

# Chemometric Analysis of Mid-Wave Infrared Spectral Reflectance Data for Sulphide Ore Discrimination

F. S. Desta<sup>1</sup> · M. W. N. Buxton<sup>1</sup>

Received: 31 October 2017 / Accepted: 21 November 2018 / Published online: 5 December 2018  
© The Author(s) 2018

**Abstract** Despite significant recent advancements in the sensor technologies, the use of sensors for raw material characterization in the mining industry remains limited. The aim of the present study was to assess the utility of applying the mid-wave infrared (MWIR) reflectance data acquired by the use of a handheld Fourier-transform infrared spectrometer (FTIR), combined with partial least squares-discriminant analysis (PLS-DA), for the characterization of a polymetallic sulphide ore deposit. In achieving the study objectives, focus was given to the MWIR portion of the FTIR dataset, as it is the least explored region of the infrared spectrum in mineral characterization studies. Three datasets—covering different wavelength ranges—were generated from the FTIR spectral data, namely the full FTIR range (2.5–15  $\mu\text{m}$ ), MWIR (2.5–7  $\mu\text{m}$ ) and long-wave infrared (LWIR: 7–15  $\mu\text{m}$ ), in order to investigate the associated information level of each defined wavelength region separately. Design of experiment was developed to determine the optimal data filtering techniques. Using the processed data and PLS-DA, a series of calibration and prediction models were developed for ore and waste materials separately. As the models applied to the MWIR data showed a successful classification rate of 86.3% for sulphide ore–waste discrimination, similarly using the full spectral FTIR dataset, a correct classification rate of 89.5% was achieved. This indicates that MWIR spectral range includes informative signals that are sufficient for classifying the material into ore or waste. The proposed approach could be extended for automating the sulphide ore–waste discrimination process, thus greatly benefiting marginally economical mining operations.

---

✉ F. S. Desta  
f.s.desta@tudelft.nl

<sup>1</sup> Resource Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

**Keywords** MWIR · FTIR · Polymetallic sulphide ore · Material discrimination · PLS-DA · Chemometrics

## 1 Introduction

In recent years, the use of sensor technologies for raw material characterization has been rapidly growing and innovative technologies are being introduced at a fast pace. However, sensor technologies are still rarely utilized for material characterization in the mining industry, due to various factors, including (1) inadequate sensor design, since most are intended for laboratory use, or are specific to a particular deposit type and operational environment; (2) the need to demonstrate the sensors' utility in the mining industry; and (3) the high initial cost related to purchasing and setting up the instruments, which may in some cases exceed the benefit to be realized. Despite the limited use of sensors, findings yielded by extant studies in this field (Buxton and Benndorf 2013; Fox et al. 2017; Goetz et al. 2009) show that the use of sensor technologies in the mining industry can improve efficiency, reduce operational cost, increase productivity, enhance safety and minimize environmental impact.

Several existing sensor technologies can be used for raw material characterization, including laser-induced breakdown spectroscopy (LIBS), Raman spectroscopy, hyper-spectral imaging, infrared technologies, and X-ray fluorescence (XRF). For example, Death et al. (2008) showed the potential for applying LIBS in online compositional determination of iron ore samples. Similarly, Kruse (1996) demonstrated that rapid acquisition of short-wave infrared (SWIR) data by the Portable Infrared Mineral Analyzer II (PIMA-II) enhances the production of drill logs and geological maps, as well as assists in the definition of alteration zones. More recently, Culka et al. (2016) investigated the potential for applying a handheld Raman spectrometer in in situ detection and discrimination of arsenate minerals at outcrops. In addition, Wells and Ramanaidou (2015) demonstrated the utility of Raman spectroscopy in automated in situ mapping of iron ore and in gangue mineralogy. According to the recent evidence, XRF analyzers can also be used for online in situ elemental analysis of bulk materials (Orbit Technologies 2017; ThermoFisher 2017). Moreover, Alov et al. (2010) demonstrated the use of XRF analyzer in iron ore mixture quality control performed directly on the conveyor belt, thus highlighting the potential for online analysis.

Infrared (IR) spectroscopy is one of the most useful analytical techniques for evaluation of organic and inorganic materials (Chukanov and Chervonnyi 2016; Griffiths and Haseth 2007; Smith 2011). It is a mature technology and provides highly reproducible analytical measurements. When infrared light interacts with a molecule, the bonds between molecule constituents selectively absorb the infrared radiation energy at specific wavelengths. The consequent changes in the vibrational energy level of the molecules can be observed through signals at specific wavelengths in the infrared spectrum. The IR region of the electromagnetic spectrum can be divided into the near infrared (NIR: 0.7–1.4  $\mu\text{m}$ ), short-wave infrared (SWIR: 1.4–2.5  $\mu\text{m}$ ), mid-wave infrared (MWIR: 2.5–7  $\mu\text{m}$ ), long-wave infrared (LWIR: 7–15  $\mu\text{m}$ ) and far infrared (FIR: 15–1000  $\mu\text{m}$ ) regions. NIR sensors can provide accurate identification of clay minerals, rock forming minerals and sulphide minerals

(Spectral Evolution 2015; Szalai et al. 2013). SWIR is one of the most widely used infrared technologies and can be applied in the identification and discrimination of phyllosilicates, sulphates and carbonates (Sun et al. 2001). In general, it is commonly employed in the identification of alteration minerals associated with mineralization. On the other hand, LWIR permits identification of rock forming minerals, whereas FIR can be employed in the rare earth mineral analyses (Clark 1999; Karr and Kovach 1969). However, to the best of the authors' knowledge, MWIR reflectance has never been used for sulphide ore discrimination. This gap in the current analytical methodology has motivated the present study.

An FTIR spectrometer simultaneously collects data in a wide spectral range in the time domain, whereby the resulting time-domain graph shows a signal changes over time. In the next step, a signal processing technique (Fourier transform) is used to convert the time-domain information to data in the frequency domain, which allows distinguishing the amount of signal within each specified frequency band over a range of frequencies (Ismail et al. 1997). Such frequency domain representation is necessary to convert the input signal into a full spectrum, which can be used to identify or quantify different materials. A FTIR spectrometer has considerable advantages over grating-based IR spectrometers. For example, it produces spectra of higher quality relative to the infrared equivalents (a higher signal to noise ratio). Its other benefits include short data acquisition time, higher accuracy, higher precision, wider scan range and high resolution (Agilent 2017; Birkner and Wang 2015; Perkins 1987; Smith 2011; Stuart 2004). Moreover, owing to the rapid technological advances, portable FTIR spectrometers can be produced, permitting their use in real-time (in situ) applications (Agilent 2017).

Multivariate data analysis (chemometrics) involves statistical and mathematical methods to process and evaluate large amounts of multivariate data (Brereton 2007; Miller and Miller 2000; Roussel et al. 2014). It includes the design of experiments and the analysis of highly complex multivariate data, in order to acquire valuable information about the entity or process under investigation. Multivariate data analysis methods are essential in understanding the different forms of relationship among variables. Therefore, as different sensors (including FTIR spectrometer) typically produce large amounts of multivariate data, chemometrics techniques can be employed to solve a wide range of problems, such as material discrimination.

In view of the preceding discussions, the objective of the present study was to develop classification and prediction models using chemometric techniques that are capable of discriminating sulphide ore and waste materials in economically suboptimal mining operations, using spectral data acquired by a handheld FTIR spectrometer. The acquired FTIR spectral data were processed to produce three datasets, namely full FTIR spectra (2.5–15  $\mu\text{m}$ ), as well as MWIR (2.5–7  $\mu\text{m}$ ) and LWIR (7–15  $\mu\text{m}$ ). This approach was adopted in order to evaluate the information level associated with each wavelength region (MWIR and LWIR) separately. In particular, as the MWIR part of the electromagnetic spectrum has been under-investigated to date due to historically limited instrument development, it was the focal point of this study.

## 2 Case Study Area, Data Acquisition and Chemical Validation Datasets

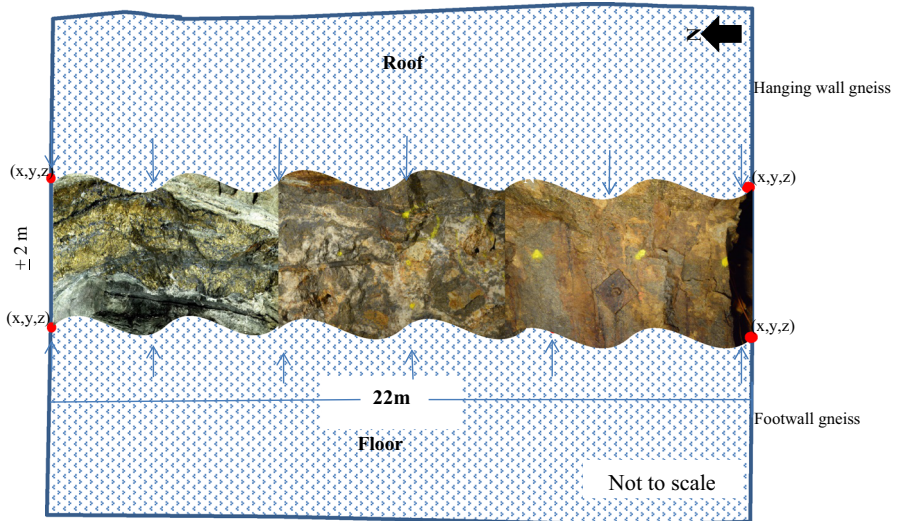
### 2.1 Case Study Area

The Reiche Zeche underground mine located in the Freiberg district, eastern Erzgebirge, Germany, served as the case study area. Originating in 1168, it is one of the oldest mines in Europe and, during its operation, it was mined for silver, copper, lead, arsenic, zinc and pyrite (Scheinert et al. 2009; Seifert and Sandmann 2006). In 1863 and 1886, the elements indium and germanium were discovered, respectively, at the local Freiberg district (Seifert and Sandmann 2006). As the mine ceased operating in 1969, the “Reiche Zeche” and “Alte Elisabeth” shafts were reconstructed in 1976, and were reopened as a research and teaching mine.

