

Article

Digital Mapping of Toxic Metals in Qatari Soils Using Remote Sensing and Ancillary Data

Yi Peng ^{1,2,*}, Rania Bou Kheir ¹, Kabindra Adhikari ³, Radosław Malinowski ⁴, Mette B. Greve ¹, Maria Knadel ¹ and Mogens H. Greve ¹

¹ Department of Agroecology, Faculty of Science and Technology, Aarhus University, Blichers Allé 20, P.O. Box 50, DK-8830 Tjele, Denmark; Cruz311@hotmail.com (R.B.K.); metteb.greve@agro.au.dk (M.B.G.); maria.knadel@agro.au.dk (M.K.); mogensh.greve@agro.au.dk (M.H.G.)

² Department of Chemistry, College of resources and environment, Zunyi Normal College, Shanghai Road 830, Zunyi 563002, China

³ Department of Soil Science, FD Hole Soils Lab, University of Wisconsin-Madison, Madison, WI 53706, USA; kabadh1@gmail.com

⁴ Space Research Centre, Polish Academy of Sciences, Bartycka 18A, Warsaw 00-716, Poland; radekmali@yahoo.com

* Correspondence: yi.peng@agro.au.dk; Tel.: +45-8715-7415

Academic Editors: Lenio Soares Galvao, Clement Atzberger and Prasad S. Thenkabail

Received: 12 September 2016; Accepted: 28 November 2016; Published: 9 December 2016

Abstract: After decades of mining and industrialization in Qatar, it is important to estimate their impact on soil pollution with toxic metals. The study utilized 300 topsoil (0–30 cm) samples, multi-spectral images (Landsat 8), spectral indices and environmental variables to model and map the spatial distribution of arsenic (As), chromium (Cr), nickel (Ni), copper (Cu), lead (Pb) and zinc (Zn) in Qatari soils. The prediction model used condition-based rules generated in the Cubist tool. In terms of R^2 and the ratio of performance to interquartile distance (RPIQ), the models showed good predictive capabilities for all elements. Of all of the prediction results, Cu had the highest $R^2 = 0.74$, followed by $As > Pb > Cr > Zn > Ni$. This study found that all of the models only chose images from January and February as predictors, which indicates that images from these two months are important for soil toxic metals' monitoring in arid soils, due to the climate and the vegetation cover during this season. Topsoil maps of the six toxic metals were generated. The maps can be used to prioritize the choice of remediation measures and can be applied to other arid areas of similar environmental/socio-economic conditions and pollution causes.

Keywords: toxic metals; cubist; multi-spectral images; spectral indices; seasonal dependence; digital soil mapping

1. Introduction

Industrialization, rapid urbanization and military activities in particular geographic locations are sources of soil pollution with toxic metals in many parts of the world [1–6]. In urban and industrial areas, human exposure to enormous amounts of toxic metals, especially arsenic (As), chromium (Cr), nickel (Ni), copper (Cu), lead (Pb) and zinc (Zn), causes serious health disorders, including disruption of the nervous and digestive systems (e.g., acne, allergies, anemia, anorexia, arthritis, sciatica, schizophrenia), the normal functioning of internal organs (e.g., lungs, kidneys, liver), damage to the circulatory system (e.g., asthma and chronic bronchitis) and the blood composition and many other unfavorable consequences [7–11].

This situation is particularly true in the case of Qatar where the population has increased from about 0.74–1.7 million from 2000 to 2010. Such a rapid increase has been associated with the erection of residential buildings in the urban coastal area of the capital (Doha City), where around 85% of the

Qatari and non-Qatari citizens reside [12,13]. For the last two decades, Qatar has had the highest per-capita carbon dioxide emissions in the world, at 49.1 metric tons per person in 2008 [14]. In 2014, the World Health Organization ranked Qatar as the second most polluted (air pollution) country in the world.

For human and ecological risk assessment, the prediction of toxic metals accumulation in soils has been determined using a variety of geostatistical methods and environmental data at various scales [15–22]. Lado et al. [16] used regression-kriging with a large number of auxiliary variables (1-km resolution) as predictors (e.g., land cover, geology, night-lights images) to model and map the distribution of eight toxic metals (As, Cd, Cr, Cu, Hg, Ni, Pb, Zn) for the topsoil in Europe. The cross-validation results showed that the overall accuracy of these maps was relatively low. They also provided two possible explanations for the low accuracy: (1) regression-kriging and ordinary kriging are not capable of dealing with many hotspots (sporadic high values); and (2) a relatively coarse resolution of the predictors is a limitation to mapping accuracy. Xie et al. [23] applied four different interpolation methods (inverse distance weighting, local polynomial, ordinary kriging and radial basis functions) to map the spatial distribution of three toxic metals (Cd, Cu, Pb) in soil in a small area (605 km²) in Beijing. The results showed that all of the interpolation methods provided a high accuracy based on cross-validation. However, compared to traditional interpolation techniques, such as kriging or inverse distance weighting, recent studies have been using data-mining tools, e.g., Cubist (www.rulequest.com), in modeling soil and environmental properties and processes [24]. One of the several advantages of Cubist is that it is able to quantify linear, as well as non-linear relationships among variables and quantifies the importance of environmental variables used in prediction.

In the last ten years, the potential of using remote sensing (RS) techniques (aerial or satellite data) for mapping toxic metals in soils has received more attention. Melendez-Pastor et al. [25] comprehensively reviewed the role of RS in soil toxic metals' detection and found it to be an efficient tool for detecting sources of toxic metals' contamination in soils. This is despite the fact that toxic metals in soils appear in low or moderate concentrations (e.g., Cr and Cu <4000 mg·kg⁻¹) and are spectrally featureless, although they can easily bind to minerals and organic matter [26–28]. Especially the reflectance spectral information of minerals in soils plays an important role in predicting toxic metals because the spectral assignments' position of minerals can change with their chemical composition and surface activity [27,29–32]. This indirect mechanism allows the relationships between spectral information and soil toxic metals to be deduced since the information of Fe-oxides, clays and organic matter in soils or vegetation cover can be extracted from the RS spectra [33–35]. Therefore, RS technology is able to map the spatial distribution of toxic metals in soils. In Spain, Kemper T. and Sommer S. [36] successfully mapped residual toxic metals in soils after a mining accident using airborne hyperspectral data. The images were collected by a HyMap sensor in 128 wavebands over the range 403–2480 nm with a spectral resolution of 13–17 nm. When Choe et al. [32] derived parameters from ground spectra associated with toxic metals in soils, the results showed significant correlations between Pb, Zn, Cu and As and different spectral absorption features (e.g., peak depth, area, asymmetry and band ratio). Afterwards, they applied these parameters to hyperspectral images (HyMAP data) to map an area affected by toxic metals. The different gradient maps of spectral parameters showed similar spatial patterns to different, kriged soil toxic metals' concentration maps. Using multi-temporal observation from satellite images to map the large-scale spatial distribution of toxic metals in soils has, to our knowledge, never been tested on arid zones. Thus, the novelty of this work is using advanced data mining tools combined with multi-temporal satellite images for digital soil mapping in a desert zone and finding the seasonal dependency of multi-temporal satellite images on soil toxic metals' determination.

In this context, the objective of the study was to facilitate the mapping of the toxic metal accumulation on a large scale and to demonstrate the toxic metals' contamination of many years of mining and industrialization in Qatar. More specifically, the aim of this study was to use as predictors the multi-temporal images of a Landsat 8 sensor (covering a span of nine months), selected

spectral indices derived from the Landsat images and environmental variables to map the spatial distribution of six chosen toxic metals (As, Cr, Ni, Pb, Cu, Zn) in Qatar. As the potential public health risk is probably associated with the intake of these six elements, we focused our study on them. Our assumption for this approach was that spatial modelling with multi-spectral images integrated with other data sources (geopedological, geomorphology anthropogenic information) could provide important information on toxic metals' accumulation in soils.

2. Materials and Methods

2.1. Study Area

The State of Qatar is a peninsula reaching into the Arabian Gulf and located approximately between the latitudes of 20°10'N and 24°30'N and the longitudes of 50°45'E and 51°38'E. It is connected to The Kingdom of Saudi Arabia via a relatively narrow strip of land to the south. The country covers approximately 11,437 km² measuring approximately 180 km from north to south and 85 km from east to west. It has a maximum elevation of 103 m above sea level and a minimum of 6 m below sea level. Topographically, Qatar is a relatively flat to undulating desert plain [37].

Qatar lies in the desert zone and is thus characterized by a hot arid climate with rainfall occurring only during the winter season from November–March. The mean annual temperature is fairly uniform, ranging from 26 to 27 °C. Although the mean summer temperature ranges from 33 to 34 °C, there are days when the temperature exceeds 43 °C. The amount of annual precipitation can vary significantly, as is the case in most desert regions. Precipitation occurs as sporadic and at times very heavy showers of short duration and in localized areas, the mean annual rates ranging from 65 to more than 100 mm/year. In general, heavy showers are followed by long dry periods. Qatar is a rocky desert area with a dominance of lower to middle Eocene beds and north-south anticlinal axes (e.g., Dukhan anticline) that contain oil fields. Limestone and dolomitic limestone including partly dolomitic marl rocks of the Upper Dammam Formation (Middle Eocene) are the most widely distributed over Qatar (around 72%). A total of eight different soil types were identified on the available soil map according to the USDA Soil Taxonomy (Qatar National Atlas, 2006) [38], of which calcids and gypsisols cover >77% of Qatar. Calcisol dominated the center and southern part of Qatar, and Gypsisols are mainly found around the coast area (IUSS WG WRB, 2014) [39]. Depressions (around 850) covered by colluvial soils made up of calcareous loam, sandy loam and sandy clay loam are common, and they have accumulated to depths ranging from 30 to 150 cm [37].

