Integrated Approach of Heavy Metal Evaluation Using Geostatistical and Pollution Assessment Index in Soil of Bauxite Mining Area

Nur Shuhada Tajudin¹*, Mazidah Zulkifli¹, Mohd Fuad Miskon², Mohamad Izzuddin Anuar³, Zulkifli Hashim³, Fikriah Faudzi⁴ and Nurul Mayzaitul Azwa Jamaluddin⁵

¹Department of Plant Science, Kulliyyah of Science, International Islamic University Malaysia, 25200 IIUM, Kuantan, Pahang, Malaysia
²Institute of Oceanography and Maritime Studies (INOCEM), International Islamic University Malaysia, 25200 IIUM, Kuantan, Pahang, Malaysia
³Malaysian Palm Oil Board, 6 Persiaran Institusi, Bandar Baru Bangi, 43000 MPOB, Kajang, Selangor, Malaysia
⁴Department of Marine Science, Kulliyyah of Science, International Islamic University Malaysia, 25200 IIUM, Kuantan, Pahang, Malaysia
⁵Laboratory of Food Safety and Food Integrity, Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

Heavy metals contamination in soil is one of the global issues, posing a threat not just to the environment but also to human health. Identifying the source and distribution of heavy metal pollutants around mining areas can provide a scientific basis for future environmental control. Distributions of the heavy metals (Cd, Cr, As, and Ni) in this study were evaluated using descriptive and multivariate statistics and further described using a geostatistical approach and pollution indices. The total content of Cr, Cd, and Ni in surface soil was observed with a higher concentration level according to the Dutch target values and the 95% Investigation Levels determined for Malaysia soil. Statistical analyses, geostatistics, and GIS mapping suggested that Cd, Cr, and Ni were derived mainly from anthropogenic sources, including mining and agricultural activities, while As could be derived from lithogenic and anthropogenic sources. Geoaccumulation index analysis demonstrated that the contamination that occurred with Cd posed the greatest risk of...
contamination, followed by Cr, Ni, and As. A spatial interpolated map showed a higher concentration of heavy metals in the vicinity of the mining area. These findings highlight the effectiveness of principal component analysis, geostatistics, and geospatial analyses in evaluating heavy metal contents in the study area. The obtained results could be used by authorities to identify areas requiring remediation management and establish scientific baseline data related to soil quality.

Keywords: Bauxite mining, geoaccumulation index, GIS, heavy metals, semivariogram

INTRODUCTION
Mining has recently emerged as one of the world’s most important industries. One of the four main methods in mining is surface mining, which involves the removal of plants and topsoils, digging down into the soil surface and potentially bedrock to obtain the resource deposits. However, various negative consequences have been observed as a result of this activity, including damage to the environment, socioeconomics, and human health. The most well-known environmental hazards that need concern are tailing dams, deposits, and slag produced by mining activities. Heavy metals such as cadmium (Cd), chromium (Cr), copper (Cu), lead (Pb), zinc (Zn), and metalloids such as arsenic (As) can be present in tailings. If mining activities are not effectively managed, it can cause detrimental effects to the environment and human life (Kusin et al., 2018). One of the most concerning issues in mining activities is the removal of the fertile soil at the top layer. According to Lee et al. (2017), unmined land such as forests has a well-developed soil profile, also known as mature soil, that can support deep-rooted trees and crops, while mined soil has a shallow depth of about 15 cm or less, known as young soil. This type of land had poor content of organic matter, low water holding capacity, low microbial activity, and was subjected to nutrient deficiency. Furthermore, removing the land’s upper layer will result in a loss of soil structure and stability, which latter will cause a serious socioeconomic implication (Prematuri et al., 2020).

The bauxite mining industry in Malaysia began in the early 2000s in Johor, without much concern for the environmental effect. Bauxite is an ore that is mined for the extraction of aluminum (Al), which is then used to make a variety of industrial goods. The alumina concentration of bauxite ore ranged from 45 to 53% (Hutchison, 2005). In Malaysia, the bauxite resources have been identified mainly in Bukit Batu, Bukit Gebong, Lundu-Sematan, and Tanjung Seberang in Sarawak, Bukit Mengkabau and Labuk Valley in Sabah, Sungai Rengit and Teluk Ramunia in Johor, and Bukit Goh in Kuantan, Pahang (Malaysian Minerals Yearbook, 2013). However, in 2014, mining activities in Kuantan, Pahang, raised a lot of debatable issues due to the environmental pollution that occurred from the mining activities (Abdullah et al., 2016). A total of 18,000 ha area in Kuantan, Pahang, has been heavily exploited for bauxite, which between 10-to-20-meter topsoil layers were scavenged
Heavy Metals Evaluation in Soil of Bauxite Mining, Kuantan (Ismail et al., 2018). In addition, oil palm trees were removed for mining purposes, causing soil erosion, sedimentation, and pollutants entering low-lying aquatic areas.

