex 2b).

After that, the Pixel Purity Index (PPI) is applied to find the most spectrally pure pixels in the image and, consequently, perform the spatial reduction of the data. The PPI finds the most spectrally pure pixels through several interactions (Wolfe and Black, 2018). In this study, PPI was performed considering 10.000 iterations and a threshold of 5.0 for WV-3 imagery. For Sentinel-2 imagery, 5000 iterations and a threshold of 2.5 were applied. Different iterations and thresholds were chosen to account for the unique spectral and spatial characteristics of each satellite imagery type, WV-3 and Sentinel-2. The justification for using PPI in this work lies in its ability to identify the most spectrally pure pixels, which can aid in accurate end member selection even in multispectral satellite data like WV-03 (Ahmad, 2013; Hafid, 2004).

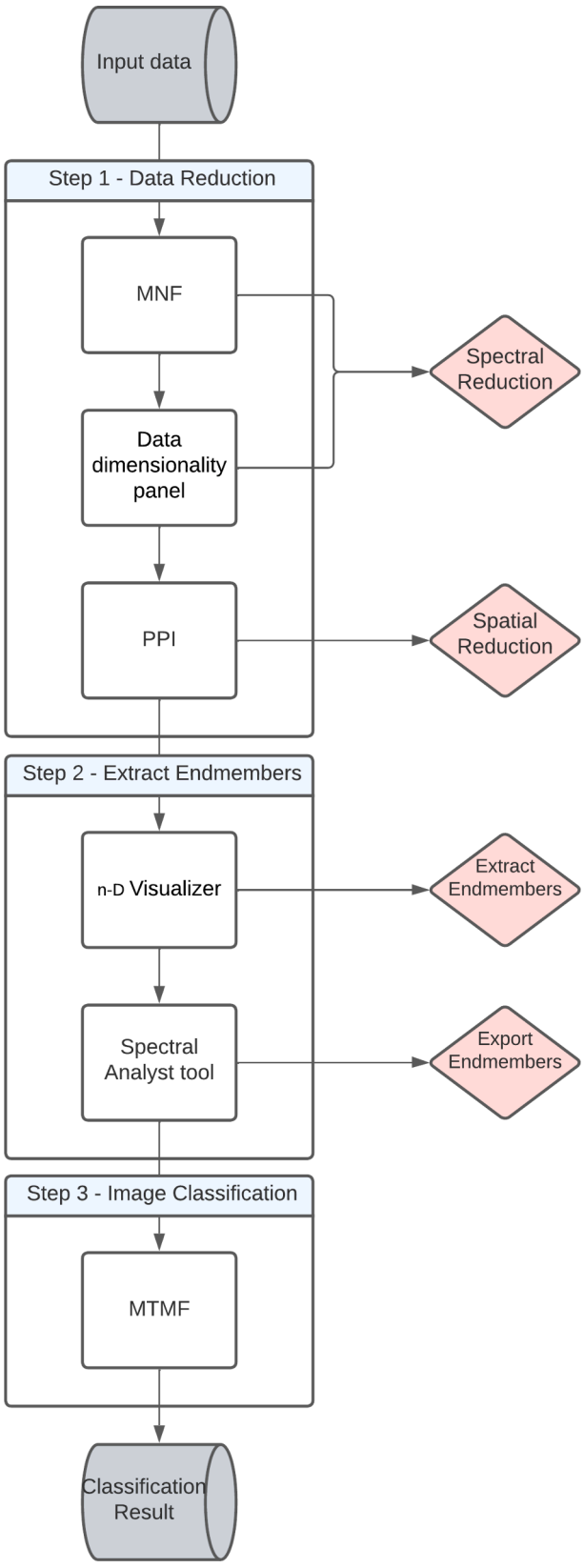

Fig. 2. Step by step of SHW workflow.

2nd step: Extract Endmembers – The PPI found in the last step is plotted in a scatter plot that can be visualized in the n-D Visualizer tool. So, it is possible to add more than two dimensions to the scatter plot and rotate it to find endmembers (Santos et al., 2023). The endmembers are

distributed at the corners of the scatter plot, so rotation is necessary until these corners are found and the endmembers are selected by clusters.

After selecting the endmembers, it's possible to extract, visualize, and export them using the spectral analyst tool. The spectral analyst tool also allows the user to compare the endmembers extracted with a known spectral library (Wolfe and Black, 2018). In previous research, this method has been ineffective (Santos et al., 2023b; Santos et al. 2022a). Therefore, in this study, an alternative approach was employed where the endmembers were used to proceed with the classification and to verify if the pegmatite information was correctly extracted.

The pegmatite endmember for Sentinel-2 imagery was designated as endmember 3, and for WV-3 imagery, it was endmember 2 (Fig. 3). Finding pegmatite endmembers in Tysfjord was not a significant challenge, thanks to the contributions from the Håkonhals and Jennyhaugen mines. These two large open-pit mines may have a substantial impact on the scatter plot, making it easier to locate the corner where the pegmatite endmembers are.

3rd step: Classification – In the third step, the extracted endmembers were used as input in the MTMF for image classification. The MTMF is a powerful classification method that yields excellent results in subpixel occurrence situations (Adiri et al. 2020; Santos et al., 2023). It's possible to observe the step by step of the SHW processes in Fig. 2.

3.7. Processing Mixture Tuned Matched Filtering (MTMF)

After extracting the pegmatite endmembers, they were used as a reference spectrum for the MTMF to classify the full images. The MTMF method is a pixel-based partial spectral unmixing method that can be employed when not all the endmembers of a scene are identified (Adiri et al., 2020; Asadzadeh and de Souza Filho, 2016; Badr, 2021; Banerjee et al., 2019). The method provides a Matched Filtering (MF) score and an infeasibility value for each endmember, allowing for the reduction of false positives from the MF processing step (Ellis and Scott, 2004; Nv5geospatial, 2024a). The SHW workflow allows the user to interpret the MTMF results through a histogram. The data of interest (pegmatite endmember) are located in the upper tail of the histogram (Nv5geospatial, 2024b). By adjusting the histogram, it is possible to decrease the number of false positives and highlight only the data of interest. After ensuring that the information of interest was selected by the histogram, we can analyze the MTMF results using the spectral behavior of Jennyhaugen and Håkonhals as ground truth. For this analysis, color slices were applied to the MTMF results. In cold colors are the pixels with a lower abundance of pegmatite, while in hot colors are the pixels with a higher abundance of pegmatite. Color slices allow for analyzing the patterns and spatial distribution of pegmatite occurrences in Tysfjord.

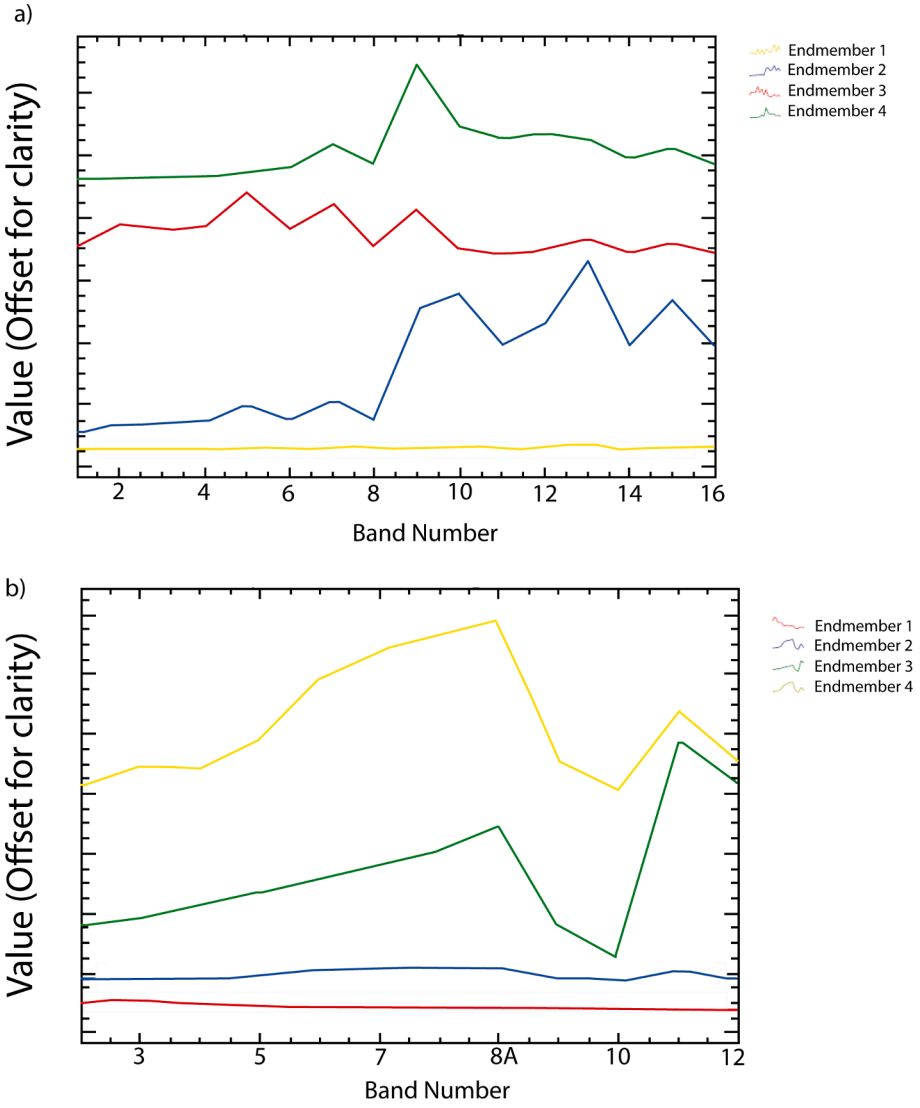

Fig. 3. Endmembers extracted from Sentinel-2 (a) and WV-3 (b).

3.8. Processing light Gradient boosting Machine (LGBM)

The LGBM algorithm is an ensemble tree-type classifier created by Microsoft in 2017 (Hajihosseinlou et al., 2023), based on the Gradient Boosting Decision Tree (BGDT) algorithm. To overcome the BGDT problem with the limit of training data size (Hajihosseinlou et al., 2023), LGBM implemented the Gradient Based One-Side Sampling (GOSS) and the Exclusive Feature Bundling (EFB) methods (Hajihosseinlou et al., 2023). In the GOSS approach, only a fraction of samples with small gradients are randomly chosen and then treated to compensate for data distribution. The EFB applies an algorithm to detect similar features on the dataset, and then bundle them so the number of features and bundled features are lower than the original features (Hajihosseinlou et al., 2023). Another difference that boosts the performance of the LGBM algorithm is that it uses a leaf-wise growth strategy instead of a level-wise technique (Hajihosseinlou et al., 2023), which can overfit the model by expanding decision trees too deep. Because of that, parameters need to be chosen carefully. The max_depth parameter prevents trees from growing too deep, as well as the num_leaves parameter, which controls the complexity of the tree model. The algorithm was implemented in Python programming language using the LGBM library and was applied for the Sentinel-2 and Worldview-3 imagery, through the following steps.

1st Step: Reconnaissance of the area – The first step of classification consisted of applying the vegetation, snow, water, and built-up indices (Normalized Difference Vegetation Index (NDVI), Normalized Difference Snow Index (NDSI), Normalized Difference Water Index (NDWI) and Normalized difference Build Index (NDBI)) for an initial reconnaissance of the area. In terms of lithology, the two predominant types (Müller et al., 2022) – AMCG (anorthositic, magnetitic, charnockitic, and granitic) rocks and the Tysfjord granite gneiss – were grouped in one Granite class. After the first analysis, four classes were defined: 1. Pegmatites, 2. Granite, 3. Water, 4. Vegetation. False positives with granite and snow were a problem in past works (Santos et al., 2023). Therefore, in this study area snow was not classified to analyze the method's effectiveness in eliminating the most challenging false positives. After the reconnaissance and definition of the classes, the training sites were chosen. The information regarding the number of pixels per class can be found in Annex 3.

2nd Step: Class separability – The second step consists of guaranteeing that the classes are well separated. For this, a class separability tool available on PCI Geomatica software was used. This tool uses the Bhattacharyya Distance method and results range between 0 and 2, being interpreted as: (i) $0 < x < 1$ – very poor separability; (ii) $1 <$ to <1.9 – poor separability; (iii) $1.9 <$ to < 2 – good separability.

The signature separability can be observed in Table 1 (WV-3) and Table 2 (Sentinel-2).

The unexpected observation of granite and pegmatite demonstrating better separability for Sentinel-2 than for WV-3 prompts a closer examination of our sampling strategy. Notably, a larger quantity of pixels were utilized for sampling in WV3. While the heightened spatial resolution of WV-3 allows for increased pixel sampling, it also introduces the potential for more confusion between pegmatite and granite, leading to a higher number of misclassified pixels. This, consequently, impacts the observed separability of classes.

3rd Step: Data pre-processing – After ensuring the classes are well separated, the third step consisted of preparing the data to be used in the Python script. ArcGIS Pro software was used to extract pixel information from the training areas of the different classes. According to Santos et al. (2022b), the bands in the reflectance peaks of the target minerals are the most suitable inputs for the classification. So, with support of the spectral analyses, and following the methodology proposed by Santos et al. (2022) the Sentinel-2 bands 4, 6, 7, and 8A were chosen as inputs while for WV-3, bands 5, 7, 9, 10, 13 and 15 were selected (see section 4.1).

Table 1
Signature separability for WV-3 imagery.

	Pegmatite	Granite	Vegetation	Water
Pegmatite				
Granite	1.89			
Vegetation	1.99	1.91		
Water	1.99	1.99	1.99	

Table 2
Signature separability for Sentinel-2 imagery.

