

Vol. 23

Original Research Paper

doi) https://doi.org/10.46488/NEPT.2024.v23i03.009

2024

Heavy Metal Contamination of Surface Sediments-Soil Adjoining the Largest Copper Mine Waste Dump in Central India Using Multivariate Pattern Recognition Techniques and Geo-Statistical Mapping

Anoop Kant Shukla*†10, Manoj Pradhan* and Onkar Nath Tiwari**

*Mining Engineering Department, National Institute of Technology, Raipur-492010, Chhattisgarh, India **Hindustan Copper Limited; Kolkata-700019, India †Corresponding author: Anoop Kant Shukla; anoop5052@gmail.com

Nat. Env. & Poll. Tech. Website: www.neptjournal.com

Received: 10-12-2023 Revised: 20-01-2024 Accepted: 23-01-2024

Key Words: Leaching Sediment contamination Soil contamination Sulphide ore mining Hierarchal cluster analysis

ABSTRACT

This detailed study assessed heavy metal contamination of sediments/soil near central India's largest copper mining area using 38 sampling sites within 10 km of the mine using atomic absorption spectroscopy. This study utilized multivariate pattern recognition methods, namely hierarchical clustering analysis (HCA) and principal component analysis (PCA), for source identification. Twelve parameters, i.e., copper (Cu), manganese (Mn), cobalt (Co), zinc (Zn), nickel (Ni), lead (Pb), organic matter (OM), cation exchange capacity (CEC), soil pH, distance (D), and elevation (E) were analyzed. The hierarchical cluster analysis (HCA) was used to analyze the sample sites with similar metal contamination and principal component analysis (PCA) was used to analyze the relationship between the parameters as well as to identify sources of heavy metal pollution. Three major pollution hotspots were detected by AHC and were classified as unpolluted/low pollution sites (UPS: mean concentration factor of 1.35 for Cu), highly polluted sites (HPS: mean concentration factor of 22 for Cu), and extremely polluted sites (EPS: mean concentration factor of 74 for Cu). PCA revealed three hidden factors/components, namely PC1 (explaining 38% of the variability), PC2 (18% of the variability), and PC3 (14% of the variability). Metals showed strong positive loading in PC1, explaining the highest variability. The mean content of Cu in soil/ sediment samples was 502.526 mg/kg. The mean copper content was 10 times higher than the natural crustal value of 45mg/kg, indicating severe pollution in several sites around the study area. Mapping of copper contamination was conducted to reveal the spatial distribution of copper contamination using QGIS. This study exposes the heavy metal contamination level in surface sediments/soil and the effectiveness of pattern recognition techniques for the assessment of multivariate datasets in discerning spatial disparities and identifying the contamination causes.

INTRODUCTION

The soil can store anthropogenic and natural contaminants. Under some conditions, soils can discharge hazardous compounds into the environment, contaminating groundwater food chains (Lu & Bai 2010, C. S. C. Wong et al. 2006). Heavy metals like Pb, Cu, Co, Ni, and others were found at elevated amounts in soils in numerous cities worldwide (Ajmone-Marsan & Biasioli 2010, Biasioli et al. 2006, Chen et al. 2005, Chirenje et al. 2004, Jarva et al. 2009, Lee et al. 2006). Location and concentration data alone cannot uncover hidden heavy metal linkages and distribution patterns. Univariate statistics such as mean, median, mode, etc., cannot discern spatial patterns in soil/sediment heavy metal distribution. Spatial impact studies use multivariate statistics such as Agglomerative Hierarchal Clustering (AHC) and Principal Component Analysis (PCA). These methods can be used as suitable tools to identify the sources of contamination (Facchinelli et al. 2001, Li et al. 2013, Tariq et al. 2008). Mining of minerals causes the removal of rock located beneath the earth's surface. When the mined mineral is brought to the surface, the minerals react with air and water (in the form of moisture), leading to various chemical interactions. If the rock mass contains sulfide minerals like iron pyrite (FeS₂), it can react with oxygen and water to produce an acidic discharge known as Acidic Mine Drainage (AMD). Large open pit porphyry copper mines produce copper sulfides (Gordon et al. 2006, Greenwood & Earnshaw 1984). Copper ore mining sites with sulfide content can, therefore, produce AMD (Shukla et al. 2018). Heavy metals in sediments, soil, and water around mining operations come from acid mine drainage (Meadows & Carpenter 1997, Swarnakar et al. 2023). This study investigates

heavy metal contamination levels in surface sediments/soil and the effectiveness of pattern recognition techniques for assessing and identifying the contamination causes around mine waste dumps.

STUDY AREA

The study area is a large central Indian mining project. It has 40% of India's copper reserves (Pandey et al. 2007). The study area covers an open-cast mine and an ore processing plant with two sites for the disposal of waste, namely a mines waste dump (MWD) and tailing storage facility (TSF), as shown in Fig. 1. Chalcopyrite (CuFeS₂) with a grade of 1%copper is the main ore mined. The open pit is 2200 meters in length and 500 m in width (Tiwari et al. 2017). The study region spans latitude 21.9406920 N to 22.0836800 N and longitude 80.6567170 E to 80.7607280 E. Fig. 2 shows two perennial rivers flowing in the neighborhood of the study area known as 'Banjar' River located in the north-east quadrant of the study area and 'Son' river located in the southern quadrant of the study area with sampling sites. Digital elevation model demarcated both the river basins and revealed that there was no cross contamination as the basins showed no intersection between the two rivers.