Heavy metals (HMs) are extremely significant among soil pollutants as some metals such as Cd, As, and Cr can be extremely carcinogenic, poisonous, and persistent in the ecosystem. Its content in soil can be irreversible and convertible to highly toxic elements that can pose harmful effects to human health and the ecosystem (Jiang et al., 2019; Liu et al., 2018; Tang et al., 2013). Due to their toxicity and long-term persistence in the environment, the study on the geochemistry of HMs has progressed significantly in recent decades (Gao et al., 2021; Krishnan et al., 2021; Zeng et al., 2021).

Monitoring the distributions of HMs involving large areas is one of the most challenging but important. Laboratory analysis limits the spatial visualization of the HMs distribution over an area. It is critical to provide precise information about HMs mobility, origin, extent, association, and biological availability at various soil depths (Yang et al., 2009a). With the development of information technology, the geostatistical approach has become an important tool for determining the distributions of soil HMs at various scales (Tóth et al., 2017). It allows quantifications of spatial features of soil properties and is able to conduct spatial interpolation with unbiased estimation (Shi et al., 2007; Webster & Oliver, 2001). The GIS tools were initially developed for data retrieval and displaying geographic information but later enhanced for spatial analysis as the research field expanded. Spatial variability of soil pollutants can be shown as a continuous map with a visually identifiable ‘hotspots’ zone using geostatistical estimators (Yang et al., 2009b). This method has been demonstrated to be a powerful tool to evaluate the background and the spatial structure of heavy metals in soils (Zhou & Xia, 2010) and soil pollution surveys (Zhang et al., 2018). It enables researchers to observe the association between heavy metal indicators and pollution sources, hence producing a guideline on establishing a cost-effective risk management policy.

Research related to monitoring the spatial distribution of HMs in bauxite mining areas is limited. A visit to the study site has found that the mined bauxite area was left forsaken without any rehabilitation actions. Thus, preliminary data on the soil properties are needed to assess the impact of mining on the surrounding area’s soil quality for future rehabilitation. The objectives of this study were to (i) determine the content of selected metals (Cd, As, Ni and Cr) in the soil of bauxite mining area; (ii) evaluate the extent of the HMs content in soil by using geoaccumulation index and principal component analysis and (iii) quantify the spatial distribution of HMs through geostatistics and geospatial analysis. This study highlights the combination of methodologies that facilitate the quantification of heavy metals distribution using multivariate analysis, geostatistical method, and environmental risk assessment. With the concern of the degraded and barren land impacted by the bauxite mining activities in Pahang, this study could provide baseline data to be used for land restoration, protection strategy, and policymaking.
MATERIALS AND METHODS

The study was conducted in Bukit Goh, Kuantan, Pahang, Malaysia and covered an area of 419 ha between 3°56’ 30” and 3°53’ 0”N latitude and 103°15’ 30” to 103° 18’ 0”E longitude. The climate is humid, with an average annual rainfall of 2000 to 3000 mm. The monthly average temperature ranged from 21°C to 31°C, respectively. Half of the study area is primarily utilized for oil palm plantations, while the other half is used for mining activities (Figure 1). Global positioning systems (GPS) were used to record the coordinates of each sampling point.

![Figure 1. Location of the study area in Kuantan, Pahang, and distribution of the soil sampling points](image)

Soil Sampling and Analysis

The sampling points were arranged systematically over the research region using a normal grid sampling method of 300 x 300 m, within the effective sampling method for geostatistical study (Webster & Oliver, 2001). A total of 64 soil samples were obtained for this purpose (Figure 1). Within each cell, three subsamples were collected from the topsoil (0 to 20 cm) and mixed at each sampling point.

Prior to the HMs analyses, the samples were pre-cleaned, air-dried, and sieved. Determination of total content of organic carbon (TOC) and pH was conducted based on Walkley-Black and KCl (1:2.5) methods (Jackson, 1958; Cai et al., 2012), respectively. For heavy metal analyses, the samples were digested with a 3.3:5.3:3.5 mixture of hydrofluoric acid (HF): nitric acid (HNO₃): hydrochloric acid (HCl) using closed digestion method (Kamaruzzaman, 1999; Tsunogai & Yamada, 1979). The digested samples were analyzed for the As, Cd, Cr, and Ni using the inductively coupled plasma-mass spectrometer (ICP-MS) Elan 9000 with standard configuration. Duplicate sample analyses and blanks were also prepared and analyzed to validate the analyses.
Heavy Metals Evaluation in Soil of Bauxite Mining, Kuantan

The Geoaccumulation Index
The geoaccumulation index ($I_{geo}$) was used to assess the level of contamination in the soils of this study. The $I_{geo}$ is calculated using Equation 1 (Muller, 1969)

$$I_{geo} = log_2 \left( \frac{C_n}{B_n} \right) \quad (1)$$

$C_n$ is the concentration of the evaluated soil HMs in this study, $B_n$ is the geochemical background concentration of metal ‘n,’ and factor 1.5 is the background matrix correction factor. Table 1 shows the $I_{geo}$ categories and their contamination level (Rahman et al., 2012).

