GEOCHEMICAL AND HYPERSPECTRAL DATA FUSION FOR DRILL-CORE MINERAL MAPPING

Cecilia Contreras, Mahdi Khodadadzadeh, Laura Tusa, Christina Loidolt, Raimon Tolosana-Delgado, and Richard Gloaguen

Helmholtz-Zentrum Dresden-Rossendorf (HZDR),
Helmholtz Institute Freiberg for Resource Technology, Germany

ABSTRACT

Hyperspectral imaging is increasingly being used in the mining industry for the investigation of drill-core samples. It provides the means to analyze a large amount of cores considerably faster than traditional methods and in a non-invasive and non-destructive manner. Traditional approaches used to analyse drill-core hyperspectral data are mainly based on visual observations and need significant human interactions. Thus, they are time-consuming and subjective. In this paper, we explore the use of supervised machine learning techniques for mineral mapping in drill-core hyperspectral data. For this purpose, we suggest to use geochemical data for generating a training set. The main contribution of this work is to fuse geochemical and hyperspectral data within a machine learning framework. Moreover, for a more complete mineral mapping task, we integrate visible near-infrared (VNIR), short-wave infrared (SWIR) and long-wave infrared (LWIR) hyperspectral data. For the extraction of input features, the traditional Principal Component Analysis (PCA) is implemented. For classification, we propose to use Random Forest (RF) because of its significant performance in hyperspectral data classification when there are few training samples available. Experimental results show that the proposed method provides comprehensive mineral maps in which the distribution and patterns of different minerals are well characterised.

Index Terms— Data fusion, mineral mapping, hyperspectral data, geochemical data, machine learning.

1. INTRODUCTION

Hyperspectral scanning is an emerging technique in the mining industry, which offers a non-invasive and non-destructive approach to image a large amount of cores in a fast turnaround time [1]. Its rich spectral information can be effectively exploited for the characterization and determination of the spatial distribution of different minerals, so called mineral mapping [2]. Minerals have different spectral responses in specific regions of the electromagnetic spectrum. These variations are related to fundamental vibrational movements of different molecular bonds. As for example, bonds in OH, H_2O, CO_3, and between Al-OH, Mg-OH, and Fe-OH are usually the ones generating absorption features in the visible to near-infrared (VNIR) and short-wave infrared (SWIR) regions of the electromagnetic spectrum. These are mainly found in mineral phases such as phyllosilicates, hydroxylated silicates, sulphates, carbonates, amongst others, frequently having a secondary origin [3, 2]. Bonds in Si-O, Si-O-Si, Si-O-Al, CO_3, SO_4, amongst others, commonly related to rock-forming minerals, have specific responses in the long-wave infrared (LWIR) region [4]. Thus, a reasonable integration of hyperspectral data in different ranges can increase the mineral mapping accuracy.

Traditional methods to map minerals in drill-core hyperspectral data are based on visual interpretations and comparisons with the available reference spectral libraries (e.g., USGS Spectral library [5]). The use of machine learning techniques is suggested to improve the accuracy, speed, and robustness of data analysis. These techniques offer automatic means to discover underlying relationships within large data sets. While for geological applications several approaches have been tested [6], for the analysis of drill-core hyperspectral data the use of machine learning techniques, especifically supervised techniques, is at an early stage of development. This is mainly because training data is usually limited and the a priori information about the mineral classes existing in each drill-core is incomplete. In our previous work, we explored the use of Scanning Electron Microscopy Mineral Liberation Analysis (SEM-MLA), performed on small thin sections of drill-cores, for training a classifier [7]. Following a similar approach, in this paper, we explore the use of a laboratory technique to generate training data. We propose to exploit geochemical data obtained from few assay measurements along drill-cores to generate the training set. Such measurements comprise the abundance of different chemical elements (e.g.,major and trace elements) in the form of proportions (e.g., weight percents or parts per million (ppm)) subject to a constant sum [8]. These elements can then be used for a whole-rock composition analysis and therefore, linked with hyperspectral data based on the different chemical bonds. To the best of our knowledge, the use of mineral

chemistry and hyperspectral data has been limited in the literature to the comparison of the results obtained independently (e.g., [9, 10]). In this work, we propose to fuse the geochemical and hyperspectral data and build a link between these two sources of information in a machine learning framework. Another important contribution of this work is that we integrate VNIR-SWIR and LWIR hyperspectral data to perform a comprehensive mineral mapping. For the classification, we suggest to use the Random Forest (RF) algorithm since it can handle classification of high-dimensional data with a limited number of training samples [11].

The rest of the paper is structured as follows: section 2 contains the methodological framework. Section 3 presents data description, experimental results, and discussions. Finally, the conclusions are drawn in Section 4.