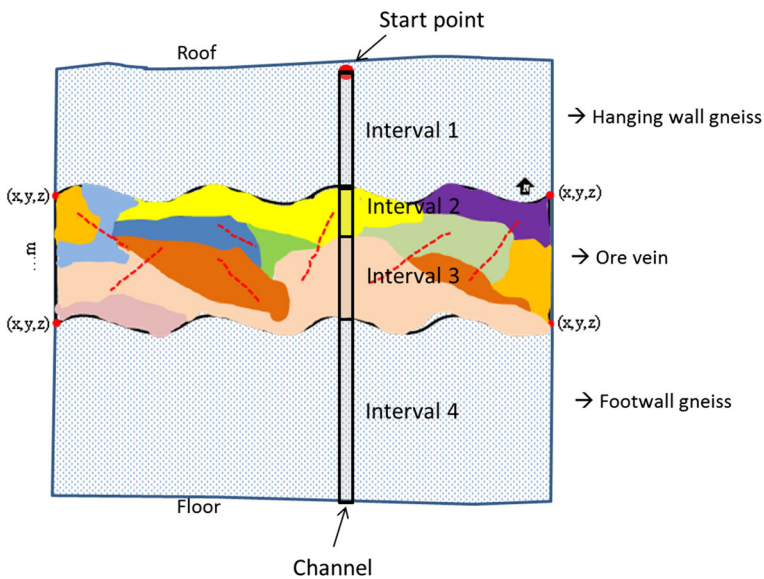
The deposit is characterized by polymetallic vein type mineralization formed by two hydrothermal mineralization events of Late-Variscan and Post-Variscan age (Seifert 2008). The Late-Variscan mineralization event dominates in the central part of the mine and mineralization is rich in sulphur, iron, lead, zinc and copper. Typical ore minerals include pyrite, galena, arsenopyrite, chalcopyrite and sphalerite, along with quartz and minor carbonate gangue. Ore minerals with a smaller Cu, Zn and Fe content characterize the Post-Variscan mineralization event. This mineralization event consists of a fluorite-barite-lead ore assemblage, mainly containing sphalerite, pyrite, galena, chalcopyrite and marcasite, as well as quartz, fluorite, carbonates and barite, as gangue (Benkert et al. 2015; Seifert 2008). In the Freiberg district, the polymetallic sulphide veins of the base metal deposits are hosted by ortho gneiss. For the purpose of the present study, ore refers to the polymetallic sulphide deposits, including sphalerite, galena and chalcopyrite. These minerals are the main sources of Zn, Pb and Cu, which are of primary economic interest. Arsenic is a penalty element in mineral processing and its presence in dust is a health concern; therefore, it is of interest as well. In the context of the current investigation, waste refers to the gangue materials, including the carbonates, fluorite, quartz and the host rock (gneiss).

### 2.2 Sample Collection

A mine face of approximately 22 m long and 2 m high was defined to test the research concept at the Reiche Zeche underground mine, as shown in Fig. 1. The defined mine face is characterized by high material variability and located at the first level of Wilhelm Stehender North, at a depth of about 150 m. The northern part of the mine face has an ore vein of thickness ranging from 30 to 100 cm, the vein consists of galena, pyrite, chalcopyrite and sphalerite minerals hosted by gneiss rock. At the central part of the mine face, ore is disseminated throughout the gneiss rock, whereas weathering of the host rock and the ore materials is observed at different locations. Moreover, small circular pores filled with calcite and ore can also be noted. Channel sampling is advantageous for capturing different lithotypes and variations in their abundance and distribution. Thus, as a part of the present investigation, channel samples were collected to address the observed spatial variability and ensure sample representativeness. For



**Fig. 1** A sketch with sample photographs that illustrates the defined mine face having ~22 m lateral extent and ~2 m height



**Fig. 2** Sketch that illustrates a channel cut at the defined mine face. The channel cross-cuts four different intervals, each belonging to a different lithotype and sampled separately

this purpose, 23 channels spaced approximately 80–120 cm apart (depending on the material variability at each channel location) were cut, and 102 samples were acquired from different intervals within each channel, as shown in Fig. 2.



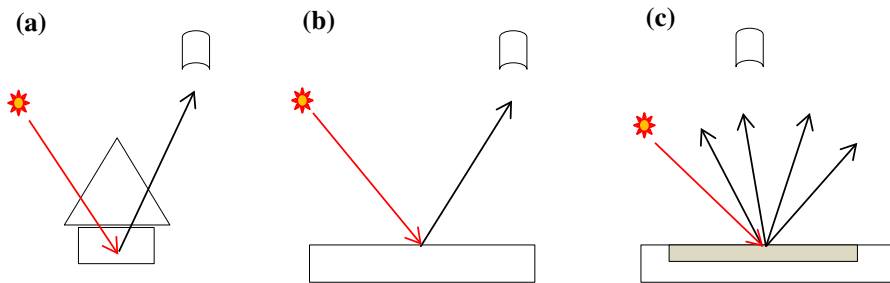
**Fig. 3** Handheld FTIR 4300 spectrometer and the three sampling interfaces

### 2.3 Data Acquisition and Instrumentation

The FTIR 4300 analyzer used in this study has three interchangeable sampling interfaces, namely external reflectance, attenuated total reflectance (ATR) and diffuse reflectance (Agilent 2017). It provides point data at a high data acquisition speed (less than 30 s). FTIR obtains full-wavelength spectra over a wide range of the electromagnetic spectrum (1.9–15  $\mu\text{m}$ ). Thus, it has a great potential for identification of various materials. The instrument is depicted in Fig. 3 and is a portable handheld device, powered by two 100/120/240 V batteries. Its compact dimensions and relatively light weight (under 2.2 kg) ensure its effective use in a wide range of in situ applications in real time. However, due to the harsh environmental conditions in the mine that served as the study site, in situ underground measurements were not attempted. Instead, samples were collected and analyzed in a laboratory.

The FTIR measurements were optimized by considering different instrument setups. This was done by interchanging the three sampling interfaces, varying the number of sample scans, modifying instrument calibration time and adjusting resolution. The external reflectance interface allows a mirror-like reflection (specular reflection) from the sample surface to be captured and is thus typically used for smooth surfaces. The ATR measures the internal reflection of the sample as the IR radiation beam passes through an ATR element in contact with the sample. Finally, the diffuse reflectance interface allows internal and external reflection to be measured simultaneously and it is usually applicable for rough surfaces. The working principles behind all three setups are shown in Fig. 4.

The performance of each of the three sampling interfaces was assessed using homogenized powdered samples to remove artefacts due to surface texture and compositional intergrowth. To obtain optimal sample scans, the influence of changing the number of sample scans on the measurement results was evaluated. The instrument was calibrated at different time intervals and the measurement results were compared. To evaluate the significance of the spectral differences, FTIR measurements were collected at 4  $\text{cm}^{-1}$ , 8  $\text{cm}^{-1}$  and 18  $\text{cm}^{-1}$  resolutions and the results were compared. The resultant optimized instrument setup comprised 64 sample scans, 126 background scans, 4  $\text{cm}^{-1}$  resolution, 15-min instrument calibration time and diffuse



