Estimating Source Apportionment of Heavy Metals Contamination in Surface Soil Based on a Positive Matrix Factorization (PMF) Model around Cerrito Blanco in San Luis Potosi, Mexico †

Arnab Saha 1,*, Bhaskar Sen Gupta 1, Sandhya Patidar 1 and Nadia Martínez-Villegas 2

1 Institute of Infrastructure and Environment, School of Energy, Geoscience, Infrastructure & Society, Heriot-Watt University, Edinburgh EH14 4AS, Scotland, UK; b.sengupta@hw.ac.uk (B.S.G.); s.patidar@hw.ac.uk (S.P.)
2 Applied Geosciences Department, Instituto Potosino de Investigación Científica y Tecnológica, San Luis Potosi 78216, Mexico; nadia.martinez@ipicyt.edu.mx
* Correspondence: as2059@hw.ac.uk
† Presented at the 4th International Electronic Conference on Geosciences, 1–15 December 2022; Available online: https://sciforum.net/event/IECG2022.

Abstract: In recent decades, heavy metal contamination in soils has caused global concern. Quantitative apportionment of heavy metal sources in the surface soil is a complex task. This study indicated a receptor model to evaluate the heavy metal concentrations of various sources for the soil and the related contamination impacts. In this study, the surface soil at the Cerrito Blanco in San Luis Potosi, Mexico was chosen as the case study location to reveal the potential pollution sources of heavy metals. The research suggested the combined use of the positive matrix factorization (PMF) model for the quantitative assessment of contamination sources as well as the spatial distribution techniques for the estimation of the pollution sources. This approach forms the basis for later soil contamination control and treatment. Throughout the study region, a total of thirty-nine samples of surface soil were collected. However, the mean concentration levels of Co, Cr, Cu, Ni, and Pb in the soils were lower than the permissible standards. It was observed that As and Cd were higher than their permissible standard values by around 5.43 and 1.19 times, respectively. The PMF findings demonstrate that the three main diverse sources of heavy metals in this study area’s soils were natural, past mining activities, and industrialisation, as well as groundwater. The concentrations of heavy metals in surface soils were considerably influenced by natural sources, which were generally the main contributing factor. The spatial distribution of soil contamination for heavy metals was also mapped using the geographic information system (GIS) technique.

Keywords: heavy metal; source apportionment; PMF; soil contamination; GIS

1. Introduction

Heavy metal background levels in soils are mostly derived from parent rocks. Heavy metals accumulate in the soil through atmospheric deposition, sedimentation, sewage irrigation, as well as other routes as a result of excessive industrial and agricultural activities [1,2]. However, there are several geological and anthropogenic sources of heavy metal contaminants, including air deposition, industrial discharges, fertiliser application, pesticide usage, etc., that might contribute to their presence in the soil in urban–rural peripheral zones [3,4]. Therefore, it is essential to evaluate heavy metal contamination along with necessary toxicological assessment.

Source apportionment is critical for adequately controlling metal contaminant inputs and determining an acceptable criterion for soil remediation is facilitated by a thorough assessment of heavy metals background concentrations [5–7]. There is a common assumption...
that the two main sources of heavy metals are natural processes and human activities [8].
The geological parent material primarily controls the natural source apportionments of
soil heavy metals [9]. Additionally, the anthropogenic sources of soil heavy metals include
mining, agriculture, vehicle emissions, ore deposits, industrial emissions, coal burning,
atmospheric deposition, etc. [4,8,10]. To determine the source identifications and the spatial
variations of heavy metals in soils, multivariate statistics are implemented in combina-
tion with geostatistical techniques [11,12]. There are also many receptor models which
are used to identify the sources of heavy metals. The most prominent and widely used
receptor models are principal component analysis/absolute principal component scores
(PCA/APCS), positive matrix factorization (PMF), and UNMIX [4,13,14]. In the early 1990s,
source identification of pollutants in the atmosphere was done using positive matrix fac-
torization (PMF) [15]. Additionally, in recent years, it has been used successfully to assign
source contributions in sediments, soils, and aquatic systems [8,16]. Therefore, this study
used the PMF model to determine the sources of seven heavy metals in the surface soils of
Cerrito Blanco, Matehuala, Mexico.

In this study, the source apportionment of arsenic (As), cadmium (Cd), cobalt (Co),
chromium (Cr), copper (Cu), nickel (Ni), and lead (Pb) in the surface soil are evaluated
using the PMF source identification receptor model. The specific objectives of this study
are as follows:

(i) to estimate the heavy metal concentrations in the surface soil around Cerrito Blanco,
    Matehuala, San Luis Potosi, Mexico;
(ii) to identify the possible pollution sources of the heavy metals using the PMF model;
(iii) to analyse the spatial distribution patterns of source factors of heavy metals.