Statistical Analysis
Multivariate statistical analyses, including principal component analysis (PCA), matrix correlation, and geostatistical analysis, were employed in this study to identify and distinguish pollution sources and HMs grouping. Common statistical analyses were processed using SPSS26.0 software, while geostatistical modeling was performed using GSWin10.0.2. Version 9.

Analysis of Geostatistics and Semivariogram
Geostatistics analysis has been widely used as part of applied statistics to analyze the spatial correlation of the dataset. It calculates the spatial dependency and structure of a measured property, such as soil parameters (e.g., pH, TOC, chemical parameters) and uses that information to calculate the unsampled locations. This approach has been effectively used to reduce the cost of sampling and lab analyses. Generally, two steps were involved in this technique which are (i) spatial modeling (variography) and (ii) spatial interpolation.

It is necessary to assume that the dataset has a spatial autocorrelation to determine the spatial correlation of the data. The correlation is expressed as semivariance. The calculation of the semivariance at h separation was calculated as in Equation 2 (distance vector) (Goovaerts, 1999):

$$\lambda(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x_i) - z(x_i+h)]^2 \quad (2)$$

$\lambda(h)$ is the number of sample value pairs within the distance interval $h$; $z(x_i)$ is the value of samples $z$ at point $x_i$ and $n(h)$ is the total pairs of sample points separated by lag distance.

Table 1

<table>
<thead>
<tr>
<th>$I_{geo}$</th>
<th>Class</th>
<th>Contamination level</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0</td>
<td>0</td>
<td>Not polluted</td>
</tr>
<tr>
<td>0-1</td>
<td>1</td>
<td>Not polluted to moderately polluted</td>
</tr>
<tr>
<td>1-2</td>
<td>2</td>
<td>Moderately polluted</td>
</tr>
<tr>
<td>2-3</td>
<td>3</td>
<td>Moderately polluted to strongly polluted</td>
</tr>
<tr>
<td>3-4</td>
<td>4</td>
<td>Strongly polluted</td>
</tr>
<tr>
<td>4-5</td>
<td>5</td>
<td>Strongly to extremely polluted</td>
</tr>
<tr>
<td>&gt;5</td>
<td>6</td>
<td>Extremely polluted</td>
</tr>
</tbody>
</table>

This relationship was characterized as a variogram and used for spatial prediction. The semivariance is charted as a distance function, and a fitted model (spherical, exponential, linear, and gaussian) was generated to express this relationship. Three parameters that can be described by the relationship between distance and the semivariance are sill, nugget, and range (Figure 2). The range is the distance when two variables become spatially independent, while variance at zero length is called nugget. The lag distance between measurements at which a variable’s value does not affect its neighbors is called sill (Lopez-Granados et al., 2002).

The HMs spatial distribution was verified for normal random distribution and analyzed using variogram analyses. Non-normal data were log-transformed before the semivariogram analysis was conducted. The best fit variogram models (linear, spherical, exponential, and gaussian) were chosen based on the smallest root mean square error (RMSE), highest R², and low nugget value (30% lower than sill). In this study, half of the maximum distance between sample locations was used as the maximum lag distance. Plotted lag intervals were adjusted short enough to obtain the best-fitted semivariogram model. Since the samples were collected based on the grid sampling method, the grid spacing of 300 m was the best for lag size intervals (Isaacs & Srivastava, 1989). The value of nugget, sill, and range in this study were determined from the fitting curves obtained. Spatial dependence in this study was defined using the nugget to sill ratio interpretation suggested by Cambardella et al. (1994) in Table 2.

### Kriging Interpolation

The next stage of geospatial analysis was to predict the HMs value in unsampled locations to produce a continuous variability map. Spatial interpolation techniques using the ordinary kriging method from ArcGIS version 10.1 (ESRI, USA) were employed for this purpose. It allows the expression of spatial soil variables to be mapped out with minimal error of predicted values (Reza et al., 2015) and provides a better approach for estimating interpolation weights. Equation 3 was used to estimate the ordinary kriging:
\[ \hat{z}(S_0) = \sum_{i=1}^{N} \lambda_i z(S_i) \]  

(3)

where the value of the predicted point (Z-hat, at location S-nought) is equal to the sum of the value of each sampled point (x, at location i) times that point’s unique weight (lambda, for location i), the accuracy of the HMs spatial maps were evaluated by cross-validation approach (Isaacs & Srivastava, 1989), interpolated mean error (ME) and mean squared error (MSE) should be close to zero and less than the sample variance, respectively. Finally, the standard mean square error (SMSE) should be near 1.