The vegetation cover of arid zones such as Qatar is characterized by openness. It is composed of a permanent framework of perennials, and in a good rainy season, many perennials succeed in flowering and producing fruits in their first season. The plant communities can be divided into non-halophytic and halophytic [37].

2.2. Soil Sampling and Laboratory Analysis

A total of 300 soil samples were collected from the top 30-cm soil layer (A horizon) in polyethylene bags in spring of 2013 and 2014. The sampling strategy followed the European soil-sampling guidelines [40], and sampling locations covered all environmental hotspots in Qatar, including industrial areas, oil and gas sites, traffic networks, harbors and airports, residential areas, solid waste dumping sites and water treatment plants. The geographical coordinates of each sampling location were recorded using a GPS (global positioning system) with 5-m precision (Figure 1). Soil samples were air-dried for three days at room temperature and sieved using a 2-mm polyethylene sieve. They were subsequently analyzed for arsenic, chromium, nickel, zinc, copper and lead concentrations. Approximately 0.5 g of soil from each sample were placed in pre-cleaned Pyrex test tubes to which 8 mL of concentrated nitric acid and 3 mL of concentrated perchloric acid were added. The mixture was heated in an aluminum block at 200 °C for 3 h until it was completely dry. After the test tubes were cooled, 10 mL of 5% HNO₃ were added and heated at 70 °C for 1 h with occasional mixing.

Upon cooling, the mixture was decanted into polyethylene tubes and centrifuged at 3500 rpm for 10 min. All of the element concentrations of the solutions were measured using inductively-coupled plasma-atomic emission spectrometry (ICP-AES; Perkin-Elmer 3300 DV). More details about the laboratory methods can be found in ISO 13196 [41,42].

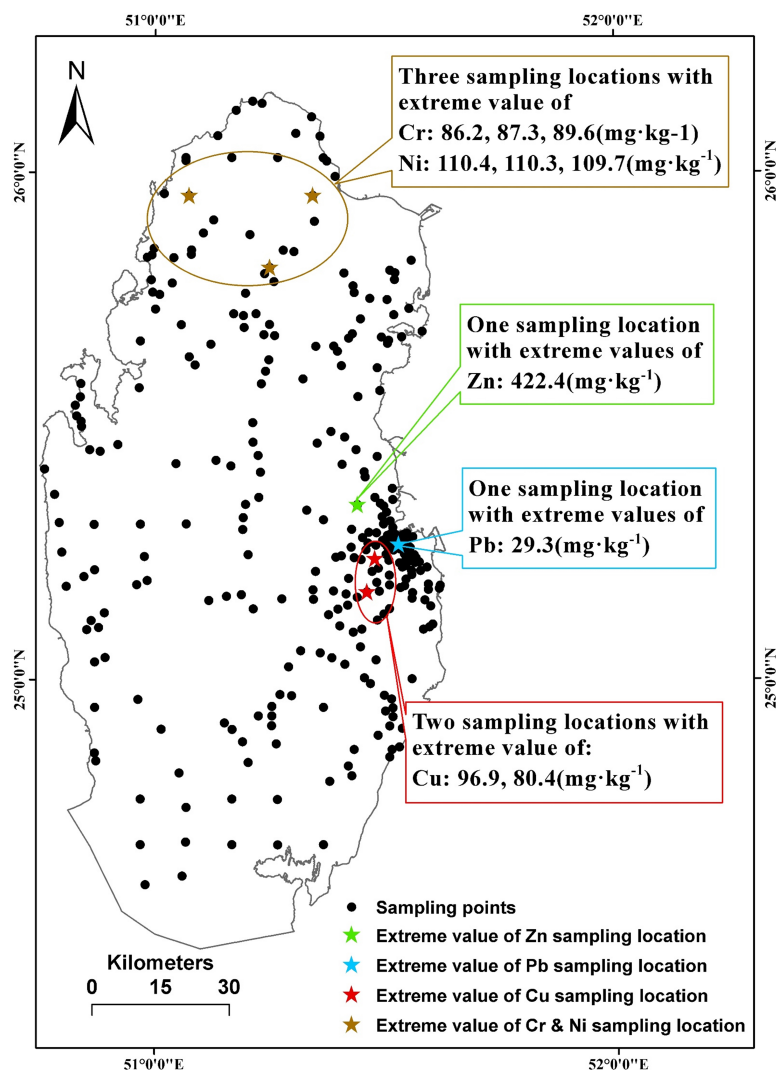


Figure 1. Sampling locations of the study area and locations with extreme values of different toxic metals in Qatari soils (the color of the text box correspond to the color assigned to the specific element).

2.3. Environmental Parameters

Environmental parameters used to predict the distribution of As, Cr, Ni, Pb, Cu and Zn included geopedological information (geology, distance to geological structures and soil type), geomorphology data (elevation, slope gradient, slope aspect, distance to drainage line) and anthropogenic information (proximity to roads, night light image and distance to environmental hotspots) (Table 1 and Figure 2). These data layers originate from various digital sources, including ancillary maps and the digital elevation model (DEM). The DEM was derived from light detection and ranging (LiDAR) technology where the LiDAR points were interpolated using the triangulation method and a fine-resolution DEM. The elevation map was resampled to a grid size of 30 m using simple mean aggregation. The selection of these parameters was based, in addition to our field observations, on their use in several predictive models of toxic metals and the variation in analogous (prevailing in arid environments) or similar environmental conditions [17,21].

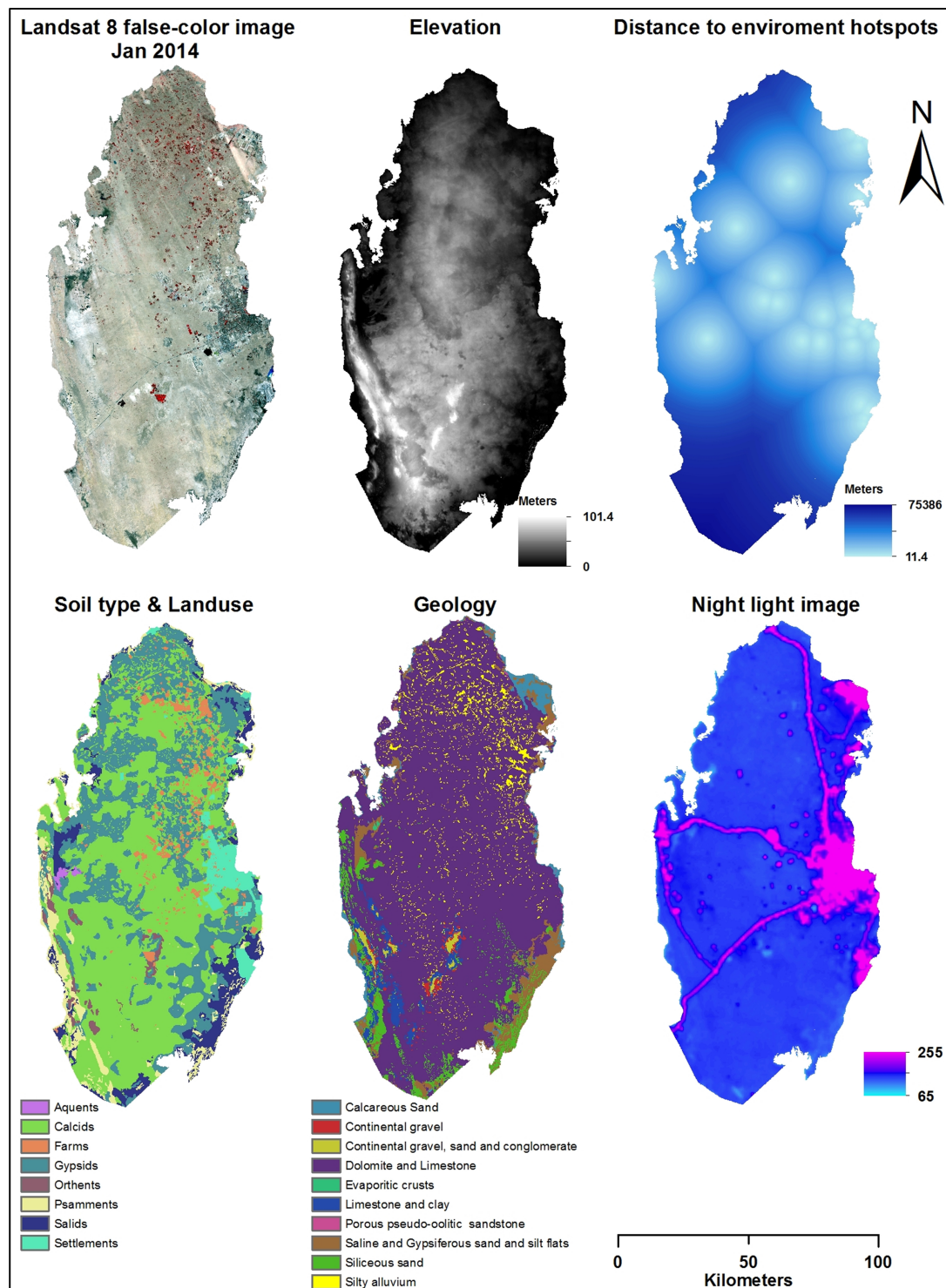


Figure 2. Selected auxiliary raster maps used as predictors for the distribution of different toxic metals in Qatari soils (25°28'61.06"N, 51°53'48.17"E).