2. Study Area

The selected study area is located between 23°40'30" N latitude and 100°35'27" W
longitude around Joya Verde soccer sports club, a non-cultivated farming area of Cerrito
Blanco, Matehuala municipality, San Luis Potosi, Mexico, with a total land area of approx.
4.84 ha (Figure 1). The soccer club is around 14 years old, and it includes three soccer fields.
It is surrounded by semi-arid vegetation and underdeveloped agricultural farmland [10,17].
The mean annual temperature in this area is 14° celsius. The rainy season lasts from June
to September, with a 300 mm average rainfall [17–19].

Figure 1. The study area map showing the distribution of soil sampling points.
3. Materials and Methods
3.1. Soil Sampling and Chemical Analysis

To evaluate the source apportionment of heavy metals in Cerrito Blanco, Matehuala area, 39 surface soil samples were collected using an auger at the depth of 0–5 cm from soccer fields and scrub areas around the fields. A proportion less than 2 mm was obtained by sieving the soil samples after they were air-dried at room temperature. According to a considerably modified version of ISO 11466:1995, surface soil samples were digested [20,21]. Next, 10 mL of aqua regia (HNO₃:HCl, 3:1) was then added to a beaker containing 1.0 gm of soil sample. The total accessible heavy metals in soils may be determined using this digestion process [17,22]. Inductively coupled plasma optical emission spectroscopy (ICP-OES) was used to determine the heavy metal concentrations in the surface soil [17,23]. For each sample site, a Garmin Etrex personal navigator GPS was used to record the location’s coordinates using the WGS-1984 coordinate system.

3.2. Positive Matrix Factorization Model (PMF) Model

PMF is a receptor model that the United States Environmental Protection Agency has recommended to analyse the different types of sources of environmental pollution [8,24,25]. This model was developed by Paatero and Tapper (1994) [15]. A matrix of speciated sample data is divided by the PMF model into two matrices: factor profiles ($f$) and factor contributions ($g$) [26]. In this research, the PMF 5.0 model was implemented to source identification of heavy metals in the surface soils.

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij}$$

where $x_{ij}$ is a quantitative matrix for the $j$th heavy metal component in $i$ samples; $g_{ik}$ denoted a contribution matrix of the $k$th source factor; for the $k$th source factor, $f_{kj}$ represents a source factor of the $j$th heavy metal component, and the residual value for the $j$th metal element over $i$ numbers of samples is denoted by the symbol $e_{ij}$.

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left( \frac{e_{ij}}{u_{ij}} \right)^2$$

where $Q$ is the difference (i.e., $e_{ij}$) between the $x_{ij}$ and the output of $g_{ik} f_{kj}$, weighted by the uncertainty measurement $u_{ij}$.

Two input files are needed for the PMF model: soil sample concentration values and soil sample uncertainty values. The following equation was used to estimate the uncertainty values:

$$u_{ij} = \frac{5}{6} \times \text{MDL (for } c \leq \text{MDL)}$$

$$u_{ij} = \sqrt{(\text{error fraction } \times c)^2 + \text{MDL}^2} \text{ (for } c > \text{MDL)}$$

where $c$ is the concentration values of soil samples, MDL is the species-specific method detection limit, and the error fraction represents the uncertainty measurement percentage.

3.3. Data Processing and Analysis

The descriptive statistical analysis of the soil samples was performed with SPSS Statistics version 28. ESRI’s (Redlands, CA, USA) ArcGIS software was used for spatial distribution mapping of source factors of heavy metals. The source identification of the heavy metals was quantitatively determined using the EPA PMF 5.0 software (United States Environmental Protection Agency - Positive Matrix Factorization (USEPA-PMF), Washington, DC, USA).
4. Results

4.1. Descriptive Statistics of Heavy Metals Concentrations in Soils

Table 1 shows the comprehensive statistical results for seven heavy metals in the surface soil of Cerrito Blanco. The average concentrations of all heavy metals increased in the following sequence: Co < Cd < Cr < Ni < Cu < Pb < As. The concentrations of all heavy metals in the soil are lower than the permissible limit or metal reference values, except As and Cd. The concentration of As in the soil is about 5 times higher than the permissible limit, which means that the level of As contamination in this soil is very high. The permissible limits of heavy metals in the surface soil were As (22 mg/kg) [17], Cd (0.80 mg/kg) [27], Co (40 mg/kg) [28], Cr (100 mg/kg) [27], Cu (36 mg/kg) [21], Ni (35 mg/kg) [27], and Pb (85 mg/kg) [21]. All the permissible limit values given in this study are taken as per the global standard values of different places published in previous studies. The reference values of heavy metals in the surface soils of Mexico were As (22 mg/kg), Cd (37 mg/kg), Co (not regulated), Cr (280 mg/kg), Cu (not regulated), Ni (1600 mg/kg), and Pb (400 mg/kg) [29]. The coefficient of variation (CV) is classified as low (CV < 16), moderate (16 < CV < 36), and high (CV > 36) [25,30]. The CV of As is 91.71, and the degree of variation is very high, which is more than other metals. However, there was greater variation and spatial dispersion in the soil of this study area for heavy metals.