2. METHODOLOGICAL FRAMEWORK

The proposed machine learning-based framework for mapping minerals in drill-core hyperspectral data is shown in Fig. 1. The approach is divided into two main phases: the training and prediction phases. For the training phase, firstly, the geochemical data is transformed using the Centered Log-Ratio (CLR) transformation. This is a logarithmic transformation scaled by the geometric mean. The CLR transformation is needed before applying machine learning algorithms to transfer the geochemical data from the compositional space to a real space. The transformed geochemical data is then clustered. For clustering algorithms, finding the appropriate number of clusters is important. In this work, the traditional Elbow method is used, in which the within-cluster sum of square errors (WSS) metric is computed to select the appropriate number of clusters [12]. For clustering, the well-known K-means clustering algorithm is applied.

After obtaining the training labels from the geochemical data, a supervised classification system is used to learn the mapping between the geochemical labels and the spectral data. More especifically, the representative spectra corresponding to the location of the geochemical measurements together with the generated labels are given to a supervised classifier for the training phase. After that, in the prediction phase, the trained model is applied on the unknown spectral data to assign unique labels. Moreover, to perform a more comprehensive mineral mapping task, VNIR-SWIR and LWIR hyperspectral data are fused.

3. EXPERIMENTAL RESULTS

To show the performance of the suggested method, we used a tray of drill-cores from the Rajapalot gold-cobalt deposit in Northern Finland. Geochemical data from 168 assay measurements acquired along a drill-core of about 328 m depth were available. This geochemical data consist of ppm concentrations of 25 different elements. These elements are As,

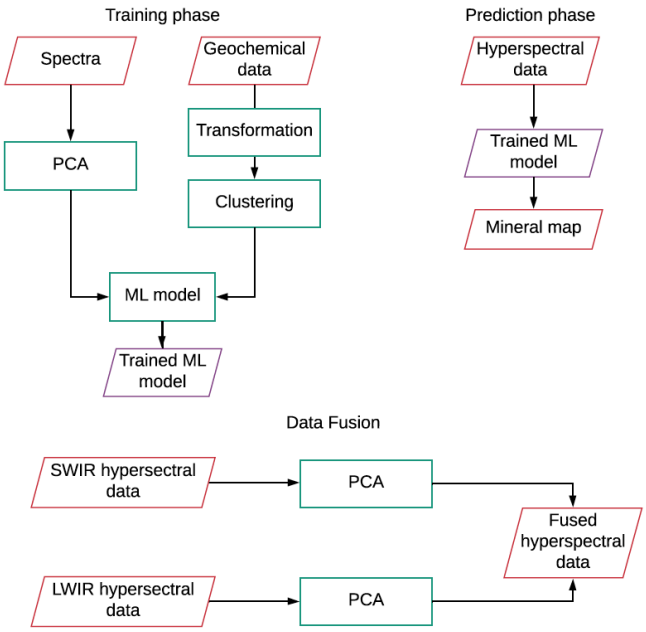


Fig. 1: Flowchart of the proposed machine learning (ML) technique to fuse geochemical data with hyperspectral data for drill-core mineral mapping.

Au, Bi, Co, Cr, Cu, Ga, Hf, Mn, Ni, Rb, S, Sc, Sr, Th, U, V, Zr, Al, Ca, Fe, K, Mg, Na, and Ti. For the clustering of the transformed geochemical data the traditional K-means algorithm was used. A total of *5 clusters* were determined with the Elbow method.

Regarding the hyperspectral data, we used VNIR-SWIR and LWIR hyperspectral data obtained with the SPECIM AisaFenix and SPECIM OWL hyperspectral cameras, respectively (see example RGB image of the drill-core tray in Fig. 2a). The spectral range of the VNIR-SWIR data covers from 463 to 2479 nm, with a total of 340 bands. While the spectral range for the LWIR hyperspectral data covers from 7500 to 11977 nm in 94 bands. Spectral samples representative of each geochemical measurement were manually collected from both hyperspectral data sets. For the feature extraction technique, we set the number of principal components of each data source as 10. For the RF algorithm, we consider an ensemble of 500 decision tree classifiers.

Fig. 2b shows the obtained mineral map following the proposed machine learning framework. By visually comparing the resultant mineral map with the RGB image of the drill-core tray, it can be observed that the main mineral distributions and alteration patterns are being characterised. For example, in the center part of the bottom section of the drill-core, sulphides identified in the RGB image are being mapped as *class 3*.

Moreover, in Fig. 3 the advantage of fusing VNIR-SWIR and LWIR data is emphasized. For example, *class 1* and *class 2* present similar spectra in the VNIR-SWIR spectral range.