Table 1. Environmental variables used to predict the distribution of different toxic metals in Qatari soils.

Environmental Variables	Type of Variable	Description	Range	Scale or Resolution	Mean	SD
Soil type map	Categorical	Scanned and registered pedological map (8 classes)	-	1:25,000	-	-
Geology	Categorical	Scanned and registered geological structures (10 classes)	-	1:36,000	-	-
Elevation (m)	Continuous	Derived from the 10-m resolution generated DEM (m)	0–101	30 m	26.2	18.0
Slope gradient (%)	Continuous	Maximum rate of change between the cells and neighbors (degree)	0–56	30 m	1.3	1.4
Slope aspect (degree)	Continuous	Direction of the steepest slope from the north (degree)	0–360	30 m	178.7	99.2
Proximity to road (m)	Continuous	The distance to road (m)	0–53,447	30 m	9554.4	9877.6
Distance to drainage line (m)	Continuous	Interpolating contour line and point elevation data (m)	0–43,930	30 m	5506.7	7270.4
Distance to environmental hotspots (m)	Continuous	Hotspots (i.e., wastewater treatment plants, solid waste dumping sites) (m)	0–86,029	30 m	24,947.3	18,561.1
Distance to geological structures (m)	Continuous	Derived from geological map (m)	0–36,519	30 m	6552.4	6800.2
Night light image	Continuous	Earth at night 2012	65–255	30 m	127.8	27.3

The scanned geological map of Qatar at the 1:36,000 scale [38] was used to delineate various lithological Qatari formations together with the geological structures. Limestone and dolomitic limestone including dolomitic marl of the Upper Dammam Formation (Middle Eocene) are most widely distributed over Qatar (around 72%) with north-south anticlinal axes (e.g., Dukhan anticline) that contain oil fields. The circumference of the delineated geological structures varies between 10 and 22,422 m. The pedological map of Qatar at the 1:25,000 scale (Qatar National Atlas, 2006) [38] was used to identify the eight different soil types.

Terrain attributes (i.e., elevation, terrain slope gradient and slope aspect) and drainage networks were derived from the 10-m resolution DEM. The latter was constructed through interpolating the contour lines using triangulation. Terrain slope ranged from 0–56°, and the slope aspect varied between 0 and 360° (distributed in nine direction classes). The distance to the drainage line varied in the studied region (i.e., urbanized and industrialized areas of Qatar) between 0 and 24,053 m (Table 1).

The proximity to roads was derived as the Euclidean distance for each raster cell to the closest road section and was between 0 and 36,804 m (Table 1). The distance to the closest environmental hotspot (i.e., wastewater treatment plants, solid waste dumping sites, desalination plants, sea harbors, airports, main roads and industrialization) ranged between 0 and 75,352 m. A night light image was created with data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS), was acquired in April 2012 and downloaded from the NASA Earth Observatory and NOAA National Geophysical Data Center website (<http://earthobservatory.nasa.gov/Features/NightLights/page3.php>). The image value ranged between 65 and 255, the high value (brightest) areas of the image being the most urbanized, but not necessarily the most populated, since highway systems appear just as bright as urban areas (Figure 2).

All of the environmental predictors were resampled to the 30-m pixel size using ‘bilinear resampling’ for continuous data and ‘nearest’ for categorical data in ArcGIS10.3.

2.4. Remote Sensing Images and Spectral Indices

We used atmospherically-corrected imagery of the Landsat Data Continuity Mission (Landsat 8) sensor (Table 2 and Figure 2) provided by the U.S. Geological Survey (USGS2014), which represents land surface reflectance. All images (30-m spatial resolution) were projected using the WGS 1984/UTM zone 39N map projection. Due to the issue of cloud cover, a set of images was selected representing only nine months of the year 2014 (January, February, April, May, Jun, July, August, September, October). In order to improve the predictive capability for toxic metals in soils and as we assumed that vegetation cover may provide important information for soil toxic metals’ spatial modeling, the following spectral indices were derived from the satellite images: BCI, biophysical composition index [43]; EVI, enhanced vegetation index [44]; LSWI, land surface water index [45]; NDVI, normalized differential vegetation index [46]; SATVI, soil-adjusted total vegetation index [47,48]; SAVI, soil-adjusted vegetation index [49]; TVI, transformed vegetation index [50]; WDV, weighted difference vegetation index [51,52]; and tasseled cap transformation including brightness, greenness and wetness [53]. The formulas used to derive these indices are shown in Table 2. ArcGIS was used to extract the values of predictors at the sampling locations.

Table 2. Description of the Landsat 8 image and derived image indices used to predict the distribution of different toxic metals in Qatar.

Landsat 8 Images and Indices ^a		Description
B2		Blue (450–510 nm)
B3		Green (530–590 nm)
B4		Red (640–670 nm)
B5		Near Infrared (NIR) (850–880 nm)
B6		SWIR1 (1570–1650 nm)
B7		SWIR2 (2110–2290 nm)
BCI ^b		$\frac{(H+1)/(2-V)}{(H+1)/(2+V)}$
Brightness	$0.3561 \times B2 + 0.3972 \times B3 + 0.3904 \times B4 + 0.6966 \times B5 + B6 \times 0.2286 + B7 \times 0.1596$	
EVI ^c		$G \frac{B5-B4}{B5+C_1B4-C_2B2+L} (1+L)$
Greenness	$-0.334 \times B2 - 0.354 \times B3 - 0.456 \times B4 + 0.6966 \times B5 - B6 \times 0.024 - B7 \times 0.263$	
LSWI		$\frac{B5-B6}{B5+B6}$
NDVI		$\frac{B5-B4}{B5+B4}$
SATVI ^c		$\frac{(B6-B4)}{(B6+B4+L)} \times (1+L) - \frac{B7}{2}$
SAVI ^c		$\frac{(B5-B4) \times (1+L)}{B5+B4+L}$
TVI		$\left(\frac{B5-B4}{B5+B4} + 0.5 \right)^{1/2} \times 100$
WDVI		$B5 - s \times B4$
Wetness	$0.2626 \times B2 + 0.2141 \times B3 + 0.0926 \times B4 + 0.0656 \times B5 - B6 \times 0.763 - B7 \times 0.539$	

^a BCI, biophysical composition index; EVI, enhanced vegetation index; LSWI, land surface water index; NDVI, normalized differential vegetation index; SATVI, soil adjusted total vegetation index; SAVI, soil-adjusted vegetation index; TVI, transformed vegetation index; WDVI, weighted difference vegetation index; ^b $H = (\text{brightness} - \text{brightness min}) / (\text{brightness max} - \text{brightness min})$; $V = (\text{greenness} - \text{greenness min}) / (\text{greenness max} - \text{greenness min})$; $I = (\text{wetness} - \text{wetness min}) / (\text{wetness max} - \text{wetness min})$; ^c $C1 = 6$; $C2 = 7.5$; $G = 2.5$; $L = 1$, $s = 1$.

2.5. Spatial Modeling and Validation

To calibrate the prediction model, the observed toxic metals data were randomly split into 75% for calibration and 25% for validation. A prediction model for each element was developed using the Cubist tool executed in R with the Cubist package (Version: 0.0.18) [54]. Cubist is a powerful data-mining tool that has been widely used in spatial modelling and mapping of a range of soil attributes in recent years [35,55–57]. The algorithm has been developed from an earlier version of C4.5 and the M5 model tree. In a conceptual Cubist model, a tree is grown where the terminal leaves contain linear regression models for prediction. There are intermediate linear models at each step of the tree, and all models are based on the variables used in previous splits. The tree is finally reduced to a set of rules that initially are paths from the top of the tree to the bottom, and the linear model is then adjusted and simplified to reduce the absolute error. Additionally, it also provides the attribute usage of each variable used in the model indicating the importance of the variables and is represented as the relative importance in percentage. The details of the Cubist algorithm can be found in Quinlan [58]. To find the optimal number of parameters during model calibration, the “trainControl” function (R package: Caret (Version: 6.0-52) [59]) with 10 iterations each with 10-fold cross-validation was applied. The final decision regarding which model to use was based on the lowest root mean square error of cross-validation.

The prediction models were evaluated using the following parameters: the coefficient of determination (R^2), root mean square error (RMSE), bias, the ratio of performance to deviation (RPD) and the ratio of performance to interquartile distance (RPIQ) [60,61].

3. Results and Discussion

3.1. General Statistics of Toxic Metals Data

A histogram and a box-plot of the elements (Figure 3) showed that the distribution of the elements is positively skewed in all cases. A wide range of variation was observed for all of the targeting elements, especially a large variation of the Zn, Cr and Ni was observed. Some of the soil samples from a few locations had extremely high values of different metals (e.g., Cr > 86 mg·kg⁻¹, Ni > 109 mg·kg⁻¹, Zn > 422 mg·kg⁻¹, Pb > 29 mg·kg⁻¹ and Cu > 80 mg·kg⁻¹). The geographical locations of those samples are shown in Figure 1. These extreme values as outliers and these sampling points most likely represent hotspots. We decided to exclude those samples because the extreme values or the local hotspots might influence the prediction based on the regression modeling [16].

