Heliyon 8 (2022) e09712

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Field hyperspectral data and OLI8 multispectral imagery for heavy metal content prediction and mapping around an abandoned Pb–Zn mining site in northern Tunisia

Nouha Mezned [a,b,*], Faten Alayet [a], Belgacem Dkhala [a], Saadi Abdeljaouad [a]

[a] Laboratory of Mineral Resources and Environment, Department of Geology, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunis, Tunisia
[b] Faculty of Sciences of Bizerte, University of Carthage, Tunis, Tunisia

ARTICLE INFO

Keywords:
VNIR-SWIR Field hyperspectral data
Landsat OLI 8 imagery
Heavy metals
Contamination prediction
PLSR
MSC SNV preprocessing

ABSTRACT

Mining and smelting releases toxic contaminants such as zinc (Zn), lead (Pb) or cadmium (Cd) into the soil thereby poisoning it and rendering it unproductive. Remotely alternatives have been widely employed in the attempt of estimating heavy metal content within soils. The present study provides a methodological approach based on VNIR-SWIR field hyperspectral data and multispectral Landsat OLI 8 imageries for the prediction and mapping of Pb, Zn and Cd heavy metal contents around the abandoned Jebel Ressas mine site in Northern Tunisia. Thus, eighty-seven soil and tailing samples were collected from the study site and VNIR-SWIR field reflectances were measured on the same collection points, as well. All samples were analysed by atomic absorption for the estimation of heavy metal concentrations. The partial least squares regression PLSR was conducted considering the measured heavy metal concentrations and using multi-scale data: VNIR-SWIR field hyperspectral data and multispectral Landsat OLI 8 imagery. Standard normal variable (SNV) and multiple scatter correction (MSC) preprocessing methods were applied for further mapping improvement. Thus, this work aims to automate the estimation of the heavy metal contents in contaminated soils, by carrying out: a modeling approach based on the PLSR using VNIR-SWIR field hyperspectral data, ii) the mapping of Pb and Zn contents thanks to the exploitation of Landsat OLI8 multispectral imagery and iii) the application of both MSC and SNV preprocessing methods to optimize the performance of the developed models, when using such spectrally and spatially degraded data.

1. Introduction

Soil contamination has been highly exaggerated by human activities. Particularly, soil is chiefly polluted by tailings left over after mining activities (Zhiyuan et al., 2014; Fashola et al., 2016). Indeed, mining and smelting releases toxic contaminants such as lead or cadmium into the soil thereby poisoning it and rendering it unproductive (Razo et al., 2004; Zhou et al., 2007; Zhiyuan et al., 2014; Fashola et al., 2016). Geology remote sensing appear as interesting tool, which its efficiency, its low cost and its rapidity have been confirmed trough previous works. Particularly, this technique was successfully tested for estimating heavy metal content within soils, the critical compartment of successful agriculture. Indeed, techniques based on reflectance spectroscopic data, multispectral and hyperspectral data have proven their efficiency for the prediction of the chemical and physical properties of soils (Gomez and al., 2008; Rossel and Webster, 2012; Hribkou et al., 2012; Laone et al., 2012; Lu et al., 2013a,b; Mohamed et al., 2020).

Several previous studies used the visible-near infrared (VNIR) reflectance spectroscopy for the effective analysis of important properties of Mediterranean soils (Laone et al., 2012). Indeed, some authors have evaluated the potential of Vis-NIR spectroscopy for the prediction of chemical and physical properties of soils, in southern Italy, such as clay, organic carbon (OC), total nitrogen (N), cation exchange capacity (CEC), and calcium carbonate ($CaCO_3$). Others have however combined laboratory spectroscopy with Hyperion imagery to analyze (SOC), total phosphorus (TP), pH and cation exchange capacity (CEC) in soils in a typical area of Chinese Gansu Province (Lu et al., 2013a,b). A recent study (Mohamed et al., 2020), was focused on the use of Visible near infrared and shortwave infrared VNIR-SWIR spectroscopy and Landsat-8 operational land imager (OLI) images to predict the content of different soil nutrients (nitrogen (N), phosphorus (P), and potassium (K)), pH, and soil organic matter (SOM) in the Wadi El-Garawla arid area (the northwest coast of Egypt). Hbirkou et al., have only used airborne

* Corresponding author.
E-mail address: mezned_nouha@yahoo.fr (N. Mezned).

https://doi.org/10.1016/j.heliyon.2022.e09712
Received 18 October 2021; Received in revised form 11 February 2022; Accepted 7 June 2022

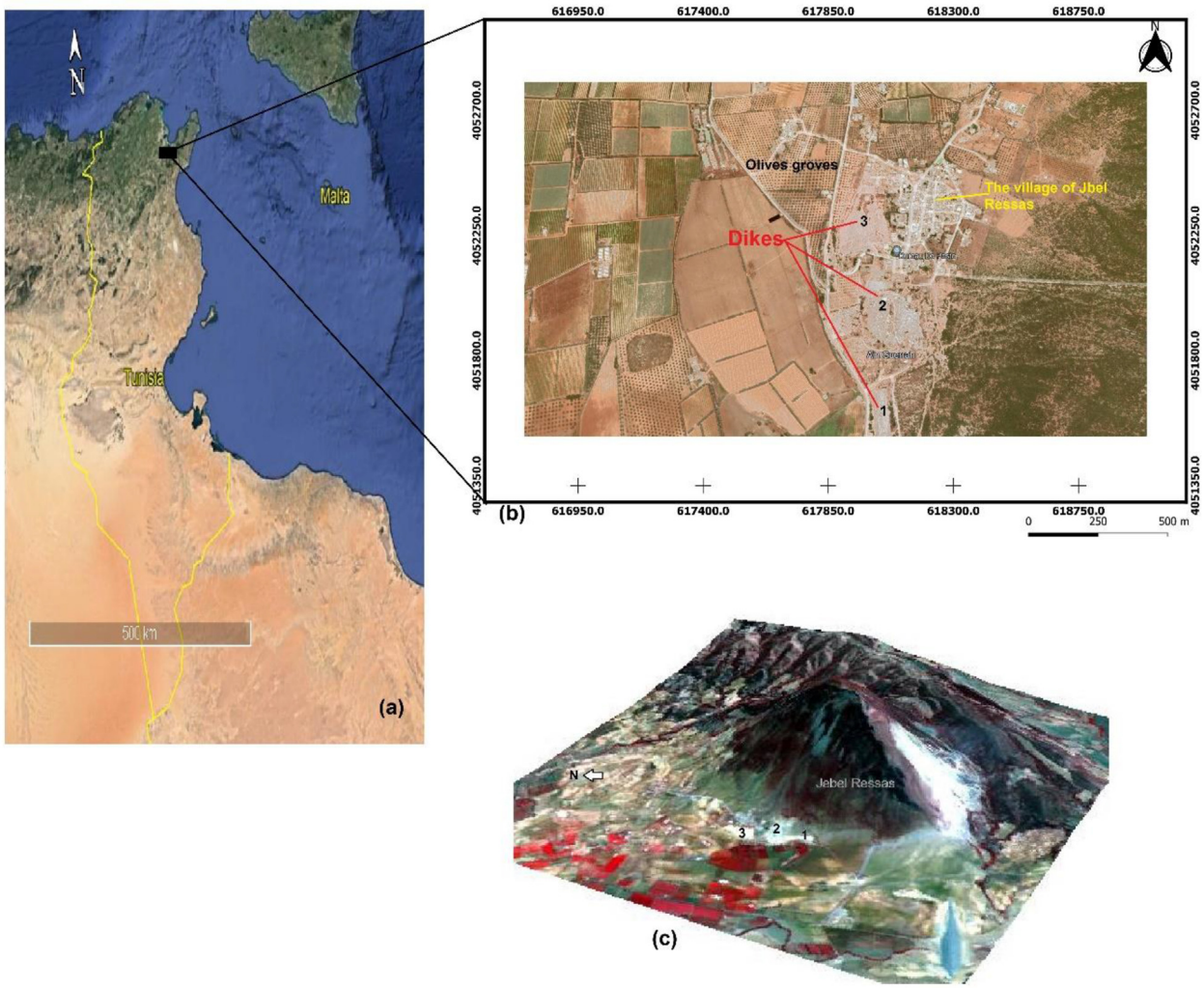

Figure 1. The location of the Jebel Ressas mine site in the North of Tunisia, showed on: (a) the Google earth map of Tunisia, (b) the zoomed window of the map and (c) the 3 D dropped false color composite image on the terrain digital model TDM, showing the topography of the study region. The black numbers (1, 2 and 3) locates the three tailing dumps (Dikes).

hyperspectral data (HyMap) for an accurate prediction of the soil organic carbon (SOC) content on cultivated fields in Germany.

Furthermore, other researchers, as human activities continues to exacerbate soil degradation, have focused on predicting heavy metal contents, such as As, Cu, Ni, Cr, Pb, Zn, Hg, Cd, etc., to assess soil contamination around mining and industrialization areas (Wu et al., 2005; Choe et al., 2008; Peng et al., 2016; Goltappeh et al., 2018; Wei et al., 2019). Laboratory VNIR (400–2500) data were well used in several studies aiming the prediction of contaminant contents, such as cadmium (Cd) and zinc (Zn) in Flood plain in Netherlands (Kooistra et al., 2001), lead (Pb) within soils in Australia (Al Maliki et al., 2014), zinc-copper-lead-chromium-nickel (Zn–Cu–Pb–Cr–Ni) in farming soils in Southern Bulgaria (Todorova et al., 2014), arsenic-lead-zinc-copper (As–Pb–Zn–Cu) in Chinese farming (Wang et al., 2014), arsenic (As) in farming soils in Italy (Stazi et al., 2014), and cadmium (Cd) (Chen et al., 2015) and zinc (Zn) (Sun and Zhang, 2017) within soils in China.