**Fig. 4** The three FTIR spectrometer interfaces: **a** ATR, **b** specular reflectance, and **c** diffuse reflectance

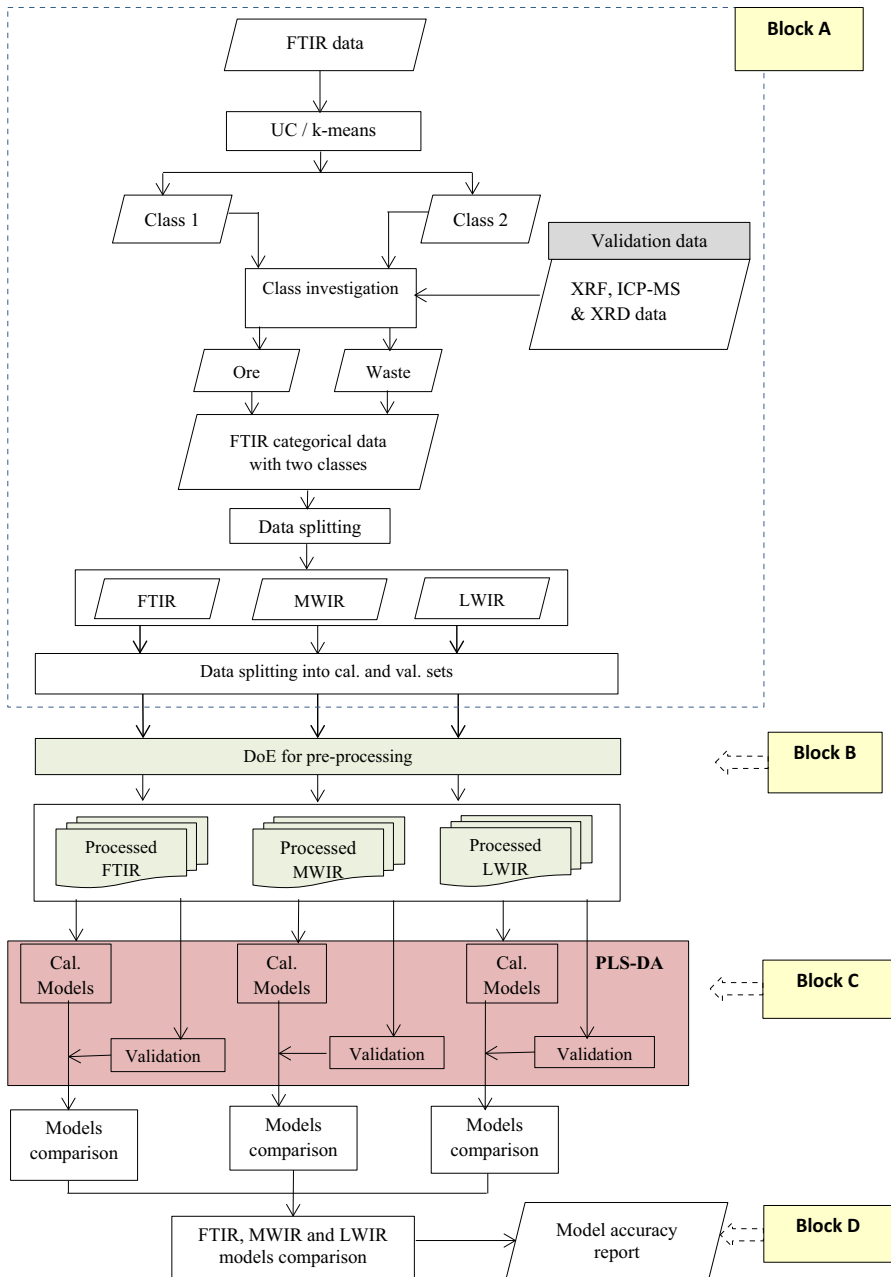
reflectance sampling interface. The FTIR spectroscopy data were collected over the  $\sim 1.9$  to  $\sim 15$   $\mu\text{m}$  wavelength range. However, for the test case materials, the spectral range from 1.9 to 2.5  $\mu\text{m}$  yielded noisy results and was excluded from further analyses. Three sub-datasets were prepared prior to modelling: the full FTIR data (excluding the range from 1.9 to 2.5  $\mu\text{m}$ ), the MWIR (2.5–7  $\mu\text{m}$ ) data and the LWIR (7–15  $\mu\text{m}$ ) data. To accommodate sample heterogeneity, multiple spectra were collected from each sample, the analysis results pertaining to 605 measurements collected using 102 samples are discussed in the sections that follow.

## 2.4 Chemical Validation Datasets

The conventional data acquisition techniques—namely X-ray diffraction (XRD), X-ray fluorescence (XRF) and inductively coupled plasma mass spectrometry (ICP-MS)—were used to obtain the data that were employed in the validation of the material discrimination results. The ICP-MS and XRF measurements were performed using 50 samples, while XRD measurements were carried out using 34 samples. The XRD data used for this study provide semi-quantitative mineralogical information, whereas the XRF and ICP-MS data provide quantitative elemental information.

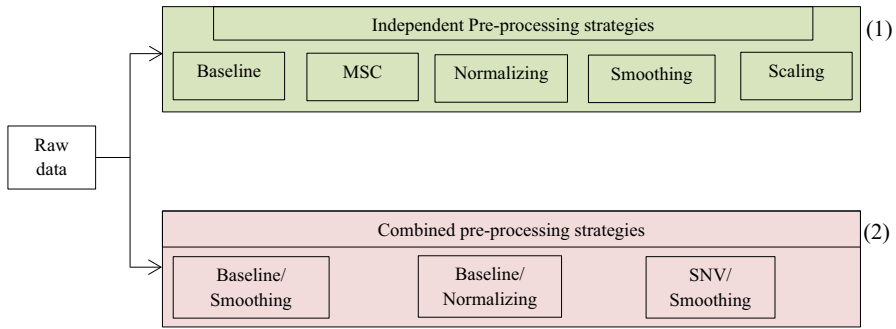
## 3 Methodology

As illustrated in Fig. 5, the material discrimination approach developed as a part of the present study is a multi-step process that incorporates data exploration, data pre-processing, data modelling and model validation. The data exploration task (denoted as block A in Fig. 5) includes pattern recognition, material identification and data splitting (e.g., into calibration and validation datasets). The design of experiment (DoE) was developed to find the optimal data pre-processing techniques (represented by block B in Figs. 5 and 6). Using the pre-processed data, a series of calibration and prediction models were developed (indicated by block C in Fig. 5). All prediction models were validated using independent datasets. The aforementioned approach was independently applied to three datasets, namely the full FTIR spectra, MWIR and LWIR. Therefore, the use of these three datasets for the discrimination of the test case



**Fig. 5** Overview of the research workflow. It includes four major steps: data exploration and preparation (block A), data pre-processing (block B), data modelling and model validation using independent datasets (block C) and model comparison (block D)





**Fig. 6** The independent (box 1) and combined (box 2) pre-processing methods that were applied to the spectral data of the full FTIR, MWIR and LWIR datasets

materials was evaluated (block D in Fig. 5). The details of each step are described below.

### 3.1 Data Exploration

#### 3.1.1 Unsupervised Classification

To identify a pattern and points of interest in the spectral data, descriptive statistics, cluster analysis and principal component analysis (PCA) were performed. Descriptive statistics, including box plots and histograms, were used to describe the basic data features. The unsupervised classification (UC) technique was applied to assess any natural patterns or groupings in the FTIR data. A highly efficient and the most commonly used UC methods is *k*-means. Thus, *k*-means with Euclidian distance was applied to examine any clustering in the spectral data. When applying this method, *n* observations are assigned to *k* clusters, using the centroid of the clusters and minimizing the sum of squared errors (Kaufman and Rousseeuw 2005) as shown below

$$J = \sum_{j=1}^k \sum_{i \in C_j} (x_i - m_j)^2,$$

where  $m_j = \sum_{i \in C_j} x_i / n_j$  denotes the cluster centroid of  $C_j$ ,  $n_j$  is the number of points in  $C_j$  and  $(x_1, \dots, x_n) = X$  represents the data matrix.

Using the full-range FTIR reflectance data and the *k*-means technique, the spectral data were classified into two classes. The UC was implemented with no a priori knowledge about potential mineral groupings. However, the number of clusters was specified in advance and two distinct classes were distinguished. The geochemical difference between the two classes was investigated using validation data (XRF, ICP-MS and XRD). The two classes exhibited variations in elemental concentrations of Cu, Zn, Pb, As and Fe. Thus, the class with a higher concentration of these elements was identified as ore, whereas the class with a lower concentration was considered as waste. In addition, the UC classification results were compared with the hand specimen

classification into the ore and waste classes. The comparison results revealed a very good agreement, confirming that the UC method can be a practical alternative to a hand specimen description for ore–waste discrimination, as the former can be automated and the latter might be subjective. Subsequently, a category variable that indicates samples belonging to ore or waste classes was appended to the full FTIR data table. Once the category variable was added to the FTIR spectral data, three datasets (the full FTIR, MWIR and LWIR) were prepared.

### 3.1.2 PCA Models

PCA is a dimension-reduction tool that transforms multiple correlated variables into a number of latent variables (uncorrelated variables). It is an effective explanatory data analysis technique that can be applied to identify the important variables (i.e., those responsible for an observed class difference) and the correlated variables (variables that contribute in the same way). Most importantly, it can be used to detect subtle outliers. In the present study, PCA models were developed using the three aforementioned datasets separately. The potential for using each dataset for separation of the two classes was assessed and compared. The loading plots of the PCA models were interpreted to identify the important (informative) variables.

### 3.1.3 Outlier Detection and Data Splitting

Different outlier detection techniques, namely Hotelling's  $T^2$ , residual map, influence plot and visual inspection of unique measurements, were considered in this study. Hotelling's  $T^2$  is a useful outlier detection tool that describes the distance to the model center, as spanned by the Principal Components (PCs). It also provides a critical limit ( $p$  value) with different statistical confidence limits. For example, for the  $p$  value of 5%, the 95% confidence ellipse can be included in the score plots of the PCA models to reveal potential outliers (i.e., data points located outside the ellipse contour). Accordingly, in the present study, the  $p$  value of 5% was adopted, the observed possible outliers were labelled, and the influence plots were inspected. The influence plot shows the sample residuals'  $X$ -variance against Hotelling's  $T^2$  and Leverage statistics. The residual statistics describe the sample distance to the model, whereas the Hotelling's  $T^2$  and Leverage indicate how well the sample is described by the model. Influence plot is a useful tool for detecting influential samples and dangerous outliers. For example, samples with high residual variance and high leverage are deemed to contain the most dangerous outliers. In the present investigation, variable residual plots (a map of residuals) were examined and the possible residuals were identified. This map is useful for determining whether samples have high residuals on few or all variables, and thus helps in outlier detection.