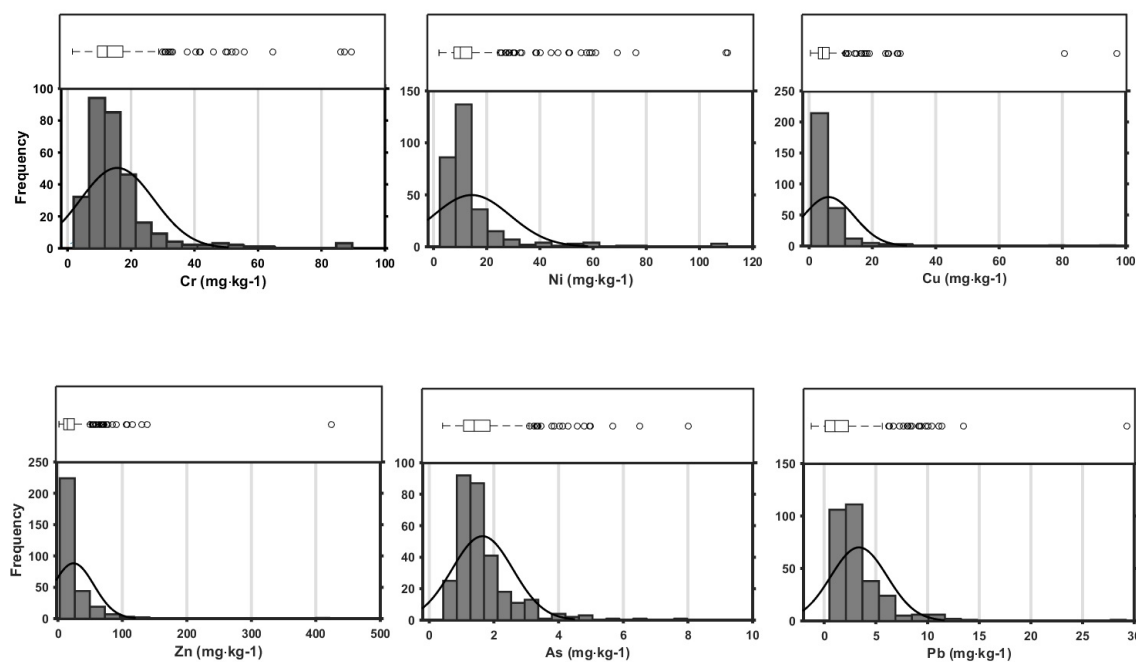


Figure 3. Histograms and box-plots of different toxic metals concentrations in the soils under study (No. of samples = 300).

The significant variability in toxic metal content is related to the different intensity levels of agriculture, urbanization and industrialization in the different areas of Qatar. Hotspots of Pb, Zn and Cu were, for example, found around the Doha City area (Figure 1), indicating that the soil contamination of these three elements was most likely affected by human activities [16,62–64]. Hotspots of Cr and Ni soils were found in the same three locations, and all located in the northern part of Qatar (Figure 1), where the parent material is partly silty alluvium and land use is dominated by irrigated agriculture. This suggests that the Cr and Ni elements in soils can be linked to various processes of soil formation, including human interference [16,62–64]. This variability in distribution affects the development of correlations between environmental parameters (natural and anthropogenic) and toxic metal pollution.

General statistics for calibration and validation sample sets are shown in Table 3. After excluding the extreme values for each element, we had 299 samples for Zn, 298 samples for Cu, 297 samples for Cr and Ni, 299 samples for Pb and 300 samples for As for modeling. Ranges in contents (mg·kg⁻¹) were: As from 0.44 to 7.99, Cr from 1.91 to 64.93, Zn from 2.89 to 130.9, Cu from 0.65 to 28.85, Ni from 0.44 to 7.99 and Pb from 0.5 to 14.1 mg·kg⁻¹ (Table 3).

Table 3. Soil toxic metals' concentrations in the calibration and validation sets.

Element (mg·kg ^{−1})	Calibration Set (75% Data)						Validation Set (25% Data)					
	n	Min	Median	Mean	Max	SD	n	Min	Median	Mean	Max	SD
Zn	224	2.8	16.4	23.5	130.9	21.6	75	4.0	14.5	20.9	77.9	17.3
Cu	223	0.6	4.4	5.8	28.8	4.6	75	1.0	4.2	5.0	28.1	3.9
Cr	223	1.9	12.7	14.9	64.9	9.2	74	3.5	12.9	14.8	50.1	8.6
Pb	224	0.5	2.7	3.4	14.1	2.3	75	0.5	2.3	2.8	12.4	2.1
Ni	223	2.3	10.1	13.4	76.1	10.8	74	3.0	10.7	13.5	55.5	10.6
As	225	0.4	1.3	1.6	7.9	0.9	75	0.4	1.5	1.7	5.0	0.8

n: number of samples; As: arsenic; Cr: chromium; Ni: nickel; Cu: copper; Pb: lead; Zn: zinc.

Table 4 shows the Pearson correlation coefficients between different toxic metals in Qatari soils. A significant correlation was found between Cr and Ni ($r = 0.95$), Cr and As ($r = 0.73$) and Ni and As ($r = 0.71$). The results showed that these three elements can be grouped together and that soil contamination of these three elements in Qatar might have similar levels and sources.

Table 4. Pearson correlation coefficients between different toxic metals in Qatari soils ($n = 300$).

	Cr	Ni	As	Zn	Pb	Cu
Cr	1	0.95 **	0.73 **	0.27	0.52 *	0.41 *
Ni		1	0.71 **	0.29	0.51 *	0.40 *
As			1	0.14	0.40 *	0.28
Zn				1	0.38 *	0.39 *
Pb					1	0.52 **
Cu						1

Levels of significance: * $p < 0.05$; ** $p < 0.01$.

3.2. Spatial Calibration and Validation

For the calibration, the Cubist model produces many linear equations for the target element based on different rules. To set the initial rule conditions for modeling Cr and Ni, for example, over 60% of the models selected soil type and geology map. For linear equation building, it is not necessary to include all of the predictors for each linear equation. Here, we show an example of the rule used for Ni predictions in the model:

Rule 1: (126 cases, mean 11.4, range 3.0–33.4, estimation error 3.9)

If:

Soil type in (1, 2, 3, 5) (the numbers indicate different soil types in Qatar)

Jan_Band2 \leq 0.1645

Then:

$$\text{Ni} = 6.8 - 13,238 \times \text{wetness (Jan)} - 10,271 \times \text{Band 6 (Jan)} - 7254 \times \text{Band 7 (Jan)} + 8651 \times \text{BCI (Jan)} - 2203 \times \text{Band 4 (Jan)} + 929 \times \text{Band 5 (Jan)} + 1305 \times \text{Band 2 (Jan)} - 379 \times \text{Band 4 (Feb)} - 255 \times \text{Wetness (Feb)} + 371 \times \text{Band 3 (Feb)}.$$

The results of the spatial prediction of the different toxic metals in Qatari soils are shown in Table 5. In general, all calibration models showed good predictive capabilities for the metals based on RMSE, R^2 , RPD and RPIQ. Of all of the prediction results, Cu had the highest R^2 and RPD values (0.74 and 1.88, respectively), followed by As > Pb > Cr > Zn > Ni. In this work, Pb and Ni gave the lowest RPD values at 1.37 and 1.36, respectively. However, their RPIQ values of, respectively, 1.1 and 1.0 were higher than for Zn (0.8). This was mainly because the calculation of the RPIQ value relies on a skewed data distribution [60]. Since all datasets in this study presented a highly skewed distribution, the RPIQ value reflected much more prediction accuracy information than the RPD value described.

Table 5. Independent validation results for different toxic metals' concentrations.

Element (mg·kg ^{−1})	Results				
	RMSE	Bias	R ²	RPD	RPIQ
Zn	12.18	0.8	0.51	1.41	0.83
Cu	2.10	0.4	0.74	1.88	1.37
Cr	5.90	0.6	0.53	1.45	1.54
Pb	1.49	0.2	0.53	1.37	1.10
Ni	5.19	0.7	0.47	1.36	1.06
As	0.58	0.05	0.60	1.53	1.48

Note: RMSE: root mean square error, RPD: ratio of performance to deviation, RPIQ: ratio of performance to interquartile distance.

Compared with previous studies, our models showed better predictive ability. Our study area was much smaller, and the resolution of all of our environmental predictors, including RS images, was finer (30 m) than in the European soil toxic metals maps (1-km resolution) made by Lado et al. [16]. This means that our models were carried out with much more accurate input data in terms of image spatial resolution. On the other hand, the climate environmental conditions between Qatar and Europe are completely different: Qatar is a desert zone; the soil parent material is also not as diverse as European soils. Therefore, the use of predictors for spatial modelling in both study areas could be different, making a comparison of modeling toxic metals in Europe and Qatar not straightforward. The variation in the European soil toxic metals database is also much higher than in the Qatar database. For example, in the European database, the median values of Zn, Cu, Cr, Pb, Ni and As were 52, 14, 22, 16, 16 and 6 mg·kg^{−1}, respectively, but the median values of the same elements in the Qatari database were 16, 4, 13, 2, 10 and 1 mg·kg^{−1}. The median value of As in the European database was almost five-times larger than in the Qatari database. In this situation, an empirical modeling approach would perform better within a less variable dataset than in a large dataset, like the European database. In the present study, very few samples of extreme values were excluded from modelling since we could not find any hard evidence to prove that these sampling sites were heavily contaminated by any heavy metals. We considered that these samples with extreme values were due to sampling errors.

