DOI: 10.5937/IOC25611K *Original scientific paper*

MULTIVARIATE STATISTICAL EVALUATION OF AGRICULTURAL SOIL CONTAMINATION NEAR THE CITY OF BOR

Vesna Krstić[1a,2], Danijela Simonović[1b], Renata Kovačević[1c], Daniel Kržanović[1d], Mirjana Šteharnik[1e], Daniela Urošević[1f], Sunčica Stanković[3a]

[1] Mining and Metallurgy Institute Bor, Alberta Ajnštajna 1, 19210 Bor, Serbia
[2] Technical Faculty Bor, University of Belgrade, V.J. 12, 19210 Bor, Serbia
[3] Faculty of Management, Metropolitan University, 11158 Belgrade, Serbia
[1a,2] vesna.krstic@irmbor.co.rs, 0000-0002-1110-0547
[1b] danijela.simonovic@irmbor.co.rs, 0000-0001-7945-693X
[1c] renata.kovacevic@irmbor.co.rs, 0000-0003-4108-0855
[1d] danijel.krzanovic@irmbor.co.rs, 0000-0003-3841-8667
[1e] mirjana.steharnik@irmbor.co.rs, 0000-0002-2733-3599
[1d] daniela.urosevic@irmbor.co.rs, 0000-0003-3490-7834
[3a] suncica.stankovic@metropolitan.ac.rs, 0000-0001-8824-1973

Abstract

This study investigates the concentrations of heavy metals in agricultural soils surrounding the Zijin Copper Bor mining complex in Eastern Serbia, focusing on four locations: Veliki Krivelj, Oštrelj, Brezonik, and Zlot. The presence of the following elements was analyzed using inductively coupled plasma mass spectrometry (ICP-MS): Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Cd, and Pb. To provide a comprehensive assessment of contamination levels and inter-element relationships, the following statistical methods were applied: descriptive analysis, principal component analysis (PCA), correlation analysis with heatmap visualization, and hierarchical cluster analysis (HCA). PCA results revealed latent factor groupings that reflect the influence of mining activities, metallurgical processes, and contaminant migration within the soil matrix.

Keywords: *soil contamination, heavy metals, ICP-MS, multivariate statistical analysis, PCA, HCA*

1. INTRODUCTION

Soil contamination by heavy metals represents one of the most pressing environmental challenges in urban and industrial areas worldwide. Mining operations and metallurgical activities—particularly in regions with intensive smelting—lead to the accumulation of potentially toxic elements in arable soils, directly threatening agroecological functions, biodiversity, and human health [1–3]. The city of Bor in Eastern Serbia, with over a century of mining tradition, is considered one of the most polluted areas in the region due to emissions of arsenic, cadmium, lead, copper, and other metals from ore processing facilities [4,5]. Numerous studies have reported concentrations of these metals far exceeding permissible threshold values, especially in zones adjacent to the smelter [6,7].

Principal Component Analysis (PCA) has proven to be an effective statistical tool for identifying and characterizing metal distributions, enabling differentiation of pollution sources and reduction of data complexity [8,9]. When integrated with Geographic Information Systems (GIS), PCA further facilitates spatial visualization of contamination risk and the delineation of critical zones for remediation [10].

The aim of this study is to apply PCA to a dataset of heavy metal concentrations in soils surrounding the Bor mining complex, with a focus on identifying dominant contamination factors and their spatial distribution. The results may serve as a foundation for defining remediation strategies and sustainable land management practices in mining-affected regions.

2. EXPERIMENTAL

Soil sampling was conducted at four representative locations: Veliki Krivelj (L1), Oštrelj (L2), Brezonik (L3), and Zlot (L4), all situated in close proximity to the mining and smelting complex. The sites were selected based on prior contamination reports, their geographic position relative to sources of mining waste, and the spatial distribution of mining infrastructure. The sampling plan was developed in accordance with ISO 18400-101:2017. All procedures were carried out by accredited personnel from the Institute for Mining and Metallurgy Bor (IRM Bor), in compliance with ISO standards 18400-102:2017, 18400-104:2018, and 18400-105:2017.

3. RESULTS AND DISCUSION

Descriptive statistical analysis of heavy metal concentrations in agricultural soils surrounding the Zijin Copper mine in Bor, Serbia—covering the locations Veliki Krivelj, Oštrelj, Brezonik, and Zlot (Table 1)—reveals pronounced variability in Fe, Mn, and Cu levels. The distributions of most metals are mildly asymmetric and platykurtic (kurtosis < 0), indicating a flattened shape relative to the normal distribution.