<table>
<thead>
<tr>
<th>Metals</th>
<th>As</th>
<th>Cd</th>
<th>Co</th>
<th>Cr</th>
<th>Cu</th>
<th>Ni</th>
<th>Pb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>119.44</td>
<td>0.95</td>
<td>0.76</td>
<td>2.96</td>
<td>20.65</td>
<td>3.20</td>
<td>36.95</td>
</tr>
<tr>
<td>Standard Error</td>
<td>17.54</td>
<td>0.10</td>
<td>0.10</td>
<td>0.37</td>
<td>1.56</td>
<td>0.30</td>
<td>3.97</td>
</tr>
<tr>
<td>Median</td>
<td>90.51</td>
<td>0.94</td>
<td>0.69</td>
<td>2.49</td>
<td>18.10</td>
<td>3.07</td>
<td>30.86</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>109.54</td>
<td>0.65</td>
<td>0.65</td>
<td>2.28</td>
<td>9.75</td>
<td>1.87</td>
<td>24.79</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>11,998.65</td>
<td>0.42</td>
<td>0.43</td>
<td>5.21</td>
<td>95.04</td>
<td>3.49</td>
<td>614.63</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>8.37</td>
<td>-1.09</td>
<td>-0.74</td>
<td>4.85</td>
<td>3.63</td>
<td>0.93</td>
<td>5.73</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.43</td>
<td>0.21</td>
<td>0.53</td>
<td>2.11</td>
<td>1.68</td>
<td>0.93</td>
<td>2.12</td>
</tr>
<tr>
<td>Range</td>
<td>578.17</td>
<td>2.18</td>
<td>2.19</td>
<td>10.82</td>
<td>47.85</td>
<td>8.13</td>
<td>126.30</td>
</tr>
<tr>
<td>Minimum</td>
<td>13.14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.28</td>
<td>7.88</td>
<td>0.24</td>
<td>8.99</td>
</tr>
<tr>
<td>Maximum</td>
<td>591.31</td>
<td>2.18</td>
<td>2.19</td>
<td>11.10</td>
<td>55.73</td>
<td>8.37</td>
<td>135.29</td>
</tr>
<tr>
<td>Sum</td>
<td>4658.01</td>
<td>37.12</td>
<td>29.73</td>
<td>115.30</td>
<td>805.17</td>
<td>124.90</td>
<td>1440.99</td>
</tr>
<tr>
<td>Coefficient of variation (CV) (%)</td>
<td>91.71</td>
<td>68.06</td>
<td>85.77</td>
<td>57.72</td>
<td>47.22</td>
<td>58.32</td>
<td>67.10</td>
</tr>
<tr>
<td>Count</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>Confidence Level (95.0%)</td>
<td>35.51</td>
<td>0.21</td>
<td>0.21</td>
<td>0.74</td>
<td>3.16</td>
<td>0.61</td>
<td>8.04</td>
</tr>
</tbody>
</table>

4.2. Source Apportionment of Heavy Metals by PMF

The PMF model was performed to better understand the sources of heavy metals in the soil along the Cerrito Blanco, Matehuala and to evaluate their apportionment to each of the heavy metals (Figure 2). The model was run 20 times with the factors set to 3, 4, and 5. The Q value was the lowest, and the majority of the residual fell between −3 and 3 [24] for the three-factor analysis. The contributions of heavy metals in factor 1 increased in the following order: As < Cr < Cu < Ni < Pb < Cd < Co (Figure 3). The Cr, Ni, and Cd concentrations, which account for 50.76%, 32.44%, and 38.75% of factor 2, respectively, are the major contributions, with limited impact from other metal sources. Factor 3 was dominated by As (94.02%), Cu (34.74%), and Cd (23.46%) (Figure 3).

Factor 1 showed a significant positive load of Co, Cd, Ni, Pb, Cu, and Cr in the source contributions and were close or below to the permissible limit standards derived from the surface soil. Therefore, the results were revealing that factor 1 is a significant lithogenic source. The concept of ionic reaction may be able to explicate the clustering characteristic resulting from PMF, since the concentration values and correlations of heavy metals may fluctuate across various parent materials. According to the spatial distribution mapping of factor 1, the study area’s western side which is surrounded by semi-arid vegetation contains the majority of the high-value locations (Figure 4). This result was in line with the
concentration–distribution patterns of Co, Cd, Cu, and Ni where significant concentrations of these heavy metals were found close to each other. The average concentrations of Co, Cd, Cu, Ni, and Pb were close to their respective permissible limits; hence, they could be of natural origin. Therefore, factor 1 might be classified as a natural source.