RESULTS AND DISCUSSIONS

The Exploratory Analysis of Soil Variables

Descriptive statistical summary of the pH, TOC, As, Cd, Cr, and Ni were presented in Table 3. The pH of the soil is categorized as an acidic type of soil, ranging from 3.75 to 4.82. Total organic carbon content was in low concentrations with an average value of 0.42%± 0.32 and ranged from 0 to 1.8%. The lower content of organic carbon was most likely due to the removal of the topsoil for mining purposes. Removal of the topsoil layers will eventually reduce the content of organic matter as biological, and decomposition activities could only be found in this layer.

For the HMs content, As, Cd, Cr, and Ni were found to be varied from 4.05 to 29.89 mg/kg, 1.19 to 3.74 mg/kg, 257.25 to 3263.87 mg/kg, and 38.89 to 906.70 mg/kg, respectively, with Cr, had the highest levels of concentrations followed by Ni, As and Cd. The total content of As, Cr, TOC, and Ni in the studied area had a great degree of variability, indicated by a substantial coefficient of variation (CV) values of 55.52%,

<table>
<thead>
<tr>
<th>Soil constituents</th>
<th>Range</th>
<th>Mean ± SE</th>
<th>Median</th>
<th>Coefficient of Variation (%)</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Other study</th>
</tr>
</thead>
<tbody>
<tr>
<td>As (mg/kg)</td>
<td>4.05 – 29.89</td>
<td>12.72± 7.06</td>
<td>10.5</td>
<td>55.52</td>
<td>1.11</td>
<td>0.23</td>
<td>29 60</td>
</tr>
<tr>
<td>Cd (mg/kg)</td>
<td>1.19 – 3.74</td>
<td>2.18± 0.57</td>
<td>2.07</td>
<td>26.21</td>
<td>0.71</td>
<td>1.17</td>
<td>0.8 0.3</td>
</tr>
<tr>
<td>Cr (mg/kg)</td>
<td>257.25 – 3263.87</td>
<td>980.81±579.75</td>
<td>904.25</td>
<td>59.11</td>
<td>1.75</td>
<td>4.01</td>
<td>100 60</td>
</tr>
<tr>
<td>Ni (mg/kg)</td>
<td>38.89 – 906.70</td>
<td>191.466±159.83</td>
<td>149.72</td>
<td>83.48</td>
<td>2.46</td>
<td>06.89</td>
<td>35 45</td>
</tr>
<tr>
<td>pH</td>
<td>3.65 – 4.82</td>
<td>4.18± 0.28</td>
<td>4.18</td>
<td>6.59</td>
<td>6.90</td>
<td>-0.91</td>
<td>- -</td>
</tr>
<tr>
<td>TOC (%)</td>
<td>0 – 1.8</td>
<td>0.43±0.32</td>
<td>0.42</td>
<td>76.19</td>
<td>1.63</td>
<td>4.05</td>
<td>- -</td>
</tr>
</tbody>
</table>

a - Dutch target values (Dutch Target and Intervention Values, 2000), b - The 95% ‘Investigation Levels’ determined for Malaysia soil (Zarcinas et al., 2004)
59.11%, and 76.19%, 83.48%, respectively. Nevertheless, Cd showed a moderate degree of variability with CV values of 26.21%. The CV explained the extent to which human activities influence environmental media (Table 4). The CV values of HMs caused by natural factors such as weathering have been low, while CV values of HMs impacted by human activities were high. It is also agreed by Jianshu et al. (2013), which stated that environmental media that is strongly affected by human activities have a higher CV value of HMs concentration. Similarities were observed in the CV value and mapping of the HMs distribution in this study. A high CV value and uneven spatial distribution of HMs concentration were found in this study indicated a high proportion of anthropogenic sources of contaminations.

For HMs comparison, the Dutch target values for soil protection (Dutch Target and Intervention Values, 2000), which has been widely used in heavy metals studies (Rudzi et al., 2018; Mohseni-Bandpei et al., 2017; Wu et al., 2018) and the 95% ‘Investigation Levels’ determined for Malaysia soil (Zarcinas et al., 2004), were also presented in Table 3. The levels of Cd, Cr, and Ni in the surface soil of the mining region were found to be greater than those values.