In almost of these research studies, several regression algorithms have been operated such as principle component regression (PCR) (Wu et al., 2005), multiple linear regression (MLR) (Choe et al., 2008), the cubist approach using spectral indices (Peng et al., 2016) and linear and non linear regression (LR & NLR) (Goltappeh et al., 2018). Among these regression approaches dedicated for interpreting spectral data to predict contents of heavy metals, the most used technique is partial least squares regression (PLSR), which handles multi-linear data and allows for the number of samples to be exceeded by that of variables (Shi et al., 2014; Wang et al., 2014; Wei et al., 2019).

Few works were however, interested in heavy metal mapping using PLSR and multispectral data. Particularly, PLSR and Landsat 8 imageries were used for mapping contents of soil arsenic (As), cadmium (Cd), nickel (Ni), and lead (Pb) (Fard et Matinfar, 2016), soil copper (Cu) (Fang et al., 2018) and soil chromium (Cr) (Khosravia et al., 2021).

In this context, the present work aims to optimize the capability of both VNIR-SWIR field hyperspectral spectroscopy and multispectral Landsat OLI 8 data for mapping heavy metal contents, particularly those of Pb, Zn and Cd, around the abandoned Jebel Ressas mine site in Northern Tunisia. Thus, this research study consists of: i) conducting PLSR algorithm for the prediction of heavy metal contents using field spectroradiometer, ii) mapping the spatial distribution of heavy metals using multispectral Landsat OLI8 data and iii) optimizing the performance of the developed models basing on standard normal variable (SNV) and multiple scatter correction (MSC) preprocessing methods.

2. Materials and methods

2.1. Study area

The district of Djebel El Ressas (Figure 1) which is considered as a pilot site (36°36′23.70″N, 10°19′05.86″E), is located 30 km South East

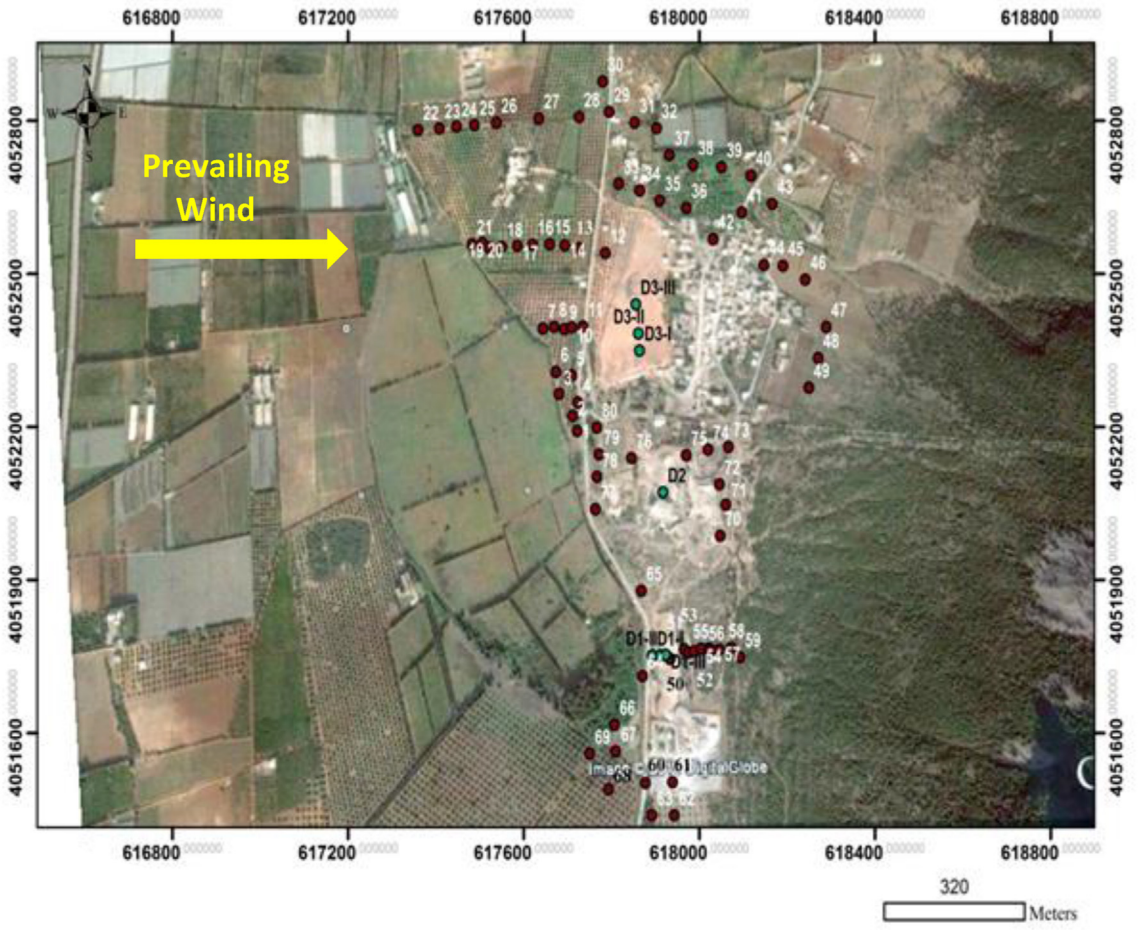

Figure 2. Map showing the point location of the eighty-seven soil (red point circles) and tailing (green point circles) samples as well as of the spectral measurement, selected around the Jebel Ressas mine. D1, D2 and D3 indicate the three tailing dumps (Dikes) Dike 1, Dike 2 and Dike 3, respectively.

from Tunis between the Ressas mountain and the Hma River, near the village of Djebel El Ressas in the North of Tunisia (Sainfeld, 1952). This abandoned mine site is characterized by a Mediterranean semi-arid climate and was exploited for the lead (Pb) and Zinc (Zn) ores. Mining wastes are stocked in three dikes, surrounded by agricultural soils, olives groves and forest. High levels of pollutants such as lead, zinc, cadmium and other heavy metals; have been reported (Mezned et al., 2020; Ghorbel, 2012; Alayet et al., 2017).

2.2. Field campaign

Eighty-seven soil samples were collected from agricultural soils around the three dykes in the Jebel Ressas mine site in 2014 (Figure 2). Samples were gathered from the first 3 cm of the soil surface, from the sampling points that were selected with respect to the contaminant transport direction. Moreover, hyperspectral field measurements were operated for all soil samples using an Analytical Spectral Device ASD Field Spec HiRes Full-Resolution spectroradiometer, under natural light between 10 AM and 2 PM (Figure 3). Three consecutive measurements of VNIR and SWIR reflectance spectra were practically taken per sample. The mean spectrum was considered after that for processing. The used instrument operates between visible (VIS) and Shortwave (SWIR) (350–2500 nm) wavelengths with a spectral resolution ranging from 3 nm to 10 nm. At a distance of 50 cm, away from each soil sample, it provides a spot diameter of 3.5 cm and an area of 38.4 cm^2, since its Field of view (FOV) is of 8°. A total number of 2151 of spectral bands were scanned for each point sample, of which 453 bands were removed in the present study, at intervals between 350-399 nm, 1340–1445 nm, 1801–1989 nm and 2401–2500 nm, due to their low instrumental signal-to-noise ratios.

2.3. Laboratory geochemical analysis

All the collected samples were placed into polyethylene bags properly labelled and transported to the laboratory, where they were air-dried, homogenized and quarted to obtain the most representative samples with about 15 g per sample as weight. After manual grinding of samples using an agate mortar, 0.3 g from each soil sample was devoted to Pb, Zn, and Cd heavy metal content determination using conventional chemical analysis. Indeed, each soil sample was the subject of firstly, a tri-acid ($HPClO_4$– HF – HNO_3-) attack and secondly, a determination of heavy metal contents by atomic absorption spectroscopy "Solar Thermo Elemental". Indeed, the tri-acid attack consists of two steps, a first cold attack, using both perchloric ($HPClO_4$) and hydrofluoric (HF) acids, was carried out on each soil sample before a hot attack using nitric acid HNO_3 (Figure 4). The extract was adjusted then with bi-distilled water. It is from this solution that the metal trace elements were determined by SOLAAR type flame Atomic Absorption Spectrophotometry.

2.4. Multispectral Landsat OLI 8 imagery

The used image for the present study was acquired by Landsat OLI 8 data on the 23th of March 2014. Landsat 8 is the most recently launched Landsat satellite and carries the Operational Land Imager (OLI) sensor that captures data with improved radiometric precision over a 12-bit dynamic range. Thus, it improves overall signal to noise performance