The potential outliers identified using the Hotelling's  $T^2$ , influence plot and residual map were visually inspected and compared. Based on the integrated findings yielded by these inspections, fifteen measurements that are possible outliers were identified and were excluded from the datasets. Subsequently, each of the three datasets was split into calibration and validation sets, ensuring approximately equal representation of each class within the two datasets (block A of Fig. 5). To avoid introducing systematic

errors, the datasets were split randomly, whereby measurements from the same samples were assigned to either validation or calibration dataset, but not both. The calibration and validation sets included 466 and 124 measurements, respectively.

### 3.2 Data Pre-processing

Data pre-processing is an integral part of multivariate data analysis, irrespective of whether it is conducted for classification, exploration or prediction purposes (Engel et al. 2013). It is performed to remove undesired variations (e.g., instrumental artefacts) and enhance the signal or variation of interest. While several data pre-processing strategies are currently available, the choice of the optimal data pre-processing method for a particular application depends on the nature of the data and the ultimate goal of data analysis. Thus, the pre-processing technique is typically selected based on trial and error, the visual inspection of the spectra and quality parameters, such as Pearson's correlation coefficient (PCC), which measures linear correlation between two variables (Engel et al. 2013). As this is largely a trial and error process, DoE is required to select the optimal independent and combined data pre-processing methods. This approach is advantageous when the aim is to analyze and understand the main effect and the interaction effect of the pre-processing techniques.

As shown in Fig. 6, baseline correction, standard normal variate (SNV), multiplicative scatter correction (MSC), smoothing (such as Gaussian filter smoothing), normalization and data scaling were the pre-processing methods considered for the present study. The upper box of Fig. 6 (labelled 1) shows the independent pre-processing techniques and the lower box (labelled 2) shows the combined pre-processing techniques. The choice of these methods was based on the fact that they are the most common artefacts of infrared data (e.g., baseline shift). In addition, the most prominent data artefacts (e.g., baseline, scatter and noise) were identified from the line plots of the reflectance spectra. Mean centering (MC) is a data scaling technique that represents variation around a mean by subtracting the sample mean from each data value (Roussel et al. 2014). With the exception of raw data, mean centering was performed in combination with each independent and combined techniques.

Baseline correction is a signal correction method that subtracts the unwanted spectral background from the main signal information. SNV normalizes by deducting the spectrum mean value from each variable in the spectrum and dividing each resulting value by the spectrum standard deviation (Fearn et al. 2009; Roussel et al. 2014). It normalizes the spectrum data to itself and minimize the light scattering effect. Normalization removes undesired intensity variation caused by multiplicative effects. It divides each spectrum based on the estimation of its spectral intensity (Roussel et al. 2014). MSC is also used to reduce multiplicative scattering effects (Fearn et al. 2009). However, smoothing is based on averaging the neighborhood points in order to minimize random noise (Roussel et al. 2014). The selected pre-processing methods were employed to develop a DoE that incorporates both independent and combined pre-processing techniques. The DoE was applied to the three datasets individually and sets of pre-processed data were generated.

### 3.3 Data Modelling and Validation

PLS-DA is a supervised classification method, used to optimize separation between different classes or groups. Once the classification model is developed, it can be used to assign unknown samples to the most probable class. In the present study, to establish a discrimination rule of the two classes (ore and waste materials), PLS-DA classification models were developed using the pre-processed data of the three datasets. PLS-DA was implemented in two steps, whereby PLS regression was followed by prediction. In PLS regression, the categorical data (in this case, the ore and waste classes) as treated as a response variable and the spectral data at each wavelength is represented as the independent variables. As shown in block C of Fig. 5, the pre-processed calibration data were used to develop a series of calibration models that were applicable to each of the two classes (ore and waste). The prediction model parameters were estimated using the calibration datasets and were subsequently validated using the independent (validation) datasets.

The performance of the calibration and prediction models when applied to the same dataset (e.g., MWIR data) after incorporating each of the previously described data filtering techniques was evaluated. The optimal data pre-processing methods were selected based on the calibration, validation and prediction statistics of the classification and prediction models, whereby lower error terms and higher predicted  $R^2$  value were deemed advantageous. In addition, confusion matrices were generated to assess the correct classification rates of individual models. The correct classification rates were computed by adding the number of true positives and true negatives and dividing their sum by the total number of the predicted samples. This approach was applied to the full FTIR, MWIR and LWIR datasets, and the material discrimination competencies of the three datasets were also assessed and compared.

## 4 Results and Discussion

### 4.1 Explanatory Data Analysis

#### 4.1.1 *k*-Means

Application of the *k*-means UC method resulted into two separate classes indicating a natural pattern or grouping in the reflectance spectral data. The observed natural groupings are based on the correlation or similarity among the measured variables (spectral data). The obtained results were reproducible for the same clustering method. To elucidate any patterns in the spectral data, ascertaining the composition of the two classes (mineral groupings) is essential. Therefore, the geochemical compositional differences between the two classes were investigated using the validation data (obtained through XRF, ICP-MS and XRD measurements). As summarized in Table 1, the two classes differ in elemental concentration of Cu, Zn, Pb and Fe, all of which are of interest for the present study due to their economic value. In addition, they also differed with respect to As, which is also of interest, since it is a penalty element in mineral processing. Thus, its early identification would be highly beneficial for optimizing product

**Table 1** Summary of the XRF and ICP-MS data showing concentrations of the indicated elements above and below the specified threshold value of the two classes

Class	Cu		Pb		Zn		As		Fe	
	Above 250 ppm	Below 250 ppm	Above 10,000 ppm	Below 10,000 ppm	Above 10,000 ppm	Below 10,000 ppm	Above 1000 ppm	Below 1000 ppm	Above 60,000 ppm	Below 60,000 ppm
Class 1 (%)	94	6	94	6	97	3	89	11	89	11
Class 2 (%)	7	93	22	78	6	94	0	100	0	100
Minimum value (ppm)	274	28	10,082	245	12,215	159	1033	57	61,400	22,500
Maximum value (ppm)	90,142	235	208,000	7014	96,400	7474	22,049	792	313,900	59,600

The % shows the proportion of samples in a given class below or above the specified threshold

**Table 2** Elemental correlation coefficient computed using 50 samples

Element pairs	Correlation coefficient
Cu–Zn	0.55
Cu–Pb	0.43
Pb–Zn	0.48
Mg–As	–0.01
Ca–As	0.17
Pb–As	–0.06
Cu–As	0.11
Zn–As	0.69

quality and eliminating risks, yielding saleable mineable material. In the material sampled from the study area for the test case, the primary sources of these elements are the sulphide ore minerals (chalcopyrite, pyrite, arsenopyrite, galena and sphalerite). Therefore, a higher/lower concentration of the elements signifies that the material is an ore/waste.

The class threshold value of each element concentration was set based on the value sequence observed in the classes and a relatively large change in elemental concentration. For example, the Cu content of 94% in the samples classified as Class 1 indicates greater than 250 ppm, as shown in Table 1. On the other hand, the Cu content of 93% in Class 2 samples indicates less than 250 ppm. Similarly, the Fe concentration of 89% in the samples categorized in Class 1 signifies greater than 60,000 ppm (6%), otherwise they are categorized as Class 2. Even though Fe has many mineral sources (e.g., sulphides, oxides and silicates), for the test case materials, Fe shows moderate to strong correlation with sulphide minerals, indicating that the main Fe sources are likely the sulphide minerals. Likewise, the Zn, Pb and As concentrations are higher in Class 1 than in Class 2 (Table 1). Therefore, based on the elemental concentration difference, Class 1 is denoted as ore, while Class 2 represents waste. In addition, the ICP-MS and XRF measurement data were used to compute the correlation coefficients of the elemental dependencies. As shown in Table 2, there is a moderate positive relationship between Cu and Zn, Cu and Pb, and Pb and Zn. Moreover, while Zn–As correlation is strong, no Pb–As correlation was observed in the data. The acquired correlation coefficients (except for a few exceptions) indicate that the elements co-occur. Therefore, a higher concentration of target elements is observed in one class than in the other.

Elemental concentration varied across samples, as shown in Table 1, where the minimum and maximum content of the indicated elements is given for both classes. The qualitative XRD measurement results show that most of the minerals assigned to the two classes are the same. This is likely because, even though their concentration varies, these minerals occur in both classes. This assertion is also supported by the XRD measurement semi-quantification results, indicating that a higher concentration of the sulphide minerals (ore minerals) was observed in Class 1 than in Class 2. It is highly likely that the acquired reflectance spectra are mixed spectra that are influenced by combined mineral signals or matrix effects. This is one of the possible reasons for not assigning 100% of the samples into one of the classes. However, with the indicated