The Cd contents in the soil were found seven times higher than the 95% ‘Investigation Levels’ determined for Malaysia soil with an average value of 2.18 ± 0.57 mg/kg. Cadmium commonly occurs in soils with its average content not exceeding 0.20 mg/kg, and its occurrence in the environment is in the form of chemical compounds (e.g., in ores of Zn, Pb, and Cu) (Jolanta & Andrzej, 2020). It has been reported that inputs from agriculture activities such as manure, sewage sludge, fertilizers, and herbicides typically enhance the total concentration of Cd in soil (Weggler et al., 2004). It is also present as an impurity in several products, including phosphate fertilizers (Swe et al., 2012). As oil palm plantations surround the study area, the higher content of Cd might be caused by long-term applications of fertilizers in the cultivated area. Considering the topography of the study area is not flat but hilly, surface soil carrying Cd might be exposed to erosion and deposited to low-lying areas by rain (Campbell, 2006). Moreover, rapid soil acidification due to the intense weathering will cause Cd to be mobile and easily leached. It is supported by the obtained CV of 26.21% for Cd, which indicated Cd mobility occurred within the study region.

The average concentration of As was found five times lower than the 95% ‘Investigation Levels’ determined for Malaysia soil and had almost two times lower than the Dutch Target value. It is in line with the findings of Kusin et al. (2018), which also reported lower As concentrations in bauxite surface sediments. Lower As found in the soil environment
indicated that its availability could be attributed primarily to the natural processes such as weathering, regardless of human anthropogenic activity (Nriagu et al., 2007).

The average value of Ni was almost four times higher than the Dutch target values for soil protection and the 95% ‘Investigation Levels’ determined for Malaysia soil. Greater amounts of these elements indicated the possibility of contamination from anthropogenic sources. However, this will be discussed in further detail in relation to spatial variability.

**Correlation of Heavy Metals in Soil**

Initial inferences on the possible sources of HMs in the soil of the bauxite mining area were made using the matrix correlation coefficients measured (Table 5). The correlation analysis showed that the total concentration of As, Cd, and Cr had a significant correlation with Ni (p<0.05). It indicated that those metals were interdependence and could originate from the same contamination source (Ali et al., 2016). Chromium also showed a significant correlation with Cd at p<0.05 (r=0.66), suggesting their availability in the soil was interdependence and might be associated with anthropogenic activities.

**Table 5**

*Matrix of correlation between heavy metals in the soil of bauxite mining area, Kuantan*

<table>
<thead>
<tr>
<th></th>
<th>pH</th>
<th>TOC</th>
<th>As</th>
<th>Cd</th>
<th>Cr</th>
<th>Ni</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>1.00</td>
<td>-0.28</td>
<td>-0.30</td>
<td>-0.11</td>
<td>-0.065</td>
<td>0.00</td>
</tr>
<tr>
<td>TOC</td>
<td>1.00</td>
<td>0.09</td>
<td>-0.21</td>
<td>-0.12</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>As</td>
<td>1.00</td>
<td>0.41</td>
<td>0.14</td>
<td>0.60**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cd</td>
<td>1.00</td>
<td>0.66**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr</td>
<td>1.00</td>
<td>0.57**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ni</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at 0.05 level (2-tailed)*

**Principal Component Analysis**

The principal component analysis (PCA) was employed to distinguish groups of HMs caused by lithogenic inputs or anthropogenic activities considering the complexity of the soil’s geochemical properties (Qiao et al., 2011). PCA showed that the original variables were separated into two factors as two eigenvalues were extracted with values higher than 1. According to the rotated component matrix, PC1 and PC2 had 43.46 and 24.78% of the total variance, respectively (Table 6). The first component matrix indicated that Cd, Cr, and Ni were interrelated, showing a strong load of those elements in the PC1. Nickel had the most significant loads with a value of 0.91, followed by Cd (0.90) and Cr (0.77). The trend was initially observed through matrix correlation analysis, where a significant relationship existed between Cd and Cr with Ni. Significant correlation between these elements indicated their common origin, which PC1 could be associated with the anthropogenic
source, notably from mining activities or geochemical weathering of minerals, which also could be derived from the long-term mining exposure (Bhuiyan et al., 2010). In addition, Cr and Ni are known to be mutually associated with a variety of rocks and commonly associated with mining activities (Spurgeon et al., 2008). However, not all HMs could be only distributed in one factor, as the As was partially in PC1 and PC2, suggesting that its availability might be controlled by more than one factor. While PC2 includes the pH and TOC, which account for 24.78% of the total variance. It revealed a different major controlling factor, most probably intrinsic or natural forming factors.

Leaching of HMs such as Cd, As, Ni, and Cr have been one of the major issues in the disposal of mineral products and mine tailings after mining operations have ceased (Dold, 2014; Schaanning et al., 2011; Smith et al., 1995). Initially, metals are stable in the ore, but their availability in the environment could increase due to oxidation, dissolution, and immobilization processes, which convert them to more soluble forms of metals. In addition, when the soil was heavily exposed to anthropogenic activities such as mining, it accelerated soil degradation.