Kemper T. and Sommer S. [36] validated Pb and As concentrations in soils after a mining accident in Spain using HyMap data in an area measuring a few thousand hectares. Their validation results showed a stronger coefficient for Pb ($R^2 = 0.85$) and As ($R^2 = 0.73$) than in our results because of their study area's very high contamination with these two elements. Therefore, the hyperspectral image was able to catch the spectral feature of the toxic metals in the soil directly from the soils. Furthermore, HyMap data provide 128 wavebands over the 403–2480-nm range with a spectral resolution of 13–17 m and a 5-m spatial resolution, which have improved the accuracy of input data for their modeling.

3.3. Variable Importance of Cubist Models for Predicting Different Elements

From the Cubist model output (usage of each variable) for the target elements, we can see that the models did not use all of the variables for calibration. For example, only RS data from January and February were chosen for all of the target elements modeled. Table 6 shows the top 10 predictors selected by the Cubist models for various toxic metals and their attribute usage ranking. In general, most of the reflectance spectral bands from January provided very high attribute usage for all of the models. Some of the spectral indices from January and February also provided high attribute usage for different elements, such as tasseled cap transformation and BCI. Furthermore, some of the geopedological, geomorphology and anthropogenic parameters were selected as important variables. For example, distance to geological structures contributed 25% of the usage for modeling Cr; the importance of distance to the drainage line for modeling Ni ranked as number 10; for the modeling of Pb, 46% and 56% of the linear models selected proximity to road and night light image as the predictor, respectively.

Table 6. Top 10 predictors selected by the Cubist calibration models for different toxic metals and their attribute usage (in percentage) ranking.

Cr	As	Ni	Zn	Pb	Cu
Jan_Band2 (84%)	Jan_Band2 (90%)	Jan_Band2 (86%)	Jan_Band5 (81%)	Feb_Band7 (88%)	Jan_Band4 (82%)
Jan_Band4 (79%)	Jan_Band4 (86%)	Jan_Band5 (81%)	Feb_Band4 (73%)	Jan_Band7 (85%)	Jan_Band2 (76%)
Feb_Band7 (73%)	Jan_Band7 (82%)	Jan_Band3 (72%)	Feb_Band3 (66%)	Feb_Band3 (80%)	Jan_Brightness (73%)
Jan_Band3 (69%)	Jan_Band6 (75%)	Feb_Band4 (70%)	Feb_Band6 (61%)	Jan_Band3 (77%)	Jan_Band5 (66%)
Jan_Band7 (69%)	Jan_Greenness (68%)	Feb_Band7 (67%)	Jan_Band5 (55%)	Feb_Band6 (70%)	Jan_Band7 (52%)
Jan_Band5 (51%)	Jan_Brightness (60%)	Jan_TV1 (58%)	Night light Image (53%)	Feb_Band4 (65%)	Jan_Band6 (52%)
Jan_LSWI (45%)	Jan_Wetness (58%)	Jan_BCI (53%)	Jan_Wetness (50%)	Jan_BCI (62%)	Jan_Band3 (45%)
Jan_BCI (41%)	Jan_EVI (45%)	Slope (42%)	Feb_BCI (45%)	Night light image (56%)	Distance to environmental hotspots (38%)
Jan_WDVI (33%)	Jan_SAVI (40%)	Distance to geological structures (37%)	Jan_Greenness (42%)	Proximity to road (42%)	Night light image (25%)
Distance to geological structures (25%)	Jan_NDVI (38%)	Distance to drainage line (28%)	Distance to environmental hotspots (35%)	Feb_Greenness (31%)	Proximity to road (20%)

From the relative importance of the predictors, we can see that Cr and Ni were highly affected by geopedological factors. Three geopedological predictors were included in the top 10 predictors for the modeling of these two elements, and two geopedological factors (soil type and geology) were selected as the most frequent rule-setting predictors for modeling Cr and Ni. On the other hand, for modeling Zn, Pb and Cu, anthropogenic parameters, such as proximity to road, night light image and distance to environmental hotspots, were of high relative importance. For example, in the Pb modelling process, the predictor “proximity of roads” showed a 42% relative importance value; this means that this variable matched almost half of the Cubist models. Besides RS data, “proximity of roads” was the second most contributed ancillary data after “night light image”. This indicates that soil toxic metals’ contamination with these three elements is highly affected by human activities. These results are much in line with previous studies [16,62–64].

3.4. Seasonal Dependency on Soil Toxic Metals’ Determination

This study found that of all of the RS images from eight different months, only images from January and February were selected in the modeling of each target element. To our knowledge, this study is the first that used RS data covering all of the seasons during a year to spatially model soil toxic metals’ concentrations. The parent material in Qatar soils is dominated by dolomite and limestone and normally has very low organic matter, clay and iron contents [38], and the toxic metals’ concentrations in Qatar soils are generally low. Therefore, it would be difficult to capture spectral features of toxic metals directly from the soils, even though most of the soil in Qatar is bare [26,27,29,32,36,65].

We suggest two possible explanations of why winter images were very important for the determination of soil toxic metals’ accumulation in Qatari soils. Both explanations are mainly based on the fact that RS images indirectly obtain soil toxic metals’ information from the vegetation cover. Figure 4 presents extracts of Landsat 8 images from different seasons in 2014 for a small area located to the north of Qatar. We can clearly see that the amount of vegetation or plants decreased dramatically after the winter season, and almost no vegetation cover was present during the summer, even for irrigated agricultural areas. This phenomenon is mainly due to the very limited precipitation and very high temperature after winter.

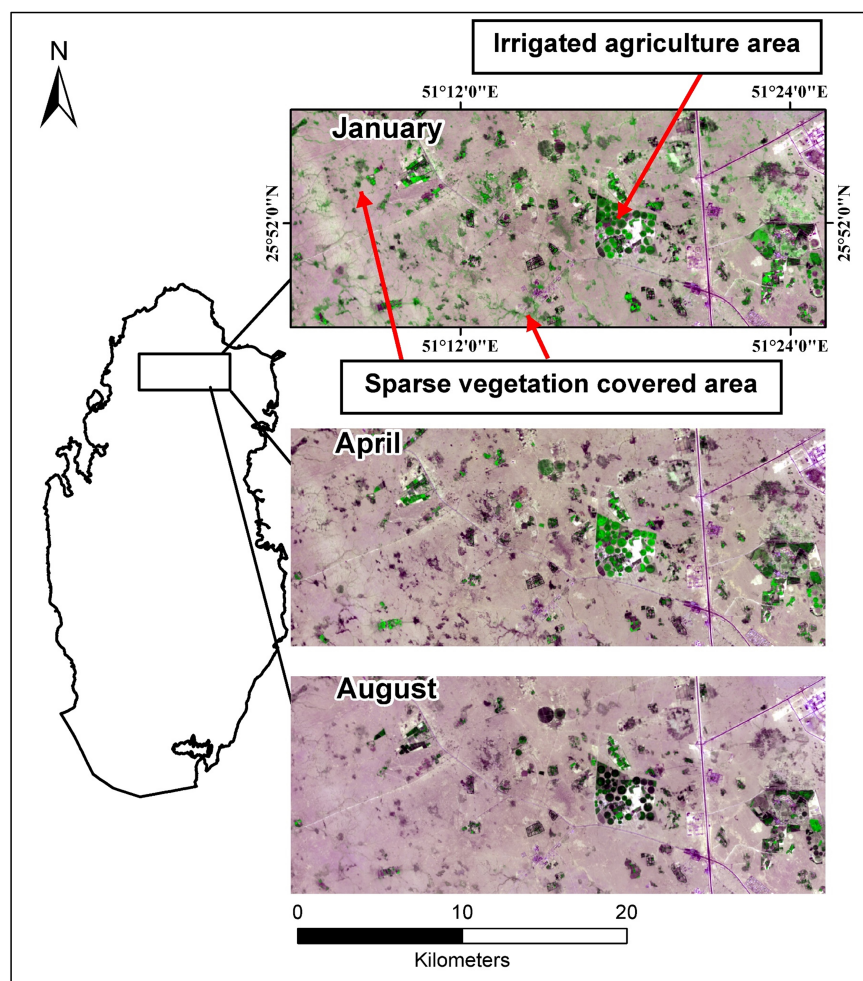


Figure 4. Extracted Landsat 8 images with red (Band 4), NIR (Band 5) and blue (Band 2) as RGB images from different seasons in 2014.

Firstly, Qatar has a relatively flat to undulating desert landscape with gently varying topography. In winter, rainwater transports some of the soil from more elevated to low-lying areas where more vegetation grows. Flooding may also occur in Qatar during heavy rainfall, and water can temporarily accumulate in lowlands. This may lead to the movement and accumulation of various toxic metals with mineral particles to lowland areas. Therefore, we observed relatively higher toxic metals' concentrations in the soils on vegetated lowland areas.

Secondly, the local climate of the study area needs to be considered. Because the rainfall occurs in the winter season and the mean temperature is around 25 °C, the conditions are best at this time of the year for germinating and growing vegetation. Some toxic metals in soil can apparently affect the germination process or plant growth. Toxic metals in high concentrations have been reported to progressively inhibit the early phases of seed germination. This also affects the reactivation of oxygen, changes in metabolic energy and the mobilization of sugars and phosphoro-organic compounds, along with changes in protein synthesis in seeds during germination [66–68]. All of these toxic metals at high concentrations in soil can inhibit germination, plant growth, chlorophyll production, reduce the germination rate and suppress yield, probably due to the toxic effects of various metabolic processes [69,70]. Different plant and vegetation phenotyping is detectable by RS data, for example the comparison between reflectance in the NIR band and the red band enable an analysis of vegetation vigor and biomass based on the chlorophyll concentration [71,72]. Therefore, the raw bands (reflectance values) are very important attributes in the modeling process (Table 6).