	Pegmatite	Granite	Water	Vegetation
Pegmatite				
Granite	1.95			
Water	2.00	1.99		
Vegetation	1.99	1.96	1.99	

4th Step: Model creation – In the model creation step, the data set was split into 25% of pixels for testing and 75% of pixels for training. This method is applied to ensure the maximum independence between training and test subsets (Cardoso-Fernandes et al., 2020). The parameters to be optimized using a grid search procedure with cross-validation were also defined in this step: "max_depth", "n_estimators", and "num_leaves".

5th Step: Model evaluation – This step involves applying different metrics to evaluate the best model identified in the previous grid search. The metrics used include precision, recall, F1 scores, and Kappa statistics. Additionally, a confusion matrix is generated based on the test subset. These evaluations are crucial for assessing the trained model's effetiveness and its ability to generalize to unseen data.

3.9. Processing Automated lineament extraction

Mineral deposits are often associated with structural controls – e.g. faults, fractures, and shear zones – which can be identified through the respective morphological lineaments that can be extracted from a variety of optical, radar, and LiDAR data (El-Desoky et al., 2022; Tripp and Vearncombe, 2004). Geological faults can be indicative of the presence of mineralized pegmatites and are, therefore, useful for identifying them. However, it is possible to hypothesize that the significant degree of metamorphism in the region may have obliterated numerous geological structures associated with pegmatites. This hypothesis would make the lineaments ineffective for use as input in this study. Nonetheless, by extracting the lineaments and creating a lineament density map (LDM) using a high-resolution hillshade (refer to Annex 4),), we were able to observe a correlation between the LDM and the locations of pegmatite mines (Fig. 4). Consequently, we opted to utilize the LDM as an input, as there is a potential connection between these lineaments and the pegmatites (Fig. 4 b). The process of extracting the lineaments can be viewed in Annex 4.

3.10. Standardized criteria for fuzzy logic

The processing method results need to be standardized before being used as inputs in the fuzzification phase. Standardization is essential to ensure that the data have a consistent and comparable scale, which helps prevent distortions in the final model outcomes (Abdelkareem and Al-Arifi, 2021; Mamouch et al., 2022; Zadeh, 1965). This process is typically performed to ensure that different variables have the same scale and variance, facilitating data comparison and processing.

To achieve this standardization, we utilized the Fuzzy Membership tool available in ArcGIS Pro software. This tool is employed to transform precise values into imprecise terms, often referred to as fuzzy values, on a scale from 0 to 1.

This process accurately reflects the relative suitability of the specified criteria value (Abdelkareem and Al-Arifi, 2021; Mamouch et al.,

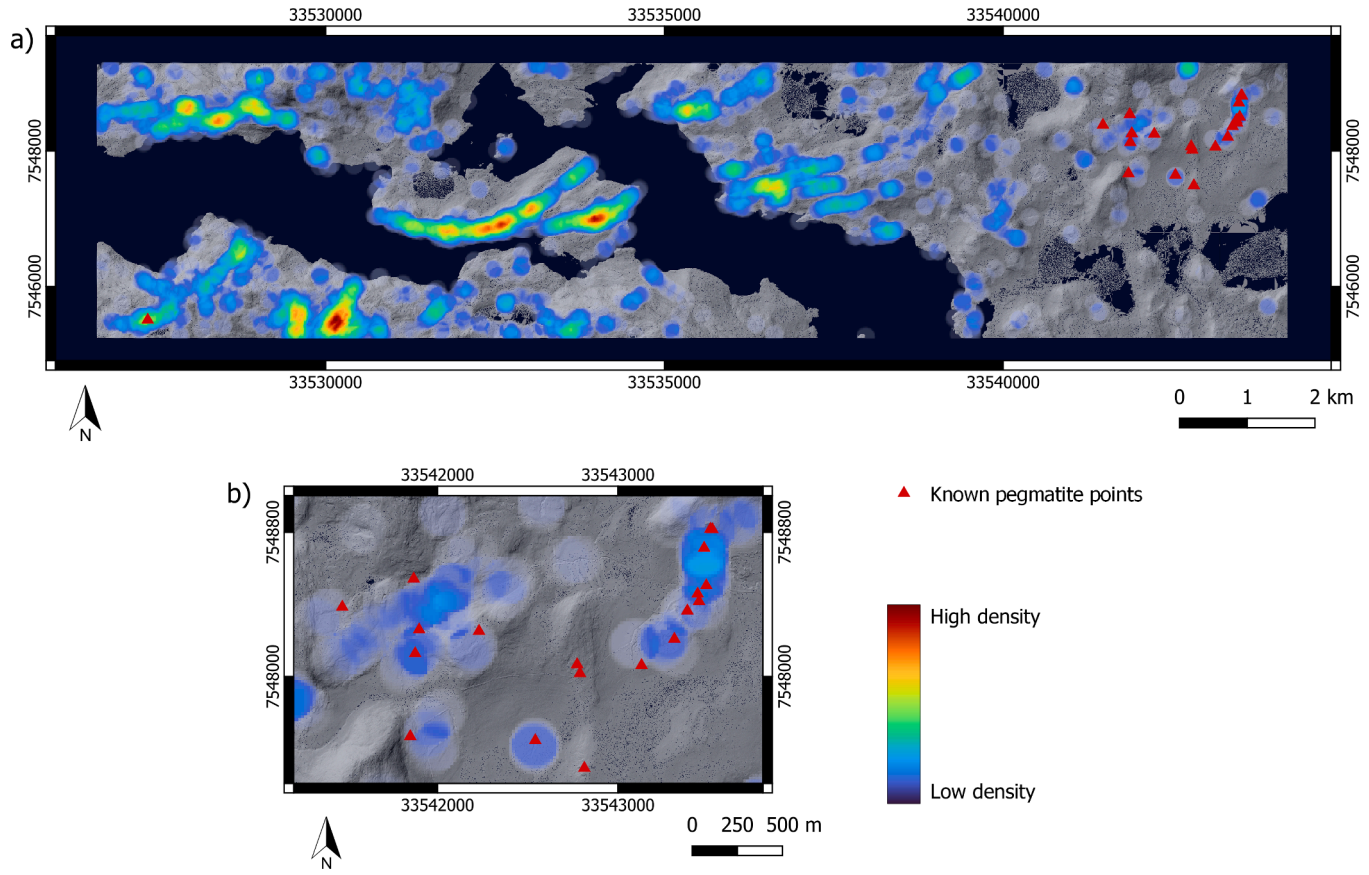

Fig. 4. Lineament Density Map – LDM (a) The LDM for Tysfjord displays the most concentrated lineaments in hot colors. The known pegmatite points are predominantly clustered in the eastern zone of the area. (b) Known points are highlighted in relation to the LDM.

2022; Zadeh, 1965). Within the fuzzy function, a range of parameters are available, including Gaussian, Small, Large, Near, MSLarge, MSSmall, and Linear. Following comprehensive testing, the Linear function emerged as the most suitable. The linear function is referred to in many works as the best option for fuzzification (Porwal et al., 2003; Yousefi and Carranza, 2015). Nevertheless, we tested all the functions to make sure we chose the most suitable one for our set of data. The Linear function computes membership through a linear transformation of the input raster, assigning a membership value of 0 at the minimum and 1 at the maximum. These parameters intricately govern the amplitude of the linear transformation applied to the input data.

3.11. *Mineral prospectivity mapping through fuzzy overlay*

After ensuring that all the input data were standardized, on a scale from 0 to 1 using the Fuzzy Membership tool, the fuzzy overlay modeling was tested in a GIS environment (ArcGIS Pro). The fuzzy logic modeling is based on the fuzzy set theory proposed by Zadeh (1965) and has been applied successfully to generate mineral prospectivity maps (Abdelkareem and Al-Arifi 2021; Mamouch et al. 2022; Sekandari et al. 2020). The real values of the inputs are included in the interval [0–1], where 0 represents pixels with no suitable location and 1 represents pixels with suitable locations (Mamouch et al., 2022). Equation (1) represents the fuzzy set theory.

$$A_{ij} = \{(X_{ij,}\mu_A)/_{ij} \epsilon X_i\}, (0 \leqslant \mu_A \leqslant 1) \quad (1)$$

where μA is called the degree of membership function of x in A and X corresponds to a set of layers Xi (i = 1, 2, 3, …, n), and each layer to r classes define.

The fuzzy overlay takes into consideration whether a particular location belongs to one or a combination of various sets (Zheng et al., 2023). Therefore, weighting the criteria is not necessary for a fuzzy overlay, as the primary focus is on the location of memberships rather than their weights (Zheng et al., 2023). Fuzzy overlay employs specific techniques to explore relative relationships and quantify interactions. The functions that can be employed are fuzzy AND, fuzzy OR, fuzzy PRODUCT, fuzzy SUM, and fuzzy GAMMA (Bonham-Carter, 1994; Lewis et al., 2014).

This study tests three fuzzification models (Fig. 5), each representing the combination of different data to generate prospecting maps for NYF pegmatites. Each model comprises at least one set, with each set containing distinct input layers used in the fuzzification process. Each set undergoes testing with various overlay types (SUM, OR, AND, PRODUCT (PROD), and GAMMA). According to (Bonham-Carter, 1994), a single map can have more than one fuzzy function, which is why the authors believe that testing all the functions is important so the reader can understand in detail the response of each function tested for our dataset.

In Model 1 (Fig. 5a), the MTMF and LGBM classification results are placed in separate sets and subjected to fuzzification combined with the LDM. Subsequently, an initial analysis of overlay-type results is conducted, with the best outcome from each test selected. The selected outcomes of each set are fuzzified again to produce the final output. Model 2 (Fig. 5b) adopts a combined approach where all three inputs are fuzzified, and each overlay type is individually tested. The optimal output is chosen as the prospectivity map. Lastly, Model 3 (Fig. 5c) follows a similar methodology to Model 2 but omits the use of LDM as input. Implementing three models allows for a comprehensive examination of approaches tailored to the specialized demands of mapping NYF pegmatites. This approach allows for evaluating effectiveness, considering several variables, reducing bias, and ensuring robust and

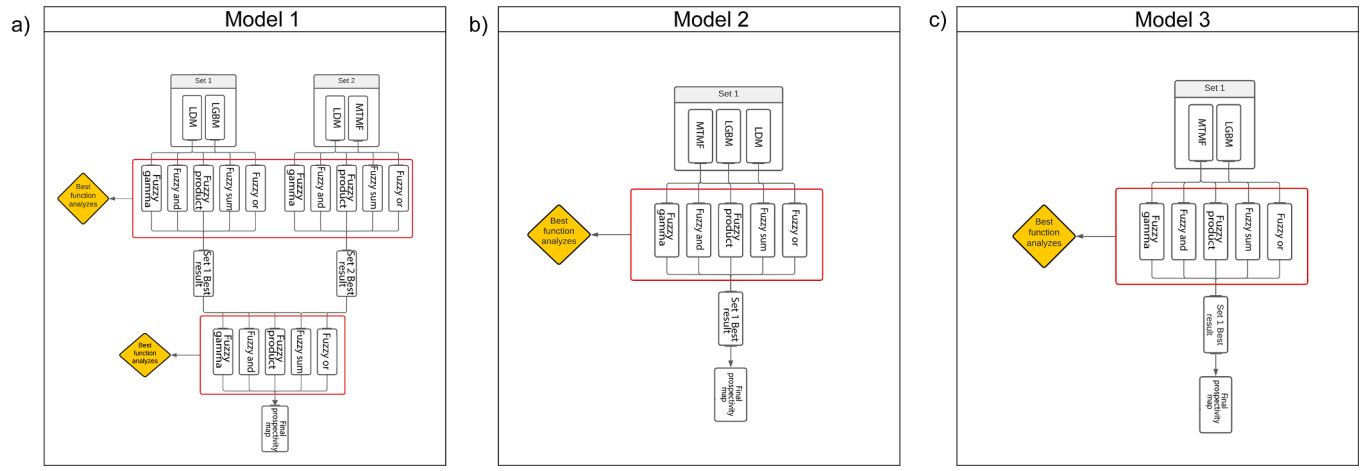

Fig. 5. Step by step of Fuzzy logic approach applied. a) Approach applied for model 1. b) Approach applied for model 2. c) Approach applied for model 3.

reliable research results. Ultimately, the results of the three models were compared to determine the most suitable approach.

To identify potential points of interest for pegmatite exploration, the final output from both satellites (WV-3 and Sentinel 2) was subjected to fuzzification. The definition of each function can be observed in Annex 5.

The authors would like to emphasize that the fuzzy overlay does not necessarily aim to obtain more accurate results than the input data, but rather to integrate and weight different pieces of information in a more flexible and holistic way (Sabbaghi and Tabatabaei, 2023). Fuzzy operators make it possible to assess the uncertainty and ambiguity associated with data, taking into account a range of variables and criteria to improve decision-making (Makropoulos et al., 2008; Popchev Ivan, 1999).

4. Results and discussion

4.1. Reflectance spectroscopy analysis

In this section, we analyzed not only the spectral behavior of the principal minerals that can be found in NYF pegmatites but also compared the spectral behavior between pegmatites and host granites in Tysfjord. Considering that false positives with granite are a great challenge in pegmatite studies, this analysis aims to understand not only the spectral behavior of pegmatites but also to understand how the classification methods can differentiate pegmatites and granites. The study of the difference in spectral behavior will help to mitigate false positives between these two rocks with similar chemical compositions.