MATERIALS AND METHODS

Sample Collection and Analysis

A Digital Elevation Model (DEM) was created using Shuttle Radar Topography Mission (SRTM) data files (N21E080. SRTMGL1.hgt and N22E080.SRTMGL1.hgt) obtained from NASA's Earthdata portal. This DEM was developed using QGIS software version LTR 3.10. Sampling locations for rivers and tributaries were chosen based on their natural topographic flow. Additionally, samples from the roadside or exposed soil were collected to analyze any additional pollution source present in the study area Fig. 2 & 3. In



Fig. 1: Study area (Courtesy: Google® Earth).

total, 38 samples were collected in duplicate at a depth of 10 cm from various river and tributary points. To prevent contamination, each sample was collected using a plastic tool and then sealed in a marked polypropylene Ziploc bag. The precise locations of these samples were recorded with GPS. Any pebbles, rock fragments, or plant matter were removed from the samples. Subsequently, the samples were pulverized using an agate mortar, finely crushed, sifted through a 200-mesh screen, and stored in polyethylene containers that had been pre-cleaned with a mixture of nitric acid and distilled water in a 3:1 ratio.

Each sample was dried to constant weight at 108°C. After drying, 1g of each sample was digested according to the technique recommended by the US EPA (Environmental Protection Agency, 1996). The leachate analysis was done using an Atomic Absorption Spectrophotometer (AAS, model AA8000 FG, Lab India). The samples were analyzed for Copper (Cu), Manganese (Mn), Cobalt (Co), Zinc (Zn), Nickel (Ni), Lead (Pb), and Iron (Fe). Soil pH, CEC, and OM were determined using standard methods (BIS 2720-22 2020, BIS 2720-24 2020, BIS 2720-26 2021).



Fig. 2: SRTM digital elevation model clipped to study area and drainage channels.





Fig. 3: (a) to (d) Affected area with green-colored deposits over soil surface adjoining a mine waste dump.

Analysis of Data Set

The data set generated using the methodology mentioned in previous section was subjected to statistical treatment. Since different metals have different crustal abundance, therefore Contamination Factor (CF) was used instead of the actual concentration. It is defined as follows:

$$Cf_x = \frac{C_i}{C_{background}}$$
 ...(1)

Where C_i is the contamination level of the metal 'i' under study, and C _{background} is the background occurrence level of the same metal 'i' (Varol 2011). The parameters with larger concentrations heavily skewed the analysis results, and the order and range of concentrations of the several physicochemical characteristics varied significantly, suggesting that the dataset is not distributed normally. The Shapiro-Wilks (S-W) test was used to test for normality, and it was found that, except for Zn, all metals failed the S-W normality test. Therefore, the Box-Cox (B-C) transformation of the data matrix was conducted to reduce the impact of outliers, and the dataset that had been B-C transformed provided the best match for the S-W normality test (Box & Cox 1964). The S-W test revealed that, with 95% confidence, all the variables for the B-C converted data were normally distributed. Kaiser-Meyer-Olkin (KMO) and Bartlett's tests assessed data appropriateness for PCA. KMO measures sampling adequacy by estimating latent factor variance in variables. Bartlett's sphericity test determines if the correlation matrix is an identity matrix, suggesting unrelated variables. The data analysis was performed in MS Excel® with XLSTAT® (Addinsoft 2023) add-in.

Cluster analysis: CA's fundamental goal is to organize data sets into clusters that are related yet distinct (Razmkhah et al. 2010). AHC is a bottom-up clustering algorithm. The contamination factor dataset was analyzed using Ward's AHC approach (Ward 1963) to find multivariate commonalities between sampling sites at different sampling points. Numerous researchers have reported using the CA approach to evaluate the quality of water (Astel et al. 2007, Hussain et al. 2008, Singh et al. 2005). Similarly, numerous studies have utilized AHC on soil data as well (Dragović et al. 2008, Lamontagne & Camire 1987, Micó et al. 2006, Navas & Machín 2002, Pîrnău et al. 2020, Tume et al. 2006).

	Cu [mg.kg ⁻¹]	Mn [mg.kg ⁻¹]	Co [mg.kg ⁻¹]	Zn [mg.kg ⁻¹]	Ni [mg.kg ⁻¹]	Pb [mg.kg ⁻¹]
Mean	502.526	702.00	24.84	24.21	28.76	23.55
Median	83.5	395.00	23.25	23.00	24.00	19.50
Min	5	10.00	7.00	5.00	7.00	5.00
Max	3408	5030.00	55.00	50.00	99.00	97.00
Std dev. (n-1)	846.809	898.23	13.20	9.67	18.47	17.23

Table 1: Basic statistics of heavy metals concentration in the study area.