### Geostatistics and Semivariogram Analysis

All the semivariogram models were constructed based on log-transformed data. Isaacs and Srivastava (1989) suggested a normally distributed dataset for kriging to work best (Lark & Lapworth, 2012; Hofer et al., 2013). Experimental semivariograms with the adjusted models of selected total metal contents are shown in Table 7. All the soil variables showed a definable spatial structure with Cd, Cr, and Ni conformed to an exponential model, while pH and As conformed to the spherical model and TOC represented by a linear model. The models’ selection and its fitting procedure are among the most important procedures for better predicting unsampled locations (McBratney & Webster, 1986). All of the models

**Table 6**

Total variance and component matrices of the soil of bauxite mining area, Kuantan

<table>
<thead>
<tr>
<th>Soil Element</th>
<th>Rotated Component Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC1</td>
</tr>
<tr>
<td>Ni</td>
<td>0.91</td>
</tr>
<tr>
<td>Cd</td>
<td>0.90</td>
</tr>
<tr>
<td>Cr</td>
<td>0.76</td>
</tr>
<tr>
<td>As</td>
<td>0.61</td>
</tr>
<tr>
<td>pH</td>
<td>0.01</td>
</tr>
<tr>
<td>TOC</td>
<td>-0.15</td>
</tr>
</tbody>
</table>
The nugget variance in this study was found to range from 0.00001 to 0.13 (Table 7). The number of point pairings decreases over short distances (Goovaerts, 1997). In this study, the nugget value perhaps better describes variability occurring within the shortest sampling interval rather than laboratory analytical errors (Goovaerts, 1999). Small nugget values indicated a relative variance and adequate sampling density to reveal the spatial features (Yang et al., 2009b). The nugget values were below 30% of the sill value except for As, indicating good spatial continuity at a shorter distance between sampling points and minimal random variation due to measurement errors (Gonzalez et al., 2001). This study could provide guidelines for understanding the spatial variation scales related to HMs monitoring.

The distance described the range in a semivariogram, where the model flattened out. Sample locations within the range are categorized as spatially autocorrelated. However, locations larger than the range are not. In this study, the range of the optimal model for the analyzed data varied between 79 to 1578 m, with the lowest range being in Cd followed by Cr, pH, TOC, Pb, and As. The range for the soil variables was above 1000 m, except for Cd and Cr, which had a value of 79 m and 225 m, respectively. These showed that the spatial structure was beyond the original sampling distance (300 m). Nevertheless, this finding provides a good indicator for the sampling density of Cd and Cr in the future, which requires shorter sampling spacing.

The nugget to sill ratio is an important indicator to reflect the extent of spatial autocorrelations of environmental factors (Cambardella et al., 1994). It reflects whether the spatial occurrence of heavy metals is due to intrinsic or extrinsic factors. The soil pH, Cd, Cr, and Ni were observed to have strong spatial autocorrelations with a value less than 25% suggesting its distributions in the study area were caused by intrinsic factors. However, As showed a moderate spatial dependence with a value of 33% reflecting its spatial distributions could be from mixed inputs of natural and anthropogenic sources.

### Table 7

<table>
<thead>
<tr>
<th>Soil Properties</th>
<th>Model variogram model type</th>
<th>Nugget variance $C_0$</th>
<th>Structural sill $C_0 + C$</th>
<th>Proportion $C_0/(C_0 + C)$ (%)</th>
<th>Range (m)</th>
<th>$R^2$</th>
<th>Residual RSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>Spherical</td>
<td>$1.0 \times 10^{-5}$</td>
<td>$4.6 \times 10^{-3}$</td>
<td>0.22</td>
<td>613.00</td>
<td>0.95</td>
<td>$3.145 \times 10^{-7}$</td>
</tr>
<tr>
<td>TOC</td>
<td>Linear</td>
<td>0.081</td>
<td>0.12</td>
<td>68</td>
<td>1262.91</td>
<td>0.52</td>
<td>$6.08 \times 10^{-4}$</td>
</tr>
<tr>
<td>As</td>
<td>Spherical</td>
<td>0.13</td>
<td>0.33</td>
<td>34</td>
<td>1578.00</td>
<td>0.00</td>
<td>$6.87 \times 10^{-6}$</td>
</tr>
<tr>
<td>Cd</td>
<td>Exponential</td>
<td>0.0061</td>
<td>0.066</td>
<td>9.2</td>
<td>79.00</td>
<td>0.64</td>
<td>$1.56 \times 10^{-5}$</td>
</tr>
<tr>
<td>Cr</td>
<td>Exponential</td>
<td>0.056</td>
<td>0.33</td>
<td>17</td>
<td>225.00</td>
<td>0.97</td>
<td>$3.64 \times 10^{-4}$</td>
</tr>
<tr>
<td>Ni</td>
<td>Exponential</td>
<td>0.065</td>
<td>0.46</td>
<td>14</td>
<td>303.00</td>
<td>0.58</td>
<td>$2.29 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
The spatial dependence obtained from semivariogram analysis showed that Cd, Cr, and Ni were aligned with the PCA results. Furthermore, it demonstrated that they were in the same set of categories, indicating that their point of contamination came from the same source. Similar results were obtained for As, which in PCA, it was partially categorized in PC1 and PC2, indicating that its source of contamination could be from both natural and anthropogenic sources.