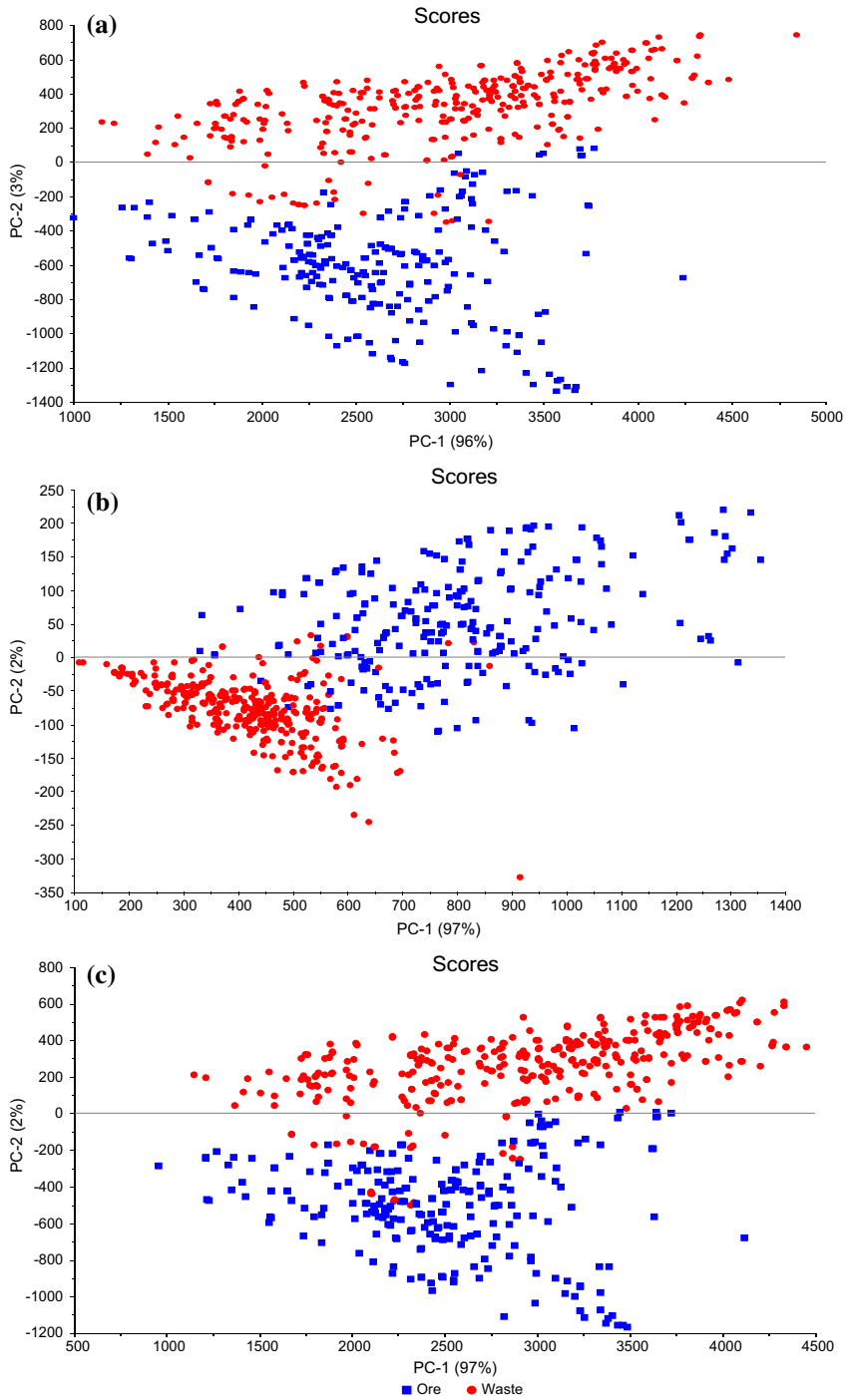
level of confidence (the % of samples in a class), samples containing the elements in quantities that are above or below the indicated value can be categorized into the two classes. Separation of the two classes based on the elemental concentration signifies presence of a signal in the spectral data that can be linked to mineralogy and hence to economic value.

The cut-off grade of a commodity is a variable that changes due to fluctuations in metal prices and mining costs. Compared to the typical current mining cut-off grades, the materials from the test case contain lower concentrations of the target elements. For example, the typical average cut-off grade for Cu in underground mining operations is above 5000 ppm (0.5%) (Calvo et al. 2016; Lundin mining 2018). However Cu content exceeded 5000 ppm (0.5%) in only two of the measured samples. Similarly, while the average cut-off grade for Zn in underground mining operations is above 5.5% (Canadian Zinc Corporation 2018; Lundin mining 2018), in 85% of the measured samples Zn content was below this value. Therefore, for the test case materials, the cut-off grade based on the current underground mining operations cannot directly be used to generically classify the ore and the waste materials. However, even when using sampled material containing elements of economic interest in low concentrations, discrimination of materials into two mineral groupings (ore and waste) was still achieved. This experimental classification result is consistent with the manual specimen classification into ore and waste. However, to avoid subjective sample classification and automate the process, application of the *k*-means technique was considered. Therefore, for the present study, the Cu, Zn, Pb, Fe and As concentration in the class specified as ore (waste) is above (below) the threshold value indicated in Table 1.

#### 4.1.2 PCA Models

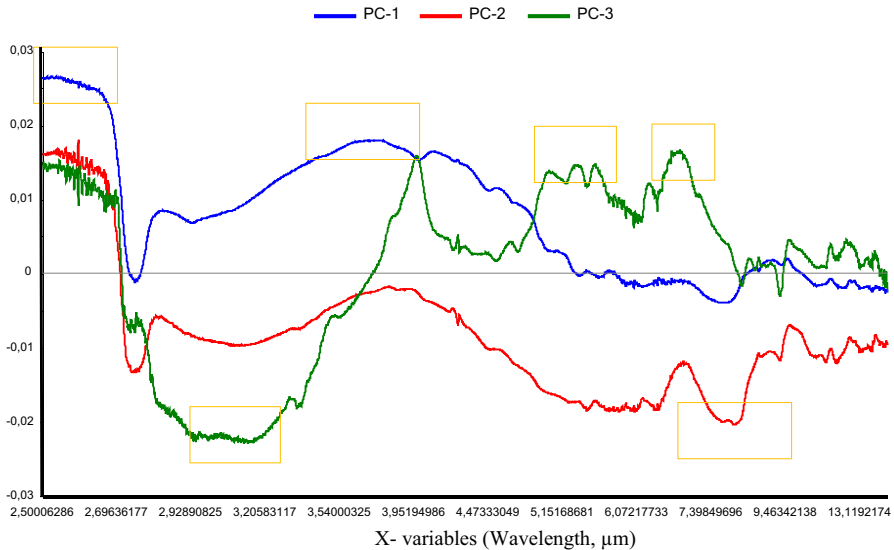
Figure 7 shows the PCA model score plots of the full FTIR, LWIR and MWIR datasets. A score plot provides valuable visual information on potential patterns in the samples. It depicts the relationship between sample differences or similarities and the data structure. As shown in Fig. 7a–c, when the models are applied to the full FTIR dataset, the two classes are better distinguished than when the individual LWIR and MWIR datasets are used, most likely because the full FTIR incorporates more informative variables that can accommodate variations in the data. The PCA models were applied to transform the spectral data of the three datasets into latent values (PCs). The loading plots of the PCs were subsequently interpreted to identify the important or informative variables.

Figure 8 shows the loading plot of the full-range FTIR, revealing that large loading coefficients or most variations are observed at different wavelength locations denoted by orange rectangles, such as at 2.5–2.6  $\mu\text{m}$ , 3.3–3.8  $\mu\text{m}$ , 7  $\mu\text{m}$ , 7.5–7.9  $\mu\text{m}$ , 8.8–9  $\mu\text{m}$ , 10.0  $\mu\text{m}$  and 10.7  $\mu\text{m}$  (note that, for the purpose of clarity, not all informative variables are indicated on the plots). As such variations are highly informative, they indicate that these variables are responsible for the separation of the two classes. When the wavelength locations of the informative variables are compared with the data contained in spectral libraries (NASA 2017), it is evident that most of the sulphide minerals show higher reflection at the identified wavelength locations. Thus, it is likely that the class separation is based on reflection signals from the sulphide minerals. For



**Fig. 7** PCA models computed using raw. **a** Full FTIR data, **b** LWIR data, and **c** MWIR data. The three datasets can be separated into two classes