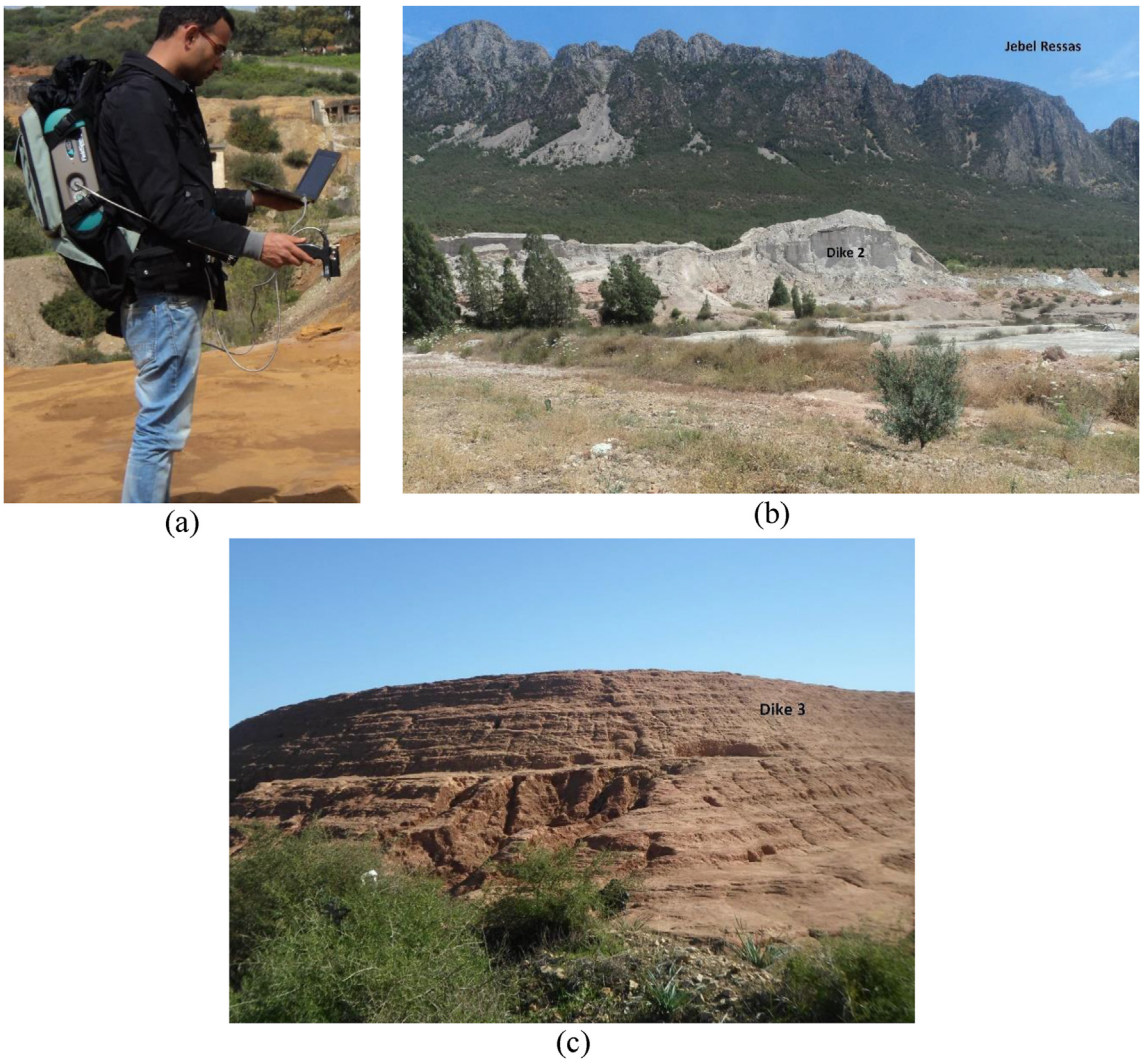

Figure 3. Photos showing: (a) the procedure of the field spectra measurement and the selected measurement points in: (b) the second tailing dump (Dike 2) and (c) the third tailing dump (Dike 3).

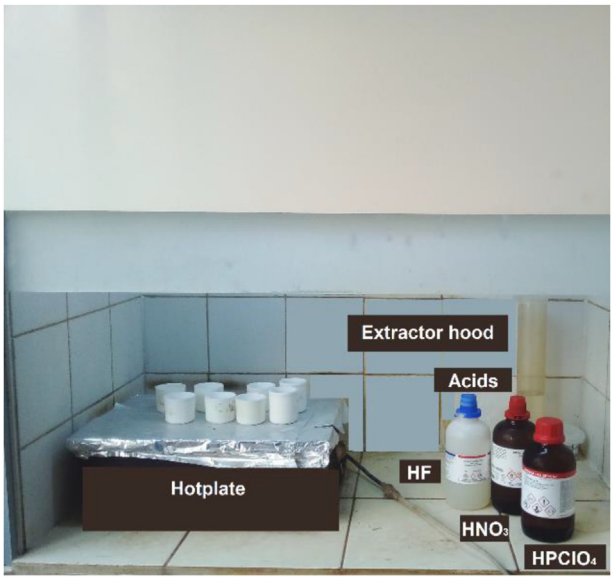

Figure 4. A photo showing the experimental tools that were used for the acid attack that was carried out on soil samples in the laboratory.

Table 1. Spectral characteristics of the Landsat OLI 8 imagery.

Band	Domain	Spectral range (nm)	Spatial resolution (m)
Band 1	Visible	430–450	30
Band 2	Visible	450–510	30
Band 3	Visible	530–590	30
Band 4	Red	640–670	30
Band 5	Near-Infrared	850–880	30
Band 6	SWIR 1	1570–1650	30
Band 7	SWIR 2	2110–2290	30

that enables improved characterization of land cover state and condition. The Level-1 data products are delivered after standard processing for each scene. In this study, we used the level L1TP image which, was radiometrically calibrated and orthorectified using ground control points (GCPs) and digital elevation model (DEM) data to correct for relief displacement. The OLI 8 data have nine spectral bands, including a pan band. Only 7 bands centered on 440 nm, 480 nm, 545 nm, 655 nm, 865 nm, 1610 nm and 2200 nm have been used in this study (Table 1). Panchromatic and cirrus bands were eliminated. All bands present a spatial resolution of 30 m.

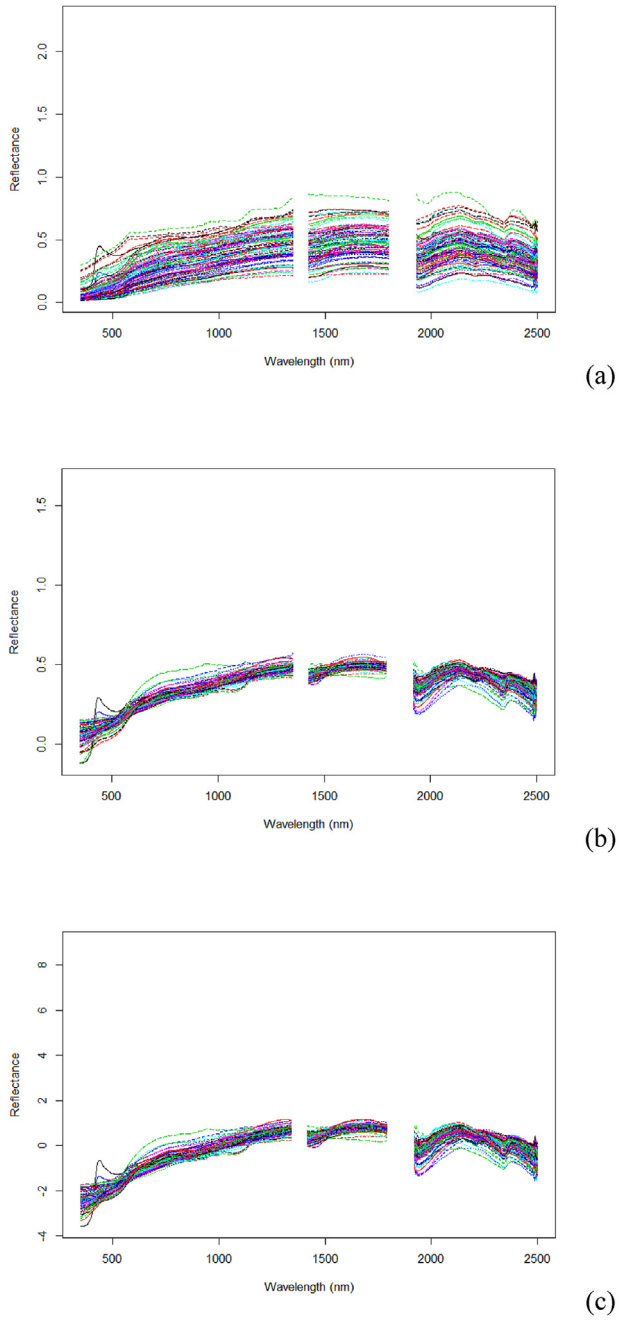

Figure 5. Field VNIR-SWIR reflectances of the 87 samples: (a) before preprocessing, (b) after MSC preprocessing and (c) after SNV preprocessing. The noises at 1400 and 1900 nm, related to atmospheric water absorptions, were removed.

Intended to eliminate the interference of vegetation areas, a mask based on NDVI calculation was applied to the Landsat OLI 8 image. All pixels having values that exceed 0.3 were removed (Lu et al., 2013a,b). The NDVI values have been calculated using band 4 and band 5 (655 nm, 865 nm) of the Landsat OLI 8 image after operating the following equation (Eq. (1)):

$$\text{NDVI} = (\text{Band 5} - \text{Band 4}) / (\text{Band 5} + \text{Band 4}) \quad (1)$$

The urban areas were also masked after identification through visual inspection. The FLAASH algorithm was after that applied for atmospherically correction of the Landsat OLI 8 image.

Table 2. Valuation parameters for performance model assessing.

Parameters	Formula	References
Coefficient of determination	$R^2 = 1 - \frac{\sum_n^{i=1}(y_i' - y_i)^2}{\sum_n^{i=1}(y_i' - \overline{y})^2}$	Kotz et al. (2006)
Root Mean Square Errors	$RMSE_p = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}}$	Douglas et al. (2018); Rossel et al. (2006); Chang et al. (2001)
Ratio of Performance to Deviation	$RPD = \frac{STD}{RMSE_p}$	Olarewaju et al. (2016); Davey et al. (2009)
Ratio of the Performance to Interquartile	$RPIQ = \frac{IQ}{RMSE_p} = \frac{Q3 - Q1}{RMSE_p}$	Bellon-Maurel et al. (2010)

y_i is the measured value of sample i; y_i' is the predicted value of sample i; $\overline{y}$ is the average of measured value; N is the number of samples; $Q1$ is the first quartile; $Q3$ is the third quartile.

2.5. Field and imagery data preprocessing

In the present study and aiming to remove unwanted variation that might impair the predictive ability of models, without excluding or altering chemically relevant variation, commonly used preprocessing techniques were applied. As carried out in several previous research studies (Zeaiter and Rutledge, 2009; Jonsson and Gabrielsson, 2009; Barnes et al., 1989), in this study, both preprocessing methods including standard normal variable SNV and multiple scatter correction MSC were particularly applied to the field hyperspectral data and to the Landsat OLI 8 multispectral imagery. The goal of conducting these methods was to remove both additive and multiplicative baseline variation without altering the shape of the spectra. Particularly, the SNV transformation, introduced by Barnes et al. (1989), was dedicated to reduce multiplicative effects of scattering and particle size. Differences in the global intensities of the signals were thus reduced. Indeed, the transformation consists in centering each spectrum and then scaled it by dividing by its standard deviation through the following equation (Eq. (2)):

$$x_{i,j}^{SNV} = \frac{(x_{i,j} - \overline{x}_i)}{\sqrt{\frac{\sum_{j=1}^{p}(x_{i,j} - \overline{x}_i)^2}{p-1}}} \quad (2)$$

Where:

$x_{i,j}^{SNV}$ is the element of the transformed spectrum,

$x_{i,j}$ is the corresponding original element of the spectrum i at variable j,

$\overline{x}_i$ is the mean of spectrum i,

p is the number of variables or wavelengths in the spectrum.