The main absorption features are located between 1350 - 1550 nm, which corresponds to the water and OH absorption window, and between 2306 - 2365 nm, which correlates with the CO_3-MgOH vibrational bonds, and only a subtle difference is present around 1400 - 1550 nm. However, in the LWIR range, the center spectra of both classes allows the differentiation between them, being *class 2* a silicates-rich class based on the feature located within 8500 - 9000 nm which corresponds to the SiO vibrational region. Hence, fusing both VNIR-SWIR and LWIR hyperspectral data represents an added value for the classification results and more comprehensive mineral maps are achieved.

4. CONCLUSIONS AND REMARKS

In this paper, we propose a machine learning approach to fuse geochemical and hyperspectral data to map minerals in drill-core samples. We suggest to use local geochemical data to generate training samples for the supervised classification of drill-core hyperspectral data. Hence, the detailed element composition analysis can be generalised along the drill-cores. Moreover, we present a data fusion approach for fusing VNIR-SWIR and LWIR hyperspectral data which complements the mineral mapping task and produces more comprehensive maps. Also, this improves the classification of minerals which are active exclusively in specific parts of the electromagnetic spectrum. The proposed approach considers a stacked vector of PCA features derived from both VNIR-SWIR and LWIR hyperspectral data cubes and extracted geochemical labels, as input for an RF classifier.

As part of our future work, we will explore different strategies for the fusion of hyperspectral and geochemical data in an advanced machine learning framework.

Acknowledgement

We would like to acknowledge and thank Mawson Resources Limited company, in particular Nicholas Cook, for providing the dataset for this work.

5. REFERENCES

[1] F. A. Kruse, "Identification and mapping of minerals in drill core using hyperspectral image analysis of infrared reflectance spectra," *International Journal of Remote Sensing*, vol. 17, no. 9, pp. 1623–1632, 1996.

[2] R. N Clark, "Spectroscopy of rocks and minerals, and principles of spectroscopy," in *Remote Sensing for the Earth Science: Manual of remote sensing*, Andrew N. Rencz, Ed., vol. 3, chapter 1, pp. 3–58. John Wiley & Sons, Inc, New York, 1999.

[3] S. Pontual, N. Merry, and P. Gamson, "Spectral Interpretation Field Manual," *AusSpec International Pty.*, vol. 1, pp. 92, 1997.

[4] J.W. Salisbury, L. S. Walter, and N. Vergo, "Mid-Indrared (2.1 - 25 um) spectral of Minerals: First Edition," Tech. Rep., U.S Geological Survey, 1987.

[5] R.F. Kokaly, R.N. Clark, G.A. Swayze, K.E. Livo, T.M. Hoefen, N.C. Pearson, R.A. Wise, W.M. Benzel, H.A. Lowers, R.L. Driscoll, and A.J. Klein, "USGS Spectral Library Version 7: U.S. Geological Survey Data Series 1035," 2017.

[6] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines," *Ore Geology Reviews*, vol. 71, pp. 804–818, 2015.

[7] I.C. Contreras Acosta, M. Khodadadzadeh, L. Tusa, P. Ghamisi, and R. Gloaguen, "A machine learning technique for drill core hyperspectral data analysis," in *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. (WHISPERS)*, Amsterdam, NL., 2018, vol. 9th, pp. 1–5, IEEE.

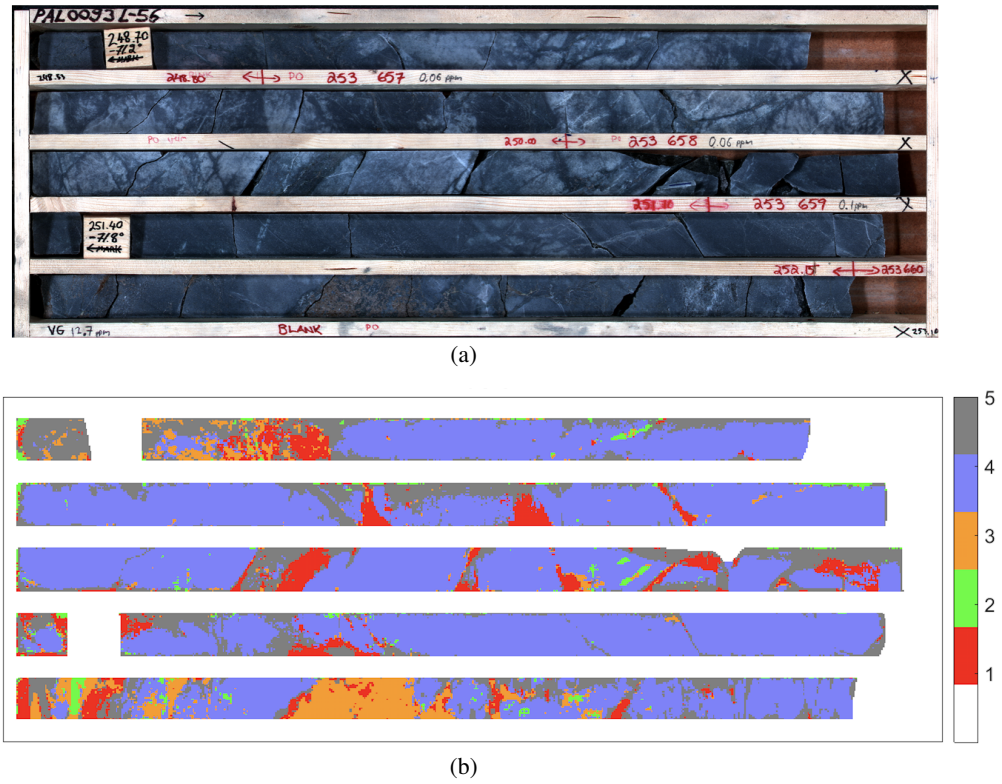
[8] A. Buccianti and E. Grunsky, "Compositional data analysis in geochemestry: Are we sure to see what really occurs during natural processes?," *Journal of Geochemical Exploration*, vol. 141, 2014.