**Fig. 8** The loading plot of the first three PCs of the full FTIR dataset

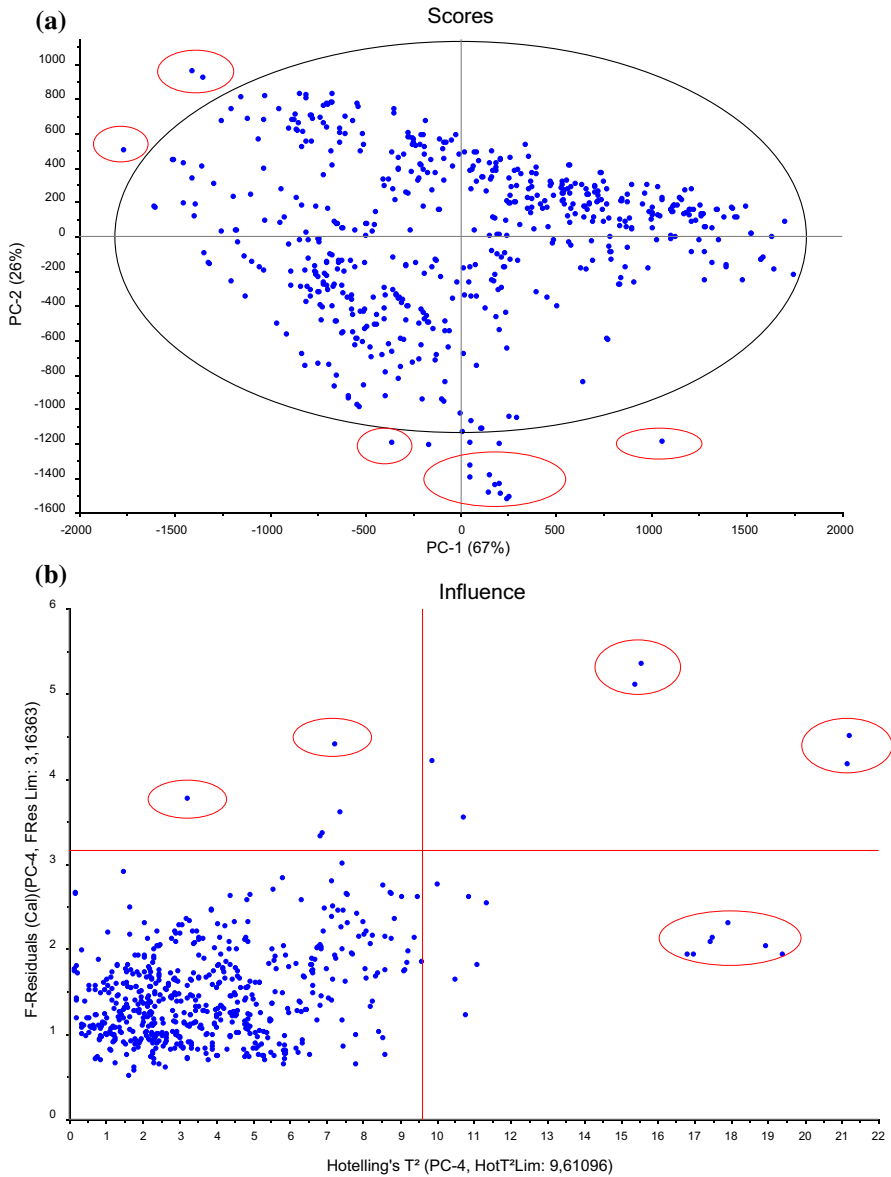
example, relatively higher reflection of galena is around  $3.5 \mu\text{m}$  (NASA 2017). These  $3.3\text{--}3.8 \mu\text{m}$  wavelengths are identified as important variables for class separation.

#### 4.1.3 Detection of Outliers

Outliers can be unique sample measurement results or noise, or might arise due to measurement errors. As mentioned in Sect. 3.1, in the present study, outliers were detected using an integrated inspection of Hotelling's  $T^2$ , residual map, influence plot and visual inspection. Figure 9 shows the possible outliers that were excluded using Hotelling's  $T^2$  and influence plots. Measurements located outside the Hotelling's  $T^2$  ellipse are potential outliers. The top right quadrant of the influence plot shows samples with higher leverage and higher residual, which are denoted as dangerous samples (as they are most likely outliers, as previously discussed). Samples in the lower right quadrant of the influence plot are influential, whereas those in the top left quadrant are poorly described by the model. Therefore, all samples in these two quadrants were carefully assessed using the Hotelling's  $T^2$  and residual map to identify those that are potential outliers (i.e., samples that are poorly described by the developed models). Therefore, to ensure proper variable description, fifteen measurements that are possible outliers were identified and excluded from the datasets.

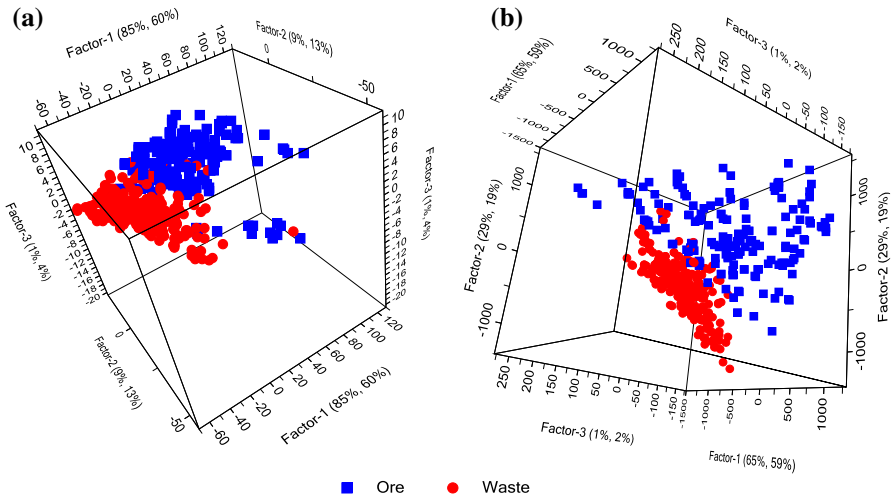
## 4.2 Data Pre-processing and Modelling

Figure 10a, b show the score plots of the PLS models of the LWIR and the full FTIR dataset, respectively, after application of the previously described data filtering techniques. The first two latent variables (Factor 1 and Factor 2) of the PLS regression



**Fig. 9** **a** A score plot with Hotelling's  $T^2$  limit ( $p$  value of 5%). **b** Influence plot with Hotelling's  $T^2$  on the  $x$ -axis and  $F$ -residuals on the  $y$ -axis. Samples within red circles are potential outliers that were excluded from the datasets

explain most of the variations in the data. For example, 94% of the variation in the LWIR data is explained by the first two factors. Moreover, the first two factors explain 94% and 73% of the variation in the spectra and the class category, respectively. This finding indicates presence of an unstructured variation in the spectral data that is not



**Fig. 10** Score plots of PLS models after data smoothing (Gaussian filter) is applied to **a** the LWIR and **b** the full FTIR dataset

related to the class category. The difference in the variation is likely due to other mineralogical information that has not been considered in the classification process. For example, sub-clustering was observed within the ore class, which could potentially be attributed to the different mineral groups that occur within the ore. Therefore, there is a high potential for further discrimination of the materials into additional classes.

Tables 3 and 4 summarize the calibration and prediction statistics of the three datasets for ore and waste prediction, respectively. Table 5 shows the correct classification rates of models applied to the three datasets, after each dataset has been subjected to individual and combined data pre-processing techniques. Therefore, the tabulated data also indicate model performance after each data pre-processing technique is applied to the data. Lower statistical error terms (RMSEcal, RMSEval and RMSEP), higher predicted  $R^2$  and higher correct classification rates signify an improved classification and predictive performance. It is evident that an enhanced prediction was obtained by applying the data pre-processing techniques to all three datasets. For example, for the full FTIR raw data, RMSEcal = 0.21, RMSEval = 0.22, RMSEP = 0.22, predicted  $R^2$  = 0.81 and a correct classification rate of 83.1% were obtained. However the error terms declined (RMSEcal = 0.18, RMSEval = 0.19 and RMSEP = 0.20), the  $R^2$  value improved (0.84) and the correct classification rate increased (89.5%) after the FTIR dataset was treated with the SNV data filtering technique. Conversely, not all data filtering techniques necessarily improved model performance. For example, the MSC filtering technique did not improve the model performance when applied to the MWIR data, as shown in Tables 3 and 4, most likely because the multiplicative effect is not pronounced in the spectral data.

For the given datasets, combining the pre-processing techniques did not result in a better prediction than that obtained when these techniques were applied individually. Moreover, the pre-processing technique that was most optimal differed for the three

**Table 3** The PLS-DA model ore class calibration and prediction statistics summary

Filtering techniques	RMSECal			RMSEVal			RMSEP			Predicted $R^2$			
	FTIR	MWIR	LWIR	FTIR	MWIR	LWIR	FTIR	MWIR	LWIR	FTIR	MWIR	LWIR	
Raw	0.21	0.23	0.23	0.22	0.24	0.24	0.22	0.22	0.22	0.23	0.81	0.80	0.78
MC	0.20	0.22	0.23	0.20	0.23	0.24	0.19	0.22	0.22	0.22	0.85	0.81	0.80
SNV	0.18	0.20	0.22	0.19	0.21	0.23	0.20	0.24	0.24	0.21	0.84	0.77	0.82
Baseline	0.22	0.22	0.24	0.22	0.22	0.24	0.23	0.19	0.23	0.23	0.79	0.85	0.79
MSC	0.20	0.24	0.22	0.21	0.26	0.23	0.25	0.28	0.23	0.23	0.76	0.69	0.79
Gaussian	0.19	0.23	0.23	0.20	0.24	0.23	0.19	0.22	0.23	0.23	0.85	0.80	0.79
Normalized	0.19	0.21	0.21	0.20	0.22	0.22	0.22	0.23	0.22	0.22	0.81	0.79	0.80
Combination of data pre-processing techniques													
Gaussian—baseline	0.22	0.22	0.24	0.22	0.22	0.24	0.23	0.19	0.23	0.23	0.79	0.85	0.80
Normalized—base-line	0.19	0.22	0.22	0.20	0.22	0.23	0.21	0.24	0.25	0.25	0.83	0.77	0.75
Gaussian—SNV	0.18	0.21	0.22	0.19	0.22	0.23	0.20	0.22	0.21	0.21	0.84	0.80	0.82