Principal component analysis: PCA minimizes data variability to reveal patterns in the dataset and emphasize dissimilarities and reduction in variables. The eigenvalues and eigenvectors in PCA are extracted from the covariance matrix, which also reflects the dispersion of the observed parameters. When we multiply the initially correlated variables by a vector of coefficients (loadings or scores), we produce new orthogonal variables called principal components (PCs). The original variables are combined in a weighted linear fashion by the PCs (Wunderlin et al. 2001). Projecting data onto a new axis creates a new variable from a PC and an eigenvector. Although there are numerous PCs used as original variables, PCs that provide details on the most important traits, summarize the entire data set, and enable data reduction with little information loss are selected (Helena et al. 2000).

RESULTS AND DISCUSSION

Agglomerative Hierarchal Clustering of Heavy Metals Dataset

The basic statistics of heavy metals contamination levels (mg.kg⁻¹) and their respective contamination factor (CF) are provided in Table 1 & Table 2. Copper and Manganese had a greater mean and standard deviation than the other heavy metals, whereas Zn had the lowest standard deviation. Cu's greater mean concentration and high standard deviation suggest abnormal distribution suggesting a multivariate statistical study. The background concentration for Cu, Mn, Co, Zn, Ni, and Pb is provided in Table 3.

The contamination factor data was subjected to the AHC routine available in XLSTAT®. The AHC algorithm clustered data using Euclidean distance for Dissimilarity and Ward's method for Agglomeration. This resulted in the clustering of the data set into three distinct clusters/groups (Fig. 4). A total of 38 sample locations were grouped into three clusters: 25 unpolluted/low pollution sites (UPS), 11 highly polluted sites (HPS) and 2 extreme pollution sites (EPS). CA indicates that one sampling point per cluster is sufficient to represent the soil quality of the whole cluster spatially. The central sample sites for each cluster are shown in Table 4.

Table 2: Sample-wise contamination factors in the study area.

Samples	CF(Cu)	CF(Mn)	CF(Co)	CF(Zn)	CF(Ni)	CF(Pb)
s1	0.11	0.43	0.42	0.21	0.32	0.65
s2	0.11	0.43	0.53	0.20	0.30	0.50
s6	0.27	0.72	1.05	0.31	0.34	0.80
s7	0.27	0.53	1.50	0.16	0.26	0.45
s8	0.29	0.65	1.11	0.18	0.20	0.40
s11	0.51	0.71	1.03	0.40	0.40	0.65
s14	1.24	1.00	1.58	0.32	0.94	1.60
s15	1.33	0.13	1.11	0.20	0.62	1.30
s16	1.27	0.83	0.71	0.14	0.40	0.85
s17	1.33	0.12	0.39	0.05	0.14	0.25
s18	1.42	0.18	1.97	0.25	0.98	1.55
s19	1.44	0.24	0.63	0.17	0.40	0.70
s21	2.40	0.23	0.39	0.08	0.14	0.25
s22	2.76	0.29	1.50	0.23	0.72	2.35
s23	16.07	1.30	1.66	0.29	0.78	1.10
s24	3.69	2.57	2.89	0.29	1.98	2.45
s26	5.73	0.54	0.87	0.27	0.70	1.45
s27	17.53	1.19	1.50	0.23	1.16	1.80
s28	9.82	0.30	0.79	0.11	0.22	0.40
s34	27.11	2.68	2.68	0.24	1.14	1.80
s35	33.07	1.83	1.11	0.23	0.68	1.65
s36	51.33	1.44	2.61	0.24	0.56	1.45
s37	72.29	0.45	1.58	0.48	0.54	0.90
s38	75.73	0.25	0.79	0.53	0.78	1.55
s3	0.11	0.43	0.74	0.21	0.30	0.30
s4	0.13	0.50	0.58	0.19	0.22	0.35
s5	0.22	0.16	0.37	0.41	0.34	0.75
s9	0.31	0.45	1.18	0.35	0.30	0.65
s10	0.44	0.48	1.50	0.26	0.36	0.70
s12	0.80	1.28	1.82	0.29	1.00	1.40
s13	1.11	1.06	1.97	0.31	0.90	1.15
s20	2.27	0.01	0.71	0.14	0.22	0.75
s25	4.09	0.08	1.26	0.23	0.52	2.00
s29	10.87	0.42	1.42	0.26	0.70	1.05
s30	12.51	0.45	1.34	0.22	0.44	1.55
s31	20.29	1.06	2.68	0.26	0.76	4.85
s32	21.04	1.07	2.45	0.41	0.34	0.45
s33	23.02	0.02	1.26	0.32	0.76	1.95

Table 3: Background concentration of metals in the study.

Metal	Background concentration [mg.kg ⁻¹]	Metal	Background concentration [mg.kg ⁻¹]
Copper (a)	45	Zinc (a)	95
Manganese (a)	850	Nickel (b)	50
Cobalt (a)	19	Lead (a)	20
(a): (Turekian & Wedepohl 1961), (b) (HS1191/HS1191: Nickel Nutrition in Plants)			

Table 4: Central sample for each cluster.

Cluster	CF(Cu)	CF(Mn)	CF(Co)	CF(Zn)	CF(Ni)	CF(Pb)	
1 (s16)	1.270	0.830	0.710	0.140	0.400	0.850	
2 (s33)	23.020	0.020	1.260	0.320	0.760	1.950	
3 (s37)	72.290	0.450	1.580	0.480	0.540	0.900	



Fig. 4: Dendrogram showing spatial similarities of sampling sites.