First, an analysis of the spectral behavior of the principal minerals found in Tysfjord NYF pegmatites was done. These analyses are important to select the most important bands to compose the remote sensing processing methods. The bands can be used as absorption or reflectance inputs according to their spectral behavior. These methods

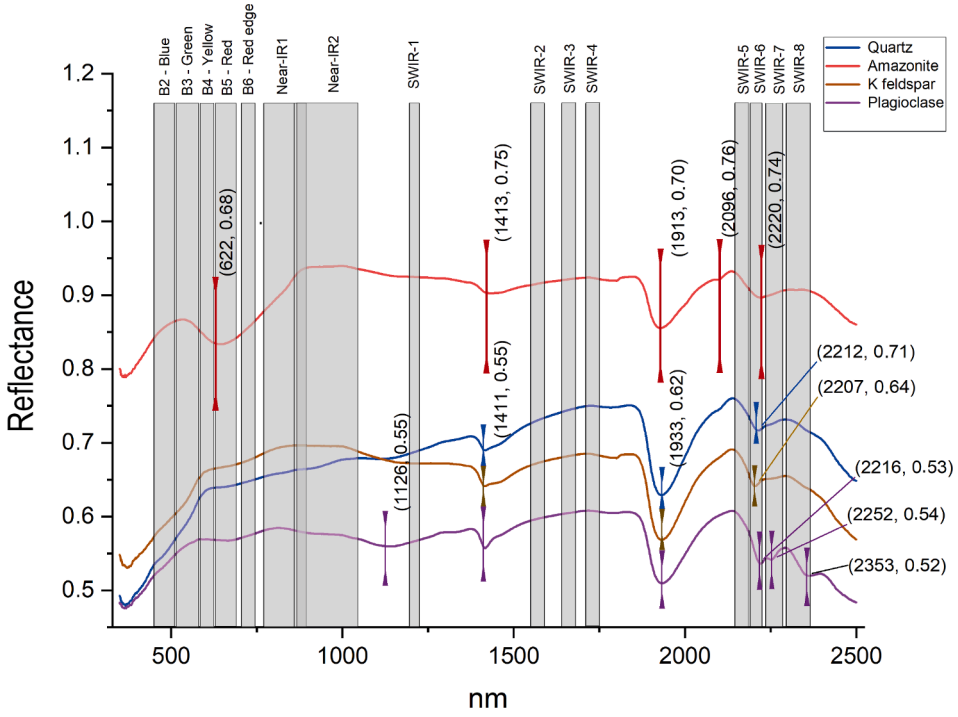

Fig. 6. Spectra from quartz, Amazonite, k-feldspar, and plagioclases from Tysfjord.

can be the most commonly used such as RGB combination, BR, and Principal Components Analyses (PCA) or even classification algorithms such as LGBM.

Fig. 6 shows the spectra of quartz, K-feldspar (including the green K-feldspar variety amazonite), and plagioclase about the WV-3 satellite spectral range. Considering that pegmatites typically contain feldspar of ceramic grade, quartz, and white mica, these three minerals are most susceptible to being picked up by the satellite sensors. These analyses provide information that allows us to apply image-processing methods focused on these specific minerals. This method can be adapted to other minerals. Concerning the amazonite spectra, it is possible to observe a iron absorption feature at (~622 nm) (Cardoso-Fernandes et al. 2023a) that is encompassed by bands 4 – yellow (B4), and B5 (red). While B6 (red edge) also includes this absorption, it is noticeable that it is not located at the center of the absorption feature. Thus, although B6 can also be used for image processing, preference should be given to the bands located closer to the center of the absorptions. An AlOH absorption feature at ~2220 nm (J Cardoso-Fernandes et al. 2023a) is encompassed by B14 (SWIR 6). We observed that B2 and B3 encompass the reflectance peak in the visible region. Both Near 1, and Near 2 also can be used as reflectance bands for amazonite, and SWIR 5 can be used as a reflectance band in the SWIR region. The spectra of K-feldspar, plagioclase, and quartz had similar absorption features and reflectance peaks with importance for WV-3 satellite bands. The plagioclase sample is the only one with an absorption (~660 nm) in the visible spectrum besides amazonite and is encompassed by B4 (red). In the SWIR region, the AlOH absorption (~2200 nm) present in the three samples (K-feldspar, quartz, and plagioclase) encompassed by SWIR 6. A FeOH absorption (~2250 nm) and a MgOH absorption (~ 2350 nm) in the plagioclase sample are covered by SWIR 7 and 8, respectively. These absorptions may indicate the presence of biotite in the sample. Quartz, and K-feldspar have similar spectral behavior for absorption features, so the bands that can be used as input are practically the same. It is essential to emphasize that none of these characteristics are exclusive to these minerals, indicating the presence of spectral mixtures even when observed on a mineral scale.

Except B4 for plagioclase and amazonite, all other bands within the visible range can be used as reflectance peaks. As same as for NIR 1 and NIR 2 the SWIR – 4, 5, 7, and 8 are the best options to be used as reflectance inputs for quartz, K-feldspar, and plagioclase in image processing methods. The only one that encompasses an absorption feature for quartz, plagioclase, and K-feldspar is the SWIR-6 band. For amazonite, the best bands to be used as reflectance inputs are NIR 1 and 2, and SWIR-5 and 8. For absorption features, we can select band 5, and SWIR-6 as the best option. It is important to emphasize the importance of carefully selecting multispectral data bands, as the lack of diagnostic features in certain WV-3 bands renders them predominantly reflectance bands, thereby diminishing their utility for target mapping.

To gain a better understanding of the spectral mixtures present in samples of quartz, K-feldspar, and plagioclase, a more thorough analysis was conducted.

In the processed quartz spectrum (Fig. 7 a), we can observe two water absorptions (~1435 nm and ~1917 nm) and one single AlOH absorption (~2207 nm). Iron Features (~486 nm and ~664 nm) were also identified. The quartz spectra analyzed from the GREENPEG spectral library had similar spectral behavior. Most of the spectra had the presence of iron features (Fe^{2+} and Fe^{3+}) and absorptions that indicate montmorillonite presence. Although, the quartz spectra of samples from Kråkmo present a deep water absorption (~1900 nm), a single sharp AlOH absorption (~2220 nm), and two diagnostic absorptions (~2300 nm and 2400 nm respectively) that can indicate Illite presence. The k-feldspar spectra also present a similar behavior. In Fig. 7 b, it's possible to observe a spectrum with a single OH absorption (~1419 nm), an AlOH absorption (~2213 nm), and two absorptions that are diagnostic of Illite (~2362nm and ~2454 nm). This set of absorption features forms a strong indicative of illite presence. In Fig. 7 c, we can observe two water absorption features (~1414 nm and ~1931 nm), one single AlOH absorption (~2204 nm), and a bending (highlighted by the red rectangle) that may indicate the presence of montmorillonite. The k-feldspar samples also present mixed spectra between montmorillonite and illite. Iron features were also detected (VNIR region) in all K-feldspar samples. The two most representative spectra were chosen to represent plagioclase samples. In Fig. 7 d, it is possible to observe a water absorption (~1418 nm), a deep water absorption (~1933 nm), and an absorption around ~2220 nm. These absorptions can indicate the presence of Illite in the sample. The absorption feature around ~2352 nm, can may indicate that this spectrum is mixed with illite. In Fig. 7 e, we can observe two water absorptions (~1414 nm and ~1933 nm) and one single AlOH absorption (~2198 nm). These absorptions, together with the bending anomaly (highlighted by the red rectangle), may indicate the presence of montmorillonite. Also identified iron features in the VNIR region (~670 nm).

4.2. Contact granite and NYF pegmatite spectral behavior

To comprehend the distinctions in spectral behavior between NYF pegmatite and granite, a spectral analysis was conducted on these rock samples. We opted for raw spectra for analysis in this section since it is the raw spectra's behavior that is sensed by orbital sensors. Thus, analyzing raw spectra is the optimal way to compare laboratory spectral measurements with orbital sensor data.

The four most representative spectra of NYF pegmatite from Tysfjord were selected for a more detailed analysis. The analysis of spectral spectral data from rock samples, as depicted in Annex 6, is pivotal for understanding the mineral composition and characteristics. Through detailed examination, characteristic absorptions were identified, providing insights into the presence of various minerals within the samples. One notable feature observed is the ramp-like Fe^{2+} absorption, which is particularly prominent in samples containing biotite (Annex 6a, Annex 6b, Annex 6c, Annex 6d). This absorption, along with others such as FeOH and MgOH, indicates the potential occurrence of biotite within the samples. Additionally, other diagnostic absorptions suggest the presence of other minerals like muscovite, chlorite, and members of the montmorillonite and white mica groups (Annex 6e, Annex 6f, Annex 6g, Annex 6h). Notably, distinct features, such as an absorption at ~2088 nm, may indicate the presence of topaz, which adds to the spectrum of minerals identified (Annex 6h).

In summary, we observed that despite biotite being present in both granite and pegmatite samples, the biotite is more spectrally influent in the granite spectral behavior which is in line with previous studies (Cardoso-Fernandes et al. 2023a). This is because biotite is evenly distributed and much more abundant in the granite gneiss than in the pegmatites. However, the optical sensors of the multispectral satellites, such as WV-3, do not have enough spectral resolution to detect these absorption features in detail. For this, a more direct comparison between the spectral behavior and the spectral range of the WV-3 satellite bands was done.

When we compare an average of a spectrum of granite from a contact zone with the range of the WV-3 satellite bands (Fig. 8a), it is noticeable that the ramp-like Fe^{2+} biotite absorption is encompassed by many bands on the VNIR region (B4, B5, B6, NIR 1 and NIR 2). The same as the VNIR region, many bands on the SWIR region (SWIR-1, SWIR-2, SWIR-3, and SWIR-4) also encompass the ramp-like Fe^{2+}. The absorption at ~2202 nm, corresponding to the AlOH feature, is encompassed by the SWIR-6 band, the FeOH absorption feature at ~2258 nm is encompassed by the SWIR-7 band, and the MgOH absorption feature at ~2357 nm is encompassed by SWIR-8.

A comparison between the average pegmatite spectra and the spectral range of the WV-3 bands was made (Fig. 8b). It is possible to observe that all the bands in the VNIR are located where the pegmatite reflects electromagnetic energy. Regarding the SWIR region, although the SWIR-1 band is not located at the center of the absorption at ~1145 nm, it can

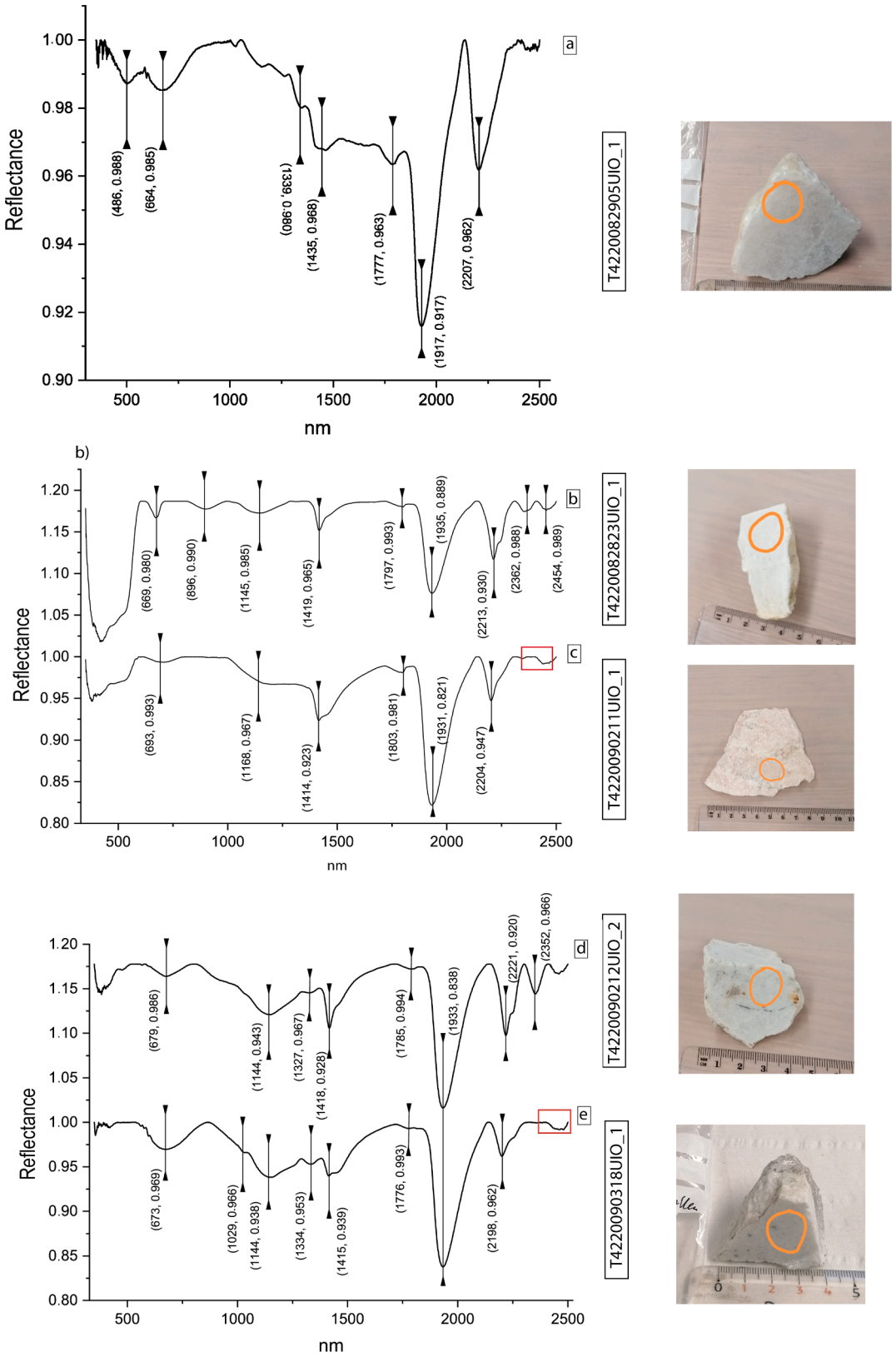

Fig. 7. Processed spectra from Quartz, k-feldspar, and Plagioclase and samples used to collect the respective spectra. The orange circles mark the location where the spectrum was collected. a) Processed spectrum of quartz (sample T4220082905UIO_1); b) Presents a possible K-feldspar spectrum mixed with illite (sample T4220082823UIO_1); c) Presents a possible K-feldspar spectra mixed with montmorillonite (sample T4220090211UIO_1). The bending highlighted by the red rectangle indicates the presence of montmorillonite. d) Presents a plagioclase spectrum with illite (sample T4220090212UIO_2). e) Presents a possible plagioclase spectrum of montmorillonite (sample T4220090318UIO_1).

