



# Monitoring soil lead and zinc contents via combination of spectroscopy with extreme learning machine and other data mining methods

Vahid Khosravi<sup>a,\*</sup>, Faramarz Doulati Ardejani<sup>b</sup>, Saeed Yousefi<sup>c</sup>, Ahmad Aryafar<sup>c</sup>

<sup>a</sup> Faculty of Mining, Petroleum and Geophysics, Shahrood University of Technology, Shahrood, Iran

<sup>b</sup> School of Mining, College of Engineering, University of Tehran, Tehran, Iran

<sup>c</sup> Department of Mining, Faculty of Engineering, University of Birjand, Birjand, Iran

## ARTICLE INFO

Handling Editor: A.B. McBratney

### Keywords:

Mine waste dump  
Toxic elements  
Spectroscopy  
Binding mechanism  
Extreme learning machine  
Geostatistical interpolation

## ABSTRACT

In order to limit pollution risk and develop proper remediation strategies, soil quality has to be controlled by rapid and sustainable monitoring measures. Visible and near-infrared reflectance spectroscopy (VisNIR) is an attractive surrogate to time-consuming and costly classical soil assessment protocols. It highly depends on selecting appropriate data mining methods for regression analysis. In this study, performance of a state of the art learning algorithm called extreme learning machine (ELM), was evaluated through comparing with the other calibration methods proposed in the literature for predicting lead (Pb) and Zinc (Zn) concentrations. Solid samples collected from a mine waste dump ( $n = 120$ ) were scanned using a Fieldspec3 portable spectroradiometer with a measurement range of (350–2500 nm) in a laboratory. Transformation of the reflectance spectra to absorbance was followed by three pre-processing scenarios including Savitzky-Golay smoothing (SG), first derivative (FD) and second derivative (SD). Partial Least Square Regression (PLSR), Support Vector Machine (SVM) and neural networks with two learning algorithms models (back propagation and extreme learning machine), were calibrated on spectral features selected by genetic algorithm, and then applied to predict soil metal concentrations. The best prediction accuracy was obtained by FD-ELM method with  $R^2$ , RMSEP, concordance correlation coefficient and RPD values of 0.93, 63.01, 0.98 and 5.92 for Pb and 0.87, 167.90, 0.91 and 5.62 for Zn, respectively. Study of the prediction mechanism proved that element sorption by spectrally active Fe-oxide and clay contents of the soil was the major mechanism by which the spectrally featureless Pb and Zn ions can be predicted. The spatial patterns of predicted toxic elements showed that FD-ELM had the most similarity with those maps obtained by interpolating measured values. Over all, it is concluded that reflectance spectroscopy combined with the ELM algorithm is a rapid, inexpensive and accurate tool for indirect evaluation of Pb and Zn and mapping their spatial distribution in dumpsite soils of Sarcheshmeh copper mine.

## 1. Introduction

As a global environmental issue, soil contamination is continuously being intensified by anthropogenic phenomena such as industrial growth and rapid urbanization. Mining is one of the core industries which plays a positive role in economic development of the countries, however, it has the potential to contaminate soil, air, and water by discharging harmful materials into the environment. Huge amounts of mining wastes are produced and then deposited at the land surface taking a vast area and consequently transforming fertile, cultivated lands into wastelands (Li, 2006). Among the various geo-environmental impacts of mining, toxic elements contamination is certainly the most significant one due to its adverse effects on soil health. Hazardous impacts of toxic elements not only cause soil health degradation but

also impose significant harms on human health as well as wildlife resources by accumulating in living organisms and affecting the food chain (Achary et al., 2016).

Before planning for an expensive and time consuming soil remediation program, it is significant to monitor toxic elements concentrations. Sampling and laboratory analyses have to be performed to evaluate natural and anthropogenic concentrations of toxic elements. Inductively coupled plasma (ICP) method and X-ray fluorescence (XRF) spectroscopy are usually used to measure total metal concentrations in soil and sediment samples (Hubert, 1983). These methods are accurate and well documented by the researchers. However, they are expensive, time consuming and laborious since requiring substantial sample preparation and digestion with abrasive chemicals (Chakraborty et al., 2015). In order to overcome the impediments of the chemical

\* Corresponding author.