**Table 4** The PLS-DA model waste class calibration and prediction statistics summary

Filtering techniques	RMSECal			RMSEVal			RMSEP			Predicted $R^2$		
	FTIR	MWIR	LWIR	FTIR	MWIR	LWIR	FTIR	MWIR	LWIR	FTIR	MWIR	LWIR
Raw	0.21	0.23	0.23	0.22	0.24	0.24	0.22	0.22	0.22	0.81	0.80	0.78
MC	0.19	0.23	0.23	0.20	0.23	0.23	0.19	0.22	0.23	0.85	0.80	0.79
SNV	0.18	0.21	0.22	0.19	0.21	0.23	0.20	0.24	0.21	0.84	0.77	0.82
Baseline	0.22	0.22	0.24	0.22	0.22	0.24	0.23	0.19	0.23	0.79	0.85	0.80
MSC	0.20	0.24	0.22	0.21	0.26	0.23	0.25	0.28	0.23	0.76	0.69	0.79
Gaussian	0.19	0.23	0.23	0.20	0.24	0.23	0.19	0.22	0.23	0.85	0.80	0.79
Normalized	0.19	0.21	0.21	0.20	0.22	0.22	0.22	0.23	0.22	0.81	0.79	0.80
Combination of data pre-processing techniques												
Gaussian—baseline	0.22	0.22	0.24	0.22	0.23	0.24	0.23	0.19	0.23	0.79	0.85	0.80
Normalized—base-line	0.19	0.22	0.22	0.20	0.22	0.23	0.22	0.31	0.25	0.80	0.59	0.75
Gaussian—SNV	0.18	0.21	0.22	0.19	0.21	0.23	0.20	0.24	0.21	0.84	0.77	0.82

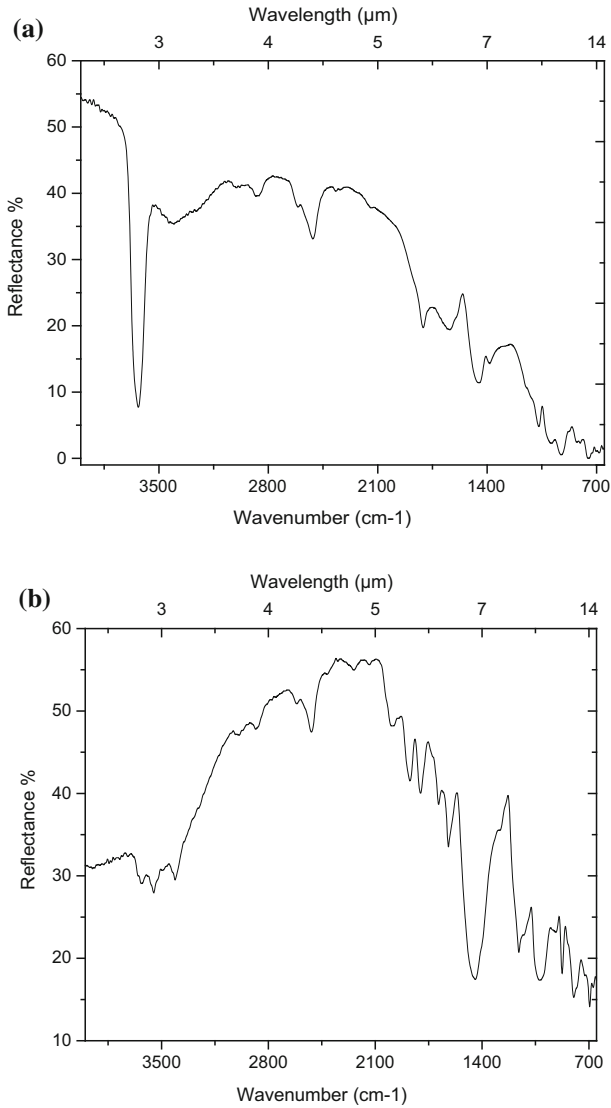
**Table 5** Summary of the correct classification (ore–waste discrimination) rates when different models are applied to the full FTIR, MWIR and LWIR datasets

Model	FTIR (%)	MWIR (%)	LWIR (%)
Raw	83.1	75.8	79.8
Normalized	87.1	85.5	84.7
Baseline	83.9	86.3	79.8
SNV	89.5	85.5	84.7
Gaussian	86.3	79.0	80.6

datasets. For example, when applied to the MWIR data, model performance improved (predicted  $R^2 = 0.85$  and 86.3% correct classification rate) after baseline correction (Tables 3, 5). On the other hand, the best results were achieved when the LWIR data were subjected to SNV ( $R^2 = 0.82$  and 84.7% correct classification rate). This finding implies that, in the LWIR dataset, the undesired intensity variation was more pronounced than in the MWIR data. These results also concur with the empirical evidence indicating that the choice of most optimal pre-processing technique is data dependent and requires a trial and error approach (Engel et al. 2013). Therefore, DoE is a crucial step in determining the most optimal data filtering technique.

Comparing the three datasets, the full FTIR data resulted in better prediction performance than the MWIR and LWIR data alone (Tables 3, 4, 5), likely because the former covers a wider wavelength region. Thus, the full FTIR dataset includes more informative variables to adequately explain the variation in the reflectance spectra than the individual MWIR or LWIR data. Overall, the three datasets showed a good potential for discrimination of the test case materials into ore and waste. Moreover, after baseline correction, SNV and normalization, the correct classification rates were higher when the MWIR, rather than the LWIR dataset, was utilized (Table 5). This is an interesting finding, since MWIR is the least explored IR region in lithological material characterization. Overall, the maximum correct classification rates achieved for the full FTIR, MWIR and LWIR datasets after data filtering were 89.5%, 86.3% and 84.7%, respectively. Owing to the limited information level in the IR spectra of the sulphide minerals and the intermediate values that obscure clear class boundaries (and thus bias model performance), the obtained accuracies are sufficient to discriminate the materials into ore and waste.

Figure 11 shows representative spectra of ore and waste material samples. Using spectral libraries (NASA 2017) and evidence reported by other authors (Ji et al. 2009; Schodlok et al. 2016), different minerals—including dolomite, muscovite, quartz, calcite and gypsum—were identified. The sulphide minerals exhibit very weak spectral features in the IR reflectance data that a direct interpretation of target minerals spectra is challenging. However, they have a higher and lower reflection points at different wavelengths. Hence, this sulphide minerals property can be transformed into valuable information using the approach proposed in this study.



**Fig. 11** Representative spectra of the two classes denoting **a** waste material and **b** ore material

### 4.3 Potential for Application for In-Situ and Real-Time Material Characterization

In situ real-time material characterization requires both appropriate instrument design and software development. The current technological advancements have enabled design and implementation of portable instrumentation, making in situ spectroscopy possible. However, for underground mining applications, care must be taken to avoid or minimize the effect of moisture and dust. With respect to software design, the

findings reported in this work demonstrated the potential of using a FTIR spectrometer for distinguishing ore from waste, which is of particular relevance for marginally economical mining operations. A prediction model was developed to classify unknown spectra into ore and waste classes. The acquired results are promising and can be improved by the inclusion of additional samples into the calibration models. Thus, a test case-specific mineral library can be developed for automated online discrimination of ore–waste materials.

The present study was carried out using powdered homogenized samples. However, the approach can be extended for whole rock applications by developing customized sampling strategies to account for inherent material variability and heterogeneity. The ongoing depletion of mineral resources, accompanied by increasing societal demand, suggests that it is likely that increasingly lower-grade ores will be extracted in future mining operations. The work presented here has demonstrated the potential of utilizing the MWIR and LWIR data for ore–waste discrimination, which could assist in greater selectivity during extraction and pre-processing, thus maximizing use of the resources while increasing economic viability.

## 5 Conclusions

Infrared spectra collected using an FTIR spectrometer were analyzed and classified using chemometric analytical methods. The utility of the obtained results for the characterization of sulphide ore and waste minerals from the selected test site was investigated. Three datasets spanning different wavelength ranges were prepared, namely the full FTIR spectra, as well as MWIR and LWIR spectra. Without a priori knowledge of the material types, the well-known *k*-means method was implemented to separate the datasets into two classes, whereby two distinct classes were identified. The mineralogical composition of the two classes was investigated using the conventional XRF, ICP-MS and XRD measurement techniques. The two classes exhibited differences in the elemental concentrations of Cu, Pb, Zn, As and Fe, and were thus defined as ore and waste. The identified categorical variables (the two classes) were inserted into the spectral data of the three datasets.

DoE was implemented to identify the optimal independent and combined data-filtering techniques for discriminating the two classes using the three aforementioned datasets. The processed data were used to make predictions about the composition of unknown samples. A series of prediction models were developed using the processed data combined with PLS-DA. Model performance was evaluated using the calibration, validation and prediction statistics in the form of an estimated prediction error. In addition, the correct classification rate was calculated for each model. The same procedure was applied for the three (FTIR, MWIR and LWIR) datasets. In general, the results showed a good agreement in model performance when applied on the three datasets. However, when the full-wavelength FTIR dataset was employed, lower prediction errors and higher correct classification rates were obtained compared to those pertaining to the MWIR or LWIR data.

Even though not all data pre-processing techniques necessary improved model performance, baseline correction, normalization and smoothing improved the clas-



sification and prediction performance of the developed models. For example, when the models were applied to the full-range FTIR dataset, 89.5% correct classification rate was achieved after subjecting the data to the SNV technique. When models were applied to the MWIR dataset, the prediction improved to 86.3% after baseline correction. Finally, after normalization of the LWIR data, an enhanced correct classification rate of 84.7% was obtained. The MWIR data alone provides sufficient information to successfully classify the samples into ore and waste. Thus, it can be posited that this region offers great potential for rock and mineral characterization. In this work, FTIR spectroscopy was successfully used to discriminate the ore and waste materials of the test case. For future research in this field, FTIR should be combined with PLS-DA to explore the potential for rapid automated online discrimination of ore and waste material. In sum, with accurate model calibration, the material discrimination process can be automated.