Alternatively, the Multiplicative scatter correction (MSC) transform, proposed by Isaksson and Kowalski (1993) and has been shown to be equivalent to the SNV transform, was carried out on each sample with a difference only in scaling factors.

The applied equation for a single spectrum is as follows (Eq. (3)):

$$x = a + b\hat{x} + e \quad (3)$$

Where:

x represents the spectrum of a sample,

$\hat{x}$ represents the average spectrum of the whole calibration data set that assumed to contain similar chemical information as x,

The applied equation using ordinary least squares regression can allowed the calculation of the corrected spectrum x^* as following (Eq. (4)):

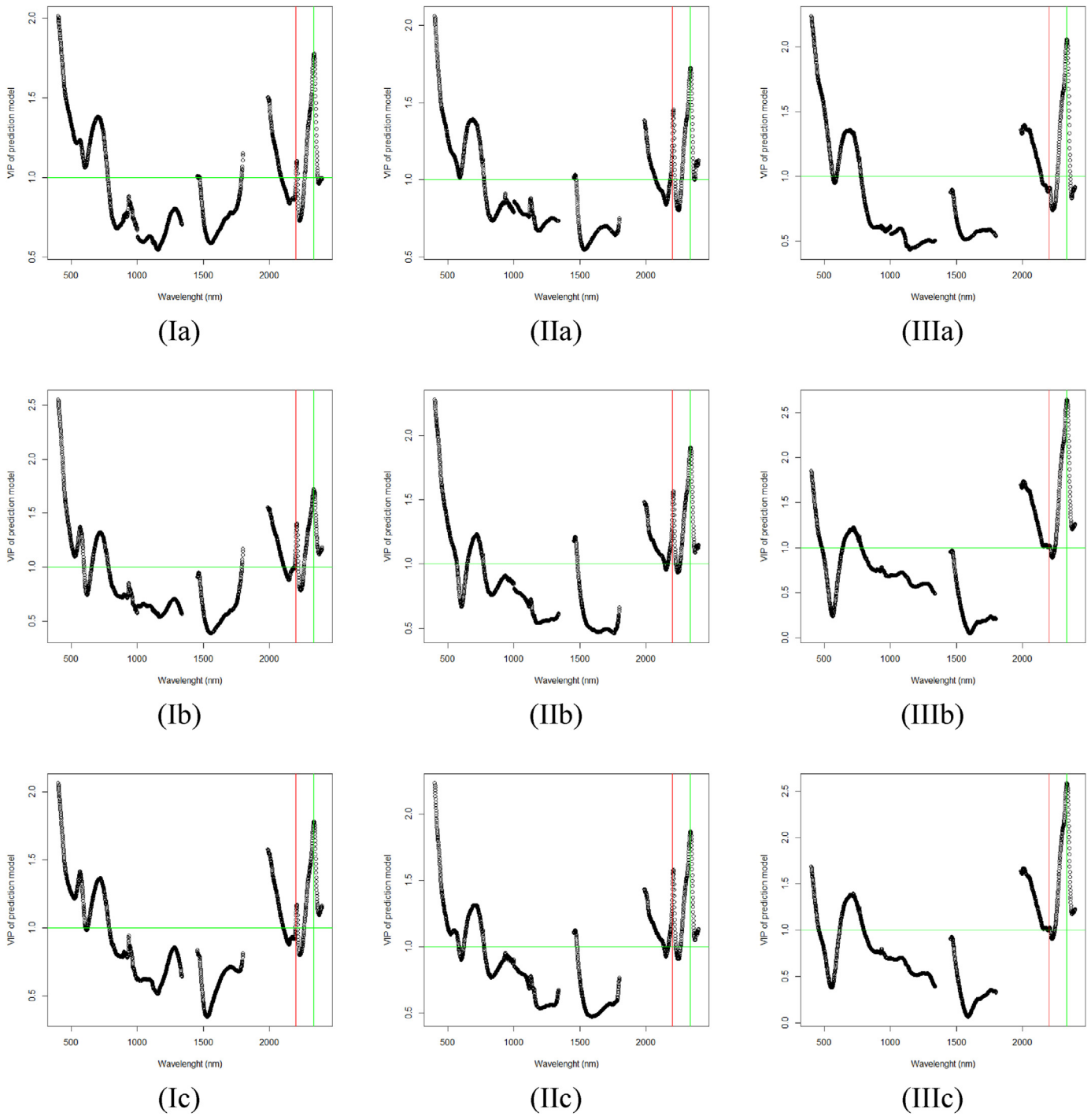

Figure 6. VIP scores showing the most important spectral features used by the PLSR regression for the prediction of heavy metal concentrations (I) Pb, (II) Zn and (III) Cd. The VNIR-SWIR data, which were used are from: (a) ASD field measured spectra, (b) ASD field measured spectra after preprocessing based on the MSC method and (c) ASD field measured spectra after preprocessing based on the SNV method. The red line indicates the band at 2200nm and the green one indicates the band at 2234nm.

$$x^* = \frac{(x - \hat{a})}{\hat{b}} = \hat{x} + \frac{\hat{e}}{\hat{b}} \quad (4)$$

2.6. Regression processing and performance assessing

The proposed methodological approach, aiming to model heavy metal contents within soils, was carried out using the PLSR regression method based on the resulting laboratory data and the VNIR-SWIR field hyperspectral data. Indeed, the resulting laboratory data consists of heavy metal contents or concentrations, which were determined by the atomic absorption method. The VNIR-SWIR field hyperspectral data consists of the VNIR-SWIR hyperspectral measurements of different soil samples, which were taken in situ in the same collection points. All spectral data were preprocessed using SNV and MSC transformations before PLSR regression step (Figure 5), in order to remove additive and multiplicative effects. The optimum number of Latent Variables were after that determined through the leave-one-out cross-validation procedure (Wold, 1978) in an attempt to verify the prediction capability of

Table 3. Heavy metal statistics of soil samples.

Parameter	Pb (mg/kg)	Zn (mg/kg)	Cd (ppm)
n	87	87	87
Min	200	500	0.9
Max	57200	80100	983.49
Mean	880	2280	119.23
Median	630	1320	22.01
Standard Deviation	10600	23500	187.29
Skewness	2.06	0.86	2.16
Kurtosis	8.29	2.43	8.49

Table 4. Heavy metal statistics correlation of soil samples.

		Pb mg/Kg	Zn mg/Kg	Cd ppm
Pb mg/Kg	Pearson Correlation	1	0.846**	0.758**
	Sig. (2-tailed)		0.000	0.000
Zn mg/Kg	Pearson Correlation		1	0.770**
	Sig. (2-tailed)			0.000
Cd ppm	Pearson Correlation			1
	Sig. (2-tailed)			

**. Correlation is significant at the 0.01 level (2-tailed).

each generated PLSR model for the calibration dataset. This procedure was repeated for all n samples to obtain heavy metal content prediction for each of them. The optimum number of Latent Variables was selected as the number resulting in the first minimum of the root mean square errors of cross-validation (RMSECV) (Viscarra Rossel et al., 2006). The resulting optimal PLSR model was then applied to the validation dataset. Considering the Mahalanobis distance (Mark and Tunnell, 1985), an H value higher than 3 was defined as spectral outliers and was removed from the calibration dataset.

In this study, calibration and validation datasets, using respectively the 75% and the 25 % of the 76 samples, were considered in PLSR regression process. A test dataset, using ten remaining samples that were selected randomly, were however, considered for the validation of the generated maps of heavy metals. According to the previous study (Dkhala et al., 2020) that have explored the difference related to the division of the whole dataset into calibration and validation sets, heavy metal concentrations were sorted in increasing order, which was showed as the optimal method.

The performance of the generated PLSR models was assessed through the comparison of predicted content values of heavy metals of the validation data set with the measured ones. Thus, the coefficient of determination in the calibration (R^2_{cv}) and the validation (R^2_{val}) sets, the root mean square errors of cross-validation (RMSECV) and the root mean square errors in the validation (RMSEP) sets, the ratio of performance to deviation (RPD) and the ratio of the performance to interquartile (RPIQ) were calculated (Table 2). The b-coefficient, larger than the standard deviation of the b-coefficients for all the spectral bands, and the variable importance in the projection (VIP), higher than 1 (Figure 6), were justified in order to have the most significant wavelengths n in the PLSR (Chong and Jun, 2005; Wold et al., 2001).

3. Results and discussions

3.1. Geochemical estimation of heavy metal contents

All concentrations of the studied heavy metals were determined basing on the atomic absorption spectroscopic results. Mean concentrations of heavy metals in all soil samples, which are considered as an indicator of the soil contamination, are summarized in Table 3. All saved values in soil samples exceed the tolerate values of soil quality guidelines (Henin, 1983). Given the tolerated value of each heavy metal, Cd show the highest concentration of the three pollutants, exceeding 0.7 ppm and being between 0.9 and 983.49 ppm. Pb show also high values exceeding 100 mg/kg (100 ppm), the tolerated concentration within soils. Zinc concentrations are ranging from 500 to 67000 mg/kg within soils and can reach 80100 mg/kg within tailing dumps and exceed thus the tolerated abundance estimated at 300 mg/kg. These revealed high concentrations of Pb, Zn and Cd heavy metals are detected in agricultural soils particularly, around the three tailing deposits, which are considered as source of contaminants. As showing in the Table 4, all trace elements are correlated according to the high R^2 coefficients, up to 0.7.