Table 5: Basic statics of each cluster.

	Statistic	CF _(Cu)	CF _(Mn)	CF _(Co)	CF _(Zn)	CF _(Ni)	CF _(Pb)
Cluster C1 (Unpolluted	No. of Samples	25					
Sample Sites – UPS)	Mean	1.35	0.56	1.11	0.23	0.52	0.97
	Standard deviation (n-1)	1.44	0.53	0.62	0.09	0.40	0.64
Cluster C2 (Highly Polluted Sample Site – HPS)	No. of Samples	11					
	Mean	22.06	1.07	1.77	0.26	0.69	1.64
	Standard deviation (n-1)	11.96	0.77	0.70	0.07	0.29	1.19
Cluster C3	No. of Samples	2					
(Extremely Polluted Samples	Mean	74.01	0.35	1.19	0.51	0.66	1.23
51(65)	Standard deviation (n-1)	2.43	0.14	0.56	0.04	0.17	0.46

Sample s16 represents cluster UPS (unpolluted sample sites with all CFs close to 1.00), Sample 33 represents cluster HPS (highly polluted sample sites with CF_{Cu} of 23.02, meaning Cu concentration was 23 times the crustal

abundance and Sample 37 represents cluster EPS with CF_{Cu} of 72.29 indicating extremely high copper accumulation in cluster C3). The summary of each cluster is presented in Table 5.

According to spatial similarity, three clusters (Table 4 & Fig. 4) are formed with increasing order of metal contamination. The UPS cluster (s1, s2, s6, s7, s8, s11, s14, s15, s16, s17, s18, s19, s21, s22, s24, s26, s3, s4, s5, s9, s10, s12, s13, s20, s25) contains points which are unaffected by the mining waste dump activity because of distance from the mine site. The HPS cluster (s23, s27, s28, s34, s35, s36, s29, s30, s31, s32, s33) contains sample locations that are located near the mine site and receive contamination via the drainage network present in the study area. The EPS cluster contains two locations, namely, s37 and s38, with extremely high copper content because these two locations are adjacent to the mine waste dump (Fig. 2). It can be observed that AHC has segregated the spatial observations in data. The clusters HPS & EPS are polluted sites in the study area and may require attention.

Principal Component Analysis of the Dataset

With the help of a few independent variables, PCA, a potent

Table 6: Results of KMO and Bartlett's test.

pattern recognition approach, can be used to explain the variation of a large dataset made up of many intercorrelated variables. (Ramadan et al. 2000). As demarcated by HCA, the effects of MWD and TSF drainages are critically evaluated through PCA. To evaluate the effects of mining waste dumps (MWD) and tailing storage facilities (TSF), respectively, PCA was used on contamination factor data sets.

Cluster HPS & Cluster EPS served as markers for polluted sites. The PCA was conducted on two data sets. The first dataset included all the clusters named as MWISP dataset (Mine Waste Including Seepage Points), and the second dataset was created by excluding the HPS and EPS clusters MWESP dataset (Mine Waste Excluding Seepage Points). Results of KMO and Bartlett's test are provided in Table 6.

PCA-spatial impact studies on cluster MWISP (mine waste including seepage points): PCA was applied to the MWISP dataset, which contains all the sample points collected in the study area. The scree plot indicated 3 PCs

	Bartlett's sphericity Test (p-Value < 0.05)		Kaiser-Meyer-Olkin measure of sampling adequacy (adequate KMO > 0.500)
MWISP Dataset	Chi-square (Observed value)	248.287	KMO = 0.628 (Test Passed)
Dataset	Chi-square (Critical value)	73.311	
	DF	55	
	p-value (Two-tailed)	< 0.0001	
	alpha	0.050	
MWESP Dataset	Chi-square (Observed value)	168.134	KMO = 0.541 (Test Passed)
Dataset	Chi-square (Critical value)	73.311	
	DF	55	
	p-value (Two-tailed)	< 0.0001	
	alpha	0.050	



Fig. 5: Scree plot for PCA on MWISP data set.

Table 7: Factor loadings for the MWISP dataset (loadings magnitude>0.5 highlighted in bold).

	PC1	PC2	PC3
CF(Cu)	0.773	-0.540	-0.013
CF(Mn)	0.526	0.481	0.341
CF(Co)	0.779	0.189	0.313
CF(Zn)	0.527	0.223	-0.499
CF(Ni)	0.864	0.247	0.130
CF(Pb)	0.837	0.006	0.068
ОМ	-0.333	0.559	-0.637
CEC	0.147	0.834	0.003
pН	-0.331	0.367	0.666
distance	-0.716	0.332	0.206
elevation	0.394	0.265	-0.415

explaining 70% of the total variance in data, although 4 PCs had eigenvalues more than 1.0. Fig. 5 shows the scree plot of PCs generated by the PCA method on the MWISP dataset.

Table 7 & Fig. 6 provides the factor loadings for the PCA analysis on different Principal Components (PCs). In order to be classified as "strong," "moderate," and "weak," loadings must correspond to absolute loading levels of

>0.75, 0.75-0.50, and 0.50-0.30, respectively, according to ((Liu et al. 2003).