3.5. RS Indices for Toxic Metals' Prediction

Some of the RS spectral indices also had a high attribute usage for modeling the target elements (Table 6). The tasseled cap transformation and BCI were the most frequently-used indices. The tasseled cap transformation ranked in the top 10 predictors for As, Zn, Pb and Cu, while the BCI index ranked in the top 10 predictors for Cr, Ni, Zn and Pb.

The tasseled cap transformation is a global vegetation index. It can be applied for disaggregation of the amount of soil brightness, vegetation and moisture content in each pixel of the Landsat data. Urbanized areas are particularly evident in the brightness image; the greater the biomass, the brighter the pixel value in the greenness image; the wetness image is related to soil features, providing information of the moisture status in the observed area. Therefore, these indices can be very good indicators for vegetation status, classification of urban area and observing soil moisture content. The advantage of using the BCI index (the development of which is based on the tasseled cap transformation) is the more accurate prediction of vegetation abundance by the ability through enhanced contrast to separate impervious surfaces from soils and vegetation [43]. In this study, we could not find any accurate land use map as the environmental parameter for spatial modelling. Therefore, this index not only provides information of vegetation cover, but also identifies urban areas (impervious surfaces).

The indices NDVI, EVI and SAVI ranked in the top 10 variables for the modeling of As only. This can be explained by the major sources of As contamination in soils being pesticides and fertilizers from agriculture [63,73]. These three indices indicate changes in the amount of green biomass. In agriculture, heavy applications of pesticides and fertilizers produce greener biomass, leaving a high concentration of As in soils. We also found a Pearson correlation coefficient value between As and each of the three indices of over 0.3 and levels of significance: $p < 0.05$. Thus, our results suggest a strong positive relationship between the level of As contamination in soils and the amount of green biomass produced on agricultural land in Qatar.

3.6. Predicted Maps of Toxic Metals in Soil

Six predicted maps of toxic metals accumulations are shown in Figure 5. Based on these maps, we can easily identify relationships between chaotic urbanization and rapid industrialization and soil accumulation of toxic metals, as confirmed also in other studies [2,10,74]. All of the predicted maps showed similar spatial distributions, such as a higher toxic metals' accumulation in the northern part of Qatar than in the south. This can be explained by anthropic activity and geopedological impact. First, most of the agriculture and industry is located in the northern part of the country (Figure 2). Second, the parent material in this area is partly silty alluvium where the soil has been eroded and reshaped by water in some form and redeposited in a non-marine setting. This influence is strongly related to the adsorption of toxic metals to sediments or soil particles that result from the alteration of parent materials.

The maps of elements Zn, Pb and Cu also show high accumulations in the eastern part of country (Doha City area). Many studies have reported accumulations of toxic metals (Zn, Pb and Cu) in urban soils, mainly derived from vehicle emissions and industrial consumption [16,62–64,75]. Higher accumulations of these three elements have also been observed closer to highways, mainly from aerial deposition of metal particulates from motor vehicles. Due to the very high correlation between Cr, As and Ni, the maps of these three elements also show very similar spatial distributions. Higher values were observed in the northern part of the country when compared to urban areas.

However, due to the issue of cloudiness, the present study included satellite images from only nine months of 2015. This problem was apparent, for example, on images from March that could not be used for spatial modelling. However, according to our findings, spring images contributed the most to calibration; therefore, the calibration model might be improved if images from March were involved in this study. The issue of cloud coverage is one of the major constraints when applying optical satellite images. Thus, alternatives for dealing with this issue might be the utilization of imagery from

the Sentinel-2 sensor that assures the most frequent revisit period and increases the chances for the collection of cloud-free images.

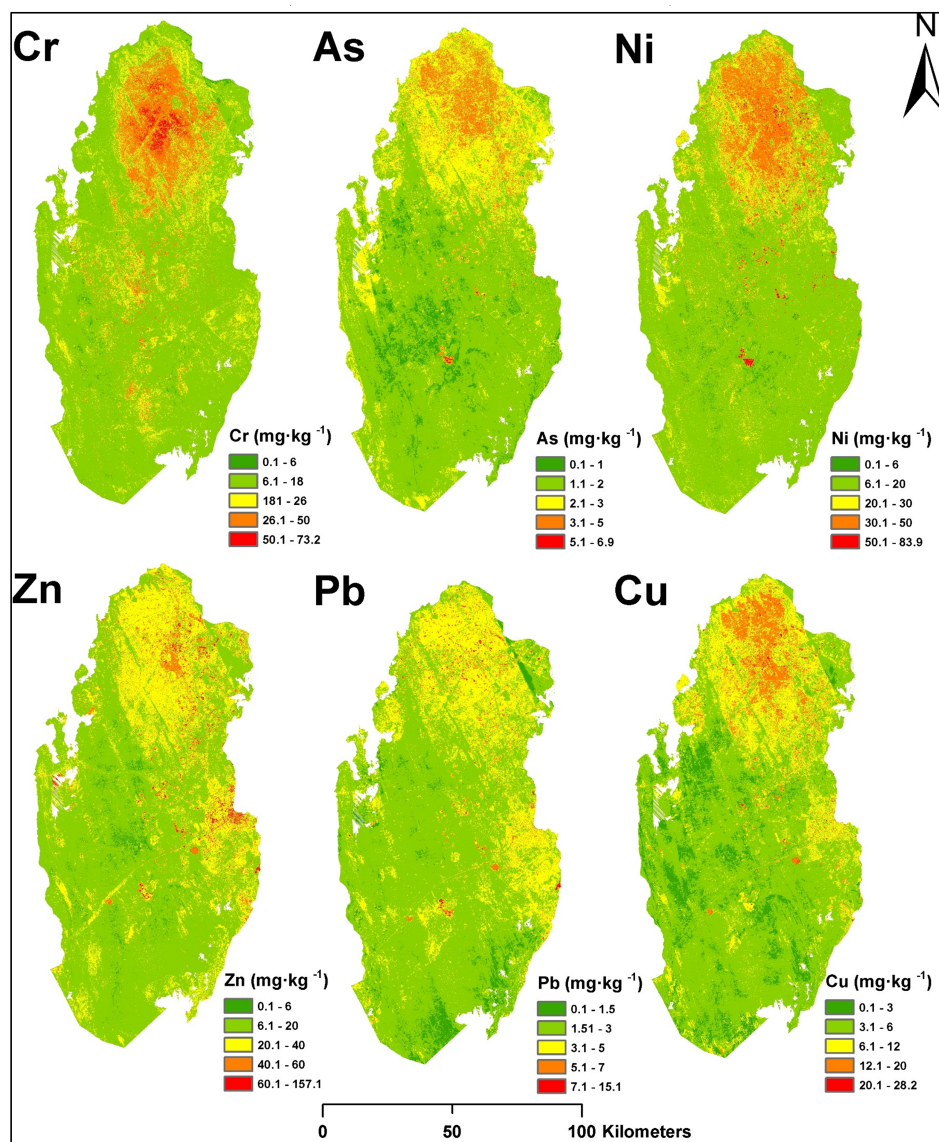


Figure 5. Predicted maps (30-m resolution) of topsoil (0–30 cm) concentrations of Cr, As, Ni, Zn, Pb and Cu using the Cubist model ($25^{\circ}28'61.06''\text{N}$, $51^{\circ}53'48.17''\text{E}$).

4. Conclusions

The use of various sets of environmental parameters (geopedological, geomorphology, anthropogenic), Landsat 8 images and spectral indices for predicting As/Cr/Ni/Zn/Cu/Pb contents in Qatari soils using Cubist regression rules showed promising predictive capabilities. One of the biggest findings of this study was that, because of the seasonal variation in the climate in Qatar, the RS images from January and February provided the most important vegetation information for predicting toxic metals' accumulation in soils. The modeling process also showed that different spectral indices, such as tasseled cap transformation and BCI, were able to enhance spectral features to indirectly reveal various toxic metals in soils. The spatial distribution of Cu, Zn and Pb was highly affected by anthropogenic parameters, such as the distance to environmental hotspots, proximity to roads and night light image. The spatial distribution of As, Cr and Ni, on the other hand, was influenced more by geopedological and geomorphology parameters.

The predictive maps can be used by environmental authorities and land managers when choosing remediation measures (preventive, protective and restorative) in urbanized and industrialized arid landscapes based on the dominant environmental cause. The maps can be considered as valuable decision-support systems (predictive tools) for all kinds of users when more sophisticated technologies and datasets are not available and do not require special skills to use. In addition, they can be applied to land evaluation, planning and management, as well as environmental protection strategies.

The overall accuracies of the spatial models can be increased by: (i) improving the scales and resolutions of the environmental parameters considered to better reflect soil pollution variation or (ii) adding other environmental parameters, such as a land use map or climate and vegetation data, if available.

Continuation of the current work relies on testing the ability of the multi-spectral images to predict other toxic metals (cadmium, Cd, mercury, Hg, uranium, U) prevailing in arid urbanized and industrialized landscapes. Furthermore, it will be interesting to compare the use of images with higher spatial resolutions or testing the fusion of images from different sensors for spatial modeling (such as WorldView, Sentinel-2, GeoEye, etc.).