Nevertheless, the fitted theoretical models suited with the existing data should not be discussed independently. The dynamic and complex behavior of HMs, influenced by various internal factors (i.e., pH, organic matter content, texture, redox state, and topography), could be the sources of the variability observed (Kabata-Pendias & Pendias, 1992). Generally, the HMs content in relation to spatial variability could be attributed to the geological origin of the rock and topography, but anthropogenic activities could further accelerate its availability in the environment. As the study area has an uneven topography with a moderate slope, the variability of the HMs will be greater, lowering the spatial autocorrelation between samples as the topsoil can be easily carried away to other locations. In addition, the mined land has been regularly disturbed for mining purposes, and bare soil without cover crops has a larger potential of erosion than flat land. It will cause the soil HMs to be distributed throughout the study area. Therefore, it was hypothesized that extrinsic soil-forming factors, parent material, and topography influenced the spatial variability and autocorrelation value of total HMs in this particular type of soil. This occurrence was further worsened by the mining activities conducted.

Assessment of Soil Contamination using Geoaccumulation Index ($I_{geo}$)

The $I_{geo}$ values for the HMs of the studied area are shown in Figure 3. The contamination levels were in the order of Cd>Cr>As>Ni. Cadmium had the highest $I_{geo}$ value with an average of 3.82±0.37 ranging from 3 to 4 and was categorized as strongly contaminated. It showed a consistent trend in this study, with its content in soil is higher than the reference value, reflecting a tendency towards higher contamination. It was in line with the results of PCA and semivariogram analysis, where anthropogenic sources could largely contribute to Cd contamination throughout the study area. Average $I_{geo}$ values for Ni were less than zero, suggesting no widespread Ni pollution in the soils. While, As and Cr showed an $I_{geo}$ value of moderate contamination.
to strongly polluted. Arsenic had a lower average value compared to the soil references and was not categorized into any component of PCA, which suggested that its occurrence could result possibly from both anthropogenic and natural factors or other sources of contamination. It is also supported by the semivariogram analysis, which had a moderate dependence showing that the source of contamination could be from both anthropogenic and lithogenic factors. However, by using $I_{geo}$, As was moderately polluted. Combinations of PCA and $I_{geo}$ analysis may assist in determining the extent of HMs contamination in soil, but the results may vary depending on the HMs as $I_{geo}$ mainly reflected the concentration levels considering the background metal.

Spatial Distribution of Soil Heavy Metal in the Study Area

The spatial distribution of pH, TOC, As, Cd, Cr, and Ni were shown in Figures 4 to 9. Similar trends can be observed among the distribution of the HMs. Higher distributions of Cd, Cr, As, and Ni was observed at the western and southwestern parts of the study area, where the land is used for bauxite mining activities. The fact that the spatial distribution pattern of soil HMs was similar showed that these pollutants were influenced by anthropogenic activities, specifically mining. However, contamination could also be caused by the long-term fertilizer usage in the surrounding oil palm plantation. It has been documented that those fertilizers such as phosphate have a trace number of HMs, and their application is significantly increased in the soil (Vogel et al., 2020; Hamid et al., 2019).

The soil pH of the study area was categorized into acidic soil types, with 100% of the study area having a pH lower than 5. This type of soil has high acidity, low effective cation exchange capacity (ECEC), and high aluminum saturation throughout the soil profiles.

![Figure 4. Spatial distribution of pH in the soil of bauxite mining area, Kuantan, Pahang](image-url)
Changes in soil pH within the study area could be caused by the nature of the parent material, micro-topography, long-term fertilizer used for the oil palm plantation, and bare land, which had been left forsaken. This type of soil has a high potential for basic cations losses through leaching and erosion. In addition, it will lead to a deficiency of essential plant nutrients such as P, Ca, K, Mg, and Mo due to Al and Mn toxicity (Dembele et al., 2016). Therefore, efforts should increase the pH, addressed through liming and effective organic matter management.