The loading plots demonstrated the association between the parameters; the closer the endpoint of variables, the more strongly the values were correlated. (see Fig. 6). It can be observed from Table 7 that the loading for Cu (0.773), Co (0.779), Ni (0.864), and Pb (0.837) in PC1 (explains 37.4 % variance) was strongly positive (factors loadings > 0.75) and for Zn (0.526) and Mn (0.527) the factor loading was moderately positive. It lies between 0.50 - 0.75. The factor loading for the distance(D) parameter was moderately negative (-0.716). This can be explained by the fact that the study area is located near an eco-sensitive zone, and no other means of heavy metal pollution are present. Therefore, being the only source in the study area, the pollution decreases as the distance from the mine site increases. Also, Table 2 shows that the contamination factor is highest for copper with a mean of 11.17, whereas CF for Mn, Co, Zn, Ni, and Pb show a mean value of 0.70, 1.31, 0.25, 0.58 & 1.18, respectively. Indicating high copper contamination in the samples. Therefore, copper dominates over other elements in the polluted samples, and it may also explain the relatively lower loading for PCs of Mn and Zn, which are essential soil elements.

PC1 explains that heavy metal pollution has occurred in the study area with copper contamination to a high degree.



Fig. 6: Loading plot of PCA on MWISP dataset.

Soil pH is negatively correlated with the metal content indicating that the acidic discharge for mine waste dump and tailing storage facility has high metallic content. PC2 explains 18.12% of the variance in the data. It exhibits moderate negative loading of Cu (-0.540), whereas Mn (0.481), Co (0.189), Zn (0.223), Ni (0.247), and Pb (0.006) had low/weak factor loadings. OM (0.559) and CEC (0.834) have moderate and strong positive loading, respectively. The capacity of a particle to exchange positive bases in response to its surroundings is known as the CEC. From the surfaces of clay minerals and organic materials, cations can swap for another positively charged ion (Aprile & Lorandi 2012). Therefore, the presence of OM can increase the CEC of soil. In PC2, negative loading of Cu with positive loading of OM and CEC indicates that acidic discharge from the mine may deplete the OM initially present in the soil and, in turn, accumulate copper in the soil matrix.

The PCA analysis can delineate the process of soil degradation by acidic mine drainage. Initially, fertile soil with high OM is attacked by the acidic discharge consuming the OM. Thus, the pH value of the discharged solution rises, which in turn precipitates heavy metal in the soil. PC3 explains 14% of the variance in the data, with pH (0.666) and OM (-0.637) showing moderate positive and negative loadings. By releasing hydrogen ions associated with organic anions or by nitrifying in an open system, organic matter typically reduces soil pH (Porter 1980). The biplot of PC1 and PC2 with sample data and parameters provides



Fig. 7: Biplot of PCs with Sample Locations.



Fig. 8: Scree plot of PCs.

	F1	F2	F3
CF(Cu)	0.441	-0.820	0.181
CF(Mn)	0.609	0.462	-0.250
CF(Co)	0.833	-0.066	-0.286
CF(Zn)	0.612	0.505	-0.055
CF(Ni)	0.933	-0.114	-0.195
CF(Pb)	0.844	-0.321	-0.088
OM	0.005	0.792	0.414
CEC	0.589	0.491	0.270
рН	0.070	0.128	-0.544
distance	-0.360	0.418	-0.639
elevation	0.364	0.145	0.680

Table 8: Factor loadings for the MWESP dataset (loadings > 0.5 highlighted in bold).

a spatial distribution of heavy metals in the study area (Fig. 7).

PCA-spatial impact studies on MWESP (mine waste excluding seepage points): PCA was conducted on a dataset

excluding the seepage points, namely MWESP (Mine Waste Excluding Seepage Points). The scree plot indicated 3 PCs explaining 70% of the total variance in data, although 4 PCs had eigenvalues more than 1.0. Fig. 8 shows the scree plot of PCs generated by the PCA method on the MWESP dataset.

Table 8 provides the factor loadings of PCs after PCA analysis on the MWESP dataset. It can be observed that Copper had a strong positive loading of 0.773 in PC1 of the MWISP dataset, whereas it falls even below 0.5 to 0.441 in PC1 of the MWESP dataset. This indicates that sample locations directly receive AMD drainage emanating from mine waste dumps and tailing storage facilities. AMD contains Copper and these locations are affected by it. The dataset MWESP contains unpolluted or very low contamination sites; therefore, typical soil behavior is observed with CEC-linked metal content in the soil with a moderate positive loading of 0.589 in PC1.

PC2 explains 21% of the variance in the dataset, where OM is positively correlated with Zn and Mn and a strongly negative correlation with Cu. High OM loadings mean un-



Fig. 9: Spatial map of copper contamination factor using inverse distance weighing method.

disturb soil with positive correlation with essential elements like Zn, Mn and Co. PC3 explains 14% of the variance in data with moderate negative loading of pH (-0.544), Distance (-0.639) and moderate positive loading of elevation (0.680). The negative correlation between elevation and pH is because the mining area is located at a relatively higher elevation; therefore, soil quality near the mine site is acidic due to the impact of acidic discharge by the mine waste sites.