Acknowledgments: This forms part of the work (National Priority Research Program Award 5-572-1-101) funded by the Qatar Foundation and executed in a collaboration between Aarhus University (Denmark) and Qatar Environment & Energy Research Institute (2012–2014). The authors would like to thank Basem Shomar from Qatar Environment & Energy Research Institute (QEERI). We are particularly grateful to Miss Rania Bou Kheir, who sadly passed away in 2014.

Author Contributions: All authors contributed to developing the overall concept of the analysis and the preparation and collection of the filed data. Yi Peng processed and analyzed the geospatial data and spatial modelling with close assistance of Kabindra Adhikari and Radosław Malinowski. Soil sampling work was carried out by Mette B. Greve and Mogens H. Greve. Most parts of the manuscript were written by Yi Peng with support from all of the remaining authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Alloway, B.J. Introduction. In *Heavy Metals in Soils*; Springer: Dordrecht, The Netherlands, 2013; pp. 3–9.
- Imperato, M.; Adamo, P.; Naimo, D.; Arienzo, M.; Stanzione, D.; Violante, P. Spatial distribution of heavy metals in urban soils of Naples city (Italy). *Environ. Pollut.* **2003**, *124*, 247–256. [[CrossRef](#)]
- Martley, E.; Gulson, B.L.; Pfeifer, H.R. Metal concentrations in soils around the copper smelter and surrounding industrial complex of Port Kembla, NSW, Australia. *Sci. Total Environ.* **2004**, *325*, 113–127. [[CrossRef](#)] [[PubMed](#)]
- Ordóñez, A.; Loredó, J.; De Miguel, E.; Charlesworth, S. Distribution of heavy metals in the street dusts and soils of an industrial city in Northern Spain. *Arch. Environ. Contam. Toxicol.* **2003**, *44*, 160–170. [[CrossRef](#)] [[PubMed](#)]
- Sterckeman, T.; Douay, F.; Proix, N.; Fourrier, H. Vertical distribution of Cd, Pb and Zn in soils near smelters in the North of France. *Environ. Pollut.* **2000**, *107*, 377–389. [[CrossRef](#)]
- Certini, G.; Scalenghe, R.; Woods, W.I. The impact of warfare on the soil environment. *Earth Sci. Rev.* **2013**, *127*, 1–15. [[CrossRef](#)]
- Arfsten, D.P.; Still, K.R.; Ritchie, G.D. A review of the effects of uranium and depleted uranium exposure on reproduction and fetal development. *Toxicol. Ind. Health* **2001**, *17*, 180–191. [[CrossRef](#)] [[PubMed](#)]
- Essentials of Medical Geology: Impacts of the Natural Environment on Public Health*; Selinus, O.; Alloway, B.; Centeno, J.A.; Finkelman, R.B.; Fuge, R.; Lindh, U.; Smedley, P. (Eds.) Elsevier Academic Press: London, UK, 2005; pp. 374–377.
- Lee, C.S.; Li, X.D.; Shi, W.Z.; Cheung, S.C.; Thornton, I. Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics. *Sci. Total Environ.* **2006**, *356*, 45–61. [[CrossRef](#)] [[PubMed](#)]
- Li, X.; Lee, S.-L.; Wong, S.-C.; Shi, W.; Thornton, I. The study of metal contamination in urban soils of Hong Kong using a GIS-based approach. *Environ. Pollut.* **2004**, *129*, 113–124. [[CrossRef](#)] [[PubMed](#)]

11. Ljung, K.; Selinus, O.; Otabbong, E.; Berglund, M. Metal and arsenic distribution in soil particle sizes relevant to soil ingestion by children. *Appl. Geochem.* **2006**, *21*, 1613–1624. [[CrossRef](#)]
12. Shomar, B.; Amr, M.; Al-Saad, K.; Mohieldeen, Y. Natural and depleted uranium in the topsoil of Qatar: Is it something to worry about? *Appl. Geochem.* **2013**, *37*, 203–211. [[CrossRef](#)]
13. Kheir, R.B. *Project Report: Sustainable Development in the State of Qatar*; National Priority Research Program: Doha, Qatar, 2012.
14. Pearce, F. Qatar to Use Biofuels? What About the Country's Energy Consumption? Available online: <https://www.theguardian.com/environment/2010/jan/14/qatar-biofuels-energy-consumption> (accessed on 14 January 2010).
15. Guo, G.H.; Wu, F.C.; Xie, F.Z.; Zhang, R.Q. Spatial distribution and pollution assessment of heavy metals in urban soils from southwest China. *J. Environ. Sci.* **2012**, *24*, 410–418. [[CrossRef](#)]
16. Lado, L.R.; Hengl, T.; Reuter, H.I. Heavy metals in European soils: A geostatistical analysis of the FOREGS Geochemical database. *Geoderma* **2008**, *148*, 189–199. [[CrossRef](#)]
17. Facchinelli, A.; Sacchi, E.; Mallen, L. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. *Environ. Pollut.* **2001**, *114*, 313–324. [[CrossRef](#)]
18. Lu, A.X.; Wang, J.H.; Qin, X.Y.; Wang, K.Y.; Han, P.; Zhang, S.Z. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. *Sci. Total Environ.* **2012**, *425*, 66–74. [[CrossRef](#)] [[PubMed](#)]
19. Lin, Y.P.; Teng, T.P.; Chang, T.K. Multivariate analysis of soil heavy metal pollution and landscape pattern in Changhua county in Taiwan. *Landsc. Urban Plan.* **2002**, *62*, 19–35. [[CrossRef](#)]
20. Saby, N.; Arrouays, D.; Boulonne, L.; Jolivet, C.; Pochot, A. Geostatistical assessment of Pb in soil around Paris, France. *Sci. Total Environ.* **2006**, *367*, 212–221. [[CrossRef](#)] [[PubMed](#)]
21. Zhang, H.H.; Chen, J.J.; Zhu, L.; Li, F.B.; Wu, Z.F.; Yu, W.M.; Liu, J.M. Spatial patterns and variation of soil cadmium in Guangdong Province, China. *J. Geochem. Explor.* **2011**, *109*, 86–91. [[CrossRef](#)]
22. Ungaro, F.; Ragazzi, F.; Cappellin, R.; Giandon, P. Arsenic concentration in the soils of the Brenta Plain (Northern Italy): Mapping the probability of exceeding contamination thresholds. *J. Geochem. Explor.* **2008**, *96*, 117–131. [[CrossRef](#)]
23. Xie, Y.F.; Chen, T.B.; Lei, M.; Yang, J.; Guo, Q.J.; Song, B.; Zhou, X.Y. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis. *Chemosphere* **2011**, *82*, 468–476. [[CrossRef](#)] [[PubMed](#)]
24. Minasny, B.; McBratney, A.B. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemom. Intell. Lab.* **2008**, *94*, 72–79. [[CrossRef](#)]
25. Melendez-Pastor, I.; Navarro-Pedreño, J.; Gómez, I.; Almendro-Candel, M. The use of remote sensing to locate heavy metal as source of pollution. In *Advances in Environmental Research*; Daniels, J.A., Ed.; Nova Science Publishers: Hauppauge, NY, USA, 2011; pp. 225–233.
26. Shi, T.; Chen, Y.; Liu, Y.; Wu, G. Visible and near-infrared reflectance spectroscopy—An alternative for monitoring soil contamination by heavy metals. *J. Hazard. Mater.* **2014**, *265*, 166–176. [[CrossRef](#)] [[PubMed](#)]
27. Wu, Y.Z.; Chen, J.; Ji, J.F.; Gong, P.; Liao, Q.L.; Tian, Q.J.; Ma, H.R. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. *Soil Sci. Soc. Am. J.* **2007**, *71*, 918–926. [[CrossRef](#)]
28. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Chapter five—Visible and near infrared spectroscopy in soil science. *Adv. Agron.* **2010**, *107*, 163–215.
29. Wu, Y.; Zhang, X.; Liao, Q.; Ji, J. Can contaminant elements in soils be assessed by remote sensing technology: A case study with simulated data. *Soil Sci.* **2011**, *176*, 196–205. [[CrossRef](#)]
30. Malley, D.F.; Williams, P.C. Use of near-infrared reflectance spectroscopy in prediction of heavy metals in freshwater sediment by their association with organic matter. *Environ. Sci. Technol.* **1997**, *31*, 3461–3467. [[CrossRef](#)]
31. Ben-Dor, E.; Irons, J.; Epema, G. Soil reflectance. In *Manual of Remote Sensing, Remote Sensing for the Earth Sciences*, 3rd ed.; Rencz, A.N., Ed.; John Wiley & Sons: New York, NY, USA, 1999; p. 111.
32. Choe, E.; van der Meer, F.; van Ruitenbeek, F.; van der Werff, H.; de Smeth, B.; Kim, K.-W. 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.* **2008**, *112*, 3222–3233. [[CrossRef](#)]