[9] L. Jackson, A. Parbhakar-fox, N. Fox, S. Me, D. R Cooke, A. Harris, and E. Savinova, "Integrating hyperspectral analysis and mineral chemistry for geoenvironmental prediction," *11th ICARD IMWA MWD Conference- " Risk to Opportunity"*, , no. 2009, pp. 1075–1080, 2016.

[10] P.C. Pinet, D. Glenadel-Justaut, Y. Daydou, G. Ceuleneer, S. Gou, P. Launeau, S.D. Chevrel, and C. Carli, "MGM Deconvolution of Complex Mafic Mineralogy Rock Slab Spectra from Visible-Near Infrared Imaging Spectroscopy : Implications for the Characterization of the Terrestrial Oceanic and Lunar Crust," in *Proceedings Whispers 8th Conference*, L.A, California, 2016, pp. 1–4, IEEE.

[11] L Breiman, "Random Forest," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

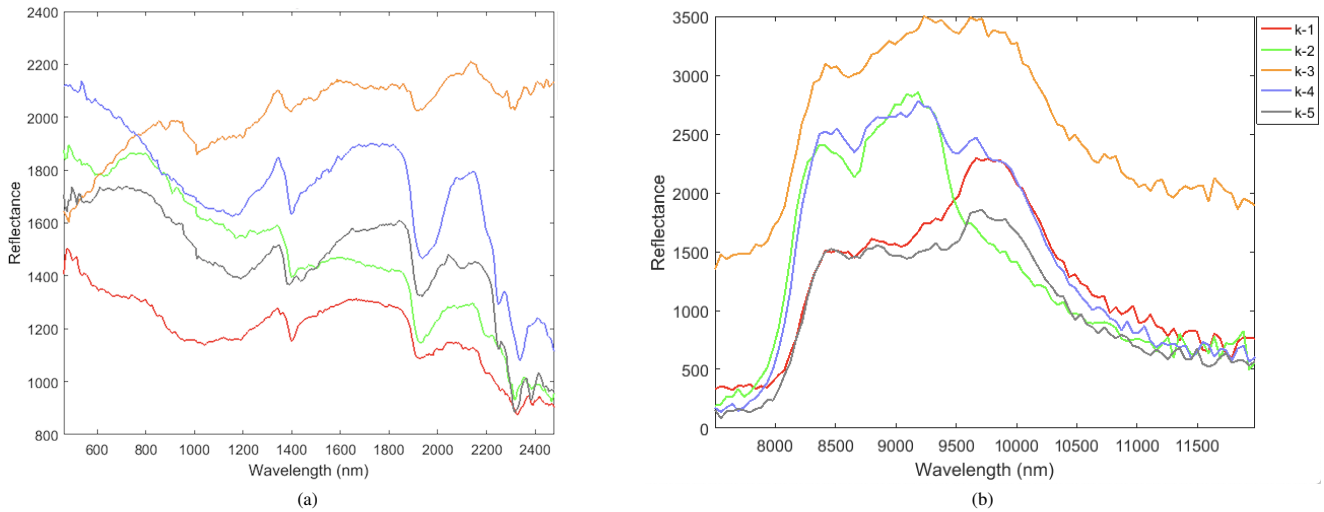
[12] D.J. Ketchen and C. L. Shook, "The application of cluster analysis in strategic management research: an analysis and critique," *Strategic Management Journal*, vol. 17, no. 6, pp. 441–458, 2002.



(a)

(b)

Fig. 2: (a) RGB image of the drill-core tray. (b) Mineral map obtained with the proposed approach.



(a) (b)

Fig. 3: Center spectra for each cluster in the resultant mineral map from the (a) VNIR-SWIR and (b) LWIR data sets.