E-mail addresses: [v.khosravi@shahroodut.ac.ir](mailto:v.khosravi@shahroodut.ac.ir) (V. Khosravi), [fdoulati@ut.ac.ir](mailto:fdoulati@ut.ac.ir) (F. Doulati Ardejani), [s.yousefi@birjand.ac.ir](mailto:s.yousefi@birjand.ac.ir) (S. Yousefi), [aaryafar@birjand.ac.ir](mailto:aaryafar@birjand.ac.ir) (A. Aryafar).

measurement methods, a rapid, inexpensive and easy-to-use method is required for routine prediction and continuous monitoring of change and variability of toxic elements in soils of a specific site.

Visible and near-infrared (VNIR, 400–1200 nm) and shortwave infrared (SWIR, 1200–2500 nm) reflectance spectroscopy (VisNIR), coupled with some modelling techniques has been developed by some researchers to analyse the spectrally active properties of sediments and soil samples such as moisture (Wijewardane et al., 2016), organic matter (Bao et al., 2017; Nocita et al., 2014), clay (Adeline et al., 2017) and Fe (Adeline et al., 2017). Toxic elements at low levels do not have spectral features within VisNIR region. However, concentrations of featureless toxic elements can be determined via correlation with above mentioned active properties of soil. Spectroscopic based prediction of toxic elements in soil and water contaminated by mining activities have been reported in some studies (Kemper and Sommer, 2002; Choe et al., 2008; Zhuang, 2009; Pandit et al., 2010; Xian-Li et al., 2012; Sun and Zhang, 2017). However, little attention has been paid to heterogeneous environments such as mine dump sites. Gholizadeh et al. (2015a), investigated the possibility of predicting some toxic elements in the soil of a coal mine dumpsite using reflectance spectroscopy. In another study on mine wastes of Jalta and Bougrine in the North of Tunisia, reflectance spectroscopy was applied to identify Fe-related and clay minerals along with quantitative characterization of some toxic metals (Gannouni et al., 2012). As another effort, Khosravi et al. (2017a), investigated possibility of mapping the distribution of arsenic and chromium in a waste dump of Sarcheshmeh porphyry copper mine using combination of VisNIR reflectance spectroscopy and geostatistical analysis. Results demonstrated that the amount and spatial variability of arsenic and chromium can be determined using spectroscopy and geostatistics. Despite the studies mentioned, more spectroscopic based investigations are needed to examine the feasibility and mechanism of predicting concentrations and distribution patterns of toxic elements which provide a basis for environmental risk assessment of waste dumps.

Due to overlapping absorptions of some components, soil VisNIR spectra are weak, non-specific and broad. Hence, chemometric methods are often required to extract the soil spectral information (Martens and Naes, 1992; Viscarra Rossel and Beherens, 2010). Considering non-linear nature of the relationship between spectral data and target soil properties, classic linear models have some deficiencies to provide complete solution and may require intricate methods to detect the relationship between response of soil attributes and spectral features. Artificial neural networks (ANN) have an inherent ability to model non-linear relationships between a set of input and output variables (Samarasinghe, 2016). Multi-layer perceptron (MLP) is one of the most widely implemented feed forward neural network architectures which is typically trained with a simple and powerful algorithm, known as back-propagation (BP). Back propagation is a supervised learning procedure which modifies the weights and threshold values by iterative back propagating the error signal from output to the input layer until it is minimised or decreased to a pre-specified level. A number of recent studies have used back-propagation neural networks (BPNN) to predict soil properties from the spectral data. Based on visible and near-infrared spectroscopy, Chen et al. (2015) approved the ability of BPNN to estimate soil cadmium in an area under long-term sewage irrigation. Furthermore, BPNNs have been used to estimate lead concentration in landfill agricultural soils (Chakraborty et al., 2015), toxic metals contamination in soils after a mining accident (Kemper and Sommer, 2002), soil salt (Farifteh et al., 2007), organic matter and extractable forms of calcium, magnesium, and potassium (Xuemei and Jianshe, 2014). The major bottleneck of the BP algorithm is that the convergence speed of the learning procedure is too slow which is generally due to: (1) using slow gradient-based learning algorithms, and (2) iterative tuning of the network parameters. Extreme learning machine (ELM) is a new significantly fast and simple learning algorithm with a better generalisation performance than the gradient-based learning