Table 1 – Descriptive statistics of heavy metal concentrations in agricultural soils at Veliki Krivelj, Oštrelj, Brezonik, and Zlot

	mean	std	min	25%	50%	75%	max	skewness	kurtosis
Cr	32.9	17.3	15.9	19.3	33.4	47.1	49.2	-0.021	-1.938
Mn	973.1	447.9	482.7	650.3	1007.5	1330.2	1394.5	-0.104	-1.817
Fe	39235.0	10805.8	29982.4	30520.3	37364.2	46078.9	52229.4	0.272	-1.641
Co	16.6	7.8	11.0	11.7	13.7	18.6	28	0.957	-0.845
Ni	14.9	5.6	9.9	10.7	14.1	18.3	21.7	0.294	-1.583
Cu	385.5	169.9	192.7	312.8	371.9	444.7	605.5	0.27	-1.016
Zn	79.5	9.9	72.9	73.7	75.5	81.3	94.1	1.052	-0.755
As	28.8	6.3	20.4	25.7	30.5	33.7	33.8	-0.544	-1.311
Cd	-0.132	0.034	-0.149	-0.149	-0.149	-0.132	-0.081	1.155	-0.667
Pb	36.8	10.8	28.8	30.2	32.9	39.5	52.6	0.962	-0.832

The strength of association between heavy metals in soils surrounding the Zijin Copper mine (Bor, Serbia) was assessed using correlation analysis. This method evaluates both the direction and magnitude of the relationship between two variables. The correlation coefficient (r) quantifies the degree of statistical interdependence between two random variables (x and y), ranging from –1 to +1. A value of zero indicates no significant association.

According to Evans [11], the absolute value of the correlation coefficient can be interpreted as follows: from 0.00 to 0.19 - very weak, from 0.20 to 0.39 – weak, from 0.40 to 0.59- moderate, from 0.60 0.79 – strong and from 0.80 to 1.00 - very strong. The results revealed very strong positive correlations between Cd–Pb (r = 0.97), Mn–Fe (r = 0.96), and Cu–Pb (r = 0.94), suggesting that these metals may originate from a common pollution source or exhibit similar accumulation behavior in soil. A moderate positive correlation was observed between As–Zn (r = 0.54), indicating partial overlap in their sources. Conversely, very strong negative correlations were found between Cr–Mn (r = –0.97), Cr–Fe (r = –0.91), and Cu–Co (r = –0.85). Strong negative associations were also noted for Co–As (r = –0.78) and Ni–Zn (r = –0.60). Very weak correlations were observed between Ni–As (r = 0.023, positive) and Fe–Ni (r = –0.006, negative), implying negligible association. These findings suggest that certain metals tend to co-occur, while others exhibit opposing behavior, likely due to differing contamination sources or specific geochemical mechanisms in the soil. Additional correlation results, including determination coefficients, are presented in Figure 1.