3.2. Spectral characterization of heavy metals

Soil samples have shown distinguished absorption peaks in the SWIR region that are centered at 2202 nm and at 2338 nm, respectively revealing the presence of clays (OH–Al) and carbonates (CO_3) (Hunt et al., 1971; Thompson et al., 1999). The presence of carbonates in the present site, was confirmed with a dominance of the carbonate calcite ($CaCO_3$), at about a mean concentration of 35 % in soils and of 75 % in dyke. Other carbonates were also identified within soil samples such as dolomite ($CaMg(CO_3)_2$), cerusite ($PbCO_3$), magnesite ($MgCO_3$) and ankerite ($FeCO_3$), but insignificantly. Previous studies (Kemper and Sommer, 2002; Hou et al., 2019) have shown that heavy metals, which are present in trace amounts within soils, are predicted indirectly, depending on their behavior. Heavy metals prediction is related to the affinity of each contaminant to different soil components, such as Fe oxides, soil organic matter SOM and above all clays, as they are particularly known as the most fixing minerals for them.

Particularly, the prediction of heavy metal concentrations such as Cd, Ni and Zn, using soil reflectance spectroscopy and statistical techniques was well operated using the band related to the presence of clays (Sun and Zhang, 2017; Sun et al., 2018; Zhang et al., 2019). In other works, heavy metals have been directly detected. Fard and Matinfar (2016) have determined prominent correlations for As, Cd and Pb, which occurred respectively in the spectral range of 1600–2400 nm, of 350–400 and 800–850 nm, and of 350–400 and 1300–1800 nm.

In this study, the considered spectral ranges in the SWIR region are of 2000–2400 nm, particularly around 2338 nm (carbonates) for Cd, and as well around 2202 nm (clays) for Pb and Zn. In the VNIR region, they are of 650–800 nm for the three heavy metals, of 400–450 and 500–600 nm for Pb and of 400–480 for Cd.

3.3. Heavy metal content modeling

3.3.1. Modeling based on field hyperspectral spectroscopy

The PLSR models were generated for the Pb, Zn and Cd trace elements using hyperspectral reflectance spectroscopy (1698 spectral bands) and their respective concentrations that were determined by atomic absorption analysis. The latent variables were fixed at 15 as maximum number.

The most important spectral region used by the PLSR for the Pb contents prediction is of 400–450 nm, 500–600 nm and 650–800 nm (Figure 7). Other spectral ranges are considered in the SWIR region around 2202 nm and 2338 nm using the MSC and SNV based preprocessed VNIR-SWIR reflectances. These spectral bands correspond to clays (Al–OH) and carbonate (CO_3) occurrences, respectively. The same spectral regions are also selected for the prediction of Zn contents, as well as others around 2000 nm and 2100 nm for both preprocessed and non-preprocessed VNIR-SWIR field spectra (Figure 7). Except the spectral absorption feature around 2338 nm, the selected spectral regions for the prediction of Cd contents, using preprocessed and non-preprocessed field spectra, are of 400–480 nm, 650–800 nm in the VNIR region and around 2000 nm and 2338 nm in the SWIR region (Figure 6).

Giving the first local minimum of the RMSECV, 2 to 11 latent variables were selected for all PLSR models, using preprocessed and non-

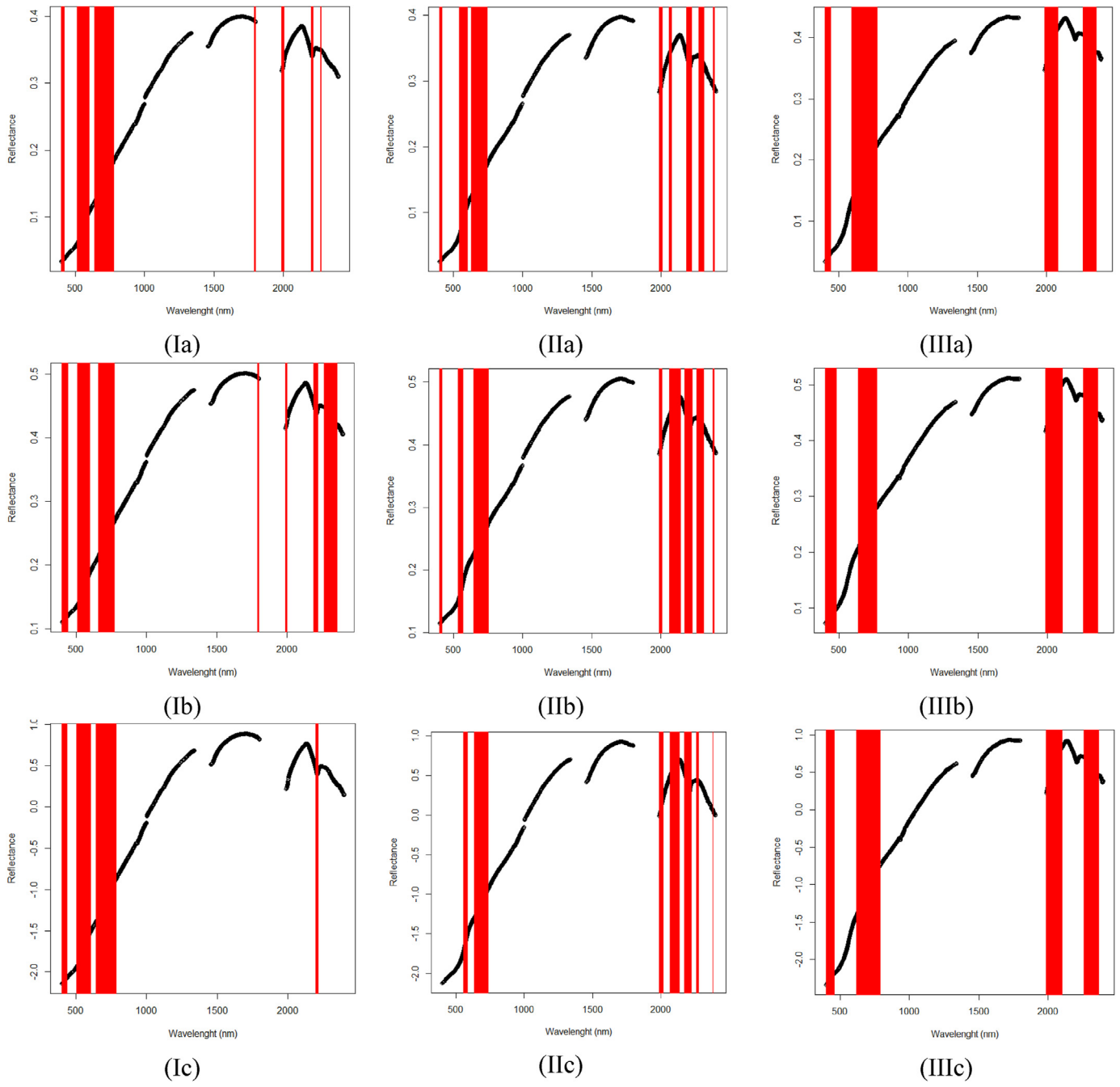

Figure 7. The most spectral bands used in the PLSR regression for the prediction of heavy metal concentrations (I) Pb, (II) Zn and (III) Cd. The VNIR-SWIR data, which were used are from: (a) ASD field measured spectra, (b) ASD field measured spectra after preprocessing based on the MSC method and (c) ASD field measured spectra after preprocessing based on the SNV method.

preprocessed VNIR-SWIR hyperspectral measurements (Table 5). The Zn content prediction model have shown the highest prediction accuracies with an R^2_{cv} equal to 0.79 and an RPD_{cv} and an RPD_{val} at 2.11 and 1.54, respectively. These prediction accuracies have been improved for the three PLSR models, using MSC and SNV based preprocessed VNIR-SWIR reflectances, showing an R^2_{cv} up to 0.8 and an RPD_{cv} that exceeds 2.18. The most interesting prediction accuracies are detected for the Zn content prediction model using the SNV based preprocessed VNIR-SWIR reflectances, with an R^2_{cv} equal to 0.82 and an RPD_{cv} and an RPD_{val} at 2.29 and 1.71, respectively. For Pb and Cd content prediction models, the most interesting prediction accuracies were revealed using the MSC preprocessing method with an $R^2_{cv} = 0.41$, an $R^2_{val} = -0.33$, an $RPD_{cv} = 1.1$ and an $RPD_{val} = 1.01$ for the Pb, and an $R^2_{cv} = 0.5$, an $R^2_{val} = 0.46$ an $RPD_{cv} = 1.13$ and an $RPD_{val} = 1.4$ for Cd. Through these results, it has been shown that MSC and SNV preprocessing methods improved the prediction accuracies of the different models that were generated for the prediction of heavy metal contents.

3.3.2. *Modeling based on Landsat OLI 8 data*

Using Landsat OLI8 data, 1 to 4 latent variables were selected for the three models for each of the Pb, Zn and Cd heavy metals. In this case study, the pixel spectra in the image, which contained the average reflectance of a 30 × 30 m, correspond to the GPS coordinates (Differential Global Positioning System DGPS) of the field sampling point.

Table 5. Results of the PLSR models of the heavy metal content prediction using the field VNIR-SWIR data and given different preprocessing methods.

Heavy Metal	Treatment	Calibration Set (56 samples)			Validation Set (19 samples)			RPIQ (test-set cal)	VL
		R^2cv	RMSECV	RPD	R^2	RMSEval	RPD		
Pb (mg/kg)	*Reflectance*	0.45	8367.8	1.12	0.04	8390.32	1.05	1.22	6
	SNV	0.38	8774.55	1.05	-0.02	8672.98	1.02	1.12	5
	MSC	0.41	8545.45	1.1	-0.33	9887.97	1.01	1.29	5
Zn (mg/kg)	*Reflectance*	0.79	10915.96	2.11	0.56	15050.12	1.54	3.47	9
	SNV	0.82	9963.97	2.29	0.64	13610.01	1.71	3.65	11
	MSC	0.82	9845.26	2.29	0.54	15306.2	1.52	3.59	9
Cd (ppm)	*Reflectance*	0.59	130.23	1.34	0.08	141.29	1.07	1.38	3
	SNV	0.49	143.19	1.12	0.54	100.5	1.51	1.45	2
	MSC	0.5	140.84	1.13	0.46	108.36	1.4	1.45	2

Table 6. Results of the PLSR models of the heavy metal content prediction using the Landsat OLI 8 data and given different preprocessing methods.