Spatial Mapping of Copper Contamination Using Geostatistical Inverse Distance Weighing Method

The Inverse Distance Weighting (IDW) method is a spatial analysis and geo-statistics technique primarily used for interpolation. IDW is based on the premise that spatial entities that are close to each other are more alike than those further apart (Shepard 1968). The general formula for IDW is as follows:

$$Z\mathbf{p} = \frac{\sum_{i=1}^{i=n} \left(\frac{Z_i}{d_i^p}\right)}{\sum_{i=1}^{i=n} \left(\frac{1}{d_i^p}\right)}$$

Here, Zp is the estimated value at point (p), Z_i is the known value at the (i)-th point, (d_i) is the distance between the point (i) and (p), and (p) is a power parameter that dictates the rate at which influence decreases with distance. The distance is typically calculated using Euclidean distance, though other distance metrics can be used. IDW is employed in various fields such as meteorology, geology, agriculture, and environmental science for interpolating data like temperature, mineral concentrations, and pollution levels. The copper contamination is shown in the red-shaded region in Fig. 9.

A spatial map of copper contamination was generated using the Inverse Distance Weighing method in QGIS software. It can be observed that a significant elevation of copper content is present in the vicinity of the Mine Waste Dump and Tailing Storage Facility within 1.25 kilometers of waste sites in the study area.

CONCLUSIONS

The study revealed that 13 out of the 38 samples (HPS & EPS Clusters) showed signs of contamination however, it is contained to flow paths. The soil samples not located on the flow paths indicated little to no contamination. The contaminated sites are situated along the drainage routes for AMD seepage emanating from the two mine waste sites (MWD & TSF), indicating that such flow paths are particularly vulnerable to Acid Mine Drainage (AMD) effects.

This study examined how mine waste dumps/storages and drainage from tailing storage facilities affected surface soil/sediments using multivariate statistical approaches. PCA and HCA were used to evaluate spatial variation in the complicated data. AHC categorization found three statistically significant soil heavy metal clusters. The PCA analysis between MWISP and MWESP datasets shows that excluding seepage spots from mining waste dumps and tailing storage facilities reduces pollution throughout the study area. PCA found several correlations: (1) Heavy metal content returns to normal with an increase in distance from the mining site. (2) Copper dominates other elements in polluted samples and lowers PC loadings of soil elements such as Mn and Zn. (3) Acidic mine discharge may deplete soil OM and increase soil matrix copper by precipitation and mechanical deposition. PCA analysis shows how acidic mine drainage degrades soil quality. The acidic discharge attacks fertile soil with high OM, consuming it and raising the pH, which precipitates heavy metals in the soil. Multivariate statistical methods helped analyze complicated data and understand their spatial variance by eliminating redundant data variables.

The TSF and MWD were identified as probable AMD sites during the field survey; therefore, mine waste is an environmental hazard and must be taken seriously. The majority of heavy metal accumulation associated with Acid Mine Drainage (AMD) in the study area was attributed to the Tailings Storage Facility (TSF) and Mine Waste Dumps (MWD). It is essential to segregate these areas from nearby drainage systems and ensure that their discharge is confined within the boundary of the mine site.

ACKNOWLEDGEMENT

Anoop Kant Shukla, Manoj Pradhan, and Onkar Nath Tiwari conceived and designed the research work. Anoop Kant Shukla and Manoj Pradhan conducted the experiments. Anoop Kant Shukla analyzed the data and wrote the manuscript. All authors read and approved the manuscript.

REFERENCES

- Addinsoft. 2023. XLSTAT statistical and data analysis solution (XLSTAT 2022.1.1). Addinsoft, New York, USA.
- Ajmone-Marsan, F. and Biasioli, M. 2010. Trace elements in soils of urban areas. Water, Air, & Soil Pollut., 213(1): 121-143. https://doi. org/10.1007/S11270-010-0372-6
- Aprile, F. and Lorandi, R. 2012. Evaluation of cation exchange capacity (cec) in tropical soils using four different analytical methods. J. Agric. Sci., 4(6): p278. https://doi.org/10.5539/JAS.V4N6P278
- Astel, A., Tsakovski, S., Barbieri, P. and Simeonov, V. 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data



sets. Water Res., 41(19): 4566-4578. https://doi.org/10.1016/J. WATRES.2007.06.030