Figure 5. Spatial distribution of TOC (%) in the soil of bauxite mining area, Kuantan, Pahang

Figure 6. Spatial distribution of Cd (mg/kg) in the soil of bauxite mining area, Kuantan, Pahang
Among the HMs, 100% of the study area had a lower As content compared to the 95% ‘Investigation Levels’ determined for Malaysia soil (Zarcinas et al., 2004). The content of As was higher at the western part of the study area, with concentrations ranging from 10–30 mg/kg. Although the levels of As were profoundly low, its availability in the environment is still a concern. Arsenic can be found in minerals predominantly as As$_2$O$_3$. It can also be obtained by processing ores containing Cu, Pb, Zn, and Ag. Arsenate in its ionic form can act as a chelate and precipitate with metals cations to form a stable complex.

**Figure 7.** Spatial distribution of Cr (mg/kg) in the soil of bauxite mining area, Kuantan, Pahang

**Figure 8.** Spatial distribution of As (mg/kg) in the soil of bauxite mining area, Kuantan, Pahang
under its favorable conditions (Bodek et al., 1988). However, at low pH, As metal can co-precipitate with or absorbed onto iron oxyhydroxides (Smith et al., 1995). Under the extreme reducing state, elemental arsenic and arsine, $\text{AsH}_3$, can be present and cause risk effects to the environment, especially when it enters the water bodies due to oxidations and transformations into more soluble forms (Simpson & Spadaro, 2016).

Distributions of Cd showed a similar trend, where higher Cd was observed at the western part of the region with higher concentration focused at the southwestern area. A total of 100% of the study area had an elevated content of Cd with a value exceeding 0.30 mg/kg. As the region is covered with both bauxite mining and oil palm cultivated area, elevated Cd might be derived from inorganic fertilizer, such as phosphate-based ores. It has been reported that 90% of Cd released to the environment was from anthropogenic resources (Hamid et al., 2019).

The distributions of Ni in the study area showed the same pattern as other HMs discussed. A total of 100% of the region had Ni content more than the suggested value of 45 mg/kg. The Cr content was higher at the western and southwestern parts of the study area, with more than 150 mg/kg. The content of Ni is reduced as it moves towards the eastern part of the studied region. It is interesting to be highlighted that the pattern in which the Ni is concentrated was at the point which bauxite mining was located (Figure 1). Therefore, it can be hypothesized that Ni distribution is related to the mining operation in the studied region.

Distributions of Cr were heterogeneous throughout the study area with values exceeding 60 mg/kg. A total of 100% of the study area had high Cr content in the soil. The distribution of Cr was similar to As, Cd, and Ni, where its content is concentrated at the southwestern
part of the study area, the center of the mining site. However, its distribution showed a
diffused trend as it moved towards the upper part of the study area. This finding is consistent
with the findings of Ismail et al. (2018), which found Cr values ranging from 176 to 703
mg/kg on tropical bauxite soil around the Kuantan region. Continuous and extensive
anthropogenic activities such as mining can eventually lead to the increase of Cr pollution
in soil (Kien et al., 2010; Krishna et al., 2013). Besides the natural weathering processes,
contaminations can be indirectly accelerated by humans through the exposure of soil
surface to the water and air. It could fasten the mineral breakdowns and transform them
into a soluble metal form. Nevertheless, researchers also reported that fertilizers, especially
phosphate fertilizer application, had become an important factor affecting Cr accumulation
in soil (Vogel et al., 2020). Long-term application of phosphate fertilizer could be one of
the factors that contribute to Cr contaminations as oil palm plantations surround it.

CONCLUSION
The average content of Cd, Cr, and Ni in the soil of the study area was higher than the
Dutch target values for soil protection and the 95% ‘Investigation Levels’ determined for
Malaysia soil. Results of I_{geo} showed conclusive evidence that Cd had the highest level of
contamination, followed by Cr, As, and Ni. Source identification based on the combined
methods of matrix correlation, PCA, and geostatistic indicated that excess Cd, Cr, and Ni
were most likely derived by the combined factors of parent materials, topography, and
anthropogenic related activities, which are the mining and long-term agricultural practices.
The combined analyses of matrix correlation, PCA, and geostatistics in monitoring HMs
distributions were effective in evaluating regional soil quality as it supports research
findings with analytical evidence. Geospatial mapping showed that the distributions of As,
Cd, Cr, and Ni were higher in the western and southwestern parts of the region, indicating
more attention should be addressed to this area. In addition, mitigation measures should be
developed to minimize secondary contamination caused by the mining activities, resulting
in soil degradation and adversely impacting oil palm productivity.

ACKNOWLEDGMENT
This research was funded by the Ministry of Higher Education through Fundamental
Research Grant Scheme for Research Acculturation of Early Career Researchers (FRGS-
RACER), (RACER/1/2019/WAB01/UIAM//1).

REFERENCES


Webster, R., & Oliver, M. (2001). *Geostatistics for Environmental Scientists Statistics in Practice*. Wiley and Sons Ltd.