Heavy Metal	Treatment	Calibration Set (54 samples)			Validation Set (18 samples)			RPIQ (test-set cal)	VL
		R^2cv	RMSECV	RPD	R^2	RMSEval	RPD		
Pb (mg/kg)	*Reflectance*	0.19	8095.7	0.63	0.24	6675.81	1.18	0.84	4
	SNV	0.26	8655.18	0.67	0.13	7139.68	1.11	0.9	2
	MSC	0.32	8328.9	0.74	0.13	7141.94	1.11	1.07	1
Zn (mg/kg)	*Reflectance*	0.11	19995.28	0.57	0.65	12419.56	1.75	0.67	4
	SNV	0.22	20808.19	0.66	0.53	14403.66	1.51	0.66	2
	MSC	0.24	20679.58	0.66	0.5	14909.09	1.46	0.66	2
Cd (ppm)	*Reflectance*	0.09	160.47	0.52	0.15	126.28	1.11	0.7	3
	SNV	0.13	163.57	0.52	0.08	131.44	1.07	0.54	2
	MSC	0.14	165.31	0.51	0.09	130.53	1.08	0.53	2

The most used spectral bands for the Pb contents prediction were the band 5 (865 nm), the band 6 (1610 nm) and the band 7 (2200 nm). The three spectral bands, band 4 (650 nm), band 5 (865 nm) and band 7 (2200 nm) were however, the most used for the prediction of the Zn heavy metal contents. For the prediction of the Cd contents, only both spectral bands were the most used, band 4 (650 nm) and band 5 (865 nm).

Although, all the generated PLSR models revealed relatively weak statistical results (Table 6). Indeed, models of the Zn content prediction seem to be the most accurate ones, considering the coefficient of determination R^2_{cv} (ranging between 0.11 and 0.24), R^2_{val} (ranging between 0.5 and 0.65), the RMSECV (around 20000), the RPD $_{cv}$ (between 0.57 and 0.66) and the RPD $_{val}$ (between 1.46 and 1.75). The PLSR model of the Zn content prediction after applying the MSC preprocessing to the imagery data shows in particular, the important results with an R^2_{cv} equal to 0.24, an R^2_{val} equal to 0.5, an RMSECV at around 2.07, an RPD_{cv} equal to 0.66 and an RPD_{val} equal to 1.46.

According to statistical coefficients, less accurate prediction performances were however, showed for Pb and even more for Cd heavy metal. These results are explained by the using of spectrally and spatially OLI 8 degraded data compared to the results revealed by the PLSR regression that was carried out on VNIR-SWIR field spectra. Furthermore, through the comparison, it was proven that Zn was still the best well predicted heavy metal, despite both spectral and spatial degradation of the used OLI 8 data. Considering its heterogeneous distribution, the high concentrations of Zn, estimated around a mean of 2280 mg/kg and with a max of 80100 mg/kg, could justify its most accurate prediction performances. Moreover, the useful spectral band around 2202 nm, caused by the combination of OH–Al bending and OH stretching modes and related to the presence of clays (Al–OH), could increase the prediction performances of such heavy metals, known by their strong binding to these minerals. Thus, the present research have shown that the performance of heavy metal content prediction using Landsat OLI 8 multispectral images decreased in order of Zn > Pb > Cd, according to an $R^2_{cv} = 0.24$ and an $R^2_{val} = 0.5$ for the Zn model, an $R^2_{cv} = 0.32$ and an $R^2_{val} = 0.13$ for the Pb model and an $R^2_{cv} = 0.14$ and an $R^2_{val} = 0.09$ for the Cd model, despite that it was better for Cd versus Pb when using VNIR-SWIR field data.

In what follows and according to the very weak performance results of the Cd prediction model (Table 6), only both Zn and Pb heavy metals are considered for the Landsat OLI 8 multispectral image mapping.

3.4. Heavy metal content mapping

Three predicted content maps of each of Zn (Figure 8Ia, Ib and Ic) and Pb (figure 8IIa, 8IIb and 8IIc) heavy metals were produced by applying the PLSR models, using Landsat OLI 8 reflectance data before and after MSC and SNV preprocessing.

The map of Zn predicted contents, given the reflectance data before preprocessing, showed high image pixel values that correspond to the three tailing dumps with an amount higher than 68000 mg/kg. Other image pixels were classified as showing significant amounts (55000 mg/kg) within olive groves surrounding the source of pollution, with respect to topographic conditions in particular, the strong slope to the South as well as the prevailing wind (Nord-West). A critical situation of contaminated soils were thus revealed around mine tailings. The predicted maps of Zn contents, given the reflectance after MSC and SNV preprocessing, showed roughly the same Zn heavy metal distribution except for the third dump in the North of the region where the contamination was more distinguished with more image pixels revealing an amount around of 68000 mg/kg.

Several image pixels with high Pb contents were showed in the predicted map, given the reflectance data before preprocessing, particularly within the first and the second tailing dumps, with an amount around 20000 mg/kg. According to these mapping results, olive groves in front of the SW versant of the Jebel Ressas were less affected by Pb heavy metal contamination (an amount less than 15000 mg/kg) emphasizing particularly, the role of the strong slope. The predicted maps of Pb contents, given the MSC and SNV based preprocessed reflectance, showed however, more contrasted regions that revealed high Pb contaminant concentrations, especially in the north, within the third tailing dump as well

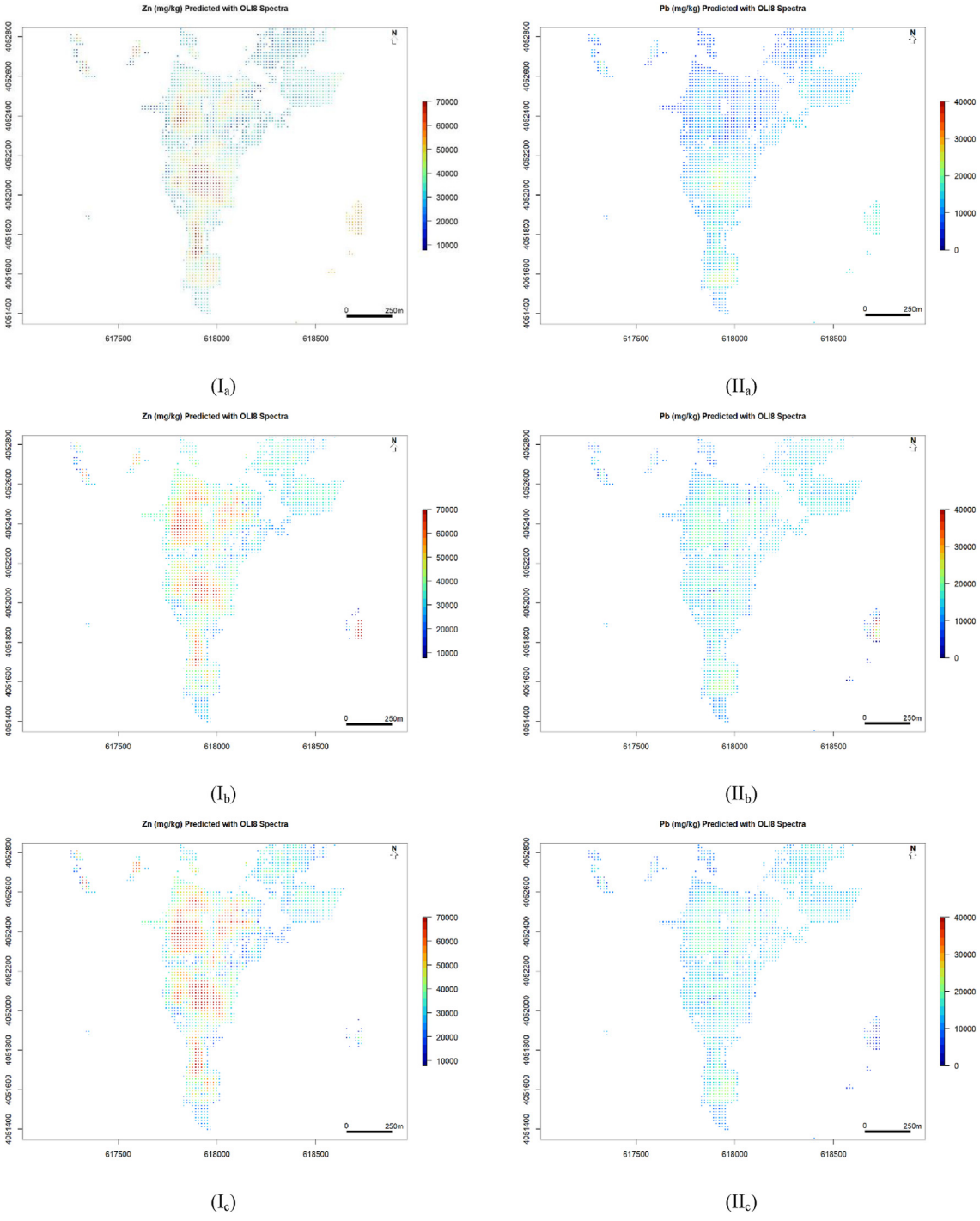

Figure 8. Maps of the predicted heavy metal contents (I) Zn and (II) Pb, which were derived by the PLSR method using Landsat OLI 8 data (30 m spatial resolution): (a) before preprocessing, (b) after preprocessing based on the MSC method and (c) after preprocessing based on the SNV method. The white color corresponds to masked vegetation using an NDVI> 0.3 and urban area.

as agricultural soils (olive groves). The predicted amount of Pb, estimating about 25000 mg/kg, is considered as an alarming value within soils that exceeding the mean value of the tolerated content.

3.5. Validation of heavy metal content maps

The accuracy of both heavy metal Zn and Pb maps was validated mainly by comparing the mapped results, using PLSR models based on Landsat OLI 8 imagery before and after MSC and SNV preprocessing, with the randomly test samples, out of the eighty-seven inputs (Figure 9).