**Acknowledgements** This research (Real-Time Mining Project) has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Grant Agreement No. 641989. We would like to thank our colleagues at TU Bergakademie Freiberg for their support during sample collection field campaigns. The comments from the two anonymous reviewers are highly appreciated, as addressing them has significantly improved the quality of this manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Agilent (2017) FTIR compact & portable systems. <http://www.agilent.com/en/products/ftir/ftir-compact-portable-systems/4300-handheld-ftir>. Accessed 2 Feb 2017
- Alov NV, Volkov AI, Usheroev AI, Ishmetsev EN, Usheroeva EV (2010) Continuous X-ray fluorescence analysis of iron ore mixtures in the production of agglomerate. *J Anal Chem* 65:169–173. <https://doi.org/10.1134/s1061934810020127>
- Benkert T, Dietze A, Gabriel P, Gietzel J, Gorz I, Grund K, Lehmann H, Lowe G, Mischo H, Schaeben H, Schreiter F, Stanek K (2015) First step towards a virtual mine—generation of a 3D model of Reiche Zeche in Freiberg. In: Proceedings of the 17th annual conference of the international association for mathematical geosciences (IAMG), vol I, II. Freiberg, Germany, Sept 2015, pp 1350–1356
- Birkner N, Wang Q (2015) How an FTIR spectrometer operates. Chemistry libretexts web. [https://chem.libretexts.org/Core/Physical\\_and\\_Theoretical\\_Chemistry/Spectroscopy/Vibrational\\_Spectroscopy/Infrared\\_Spectroscopy/How\\_an\\_FTIR\\_Spectrometer\\_Operates](https://chem.libretexts.org/Core/Physical_and_Theoretical_Chemistry/Spectroscopy/Vibrational_Spectroscopy/Infrared_Spectroscopy/How_an_FTIR_Spectrometer_Operates). Accessed 5 July 2017
- Breteron RG (2007) Applied chemometrics for scientists. Wiley, Chichester
- Buxton MWN, Benndorf J (2013) The use of sensor derived data in optimization along the mine-value-chain: an overview and assessment of techno-economic significance. In: Proceedings of the 142nd SME annual meeting and exhibit, Denver, CO, USA, Feb 2013, pp 13–38
- Calvo G, Mudd G, Valero A, Valero A (2016) Decreasing ore grades in global metallic mining: a theoretical issue or a global reality? *Resour* 5:36. <https://doi.org/10.3390/resources5040036>
- Canadian Zinc Corporation (2018) Mineral resources and reserves. <http://www.canadianzinc.com/projects/prairie-creek/resources>. Accessed 20 March 2018
- Chukanov NV, Chervonnyi AD (2016) Infrared spectroscopy of minerals and related compounds. Springer, Basel
- Clark RN (1999) Spectroscopy of rocks and minerals and principles of spectroscopy. In: Rencz AN (ed) Manual of remote sensing remote sensing for the earth sciences, vol 3. Wiley, New York, pp 3–58

- Culka A, Kindlova H, Drahotka P, Jehlicka J (2016) Raman spectroscopic identification of arsenate minerals in situ at outcrops with handheld (532 nm, 785 nm) instruments. *Spectrochim Acta A Mol Biomol Spectrosc* 154:193–199. <https://doi.org/10.1016/j.saa.2015.10.025>
- Death DL, Cunningham AP, Pollard LJ (2008) Multi-element analysis of iron ore pellets by laser-induced breakdown spectroscopy and principal components regression. *Spectrochim Acta B At Spectrosc* 63:763–769. <https://doi.org/10.1016/j.sab.2008.04.014>
- Engel J, Gerretzen J, Szymanska E, Jansen JJ, Downey G, Blanchet L, Buydens LMC (2013) Breaking with trends in pre-processing? *Trends Analyt Chem* 50:96–106. <https://doi.org/10.1016/j.trac.2013.04.015>
- Fearn T, Riccioli C, Garrido-Varo A, Guerrero-Ginel JE (2009) On the geometry of SNV and MSC. *Chemom Intell Lab Syst* 96:22–26. <https://doi.org/10.1016/j.chemolab.2008.11.006>
- Fox N, Parbhakar-Fox A, Moltzen J, Feig S, Goemann K, Huntington J (2017) Applications of hyperspectral mineralogy for geoenvironmental characterisation. *Miner Eng* 107:63–77. <https://doi.org/10.1016/j.mineng.2016.11.008>
- Goetz AFH, Curtiss B, Shiley DA (2009) Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. *Miner Eng* 22:490–499. <https://doi.org/10.1016/j.mineng.2008.12.013>
- Griffiths PR, Haseth JA (2007) Fourier transform infrared spectrometry. *J Am Chem Soc* 129:13358–13358. <https://doi.org/10.1021/ja076968d>
- Ismail AA, van de Voort FR, Sedman J, Paré JRJ, Bélanger JMR (1997) Chapter 4 Fourier transform infrared spectroscopy: principles and applications. In: *Techniques and instrumentation in analytical chemistry*, vol 18, pp 93–139. [https://doi.org/10.1016/S0167-9244\(97\)80013-3](https://doi.org/10.1016/S0167-9244(97)80013-3)
- Ji J, Ge Y, Balsam W, Damuth J, Chen J (2009) Rapid identification of dolomite using a Fourier transform infrared spectrophotometer (FTIR): a fast method for identifying Heinrich events in IODP site U1308. *Mar Geol* 258:60–68. <https://doi.org/10.1016/j.margeo.2008.11.007>
- Karr C, Kovach JJ (1969) Far-infrared spectroscopy of minerals and inorganics. *Appl Spectrosc* 23:219–223. <https://doi.org/10.1366/000370269774380932>
- Kaufman L, Rousseeuw PJ (2005) In: Nj U (ed) *Finding groups in data: an introduction to cluster analysis*. Wiley, Hoboken
- Kruse FA (1996) Identification and mapping of minerals in drill core using hyperspectral image analysis of infrared reflectance spectra. *Int J Remote Sens* 17:1623–1632. <https://doi.org/10.1080/01431169608948728>
- Lundin Mining (2018) Reserves and resources. <https://www.lundinmining.com/operations/overview/>. Accessed 20 March 2018
- Miller JC, Miller JN (2000) *Statistics and chemometrics for analytical chemistry*, 4th edn. Prentice Hall, Harlow, New York
- NASA (2017) ASTER spectral library. <https://speclib.jpl.nasa.gov/>. Accessed 1 Sept 2017
- Orbit Technologies (2017) Online XRF analyzer. <http://www.orbitindia.com/online-xrf-analyzer/>. Accessed 27 Oct 2017
- Perkins WD (1987) Fourier transform-infrared spectroscopy: part II. Advantages of FT-IR. *J Chem Educ* 64:A269. <https://doi.org/10.1021/ed064pA269>
- Roussel S, Preys S, Chauchard F, Lallemand J (2014) *Multivariate data analysis (Chemometrics). Process analytical technology for the food industry*. Springer, New York, pp 7–59
- Scheinert M, Kupsch H, Bletz B (2009) Geochemical investigations of slags from the historical smelting in Freiberg, Erzgebirge (Germany). *Chem Erde* 69:81–90. <https://doi.org/10.1016/j.chemer.2008.03.001>
- Schodlok MC, Green A, Huntington J (2016) A reference library of thermal infrared mineral reflectance spectra for the HyLogger-3 drill core logging system. *Aust J Earth Sci* 63(94):1–949. <https://doi.org/10.1080/08120099.2016.1234508>
- Seifert T (2008) Post-collisional magmatism and its relationship to late-variscan ore forming processes in the Erzgebirge (Bohemian Massif) Germany. IOS Press BV, Amsterdam, pp 1–128
- Seifert T, Sandmann D (2006) Mineralogy and geochemistry of indium-bearing polymetallic vein-type deposits: implications for host minerals from the Freiberg district, Eastern Erzgebirge, Germany. *Ore Geol Rev* 28:1–31. <https://doi.org/10.1016/j.oregeorev.2005.04.005>
- Smith BC (2011) *Introduction to infrared spectroscopy. Fundamentals of fourier transform infrared spectroscopy*, 2nd edn. CRC Press, New York, pp 1–17
- Spectral Evolution (2015) UV-VIS-NIR spectrometers. Spectral evolution web. [http://www.spectralevolution.com/spectrometers\\_mining.html](http://www.spectralevolution.com/spectrometers_mining.html). Accessed 25 June 2015

- Stuart BH (2004) *Infrared spectroscopy: fundamentals and applications*. Wiley, Chichester
- Sun Y, Seccombe PK, Yang K (2001) Application of short-wave infrared spectroscopy to define alteration zones associated with the Elura zinc–lead–silver deposit, NSW, Australia. *J Geochem Explor* 73:11–26. [https://doi.org/10.1016/S0375-6742\(01\)00167-4](https://doi.org/10.1016/S0375-6742(01)00167-4)
- Szalai Z, Klaudia K, Jakab G, Sipos P, Belucz B, Németh T (2013) The use of UV-VIS-NIR reflectance spectroscopy to identify iron minerals. *Astron Nachr* 334:940–943. <https://doi.org/10.1002/asna.201211965>
- ThermoFisher (2017) X-ray fluorescence (XRF). <https://www.thermofisher.com>. Accessed 30 Aug 2017
- Wells MA, Ramanaidou ER (2015) Raman spectroscopic core scanning for iron ore and BIF characterization. In: *Proceedings of the 11th international congress for applied mineralogy (ICAM)*, Mianyang, China, July 2013, pp 387–396