Figure 2. Source profiles and source contributions of heavy metals in the soil from PMF.

Figure 3. Factors fingerprint of each heavy metal contributing to the different pollution sources.

Factor 2 was associated with Cr, Ni, and Cd (Figure 3). According to the results, these three metals contributed to this source factor more significantly than the other four metals. Previous research revealed that non-ferrous metal smelting, smelter slags, mining, geochemical exploration, and other industrial operations were the sources of the heavy metal contamination found in Mexican soils [10,31,32]. The major sources of Cd have been identified as of metallurgical origin [8]. Industrial emissions are the only source of Cr and Ni. The spatial distribution of factor 2 shows that the middle portion of the study area of the soccer field has the majority of the high-value locations (Figure 4). As a result, it is possible to confirm that factor 2 represents the mining and industrial sources.

Figure 3. had more significant contributions of As, Cu, and Pb than the other three sources. The As is mostly associated with irrigation water and is observed in a considerable area of soccer fields and the surrounding study area. Studies conducted earlier revealed that the water in the study area had incredibly high quantities of heavy metals, which are
caused by the dissolution of metallurgical waste from an abandoned smelter upstream of Matehuala [17–19]. Unfortunately, this contaminated water is the main water source for local agriculture [17,20]. Inorganic As, Cu, and Pb chemicals such as sodium arsenate, lead arsenate, calcium arsenate, cattle dung, animal waste, and fertilisers were utilised as herbicides or pesticides in the surrounding farmland, which probably has an impact on the groundwater. The spatial distribution pattern of factor 3 revealed that the soccer field area with high local concentrations is mostly affected by irrigation (Figure 4). Therefore, it might be concluded that factor 3 reflected the groundwater sources.

Figure 4. Spatial distribution of PMF source factors in the surface soil: (a) Factor 1; (b) Factor 2; (c) Factor 3.

5. Conclusions

This study measured the contamination risks by descriptive statistical analysis approach, identification of various sources of heavy metals in the surface soil using the PMF model, and observed the spatial distribution patterns of source factors based on an IDW interpolation technique. This study concluded that the mean concentrations of Co, Cr, Cu, Ni, and Pb in the surface soils in Cerrito Blanco were lower than their permissible limits. But the mean concentrations of As and Cd were higher than their permissible limits. The PMF results showed that the identification of source contribution for heavy metals in the surface soil was as follows: 53% for natural sources, 23% for mining activities and industrialization sources, and 24% for irrigational groundwater sources. However, the spatial distribution patterns of source factors and model outcomes revealed that Co, Cd, Cu, Ni, and Pb originated from natural sources; Cr, Cd, and Ni may be obtained from past mining and industrialization; and As and Cu mainly came from groundwater sources. The results of this study are relevant to the growing concern for the human risk of As exposure from the surface soil in recreational and agricultural areas.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/IECG2022-13746/s1.
Author Contributions: Conceptualization, A.S. and B.S.G.; methodology, A.S. and B.S.G.; software, A.S.; formal analysis, A.S., B.S.G., S.P. and N.M.-V.; investigation, A.S. and B.S.G.; resources, A.S., B.S.G. and S.P.; data curation, B.S.G. and N.M.-V.; writing—original draft preparation, A.S.; writing—review and editing, A.S., B.S.G., S.P. and N.M.-V.; visualization, A.S.; supervision, B.S.G. and S.P.; project administration, B.S.G.; funding acquisition, B.S.G. and N.M.-V. All authors have read and agreed to the published version of the manuscript.

Funding: The project has received funding from the Engineering and Physical Sciences Research Council (EPSRC-IAA), UKRI through Project 748502: Spreading the Chemical Free Arsenic Re-moval Technology to Various Communities in India, Bangladesh, and Mexico.

Data Availability Statement: The data and information cited in the paper can be found in the re-reported literature as references.

Acknowledgments: The authors are thankful to the Institute of Infrastructure and Environment, The School of Energy, Geoscience, Infrastructure and Society (EGIS), Heriot–Watt University, Edin-burgh for providing support and a student bursary to the first author for doctoral research through the James Watt Scholarship. The authors also thank IPICyT, San Luis Potosi, Mexico for providing the necessary data, feedback, and support. We used some heavy metals data and GIS shapefile data while working on this study, and the data has been published previously in the International Journal of Environmental Research and Public Health, MDPI in 2018 and Journal of South American Earth Sciences, Elsevier in 2022 for a different type of study on toxic metal contamination. References to these studies are cited in the inner text.

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

References


27. Dí, O.; Aa, A.; Oe, O. Heavy Metal Concentrations in Plants and Soil along Heavy Traffic Roads in North Central Nigeria. *J. Environ. Anal. Toxicol.* 2015, 5, 1. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.