Thus, predicted (Landsat OLI 8 map) and observed (test set) data, corresponding to the same sampling point, were compared for the evaluation of the degree of similarity. The Zn predicted and observed concentrations in the scatter plot show relative high convergence, especially when applied PLSR models, which using the preprocessed data given the MSC and SNV transformations, with an R^2 equal to 0.724 and 0.730, respectively. For Pb content maps, the comparison of predicted and observed (tested) concentrations reports R^2 values around 0.836 and 0.770, particularly using PLSR model, which consider preprocessed reflectances through the MSC and SNV transformations, respectively.

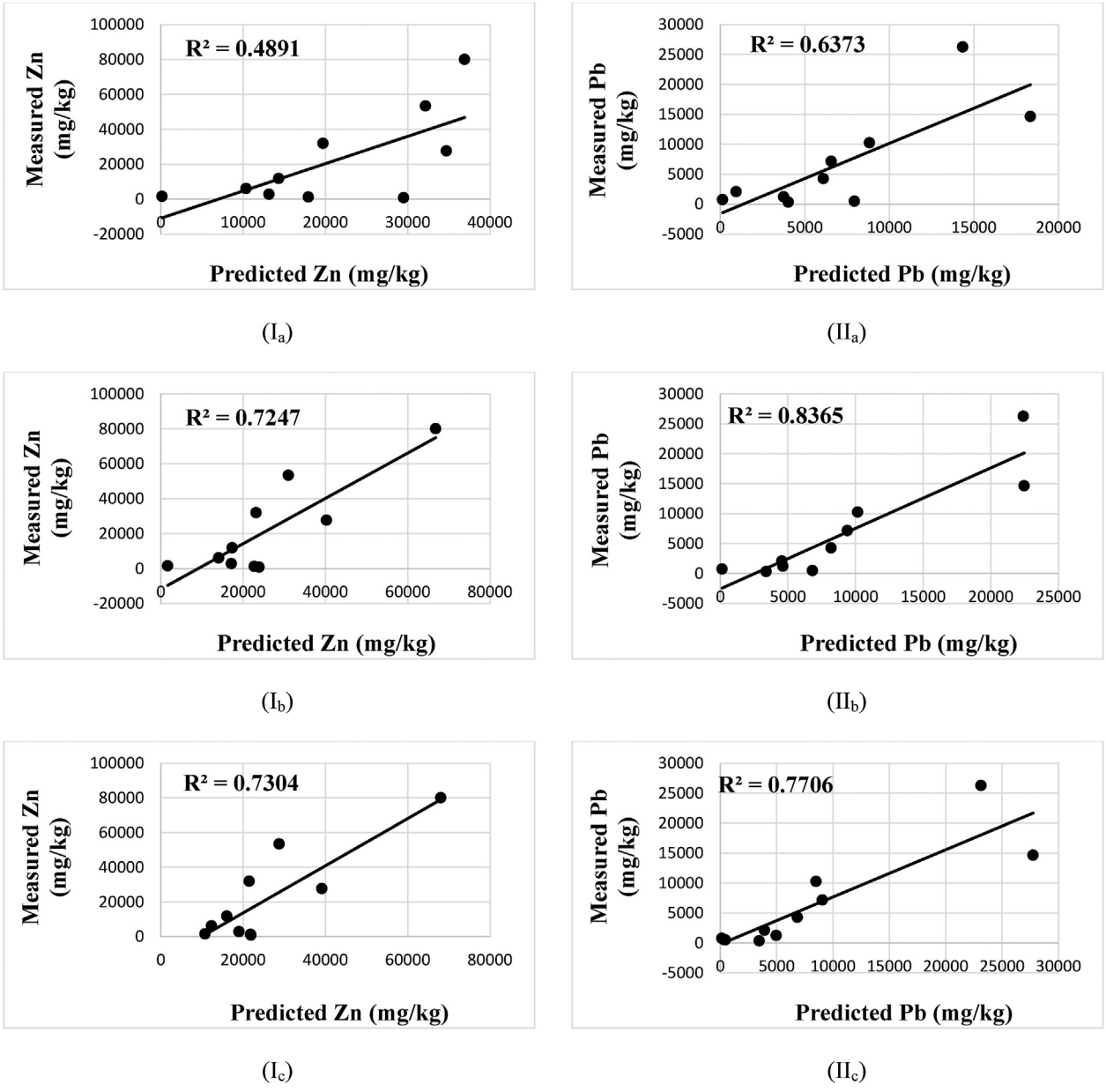

Figure 9. Scatter plots showing the comparison between the observed (Test set data) and the predicted (from maps) data derived from the PLSR models dedicated for the prediction of heavy metal concentrations (I) Zn and (II) Pb. The predicted data, which were used are from: (a) Landsat OLI 8 mapping, (b) Landsat OLI 8 mapping after preprocessing based on the MSC method and (c) Landsat OLI 8 mapping after preprocessing based on the SNV method.

These results showed that PLSR models could be applied for mapping Zn contents and even less for Pb contents within dumps and soils, using Landsat OLI 8 multispectral data. The PLSR performance results are still improved especially when conducting MSC and SNV preprocessed methods versus non-preprocessed reflectances.

4. Conclusions and perspectives

The present study aimed to use hyperspectral terrain spectroscopy and Landsat OLI 8 Multispectral data, before and after preprocessing, for the indirect content estimation and mapping of the Zn, Pb and Cd heavy metals for the first time around a particular Tunisian context of a mining site. The used approach, basing on PLSR regression method, highlights the usefulness of MSC and SNV preprocessing to improve prediction performances of heavy metal contents.

Results have shown that Zn was still the best well predicted heavy metal, using field data (with an $R^2_{cv} = 0.82$, an $R^2_{val} = 0.64$, an $RPD_{cv} = 2.29$, an $RPD_{val} = 1.71$) and despite both spectral and spatial degradation of the used OLI 8 data (with an $R^2_{cv} = 0.22$, an $R^2_{val} = 0.53$, an $RPD_{cv} = 0.66$, an $RPD_{val} = 1.51$). Furthermore, it was proven that although the performance of prediction using VNIR-SWIR field spectra is better and more accurate for Zn and Cd contents ($R^2_{cv} = 0.82$ and $R^2_{cv} = 0.49$) rather than for Pb contents ($R^2_{cv} = 0.38$). The prediction performance was however, decreased in the order of Zn > Pb > Cd using Landsat OLI 8 multispectral images ($R^2_{cv} = 0.24$ and $R^2_{val} = 0.5$; $R^2_{cv} = 0.32$ and $R^2_{val} = 0.13$; and $R^2_{cv} = 0.14$ and $R^2_{val} = 0.09$). The decrease of the prediction performance of Cd content has been related to the spectral and spatial degradation of the used OLI 8 image data. In almost all results, both MSC and SNV preprocessing methods improved the prediction performances of heavy metal contents of PLSR models, especially when using Landsat OLI 8 multispectral images. It has been shown that preprocessing methods could ameliorate the PLSR regression results in presence of degraded data spatially and spectrally.

Furthermore, the accuracy of the Zn and Pb concentration maps was validated through the comparison between extracted concentrations

from each heavy metal map, resulting from PLSR model processing, and measured concentrations of the ten randomly samples (test set out of the eighty-seven inputs), resulting from atomic absorption spectroscopic analysis. The calculated R^2 coefficients have shown that PLSR regression method using MSC and SNV based preprocessed Landsat OLI 8 data is promising for heavy metal content prediction and mapping, despite the spectral and spatial data degradation, especially for Zn heavy metal when applying both MSC and SNV preprocessing methods (with an $R^2 = 0.724$ and an $R^2 = 0.730$) and for Pb heavy metal, as well (with an $R^2 = 0.836$ and an $R^2 = 0.770$). The conducted approach revealed alarming amounts of Pb, Zn and Cd heavy metals within soils around the Jebel Ressas mine site, that exceeding the tolerated mean values. Their distribution, from dumps towards soils, is conditioned by the topographic conditions of the site as well as the erosive action of the prevailing wind.

A thorough study will be conducted in the future with the aim of further improving the mapping of heavy metals through the exploration of different multispectral image data, such as SENTINEL-2 and Aster data, which have been tested with success for the mapping of the contents of minerals with high pollution potential (Dkhala et al., 2020) and the mapping of phosphate mineralization (Mezned et al., 2020), respectively. Several preprocessing methods will be also tested to evaluate the prediction performances of different regression models.

Declarations

Author contribution statement

Nouha Mezned: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Faten Alayet: Performed the experiments; Analyzed and interpreted the data.

Belgacem Dkhala: Performed the experiments.

Saadi Abdeljaouad: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

The authors do not have permission to share data.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Alayet, F., Mezned, N., Sebai, A., Abedljaoued, S., 2017. Continuum removed band depth analysis for carbonate mining waste quantification using XRD and Hyperspectral spectroscopy in the north of Tunisia. J. Appl. Remote Sens. 11 (1), 016021. SPIE.

Al Maliki, A., Bruce, D., Owens, G., 2014. Prediction of lead concentration in soil using reflectance spectroscopy. Environ. Technol. Innovat. 1.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. 43 (5), 772–777. Research Article.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.M., McBratney, A., 2010. Prediction of soil attributes by NIR spectroscopy. A critical review of chemometric indicators commonly used for assessing the quality of the prediction. Trac-Trends Anal. Chem. 29, 1073–1081.

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh Jr., C.R., 2001. Near-infrared reflectance spectroscopy—principal components regression analysis of soil properties. Soil Sci. Soc. Am. J. 65, 480–490.

Chen, T., Chang, Q., Clevers, J.G.P.W., Kooistra, L., 2015. Rapid identification of soil cadmium pollution risk at regional scale based on visible and near-infrared spectroscopy. Environ. Pollut. 206, 217–226.

Choe, E., Meer, F., Ruitenbeek, F., Werff, H., Smeth, B., Kim, K.W., 2008. Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: a case study of the Rodalquilar mining area, SE Spain. Remote Sens. Environ. 112, 3222–3233.