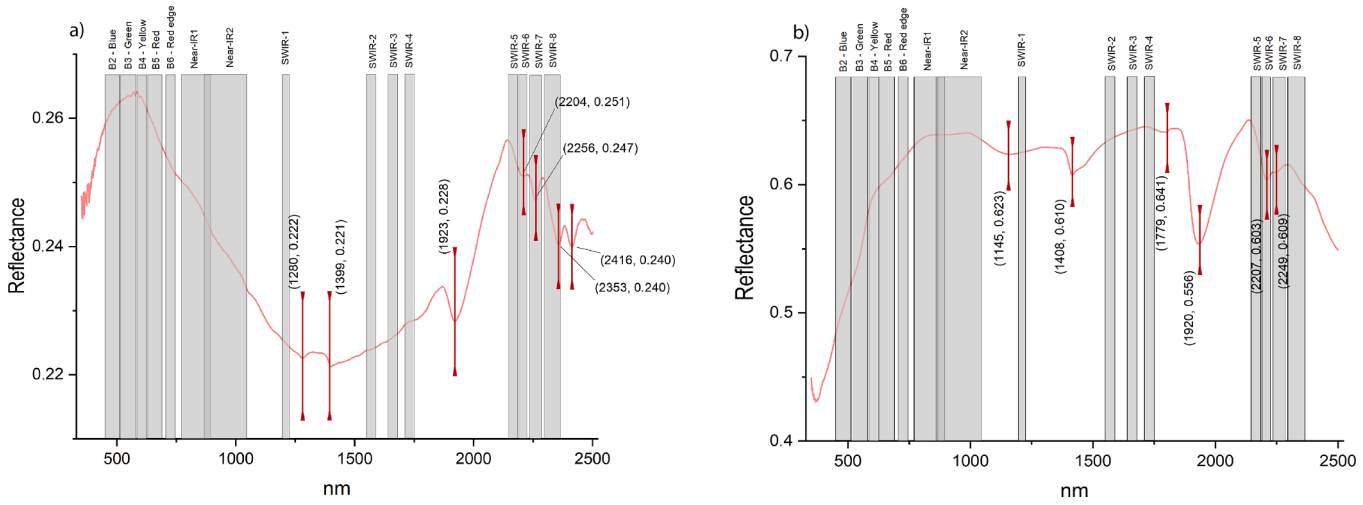

Fig. 8. A) average spectrum from granite samples in contact zone with pegmatite from tysfjord. b) average spectrum from pegmatite sample from tysfjord.

be useful for image processing. Moreover, SWIR-6 encompassed the AlOH absorption feature at ~2207 nm while SWIR-7 encompassed the FeOH absorption feature at ~2249 nm. The SWIR-2, SWIR-3, SWIR-4, and SWIR-5 are on the reflectance behavior of the pegmatite.

After analyzing and comparing the spectral behavior of granite and pegmatites with the WV-3 bands, it is noticeable that the absorption features in the SWIR region are not enough to differentiate pegmatites and granites. Although chlorite, muscovite, and biotite have important absorption features in the last bands of the SWIR region, these absorption features will be detected in both granite and pegmatite with poor detail to allow the multispectral satellite sensors to differentiate them. On the other hand, the ramp-like Fe^{2+} present in biotite can be detected by many bands on VNIR and the first bands of the SWIR region and can be the principal factor in discriminating granites and pegmatites in Tysfjord. In contrast to the diagnostic absorptions located at the end of the SWIR regions, the ramp-like Fe^{2+} is detected by sufficient bands permitting the ramp-like Fe^{2+} behavior to be captured even on multispectral satellites.

To ensure that the ramp-like Fe^{2+} absorption feature is picked up by the sensors, the average granite and pegmatite spectra were resampled to WV3 and Sentinel-2 resolution. In Fig. 9, the ramp-like Fe^{2+} absorption feature is highlighted by the red rectangle, confirming that the ramp-like Fe^{2+} absorption can be identified by multispectral satellites, even by sensors with lower spectral resolution such as Sentinel-2 (Fig. 9 B). As expected, the ramp-like Fe^{2+} absorption is more detailed in the WV3 resampled spectrum (Fig. 9 A). However, the average pegmatite spectral behavior doesn't show the ramp-like Fe^{2+} absorption and, for both sensors' spectral resolution, there is a crescent reflective behavior that starts ~1400 nm. Thus, the granite's ramp-like Fe^{2+} absorption can be used to differentiate pegmatite from granite in contact zones reducing false positives in classification results. As predicted, in Fig. 9, the SWIR 6 and SWIR 7 bands of WV3 detected the AlOH absorption, while the

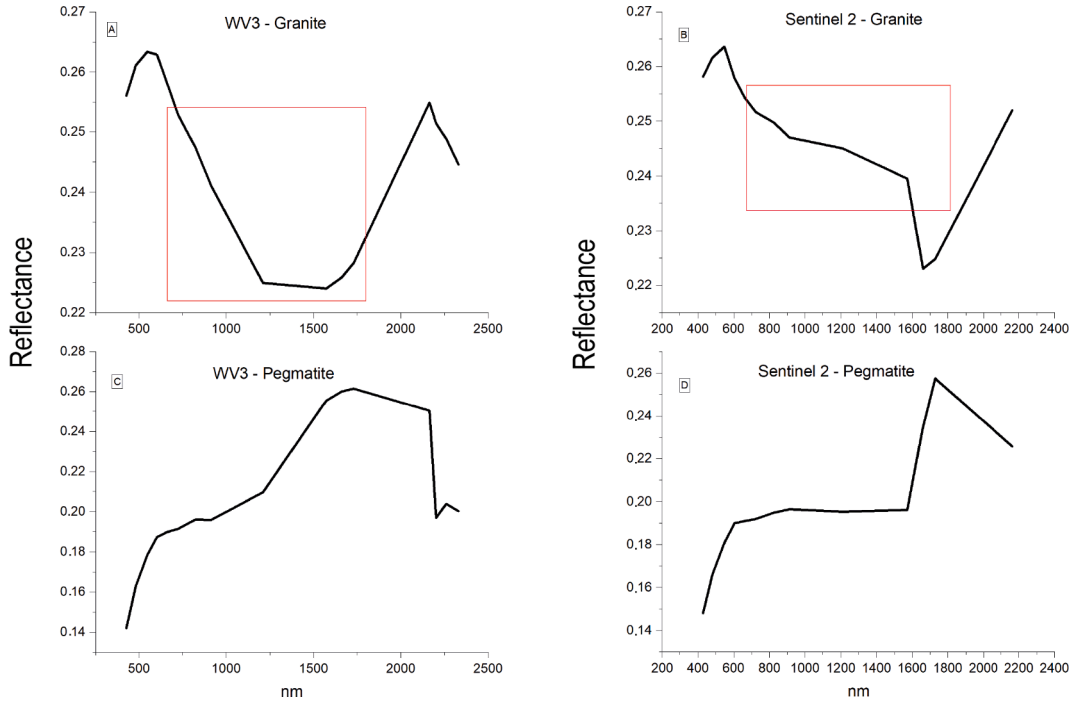

Fig. 9. Comparison of pegmatite and granite spectral behavior. Average spectral behavior of granite resampled for WV-3 (A) and Sentinel-2 (B) resolutions, and of pegmatite resampled for WV-3 (C) and Sentinel 2 (D) resolutions. The ramp-shaped Fe^{2+} is indicated by the red rectangle.

MgOH absorption is reflected in bands SWIR 7 and SWIR 8 (Fig. 9 D). This indicates that the WV3 satellite can detect mineral compositions with its spectral resolution. The same cannot be stated for Sentinel-2: while the biotite absorption was detected (Fig. 9 B), band 12 (Fig. 9 D) is not detailed enough to distinguish AlOH, FeOH, and MgOH features with the Sentinel-2 spectra, as shown in previous studies (van der Meer et al., 2014). Overall, these results show that the WV3 data is a powerful tool for mineral spectral analyses in comparison with the open-source multispectral satellites.

The spectral analyses have been important to understand the minerals influencing the signal that is captured by the satellite sensors. Although, according to Santos et al. (2022) the analyses of spectra collected directly from the satellite images are the best way to select the bands to be used as input in the image processing. This comparison can help select key bands for multispectral data processing and can be applied in greenfield areas or brownfields when there are no known pegmatite outcrops to be used as ground truth.

As expected for minerals, it is possible to observe (Annex 7) a stronger spectral response in the SWIR region. On the SWIR region, the reflectance peaks are almost the same for both mines (SWIR 1, SWIR 2, SWIR 5, and SWIR 7). Although in Jennyhaugen mine, SWIR 1 reflects more than SWIR 2, while in Håkonhals mine we observed the opposite. The absorption bands on the SWIR region are the same for the two mines (SWIR 3, SWIR 6, and SWIR 8). Despite the spectral response on the VNIR region being especially weak in Jennyhaugen, the reflectance and absorption bands are the same between the two mines.

A comparison was also made between the pegmatite spectrum (Annex 8 a) and a possible granite spectrum (Annex 8c) directly on the WV3 image. It is observed that the ramp-like Fe^{2+} absorption is not evident in the spectrum analysis, but the pegmatite exhibits more pronounced reflection peaks compared to the granite. This difference could be attributed to the presence of biotite in the granite, which may have influenced its less reflective behavior in the spectrum.

4.3. LGBM classification

The inputs used in LGBM for the Sentinel-2 data were the bands 4, 6, 7, and 8A. The classification was performed using all inputs previously described. For WV-3, bands 5, 7, 9, 10, 13, and 15 were used as inputs. The best parameters, after cross-validation, for the Sentinel-2 data were max_depth = 10, n_estimators = 45, and num_leaves = 15. The kappa statistic was 0.97. The best parameters for the WV3 data were max_depth = 10, n_estimators = 500, num_leaves = 15. The kappa statistic was 0.99. Other metrics can be observed in Table 3. As it is possible to observe in Table 4, WV3 has better scores than Sentinel-2. This phenomenon may be attributed to the greater quantity of sampled pixels acquired from WV3, a possibility facilitated by its superior spatial resolution (Alwosheel et al., 2018; Krizhevsky et al., 2017).

The most important input features weighing on the final prediction were also analyzed (Table 4). Regarding Sentinel-2, band 4 had the highest contribution, while bands 7 and 8A had the lowest. When compared with previous results obtained by Santos et al. (2022a), band 4 shows the highest reflectance among the visible bands in NYF pegmatites. Despite pegmatites reflecting more electromagnetic energy in bands 6 and 7 than in band 4, in that region, there is no clear reflectance peak like in band 4. This can be indicative that the bands that are at the top of reflectance peaks had more influence in the classification than others.

Table 3
Score report for LightGBM algorithm.

	Precision		Recall		F1 score	
	Sentinel-2	WV-3	Sentinel-2	WV-3	Sentinel-2	WV-3
Pegmatites	0.97	1.00	0.97	0.98	0.97	0.99
Granite	0.95	0.98	0.98	1.00	0.96	0.99
Vegetation	1.00	1.00	0.99	1.00	0.99	1.00
Water	1.00	1.00	1.00	1.00	1.00	1.00

Table 4
Feature importance.

Sentinel-2 Inputs	Importance %	WV3 Inputs	Importance %
Band 4	40	Band 5	7
Band 6	26	Band 7	7
Band 7	17	Band 9	23
Band 8A	17	Band 10	18
		Band 13	23
		Band 15	22

The most important inputs for WV3 classification were band 9 (SWIR 1) and band 13 (SWIR 5). Both SWIR 1 and SWIR 5 are the bands with higher reflectance peaks among all the bands. This indicates that choosing bands where NYF pegmatites have the highest reflectance can be an assertive method. Band 15 (SWIR 7), and band 10 (SWIR 2) also encompass reflectance peaks for pegmatite. The lowest contribution is from the VNIR bands, contrasting the highest contribution of the SWIR bands, showing the importance of the SWIR region for mineral studies. Using Sentinel-2 as an example, in the absence of more bands in the SWIR region, bands in the VNIR region can be used as an alternative. This becomes evident when observing the number of bands that can sense the Fe ramp-like absorption characteristic of biotite in the VNIR bands, as was noted in the previous section.

In Fig. 10, is possible to observe the classification result for both satellites. Fig. 10 a) shows us the WV3 classification result for Håkonhals, where all the mine area was classified as pegmatite, including the mine paths that may contain pegmatite traces. Fig. 10 b) shows the WV3 classification result for Jennyhaugen. In this case, the algorithm didn't identify all the mine areas as pegmatite. This result is more accurate with the location where the pegmatite is more exposed in the mine. Sentinel-2 didn't have a result as refined as WV-3, but both mines were correctly identified.

As expected, the WV3 classification gets better results in the identification of the ground truth sites, especially in Jennyhaugen. However, false positives still appear in satellite results, on coastal lines and urban areas. The WV-3 had more false positives with urban areas, especially roads. It is worth noting that WV-3 classification identified snow as pegmatite. These false positives with snow were not as prominent for Sentinel-2, as snow pixels were mistakenly classified as vegetation. Comparing the results of Sentinel-2 with the true-color composite image reveals that in certain areas where vegetation, rocks, and snow share the same space, the spatial resolution of Sentinel-2 may have been a limiting factor. In these areas, vegetation might have been more dominant than other classes, suppressing information from the others. WV-3, with higher spatial resolution, was better able to distinguish between elements in the same area. False positives with snow/ice were expected, allowing us to assess the method's efficiency in eliminating these false positives in the prospectivity map.