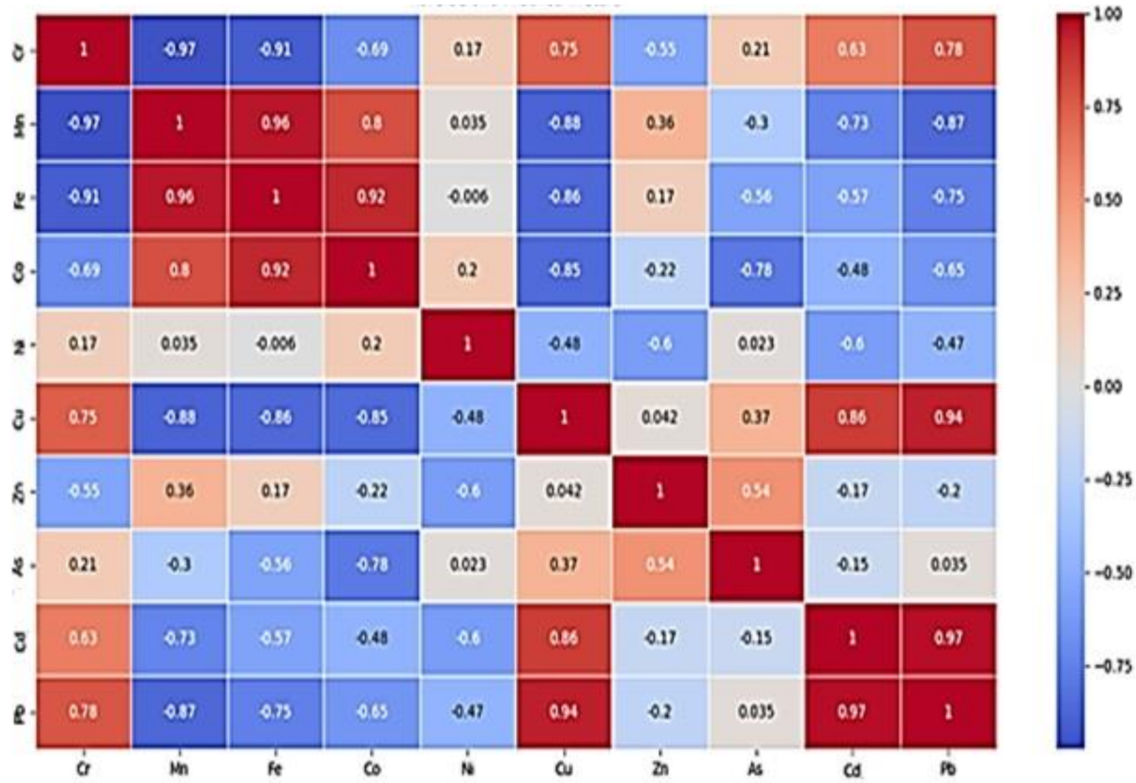

Figure 1 – Correlation Matrix

To identify relatively homogeneous clusters of elements, exploratory factor analysis (EFA) was applied. Factor extraction was based on Kaiser's criterion, which retains only components with eigenvalues greater than 1 [12]. The EFA revealed three components with eigenvalues exceeding 1: the first component accounted for 60.47% of the total sample variance, the second for 21.50%, and the third for 18.03%, as shown in Table 2.

Table 2 – Results of Exploratory Factor Analysis

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.047	60.470	60.470	6.047	60.470	60.470
2	2.150	21.497	81.968	2.150	21.497	81.968
3	1.803	18.032	100.000	1.803	18.032	100.000

The selection of a three-component solution as optimal was further supported by the scree plot criterion (Figure 2), based on Cattell's scree test. The plot displays a sharp decline in eigenvalues between the first and third components, while the remaining components exhibit negligible values (≈ 0), indicating minimal explanatory power. The curve shows a clearly defined "elbow point" after the third component, which reinforces the decision to retain only the first three components.

The results of the confirmatory factor analysis (Principal Component Analysis, PCA) confirmed that the heavy metals are grouped into three distinct components. Each component explains a specific portion of the total variance in the dataset. Component 1: factor loadings range from –0.965 to 0.919. Component 2: loadings range from –0.710 to 0.994. Component 3: loadings range from –0.974 to 0.733. These results are presented in Table 3 – Rotated Component Matrix.

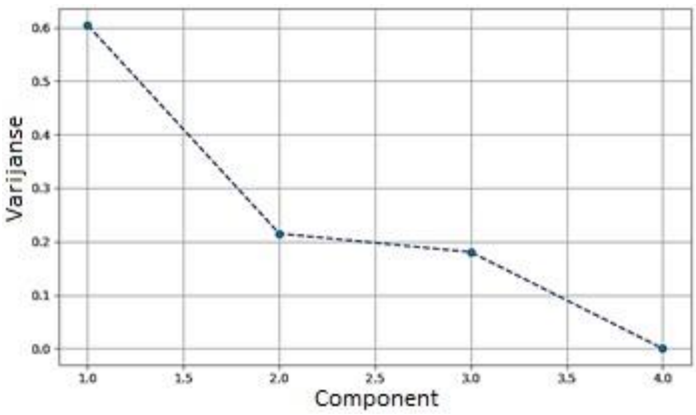

Figure 2 – Scree Plot (Cattell's Criterion)

Table 3 Rotated component matrix

	Component		
	1	2	3
Mn	-.965		
Pb	.963		
Cr	.919		
Cu	.914		
Cd	.879		
Fe	-.863		
As		.994	
Co		-.710	
Ni			-.974
Zn			.733

It can be concluded that all metals within the first component exhibit high factor loadings, suggesting a common origin or source of contamination. This component likely reflects a shared process or pollution source—such as mining-related activities. As shown in Table 3, the first component includes metals directly associated with copper mining operations (Pb, Cu, Cd, Cr). Mn and Fe display negative loadings, indicating an inverse relationship with the other metals in this component, which may suggest their affiliation with the natural geochemical background. The second component is dominated by As (0.994) and Co (–0.710), both showing strong loadings. This component likely represents a distinct source or process linked to arsenic and cobalt contamination. The negative loading of Co implies an opposing behavior relative to As, possibly due to differing geochemical pathways or mobility in soil. The third component comprises Ni (–0.974) and Zn (0.733). Led by Ni, this component may reflect diffuse migration processes of metals into deeper soil layers. The PCA results indicate that Cu, Pb, Cd, and Cr are strongly associated with the first component, Zn and As with the second, and Ni with the third. This separation suggests the presence of multiple contamination factors: mining origin, industrial emissions, and geochemical behavior of metals in soil.