- Biasioli, M., Barberis, R. and Ajmone-Marsan, F. 2006. The influence of a large city on some soil properties and metals content. Sci. Total Environ., 356(1-3): 154-164. https://doi.org/10.1016/J. SCITOTENV.2005.04.033
- BIS 2720-22. 2020. BIS 2720-22. Bureau of Indian Standards. Retrieved from https://archive.org/details/gov.in.is.2720.22.1972
- BIS 2720-24. 2020. BIS 2720-24. Bureau of Indian Standards. Retrieved from https://archive.org/details/gov.in.is.2720.24.1976
- BIS 2720-26. 2021. BIS 2720-26. Bureau of Indian Standards. Retrieved from https://archive.org/details/gov.in.is.2720.26.1987
- Box, G.E.P. and Cox D.R. 1964. An analysis of transformations. J. R. Stat. Soc. B., 26. https://www.jstor.org/stable/2984418
- Chen, T.B., Zheng, Y.M., Lei, M., Huang, Z.C., Wu, H.T., Chen, H., Fan, K.K., Yu, K., Wu, X. and Tian, Q.Z. 2005. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. Chemosphere, 60(4): 542-551. https://doi.org/10.1016/J. CHEMOSPHERE.2004.12.072
- Chirenje, T., Ma, L.Q., Reeves, M. and Szulczewski, M. 2004. Lead distribution in near-surface soils of two Florida cities: Gainesville and Miami. Geoderma, 119(1-2): 113-120. https://doi.org/10.1016/ S0016-7061(03)00244-1
- Dragović, S., Mihailović, N. and Gajić, B. 2008. Heavy metals in soils: distribution, relationship with soil characteristics and radionuclides and multivariate assessment of contamination sources. Chemosphere, 72(3): 491-495. https://doi.org/10.1016/j.chemosphere.2008.02.063
- Environmental Protection Agency, U. 1996. Method 3050b acid digestion of sediments, sludges, and soils 1.0 scope and application. Method 3050B: Acid Digestion of Sediments, Sludges, and Soils. https://www.epa.gov/ sites/default/files/2015-06/documents/epa-3050b.pdf
- Facchinelli, A., Sacchi, E. and Mallen, L. 2001. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. Environ. Pollut., 114(3): 313-324. https://doi.org/10.1016/S0269-7491(00)00243-8
- Gordon, R.B., Bertram, M. and Graedel, T.E. 2006. Metal stocks and sustainability. Proc. Natl. Acad. Sci. U.S.A., 103(5): 1209-1214. https:// doi.org/10.1073/pnas.0509498103
- Greenwood, N.N. and Earnshaw, A. 1984. Chemistry of the Elements (1st Edition), 268-327.
- Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J.M. and Fernandez, L. 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res., 34(3): 807-816. https://doi.org/10.1016/S0043-1354(99)00225-0
- HS1191/HS1191: Nickel Nutrition in Plants. Retrieved November 28, 2023, from https://edis.ifas.ufl.edu/publication/HS1191
- Hussain, M., Ahmed, S.M. and Abderrahman, W. 2008. Cluster analysis and quality assessment of logged water at an irrigation project, eastern Saudi Arabia. J. Environ. Manage., 86(1): 297-307. https://doi.org/10.1016/J. JENVMAN.2006.12.007
- Jarva, J., Tarvainen, T., Lintinen, P. and Reinikainen, J. 2009. Chemical characterization of metal-contaminated soil in two study areas in Finland. Water, Air, Soil Pollut., 198(1-4): 373-391. https://doi. org/10.1007/S11270-008-9852-3
- Lamontagne, L. and Camire, C. 1987. Soil analysis and numerical classification of the Lanoraie delta, Quebec. Can. J. Soil Sci., 67(3): 417-432. https://doi.org/10.4141/cjss87-041
- Lee, C.S.L., Li, X., Shi, W., Cheung, S.C.N. and Thornton, I. 2006. Metal contamination in urban, suburban, and country park soils of Hong Kong: a study based on GIS and multivariate statistics. Sci. Total Environ., 356(1-3): 45-61. https://doi.org/10.1016/J.SCITOTENV.2005.03.024
- Liu, C.W., Lin, K.H. and Kuo, Y.M. 2003. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in

Taiwan. Sci. Total Environ., 313(1-3): 77–89. https://doi.org/10.1016/ S0048-9697(02)00683-6