Introducing a poor class separability between pegmatite and granite, it's not surprising that the WV-3 had more misidentified pixels between these two classes. The confusion matrix can be observed in Fig. 11.

4.4. MTMF abundance mapping

The MTMF process generates a comprehensive raster rule by amalgamating all MF bands and providing a visual representation of each MF band separately in grayscale. Upon inspecting each MF band, a histogram emerges. Fine-tuning this histogram enables the precise selection of pixels containing the target material. Analyzing the histogram (Fig. 12) is of great importance to ensure that the target information was correctly selected. Knowing the target (endmember) is in the upper tail of the histogram, it is possible to adjust the histograms to maximize target information and minimize false positives.

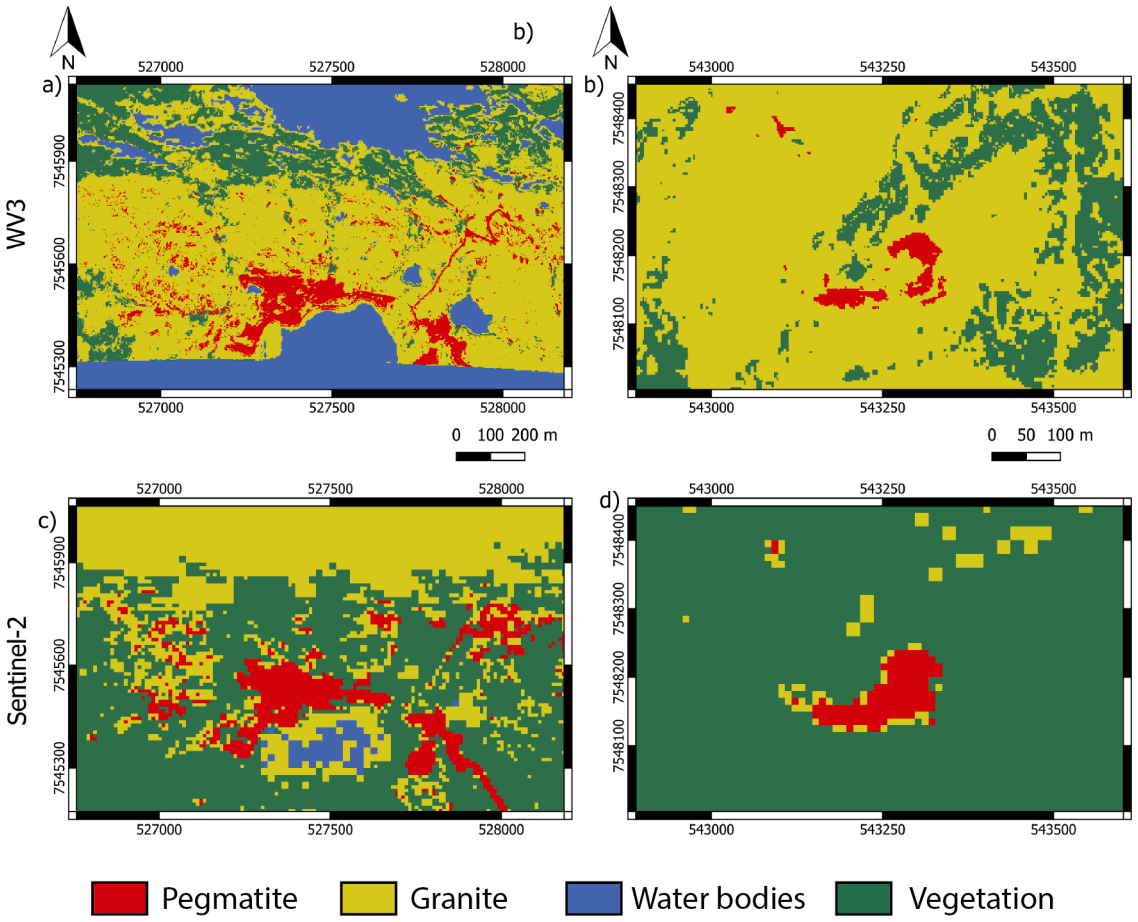

Fig. 10. LightGBM classification result for WV-3 and Sentinel-2 satellites. a) WV-3 classification results highlighting Håkonhals. b) WV-3 classification results highlighting Jennyhaugen. c) Sentinel-2 classification results highlighting Håkonhals. d) Sentinel-2 classification results highlighting Håkonhals.

Fig. 12 shows the best stretch of the lines that best correspond to each satellite. We can observe in Fig. 12 a) that the best stretch for WV-3 was between the values 0.253 and 0.587. Fig. 12 b) shows us that the stretch between 0.228 and 0.638 was the best option for selecting the pegmatite information for the Sentinel-2 data.

As well as being able to see the results of the MTMF, we were also able to see the relationship between pixels with an abundance of pegmatite and a lack of lineaments. This analysis is important to justify the structural control on pegmatite emplacement, but first, let's analyze the MTMF result. The MTMF shows a similar result in Håkonhals for both WV-3 (Fig. 13 a) and Sentinel-2 (Fig. 13 ff). In both satellites, the MTMF was capable of correctly identifying where the pegmatite is more exposed (240240 to 255 values). The spatial patterns observed in Håkonhals predominantly exhibit pixels with medium to maximum values (144144 to 255 values), surrounded by clusters of pixels with the lowest values (00 to 127 values). In Jennyhaugen, a more distinguishable spectral behavior is observed. For WV-3, the MTMF was capable of correctly identifying where the pegmatite is more exposed (Fig. 13 hh), as well as traces of pegmatite in the middle of the Jennyhaugen mine. For Sentinel-2, the MTMF identified the Jennyhaugen mine with lower values than for the WV-3 data (Fig. 13 cc). The location where the pegmatite is more exposed was assigned medium values (128128 to 143) and the other parts of the mines have even lower values (00 to 127). This is likely due to the absence of distinct spectral bands in Sentinel-2 for identifying the areas where pegmatite is more prominently exposed in the Jenny Haugen mine.

Given that the maximum values (240 to 255) in WV-3 satellite results indicate pixels where the pegmatite is more exposed, normally, at the mine and recognizing that this scenario is less common in nature, we are specifically examining spatial patterns in 144 to 159 values, 160 to 175 values, 176 to 191 and 192 to 207 values to identify pegmatites. These pixels can be surrounded by clusters with lower values than the pixels with a lower abundance of pegmatite. This spectral behavior is more specific to WV3, for Sentinel-2 we observed clusters with maximum values (up to 255 values) in areas other than mines. In general, the MTMF showed good performance for both WV-3 and Sentinel-2 satellites. The WV-3 satellite's highest GSD gives it an advantage over Sentinel-2 results. However, the MTMF, when applied through the SHW workflow, can still produce acceptable results for multispectral satellites.

The study area encompasses 30 known pegmatite points, out of which six exhibit detectable exposure through orbiting satellites (excluding the Håkonhals and Jennyhaugen mines). The MTMF applied to the WV-3 satellite successfully identified all of these six points. In contrast, the application of the same method to the Sentinel-2 bands identifies only two points and with a low distribution of pegmatites. Additionally, the LGBM applied to WV-3 bands identified 4 of the known pegmatite points, while the Sentinel-2 bands managed to identify 3 of the known points.

Regarding the connection between pegmatites and linear structures, some points mapped by the MTMF were chosen for comparison between the pegmatite distribution and the LDM. We can see a relationship between the Håkonhals (Fig. 13b, g) and Jennyhaugen (Fig. 13c, h) mines for both satellites. While Håkonhals is in an area of high lineament density, Jennyhaugen is in an area of low density. Concerning Sentinel-2, we can highlight two points. One, 3.5 km northeast of Håkonhals (Fig. 13 d), is a cluster of clusters indicating the abundance of pegmatites. These clusters are in a location with a medium to low density of

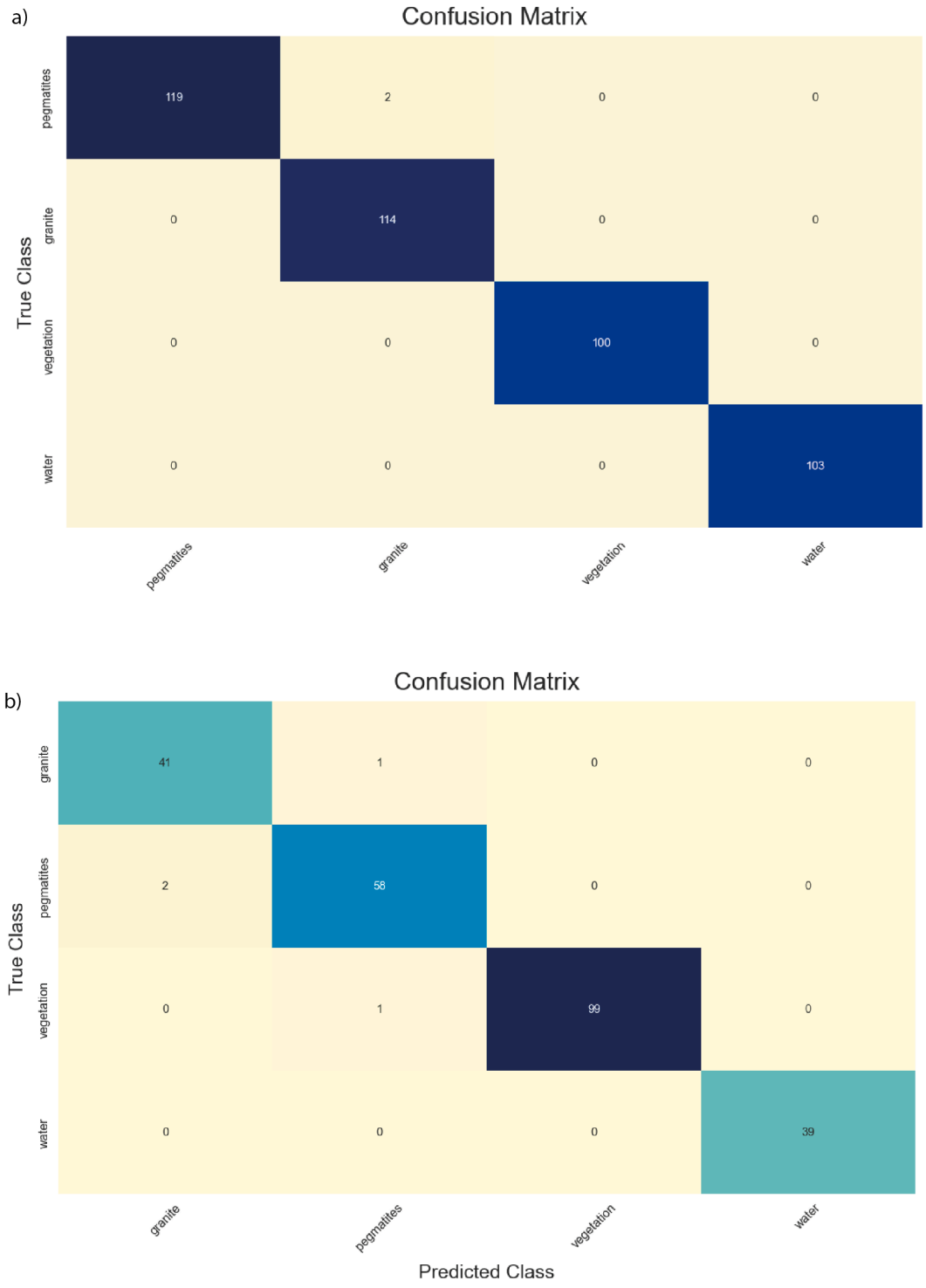

Fig. 11. Confusion matrix using: a) WV-3 satellite data; b) Sentinel-2 satellite data.

lineaments. Another point is 3.8 km from Håkonhals (Fig. 13e) and is also in an area of medium lineament density. Concerning WV3, we can highlight the point 3.5 km from the Håkonhals (Fig. 13i), this point was also highlighted by Sentinel-2 (Fig. 13d).

It is possible to validate the precision of the MTMF results by examining a 2D scatter plot that displays the distribution of mapped endmembers (Fig. 14). In this plot, background noise is represented by points with MF scores around or below zero. False-positive pixels are characterized by high infeasibility (IF) values, while correctly mapped pixels present low IF values coupled with high MF scores. This analysis reveals that correctly identified endmembers are clustered over Håkonhals, one of the ground truth areas, thus confirming the precision of endmember selection within the workflow for WV-3.

The final prospectivity map successfully identified three of the six known pegmatite points in the study area, excluding the Håkonhals and Jennyhaugen mines, while the MTMF method alone identified all six points. However, the final map significantly reduced the number of false positives compared to the MTMF results. This reduction in false positives is a crucial improvement, enhancing the efficiency of the mapping process for pegmatite prospecting by minimizing confusion with urban

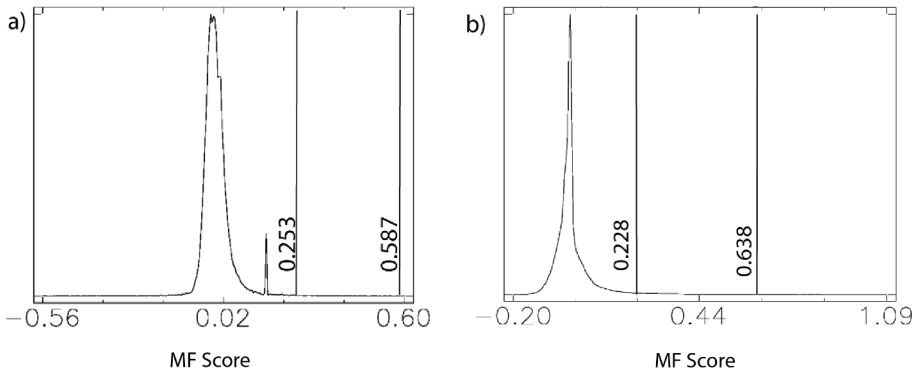

Fig. 12. MTMF histogram. for WV-3 (a) and Sentinel-2 (b) imagery.