33. Grunwald, S.; Vasques, G.M.; Rivero, R.G. Chapter one—Fusion of soil and remote sensing data to model soil properties. *Adv. Agron.* **2015**, *131*, 1–109.
34. Mulder, V.L.; de Bruin, S.; Schaepman, M.E.; Mayr, T.R. The use of remote sensing in soil and terrain mapping—A review. *Geoderma* **2011**, *162*, 1–19. [[CrossRef](#)]
35. Peng, Y.; Xiong, X.; Adhikari, K.; Knadel, M.; Grunwald, S.; Greve, M.H. Modeling soil organic carbon at regional scale by combining multi-spectral images with laboratory spectra. *PLoS ONE* **2015**, *10*, e0142295. [[CrossRef](#)] [[PubMed](#)]
36. Kemper, T.; Sommer, S. Mapping and monitoring of residual heavy metal contamination and acidification risk after the aznalcóllar mining accident (Andalusia, Spain) using field and airborne hyperspectral data. In Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy, Herrsching, Germany, 13–16 May 2003; pp. 333–343.
37. Scheibert, C.; Stietiya, M.; Sommar, J.; Abdalla, O.; Schramm, H.; Al Memah, M. *The Atlas of Soils for the State of Qatar*; Ministry of Municipal Affairs and Agriculture: Doha, Qatar, 2005; pp. 19–36.
38. Qatar, T. *Qatar National Atlas*; The Planning Council of The State of Qatar: Doha, Qatar, 2006.
39. IUSS Working Group. *World Reference Base for Soil Resource 2014, International Soil Classification System for Naming Soils and Creating Legends for Soil Maps*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2014.
40. Theocharopoulos, S.P.; Wagner, G.; Sprengart, J.; Mohr, M.E.; Desaulles, A.; Muntau, H.; Christou, M.; Quevauviller, P. European soil sampling guidelines for soil pollution studies. *Sci. Total Environ.* **2001**, *264*, 51–62. [[CrossRef](#)]
41. Kheir, R.B.; Greve, M.H.; Abdallah, C.; Dalgaard, T. Spatial soil zinc content distribution from terrain parameters: A GIS-based decision-tree model in Lebanon. *Environ. Pollut.* **2010**, *158*, 520–528. [[CrossRef](#)] [[PubMed](#)]
42. Kheir, R.B.; Greve, M.H.; Derooin, J.P.; Rebai, N. Implementing GIS regression trees for generating the spatial distribution of copper in mediterranean environments: The case study of Lebanon. *Int. J. Environ. Anal. Chem.* **2013**, *93*, 75–92. [[CrossRef](#)]
43. Deng, C.; Wu, C. BCI: A biophysical composition index for remote sensing of urban environments. *Remote Sens. Environ.* **2012**, *127*, 247–259. [[CrossRef](#)]
44. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
45. Xiao, X.; Hollinger, D.; Aber, J.; Goltz, M.; Davidson, E.A.; Zhang, Q.; Moore, B., III. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sens. Environ.* **2004**, *89*, 519–534. [[CrossRef](#)]
46. Rouse, J.W.; Haas, J.R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with erts. In Proceedings of the Third ERTS Symposium, Washington, DC, USA, 1974; pp. 309–317.
47. Marsett, R.C.; Qi, J.; Heilman, P.; Biedenbender, S.H.; Watson, M.C.; Amer, S.; Weltz, M.; Goodrich, D.; Marsett, R. Remote sensing for grassland management in the arid southwest. *Rangel. Ecol. Manag.* **2006**, *59*, 530–540. [[CrossRef](#)]
48. Qi, J.; Marsett, R.; Heilman, P.; Bieden-bender, S.; Moran, S.; Goodrich, D.; Weltz, M. RANGES improves satellite-based information and land cover assessments in southwest United States. *Eos Trans. Am. Geophys. Union* **2002**, *83*, 601–606. [[CrossRef](#)]
49. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
50. Nellis, M.D.; Briggs, J.M. Transformed vegetation index for measuring spatial variation in drought impacted biomass on Konza Prairie, Kansas. *Trans. Kans. Acad. Sci. (1903)* **1992**, *95*, 93–99. [[CrossRef](#)]
51. Clevers, J.G.P.W. The derivation of a simplified reflectance model for the estimation of leaf area index. *Remote Sens. Environ.* **1988**, *25*, 53–69. [[CrossRef](#)]
52. Richardson, A.J.; Weigand, C. Distinguishing vegetation from soil background information. *Photogramm. Eng. Remote Sens.* **1977**, *43*, 1541–1552.
53. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd ed.; Pearson Education India: New Delhi, India, 2009.
54. Kuhn, M.; Weston, S.; Keefer, C.; Coulter, N.; Quinlan, R. *Cubist: Rule- and Instance-Based Regression Modeling*; Package Version 0.0.18; CRAN: Wien, Austria, 2014.

55. Quinlan, J.R. *C4. 5: Programs for Machine Learning*; Morgan Kaufmann: Burlington, MA, USA, 1993.
56. Adhikari, K.; Hartemink, A.E.; Minasny, B.; Kheir, R.B.; Greve, M.B.; Greve, M.H. Digital mapping of soil organic carbon contents and stocks in Denmark. *PLoS ONE* **2014**, *9*. [[CrossRef](#)] [[PubMed](#)]
57. Adhikari, K.; Kheir, R.B.; Greve, M.B.; Bocher, P.K.; Malone, B.P.; Minasny, B.; McBratney, A.B.; Greve, M.H. High-resolution 3-D mapping of soil texture in Denmark. *Soil Sci. Soc. Am. J.* **2013**, *77*, 860–876. [[CrossRef](#)]
58. Quinlan, J.R. Learning with continuous classes. In *5th Australian Joint Conference on Artificial Intelligence*; Word Scientific: Singapore, 1992; pp. 343–348.
59. Kuhn, M. *Caret: Classification and Regression Training*; R Package Version 6.0-24; CRAN: Wien, Austria, 2014.
60. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.M.; McBratney, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trend Anal. Chem.* **2010**, *29*, 1073–1081. [[CrossRef](#)]
61. Chang, C.W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [[CrossRef](#)]
62. Sun, Y.; Zhou, Q.; Xie, X.; Liu, R. Spatial, sources and risk assessment of heavy metal contamination of urban soils in typical regions of Shenyang, China. *J. Hazard. Mater.* **2010**, *174*, 455–462. [[CrossRef](#)] [[PubMed](#)]
63. Wei, B.; Yang, L. A review of heavy metal contaminations in urban soils, urban road dusts and agricultural soils from China. *Microchem. J.* **2010**, *94*, 99–107. [[CrossRef](#)]
64. Kheir, R.B.; Shomar, B.; Greve, M.B.; Greve, M.H. On the quantitative relationships between environmental parameters and heavy metals pollution in mediterranean soils using GIS regression-trees: The case study of Lebanon. *J. Geochem. Explor.* **2014**, *147*, 250–259. [[CrossRef](#)]
65. Shi, T.; Wang, J.; Chen, Y.; Wu, G. Improving the prediction of arsenic contents in agricultural soils by combining the reflectance spectroscopy of soils and rice plants. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 95–103. [[CrossRef](#)]
66. Ahmad, M.S.A.; Ashraf, M. Essential roles and hazardous effects of nickel in plants. *Rev. Environ. Contam. Toxicol.* **2011**, *214*, 125–167. [[PubMed](#)]
67. Ashraf, M.Y.; Sadiq, R.; Hussain, M.; Ashraf, M.; Ahmad, M.S.A. Toxic effect of nickel (Ni) on growth and metabolism in germinating seeds of sunflower (*Helianthus annuus* L.). *Biol. Trace Element Res.* **2011**, *143*, 1695–1703. [[CrossRef](#)] [[PubMed](#)]
68. Siddiqui, M.H.; Al-Whaibi, M.H.; Basalah, M.O. Interactive effect of calcium and gibberellin on nickel tolerance in relation to antioxidant systems in *Triticum aestivum* L. *Protoplasma* **2011**, *248*, 503–511. [[CrossRef](#)] [[PubMed](#)]
69. Pourrut, B.; Shahid, M.; Dumat, C.; Winterton, P.; Pinelli, E. Lead uptake, toxicity, and detoxification in plants. *Rev. Environ. Contam. Toxicol.* **2013**, *213*, 113–136.
70. Mohamed, H.I. Molecular and biochemical studies on the effect of gamma rays on lead toxicity in cowpea (*Vigna sinensis*) plants. *Biol. Trace Element Res.* **2011**, *144*, 1205–1218. [[CrossRef](#)] [[PubMed](#)]
71. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B* **1994**, *22*, 247–252. [[CrossRef](#)]
72. Kim, J.; Grunwald, S.; Rivero, R.G. Soil phosphorus and nitrogen predictions across spatial escalating scales in an aquatic ecosystem using remote sensing images. *IEEE Trans. Geosci Remote Sens.* **2014**, *52*, 6724–6737.
73. Chibuike, G.; Obiora, S. Heavy metal polluted soils: Effect on plants and bioremediation methods. *Appl. Environ. Soil Sci.* **2014**, *2014*, 752708. [[CrossRef](#)]
74. Li, Y.; Li, C.K.; Tao, J.; Wang, L.D. Study on spatial distribution of soil heavy metals in Huizhou city based on BP-ANN modeling and GIS. *Procedia Environ. Sci.* **2011**, *10*, 1953–1960. [[CrossRef](#)]
75. Maas, S.; Scheifler, R.; Benslama, M.; Crini, N.; Lucot, E.; Brahmia, Z.; Benyacoub, S.; Giraudoux, P. Spatial distribution of heavy metal concentrations in urban, suburban and agricultural soils in a mediterranean city of Algeria. *Environ. Pollut.* **2010**, *158*, 2294–2301. [[CrossRef](#)] [[PubMed](#)]