- Li, X., Liu, L., Wang, Y., Luo, G., Chen, X., Yang, X., Hall, M.H.P., Guo, R., Wang, H., Cui, J. and He, X. 2013. Heavy metal contamination of urban soil in an old industrial city (Shenyang) in Northeast China. Geoderma, 192(1): 50-58. https://doi.org/10.1016/J.GEODERMA.2012.08.011
- Lu, S.G. and Bai, S.Q. 2010. Contamination and potential mobility assessment of heavy metals in urban soils of Hangzhou, China: relationship with different land uses. Environ. Earth Sci., 60(7): 1481-1490. https://doi.org/10.1007/S12665-009-0283-2
- Maas, S., Scheifler, R., Benslama, M., Crini, N., Lucot, E., Brahmia, Z., Benyacoub, S. and Giraudoux, P. 2010. Spatial distribution of heavy metal concentrations in urban, suburban and agricultural soils in a Mediterranean city of Algeria. Environ. Pollut. (Barking, Essex: 1987), 158(6): 2294-2301. https://doi.org/10.1016/J.ENVPOL.2010.02.001
- Meadows, A.E. and Carpenter, A.A. 1997. Acid Mine Drainage and Groundwater Pollution. Groundwater Pollut. Primer, CE, 4594.
- Micó, C., Recatalá, L., Peris, M. and Sánchez, J. 2006. Assessing heavy metal sources in agricultural soils of a European Mediterranean area by multivariate analysis. Chemosphere, 872-863 :(5)65. https://doi. org/10.1016/j.chemosphere.2006.03.016
- Navas, A. and Machin, J. 2002. Spatial distribution of heavy metals and arsenic in soils of Aragón (northeast Spain): controlling factors and environmental implications. Appl. Geochem., 17(8): 961-973. https:// doi.org/10.1016/S0883-2927(02)00006-9
- Pandey, P.K., Sharma, R., Roy, M. and Pandey, M. 2007. Toxic mine drainage from Asia's biggest copper mine at Malanjkhand, India. Environ. Geochem. Health, 29(3): 237-248. https://doi.org/10.1007/ s10653-006-9079-4
- Pîrnău, R.G., Patriche, C.V., Roşca, B., Vasiliniuc, I., Vornicu, N. and Stanc, S. 2020. Soil spatial patterns analysis at the ancient city of Ibida (Dobrogea, SE Romania), via portable X-ray fluorescence spectrometry and multivariate statistical methods. Catena, 189: 104506. https://doi. org/10.1016/j.catena.2020.104506
- Porter, W.M. 1980. Soil acidity: is it a problem in Western Australia? J. Dept. Agric. West. Australia, Ser. 4, 21(4): 126-133. https://researchlibrary. agric.wa.gov.au/journal_agriculture4/vol21/iss4/8
- Ramadan, Z., Song, X.H. and Hopke, P.K. 2000. Identification of sources of Phoenix aerosol by positive matrix factorization. J. Air Waste Manag. Assoc. (1995), 50(8): 1308-1320. https://doi.org/10.1080/10473289.2 000.10464173
- Razmkhah, H., Abrishamchi, A. and Torkian, A. 2010. Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: A case study on Jajrood River (Tehran, Iran). J. Environ. Manage., 91(4): 852-860. https://doi.org/10.1016/j.jenvman.2009.11.001
- Shepard, D. 1968. A two-dimensional interpolation function for irregularlyspaced data. Proc. 1968 23rd ACM Natl. Conf. On, 517-524. https://doi. org/10.1145/800186.810616
- Shukla, A.K., Pradhan, M. and Tiwari, O.N. 2018. Impact on sediments and water by release of copper from chalcopyrite bearing rock due to acidic mine drainage. AIP Conf. Proc., 1952(1): 020064. https://doi. org/10.1063/1.5032026
- Singh, K.P., Malik, A. and Sinha, S. 2005. Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques - A case study. Anal. Chim. Acta, 538(1-2): 355-374. https://doi.org/10.1016/J.ACA.2005.02.006
- Swarnakar, A.K., Bajpai, S. and Ahmad, I. 2023. Performance evaluation of different soil media by batch-operated pilot-scale horizontal subsurface flow constructed wetlands for wastewater treatment. Pollut., 4(9): 1567-1578. https://doi.org/10.22059/POLL.2023.357480.1855
- Tariq, S.R., Shah, M.H., Shaheen, N., Jaffar, M. and Khalique, A. 2008. Statistical source identification of metals in groundwater exposed to industrial contamination. Environ. Monit. Assess., 138(1-3): 159-165. https://doi.org/10.1007/S10661-007-9753-8

- Tiwari, O.N., Shukla, A.K., Pradhan Manoj, and Nandy Tapas. 2017. Migration of copper to nearby adjoining surface water bodies from tailing impoundment of an open pit copper mine. Ecol., Environ. Conserv., 24(1): 306-314.
- Tume, P., Bech, J., Longan, L., Tume, L., Reverter, F. and Sepulveda, B. 2006. Trace elements in natural surface soils in Sant Climent (Catalonia, Spain). Ecol. Eng., 27(2): 145-152. https://doi.org/10.1016/j. ecoleng.2006.01.004
- Turekian, K.K. and Wedepohl, K.H. 1961. Distribution of the elements in some major units of the earth's crust. GSA Bull., 72(2): 175-192. https://doi.org/10.1130/0016-7606(1961)72[175]2.0.CO;2
- Varol, M. 2011. Assessment of heavy metal contamination in sediments of the Tigris River (Turkey) using pollution indices and multivariate statistical techniques. J. Hazard. Mater., 195: 355-364. https://doi. org/10.1016/J.JHAZMAT.2011.08.051
- Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc., 58(301): 236-244. https://doi.org/10.1080/01621 459.1963.10500845

- Wong, C.S.C., Li, X. and Thornton, I. 2006. Urban environmental geochemistry of trace metals. Environ. Pollut., 142(1): 1-16. https:// doi.org/10.1016/J.ENVPOL.2005.09.004
- Wong, S.C., Li, X.D., Zhang, G., Qi, S.H. and Min, Y.S. 2002. Heavy metals in agricultural soils of the Pearl River Delta, South China. Environ. Pollut., 119(1): 33-44. https://doi.org/10.1016/S0269-7491(01)00325-6
- Wunderlin, D.A., María Del Pilar, D., María Valeria, A., Fabiana, P.S., Cecilia, H.A. and María De Los Ángeles, B. 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A Case Study: Suquía River basin (Córdoba-Argentina). Water Res., 35(12): 2881-2894. https://doi.org/10.1016/S0043-1354(00)00592-3

ORCID DETAILS OF THE AUTHORS

Anoop Kant Shukla: https://orcid.org/0000-0003-2135-9672