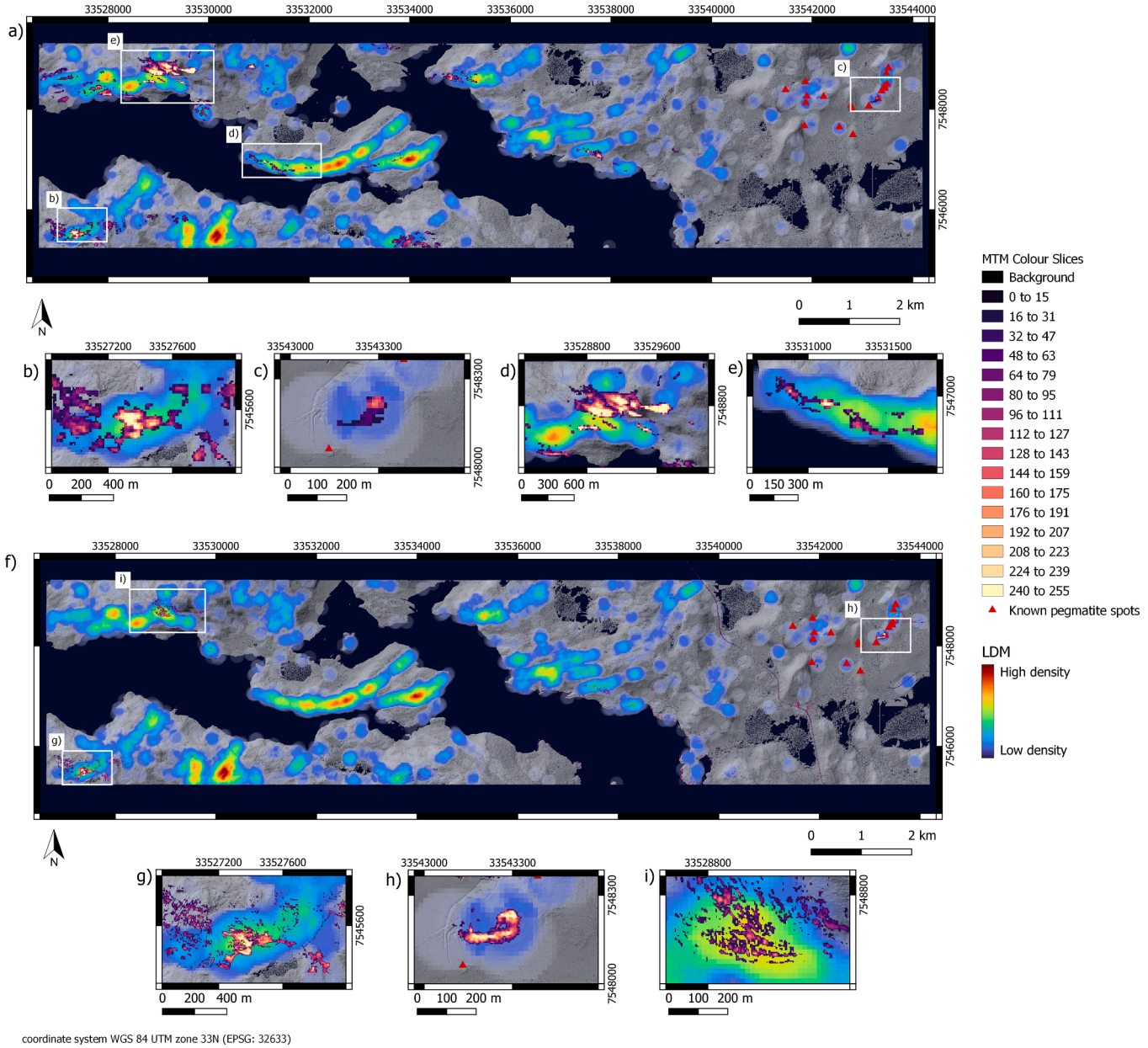

Fig. 13. MTMF results and LDM results allow the structural control on pegmatite emplacement to be observed. a) Results for Sentinel-2 satellite. b) Results for WV3 satellite.

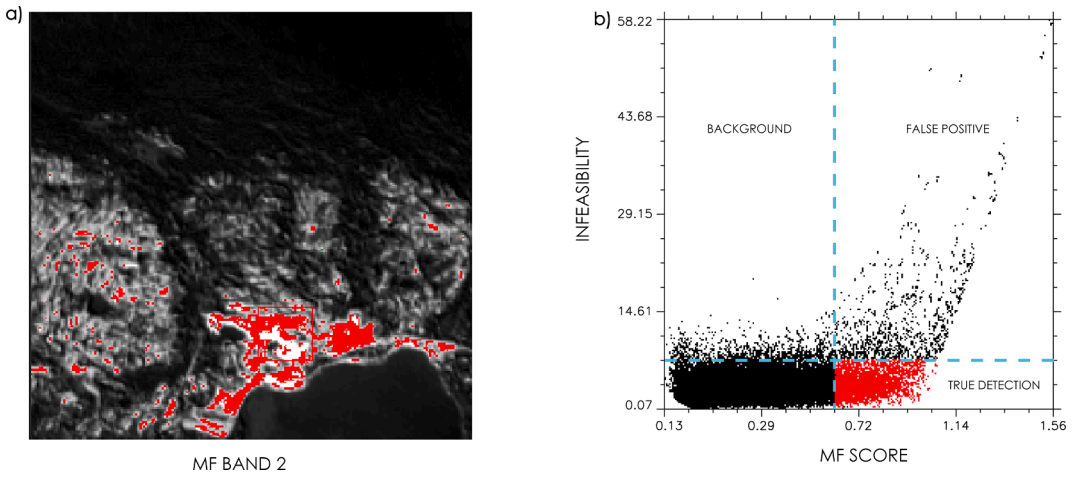

Fig. 14. Scatter plot for detecting correctly mapped pixels in relation to the WV3 satellite. a) MF 2 band in grayscale. Correctly mapped pixels are highlighted in red. b) Scatter plot between MF Score and MTMF infeasibility result.

areas, coastal lines, and other geological formations that traditionally introduce noise into the data.

4.5. Mineral prospectivity mapping

4.5.1 Model 1 results

Model 1 demonstrates variability among the outputs of each set, despite utilizing satellite images of the same study area. This underscores the importance of thoroughly testing each function and selecting the most suitable one. The results indicate that the outputs of these functions vary depending on the input data. In other words, even when employing the same function, the results may differ if the input data varies.

As observed in Table 5, the best first fuzzy operator for Model 1 was the "Prod" operator, while in the second operator, the "Gamma" operator was optimal for both the WV-3 and Sentinel-2 data. For set 2, the best fuzzy operators were the "Prod" operator for WV-3 and the "Gamma" operator for Sentinel-2. The best outputs from both sets were used in synergy to create the final prospectivity map, with the "Gamma" operator proving to be the most effective function for both the WV-3 and Sentinel-2 data.

In the obtained results, pegmatites exhibit values of $\geq$ 0.59 on Sentinel-2, on a scale from 0 to 0.92, whereas on the WV3, pegmatite points show values of $\geq$0.18, on a scale from 0 to 0.62. The Sentinel-2 output successfully eliminated all false positives associated with snow and coastal regions. The Sentinel-2 satellite is also effective in reducing false positives in urban areas, outperforming WV-3 in specific locations. However, when considering the overall context, there is no significant difference between the two satellites in terms of false positive elimination in urban areas. Furthermore, WV-3 exhibits pixels along coastal areas with values ranging from 0.001 to 0.11, where these low values assist in distinguishing them from potential pegmatite points.

Table 5
Statistical data for Håkonhals and Jennyhaugen.

	Jennyhaugen		Håkonhals		
Pixels Count	7		101		
Mean	3.2		5.2		
Standard Deviation	1.3		0.8		
Fuzzy Value	Count	Percent		Count	Percent
0.56–0.63	1	14.29 %	–		–
0.69–0.76	3	42.85 %	2		1,98 %
0.76–0.82	2	28.57 %	16		15.84 %
0.82–0.89	1	14.29 %	34		33.67 %
0.89–0.95	–	–	49		48.51 %

When comparing the results of Model 1 with the inputs utilized in each set (MTMF and LGBM), it becomes apparent that the methodology employed effectively reduced false positives, especially concerning the Sentinel-2 data. This underscores the substantial potential of Sentinel-2 for pegmatite prospecting using the proposed method, despite being free data with lower special resolution than WV-3. However, the fuzzification on WV-3 managed to identify more potential pegmatite spots, which may be related to its higher spatial and spectral resolution.

4.5.2. Model 2 results

Model 2 combines all processing data into a single set. For the Sentinel-2 data, much like in Model 1, the gamma fuzzy operator yielded the best results, eliminating false positives. The overall outcomes of Model 2 for Sentinel-2 data closely resemble those of Model 1, highlighting the same potential points with minimal differences. Pegmatite points in the Sentinel-2 output have values $\geq$ 0.76 (on a scale from 0 to 0.95). Moving to the WV-3 data, the "Prod" operator produced the best results, with pegmatite points having values $\geq$ 0.16 (ranging from 0 to 0.36). However, false positives with urban areas persisted, and coastal lines appeared with very low values (close to 0). Similarly, to the Sentinel-2 data, the WV-3 output of Model 2 closely resembles the final output of Model 1. However, Model 2 identified fewer pixels as pegmatites in the WV-3 data. Comparing the two models, it is apparent that Sentinel-2 performed better in Model 2, while WV-3 produced better results in Model 1. This suggests that the method proposed in Model 1 may be more suitable for handling more robust data, such as that from WV-3.

The application of additional fuzzy operations in Model 1 may explain the difference in results for the WV-3 data. In Model 1, the "Prod" function was applied at least once to the WV-3 data before the "Gamma" operation in both sets 1 and 2. In contrast, Model 2 applied only one fuzzy operation (Gamma) to the entire set. Regarding the Sentinel-2 data, the "Gamma" function was applied twice in Model 1: once in Set 2 and again in the final step to compose the prospecting map. The similar output of the Sentinel-2 data for Models 1 and 2 can be attributable to both models having a greater influence from the "Gamma" operator.

4.5.3. Model 3 results

Model 3 performs the fuzzification of processing data without the LDM input. The best output for Sentinel-2 data is obtained with the "And" operator. Pegmatites receive a value $\geq$ 0.80 (on a scale from 0 to 1). There is a greater influence from the MTMF classification. False positives with snow have been eliminated; however, false positives with

coastal lines and urban areas such as roads still appear but with moderate values. There is not much difference between this result and the MTMF classification used as input. The best operation for WV-3 data was the "Gamma" operator. Pegmatites receive a value ≥ 0.86 (on a scale from 0 to 1). It eliminates false positives with urban areas and reduces false positives with coastal areas compared to the LGBM classification; however, the result is very similar to the MTMF classification used as input.

The absence of LDM results in outcomes that closely resemble only one of the inputs used, providing little enhancement in false positive elimination when compared with Models 1 and 2. This underscores the significance of incorporating LDM in fuzzification for NYF pegmatite prospecting.

An important observation from the results is that, across all models, the operations "And", "Gamma," and "Prod" yielded similar outcomes, aligning well with the study's objectives. On the other hand, the "Or" and "Sum" operations also produced similar results, but they diverged the most from the study's goals. These operations generated maps with aggregated information from all inputs, resulting in visually cluttered outputs that hindered the clear and objective identification of potential pegmatite points in the study area.

The best function for the fuzzification of the final outputs from each satellite was the "sum" function. This function ensured the combination of data from both WV-3 and Sentinel-2 in Model 1 without any loss of information. This allowed for the analysis of results from both satellites in a single raster.

5. Points of interest for exploration

An analysis of the fuzzy scores of Håkonhals and Jennyhaugen was conducted to select the points of interest based on the fuzzy logic approach to mineral prospectivity mapping. As indicated in Table 5, Håkonhals demonstrates notably higher values, ranging from a minimum of 0.69 to a maximum of 0.95, with an average of 5.2 units. Overall, Håkonhals accounted for 101 pixels corresponding to pegmatites. Among these 101 pixels, 48.51% exhibit values classified as very high to maximum, 33.67% as very high, 15.84% as high to very high, and 1.98% as high. When focusing solely on Håkonhals, point selection would be limited to values above 0.69 units. However, Jennyhaugen exhibits lower values, ranging from a minimum of 0.56 to a maximum of 0.76, with an average of 3.28 units. Only seven pixels were identified as pegmatites in Jennyhaugen. Of these pixels, 14.29% correspond to medium values, 42.85% correspond to high values, 28.57% to very high values, and 14.29% to very high to maximum values. Unlike Håkonhals, no pixels were classified with the maximum value (0.95). Observing the fuzzy scores behavior of Håkonhals and Jennyhaugen, it was possible to determine that values between 0.56 and 0.89 correspond to the highest abundance of pegmatites.

Applying a color slice allowed for the selection of values of interest (between 0.56 and 0.89), which were then analyzed in a GIS environment to determine the main points of interest.

Fig. 15 presents a graph illustrating the relationship between the quantity of pixels and their corresponding fuzzy values. Upon analyzing these statistical data from the final prospecting map, we observe that pixels with very high to maximum values (between 0.89 and 0.95) are the least represented in the image (13% of the total pixels), whereas pixels with values between 0.69 and 0.76 are the most represented (19.8%). This was expected, as pixels with values closer to the maximum are more likely to be pegmatites. Consequently, few pixels are classified with maximum values, enabling clear and objective identification, and facilitating the differentiation of false positives with granites. Thus, in selecting potential prospecting points, priority was given to pixels with values ≥0.76 and ≤0.95. Pixels with values <0.76 were then considered false positives with granite.