Thus, the PCA identifies three components that explain the presence and behavior of heavy metals in the soil: mining influence, secondary industrial processes, and subsurface migration of contaminants. These findings enable differentiation of impact zones, which is essential for targeted ecological monitoring and land management in the vicinity of the mining complex.

The dendrogram (Figure 3) reveals two distinct clusters: locations L1 and L2 form one group, while L3 and L4 form another. These clusters are clearly separated. L1 and L2 likely share a similar contamination profile, possibly due to a common pollution source such as similar geological substrate or agricultural practices. In contrast, L3 and L4 are internally similar but differ from the first group—potentially influenced by other environmental factors such as wind dispersion or topography.

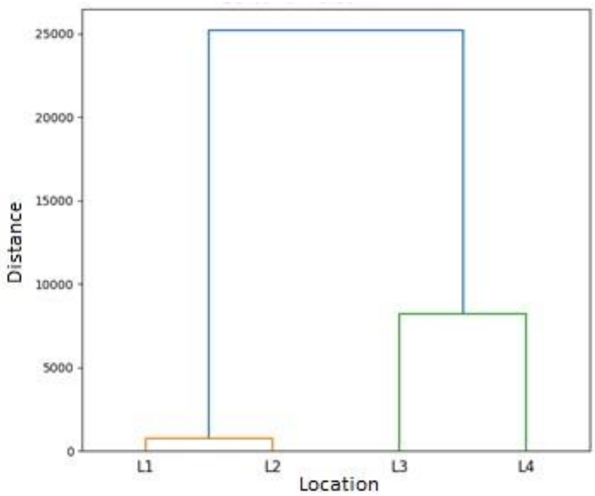

Figure 3 – Dendrogram of Sampling Locations

4. CONCLUSION

The application of multivariate statistical methods enabled the identification of heavy metal contamination in agricultural soils surrounding the Zijin Copper mining complex in Bor, Serbia. Principal Component Analysis (PCA) revealed distinct factor groupings indicative of dominant mining and metallurgical influences, while correlation and cluster analyses confirmed inter-element relationships and spatial differentiation of contamination patterns. The findings underscore the necessity for continuous environmental monitoring and integrated risk assessment in agroecological zones exposed to industrial pressure.

ACKNOWLEDGEMENTS

The research presented in this paper was done with the financial support of the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, within the Contract No. 451-03-136/2025-03/200052.

REFERENCES

[1] B. Tang, C. Christia, B.-X. Mai, Environ. Pollut., 306 (2022) 119469.
[2] O. Adedeji, A. Olusola, G. James, H.A. Shaba, I.R. Orimoloye, S.K. Singh, S. Adelabu, Environ. Monit. Assess., 192 (12) (2020) 798.
[3] Y.B. Sun, Q.X. Zhou, X.K. Xie, R. Liu, J. Hazard. Mater., 174 (1-3) (2010) 455-462.
[4] Đ. Nikolić, N. Milošević, Ž. Živković, I. Mihajlović, R. Kovačević, J. Serb. Chem. Soc., 75 (4) (2010) 625-641.
[5] V. Tasić, R. Kovačević, B. Maluckov, T. Apostolovski-Trujić, B. Matić, M. Cocić, M. Šteharnik, Water Air Soil Poll., 228 (6) (2017) 230
[6] R. Kovačević, V. Tasić, T. Apostolovski-Trujić, A. Ivanović, Atmosphere, 15(12) (2024) 1549.
[7] S. Šerbula, T. Kalinović, J. Kalinović, A. Ilić, ESPR, 20 (12) (2013) 8919-8929.
[8] Gergen, M. Harmanescu, hem. Cent. J.6 (1) (2012) 156.
[9] C. Kahangwa, J. Munganga, B. Mbuyu, GEP, 10 (4) (2022) 1-14.
[10] Y. Gonzales-Castanedo, T. Moreno, R. Fernández-Camacho, et al., Atmos. Environ., 92 (2014) 130-137
[11] J.D. Evans, Straightforward statistics for the behavioral sciences. Thomson Brooks/Cole Publishing Co. (1996).
[12] B.G. Tabachnick, L.S. Fidell, Using Multivariate Statistics (6th Edition). NY: Pearson Education, Inc. (2013).