Chong, I.G., Jun, C.H., 2005. Performance of some variable selection methods when multicollinearity is present. Chemometr. Intell. Lab. Syst. 78, 103–112.

Davey, M.W., Saeys, W., Hof, E., Ramon, H., Swennen, R.L., Keulemans, J., 2009. Application of visible and near-infrared reflectance spectroscopy 574 (vis/nirs) to determine carotenoid contents in banana (musa spp.) fruit 575 pulp. J. Agric. Food Chem. 57, 1742–1751.

Dkhala, B., Mezned, N., Gomez, C., Abdeljaouad, S., 2020. Hyperspectral field spectroscopy and SENTINEL-2 Multispectral data for minerals with high pollution potential content estimation and mapping. Sci. Total Environ. 740, 18 art. 140160; ISSN 0048-9697.

Douglas, R.K., Nawar, N., Alamar, M.C., Mouazen, A.M., Coulon, F., 2018. Rapid prediction of total petroleum hydrocarbons concentration in contaminated soil using vis-NIR spectroscopy and regression techniques. Sci. Total Environ. 616–617, 147–155. March 2018.

Fard, R.S., Matinfar, H.R., 2016. Capability of vis-NIR spectroscopy and Landsat 8 spectral data to predict soil heavy metals in polluted agricultural land (Iran). Arabian J. Geosci. 9 (20), 1–14.

Fang, Y., Xu, L., Peng, J., Wang, H., Wong, A., Clausi, D.A., 2018. Retrieval and mapping of heavy metal concentration in soil using time series Landsat 8 imagery. In: ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-3, pp. 335–340, 2018.

Fashola, M.O., Ngole-Jeme, V.M., Babalola1, O.O., 2016. Heavy metal pollution from gold mines: environmental effects and bacterial strategies for resistance. Int. J. Environ. Res. Publ. Health 13 (11), 1047.

Ghorbel, M., 2012. Contamination métallique issue des déchets de l'ancien site minier de Jbel Ressas. Modélisation des mécanismes de transfert et conception des cartes d'aléa post-mine dans un contexte carbonaté sous un climat semi-aride, évaluation du risque pour la santé humaine. université de Toulouse III, p. 196. Thèse de doctorat.

Goltappeh, F., Zeaiean Firouzabadi, P., bakhtyari, H.R.R., 2018. Evaluation of OLI sensor data, the capabilities of ALTA reflectance spectrometer and using the concept of virtual StationsMapping the distribution of heavy metals in soil. J. Geospat. Informat. Technol. 5 (4), 113–145.

Gomez, C., Lagacherie, P., Coulouma, G., 2008. Continum Removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectal measurements. Geoderma 141–148.

Henin, S., 1983. Les éléments traces dans le sol. Bulletin A.F.E.S. n°2, pp. 67–71.

Hou, L., Li, X., Li, F., 2019. Hyperspectral-based inversion of heavy metal content in the soil of coal mining areas. J. Environ. Qual. 48, 57–63, 2019.

Hribkou, K., Pätzold, S., Mahlein, A.-K., Welp, G., 2012. Airbone hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale. Geoderma 175–176, 21–28.

Hunt, G.R., Salisbury, J.W., 1971. Visible and near infrared spectra of minerals and rocks. II. Carbonates, Mod. Geology 2, 23–30.

Isaksson, T., Kowalski, B., 1993. Piece-wise multiplicative scatter correction applied to near-infrared diffuse transmittance data from meat products. Appl. Spectrosc. 47 (6), 702–709.

Jonsson, H., Gabrielsson, J., 2009. Comprehensive Chemometrics, Methods, Chemistry. Molecular Sciences and Chemical Engineering, pp. 199–206, 2009.

Kemper, T., Sommer, S., 2002. Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. Environ. Sci. Technol. 36, 2742–2747, 2002.

Khosravia, V., Ardejanib, F.D., Gholizadeh, A., Saberioon, M., 2021. Satellite Imagery for Monitoring and Mapping Soil Chromium Pollution in a Mine Waste Dump, 2021.

Kooistra, L., Wehrens, R., Leuven, R.S.E.W., 2001. Possibilities of visible-near-infrared spectroscopy for the assessment of soil contamination in river floodplains. Anal. Chim. Acta 446 (1), 97–105.

Kotz, S., Campbell, B., Balakrishnan, N., Vidakovic, B., 2006. Encyclopedia of Statistical Sciences, 16. Wiley.

Leone, A.,P., Viscarra Rosse, R.,A., Amenta, P., Buondonno, A., 2012. Prediction of soil properties with PLSR and vis-NIR spectroscopy: application to mediterranean soils from southern Italy. Curr. Anal. Chem. 8 (2), 283–299.

Lu, D., Li, G., Moran, E., Hetrick, S., 2013a. Vegetation change detection in the Brazilian Amazon with multitemporal Landsat images. Remote Sens. Nat. Resour. 127–140.

Lu, P., Wang, L., Niu, Z., Li, Linghao, Zhang, W., 2013b. Prediction of soil properties using laboratory VIS-NIR spectroscopy and Hyperion imagery. J. Geochem. Explor.

Mark, H.L., Tunnell, D., 1985. Qualitative near infrared reflectance analysis using Mahalanobis distances. Anal. Chem. 57 (7), 1449–1456.

Mezned, N., Fatnassi, A., Abdeljaouad, S., 2020. Potential of ASTER's VNIR, SWIR and TIR bands for phosphate exploration in th center-west of Tunisia. J. Appl. Remote Sens., Society of Photo-Optical Instrumentation Engineers SPIE, 014509.

Mohamed, E.S., El Baroudy, A.A., El-beshbeshy, T., Emam, M., Belal, A.A., Elfadaly, A., Aldosari, A.A., Abdelraouf, M.A., Lasaponara, R., 2020. Vis-NIR spectroscopy and satellite landsat-8 OLI data to map soil nutrients in arid conditions: a case study of the northwest coast of Egypt. Rem. Sens. 12, 3716, 2020.

Olarewaju, O.O., Bertling, I., Magwaza, L.S., 2016. Non-destructive 683 evaluation of avocado fruit maturity using near infrared spectroscopy and 684 pls regression models. Sci. Hortic. 199, 229–236.
Peng, Y., Bou Kheir, R., Adhikari, K., Malinowski, R., Greve, M.B., Knadel, M., Greve, M.H., 2016. Digital mapping of toxic metals in Qatari soils using remote sensing and ancillary data. Rem. Sens. 8, 1003, 2016.
Razo, I., Carrizales, L., Castro, J., Díaz-Barriga, F., Monroy, M., 2004. Arsenic and heavy metal pollution of soil, water and sediments in a semi-arid climate mining area in Mexico. Water Air, Soil Pollut. 152, 129.
Rossel, R.A. Viscarra, Walvoort, D.J.J., Mcbratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131 (1), 59e75.
Rossel, R.A.V., Webster, R., 2012. Predicting soil properties from the Australian soil visible-near infrared spectroscopic database. Eur. J. Soil Sci. 63, 848e860.
Sainfeld, P., 1952. Les gîtes plombo-zincifères de Tunisie. Annales des mines et de la géologie N°9.
Shi, Z., Wang, Q., Peng, J., Ji, W., Liu, H., Li, X., Viscarra Rossel, R., 2014. Development of national VNIR soil-spectral library for soil classification and the predictions of organic matter. Sci. China Earth Sci. 57, 1671–1680.
Stazi, S.R., Antonucci, F., Pallottino, F., Costa, C., Marabottini, R., Petruccioli, M., Menesatti, P., 2014. Hyperspectral visible–near infrared determination of arsenic concentration in soil. Commun. Soil Sci. Plant Anal. 45, 22.
Sun, W., Zhang, X., 2017. Estimating soil zinc concentrations using reflectance spectroscopy. Int. J. Appl. Earth Obs. Geoinf. 58, 126–133.
Sun, W., Zhang, X., Sun, X., Sun, Y., Cen, Y., 2018. Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals. Geoderma 327, 25–35.
Thompson, A.J.B., Hauff, P.L., Robitaille, A.J., 1999. Alteration mapping in exploration: application of short-wave infrared (SWIR) spectrocsopy. Econ. Geol. Newlsett. 39.
Todorova, M., Mouazen, A., Lange, H., Atanassova, S., 2014. Potential of near-infrared spectroscopy for measurement of heavy metals in soil as affected by calibration set size. Water Air Soil Pollut. 225.
Viscarra Rossel, R.A., Walvoort, D.J.J., McBrateney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131, 59–75.
Wang, J., Cui, L., Gao, W., Shi, T., Chen, Y., Gao, Y., 2014. Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy. Geoderma 216, 1–9.
Wei, L., Yuan, Z., Yu, M., Huang, C., Cao, L., 2019. Estimation of arsenic content in soil based on laboratory and field reflectance spectroscopy. Sensors, MDPI 19 (18), 3904.
Wold, S., 1978. Cross-validatory estimation of the number of components in factor and principal components models. Technometrics 20, 397–405.
Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of Chemometrics. Chemometr. Intell. Lab. Syst. 58, 109–130.
Wu, Y.Z., Chen, J., Wu, X., Tian, Q.J., Ji, J.F., Qin, Z., 2005. Possibilities of reflectance spectroscopy for the assessment of contaminant elements in suburban soils. Appl. Geochem. 20 (6), 1051–1059.
Zeaiter, M., Rutledge, D., 2009. Preprocessing Methods Comprehensive Chemometrics, pp. 121–231.
Zhang, X., Sun, W., Cen, Y., Zhang, L., Wang, N., 2019. Predicting cadmium concentration in soils using laboratory and field reflectance spectroscopy. Sci. Total Environ. 650, 321–334.
Zhiyuan, L., Zongwei, M., Tsering, J.K., Zengwei, Y., Lei, H., 2014. A review of soil heavy metal pollution from mines in China: pollution and health risk assessment. Sci. Total Environ. 468–469, 843–853.
Zhou, J., Dang, Z., Cai, M., Liu, C., 2007. Soil heavy metal pollution around the dabaoshan mine, guangdong Province, China. Pedosphere 17 (5), 588–594.