A total of 19 points of interest and two areas of interest were identified on the final map. These areas of interest correspond to locations

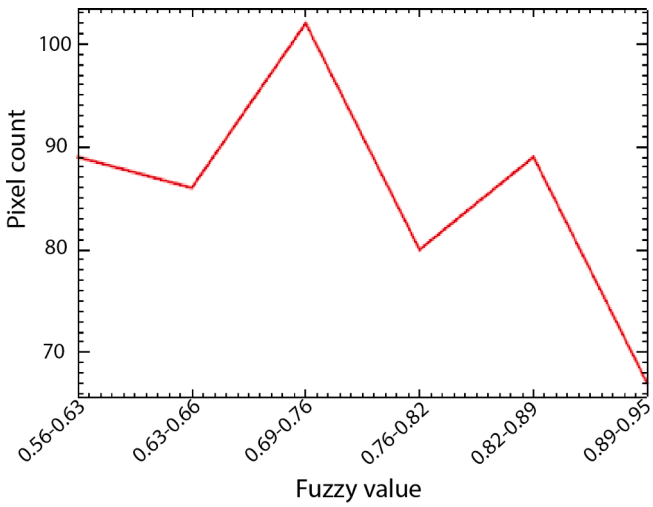

Fig. 15. Relationship between the quantity of pixels and their corresponding fuzzy values in the final prospecting map.

exhibiting notable clustering of points. The points and areas of interest are depicted in Fig. 16. **Points 1 to 5 (**Fig. 16 **d-g):** surround the Håkonhals mine, in the southwest zone of the study area. **Points 6, 8 to 13 (**Fig. 16 **h, j-n)**: are located 3 km north of the Håkonhals mine and 3.5 km south of the village of Presteid. **Points 7, 14, and 15 (**Fig. 16 **i, o):** are located 3.5 km north of the Håkonhals mine and 2 km southwest of Kaldvag village. **Point 16 (**Fig. 16 **p):** is located 1 km **southwest** of Kaldvag village.

Point 17 (Fig. 16 **q):** is located 500 m southeast of Møllnvatnet Lake. **Points 18 and 19 (**Fig. 16 **r, s):** Both spots are located 5.6 km distance from Jennyhaugen mine and near Klottervatnet lake. **Area 1:** This is located near form points 14, 15 and 7 and 3.5 km north of the Håkonhals mine and 2 km southwest of the Kaldvag village. **Area 2:** Is located 2 km northeast of the Kaldvag village.

6. Conclusion

Working with different fuzzification models was crucial to analyze the effectiveness of each approach regarding the specific challenges of NYF pegmatite prospecting and allowed us to identify which one achieved the best result. All inputs utilized in the tested models can be derived through data processing methods that do not require extensive and complex data from the study area, such as geochemical and geophysical data. This enables application of this method in greenfield sites facilitating exploration in previously unexplored regions with limited available data. The ramp-like Fe^{2+} absorption marks the presence of biotite, and it is marked in granite samples showing to be a potential key to distinguish between granite and pegmatite. The ramp-like Fe^{2+} can be detected by enough bands in the VNIR and SWIR region to be detected even by multispectral satellites. Comparing the results of the classification methods, it is notable that MTMF achieved more refined results, managing to distinguish different concentrations of pegmatites within the Jennyhaugen and Håkonhals mines. However, the LGBM algorithm was able to identify both mines perfectly, also demonstrating itself to be a powerful method in pegmatite identification Model 1 achieved good performance in both satellites. The application of more fuzzy operators in Model 1 may have been the key to its success. Model 2 had a very close result to model 1, however, model 1 stands out in terms of reducing false positives. The final prospecting map was generated by fuzzifying the results from Sentinel-2 and WV-3 in Model 1. In this map, the pixels of interest (potentially corresponding to pegmatites) have values ≥0.76 and ≤0.95. Nineteen points of interest and two areas of interest were selected for validation by fieldwork. In this study, the focus was on a method that can be adapted to greenfield areas.

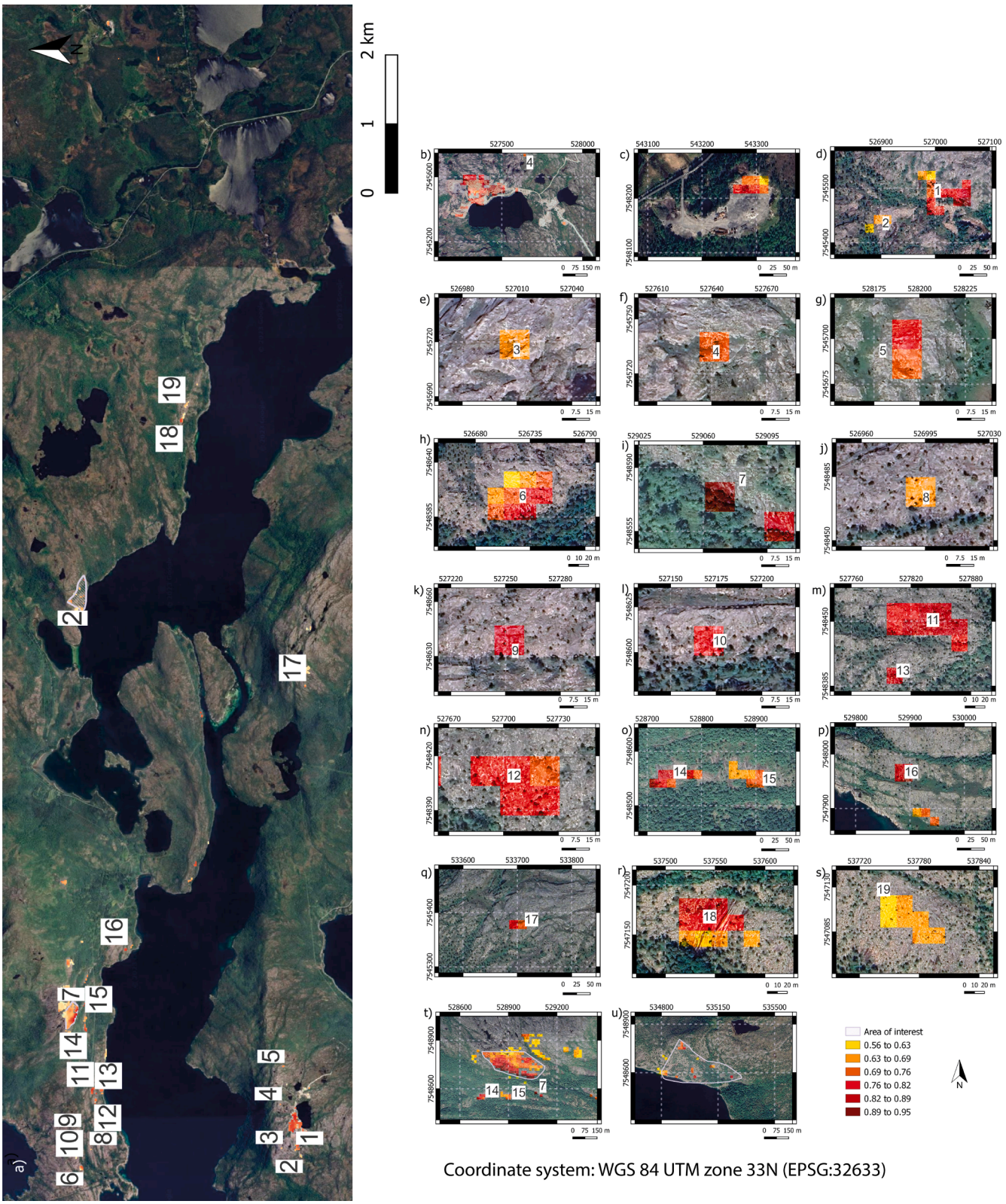

Fig. 16. Distribution of Points and Areas of Interest in the Study Area. a) Overview of the study area showing the identified points of interest and areas of interest. Notable clustering of points is observed around specific geographic features. b) Fuzzy scores in the Håkonhals mine. c) Fuzzy scores in the Jennyhaugen mine. d) Points of interest 1 and 2. e) Point of interest 3. f) Points of interest 4. g) Point of interest 5. h) Point of interest 6. I) Point of interest 7. j) Point of interest 8. k) Point of interest 9. L) Point of interest 10. m) Points of interest 11 and 13. n) Point of interest 12. o) Points of interest 14 and 15. p) Points of interest 16. q) Point of interest 17. r) Point of interest 18. s) Point of interest 19. t) Area of interest 1. u) Area of interest 2.

However, this method can be evaluated with other inputs for Brownfield areas with more known data.

Declaration of generative AI in scientific writing

During the preparation of this work, the author(s) used Grammarly (2023 © Grammarly Inc.) and ChatGPT (ChatGPT 3.5) to improve the readability and flow of the writing. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

CRediT authorship contribution statement

Douglas Santos: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Visualization, Project administration. **Antônio Azzalini:** Methodology, Software, Formal analysis, Data curation, Writing – review & editing, Visualization. **Ariane Mendes:** Formal analysis, Writing – review & editing, Funding acquisition. **Joana Cardoso-Fernandes:** Formal analysis, Writing – review & editing, Visualization, Project administration. **Alexandre Lima:** Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Axel Müller:** Resources, Writing – review & editing, Project administration. **Ana C. Teodoro:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Douglas Santos reports financial support was provided by Institute of Earth Sciences. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The work is co-funded by Institute of Earth Sciences by FCT, 632 Projects UIDB/04683/2020 (https://doi.org/10.54499/UIDB/04683/2020) and UIDP/04683/2020 633 (http s://doi.org/10.54499/UIDP/04683/2020). Douglas Santos was financially supported by Portuguese national funds through FCT (Grant: UI/BD/154412/2023). The work was also supported by European Union's Horizon 2020 innovation programme under grant agreement No 869274, project GREENPEG: New Exploration Tools for European Pegmatite Green-Tech Resources. Institute of Earth Sciences by FCT, 632 Projects UIDB/04683/2020 (https://doi.org/10.54499/UIDB/04683/2020) and UIDP/04683/2020 633 (https://doi.org/10.54499/UIDP/04683/2020).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oregeorev.2024.106347.

Data availability

Data will be made available on request.

References

Abdelkader, M.A., Watanabe, Y., Shebl, A., El-Dokouny, H.A., Dawoud, M., Csámer, Á., 2022. Effective delineation of rare metal-bearing granites from remote sensing data using machine learning methods: A case study from the Umm Naggat Area, Central Eastern Desert, Egypt. Ore Geol. Rev. 150, 105184. https://doi.org/10.1016/j.oregeorev.2022.105184.

Abdelkareem, M., Al-Arifi, N., 2021. Synergy of remote sensing data for exploring hydrothermal mineral resources using GIS-based Fuzzy logic approach. Remote Sens. https://doi.org/10.3390/rs13224492.

Adiri, Z., Lhissou, R., El Harti, A., Jellouli, A., Chakouri, M., 2020. Recent advances in the use of public domain satellite imagery for mineral exploration: A review of Landsat-8 and Sentinel-2 applications. Ore Geol. Rev. 117. https://doi.org/10.1016/j.oregeorev.2020.103332.

Ahmad, F., 2013. Pixel Purity Index Algorithm and n-Dimensional Visualization for ETM + Image Analysis: A Case of District Vehari.

Ali, H.F., Abu El Ata, A.S.A., Youssef, M.A.S., Salem, S.M., Ghoneim, S.M., 2023. A newly-developed multi-algorithm integration technique for mapping the potentially mineralized alteration zones, Egypt. J. Remote Sens. Sp. Sci. 26, 691–711. https://doi.org/10.1016/j.ejrs.2023.07.008.

Alwosheel, A., van Cranenburgh, S., Chorus, C.G., 2018. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. J. Choice Model. 28, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002.

Andersson, U., Gorbatschev, R., A.Wikström, Högdahl, K., M.Ahl, Nyström, J.-O., Sjöström, H., Bergman, S., Eklund, O., Claeson, D., Mansfeld, J., Stephens, M., Wahlgren, C.-H., Lundqvist, T., Smeds, S.-A., Sundblad, K., Öhlander, B., 2004. The Transscandinavian Igneous Belt in Sweden: a review of its character and evolution.

Asadzadeh, S., de Souza Filho, C.R., 2016. A review on spectral processing methods for geological remote sensing. Int. J. Appl. Earth Obs. Geoinf. 47, 69–90. https://doi.org/10.1016/j.jag.2015.12.004.

Badr, Y.S., 2021. Integrated spectral analysis of ASTER and gamma-ray spectrometric data to delineate alteration minerals for uranium exploration in Gabal Khashm El-Risha area, North Eastern Desert, Egypt. Remote Sens. Appl. Soc. Environ. 24, 100632. https://doi.org/10.1016/j.rsase.2021.100632.

Bai, Y., Wang, J., Zhou, K., Wang, S., Jiang, G., Zhou, S., Cui, S., Yan, J., Wu, M., Ma, X., Fan, X., 2024. Mapping lithium-bearing pegmatite with short-wave infrared (SWIR) hyperspectral imaging of Jingerquan Li–Be–Nb–Ta pegmatite deposit, eastern Tianshan, NW China. IEEE Trans. Geosci. Remote Sens. 62, 1–15. https://doi.org/10.1109/TGRS.2023.3341496.

Banerjee, K., Jain, M.K., Jeyaseelan, A.T., Panda, S., 2019. Landsat 8 OLI data for identification of hydrothermal alteration zone in Singhbhum Shear Zone using successive band depth difference technique - A new image processing approach. Curr. Sci. 116, 1639–1647. https://doi.org/10.18520/cs/v116/i10/1639-1647.

Boadi, B., Sunder Raju, P.V., Wemegah, D.D., 2022. Analysing multi-index overlay and fuzzy logic models for lode-gold prospectivity mapping in the Ahafo gold district – Southwestern Ghana. Ore Geol. Rev. 148, 105059. https://doi.org/10.1016/j.oregeorev.2022.105059.