algorithms such as back propagation. It has recently attracted considerable attention due to its successful use in a wide variety of spectroscopic applications including regression and classification problems (Chen et al., 2012; Guo et al., 2015; Sun et al., 2016; Wu et al., 2017). However, according to our reviews, prediction of soil toxic elements by combination of ELM and VisNIR spectroscopy has not been investigated in the previous studies.

The aims of our study were: a) Comparing the performance of ELM with other frequently used calibration techniques -PLSR, SVM and BPNN- to predict Pb and Zn concentrations in samples collected from one of the Sarcheshmeh porphyry copper mine waste dumps. b) Identification of the prediction mechanism by exploring the relationship between spectrally active binding sites and metal concentrations in waste samples. c) Comparison between the spatial distribution maps of waste Pb and Zn concentrations obtained by chemical measurement and prediction models.

## 2. Materials and methods

### 2.1. Study area

Sarcheshmeh is the largest open pit porphyry copper mine in Iran and one of the largest Oligo-Miocene copper deposits in the world. It is located in the Band Mamazar-Pariz mountains about 160 km southwest of Kerman, Kerman province, Iran (Fig. 1) (Aryafar et al., 2012).

Sarcheshmeh has a semi-arid climate with temperature varies between  $-20^{\circ}\text{C}$  to  $+35^{\circ}\text{C}$  in winter and summer, respectively (Doulati Ardejani et al., 2008). Basic to intermediate volcanic trachybasalt, trachyandesite  $\pm$  andesite rocks of Eocene age, form about 70% of the geological basis of the area (Khorasanipour and Aftabi, 2011). The Miocene stockwork-vein porphyry copper mineralization is a member of complex series of magmatic intrusives emplaced in the granitoid, trachybasalt and trachyandesite (Ghorashi-Zadeh, 1978).

The oval shaped ore body of Sarcheshmeh with dimensions of 1200 and 2300 m, contains nearly 1200 million tons of ore with average copper and molybdenum contents of 0.9% and 0.03%, respectively (Banisi and Finch, 2001). About 50 years of mining activity, has produced 31 active and inactive waste dumps containing over 400 million tons of wastes at the vicinity of the mine. There are some environmental problems associated with wastes dumps including occupying large areas and damaging landscapes, ground and surface water contamination and affecting ecosystems, vegetation, fauna and humen (Fig. 2). Contacting atmospheric oxygen and humidity with metal sulphides in the mine waste dumps generates acid mine drainage (AMD) (Yousefi et al., 2015). Toxic elements can be dissolved, mobilised, transported and released into the environment by these acid waters (Shoja and Salari, 2015).

### 2.2. Field sampling and laboratory measurements

A flowchart of the research is shown in Fig. 3. One hundred and twenty solid samples were gathered from dump No. 31, located at northeastern part of the main pit. They were collected from 0 to 10 cm depth of the surface, meanwhile, coordinates of the points were determined using a global positioning system (GPS) with an accuracy of about  $\pm 5$  m. To take representative samples, nearly 4 kg of waste material, sieved by screen 4 mesh, were collected at each sampling location. A stainless steel shovel was used to take samples and place them in air-tight polyethylene plastic bags (Khosravi et al., 2017b). Samples were transported to the laboratory, air-dried and their stones were removed. They were then crushed and sieved with a 2-mm sieve and divided into two subsamples, with one part used for chemical analysis of waste properties and another for spectral measurements after grounding to  $< 200\ \mu\text{m}$  to minimise the impact of particle size on samples' spectra (Kooistra et al., 2001). Concentrations of Pb and Zn, clay content, Fe oxides and paste pH were determined for laboratory

Download English Version:

<https://daneshyari.com/en/article/8894178>

Download Persian Version:

<https://daneshyari.com/article/8894178>

[Daneshyari.com](https://daneshyari.com)