Bonham-Carter, G., 1994. Geographic Information Systems for Geoscientists: Modelling with GIS, Computer methods in the geosciences. Elsevier Science & Technology Books.

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

Cardoso-Fernandes, J., Teodoro, A.C., Santos, D., de Almeida, C., Lima, A., 2023b. Spectral Library of European Pegmatites, Pegmatite Minerals and Pegmatite Host-Rocks – The Greenpeg Database. https://doi.org/10.5281/zenodo.7938407.

Cardoso-Fernandes, J., Teodoro, A.C., Lima, A., 2019. Remote sensing data in lithium (Li) exploration: A new approach for the detection of Li-bearing pegmatites. Int. J. Appl. Earth Obs. Geoinf. 76, 10–25. https://doi.org/10.1016/j.jag.2018.11.001.

Cardoso-Fernandes, J., Teodoro, A.C., Lima, A., Roda-Robles, E., 2020. Semi-automatization of support vector machines to map lithium (Li) bearing pegmatites. Remote Sens. 12. https://doi.org/10.3390/rs12142319.

Cardoso-Fernandes, J., Santos, D., de Almeida, C., Lima, A., Teodoro, A.C., project team, G., 2023a. Spectral Library of European Pegmatites, Pegmatite Minerals and Pegmatite Host-Rocks – the GREENPEG project database. Earth Syst. Sci. Data 15, 3111–3129. https://doi.org/10.5194/essd-15-3111-2023.

Cerny, P., 1991. Rare-element granitic pegmatites. Part I: anatomy and internal evolution of pegmatite deposits. Geosci. Canada 18, 49–67.

Chaves, C., Pereira, E., Ferreira, P., Guerner Dias, A., 2021. Concerns about lithium extraction: A review and application for Portugal. Extr. Ind. Soc. 8, 100928. https://doi.org/10.1016/j.exis.2021.100928.

Chirico, R., Mondillo, N., Laukamp, C., Mormone, A., Di Martire, D., Novellino, A., Balassone, G., 2023. Mapping hydrothermal and supergene alteration zones associated with carbonate-hosted Zn-Pb deposits by using PRISMA satellite imagery supported by field-based hyperspectral data, mineralogical and geochemical analysis. Ore Geol. Rev. 152, 105244. https://doi.org/10.1016/j.oregeorev.2022.105244.

Dao, R.A.I., Ilboudo, H., Naba, S., 2022. Nouvelles données lithologiques et structurales du secteur de Nindangou dans le prolongement Est de la ceinture de Goren (Burkina Faso-Afrique de l'Ouest). Bull. L'institut Sci. 93–103.

Ding, W., Ding, L., Li, Q., Li, J., Zhang, L., 2023. Lithium-Rich Pegmatite Detection Integrating High-Resolution and Hyperspectral Satellite Data in Zhawulong Area, Western Sichuan, China. Remote Sens https://doi.org/10.3390/rs15163969.

El-Desoky, H.M., Tende, A.W., Abdel-Rahman, A.M., Ene, A., Awad, H.A., Fahmy, W., El-Awny, H., Zakaly, H.M.H., 2022. Hydrothermal Alteration Mapping Using Landsat 8 and ASTER Data and Geochemical Characteristics of Precambrian Rocks in the Egyptian Shield: A Case Study from Abu Ghalaga, Southeastern Desert, Egypt. Remote Sens https://doi.org/10.3390/rs14143456.

Ellis, R.J., Scott, P.W., 2004. Evaluation of hyperspectral remote sensing as a means of environmental monitoring in the St. Austell China clay (kaolin) region, Cornwall, UK. Remote Sens. Environ. 93, 118–130. https://doi.org/10.1016/j.rse.2004.07.004.

Forson, E.D., Menyeh, A., Wemegah, D.D., 2021. Mapping lithological units, structural lineaments and alteration zones in the Southern Kibi-Winneba belt of Ghana using integrated geophysical and remote sensing datasets. Ore Geol. Rev. 137, 104271. https://doi.org/10.1016/j.oregeorev.2021.104271.

Gemusse, U., Cardoso-Fernandes, J., Lima, A., Teodoro, A., 2023. Identification of pegmatites zones in Muiane and Naipa (Mozambique) from Sentinel-2 images, using band combinations, band ratios, PCA and supervised classification. Remote Sens. Appl. Soc. Environ. 32, 101022. https://doi.org/10.1016/j.rsase.2023.101022.

Ghoneim, E.-S., 2018. Application of remote sensing techniques on aster data for alteration zones extraction and lithological mapping of elfawakhir – el-sid area, central eastern desert, egypt: an approach for gold exploration. Egypt. J Geol.

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

Hafid, A.A., 2004. Multispectral Analysis Of Satellite Images.

Hajihosseinlou, M., Maghsoudi, A., Ghezelbash, R., 2023. A novel scheme for mapping of MVT-Type Pb–Zn prospectivity: lightGBM, a highly efficient gradient boosting decision tree machine learning algorithm. Nat. Resour. Res. 32, 2417–2438. https://doi.org/10.1007/s11053-023-10249-6.

Husdal, T., 2008. The minerals of the pegmatites within the Tysfjord granite, northern Norway. Bergverksmuseets Skrift. nor. Bergverksmus. Skr. 38, 5–28.

Jiang, Q., Dai, J., Wang, D., Wang, C., Tian, S., 2023. Lithium-bearing Pegmatite Exploration in Western Altun, Xinjiang, using Remote-Sensing Technology. Acta Geol. Sin. - English Ed. 97, 681–694. https://doi.org/10.1111/1755-6724.15025.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet Classification with Deep Convolutional Neural Networks. Commun. ACM 60, 84–90. https://doi.org/10.1145/3065386.

Lewis, S.M., Fitts, G., Kelly, M., Dale, L., 2014. A fuzzy logic-based spatial suitability model for drought-tolerant switchgrass in the United States. Comput. Electron. Agric. 103, 39–47. https://doi.org/10.1016/j.compag.2014.02.006.

London, D., 2016. Rare-Element Granitic Pegmatites. Rare Earth Crit. Elem. Ore Depos. https://doi.org/10.5382/Rev.18.08.

Makropoulos, C.K., Butler, D., Maksimovic, C., 2008. In: Building Decision Support Systems Based on Fuzzy Inference BT - Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications. Springer, Berlin Heidelberg, pp. 215–228. https://doi.org/10.1007/978-3-540-79881-1_16.

Mamouch, Y., Attou, A., Miftah, A., Ouchchen, M., Dadi, B., Achkouch, L., Et-Tayea, Y., Allaoui, A., Boualoul, M., Randazzo, G., Lanza, S., Muzirafuti, A., 2022. Mapping of Hydrothermal Alteration Zones in the Kelâat M'Gouna Region Using Airborne Gamma-Ray Spectrometry and Remote Sensing Data: Mining Implications (Eastern Anti-Atlas, Morocco). Appl. Sci. 12. https://doi.org/10.3390/app12030957.

Mashkoor, R., Ahmadi, H., Rahmani, A.B., Pekkan, E., 2022. Detecting Li-bearing pegmatites using geospatial technology: the case of SW Konar Province, Eastern Afghanistan. Geocarto Int. 37, 14105–14126. https://doi.org/10.1080/10106049.2022.2086633.

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

L3Harris, n.d. Spectral Hourglass [WWW Document]. URL https://www.nv5geospatialsoftware.com/docs/SpectralHourglassWorkflow.html (accessed 10.7.22).

Nv5geospatial, n.d. Mixture Tuned Matched Filtering [WWW Document]. URL https://www.nv5geospatialsoftware.com/docs/MTMF.html (accessed 11.2.21).

Popchev Ivan, V.P., 1999. Fuzzy logic operators in decision-making. Cybern. Syst. 30, 725–745. https://doi.org/10.1080/019697299124966.

Porwal, A., Carranza, E.J.M., Hale, M., 2003. Knowledge-driven and data-driven fuzzy models for predictive mineral potential mapping. Nat. Resour. Res. 12, 1–25. https://doi.org/10.1023/A:1022693220894.

Rajan Girija, R., Mayappan, S., 2019. Mapping of mineral resources and lithological units: a review of remote sensing techniques. Int. J. Image Data Fusion 10, 79–106. https://doi.org/10.1080/19479832.2019.1589585.

Rajesh, H.M., 2004. Application of remote sensing and GIS in mineral resource mapping - An overview. J. Mineral. Petrol. Sci. 99, 83–103. https://doi.org/10.2465/jmps.99.83.

Roger, Clark, 1999. Spectroscopy of Rocks and Minerals , and Principles of Spectroscopy, in: Remote Sensing for the Eartb Sciences: Manual of Remote Sensing, 3 Ed., Vol. 3. US. Geological Survey, Denver, Colorado, pp. 3–58.

Romer, R., Kjøsnes, B., Korneliussen, A., Lindahl, I., Skyseth, T., Stendal, M., Sundvoll, B., 1992. The Archaean–Proterozoic boundary beneath the Caledonides of northern Norway and Sweden: U–Pb, Rb–Sr and Nd isotopic data from the Rombak–Tysfjord area.

Sabbaghi, H., Tabatabaei, S.H., 2023. Data-driven logistic function for weighting of geophysical evidence layers in mineral prospectivity mapping. J. Appl. Geophys. 212, 104986. https://doi.org/10.1016/j.jappgeo.2023.104986.

Sabins, F.F., 1999. Remote sensing for mineral exploration. Ore Geol. Rev. 14, 157–183. https://doi.org/10.1016/S0169-1368(99)00007-4.

Santos, D., Cardoso-Fernandes, J., Campos de Lima, A.M., Teodoro, A.C.M., 2022a. The Potential of Spectral Unmixing Method Applied to PRISMA Hyperspectral Images in the Identification of Li Minerals: an Evaluation for Prospecting Purposes 43, 1226811. https://doi.org/10.1117/12.2636034.

Santos, D., Cardoso-Fernandes, J., Lima, A., Müller, A., Brönner, M., Teodoro, A.C., 2022b. Spectral analysis to improve inputs to random forest and other boosted ensemble tree-based algorithms for detecting NYF Pegmatites in Tysfjord, Norway. Remote Sens. 14, 3532. https://doi.org/10.3390/rs14153532.

Santos, D, Mendes, A., Azzalini, A., Cardoso-Fernandes, J., Lima, A., Teodoro, A.C., 2023a. Automated lineament extraction applied to high-resolution imagery Worldview-3 and LiDAR data for pegmatite mineral exploration, in: Proc.SPIE. p. 1273414. https://doi.org/10.1117/12.2679367.

Santos, Douglas, Mendes, A., Azzalini, A., Cardoso-Fernandes, J., Lima, A., Teodoro, A.C. M., 2023b. Mapping NYF pegmatite outcrops through high-resolution Worldview-3 imagery 37. https://doi.org/10.1117/12.2675815.

Sekandari, M., Masoumi, I., Beiranvand Pour, A.M., Muslim, A., Rahmani, O., Hashim, M., Zoheir, B., Pradhan, B., Misra, A., Aminpour, S.M., 2020. Application of Landsat-8, Sentinel-2, ASTER and WorldView-3 Spectral Imagery for Exploration of Carbonate-Hosted Pb-Zn Deposits in the Central Iranian Terrane (CIT). Remote Sens. 12. https://doi.org/10.3390/rs12081239.

Tripp, G.I., Vearncombe, J.R., 2004. Fault/fracture density and mineralization: a contouring method for targeting in gold exploration. J. Struct. Geol. 26, 1087–1108. https://doi.org/10.1016/j.jsg.2003.11.002.

van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., 2014. Potential of ESA's Sentinel-2 for geological applications. Remote Sens. Environ. 148, 124–133. https://doi.org/10.1016/j.rse.2014.03.022.

Wolfe, J.D., Black, S.R., 2018. Hyperspectral Analytics in ENVI.

Yamusa, I.B., Yamusa, Y.B., Danbatta, U.A., Najime, T., 2018. Geological and structural analysis using remote sensing for lineament and lithological mapping. IOP Conf. Ser. Earth Environ. Sci. 169, 12082. https://doi.org/10.1088/1755-1315/169/1/012082.

Yousefi, M., Carranza, E.J.M., 2015. Fuzzification of continuous-value spatial evidence for mineral prospectivity mapping. Comput. Geosci. 74, 97–109. https://doi.org/10.1016/j.cageo.2014.10.014.

Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8, 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X.

Zhao, Z.-F., Zhou, J.-X., Lu, Y.-X., Chen, Q., Cao, X.-M., He, X.-H., Fu, X.-H., Zeng, S.-H., Feng, W., 2021. Mapping alteration minerals in the Pulang porphyry copper ore district, SW China, using ASTER and WorldView-3 data: Implications for exploration targeting. Ore Geol. Rev. 134, 104171. https://doi.org/10.1016/j.oregeorev.2021.104171.

Zheng, X., Sarwar, A., Islam, F., Majid, A., Tariq, A., Ali, M., Gulzar, S., Khan, M.I., Sardar Ali, M.A., Israr, M., Jamil, A., Aslam, M., Soufan, W., 2023. Rainwater harvesting for agriculture development using multi-influence factor and fuzzy overlay techniques. Environ. Res. 238, 117189. https://doi.org/10.1016/j.envres.2023.117189.

Zhou, H., Müller, A., Augland, L.E., Kristoffersen, M., Erambert, M., 2022. Titanite links rare-element (meta-)pegmatite mineralization to Caledonian metamorphism. Geochim. Cosmochim. Acta 332, 285–306. https://doi.org/10.1016/j.gca.2022.06.014.

Zhou, H., Müller, A., Berndt, J., 2023. Quartz chemistry fingerprints melt evolution and metamorphic modifications in high-purity quartz deposits. Geochim. Cosmochim. Acta 356, 179–195. https://doi.org/10.1016/j.gca.2023.07.015.